PONTIFICIAE ACADEMIAE SCIENTIARVM SCRIPTA VARIA
92

HUMAN GENOME

ALTERNATIVE ENERGY SOURCES FOR DEVELOPING COUNTRIES

FUNDAMENTAL PRINCIPLES OF MATHEMATICS

ARTIFICIAL INTELLIGENCE

PROCEEDINGS

Plenary Session of the Pontifical Academy of Sciences
25-29 October 1994

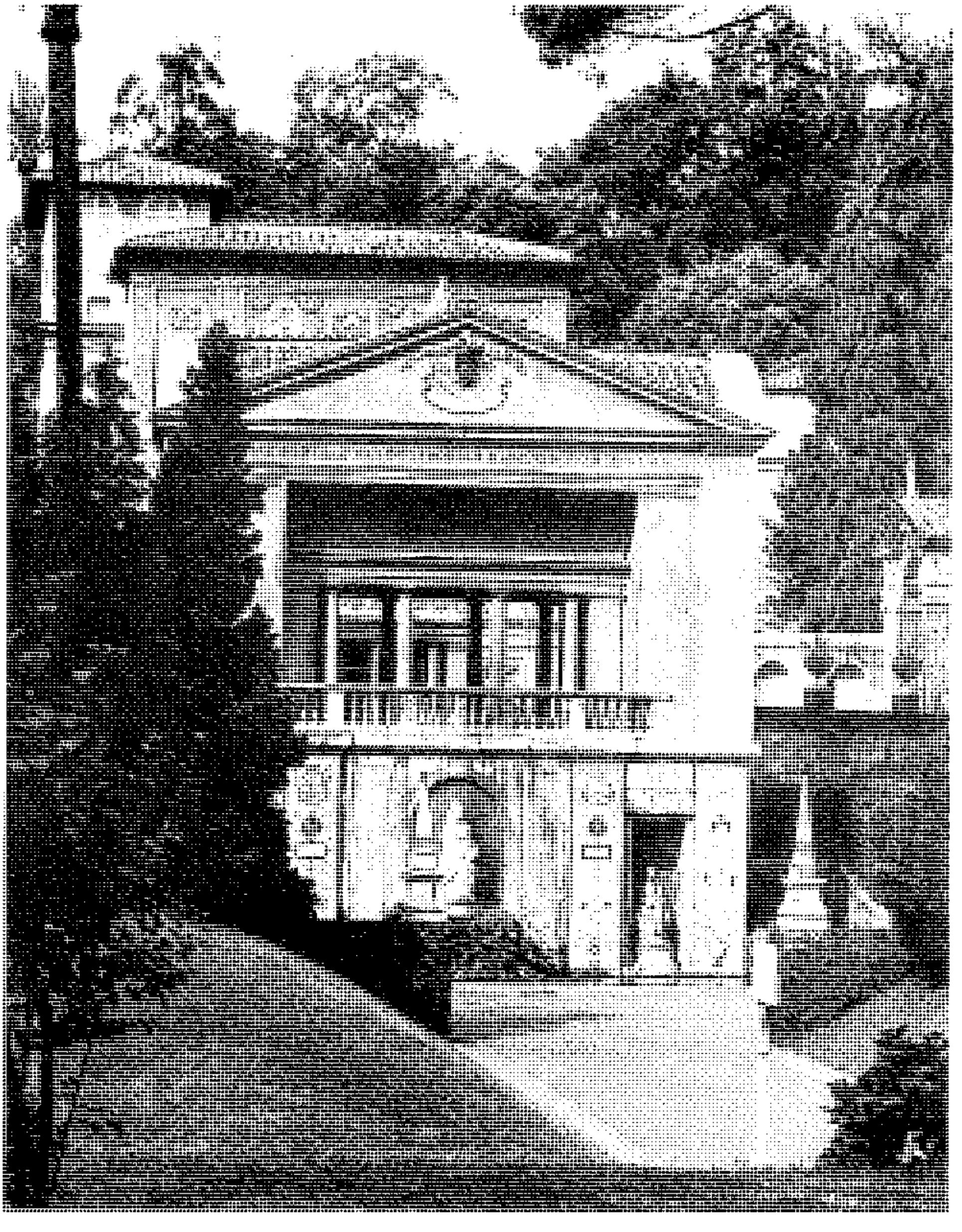
PONTIFICIA
ACADEMIA
SCIENTIARVM

EX AEDIBVS ACADEMICIS IN CIVITATE VATICANA

MCMXCVIII

The opinions expressed with absolute freedom during the presentation of the papers and in the subsequent discussions by the participants in the Plenary Session, although published by the Academy, represent only the points of view of the participants and not those of the Academy.

ISBN 88-7761-066-2



CONTENTS

List of Participants 13

Allocution of the Holy Father John Paul II 17

Address of the President of the Pontifical Academy of Sciences at the Solemn Pontifical Audience 25

Opening Address by the President of the Pontifical Academy of Sciences (Nicola CABIBBO) 27

SCIENTIFIC PAPERS

HUMAN GENOME

B. PULLMAN: Introduction 31

I. *Scientific Aspects*

C. HÉLÈNE: Nucleic Acids Structures and the Control of Gene Expression 37

J. WEISSENBACH: The Human Genome Project 47

L. CAVALLI-SFORZA: The Human Genome Diversity Project 61

M.F. SINGER: From Genomic Junk to Human Disease 77

II. *Legal Aspects*

R. DE MENDIZABAL ALLENDE: La investigación genética en el marco de la constitución española: sobre el derecho a la intimidad . 95

S. GRISOLIA: The Human Genome Project and the Third World . 105

C.M. ROMEO-CASABONA: Legal Protection of Values by the Human Genome Research 119

III. *Ethical Aspects*

J.-M. MALDAMÉ: Identité humaine et génétique 157

W.B. FRIEND: Pastoral Concerns About the Human Genome Initiative 171

ALTERNATIVE ENERGY SOURCES
FOR DEVELOPING COUNTRIES

A. BLANC-LAPIERRE: Introduction 193

U. FARINELLI: Alternative Energy Sources for the Third World: Perspectives, Barriers, Opportunities 195

M. MOSHINSKY: Energy Resources: Present and Future in Mexico . 229

J. DÖBEREINER: Alternatives for Biofuel Production: the Brazilian Example 239

C.N.R. RAO: Alternative Energy Strategies for Sustainable Development of India 245

F.M. WA KALENGA: African Energy Problems, Optimal Energy Policy and the Pertinence of the Concept of Alternative Energies 249

FUNDAMENTAL PRINCIPLES OF MATHEMATICS
AND ARTIFICIAL INTELLIGENCE

I. *Fundamental Principles of Mathematics*

E. DE GIORGI: Fundamental Principles of Mathematics 269

L.A. CAFFARELLI: Modelling Analysis and Numerical Simulation . . 279

II. *Artificial Intelligence*

J.L. Lions: Contrôle et Intelligence Artificielle 285

A.M. Battro: Intellectual Prostheses. Theory and Practice 293

B. Gulyás: Functional Brain Activation Studies in Man with Advanced Neuroimaging Techniques: the Experimental Basis for Computational Theories 305

S. Ullman: Artificial Intelligence and Computer Vision 325

PONTIFICAL ACADEMY OF SCIENCES' ACTIVITY CONCERNING THE ANALYSIS OF THE IMPACT OF SCIENCE ON ECONOMIC AND SOCIAL LIFE

B.M. Colombo and A. Golini: Resources and Population 335

C. Ramel: Chemical Hazards in Developing Countries 341

M. Oda: From the Exploration of the Universe to the Exploration of the Human Brain 347

On the previous page:
Casina Pio IV, Vatican City, Seat of the Pontifical Academy of Sciences.

Participants in the Plenary Session.

LIST OF PARTICIPANTS

I - ACADEMICIANS

ANFINSEN Christian B.	(U.S.A)
ARBER Werner	(Switzerland)
BLANC-LAPIERRE André	(France)
CABIBBO Nicola	(Italy)
CAFFARELLI Luis A.	(U.S.A.)
CAVALLI-SFORZA Luigi L.	(U.S.A.)
COLOMBO Bernardo M.	(Italy)
COTTIER Georges M.M.	(Vatican City)
COYNE George V.	(Vatican City)
CROMBIE Alistair C.	(United Kingdom)
DALLAPORTA Nicola	(Italy)
DE GIORGI Ennio	(Italy)
DÖBEREINER Johanna	(Brazil)
GERMAIN Paul	(France)
HELLER Michael	(Poland)
JAKI Stanley L.	(U.S.A.)
KEILIS-BOROK Vladimir I.	(Russia)
LEVI-MONTALCINI Rita	(Italy)
LICHNEROWICZ André	(France)
MALU Felix Wa Kalenga	(Zaire)

MOSHINSKY Marcos	(Mexico)
MURADIAN Rudolf	(Armenia)
ODA Minoru	(Japan)
PAVAN Crodowaldo	(Brazil)
PONNAMPERUMA Cyril A.	(U.S.A)
PULLMAN Bernard	(France)
PUPPI Giampietro	(Italy)
RAO Chintamani N.R.	(India)
REES Martin J.	(United Kingdom)
RUNCORN Stanley K.	(United Kingdom)
SELA Michael	(Israel)
SINGER Maxine F.	(U.S.A.)
SOUTHWOOD Richard	(United Kingdom)
SZCZEKLIK Andrzej	(Poland)
TOWNES Charles H.	(U.S.A.)
TUPPY Hans	(Austria)
WHITE Robert J.	(U.S.A.)

II - EXPERTS

Dr. Antonio M. BATTRO: *Battro & Denham*, Libertad 1365 - *Buenos Aires* 1016 (Argentina).

Rev. P. Hervé CARRIER, S.J.: *Centre Justice et Foi*, Maison Saint Robert Bellarmin, 25, rue Jarry Ouest - *Montreal* H2P 1S6 (Canada).

Prof. Ugo FARINELLI: *Ente per le Nuove Tecnologie, l'Energia e l'Ambiente (ENEA)*, Viale Regina Margherita, 125 - I-00198 *Roma* (Italy).

H.E. Msgr. William B. FRIEND, Bishop of Shreveport: *Diocese of Shreveport*, 2500 Line Avenue - *Shreveport*, LA 71104-3043 (U.S.A.).

Prof. Santiago GRISOLIA (Fundación BBV *): *Fundación Valenciana de Investigaciones Biomédicas*, Instituto de Investigaciones Citológicas, Amedeo de Saboya, 4 - 46010 *Valencia* (Spain).

Prof. Balázs GULYAS: *Karolinska Institute*, Nobel Institute of Neurophysiology, Laboratory for Brain Research and PET - S-171 77 *Stockholm* (Sweden).

Prof. Claude HELENE: *Muséum National d'Histoire Naturelle*, Laboratoire de Biophysique, INSERM unité 201 - C.N.R.S. unité associée 481, 43, rue Cuvier - F-75231 *Paris* Cedex 05 (France).

Rev. P. Jean-Michel MALDAMÉ, O.P.: *Ecole de Théologie*, 1, avenue Lacordaire - F-31078 *Toulouse* (France).

Dr. Rafael DE MENDIZABAL ALLENDE (Fundación BBV*): *Tribunal Constitucional*, Domenico Scarlatti, 6 - *Madrid* (Spain).

Prof. Claes RAMEL: *Stockholm University*, Department of Genetic and Cellular Toxicology, Wallenberglaboratoriet - S-106 91 *Stockholm* (Sweden).

Prof. Carlos M. ROMEO-CASABONA (Fundación BBV*): *Universidad de Deusto*, Cátedra de Derecho y Genoma Humano, Avda. Universidades, 24, Apartado 1/E 48080 - 48007 *Bilbao* (Spain).

Prof. Shimon ULLMAN: *The Weizmann Institute of Science*, POB 26 - 76100 *Rehovot* (Israel).

Prof. Jean WEISSENBACH: *Genethon*, 1, rue de l'Internationale - B.P. 60 - F-91002 *Evry* Cedex (France).

H.E. Msgr. Jozef ZYCINSKI, Bishop of Tarnow: ul. I. Moscickiego 9 - PL-33-100 *Tarnow* (Poland).

III - INVITED OBSERVERS

BASTI Gianfranco, Pontificia Università Lateranense.

CARBONELL Rafael, Pontificia Università Gregoriana.

* Fundación Banco Bilbao Vizcaya (Spain).

CUMO Maurizio, ENEA, Viale Regina Margherita 125 - 00198 *Roma* (Italy).

GAHL Robert A., Ateneo Romano della Santa Croce.

GATTI Guido, Università Pontificia Salesiana.

KIELY Bartholomew, Pontificia Università Gregoriana.

Madame Birthe LEJEUNE, 31, rue Galande - F-75005 *Paris* (France).

MARTINEZ Rafael, Ateneo Romano della Santa Croce.

MASELLA Alessandro, Pontificia Università Lateranense.

PELLEREY Michele, Università Pontificia Salesiana.

POLIZZI Vincenzo, Università Pontificia Salesiana.

SANGUINETI Juan José, Ateneo Romano della Santa Croce.

SANI Luciano, Università di Roma, Piazza S. Pietro in Vincoli, 10 - 00153 *Roma* (Italy).

SANNA Ignazio, Pontificia Università Lateranense.

SARTORI Lino, *Roma* (Italy).

TANZELLA NITTI Giuseppe, Ateneo Romano della Santa Croce.

VILLANUEVA Javier, Ateneo Romano della Santa Croce.

WELTEN Willibrod, Pontificia Università Gregoriana.

SOLEMN PONTIFICAL AUDIENCE AT THE PLENARY SESSION OF THE PONTIFICAL ACADEMY OF SCIENCE

On 28 October Pope John Paul II received the Pontifical Academicians at a solemn audience.

ALLOCUTION OF THE HOLY FATHER JOHN PAUL II

Excellences,
Messeigneurs,
Mesdames, Messieurs,

1. C'est pour moi une grande joie de vous rencontrer au cours de la Session plénière annuelle de l'Académie pontificale des Sciences. J'adresse à chacun de vous un salut déférent et cordial, en vous renouvelant l'assurance de mon attention et de mon estime pour vos activités au sein de l'Académie.

Au début de notre entrevue, je voudrais tout d'abord honorer la mémoire des sept membres illustres de votre assemblée qui sont morts au cours de l'année écoulée. Je prie le Seigneur de leur accorder la récompense éternelle, en souhaitant que leurs contributions au travail de l'Académie demeurent des points de repère et soient une invitation à poursuivre inlassablement la recherche, pour le service de la vérité et pour le service de nos frères, car c'est de la vérité que découle la dignité humaine (cf. *Veritatis splendor*, n. 63).

2. Votre session plénière est l'occasion de publier la nomination des nouveaux académiciens, appelés à participer à la vie de l'Académie grâce à leurs compétences et à leur travaux largement reconnus. Je suis heureux de saluer leur arrivée, qui accentue la dimension internationale de votre assemblée, ouverte ainsi à de nouvelles disciplines scientifiques. Cela vous permet d'être davantage à l'écoute des techniques et des sciences qui ne cessent de progresser dans tous les continents.

Car les interrogations auxquelles notre société est affrontée requièrent de plus en plus l'éclairage des sciences, qui sont une des grandes richesses de notre monde sans cesse en développement et en mutation.

Mais, dans le même temps, on ne doit pas perdre de vue que la science ne peut prétendre rendre compte à elle seule de l'origine transcendante et de la finalité ultime de l'existence humaine; tout chercheur est invité à tenir compte des interrogations métaphysiques et morales, qui se font plus pressantes lorsque la certitude obtenue par la science est confrontée à la vérité intégrale sur l'homme.

3. Dans le programme de travail de la présente session comme dans vos précédentes réunions, vous accordez une place importante à la question du génome humain, qui est un enjeu essentiel pour l'avenir des personnes et de l'humanité. J'apprécie que, face à une telle interrogation, vous poursuiviez inlassablement la réflexion, afin de proposer à nos contemporains une analyse où se lient, sans contradiction, le constat scientifique et la vérité intégrale de ce qu'est objectivement l'homme.

La découverte progressive de la carte génétique et les précisions de plus en plus fines du séquençage du génome, investigations qui prendront encore plusieurs années, sont une avancée dans les connaissances scientifiques qui suscite tout d'abord un émerveillement légitime, en particulier en ce qui concerne la reconstitution de la chaîne d'A.D.N., base chimique des gènes et des chromosomes. Il semble désormais acquis que, pour toutes les espèces vivantes y compris l'homme, l'A.D.N. soit le support des caractères héréditaires et de leur transmission à la descendance. Les multiples conséquences pour l'homme, qui ne peuvent être encore totalement discernées, sont porteuses de promesses.

En effet, dans un avenir désormais assez proche, on peut raisonnablement envisager que le séquençage intégral du génome offrira de nouvelles voies à la recherche à finalité thérapeutique. Ainsi, des malades qui ne pouvaient pas être soignés de manière adéquate, par suite de pathologies héréditaires souvent létales, pourront désormais bénéficier des traitements nécessaires à l'amélioration de leur état et à une éventuelle guérison. En agissant sur les gènes malades du sujet, on pourra aussi prévenir la manifestation de maladies génétiques et leur transmission.

La recherche sur le génome permettra à l'homme de se comprendre lui-même, à un niveau jusqu'alors jamais atteint. En particulier, on pourra ainsi mieux percevoir les conditionnements génétiques, et les distinguer de ceux qui proviennent de l'entourage naturel et culturel et de ceux qui sont liés à l'expérience propre de l'individu. De plus, en mettant en lumière les réseaux de conditionnements dans lesquels se déploie la liberté de l'homme, nous parviendrons à en saisir plus clairement la réalité mystérieuse.

Certaines personnes seront peut-être tentées de rechercher une explication uniquement scientifique de la liberté humaine, et de la tenir pour suffisante. Une telle explication reviendrait à nier ce qu'elle tend à expliquer; elle irait à l'encontre de l'évidence intime et irréfutable que notre moi profond ne se réduit pas aux conditionnements dont il peut être tributaire, mais qu'il demeure en définitive le seul auteur de nos décisions.

Des progrès scientifiques comme ceux qui portent sur le génome honorent la raison de l'homme, appelé à être seigneur de la création, et ils honorent le Créateur, source de toute vie, qui a confié à l'humanité la gestion du monde. Les découvertes de la complexité de la structure moléculaire peuvent inviter les membres de la communauté scientifique, et plus largement l'ensemble de nos contemporains, à s'interroger sur la causalité première, sur Celui qui est à l'origine de toute existence et qui a façonné chacun de nous dans le secret (cf. *Ps* 139, 15; *Pr* 24, 12).

4. En ce qui concerne les interventions sur le séquençage du génome humaine, il convient de rappeler quelques règles morales fondamentales. Toute action sur le génome doit s'effectuer dans le respect absolu de la spécificité de l'espèce humaine, de la vocation transcendantale de tout être et de son incomparable dignité. Le génome représente l'identité biologique de chaque sujet; plus encore, il exprime une part de la condition humaine de l'être, voulu par Dieu pour lui-même, grâce à la mission confiée à ses parents.

Le fait de pouvoir établir la carte génétique ne doit pas conduire à réduire le sujet à son patrimoine génique et aux altérations qui peuvent y être inscrites. Dans son mystère, l'homme dépasse l'ensemble de ses caractéristiques biologiques. Il est une unité fondamentale, dans laquelle le biologique ne peut être séparé de la dimension spirituelle, familiale et sociale, sans courir le risque grave de supprimer ce qui est la nature même de la personne et de n'en faire qu'un simple objet d'analyse. La personne humaine, par sa nature et par sa singularité, est la norme de toute recherche scientifique. Elle «est et doit rester le principe, le sujet et la fin» de toute recherche (Conc. Vat. II, *Gaudium et spes*, n. 25).

A ce propos, on se réjouit du refus de nombreux chercheurs de considérer que les découvertes effectuées sur le génome puissent constituer des brevets susceptibles d'être enregistrés. Parce que le corps humain n'est pas un objet dont on peut disposer, les résultats des investigations sont à diffuser à l'ensemble de la communauté scientifique et ne peuvent pas être la propriété d'un petit groupe.

La réflexion éthique doit aussi porter sur l'utilisation des données médicales concernant les individus, spécialement celles qui sont contenues dans le génome et qui pourraient être exploitées par la société au détriment des personnes, par exemple en éliminant les embryons porteurs d'anomalies chromosomiques ou en marginalisant les sujets affectés de telle ou telle maladie génétique; on ne peut pas non plus violer les secrets biologiques de la personne, ni les explorer sans son consentement explicite, ni les divulguer pour des usages qui ne seraient pas strictement d'ordre médical et à finalité thérapeutique pour la personne considérée. Indépendamment des différences biologiques, culturelles, sociales ou religieuses qui distinguent les hommes, il y a en effet pour chacun un droit naturel à être ce qu'il est et à être le seul responsable de son patrimoine génétique.

5. Cependant, il ne faut pas se laisser fasciner par le mythe du progrès, comme si la possibilité de réaliser une recherche ou de mettre en œuvre une technique permettait de les qualifier immédiatement de moralement bonnes. La bonté morale de tout progrès se mesure au bien authentique qu'il procure à l'homme, considéré selon sa double dimension corporelle et spirituelle; ainsi, on rend justice à ce qu'est l'homme; en ne reliant pas le bien à l'homme, qui doit en être le bénéficiaire, il serait à craindre que l'humanité ne coure à sa perte. La communauté scientifique est sans cesse appelée à maintenir l'ordre des facteurs, en situant les aspects scientifiques dans le cadre d'un humanisme intégral; elle tiendra ainsi compte des questions métaphysiques, éthiques, sociales et juridiques qui se posent à la conscience et que les principes de la raison sont à même d'éclairer.

Dans le programme de votre présente session, je me réjouis que vous ayez eu le souci, comme hommes de science, de mettre vos connaissances au service de la vérité morale, en réfléchissant aux implications éthiques et aux adaptations législatives qu'il serait nécessaire de proposer aux gouvernements et aux équipes scientifiques. Il est souhaitable que votre voix autorisée contribue à l'élaboration d'un consensus international dans un domaine aussi délicat, consensus fondé sur la vérité objective de l'homme, appréhendée par la droite raison. A partir de là, il faut espérer que les institutions concernées s'attacheront à favoriser une réflexion approfondie, pour que chaque pays puisse se doter des réglementations qui protégeront la personne humaine et son patrimoine génétique, tout en stimulant la recherche fondamentale et la recherche appliquée à la santé des individus.

6. Ce n'est pas en raison d'une compétence scientifique particulière que le Magistère s'intéresse aux domaines qui font l'objet de vos recherches; l'existence même de l'Académie montre que l'Eglise respecte l'autonomie des disciplines scientifiques. De plus, «loin d'opposer les conquêtes du génie et du courage de l'homme à la puissance de Dieu, [...] les chrétiens sont au contraire bien persuadés que les victoires du genre humain sont un signe de la grandeur divine et une conséquence de son dessein ineffable» (Conc. Vat. II, *Gaudium et spes*, n. 34). L'Eglise n'intervient qu'en vertu de sa mission évangélique: elle a le devoir d'apporter à la raison humaine la lumière de la Révélation, de défendre l'homme et de veiller sur «sa dignité de personne dotée d'une âme spirituelle, de responsabilité morale, et appelée à la communion bienheureuse avec Dieu» (Congr. pour la Doctrine de la Foi, *Donum vitae*, n. 1).

Dès que l'homme est en cause, les problèmes dépassent le cadre de la science, qui ne peut rendre compte de la transcendance du sujet ni édicter les règles morales découlant de la place centrale et de la dignité primordiale du sujet dans l'univers. Dans cet esprit, l'existence de comités d'éthique est à encourager, pour aider la science à évaluer les aspects moraux des recherches et à en déterminer les conditions éthiques.

7. Parmi les autres thèmes que vous abordez, il y a celui des énergies de substitution pour les pays en voie de développement, thème dont on mesure l'intérêt pour l'avenir de l'humanité en cette

période où les questions liées à la démographie font l'objet de graves débats. Pour favoriser le dynamisme économique du monde, il est important de faire l'inventaire des solutions réalistes pour remplacer les ressources actuelles, qui risquent de s'épuiser un jour. Plus que toute autre, la génération présente a la responsabilité et le devoir de ne pas gaspiller inutilement ses richesses énergétiques. Les décisions en ce domaine doivent aussi tenir compte des générations futures. Les ressources énergétiques de notre planète sont des richesses qui doivent permettre à tous les peuples de se développer et d'avoir les moyens matériels d'une vie digne, en évitant de créer des déséquilibres économiques et écologiques. Ces ressources ne peuvent pas être utilisées par un petit nombre de pays au détriment des autres. La répartition des biens sur le sol de la planète est inégale. La solidarité et le partage sont indispensables pour créer des relations équitables entre les pays producteurs et les pays consommateurs.

8. A côté de la notion de «certitude mathématique», les recherches entreprises sur les «Principes fondamentaux en mathématique» ont conduit à reconsidérer la démarche épistémologique que les mathématiciens doivent suivre pour respecter les exigences propres à leur science, telles que la clarté, la cohérence, l'honnêteté intellectuelle et la confiance dans les capacités rationnelles de l'homme. En lien avec cette réflexion, a été forgé le concept clé d'«intelligence artificielle». Il convient cependant de rappeler que la machine reste un instrument au service de l'homme. Son «intelligence» est limitée, car il ne s'agit pas de la raison au sens plein du terme, raison qui permet à l'homme de se penser comme créature, d'appréhender le bien, le vrai et le beau, de diriger sa vie et de se mouvoir vers sa fin grâce à l'acte volontaire.

Vous évoquez à ce propos l'importance de l'étude des corrélations entre le cerveau humain et les systèmes électroniques dans le domaine des neurosciences, pour que la machine puisse suppléer un certain nombre de déficiences humaines et améliorer la qualité de la vie de personnes handicapées. C'est la grandeur de la science que d'être particulièrement au service de ceux de nos frères qui ont le plus besoin d'aide pour mener une existence conforme à leur nature et à leur incomparable dignité.

9. Alors que nous approchons du soixantième anniversaire de la refondation de cette illustre institution par Pie XI, on peut affirmer qu'elle remplit les fonctions qui avaient été assignées aux scientifiques: désignés en fonction de leur compétence, sans discrimination d'origine ou de religion, ils sont appelés à agir librement. Dans un souci de meilleure efficacité, vous avez mis à l'étude votre règlement interne, pour pouvoir remplir de manière plus adéquate la mission inscrite dans vos statuts: la participation aux progrès des sciences et l'approfondissement de la nature de la connaissance scientifique.

Au terme de notre rencontre, permettez-moi de vous remercier pour les contributions que vous apportez au Saint-Siège, sur des questions nouvelles et significatives qui requièrent des connaissances approfondies. Dans les immenses progrès du monde contemporain, il appartient à la communauté tout entière d'être particulièrement vigilante à promouvoir un humanisme intégral. C'est le sens même de l'homme qui est en cause. Je confie au Très-Haut vos efforts et vos recherches toujours ouverts aux exigences de cet humanisme.

ADDRESS OF THE PRESIDENT OF THE PONTIFICAL ACADEMY OF SCIENCES AT THE SOLEMN PONTIFICAL AUDIENCE

Holy Father,

Allow me to express my devoted gratitude for having granted this Audience to the Pontifical Academy of Sciences, on the occasion of its Plenary Session.

Seven new Academicians, recently appointed by Your Holiness, are participating in this meeting. Their appointment has enlarged the range of countries represented in the Academy and at the same time has enriched the diversity of scientific disciplines.

The Plenary Session of the Academy has been dedicated to arguments of great importance, namely, the Human Genome, the Alternative Energies for Developing Countries, the Fundamental Principles of Mathematics and Artificial Intelligence.

These are arguments of great relevance for the well-being of humanity and, at the same time, are the object of intense scientific investigation.

The proceedings of the Session have been helped by the generous participation of experts from many countries who took part in the meeting.

One of these themes, the project to investigate the structure of the Human Genome, has raised many questions, particularly among scientists, who realize how their own researches may become, on the one hand, of great utility to the improvement of the conditions of life and of men's health, but on the other hand, they may lead to serious and complex problems due to the implications (that) the new scientific discoveries have with ethical, social and juridical arguments.

On the occasion of the Plenary Session, the Academy has been involved in a profound discussion on its Internal Regulations to find the most fitting formulation which will allow it to face the challenge of the beginning of this third millennium.

In this way, the scientific Organism of the Holy See, which Your Venerable Predecessor Pius XI liked to call "his Scientific Senate", will be able to offer answers to the questions of mankind, and to establish a suitable link between the world of science and the Church. The Academy feels in this respect a great responsibility which is increased by the attention (that) the Holy See has always manifested for its work.

In front of the new scientific frontiers, it becomes more and more necessary for researchers to reflect with particular attention upon the deeper meaning of their discoveries and upon their wide implications concerning the future of human development.

Holy Father, we shall all be deeply grateful if you will kindly enlighten us with your teachings, which will allow us to find the way to be followed with renewed strength by scientists in the awareness of working for the good of the whole of mankind.

OPENING ADDRESS BY THE PRESIDENT OF THE PONTIFICAL ACADEMY OF SCIENCES

NICOLA CABIBBO

Dear colleagues and friends of the Academy who are here present in this meeting of the Academy: first of all, welcome to everybody.

I have a few announcements from the Academy. The first is that Professor Bernard Pullman has been nominated member of the Council of the Academy, and I am looking forward to working with him and with the other members of the Council. As you know, during the last three years we lost seven members of the Academy, two of them astronomers: William Morgan from the United States, and Jan Oort from the Netherlands. We lost Salimuzzaman Siddiqui, from Pakistan, a chemist, and four members who are scientists of life sciences: Professor Jérôme Lejeune, who was also a member of our Council, Severo Ochoa, Roger Sperry from the United States, and Janos Szentágothai from Hungary.

Today we will hear commemorations of these members, except for Professor Roger Sperry, whose death occurred only a few weeks ago, so we will commemorate him at our next meeting.

On this occasion, we have Mrs. Lejeune and Professor Lejeune's son with us. We will also hear brief testimonials from Professor Grisolia, who was a close friend of Professor Ochoa, and Professor Gulyas, who was a student of Professor Szentágothai.

We have seven new academicians: Professor Luis Caffarelli, Professor Luca Cavalli-Sforza, Alistair Crombie, Vladimir Keilis-Borok, Rudolf Muradian, Cyril Ponnamperuma, Andrej Szczeklik and Robert White.

Professor Keilis-Borok is not present today but he will come in two days because of problems with plane reservations.

Today, as you see from our programme, we will have a closed session where we will present a sort of draft regulation for the Academy's operation. This was prepared by the Council on the basis of many observations from the academicians. This can be discussed today; it will be open for discussion.

I recall that our Statute, the founding paper for the Academy, foresees that the regulations will be written by the President and the Council.

However, we would really like to do this only after having heard the wishes and advice of all the academicians. We will do this in the afternoon.

As for the programme of the plenary session today, as you can see, our decision for the session was essentially to concentrate on the presentation of the Academy's activities during the past two years.

As you know, the Academy only meets every two years in plenary session but in the intervening time we hold a series of activities, study weeks, special meetings, etc., and I think it is very interesting for the academicians to hear reports on these activities.

In particular, the central theme will be the human genome, scientific and legal aspects, and there will be a presentation of this theme but also the presentation of the results of a meeting that we held as Academy last year; there are ethical aspects here also.

Tomorrow we will have a discussion on alternative energy sources for developing countries. This theme was suggested by some academicians; in particular, I think Professor Blanc-Lapierre was the one who proposed this theme.

On Friday we will have a discussion on the fundamental principles of mathematics and artificial intelligence. This is a theme which was proposed by Professor Lejeune and Professor De Giorgi. Finally, we will have a discussion on the activities concerned with the analysis of the impact of science and technology on economic and social life. As you know, our Academy has been periodically involved in these problems; in particular, last year we held an important meeting on chemical hazards in developing countries, and this will be reported by Professor Ramel from the Royal Swedish Academy of Sciences who organized that meeting. There will be a report by Professor Rees on the epoch of galaxy formation, another recent meeting of the Academy, and finally on resources and population. This has attracted a lot of attention. I have received a lot of mail on this subject, and I think we should discuss it. On Friday, we have been granted an audience by the Holy Father, which will take place in the morning.

After the commemoration we will distribute a folder with the documents related to the meeting. I am sorry for not having distributed it before.

SCIENTIFIC PAPERS

on:

HUMAN GENOME

INTRODUCTION

BERNARD PULLMAN *

Monsieur le Président, Monseigneur, Mesdames et Messieurs,

Dans l'histoire de la Science et peut-être même de l'Humanité, l'un des événements les plus importants de la fin du deuxième et du commencement du troisième millénaire risque d'être l'immense entreprise de décryptage du génome humain qui, selon toute probabilité, sera réalisée pendant cette période.

La portée pratique, mais aussi symbolique, de cet événement est très grande. L'homme va pouvoir connaître les détails des structures qui sont à la base des processus héréditaires déterminant son essence et sa spécificité. Or, la connaissance, c'est le pouvoir. En effet, il ne fait guère de doute que ce décryptage lui donnera une possibilité très large d'agir, s'il le désire — *et il le désirera* — sur ces structures. Il est alors évident que cette capacité nouvelle ouverte à l'homme pose des problèmes éthiques, juridiques et sociaux inédits et importants. Leur examen est le but de la longue session que le Conseil de l'Académie Pontificale des Sciences a estimé approprié de consacrer à ce sujet à l'occasion de sa Réunion Plénière de cette année.

Avant de donner la parole aux conférenciers qui ont aimablement accepté de nous présenter les différents aspects que je viens d'énoncer du problème du génome humain, je voudrais seulement rappeler que l'organisation de la réunion d'aujourd'hui a une histoire. Au point de départ se trouve un Colloque International sur, justement, les aspects éthiques et juridiques liés au projet du génome humain, qui s'est tenu à Bilbao en Espagne, du 24 au 26 Mars 1993, *sous le patronage de la Fondation BBV*, et auquel ont participé une centaine de savants éminents venus du monde entier. A l'issue de ce Colloque, la Fondation BBV et

---

* Académicien Pontifical.

en particulier les collègues espagnols qui ont joué un rôle actif dans l'organisation et la tenue de cette réunion, conscients de l'intérêt que l'Académie Pontificale des Sciences et les hautes autorités de l'Eglise portent à ce sujet, ont fait connaître leur désir de communiquer à notre Compagnie les conclusions essentielles des débats qui se sont tenus à Bilbao. A la suite de cette démarche, j'ai eu l'honneur d'être chargé par le Président et le Chancelier de l'Académie Pontificale d'organiser une rencontre répondant à cette demande. Celle-ci a eu lieu sous forme d'un Groupe de Travail qui s'est réuni au siège de notre Académie les 19 et 20 Novembre 1993 et auquel ont participé, outre notre Président et notre Chancelier, les Professeurs Cottier, Lejeune, Sela et moi-même représentant l'Académie, les Professeurs Mendizabal Allende, Pardo Garcia, Romeo Casabona, Grisolia Garcia, Villar Palasi et Gerin en tant qu'organisateurs de la réunion de Bilbao, et quelques personnalités invitées parmi lesquelles, en particulier, Monseigneur Carrasco de Paula, Recteur de L'Ateneo Romano della Santa Croce et Monseigneur Elio Sgreccia, Secrétaire du Conseil Pontifical pour la Famille. Les débats animés et approfondis qui ont eu lieu à cette réunion préliminaire ont servi de base pour la structuration de notre session actuelle.

Je suis particulièrement heureux de revoir ici, aujourd'hui, nos distingués collègues espagnols, les Professeurs Mendizabal, Grisolia et Casabona, et d'adresser par leur intermédiaire nos très vifs remerciements à la Fondation BBV et tout particulièrement à son Président, le Professeur José Angel Sanchez Asiain, pour l'aide précieuse, aussi bien intellectuelle que matérielle, qu'ils nous ont apportée, dans la préparation et l'organisation de ces importantes réunions. L'implication de la Fondation BBV dans nos travaux témoigne de l'importance que cette Fondation attache au thème de nos délibérations et de sa disponibilité pour venir en aide à ceux qui recherchent la «bonne» issue aux problèmes posés.

Je ne veux pas retarder plus longtemps la présentation du sujet par les experts désignés et le débat auquel vous êtes tous conviés à prendre part. J'ai eu l'occasion d'exposer, dans mon Introduction à la réunion du Groupe de Travail de Novembre 1993, mes vues sur quelques aspects du problème, entre autres sur la question du «droit» de l'homme à manipuler le génome, question fréquemment posée sous la forme: «Le génome est-il sacré?». Pensant que ces remarques sont susceptibles d'intéresser certains d'entre vous, je me suis permis de vous faire distribuer une copie de cette Introduction. Ici, je voudrais seulement citer quelques phrases terminales de ce document. Je me cite:

«Le projet du génome humain est en marche et il ne fait aucun doute qu'il sera mené à son terme. C'est là une perspective incontournable, qu'on s'en réjouisse ou qu'on le déplore Pour ceux qui croient que le génome est sacré, en fait, aussi pour ceux qui ne le croient pas, il doit être essentiel de veiller à ce que cette grande aventure humaine soit menée pour le bien des hommes, biens matériels, certainement, comme, par exemple, la prévention ou la guérison des maladies héréditaires — qui suppose de toute évidence une intervention humaine sur le génome — mais surtout ses biens spirituels dont dépend leur dignité d'homme, donc leur avenir. Après tout, dans le génome, de toute manière, ce ne sont pas les molécules qui sont sacrées. Ce qui est sacré, c'est le message qu'il porte, ce sont certaines valeurs qu'il représente et auxquelles nous tenons et parmi lesquelles je placerai en premier lieu le respect de la dignité humaine. Jamais, peut-être, l'homme ne s'est trouvé placé d'une façon aussi dramatique, par le fait du progrès scientifique, face à une vision prometteuse d'une part, mais apocalyptique de l'autre, de son avenir. C'est pourquoi les discussions que nous allons ouvrir maintenant, ici, sur les aspects éthiques, juridiques et sociaux liés au projet du génome humain et qui serviront de base, je crois, aux discussions sur ce sujet à la prochaine réunion plénière de l'Académie Pontificale, prévue pour Octobre 1994 — nous y sommes — ont tant d'importance».

Monsieur le Président, Monseigneur, Mesdames et Messieurs,

La réunion *d'aujourd'hui* comporte, comme vous le savez, trois parties, devant traiter, respectivement, des aspects scientifiques, juridiques et éthiques, du projet du génome humain. La matinée de *demain* sera consacrée à une discussion générale du projet et des problèmes qu'il soulève. J'ouvre maintenant la session consacrée aux aspects scientifiques du projet.

I.

SCIENTIFIC ASPECTS

NUCLEIC ACIDS STRUCTURES AND THE CONTROL OF GENE EXPRESSION

CLAUDE HÉLÈNE *

Introduction

The *diversity* of living organisms on our planet hides a remarkable *unity*. The basic unit is the living cell. Microorganisms are made of a single cell whereas higher organisms may contain thousands of billions of cells. A living cell in eukaryotes consists of an external envelope (a membrane made of a phospholipid bilayer), a large number of organelles within the cytoplasm (e.g., mitochondria specialized in energy production, the endoplasmic reticulum where proteins are assembled ...) and a nucleus which contains the genetic information. In (unicellular) prokaryotes, there is no separation between the genetic information and the rest of the cell (no nucleus) in contrast to eukaryotes where a nuclear membrane is present. A plant cell differs from, e.g., a human cell, by the presence of a cell wall which makes the cell more rigid, a specialized organelle called a chloroplast where photosynthesis takes place and a large vacuole which captures a large part of the cell products.

The genetic information of all living organisms is contained within long macromolecules of deoxyribonucleic acid (DNA) constituted by a double helix of two polynucleotide chains. This information is first transcribed into ribonucleic acids (RNA). Each unit of information, called a *gene*, gives rise to one RNA transcript which is then processed to give a mature *messenger* RNA. The latter is translated by ribosomes into a chain of amino acids, called a polypeptide or a protein. The polypeptide chain folds on itself to confer upon the protein a specific function (enzymes, cell

* Membre de l'Académie des Sciences. Laboratoire de Biophysique, Muséum National d'Histoire Naturelle.

surface receptors, receptor ligands, hormones, proteins maintaining the architecture of cells, ...).

GENE ORGANIZATION

A gene is a portion of cellular DNA which contains coding sequences (exons) separated by non-coding intervening sequences (introns), and a control region (promoter) located upstream of the coding sequences (Figure 1). The control region dictates whether the cell should synthesize the messenger RNA coded for by a specific gene. Some genes are transcribed in all tissues; others have a much more limited expression, some of them being expressed in a single specific tissue (Figure 2).

In prokaryotes, where the DNA is not isolated in a specific

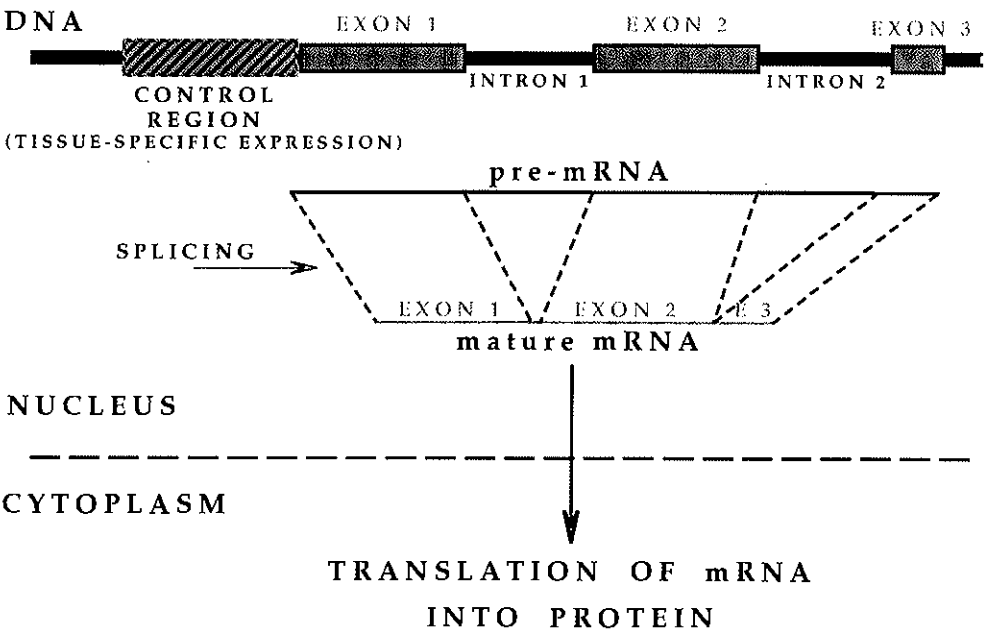


FIG. 1. *Gene organization.* A gene is a double-helical DNA fragment that contains a control region, several coding sequences (exons) and intervening sequences (introns). After transcription of one of the DNA strands into a pre-messenger RNA, the RNA introns are eliminated by a process called *splicing* which brings the coding sequences (exons) in continuity. A mature RNA is obtained after splicing and other post-transcriptional modifications (capping, addition of a polyA tail ...). After maturation tbe messenger RNA moves to cytoplasm (through the nuclear membrane) where it is translated into a chain of amino acids, a protein.

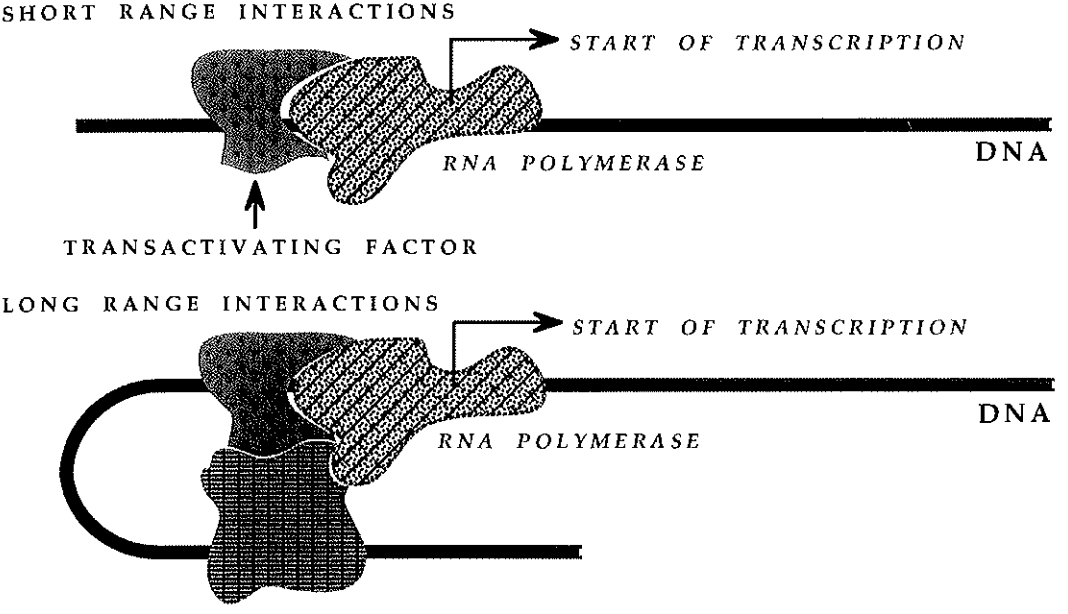


FIG. 2. *Control of gene expression.* The control of gene expression at the level of transcription involves proteins (called transcription factors) that bind upstream of and establish contacts with RNA polymerase. Short-range and long-range interactions have been described. The binding of transcription factors is sequence-specific and determines the tissue specificity of gene expression.

compartment, transcription and translation take place in a coupled process. In eukaryotes, the DNA is located within the nucleus of cells. Transcription of one strand of the double helix gives rise to a pre-messenger RNA which is processed within the nucleus: non-coding (intronic) sequences are removed in a process called splicing, a poly A tail and a cap are attached. Then the mature messenger RNA leaves the nucleus and reaches the cytoplasm where the translation machineries (the ribosomes) are located. Therefore translation is decoupled from transcription in eukaryotes (Eick *et al.*, 1994).

SEQUENCING THE HUMAN GENOME

The human genome consists of 23 pairs of DNA molecules (23 are provided by the mother, 23 by the father, during fertilization of an ovum by a spermatozoid). The DNA macromolecules are associated with proteins, especially histones, which allow the long filament of the double helix to adopt a compact structure, called a chromosome. The total DNA of the 46 chromosomes represents a chain of three billion units, called nucleotides.

There are only four different nucleotides which differ by the nature of the nucleic acid bases: adenine (A), thymine (T), cytosine (C) and guanine (G). The two strands of the DNA double helix are complementary: A on one chain is always associated to T on the other one, G on one chain to C on the other one. These A.T and G.C base pairs are held together by hydrogen bonds. The base pairs open during replication, the enzymatic process which allows a cell to duplicate its genetic information before cell division. The ultimate goal of the *human genome project* is to decipher the order of the nucleotide units in the two chains of the DNA double helix. The sequence of the nucleotides should give all the information required to determine the sequence of the amino acids of all proteins involved in the various functions of a living organism.

The genetic code allows us to establish the sequence of amino acids in a protein from the nucleotide sequence of a messenger RNA. An intermediate step in the human genome project is to sequence the messenger RNAs which represent only a small fraction of the DNA sequence but contain the information for protein sequences. This can be achieved by reverse transcribing mRNAs into cDNAs which are then amenable to "classical" DNA sequencing procedures (Mural, 1994). By sequencing cDNAs from different organs, one obtains information on the tissue specificity of gene expression. However, the control regions located upstream of the genes are not obtained from cDNA sequence information. The procedures developed to tag long DNA molecules at specific sites and the technologies presently available to rapidly sequence long DNA fragments are described in greater detail in Dr. Jean Weissenbach's presentation.

It should be noted that there is not *a* human genome. Except for "true" twins, there are variations in the sequence of the genome from one individual to another. Some of these differences have no consequence on the sequence of the protein products because there is some degeneracy in the genetic code [there are 64 "words" of three nucleotides each in the nucleic acid language but only 20 "words" (aminoacids) in the protein language]. Other differences introduce minor variations in protein functions. But some of them may have a dramatic influence on human life. Gene mutations or alterations have been characterized that are responsible for genetic diseases or for predisposition to certain types of pathological disorders. In addition the human genome is a dynamic system with possible reorganizations during development and cell division (see the presentation by Prof. Maxine Singer).

TRANSGENIC ORGANISMS

The identification, isolation and sequencing of genes has led to new areas of research dealing with gene transfer and expression in foreign organisms. It has become possible to engineer DNA vectors that can express genes in a controlled manner. This "molecular surgery" makes use of restriction enzymes as "molecular scissors" and circular DNA (plasmids) or viral DNA (or RNA) as vectors.

Transgenic microorganisms or cells in culture have been used to produce proteins in large quantities. These proteins may then be utilized: i) in fundamental studies on protein structure and function (e.g., X-ray crystallography, nuclear magnetic resonance and other spectroscopic studies ...); ii) for their catalytic activity in industrial applications, and iii) for administration to human beings to cure pathological disorders (e.g., interferons, tissue plasminogen activator, erythropoietin ...).

Transgenic animals can be used: i) to analyze the role played by a particular gene during development (Gossen and Vijg, 1993); ii) to create animal models for human pathological disorders (e.g., in cancer and atherosclerosis research or in degenerative diseases of the central nervous systems ...) (Smithies, 1993); iii) to produce specific proteins for therapeutical applications.

Transgenic plants have been created: i) to get a better knowledge of specific genes during plant development (Benfey and Chua, 1989; Tabler, 1993); ii) to improve the agronomical traits of plants important in food production; iii) to produce foreign proteins in large amounts; iv) to make plants resistant to herbicide activity, to insect attack or to viral and fungal infections (Bejarano and Lichtenstein, 1992; Tepfer, 1993).

In transgenic animals and plants, a gene can be introduced under the control of a tissue-specific promoter so that gene expression is limited to a particular organ or tissue. Alternatively, a gene can be knocked out by homologous recombination so that its effects on embryonic or adult development can be analyzed.

The development of transgenic species concerns important areas of activities: fundamental research, industrial applications, food production, human health. This is made possible through a better knowledge of gene organization and sequencing, control of gene expression, identification of genes involved in pathological disorders ...

GENE-SPECIFIC THERAPEUTICAL APPROACHES

One of the objectives of sequencing the human genome is to get access to the genes which are involved in a large number of genetic diseases. Once these genes are identified, they provide an invaluable information to develop diagnostic tools and to design therapeutical interventions. The gene product (a protein) can be chosen as a target for new drug research. Alternatively the gene rather than its product can be chosen either as a target for new drugs (Hélène and Toulmé, 1990; Hélène, 1991; Hélène and Saison-Behmoaras, 1994, Thuong and Hélène, 1993), or as a drug itself (Kahn, 1992; Culver and Blaese, 1994; Gilboa and Smith, 1994; Friedman, 1994, Mark and Woo, 1994). For example, if a disease is associated with deficiency of a single gene, gene therapy protocols can be designed to replace the missing function. If a disease arises from a mutated gene which gives rise to a protein with an altered function, strategies can be developed to block the expression of this particular gene.

Small nucleic acid fragments, called oligonucleotides, can be used to control gene expression in a highly sequence-specific way. In the "antisense" strategy, an oligonucleotide complementary (in the Watson-Crick sense) to a messenger RNA can be used to block translation of the message into a protein (Hélène and Toulmé, 1990; Hélène and Saison-Behmoaras, 1994). In the "antigene" strategy, the oligonucleotide can be targeted to the gene itself in order to block the transcription of a single gene within the living cell (Hélène, 1991; Thuong and Hélène, 1993) (Figure 3).

Gene therapy approaches require DNA vectors to transfer and express a specific gene. Viral vectors (such as adenovirus, adeno-associated virus, herpes virus, retroviruses ...) or non viral vectors (polymers, nanoparticles ...) can be used. The DNA vector can be administered directly to human beings, e.g., injections into tumors or nasal spray for adenovirus carrying the CFTR gene to cure cystic fibrosis. Alternatively, cells from a patient (tumor cells, hematopoietic cells, ...) can be treated *in vitro* with the DNA vector and then re-infused to the patient. Both gene therapy and cell therapy protocols are undergoing clinical trials in pathological disorders such as cancer and cystic fibrosis.

The pathology of cancer provides examples of the role that gene identification and sequencing is playing — and will play — in the development of new therapeutical strategies. The transformation of normal cells into malignant cells is a multi-step process which involves the activation of cellular proto-oncogenes into oncogenes and the inactivation of tumor-suppressor genes (Volgenstein and Kinzler, 1993) or of DNA repair genes (Jirigny, 1994). Oncogenes can be chosen as targets for inhibitors of their function either at the gene (Hélène, 1991; Hélène and Saison-Behmoaras, 1994) or at the protein levels (McCormick and Myers, 1994). Alternatively, replacing tumor suppressor or DNA repair genes via gene therapy protocols might lead to an arrest of cell proliferation and

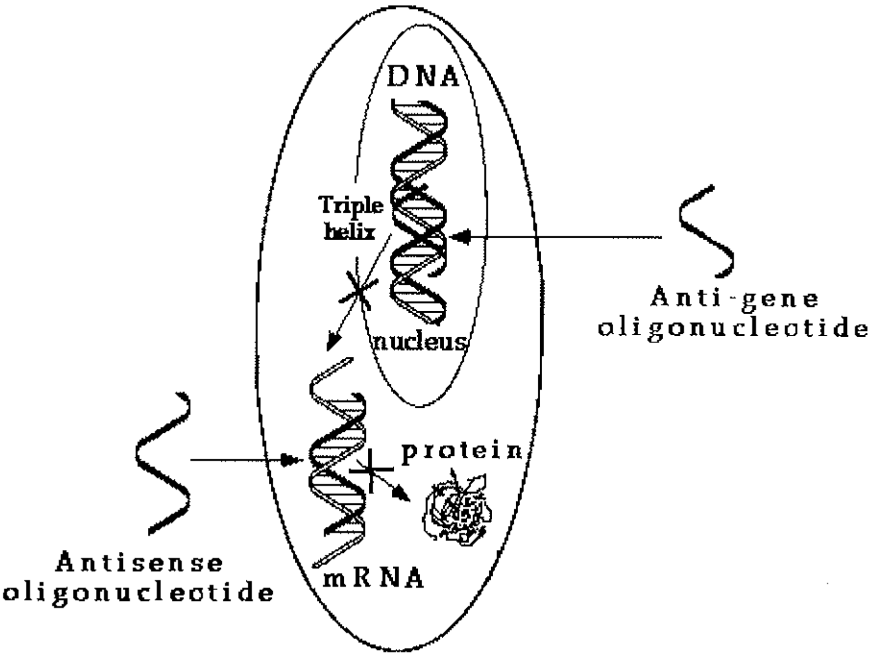


FIG. 3. *Strategies to inhibit gene expression with synthetic oligonucleotides.* An *oligonucleotide* is a short, single-stranded fragment of nucleic acid (12-20 nucleotides) which can bind to a messenger RNA (*antisense* or *ribozyme*, lower left of the figure) or to a gene in the nucleus (*antigene*, top part). Gene expression is blocked at the translational or transcriptional level, respectively. Ribozymes are a special class of oligoribonucleotides that can induce sequence-specific cleavage of a mRNA. The antisense, ribozyme or antigene oligonucleotide sequence can be chosen in such a way as to block expression of a single gene within a living cell.

tumor growth (Culver and Blaese, 1994). It has been observed that expression of a histocompatibility gene (HLA-B7) (Nabel *et al.*, 1993) or that of an antisense against the growth factor IGF-1 (Trojan *et al.*, 1993), could induce a cellular immune response which attacks and destroys the original tumor.

Research in gene mapping and sequencing is providing advances in medical and biological knowledge on which new therapeutic strategies can be based. It also permits the development of accurate tests to screen for predisposition to certain types of diseases. This also raises numerous ethical and practical questions that must be addressed by our society from both a scientific and a social point of view.

REFERENCES

Bejarano, E.R., and Lichtenstein, C.P., 1992, Prospects for engineering virus resistance in plants with antisense RNA, *Trends in Biotechnology*, **10**, 383-387.

Benfey, P.N., and Chua, N.H., 1989, Regulated genes in transgenic plants, *Science*, **244**, 174-181.

Culver, K.W., and Blaese, R.M., 1994, Gene therapy for cancer, *Trends in Genetics*, **10**, 174-178.

Eick, D., Wedeland, A., and Heumann, H., 1994, From initiation to elongation: comparison of transcription by prokaryotic and eukaryotic RNA polymerases, *Trends in Genetics*, **10**, 292-296.

Friedmann, T., 1994, Gene therapy for neurological disorders, *Trends in Genetics*, **10**, 210-214.

Gilboa, E., and Smith, C., 1994, Gene therapy for infectious diseases: tbe AIDS model, *Trends in Genetics*, **10**, 139-144.

Gossen, J., and Vijg, J., 1993, Transgenic mice as model systems for studying gene mutations *in vivo*, *Trends in Genetics*, 9, 27-31.

Hélène, C., and Saison-Behmoaras, E., 1994, La stratégie antisens: nouvelles approches thérapeutiques, *Médecine/Sciences*, **10**, 253-273.

Hélène, C., 1991, Rational design of sequence-specific oncogene inhibitors based on antisense and antigene oligonucleotides, *European Journal of Cancer*, **27**, 1466-1471.

Hélène, C., and Toulmé, J.J., 1990, Specific regulation of gene expression by antisense, sense and antigene nucleic acids, *Biochimical Biophysical Acta*, **1049**, 99-125.

Jirigny, J., 1994, Colon cancer and DNA repair: have mismatches met their match, *Trends in Genetics*, **10**, 164-168.

Kahn, A., 1992, Tbérapie génique, *Médecine/Sciences*, **8**, 3-33.

Mark, A.K., and Woo, S.L.C., 1994, Gene therapy for metabolic disorders, *Trends in Genetics*, **10**, 153-257.

McCormick, F., and Myers, P., 1994, From genetics to cbemistry: tumor suppressor genes and drug discovery, *Chemistry and Biology*, **1**, 7-9.

Mural, R.J., 1994, Finding one's way through DNA, *Trends in Biotechnology*, **12**, 393-394.

Nabel, G.J., Nabel, E.G., Yang, Z.Y., Fox, B.A., Plautz, G.E., Gao, X., Huang, L., Shu, S., Gordon, D., and Chang, A.E., Direct gene transfer with DNA-liposome complexes in melanoma: expression, biologic activity, and lack of toxicity in humans, *Proc. Natl. Acad. Sci. USA*, **90**, 11307-11311.

Smithies, O., 1993, Animal models of buman genetic diseases, *Trends in Genetics*, 9, 112-116.

Tabler, M., 1993, Antisense RNA in plants: a tool for analysis and suppression of gene function, in *Morphogenesis in Plants*, Roubelakis-Angelakis and Tran Tranh Van (Eds), Plenum Press, New York, pp. 237-258.

Tepfer, M., 1993, Viral genes and transgenic plants, *Bio/Technology*, **11**, 1125-1132.

Tbuong, N.T., and Hélène, C., 1993, Sequence-specific recognition and modification of double-heoical DNA, *Angewandte Chemie, International Edition in English*, **32**, 666-690.

Trojan, J., Johnson, T.R., Rudin, S.D., Ilan, J., Tykocinski, M.L., and Ilan, J., 1993, Treatment and prevention of rat glioblastoma by immunogenic C6 cells expressing antisense insulin-like growth factor I RNA, *Science*, **259**, 94-97.

Volgenstein, B., and Kinzler, K., 1993, The multistep nature of cancer, *Trends in Genetics*, 9, 138-141.

THE HUMAN GENOME PROJECT

JEAN WEISSENBACH *

THE HUMAN GENOME

All of the biological functions which are necessary for the maintenance of a species in its environment are determined by the expression of a hereditary program which is stored in the chemical form of nucleic acids and designated by the term genome. The acquisition of the information contained in this program is therefore fundamental for biologists.

Using the techniques which are presently available, this acquisition is possible at present for genomes of modest size –100 megabases at the most. However, this project requires a mobilization of resources without precedent in biology for genomes in the 100 megabase range and is not yet really possible for large genomes such as that of human beings. This difficulty is partly due to the splitting of the genes of higher organisms into introns and exons, and to the accumulation of a fairly large number of superfluous, repetitive sequences in the genome by diverse biochemical mechanisms. These repeated sequences may number several hundred thousands of copies per haploid genome. The useful part of the genome (that part which has a functional role) is thus much smaller than the total size of the genome.

In humans, for example, the genes, which are estimated to number between 50,000 and 100,000, only represent 5% of the totality of the genome (3 billion base pairs). The rest of the DNA consists mainly of useless sequences and a small proportion of sequences for which the functions are poorly or not at all understood. There are also sequences which are involved in maintaining the integrity of chromosomes and ensuring their proper separation during cell division. The number of known

---

* GENETHON, France.

genes is still very small, although it is increasing rapidly (between 3,000 and 4,000 in 1994).

With the computer programs which are currently available, it is possible to recognize the majority of exons present in a human DNA sequence and therefore to deduct the sequence of the protein (Uberbacher and Mural, 1991). On the other hand, the function of a protein can only rarely be predicted from its sequence. Therefore, it is necessary to resort to biological experimentation in order to understand the function of each gene. Even when the sequence of the genome will be known in its totality and the complete catalog of the genes will be available, a considerable amount of time will still be required to understand the functions that the genes encode.

GENESIS OF THE GENOME PROGRAM

In the middle of the 1980s, progress in the methods of molecular biology, notably in sequencing techniques, led a number of biologists to propose an ambitious program with the goal of determining the complete sequence of the human genome. This proposition provoked contradictory reactions in the scientific community and led to passionate debates. Commissions of experts composed of partisans and adversaries of complete sequencing held meetings in the United States. Surprisingly, a consensus was reached in 1988, which resulted in a convergence of the objectives of fundamental biology which was aiming to establish an exhaustive catalog of the biological functions of living organisms and those of medical genetics which was attempting to discover the causes of genetic diseases (Committee on Mapping and Sequencing the Human Genome and Council, 1988). In fact, knowledge of the complete sequence of the genome should make it possible to produce a catalog of human genes and those of other species, and to find the genes responsible for genetic diseases much more rapidly.

This program can be separated schematically into two phases. The first involves the mapping of large genomes (humans and a few model organisms) and the sequencing of the genomes of a few model organisms such as *E. coli*, yeast, and a small nematode. In parallel with this, more efficient sequencing techniques will be developed which will make it possible to carry out the sequencing of large genomes in the second phase of the program, in a reasonable time period and at lower cost (Committee on Mapping and Sequencing the Human Genome and Council, 1988; Watson, 1990).

Medical benefits

Whereas the mapping phase is only a technical step for the numerous genomes of model organisms, it takes on special importance in humans on account of its medical implications. The genes for numerous genetic diseases can be found much more rapidly by using good genome maps than by waiting for the complete sequence to become available. Furthermore, even when knowledge of the complete sequence is available, we will still have to use the genetic mapping approach to localize the genes for genetic disorders.

The urgent nature of this program is not only due to the severity of certain genetic diseases, but also to the fact that the field of human genetics has been in stagnation for the major part of this century. The impossibility of performing experiments in humans for obvious ethical reasons has restricted this discipline to the realm of pure observation, limited by the poverty of analytical methods available. After laborious beginnings lasting for many decades but occasionally punctuated by a few essential technical innovations, the field of human genetics has experienced unprecedented growth since the beginning of the 1980s, due to the advent of recombinant DNA technology. The ability to isolate, amplify, and identify specific fragments of DNA, and the techniques of *in vitro* recombination have greatly increased our capacity for observation of genomes and in particular, the human genome. These developments have been especially accelerated by the appearance of the polymerase chain reaction (PCR) for genome analysis (Saiki *et al.*, 1985). These successive technological revolutions have led to considerable change in the strategies utilized for the study of genomes. The ability to analyze our genes has led to an acute resurgence of the old problem of eugenism and has created new questions for society.

Although the human genome is not a priority for basic biology because the human species does not lend itself to experimentation, it is, on the other hand, of fundamental importance for medicine. Human genome analysis is most importantly intended as a tool for the identification of genes responsible for genetic diseases, which represents the driving force of this project. Several thousand genetic diseases due to a defect in a single gene are known today (McKusick, 1992). The functional defects, however, remain unknown in the majority of cases. An understanding of the defects which are responsible for these disorders is the first step in a rational approach to their treatment.

A strategy for the identification of the genes responsible for genetic diseases was elaborated during the 1980s. This process, which is known as positional cloning, is experimentally very arduous, and may take 5 to 10 years for a single disorder. The task to be accomplished for the numerous diseases that remain to be studied is therefore enormous and is ample justification for a program whose aim is to explore the human genome in a systematic manner. This systematic research requires an extensive infrastructure. It should, however, entail a reduction in cost and an increase in the efficiency of the discovery of disease genes. The creation of this infrastructure is in fact one of the first outcomes of the genome project.

Objectives and implementation of the project

Although several countries are contributing to the project, the United States has been the driving force behind it and is the only country which has elaborated a program commensurate with the scientific challenge that this enterprise represents. The project, as it was initially conceived in the United States, sets the following goals (U.S. Congress, 1988; U.S. Department of Health, 1990):

— mapping and sequencing of the genomes of evolutionarily simple organisms as models for study and analysis;
— development of procedures for the storage and diffusion of information in computerized databases;
— development and transfer of new technologies;
— mapping and sequencing of the human genome;
— a study of the ethical, legal, and social consequences that might result from the acquisition and utilization of this knowledge.

This program was to be achieved over a period of 15 years, with an annual budget of 200 million dollars. Extensive international cooperation was anticipated, as well as the participation of private industry. Specialized centres for the study of the genome were to be created, and supervision was to be shared by two government agencies: the Department of Energy (DOE) (3-5 centres), and the National Institutes of Health (NIH) (10 to 20 centres). After acceptance and financing by the American Congress, the program officially began on October 1, 1990, for an initial period of 5 years. Because of the early results and new technological advances, the 5-year goals have been revised for the period from 1993 to 1988 (Collins and Galas, 1993).

Other programs for the study of the genome have begun recently in the United Kingdom, Japan, Russia, Italy, France, and Germany. There is no real coordination between these programs, but unnecessary competition and redundancy are avoided to a certain extent. Other public or private initiatives which had commenced prior to the launching of the project, such as those of CEPH (Centre d'Etude du Polymorphisme Humain) and GDB (Genome Data Base), have the vocation of making a certain number of mapping tools available to the ensemble of the scientific community. CEPH distributes DNA samples used as a reference for the establishment of the human genetic map to laboratories all over the world, and maintains a database of the results obtained. GDB collects and coordinates mapping data which has been published all over the world. HUGO (the Human Genome Organization) is a society of scientists which includes the principal actors in the field of genomic research. HUGO organizes and coordinates scientific meetings to summarize current progress, especially concerning the map, as well as seminars for reflection on the ethical and social problems which may result from advances in human genetics.

Mapping of the genome

It has already become clear that the availability of detailed maps of the human genome already constitutes a remarkable tool for positional cloning strategies. Different kinds of maps are now being developed, which correspond to different requirements.

Historically, the most well-known of these is the genetic map. Its principle was established in the 1920's by T.H. Morgan's group who were studying the transmission of hereditary characteristics in the fruit fly. They were able to establish a parallel between this genetic map in the fruit fly, and the observation of its chromosomes in the microscope. Chromosomal abnormalities appeared in the same order as the mutations on the genetic map and therefore this representation of the microscopic appearance of the chromosomes was called a cytogenetic map. The techniques for cytogenetic observations have improved considerably since those first experiments. Later, other physical maps on a molecular scale began to appear with the development of techniques for manipulation of DNA (Nathans, 1979). They display various degrees of resolution. At the present time, considerable efforts are being expended to construct a functional map on which genes are placed on their structural support.

Genetic Mapping

In human beings it is rare to find a genetic disease associated with a chromosome abnormality that is visible in the microscope. This type of observation, however, enables scientists to obtain a first clue to the localization of the gene responsible for the disease. In all other cases, disease gene localization begins with a genetic mapping phase based on analysis of the transmission of the disease in affected families. Genetic diseases are transmitted according to Mendel's laws. This is a direct consequence of the way chromosomes are transmitted during reproduction. The chromosomes that we inherit from our parents are composed of regions coming from either one or the other of our grandparents. Since we can distinguish between the chromosomes which come from each of our parents, it is possible to follow the transmission of chromosomes from one generation to the next, and to define the parental origin of each chromosomal region.

By following the transmission of chromosomes from one generation to the next in a large number of families affected by the same genetic disorder, it is possible to identify a chromosomal region which is common to all individuals who have the disease (Fig. 1). In order to effect such an analysis of transmission, a genetic map is needed for reference; it should consist of reference points, or markers, which enable the investigator to distinguish between the copies of chromosomes from each parent and, if possible, each grandparent.

More or less detailed genetic maps have been developed for numerous animal and plant species which are useful as experimental models or are of economic importance. In humans, however, analysis of genetic transmission of hereditary characteristics has only been possible since the beginning of the 1980s, when the first polymorphic DNA markers were described (Botstein *et al.*, 1980). These polymorphismes serve as markers, which make it possible to determine the parental origin of each chromosomal region. Since then, a large number of these DNA polymorphisms have been placed on the genetic reference map. This reference map is being produced by a group of international collaborators who analyze the DNA from a limited number of large families assembled by CEPH (see above), which assembles the results obtained in the form of a database which is accessible for international collaboration (Dausset *et al.*, 1990).

Until 1992, the human genetic map produced by the CEPH collaborators suffered from several major drawbacks: (1) Most of the polymorphic markers were biallelic systems which were frequently present in the homozygous state and hence not informative regarding their transmission in the majority of the families studied (Donis-Keller *et al.*, 1987). (2) The distribution of markers along the chromosome maps was still very irregular, and some regions were practically devoid of polymorphic markers.

Because of the low informativity of the biallelic markers, a second generation genetic map consisting of multiallelic markers was produced in 1992 (Weissenbach *et al.*, 1992) and an updated version in 1994 (Gyapay *et al.*, 1994). These multiallelic markers are short tandem repeats (10 to 30

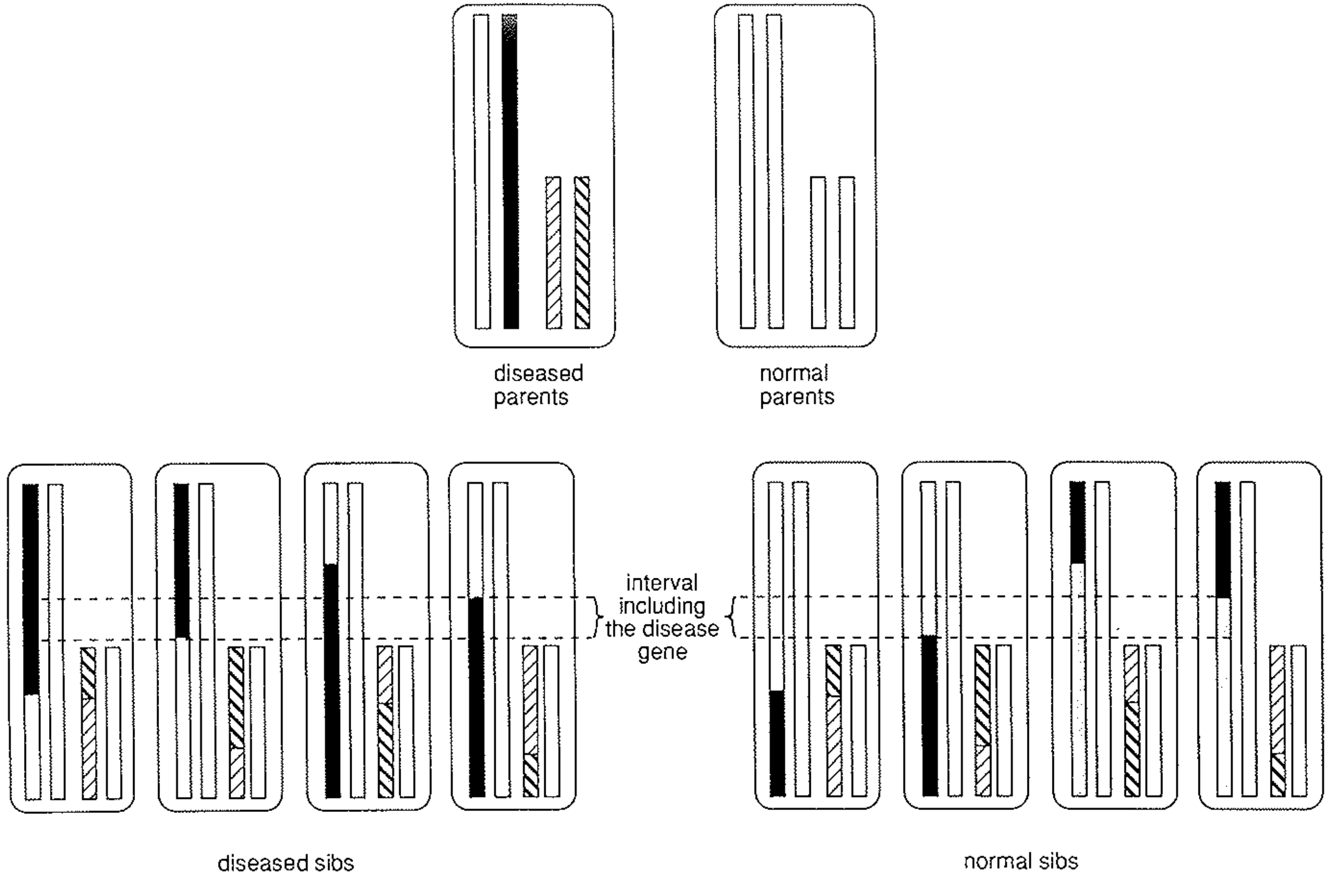


FIG. 1. Family analysis of transmission of an autosomal dominant disorder. The genome of members of affected families is represented by two pairs of chromosomes. The disease gene is carried on one of the large chromosomes (the black chromosome) of the affected parent whereas the small chromosome pair is normal. Each chromosome of the affected parent can be distinguished by its shading pattern; the white chromosomes of the unaffected parent need not be taken into consideration. Each parent transmits one chromosome of each pair to the children. The children's chromosomes are composites of fragments of parental chromosomes resulting from one or more recombinations between homologous parental chromosomes which occur during meiosis in the reproductive cells of each parent. Analysis of these composite chromosomes in healthy and affected siblings permits the identification of a region (dotted lines) which is shared by all affected individuals but absent in their healthy siblings. The other regions of the large black chromosome, the large grey chromosome, and the small chromosomes are not present in all the affected siblings, and are therefore not associated with the disease.

repetitions) of very simple nucleotide sequences (di-, tri-, or tetra-nucleotides) known as microsatellites or STRs (simple or short tandem repeats): the allelic variations correspond to changes in the number of times the motif is repeated.

On account of the intensive nature of the experiments that have to be done, the most complete and precise genetic maps are produced within the context of large mapping projects such as Généthon, a laboratory supported by the AFM (French Muscular Dystrophy Association) in France or the CHLC (Cooperative Human Linkage Center) in the United States (Murray, 1994). A map of over 5,000 highly informative markers will be produced by Généthon in 1995.

Physical Mapping

Physical mapping attempts to describe the physical support of the hereditary material (the chromosomes) on a scale which depends on the method of observation used. The first physical maps merely described the chromosomes as they appeared microscopically (Fig. 2). This permitted observation of anomalies in the number of chromosomes (such as trisomy 21 associated with Down's syndrome) or anomalies in the structure of chromosomes such as deletions, insertions, inversions, and translocations. These structural abnormalities are frequently associated with diseases. The resolution of this cytogenetic map is relatively weak, however; the chromosome bands which can be observed using the light microscope by routine methods (on metaphase chromosomes) are of the order of several million base pairs. Other techniques can be used for observation of chromosomes at mitotic stages in which they are more elongated (prometaphase). The techniques of microscopic observation have recently been revolutionized due to the observation of elongated fibers of chromatids using fluorescent DNA probes. These techniques for stretching allow direct observation of chromatids at kilobase resolution (see (Houseal and Klinger, 1994)). Since these techniques also permit simultaneous observation of several hundred kilobases in the same field, they will progressively take the place of indirect molecular analytical procedures.

The present objective of physical mapping is to establish an ordered collection of cloned DNA fragments covering the totality of the genome (Fig. 2) (U.S. Congress, 1988; U.S. Department of Health, 1990). This collection will eventually serve as a source of material to be sequenced, but has a more immediate use in the positional cloning of disease genes. When the genetic localization is sufficiently precise, and permits delimitation of a small interval on the chromosome within which the disease gene is situated, one attempts

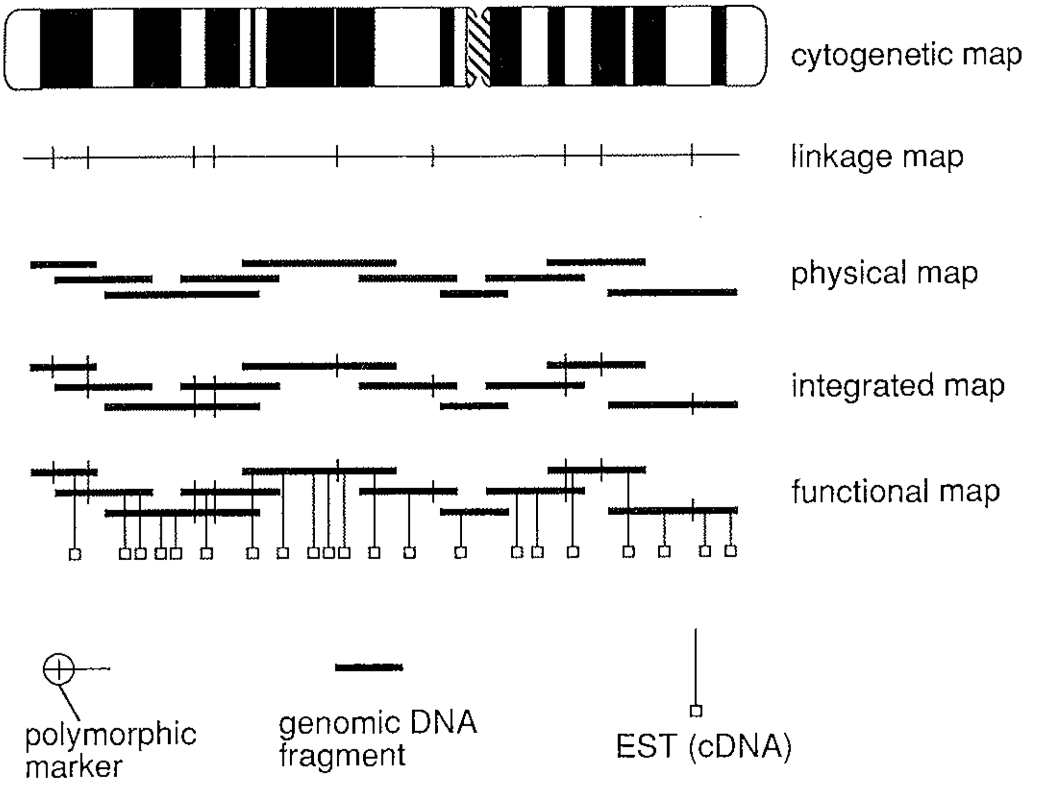


FIG. 2

Cytogenetic map. This schematic map represents the chromosome bands observed in the light microscope after staining of the chromosomes. The alternation of dark and light bands corresponds to the pattern of stained and unstained zones seen in the microscope. Each chromosome has a specific banding pattern, which makes it posible to identify each one of the 23 pairs of human chromosomes.
Genetic map. Map of markers which correspond to variations in the sequence of DNA (polymorphisms). These variations are stable, transmissible from one generation to the next, and permit differentiation between individuals as well as between the two copies of each chromosome in a single individual. The order of these markers and the distances which separate them are determined by a genetic method based on statistical analysis of the transmission of the different observed variations in polymorphisms in human families. The genetic map makes it possible to determine the parental and grandparental origin of each chromosome segment and thus to associate the transmission of a given segment with a genetic trait (such as a disease).
Physical map. There are several different ways of representing the physical map of the chromosomes. In this example it consists of an alignment of overlapping segments of cloned DNA. The ordering of these overlapping fragments represents a reconstitution of the DNA molecule of a chromosome.
Integrated map. This map combines elements of different origin. In this example, markers from the genetic map have been placed on segments of DNA from the physical map. Integration of these maps allows verification of whether the orders obtained using independents methods are identical.
Functional Map or map of the genes. This involves positioning the known genes as well as parts of unknown genes called ESTs (Expressed Sequence Tags) on a physical map, such as the ordered collection of overlapping segments above.

to cover this region with the help of cloned fragments. The efficiency of "walking" toward disease genes can be considerably enhanced by ordered collections of overlapping DNA fragments which are arranged according to their positions on the chromosomes. Such collections already exist for the small genomes of certain model organisms (yeast, fruit fly and nematode).

For several years now it has been possible to obtain and propagate large DNA fragments (from several hundred kilobases to over one megabase) as yeast artificial chromosomes (YAC) (Burke *et al.*, 1987). To arrange these fragments in order it is necessary to find parts which are common to more than one fragment so that the overlapping regions can be recognized. Therefore, methods for the identification of these fragments must be used. The PCR is one technique which is commonly used for this identification process. It enables scientists to amplify both small unique fragments known as sequence-tagged sites (STS) using unique primers (Olson *et al.*, 1989), and sequences found between the repeated elements which are very numerous in the genome (Nelson *et al.*, 1989). In the latter case it is unnecessary to know the sequence that one wishes to amplify. Other methods attempt to produce a fingerprint for each cloned fragment: the sub-fragments resulting from digestion of the cloned fragments by one or more restriction enzymes are then characterized by their size and part of their content (this latter is defined by the presence or absence of small sequences such as restriction sites [Coulson *et al.*, 1986] or presence of repeated sequence elements in sized fragments [Stallings *et al.*, 1990]). Fingerprints of all the digestions which have been carried out are then compared with each other using computer programs, in order to identify the regions which are common to different clones (Bellanné-Chantelot *et al.*, 1992). As with the genetic map, the construction of a physical map of the genome by constituting an ordered collection of large DNA fragments in the form of yeast artificial chromosomes requires considerable experimental work which is also carried out in specialized laboratories which have a large degree of automation and substantial material and human resources. A first physical map of the ensemble of the human genome was produced by the CEPH and Généthon groups at the end of 1993 (Cohen *et al.*, 1993).

Functional map

In addition to the 3,000 to 4,000 known genes, fragments of cDNAs of several tens of thousands of messenger RNAs which correspond to the parts of the genes which are expressed (exons) have now been sequenced (Expressed Sequence Tags or ESTs) (Adams *et al.*, 1991). The localization of the majority of these ESTs on the chromosomes is unknown. However, prior mapping of these sequences should considerably facilitate the identification of the genes responsible for genetic diseases. There are plans to place these ESTs on an existing map. They will be positioned on an ordered collection of YACs when possible (Fig. 2), but since this collection only partially covers the genome, they will be localized on a map produced using a collection of radiation hybrids which were previously characterized with regularly-spaced markers from the genetic map. In this way the ESTs can be rapidly analyzed to intervals localized by markers from the genetic map. Mapping of expressed sequences should not be expected to lead to a complete inventory of our genes, however. This goal can only be reached by genomic sequencing.

Sequencing

The goal of the program for the study of the human genome is to obtain the complete sequence of each chromosome. Unfortunately, this goal will be impossible to attain in the allotted time period and at reasonable cost using the techniques and methods presently available, in spite of automation. On the other hand, studies are presently being carried out on small genomes such as that of yeast (16 million nucleotides and 7,000 genes) and a small nematode (100 million nucleotides and 15,000 to 20,000 genes). This work has already been completed for 4 chromosomes of yeast by a network of European laboratories (3 chromosomes) and one American group (one chromosome). The results have led to the conclusion that about half of the sequenced genes do not correspond to any known biological function. An Anglo-American nematode-sequencing project has the ambition of sequencing 20 million base pairs per year, and to finish the sequence of this genome by 1998. In Europe, the sequencing of the yeast genome is being accomplished in laboratories of a classical biology scale, whereas intensive sequencing units equipped with dozens of automatic sequencers have been created for the nematode program and by private companies for the sequencing of human ESTs.

If sequencing technology progresses with sufficient rapidity, the complete sequence of the genome will be available before all the disease genes have been identified by positional cloning. This latter technique will then be essentially replaced by computerized analysis of the intervals delimited by genetic mapping in which the previously-identified genes will be considered candidates according to predictions of function suggested by sequence analogies with known genes.

Conclusion

Investigations of genomes are in the process of causing profound changes in the organization of biological research. Private investors have also shown a growing interest in projects for the massive acquisition of knowledge in the field of genetics and genomics (sequencing of ESTs; creation of companies to search for genes implicated in common diseases such as diabetes and cardiovascular disease, etc.).

The mapping and sequencing program for the human genome and the genomes of model organisms should lead to a complete inventory of the genes of living organisms. Moreover, it will generate a substantial number of new questions in the domain of biology and greatly stimulate research for the coming century. The identification of the genes responsible for a large number of genetic diseases will itself lead to a new era of medicine in which we will attempt to correct the errors that are inscribed in the genomes of diseased individuals.

REFERENCES

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnik, M., Polymeropoulos, M.H., Xiao, R., Merril, R.H., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R. and Venter, J.C., 1991, «Complementary DNA sequencing: expressed human sequence tags and human genome project», *Science* 252, 1651-1656.

Bellanné-Chantelot, C., Lacroix, B., Ougen, P., Billault, A., Beaufils, S., Bertrand, S., Georges, I., Glibert, F., Lucotte, G., Susini, L., Codani, J.J., Gesnouin, P., Pook, S., Vaysseix, G., Lu-Kuo, J., Ried, T., Ward, D., Chumakov, I., Le Paslier, D., Barillot, E. and Cohen, D., 1992, «Mapping the whole genome by fingerprinting yeast artificial chromosomes», *Cell* 70, 1059-1068.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W., 1980, «Construction of a genetic linkage map in man using restriction fragment length polymorphism», *Am. J. Hum. Genet.* 32, 314-331.

Burke, D.T., Carle, G.F. and Olson, M.V., 1987, «Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors», *Science* 236, 806-812.

Cohen, D., Chumakov, I. and Weissenbach, J., 1993, «A 1st-generation physical map of the human genome», *Nature* 366, 698-701.

Collins, F. and Galas, D., 1993, «A new five-year plan for the U.S. genome project», *Science* 296, 43-46.

Committee on Mapping and Sequencing the Human Genome, National Research Council, 1988, *Mapping and sequencing the human genome*. (Washington: National Academy Press).

Coulson, A., Sulston, J., Brenner, S. and Karn, J., 1986, «Toward a physical map of the genome of the nematode *Caenorhabditis elegans*», *Proc. Natl. Acad. Sci. USA* 83, 7821-7825.

Dausset, J., Cann, H., Cohen, D., Lathrop, M., Lalouel, J.M. and White, R., 1990, «Program description: Centre d'Etude du Polymorphisme Humain (CEPH): collaborative genetic mapping of the human genome», *Genomics* 6, 575-578.

Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T.P., Bowden, D.W., Smith, D.R., Lander, E.S., Botstein, D., Akots, G., Rediker, K.S., Gravius, T., Brown, V.A., Rising, M.B., Parker, C., Powers, J.A., Watt, D.E., Kauffman, E.R., Bricker, A., Phipps, P., Muller-Kahle, H., Fulton, T.R., Ng, S., Schumm, J.W., Braman, J.C., Knowlton, R.G., Barker, D.F., Crooks, S.M., Lincoln, S.E., Daly, M.J. and Abrahamson, J., 1987, «A genetic linkage map of the human genome», *Cell* 51, 319-337.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J., 1994, «The 1993-1994 Généthon human genetic linkage map», *Nature Genetics* 7, 246-339.

Houseal, T.W. and Klinger, K.W., 1994, «What's in a spot», *Human Molecular Genetics* 3, 1215-1216.

McKusick, V., 1992, *Mendelian inheritance in man. Catalogs of autosomal dominant, autosomal recessive, and X-linked phenotypes*, 10th ed. (Baltimore: The Johns Hopkins University Press).

Murray, J. et al., 1994, «A comprehensive human linkage map with centimorgan density», *Science* 265, 2049-205.

Nathans, D., 1979, «Restriction Endonucleases, Simian Virus 40, and the new genetics», *Science* 206, 903-909.

Nelson, D.L., Ledbetter, S.A., Corbo, L., Victoria, M.F., Ramírez-Solis, R., Webster, T.D., Ledbetter, D.H. and Caskey, C.T., 1989, «Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources», *Proc. Natl. Acad. Sci. USA* 86, 6686-6690.

Olson, M., Hood, L., Cantor, C. and Bostein, D., 1989, «A common language for physical mapping of the human genome», *Science* 245, 1434-1435.

Saiki, R.K., Scharf, S.J., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N., 1985, «Enzymatic amplification of beta-globin sequences and restriction site analysis for diagnosis of cycle cell anemia», *Science* 230, 1350-1354.

Stallings, R.L., Torney, D.C., Hildebrand, C.E., Longmire, J.L., Deaven, L.L., Jett, J.H., Doggett, N.A. and Moyzis, R.K., 1990, «Physical mapping of human chromosomes by repetitive sequence fingerprinting», *Proc. Natl. Acad. Sci. USA* 87, 6218-6222.

U.S. Congress, Office of Technology Assessment, 1988, *Mapping Our Genes-The Genome Projects: How Big, How Fast? OTA-BA-373.* (Washington: U.S. Government Printing Office).

U.S. Department of Health and Human Services and U.S. Department of Energy, 1990, *Understanding our Genetic Inheritance-The U.S. Genome Project: The First Five Years FY 1991-1995.* (Springfield, Va: National Technical Information Service, U.S. Department of Commerce).

Uberbacher, E.C. and Mural, R.J., 1991, «Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach», *Proc. Natl. Acad. Sci. USA* 88, 11261-11265.

Watson, J.D., 1990, «The Human Genome Project: Past, Present, and Future», *Science* 248, 44-49.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. and Lathrop, M., 1992, «A second-generation linkage map of the human genome», *Nature* 359, 794-801.

THE HUMAN GENOME DIVERSITY PROJECT *

LUCA CAVALLI-SFORZA **

The Human Genome Diversity Project (HGD Project) is an international anthropology project that seeks to study the genetic richness of the entire human species. This kind of genetic infomation can add a unique thread to the tapestry of our knowledge of humanity. Culture, environment, history, and other factors are often more important, but humanity's genetic heritage, when analyzed with recent technology, brings another type of evidence for understanding our species' past and present. The Project will deepen our understanding of this genetic richness and show both humanity's diversity and its deep and underlying unity.

The HGD Project is still largely in its planning stages, seeking the best ways to reach its goals. The continuing discussions of the Project, throughout the world, should improve the plans for the Project and their implementation. The Project is as global as humanity itself; its implementation will require the kinds of partnerships among different nations and cultures that make the involvement of UNESCO and other international organizations particularly appropriate.

I have spent most of my career studying human genetic diversity, and I am the chair of the International Executive Committee of this proposed Project. In my talk, I will briefly discuss the Project's history, describe the Project, set out the core principles of the Project, and demonstrate how the Proiect will help combat the scourge of racism.

A Short History of the HGD Proiect

The Human Genome Diversity Project began, formally, in 1991. It resulted from my interaction with Allan Wilson, professor of Biochemistry

---

* Presented at UNESCO, Paris, September 21, 1994.

** Pontifical Academician. Professor of Genetics, emeritus (active), Stanford California, U.S.A.

at the University of California, Berkeley. Allan was responsible for the most interesting developments in the evolutionary analysis of mitochondrial DNA. My own involvement in the study of human evolution began in 1951. Allan and I began collaborating in 1990, through the offices of Mary-Claire King, professor of genetics at Berkeley who had graduated with him in 1975 and who collaborated with me in the early eighties. The three of us, along with Charles Cantor and Bob Cook-Deegan, a historian of the human genome exploration, published an appeal in *Genomics* in 1991, making known to molecular geneticists the need for a systematic study of the human species. Sir Walter Bodmer, director of the Imperial Cancer Research Fund of London and a long-time collaborator — we wrote together two relevant books, *The Genetics of Human Populations* and *Genetics, Evolution, and Man* — was at the time Chairman of the international Human Genome Organization, HUGO. In this capacity he named a committee to foster the development of the Human Genome Diversity Project, which I was asked to chair. Unfortunately, Allan Wilson died soon after.

In the fall of 1991, I convened a meeting at Stanford of professors Ken Weiss, Marc Feldman, and Mary-Claire King (we kept in touch with Ken Kidd by phone) to discuss applying for funds with which to organize workshops designed to plan the activity of the HGD Project. Funds were obtained through the U.S. National Science Foundation, with the National Institutes of Health and the U.S. Department of Energy also participating. Various aspects of the project were discussed in four meetings:

1. July 1992, Stanford, California — populations genetics and statistical aspects;
2. October 1992, Pennsylvania State University, Pennsylvania — anthropological aspects, especially sampling;
3. March 1993, Bethesda, Maryland — molecular, biological, and ethical aspects; and
4. September 1993, Alghero, Sardinia, Italy — first international meeting, and formal constitution of the organs of the HGD Project: International Executive Committee, Regional Committees, and International Forum.

These workshops have helped give substance to the goals and methods of the HGD Project. Their conclusions are described below.

In addition to these formal planning meetings, the Project has been discussed in many other settings. In the first South-North Human Genome meeting sponsored principally by UNESCO in May 1992, held in Caxambu, Minas Gerais, one day was dedicated to Human Genome Diversity. It was the subject of a U.S. Senate hearing, convened by Senator

Daniel Akaka of Hawaii in April 1993. It was also the subject of a special symposium of anthropologists, convened in November 1993 by the Wenner-Gren Foundation. In January 1994, HUGO (the Human Genome Organisation) formally adopted the Human Genome Diversity Project as one of its projects. In May 1994, I met in Paris, along with Professor Henry T. Greely, chair of the ethics subcommittee of the North American Regional Committee of the HGD Project, with Justice Noëlle Lenoir and others to discuss this Project. Throughout the past several years, representatives of the Project have met with many different groups interested in this topic.

Main Decisions taken at the Meetings of the HGDP

From the outset the HGD Project has been planned on a global basis with world-wide coordination of sampling and testing. Local participation in all areas of the world will be essential and the success of the project will also be entirely dependent on international collaboration and cooperation.

It is recognised that not all regions of the world are experienced in the techniques of molecular biology and genetics and that some countries will not, in the foreseeable future, acquire the 'cutting edge' technology that is needed for the mapping and sequencing work of the Human Genome Project. However, it is feasible for the more limited technological demands of the HGD Project to be met by most countries, given training of laboratory staff and help with techniques. One of the exciting aspects of the HGD Project is that it offers all countries a unique opportunity to become involved in, and contribute to, the global human genome initiative by undertaking the collecting and typing of samples from their own region as well as other studies of local interest.

At the first workshop, in Stanford in July 1992, population geneticists and statisticians were invited to discuss questions of scientific feasibility and sampling. Two main conclusions were reached.

(1) Samples should be taken from about 400-500 "populations", a term indicating ethnic groups defined by a self-selected name, chosen by some criteria that would make them representative of the world's populations. An alternative sampling criterion championed by Wilson — namely, taking single individuals at regular geographic distance from each other — was not completely excluded, although it was acknowledged that it presents practical difficulties.

(2) Two types of samples from each population are needed:

(a) From a smaller number of individuals (about 10 to 50, possibly 20-25), blood samples would be taken rapidly to nearby laboratories where white cells could be transformed, to grow in test tubes and produce DNA for the analysis of genetic material. This would ensure that material for biological study would be available for indefinite periods of time.

(b) From a larger number of individuals, 100-200 persons, small amounts of blood or other biological fluids like saliva or specimens like hair are to be collected, from which smaller, finite amounts of DNA would be available for studies which statistically require larger numbers of individuals from each population.

This work allowed us to form a first estimate of the cost of the project. Most of the specific expenses would be due to collecting the samples and preparing transformed cell lines. The latter are estimated to cost about $500 per individual; with 10,000 cultures to be made in a period of, say, five years, for a total cost of about $5 million dollars. It is anticipated that the collection of samples would cost about the same. It is more difficult to estimate the cost of testing (that is, of ascertaining the DNA variation among the sampled people); this is open-ended, since the material should be available for a long time and stimulate interest in new genetic markers as they develop from future research. It was estimated that the costs of testing and of administration might bring the total for five years to the order of $25-$30 million. This is about 1% of the anticipated cost of the Human Genome Project itself, which is now underway and which, as is well known, does not plan to study more than one genome; that is, the Human Genome Project, unlike the Diversity Project, will not include variation. It is a modest cost, compared with the majority of biological projects.

The second workshop, at Penn State, considered which coverage of the world's genetic variation was most appropriate. The scope of the exercise was to record examples of populations that were considered of special interest by anthropologists. Since that time, we have decided to broaden our investigations. For instance, we found that African-Americans have considerable interest in their origins, which are usually unknown since relevant historical information is difficult to find. In fact, African-Americans in the United States had started independent research on the subject. We would like to include in our research all ethnic groups who declare an interest in our work

The third workshop, at Bethesda, Maryland, considered technical and ethical problems. There is little worth reporting here on the former, the latter are the main subject of the present talk, and considerations suggested by discussion at the workshop are incorporated in the balance of these remarks.

In the fourth workshop, held at Alghero, there was for the first time a representation from 24 different countries. It was decided to enlarge the

International Executive Committee to include representatives from India, Africa and Japan. It was decided that there should be an International Forum, in which funding agencies would be represented and each region would be represented by one or more members. Regional committees should form of their own initiative. The European regional committee was already in existence and has been funded by the European Economic Community since 1992. It has elected Sir Walter Bodmer as its chairman. In the course of the last year, other regional committees have begun to form. A North American committee has formed and met twice at Stanford, in January and in August 1994. Its chairman is Professor Ken Weiss of Pennsylvania State University, and it is the only region so far to have proposed bylaws for its own functioning. A South American Committee has been formed, with Professor Sergio Pena in the chair. A Chinese committee has been formed, with Professor Du Ruofo of Beijing Academia Sinica in the chair. Other regional committees are forming in other parts of the world.

The International Executive Committee met in Alghero, for the first time, and on September 16, in London.

Specific Aims, Design, and Methods of the HGDP

The HGDP is a project designed to take a sample of the human species for the purpose of genetic study of individual variation. The number of individuals to be sampled will be between 10,000 and 100,000. For most of these individuals, very small quantities of their genetic material, DNA, will be stored in suitable repositories; for a smaller fraction, perhaps 10% of individuals, a more permanent source of DNA will be stored, in the form of transformed cell lines. This is a well-tested method for the conservation of live cells for indefinite periods, under conditions in which they can multiply, if desired, for producing DNA. This DNA will for the most part be identical to that of the individual of origin and will be made available for study to qualified scientists around the world. Information thus obtained will be made available for scientific purposes, including the history and origins of peoples. Knowledge of potential health importance may be acquired, for instance, concerning the incidence of inherited diseases, sensitivity and resistance to infectious agents or to diet or to environmentally-induced diseases, or concerning ways to optimize vaccines for specific populations and diseases.

It should be stressed that the research involved is not new. There is already a vast amount of genetic information on human populations collected since World War I. A recent review of the available data is contained in a book entitled *The History and Geography of Human Genes*, authored by L. Cavalli-Sforza, P. Menozzi, and A. Piazza (Princeton University Press, 1994). This book contains about 500 pages of text and tables and 500 maps giving information on the geographic distribution of 110 genes in the world and by continent. The data used in the book come from over 76,000 records found in the scientific literature and published before 1987. They refer to a total of 58 genetic systems and over 300 genes, and are based on blood samples, collected in about 3,400 different locations, published by several thousand investigators.

What is new in the HGDP is the fact that genes can now be studied in greater depth, at the level of genetic material itself, DNA. This ensures a precision and completeness not possible until now. Moreover, techniques can be standardized, and DNA can be stored more easily than the biological materials used before (red or white blood cells, serum, plasma, etc.). Genetic markers are the inherited characters that can be revealed by genetic analysis. A genetic marker is defined by two or more different forms that a given gene can take in different individuals. In general, human populations differ, with respect to a genetic marker, only in the relative frequencies of the different forms. An example of a genetic marker is the ABO blood group system, in which the gene determining the system can exist in three forms: A, B and O. The frequencies of A, B and O vary in different populations. There are other blood-group systems, all of which can be detected in the laboratory by immunological techniques. Other physical and chemical techniques are used to distinguish other types of markers, which have been known for decades. The investigations made by thousands of research workers with these "classical" procedures, known before the direct investigation of DNA was possible, were done in a haphazard manner, and there is considerable waste of the information then collected. For instance, population 1 might have been investigated for genes A and B; population 2, studied by a different team, for genes A and C; population 3 for C and D, limiting the number of direct comparisons among populations for the same genes to a small fraction of all the data collected over sixty or seventy years. To reach reproducible conclusions, it is necessary to use comparisons over many different genes, and therefore much of the existing classical information cannot be employed, or at least not in an optimal manner. Nevertheless, some problems have been solved, but there remain many puzzling observations and questions, which can now be easily solved by a systematic study using modern molecular methods, at the level of DNA.

A certain amount of material has already been analyzed with modern molecular techniques, and the comparison with conclusions obtained with classical methods shows no major discrepancies. We do not expect, therefore, any change in the basic conclusions that have already been reached with classical genetic markers, but the advantages to be obtained with a systematic collection of DNA and its analysis with standard methods are a considerable increase in the precision of the studies and an answer to the problems that escaped analysis before DNA analyses were possible.

The project will be carried out largely as an effort by each affiliated region. Supervision by the International Executive Committee and the International Forum will ensure that the scientific effort will be coordinated and that agreed-upon rules will be followed by all regional bodies. The molecular techniques employed in the project are the same as those used for molecular genetics applied to diagnostic problems in medicine or in other fields. Therefore, our insistence on carrying out the research directly in every country or region will allow these new advanced technologies to be introduced wherever biological, medical, or industrial laboratories exist. Technology transfer to the South is thus going to be a major benefit of the project.

For maximum efficiency, there should be some central storage of DNA and of the data collected. We are, however, far from having reached this point, and before this storage can be accomplished, we need to clarify minimum rules that each region must follow. A major current priority is the establishment of strong ethical rules that can be accepted generally by the many ethnic groups whose cooperation is necessary if the project is to succeed. In what follows I will summarize our current thoughts on these issues. Reactions from UNESCO's International Bioethics Committee will be very useful, especially as the project is still in its formative stages.

Basic Principles Guiding the HGDP

Minimum ethical principles must be formulated, recognizing that, in the process of collecting, the human rights of people in participating populations should be respected and that these people are partners in the work rather than merely subjects of it. Any particular region may have broader ethical concerns than those addressed here, which should be considered only as a common minimum. The principles here summarized address six points.

1. *Ethical duties of regional committees*

(a) Take steps to ensure that all samples are collected with meaningful informed consent, obtained from the government of the country, the local official authorities, the population sampled as a group, and also from every sampled individual. Informed consent means that all those involved have been informed about the process of taking samples, including risks, if any, and the general purposes for which the samples will be used. The language employed, and the formal embodiment of the consent, may have to vary depending on regions, situations, languages, customs, and applicable laws.

(b) Take steps to ensure that privacy of individuals participating in the project is protected. The individual identities must not be made known to the investigators handling the samples and must be respected in culturally appropriate ways.

(c) Encourage the full participation of the sampled population, as may happen, for instance, by giving the population a role in designing the questions to be studied, physical participation in the process (ideally, when possible by carrying on as much of the investigation locally as possible, with the maximum possible participation by local scientists and personnel, and in the process promoting technology transfer), and later keeping the population informed about results. All this will be more feasible if, for each population, an anthropologist who is in continuous contact with that population is involved in the process.

(d) Ensure the existence of a body that reviews ethical issues concerning the collecting activity that it promotes, for example, an ethics subcommittee of the regional committee.

2. *Rights of sampled populations concerning commercial use of samples or data*

It is not clear if the HGD Project will lead to the development of any commercial products, but it should be ready to handle this event if it arises, under the provisos that it will not benefit financially from such projects and that some reasonable share of financial benefits shall return to the sampled population. Each regional committee, therefore, must protect the rights of the sampled populations concerning the commercial use of the samples, of data derived from the samples, and of the results of the work carried out using the samples.

A possible practical method could be based on contracts between the region and those who collect the samples under its auspices, those who use the samples or data stored in repositories, and those who use information developed from those samples or the data. The contracts could provide that no commercial use of the samples or data, or information developed from them, could be made except under conditions specified by the regional committees.

Should financial benefits develop, a share of them shall return to the sampled populations. This may happen in a variety of ways; one possibility could be that a royalty on any commercial product be paid to a third party, which could be an appropriate international organization, recognized by the United Nations or UNESCO, to be used for the benefit of the populations whose samples contributed significantly to the development of the product. Alternatively, the regional committee might require that no commercial use be made without the interested party's negotiating an agreement directly with the relevant populations. In any case, before deciding on the method to adopt, the regional committee should consult with the population concerned. The method may vary from population to population.

3. *Populations to be sampled preferentially*

Obviously, only populations that are interested in being sampled will participate, and those that have been more intensively studied anthropologically have some priority. On the other hand, the coordination of efforts requires that resources of the project be used economically and the main aim of the project be respected: namely, sampling populations that, taken as a whole, represent the whole world. Within these limitations, regional committees may choose samples from populations that meet one or more of the following criteria:

(a) populations that can answer specific questions concerning the processes that have had the greatest impact on the genetic composition of contemporary "ethnic groups", language groups, and cultures;

(b) populations that are anthropologically unique, exhibiting unique cultural or linguistic attributes that distinguish them from most or all others;

(c) populations that might be especially informative in identifying the genetic etiology of important diseases;

(d) populations that are in danger of losing their identity as genetic, cultural, or linguistic units;

(e) populations that are geographically, linguistically, culturally, or historically related to populations meeting one or more of the criteria above; or

(f) populations from regions, language groups, or cultural types that have not otherwise been sampled.

4. *Minimum information to be obtained at each collection site*

The HGD Project is not a "genetics" project but one that seeks to combine genetic information with other sources of knowledge about our species. One important consequence is that the surveys done for the Project must include not just genetic samples but also broader information about the communities that participate. Each regional committee should ensure that for each surveyed community, the following kinds of information are collected:

(a) language;
(b) general aspects of its "culture" sufficient to place it on the world culture map (e.g., residency and marriage systems, patterns of adoption or fostering);
(c) the way in which the population was sampled (e.g., by household, by lineage, at a gathering place, clinic, etc.; these data should be collected by anthropologists familiar with the group).

For the individual:

(a) genealogical relationships among sampled individuals to the extent known within the population;
(b) birthplace and place of residence of the individual and of his/her parents;
(c) if applicable, clan of individual, father, and mother;
(d) birthplace, ethnic affiliation, and clan of spouse, unless both husband and wife have been sampled (in the records, localities and clan names may not be written explicitly but must be coded for necessity of privacy).

5. *General availability of results*

Each region must be committed to the general principle of sharing data and open access to all qualified researchers. Until repositories of cell lines and DNA samples, along with computerized data bases, are created and funded, however, and perhaps thereafter, absolutely free access may not be possible. Data collected by the local research workers may require that special agreements be made with the owners of those data about publication rights, if the data have not yet been published or accepted for publication. Other special provisions may be motivated by requests from regional committees for reimbursement of reasonable expenses in providing samples or data.

6. *Technology transfer*

New methods for biological research and development may prove to be the greatest technical breakthrough of the century. Knowledge of and access to those methods are spread unevenly across the world today. Because of its global scope and the important role of local and regional scientists in the Project, the HGD Project can be a useful tool for spreading the new biotechnology knowledge and methods around the world. Each regional committee must consider how it can diffuse that knowledge, through training local scientists and technicians, through donating equipment and supplies, and through other means.

THE HGD PROJECT AS AN ATTACK ON RACISM

People (including scientists) who are unfamiliar with our discipline, population genetics, are frequently frightened by the idea that the study of human diversity will encourage racism. Nothing is more false; in fact, population genetics has given the best proofs we have that racism is wrong.

Racism, one of the scourges of humanity, is the belief in the inherent, biological (genetic) superiority of a race, usually one's own. Unfortunately, it is again rampant in much of the world. It is usually coupled with the erroneous conceptions that races are "pure" and that a mixture of races is deleterious. These ideas developed especially in the nineteenth century and were responsible for the most odious hatred and persecutions, even genocide, especially in the first half of the twentieth century. They are far from gone and are easily revamped when an increase in foreign immigration brings people in frequent contact with the unfamiliar faces of foreigners and their diverse customs. In the minds of most people, there is also an unjustifiable, but often complete, confusion between differences in customs, which are easily changed if necessary, and biologically inherited differences, which are not. Some of the latter are very conspicuous, like skin, hair and eye color, and facial and bodily traits. They seem to be homogeneous characteristics of populations of specific origins; in fact, they often allow us to guess with little error where a person's ancestors may have come from. I believe it is this fact which leads people to assume, even if only subconsciously, that there are only a few different, well-characterized races on earth and, moreover, that races can be "pure" (i.e., homogeneous).

Another problem is that people are easily misled into believing that a great number of peculiar behaviors, which are simply acquired during upbringing, are also inherited like genetic traits, skin color, etc., and are immutable. Customs vary greatly from one human ethnic group to another. Why do people frequently tend to believe that customs may be inherited genetically? I assume it is because they are automatically associated with the physical differences we observe among "races". But a great deal of our behavior is learned, and customs are patterns of behavior that are carried over the generations by cultural mechanisms of transmission, which are poorly understood but have nothing to do with genetics. Many of these cultural differences are valuable and represent adaptations of great interest. Other cultural differences, however, may cause part of the resistance encountered by immigrants to foreign countries or by foreigners living in a country who are visibly different from the regular residents. It is my belief that difficulties in accepting foreigners are usually exacerbated by the very common misperception that these differences are part of the biological background and are therefore immutable.

These superficial, erroneous conclusions are thoroughly entrenched in the beliefs of many people and are in part responsible for generating hatred. To this, one may add that absurd ideas of racial superiority have often been taken as an excuse for occupying territories and whole regions previously settled by people who are less advanced economically. To these prejudices one can also add the totally unfounded, but still common, conviction that racial mixture is a cause of human racial "degeneration".

The truth is that there is no documented biological superiority of any race, however defined. Nowhere is there purity of races, except in plants and in some domestic animals that have undergone a special inbreeding process for laboratory purposes. Even in these cases there is no absolute "purity", which is very difficult to achieve. No damage is caused in humans by racial mixture. Humans thrive by remaining individually different one from the other. In fact, the concept of race can hardly be given a scientific, careful definition.

Why, then, do we see conspicuous differences among people originating from, say, Africa, Europe, or East Asia or among Native Americans or Australians? An explanation accepted by many scientists is that these differences are the result of physical adaptations to different climates, adaptations that developed only very recently in the history of humans, following their spread to continents, most probably starting from Africa between 60,000 and 100,000 years ago. Moving from hot climates to temperate or cold and extremely cold ones, or changing diets because the flora and fauna were different, has mediated such adaptations. It is likely that these differences are superficial, because the surface of the body is the place at which the interior of the body and the external environment (and the climate with it) come in contact. It is very easy to understand that adaptation to strong ultraviolet radiation from the sun has favored a development of protective pigmentation, hence dark skin. On the other hand, white skin helps in producing vitamin D. Not enough of this vitamin is present in the cereal diet to which Europeans have adapted in the last ten thousand years. Were the skin dark, not enough solar ultraviolet would penetrate it to allow the production of vitamin D from precursors present in cereals. Rickets would develop. Thus, at high latitudes, where the solar radiation is weak (unless the diet includes much fish and meat), the skin must evolve toward a lighter color. All other traits convince us, upon closer observation, that the superficial differences among "races", such as the shape of the face, of the eyes, of the nose, and the size and shape of the body, have most likely been engineered by nature so as to provide better adaptation to the different climates.

Thus, the characteristics we see with the naked eye that help us to distinguish individuals of the different continents are, in reality, "skin deep", and whenever we look under the veneer we find that the differences that seem so conspicuous to us are really trivial. The study of human population genetics has produced a most important result: apart from these superficial differences, which are few, individual humans are genetically quite diverse, but the average differences among human groups are small. They are much smaller than the superficial, skin-deep differences would lead us to believe, and they are also relatively small compared with the differences existing among individuals within the groups. When I say "groups", I am trying to replace the more common word "race", which is misleading because it is impossible to define accurately (and it is full of unpleasantness). A better word is "population", however defined: that of a village, a town, a region, or even a continent. Even a remote village is highly heterogeneous genetically when you look at truly inherited characteristics under the skin. Just to give some idea, a remote, highly isolated village may have a heterogeneity among individuals that is 50% or 60% as great as that of the whole world. Nevertheless, by a sufficiently detailed study we can find differences and similarities everywhere. Villages that are geographically close also tend to be similar genetically, but we can show that they differ, although very slightly, provided that we examine enough genes and individuals. Genetic analysis can be used to define close relatives or even individuals, given knowledge of their relatives. This is because genetic differences among *individuals* are of great magnitude. On the other hand, there has not been enough time to build up great differences among human *groups*. The greater the time since the divergence among groups began, the greater genetic diversity is expected. But humans started diverging only recently. Biological evolution is very slow, and there has been insufficient time to generate great differences among human regional groups. Those traits we see superficially are very noticeable, but the other characters that we see only by laboratory techniques are the greatest part of the differences among us, and these show small differences among groups. There has been no special reason for them to increase as dramatically as differences in skin or eye or hair color and the other superficial adaptations to different climates. The latter must of necessity affect the surface of the body and are, of necessity, very conspicuous. We should not be misled by this association into thinking that the hidden genetic differences are also as great.

Conclusion

Science per se, done in the interest of intellectual curiosity and in respect of life of humans and other species, is morally neutral. Its applications must be, however, under the control of society, because they can be good or bad. It is mostly because of such applications — or because of disagreement with religious dogmas — that science has come under fire, often for the wrong reasons. Ignorance can breed fear and hate, but I have discovered that it is most dangerous when mixed with the personal political agenda of science haters. We must take care to communicate about our project to a wide variety of audiences, many of which do not understand scientific methodology.

I am not suprised when I am attacked by racists because of my ideas; it is perfectly understandable. We cannot share any common ground. The confusion made by some others between our study and eugenics is less easy to understand or to forgive. The data and theory of population genetics show how difficult, or impossible, it is to carry out a eugenics program. It has even been asserted that the knowledge we generate can be used for the genocide of indigenous people. The vast genetic differences among individuals and the small differences between populations make any such application impossible. These kinds of baseless allegations stem largely from a general lack of understanding of human genetics. The Project will make strong efforts to combat this.

The information on genetic variation among individuals collected in the last sixty or more years, when there was no plan to coordinate the project, has already generated important knowledge. The new technologies applied according to the plan that we envisage will allow enormous improvement in our current understanding. Not only will intellectual curiosity be better satisfied, but also useful medical applications will improve welfare, and knowledge developed in this way will help further to reduce the impact of racism. The project is also an easy way to introduce important new methodologies to enable technology transfer to countries of the South. If any financial benefits come from pharmaceutical applications of the knowledge thus acquired, the HGD Project is also one organization that recognizes the rights of the original donors of biological material to participate in the benefits.

The Educational, Scientific, and Cultural importance of this project, as in the third, fourth, and fifth letters of the acronym of UNESCO, should be clear. We hope that we can establish useful links with UNESCO's International Bioethics Committee to further the Project's important work. I take from a section of the Alghero report, which is included in its entirety as an Appendix, the following sentences which are a good summary of the project:

> The primary case, therefore, for the Human Genome Diversity Project is cultural. The study of genetic polymorphism in human populations creates a unique bridge between the science of human genetics and the humanities, including anthropology, archaeology, history and linguistics, and presents scientists with a unique opportunity to contribute to the world's cultural heritage. There is a cultural imperative for us to respond to that opportunity and use the extraordinary scientific power that has been created through the development of DNA technology to generate — for the benefit of all people — information about the history and evolution of our own species.

FROM GENOMIC JUNK TO HUMAN DISEASE *

MAXINE F. SINGER **

Electron micrographs show DNA as a spaghetti-like mass of strings when a chromosome is "exploded" by stripping off its proteins [1]; the thin fibers are a single continuous DNA double helix. Similar helical molecules are in every chromosome, in every living organism. If the atomic details were visible, the double-helix would look like the now familiar molecular models [2]. The two chains of alternating phosphate and deoxyribose moieties are wound about each other with their purine and pyrimidine bases, usually abbreviated A, G, T, and C, forming pairs in the middle of the helix: A with T, G with C. In the nuclei of human cells, the individual chromosomes have, on the average, 120 million such base pairs of DNA, or altogether about 3 billion base pairs in each haploid genome.

Genes, the bits of genetic information, are stretches of DNA hundreds or thousands of base pairs long. Thus, each gene is a sequence of base pairs that contains biologically useful and interpretable information encoded in the order of the four base pairs. Such gene sequences are decoded by cellular mechanisms resulting in the formation of products useful to a cell — proteins or RNA molecules. Less well appreciated is the fact that estimates of the number of base pairs needed to encode the 100,000 or so genes that are likely to exist in the human genome are at most about 100 million base pairs, or less than 10% of the total. The remainder is often referred to as "junk" or non-sense DNA.

"Junk" has occasionally proved useful in some human endeavors. Picasso and others made wonderful sculptures of junk, though they dignified the trash with the term "found objects". Excursions in the junk yard of the human genome have also been rewarding, as described in this paper.

---

* Read 8 November 1991. Published in Proceedings of the American Philosophical Society, vol. 138, n. 1, 1994.

** Pontifical Academician. President, Carnegie Institution of Washington Scientist Emeritus, National Institutes of Health.

Genes and Mutations

Mendel and, later, others were able to discern the existence of what we now call genes because each gene can take several forms and can even change form, sometimes in unfortunate ways, thereby causing physiological malfunction and disease. The science of genetics flourished by taking advantage of this process of gene change — or mutation.

What is meant by different forms or versions of a gene — or by mutation? In molecular terms, a mutation is a change in the sequence of bases in DNA that alters the encoded information. One cause of mutation is a change from one base pair to another. In one well-known example, the change of a single base pair from an AT pair to a TA pair, alters one amino acid in the encoded hemoglobin β-polypeptide and results in sickle cell disease (Figure 1). Similarly, the deletion of one or more base pairs can cause a mutation, as can small rearrangements of the order of the base pairs along chromosomal DNA.

These kinds of changes take place more frequently than might have been thought. Consequently, most genes exist in several different forms in the members of any particular species, including humans. We could refer to these different forms as mutant forms, but that might erroneously imply that one form is the "right" one and the others are all defective. Thus, geneticists prefer a more neutral term, the term alleles, to describe the different versions of a particular gene. Some alleles may be defective (as is the allele for the gene encoding the β-polypeptide of hemoglobin in sickle cell disease), but others may yield a slightly different but functional protein which still allows the development of a healthy organism.

a.	ATG	GTG	CAC	CTG	ACT	CCT	GAG	GAG	AAG
	TAC	CAC	GTG	GAC	TGA	GGA	CTC	CTC	TTC
	Met	Val	His	leu	Thr	Pro	Glu	Glu	Lys
b.	ATG	GTG	CAC	CTG	ACT	CCT	GTG	GAG	AAG
	TAC	CAC	GTG	GAC	TGA	GGA	CAC	CTC	TTC
	Met	Val	His	Leu	Thr	Pro	Val	Glu	Lys

Fig. 1. The first few base pairs (A, T, G, etc.) and amino acids (Met, Val, etc.) in the gene and polypeptide corresponding to β-globin in a), the normal and b) the sickle alleles.

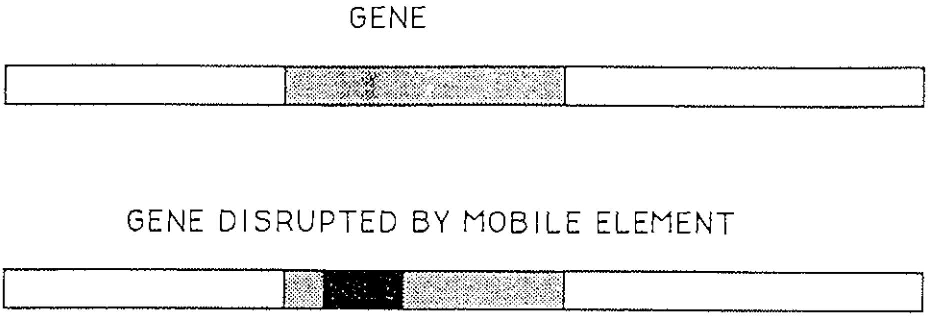


FIG. 2. Schematic diagram showing the disruption of a gene by a mobile element. The bar represents a double-helical segment of DNA.

Another well-known pair of alleles yield, respectively, the wrinkled and smooth peas that Mendel studied and that started the whole advance of modern genetics. A photograph of a pod containing both wrinkled and the more usual smooth peas appeared recently on the cover of the journal *Cell*, accompanying a report by Bhattacharyya and co-workers [3]. The difference between the wrinkled and the smooth peas is a change in a single gene; it is mutated in the wrinkled peas to a form an allele whose information is essentially silent.

Thanks to the ability to purify genes by recombinant DNA techniques, even though they are embedded in billions of base pairs of DNA, the nature of the allele causing wrinkled pea formation was defined. Here, the alteration was neither a base pair change, nor a deletion nor a small rearrangement. Instead it was the insertion, into the DNA of the gene, of another piece of DNA, several thousand base pairs long (Figure 2). This interruption of the gene's information makes it impossible for the cell and organism to decode the gene even though its sequence of bases remains in the genome.

MOBILE DNA ELEMENTS

In 1990, the discovery of such an interrupted allele was not surprising. Mobile or transposable DNA elements, popularly called "jumping genes", had by then been recognized for over forty years. In the mid-1940s, before it was known that genetic information is encoded in DNA, indeed even before the structure of DNA had been described, Barbara McClintock had concluded that there were jumping genes in corn — she called them "controlling elements" [4]. Very few people could follow her experiments or her logic and many people either actively disbelieved or simply ignored the mounting body of evidence she accumulated. At one point, she gave up publishing or lecturing. Today, even with hindsight, it remains difficult to follow McClintock's descriptions of her complex experiments and reasoning. But many investigators have now established that mobile elements are ubiquitous in nature and their analysis by modern techniques has made it possible to understand how they operate [5]. Using today's language, we can consider some of the corn kernels that McClintock studied.

Indian corn kernels are normally a deep red. A mutation in a gene required to synthesize the red pigment results in yellow kernels (and a healthy corn plant). The mutation is caused by a mobile element which lands inside the "pigment" gene and disrupts it. Occasionally, the element jumps out again. If this happens early in the growth of the kernel when there are only a few cells, almost the whole mature kernel will look red; it will have only a few yellow patches. But if the escape of the mobile element occurs late during kernel formation, then only a few cells will have the normal allele of the gene and the kernel will have only a few red spots. Intermediate markings also occur.

A variety of colored markings on snapdragon flowers are, similarly, a consequence of the activity of a mobile element [6]. Darwin himself described white snapdragon blossoms with red spots and stripes, although he was, of course, ignorant not only of jumping genes but of genes themselves.

Starting about twenty years ago, mobile elements began to be found in every kind of living organism — bacteria, invertebrates and vertebrates, including mammals. This vindicated McClintock and made plain just how profound was her discovery. For instance, mobile segments of DNA are responsible for the spread of much antibiotic resistance among bacteria; here, the elements' DNAs contain within them genes that make the bacteria resistant to drugs such as penicillin. Yeast and other fungi have mobile elements. Flies such as *Drosophila melanogaster*, appear to have a superabundance of them; more than twenty different kinds of mobile elements have been found in the genomes of these flies. Many of the *D. melanogaster* alleles that were used to discover the basic rules of genetics early in the twentieth century are now known to have been caused by mobile elements.

A Human Mobile Element

Until quite recently, the inability to carry out experimental breeding made the study of human genetics difficult. Geneticists had to count on chance mutations that gave obvious consequences, mainly inherited diseases, and on rare, very large multigenerational families. These investigations made it plain that human genetics works by the same general rules that apply to peas and flies. But the analysis of the nature of detrimental alleles developed only in the last twenty years, made possible by our contemporary ability to manipulate the DNA isolated from a small sample of human blood cells and to study its properties in great detail. Not surprisingly, mobile elements were only proven to be active in the human genome in 1988.

Haig Kazazian and his colleagues at The Johns Hopkins Medical School, studied a group of boys with the usually inherited disease, hemophilia A [7]. The failure to clot blood properly, which typifies this disease, is associated with mutated alleles of a gene that encodes a protein called Factor VIII. As most students of the inbred royal families in European history know, hemophilia A generally affects males but is inherited from their healthy mothers, not their fathers. This is because the gene for Factor VIII occurs in humans on the X chromosome. Females have a pair of X chromosomes; the lack of a functional Factor VIII gene on one of the pair is made up for by the functional gene on the other. If such a healthy carrier female mates with a healthy male, half of their daughters will be carriers (on the average) and half of their sons will be diseased. This is because each offspring gets one or the other of the mother's X chromosomes and either an X or a Y chromosome from the father; two Xs make for a female offspring, and X and a Y for a male. The male, lacking a second X chromosome and thus a functional Factor VIII allele, will be diseased.

Among the boys studied by Kazazian and his colleagues were two who were diseased although there was no history of the disease in their families; the boys were unrelated. Their illnesses appeared to be the result of new mutations. When the structure of the DNA of the Factor VIII genes was determined, it turned out that an unrelated DNA segment had inserted in the coding region of the genes. No such interruption was observed in the Factor VIII genes on the mothers' X chromosomes. The debilitating interruptions were new. They must have occurred during the maturation of the eggs that formed these boys, or very early in embryonic development. Kazazian and his colleagues determined the DNA sequence of the insertions in both the Factor VIII genes and found that they were very similar to one another. A computer search of the DNA sequence data bank into which scientists worldwide place the DNA sequences determined in their laboratories revealed that the insertion had base pair sequences almost identical to the sequence of some of the human junk DNA that we had been studying in my laboratory [8]. By then, it was suspected that members of this particular class of junk, named LINE-1, might be copies of a mobile DNA element. Kazazian's experiments proved this speculation and established that mobile DNA elements exist and jump in human chromosomes.

Genes and chromosomes have evolved over millennia to give well-functioning individuals of many species. It is difficult for many people to imagine that chromosomes, which provide elegant blueprints for the construction of marvelously functioning organisms of all kinds, including humans, have retained elements that can disrupt the genetic program. Other kinds of mutations appear to arise from random errors caused either by malfunctioning metabolic processes or by environmental agents such as radiation or chemicals. But mobile DNA elements are intrinsic to genomes. Why, in the course of evolution, were such disruptive DNA sequences retained?

Satisfactory answers to this question are not yet available. But a variety of possible answers are being considered and tested in many laboratories. Consequently a substantial amount of information about mobile DNA elements has been acquired [5]. Thus, all mobile DNA elements contain genes of their own. These genes encode proteins that enable the elements to move. So self-provided, and by poaching on some processes and molecules required by cells for other purposes, the mobile elements are virtually independent agents within the genome.

The Mechanism of Mobile Element Action

The known mobile DNA elements can be classified in a variety of ways. One way, which is appropriate regardless of species, divides them into two groups on the basis of the mechanism utilized for the insertion of the element into new chromosomal positions. Members of one group move within chromosomes by what has been called a "cut-and-paste" mechanism; the DNA of the mobile element is "cut" out of one place on the chromosomal DNA and "pasted" in another. The second group operates more like a FAX machine; a copy of the jumping segment is made and the copy appears in a new chromosomal position. One of the striking differences between the "cut-and-paste" and "FAX" mechanisms is that in the former, the total number of copies of the mobile DNA segment doesn't change while in the latter, one more copy of that DNA sequence is added to the genome with each new insertion; one active element can "FAX" many copies to many different chromosomal locations, particularly if it is active over millions of years of evolution.

The only known human mobile element, the LINE-1 DNA, is the

"FAX" type. It seems to have been churning out copies for a long time, at least since the days of our last common ancestor with other contemporary primates: even Old World monkeys have very similar LINE-1 DNAs. Indeed, LINE-1 elements are even more ancient than that. Very closely similar elements occur in all mammals and in marsupials. In human DNA, as in other mammals, there are about 100,000 LINE-1 sequences, which again suggests that they have been active for a long time. This means that LINE-1 sequences are between 3 and 5 percent of total human DNA. However, because nothing is known about the rate at which LINE-1 elements make and insert new copies, we have no idea how long it might take to accumulate 100,000 copies.

The observations on the hemophilia A cases mentioned earlier, suggest that LINE-1 segments must be active in germ line cells — developing eggs or sperm, or in very early embryos, before germ line cells are sequestered. LINE-1 segments are also active in somatic cells, but such activity could not lead to the accumulation of more and more LINE-1s in the human genome because such cells do not contribute to the formation of succeeding generations. The available experimental results support the hypothesis that LINE-1 elements are especially active in very early embryonic cells.

LINE-1 elements as well as a variety of other mobile elements of the FAX type are copied by a process called reverse transcription. In the first step, one strand of the LINE-1 DNA sequence is copied into a single strand of RNA [9]. RNA is chemically similar to a single DNA strand and the copying mechanism preserves the information in the DNA — that is, it preserves the sequence of bases. This process is called transcription: it leaves the original DNA intact and it is the same as the first step used to decode all genes in all organisms. Some of this RNA probably functions as a typical messenger RNA in the normal decoding process, thereby allowing synthesis of the proteins that are encoded within the element and are required for LINE-1 activity. The LINE-1 DNA itself contains special sequences of bases that serve as signals and markers for the start and stop of transcription, thus distinguishing the DNA segment from the surrounding DNA [10].

In the next step, which is *not* typical of normal genes, the RNA is copied back into DNA. This is the reverse of transcription. This DNA copy can then be inserted in a new chromosomal location. It is typical of mobile elements of the FAX type (and some viruses mentioned later) that they themselves carry the information to make the enzyme, reverse transcriptase, that carries out this process. In other words, a gene encoding reverse transcriptase is within the LINE-1 element.

A LINE-1 element encompasses about 6,000 base pairs of DNA

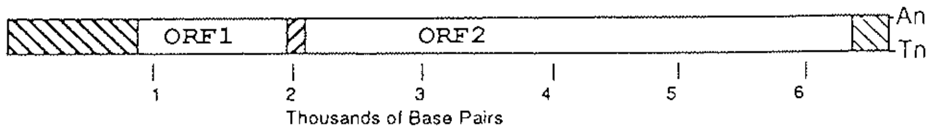


FIG. 3. Schematic diagram of a human LINE-1 element.

(Figure 3). At one end, the start, called the 5′ end (at the left in the figure), there are about 900 base pairs of DNA that do not appear to encode anything. This region is followed by about 1,100 base pairs that encode a polypeptide called p40, whose function is unknown [11]. A stretch of DNA that can be deciphered by the genetic code to yield a polypeptide is called an open reading frame; these 1,100 base pairs are called open reading frame 1. After a short DNA segment, there is another coding region, open reading frame 2, almost 4,000 base pairs long, which includes the segment that encodes the reverse transcriptase [12]. This is followed by another noncoding region of 200 base pairs. At the very end, the 3′ end at the right, is a region that is notably full of one kind of base pair, AT, with one strand mainly A, the other mainly T.

The hypothesis that LINE-1 sequences are in fact mobile elements of the FAX type, was tested some years ago by looking for the predicted RNA molecules [9]. Many human and monkey cell types were analyzed. The RNA was detected in only one kind of cell in culture: human teratocarcinoma cells. Such cells are obtained from tumors that arise from germ line cells. These experiments suggested that LINE-1s are especially active in relatively undifferentiated early embryonic-type cells. Since that time, this then tentative conclusion has been confirmed in several ways. For example, the p40 protein encoded by open reading frame 1 occurs in such cells but is not detectable in certain other cell types [11]. Also, reconstructed DNA segments carrying the 5′ and 3′ noncoding LINE-1 sequences and an easily measured "reporter" gene, are active in teratocarcinoma cells but not certain other cell types [10]. Recently, the p40 protein has been detected by immunological techniques in a few, but by no means all, tumor cell lines. Perhaps most interesting, p40 is readily detectable in a large percent of testicular tumors [13]. It has been suggested that p40 may be a useful diagnostic marker for tumors of testicular origin [13].

There are several unusual aspects to the transcription of LINE-1 elements into RNA and the translation of that RNA into p40 and polypeptide encoded by open reading frame 2. For example, the translation of open reading frame 1 into p40 begins 900 base pairs downstream from the beginning of the RNA — a very long distance. In most eukaryotic genes, the coding region begins shortly after the start of the RNA. Current evidence [13a] suggests that translation of open reading frame 2 into protein involves a fresh start, even further down the RNA, and is independent of whether or not open reading frame 1 is translated — again an unusual situation for eukaryotic protein synthesis. Transcription of LINE-1 elements begins at the first residue at the 5′ end of the element [14]. Yet, unlike most protein-encoding genes, all the DNA segments that regulate transcription — turn it on and off and specify the starting point — are *inside* the element, within the first 660 base pairs [10]. These include the DNA signals that limit transcription to only certain cell types.

This means that, at least in principle, each newly copied and inserted element could be independently competent to generate additional LINE-1s by transcription, reverse transcription, and so forth. If that were true, then LINE-1 elements could infest the entire genome, ruining many genes with lethal consequences for the organism and the species. However, it appears that most LINE-1 insertions are of incapacitated or incomplete copies of LINE-1. This is true of the two newly observed insertions that caused the hemophilias. And it also is true of the bulk of LINE-1 elements in the human genome.

Of the estimated 100,000 LINE-1 elements in the human genome, all but about 4,000 are incomplete, generally lacking DNA sequences from the left end — the "beginning" of the element. Many also have internal deletions or rearrangements. Among the full length LINE-1s many are incapable of acting as mobile elements because the reading frames encoding the polypeptides are ruined by base changes that interrupt translation — that is, by mutations.

Reverse Transcriptase

It is provocative to speculate about the significance of the reverse transcriptase enzyme encoded within an active LINE-1 element. From the time in the early 1950s when it was proven that genes are DNA sequences and the double-helical structure of DNA was proposed, through the 1960s when the genetic code was elucidated, it was believed that the flow of genetic information was unidirectional. The sequence of a gene in DNA could be transcribed into an RNA copy and the RNA then translated into a polypeptide.

In 1970, Temin and Baltimore independently first demonstrated that transcription can also be reversed; RNA can be copied into DNA [15, 16].

The discovery pertained to certain viruses — the retroviruses — and it was widely assumed that such retroviruses had an exclusive license for reverse transcriptase. But, as often happens, the earlier generalization had to be modified. Now it is known that several other kinds of viruses as well as two types of mobile elements encode and utilize reverse transcriptase. The mobile elements of one type are like retroviruses in that they have the same long sequence of DNA at each end of the elements' DNA, the so-called long terminal repeats. Those in the second group are similar to LINE-1 and have no long terminal repeats. Such viruses and mobile elements were found associated with yeast and other fungi, plants, invertebrates, and mammals. It was then concluded that while the presence of reverse transcriptase is not an exclusively viral activity, it is restricted to eukaryotes. This generalization also failed when it was discovered in 1989 that some (but not all) strains of *E. coli* and most, (if not all) strains of *Myxococcus xanthus* (another Gram-negative bacteria) also encode a reverse transcriptase, this one associated with the presence of a peculiar RNA connected by a 2′-OH branch to a single strand of DNA [17].

The generalization that then seemed secure concerning distribution of reverse transcriptase was that the enzyme is associated only with unusual genetic entities such as mobile elements and viruses that are not essential to cell function. But even this conclusion was short-lived. Recent evidence indicates that telomerase, the enzyme required presumably by all eukaryotic organisms to synthesize telomeres — the very ends of chromosomal DNA molecules — is also a reverse transcriptase [18]. These ends contain consecutive repetitions of a base sequence such as TTTTGG on one strand, and the complementary sequence CCAAAA on the other. Here, the RNA to be copied is associated with the enzyme protein, and, at each end of a chromosome, the 3′ end of one of the two DNA strands provides a primer onto which the new telomeric DNA is added. Thus, reverse transcriptase activity is apparently essential, at least to eukaryotes, whose linear chromosomal DNAs have ends. Moreover, telomerase must be quite ancient because the telomeres in all eukaryotes are very similar.

What, if any, is the relation of telomerase to the reverse transcriptases encoded by the mobile elements or viruses? The fact that LINE-1 elements occur in all marsupials and mammals does make it conceivable that their reverse transcriptase is required for *some* essential function, at least in these organisms, but there is now no reason to suggest that telomerase function is provided by LINE-1 elements. However, there is an intriguing observation which suggests that there may be at least an evolutionary relation between telomerase and the LINE-1 reverse transcriptase.

Mutation of a yeast gene, called *EST*1, results in successive shortening of yeast telomeres with successive cell divisions and ultimately, in cell senescence [19]. The structure of the gene predicts a protein with an amino acid sequence characteristic of reverse transcriptase. Thus, it is likely that *EST*1 encodes a protein portion of yeast telomerase. When the amino acid sequence of the *EST*1 protein was compared with that of various retrovirus and mobile element reverse transcriptases, the similarity to the human LINE-1 sequence was among the most notable.

In humans (but curiously not in mice) average telomere length (approximately 10,000 base pairs of repeating AGGGTTs) appears to decrease as individuals age or with serial divisions of primary cells in culture. And it has been suggested, though not demonstrated, that telomerase may be only at low levels in mammalian somatic cells [18]. The consequent loss of telomere length could then, according to these ideas, contribute to chromosome instability and aging in humans, as it does in yeast *EST*1 mutants. If this is so, then the relative levels of telomerase may turn out to correlate with LINE-1 expression.

Structural and Evolutionary Considerations

An analysis of the distribution LINE-1 sequences has led to suggestions that the element plays a role in the structure and, perhaps, function of human chromosomes [20, 21, 22]. Individual chromosomes in a cell can be routinely identified on the basis of the pattern of bands seen after staining with certain dyes. There is a reasonably good, though not perfect concordance between the so-called G-dark bands and the distribution of LINE-1 sequences as determined by *in situ* hybridization of LINE-1 DNA sequences to intact chromosomes. Thus, LINE-1s tend to concentrate in G-dark bands. Moreover, it is also known that the long stretches of DNA folded into these G-dark bands are relatively rich in AT compared to GC base pairs, and that G-dark bands replicate rather late during S-phase. The concentration of LINE-1 sequences in these regions may be fortuitous, but it is also possible that it contributes to the structural and functional specificities of the G-dark bands. If so, this could hint at one reason why human and other mammalian chromosomes have tolerated the high abundance of LINE-1 elements — namely that they play a role in chromosome structure [21].

There is yet another, quite different way to explain the frequency of LINE-1 like elements in eukaryotes and their special persistence and abundance in mammals, namely, on the basis of evolutionary mechanism. Ever since the discovery of what has been called "junk" DNA, a portion of which we now know to be mobile elements and their non-functional progeny, there has been speculation as to how and why, such sequences evolved. One speculation was put to rest succinctly by Ernst Mayr, in his book, "The Growth of Biological Thought" [23]: "The teleological answer that the seemingly functionless DNA is stored up in order to have it available in future times of need is altogether unsatisfactory". Mayr refers here to the fact that no organism has any way of "knowing" what may or may not be useful in the future. This does not mean that such seemingly functionless DNA may not turn out to be useful in the future, but only that the possibility can not explain its preservation. Mayr goes on to point out that Orgel and Crick, in 1980, made a proposal that has caught on with many, "that (the) extra DNA is, so to speak, parasitic, the (host) organism being helpless to prevent its replication and accumulation" [24]. A related idea is embodied in the term "selfish" DNA as it is applied to the repeated DNA sequences. Mayr is not enthusiastic about this interpretation, saying "it is intuitively distasteful to a Darwinian. Surely natural selection, a Darwinian would say, should be able to come up with a defense mechanism against such an expensive type of parasitism". But, of course all genomes are selfish in this respect, including both host and parasite. For example, evolutionary changes in a species' genome are seen to be defensive if they help protect the organisms from viral infection. But similarly, evolutionary changes in viral genomes are defensive if they counteract the organism's defense mechanisms. Thus, successful parasitism represents a stand-off.

We can then imagine that the ecological niche for LINE-1 elements is someone (or something) else's genome, and that for some now obscure reason, mammalian genomes may be an especially fit environment. As already pointed out, newly inserted copies of LINE-1 elements are usually — perhaps always — nonfunctional. One can imagine that without some such compromise the LINE-1's ecological niche would have been destroyed — along with the organism. However, each time the genome is duplicated, any active LINE-1s will also be duplicated thereby assuring the continuity of the element. Thus, LINE-1 elements appear to be parasitic on the human genome. If it were to turn out that the LINE-1 elements do indeed provide an important function to mammals, then the term symbiotic may be more appropriate than parasitic.

The idea of a genome as the niche for an unrelated segment of DNA is not new — that is what happens with lysogenic bacterial viruses like lambda bacteriophage in *Escherichia coli* or with endogenous retroviruses. Such situations are one possible result of infection of cells by these viruses.

Another question about mobile elements like LINE-1s concerns their origins. One possibility is that they arose from normal genomic sequences by processes of recombination and mutation that yielded a competent mobile element. This could have happened once during evolution followed by the spread of the elements to many species. Alternatively, similar elements might have emerged, independently, in fungi, plants, invertebrates, and vertebrates. Either way, this mechanism implies the prior existence of an independent reverse transcriptase in all cell types. Perhaps the ancient telomerase was a source of a reverse transcriptase gene.

Another possibility is that LINE-1 elements arose by mutation from an endogenous retrovirus or a retrovirus-like mobile element, again either once during evolution or repeatedly. This implies that retroviruses and the retrovirus-like mobile elements are evolutionarily more ancient than LINE-1 like elements. Several theoretical arguments can be raised against this. One is that the complex structure of the long terminal repeats of retrovirus-like entities and the critical role played by the long terminal repeats in the replication of retrovirus-like elements suggest a later, not an earlier origin than LINE-1 elements. Similarly, the specific endonuclease (or integrase) encoded in retrovirus-like elements is another complexity lacking in LINE-1 elements. Another argument is derived from evolutionary trees constructed on the basis of the similarities among the reverse transcriptases encoded by retroviruses, DNA viruses, retrovirus-like mobile elements, and LINE-1 elements; these considerations branch LINE-1 elements from the tree at a relatively very early time [25]. Thus, a third possibility is that the ancestors of LINE-1 elements are very ancient and were passed vertically during evolution of all eukaryotes. The presence of LINE-1 like elements in many different kinds of eukaryotes is consistent with this.

LINE-1 elements may also have arisen by the first or second possible mechanism in some organism and then been transmitted horizontally from one species to another. There is good evidence that a cut-and-paste type of mobile element in *Drosophila*, called P, was transmitted horizontally; a very recent paper suggests that the transmitting agent was a mite [26]. Also it is likely that a *Drosophila* LINE-1 like element named *jockey* was transmitted horizontally, probably from *Drosophila melanogaster* to *D. funebris* [27]; just how such horizontal transfer was achieved is unknown. This finding indicates that horizontal transfer must be considered in investigating the origins of the LINE-like elements. Most likely, many of these mechanisms as well as others have contributed to the origin of the LINE-like elements in contemporary genomes.

Concluding Remarks

What seemed to some uninteresting junk in the human genome has turned out to have implications for a broad range of biological questions — human development and disease, the origin and distribution of reverse transcriptase, and evolution itself. For the near future, it is of special importance to continue to investigate the detailed mechanism by which LINE-1 moves in the human genome and the significance of these events to human biology.

REFERENCES

[1] J.R. Paulson and U.K. Laemmli. 1977. The Structure of Histone-depleted Metaphase Chromosomes. *Cell* 12: 817-828.
[2] J.D. Watson and F.H.C. Crick. 1953. Molecular Structure of Nucleic Acids: A Structure for Deoxynucleic Acids. *Nature* 171: 737-738.
[3] M.K. Bhattacharyya, A.M. Smith, T.H.N. Ellis, C. Hedley, and C. Martin. 1990. The Wrinkled-Seed Character of Pea Described by Mendel is Caused by a Transposon-Like Insertion in a Gene Encoding Starch-Branching Enzyme. *Cell* 60: 115-122.
[4] B. McClintock. 1987. Collected Papers: The Discovery and Characterization of Transposable Elements. Garland Publishing, Inc., New York.
[5] D.E. Berg and M.M. Howe (eds.) 1989. Mobile DNA. Am. Soc. for Microbiol. Washington, D.C.
[6] E.S. Coen, T.P. Robbins, J. Almeida, A. Hudson, and R. Carpenter. 1989. Consequences and Mechanisms of Transposition in *Antirrhinum majus*. Ibid., 413-436.
[7] H.H. Kazazian, Jr., C. Wong, H. Youssoufian, A.F. Scott, D.G. Phillips, and S.E. Antonarakis. 1988. Hemophilia A Resulting from *de novo* Insertion of L1 Sequences Represents a Novel Mechanism for Mutation in Man. *Nature* 332: 164-166.
[8] T.G. Fanning and M.F. Singer. 1987. LINE-1: A Mammalian Transposable Element. *Biochim. Biophys. Acta* 910: 203-212.
[9] J. Skowronski and M.F. Singer. 1985. Expression of a Cytoplasmic LINE-1 Transcript is Regulated in a Human Teratocarcinoma Cell Line. *Proc. Natl. Acad. Sci. U.S.A.* 82: 6050-6054.
[10] G. Swergold. 1990. Identification, Characterization, and Cell Specificity of a Human LINE-1 Promoter. *Mol. Cell. Biol.* 10: 6718-6729.
[11] D.M. Liebold, G.D. Swergold, M.F. Singer, R.E. Thayer, B.A. Dombroski, and T.G. Fanning. 1990. *Proc. Natl. Acad. Sci. U.S.A.* 87: 6990-6994.
[12] S.L. Mathias, A.F. Scott, H.H. Kazazian, Jr., J.D. Boeke, and A. Gabriel. 1991. Reverse Transcriptase Encoded by a Human Transposable Element. *Science* 254: 1808-1810.
[13] G.L. Bratthauer and T.G. Fanning. 1992. Active LINE-1 Retrotransposons in Human Testicular Cancer. *Oncogene* 7: 507-510.
[13a.] J.P. McMillan and M.F. Singer. 1993. Translation of the Human LINE-1 Element, L1Hs. *Proc. Nat'l. Acad. Sci.* U.S.A. 90: 11533-11537.
[14] J. Skowronski, T.G. Fanning, and M.F. Singer. 1988. Unit-Length LINE-1 Transcripts in Human Teratocarcinoma Cells. *Mol. Cell. Biol.* 8: 1385-1397.
[15] D. Baltimore, 1970. Viral RNA-Dependent DNA Polymerase. *Nature* 226: 1209-1211.
[16] H.M. Temin and S. Mizutani. 1970. Viral RNA-Dependent DNA Polymerase. *Nature* 226: 1211-1213.
[17] B.C. Lampson, M. Viswanathan, M. Inouye, and S. Inouye. 1990. Reverse Transcriptase from *Escherichia coli* Exists as a Complex with msDNA and Is Able to Synthesize Double-stranded DNA. *J. Biol. Chem.* 265: 8490-8496.
[18] E.H. Blackburn. 1991. Structure and Function of Telomeres. *Nature* 350: 569-573.
[19] V. Lundblad and E.H. Blackburn. 1990. RNA Dependent Polymerase Motifs in *EST1*: Tentative Identification of a Protein Component of an Essential Yeast Telomerase. *Cell* 60: 529-530.

[20] G. Bernardi. 1989. The Isochore Organization of the Human Genome. *Annu. Rev. Genet.* 23: 637-661.
[21] L. Manuelidis. 1990. A View of Interphase Chromosomes. *Science* 250: 1533-1540.
[22] J. Korenberg and M.C. Rykowski. 1988. Human Genome Organization: Alu, Lines, and the Molecular Structure of Metaphase Chromosome Bands. Cell 53: 391-400.
[23] E. Mayr. 1982. The Growth of Biological Thought. The Balknap Press, Cambridge, MA. p. 579 *et seq.*
[24] L.E. Orgel and F.H.C. Crick. 1980. Selfish DNA: the Ultimate Parasite. *Nature* 284: 604-607.
[25] Y. Xiong and T.H. Eickbush. 1990. Origin and Evolution of Retroelements Based on their Reverse Transcriptase Sequences. *EMBO J.* 9: 3353-3362.
[26] M.A. Houck, J.B. Clark, K.R. Peterson and M.G. Kidwell. 1991. Possible Horizontal Transfer of *Drosophila* Genes by the Mite *Proctolaelaps regalis. Science* 253: 1125-1129.
[27] L.J. Mizrokhi and A.M. Mazo. 1990. Evidence for Horizontal Transmission of the Mobile Element *jockey* Between Distant *Drosophila* Species. *Proc. Natl. Acad. Sci. U.S.A.* 87: 9216-9220.

II.

LEGAL ASPECTS

LA INVESTIGACIÓN GENÉTICA EN EL MARCO DE LA CONSTITUCIÓN ESPAÑOLA: SOBRE EL DERECHO A LA INTIMIDAD

RAFAEL DE MENDIZABAL ALLENDE *

A) En la jurisprudencia constitucional

1. *Concepto*

Habida cuenta de que no existe un pronunciamiento concreto del TC sobre la incidencia del derecho a la intimidad en relación a la ingeniería genética, me parece adecuado como método de trabajo, examinar los casos más relevantes en los que este Tribunal se ha pronunciado sobre el contenido del derecho a la intimidad, y los problemas que plantea, para extraer las consecuencias oportunas.

STC 110/84. Ponente D. Ángel Latorre

- Asunto: se impugnó una resolución de la Inspección Tributaria que conminaba a un contribuyente (Garrido Falla) a presentar sus cuentas bancarias ante la Administración Tributaria.

- FJ 3:

— La idea originaria del derecho a la intimidad, expresamente reconocido en escasos textos constitucionales, es el respeto a la vida privada, lo que tienen su engarce en las libertades tradicionales.

— El avance de la tecnología y los medios de comunicación ha obligado a mantener la protección de la intimidad, más allá del aseguramiento del domicilio y del respeto del derecho a la correspondencia.

— De esa situación se deriva la necesidad del reconocimiento del derecho a la intimidad o la vida privada que abarque a las intromisiones

* Magistrado del Tribunal Constitucional.

que por cualquier medio puedan realizarse en el ámbito reservado de las personas.

• FJ 4:

— se plantean dos cuestiones:

a) ¿en qué medida el conocimiento de las cuentas bancarias por la Administración a efectos fiscales se encuentra protegido por el derecho a la intimidad?

b) ¿en qué medida puede conocerse por la vía de la Inspección Tributaria, hechos pertenecientes a la estricta vida personal y familiar del individuo? Una cuenta corriente es «como la biografía personal en números».

• FJ 5 y ss:

— Respecto de la primera cuestión, se recuerda que los derechos fundamentales no son ilimitados, y concurre en este caso la necesidad de satisfacer un interés constitucionalmente protegido, como es la contribución al sostenimiento de las cargas públicas. Al figurar en los extractos bancarios la causa genérica de cada partida, pero no la causa concreta, la limitación del derecho no se plantea en toda su intensidad.

— Respecto de la segunda, se razona que tanto el artíc. 17.1 del Pacto de Derechos Civiles y Políticos, como la Ley 1/82 de Protección del Honor, Intimidad y de la propia Imagen, protegen la intimidad frente a injerencias arbitrarias o ilegales, sin que se consideren como tales, las actuaciones acordadas por la autoridad de acuerdo con la ley, que solo limitará el derecho en cuestión por razones de interés público.

— En cualquier caso, en el FJ 6, se destaca el sistema de cautelas que impone la legislación a quienes acceden a estos datos y las sanciones que pueden imponerse a los que revelen los mismos (FJ 9). Ese sistema de cautelas, garantizaría el derecho del recurrente a que no se revelarán los datos obtenidos, en ningún caso.

STC 197/91

• Asunto: supuesta violación del derecho a la intimidad al afirmar un medio de comunicación que la madre natural del hijo adoptado por Sara Montiel, trabajaba en una barra americana.

• FJ 3:

— El derecho a la intimidad está estrictamente vinculado a la dignidad de la persona que reconoce el artíc. 10 CE e implica la existencia de un ámbito propio y reservado frente a la acción y conocimiento de los demás, necesario para mantener una calidad mínima de vida. (Reitera el mismo concepto que en la STC 231/88, sobre la distribución del video sobre la muerte de Paquirri).

— A diferencia de lo que ocurre en el derecho al honor, en las cuestiones relativas a la intimidad, no tiene cabida la exceptio veritatis.

STC 143/94.

• Asunto: NIF.

Insiste en la idea de la STC 110/84, básicamente en que los derechos fundamentales no son ilimitados y en que la supuesta invasión en la intimidad por la acumulación de datos en poder de la Administración, tiene su contrapeso en el fuerte sistema de cautelas, garantías y en su caso sanciones al funcionario que revela la información, que garantizan la opacidad de la información y el destino exclusivo a los fines para los que se reclama.

2. *Manifestaciones*

• Domicilio:

STC 22/84:

• Asunto:

Acuerdo del Ayuntamiento de Murcia ordenado la entrada en un inmueble para su derribo, sin mandamiento judicial amparado por Sentencia Judicial firme.

Sobre este tema es conocida la doctrina posterior y el cambio de criterio del TC (las cito con detalle en el proyecto de sentencia remitido a V.E. en el recurso 709/91). Sin embargo esas resoluciones posteriores no añaden nada nuevo respecto a la noción de intimidad que se contiene en la primeramente citada. (Solo la STC 137/85, extiende la protección a las personas jurídicas).

• FJ 5:

Tras delimitar el concepto de domicilio, añade que garantiza la privacidad de su titular, en un ámbito limitado que la propia persona elige y que tiene como característica quedar exento o inmune a las invasiones o agresiones exteriores de otras personas o de la autoridad pública.

Es un espacio donde el individuo vive sin estar sujeto necesariamente a los usos y convenciones sociales y ejerce su libertad más íntima. No solo protege el espacio físico en sí mismo considerado, sino que en él hay una emanación de la persona y de su esfera privada. Impone una extensa serie de garantías y facultades en las que se comprenden las de vedar toda clase de invasiones, incluidas las que puedan realizarse sin penetración de directa, por medio de aparatos mecánicos, electrónicos u otros análogos.

- Supuestos de sujeción especial:

- Presidiarios:

STC 73/83: ponente Diez de Velasco

- Asunto: La Dirección de la prisión de Herrera de la Mancha ordenó la suspensión de las comunicaciones de los internos y sus abogados, así como la intervención de su correspondencia.

- FJ 7:

— Las comunicaciones de los internos han de realizarse con el máximo respeto a la intimidad y solo pueden ser suspendidas por orden de la autoridad judicial con carácter general, si bien en los supuestos de terrorismo además podrá acordar la suspensión el Director de establecimiento dando cuenta a la autoridad judicial.

— No elabora un concepto de intimidad aunque se refiere a ella.

STC 89/87: ponente Rubio Llorente.

- Asunto: Derecho de los presos a mantener relaciones íntimas en prisión.

- FJ 2:

El mantenimiento de relaciones sexuales por un interno en un centro penitenciario, forma parte del derecho a la libertad a secas, pero no del derecho a la intimidad.

Es inherente a la privación de la libertada la limitación en la privacidad sexual. Con ello no se viola ningún derecho fundamental.

STC 120/90: La STC 137/90, confirmó el criterio.

- Asunto: Alimentación forzosa de terroristas del GRAPO en huelga de hambre. Dos votos particulares cuestionan el debilitamiento del derecho al suicidio por razones de sujeción especial.

- FJ 12:

La intimidad corporal garantiza la inmunidad frente a toda indagación o pesquisa que sobre el cuerpo quisiera imponerse contra la voluntad de la persona. Al ser alimentados forzosamente, ésta no se viola, atendiendo a los medios empleados, la parte del cuerpo afectada y la finalidad perseguida. En definitiva, al tratarse de reclusos, mantienen con la Administración una relación de sujeción especial que justifica la debilitación de un derecho, en este caso a la muerte por voluntaria inanición, que tendrían si estuvieran en libertad.

• protección de la integridad corporal:
STC 37/89:

• Asunto: mujer acusada de aborto se niega a obedecer la orden de un juez que en resolución no motivada la somete a una prueba pericial médica para determinar si provocó un aborto.

• FJ 7:
La CE garantiza la intimidad personal, de la que forma parte la intimidad corporal, de principio inmune frente a toda indagación o pesquisa que sobre el cuerpo quisiera imponerse contra la voluntad de la persona, cuyo sentimiento de pudor queda así protegido por el ordenamiento en tanto responda a criterios arraigados en la cultura de la comunidad. La protección de la intimidad reclama que la decisión judicial que ordena su restricción, se adopte en forma adecuada, con motivación suficiente, y se aprecie razonablemente por la autoridad actuante la relación entre el sujeto y la adopción de la medida.

3. *Las tensiones del derecho a la intimidad en relación con otros derechos:*

— Esencialmente se plantea con relación al artíc. 20 CE, que protege la libertad de expresión.

— Como conclusión de las diferentes sentencias analizadas (STC 20/92, difusión del padecimiento del SIDA por un arquitecto de Palma de Mallorca, 197/91, antes citada sobre el hijo adoptivo de Sara Montiel), y otras resoluciones esenciales como las SSTC 231/88, sobre la difusión del video de la muerte de Paquirri, 171/90, sobre revelación de aspectos de la vida sexual del piloto del avión siniestrado en el accidente de Sondica, 241/91, sobre revelaciones periodísticas respecto de las relaciones extramatrimoniales del ex-magistrado Lorenzo Penalva, se extrae la conclusión de que si bien en una sociedad democrática debe primarse esencialmente el derecho a la información en la medida en que contribuye a la formación de la voluntad de la propia sociedad, ese derecho alcanza su límite, cuando se pretende la difusión de noticias, aún ciertas, sobre hechos íntimos de la vida de personajes públicos que nada aportan a la formación de esa voluntad social y que se vierten con carácter innecesario. Como dice en su inciso final la STC 20/92, el artíc. 20 de la CE no protege la curiosidad ajena.

— Un caso singular de colisión de derechos es el resuelto por la STC 170/87. (El TEDDH, resolvió otro caso similar, en el mismo sentido. Asunto Sutter).

— Un camarero fue despedido por dejarse la barba, lo que a juicio del empresario perjudicaba la imagen higiénica del establecimiento.

En el FJ 4 de esta resolución, se dice que la Magistratura, en el marco de sus atribuciones, calificó como legítima la decisión empresarial de obligar al recurrente a afeitarse. Al no cuestionarse la decisión personal sobre la apariencia física del recurrente, sino la legalidad de la orden empresarial que condiciona la apariencia externa del sujeto, el tema planteado carece de relevancia constitucional.

A) En el Tribunal Supremo Español

No son muchas las resoluciones dictadas en elación a la preservación del derecho a la intimidad.

En relación con el tema que se solicita, me parece interesante la de fecha 13 de marzo de 1989, az 2040, ponente Latour Brotons.

— Una atleta que se somete voluntariamente a una prueba biológica, ante las dudas planteadas acerca de su sexo. La Federación Española de Atletismo difunde públicamente el resultado de los análisis, que detectaron una malformación cromosómica.

En los FJ se indica que forma parte del derecho a la intimidad, la preservación de los datos analíticos y biológicos de las personas. Mantiene que ese reducto de intimidad fue violado por la difusión de la noticia, ya que podía haberse acordado la baja de la atleta del equipo nacional, con mayor discreción. Prima en este caso el derecho a la intimidad sobre la libertad de información.

La primera Sentencia del TS en esta materia fue la de 28 de octubre de 1988 (caso Paquirri), az 6015, ponente Serena. Es interesante por su argumentación en relación con el límite que la persona pública tiene en relación con el derecho a la intimidad. Fue anulada por la STC 231/88.

B) El Tribunal Europeo de Derechos Humanos

Si bien no he consultado personalmente las resoluciones debido a la escasez de tiempo, en el libro «El derecho a la protección de la vida privada en la jurisprudencia del TEDH», 1994. (Sig. biblioteca 22104) se hace una detallada referencia al concepto y particularidades en la jurisprudencia del Tribunal. No consta en el libro ninguna referencia a pronunciamiento expreso sobre temas de ingeniería genética.

Se cita en primer lugar el Convenio 108 del Consejo de Europa 1981, ratificado por España y publicado en el BOE el 15 noviembre de

1989, relativo a la protección de la intimidad amenazada por el almacenamiento de datos. También se cita como antecedente jurisprudencial Europeo la Sentencia del Tribunal Constitucional Alemán de fecha 15 de diciembre de 1983, sobre la elaboración del censo.

Entre otras resoluciones del TEDH, que con detalle se traen a colación en el proyecto de Sentencia remitido a V.E núm 709/91 (casos Funke, Chappel, Niemitz), se cita especialmente el caso MARKX, para concluir que si bien el objeto del artíc. 8 del Convenio Europeo es esencialmente la protección del individuo frente a interferencias arbitrarias de los poderes públicos, el Estado no solo debe abstenerse de esas injerencias, sino que está obligado a una conducta positiva para garantizar el respeto a la vida familiar.

Este cambio de criterio del TEDH, fue posteriormente modulado en las resoluciones ABDRILAZIS Y REES. Se trataba de determinar si el individuo que ha cambiado de sexo, tiene derecho a borrar de los registros públicos, cualquier referencia a su pasado biológico. En definitiva, tras condenar al Estado infractor se le deja a este libertad para paliar la situación conforme considere más oportuno con arreglo a su propia legislación.

C) En la Jurisprudencia Norteamericana

Dado que el concepto de privacidad tuvo su primera exposición en el derecho de ese país («The right to privacy» por Brandeys y Warren, en *Harvard Law Review*, 1890) y que la tecnología por ellos desarrollada hace razonable pensar que las primeras resoluciones expresas se producirán en los Estados Unidos, me pareció oportuno tratar de conocer la situación de dicha cuestión.

1. En el libro «The Code of Codes», sig. Biblioteca 21313, desde un punto de vista no jurídico, se expone la realidad de la ingeniería genética, los conceptos básicos, y los problemas que empieza a plantear jurídicamente a la sociedad americana.

Tras analizar las aplicaciones de dichas técnicas, con especial referencia al ADN y al PET, se pregunta, no ya si terceros pueden acceder a esa información, sino si es lícito darla al propio interesado, en la medida en que el ser humano se convierte en «un hombre de cristal», siendo impredecibles las reacciones del sujeto ante la lectura de lo que va ser casi con toda seguridad su vida futura. El diagnóstico anticipado de enfermedades incurables, el saber como se va a proceder ante determinados estímulos, conocer la propensión del individuo al consumo de drogas, etc. El conocimiento anticipado de esta información y la posibilidad de ser exigidas en el marco de la contratación laboral, o en la suscripción de pólizas de seguros, plantean agudos problemas, y la tensión entre el derecho del individuo a que se respete su intimidad, que incluso él mismo pudiera querer ignorar y las exigencias de quien al contratar quiere hacerlo sobre la base de la mayor información posible. Su utilización en el derecho penal, pruebas de ADN, y la obligación del encartado de someterse a las mismas, plantea también enconados problemas. Asimismo en las decisiones sobre el aborto eugenésico. Se refiere el citado libro (pág. 173) a una serie de estudios jurídicos, encargados al National Research Council-National Academy of Study on DNA fingerprinting, elaborados por jueces y abogados de los Estados Unidos.

2. En el artículo de Jed Rubenfeld, publicado en la Revista de Derecho de Harvard, febrero 1989, número 4, se analiza la evolución del concepto de «privacy» en la jurisprudencia del Tribunal Supremo.

Advierte el autor que a partir del caso GRISWOLD V. CONNECTICUT (381 US 479 1965), se elabora un nuevo concepto de «privacy». Pasa de ser un límite a la facultad de un tercero para obtener o difundir una información relativa a un sujeto particular, a conceptuarse como un derecho sustantivo de carácter reaccional frente a las injerencias de un Estado que trata por la vía de la coacción imponer un estilo de vida o de modelar la conducta de los individuos en sociedad. En este concreto caso, prohibiendo la venta de anticonceptivos. Como se indica en dicha resolución, el derecho a la intimidad incluye la libertad de las parejas casadas de decidir por ellas lo que deben hacer en la intimidad de su habitación.

Según el citado autor, la «Privacy», es el derecho fundamental de no tener una vida determinada por una progresiva y lógica intervención estatal. La concibe como un concepto político que garantiza el último reducto del ciudadano referente a un Estado que cada vez aparece como más totalitario. El concepto difiere de la concepción alemana en la que la intimidad se analiza desde el ámbito de la invasión de determinadas esferas donde el individuo desarrolla su vida privada.

Cita como antecedentes jurisprudenciales los casos MEYER Y PIERCE de 1923 y 1925, relativos al derecho a la educación, (anula leyes que prohibían la educación en lengua extranjera o en que imponían la asistencia a centros públicos, en determinada edad).

D) CONCLUSIÓNS

De las restantes lecturas efectuadas (X. O'Callaghan: La Libertad de expresión y sus límites: Honor, intimidad e imagen. Pablo Salvador

Coderch: el Mercado de las ideas. José Martínez de Pisón: El derecho a la intimidad en la jurisprudencia constitucional) y especialmente de la doctrina del Tribunal Constitucional, destaco los siguientes puntos:

1. Existe una gran dificultad en definir el concepto intimidad. Tanto en las definiciones doctrinales como jurisprudenciales suele incluirse una referencia a lo íntimo, en clara petición de principio. Se tiende a dar una definición negativa del concepto.

2. Parece subyacer como idea central, la famosa frase del juez Cooley «The right to privay is the right to be let alone», el derecho a la intimidad es el derecho a ser dejado solo.

3. Tras remitirse a la doctrina general de carácter relativo de los derechos fundamentales, justifica el Tribunal Constitucional la pérdida de la intimidad del particular que se ve obligado a suministrar datos a la Administración, por el rígido sistema de controles y garantías que legalmente se establece. En definitiva se trata de que esa información no se utilice fuera del ámbito para el que se solicitó y que en cualquier caso no se someta a la publicidad gratuita. En este sentido debe citarse la legislación española sobre protección de datos informáticos de 29 de octubre de 1992, y la que en las sentencias 143/94, y 110/84 se citan, en relación a las concretas materias que se desarrollan: NIF y cuentas corrientes (OM 14 enero 1978, ley 50/77).

4. La colisión entre la intimidad y otros derechos, debe resolverse mediante una resolución ponderada, con arreglo a los valores culturales rectores de la sociedad en el momento concreto en que se plantee el conflicto, prestando una especial atención al fin perseguido. Se destaca la idea de que la exceptio veritatis no puede oponerse como causa de justificación, cuando la revelación de datos íntimos de un tercero se haya hecho de forma innecesaria, bien por el modo en que se expresó o por la falta de interés social el contenido de la información.

El núcleo de la cuestión se plantea entre la protección de lo íntimo y la demanda de información por el Estado o por un tercero para su propio uso, cuando se alegue para ello un fin que pueda entenderse constitucionalmente legítimo. Con especial intensidad aparece el conflicto, en los supuestos de vinculación entre el individuo y el Estado por relaciones de sujeción especial, si se examina a la luz de la doctrina de este Tribunal que justifica un debilitamiento de los derechos fundamentales en esos casos.

La solución que se de, será fruto de la concepción política vigente, en el sentido más amplio de la expresión, conectando con la, en mi opinión, atractiva idea de Rubenfeld, de que tras la intimidad se esconde la idea de lucha política individual, frente a una presencia cada vez mayor de un Estado con ansias totalitarias en el seno de la sociedad democrática.

En cualquier caso destaco la cautela que la STC 110/84, extremó al justificar la invasión de la intimidad, resaltando que la Administración solo iba a conocer los movimientos globales de las cuentas corrientes, pero no las concretas partidas; así se justificaba la preservación del derecho a la intimidad. Por contraste, no cabe duda que la información genética sitúa al individuo en una situación de absoluta-transparencia frente al Estado.

Si como Garrido Falla, decía, según el FJ 6 de la STC 110/84, «una cuenta corriente es la biografía personal en números», con fundado temor la doctrina califica de «hombre de cristal», al ser humano frente al manejo de la información genética.

THE HUMAN GENOME PROJECT AND THE THIRD WORLD

SANTIAGO GRISOLIA *

Following the Third International Workshop dedicated to Cooperation for the Human Genome Project in which I had a major role, and which was dedicated to legal aspects, I had the opportunity, thanks to the generosity of the Pontificial Academy, to discuss some of the general aspects of the so-called Human Genome Project here about a year ago. I will be happy to send to anyone, if this has not been published, the remarks I presented at that time. Because of this, and also because I saw in the Program such prestigious participants as Professors Singer, Cavalli-Sforza, Weissenbach, Helena and Perutz, I thought to restrict my remarks today to the topic of Gene Therapy. As I indicated in my final resumé last year, this is an area of much interest both because of its revolutionary medical implications and because it will serve as a paradigm of the close interactions between scientific development, ethics and legal aspects of the Human Genome Project. Since I am Chairman of the UNESCO Scientific Coordinating Committee for the Genome Initiative and also a member of the UNESCO Bioethics Committee, I shall also stress some of these implications and will try to clarify possible misunderstandings on the topic of Gene Therapy, with special emphasis on the Third World. Since this is a mixed audience I will avoid technical terms as much as possible.

Genes are the portions of DNA which contain the instructions which give origin to the structure and function of our bodies. In other words, genes regulate and are responsibile for the physical characteristics of our body height, color of our skin, eyes, hair, etc. Of course, genes, normal or abnormal, are the patrimony which parents transmit to their children. Changes of genes can produce alterations in the physical characteristics of functions, which on occasion result in disease.

---

* BBV - Foundation. Fundación Banco Bilbao Vizcaya, Spain.

This inheritance, which up to now was immutable, can now begin to be controlled. Indeed, it is already possible to correct some of these changes which result in undesirable effects by introducing a normal gene in an organism to replace the function of a defective gene: this is what is known as Gene Therapy. Without doubt, in the relatively near future, the possibility of changing human heredity via germ line therapy may be considered and perhaps initiated.

1. Gene Theraphy

The term 'human gene therapy' came into use because it identified, as presumptively beneficent, technologies that might have provoked more opposition if called human genetic engineering, particularly in the 1970s, when any development or means for altering the genetic capacity of human cells was hotly debated and bitterly opposed by some. The question, if asked, "What are the good uses of human genetic engineering?", might have produced a negative response. Human gene therapy is and sounds benign.

On September 14th 1990, a little girl four years old, with a deficiency of adenosine deaminase (ADA), was injected with 10^9 of her own T cells in which had been inserted a normal ADA gene. This was done in the Childrens Unit of Intensive Care of the Clinical Center of the National Institute of Health (NIH) in Bethesda, Maryland. That was the beginning of gene therapy in humans. Although gene therapy is a rational and logical treatment for certain types of diseases in humans, until rather recently the use of this type of technology was criticized as something very far in the future and almost science fiction. As a consequence, there was very extensive debate before the protocol for ADA deficiency was approved. Nevertheless, once it was shown to be possible, most people were favorably impressed. Thus, in spite of the few people who suffer from ADA deficiency in the whole world, the protocols used to correct this deficiency played a key role in the demonstration of the feasibility of gene therapy, and therefore established the first step towards a new Medicine. The Advisory Committee of the NIH (RAC) has approved some 80 protocols for several diseases, whereas the first gene therapy protocol of Anderson in 1989 for ADA was reviewed fifteen times by seven different regulatory bodies. A mere five years later most protocols can be approved on a standard FDA review. Also Great Britain, France, Italy, Holland and China have started to use some forty protocols in quite a number of patients, as of September 1994. Interestingly the majority, $\approx 80\%$, are for cancer, or cancer related protocols.

This is not surprising since, indeed, it has become evident in the last two years that in addition to the approximately 4,000 monogenetic diseases known, potentially all diseases have a genetic origin. That is to say that many diseases have a polygenic origin, for example, a large portion of persons affected with diabetes, and that the predisposition to a number of diseases also has a genetic basis. In all likelihood, a great deal of emphasis in the future will be given to treat the more common diseases, and precisely by their being more common they will be of more interest to society and thus to industry. This will have great benefits for many people including those of the Third World, but may carry the paradoxical danger of postponing or forgetting the treatment of certain monogenic diseases, particulary the less common ones, although these were the ones which originated gene therapy!

It has been pointed out (Jolly, 1994) that it is important to be realistic about the state of the development since apparently attractive factors of less characterized systems may have to be forgotten during the serious business of making useful, clinically acceptable vectors. Examples of this include the decreases in titers of many vectors as the inserted genes become more complex, and the lack of consistent, insertional targeting specificity previously claimed for adeno-associated viruses (AVV) as a vector.

It is important to realize that there are several precedents for related forms of therapy. The oldest of these is the vaccinia vaccine for smallpox. Allogeneic organ transplants and attenuated viral vaccines such as polio, measles, mumps, and rubella all involve transfer of "foreign" genetic material into patients.

2. Types of Gene Transfer

There are two main types of methods for gene transfer: those based on viral vectors and those based on the introduction of DNA into cells or tissues by chemical or physical means. These include the direct injection of DNA, or the formation of DNA-protein complexes with affinity for specific cell receptors. The methods which do not use viruses have, thus far, a low transfer capacity, and/or result in rapid degradation of the DNA. While the viral systems are more effective, the main concern regarding the use of viral systems is safety. The design of the majority of the viral vectors is done in such a way as to minimize the possibility of transfer of the complete virus or that of gene recombination which could give rise to the complete virus.

A. *Viral Systems*

They can be divided into those which adapt their sequences to the cell genome of the recipient in a stable form and those which remain as episomes. The genomic integration has the advantage that the therapeutic genes are inherited by subsequent cell generations. However, they require cell division for their integration. The prototypes are the retroviruses. Retroviruses are diploid positive-strand RNA viruses that replicate through an integrated DNA intermediate. The incoming RNA of the virus is reverse-transcribed into DNA by a virally encoded reverse transcriptase that is carred as a protein in each viral particle. The viral DNA is integrated pseudo-randomly into the host genome. The normal cellular machinery is then utilized to make RNA copies of the genome and a spliced version that encodes the retroviral envelope. The produced RNA encodes the major gag, pol, and env proteins which are all part of the viral particle. The virus then assembles around the two RNA genomes and buds off from the cell carrying a piece of the outer lipid membrane and embedded envelope proteins.

Among the viruses which remain as episomes are the adenoviruses which have the advantage of efficiently infecting non-dividing cells. However, since they do not become part of the cell genome, and since they have been modified to inhibit replication, they are eliminated when the cell divides. Because of the above considerations, a great number of gene therapy protocols have used retroviral vectors. A great deal of effort has been spent in the construction of these vectors so that they will not be able to complete themselves; otherwise they might cause a number of untoward effects such as tumors. The easy and more used have been the murine leukemia virus (MVL) and similar viruses. In these viruses, the packing of RNA in virions is regulated by one sequence of the viral genome known as the packing signal psi (ψ). Thus the retroviral vectors need the presence of the packing sequence psi (ψ) for encapsidation but permit the elimination of other genes necessary for the formation of the viron. To make possible the packing of the RNA and to form a viron these genes have to be offered by another genome of a packing cell line. To diminish recombination, the homology between codifying sequences of the packing cell line and that of the psi of the vector is reduced as much as possible.

Also the sequences which codify the viral proteins have been separated in two different plasmids in order to reduce even more the possibility of viral recombination. All these precautions have made the retroviral vectors very safe instruments for gene tranference. Approximately 100 patients have been treated so far and no significant treatment-related adverse events have been noted. This promising initial clinical experience forms a valuable foundation for further trials.

It is worth noting that building retroviral vectors is still a mixture of art and science. Many creative ideas require multiple design attempts before performing anywhere close to that desired.

As already said, retroviral vectors are not useful for transduction of cells which cannot replicate. In these cases, however, DNA viruses are very useful. Thus for example, DNA viruses have tropism for lung cells and can be used for diseases of the lung.

Adenoviruses are a family of viruses (over forty serotypes) that cause benign respiratory tract infections in humans. The major incentives to develop adenoviral vectors included the promise of higher titers than retroviral vectors, the expected benign clinical profile (based on its long-term use as a vaccine for adenovirus respiratory tract infections by the US Army), and its likely ability to result in high levels of transgene expression. There are many strains of adenovirus that infect humans but strains 5 and 2 from the C subgroup are the ones that have been predominantly used to make vectors because their molecular composition is the best characterized. However, the clinical experience with adenoviral vaccines refers mainly to adenoviruses of strains 4 and 7 (Jolly, 1994).

Nevertheless, the use of vectors of DNA virus in humans has several potential problems. It is not clear if gene transference with this vector will be effective in persons who have been exposed to a previous adenoviral infection. Perhaps to make effective successive innoculations, it will be necessary to use immunosuppressors, which in some diseases such as cystic fibrosis could make matters even worse. It may possibly be necessary to repeat the administration of the recombined genes to compensate the episomic instability in the cells.

Another type of virus which has been studied is the herpes simplex which has a tropism for the CNS. Other viruses used are the virus adeno-associated (VAA). The VAA are parvoviruses, not pathogenic, without autonomic capacity of replication and which integrate specifically in the cell genome. The native VAA integrate normally in the distal portion of the large arm of chromosome 19, but it can be made unspecific via selection to resistance to neosporin. At any rate, this is an area in which although much is known, much still needs to be done.

B. *Non-viral Systems*

Some of these systems are based on the synthesis of DNA-protein complexes which may transfer the needed genes to a number of cells via endocytosis. These complexes can suffer degradation by lysosomes. However, because these complexes cannot escape the cell vesicles, their efficiency is limited. Nevertheless, since many viruses which enter the cell using their receptors have tbese mechanisms, systems have been worked out using co-internalization of adenovirus and DNA-protein conjugates. These systems are much more effective, nearly 10 times, and they are under extensive investigation. The direct injection of DNA as a plasmid has been studied and it appears that heart and skeletal muscle may be treated in this way. However, the main problem in humans is the difficulty in maintaining a long-term expression.

3. Applications

A. *Deficiency of adenosine deaminase*

This has already been discussed. This deficiency is due to the accumulation of products of degradation of DNA, which act as substrates for AD. In absence of AD they destroy T and B lymphocytes. Thus far the therapy with retrovirus is only a palliative and temporary treatment which necessitates periodic administrations. However, at present there are protocols using pluripotent stem cells which may really cure the disease.

It should be noted, as already indicated, that although the retroviruses are very good vectors, their integration into the genome of other cells requires that they be replicating. This has been a major problem for the use of bone marrow, because its stem cells are usually at rest: nevertheless, some success has been obtained and two children have been thus treated with their own modified stem cells. However, as already pointed out, it is important to stress that the literature sometimes does not give sufficient emphasis to the fact that many of the treatments thus far tested require repeated administration. Often commentators, particulary non-scientific, assume or give the impression that the therapy "cures" forever. This might well happen with germ line techniques, if ever used as we will discuss later on, but certainly unless the treatments with stem cells are successful, and keep producing modified cells for cell replacement, all protocols thus far have necessitated repeated treatments.

B. *Cancer*

Among the variety of procedures directed to the treatment of cancer, it is important to mention the technique known as "adaptive immunotherapy", that is, the transfer of genes to the immunological system to stimulate its function, or the transfer of genes to tumor cells to facilitate their identification and to stimulate rejection. Other strategies are directed to restoration of the normal phenotype in the tumoral cells via the transference of suppressor genes or the improvement of bone marrow transplant for lymphoid tumors. The most recent treatment is the induction of cell susceptibility via the use of the misnamed suicide genes.

(1) *Adaptive Immunotherapy.* These procedures are directed to stimulate cell immunity and are based on the insertion in tumor cells or in lymphocytes of cytokine genes. For example, cytotoxic T lymphocytes (CTL), stimulated by IL-2, will fight foreign tissues. This technique using LAC cells has been used in melanoma patients and has been partially successful. In a few cases there was complete regression, particularly when the T lymphocytes were obtained from the tumor (tumor infiltrating lymphocytes). The use of the tumor necrosis factor (TNF) has a significant effect on tumors. Unfortunately, the dose needed is toxic for humans. Nevertheless, the attempt has been made to obtain high concentrations in the tumor using lymphocytes from the tumor treated with TNF. Another type of immunotherapy is based on modification of tumor cells to produce cytokines or to express surface antigens. It has been used for melanomas and carcinomas from colon and mammary glands. They have used IL–2, –4, –5 and –6 and the granulocyte stimulating factor CD–2 and interferon. All these treatments are in progress and they appear to be more effective in primary tumors than in metastases.

(2) *Therapy with Tumor Suppressor Genes.* This treatment is theoretically possible although difficult and far in the future. The restoration of the suppressor gene of retinoblastoma (RB) has been tested experimentally in this tumor and in prostatic carcinoma, and the p53 in mammary carcinoma, in hepatomas and in T cell leukemia. Although the technique inhibits cell growth, little is known about the effect *in vivo* on a tumor already established. There are other difficulties because of the complex character of tumor proliferation due to the implication of other genes (oncogenes) and the need that the transfer will affect all tumor cells.

(3) *Cell Marking in Bone Marrow Transplant.* Gene transfer can be very useful to follow cells *in vivo.* Although bone marrow transplant is very useful, there are problems, such as the difficulty in obtaining histocompatible donors, and this has necessitated using alternatives to the autogenous bone marrow, hopefully free of cancer cells, but rarely so, for some hematological conditions. Here is where the marking could be useful to avoid the transplantation of bone marrow still containing tumor cells, to separate stem cells or to locate the origin of a recurrency. There are several protocols that use marking genes for myeloblastic leukemia, other types of leukemia, neuroblastoma, multiple myeloma and metastatic breast cancer.

(4) *Induction of Cell Susceptibility*. This is know by the unfortunate title of the use of suicide genes. The procedure first tried consists in the introduction in tumor cells of the thymidine kinase gene from the herpes virus. The protein codified by this gene makes the cells which contain it susceptible to gancyclovir. The first application in humans has been the treatment of brain tumors. Due to the absence of proliferative activity in the normal brain tissue, the retroviral vectors will be integrated only in the genome of the tumor cells. Interestingly, there is the so-called by-stander effect, whereby even the tumoral cells in the vicinity of those infected by the virus are also destroyed. Several variations of this technique have been proposed using cytosine deaminase and 5-fluorocytosine. This latter combination results in the production within the tumor cell of 5-fluorouracil, a well known agent in cancer chemotherapy. Another combination uses a gene derived from the varicela zoster virus and new anti-herpes drugs.

C. *Cardiovascular Diseases*

Genes can be transferred *in vitro* to the cells of the circulatory system, both endothelial and those of the smooth muscle. The latter which are in a larger number, have the peculiarity that their proliferation can be regulated *in vivo* via the autocrine or paracrine production of growth factors. The modified cells may be implanted later on *in vivo* either in a vascular transplant or in a native artery so that the proteins synthesised by them may enter the circulatory system.

A recently proposed and approved procedure for gene therapy would employ the gene for the vascular endothelial growth factor (vegf) to stimulate the growth of collateral vessels, in effect a by-pass, in the area of an arterial blockage supplying blood to the leg. A tiny angioplastic balloon coated with vegf genes would be inflated near the arterial blockage where the genes should be taken up by the cells of the arterial wall which should produce vegf proteins, and in turn induce new blood vessels to form aorund the blockage. And there is the possibility, if it works, of developing it for blockage of arteries of the heart.

As pointed out recently (Ohno *et al.*, 1994), injury of the arterial wall increases the production of molecules that stimulate smooth muscle cell migration and proliferation, which may lead to intimal hyperplasia, which contributes to the pathogenesis and atherosclerosis of some other cardiovascular disorders. In other words, there may be a reactive cellular proliferative response with local regrowth of cells which will diminish blood flow. This process, called re-stenosis, has resisted conventional treatments such as antiplatelet agents, angiotensin-converting enzyme antagonists, or cytotoxic drugs in humans. To limit cellular proliferation at specific sites via gene therapy in the arterial wall could help to understand the pathogenesis and the rational treatments of vascular proliferative diseases. As already mentioned for the selective elimination of dividing cells, an interesting approach is to express a recombinant gene for thymidine kinase (tk), which converts the nucleoside analog gancyclovir into an active toxic form in transduced cells. The tk enzyme phosphorylates gancyclovir *in vivo* and its subsequent incorporation into DNA, induces chain termination in dividing cells, causing cell death.

This so-called prodrug approach, in which high levels of recombinant tk gene are expressed locally, seems a good and promising procedure. This method provides for sustained expression and conversion of the prodrug to its toxic form locally at the time of peak cell proliferation at concentrations that cannot be achieved by drug delivery. In other words, local delivery of an antiproliferative agent during the peak of smooth muscle cell division after balloon injury might limit intimal hyperplasia. Failure to accomplish this with cytotoxic drugs is probably due to rapid removal of the drugs by the arterial circulation.

D. *Hematological Diseases*

The manipulation of stem cells of the bone marrow is essential to correct some hematological problems, for example, thalassemias or when one requires a large amount of enzyme or protein for therapy. Since thus far gene transfer to stem cells of bone marrow is very difficult, the co-expression of the gene of interest with a gene for resistance to multiple drugs (MDR) has been tested. This gene codifies a protein which is able to pump out from the cell a number of drugs which are toxic. There are some experiments using Taxol in murine models, which need to be evaluated and extended before testing it in humans. The hemoglobin alterations present very serious problems in finding efficient vectors.

E. *Lung Diseases*

There are two of these which are of great current interest: cystic fibrosis, which is the most common monogenic disease among us, and the deficiency of antitrypsin (AT). The main difficulty for the latter is that it is necessary to obtain a blood level of approximately 2 mg/ml to avoid accumulation of elastase in the neutrophils in the lung fluids which results in a progressive deterioration of the lung epithelium. As for cystic fibrosis (CF), the problem resides in the genetic codification of the transmembrane conductance regulator. The type of the lung epithelium is not yet known in which the gene is normally expressed or how to carry on a selective transference. Because of the slow replacement of the respiratory epithelium, several vectors derived from adenoviruses have been used instead of retroviruses. Thus the genes for CF and AT have been transferred to the bronchial epithelium using adenovirus. Although there are some problems and unknowns regarding concentrations and length of treatments, the CF in humans is now being treated. The first protocol used bronchoscopy but the possibility of using aerosols is also being tested. Cationic liposomes have been also tested in humans.

The recent work of Crystal and colleagues, as pointed out recently (Alton and Geddes, 1994), marks an important step in the progress of human gene therapy trials for cystic fibrosis. The current life expectancy of CF patients is approximately thirty years, and the main problems of morbidity and mortality result from recurrent respiratory infection. Pathology typically originates in the more peripheral airways of CF lungs.

In theory, several problems may occur following the use of adenoviruses, including the induction of cellular immunity, recombination to produce a replication-competent virus and the production of neutralizing antibodies reducing the efficacy of further applications. In their study of four CF patients in whose lungs varying doses of CFTR were administered, Crystal *et al.* found no evidence for recombination and no evidence for serum neutralizing antibodies following this single application.

More extensive deletions of the adenoviral genome are also likely to help. The first generation of recombinant adenoviruses (Crystal, 1994) included deletions of sequences spanning E1a and E1b to prevent viral replication. Finally, cationic liposomes are probably relatively inefficient at gene transfer in comparison with viral vectors but they are less likely to provoke inflammation or to initiate an immune reaction on repeated administration. However, their complex physico-chemical properties with respect to combination with both DNA and the surrounding ionic environment are poorly understood, and this probably contributes to their variable transfection efficiency in gene transfer studies. Gene therapy for CF is moving rapidly, but gene therapy is still a young discipline.

F. *Liver Diseases*

This organ has an inherent advantage for gene transfer because of its extensive circulatory system. Although it has a great metabolic activity, it has a low proliferative capacity, thus gene transfer via retrovirus is low. In tissue culture, hepatocytes may divide two or three times which has permitted transfer with an efficiency of up to 30%. The first protocol for the liver was concerned with the treatment of familial hypercholesterolemia (FH). FH is an autosomic dominant problem caused by a defect in the gene for the receptor of the lipoproteins of low density (LDL). The absence of the gene results in the accumulation of cholesterol, of LDL, and then in atherosclerosis and coronary disease. The homozygote type cannot be treated by either diet or medication. After experiments in the Watanabe rabbit, primates were treated with a retroviral vector which included the DNA complementary for the normal gene of the LDL receptor plus a hybrid promotor to activate the transcription. Following these studies in primates, the protocols have been used in humans successfully.

G. *Acquired Immune Deficiency Syndrome (AIDS)*

Among the strategies used for AIDS there are protocols of gene transfer directed to block the viral infection and/or interfere with the expression of the genes. Although there have been many variations of this, including the use of the so-called suicide genes in CD8+ lymphocytes, the results have thus far been disappointing.

H. *Gene therapy and the Third World*

UNESCO is an institution concerned with securing for the peoples of all the world fair participation in the benefits that flow from scientific and technical advances. How will that be accomplished with technologies such as gene therapy? There is a strong claim to share the benefits from genetic information.

Many in the developing world, however, look on the genome project with conflicting views, particularly that aspect which contemplates genetic therapies, despite a claim to fair participation. Is it right to devote so many resources to therapies that are likely to be so expensive? Is it plausible that they will be affordable only for the wealthiest societies? How many people will they help? Are there relationships between these technologies and others so that public health may be more effective?

If we are correct that gene therapy has a chance to be a technology of widespread use, then the oft-expressed concern that the economically developed world is spending its medical resources wastefully in developing sophisticated procedures, is misguided.

It is true, of course, that the first uses of any complex technology are extraordinarily expensive. It costs a great deal to engineer a new automobile or the generation of a computer chip. The development of gene therapy is also complex. However, as with all technologies, costs drop with more use. Moreover, improvements in strategy and design may sharply reduce cost through improvement in product. Computers more powerful than any that existed in 1960 are now inexpensive and widely used. It is not unlikely that present gene therapies will be subject to comparable reductions in costs for the more common diseases within a reasonable time.

I. *Germline Gene Therapy*

Germline gene therapy has attracted a great deal of attention. Imagining a world in which some people — the state, physicians, parents — have authority to select the genetic characteristics of the next generation, to produce desired traits, provides occasion to reflect on the true nature of individual rights and human dignity.

All major statements about germline gene therapy condemn its present use. That position is clearly correct. That genes can be put into animal germlines, and made to express, does not begin to answer safety issues for human use. Moreover, there are no animal models that can predict the impact on the human brain. Enormous technical problems would have to be solved to make the technology realistic in light of the risks, particularly the control of gene expression throughout the process of cellular differentiation. Moreover, one must have a basis for predicting confidently the consequences of novel or altered genetic material in the workings of each and every cell types. So far as we know, there have been no efforts anywhere to attempt germ line therapy on human beings.

The prohibition of germline therapy is a matter of formal legislation in some nations (e.g. Sweden) and is accomplished through regulatory controls in many others, for example, Great Britain and the United States. Present prohibition does not deny the possibility of future use. Yet, there are important European documents that condemn germline gene therapy unequivocally. "Any form of therapy on the human germinal line shall be forbidden". (Council of Europe Recommendation 1100).

In the United States, a number of prominent commentators believe that discussion should begin about developing germline gene therapy. Germline gene therapy sounds most attractive as a technology for permitting a couple to spare their descendants the burden of genetic disease. So, for example, a person at risk for a disease caused by a dominant gene might seek the procedure to guarantee that his future children would be free of it. The question posed is that if it is moral to remedy his condition by somatic cell gene therapy, why not cure it once and for all by germline techniques?

The probable starting point for germline therapy is the human zygote at the four or eight cell stage, obtained through *in vitro* fertilization. Animal models suggest that gene insertion that will differentiate into every cell is feasible. But if a particular zygote can be identified as carrying an undesirable gene, and therefore be appropriate for treatment, why would one try the extraordinary procedure of repair rather than selecting for implantation a zygote that did not have the undesirable gene? Such screening procedures are in use, on an experimental basis, in a few hospitals around the world for cystic fibrosis.

There are many who strongly oppose *any* selection or discarding of zygotes, and for these persons, that opposition is itself a sufficient basis to protest against the development of germline gene therapy. Opposition to selecting zygotes goes hand-in-hand with opposition to non-therapeutic experiments on zygotes that will never be implanted. For example, German law prohibits non-therapeutic experiments on the human zygote. While all acknowledge the special respect owed to the fertilized egg throughout its development, what actions are required or prohibited by that respect is an issue on which the world community is deeply divided.

There may be couples in which each partner is homozygous for a particular recessive gene, say the cystic fibrosis gene. For this couple, no "normal" children are possible. There may be such couples, but the prospect of launching a technological enterprise so vast to serve them seems too implausible to justify relaxation of the prohibition on developing germline therapies.

However, it seems appropriate to consider whether there are sound reasons for categorically barring germline gene therapy. Do concepts of human dignity and respect for human life require categorical condemnation of germline therapy?

REFERENCES

This text represents largely an abstract and modification of the article «Terapia Génica: La nueva frontera» hy Rubén F. Moreno-Palanques and Santiago Grisolia, published in «Farmacología y Terapéutica»: Editorial Médica Panamericana, and the report of the Subcommittee on Human Gene Therapy of the International Bioethics Committee, UNESCO, September 1994, prepared by Harold Edgar and others.

Other references used were:

— Alton, E. and Geddes D., «A mixed message for cystic fibrosis gene therapy», *Nature Genetics*, 8, 8-9 (1994).

— Crystal, R.G. *et al.*, *Nature Genetics*, 8, 42-51 (1994).

— Jolly, D., «Viral vector systems for gene therapy», *Cancer Gene Therapy*, 1, 51-64 (1994).

— Ohno, T. *et al.*, «Gene therapy for Vascular Smooth Muscle Cell Proliferation After Arterial Injury», *Science*, 265, 781-784 (1994).

LEGAL PROTECTION OF VALUES BY THE HUMAN GENOME RESEARCH

CARLOS M. ROMEO-CASABONA *

1. Introduction

The massive investment made in the Human Genome Project and similar research programmes, together with the involvement of not just national governments but also — in some cases — bodies of the stature of UNESCO and the European Union, are indicative of the fact that man at the end of the twentieth century has embarked upon one of the most ambitious undertakings in the recent history of mankind. It is no coincidence either that the Nobel Prizes for Medicine and Chemistry in 1993 have once again recognised the efforts of researchers in the field of human genetics.

The applications and results of research on the human genome will afford a wealth of benefits for the health and well-being of mankind. Research of this kind also requires collective support from all quarters. However, at the same time, the discovery that genetic manipulations can affect the identity and the individual and unrepeatable nature of the human being, and can lead to the irreversible modification of the genetic heritage of the human species, its integrity and diversity, and the knowledge that such manipulations are an instrument of incalculable consequences for eugenic practices which we had thought had been done away with once and for all, now show the ambivalent side of this type of research, which is akin to a new two-faced Janus. I have had occasion to remark that New Information and Communications Technologies are opening up a new age: the information age. The knowledge afforded by human genome research

* Director, Chair in Law and the Human Genome, BBV Foundation - Provincial Government of Biscay, University of Deusto, Bilbao (Spain).

will have a multiplying effect on the characteristics of this new information age. However, it should be noted that whereas New Information Technologies basically enable vast amounts of information to be gathered, accessed instantly and processed readily — information which already exists —, the Human Genome Project goes much further since it creates knowledge, ie information on unreachable aspects of our very being, aspects of which we ourselves are ignorant; it also provides information concerning our collective origins and predictions of our likely future as individuals. Thus, although there is much truth in the old saying that knowledge is power, especially in the hands of someone who has access to such knowledge and can make use of it, there is little need to stress the fearsome tool which might result from the combination of New Information and Communications Technologies and knowledge of the human genome, a tool which would facilitate intrusion into the innermost lives of individuals and may lead to discrimination throughout the different walks of social life. I am convinced that we are on the verge of a new civilisation, the tools of which are the two formidable technological and scientific manifestations just mentioned, of which man will have to be both maker and designer. I say maker because wishful optimism moves us to dismiss the idea that he may become victim of himself. This, therefore, is our great collective and generation-to-generation responsibility towards civilisation, towards mankind: to pass on a better world in accordance with our own beliefs and values, but leaving future generations free to decide, without being conditioned previously, as to their best interests, in accordance with their beliefs, their values and their needs.

Given that human genetic research can affect all of society, it is only right that society, not just scientists, should discuss and decide what it is prepared to take on or to reject. Democratic society manifests itself through the Law, and it is through the Law that rules of behaviour are established by which we all agree to abide. However, unlike other fields of thought, the Law must provide a single answer to each specific case, and this necessarily involves a choice between possible alternative solutions and the rejection of those deemed to be undesirable or inappropriate in each situation. This is the task entrusted to legislators, who must base themselves on the Constitution and the ethical and social values in force at a given time, provided that these are in keeping with the former; judges must act in a similar manner when applying the Law. Occasionally, however, the Law is not equipped to offer clear-cut and precise answers to the new problems raised by uncontainable scientific and technological advances, which are occurring at such speed that they are difficult to assimilate, as is often the case with human genome-related research and applications. The Bilbao

International Meeting on "The Law and the Human Genome Project" held in 1993 in the University of Deusto, under the auspices of and sponsored by the BBV Foundation, in conjunction with the Diputación Foral de Bizkaia (Provincial Government of Biscay), highlighted and discussed the most pressing current and short-term legal problems: the repercussions on individual freedom and privacy and on other fundamental rights; the implications in the field of insurance and labour relations; discussion on the patentability of human genes and sequences; striking a balance between the right to engage in research and the rights of people, as well as the limits to genetic manipulation in relation to the aforementioned values; the juridico-procedural relevance of so-called "genetic fingerprinting" and biological paternity tests as a means of identification, etc.

For the above reasons it is necessary to promote reflection and debate by society in parallel with the stream of discoveries and applications derived from human genome research. This reflection must first and foremost be multidisciplinary, that is, it must be undertaken jointly with scientific researchers in Molecular Biology, Medicine and other related sciences, but also with the clinicians whose job it will be to apply the results. Secondly, it must be interdisciplinary, through the pooling of efforts by experts from the various fields of Law most directly involved, by scholars of Bioethics and other fields of knowledge who are in a position to enrich the dialogue and reflection, and by those sectors of the population which have a legitimate interest in the subject. Cooperation and international debate are considered nowadays to be the only path towards uniform and harmonious criteria for the answers which the Law has to provide. Only if this path is followed will it be possible to prevent the appearance of "genetic paradises", ie to prevent activities generally considered to be undesirable from being carried out in countries with no regulations or — at best — highly permissive ones; only in this way can we prevent advances in science (genetics, in this case) from becoming yet another factor of discrimination and isolation in developed and developing countries, and ensure that the latter benefit equitably from said advances.

In conclusion, society is faced with a new challenge which must be addressed without prejudices or irrational fears, so that the acquisition of scientific knowledge and the potential applications thereof encounter no other obstacles than the logical prevention of prejudice towards individuals or groups.

As is known, the Chair in Law and the Human Genome of the BBV Foundation — Provincial Government of Biscay at the University of Deusto, has been promoted and encouraged by the BBV Foundation as a part of a broad and ambitious programme of actions relating to the Human Genome

Project, in which the Foundation has been involved since the Project began; the Chair was set up by virtue of an agreement concluded jointly by the Foundation, the Provincial Government of Biscay and the University of Deusto. At this point, it is only right that I should recall that my own involvement in the imaginative project which has been opened up by such a unique Chair was rendered possible thanks to close cooperation between the Universities of La Laguna and Deusto, that have used the appropriate instruments to provide an adequate vehicle for this joint effort, the importance of which they are fully aware. It is against this backdrop that the Chair in Law and the Human Genome, the first Chair of its kind in the world, was conceived by the BBV Foundation as a forum for investigation, special teaching, study and reflection, and also as a meeting place to promote the common dialogue which is required.

It has often been repeated and is well known that advances in biomedicine, that is, medicine and biotechnology, are making huge contributions to the wellbeing of individuals and humanity as a whole. At the same time, however, as the inevitable other side of the coin, the undesired effects of these advances are awakening concern over the potential hazards they represent to the individual and society.[1] Issues relating to organ transplantation, assisted reproduction and genetic manipulation are seen as particularly problematic,[2] as is human experimentation in the quest for attaining biomedical progress.

As regards human genetics, it should be remembered that Spain has hosted an ongoing debate on the Human Genome Project for some years. Two important international workshops on "International Cooperation for the Human Genome Project" were held in Valencia in 1988 and 1990, and have served as valuable precedents for the last International Workshop: "The Human Genome Project. Legal Aspects", organized the last year also by the BBV Foundation with the support of the Provincial Government of Biscay (Diputación General de Bizkaia) at the University of Deusto, Bilbao.[3]

Investigation into human genetics offers the most recent (but not only) example of the lively ethical, philosophical, sociological and legal debates awakened or rekindled by modern science, debates which are far from settled. Within the legal field the discussion primarily turns on whether the

---

1 Romeo Casabona, Carlos M: «La persona entre la Biotecnología, la Bioética y el Derecho» [The person between biotechnology, bioethics and the law], *Folia Humanística*, no. 276, 1984, p. 4.

2 Association Internationale de Droit Pénal: «Droit Pénal et techniques modernes (Resolutions)», *Revue Internationale de Droit Pénal*, v. 59, 1988, p. 1327.

3 See *Human Genome Project: Ethics*, published by Fundación BBV, Madrid, 1991.

new situations call for adaptation of existing legal provisions or for the promulgation of new statutory controls on these activities. It is therefore essential from the legal standpoint to resolve a preliminary question: how should the law intervene, to what effect and on which aspects? As can be seen, this is initially a methodological problem which once resolved will contribute to developing appropriate responses for each particular case. This approach, by analyzing the legal principles and values involved will, first, permit us to determine what must be supported, guaranteed and protected; second, delimit what needs to be regulated and limited; lastly, indicate what must be prohibited or punished, how severely and by means of which public instruments.

Although only partially and for certain spheres of activity, Spanish law already provides regulations directly or indirectly applicable to genetic innovations and human genome study.[4] However, Spanish criminal law, with its characteristic function of invoking penalties to protect certain legal values from serious or intolerable aggression, does not yet address these activities. Whether or not the criminal code should have such provisions is the fundamental question I shall address in this paper. The Criminal Code Bill of 1992, however, does envisage legal limits on different forms of human genome intervention or experimentation, and for such purpose lays down various criminal categories. Whatever the merits of the proposed regulations from a technical, legal and criminal justice standpoint, and irrespective of whether or not the proposed Code is eventually approved, these new provisions are clear signs that the legislator has felt the need to intervene.

From another perspective, the right to scientific research also appears as an interest obviously worthy of protection but at times contraposed to other individual or societal values. The right to engage in scientific research is regarded as a projection of the right to scientific creation and as such primarily responds to the interests of the scientist or researcher, but also to the collective interest in promoting scientific progress and its general benefits for society. There thus arises another aspect which must be likewise guaranteed and protected: the dissemination and circulation of scientific knowledge and information. Article 20.1b and 20.1c of the Spanish Constitution recognize and protect the right to scientific production and

[4] These and other human genetics investigation and experimentation are dealt with in more depth in the report I prepared together with Professor Juan Felipe Higuera Guimera for the Spanish Ministry of Health and Consumer Affairs, entitled «El Derecho ante los Avances y Conocimientos de Ingeniería Genética» [The Law and Genetic Engineering Advances and Knowledge], 1992.

creation and to academic freedom, but establishes as a limit thereupon the respect for the rights recognized in Title 1, by the legal provisions implementing it, and especially by the right to honour, privacy, to personal reputation, and to the protection of youth and childhood (article 20.4). Consequently, although the legitimacy of scientific research is recognized, this freedom — like all freedoms — is not unlimited. This once again reminds us of the need to determine those limits and to locate them among the collective as well as individual interests. In keeping with this dual vision, article 44.2 of the Constitution, under the chapter on the principles governing economic and social policy, imposes on government authorities responsibility for promoting science and scientific and technical research in the general interest. It has been pointed out that the unbreachable limit is found in the human being, his development and immediate setting.[5] Along those lines, Eser has stated that the limits to freedom of research must in any case be found where a medical or scientific procedure gives rise to violations of the civil or criminal code, clashing with the protection of a constitutionally or statutorily guaranteed interest.[6]

2. Criteria for the intervention of law

As Eser[7] has already pointed out, within this ambit we can imagine different systems of step-by-step regulatory control of widely variable intensity and scope. The different steps range from an initial level of deontological self-regulation by the research community, to administrative law provisions, to the introduction of civil code protective statutes — or the strengthening of such instruments, we might add — and even, where necessary, to criminal prohibitions. Indeed, other areas of human activity are subject to this multi-faceted regulation with scaled levels of intensity, and I do not think such a regimen is any less apropos for biotechnology and the biomedical sciences in general as a means of providing the greatest

---

[5] Barbero Santos, Marino: «Fecundación asistida e ingeniería genética. Consideraciones jurídico-penales» [Assisted fertilization and genetic engineering. Criminal law considerations], *Ingeniería genética y reproducción asistida*, Madrid, 1989, p. 329.

[6] Eser, Albin: «La moderna medicina de reproducción e ingeniería genética. Aspectos legales y sociopolíticos desde el punto de vista alemán» [Modern reproduction medicine and genetic engineering. Legal aspects from the German viewpoint], *Ingeniería genética y reproducción asistida, op. cit.*, p. 274 and ff.

[7] Eser, Albin: «Genética humana desde la perspectiva del Derecho alemán» [Human genetics from the perspective of German law], translated into Spanish by C.M. Romeo Casabona, *Anuario de Derecho Penal y Ciencias Penales*, 1985, p. 363 and ff.

possible room for the pursuit of research freedom while at the same time effectively curbing any socially undesirable consequences.

2.1. *Ethical or deontological self-regulation: an exclusively complementary function*

There seems to be a common feeling that scientific activity should be pursued in such a way that ethical and moral considerations will not be ignored. Sensitive to this belief, professor Jean Dausset has proposed that the *Universal Declaration of Human Rights* include the following article: "Scientific knowledge should only be used in the service of man's dignity, integrity and future, but nobody can impede its acquisition". Briefly put, the proposal suggests that while no limitation can be placed on scientific knowledge per se or on its acquisition, its applications and derivations can be limited as required by individual and collective interests.

Appeals to ethics normally lead us to consider the area of professional ethics. It is not uncommon to finds attempts to invoke the latter as the exclusive vehicle for resolving potential problems, thereby keeping their control within the profession and away from the larger community. On the other hand, the situation is much graver when professional ethics are codified and imposed on professionals from outside. I say this imposition is external to the extent that the codes arise from professional organizations or associations in which membership is a prerequisite for practising the profession in question. Such situations imply the not necessarily voluntary acceptance of such deontological rules. This is particularly true when the professional ethics are not legally codified, as is the case with the Spanish Code of Medical Deontology and Ethics approved by the Spanish General Council of Medical Associations in 1990, substituting the previous Code of 1979.[8] Nevertheless, irrespective of the question of the legal nature of deontological norms or the ethical principles regulating the most conflictive aspects of a profession, it is arguably unacceptable that regulation of professional issues directly affecting the society at large be left to the decision of the professional, either individually (personal ethics) or through the professional group to which he or she belongs (deontological norms). Government must therefore not abdicate its legitimate oversight functions in this terrain, and adopt regulatory controls if necessary, as it does in connection with other activities affecting society.

The foregoing is not inconsistent with recognizing the important role of

---

[8] See Iglesias, Tomás: «El discutible valor jurídico de las normas deontológicas» [The debatable legal value of deontological norms], *Jueces para la Democracia*, no 12, 199, p. 53 and ff.

bioethics in its attempt to construct appropriate guidelines for resolving ethical conflicts on the basis of three fundamental principles: autonomy, beneficence and justice. Nor is there any contradiction in advocating the coexistence of legal regulation with professional ethical principles that reinforce interests and values deemed worthy of protection, such principles being conceived as guidelines or prescriptions for professional conduct or even as the highest ideals which ennoble the profession, more than the deontological code. Over the years we have had international examples of such an outlook in the form of formal declarations on specific subjects.[9] And there are examples at the national level, such as the guidelines or recommendations of the scientific societies of physicians, as is frequent in the United States and the Federal Republic of Germany. These initiatives has thus far yielded good results, in that they have been usually accepted as obligatory ethical references.

Along the same lines, biomedical ethics committees can fulfil an important and helpful function in the decision-making process with respect to specific conflicts, similar to the role of clinical testing committees in the investigation of pharmaceutical and similar products. While these committees have a long tradition in some countries (particularly in the Anglo-Saxon world, perhaps favoured by the common law system and the importance it attaches to judicial case-law and precedents, as opposed to the continental European system, primarily based on normative provisions), in Spain their creation is just now under consideration in some hospitals.

Lastly, the so-called National Bioethics Committees[10] are most appropriate for spotting new problems deriving from the most important biomedical advances and their application in human beings. They also work to offer guidelines and advice, which, though not binding, do carry special moral weight for the professional groups concerned and the general population, and call on government authorities to assume their responsibilities and take the appropriate institutional initiatives. Spain's National Commission on Assisted Reproduction, though limited in scope and sector, could be patterned on this model. The Commission, whose creation was provided for by Act 35 of November 22nd, 1988 on Assisted Reproduction

---

[9] See for example the Helsinki Declaration on human experimentation at the 18th World Medical Assembly (1964), revised by the Tokyo (1975) and Rome (1983) Assemblies.

[10] Numerous committees of this kind have been formed in different countries, varying in function, structure, makeup and lines of dependence. The first was created in France by the President of the Republic: the *Comité Consultant National d'Ethique pour les Sciences de la Vie et de la Santé*. At the Madrid International Conference on National Bioethics Committees (March 1992), a proposal was put forward for the creation of an international committee.

Techniques, will be charged with advisory functions as well as other competencies of an administrative nature. The Act indicates that the Commission shall be "aimed at providing guidance on the use of these techniques, at collaborating with the government as to compiling and updating scientific and technical knowledge, and at formulating operating criteria for the Centres or Services where assisted reproduction techniques are performed, in order to facilitate their improved use" (section 21.1). Reinforcing the impression that the National Commission is modelled on the National Bioethics Committees is the Act's provision that committee members represent *inter alia* a "broad social spectrum" (section 21.3), thus indicating that the Commission is not to have a strictly technical character.[11]

2.2. *Health administration control as a preventive measure*

Government intervention in public health matters is fully accepted in advanced societies. We are witness to an increasing control by public authorities of health-care activity understood in the broadest sense of the term. This applies not just to the executive branch of government, which I shall take up later, but also to the legislative and judicial powers. This process is due to various concurrent factors, notably: 1) the health of the citizenry represents one of the prime policy goals of the present-day social or welfare state, which allocates sizable human, economic and material resources for this purpose; 2) the preservation of the individual's health must embrace the quality of public health care and diminish the risks involved, above all, from biotechnological development and the availability of new, complex, and potentially more harmful treatments; 3) this intervention is even more justified and even imperative in relation to health-care practices that possibly affect fundamental individual rights, such as life and personal integrity, liberty, privacy, the family, etc.

The public sector's role in medicine has consequently expanded, either as direct provider of health-care services or indirectly, in the form of collaboration with private sector medicine (by means of outsourcing contracts and agreements) or by financing private health care and research.

---

[11] For its part, Spanish legislative Act 42 of December 28th, 1988 on the donation and use of human embryos and fetuses or their cells, tissues or organs, provided for the creation of a National Commission to oversee the donation and use of human embryos and fetuses (in its first additional provision, sub-paragraph f). As in the case of the National Commission on Assisted Reproduction, the government has not yet approved the Royal Decree constituting the Commission even though both laws provided for this purpose a period of six months reckoned from the date of their respective enactment.

In any event, whatever health-care system is adopted, new administrative regulatory channels have been instituted for authorizing certain health-care services. These include personnel and material certification in health centres, professional certification for implementing or pursuing certain practices, follow-up of new treatments or of those still in the research and experimental stages, the stipulation of protocols, etc. These are more or less indirect procedures that allow the establishment of preventive measures whereby assurances can be provided in connection with some of the more delicate and conflictive activities. On the other hand, when health-care professionals are directly employed by the government, the authorities have a highly effective regulatory tool: the possibility of disciplinary action, with the application of sanctions and other corrective measures for the more serious infractions.

In Spain, this process has been gradually incorporated into the country's health-care system. The European Community recognizes the right to the protection of one's health, with such protection being incumbent upon the public authorities (section 43). This right was further implemented by the General Health Act (Act 14 of April 25th, 1986) as to health-care delivery and the creation of the National Health Care System. In addition, there are sectorial laws which regulate specific medical treatments and practices, such as organ extraction and transplantation, clinical autopsies, assisted reproduction techniques, clinical trials, etc. The regulations implementing these laws have conferred upon Spanish government health authorities an ever more significant and effective capacity to intervene in these matters. Taken as a whole they display the above-cited monitoring and intervention features: regulation of the lawful scope of activity, oversight to assure the quality of professional practice, and decisions aimed at safeguarding the relevant individual rights and legally protected interests.

In conclusion, if this growing protagonism has taken root in numerous complex biomedical activities, nothing should impede its extension to activities and investigation relating to human, animal or plant genetics. Such intervention must be understood to be fundamentally justified as a means of forestalling the potential hazards entailed in these and other biomedical science activities and not as a lever for government-run science or as a constraint on the freedom of investigation. The latter reservation does not argue against the government's legitimate function of encouraging and promoting those scientific branches most deserving support in the pursuit of general interests deemed opportune, or of restricting or eliminating financial support for research running counter to regulatory control mechanisms or having undesirable aims.

Two laws with a direct bearing on regulation of activities in human genetics have to be considered within this framework: legislative Act 42 of December 28th, 1988 on the *donation and use of human embryos and fetuses or their cells, tissues or organs*, and Act 35 of November 22nd, 1988 on *Assisted Reproduction Techniques*. The first is concerned with the donation and use of human embryos and fetuses or their cells, tissues or organs for diagnostic, therapeutic, research or experimental purposes (section 1) and specifically with genetic technology. The second regulates diverse aspects relating to assisted reproduction techniques, such as diagnosis, therapy, research and experimentation with gametes and pre-embryos or preimplantation embryos.[12] They attribute important oversight competencies to the government in connection with these activities, particularly to the commissions envisaged for such purpose. However, in some cases, the commissions are given excessive discretion as to authorization for the pursuit of certain insufficiently delimited activities (eg, sections 16.1.k and 20.2.B.n) and r) of Act 35/1988). The cases and circumstances in which authorization is discretional with the authorities or commissions should be expressly mentioned and defined by the Act, otherwise the scope of what is actually permitted or prohibited by the law is overly diluted.

As an ultimate reference, both laws contain a catalogue of prohibited practices (serious and very serious offences), including some which are directly tied to genetic manipulation, and which are subject to administrative sanctions, that is, non-penal sanctions (section 20 of Act 35/1988 and section 9 of Act 42/1988). The different violations may be classified into several groups: a) those involving noncompliance with health administration rules; b) infringement of the rights or interests of persons involved in assisted reproduction techniques (users and donors); c) those involving protection of human biological "material" (gametes and pre-embryos); d) application of assisted reproduction techniques for purposes other than those envisaged under the Act; and e) prohibited genetic manipulation or experimentation.

As can be seen, Spanish lawmakers have decided for now not to establish new criminal categories in these Acts under which to impose criminal sanctions on violators of its regulatory provisions. With respect to some of the serious offences contained in the Act on Assisted Reproduction Techniques, doubts are raised, not as to their prohibition (something that appears obvious) but as to which legal instrument should be used — penal

[12] Both acts were challenged as unconstitutional before the Spanish Constitutional Court in 1989 (no judgement has yet been handed down) by the Popular Party Group in the Congress of Deputies, which argued that the laws did not sufficiently protect certain fundamental rights.

or administrative. This uncertainty specifically arises in connection with the following conduct: the creation of human beings by cloning or other procedures aimed at racial selection (section 20.2.B.k); creation of identical human beings by means of any type of cloning or other procedure able to originate several identical human beings (section 20.2.B.l); parthenogenesis, or stimulating development of an ovule by thermal, chemical or physical means without its fertilization by a spermatozoon, which would give rise to only female offspring (section 20.2.B.m); sex selection or genetic manipulation for non-therapeutic or unauthorized therapeutic purposes (section 20.2.B.n); creation of pre-embryos of persons of the same sex for reproductive or other purposes (section 20.2.B.o); fusion of pre-embryos or any other procedure aimed at producing chimeras (section 20.2.B.p); human genetic exchange, or recombination with other species, for the production of hybrids (section 20.2.B.q); unauthorized transfer of human gametes or pre-embryos to the uterus of another animal species, or the inverse operation (20.2.B.r); ectogenesis, or the creation of an individual human being in a laboratory (section 20.2.B.s); creation of pre-embryos with sperm from different individuals for transfer to a uterus (section 20.2.B.t); simultaneous transfer to a uterus of pre-embryos originating from ova of different women (section 20.2.B.u); use of genetic engineering and other procedures for military or other purposes to produce biological arms or human exterminating weapons of any type whatsoever (section 20.2.B.v). The Act on donation and use of human embryos and fetuses or their cells, tissues or organs also raises doubts as to the criminal nature of some of the prohibitions it sets down, such as on all acts aimed at altering non-pathological human genetic inheritance (section 9.2.B.a).

The question thus inevitably arises of whether these legal instruments are sufficient or, on the contrary, others (such as criminal statutes) are needed in order to achieve adequate legal protection of the interests and values at stake. The creation of criminal categories in respect of the conducts prohibited by these laws (administrative offences) does not imply violation of the principle of *ne bis ni idem*, for only one type or the other would be applicable (the criminal or administrative sanction). Where such conflict arises under Spanish law it is clearly resolved in favour of preference for the criminal sanction.

2.3 *Liability for harmful or hazardous acts*

The legally available means for redressing these situations is the reparation of the damages occasioned, which pertains to the realm of civil law, and repression for that damage, fundamentally attributed to criminal law. This is not to say that the repressive function is exclusive to the latter or to minimize the important deterrent function that can be fulfilled both by civil and criminal law.

2.3.1. Civil liability

As regards reparation of damages under civil law (there is now discussion of incorporating reparation of damages as a concrete penal instrument in criminal law), that is, of finding civil liability, three essential prerequisites must be met: 1) culpability; 2) the occasioning of real, economically measurable damages; and 3) a causal link between the first two prerequisites. Legal doctrine and practice in North America — and more recently in Germany and France as well — hold that an economically compensable damage exists in cases of children born with hereditary malformations in families with some known risk in their medical history but whose parents were nonetheless not offered the medically indicated prenatal genetic screening and corresponding genetic counselling; or where such tests or counselling were deficient (wrongful life).[13] Judicial questions of this type have not yet been raised in Spain, although a similar complaint was recently reported (apparently a criminal complaint, which would not be likely to succeed), and is awaiting decision.

2.3.2. The intervention of criminal law

The role of criminal law in relation to a field as new as the one we are studying here should be guided by its traditional governing principles, but as those principles are currently understood,[14] that is, consistent with the function of criminal law and the principles of minimum intervention and *ultima ratio*.[15] It was recalled above that criminal law has as its exclusive responsibility the protection of legal "goods", that is, of the fundamental principles, values and interests of the individual and community, from conducts that harm or endanger them whenever such conducts involve serious offence to the society's prevailing socio-ethical norms. Nevertheless,

---

[13] Crammer, Stephan V.: «Pränatale Diagnostik und Fetaltherape» in *Medizinrecht*, 1991, p. 14 and ff, and Knoppers, Bartha M.: *Human Dignity and Genetic Heritage*, Law Reform Commission of Canada, Ottawa, 1991, p. 54 and ff.

[14] For more on what follows, see Cerezo Mir, José: *Curso de Derecho Penal. I Parte General*, 3rd edition, Tecnos, Madrid, 1985, p. 11 and ff.

[15] See Lathi, Raimo: «Criminal Law and Modern Bio-Medical Techniques. General Report», *Revue Internationale de Droit Pénal*, v. 59, 1988, p. 611.

resort to criminal law should be reserved for situations of intolerable attack on the most important legal values and only when made strictly necessary by the insufficiency of other, non-penal legal instruments (administrative law, civil law, etc.). These criteria should also be applied in relation to the law's attitude toward human genetics, such that the place of criminal law in regard to that area should usually be accessory or nonexistent, especially when other branches of law can fully satisfy the need to protect values and interests and limit undesirable activities. On the other hand, recourse to criminal law will be inevitable — but always exceptional — when dealing with particularly grave conducts in relation with interests or values worthy of heightened protection. Consequently, as indicated further above, the first limiting and sanctioning filter should be obtained from the laws regulating, curbing and sanctioning procedures for assisted reproduction and the use of human gametes and embryos — in Spain in the aforementioned Acts on Assisted Reproduction Techniques and on the donation and use of human embryos and fetuses or their cells, tissues or organs. This is in addition to civil reparations of the damages occasioned.

Although these criteria are clear, they are in practice difficult to specify. As mentioned earlier, they should reflect the community's dominant socio-ethical conceptions in relation to the different consequences of human genetics discoveries and activities, and at the moment it is not easy to find widely accepted points of agreement. In addition, varying perspectives on the role of criminal law — though generally accepting the core criteria set out above — can delay or accelerate its application. Thus under a conception of criminal law accentuating its preventive function, an ends-based rationale (the utilitarian conception) could lead to criminal sanction of conducts that endanger the principles of law whose protection is sought, even when such danger is merely potential, without the need for a specific case to have arisen (crimes of abstract danger); whereas a values-based rationale will emphasize the rights of the individual and the ideals of humanity and justice,[16] thus tending to limit the application of criminal categories. Other sectors appeal to criminal law's symbolic function of reinforcing moral criteria in the face of deviant conducts, that is, demonstrating the moral error of the conduct, and emphasizing its declaratory function in asserting the society's limits of tolerance.[17] Finally, others place more importance on the regulatory rather than the punitive function of criminal law (the latter itself would fulfil that regulatory

[16] *Ibid.*, p. 611.
[17] *Ibid.*, p. 612.

function). In this approach, certain conducts are not directly outlawed, but instead made conditional on compliance with determined conditions and procedures, whose infringement could give rise to criminal sanctions (similar to what occurs in economic criminal law).[18]

The foregoing undoubtedly confirms the complexity and difficulty of finding appropriate criminal justice criteria, and at the same time underscores the risk of expanding criminal law away from its essential function and principles of intervention, from which it should not stray, without forgetting the aspiration to the above-mentioned ideals of justice and humanity. Hence, having established this basic framework, further advances will require detecting and delimiting the legal principles or values or interests that could be affected by applications in the field of human genetics. Only after having identified those values will it be possible to analyze the conducts which could represent a threat thereto, assess their gravity, and tailor criminal policy accordingly. In any event, we cannot yet aim for definitive solutions; the possible consequences of human genome knowledge are still not sufficiently known and, second, certain consequences which today may be deemed as undesirable or deviant could perhaps be viewed differently in the future, particularly if the undesirable effects can be prevented. In addition, it should be borne in mind that much of the genetic experimentation described in scientific literature is regarded as theoretically possible but not yet technically practicable, which raises the question as to whether criminal law's preventive function (where sanctions are deserved and needed) should be pursued even before such possibilities are realized.

3. Legal values or goods meriting protection or heightened protection

Human genetics investigation and its potential applications may affect legal principles, or individual or collective interests or values. The task is to identify these, the protection they enjoy under our laws, and the gaps in that protection and to decide which legal instruments should be employed to guarantee adequate protection, according to their importance and the types of aggression to which they are exposed, without the recourse to criminal law being always necessary.

[18] *Ibid.*

3.1. *Individual goods or interests*

Humanist currents of recent decades have accentuated the recognition of human dignity and the free development of one's personality as individual values of the first order. These values often appear at the centre of many aspects of the debate concerning human genetics. The Spanish Constitution of 1978 expressly lays down in article 10.1 that "Human dignity, man's inviolable and inherent rights, the free development of his personality, respect for the law and for the rights of others are essential to political order and social peace". Frequent, indiscriminate and sometimes abusive appeals to these values and the lack of attention to studying their overall place and precise function in the Constitution should not lead us to lose sight of their intrinsic importance. What I wish to stress here is the projection of these inherent rights onto the other fundamental rights and public liberties enshrined in the Spanish Constitution (similar to that of other similarly structured Constitutions), in the sense that the former give the latter greater constructive sense and cohesion, more than their conception as autonomous fundamental rights.[19] The principle of equality before the law, and, above all, its consequent ban on discrimination by reason of birth, race, gender, religion, opinion or any other personal or social condition or circumstance (article 14 of the Spanish Constitution), are also brought to the forefront by the discriminatory potential of certain human genome applications.

Most of the individual rights which could be affected by human genetics applications are specifically recognized in modern-day constitutions and international human rights declarations and conventions, and they have traditionally been conferred a special protection by criminal law. This is the case with human life, integrity (physical and mental), freedom of decision and self-determination, all of which are legal values directly protected in criminal law with respect to almost all forms of aggression. This, however, is not always the case, as I will attempt to demonstrate further below. There has been some flux in the protection afforded other legal interests, but the trend in recent years with respect to human rights is that they merit heightened protection in the face of new forms of aggression

[19] In this sense see Vale Muñiz, José M. and González González, Marisé: «Utilización abusiva de técnicas genéticas y Derecho Penal» [Abusive use of genetic techniques and criminal law], in *Poder Judicial*, no. 26, 1992, p. 126 and ff. Also, Mateo Pardo, Regino: «La 'dignidad de la persona human' y su significación en la Constitución española de 1978 a través de la jurisprudencia del Tribunal Constitucional» [Human dignity and its significance in the Spanish Constitution of 1978 in the case-law of the Constitutional Court] in *Escritos Jurídicos en Memoria de Luis Mateo Rodríguez*, Universidad de Cantabria, 1992, 341 and ff (348 and ff).

that were previously unthinkable or regarded as unimportant, such as with the life and bodily and mental (future) integrity of the fetus.

3.1.1. Life and integrity after birth. Genetic integrity

1. Just as occurs with some medical activities (diagnostic, therapeutic) that are new, still in the experimental stage or not yet consolidated — particularly the more aggressive ones or those affecting especially vital parts of the person — some aspects of the application of genetic techniques pose grave potential danger to the life, health or integrity of persons, including their genetic integrity. In general, criminal categories sanctioning homicide and bodily harm can satisfy the need for protection with respect to these conducts, including in very exceptional cases those involving wilful misconduct (second degree or even first degree). In any case, we must ascertain whether or not criminal categories regarding injuries afford adequate protection from certain genetic actions on human beings after birth.

2. As is commonly known, genetic engineering is giving rise to a variety of manipulations of human genetic components. Most noteworthy among these is the potential offered by gene therapy for curing serious genetically caused diseases or defects by means of the addition, modification, replacement or removal of genes. Distinct types of genetic defects are involved: hereditary, when transmitted by the parents' genes; non-hereditary, when anomalies are produced by unforeseen errors in the formation of sex cells; and congenital, when they occur during fetal development. For the moment these techniques are aimed at correcting defects of a monogenic origin. A distinction should also be made between manipulation of somatic and of germ-line cells, and between the different genetic diagnostic tests, including antenatal diagnosis (preconceptive and prenatal). The latter tests, in the context of possible subsequent therapy (or of abortion for eugenic reasons, where permitted under the country's laws), should be considered as lawful, provided they respond to the needs arising from the *lex artis* in each specific case, as I will further expound ahead.

Genetic manipulation of somatic cells entails specific actions on those cells that do not affect the person's genetic pattern, because such cells do not transmit the alterations made to them to the person's offspring (eg, in the pancreatic islets to combat diabetes). Nevertheless, somatic cells and their components (including genetic components) form part of the individual's personal integrity (physical or mental), within a subcategory we could term as "genetic integrity", and are consequently entitled to the protection afforded said legal interest under criminal law. The following conclusions can therefore be established: 1) criminal law should regard somatic cell gene therapy as it does any other treatment (without prejudice to the relevant special considerations in respect of new or experimental treatments), that is, it is what is known as therapeutic experimentation, which implies submission to the general guidelines and limitations commonly accepted for this therapeutic category (namely the weighing of risks and benefits, and the patient's informed consent); 2) somatic cell gene therapy when in conformity with the *lex artis* is therefore lawful, and, moreover, does not even come under the bodily injuries criminal category; 3) any other, non-therapeutic action entailing alteration of the genetic components of a person's somatic cells does qualify as a bodily injuries criminal offence to the extent that it impairs the person's physical integrity or physical or mental health (section 420 of the Spanish Criminal Code), and the exact offence will depend on the medical care required, where such care is possible (sections 420, 421 and 582 of the Spanish Criminal Code), or on how these offences are configured in the applicable criminal code. Nevertheless, these actions will be lawful if they are based on the informed consent of the interested party (the holder of the legally protected interest) and if the laws of the country in question recognize the efficacy of such consent in relation to physical integrity and health issues (under Spanish law, section 8.11 of the Criminal Code, concerning the legitimate exercise of a right or profession, is frequently cited in this sense).

With respect to germ-line manipulations (germinal cells: gametes and zygotes), the same criteria as those proposed for somatic cell manipulation should, in principle, be applicable. Bear in mind that the question we are dealing with here is not that of protecting gametes and zygote, but the reproductive capacity of individuals presenting abnormalities in their reproductive cells or which manifest themselves immediately after their union. Germ-line gene therapy, however, raises other, more serious ethical and legal problems. Although it will surely contribute in the future to eradicating defects from the manipulated lineage, it will also bring about permanent genetic modifications that are transmitted to later generations and whose possible impact on the human species is not yet precisely understood, thus making it difficult to control the potential negative effects. The misgivings awakened by such as yet unknown effects have led some specialists to propose an outright ban on germ-line cell therapy and others to seek a moratorium until a fuller understanding is achieved. Lastly, there are those who hold that this therapy should not be rejected, because it will not be possible to see any actual risks to the human species as such until several generations have passed and the alterations are reproduced in other human beings.

The foregoing begs the resolution of two questions. First, to determine in these cases what should be understood by therapy in the strict sense and its possible differentiation from manipulations aimed at genotypic or phenotypic improvements and not at correcting a clearly pathological condition. Second, given that other non-therapeutic germ-line manipulations will be genetically transmitted to offspring, to determine if these cases involve some other legal interest which transcends the collective meriting protection and, if so, to identify that interest. I will explore this second aspect further below. Lastly, in light of the second consideration, to determine whether the law should allow germ-cell gene manipulations of any kind, and, if not, should their prohibition regard such conducts as criminal offences? This question will also be explored further ahead (although I can now tell you that the answer is affirmative). In view of the current situation, the most judicious course would seem to be support for a moratorium insofar as concerns germ-line gene therapy only.[20] For the time being, criminalization of such therapies does not seem called for. Their curtailment should be a matter of administrative law and serve as a criteria for the granting or denial of public funds for related research or applications, without prejudice to the possibility of allowing use for therapeutic purposes, according to the merits of each specific case and a weighing of the safeguards against undesirable mutations or aberrations, all as evaluated by a committee of experts.

In Spanish law, the Act on Assisted Reproduction Techniques allows gene therapy (section 1.3), specifically for preimplantation embryos (pre-embryos *in vitro* or in the uterus), embryos and fetuses in the uterus on certain conditions: the national government must have issued a royal decree setting down a list of diseases for which therapy is possible with strictly scientific criteria (First Final Provision, d), the therapy must not influence non-pathological hereditary characteristics, and the aim is not race or individual selection (section 13.3, c and d, respectively). This broadly worded last condition suggests that Spanish law does not, in principle, exclude, germ-line gene therapy, without prejudice to the restrictions that could be introduced in this respect when the aforementioned list of diseases is published. Considered as therapy-related are sex selection for therapeutic purposes, that is, as a means of impeding the transmission of hereditary diseases linked to the sex chromosomes, and the surgical creation of

[20] In this regard see Resolution 6.8, Section II of the International Criminal Law Association's XIV International Congress on Criminal Law, in *Kongressakten*, Vienna, 1989, p. 285 and ff, which envisages in support of such a moratorium the establishment of, as minima, deontological guidelines and/or a restrictive authorization policy.

beneficial genetic mosaics by grafting cells, tissues or organs of embryos or fetuses to patients in which they are biologically or genetically dead or missing (section 8 of Act 42/1988).

3.1.2. Life and integrity pre-birth. Protection of the preimplantation embryo

1. Abortion advocates have gained steady support and legislative backing for their positions, resulting in possible inattention to the legal protection of the unborn from their earliest stages of development. On occasion this disregard is deliberate and "consistent" with the limits sought on the protection of the unborn in connection with abortion. I will not enter into this controversy or the ideological and sociological issues that fuel it. What is certain is that biotechnology forces a reconsideration of the protection of the unborn in other aspects unrelated to voluntary abortion.[21] In any event, human life after conception and prior to birth, the life of the unborn, has always merited special penal protection from abortion with *mens rea* by a third party (again, I repeat, without here considering abortions performed with the consent or at the request of the mother and the circumstances in which voluntary abortion is decriminalized), but the same cannot be said with respect to negligent abortion. Genetic manipulations (and medical manipulation in general) can give rise to an increased number of the latter type of abortions, which should be prevented. In fact, negligent abortion performed by a third party should be punished, provided the presence of reckless negligence, regardless of the means used, or whether the result was foreseen.[22]

2. Although legal professionals are not always aware of its importance, reducing the number of embryos within a multiple pregnancy also has criminal justice implications. By multiple pregnancy I am referring to when the woman is impregnated with various embryos by means of assisted reproduction techniques, whether by *in vitro* formation and transfer to the woman of more embryos than necessary or by hormone-induced ovular hyperstimulation of the woman. A multiple pregnancy achieved by these

---

[21] On Constitutional and penal law protection of the beginnings of human life, see Romeo Casabona, Carlos M.: «Human Life as a Value Protected by Penal Law», in *Le Droit face aux dilemmes moraux concernant la vie et le mort*, Proceedings of the XXth Colloquy of European Law, Glasgow, 10-12 September 1990, Council of Europe, Strasbourg, 1993, p. 135 and ff.

[22] This proposal seeks to underscore the scarce punitive margin allowed by the current regulatory provisions of the Spanish Criminal Code, specifically section 412 as it now stands, if they are construed as only covering the negligent case and section 565 (and the fault set down in section 586 bis) is considered as not applicable in these cases.

means can actually endanger its eventual coming to term, due to the increased mortality of each of the embryos and to the risks such pregnancies represent for the health and even life of the pregnant woman. Several legal problems are involved. The destruction of the "excess" embryos occurs after their implantation in the maternal endometrium, and therefore it cannot be said that such action does not constitute a punishable abortion even if we appeal to the proposed criteria of implantation as the lower threshold of penal protection. We must consequently conclude that this is an abortion offence,[23] understood as the destruction of human life during gestation and not just as the mere interruption of the pregnancy, for the pregnancy continues, although the number or embryos has been reduced. Any other interpretation would lead to having to accept as lawful and not punishable the destruction of an embryo in an unprovoked multiple pregnancy without any legal authorization. The only legal procedure for rendering such action lawful would be to cite the justifiable grounds for abortion in those systems allowing abortions in certain circumstances and conditions (time period or indication).

As regards Spanish law, this would involve determining whether the specific case involves a therapeutic or a eugenic indication (section 417 bis 1.1 and 1.3 of the Criminal Code).[24] Both situations would require a finding by professionals different from those who are to perform or direct the abortion, along with the other general and specific requirements for each of the indications. But the indication would have to come first, that is, serious danger for the life or physical or mental health of the mother, or where it is presumed that the fetus will be born with grave physical or mental handicaps. The usual situation, however, is that neither of these two cases arises. Moreover, the destruction of the "excess" embryos is not selective, but random, depending on which are more easily accessible to the ultrasonic puncture or injection of potassium chloride in the thoracic or cardiac cavity. In other words, healthy embryos are eliminated to favour the survival of other equally healthy embryos (though with common risks should they all be maintained). The paradox is that the indication could be extended to all of them as a group (given that they all present a risk of defective or pathological development if gestation continues with all of them), but to no single one of them. Certainly, though, the lesser evil is to save some to the detriment of others,[25] and that upon applying the indication to the

[23] See Eser, Albin: «Amenazas a la vida humana en su comienzo» [Threats to human life at its commencement], in *Jornadas sobre la Reforma del Derecho Penal en Alemania*, Consejo General del Poder Judicial, Madrid, 1991, p. 14 and ff.

[24] For German law in this regard, see Eser, *ibid.*, p. 15.

[25] Eser, *ibid.*

eliminated embryos such indication no longer holds for the surviving ones (as the risk of defective development also disappears).

There is also the unresolved question of whether or not the situation in which a physician consciously and deliberately induces a multiple pregnancy, knowing that some of the embryos will later have to be eliminated, should be considered as an *actio illicita in causa* and the physician therefore excluded from issuing the indication for abortion (posing the added imputability if he or she is not the one who performs or directs the abortion). For this reason, given that the aim of this procedure is often to assure at all costs that the "client" becomes pregnant, the appropriate course of action would be for the physician to only transfer the number of embryos strictly necessary for assuring impregnation or that the already available procedures for more controlled and measured ovular stimulation be used to avoid such situations.

3. Some aspects of prenatal testing are also of interest from a criminal law perspective.[26] Prenatal testing can be understood to refer to information on possible congenital defects of the fetus obtained by means of a set of procedures that biomedicine makes possible for such purpose. Or, put in more technical terms: "All those prenatal actions aimed at diagnosing congenital defects, understood as any external or internal, familial or sporadic, hereditary or non-hereditary, singular or multiple anomaly in the morphological, structural, functional or molecular development present at birth (even though not manifested until later)".[27] These screening techniques can detect the presence of disease or malformation in the fetus, and even predict the fetus' sex. Couples with a high risk of offsprings suffering abnormalities can also use it to confirm or rule out the presence of such abnormalities in the fetus. Thus, the fetus is the subject of prenatal diagnostic techniques, as opposed to preconceptive testing. Prenatal testing, like the latter, generally leads to some form of genetic counselling. Today

---

26 See the report of the Comité Consultatif National d'Ethique pour les Sciences de la Vie et de la Santé: «Rapport: le diagnostic prénatal et périnatal. Le diagnostic d'une prédisposition», Paris, 1985, p. 4 and ff; Fraser, F. Clarke: «Diagnostic prénatal des désordres génétiques», in *Cahiers de Bioéthique*, no. 2, Quebec, 1980, p. 3 and ff; Powledge, Tabitha M. and Fletcher, John: «Recommandations concernant les problèmes moraux, sociaux et juridiques relatifs au diagnostic prénatal», also in *Cahiers de Bioéthique*, no. 2, Quebec, 1980, p. 92; Romeo Casabona, Carlos M.: «El diagnóstico antenatal y sus implicaciones jurídicos-penales» [Antenatal diagnosis and its implications for criminal law], in *La Ley*, no. 1751, 1987, p. 6.

27 Carrera, José M.: «Diagnóstico prenatal: Un concepto en evolución» [Prenatal testing: An evolving concept], *Diagnóstico Prenatal*, edited by J.M. Carrera, Salvat, Barcelona, 1987, p. 5, following the definition made under the auspices of the WHO.

there are a variety of medical techniques which provide early detection of fetal abnormalities. Amniocentesis is the one which currently allows diagnosis of the greatest number of congenital diseases. It entails certain risks, such as loss of the fetus, fetal injuries and infection, neonatal disorders, and complications for the mother (visceral perforation, premature detachment of the placenta, premature breaking of waters, infection, fainting and death of the mother, uterine contractions and premature labour, postpartum haemorrhaging). Nevertheless, the test has undergone continual improvement and its inherent dangers have been steadily reduced.

Prenatal screening offers, or is well on the way to offering a wide range of diagnostic possibilities. Some of the main aims they serve are to: a) reassure high-risk parents that the fetus does not present any disease or malformation; b) allow treatment (surgical, medicinal, genetic therapy) of the fetus in order to cure or alleviate determined anomalies, or fetal therapy; c) indicate the optimum delivery method in view of the malformations suffered by the fetus (eg, a caesarean section); d) determine the post-delivery or later treatment to be followed with the newborn (preventive medicine); e) adopt a decision as to eugenic abortion where permitted by law (in Spain by section 417 bis number 1.3 of the Criminal Code, and in countries where abortion is allowed by medical indication and/or in the earlier stages of the pregnancy); g) accept the likely handicapped child or prepare the legal proceedings for his or her adoption by others or admission into a centre for abandoned children. The decision as to these options will depend on the personal convictions of the parents and the relevant legal regulations. In any event, it has been pointed out that this range of possibilities offered by prenatal screening, which after all will permit improvements in obstetrical and neonatal care, should not imply an indiscriminate increase in abortions, thanks to the advances achieved in perinatology over recent years.[28]

There is no objection to classifying prenatal testing as a therapeutic measure, as with any other diagnostic procedure, from the moment it is directed toward the *treatment* of the fetus to the fullest possible extent.[29] It must be considered that such diagnosis does not constitute a bodily injuries criminal offence with respect to the mother or embryo insofar as it represents a potential benefit for both. This requires verifying the medical indication for screening, weighing the risks and advantages and acting in accordance with *lex artis*. This category is understood to include diagnostic

---

[28] Carrera: «Diagnóstico prenatal: Un concepto en evolución», *ibid.*, p. 8.

[29] See Eser: «Genética humana desde la perspectiva del derecho alemán», *op. cit.*, p. 351.

techniques that still have a certain experimental component (therapeutic experimentation) and hence some risk (as occurs to a point with amniocentesis in light of the already described risks it entails). If the screening is performed with a view to a eugenic abortion, the injuries produced by the diagnostic tests (in Spanish law generally equivalent to a bodily injuries misdemeanour and not criminal offence) are so minor as to be covered by the consent of the pregnant woman, who suffers them in the first place.

4. A comparison of various criminal justice systems (and the current Spanish Criminal Code) shows there are also gaps in the legal protection of the health and physical integrity (and later mental integrity) of the unborn, who are much more vulnerable as a consequence of the development of biomedical sciences (without prejudice to the necessary reminder that these sciences are also providing a greater understanding and better techniques to nurture the unborn child's development). The ensuing question is therefore whether or not the integrity of the embryo or fetus merits protection under criminal law. In my opinion it does,[30] both as a consequence of the intrinsic importance of the legal interests involved and because, at least at present, most of the effects and handicaps of these injuries are difficult to treat and correct and will therefore persist throughout the life of the future child. In support of this view, I can cite the famous April 11th, 1985 judgement number 53 of the Spanish Constitutional Court on the constitutional challenge to the law which decriminalized abortion in certain situations and medical indications. The high court held that the life of the *nasciturus* embodies a value that is constitutionally protected under article 15 of the Constitution ("All have the right to life and to physical and moral integrity ..."). I think this criterion can be properly construed as extending to the integrity of the unborn as well.

5. Today's broadened understanding of the different stages of embryonic and fetal development and the fact that modern assisted reproduction techniques allow embryos to be produced in laboratories and outside the uterus (pre-implantation embryos), together with the knowledge

[30] As I outlined in 1981 in *El Médico y el Derecho Penal. La actividad curativa*, Bosch, Barcelona, 1981, p. 279 and ff, and developed in later studies such as «La protección jurídica del concebido. El feto como paciente» [Legal protection of the unborn. The fetus as patient], in *Revista Jurídica de Castilla-La Mancha*, no. 7, 1989, 445 and 448. See also Cuerda Riezu, Antonio: «Límites jurídico-penales de las nuevas técnicas genéticas» [Criminal law limits of the new genetic techniques] in *Anuario de Derecho Penal y Ciencias Penales*, 1988, p. 442 (although from a different standpoint); Vale Muñiz and González González: «Utilización abusiva de técnicas genéticas y Derecho Penal», *op. cit.*, p. 144 (argued on grounds similar to those presented in this text).

that they can be manipulated and are not necessarily destined to be implanted and develop in a woman, has led to two contradicting situations. First, there is a tendency to not consider as criminal abortion the destruction of "excess" embryos (pre-implantation embryos or pre-embryos) produced by means of *in vitro* fertilization and used as research material. For such an offence to exist the destroyed embryo must have been previously implanted in the mother's womb or expelled (naturally or artificially) therefrom after implantation, unless one wishes to incur in an application of law by analogy, which in Spanish criminal law is generally considered to violate the "principle of legality" — there is no crime without a law, no punishment without a law — and therefore not allowed. This issue is directly connected to the question of at what moment legal protection of the embryo has effect. The present trend is to consider the moment of implantation in the uterus for such purposes.[31] The solution is certainly not easy when someone destroys embryos or gametes that were obtained from a couple for later use and are irreplaceable (ie, when no more gametes are available), in order to prevent the couple from bearing offspring, or in order to prevent a researcher in possession of the gametes to continue his or her work. The latter case could constitute a crime against property but this would imply recognizing a reification of the human reproductive element or of the embryo itself. Such reification would surely be inadmissible in relation to the first case; perhaps criminal coercion could be applied to that case, and if no criminal liability is found, action for civil liability could be pursued.

Second, given the shortcomings pointed out above, some sectors have called for the need to grant some form of legal protection, even penal protection in extreme cases, to the non-implanted embryo in order to prevent its manipulation (or some types of manipulations: those that are not aimed at benefiting the embryo, ie, diagnostic or therapeutic measures aiming at its later implantation; see further above for the remarks on somatic and germ-cell gene therapy). Advocates of this position invoke different arguments, defining the scope of sanctionable conducts more broadly or narrowly. Some base their arguments on the view that the pre-implantation embryo is inherently worthy of protection, others point to the

[31] For more on the state of this question, see Romeo Casabona, Carlos M.: «La reforma del aborto: Límite mínimo, figuras delictivas y sistema de indicaciones» [Abortion reform: lower limit, criminal categories and medical-indication system], in *Propuestos para la Reforma Penal*, Centro de Estudios Criminológicos, Universidad de La Laguna, 192, p. 138 and ff, and in *Actualidad Penal*, 1991, p. 137.

right to free development of the future person,[32] and others appeal to the interests of humanity as a whole. In any event, proposals for criminal-law protection of the preimplantation embryo do not necessarily contradict legal systems in which abortion has been decriminalized, above all if the legal value whose protection is actually sought is correctly profiled. Nevertheless, jurists seem to agree as to the prohibition of obtaining embryos *in vitro* for purposes other than procreation, such as for research or industrial use, and the prohibition, backed up by the threat of penal sanction, on implanting embryos which have been previously subject to any of these manipulations (that is, those whose purpose is not diagnostic or curative with respect to the embryo).[33]

3.1.3. Genome analysis and individual freedom

1. The Human Genome Project and genome analysis will allow a person's genetic map to be completely known, making possible the prediction of future genetic diseases, even those with onset many years after the analysis is performed. As is known, this capability will allow a variety of preventive or curative measures to be taken before the disease manifests itself, particularly gene therapy measures, depending on the state of these techniques at each point in time. This knowledge, however, can suggest the adoption of certain measures which could gravely affect the individual's freedom or bring about discriminatory situations.

2. In this medical area the patient's right to be informed enjoys the same legal recognition as has been granted for medicine in general since years ago, specifically embodied in various charters of patient's rights and health legislations (in Spanish law, principally the General Health Act of April 25th 1986, section 10.5). The penal implications of omitting such

---

[32] Vale Muñiz and González González: in «Utilización abusiva...», *op. cit.*, p. 124 and ff (particularly p. 125 and ff and 132 and ff) maintain a restrictive position. Conversely, Germany's Embryo Protection Act (Gesetz zum Schutz von Embryonen-Embryonenschutzgestez) of December 13th, 1990 has opted for broad criminal coverage, as will be seen further ahead.

[33] See Eser: «La moderna medicina de la reproducción...», *op. cit.*, p. 288; Romeo Casabona, Carlos M.: «La utilización de embriones y fetos humanos con fines de investigación genética u otros fines no terapéuticos» [The use of human embryos and fetuses for genetic investigation or other non-therapeutic purposes], in *Eguzkilore*, Cuaderno del Instituto Vasco de Criminología, San Sebastián, no. 5, 1992, p. 151 and ff; Schreiber, Hans-Ludwig: «Der Schutz des Lebens durich das Recht an setnem Beginn und an setnem Ende», *Festchrift fur Günter Schewe*, Springer Verlag, Berlin, 1991, p. 127; Vale Muñiz and González González: «Utilización abusiva...», *op. cit.*, p. 134, although they hold that penal intervention is not justified.

information are very limited, however (and mainly turn on the fact that a properly informed patient is a prerequisite for the consent of the patient to have legal force in those offences for which such consent releases a third party from criminal liability), in some cases because there are no specific penal provisions addressing such omission and in others because it is seen as justifiable, without prejudice to any civil liability which could arise in connection therewith.

3. This context also brings about new issues regarding personal freedom of decision or autonomy. Individuals will now be faced with the decision of whether or not to submit to genome analysis, particularly when some genetic disease or future disposition to such a disease is suspected (eg, to enter into an employment contract or obtain positions entailing responsibility or risks, to take out a life insurance policy, obtain administrative authorizations or licenses, and other situations which will not be covered in this paper).[34] The question can be problematic when genome analysis is needed in order to accurately diagnose an embryo whose mother has a family history of genetic disorders, either as a precondition for abortion (where abortion is only allowed on certain medical indications, the most precise diagnosis possible is required; where abortion is not prohibited within the early stages of gestation, such diagnosis is not required) or in order to give the embryo "in utero" treatment. Lastly, the use of assisted reproduction techniques has the potential to produce pregnancies without the consent of the mother or pregnancies on terms not authorized by the mother (eg, the use of gametes or embryos from persons other than the couple, sex selection, implantation of a greater number of embryos than planned or needed).

4. In North America there is ongoing debate over whether antenatal screening (preconceptive and prenatal) screening should be compulsory for at-risk individuals and, if prenatal tests confirm the presence of fetal abnormalities, whether the mother should be obliged to abort, or whether this could be imposed as a condition for performing the prenatal tests, and whether such tests should be generally performed on at-risk population groups. These are negative eugenic proposals, which are only acceptable if the interested parties freely give their consent. Neither public authorities nor private parties can impose such measures coercively as they touch on the essence of a fundamental right.[35]

---

[34] See the Danish Council of Ethics: «Genetic Screening», *Ethics and Mapping of the Human Genome*, 1993, p. 43 and ff.

[35] See Eser: «La moderna medicina ...», *op. cit.*, p. 293.

5. Fetal therapy (or embryonic therapy, where such is the case) necessarily entails intervention in the mother's body in order to access the fetus (medications, transfusions, fetal surgery), even when the operation is performed *ex utero*, that is, extracting the fetus from the womb and later reintroducing it once the operation has been completed. The mother's consent is necessary for such therapeutic measures, but this raises the question of the legal implications of a mother's refusal to allow such treatment (for religious reasons, for example). The answer must consider the legal nature of the conflict involved, weigh the competing interests (conceived more broadly than when weighing the legally protected values involved), and judge the risks incurred by the mother if the intervention is performed as opposed to those which arise for the fetus if not. The conflict could develop as one between the mother's freedom, and in some cases her right to health and integrity, and the life of fetus, in extreme cases, or at least its health and integrity in others. Criminal law has mechanisms for solving such conflicts (which can give rise to conflicting duties), when depending on the decision adopted a criminal offence is committed which affects the above-mentioned legally protected values: the defence of acting on necessity or in fulfilment of a duty. As is known, both defenses are expressly recognized in the Spanish Criminal Code (section 8.7 and 8.11, respectively).[36]

6. A pregnancy brought about by assisted reproduction techniques (artificial insemination or embryo implantation) without the consent of the mother does not per se imply a physiological aggression against the physical integrity or health of the woman, unless it is contraindicated for her physical or mental health state. But it is hard to imagine any therapeutic purpose, or at least any primarily therapeutic purpose, for such an act, which would most likely be driven by spurious motives which need not be examined. It does represent, however, a serious aggression against the mother's freedom (and that of the couple), not only due to the coercive nature of the intervention itself but also because of imposition of an unwanted pregnancy and motherhood (and paternity).

7. All these hypothetical, unjustified, compulsive interventions are an attack on individual liberty and merit penal sanction. Depending on the circumstances, they may constitute an offence of illegal detention, threat or coercion, with the last one being the most common. They do not always

---

[36] For more on this see Romeo Casabona, Carlos M.: «La protección jurídica del concebido. El feto como paciente», *op. cit.*, p. 444 and ff.

constitute a criminal offence, however; the compulsion required by the respective criminal category (in Spain, section 496 of the Criminal Code) may not always be present, for in these cases it is particularly easy for the intervention to be attributed to deception or error, thereby removing criminal liability, leaving only the possibility of civil or administrative liability. Second, it is very doubtful that these criminal categories can satisfactorily address the seriousness of the acts involved. Consequently, specific criminal categories should be provided for the more serious acts of disregard for the consent of the interested party in relation to the imposition of certain situations (eg, an imposed pregnancy).

8. The above-mentioned freedom of decision and other fundamental rights (such as personal integrity, privacy, the right not to declare against oneself) also have bearing on the role of the genome in civil paternity suits or criminal cases. The results of genetic analysis can provide decisive evidence for confirming or refuting accusations. The question is therefore whether a person can be made to submit to such testing against his or her will, how should refusal be punished if at all, and the judicial value of such evidence. Of special interest is the investigation of the DNA of a suspected criminal offender (just as such methods are extremely useful for identifying disfigured crime or accident victims), the results of which could help to confirm the accusations and serve as decisive evidence (the so-called "genetic fingerprint"). But several major problems arise. First, there are questions as to the legitimacy of the procedure itself and thus of the probative value of evidence obtained coercively, that is, without the consent of the person. Even where a legal obligation exists to furnish such evidence, the person's right to self-determination with respect to the information is involved, as is his or her right to personal integrity. Second, in both North America and Germany criticism has been voiced of the facility with which courts of justice accept this evidence as irrefutable when serious doubts have been raised in many cases as to their credibility and reliability (because of the techniques employed, the state of the specimens, etc).[37] In other

[37] Nevertheless, section 81 of the German Criminal Procedure Act (StPO) has led the Supreme Court of that country (BGH: eg, judgements of August 21st, 1990 and July 3rd, 1990) to consider this evidence as well as the required blood sample as legitimate. The said article provides: «I. A corporal investigation of the accused may be conducted in order to determine such facts as are relevant to the proceedings. Blood samples and other corporal interventions performed without the consent of the defendant by a doctor for exploratory purposes in accordance with the rules of medical technique, provided they do not represent any danger to his health». See Jung, H.: «Zum genetischen Fingerabdruck», *Monatschrift für Kriminlogie und Strafsrechtsreform*, 1989, p. 103 and ff.

parts this concern has led to recommendations that in order to conduct such tests the laboratories must be accredited as conforming to certain conditions and guidelines that assure their reliability.[38] It is believed that these techniques are not currently defendable and their forensic application is premature due to the possibility of major errors. Further scientific discussion is needed before an acceptable standard is found.[39]

This type of investigation should be regulated by law,[40] on the basis of respect for the freedom and privacy of the person, without prejudice to the legal consequences which may arise from his or her actions.

3.1.4. Genome analysis and the protection of "genetic privacy"

Problems are raised concerning the use of information obtained or which could be obtained from the analysis of a person's genome. In particular, I will highlight two important issues of criminal law:

1. Knowledge of this information will reveal very important aspects of the person which directly affect his or her privacy. Disclosure would represent a serious danger, first of all, because of the danger of converting the person into a "transparent" or "glass" citizen. There is also the possibility, as we shall see further ahead, for discrimination of all kinds: personal, familial, employment, life insurance or pension plan, for obtaining certain official permits or authorizations, etc. Furthermore, the vast quantity of information generated will require its computerization, with all the

---

[38] See the Council of Europe *Recommandation No. R(92)l, du Comité des Ministres aux États membres sur l'utilisation des analyses de l'acide désoxyribonucléiques (ADN) dans le cadre du systéme de Justice pénale,* point 6. Also very interesting are its recommendations on the use of these samples and the information they provide (point 3), and on the protection of said information (point 7).

[39] In Germany, Rademacher, C., in «Neue Juristische Wochenschrift», 1991, p. 735, and Jung, in «Zum genetischen...», *op. cit.*, p. 105, are critical with the criterion cited by the BGH and hold that the legitimacy of the said tests cannot be supported solely by section 81a. The latter author believes that additional regulation is needed in view of the numerous problems posed by information obtained on the person (sensitive data). Conversely, Kleinknecht, Theodor and Meyer, Karlheinz, in *Strafprozessordnung Kommentar*, 39th ed., C.H. Beck, München, 1989, p. 275, support the high court's finding that the purpose of genetic analysis is legitimated by section 81a. On these issues in Spanish law, see the recent study by Peris Riera, Jaime M.: «La identificación genética y los derechos fundamentales [Genetic identification and fundamental rights], *Arbor*, no. 564, p. 45 and ff.

[40] See International Criminal Law Association, XIV International Congress on Criminal Law, Resolution 6.6.

advantages entailed therein but also with the increased vulnerability of the processed data to uncontrolled or unauthorized dissemination.[41]

Hence, assurances of confidentiality as a means for protecting the individual's privacy become of paramount importance. The requirements of legal protection of data and of professional secrecy also embrace the genetic information of each individual, who shall decide to whom, when and to what extent such information may be disclosed. There must therefore be a prohibition on the disclosure of a person's genome information to third parties without the consent of the person or his or her legal representatives, even where the information is requested by a family member seeking to determine his or her possible predisposition to a pathological gene discovered in the analyzed person.[42]

Legal specialists in Spain unanimously underscore the insufficiency of the current criminal justice provisions for the protection of secrecy, although the protection is somewhat more satisfactory when a public employee is involved (sections 367 and 368 of the Spanish Criminal Code). Protection under civil law is available thanks to Organic Act 1 of May 5th, 1982 on the civil protection of the right to personal and familial privacy, honour and personal reputation. The said Act regards as wrongful intrusion the disclosure of an individual or family's private information obtained in the course of a professional or official activity by the person responsible for the disclosure (section 7.4). The recently promulgated Organic Act on the regulation of the automated processing of personal information of October 29th, 1992 strengthens the protection of health information and sets down conditions for its automated processing by health professionals and public and private health-care centres. These provisions basically defer to current health legislation regulating access to information and confidentiality (sections 7.3 and 8). This will soon be seen to be insufficient given the nature of this information and of the persons who could access it. One of the laws which this Act cites as reference is the General Health Act of 1986, which recognizes confidentiality as a right of the patient (sections 10.3 and 61).

2. As a result of the advances in biomedical sciences in general and, in particular, of human genetics and its predictive possibilities, there has arisen, along with the right to information, or the right to know, the right

---

[41] See Romeo Casabona, Carlos M.: *Poder informático y seguridad jurídica*, Fundesco, Madrid, 1988, p. 12 and ff.

[42] For more on these problems from an ethical and legal standpoint, see The Danish Council of Ethics, «Protection of Sensitive Personal Information. With Special Reference to Genetic Data», in *Ethics and Mapping of the Human Genome*, *op. cit.*, p. 5 and ff.

"not to know".[43] It is considered legitimate and respectable for a person to wish not to be informed as to the future onset of a disease, particularly in connection with mortal diseases for which effective therapy is not yet available, in order that such information cannot condition his or her personal expectations. Infringements of this right should in the sphere of criminal law be treated as injuries to the individual's personal liberty and it must be ascertained whether the corresponding criminal offences are adequately defined so as to cover the means used for such infringements.

3.2. *Collective interests: the genetic integrity (and other similar values) of humanity*

Although genetic manipulation and genetic engineering directly affect specific individuals, such actions on individual human genomes clearly have the transcendent capacity to affect the human species as a whole, its integrity, identity, individuality and inalterability. Hence the misgivings they awaken. Positive eugenics, genetic manipulation for racial or racist purposes, and those that in any way degrade the human species, say, by mixing it with animals, are widely rejected. There is concern with staving off any "modernized" resurgence of the eugenic trends which arose in the Anglo-Saxon world at the beginning of the century, and the abuses committed under the German national-socialist regime. Even therapeutic genetic interventions, the legitimacy of which is generally accepted (with all the reservations mentioned above when affecting the germ-line) do not fail to stir concern and doubts as to their potential impact on altering the genome of the human species. That potential is openly discussed in relation to germ-line therapy. There is discussion of its possible use to fortify human resistance to certain diseases (eg, cancer and other, viral diseases) or for negative eugenics, ie, the elimination of pathological genetic material.[44] In keeping with the spirit of the above-mentioned moratorium until the consequences and effects of germ-line therapy are better known, it also seems advisable to refrain from definitive pronouncements on this eugenic proposal until we are in a better position to ascertain its possibilities and effects, and to control any undesirable effects if necessary. Judged in the hypothetical realm, however, the idea should be seen in a favourable light.

But the question which still stands is: are there new collective interests which merit legal protection and, if so, should the conducts that seriously

---

[43] See Rodota, Stefano: «Le droit aux dilèmmes de la vie et de la mort», in *XXth Colloquy on European Law*, Glasgow, 1990, Council of Europe, Strasbourg, 1992, p. 14.

[44] For more on this see Eser: «La moderna medicina de la reproducción ...», *op. cit.*, p. 294.

endanger them be subject to penal sanction? We could cite certain interests which concern humanity as a whole, such as has been traditionally understood in connection with the right of peoples and the crime of genocide. These generic values would serve to protect the inalterability of certain characteristics of the human species, its genetic variety and plurality and, in the most extreme cases, the very survival of the human species, from the possible threats posed by eugenics or other biotechnological or genetic engineering propositions.

Although this line of argument needs to be explored more deeply, as a preliminary step I will outline some considerations which might be taken into account:

1. The inalterability and intangibility of the human genetic heritage would give rise to the following prohibitions:[45] human genetic exchange (formation of hybrids); interspecies fertilization or interspecies transfer of embryos (from animals to human beings or vice versa) for any purpose or at any stage of development, except for certain time-limited diagnostic tests; the fusion of human pre-embryos or of interspecies pre-embryos (chimeras) for any purpose or at any stage of development; other actions aimed at altering human, non-pathological genetic heritage. The last prohibition once again points to the need to distinguish between conditions which are pathological and those which are merely anomalous, different, rare or infrequent or slightly deviant.[46] These allusions to the normal and the pathological, which can only be mentioned here, are of pivotal importance to a reflection on future applications of genetic engineering. This reflection must not be confined to the sphere of biomedical analysis but taken up on the philosophical, sociological, ethical and legal terrain. The proposals put forward above with regard to germ-line gene therapy should also be taken into account in this connection, but I will here call special attention (as a clear illustration of the valuative and dialectical tensions which are potentially present and compel further reflection and discussion) to the fact that this emerging right to an unaltered genetic heritage could be opposed by a no less legitimate right to a healthy genetic heritage. Is this conflict real or apparent? The latter is most likely true; in reality both interests can be tended to, conciliated and even harmonized, but, as I have held

---

[45] Vale Muñiz and González González: «Utilización abusiva de técnicas genéticas y Derecho Penal», *op. cit.*, p. 131 and ff.

[46] On this last aspect see Eser: «La moderna medicina de la reproducción e ingeniería genética. Aspectos legales y sociopolíticos desde el punto de vista Alemán», *op. cit.*, p. 291; Knoppers: «Human Dignity and Genetic Heritage», *op. cit.*, p. 43 and ff.

throughout, only when future knowledge has allowed us to more precisely understand the effects of intervention in the germinal line.

In connection with some of these questions, the Council of Europe's Recommendation of January 26th, 1982[47] asked that the catalogue of human rights include "the intangibility of genetic inheritance vis-a-vis artificial interventions" and that the necessary protection be provided by the adoption of the corresponding laws.[48] The International Criminal Law Association, for its part, proposes that this right be protected by law, and that the transfer of human gametes for non-therapeutic purposes be prohibited. The association expressly advocates criminalization of experiments aimed at generating hybrids or chimeras through the fusion of human and animal cells.[49]

2. The identity and uniqueness of human beings, as the right to individuality, to be oneself and different from others, accepting the complexity and variability of human nature: the creation of identical human beings by means of cloning[50] or other genetic procedures for any purpose (race selection, creation of "specialized" human beings).

3. Privation of the double genetic endowment, and thus of the male and female genetic line, to which could be added the potential risks to the health of the person so procreated: parthenogenesis, or the generation of a human being from only one gamete (egg); generation of embryos from gametes of persons of the same sex (xx, female: xx, xy, male); and from a different standpoint, ectogenesis, the development of an individual in a laboratory, although today this seems impossible.

4. Protection of the survival of the human species: the creation of biological or other weapons using genetic engineering; alterations of the environment by means of genetic engineering, particularly through the release of genetically modified microorganisms, so as to make it dangerous for or incompatible with human life.

Without preempting the further discussion and development warranted by these issues (including the identification of the legal interests involved), in principle, all these conducts are seriously detrimental to humanity and

---

[47] Council of Europe Recommendation number 934 of 1982.

[48] Knoppers, in «Human Dignity and Genetic Heritage», *op. cit.*, p. 29 and ff, however, alludes to the risks which could arise from the assumption of that right.

[49] International Criminal Law Association, in Resolutions 6.1, 6.8 and 6.10, respectively, adopted at its XIV International Congress on Criminal Law.

[50] In this regard see the International Criminal Law Association, Resolution 6.9 of the XIV International Congress on Criminal Law.

should be criminalized. I will not take up here whether the corresponding criminal categories should be configured as injurious offences or, in some cases, as hazardous offences. Nor will I study whether the Criminal Code should include provisions in respect of all these conducts now or wait until they are scientifically possible.

4. Final considerations

The ideas, reflections, judgements, suggestions and conclusions presented throughout this paper must of necessity be considered as only provisional. The area before us is too complex for it to be otherwise. Despite the rapid advances and understanding attained, this field of exploration is still in its preliminary stages and hence our grasp of its potential and possibilities is likewise only beginning. Such precaution is further warranted by the fact that the future of humanity is more obviously affected by these developments than by those in any other area.

Hence the fundamental task before us is to remain attentive to the ongoing investigations and their results, to persist in interdisciplinary debate, to be cautious in legislative policy and open and willing to review and revise as necessary the criteria thus far adopted. If we are to assume that investigation of the human genome should not be curbed, we must also accept that it has to be controlled and that its possible applications and derivations must be analyzed in each case. This should not necessarily stir any misgivings from the scientific community, for the future of humanity concerns us all.

III.

ETHICAL ASPECTS

IDENTITÉ HUMAINE ET GÉNÉTIQUE

JEAN-MICHEL MALDAMÉ *

Le progrès des sciences au cours de ces dernières années est lié à l'émergence de nouvelles méthodes et de concepts aptes à saisir la complexité.[1] C'est dans ce nouvel horizon de la pensée scientifique que se situe la génétique; elle dépasse les modèles d'explication mécaniste (où la seule causalité est un jeu de forces déterministes — la causalité efficiente), pour réintroduire la finalité grâce à la notion scientifique de plan d'organisation on encore de téléonomie.[2] Les concepts de la génétique réintroduisent également le sens de l'unité de l'homme exprimée par le concept d'organisme.[3]

Les deux concepts majeurs sont aujourd'hui ceux d'organisation et d'information.[4] Ils donnent son originalité à la notion de gène, dans la

---

* Ecole de Théologie, France.

[1] Cf. François Jacob, *La Logique du vivant. Une histoire de l'hérédité*, Paris, Gallimard, 1970; Jean-Louis Le Moigne, *La Modélisation des systèmes complexes*, Paris Dunod, 1990. François Jacob écrit: «L'hérédité se décrit aujourd'hui en termes d'infomation, de messages, de code. La reproduction d'un organisme est devenue celle des molécules qui le constituent. Non que chaque espèce possède l'aptitude à produire des copies d'elle-même. Mais parce que la structure des macro-molécules est déterminée jusque dans le détail par des séquences de quatre radicaux chimiques contenus dans le patrimoine génétique. Ce qui est transmis de génération en génération, ce sont les «instructions» spécifiant les structures moléculaires. Ce sont les plans d'architecture du futur organisme. Ce sont aussi les moyens de mettre ces plans à exécution et de coordonner les activités du système. [...] Chaque oeuf contient donc, dans les chromosomes reçus de ses parents, tout son propre avenir, les étapes de son développement, la forme et les propriétés de l'être qui en émergera. L'organisme devient ainsi la réalisation d'un programme prescrit par l'hérédité. A l'intention d'une Psyché s'est substituée la traduction d'un message» (p. 9-10).

[2] François Jacob, *La Logique du vivant. Une histoire de l'hérédité*, Paris, Gallimard, 1970.

[3] Cf. Georges Canguilhem, *La Notion d'organisme*, Paris, Vrin, 1972.

[4] Cf. François Jacob, *op. cit.*: «Avec le concept de programme appliqué à l'hérédité, disparaissent certaines des contradictions que la biologie avait résumées par une série d'oppo-

mesure où ils ne font pas du gène une chose isolée, car ils envisagent sa fonction en corrélation avec les autres parties de la cellule qui est la base de l'être vivant.[5] En effet, les gènes président au développement et à l'organisation du corps humain. Ils en assurent la permanence.[6] Les scientifiques parlent de l'action du génome en termes de programme ou de codage.[7] C'est là un nouvel ordre de conceptualité scientifique.

Au sein de cette nouvelle vision de la science, les gènes expliquent la permanence de l'espèce, mais aussi le mécanisme des variations individuelles qui préside à l'irréductible singularité humaine.[8] Il y a un génome humain bien défini et, dans cet ensemble, il y a place pour des combinaisons selon un agencement unique. Il y a des gènes pour chaque détail: la couleur de ses yeux, la taille, la morphologie générale, ... Mais aussi la santé, la capacité de résistance à certaines agressions, la fragilité face à certaines maladies. La question se pose alors: peut-on étendre cette explication à ce qui en l'homme est irréductible à l'animalité, l'activité intellectuelle ou spirituelle?

Ainsi, au terme de l'heureuse révolution scientifique dont peut bénéficier la médecine, la question revient avec force: qu'est-ce que l'homme? La question se pose d'une manière nouvelle à partir de la génétique, parce que celle-ci apporte des connaissances sur le patrimoine transmis de

---

sitions: finalité et mécanisme, nécessité et contingence, stabilité et variation. Dans l'idée de programme viennent se fondre deux notions que l'intuition avait associées aux êtres vivants: la mémoire et le projet. Par mémoire, s'entend le souvenir des parents que l'hérédité trace à l'enfant. Par projet, le plan qui dirige dans le détail la formation d'un organisme» (p. 10).

[5] Cf. François Gros, *Les Secrets du gène*, Paris, Seuil, 1986: «Ce qui caractérise le mieux le gène aujourd'hui, ce n'est pas sa matérialité physique et chimique au niveau de l'ADN [...], ce sont bien davantage les produits qui résultent de son activité: ARN messager cytoplasmique et protéine» (p. 297).

[6] «Le programme génétique [...] est constitué par la combinatoire d'éléments essentiellement invariant. Par sa structure même, le message de l'hérédité ne permet pas la moindre intervention concertée du dehors. Chimiques ou mécaniques, tous les phénomènes qui contribuent à la variation des organismes et des populations se produisent en toute ignorance de leurs effets. [...] La nature même du code génétique empêche tout changement délibéré du programme sous l'effet de son action ou du milieu» (François Jacob, *op. cit.*, p. 11).

[7] Cf. François Gros, *Les Secrets du gène*, Paris, Seuil, 1986, chapitre IV: «La cybernétique des gènes et chapitre V: «Code et messager».

[8] Cf. P. van Gansen, *L'Homme et ses gènes*, Bruxelles, Vigot, 1992: «A l'exception des jumeaux, tous les individus d'une même espèce sont génétiquement distincts [...]. Il y a 2^{23}, soit plus de 5 millions de formules génétiques différentes dans les cellules sexuelles. Au moment de la fécondation deux cellules sexuelles fusionnent pour engendrer un embryon diploïde. Il y aura plus de 25.000 milliards de combinaison génétique pour cet embryon. Par conséquent, deux enfants d'un même couple, à moins d'être jumeaux vrais, n'ont aucune chance d'avoir la même formule génétique» (p. 31).

génération en géneration. Des magazines, des revues de vulgarisation — comme *Science et vie* — et même des revues scientifiques — comme *Nature* — annoncent que la connaissance du patrimoine génétique permettrait d'expliquer non seulement des maladies somatiques, mais des comportements sociaux ou sexuels, tels que l'homosexualité ou la délinquance.[9]

Son destin est-il inscrit une fois pour toutes dans ses gènes? Si tel est cas peut-on encore parler d'âme, de responsabilité et de liberté?

I. L'homme, un être en humanité

Pour répondre à la question de l'identité humaine en respectant les nouvelles connaissances, il serait maladroit de partir d'une définition purement spirituelle, car ce serait cautionner une anthropologie dualiste. Je resterai donc au plus près de la biologie. Plutôt que de parler *a priori* d'âme rationnelle et immortelle, j'emploierai le terme médical de psychisme, même si ce terme est plus pauvre que celui d'âme. Dans le prolongement du fait que les gènes constituent un patrimoine reçu qui informe l'organisme humain, je dois considérer l'ontogénèse du psychisme: est-elle liée au génome? Si oui, comment?

Pour répondre à cette question, je propose une réflexion sur la constitution du corps humain. Dès la fécondation de l'ovule, l'être vivant forme une unité; mais les organes et les systèmes qui le composent ne s'intègrent en un tout fonctionnel que progressivement, comme le montrent les résultats de l'embryologie — n'étant pas spécialiste de ces questions, je les reprends de mes lectures, en particulier de l'ouvrage écrit par le Professeur André Bourguignon.[10] C'est en étudiant la constitution progressive de l'être humain que l'on pourra voir ce qui est spécifiquement humain.

1. *Place du système nerveux central*

Tout le monde sait qu'au cours de son développement, l'être vivant ne grandit pas passivement. C'est lui qui se construit psychiquement à partir de ses besoins, de ses affects et de ses actions, en fonction égale-

---

[9] Cf. Benno Mueller-Hill, «Le spectre de l'injustice génétique», *La Recherche*, décembre 1993, p. 1386-1389.

[10] André Bourguignon, *Histoire naturelle de l'homme*, t. 2: *L'homme fou*, Paris, PUF, 1994, chapitre II: «L'ontogénèse du psychisme humain», p. 75-112, cf. *Histoire naturelle de l'homme*, t. I: *L'homme imprévu*, Paris, P.U.F., 1989.

ment de ce qui lui est offert.[11] Toujours animé par l'envie constante de dépasser ses capacités du moment, le «petit d'homme» anticipe sur ses possibilités. L'envie de marcher précède la marche; l'envie de parler, le langage, ... Ceci est lié au système nerveux central. Aussi pour répondre à la question posée sur l'identité humaine, l'étude de la constitution du système nerveux central apparaît comme nécessaire.

Pour que l'être vivant se développe de manière autonome en maintenant sa cohérence interne et externe, il faut que les circuits neuronaux correspondant aux fonctions et aux besoins élémentaires soient stabilisés en temps opportun tout en gardant une grande souplesse de fonctionnement. Comment ceci est-il mis en place? Quel est dans ce processus le rôle du génome humain? Tous les circuits neuronaux sont-ils déterminés par le génome? Sinon, comment sont-ils stabilisés?

Il est d'observation commune que l'écart entre le psychisme adulte et celui du nourrisson est immense et que l'enfance et l'adolescence sont des temps d'épreuve difficiles. De manière plus scientifique, on peut dire que le système nerveux central humain est caractérisé par la précocité relative de sa mise en place, la lenteur de son développement et la complexité structurelle et fonctionnelle de son organisation.[12] Pour répondre à la question posée, il faut donc interroger la science.

2. *Ontogénèse du système nerveux central humain*

Dès le début de la vie intra-utérine du foetus humain, les neurones prolifèrent[13] et migrent vers le territoire qui leur est assigné. Une fois mis en place, les groupes de neurones entrent en connexion avec leur structure cible et à ce moment beaucoup disparaissent afin d'ajuster leur nombre aux besoins fonctionnels de leur territoire cible. Dès l'âge de 8 semaines, le cortex du foetus est différencié, mais immature; à quatre mois les hémisphères cérébraux sont encore au stade de la migration des neuroblastes. Ce n'est qu'à trente-quatre semaines que la glie se sépare des neurones. La formation des synapses débute dans le tronc cérébral et les noyaux gris centraux, pour s'achever dans le cortex, en s'accélérant à la 35^{e} semaine. A la naissance, le cerveau a pratiquement acquis tous ses neurones et la neurogenèse ne se poursuit que dans quelques structures.[14]

---

[11] *Op. cit.*, p. 75.

[12] *Op. cit.*, p. 76.

[13] A un rythme de l'ordre de 250.000 par minute!

[14] La neurogénèse ne se poursuit que dans quelques structures comme le cervelet et l'hypocampe; la myélinisation des axones corticaux se poursuit pendant deux ans.

Pendant les deux premiers mois après la naissance, le néocortex est peu développé, ce qui n'autorise que des mouvements réflexes et des mouvements de poursuite visuelle et auditive. A trois mois, certaines parties du néocortex deviennent fonctionnelles, comme le prouvent la sensibilité aux formes et la possibilité de différer une réaction à un stimulus interne comme la faim. Les deux premières années sont consacrées au développement des aires de projection primaires, ce qui limite les processus cognitifs à une activité assez élementaire. Ce n'est qu'à la fin de l'âge préscolaire que l'enfant peut faire des choix indépendants de la situation présente. Il faut 15 à 20 ans pour que le cerveau humain achève sa maturation, mais son aptitude au changement persiste jusqu'à un âge avancé.[15]

Je relève qu'un tel processus met en oeuvre une double action, celle imposée par le génome, dont la moitié des gènes commandent la synthèse des protéines cérébrales, et celle imposée par l'environnement dans lequel grandit et se développe l'individu. Je relève donc deux aspects, une détermination génétique et une détermination épigénétique.

3. *La détermination génétique*

La détermination génétique est responsable de l'organisation générale du système nerveux central, de son câblage et, pour une part, de sa spécification fonctionnelle. Elle assure la reproduction relativement invariante des individus d'une même espèce. Parce que la distribution des gènes est aléatoire, les individus d'une même espèce diffèrent les uns des autres. C'est donc la variabilité génétique qui explique pourquoi les nouveau-nés sont différents les uns des autres dans leurs attitudes fondamentales face au stimuli, pour le sommeil, le tonus, l'appétit.[16]

La détermination génétique explique aussi l'entrée en activité de ceux des neurones qui participent aux fonctions vitales et aux comportements nécessaires à la survie. Mais cette intégration à détermination génétique ne concerne que les fonctions élémentaires. En effet, pour spécifier, puis intégrer les fonctions nerveuses supérieures, il faut l'intervention de l'environnement, lors d'expériences sensori-motrices et psychiques, cognitives et socio-affectives.

Au risque de simplifier, on peut quantifiter et relever une certaine disproportion entre le génome et ses cent mille gènes (100.000), et

[15] *Op. cit.*, p. 76.
[16] *Op. cit.*, p. 78.

la tâche requise de spécifier cent mille millions de neurones (10^{11} ou 100.000.000.000) et un million de milliards de synapses (10^{15} ou 1.000.000.000.000.000).

4. *La détermination épigénétique*

La détermination épigénétique intervient dès la conception,[17] mais c'est surtout après la naissance que ses effets amplificateurs ou réducteurs des différences se font, car non seulement le génome contient trop peu d'instructions pour ajuster au milieu les fonctions neuropsychiques complexes de la vie, mais surtout il ne peut prévoir dans quel milieu le nouveau-né achèvera son développement. C'est pourquoi la détermination épigénétique doit nécessairement compléter les informations génétiques d'un cerveau immature.[18]

A la naissance, les circuits neuronaux sont en place, mais en dehors de ceux qui assurent les fonctions vitales (respiration circulation, ...), seuls quelques circuits indispensables à la survie sont déjà spécifiés par les gènes. Les autres attendent du milieu leur spécification fonctionnelle. Leurs synapses qui sont à l'état labile ne seront stabilisées et fixées dans leur fonction qu'après leur entrée en activité.

Il est important de relever l'importance du temps, car il existe donc pour chaque structure neurospyschique labile une période sensible, génétiquement déterminée quant à son apparition et à sa durée, pendant laquelle s'établit la fonction.[19] Le génome met en place un réseau peu ou pas spécifié, dans lequel l'environnement sélectionne les synapses les plus

---

17 L'environnement intra-utérin exerce une action sur le système nerveux central, dans la mesure où certaines variations survenant dans le corps maternel peuvent se répercuter sur l'embryon et le foetus: agents biologiques, mais aussi variations émotionnelles imposées à la mère par son propre environnement. *Op. cit.*, p. 81.

18 On caractérise l'état immature du cerveau par les traits suivants: 1. le petit nombre de neurones activés ou activables; 2. la lenteur de la conduction des signaux neuronaux; 3. la faiblesse du débit des impulsions neuronales en rapport avec une transmission synaptique malaisée; 4. la sensibilité des neurones à l'environnement, à une certaine phase de leur maturation. Il résulte que le cerveau à la naissance a moins de fiabilité et de redondance qu'à l'âge adulte et surtout qu'il possède une plasticité et une vulnérabilité plus grande. *Op. cit.*, p. 80.

19 La plasticité se manifeste pour certains neurones, lors de périodes dites critiques, ou plus exactement sensibles, correspondant à l'apparition plus ou moins rapide de constituants chimiques, de structures microscopiques et d'activité spontanée ou provoquée. Pendant ces périodes sensibles, le cerveau est particulièrement réceptif aux influences environnementales, c'est-à-dire qu'il est alors éminemment vulnérable. Toutes les structures et fonctions ne se développant ni en même temps, ni au même rythme, il y a une multitude de périodes sensibles. *Op. cit.*, p. 81.

actives et trace ainsi les circuits spécifiques, tandis que les synapses inactives dégénèrent. Ceci est l'objet d'observations médicales obtenues souvent à partir de traumatismes. Mais cela recoupe l'observation heureuse de la croissance des enfants, de l'apparition du sourire, de la vision, du langage, qui font l'émerveillement devant la vie.

Dans l'examen du dossier médical, je relève aussi que le développement n'est pas continu. Il y a des remaniements qui sont de refontes globales des circuits neuronaux.[20] On observe également que les périodes sensibles sont d'autant plus longues que la croissance est plus lente et que la fonction se situe à un plus haut niveau d'intégration.[21]

5. *Une vie humaine*

Le psychisme a donc une histoire. Il se développe dans une interaction étroite entre le petit d'homme et son environnement. Les principales étapes sont au coeur de toute pédagogie et font partie de la vie des familles et des groupes où nous vivons. Le petit d'homme a besoin de l'adulte pour vivre. Il est élémentaire de constater que l'homme ne vit pas que de pain, mais de présence, d'affection et de paroles à lui adressées. La structuration de son psychisme est corrélative de cette activité. L'étude de la croissance permet de relever quelques étapes où l'épigénèse est essentielle pour actualiser les dispositions génétiques.

De la conception à la naissance, l'enfant est dans le sein maternel. Il n'est pas dépourvu de sensibilité: la main, la plante des pieds, puis toute la peau deviennent sensibles dès la 14^{e} semaine. A la 26^{e} semaine, l'enfant entend les sons émis par sa mère. Le goût est développé. Mais c'est très peu de chose. Tout commence vraiment à la naissance.

Dans les premiers mois, les activités s'accompagnent d'un engagement global de tout l'organisme (mouvement respiratoire, rythme cardiaque, ...). Il y a peu à peu une sélection corrélative de l'attitude des adultes. Les perceptions sont reprises dans une activité qui intègre les informations venues sur le monde. Vers trois mois, le contrôle des membres est acquis et l'enfant fait la différence entre ce qui lui est intérieur et ce qui lui est extérieur. A la fin de la première année, l'enfant se déplace et donc se met lui-même face à des situations non prévues. Tout

---

[20] La marche soutenue disparaît à 8 semaines et réapparaît au 9^{e}-12^{e} mois; la coordination oeil-main cesse entre la 4^{e} et la 20^{e} semaine; l'imitation de la mimique présente à 2-3 semaines prend fin très vite pour ne revenir qu'à la fin de la première année ... *Op. cit.*, p. 89.

[21] *Op. cit.*, p. 89.

ceci prend une forme nouvelle avec l'acquisition du langage qui est la marque d'une vie psychique spécifiquement humaine.[22]

Les relations humaines sont pleinement humaines. Car le langage est porteur d'une vision du monde dans une culture donnée. Il contribue à la fois au refoulement de l'inconscient et à l'instauration de la loi qui marque des interdits et définit une maîtrise de l'espace, du temps et de la relation à la culture. Le langage une fois instauré, la poursuite de la croissance dépend à la fois de facteurs individuels et de facteurs socio-culturels.[23]

Le fonctionnement neurologique du cerveau montre que le petit d'homme qui vient au monde est apte à entrer en relation avec son environnement et d'une manière privilégiée à l'environnement humain. Paradoxe de l'être humain qui naît inachevé et qui doit devenir lui-même par lui-même! Défi de l'humanisation, car l'humanité n'est humaine qu'en relation. Ce qui est inscrit dans la chair même, n'est pas réductible aux échanges décrits par la science. L'homme n'est pas un système biologique constitué indépendamment d'une activité spirituelle. Les recherches modernes sur l'âme qui vont dans ce sens sont fécondes, tant dans la perspective de la phénoménologie que dans la perspective d'un renouveau de la tradition aristotélicienne, illustrée par Paul Coutagne.[24]

III Anthropologie et dignité de l'homme

La réflexion faite à partir des données de la médecine montre qu'il y a un psychisme irréductible au déterminisme génétique. Est-ce suffisant pour répondre à la question de l'identité de l'homme? Non, car s'il est acquis que l'on ne puisse accorder aux gènes toute la responsabilité de la mise en place des circuits neuronaux, il faut se demander si le fonctionnement des circuits neuronaux du cerveau peut rendre raison de ce qui est spécifiquement humain. Je poserai cette question à partir des travaux de l'*Homme neuronal* de Jean-Pierre Changeux,[25] qui s'incrivent

---

22 *Op. cit.*, p. 109.

23 *Op. cit.*, p. 110.

24 Cf. Paul Coutagne, «L'homme neuronal a-t-il une âme?», dans «Le philosophe peut-il encore parler de l'homme?», *Bulletin de l'Institut catholique de Lyon*, n° 83, juillet-septembre 1987, p. 29-38; «L'homme au prisme des sciences de la vie», dans *Après Galilée*, sous la direction de Paul Poupard, Desclée de Brouwer, 1994, p. 159-196; cf. également l'ouvrage collectif, *La Peau de l'âme*, Paris, le Cerf, 1994.

25 Jean-Pierre Changeux, *L'Homme neuronal*, Paris, Fayard, 1983 — paru ensuite en édition de poche Fayard-Pluriel, 1984; en collaboration avec Alain Connes, *Matière à pensée*, Paris, Odile Jacob, 1989; *Raison et plaisir*, Paris, Odile Jacob, 1994. Dans *L'Homme neuronal*,

dans la perspective de mon approche, attentive à ne pas dissocier a priori l'esprit et le corps, le somatique et le psychique. Cet ouvrage n'ignore pas l'importance de l'épigénèse; mais parce que, pour son auteur, ce n'est qu'un processus d'individuation,[26] il n'y a pas en lui les éléments suffisants pour une reconnaissance de ce qui est spécifiquement humain: sa liberté et sa responsabilité.

1. *Liberté et conscience morale*

Pour J.-P. Changeux, l'esprit est inutile.[27] Aussi de même que l'analyse de la genèse du psychisme a montré que la complexité du fonctionnement du cerveau ne peut être réduite à l'influence exclusive des gènes, je dois montrer que la complexité de la vie de l'esprit ne peut être réduite au fonctionnement du système nerveux central. Je reste ainsi fidèle à la méthode employée jusqu'alors, dans le prolongement des travaux scientifiques proposant des modèles soumis ensuite à vérification.

Au terme de son livre, J.-P. Changeux donne comme exemple, vérifiant la justesse de son interprétation, la naissance de l'écriture,[28] où il use

---

J.-P. Changeux se propose de présenter le «phénomène humain» (p. 358-360/330-331) à partir des sciences du système nerveux, sciences rigoureuses, par opposition aux sciences humaines. Son travail récuse la séparation entre «l'organisation fonctionnelle du système nerveux et son organisation neurale» (p. 364/335). La première édition de l'*Homme neuronal* étant épuisée, nous donnons une double référence de pagination, celle de 1983, puis celle de 1984.

[26] Cf. *op. cit.*, chap. VII «épigénèse». En particulier, les paragraphes sur «l'empreinte culturelle», qui se conclut ainsi: «En fait, si la théorie [celle d'une épigénèse par stabilisation sélective des neurones et synapses en développement] s'avère exacte, l'activité (spontanée ou évoquée) ne travaille que sur des dispositions des neurones et de connexions qui pré-existent à l'interaction avec le monde extérieur. L'épigénèse exerce sa sélection sur des agencements synaptiques préformés. Apprendre, c'est stabiliser des combinaisons synaptiques établies. C'est aussi *éliminer* les autres» (p. 239-304).

[27] Cf. *op. cit.*: «Cet ensemble d'observations et de réflexions conduit non seulement à prendre en compte les mécanismes internes du comportement, mais à adopter vis-à-vis d'eux un point de vue déterministe. Rien ne s'oppose plus sur le plan théorique à ce que les conduites de l'homme soient décrites en termes d'activités neuronales» (p. 169-159); «A quoi bon parler d'"Esprit"?» (p. 364-335); «L'homme n'a dès lors plus rien à faire de l'"Esprit", il lui suffit d'être un homme neuronal» (p. 227-211).

[28] Cf. *op. cit.*: «Dans son exploration du monde, le scientifique procède en deux étapes. Il focalise son attention sur un objet, élabore, avec plus ou moins de succès un "modèle" de cet objet et s'accorde à considérer celui-ci comme une "représentation" simplifiée et formalisée de l'objet réel qui lui est extérieur. Mais, comme chacun de nous en a l'amère expérience, tous les modèles envisagés à un moment donné ne sont pas nécessairement retenus par la suite. La confrontation avec les résultats de l'expérience, parfois valide le modèle, parfois l'anéantit! Par exemple, le modèle proposé s'applique-t-il à une invention qui a changé la face du monde, celle de l'écriture?» (p. 368-338).

de la notion d'image, liée à la notion de représentation. Pour l'*Homme neuronal*, l'activité du cerveau consiste à produire des représentations. «Percepts, images de mémoire et concept constituent des formes ou des états divers d'unités matérielles de réprésentation mentale» qui sont regroupées sous le terme général d'«objets mentaux».[29] Le langage se réduit à la gestion des représentations et des images qui sont des objets mentaux.

Pour l'*Homme neuronal*, l'esprit n'est rien de plus que l'association de représentations: «L'aptitude fondamentale de l'encéphale des vertébrés supérieurs, et en particulier de l'homme, est de construire des "représentations", soit à la suite d'une interaction avec l'environnement, soit, spontanément, par focalisation "interne" de l'attention. Si on adopte la théorie proposée, ces représentations s'échaffauderaient par mobilisation des neurones dont la répartition au niveau des multiples aires corticales déterminerait le caractère figuratif ou "abstrait"».[30] Cette conception est-elle satisfaisante? N'ignore-t-elle pas trop de ce qui est humain? Pour y répondre il faut poser la question de la conscience, comme activité cérébrale.

2. *La conscience de soi*

La conscience est une activité de connaissance. En acceptant de relever dans le processus de connaissance uniquement ce qui concerne la perception et la représentation comme activité cérébrale, il y a lieu de se poser la question de l'auto-représentation.

Le système cérébral peut-il se percevoir lui-même? Peut-il être conscient de lui-même? Qu'il y ait une perception d'ensemble des activités et une auto-régulation dans le fonctionnement d'un système qui se forme autour d'un équilibre est une chose incontestable, mais peut-on pour autant identifier la perception avec la conscience, définie comme une saisie de soi par soi? Non!

D'abord pour des raisons logiques. Lors de l'élaboration de la théorie des ensembles, on a relevé qu'il était contradictoire de parler d'un «ensemble de tous les ensembles». Il en va de même pour un système qui est formalisé scientifiquement comme ensemble d'éléments en relation les uns avec les autres selon des règles précises et structurées. Aussi la fonction de régulation que le système nerveux central remplit à l'égard

29 *Op. cit.*, p. 181-169; une référence philosophique est faite à l'oeuvre de Wittgenstein; cette citation réduit le concept à un «jeu de langage» et s'oppose à *Matière et mémoire* de Bergson.

30 *Op. cit.*, p. 366-336.

de lui-même ne saurait-être appelée conscience, parce que celle-ci ne prend pas pour objet son propre fonctionnement. La conscience n'est pas l'auto-régulation décrite par la biologie. De fait, s'il y a une science du cerveau, elle n'est pas le fait du cerveau, mais de l'homme en son entier — dans la complexité de son corps, inséparable du langage et des relations à autrui.

En second lieu, nous constatons ici un dépassement analogue à celui que nous avons relevé plus haut: de même que les gènes ne peuvent présider à toute la constitution du cerveau, de même l'auto-régulation neuronale ne peut présider à la vie consciente d'un sujet humain qui se réfère à la totalité de son corps et de ses activités corporelles.

La conscience n'est pas la redondance des connexions neuronales ou le bouclage du cerveau sur lui-même. Elle est une saisie de la totalité de l'activité corporelle et langagière de l'homme. Elle ne peut être le fait que d'une instance qui ne se réduit pas à un organe. Il y a conscience de soi par un sujet humain et non pas la connaissance qu'une partie de l'organisme prendrait des autres parties et d'elle-même. Le cerveau fonctionne aveuglément. Il ne peut dire qui il est.

L'écriture du livre de J.-P. Changeux le montre: l'étude du cerveau suppose une distance qui ne réduit pas la connaissance de soi à une auto-régulation cérébrale. Il y a dans la connaissance une intentionalité, et plus encore chez l'homme: la réflexion.

3. *La question du sens*

La réflexion est-elle seulement une régulation interne? Certes, le modèle cybernétique utilisé en biologie permet de dépasser le mécanisme de l'âge classique. Il permet de justifier un équilibre par la prise en compte des divers éléments et donc d'une auto-régulation, puisque l'auto-régulation suppose une gestion des perceptions et des affects. Mais la connaissance ne se réduit pas à cela: elle exige une évaluation. Connaître, c'est séparer ce qui relève du bruit et de l'objet de connaissance. Or cette séparation suppose une référence à des normes extérieures au fonctionnement lui-même. D'avantage encore, la réflexion suppose une séparation, une distinction entre les perceptions et une hiérarchisation. Pour cette raison, elle est créatrice de nouveauté.

La notion de réflexion introduit en outre à une dimension nouvelle: celle du sens. Il ne suffit pas de manier des images pour connaître, ni de prononcer des mots pour parler. Il doit y avoir une saisie globale de la réalité. La parole est au service d'une pensée et d'une communication.

La pensée est au service de l'intelligence qui est mise en relation avec une vérité objective, dans un rapport qui brise avec le solipsisme.

On accède au sens par une pensée qui accepte de renoncer à son pouvoir de représentation pour laisser être l'objet connu et se rapporter à lui en vérité. L'exigence de vérité ne peut être que celle d'un sujet qui, pour connaître, ne s'enferme pas dans ses mécanismes d'auto-régulation ou d'auto-contrôle.

En restant au plus près du donné scientifique, il faut donc relever que toute connaissance se réfère à un sujet dont on ne peut pas séparer les éléments. C'est ce que signifie la notion de corps.

Le terme corps ne désigne pas seulement un contenant, l'enveloppe de la pensée. Le corps c'est l'être humain dans sa capacité de donner sens à des activités, à des relations, à une activité dont il a la responsabilité. La donnée fondamentale est le corps, sans lui rien ne serait, ni commencement d'existence, ni croissance, ni perception, ni affects, ni sentiment, ni pensée. La phénoménologie a développé cette dimension.[31]

4. *L'inscription charnelle*

Pour rester fidèle à la ligne austère et modeste que nous avons choisie, nous nous interrogerons sur la singularité. En effet, si le vivant peut être l'objet de l'étude biologique, c'est qu'il a une unité. La biologie le postule pour se différencier des autres disciplines scientifiques, physique ou chimie. Cette régulation unifiante est-elle correctement désignée par la notion de système auto-régulé et auto-référé?

La tradition philosophique distingue entre les causalités.[32] Elles ne peuvent se réduire à une seule instance, celle de la biologie, même si celle-ci use des modèles systémiques et cybernétiques. La distinction entre diverses instances de causalité est fondamentale. Une action ne saurait être dite sans référence à une pluralité de principes d'action.

De manière plus précise, il faut reconnaître que la valeur des nouveaux concepts scientifiques et la construction de modèles cybernétiques ou systémiques, ne permettent pas d'épuiser la richesse de la réalité, car les modèles qui expriment l'unité sont toujours mathématiques. Ils ne relèvent que la structure formelle. Celle-ci ne dit pas toute la richesse de sens du mot forme. En effet, pour parler d'un vivant, l'unité formelle ne

31 Cf. Xavier Lacroix, *Le Corps de chair*, Paris, Le Cerf, 1992.

32 On connaît le célèbre texte de Platon où il montre que la vraie cause de sa présence à Athènes n'est pas dans son corps, mais dans son âme. Platon, *Phédon*, 97c-99b, trad. Léon Robin, Paris, «Les belles Lettres», Budé, p. 70-71.

saurait être entendue seulement au plan mathématique. Elle pourrait n'être alors que la résultante d'un jeu de forces ou d'interactions. L'unité est un principe actif. Elle donne cohésion à ce qui sans elle se dégraderait. Or un principe actif d'unité ne saurait-être du même ordre que ce qu'il rassemble.

Pour le manifester, on peut analyser ce qui sert de paradigme pour les modèles biologiques utilisés en génétique, le langage. De même que le sens d'une phrase ne se réduit pas à la juxtaposition des mots, ainsi la forme qui fait l'unité du vivant ne se réduit pas aux échanges énergétiques nécessaires à la vie. Le sens n'est pas ailleurs que dans les mots, mais il ne s'y réduit pas. La traduction montre qu'il n'y a de sens qu'en excès par rapport à la syntaxe et à la grammaire — plus encore, une transcendance de la pensée par rapport au dire. Comment ceci est-il possible, sinon parce que le sens unique de la phrase n'est pas confondu avec la multiplicité des mots? De même pour la vie! Le principe unifié de la vie humaine est irréductible aux échanges d'énergie qui sont analysés par la biologie. Etre irréductible ne signifie pas être indépendant. De même que la parole et la pensée ne sont pas seulement une combinaison de mots, la vie ne se réduit pas à une interaction entre des unités biochimiques. Mais, comme dans le langage, où l'absence de mots tue la pensée, l'absence et le dysfonctionnement des éléments qui font la vie est source de mort.

L'ontogénèse et l'épigénèse fondent la possibilité de la vie d'un sujet humain libre. Sans elles rien ne pourrait être. Limitées, elles n'ouvrent pas sur l'infini; mais ce sur quoi elles ouvrent n'est pas strictement déterminé. Le génome est ce sans quoi rien ne fonctionnerait; il donne des possibilités et des limites. Pourtant l'individu humain est davantage. Davantage, mais pas ailleurs que là où génome et système neuronal opèrent. L'individu humain est une personne. Ceci n'est pas au terme de son achèvement, mais dès son principe, car sa vie lui a été donnée en humanité.

Pour conclure: une philosophie de l'unité pour le corps

Les réflexions philosophiques permettent de libérer l'esprit de tout réductionisme, en montrant où est l'humanité de l'homme. Elle est dans l'unité irréductible de la personne humaine. Cette conclusion philosophique ne va pas à l'encontre de la pratique médicale, car le soignant sait que le corps qu'il soigne n'est pas un contenant, enveloppe d'un esprit. Certes, comme corps, il est dans l'espace-matière-temps qu'étudie la science avec précision et compétence. Mais le corps n'est jamais sans principe d'unité. Le corps lui-même assure son unité, par un principe qui lui est immanent, l'âme,[33] principe de son unité.

Il en résulte qu'un corps humain, comme corps, est porteur d'un projet de vie. Il est porteur d'un désir de vivre et d'exigences qui fondent son droit que doit reconnaître la morale. Ces exigences sont internes et président à son développement, bien avant que la réflexion ne soit possible. Le moraliste dira que le corps doit être reconnu pour lui-même, à partir du dynamisme qui l'habite et préside de l'intérieur à sa croissance et à son développement.

Au terme de cette réflexion, j'aimerais pouvoir dire la grandeur de l'homme, à partir de ses oeuvres spirituelles, artistiques, scientifiques ou philosophiques. Qu'il me suffise d'avoir montré qu'elles ne sont pas que le fruit d'une détermination génétique ou cérébrale et que l'action du génome est ouverte sur un possible où l'humanité advient à elle-même, en humanité qui pour moi s'étend non seulement au jeu analysé par la sociologie, mais à la communion des saints!

J'ose encore une remarque. A cause de la tension vers l'avant qui caractérise le vivant, il est possible de prolonger la réflexion philosophique par une remarque théologique. Si on se place dans la foi biblique en la création, on peut légitimement parler de la vie comme d'un don — un don de Dieu. Or tout don est porteur d'une intention. On peut donc reconnaître que, dans le corps et l'âme ensemble qui font l'être humain, sont inscrits une intention et un projet. Lorsque le corps s'accroît, se complexifie, se perfectionne, noue des liens et s'unifie, un dessein se réalise; il est inscrit dès les infimes débuts de la vie, même s'il n'apparaît pas clairement au tout premier commencement de la vie, et qu'il apparaîtra pleinement lors de l'achèvement. Pour les croyants, cet achèvement s'est déjà manifesté dans la résurrection de Jésus-Christ. Là un corps (celui qui est né de Marie) a été transfiguré et glorifié par son union à Dieu. Commencement d'un monde nouveau et lumière sur le sens de notre destinée: rencontrer Dieu lors de l'universelle résurrection et être transformé par lui à son image.

---

[33] Le terme d'âme reste malheureusement équivoque, car il est trop marqué par le dualisme. Au sens où nous l'employons ici, il convient pour dire le principe actif d'unité d'un être vivant.

PASTORAL CONCERNS ABOUT THE HUMAN GENOME INITIATIVE

WILLIAM B. FRIEND *

I wish, first of all, to express my thanks for this opportunity to offer a few remarks in light of the agenda of this plenary session of your esteemed Pontifical Academy of Sciences.

My presentation will offer a pastoral perspective to the topic of the Human Genome Initiative for that is what I am, a pastor of a local church. My remarks will be divided into (I) a brief introduction of the Human Genome Initiative (HGI), sometimes referred to as the Human Genome Project (HGP), and its significance; (II) an overview of Catholic ethical-moral perspectives related to the Initiative; (III) reflections on ethical issues associated with genetic counseling; and (IV) consideration of issues of public policy and education. I close with a few comments in summary concerning what can be some of the next considerations for further study. Given the complexities of the initiative this presentation makes no pretense to completeness.

Presuming to speak for many pastors, permit me to say that those of us who are pastors appreciate the work of scientific research, clinical practice and technological development which is truly of benefit to human beings. We are becoming more aware of the intrinsic limitations of genetics and the distressing ambiguities that sometime remain after diagnosis. Our pastoral experience acquaints us with the sometimes unreasonable expectations people place on scientists and medical practitioners, and, conversely, the expectations artificially created by the advent of newer technologies. Our practice also reassures us that most people want to do the right thing. We are often with people as they struggle in making difficult decisions of major import. The norm given for such activities is that whatever we do should be

* Bishop of Shreveport, U.S.A.

in accordance with God's design and will, and be for the benefit of the human race so that people are enabled, both individually and in society, to pursue their total calling as human persons. Pastors join you in prayer and in the discovery of answers.

I. Significance of the Human Genome Initiative

As most of you know, the Human Genome Initiative holds the promise of being one of the most significant undertakings of science and medicine in our lifetime. This multinational, multibillion dollar, collaborative research effort being conducted over a period of fifteen years or more will enable us all to have a more profound vision of our human nature and some of its component parts.

The fundamental goal is to understand the molecular nature or make-up of genetic diseases and genetic predispositions, to map and sequence the human genome, and to provide the tools necessary to map the genes for thousands of genetic diseases and to link genetic predispositions to diseases. It will enhance our understanding of human biology, foster the development of new technologies serving society through new generations of scientists with new and marvelous scientific insights and medical advancements. But what of its intergenerational impact?

Ethical and moral considerations are challenged by the magnitude and complexities of this initiative. A joint effort among philosophers, theologians, legal scholars, social scientists, biologists and other scientists, clinicians and ethicists, as modeled in this Academy's Plenary Session, *but on an ongoing basis*, will be needed in order to develop the moral, ethical and legal maps that will help to provide standards, policy and direction during the scheduled years of intensive research, and for the use of outcomes of this initiative. Ethical and moral considerations will be enlarged far beyond the scale of what serves as adequate treatment today. The outcomes of the research may indicate, for example, that each kind of genetic defect, such as in breast cancer, hypercholestrolemia or Tay-Sacks, will require a distinctive set of ethical guidelines (Goldstein, 1993). In practical terms ethical-moral studies not only will encompass many more and different applications, but a more organized study and evaluation will be needed. Such studies will most likely have to be as large in scope as the Human Genome Initiative itself, if at all possible. At this time, 5% of the budget is allocated for this purpose.

II. Catholic Ehical-Moral Perspectives Regarding Genetic Research and Therapy

The 1987 document *Donum vitae*, published by the Congregation for the Doctrine of Faith, presented a fundamental criterion for the uses of science and technology: the integral good of the human person. The Instruction also acknowledged the constructive role science and medicine can play in aiding human persons. It cautioned against a blanket rejection of newer technological possibilities on the grounds of mere artificiality. Its anthropological assumptions are spelled out. The *Catechism of the Catholic Church* incorporates teachings of *Donum vitae* as they pertain to life, its care and scientific research, as will be noted in specific ways in the paragraphs to follow. These teachings, along with the works of ethicists and theologians, rooted in the whole framework of Catholic moral reasoning, offer pathways to ascertaining the *true* and the *good* in our time as we ponder the Human Genome Initiative.

The moral theologian Benedict M. Ashley argued that the compassion of the Church for the sick promotes development and application of technology in the pursuit of God-given, truly human goals. Genetic reconstruction is desirable to remedy genetic defects. But serious ethical questions begin to arise at the borderline where the aim of genetic reconstruction shifts from therapy to the "creation" of new human types (Cassidy and Pellegrino, 1993).

As with every other medical therapy, the basic ethical *questions* which are to be assessed at least include *answers* to the following questions: (1) Is the therapy in the best interest of the patient? (2) Do the benefits outweigh the risks? (3) Is there informed consent by the patient? (4) Are the basic principles of beneficence, nonmaleficence, prudence, justice and charity and respect for patient autonomy operative in the undertaking?

Generally speaking, there is nothing in the Catholic tradition that forbids genetic experimentation *per se*. We are sensitive especially to the enormous potential for good or harm. But it is the abuses of genetic manipulation, not their therapeutic applications, which are our concern as Catholics. The central ethical-moral question in the practical order, then, is: How, and by what criteria, do we ascertain what is *morally licit*, what is *morally dubious*, and what is *morally illicit* (Cassidy and Pellegrino, 1993).

From Catholic moral-ethical teachings one can develop a checklist or outline of principles which assists in decision-making as it relates to the Human Genome Initiative. Such a checklist or guidance system includes at least the following fifteen considerations:

1. The human being must be absolutely respected as a person from the very instant of his or her conception (*Gaudium et Spes*, 51; *Catechism*, 1994).

2. If prenatal diagnosis respects the life and integrity of the embryo and the human fetus, and is directed toward its safeguarding or healing as an individual, prenatal diagnosis is morally licit (*Donum vitae*, 1987; *Catechism*, 1994).

3. A strictly therapeutic intervention whose explicit objective is the healing of various maladies such as those stemming from chromosomal defects will, in principle, be considered desirable, provided it is directed to the true promotion of the personal well-being of the individual without doing harm to his or her integrity or worsening his or her condition of life. Such an intervention would indeed fall within the logic of the Christian moral tradition (Pope John Paul II, 1983; *Catechism*, 1994).

4. Therapeutic genetic interventions require the free and informed consent of the parents (*Donum vitae*, 1987).

5. In cases of clinical experimentation in which the embryo is living, whether viable or not, respect must be shown just like for any other human person. The informed consent ordinarily required for clinical experimentation on adults cannot be granted in non-therapeutic situations by the parents, who may not freely dispose of the physical integrity or life of the unborn child. Human embryos or fetuses are not to be treated *merely* as a means to an end or as a mere object of scientific investigation (*Donum vitae*, 1987).

6. Research, even when limited to the simple observation of the embryo or fetus, would become illicit were it to involve risk to the embryo's/fetus' physical integrity or life by reason of the methods used or the effects induced (*Donum vitae*, 1987).

7. In cases of experimentation one has to distinguish between experimentation for purposes which are not directly therapeutic and experimentation which is clearly therapeutic for the patient. If the embryos are living, whether viable or not, they must be respected just like any other human person; experimentation on embryos which is not directly therapeutic is illicit (*Donum vitae*, 1987).

8. The practice of keeping alive human embryos *in vivo* or *in vitro* for experimental or commercial purposes is totally opposed to human dignity (*Donum vitae*, 1987).

9. Experimentation that is clearly therapeutic, *i.e.*, when it is a matter of experimental forms of therapy used for the benefit of the embryo itself in a final attempt to save its life, and in the absence of other reliable forms of therapy, recourse to drugs or procedures not yet fully tested can be licit (*Donum vitae*, 1987; Pope John Paul II, 1982; *Declaration on Euthanasia*, 1980).

10. Biological and genetic manipulation of human embryos which involves attempts or plans for fertilization between human and animal gametes and the gestation of human embryos in the uterus of animals, or artificial uteruses, are contrary to the human dignity proper to the embryo. And at the same time they are contrary to the right of every person to be conceived and to be born within and from marriage (*Donum vitae*, 1987).

11. Attempts or hypotheses for obtaining a human being without any connection with sexuality through "twin fission", cloning or parthenogenesis are to be considered contrary to the moral law, since they are in opposition to the dignity both of human creation and of the conjugal union (*Donum vitae*, 1987).

12. The freezing of embryos, even when carried out in order to preserve the life of an embryo (cryopreservation), constitutes an offense against the respect due to human beings by exposing them to grave risks of death or harm to their physical integrity; depriving them, at least temporarily, of maternal shelter and gestation; and placing them in a situation in which further offenses and manipulation are possible (*Donum vitae*, 1987).

13. Chromosomal or genetic manipulation to influence inheritance, which is not therapeutic, but aimed at producing human beings selected according to sex or other predetermined qualities is contrary to the personal dignity of the human being and his or her integrity and identity. In no way can this be justified on the grounds of possible beneficial consequences for future humanity (*Donum vitae*, 1987; *Catechism*, 1994).

14. It is immoral to produce human embryos intended for exploitation as disposable, biological material (*Donum vitae*, 1987; *Catechism*, 1994).

15. It is an illusion to claim moral neutrality in scientific research and its applications. On the other hand, guiding principles cannot be inferred from simple technical efficiency, or from the usefulness accruing to some at the expense of others or, even worse, from prevailing ideologies. Science and technology by their very nature require unconditional respect for fundamental moral criteria. They must be at the service of the human person, of his/her inalienable rights, of his/her true and integral good, in conformity with the plan and will of God (*Catechism*, 1994).

Two observations are added to the citation of these fifteen principles. First, W. French Anderson, (1989) a pioneer in gene therapy experimentation, maintains that a line should be drawn to use gene transfer only for the treatment of serious diseases and not for any prevention of disease or other eugenic purposes. His distinctions between therapy, prevention and eugenic purposes are helpful. There appears, however, to be a new twist involving the principle of double effect and secondary consequences being permitted when he includes both somatic and germ-line therapy for serious diseases as morally acceptable actions. This approach using germ cell engineering as a secondary consequence of somatic cell therapy under the principle of double effect would lead to a whole new set of ethical problems. Will it be so easy to separate therapeutic interventions from non-therapeutic genetic techniques if both somatic and germ line therapies are permitted to fuse? Lappé (1993) appears implicitly to appeal to the principle of the double effect concluding that "germ line engineering as a directed attempt to change the genotype of future generations cannot be ethically justified. However, when such charges arise from an indirect and otherwise unavailable consequence of an approved form of somatic cell engineering, they are morally acceptable". The question remains how radical and harmful is any such consequence, albeit secondary or tertiary? *Can, should, ought* this still be permitted merely because it only appears to be a by-product of some primary, so-called responsible action? We may have to re-think our understanding of the principle as it relates specifically to these new genetic interventions. Anderson (1989) limits genetic manipulation to serious diseases only because of the danger of inadvertently modifying the human genome in ways that could alter the nature of the person.

Second, as early as 1972, Karl Rahner maintained that even though humans have an essential freedom to manipulate themselves, there is a limit to that freedom when the actions could modify or even destroy the very nature that is the root of the human essence. He observed there is an ultimate moral commission to refrain from those acts which will change the nature of our humanity. At the same time there is the challenge to the theologian of continuing to search out what constitutes human nature so that the horizon and limitations of self manipulation are clearly understood. The acceptable interventions then are the ones that will contribute to the preservation of our God-given human nature (Rahner, 1972).

III. Ethics and Genetic Counseling

Genetic testing and counseling generally occur in primary care settings in the present experience. Usually the testing occurs when there is an obvious underlying disease, when particular family histories are remarkable for rare Mendelian conditions with grave implication for morbidity and mortality. It is done when developmental delay or abnormality in children raises suspicions of disease with unidentifiable etiology, or as part of limited screening among practitioners who engage in prenatal and obstetrical care (McCrary and Allen, 1994).

While these uses of genetic testing are not uniformly applied because of different practices among clinical and practice settings, more frequent and universal usage will probably come into play as new testing technologies become cost effective and readily available for use in primary care settings.

The trend toward more genetic testing introduces the questions: *How will the patient be informed? In what way will the patient or the parents be involved in the decision-making? What is the role of genetic counselors, particularly in a Catholic health care facility?*

The Houston Experience

The Houston Working Group of the National Conference on Genetics, Religion, and Ethics in 1992, explored developing appropriate guidelines for physicians and genetic counselors as they seek to apply newly gained genetic insights from the Human Genome Initiative. Educational models were explored also (Lustig, *et al.*, 1992).

Through the use of interviews of parents of patients, the researchers focused primarily upon issues in prenatal and neonatal diagnosis. In this particular focus the researchers sought the model of how tangible decisions were made by individuals as they dealt with information relating to medicine and genetics, and religious, ethical and moral teachings.

A summary review of the study's findings indicated (1) the need for improvement in the initial counseling session with regard both to the initial diagnosis of the genetic condition and to information about likely prognosis; (2) the sense of timing was a critical issue in clinical ethics; (3) dealing with a married couple as a team was very important; as was (4) the counselor's attentiveness to compassion, care and personal values. Eric Cassel's (1990) judgment extended to this study seemed verified, *i.e.*, the need for medicine to learn to deal with the reality of a patient's suffering, rather than simply to treat the patient's physical condition. Lustig *et al.* (1992) observed in this Houston Study that "while generalizations concerning biology and pathology underlie the epistemology of allopathic medicine and related medical research, the function of the practitioner is to abstract essential information from these generalizations which are relevant to individual patients". The language of "disease" is the vocabulary of the medical specialist. As such, ironically, it may be a generalization based upon a nosology of disease categories rather than upon the particular circumstances of the patient. Improvement is called for in developing a description of unique personal events and bridging the gap between diagnostic working generalizations and the patient's unique experience.

The interviewers discovered, as could be expected, that most interviewees identified their spouses as their primary locus of support and solace. Immediate family and close friends were also mentioned. Less common sources of support were relatives geographically distant, "networks", and pastors or chaplains. Overall, few subjects recalled their pastors as being sources of moral insight in the hard choices they confronted.

Social scientists, philosophers and theologians could perhaps help clergy become more useful in such situations. The challenges facing pastors are the realities of a culture that promotes privatized religion; the parish congregation not being perceived as a crucible of moral meaning or practical judgment; the compartmentalization of thinking about life and ethics; and scientism, the confusing of technical expertise with moral authority, inclining patients to listen only to the clinician. There is need also to include pastors in the interdisciplinary team at the time of counseling and to equip clergy with more theodicy, *i.e.*, explanations meant to reconcile the presence of suffering and evil with the infinite love and the infinite power of God.

Virtually all interviewees had been raised in a religious tradition. They represented a wide variety of identifiable religious values and ideas. Clear evidence was given that parents were aware of the moral gravity of their choices. Parents who developed their moral judgments in a religious light generally concluded that improving the situation of children afflicted with genetic anomalies is a clear implication of faith. They, too, are fearful of the moral "slippery slopes" and the return of eugenics.

These data would appear to indicate that further development will be needed in the preparation of and protocols for individuals who provide genetic counseling. They appear also to indicate the linkage of the actual *genetic counseling* with the other *associated counseling* needed by the family and individuals who offer caring and support.

Using Test Results in Counseling

Prior to prenatal diagnosis, the controversial indication and research have to be considered. McCrary and Allen (1994) observed that in genetic testing the issues of predictive value become critically important. For particular tests to be professionally and morally acceptable, it will be necessary to establish minimum levels of sensitivity and specificity. This outcome will help to avoid the emotional connotations of and potential adverse health effects of false positives. Issues of incomplete penetrance, *i.e.*, the proportion of individuals of a particular gene type that expresses a phenotypic (properties) effect in a given environment and variable expressivity *i.e.*, the relative capacity of a gene to affect the phenotype of the organism of which it is a part, also have to be considered. Physicians are faced with a different task in counseling the patient to understand that common diseases are multifactorial in etiology (causes), requiring numerous adverse genetic and environmental events for disease to be expressed (penetrant).

Free and Informed Consent

Before prenatal diagnosis, the genetic counselor needs to deal with informed consent in a manner which encompasses the disclosure of the appropriate information with competence and the voluntariness of the patient being recognized. These basic elements constitute informed consent ethically processed. Informed consent remains a needed process throughout diagnosis and therapy (A.R. Jones *et al.*, 1986).

Once the diagnosis has been made the clinician or counselor often wonders how much information should be shared above and beyond what is needed for informed consent. One helpful answer was provided by Nancy Wexler in U.S. Senate Hearings in 1993 when she stated the following:

1. I have the right not to know my genetic status, if I choose not to;
2. If I do know my genetic status, I do not need to share that information with others;
3. And I have the right to use information to make decisions based on my own values and without coercion from others.

This testimony by Wexler reflects the principle of genetic autonomy or self-governance, the "Points to Consider" and the just respect for the privacy of the patients and the confidentiality of their medical records.

Counseling Challenge: Gaps Between Diagnosis and Therapy

As science enters deeper into the mapping of the genome, capacities for diagnosing various hereditary anomalies will increase. Concomitantly we can expect a genetic therapeutic gap for some time. Experimental cures, gene therapies and various ameliorative interventions will be difficult to develop and expensive to use. This reality will increase challenges to beneficence and justice in counseling which will in turn require assistance from ethicists and policy resources, and open the way for pastoral guidance and support systems.

Sharing Genetic Information

Physicians and counselors will have to be very concerned about the control and custody of genetic information as it pertains to insurance, employment and other implications for patients. How much genetic information should be shared? Should genetic information be placed in the medical record? Should we ever permit genetic information to be classified as "pre-existing conditions" for exclusionary insurance practices? What will the counselor do about sharing information regarding the implications for lineal descendants? Extended family members? What are the other disclosure dilemmas and their ethical assessments? How will the results of general screening programs be managed ethically (McCrary and Allen, 1994)?

Value Neutrality in Counseling

In the 1980s, sworn "moral neutrality" was the ideology or working ethos of professional genetic counselors. In such a practice the counselor used the so-called "value neutral", nondirective ethos in disseminating information to their patients or clients. This approach is being increasingly challenged as patients who are faced with more difficult decisions search for a prescriptive recommendation, and in fact do call on their own values and in many cases their particular religious beliefs (Caplan, 1993). Moreover, disease itself is not value free. It is thought to be bad, or something to avoid, to treat, to overcome. Information about disease has, therefore, a built-in value tone which is reinforced in culture, history and religion.

Caplan says it is time to abandon the ethos of value neutrality in genetic counseling. He observed: "value neutrality discourages those in the field from coming to grips with the central question that now confronts the field — how to define genetic disease and disorder in order to lay out appropriate targets for testing and counseling ... as more and more information becomes available ... this question can no longer be avoided". Moreover, Caplan observed, "the outdated stance of moral or value 'neutrality' leaves the counselor powerless in the face of what may be immoral requests on the part of clients". Finally, he warns that "a stance of value neutrality may foster a sense of irresponsibility about reproduction and genetics. If reproductive choice is simply a matter of individual choice and nothing more, if no moral agreement or persuasion is considered licit, then the message the profession is sending to its clients and the public is that you can do with your gametes as you please ... value neutrality, whether possible or not ... is simply not an adequate ethic in the public policy arena. It is precisely the information sought by those involved in the genome project that points toward the inadequacies of the current ethic in genetic counseling" (Caplan, 1993).

Genetic counseling seems to be in need of refinements, therefore. It is challenged by the rapidity of discoveries of new information, the personal needs and values of a variety of patients, the complexities of the new scientific discoveries, the time gap between discovery and therapy and the slowness of developing applicable ethical guidelines. In addition, the number of genetic counselors is in short supply. For example, in the U.S.A. there are at present only about 1,000 genetic counselors.

Religion and Genetic Counseling

Religious leadership is challenged by all of this. There is a need to help provide essential ingredients for the moral, ethical and pastoral dimensions associated with genetic diagnosis and therapy. Continuing scholarly work on the nature of the human person, theodicy, the coping with evolutionary theory and deeper insights in some areas of ethics and moral discernment are needed. Local church congregations need to discover how better to serve families who face genetic decision-making dilemmas. Other questions emerge in this regard. What guidelines can be developed for families? What advocacy can and should be offered? What will effectively replace the weaknesses in contemporary genetic counseling, *e.g.*, value "neutrality"? How best can issues of genetic decision-making for families be incorporated into catechetics and marriage preparation? How can pastors better serve supportive counseling needs that will complement genetic counseling?

IV. Public Policy Development and Education

Standards

Juengst and Watson (1994), in their study of the Human Genome Initiative and the responsible use of genetic knowledge, address issues involved in the integration of new genetic tests into health care. Social policy problems involved in integrating these tests effectively into medical practice include developing standards for a number of dimensions of health care, such as the following:

1. Ensuring the accuracy and quality control of genetic tests;
2. Defining the indications for testing and the design of testing protocol;
3. Establishing the credentials for clinicians who perform tests;

4. Protecting the confidentiality of information obtained from testing;
5. Controlling access to and use of test results by third parties, e.g., insurers, employers, state health care agencies and bureaucracies; and
6. Providing reimbursement for testing and test related counseling; determining the payment plan.

Concern about these questions is discussed against a background of previous national experiences with widespread screening for genetic conditions, as for example in sickle cell disease, spina bifida, et cetera. The general public and the religious communities have not always received full reports on these earlier discussions. More open discussion is needed so that public policy can be in place before the clinical assessments of new genetic therapies and services. Moral norms and ethical standards are needed beyond minimalist principles since intergenerational benefits and risks are at issue.

Equality - Justice

In 1992 the Chicago Working Group of the National Conference of Genetics, Religion and Ethics cited Murray (1991) who said that "the sciences of inequality, with genetics at the forefront, will force us to reinterpret what equal treatment and equal regard mean in an enormous range of contexts. But they need not threaten the ethical care of that commitment". Various commentators predict the Human Genome Initiative will strain society's commitment to equality. One effect of a growing public knowledge may be to note ethnic differences among human beings, as in the period roughly from 1918 to 1945 which led to the so-called efforts of "ethnic cleansing", recurring recently in Eastern Europe.

Already debates about freedom and determinism are entering into public conversation (Brock, 1991 and Wright, 1994). Popular interest is being heightened through television programs such as *Perfect People* broadcast by Cable News Network (CNN) September 4, 1994. Keen contemporary interest in these questions is likely to influence public policy on education, child care issues, pursuit of the so-called "level playing field" and crime and punishment. Eugenics distinguished by the "perfect child" syndrome will most likely surface again to challenge social justice and equality.

Religious leaders and communities committed to human dignity, freedom, equality and social justice need to address the difficult questions about the possibilities and discoveries which are being added to the bank of genetic knowledge. Our efforts as spiritual leaders may not keep us abreast of all the rapid discoveries and issues, but we need to be made so. Our world cannot afford another perversion using unscientific theories of genetic determinism that may be manipulated to categorize some lives as Lebensunwertigen leben, "life unworthy of living". Racism, classism and sexism have all regularly paraded under the guise of objective scientific analysis, observed Lewontin *et al.*, (1984).

Such possibilities of grave societal threats can arise out of the Human Genome Initiative if an intensification of determinist ideologies is allowed to be promoted in our cultures.

Education

Religious leaders and communities, therefore, need to collaborate with scientists in order to pursue educational efforts at least in two areas:

1. The provision of accurate information about what is known and what is unknown about genetics. *The influences and interactions of environment and experience have to be studied along with genetic knowledge.*

2. Teaching people to understand how ideological biases regularly infect the interpretation of scientific findings. Gross (1994) has identified current radical attacks on science from deconstructionists. A given genetic endowment is not inherently good or bad. Judgments concerning the relative (as opposed to the priceless and incalculable) worth of a given human being depend upon interaction with religious teachings, personal values, environment and culture. As Dobzhansky (1973) observed "human quality is an ethical precept, not a biological phenomenon".

The Chicago Working Group (1992) reminds us that "religious, philosophical and social commitments are not determined by knowledge of genetics; rather, these commitments define what a genuinely human response to hereditary diversity should be". The Human Genome Initiative calls religious communities to attend to the relationship between cultural ideologies and understandings of heredity.

Public policy and education will have to attend to the Initiative and the issues of health care, access, insurance and employment discrimination. As Hillary Rodham Clinton observed in the health care reform debates in the U.S.A., every citizen could experience having "previously existing conditions" that prevent coverage or dramatically increase insurance premium payments. While insurance actuaries say they can assess the risks of genetic conditions just as any other risk and spread costs, this remains to be seen. A better public policy may be to keep confidential the genetic information about an individual, but open access to genetic therapy.

Religious Liberty

Public policy will need to respect religious communities that dissent from the prevailing teachings of modern technological science and medicine. The freedom of families to practice their own faith must be upheld. Society cannot assume that it is mature and knowing enough to determine what is normative for human nature, and then enforce its assumptions or conclusions. The very questions of human identity are involved in what happens to the family. How important will the role of the state become in these areas for those who are dependent on state help for these same services? How coercive will the state become relative to most genetic screening? Will there be an unwarranted expansion of state power over certain, more vulnerable groups of our society?

Public Service and Advocacy Vs. The Technological Imperative

Religious communities would serve well if some coordinated process were developed to examine more closely the way society and government supervise or fail to supervise scientific practice, especially as it refers to genetic research and application. Because something can be done scientifically, does not mean that it ought to be done morally. How will monitoring be improved, while at the same time respecting scientific inquiry?

As in all health care, issues of social justice will be raised from the Human Genome Initiative as its findings are applied in various therapies and technologies. Lebacqz (1983) recommends to churches that in addition to supporting families and education they "be involved as advocates on public policy issues related to new genetic technologies — *e.g.*, supporting legislation to aid disabled persons, examining proposals for genetic screening programs in the neighborhood and giving a 'prophetic voice' to issues raised for parents and others".

In relation to the Human Genome Initiative, our advocacy should address issues of confidentiality, equal rights for the disabled, social justice for persons who suffer from genetic disorders and a strong statement against the ideological nature of giving out information about one's heredity.

Consumerism

In market driven economics there is a danger that genetics will be perceived as just another product for consumers and lead to consumerism. As Collins has reported, couples occasionally come into his clinic with a new-car mentality. They say, "It's got to be perfect, and if it isn't you take it back to the lot and get a new one" (Cowley, 1990). Parents of patients cannot always face the reality that the life of a child is at stake. The issues become even more complicated when questions about therapy and eugenics are raised in connection with germ line manipulation and gene enhancement therapy, *e.g.*, in germ line therapy, the sciences and the outcomes are not known, while enhancement experiments cannot be justified due to medical risks lacking scientific merit. There are emerging controversies that will have pastoral ethical implications, *e.g.*, the use of a genetically engineered human growth hormone in children who are smaller in stature.

On the positive side numerous genetic support groups have developed and they reflect an impressive (and highly desirable) consumerization of medical genetics. They will have their effect on forcing public policy as they grow and mature.

Commercialization

McKusick (1993) observed that it is in the area of laboratory diagnostic services that the commercialization of medical genetics is most evident. New technologies are speeding up laboratory studies at a dramatic rate. Two dangers arise: (1) that there can be inappropriate counseling in connection with the rendering of laboratory results and (2) that such research becomes merely a business for profit. Policy guidelines for regulation of such business will be needed.

Public Policy Formulation

Collins (1993) observed that most legislators in the U.S.A. know little about genetics, and that policy is lagging behind progress in scientific technology. In the U.S. Congress, *e.g.*, health aides to Congressmen attempt to understand the complexities of the scientific questions and make recommendations to their Congressmen. While this is not as it should be, this is the way it is. Brown (1993) pointed out another danger in that social policies designed for one group of patients may actually harm another group of patients who are genetically susceptible to whatever manipulation occurs.

Brown also observed that there are no policies or regulatory guideline in the U.S.A. under which consideration for availability of genetic testing, quality control, and test-related counseling is regulated. The agency in charge of performing proficiency testing is overburdened, underfunded and undersupervised. In Great Britain geneticists have been able to corner the market for profit in the case of some genetic discoveries because of the slowness of governmental guidelines.

Wexler (1992) stated that another one of the problems is that it is technically possible to do multiple testing, which looks for a number of genetic abnormalities at the same time. How will a patient give free and informed consent for each test? Just because these tests can be done together does not mean they should be stacked. Some states routinely do multiple testing on single blood samples from newborns already. Doing the same for children and adults may be seen as a simple extension, done without much consideration of the ethics. Public policy will need to attend to standards for this phenomenon as well.

International Cooperation

Indications are that overall international cooperation is very good. The international Human Genome Organization (HUGO) serves as a clearinghouse for mapping information. Johns Hopkins University maintains another mapping database. The human genome center at Los Alamos, California, maintains GenBank, a database of DNA sequences worked out by researchers worldwide. GenBank's sequence database passed seventy million base pairs by the beginning of 1992; with improving technology and increasing resources that number is doubled every two years (Erickson, 1992).

As is the case in most scientific research, the flow of information is very good when a discovery is twenty megabases away, but when a discovery is a half a megabase away there is less cooperation and information sharing. This allows a researcher time to follow up and complete the scientific process (Watson, 1993).

There are ethical problems on the horizon that pertain to information sharing and will have to be discussed, *e.g.*, ownership of discoveries; recognition of authorship and patents, especially in computer network developed materials; invasions of privacy; and the identification of what can be commercialized and what should not be so designated. Nations will indeed have much to explore in these future policy areas and will need to establish some kind of National Bioethics Commission, if they have not done so already.

While there appears to be very good international collaboration among scientists in the Human Genome Initiative, there is yet to be developed an organized and comparable effort among ethicists and theologians who specifically address the panorama of issues associated with the Human Genome Initiative along with the various implications for the different world religions.

V. Summary and Conclusions

The Human Genome Initiative is a major scientific undertaking which will change many understandings about human life. It will be a great force for the good of humankind, but, as in all human endeavors, it will need a continuous examination of the ends to which this initiative and its applications are directed and the means which are employed.

As the Boston Working Group (1992) explored the Human Genome Project and Prenatal Diagnosis they observed "the holistic, communal, growth-oriented early Christian approach to human life and purpose, with its focus on God-likeness, sets a tone and provides an orientation that can help modern bioethics to avoid distorted orientations and to point its endeavors in the direction of a therapy that is coherent with authentically God-like human existence".

First and foremost among pastoral concerns is the need for continuing conversation about the status and findings of genetic research so that medical research leading to a therapy is discussed openly, not only medically, but also from the perspective appreciative of the societal, ethical, moral and pastoral consideration.

Among the concerns shared in such conversations, I suggest, should be (1) the criterion for judging genetic interventions; (2) the values involved; and (3) the procedural methods for working together for the benefit of the patient and the family.

Such a dialogue can be facilitated when pastors and others are shown how, for example, "Points to Consider in the Design and Submission of Human Somatic-Cell Gene Therapy Protocols" have been addressed, and when the findings of ELSI (National Center for Genome Research [NCHGR] Program on Ethical, Legal and Social Implications) can be shared more fully.

Continuing dialogue and public conversation about the ethical-moral issues of the Human Genome Initiative are vitally important. Pellegrino (1991) observed that the ethics of clinical research need to be viewed from three different perspectives: (1) the process of acquiring new knowledge; (2) the moral use of the knowledge acquired; and (3) the ethics of the investigator seeking this knowledge. This observation serves to remind us that in the final analysis it is the personal morality of the clinical investigator, that is the ultimate safeguard of the safety of the experimental subject. Likewise, it is the genetic counselor, the physician and the nurses to whom patients and families turn in times of difficult medical and moral choices. They bear the serious responsibility to encourage the good and that which leads to the total calling of the human person. It is their integrity and ethical conduct that will come into the foreground of any national debate when certain undesirable or questionable outcomes become obvious.

Fortunately, scientists and practitioners do have ethical principles which are available for their guidance and exploration. In this effort we serve good research and medicine, and we serve humankind.

The good to be pursued will be the subject of intense debate over new genetic technologies, particularly in pluralistic societies. Our values will need to be made known clearly, appreciated on their own merits, and/or otherwise be offered for "guidance and exploration" in the pursuit of furthering the dialogue between scientists and the Church.

Again, I thank you for addressing the topic in your assembly and for having so kindly invited me to share it with you.

REFERENCES

Anderson, French, 1989. «Human Gene Therapy: Why Draw A Line?», *Journal of Medicine and Philosophy 14*: 681-693.

Ashley, Benedict M., 1987. «Constructing/Reconstructing the Human Body», *The Thomist. 51*, 3, 501-520-992.

Boston Working Group of the National Conference on Genetics, Religion and Ethics. *The Human Genome Project and Prenatal Diagnosis.* Working Paper presented at the Genetics, Religion and Ethics Conference (March) 1992.

Broch, Dan W. *The Genome Project and Human Identity.* Conference on Legal and Ethical Issues Raised by the Human Genome Project. Houston (March 8) 1991.

Brown, Michael S.; Collins, Francis S.; Goldstein, Joseph L.; Watson, James D.; Wexler, Mary Sabin. 1993 «Roundtable: The Human Genome Project», *Issues In Science and Technology*, X, 1.

Cable News Network (CNN), 1994. «Perfect People», *Inside Story.* (A television Broadcast in U.S.A., September 4).

Caplan, Arthur L., 1993. «Neutrality and Not Morality: The Ethics of Genetic Counseling». In: Bartels, Dianne M., LeRoy, Bonnie S., Caplan, Arthur L. eds. *Prescribing Our Future.* (New York: Aldine de Gruyter) 149-165.

Cassel, Eric. J., 1990. *The Nature of Suffering.* (New York: Oxford University Press).

Cassidy, Joseph D. and Pellegrino, Edmund D., 1993. «A Catholic Perspective on Human Gene Therapy», *International Journal of Bioethics. 4*, 1 (March) 11-18.

Catechism of the Catholic Church, 1994. Libreria Editrice Vaticana. (Washington, D.C. United States Catholic Conference) 547-552.

Chicago Working Group of the National Conference on Genetics, Religion and Ethics. (1992) *The Human Genome Project: Issues for the Religious Communities.* Working Paper Presented to the Genetics, Religion and Ethics Conference (March).

Clinton, Hillary Rodham, 1994. «Health Care Reform». Interviews Broadcast on CNN News Television (March through July).

Collins, Francis S. in Brown, Michael, et al., 1993. «Roundtable, The Human Genome Project», *Issues In Science and Technology, X*, 1.

Congregation for the Doctrine of the Faith, 1980. *Declaration on Euthanasia.* 10-11.

Congregation for the Doctrine of the Faith, 1987. *Donum vitae. Instruction on Respect for Human Life In Its Origin and On The Dignity of Procreation.* 1-32.

Cowley, Geoffrey, 1990. «Made-to-Order-Babies», *Newsweek* (Special Issue: the 21st Century Family), 94-100 (Winter/Spring).

Davis, Joel, 1970. *Mapping The Code: The Human Genome Project and the Choice of Modern Science.* (New York: Wiley & Sons).

Dobzhansky, Theodosius G., 1973. *Genetic Diversity and Human Equality.* (New York: Basic Books).

Erickson, Deborah, 1992. «Hacking the Genome», *Scientific American* (April) 128.

Goldstein, Joseph F. 1993. In Brown, Michael S., et al., «Roundtable, The Human Genome Project». *Issues in Science and Technology, X*, 1.

Gross, Paul, 1994. *Higher Superstition*: (The Academic Text's Quarrel with Science). Soon to be published.

Houston Working Group of the National Conference on Genetics, Religion and Ethics, 1992. *Genetics, Religion and Ethics: The Clinical Connections.* Working Paper Presented to the Genetics, Religion and Ethics Conference (March).

John Paul II, 1983. «Discourse to the Participants in the 35th General Assembly of the World Medical Association», October 29, 1983: *AAS* 76(1984) 392.

Jones A.R., et al., 1986. *Clinical Ethics* (2nd Ed.) (New York: Macmillan Publishing).

Juengst, Eric J. and Watson, James D., 1994. «Human Genome Research and the Responsible Use of New Genetic Knowledge», In Monagle, John F. and Thomasma, David C. *Health Care Ethics: Critical Issues.* Gaithersburg, Maryland: Aspen Publishers, Inc. 3-5.

Lappé, Marc, 1993. «Ethical Issues in Manipulating the Human Germ Line», in Shannon Thomas A. (ed) *Bioethics.* (Fourth Edition) (Mahwah, New Jersey: Paulist Press).

Lewontin, Richard C., Rose, Steven and Kanin, Leon J., 1984. *Not In Our Genes: Biology, Ideology and Human Nature.* (New York: Pantheon Books).

Lustig, B. Andrew, Bouma III, Hessel, and Ledley, Fred., 1992. *Genetics Religion and Ethics: The Clinical Connection.* Working Paper Presented to the Genetics, Religion and Ethics Conference (March).

Lebacqz, Karen, 1983. *Genetics, Ethics and Parenthood.* (New York: Pilgrim Press).

McCrary, S. Van and Allen, William L., 1994. «The Human Genome Initiative and Primary Care» in Monagle, John F. and Thomasma, David C. *Health Care Ethics: Critical Issues.* (Gaithersburg, Maryland: Aspen Publishers, Inc.) 305.

McKusick, Victor A., 1993. «Medical Genetics — A 40-Year Perspective on The Evaluation of A Medical Specialty From A Base Science», *Journal American Medical Association* 27:19 (November) 2355.

Murray, Thomas, 1991. «Ethical Issues In Human Genome Research». The FASEB *Journal 5* (1): 55-60.

Pellegrino, Edmund D., 1991. «Beneficence, Scientific Autonomy, and Self-Interest: Ethical Dilemmas in Clinical Research», *Georgetown Medicine.* 1:1 21-28. Also, see Monagle, John F. and Thomasma, David C., *Health Care Ethics: Critical Issues.* (Gaithersburg, Maryland: Aspen Publishers, Inc.). 325-333.

Rahner, Karl, 1972. «The Experiment With Man» and «The Problem of Genetic Manipulation», *Theological Investigations, IX*, 205-252.

Second Vatican Council. *Gaudium et Spes*, 51.

Watson, James D., 1993 with Brown, Collins, Goldstein, Wexler. «Roundtable: The Human Genome Project», *Issues In Science and Technology, X*, 1.

Wexler, Nancy Sabin, 1993. with Brown, Collins, Goldstein, and Watson. «Roundtable: The Human Genome Project», *Issues In Science and Technology, X*, 1.

Wright, Robert, 1994. *The Moral Animal.* (New York: Pantheon Books) [The New Science of Evolutionary Psychology].

SCIENTIFIC PAPERS

on:

ALTERNATIVE ENERGY SOURCES FOR DEVELOPING COUNTRIES

INTRODUCTION

ANDRÉ BLANC-LAPIERRE *

L'Académie Pontificale des Sciences a souvent porté son attention sur des problèmes relatifs à l'énergie. Je citerai, en particulier, dans cette perspective, trois semaines d'études parfois prolongées par l'activité de groupes de travail:

— *10-15 Novembre 1980*: Humanité et énergie: besoins, ressources, espoirs.
— *11-14 Juin 1984*: L'Énergie pour la Survie et le Développement.
— *17-22 Novembre 1991*: Ressources et population.

Aujourd'hui, il s'agit plus précisément d'énergies de substitution (premier point) pour les pays en développement (deuxième point).

Probablement, la notion d'énergies de substitution — qui a des liens mais ne s'identifie pas avec celle d'énergies renouvelables — aura-t-elle besoin d'être précisée, surtout pour des pays en développement. Comment remplacer une énergie défaillante? Comment faire face à la pénurie? La réponse à ces questions dépend énormément des contextes locaux, de l'échelle géographique ou de l'échelle de temps, des moyens de transport ou de développement sur place, des possibilités financières, des conditions politiques ...

Je crois intéressant, dans ces quelques mots d'introduction, de comparer, en ce qui concerne l'énergie dans le monde, la situation correspondant à la première semaine d'études mentionnée et celle que nous connaissons aujourd'hui, quatorze ans après. Le marché de l'énergie s'est fortement modifié, rendant ainsi caduques certaines des conclusions adoptées en 1980.

Le marché de l'énergie continue de dépendre du marché du pétrole. De 1973 à 1980, les prix pétroliers avaient été multipliés par 7 ou 8

---

* Académicien Pontifical. Académie des Sciences, France.

[croissance économique et politique restrictive de l'OPEP]. À partir de 1980, la demande à commencé à faiblir [croissance économique ralentie et politiques énergétiques nouvelles adoptées par les pays occidentaux: économies d'énergie, substitution au pétrole, du gaz, du charbon, du nucléaire et, plutôt symboliquement, des énergies renouvelables]. Parallèlement, l'offre du pétrole a crû a partir de 1980 [développement de gisements nouveaux et même production accrue de certains pays de l'OPEP pour maintenir leurs revenus]. Les prix se sont effondrés en 1986. Actuellement, en dollars constants, ils sont revenus entre 2 et 3 fois les prix de 1972.

En principe, cette situation aurait dû être favorable à l'ensemble des pays consommateurs. En fait, elle ne l'a été que pour les pays industriels.

La situation énergétique des pays en développement s'est aggravée. La dépendance des importations a crû, ainsi que la pénurie d'approvisionnements due aux manque de moyens de paiement et d'investissement. La crise économique a fait baisser les exportations à partir des pays en développement ainsi que les investissements provenant des nations industrialisées. Parallèlement, la demande d'énergie importée par les pays en développement s'est accrue (essence, gazole, butane, ...).

Ces brèves remarques montrent l'intérêt du sujet qui va être traité. Il s'agit de faire le point sur la situation actuelle relativement à l'apport que certaines sources d'énergies de substitution peuvent offrir aux pays en développement, sur l'évolution possible de cet apport et sur les conditions réalistes qui permettraient de le faire croître.

ALTERNATIVE ENERGY SOURCES FOR THE THIRD WORLD: PERSPECTIVES, BARRIERS, OPPORTUNITIES

UGO FARINELLI *

SUMMARY - Energy will continue to play an important role in the future world. A rapidly increasing availability of energy at reasonable price is one of the key issues for the development of Third World countries.

Fossil fuels are not running short in the near future. Prices have shown a remarkable stability in the last years. On the long term, however, the resources of fossil fuels (particularly oil) will be depleted.

The deterioration of local and global environmental conditions and the increasing recognition of the threats posed by man-produced enhancements of the greenhouse effect have become the main driving force towards the development of energy alternatives and of improvements in the efficiency of energy use.

Nuclear energy will continue to play a role in energy supply, particularly important for industrialised countries. Its share in the energy budget, however, is not expected to rise in the next decades, and in particular it is doubtful that it will contribute appreciably to the increasing energy needs of developing countries.

Many technologies for the utilisation of renewable sources of energy offer solutions that are already cost-effective in several situations of Third World countries: they include hydroelectricity, biomass, solar, geothermal and wind energy.

Once the obstacles to the diffusion of renewable energies are removed, they should become a substantial part of the energy supply in both industrialised and developing countries. Such obstacles are mostly of a non-technical nature, and derive from incentives for the use of fossil fuels, from energy prices that do not reflect the costs paid by the community (indirect and social costs, damage to the environment etc.), and from insufficient information. If put on an equal footing with other energy sources (the so-called level playing field), renewables should be able to compete on the market even with today's technologies; an effort of R&D to improve these technologies or to develop new ones would make this competition even more favourable to renewahle energies.

* ENEA, Italia.

However, the future energy budget of the world (and of developing countries in particular) can be expected to depend not on one or another hegemonic source as in the past, but rather on a much greater effort toward increased efficiency of energy utilisation and, on the supply side, on the co-existence of several sources, to be used in different proportion according to local needs, opportunities and applications. Developing countries, which are facing a major energy transition while less bound by investments made in the past, have the opportunity of greater flexibility.

1. Introduction

Throughout the history of humankind, energy has been an important factor of development; the availability and exploitation of new energy sources has accompanied and made possible major economic and social change [1]. In the Palaeolithic age, the use of fire allowed cooking and conserving food, and protection against winter cold: this produced a less day-by-day life and hence a more complex and stable social organisation; it also brought to an extension of human habitat. Ten thousand years ago, the use of animal power was an important component in the advent of agriculture. In the Renaissance, the use of wind for navigation and for windmills contributed to a new development of agriculture and to broader horizons for commercial and cultural exchanges.

The industrial revolution at the beginning of the nineteenth century was made possible by the use of hydropower and more extensively of coal as an energy source. This involved preferential development of countries having large coal reserves (such as Britain and Germany) and, within them, to the birth of new large cities that rose in the vicinity of coal deposits. The rebirth of the economy after World War II and the great development of industrialised countries were made possible, and at the same time deeply conditioned, by the availability of oil: an abundant, flexible and — most of the time — cheap energy source.

In the past, at each time one form of energy has been predominant over the others (Fig. 1). As a new, more convenient form of energy became available it gradually took the place of the former. Will this be the case also for the future? Although natural gas, nuclear and renewable energies have been indicated in turn by their supporters as the inevitable hegemonic energy source of the future, it is our opinion that we are going towards an era of more open options, with several sources coexisting, to be used in a greater or smaller proportion according to local needs and opportunities, to changing conditions over time, and to different applications; the key to such a future is increased flexibility. This consideration is particularly true

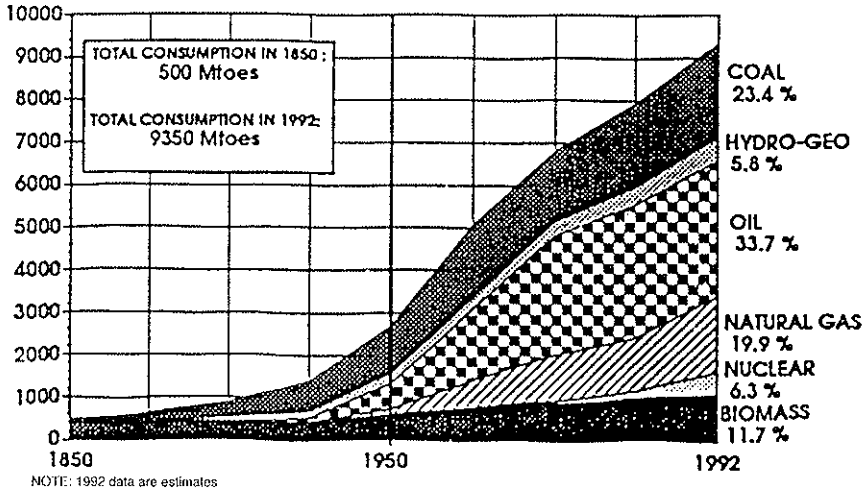


FIG. 1. Historical world energy consumption by source.

for developing countries, which, as we shall see, are facing a major energy transition although they are less bound by investments made in the past.

2. MOTIVATIONS FOR ALTERNATIVES

Fossil fuels account for the great majority of energy consumption at the world level (about 78% of the total energy consumption, including non-commercial sources). The famous report by the MIT for the Club of Rome, "Limits to Growth", published in 1972 [2], pointed to the finite nature of fossil fuel and mineral resources (as well as to the limited carrying capacity of the environment). When the first oil crisis shocked the world just one year later, it seemed to many people that the gloomy predictions of that report were already taking shape.

We now know that the oil crisis was not due to physical exhaustion of the reserves. Although fossil fuels are non-renewable, finite resources, the threat of their physical exhaustion has not yet materialised, and their known reserves are higher than at all times. This can be better perceived in looking at the ratio between assured reserves and production for oil, for gas and for coal (Fig. 2). As the need increases, new reserves are found, thanks also to new and improved methods for exploration and for the exploitation which allow economical recovery of more fuel from a given deposit. Concern about the assurance of the availability and stability of prices of fossil fuels

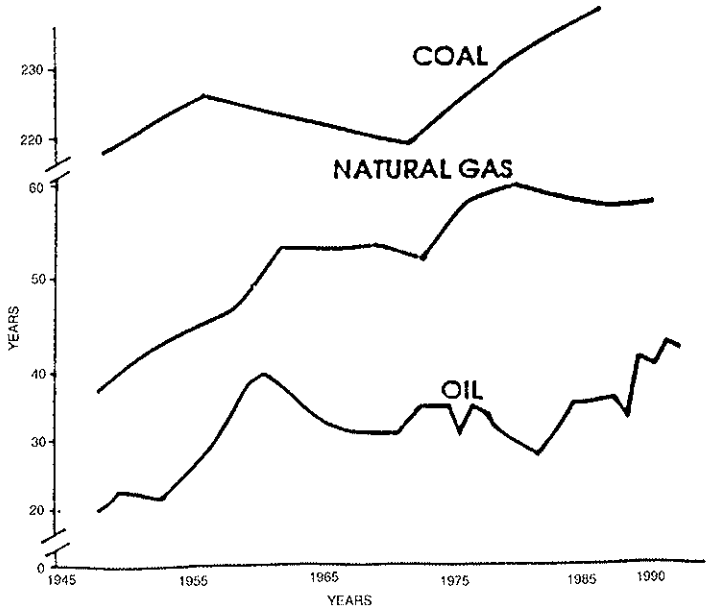


FIG. 2. Ratio of resources to yearly consumption for fossil fuels.

(in particular oil), at least for the short and medium term, is less incumbent at present than in the past. In principle, situations of price instability could occur again in the future. Reserves in the Middle East (the largest at the world level) have extraction costs that are much lower than those of other deposits and can therefore dictate to a large extent the price of oil on the international market. The two oil crises of 1973 and 1979-1980, with the consequent world-wide economic depression, and the countershock of 1985 have shown how the rigidity that affects both demand and supply can lead to instabilities in the market. Prices (in particular of oil) have shown a marked volatility which, however, seems to reduce with time. Oil prices are not very different today (in real terms) from what they were before the first energy crisis.

Scarcity of foreign currency has been another driving force for alternatives. However, in most cases the alternatives are capital intensive, and they often require investments in hard currency. Only for certain alternatives and for countries with an advanced industrial infrastructure (Brazil, India, China ...) this motivation can be valid.

The most important limit to the use of energy today derives from preoccupation on environment and climate. Many negative effects on the environment derive, directly or indirectly, from the energy cycle. In recent times, the local effects (which have always been recognised) are accompanied by increasing preoccupation on regional and global effects [3].

Acid rains, deriving from the release of sulphur and nitrogen oxides in the combustion of fossil fuels (particularly, but not only, coal) occur up to many hundred kilometres away from the polluting source and have adverse effects on agriculture, on forests, on lakes and on the conservation of manufacts and works of art. Although acid emissions have been drastically reduced in industrialised countries, they are increasing in developing countries, especially those (like China) that have ample domestic resources of coal. Reducing the emissions is feasible (although at non-negligible costs) for large installations, such as power plants; it is much more difficult for smaller users, such as domestic heating or cooking.

Other energy-related environmental problems derive from oil spills from tankers, which are often cause of major ecological disasters; from nuclear accidents, which in the case of Chernobyl have spread radioactive products over half a continent; and the contribution to deforestation and desertification deriving from the non-sustainable use of firewood.

The most striking case of adverse effects of the energy cycle is that of possible global climate changes associated with man-produced greenhouse warming [4]. Although still uncertain and debated in its magnitude and in the distribution of its effects, the consequences of the global warming associated with the increasing concentration of greenhouse gases (GHG) in the atmosphere could be so disruptive to cause great preoccupation. Even in a situation of uncertainty, there is thus today a general consensus (underlined by the Climate Convention at the Earth Summit of Rio de Janeiro in 1992) that precautionary measures to contain GHG emissions are well justified as an insurance policy for the future.

Thus, the main motivation for energy alternatives is today to be found in environmental and climatic worries. For developing countries (DC), this plays in two ways. Directly, because the degradation of the environment is obvious in many of these countries, it has negative effects on the quality of life that are easily perceived and sometimes interferes with productive activities. Indirectly, because the sensitivity to global environmental and climatic problems is high in advanced countries and this tends to condition aid-to-development programmes, giving priority to initiatives that take into account the needs of sustainable development.

Finally, consideration has been given to the role of energy as a cause of conflicts (or to its possible role in maintaining peace) [5]. Wars have been fought for energy sources, particularly oil, even recently. Energy-related lack of water has been a major cause of conflicts in many regions, including the Middle East. In the last period, the share of gas, coal and even electricity sold on the international market has been steadily increasing, while that of oil has always been high (Fig. 3). Links are thus created between countries (in particular IC and DC) which may have stabilising as well as destabilising effects. The question whether indigenous sources, as most renewables, would alleviate causes of conflict, or on the contrary remove relations that may play a positive role, is still wide open [6].

We will not consider in the following the shift from one fossil fuel to another, although this may have some importance in particular cases. Actually, two opposite trends can be observed. Countries having indigenous resources of coal tend to extend its use, in order to curb the need to import hydrocarbon fuels (and oil in particular); the use of coal (which is to be found in many areas of the world, including USA, Canada, Australia, South Africa, China and Russia) has been promoted also in countries with little or no domestic deposits, in order to improve the security of energy supply. On the other hand, the use of natural gas as a replacement of oil and especially of coal, is one way of reducing the emission of carbon dioxide (the major GHG) and local or regional pollution. With new combined-cycle natural gas power plants reaching net efficiencies above 55%, the electricity production from gas becomes competitive in many situations with that from coal-based plants.

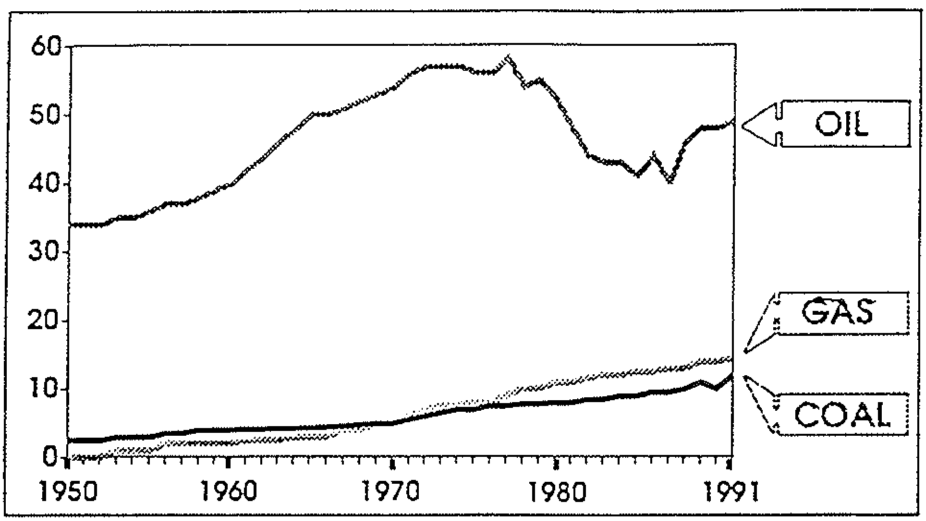


FIG. 3. Share of energy sources on the international market.

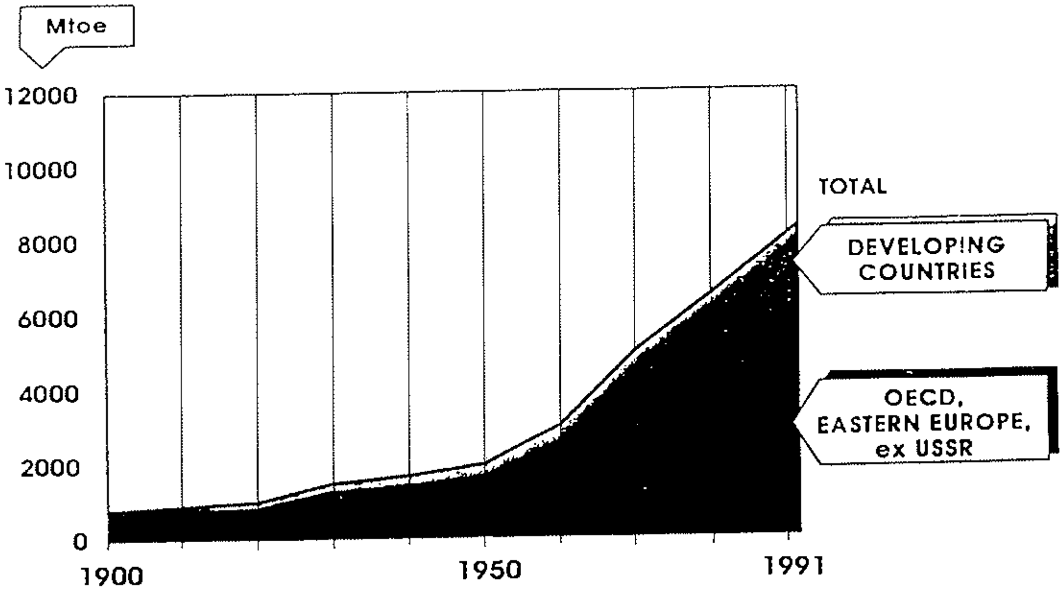
3. Energy needs of developing countries

The gap between energy consumption of industrialised countries (IC) and DC is very big. With 1/4 of the population, the industrialised world consumes 4/5 of the commercial energy (Fig. 4). The average inhabitant of an IC can dispose of ten times as much the energy of his counterpart in a DC. This ratio becomes even much bigger if we consider particular countries (Canada and Bangladesh have a ratio of nearly 100). If we consider further the gap between the more affluent and poorest layers of the population (which is generally much higher in DC) the ratio between the world's rich and poor in terms of energy availability is enormous.

However, DC use energy much less efficiently than IC (Fig. 5). This is not only due to their stage in the industrialisation process (much energy and energy-intensive materials are needed to build infrastructures), but also to the unavailability in DC of the most efficient technologies (aid-to-development programmes in IC generally prefer the transfer of technologies which are already obsolete).

If the Third World is going to industrialise (and this is clearly under way at least in the major countries of Asia and South America), much more energy will be needed in the future by the development of their economies. Whicbever the way in which this energy will be produced, it is essential to limit this growth by adopting the most efficient technologies which are available today in IC and even by "leapfrogging", i.e. employing technologies which have not been tested on a full scale in IC, provided they are adapted to the actual needs and conditions of the DC in which they are used [7]. However effectively these improvements in energy efficiency are applied, the energy demand of DC is going to increase substantially and rapidly in the next decades.

As usual when dealing with DC, lumping them together under one definition is generally misleading. Some DC are practically completing their full industrialisation, both in South East Asia (such as South Korea, Taiwan, Singapore, Malaysia) and in Latin America (Brazil, Argentina, Chile, Mexico). Others are experiencing a sustained and fast economic growth, notably China and to a lesser extent India, Indonesia and others. Another group of countries is rich in fossil energy resources (Saudi Arabia and the Gulf countries, Venezuela, Indonesia, Libya, Algeria, Nigeria ...) and can count on revenues from exporting energy products to fuel the growth of their economies. Sub-Saharan and Central African countries are those which share with some other selected countries (such as Afghanistan and Bangladesh) the worst economic and living conditions, with substantially no sign of improvement in the last decades. This distinction will be kept in



Source: World Energy Conference

FIG. 4. World Energy Consumption: Historical Data.

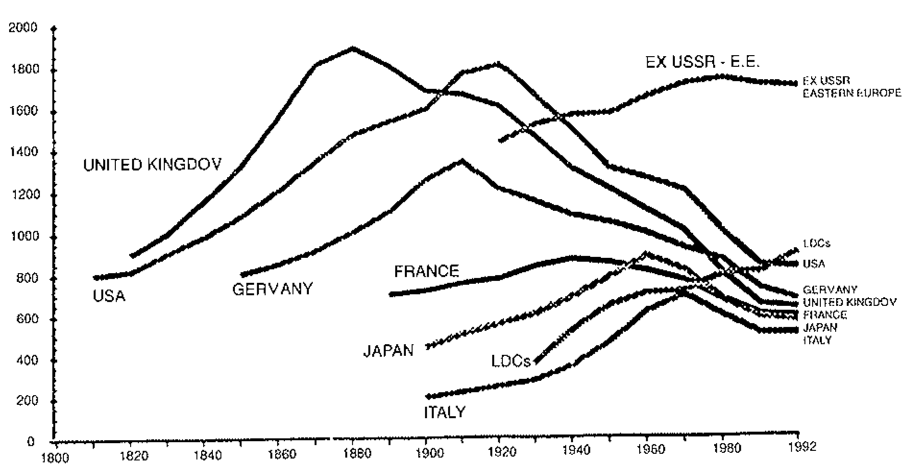


FIG. 5. Historical trends of (commercial) energy intensity of GNP for several countries.

mind in the following, although reasons of space will not allow to consider separately countries or groups of countries for most of the energy alternatives or applications.

4. Nuclear Energy

One form of energy which represents an alternative to fossil fuels is nuclear energy. Nuclear energy in normal operating conditions releases very little pollutants, radioactive or other. It does not contribute greenhouse gases to the atmosphere.

At the world level, nuclear energy provides today 17% of the electricity, or 6% of the total (commercial) primary energy. Most of it is in IC: nuclear accounts for 23% of the electricity (9% of total primary energy) in the OECD countries, and 33% of electricity (12% total) in the European Union. In DC, nuclear energy contributes about 3% of the electricity (less than 1% of total energy): however, most of the nuclear contribution is concentrated in three countries that are at the threshold of entering the group of industrialised countries: South Korea (where 42% of the electricity is of nuclear origin), Taiwan (37.4%), Argentina (14.4%).

Nuclear energy is experiencing a difficult period in most parts of the world. Although in a number of countries (such as France and Belgium) nuclear power accounts for well over one-half of electricity production, and in many other countries it supplies a sizeable share of power in good safety and economic conditions, very few new orders for nuclear power plants are being placed, so that it is unlikely that in the next ten to twenty years the installed nuclear power will grow at the world level, since old plants being decommissioned may more than compensate for the new plants coming into operation.

The reasons for this situation of near stalemate are many. The growth of energy demand in developed countries has proceeded at a much lower rate than previously predicted. France has postponed its plans for the construction of new nuclear plants (apart from the completion of the four units presently being built) until after the year 2000, since the very high percentage of nuclear in the electricity production mix is already causing some problems in load-following. In other industrialised countries, worries on safety and radiation hazards by the general population, especially after the Chernobyl accident, have played a negative role; however, the main motivation is the lack of economic appeal. With a price of fuels on the international market remaining basically low, with advanced natural gas power stations having efficiencies well exceeding 50% (58% in the last reported case), with little short-term preoccupation on the availability of fossil fuels, nuclear energy does not present an economic advantage except in special situations. In particular, in many countries, safety considerations have negative economic impacts: not only because they introduce extra safety equipment in the plant, but also because they make licensing and authorisation procedures longer and more uncertain, introducing higher capital costs during construction and making it difficult to perform a reliable investment plan in advance.

Moreover, the process of privatisation of the electric power industry has made it necessary for the utilities to find capital for investments for power stations on the market rather than from government allocations. The high capital intensity of nuclear stations, the large scale intrinsic in nuclear power and the uncertainty of the times of return of the capital (which is in any case very long) constitute a comparative disadvantage for investments in nuclear power, even when the long-term economic perspective would be favourable.

Similar considerations apply to developing countries, where lack of capital is a constant, where the grid is often too small to accommodate large-scale plants, where the scarce distributed technological knowledge is an obstacle to such high technology applications and where many necessary infrastructures are missing.

Possible exceptions to this picture are East Asian countries. The Japanese government has reaffirmed its commitment to nuclear power, planning to build 20 new nuclear power units by the year 2010 [8]. This will bring Japan to produce 32.2% of its electricity by nuclear. South Korea has nine operating power reactors in four sites, for a total of 7,220 MW_e, supplying 42.2% of the country's electricity; new constructions are not foreseen for the time being [9]. Six nuclear power units provide 37.4% of Taiwan's electricity. The construction of a seventh plant was suspended in 1986 due to public concerns over safety and waste management but then resumed in 1993 for completion by the year 2000.

China is at the beginning of its nuclear programme, with a 300 MW_e demonstration power plant at Quinshan and two PWR units of 900 MW_e each at Daya Bay now in operation; two more units of 600 MW_e are being built at Quinshan, and there are several other projects at various stages along the pipeline. Nuclear energy is perceived as an important contribution to energy supply in the coastal region, where most of the industry is concentrated and which are far away from the coal deposit. A reference nuclear plan by the Central Government considers 32 GW_e in the coastal provinces and 6 GW_e inland as a minimum scenario. However, the lack of capital, the decentralisation of the decision to build power plants to the provincial and utility levels, and the lack of adequate infrastructure make the implementation of this programme rather uncertain [10].

Altogether, a general return of massive nuclear constructions does not seem probable for the time being. In particular, it is difficult to see nuclear energy as an important component of the energy future of DC except maybe for a few countries. Advantageous would be the development of smaller-size units, of simpler systems with intrinsic safety features, the development of internationally accepted standardised safety criteria.

For the very long term, entirely new fission reactors, accelerator based systems or fusion devices [11] might play a role. However, the perspectives are so uncertain and the time frame so long (especially for DC) that they will not be considered here.

5. Renewable energies: characteristics

Renewable energies are by definition sustainable. Although they are not free of environmental concerns, their effects on the environment are generally smaller than those of traditional energy sources. It is therefore to be expected that the long term energy scenario shall rely heavily on renewable energy sources (RES).

The total availability of RES is very large. In principle, they could supply many times the total energy needs of the world. In practice, their availability is limited by a number of constraints, especially of an economic nature [12]. We shall give in the following some data on economically accessible RES, which will show that even if economic considerations are taken into account, the potential of RES remains very high. A summary of these considerations is given in Table 1, taken from ref. [13].

A characteristic of RES is their wide-spread distribution. Even if solar irradiation, wind energy, biomass etc. differ very much from one-country to another, no region of the world is deprived of one or another form of RES in quantities that could contribute very significantly to their energy budget.

Although some applications of RES are based on simple, locally available technology, in many cases they require advanced and sophisticated technology. In other cases, the best solution can derive from the blending of traditional and advanced technology. Research in the field of RES is therefore of great importance, and can determine quite different routes of development; this is particularly true for some innovative approaches, such as photovoltaics or the advanced uses of biomass, but even more traditional sources of energy, like hydroelectricity, are still evolving and their further diffusion will involve improvements and adaptation to different and new situations.

Table 1 - *Res Annual Accessible Potential by Region & Technology* (Mtoe)

WITH PRESENT TECHNOLOGIES	HYDRO	GEO-THERMAL	SOLAR	WIND	WOOD, COMMER.	WOOD, NON COMMER.	BIOMASS, ENERGY CROPS	BIOMASS, WASTE	TOTAL
NORTH AMERICA	169	3.5	17.5	16.4	204	7.5	13.6	67	498
WESTERN EUROPE	174	3	6.5	14	43	5	5.3	52	303
JAPAN AUSTRALIA NZ	34	3.6	6.5	9.8	17	0	3.9	16	91
EASTERN EUROPE	180	3.1	7.5	1.8	218	15.5	15.4	56	497
TOTAL NORTH	**557**	**13.2**	**38**	**42**	**482**	**28**	**30**	**191**	**1389**
LATIN AMERICA	144	1.5	17	1.8	324	158	15.7	55	717
N. AFRICA MIDDLE EAST	7.5	0.1	11	0.1	10	11	0.9	6	47
AFRICA	24.5	0.3	19.5	1.2	70	200	3.9	24	343
INDIA	17.5	0	43	5.3	9	66	5.2	69	215
CHINA	71.5	2.1	32	6.3	17	48	3.2	84	264
ASIA OCEANIA	56	2.7	39.5	3.3	68	162	2.8	65	399
TOTAL SOUTH	**321**	**6.8**	**162**	**18**	**498**	**645**	**32**	**313**	**1996**
TOTAL WORLD	**878**	**20**	**200**	**60**	**980**	**673**	**70**	**504**	**3383**
Rel. Share	26%	0.6%	6%	1.8%	29%	20%	2%	15%	

6. A Review of Renewable Energy Technologies

6.1 *Hydropower*

Large scale hydropower plants are based on mature technologies and have been built with varying degree of success in all the regions of the world [14]. Hydropower is the largest renewable energy source in most industrial countries. At the world level, it provides about 15% of the generated electricity. In 1988, the total installed capacity of 549,000 MW_e have generated 2040 TWh. Estimates of the technically exploitable potential indicate a figure of about 15,000 TWh. Pumped hydropower plants also play a major role as a means of storing energy for peak load demands.

The economy of large scale hydropower projects depends very much on the site, and in some cases it can be very favourable. Their operation is relatively simple.

The number of appropriate sites is limited, and in industrial countries the most favourable ones have already been used; a large potential remains in developing countries, where the implementation is limited by the availability of capitals, by adverse environmental effects and by the lack of a large scale distribution grid. Recent advances in electricity transmission, especially high voltage direct current (HVDC) lines, have increased the economically acceptable distance between production and utilisation of electricity; for some of the otherwise more favourable sites, however, insufficient presence of potential users (particularly industries) within a few hundred kilometres and technical, financial and environmental problems involved in HVDC lines have discouraged the construction of large scale plants.

Large scale hydroelectric plants, moreover, may have a negative impact on alternative uses of water (in particular for irrigation) and of land, and may require displacements of population. Environmental concerns have also been increasingly voiced; they include erosion, sedimentation, damage to flora and fauna, climatic effects, seismic aspects and diffusion of certain diseases. Finally, in developing countries geological complications have often created problems and caused cost overruns.

In order to improve prospects for new hydropower plants, it is necessary to devote increased attention to the integration of the plant with other water uses: a dam can actually be beneficial for water management and irrigation, if properly designed and operated. It is also necessary to develop adequate capabilities of assessing and predicting environmental and other effects.

Small scale and mini-hydro plants (up to a few MW) pose fewer problems of such a nature. Their diffusion in the last 10 years has been smaller than predicted, but still important. There has been a considerable effort by producers of hydraulic and electric equipment to improve the quality and the economics of the components; such improvements include design, standardisation, manufacturing methods, choice of materials, automation and remote operation. There has been a certain concentration of the industries of this sector, with positive effects on costs and standardisation. Remarkably, some of these industries of international importance are located in developing countries (China, South America).

Innovations in this field have concerned many aspects of the plant, that range from new dynamic seals to inflatable rubber transverses, from simple control systems to devices that facilitate fish transit. In some cases, traditional concepts have been revived with new materials and improved design. Many plans for integrating water uses have been implemented.

Micro-hydro plants (stand-alone plants, i.e. not grid-connected, and generally below 100 kW) have seen less development, especially as components are concerned. Components meant for other uses (such as reversed pumps, or automobile generators) are often employed. A positive development concerns electronic load controllers to replace expensive and unreliable mechanical governors.

6.2 *Fuelwood and Charcoal*

Fuelwood continues to play a major role in many of the less developed countries, where it represents the largest contribution to the energy budget of the rural populations (although generally outside the commercial channels). The low efficiency in the use of wood (often below 10%, both when it is burnt directly and when it is transformed into charcoal) has prompted a number of initiatives to develop and diffuse wood stoves for cooking and heating that have a much higher efficiency of energy use [15]. Although these implements are simple and often derive from modifications of traditional stoves, advanced design methods have helped in looking for optimal designs and in confirming empirical suggestions.

Some of the programmes to introduce advanced stoves have been successful, especially those that took into due account local social and cultural barriers. In China, 120 million fuel-saving stoves are reported. In other countries, a reasonably successful application of better stoves concerns charcoal stoves for urban areas, where, unlike in rural areas, both fuel and stoves are commercial commodities, often costly in relation to incomes.

However, despite these positive developments, on the whole the situation is today not much better than it was 10 years ago. The trend towards deforestation and desertification has continued, especially in semi-arid and mountainous regions. Further improvement in fuelwood uses should be based on an assessment of past programmes, learning from past failures. Moreover, such programmes should be framed in a broader context, including programmes for reforestation, consideration of fuel collection and storage problems, use habits, barriers to commercialisation and market penetration. Many difficulties have to be overcome in order to implement such integrated programmes, including problems of tenure of land, trees and wood related to reforestation programmes; the use of wood for other productive activities; and the need to formulate effective policies for managing resource substitution and energy transition.

In some cases, where fuelwood is increasingly scarce and inversion of deforestation cannot be readily attained, it may be necessary to shift to different energy sources altogether. Agricultural residues may be one alternative, which is already extensively exploited. In other cases (as in some examples in Latin America) where liquefied gas (LPG) is available, its high intrinsic efficiency in small scale applications may justify its penetration for the basic uses of rural families. Since the present use of fuelwood can hardly be classified as a renewable energy (as it brings to a progressive depletion of the source on which it is based), it is more convenient from an environmental and climatic point of view, and in terms of conservation of resources, to rely on a more efficient, albeit explicitly non renewable, source of energy such as LPG.

6.3 *Biomass*

Biomass accounts today for about 15% of world energy use and 38% of energy use in developing countries [16]. Many evaluations have been made on the potentials of energy biomass; the results are so dependent on assumptions of various types that it is difficult to come to definite figures on which agreement can be found [17]. However, biomass will remain, at least for three decades, by far the most important renewable energy source, especially in DC.

The energy utilisation of biomass can be seen under two complementary points of view: the source of biomass and the technology used for its transformation and utilisation. Biomass for energy use can derive from residues or by-products of food or other production, or it can be cultivated ad hoc (energy crops).

While at least a part of agricultural residues should be recycled on the place to improve soil fertility, many organic residues and wastes of agricultural, urban and industrial origin are in any case available and their energy utilisation often solves an environmental problem.

Food industries, which are often the first step towards industrialisation in developing countries, produce organic wastes that may be transformed into methane (via anaerobic digestion) or used for heat and/or electricity, which in turn can be used at least in part by the same industry. The sugar (and the ethanol) producing industries, for instance, generate large quantities of bagasse, which up to now has been only partially utilised; the same applies to parts of the plant that are not directly employed in the food processing. Their use in an efficient cogeneration plant, for instance based on biomass gassification and on steam injected gas turbines, could produce the heat necessary to the industry, and electricity much in excess of its needs. The stillage from the same industries can generate important quantities of biogas, and the residue is more suitable for fertilisation than the original liquid. Similar situations exist, for instance, in the fruit processing and in the vegetal oil producing industries.

The rapid growth of large cities in the Third World generates great amounts of solid and liquid wastes, that, if properly collected and treated, could yield valuable energy for urban use. The production of gas from municipal solid waste (MSW) landfills, or the burning of MSW directly or through the production of residue-derived fuel (RDF) for electricity generation are useful options. Much of the combustible waste, however, is often collected and used by poor parts of the population before it reaches the incinerator. Anaerobic fermentation of liquid wastes or sludge, with biogas production, can at least reduce or eliminate need for external energy in the waste water treatment. It should also be mentioned in general that processes that avoid uncontrolled fermentation of wastes, with release of the produced gas to the atmosphere, are doubly important in terms of global climate: for their energy value (and therefore for the avoided effect of the fuel they replace) and for removing a source of emission of methane, which is a greenhouse gas at least 20 times more effective than carbon dioxide.

Crops specifically meant for energy production compete with food for land utilisation. In particular circumstances, however, especially when excess productive land is available, energy crops may be justified. The use of marginal lands for energy crops should also be considered. The very large scale production of ethanol fuel from sugar cane in Brazil (the so-called ProAlcool programme) has shown that such a process is indeed possible and could become economically interesting with higher oil prices and with improved technologies. It should be mentioned that this programme, implemented with a considerable speed, has made use of technologies that were already available, with limited improvements. Effectiveness of the process could presumably be greatly increased by developing new advanced technologies. Environmental aspects as well as land utilisation problems have to be considered more accurately, but they appear to be solvable. The use of oleaginous plants to produce oils to be used in Diesel engines is another possibility which is receiving increased attention.

Technologies for the transformation and utilisation of biomass cover a wide range, from well-established technologies such as direct combustion to those in the research stage, like enzymatic hydrolysis of cellulosic materials to produce ethanol or biomass gasifiers coupled to combustion turbines to produce electricity. All these technologies have witnessed improvements in the last years, and the cumulated experience is today much greater than ten years ago; for instance, the United States alone has more than 8000 MW_e in biomass-powered electric generating capacity [18]; however, most of the advanced biomass conversion technologies with the greatest promise still remain in the developmental stage.

Very simple technologies based on local materials and capabilities have not always lived up to expectations. One example is anaerobic digestors: the simple, backyard type widely diffused in some countries (notably China) have shown problems in operation and endurance, and have in many cases been replaced by systems which are still simple, but use somewhat more sophisticated technologies. The blending of improved and even advanced technologies with traditional concepts is proving a successful approach in this field [19].

To take up just a few examples of diffusing or emerging technologies, gassification of residues such as rice husks is receiving some diffusion in South East Asia and Latin America countries, as a more effective use than direct combustion. A concept that has raised some interest, but is not yet diffused in field use, is the coupling of gassifiers with gas turbines (possibly with steam injection and intercooling) for electricity generation. Another example of a technology with promises for the longer term is the treatment of cellulosic materials by steam explosion, which may be followed by biological or chemical hydrolysis to produce ethanol or other fuels. Small stirling (hot air) engines capable of using wood or agricultural residues as fuel have recently received renewed interest after a period of neglect. Pyrolysis of various biomass products is being investigated, particularly in the European Community, to generate liquid biofuel or a biooil/biocoal slurry that could replace fuel oil in existing power plants.

The whole field of biomass for energy use is subject to important and perhaps unforeseen developments. The very efficiency of conversion of solar energy into energy-rich organic compounds by means of photosynthesis is being investigated in its fundamentals. It is conceivable that modern genetic engineering methods could bring to significant improvements in the photosynthetic process; these should be coupled with corresponding improvements in the successive "dark" phases of plant synthesis to allow greater yields than possible today, without having to recur to a massive use of fertilisers.

6.4 *Solar thermal and thermodynamic*

Direct thermal uses of solar radiaion, or its indirect uses through thermodynamic cycles, involve a large number of different technologies and applications. Some of them have made good progress in the last decade, some are still of uncertain success, others are being abandoned.

Passive heating and cooling of buildings is a practice that has traditional roots in many countries, and that has been revived in recent years by applying modern analysis and computing methods. Efforts have recently been made of blending the new methods with local materials and traditional building practices. Such schemes could have an important effect on the comfort and on the energy consumption both in rural and in urban areas. The main need in this area is to introduce environmental comfort, and the associated energy consumption, as basic parameters of architectural design at all stages, just as the structural aspects or the distribution of spaces. This involves not only the preparation of appropriate, user-oriented design tools that integrate these aspects, but also a major change of attitude in architects and building engineers.

Solar water heaters have made steady technological progress, but their diffusion is slower than anticipated. They provide a commodity (hot water) which is not common in many developing countries (but still widespread in middle and high income strata of the population and in tourist applications), at prices which are competitive with production from electricity, especially with respect to prevailing inefficient heaters, but not yet with natural gas where this is available, distributed and correctly used. In several countries, competition is not easy with LPG (often the most common source for water heating) because of the high subsidies received by LPG. However, 1.8 million m^2 of solar collectors for water heating are reported from China. Technical improvements concern more economical methods for producing the collectors, better and simpler control systems, more rugged and reliable components. Efforts are necessary in most countries to provide cheaper and better installation and maintenance services.

Solar dryers for crops are a promising technique that still needs development. They can be built according to simple schemes and using inexpensive materials, but should include a minimum of control and temperature regulation in order to achieve optimum results on the various types of crops. They can effectively replace oil burning dryers, and help avoiding post-harvest losses with respect to the traditional open air natural drying practices. Year-round applications, such as timber drying, or integration of different applications, favour the return of the investment. When waste products are present that can be used as fuel (e.g. burning of rice husks for hot air drying of rice) this can represent a more practical approach.

Solar cookers had drawn much attention and raised hopes as a substitute for inefficient wood-stoves. The results have so far been mostly disappointing, especially for the difficulty of coping with local habits (such as cooking at early morning or at night, or cooking in-door) and for the lack of availability in cloudy weather. As a consequence, interest in this technology is generally declining.

By concentrating solar radiation and/or using special materials, temperatures well above 100 °C can easily be reached. Fluids at such temperatures can be used in a number of applications, including refrigeration (by means of absorption coolers), desalination (e.g. by a multiflash process), process heat or steam for industrial uses etc. Applications of all these types have been reported, and in many cases they have been a technical success. However, their diffusion is for the time being very limited essentially due to economic reasons.

Solar ponds are a technique that not only allows collecting solar energy but also storing it as heat for relatively long periods. They may be economically interesting when rather special circumstances occur: in particular, when salt is freely available and natural enclosures can be used for the pond. A number of solar ponds exist or are being built in several countries. The largest in operation is in Israel, on the borders of the Dead Sea, and generates 5 MW electric power. Others are being used for desalination and for heat storage.

Several different technologies have been proposed and tested for producing electricity from solar energy through thermodynamic cycles operating at relatively high temperatures. This temperature can be obtained through different types of concentrators: mirrors aiming solar radiation at a central tower, parabolic dish collectors and parabolic troughs. Only the last have met with success so far. An Israeli-US venture (presently unfortunately disappeared because of financial difficulties) has built a number of power generating solar installations, now totalling over 350 MW_e, in Southern California, selling electricity to the grid at competitive prices.

Scientific research has played a role in the development of solar thermal and thermodynamic energy: selective coatings, high insulation panels, advanced heat-transfer systems have brought about continuous improvements with important results on the final product. Today it seems unlikely that scientific breakthroughs could change drastically the perspectives of this RES; it is more probable that future progress will continue to be based on incremental improvements, on better systems design, on the adoption of better and cheaper materials and on the adaptation of systems to the particular conditions of utilisation.

6.5 *Solar photovoltaic*

Photovoltaic (PV) direct conversion of solar energy into electricity, although it is much more expensive in general than other technologies, has long been considered a very promising source because of two reasons: it is a highly innovative technology which is not based on mature components, and therefore has a high potential of cost reduction; secondly, because of its near independence on scale (which makes even very small plants relatively interesting), and of the fact that it requires very little, if any, intervention for operation and for maintenance (which makes it ideal for remote applications in isolated sites). PV has an intermediate market: even if its cost is today much higher than conventional sources for bulk electricity production, there are many applications for which it is already competitive, and the size of this intermediate market increases rapidly as the price of PV systems decreases (Fig. 6). In this way, since the potential market curve was projected to be consistently higher than the supply curve, there would be a steady penetration of PV even without incentives; public funds could be used to speed up diffusion rather than to fill up economic gaps.

These considerations remain substantially valid. In particular, the cost reduction as a function of production volume has been confirmed, and in some cases it has even surpassed expectations. The diffusion of PV energy, on the other hand, has been much slower than anticipated. The PV shipment has grown steadily, doubling its size every 4 to 5 years to the present value of about 60 MW/year. This is indeed a remarkable success, if compared with the diffusion rates of other innovative technologies, but due to the very low starting value it will take between 30 and 40 years at this rate of growth before PV takes up a significant share of the world energy budget. In order to keep up with this rate and possibly accelerate it, a great effort will be needed in the future, involving governments, industries and utilities.

The next class of applications to be addressed as the price of PV decreases was expected to be the supply of power to locations not connected to the grid and for which PV is often competitive with Diesel generators. Such locations are mostly to be found in the rural areas of developing countries, where people often can hardly afford a Diesel generator and little capital is available for the high initial investment needed for a PV system. Hybrid systems, combining PV and Diesel generators, with PV providing between 35 and 70% of the energy, have had a certain

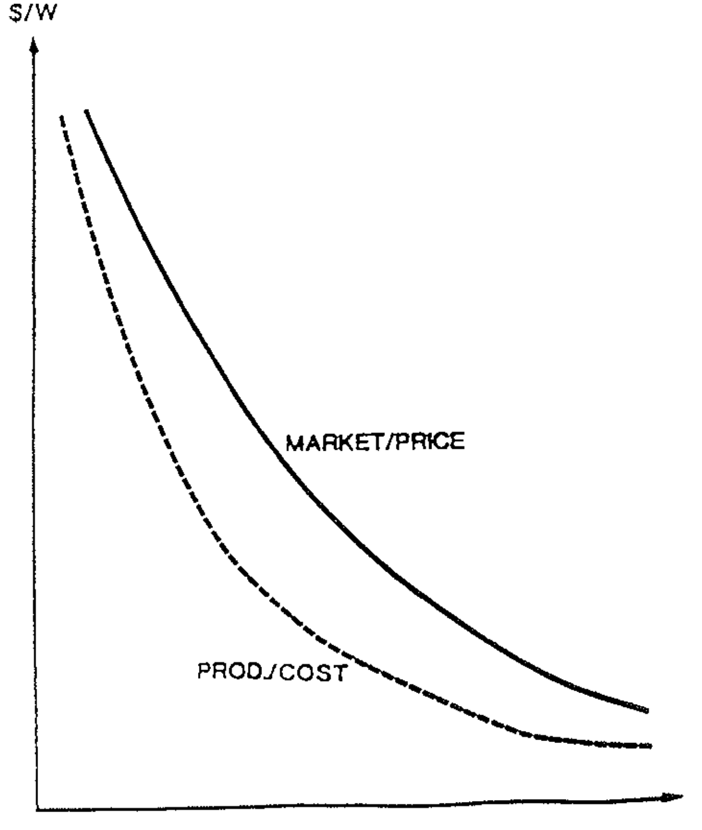


FIG. 6. Comparison between Supply and Demand for Photovoltaic Systems.

diffusion in the last times; such systems can greatly reduce the need for electric storage. Other kinds of hybrid systems, like PV and wind or PV and biomass, are also being explored.

On the other hand, promising applications have recently been identified in special situations of large grids [20]. In parts of the distribution network far from generating stations, in the presence of high insolation, of a good match between insolation and load, and of high costs of alternative generating solutions, grid reinforcements by PV systems may be cost-effective even in the near future. If utilities will adopt a more detailed, space-dependent system of cost analysis that takes in due account demand-side management, the short and medium term perspectives of PV are likely to be positively re-evaluated. On the longer term, increases in conventional fuel prices, more stringent environmental regulations, availability of capitals at low interest rates are factors that would facilitate the penetration of PV systems. It can therefore be envisaged that a diffusion of PV systems in grid-connected applications in industrial countries can proceed in parallel with the stand-alone applications in developing countries, thereby supplying the increased market necessary to cut down production costs.

A further important decline of the price of PV systems is to be expected in the future. This decline should result both from a scale factor (the largest industrial plants now produce a few MW per year, and a scale up of a factor 10 would reduce cell production costs of about a factor 2) and from a shift to new technologies. In order to reduce the cost of material for the crystal silicon (X-Si) technologies, it was necessary to rely on thin films, the most obvious candidate for which was amorphous silicon (a-Si). Actually, a-Si has production costs much lower than X-Si, but its characteristics of low efficiency and initial degradation have made it unsuitable for the general energy market, and its applications, therefore, have so far concerned a quite separate market (consumer electronics and gadgets). It would now be necessary to improve its performance in terms relevant for power systems so that it can contribute to the diffusion of PV into the energy market.

PV cells based on the superposition of various layers of a-Si have shown both higher efficiencies and improved stability. Still higher efficiencies can be obtained by using layers of different materials (for instance a-Si and a-Ge, or different Si-Ge alloys), in which each material optimises the utilisation of the various components of solar radiation.

Other materials are being investigated as candidates for PV cells; among them, gallium arsenide (GaAs) which is also of interest for electronic applications and which could be manufactured in thin films. Copper indium diselenide (the so-called CIS) and cadmium telluride are the other materials that appear more promising and receive most attention today.

Many other possibilities have been investigated in the past, with less success. There is no reason why entirely new PV materials should not be identified and developed in the future. The photovoltaic effect is fairly common, and its full theoretical implications are far from being explored. Scientific breakthroughs could well be expected in this area.

The whole conception of PV systems is in turn not frozen. While most of the effort today is directed towards abating costs of components and structures, alternative routes are possible and have been proposed and in some cases experimented. The most common one is concentrating the sun on the PV cell by means of sun-tracking lenses or mirrors; the surface of the cell is greatly reduced, and one can employ more sophisticated and more efficient materials; the cost is shifted to the concentrating device. This alternative is of interest in climatic conditions in which the sky is very clear, and the direct radiation from the sun is much greater than the diffused radiation from the sky; however, the presence of moving mechanical parts may involve additional needs for operation and maintenance, and the applicability of concentrating systems in remote areas and/or in countries with limited technological capabilities is often questioned.

An even much more innovative approach, which has been proposed several years ago and never experimented, but which is receiving some renewed attention lately is that of orbiting PV stations. PV systems based on stationary-orbit satellites could receive full solar radiation, without attenuation, 24 hours a day and relay the power to an earth based station by means of a microwave beam. Recent developments of this concept have also proposed a moon-based station. Although these concepts are clearly a long way from present technologies, and their economic (and also environmental) implications are still difficult to assess, they are useful as examples of the great variety of solutions which could evolve from RES applications of today.

6.6. *Geothermal energy*

Geothermal energy is steadily extending its contribution to the world energy balance [21]. Since 1980, geothermal electricity generation capacity at the world level has passed from 2000 to 5830 MW_e (from 500 to 2000 MW_e in developing countries). Important developments have taken place in particular in the Philippines (1000 MW_e), Mexico (700 MW_e), Nicaragua and El Salvador. The use of lower enthalpy geothermal fluids for heating purposes (including urban heating, greenhouses, aquaculture and animal breeding) has extended mostly in industrial countries, because of the relatively lower demand for space heating in Third World countries and of the lack of favourable situations (such as consistent reservoirs close to large cities). The present installed capacity for direct thermal uses of geothermal energy in the world is now about 10,800 MW_t.

At the world level, projections for the next 10 years indicate an annual growth rate of the order of 6% for power production and 4% for direct uses.

Although geothermal energy is close to being a mature technology (its first applications for electricity generation are nearly a century old), constant progress is being achieved. Technical improvements of the last decade concern, for instance, the ability of modulating the load in an economic and efficient way in both steam and water dominated situations, even for relatively small plants; the building of simple and reliable single and double flash systems at moderate cost installed in remote locations (such as in

Tibet, Bolivia and Hawaii); and at the higher range of the technology, the development of techniques quite different from those of the oil industry for drilling deep wells (4000 m and beyond) through hard, hot rocks; the improvement of materials and sealing techniques; the possibility of drilling multiple wells branching off from a main well at high depths; the drilling of quasi-horizontal wells; the improvement of geophysical and geochemical prospecting and evaluation methods.

Environmental considerations are playing a much greater role than in the past; re-injection of water and dissolved wastes is now a universal practice; abating gaseous pollutants (such as H_2S) is also being achieved with increasing efficiency.

Techniques for a better and more rational utilisation of geothermal reservoirs are being developed and diffused; they include, for instance, the use of geothermal heat for chemical processing and, in perspective, the production of pure water from the plant (salt water can be substituted for re-injection).

For the longer term, the hot dry rock (HDR) technology has made some steps forward, although a significant demonstration is still lacking (and hoped for the next decade). This technology consists in injecting water in a deep, hot and dry formation through one well and recovering it as steam through another well. The major difficulty consists in connecting the two wells through a system of fractures, having large surfaces for heat exchange but little dispersion of the injected water. If this challenge can be met (as theoretical considerations and some small scale experiment would indicate possible), the availability of HDR technology would allow extending the resource bases for geothermal energy to many more countries, since dry geothermal deposits are much more common than wet ones. DC could well profit from this development.

6.7 *Wind energy*

Wind is the renewable energy source which is in closer competition with conventional energy sources for the production of electricity and mechanical power, at least where appropriate wind regimes are available. Even when environmental, land-use, and systems constraints are taken into account, wind power could accommodate a substantial portion of global electricity demand, perhaps 20 percent [22]. At present (1994) [23)], the installed wind generating capacity is 3,350 MW_e in the world, of which 1,700 in North America, 1,550 in Europe; the contribution of DC is for the time being negligible. The installed power is rapidly increasing, as well as the energy generated (6.1 TWh expected in 1994 against 5.4 TWh in 1993 and 4.4 in 1992).

Since the power that is carried by the wind per unit area is proportional to the third power of its velocity, the convenience of wind generation depends very much on the velocity of the wind that can be expected on the average in a given location. The ideal situation involves high average wind velocities (e.g. 8 m/s), constant or slowly varying wind direction, absence of sudden gusts. Wind regimes are very variable, not only on a regional scale but also at the local scale: the configuration of the ground around the wind machine and the presence of vegetation have important effects. It is therefore necessary to have accurate evaluations and measurements of the actual conditions on each site considered for the installation.

Wind power technology has evolved very notably in the last decade; most of this evolution and of the new applications of wind machines is concentrated in industrial countries and adapted to their conditions. Wind farms providing electricity to the grid have been built in California (over 1500 MW), in Denmark (300 MW) and in some other locations; the cost of the electricity produced is close to (or smaller than) the avoided cost of providing extra power by conventional sources, in particular coal-fired stations with state-of-the art pollution abatement. The average size of the machines employed has risen steadily in the last ten years from a few tens of kW to a few hundred kW. Large scale machines, one or several MW of power, are still at an experimental stage, and it is not sure whether they will eventually become more economical than medium-sized ones, except perhaps where space is scarce or in off-shore applications.

Recent technological developments include advanced airfoils, variable speed machines that employ solid state advanced electronic power conditioning systems, and continuing development of new materials that yield lighter, stronger components. Energy capture has been increased 10% or more, while capacity factors that ten years ago were in the range of about 15% are now as high as 35%. All these improvements are directly reflected in lower generating costs.

In order to extend the use of wind energy in DC (at least where the required climatic conditions are present) it is necessary to develop or adapt the technology to their special needs: in particular, stand-alone systems may be required rather than grid connected systems; environmental conditions may be more severe (as where sand-carrying winds are common); average wind velocities to be considered may be smaller; and reliability is even more important than in IC. However, it is convenient to draw from the experience already gained by industrial countries, where machines tested for several years in the context of frequent exchange of information among users have shown to be the most successful for further diffusion, rather than starting an entirely new process of development.

Smaller machines, up to a few kW, have long been used in many countries to pump water for human and cattle use and for irrigation and to charge batteries. In IC, this practice has declined with the diffusion of rural electrification. The interest for these applications is large in DC. Wind pumps are economically attractive, at least for drinking water, even when wind regimes are not very favourable (e.g. 3.5 m/s average velocity). About 110 thousand wind generators and 160 thousand wind pumps are reported to be operating in China.

Modern technologies developed in IC and in some DC are more reliable, lighter and less expensive than traditional models. However, positive results have also been obtained by using locally available materials (wood, bamboo mats etc.) and variations of traditional schemes.

6.8 *Ocean energy*

Several different sources of energy are present in oceans: tides, currents, waves, thermal gradients, salinity gradients [24]. They have originated experimental and prototype systems at different degrees of development and with various perspectives of diffusion. All these forms of energy are of course available only at shore locations (or just off-shore), in particular at islands; they are also very variable with location.

Plants based on tides require a high tidal difference (typically more than 3 m); special coast configurations are also needed (such as nearly closed bays or river estuaries) in order to reduce barrage costs. Technologies used in these plants are mostly conventional. In addition to the 240 MW_e plant at La Rance in France, which has been operating for many years, a 400 kW_e plant has been in operation since 1968 off Murmansk in Russia; a 17.8 MW_e is being commissioned at Annapolis, Canada, as well as a 3.2 MW_e plus a number of smaller tidal plants for a total power of 1.8 MW_e in China. Three major projects have been considered for the future: one in the Bay of Fundy in Canada (for an initial power of 20 MW_e) and two very large ones (several thousand MW_e) on the Eastern coast of the Russian Federation and on the estuary of the Severn River in the UK respectively. Several other projects have been investigated, for a total of about 70,000 MW_e but it is not expected that positive decisions may be taken on any of the major projects, at least for the next several years.

Wave power is also unevenly distributed, with a bias in favour of high latitudes. Many different schemes to harness this potentially abundant source of energy have been studied; up to now, results have not been encouraging from the viewpoint of economics, although some further research and development seems justified.

Ocean currents carry large quantities of energy; however, their utilisation is in general extremely expensive. Only when special conditions are present (such as the existence of a bridge on which to fix turbines across a channel with high and steady current) could a plant be economically justified.

The exploitation of thermal gradients of oceans (OTEC: Ocean Thermal Energy Conversion) is in principle possible, at least in tropical regions where such gradients are important. A number of small experimental OTEC plants exist, especially in the Pacific. The economics of such systems is difficult, because of the low efficiency of the turbines, of the large heat exchange surfaces implied by the low temperature difference (typically less than 20 °C) and of the need to pump large quantities of water from depths of a few hundred meters. Both closed cycle and open cycle systems have been studied, the closed cycle being at a more advanced stage of development. Some side advantage could be gained from using the deep sea water pumped to the surface for aquaculture, since it contains much greater concentrations of nutrients than surface water; evaluations carried out in Taiwan indicate that this application could make OTEC attractive in that region. The potential issue of release of CO_2 dissolved in the sea water has been mentioned as an adverse side-effect.

Apart from some local situations, it appears unlikely that energy from the oceans will have an important role in the foreseeable future.

6.9 *Draught animal and human power*

The traditional practice of using animals for agriculture, for transportation, for water pumping and for other mechanical work provides an important contribution to the energy balance of rural areas in less developed countries [25]. Although economic development can be expected to displace this primitive source of power, such practices can be of great help in a transient period, and it is therefore useful to exploit it in the most effective way. The use of animal power can also be advantageous in swamps and where soil could be damaged by machinery. Draught animal power can be improved by adopting more efficient harnessing devices, more appropriate tools and implements (such as pumps), better breeding, health care and feeding programmes. The present or recent past experience of industrial countries may help in this connection, although it does not cover all possibilities (e.g. the use of elephants).

Human power remains an important contribution, and it can be used more effectively, just as in the case of animal power. It is sufficient to recall the use of bicycles, which is not only of great importance in Third World countries, but is experiencing a revival in many urban and suburban areas of the industrial world. In the centres of many crowded metropolitan areas of both the North and the South, the average velocity of motor vehicles in the traffic is not higher than that of bicycles, while the environmental conditions are close to being unbearable.

Hand pumps for drinking water may provide a cheap and appropriate interim solution where the water head is low. Surprisingly, reliable, efficient and cheap hand pumps are not easy to find, especially in developing countries. Programmes for the development and diffusion of appropriate pumps have been carried out by UNICEF, by the UNDP and the World Bank and more recently by IRDC in many LDC's with positive results.

7. *Applications of renewable energies*

RES can be used for different applications, that include practically every facet of final energy use. The product or technology considered can yield directly electricity, or heat, or mechanical power, or chemicals to be used as fuels. In turn, electricity can be used for a very large number of applications, ranging from lighting and telecommunication to power for industrial uses; in this presentation, concentration will be on the scale of electricity production and distinguish among sources more suitable for feeding a single small appliance (e.g. a radio, a light, a refrigerator), typically below 100 W; those for a household, typically between 100 W and 1 kW; those providing power at the scale of a village, from tens of kW to hundreds of kW; and those that feed a large grid. Heat is often classified according to temperature (low, below 100 °C; medium, between 100 and about 250; and high, above 250 °C); heat can be used to provide sanitary hot water, or for drying crops, or for space heating; it can be converted into electricity; or it can be used in industrial processes, or in order to operate refrigerators, to desalinate water, to move engines or to activate chemical reactions, that may in turn yield fuels. Fuels can be used to produce heat, or to operate engines that produce mechanical power, that in turn can be used in transportation, or to operate tools, to pump water, or to generate electricity. In other words many pathways are possible, based on successive transformations of one form of energy into another. Without entering these rather complicated schemes (which involve more or less conventional technologies and include the problems of storage and transportation of the various forms of energy), the most commonly used or envisaged applications of the various renewable sources of energy are presented in Table 2.

8. Conclusions

Energy will continue to play an important role in the future world. Its availability at a reasonable price is one of the key issues for the development

Table 2 - *The End-use Applications Available for the Various and Renewable Energy Sources, Technologies and Fuels, with Special Reference to the Tasks Relevant in Developing Nations.*

	Cooking and food processing	Drying, water and space heating	Industrial steam	Cogeneration	Telecommun-ications	Household electricity	Village power	Large-scale grid	Sterilisation and disinfection	Desalination and distillation	Water pumping	Grinding, threshing and wood cutting	Space cooling and refrigeration	Traction power for agriculture	Transportation
Solar photovoltaic					X	X	X	X	X	X	X	X	X		
High-temperature solar			X	X			X	X		X	X		X		
Low-temperature solar	X	X	X						X	X	X		X		
Mini-hydro						X	X	X			X	X	X		
High-temperature geothermal			X	X			X	X		X					
Low-temperature geothermal		X													
Wind						X	X	X	X		X		X	X	
Biomass combustion	X	X	X	X			X	X		X					
Biomass gasification						X	X	X			X	X	X	X	X
Biogas	X	X	X	X		X	X	X	X	X	X	X	X	X	
Biomass-derived liquid fuels														X	X
Ocean							X								
Animal and human power											X			X	X

of Third World countries. The rapid increase of population, especially in developing countries, which will not level off before the middle of the next century, and the expected improvement of their quality of life will demand more and more energy, much as we can do to improve the efficiency of energy utilisation.

Fossil fuels are not running short in the mear future. Although their non-renewable resources are eventually bound to finish, this is not going to happen soon for coal or natural gas, and not even for oil. Prices have shown a remarkable stability in the last years, as compared with the volatility of the previous period. Price surges and crises of availability are certainly possible in the future, especially as a consequence of unforeseen political events, and such scenarios have recently been discussed lately. However, as long as decisions remain rational, they should be less dramatic than in the past: both producers and users have found out that tensions and confrontations produce negative effects for both sides.

Nuclear energy will continue to play a role in energy supply, particularly important for IC. Its share in the energy budget is not expected to rise in the next decades, and in particular it is doubtful that it will contribute appreciably to the increasing energy needs of DC.

At the beginning of the 1980's, the International Conference on New and Renewable Sources of Energy convened by the United Nations in Nairobi adopted a positive and somewhat optimistic view on the future of renewables. The emphasis was then on the availability of energy, and renewable sources of energy were seen mostly as a way out of the predicted scarcity and soaring prices of fossil fuels. A decade later, the United Nations Conference on Environment and Development in Rio de Janeiro, although it regrettably devoted little attention to the specific solution offered by renewable energies, set a completely different scenario. The emphasis is now on the protection of the environment and of global climate: the deterioration of local and global environmental conditions, the increasing recognition of the threats posed by man-produced enhancements of the greenhouse effect have become the main driving force towards the development of energy alternatives, in particular renewable sources of energy and improvements in the efficiency of energy use.

Renewables will not diffuse as a result of technology push. Only a response to an actual demand of the market will change the situation. It remains to be demonstrated that once the obstacles to the diffusion of renewable energies are removed, they can become a substantial part of the energy supply in both industrialised and developing countries. Such obstacles are mostly of a non-technical nature, and derive from legislation that, more or less consciously, incentivates the use of fossil fuels, from energy prices that do not reflect the costs paid by the community, from insufficient information. If really put on an equal footing with other energy sources (the so-called level playing field), renewables should be able to compete on the market even with today's technologies; an effort of R&D to improve these technologies or to develop new ones (which is fully justified in terms of public support and industrial commitment) would make this competition even more favourable to renewables.

This view has some non-negligible strategic implications. Much of the talk about applications of renewables (for instance, solar photovoltaics) has concerned niche markets. Niche markets (i.e. limited or specialised applications for which RES have no competitor, or certainly perform better and are cheaper than alternative solutions) are indeed important in providing an assured basis for the expansion of the market, but the attention should be concentrated on large-scale applications where renewables directly compete with traditional energies.

Another consequence is that the past emphasis on using renewables to alleviate the living conditions of the world's poor (such as supplying small amounts of energy to backward rural villages not connected to the grid) may have been excessive. In such cases, the lack of entitlement of those who directly benefit from the application puts them outside the market. Even if it is certainly beneficial from a social point of view to devote international aid funds and national investments for social support to this kind of projects, more attention should be given to the cases in which the final user either has money to buy, or may make enough money just through the availability of energy which derives from the project, so as to establish a correct market approach. In other words, more emphasis should be given to the possible role of renewables in supporting profitable activities and in creating revenues. If applied for instance to Africa, such considerations may point to a quite different strategy from that followed in the past.

Responsibility is with the industries, of both DC and IC, to make available products and technologies for the production and the rational use of energy which are suited to the conditions of DC and which protect in the best possible way environment and global climate. This will ultimately prove to be in the interest of everybody.

Acknowledgement

The author is grateful to Prof. André Blanc-Lapierre for stimulating discussions and constructive suggestions.

REFERENCES

[1] U. Farinelli, «A setting for the diffusion of renewable energy» in A.A.M. Sayigh ed., Proceedings of the Third World Renewable Energy Congress, Renewable Energy, Pergamon, New York 1994.

[2] D.H. Meadows, D.L. Meadows, J. Randers and W.W. Behrens III, «The Limits to Growth», A Report for the Club of Rome's Project on the Predicament of Mankind, Universe Books, New York, 1972.

[3] S. Postel, «Carrying Capacity: Earth's Bottom Line», in L.R. Brown, C. Flavin and S. Postel, eds., «State of the World 1994», pp. 3-21, a Worldwatch Institute Report on Progress Towards a Sustainable Society, W.W. Norton & Co., New York - London, 1994.

[4] U. Colombo, U. Farinelli, «Energy, environment and global climate» *Perspectives in Energy*, Vol. 3, 1994-1995, in press.

[5] U. Farinelli, P. Valant, «Energy as a source of potential conflicts», *International Journal of Global Energy Issues*, Vol. 2, No. 1, pp. 31-40, 1990.

[6] U. Farinelli, B. Ischinger and H.Z. Tabor, «Solar Energy - Energy for Peace», World Solar Summit, UNESCO 1993.

[7] J. Goldemberg, T.B. Johansson, A.K.N. Reddy and R.H. Williams, «Energy for a Sustainable World», pp. 285-286, Wiley Eastern Ltd., New Delhi, 1988.

[8] NucNet, The World's Nuclear News Agency, News N. 447/94, September 16, 1994.

[9] EIA (Energy Information Administration), US Department of Energy, «World Nuclear Capacity and Fuel Requirements 1993», DOE/EIA-0463(93), Washington D.C November 30, 1993.

[10] U. Farinelli, Zuo Hu, «The nuclear energy programme of China: perspectives and elements of international relevance», report to the China Council for International Co-operation on Environment and Development, Beijing, July 1994.

[11] U. Colombo, U. Farinelli, «Progress in Fusion Energy», *Ann. Rev. Energy Environ.* Vol. 17, pp. 123-159, 1992.

[12] B. Dessus, B. Devin, F. Pharabod, «World Potential of Renewable Energies», *La Huille Blanche*, vol. 1, pp. 1-50, 1992.

[13] United Nations Solar Energy Group for Environment and Development (UNSEGED), «Solar Energy: A Strategy in Support of Environment and Development», United Nations, New York 1992.

[14] J.R. Moreira, A.D. Poole, «Hydropower and its Constraints», in T.B. Johansson, H. Kelly, A.K.N. Reddy and R.H. Williams editors, «Renewable Energy Sources for Fuels and Electricity», pp. 73-119, Island Press, Washington D.C., 1993.

[15] L.A. Kristoferson and V. Bokalders, «Renewable Energy Technologies — their applications in developing countries», pp. 55-91, Intermediate Technology Publications, London 1991.

[16] D.O. Hall, F. Rosillo-Calle, R.H. Williams and J. Woods, «Biomass for Energy: Supply Prospects», in T.B. Johansson, H. Kelly, A.K.N. Reddy and R.H. Williams editors, «Renewable Energy Sources for Fuels and Electricity», pp. 593-651, Island Press, Washington D.C., 1993.

[17] E.D. Larson, «Technology for Electricity and Fuels from Biomass», Annual Rev. Energy Environ. Vol. 18, pp. 567-630, 1993.

[18] J. Tumbull, «Strategies for Achieving a Sustainable, Clean and Cost-Effective Biomass Resource», Electric Power Research Institute (EPRI), Palo Alto, CA, 1993; J. Turnbull, «Use of biomass in electric power generation: the California experience», Biomass and Bioenergy, Vol. 4, no. 2, pp. 75-84, 1993.

[19] Department of Environmental Protection and Energy, Ministry of Agriculture, People's Republic of China, «Biogas and Sustainable Agriculture», BORDA, Bremen (FRG), 1993.

[20] Proceedings of the Executive Conference on Photovoltaic Systems for Electric Utility Applications: «Opportunities, Critical Issues and Development Perspectives» (Taormina, 2-5 December 1990), International Energy Agency OECD/EA, Paris, 1991.

[21] C.G. Palmerini, «Geothermal Energy», in T.B. Johansson, H. Kelly, A.K.N. Reddy and R.H. Williams editors, «Renewable Energy Sources for Fuels and Electricity», pp. 549-591, Island Press, Washington D.C., 1993.

[22] M.J. Grubb, N.I. Meyer, «Wind energy: Resources, Systems, and Regional Strategies» in T.B. Johansson, H. Kelly, A.K.N. Reddy and R.H. Williams editors, «Renewable Energy Sources for Fuels and Electricity», pp. 157-212, Island Press, Washington D.C., 1993.

[23] P. Gipe, «Overview of World-wide Wind Generation», in A.A.M. Sayigh ed., Proceedings of the Third World Renewable Energy Congress, Renewable Energy, Pergamon, New York 1994.

[24] J.E. Cavanagb, I.H. Clarke and R. Price, «Ocean Energy Systems», in T.B. Johansson, H. Kelly, A.K.N. Reddy and R.H. Williams editors, «Renewable Energy Sources for Fuels and Electricity», pp. 513-547, Island Press, Washington D.C., 1993.

[25] L.A. Kristoferson and V. Bokalders, «Renewable Energy Technologies — their applications in developing countries», pp. 119-131, Intermediate Technology Publications, London 1991.

ENERGY RESOURCES:
PRESENT AND FUTURE IN MEXICO

MARCOS MOSHINSKY *

This talk is one of those included in the session on "Alternative Energy Sources for Developing Countries" and the speakers concentrate on the problem of the Third World on this subject, particularly for those countries that have little oil or gas. The only exception is my own country, Mexico, that at the present rate of consumption and exploration could have it for 30 or 40 years. This lapse of time may be considered as an instant in the life of humanity, if it should continue to exist for a period as long as that from the first lumanoids to the present, or it could be considered as infinity, if civilization, as we know it, disappears in that period, which is not an improbable occurrence if we look around us today.

In any case, 30 or 40 years is of the order of a life-time, and certainly much longer than the average period in which politicians are in power. Thus the search for alternative energy sources in Mexico is of interest to the country, as it should be for any other country in the world, but it is not a cause of anxiety.

Some of the people I have spoken with on this subject, when researching material for this talk, were even cynical about this subject. Why, they asked, are developed countries so anxious for the underdeveloped or developing nations to look for alternative energy sources? Are they afraid that their own access to oil and gas will be jeopardized, or that the price will become excessive if the underdeveloped try to get their share of the most common source of energy in the world today?

I do not share this cynical attitude, and in fact in this talk I will indicate to you the efforts being made in Mexico in relation with alternative

---

* Pontifical Academician. Universidad Nacional Autónoma de México, Instituto de Física, México.

energy sources. But before getting into this matter, let me show you a few slides regarding the energy situation in Mexico today.

My data is a bit old, because this type of information is always published in delayed form, due to the difficulty of recompiling it from different sources. I will start with slide N° 1 which indicates the amount of energy produced in Mexico in the years 1991 and 1992 and the variation in them in those years. Note that the amount is given in petacalories (10^{15}) and some of the sources of energy are subdivided, but in a nutshell, and in percentage points, in 1992 coal provided 1.4% of the energy, hydrocarbons mainly oil and gas, 90%, electricity from nuclear, geo and hidro sources 4.3%, and biomass, mainly from sugar cane bagasse and wood from our dwindling forests, 4.3%. Except for non petroleum electricity, the variations between 1991 and 1992 do not exceed 5%.

This imbalance in favor of hydrocarbons can be seen more clearly in the second slide. Most of our oil and gas come from the Gulf of Mexico or the land close to it in the states of Tabasco and Campeche close to the Yucatan Peninsula. Of the total energy we produce, we export about 28% mainly in the form of oil.

Our internal consumption of energy is shown in graph #3, where we see that hydrocarbons (crude oil, condensates, gas) still cover close to 84% of our internal energy need. This oil goes mainly to refineries or gas plants and only about 10% of it is used to produce electricity directly.

In our final slide about the present situation, we show the origin and final destination of all the energy we produce.

Does this abundance of energy resources, particularly in its most desirable form, oil and gas, imply that Mexico has left the Third World? The answer is not yet and the culprit is the population explosion. In 1940 Mexico, the whole country, had a population of 20 million; 50 years later it was 85 million, and it will top 100 million by the year 2000. Mexico City is now the largest in the world with a population close to that of the whole country in 1940.

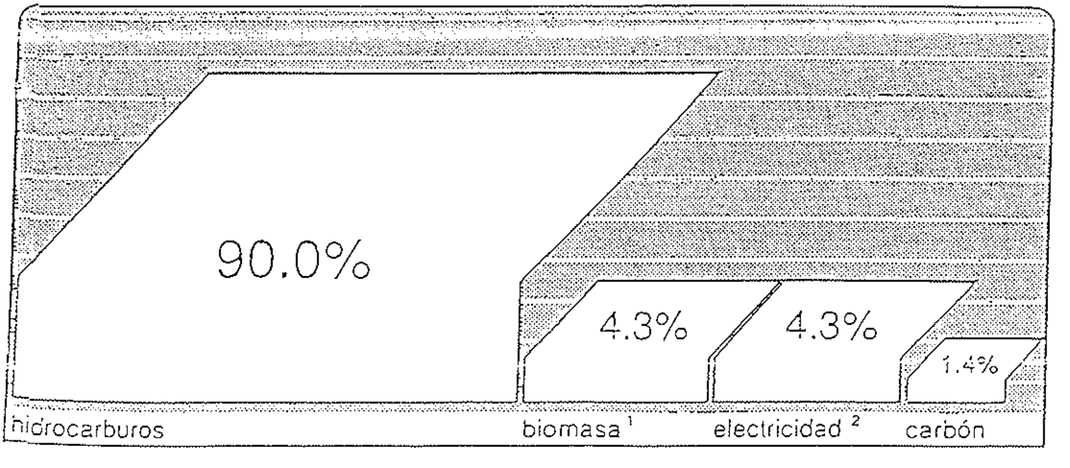
The standard of living in any country is closely correlated with the energy available per capita, which may be 440 kilograms of oil equivalent (KOE) in India or China and some parts of Africa. It is close to 1300 KOE in Mexico but still far from the 7000 KOE or more in the United States or some European countries as indicated in the "World Development Report 1992".

We need more energy, but not necessarily in the form we are getting it now which, in particular, has made Mexico City one of the most polluted in the world, mainly from internal combustion engines in cars, buses and trucks.

Cuadro 1: *producción de energía primaria.*

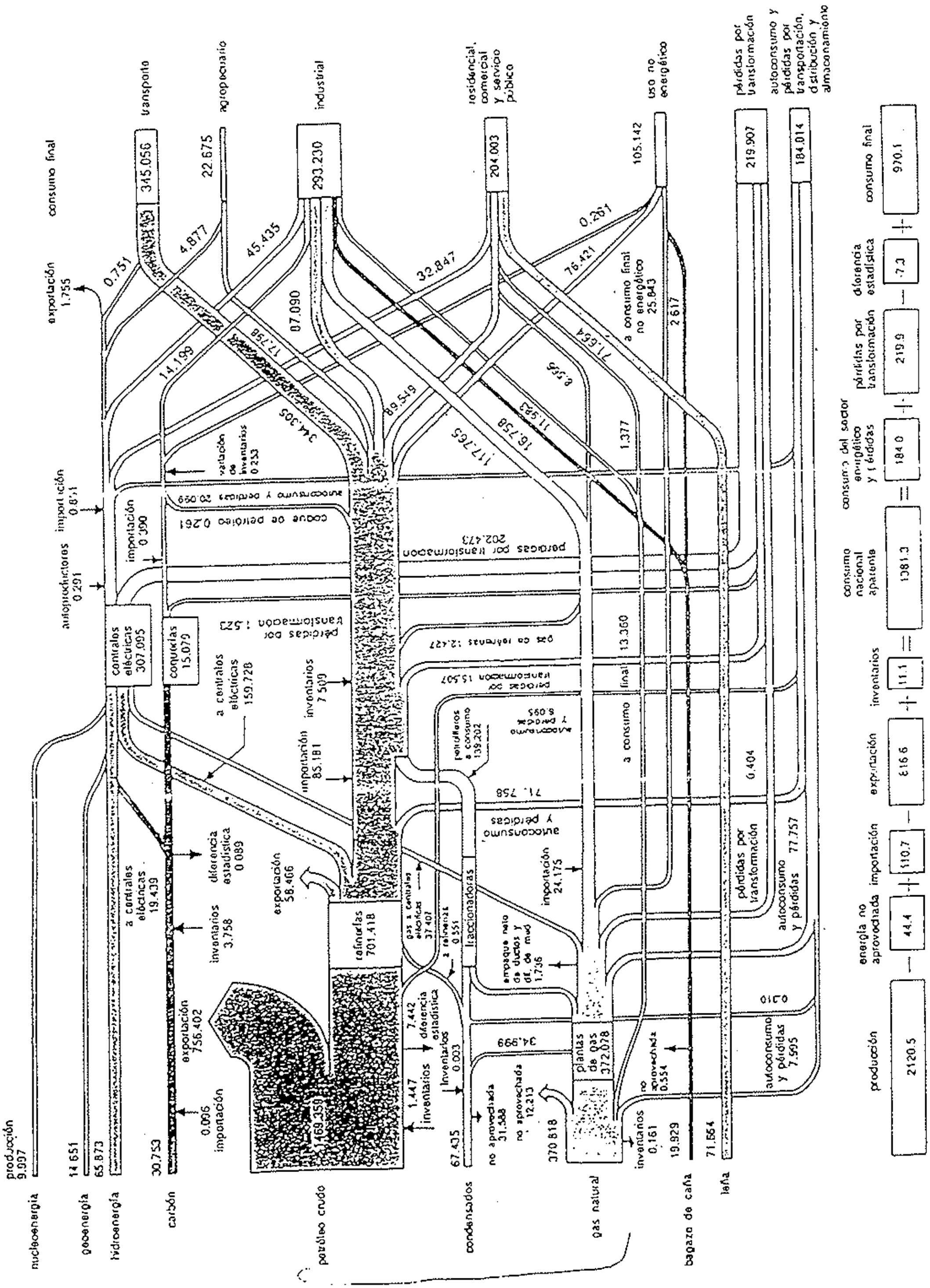
	1991		1992		variación porcentual 1992-1991
	petacalorías	%	petacalorías	%	
total	2118.830	100.0	2120.479	100.0	0.1
carbón	32.363	1.5	30.753	1.4	-5.0
hidrocarburos	1913.835	90.3	1907.612	90.0	-0.3
petróleo crudo	1471.940	69.5	1469.359	69.3	-0.2
condensados	64.609	3.0	67.435	3.2	4.4
gas no asociado	61.888	2.9	58.533	2.8	-5.4
gas asociado	315.398	14.9	312.285	14.7	-1.0
electricidad	80.450	3.8	90.521	4.3	12.5
nucleoenergía	10.969	0.5	9.997	0.5	-8.9
geoenergía	13.898	0.7	14.651	0.7	5.4
hidroenergía	55.583	2.6	65.873	3.1	18.5
biomasa	92.182	4.4	91.593	4.3	-0.6
bagazo de caña	20.939	1.0	19.929	0.9	-4.8
leña	71.243	3.4	71.664	3.4	0.6

Lámina 1: *producción total de energía* (2120.479 petacalorías).



[1] Incluye leña (3.4%) y bagazo de caña (0.9%).
[2] Incluye hidroenergía (3.1%), geoenergía (0.7%) y nucleoenergía (0.5%) evaluados en su equivalente de energía.

Diagrama 2: *Balance nacional de energía, 1992* (petacalorías).



Cuadro 4: *Oferta interna bruta de energía primaria.*

	1991		1992		variación porcentual 1992-1991
	petacalorías	%	petacalorías	%	
total	1326.138	100.0	1324.865	100.0	– 0.1
carbón	34.353	2.6	34.607	2.6	0.7
hidrocarburos	1119.294	84.4	1108.698	83.7	– 0.9
petróleo crudo	716.597	54.0	714.404	53.9	– 0.3
condensados	36.485	2.8	35.850	2.7	– 1.7
gas no asociado	61.595	4.6	58.161	4.4	– 5.6
gas asociado	304.617	23.0	300.283	22.7	– 1.4
electricidad	80.450	6.1	90.521	6.8	12.5
nucleoenergía	10.969	0.8	9.997	0.7	– 8.9
geoenergía	13.898	1.1	14.651	1.1	5.4
hidroenergía	55.583	4.2	65.873	5.0	18.5
biomasa	92.041	6.9	91.039	6.9	– 1.1
bagazo de caña	20.798	1.6	19.375	1.5	– 6.8
leña	71.243	5.3	71.664	5.4	0.6

The main effort made in Mexico has been to get less pollution from traditional energy sources, like improving the quality of fuels and moving out of Mexico City, the biggest refinery, and forcing car owners not to use thier vehicles one day a week, or to use vehicles that utilize gas rather than gasoline. The ideal would be an electric vehicle, and at the same time to situate all the thermal electricity plants far from Mexico City, but this obviously is a global problem, which would require years of research in developed as well as developing countries, and if it is realized at all, it will become feasible well into the twenty-first century.

But then another question crops up and that is for how long can we rely on oil and gas, or even on coal, to provide for our energy needs. All these products were produced from plants that grew because of the presence of the sun, millions, maybe billions of years ago. If the sun could produce such condensed forms of energy in the past, why could we not use it in the present, as astronomy tells us that the sun is likely to be there for billions of years more?

Well, some form of use of the sun has been made by mankind since ancient times. Agriculture, for example, would be impossible without the sun, which even provides the water from evaporation of the seas and its later condensation as rain. It is even possible now to grow biomass, not for direct human consumption, but to transform it into a fuel that can be used in cars, as our Brazilian colleague, Dr. Dobereiner, will tell us later. But the real challenge, besides agriculture, is not to use the sun to heat houses or water, but directly obtain from it electricity in a concentrated form, or fuels for all forms of transportation to which we have gotten accustomed.

In principle this is already possible through the use of photovoltaic cells, but at present it is not yet a competitive source of electricity and besides it is far from providing power in the enormous quantities in which present-day society needs it.

When I was a graduate student in Princeton, more than forty years ago, the dream was that before needing to ask the sun directly for its energy, we could use heavy elements, not produced in our sun, but possibly originally available in the galaxy through star supernova explosions, to get nuclear energy. That dream has in part been realized, for example, in France which gets 5% of its electricity from nuclear sources. However, the accidents in nuclear reactors, Chernobyl in particular, have made people wary of this source of energy, particularly as the radioisotopes that are produced in the process are very dangerous and we do not know where to deposit them safely. Nevertheless, nuclear energy is an option, and it leads to amusing situations. In the same way that cocaine barons launder money, there is also the laundering of energy. Italy, for example, who has stopped producing nuclear energy, refuses to buy it from France that does use this process, but accepts it from Switzerland, which in turn buys it from France. Incidentally, in Mexico there is one nuclear reactor that has been in operation for several years of 650 megawatts that produces about 4% of our electricity needs. Another one, with the same characteristics, is now in its final testing runs. Both are in Laguna Verde on the coast of Veracruz State, facing the Gulf of Mexico.

Another dream of my youth was to obtain energy from fusion instead of fission with, in principle, less undesirable consequences, but fifty years later this is still a dream.

What is clear from all the above exposition is that before we have alternative sources of energy, we have to do a lot of research and development. The advanced countries have the best resources to do this type of work and my advice in Mexico would be to create what is known, in espionage language, as "moles" in the institutions of the more advanced countries to participate in this type of research. This implies training students to the highest level possible in our countries in mathematics, physics, chemistry, biology, etc., and sending them openly to work in this type of institution for many years, with periodic visits to their home countries to transmit to others the information they have received. Once a critical mass in a given field related to alternative energy sources has been achieved in one of the developing countries, an appropriate institution should be created in order to have these people cooperate in the development at the same level as in the advanced countries.

Once I have given you the information about the situation in Mexico and my thoughts about how one should go about participating in the search of alternative energy sources, I would like to take the last minutes of this talk to mention what is actually being done in Mexico in this field.

First, we will not move into the biomass business, because in most of the country we have too many mountains and too little water; therefore, the fertile, arable land should be kept for agricultural products for human or animal consumption, with the latter to be turned later also into food.

We are greatly concerned with the problem of more efficient and less polluting use of conventional energy sources and in particular of oil and gas. The Instituto Mexicano del Petroleo, and Pemex, the national oil company, are working on the problem, both in the larger cities and in the regions where oil and gas are produced. They are also being aided by transnational car manufacturers that are designing more efficient engines and catalytic converters to lower the pollution.

Nowadays about 90% of the Mexican people are being served by the national electricity network and are well connected through pipelines and highways to receive the oil products, particularly gasoline, and gas for cooking and heating, that they require. This leaves out 10% of the population: about 9 million who live in nearly 80,000 communities of less than 2,500 each. In fact, 1.7 million live in communities of 100 people or less, and all of them are not only outside the electric grid work but they are also difficult to reach through roads on which only vehicles with four-wheel power can travel. These would be the segment of the population that could benefit most from alternative energy sources. It turns out, however, that several problems stand in the way of the process of bringing power to rural areas, and they are not so different from problems related with other services such as education, medical care, safe drinking water etc.

The first requirement in bringing alternative energy sources to these areas is that the population should really want them and be able to train some of its members on how to maintain them. Wind power, for example, is abundant in some of the mountain or coastal areas of Mexico, but installing the equipment and batteries required to store it will prove useless unless some of the local inhabitants are willing and able to sustain it. Solar energy using photovoltaic cells has even stronger requirements for sustainability, and this applies also to small hydroelectric projects. Incidentally, Mexico makes use, in a simple way, of enormous amounts of solar energy in Guerrero Negro, Baja California. This place, on the Pacific side of the peninsula, has shallow lagoons which, when their contact with the ocean is cut out, dry up, with the salt collecting at the bottom. Thus Mexico is one of the biggest salt exporters in the world and the energy involved, courtesy of the sun, is of the same order as the total hydroelectric energy used in the country. Ocean energy is not being considered in Mexico at present although in the north end of the Gulf of California, in a place called Punta Pañasco, there are some of the highest tides in the world.

What has been successfully pursued is geothermal energy, available in several parts of the country and particularly in the northwest. This source has become so important that sizable towns have been created near the place where it has become available, and later the energy, transformed into electricity, has been connected to the national grid.

A basic aspect of energy sources, renewable or not, in these isolated communities is financing. Usually the people living in them are the poorest of a developing country, and thus providing the service is a subsidy not an investment. A programm inaugurated in Mexico called "National Program of Solidarity", PRONASOL in Spanish, has precisely the objective of helping the poorest people of the country to help themselves. In any of the PRONASOL projects, it provides 50% of the cost, the rest being contributed by the state government, and by labor from the community itself. Through these processes, wind and solar projects have been installed in 13,000 individual houses mainly for lighting and the use of small appliances and communications and, on a large scale, for several small communities, in which small hydro is also used.

For the larger towns in Mexico and in particular, in Mexico City, the best for an alternative source of energy is to employ in some way the enormous amount of trash that these cities produce, which now causes a big storage problem. If in some way the immense trash heaps could become a source of methane that later could be used in household or industrial applications, or if the inhabitants of large, Mexican cities could be disciplined, as in Germany, to make separate heaps for metal, glass and all other materials, with the latter to be converted into combustible pellets and burned far from the cities, one could achieve a dual purpose: on the one hand, to get rid of much of the trash and, on the other, to obtain a substantial amount of energy.

What is clear from the previous remarks is that except for geothermal energy, none of what is called alternative energy sources in international publications will contribute substantially to Mexico's needs in the next twenty and maybe even fifty years. This does not diminish the need for research in this field, because when the change to these alternative sources comes about, either because the reserves of oil and gas become quite low, or for environmental reasons, the replacement is likely to be quite abrupt. An example is the case of computers. I was around as a student in Princeton when John von Neumann, possibly the greatest mathematician of this century, built, with the help of outstanding scientists and engineers, one of the first electronic computers which occupied a small building. For several decades afterwards, while computers were developing, one could see them only in selected places where scientific or business needs required them. Today any secretary or airline attendant has at her disposal a computer probably more powerful than the one von Neumann had, though they use them only for routine matters. But this has happened in the last fifteen or twenty years while von Neumann did his main work in this field fifty years ago.

A similar situation is likely to happen in connection with alternative energy sources. After a period in which they are likely to be dormant, but in which research on them will continue, they may bloom suddenly, and the countries that do not use them may be at a severe disadvantage, in the even more competitive and interrelated world of the future.

I would like to end by stressing that the main alternative source of energy is the human brain, not because you can run a car with the heat of a thousand or million heads, but because it has been the source of all the knowledge we have. Thus one of the main problems of Third World countries should be the development of the brain power of the population, through effective education from infancy up to graduate school, and then to its constant renewal throughout the whole lifetime of the individual

REFERENCES

Bauer, M. and Quintanilla, J., Development and Energy Efficiency: An Alternative Scenario, *Proceedings of the International Association of Energy Economics 15th Annual North American Conference*, Seattle, Washington 1993.

Martínez, A.M., Huacuz, J.M., Bauer, M., Renewable Energy Systems in Mexico, *Proceedings of the U.N. Castelgandolfo II on Renewable Energy for Environment and Development, 1993.*

Bauer, M., Energías no Convencionales, (Unpublished).

Balance Nacional de Energía, Secretaría de Energía, Minas e Industria Paraestatal, México 1992.

World Development Report 1992, Development and the Environment, *Published by the World Bank, Oxford University Press 1992.*

Renewable Energy Resources: Opportunities and Constraints 1990-2020, *World Energy Council Report 1993*, London 1993.

ALTERNATIVES FOR BIOFUEL PRODUCTION: THE BRAZILIAN EXAMPLE

JOHANNA DÖBEREINER *

Introduction

Alternative energy sources use solar energy instead of fossil fuels and return only the amount of CO_2 to the atmosphere which the crop has taken out to produce the biomass. Due to the climatic conditions, agriculture in tropical regions can catch much more sun energy than crops in temperate regions. In addition, tropical plants from the *Gramineae* family have developed a much more efficient photosynthetic pathway, the C_4 photosynthesis, which can capture more than twice the amount of sun energy than C_3 plants prevailing in temperate regions. It is therefore understandable that the major alternative biofuel, which one day will have to replace fossil fuels, will have to come principally from tropical regions. Brazil, for various and different reasons, has become the world leader in bioethanol production and in this paper we will give some explanations for the circumstances which have lead to this.

In a paper we presented to the Pontifical Academy in 1981, it was already suggested that ethanol produced from sugar cane seems to be the most promising alternative with yields about twice as high (more than 4000 $L.ha^{-1}$ yr) than the amount produced from other sugar producing crops or starch crops (Dobereiner *et al.*, 1981). The only other biofuel alternative suggested at the time for the replacement of Diesel oil was the African oil palm (*Elaeis guineensis*) which yields 18 to 38 tons per year of oil corresponding to 76% higher energy yields than sugar cane ethanol (Serra *et al.*, 1979).

This crop is widely grown in various regions of the Amazon basin

* Pontifical Academician, EMBRAPA CNPAB, Rio de Janeiro, Brazil.

and has the advantage of being able to grow within the forest or at least in mixed culture with other tree crops. Therefore, it contributes much less to the erosion of Amazon soils than crops which need cleaning or plowing of the soil every year. It has been estimated that 6 million ha of Amazonian forest area representing a little more than one percent of the total area, or 15 to 18% of the area already deforested, would be enough to produce sufficient oil to replace all Diesel oil (460,000 barrels per day) in Brazil (Boddey, 1993). Unfortunately, this biofuel alternative is not in use yet, partly because there is no financial support as of yet for the necessary research to put it into practice. Apart from relatively high planting costs, the remaining costs are relatively low, as harvesting is done mainly by hand and using mules. Processing fuel is supplied by burning the fibrous residues. Oil palms can be harvested continuously throughout the year and the fuel is obtained by simple pressing, leaving much less effluents to dispose of than other biofuel processes.

The main reason for the little use of this biofuel in the world may be the high price of edible oils, which are about three times the price of diesel oil today. Thailand, for example, which has large plantations of this crop, uses it as one of its major export crops as edible oil to Europe.

The Brazilian Ethanol Programme

This programme started as a strategic programme in the early seventies as a response to the oil crisis of 1973, but it only came to full effect in 1980 with the second oil crisis. By 1985, 90% of all new motor cars, totaling 4 million, were running on 95% ethanol from sugar cane. This was due to the development of specific modified motors which are each time becoming more efficient and are now using little more of this alcohol than cars running on gasoline. More than one million jobs were created in the interior of the country and 200,000 barrels a day of petrol were economised by the country. Unfortunately, during the second half of the eighties, Brazil followed the international tendencies succumbing to the transitionally low petrol prices and deaccelerated the programme. Unlike most other countries, Brazil has land soils and climate which permit further expansion of the sugar cane crop without any detriment to food production as only 8% of the areas used for agriculture so far are planted with sugar cane.

The total production of ethanol in Brazil could easily be increased from the current 11 billion litres per year to 18 billion with minimal additional investments, simply by optimizing industrial installations and by increasing yields.

Recent research in Brazil has shown that the sugar cane varieties in use in the country today have developed symbiotic associations with endophytic nitrogen fixing bacteria, which can furnish all nitrogen necessary for high yields (Urquiaga *et al.*, 1992; Dobereier *et al.*, 1993; James *et al.*, 1994). These bacteria colonize in high numbers (up to 10^6 cells.g^{-1}) the whole stem and also leaves of the Brazilian sugar cane genotypes where more than ample sugar and malate are available to furnish the ATP necessary for the bacteria to convert the air nitrogen into forms the plant can use.

This is due to the selection during the last century of varieties or genotypes of this crop, which produce high yields with very low N fertilizer levels, much below the amounts accumulated by the crop. In addition, the elimination of burning the cane leaves before harvest has been shown to increase yields and to conserve soil fertility (Oliveira *et al.*, 1994). Due to these results we are now recomending to the farmers to eliminate completely all nitrogen fertilizers and to stop burning but to use the economy of more than 70% for irrigation and, where necessary, for increased phosphate and potassium fertilization. This will increase yields by 30% without any additional costs.

The cost of ethanol produced in Brazil currently is around US$ 0.25 and its retail price US$ 0,53, while the gasohol retail price is US$ 0.62. This difference is created by the government in order to compensate for the 20% higher fuel consumption by ethanol cars. All fuel stations in the country now sell both ethanol and gasohol which contains 22% absolute ethanol, while pure gasoline is not available any more in Brazil. There are now more than 4 million motor cars running on 95% ethanol, about half of the total fleet and the government, now realizing the ecological advantages of the Proalcool, has resumed its support for the programme, even with the present, very low petroleum prices.

The Energy Balance

The main reason for the success of the Brazilian Proalcool Programme, while most other countries in the world are unable to introduce viable biofuels, is the energy balance. As stated before, the Brazilian sugar cane genotypes have been grown (cultivated) or selected for long with low fertilizer levels. This is due to the lack of subsidies of the nitrogen fertilizers and consequent high prices. This way, without knowing until recently, sugar cane as also other crops in Brazil, were cultivated for high yields with low fertilizer levels, much in contrast to the green revolution. Also due to the relatively low use of mechanization in this crop production, ethanol produced today in Brazil uses about three times less energy than is produced by sugar cane. With the complete elimination of nitrogen fertilizers and the elimination of cane burning before harvest, the energy balance will go up to 4 or 5, that means five times more energy can be produced in Brazil than is used to produce the ethanol. The use of sugar cane bagasse to cover all factory power for ethanol production further can increase the energy balance to 5.8 (Boddey, 1993). This contrasts with the various ethanol or oil producing attempts in Europe or in the US which use so much fertilizer and mechanization for growing crops, that the energy balance is practically one: that means as much energy is needed to produce the biofuel as is produced. This makes any biofuel programme senseless. The decisive key to low fertilizer use is only now being identified as being due mainly to plant breeding, which in the North has never been done and would take many years to reach the advantage Brazil reached without knowing it for so many years.

Ecological Impacts of the Proalcool Programme

In addition to the strategic reasons which created the Proalcool, it is now becoming clearer and clearer that the Programme has great impact on the environment. The use of ethanol cars in Brazil reduced by 75% the lead content of the air in the large cities and by 57% the carbon monoxide. Such cars produce 64% less hydrocarbons and 13% less NOx when compared with cars using gasoline. In addition ethanol is the only vehicle fuel used on a large scale which only recycles CO_2 obtained from the atmosphere by photosynthesis and therefore will reduce the greenhouse effect. The Brazilian technology of producing alcohol from sugar cane grown without any nitrogen fertilizer will not only contribute to reducing ethanol production prices and increasing the energy balance, it will also avoid contamination of ground waters with nitrates.

With the considerations discussed in this paper, the Brazilian Proalcool Programme must be considered a great environmental and social programme, which only recently has been recognised as such.

REFERENCES

Bodey, R.M. 1993, «Green energy from sugar cane» *Chemistry and Industry* 17 (May) 355-358.

Döbereiner, J., Araujo-Neto, J.S. and Arkcoll, D.B. 1981, «Energy alternatives from Agriculture» *Pontificiae Academiae Scientiarum Scripta Varia 46*, 432-458.

Döbereiner, J., Reis, V.M., Paula, M.A. and Livares, F.L. 1993, «Endophytic diazotrophs in sugar cane, cereals and tuber plants». In: Palacios, R., Mora, J. and Newton, W.E. eds. *New horizons in nitrogen fixation*. Doderecht: Kluwer Academic Publishers, pp. 671-676.

James, E.K., Reis, V.M., Olivares, F.L., Baldani, J.I. and Döbereiner, J., 1994, «Infection of sugar cane by the nitrogen-fixing bacterium *Acetobacter diazotrophicus*». *Journal of Experimental Botany* 45, 757-766.

Oliveira, O.C., Urquiaga, S., and Boddey, R.M., 1994, «Burning cane: the long term effects». *International Sugar Journal* 96, 272-275.

Serra, G.E., Goldemberg, J., Carvalho, C.M. and Moreira, J.M., 1979, «The energetics of alternative biomass sources for ethanol production in Brazil» III *International Symposium on Alcohol Fuels Tecnology*, California.

Urquiaga, S., Cruz, K.H.S., and Boddey, R.M., 1992, «Contribution of nitrogen fixation to sugar cane: Nitrogen-15 and nitrogen balance estimates». *Soil Science Society of America Journal*, 56, 105-114.

ALTERNATIVE ENERGY STRATEGIES FOR SUSTAINABLE DEVELOPMENT OF INDIA

C.N.R. RAO *

I am not a specialist in any aspect of energy, except that I have been involved in science policy in India for some time. I had also the opportunity to be associated with national planning. In this capacity, I studied the energy options for India in some detail. In this presentation, I shall attempt to show how complex the situation is in a developing country. There is not a single source of energy which can answer the demands of a growing developing country with a large population like India.

First of all, we have to remember what the energy consumption is like in a poor country like India. The per capita energy consumption in India is very small indeed. It is almost negligible compared to that of the advanced countries. If India therefore has to become a prosperous modern country, it has to look into energy demands in a proper way. It is known that the per capita GNP varies almost linearly with the per capita electricity consumption. You can see that India is very low compared to Japan or the U.S. One can say that there has been a slow increase in energy consumption in the last few decades, but this is far from staisfactory. How do we go from this lowly position to something reasonable where the energy requirements of its vast population are met adequately? The same kind of linear plot is obtained if one plots life expectancy against per capita electricity production. India's life expectancy is increasing because of modern medicine. Here again when compared to Japan and other advanced countries, we are below par. The problem of a developing country like India is that we have a large population with large requirements and tremendous aspirations.

The compound growth of coal and electricity consumption in India is

* Pontifical Academician, Jawaharlal Nehru Centre for Advanced Scientific Research Bangalore, India.

interesting. The increase was 3% in coal, 7% in oil, 10% in electricity in 1987. This was over seven years ago. The installed capacity for power generation was supposed to be about 65,000 megawatts. For the last five years, according to the Five Year Plan of India, we planned for about 38,000 megawatts (MW), of which 7,000 were hydro, 30,000 thermal and 700 nuclear, but we still have a big power shortage, although the potential itself is different. The potential for hydro power is quite large. There is the possibility for mini-hydro projects as well (say around 5,000 MW). What about the installed capacity? The situation is different and well below our requirements. Can we go on increasing the thermal plants that we are planning, using gas, diesel and coal? I am not happy with this, partly because of the pollution problems and partly because I believe that we should not indiscriminately burn petroleum crude and its products for power. They are needed for posterity to produce pharmaceuticals, fertilizers etc.

We have a lot of coal. I think that the coal available in India is around 170 billion tons. The annual production is about 250-270 million tons. But we cannot burn coal the way we are doing without causing tremendous pollution. Already Indian air is not good to breathe, and we have to worry about these thermal plants coming up in larger number, if they use coal. One solution is to improve coal technologies, improve gasification and hydrogenation of coal (liquefaction of coal).

Nuclear and coal technologies can have comparable costs. If one produces power near the pit head it costs something, and if far away from the pit head, it costs very much more, depending on the distance and so on. Unfortunately, India does not have enough money to invest in nuclear energy. I remember that Mr. Rajiv Gandhi was planning to install additional nuclear power plants for 10,000 MW. Then the target became 6000 MW and later 3000 MW. We may now actually succeed in getting 1000 MW or less of nuclear power. Nuclear energy, when it will be there, will not be a major component of our energy. We have to look at other possibilities.

People who say newer energy sources are all for the books should not think so. We have to seriously consider solar thermal plants and other possibilities. What we can have is a kind of mixed energy economy which would include in addition to hydropower, nuclear power and thermal power, the whole biomass, wind farms and small hydroelectric stations. Solar photovoltaics would be useful for remote villages, but the total power from this source will be small (say a few hundred MW). We have to look into all these possibilities, even though they may not serve all the people in all places.

A typical solar thermal electric power station appears to involve an installed cost of 2,000 $ per kilowatt; and the generation cost is as low as that of any other energy. I think that it is from 1.5 to 2 rupees per kilowatt. The potential of windmill power in India is a few thousand MW. Our capacity for tidal power is also not negligible. In principle, we can think of tidal power in at least two or three locations and wind power in three or four areas in India, in addition to some nuclear energy. This will still not answer India's needs, because the population is so large.

Solar photovoltaics are small capacity systems. It is useful to create a one-kilowatt, five-kilowatt (maybe one 500 KW) stations in isolated regions far away from the grid system, so that we can help the local farmer to pump water, have a television set and some street and home lighting. Some years ago, we planned for 100 MW of solar photovoltaic power in a five year period. We could only reach a 5 MW target. I think that India presently has a total of 25 MW of solar photovoltaic power at best. I feel that we should go for about 1000 MW in the next five years. Although the technology is not good enough to make it cheap at present, it may become cheap soon through amorphous silicon solar cells.

The biomass problem has to be looked into in two different ways. I want to give the optimistic view which the biomass enthusiasts talk about. Although the total land area available in India is about 320 million hectares, the degraded forest and various kinds of wasteland is about 137 million hectares. The wasteland can be used for biomass production. This is a tremendous responsibility. How do you make a reasonable calculation? First, we have to get used to large numbers. Our rural population receives no energy from the government. (All the energy we produce is for the city people who constitute a small percentage of the total population). There are 627 million rural population, and they are the ones for whom we have to provide the energy. If we take 120 units per person in the villages (which is very little), we require 120 multiplied by 627 million units of electric power. In principle, we can produce the required biomass using about 20 million hectares out of the 137 million hectares available. Biomass option should therefore not be thrown out of the window. As far as the costs are concerned, they are comparable to coal (in fact, it can be cheaper than that of coal). Unfortunately, nothing much has happened. The installation costs for biomass-based energy units are lower than solar photovoltaic, microhydro, wind power and so on.

The Johannes-Reddy study shows that renewable energy can indeed provide a lasting solution. In principle, we can have a different world. We can have a world by the middle of the 21st century where the biomass and the renewables can take care of more than 13% of the total electricity production. Such a world is possible, because we cannot go on burning oil. An Academy of this kind, where morality and ethics are important issues of concern, should be concerned about this matter. I think that we as a civilization cannot burn oil and gas that God Almighty has given us, because they are required for the future for the production of pharmaceuticals, drugs and fertilizers. They are required for our great grand-children and their great grand-children. I think that this is a very important responsibility. I advise against burning gas indiscriminately, however much gas we may have in the world. Countries like Brazil and others in Latin America and Africa can become exporting countries for energy. They could export alcohol in very big quantities if they have suitable agriculture policy. Today, the West is importing a lot of gasoline. Instead, they could be importing alcohol, if a new world order based energy requirements on agriculture and biomass.

In an Academy like this we should also look at the scientific part of the energy problem. I am glad that Johanna talked about some aspects of the science. I would like to say one or two things that we have to do in the scientific area. What are those research and development priorities that are emerging from renewable alternative energy sources in the intensive global scenario? We have to do many things. We have to demonstrate for example, the efficacy of a new generation of turbines, of biomass energy and raising of tailored crops in different regions. We have to discover a good fuel cell for methanol which uses a good proton conductor. There is need for advances in storing hydrogen and biomass-integrated gasifier fuel systems. We have to worry about how to convert inorganic carbon into organic carbon, a problem of relevance by the year 2050. I think that this is a nice area of research since by 2050, we may be really running short of organic carbon.

In conclusion, what energy scenario should we plan for a country like India for the next five to ten years? I suggest that by the year 2010 we should add another 20,000 MW of hydroelectric power and should go for additional 2000 MW of nuclear. We should add 7000 MW of wind and ocean based energy. I strongly feel that to these we should add 20,000 MW of biomass based energy. I would go slow on thermal power (maybe 10,000 MW in the next 5-10 years), but I would certainly suggest 5000 MW of solar thermal and 1000 MW of solar photovoltaic.

I trust that I have been able to show how difficult it is to plan for a vast country which is not only energy poor, but is also economically poor, though we may have the scientific capability. It is not only the question of money, but also the difficulty of making the right choices. I feel that we cannot have a simple solution based on one option, but one based on a mixture of technologies.

AFRICAN ENERGY PROBLEMS, OPTIMAL ENERGY POLICY AND THE PERTINENCE OF THE CONCEPT OF ALTERNATIVE ENERGIES

FÉLIX MALU WA KALENGA *

I. Introduction

The last decade of the twentieth century is overshadowed hy the efficiency with which man, using very often the powerful tools provided by science and technology, is destroying the bio-physical environment on which resources all living species depend. Rapid population growth, increased mohility of people, large-scale urbanization, massive industrialization and the large-scale use of chemicals in food production are producing, among other things, climatic changes with life-threatening effects, a mere 20 to 40 years hence, according to various simulation results. Finding adequate responses to this threat necessitates individual as well as societal awareness, and full recognition of the fact that a collective will and effort of all countries are needed.

This is in essence the view taken by the "World Commission on Environment and Development (WCED)" in its report called "Our Common Future" [1]. The report, also known as the Brundtland report, stresses the concept of "sustainable development" defined as a "development that meets the needs of the present without compromising the ability of future generations to meet their own needs". The above definition contains two key concepts. The first one is the concept of "need", particularly "basic need". The second implicit concept is the limitation imposed on the environment's ability to meet present and future needs by the state of technology and of the socio-political organization.

Sustainable development cannot be attained without providing all men

---

* Pontifical Academician. Commissariat Général à l'Energie Atomique, Zaïre.

with such essential things as food, shelter, transportation, education and health service. To cater to all these "basic needs" requires, as do all human activities, energy in the right form. This entails, more often than not, burning fossil fuels that contribute to the "greenhouse effect". This effect is accentuated by the large-scale destruction and burning of rainforests in the tropics for agricultural and energy purposes. The threat of the "greenhouse effect" is the most pressing problem one has to deal with when devising energy policy to-day. The case of the African continent is a good example.

II. Problems and Constraints of the African Energy Sector

At an equivalent of 0.1 ton of oil, Africa has the lowest per capita consumption of energy in the world. The African low consumption comes with many disturbing phenomena. Indeed the most important primary energy resources used are fuelwood and derivatives (70% of the commercial and non-commercial energy consumption). Population growth strains fuelwood supplies especially when forests are displaced by farmlands. The destruction and burning of forests contribute to the greenhouse effect and desertification. As fuelwood becomes more expensive it is replaced by oil. The increasing cost of oil over the years, since the first oil shock of 1973, further strains traditional ligneous fuel supplies. Non-oil producing African countries, particularly Sub-Saharan ones, are thus confronted with a vicious circle of steady environmental degradation and a reduced standard of living: deforestation and heightened burning of wood and waste deplete the soil of nutrients and increase soil erosion, reducing in the process the African food and energy production capacities; increased food and energy imports that result in straining financial resources; reduced purchasing power leads to a shortage of commercial energy and thus increases the demand for non-commercial traditional ligneous fuels; the rate of consumption of ligneous fuel becomes greater than the rate of natural growth leading to a serious shortage of fuelwood and derivatives. To break the vicious circle necessitates shifting wood resources from a noncommercial status to a commercial one by way of energy plantation. Indeed, at the present rate of population growth alone, the demand for this type of fuel in 2020 is estimated at three times the 1990 level. Energy plantation must be carried out in such a way as not to induce another vicious circle, this one concerning food production.

Notwithstanding the increasing comsumption of ligneous fuel, the demand for oil is bound to rise with any rise in the standard of living, since traditional fuel is not suited to power production, pumping and transport. Oil accounts for 65% of the total consumption of commercial energy in

Africa and for 50% of the total commercial energy expenditure. It places a significant burden on the balance of payments of the African non-oil producing countries.

The future outlook of the energy sector of African countries, particularly Sub-Saharan ones, is none too promising. The per capita consumption of commercial energy has increased at the rate of only 0,9% p.a. between 1980 and 1990, while the rate of population growth stands at 3% over the years. The shortfall has been compensated by a rapid increase in the rate of consumption of non-commercial forms of energy as four-fifths of the population use ligneous fuel wholly or in part to meet their energetic needs.

On the financial front, the total investment in the energy sector in Africa was of the order of $2 billions in 1990. Based on the historical 5% per year increase in the production of energy in the world, Africa must invest at least $4.7 billions in the year 2000 (in 1989 dollars) just to keep even standard-wise, taking into account the galloping demography.

A precise assessment of African energetic prospects is difficult to make because of the lack of reliable information on end-use energy, and on government and private plans for housing, transportation and industrial plants. What can be said is that a realistic, long-term forecast for the African energy sector must by necessity involve the combination of many factors, technological, human, socio-economical. Assuming that the offer of energy will increase at a rate of 5 to 6% up to the year 2020, the part of oil in the total offer of commercial energy in Sub-Saharan countries will rise to 70%. The part of natural gas will increase to 15%, while the part of coal will increase to 5%. At 50%, ligneous fuel will remain the most important source of energy in rural as well as urban areas.

Positive responses to African energy problems may take a variety of forms, which we shall now discuss in a more general context.

III. Terms of Reference of an Optimal Energy Policy

The aim of an energy policy is to determine the way the supply and the demand of energy can be matched at the lowest marginal or opportunity cost, taking into account all the variables that affect the energy equation: that is, the technology available, the capital resources that can be mobilized, the sources of primary energy that can be put into contribution, the socio-political and socio-economical conditions at the moment.

To be optimal an energy policy must conform to the call for a development that is sustainable. This means that one must be able to provide in a sustainable way sufficient energy that conforms to the criterion of minimizing various constraints and risks, particularly the "greenhouse effect".

The problem is to know if such an optimal energy policy is technically possible, pragmatically achievable and economically sound, taking into account the freely established socio-economic goals of advanced and Third World countries alike. To express it in other words, is it possible to devise a sustainable energy policy supporting the sustainable development of all the countries of the world without tempering with the freely established socio-economic goals of the rich and the poor countries alike?

The first answer to this question is to emphasize the conservation of energy. The safest, cheapest, surest way to match the supply and the demand of energy, while minimizing various risks, is through greater energy use efficiency. One unit of energy saved is equivalent to one unit of energy produced. Energy efficiency improvement can be obtained by the right combination of technology, market incentives, and changes in lifestyle.

The systematic reduction of the energy intensity of a unit of GNP is thus the first requisite of a sound energy policy. The outcome of such an energy policy is a low energy path that minimizes expenditure of resources, including environmental ones.

But the decrease of the energy intensity of economical activities will not suffice in itself to support sustainable development. Indeed, if the traditional 2 to 4% p.a. rate of growth of economical activities in the advanced countries is maintained, their GNP will double every twenty years. To sustain the same level of energy production requires doubling every twenty years the energy efficiency of economical activivies. By the year 2050 the energy use efficiency should be at least eight times the present level, and must continue to double each twenty years thereafter. Obviously, this is impossible to sustain indefinitely. In fact, the most that can be achieved is probably a four times increases in the energy efficiency by the middle of the next century [3].

If improvement in energy efficiency is the first requisite of a sound energy policy, whose outcome is a low energy path, it is not sufficient, obviously, to support sustainable development at the traditional rate of growth of the economy in the developed world. One reaches a similar conclusion if one considers the rate of growth of the population in the Third World. A significant increase in the supply of energy is unavoidable. This must be done using preferably benign, renewable energy sources. The second term of reference of an optimal energy policy is, therefore, the aggressive and large-scale promotion of benign, renewable energy sources; that is soft solar energy, wind energy, hydro-power, biofuel, geothermal energy.

Renewable energy has its drawbacks. It is, more often than not, a decentralized and deconcentrated form of energy. Furthermore, according to some estimations, renewable energy sources, with solar energy used locally only (soft solar energy), can at best provide about 1 Terra-Watt per annum [4]. Renewable sources of energy, although useful, and bound to grow in importance in the future, cannot, accordingly, be considered in the mid term, as a replacement option for fossil fuels but as complementary sources of energy. This means that the consumption of fossil fuels is bound to grow, unless one increases the use of nuclear energy and of hard solar energy.

From the above discussion the fact emerges that it is impossible to find a low sustainable energy path that can support sustainable development while maintaining the same rate of economic growth in the developed countries, and the same rate of population growth in the developing countries. In other words, the provision of energy that is at the same time sufficient, safe, economical, benign to the environment, cannot be sustained at the present rate of growth of economic activities in the advanced countries, and the present rate of population growth in the Third World. To support this assertion it suffices to note that one will need to provide a minimum of 20 TWa in the year 2050 for a population of 10 billions if the 1985 economic activities and population growth rates are adopted. This amount of energy cannot be produced using only benign sources of energy.

The question then is to know how to meet in a sustainable manner, the future increase of the energy demand, estimated between 20 to 40 TWa in 2050, forecasted by various studies, such as those of the WCED, the "World Energy Conference (WEC)", the "International Energy Agency (IEA)", the "Organisation for Economic Co-operation and Development (OECD)", the "Commission of the European Community (CEC)", the "International Institute for Applied Systems Analysis (IIASA)", while minimizing the damages to the environment, such as global warming? The answer to this question constitutes the third term of reference of an optimal energy policy.

The fourth term of reference of an optimal energy policy derives from a normative assessment of the state and the impact of alternative energy technologies as measured against reference energy system. Three alternative technical scenarios can be considered taking into account available fuel resources and energy technologies [5]. They are :

1. Large-scale electrification based on both coal and nuclear fuel (large-scale electrification scenario);
2. Fossil-based systems stressing enhanced conventional oil and gas as well

as synthetic, with decreased emphasis on nuclear electric (fossil fuel based scenario);
3. Renewable energy systems stressing soft and hard solar energy, biofuel and cogeneration (renewable scenario).

In the large-scale electrification scenario, the demand for energy is satisfied by increasing the use of electricity in place of the direct use of fossil fuel. One can, for example, aim at electricity taking up 55% of total commercial energy demands. Advanced inherently safe nuclear power reactors, advanced coal-based electric generating technologies (combined cycle cogeneration, circulating fluidized bed reactor, advanced turbine such as steam injected gas turbine ...) are becoming available at affordable capital costs. This scenario corresponds to the largest primary energy consumer because of electric generation lost, which is, by the way, extensive in African electrical systems.

In the fossil fuel scenario, the demand for energy is satisfied mainly by using oil and gas. In the long term it relies on the substitution of coal and gas for oil, limited electrification and enhanced oil recovery.

In the renewable scenario, the demand for energy is satisfied using mainly renewable resources (solar, geothermal, hydro-power, biofuel, OTEC ...). Electricity is produced preferentially from decentralized generating facilities (low-head hydro-power, soft solar, cogeneration). This scenario enhances the penetration of renewable resource technologies, while maintaining historical fuel use patterns.

Any sound energy policy is by necessity a mix of the three technical scenarios described above. For example, one can call:

1. In the short term:

— For a larger promotion of soft solar energy;
— For a better management of biomass farm and tropical rainforests;
— For an aggressive inter-fossil fuels substitution in order to lower the emission of greenhouse gases (switching, as much as possible, from coal to lower CO_2 emitting fossil fuels, such as oil and natural gases);
— For the combustion and energy recovery of non-recyclable waste, refuse and biomass;
— For the implementation of advanced biomass conversion technologies (biomass based cogeneration, biotechnology based power plant ...).

2. In the medium term:

— For the use of hydrogen as a secondary carrier of energy, using production technologies applicable in the solar-hydrogen concept [6];
— When necessary, for a cleaner and more efficient use of coal in the production of electricity.

The combination of these measures can add, according to some estimations, about 4 to 5 TWa to the present 10 TWa produced [7]. This will suffice for the medium term, that is, up to the year 2020 or 2030, depending on the rates of economic and population growth considered. Beyond this, excluding a massive use of coal, one has to rely on hard solar energy, fission and fusion energy.

IV. The Merits and Drawbacks of the Nuclear Option

One cannot avoid considering the fission and fusion energy options when studying the question of the long-term sustainable energy supplies to support sustainable development.

As perceived by the general public, the nuclear option is beset by three main problems:

1. The safety of nuclear installations;
2. The safe handling and disposal of radio-active wastes;
3. The proliferation of nuclear weapons.

The most important and pressing issue concerns the safety of nuclear installations, as demonstrated by the Three Miles Island accident and the Chernobyl disaster. The safety of a nuclear installation depends on its design, setting, construction and operation. Concerning the design stage, the concept of "Inherently Safe Nuclear Reactor (ISR)" is now emphasized. The first aim of the ISR concept is to foster a new safety philosophy in power reactor technology by making the safety requirement the first and prevailing requisite of the designing process. This means that the function of any safety arrangement, be it passive or active, must be designed into safety margins and integrated, right at the beginning of the conceptual process, into the nuclear reactor. This must be done in such a way as to put the nuclear reactor in a position where it cannot experience a destructive accident, principally «Loss of Coolant Accident, (LOCA)", associated or not with a "Super-Prompt Critical Reactivity Accident (SPCRA)" [8].

The second aim of the ISR concept is to render less cumbersome the licencing, designing, building and operating process of a given type of nuclear power reactor. The third aim of the ISR concept is to counter LOCA and/or SPCRA by relying on inherent safety properties that manifest themselves through usually interacting nuclear, physical, chemical, mechanical, thermo-hydraulic properties, such as Doppler effect, density variation of core components and structural materials [9]. The emphasis on passive safety arrangements utilising inherent safety properties, solves the thorny issue of man-machine interactions in the operation of nuclear reactors.

Finally, it is appropriate to judge the drawbacks of the nuclear option against its merit. Barring a major nuclear accident, whose probability is small, the nuclear option is characterized by a minimal impact on the environment, and virtually no direct contribution to the greenhouse effect [10, 11].

V. The Global energy Village (GEVI)

The pervasive and critically important role of energy in support of sustainable development is made all the more obvious when one considers the villages of rural, Sub-Saharan countries. It is in this context that one can readily implement the call for the rational uses of ligneous fuel as a benign source of energy, and for the preservation of rainforests to alleviate, somewhat, the threat of global warming. The identification of energy issues at the village level is, therefore, of critical importance, although the village economy does not count for much in the national monetary economy.

To become sustainable, the development process at the village level must start by being appropriate first. Appropriate development at the village level requires that resources be used to make villagers more self-sufficient. Appropriate development thus adds little or nothing to GNP. Or better still it reduces it by making villagers more independent of the national economy [12].

To get Africa going again, it is advisable to make the villagers rise by their own bootstraps, out of the present subsistence condition and into a modern economy, by promoting first some kind of autarchical village economy at the beginning. Indeed, it is now largely obvious that the "indiscriminate market growth and trickle down" approach to Third World development has failed miserably in Sub-Saharan countries at least. Certainly, some improvements in life expectancy, infant mortality, literacy and GNP have been made over the four UN decades of development since the "Truman design" of 1949. But the benefits have been uneven. The poor, particularly in rural communities, have experienced virtually no improvement in their living conditions.

The fact that this dismal record has been achieved during the "glorious thirty years (1950-1980)" of the history of capitalism is a cause for additional concern. Indeed, in the recessed 1980s, one has direct cognizance of a turn for the worse. African debt has risen to unmanageable levels. Sub-Saharan eco-systems are being degraded at an alarming rate. Conventional economical development plans have registered negative growth. The trickle down mechanism of capitalistic development philosophy has not worked well, to say the least, in Sub-Saharan countries. Basic needs of the rural poor are no longer being provided at the required level. To change all this, one must involve the poor themselves as first actors of their own development. This means that things that add little to GNP must be produced first in the village, such as cheap tools, cheap housing, cheap clean water, cheap transportation [13].

For that purpose, the provision of cheap and sustainable energy supplies is of paramount importance. To provide sufficient, sustainable energy at a reasonable and affordable price to the villagers means solving two intertwined aspects of the energy crisis affecting African village economy. The first aspect is the oil crisis which still besets rural Africa. This aspect of the village energy crisis is rather well understood. The second aspect is the fuelwood and charcoal availability problem. This second aspect of the village energy crisis is not adequately studied because of the lack of reliable surveys of fuelwood and charcoal production and consumption [14].

To solve the village energy crisis in rural Africa, one must consider the supply and the demand side of the energy equation. As regards the supply side, an interesting solution is proposed by the "Global Energy Society" in line with the call for sustainable alternative development [15]. Basically, the idea is to promote well-known renewable energy technologies, such as biogas plant, biomass gasifier, soft solar thermal system, windmills, in such a way as to make the village community an "Integrated Self-Sufficient Rural Energy Centre" called the "Global Energy Village (GEVI)". The GEVI concept has evolved out of the need to go from energy dependence to energy independence, from outside to internal control, from non-renewable sources of primary energy to renewable ones. GEVI is set up using the right approach: that is, by involving the rural community at all stages of the project conception and execution, and by emphasizing the utilisation of locally available resources [16].

VI. The Pertinence of the Concept of "Alternative Energy"

By stressing the use of renewable sources of energy the GEVI concept is in keeping with the general pattern of alternative energy and alternative development. The association of the two concepts may lead to misunderstanding. That is why it is appropriate to dwell on the implications of both of them.

The concept of alternative energy has come to the forefront of the energy problem following the OPEC oil embargo of 1973. The concept was widely adopted due to the realization that the hydrocarbon fuel resources are finite, while the demand for energy was rising at a rate of 5.3% p.a. since the end of World War II. For all practical purposes, the concept means renewable energy systems: that is, solar energy, hydro power, biofuel, geothermal, wind, OTEC.

Of course, man has always used "alternative energy resources": firewood, crop drying, wind power. What is new is the magnitude of the demand for energy, and the increasing sophistication and cost of the technology to be used to meet it. This remark makes it all the more astonishing to discover that the concept of "alternative energy" is more often than not associated with the concept of "developing contries". The association is questionable. It conveys the idea that "alternative energy" is by necessity the only kind of energy the developing countries can afford, regardless of the form of energy (primary source as well as final energy form) that best suits their needs. In other words "alternative", which carries the notion of "choice", means in practice "no choice".

A more reasonable and correct association must tie positively together the concept of "alternative energy" and the concept of "alternative development", as is done in the GEVI concept. To elaborate further on this relationship one needs to consider the basic tenets of an acceptable, alternative development.

VII. Alternative Development Tenets

Today one witnesses a convergence of national policies due to the globalisation of world affairs. This tendency presents three major problems:

— How to create an international framework enabling peoples and countries to manage interdependence?
— How to manage interdependence not only in an efficient but in an equitable way?
— How to be interdependent while preserving the specificity of one another's culture?

Worldwide integration of human affairs has different effects on countries and thus on people. Marginalization tends to be the norm rather than the exception. The call for the respect of human rights has yet to be reconciled with the need for social justice, the call for solidarity with the need to avoid uniformity. What is needed, and still lacking, is a long-term oriented, broadly based sustainable development centered on human beings, which is ecologically sound, capable of widening people's options, not of reducing them to the prevailing vision of progress centered mainly on economic productivity.

The failure of four UN decades of development for the Third World since the "Truman design" of 1949 is probably due to the fact that it was exclusively defined as economic development. That is, the progress of a society was reduced to its ability to produce goods and services. Such a truncated concept of development was considered possible only by emulating the aspirations, values, culture and technology of the developed countries. Since the benefits of the economic development have not even trickled down to the masses of the developing countries, an increasing number of people are calling for an alternative development paradigm based on three premises [17]:

— endogeneity of development;
— non-hierarchical concept of human relations;
— generation of knowledge relevant for integral development.

The call for an alternative development paradigm is motivated, and made more urgent, by the fact that Third World countries are becoming ungovernable, one by one, whatever the complexion of governments, that is, right wing, left wing, military or democratic. They are all becoming incapable of adequately responding to the expectations of their people.

On the whole, the economic, social, moral and ecological crises which the world is facing today are acute enough to warrant a revision of policies and strategies. The majority of the population of many countries of the Third World has reached an acute state of generalized poverty. To reverse the trend, one must somehow reconcile the modern and traditional ways of life in the Third World. The scientific community can help in this endeavour.

As regards specifically the energy sector, constraints and opportunities of the sufficient provision of energy based on a minimal use of material and capital intensive inputs must he assessed. Two contrasting approaches to this problem can be presented. The first approach is the "technological approach". It is based on an ethic of conquest, of competition, on the control of nature, on the belief of a never ending, increasing material benefit. The second approach is the "ecological approach". It emphasizes the quality of life with socio-political institutions centered on the individual. It stresses the primacy of humanistic and spiritual values. It encourages communal activities, and the integration of activities related to work, study and leisure.

Since the importance of the demand for energy is tied to values, life styles and modes of consumption, the use of energy is bound to grow in proportion to the growth of the economy in the first approach. All things being equal, the technological approach necessitates the development at a rapid rate of new sources of energy. The cost of these is bound to increase due on the one hand to the increasing scarcity of non-renewable fossil fuels (gas, oil), and, on the other hand, to various negative impacts on the environment. The ecological approach requires a much lesser amount of resources, as shown by a case-study conducted in California [18]. The priority in the first approach is to meet a continuously increasing demand for energy at any price. The priority in the second approach is to conserve energy. Apart from that, the energy strategy has to be divised according to the necessity to optimize the availability of energy taking properly into account the constraints that affect the supply and demand of energy at a given location. Each energy strategy has to be translated into energy policy decisions. The latter affects human, and hence, societal values and lifestyle, which in turn affects the demand for energy.

Whatever the approach, it should be stressed that the energy future of mankind is made of hard and painful choices. That is why an increasing number of people are striving for a greater decentralization of institutions, governance, and economic activities, for technology to become less domineering, more decentralized, more environmental-friendly and resource-conserving. One of the results of this tendency is a steady reduction in the per capita energy requirements in developed nations since the first oil shock of 1973.

VIII. Conclusion

Science and technology have been enormously effective in the past in improving man's knowledge about his environment, and in creating new resources to replace scarce ones. However, science and technology alone cannot be relied upon to solve all the emerging global problems facing mankind, particularly on the energy front. Three reasons at least justify this assertion. The first one is that technological innovation responds to market forces and to profit opportunities. Global problems, particularly environmental ones, do not necessarily create such opportunities. The second reason is the capital constraints, which inhibit implementation of large projets. The third reason is that some environmental problems are simply beyond the reach of science and technology, such as recreating extinct species. In spite of these restrictions, technological progress can help enormously to sustain alternative development strategies.

Different world fora have shown convincingly that the current world development trajectory, driven by rapid population growth in the Third World and by income growth in the developed countries, is excessively and unsustainably energy/material intensive. Changing this state of affairs necessitates concerted actions on institutional policies, on mechanisms and frameworks for sharing new technological applications, on development of human resources, on strategies that widen people's options, not reducing them to the prevailing tenets of industrial countries. That is the only way to insure, at the same time, socio-economic and ecological sustainability.

Whatever the pertinence of the views expressed above it remains, however, that few societal decisions rival in complexity and difficulty decisions on energy. This is due to the importance of energy for everyday life. A decision to build or not to build a major energy plant, to alter price relationship between primary energy sources, or to stimulate or refrain energy demand, affects in one way or another the final energy system and thus societal patterns. An optimal energy policy thus must consider at the same time the offer and demand of energy, and the social implications of energy decisions.

ANNEX

1. *Table A*: Offer of primary energy in Sub-Saharian countries in percentage of the total energy supply.

Source of energy (commercial energy)	1986	2020 (projection)
Oil	24,0	35,0
Gas	3,0	7,5
Electricity	3,0	5,0
Coal	4,0	2,5
Total	34,0	50,0
Ligneous	66,0	50,0
Total	100,0	100,0

Hydro and geothermal electricity. Reference: UN statistical bureau.

3. *Constraints of various primary energy resources.*

Table 1: Estimated potential of world renewable energy supply.

Source	Potential (TW)		
	Technical	Realizable	Constraint
Fuel and forest farms	7,5	2,5	ecological climatological
Soft solar; soil storage; heat pump	5,0	1,0	economic technological
Hydropower	2,9	1,5	ecological social
Wind	3,0	1,0	economic
OTEC	1,0	0–1	ecological climatological technological
Geothermal	0,4	0,2	economic
Organic waste	0,1	0	balanced
Glacier power	0,1	0	technological
Tidal	0,04	0	computational
Hard solar	—	2	material and land
Total	20,00	8,3 - 9,3	

Reference: United Nations University (1971), Renewable energy prospects; (Pergamon Press).

REFERENCES

[1] World Commission of Environment and Development (WCED), 1987, Our Common Future (Oxford: University Press).

[2] Malu Wa Kalenga, 1988, «Interconnexion energy-nourriture: l'Afrique peut-elle atteindre l'autosuffisance alimentaire?», Agriculture and the quality of life: new global trends, ed. G.B. Bettolo (Vatican: Pontifical Academy of Sciences) 239-250.

— Malu Wa Kalenga, 1994, The complex food-energy nexus: how to achieve sustainability; (Bruxelles: Compte rendu de l'Académie des sciences d'Outre-Mer de Belgique); Sous presse.

[3] Ted Trainer, May-June 1990, A rejection of the Brundtland report (IFDA dossier 77) 71-84.

[4] Hafele, W., 1989, L'énergie et l'environnement: la transition (*AIEA bulletin*, vol. 31, n° 2) 5-11.

[5] Beller M. *et al.*, 1979, Impact of new energy technologies as measured against reference energy system; in UN renewable energy prospects (Oxford: Pergamon Press) 891-900.

[6] Friberg, R., 1993, General survey of technologies applicable within the solar-hydrogen concept (*International journal of hydrogen energy*, New York: Pergamon Press) 853-882.

[7] IAEA, May 1989, IAEA's contribution to sustainable development (IAEA Vienna), 35 p.

[8] Malu Wa Kalenga, June 1987, Small and medium size inherently safe nuclear reactors (*Zaïrian journal of nuclear science*, vol. 8, n° 1) 1-14.

[9] See reference [8].

[10] See reference [4].

[11] See reference [7].

[12] See reference [3].

[13] See reference [3].

[14] Malu Wa Kalenga, 1985, Incidence énergétique de l'exode rural et de l'explosion démographique en Afrique au Sud du Sahara (Presse Universitaire du Zaïre), 256 p.

[15] Malu Wa Kalenga, June 1988, Global Energy Village (GEVI): constraints and perspectives; Paper presented at the "Energy for Africa international conference", (Santa Clara, U.S.A.). 9 p.

[16] See reference [15].

[17] See among others:

— Gavani C., 1990, An international development strategy for the 90s. Global stakes require a new consensus (Switzerland: IFDA dossiers 78) 73-86.

— Amsur Rahman, 1991, Toward an alternative development paradigm (Switzerland: IFDA dossiers 81) 18-27.

[18] Carlson C. *et al.*, 1982, Energy futures, human values and lifestyles (Boulder Colorado: Westview Press) 191 p.

SCIENTIFIC PAPERS

on:

FUNDAMENTAL PRINCIPLES OF MATHEMATICS AND ARTIFICIAL INTELLIGENCE

I.

FUNDAMENTAL PRINCIPLES OF MATHEMATICS

FUNDAMENTAL PRINCIPLES OF MATHEMATICS

ENNIO DE GIORGI *

In questa relazione sono esposte alcune idee riguardanti una "teoria base" dei fondamenti della matematica alla quale da alcuni anni lavoro insieme a un piccolo gruppo di colleghi ed amici (cfr. bibliografia). La prima idea della "teoria base" è l'idea che la Matematica può considerare oggetti *qualitativamente* e non solo *quantitativamente* differenti. Per seguire la tradizione e per evitare che i primi assiomi sulle qualità rassomigliano a dei giochi di parole usiamo spesso la parola "proprietà" come perfetto sinonimo della parola "qualità". Inoltre quando q indica una qualità, mentre x indica un oggetto di qualsiasi specie, scriviamo

$$qx$$

per dire che x gode della proprietà q.

La prima qualità che consideriamo e che indichiamo con il simbolo *Qqual* è la proprietà di essere una qualità; quindi scriviamo *Qqual* z per dire che z è una qualità. Il primo assioma della nostra teoria può essere quindi scritto nella forma

$$\mathit{Qqual}\ \mathit{Qqual}$$

che equivale all'affermazione: *Qqual* gode della proprietà *Qqual*. Dopo *Qqual* introdurremo *Qrel*, cioè la proprietà di essere una relazione, *Qop*, cioè la qualità di essere un'operazione, *Qcoll*, cioè la qualità di essere una collezione, *Qnum*, cioè la qualità di essere un numero, *Qprop*, cioè la qualità di essere una proposizione. *Qpred*, cioè la qualità di essere un predicato, *Qver*, cioè la qualità di essere una proposizione vera, *Qfals*,

* Accademico Pontificio. Scuola Normale Superiore, Italia.

cioè la qualità di essere una proposizione falsa. Per ognuna di queste specie scegliamo poi alcuni oggetti che ci sembrano fondamentali, per esempio tra le collezioni abbiamo scelto *V*, collezione di tutti gli oggetti collezionabili, *Coll*, collezione di tutte le collezioni, *Ins*, collezione di tutti gli insiemi, *N*, collezione dei numeri naturali (cioè dei numeri 0, 1, 2, ...), $\emptyset$, collezione vuota. Oltre agli oggetti fondamentali delle varie specie occorre scegliere un certo numero di assiomi fondamentali, da cui seguiranno poi i teoremi che insieme agli assiomi dovranno costituire l'edificio, che speriamo coerente ed armonico, della teoria.

La libertà con cui in matematica si scelgono gli assiomi può sconcertare chi giustamente si fida delle "certezze matematiche", le ritiene solidi punti fermi anche in un secolo che ha visto molte "rivoluzioni scientifiche" e perfino il successo in epistemologia dell'idea che le teorie scientifiche non sono "dimostrabili", ma sono "falsificabili". Io credo che le «certezze matematiche» siano realmente incrollabili, ma ricordo che si tratta in generale di certezze "ipotetico-deduttive", che gli enunciati dei teoremi sono del tipo "se vale l'ipotesi *A* allora vale la tesi *B*", "se si accetta l'assioma *X* si deve accettare la conseguenza *Y*".

Io credo che il matematico debba usare ampiamente la libertà di scegliere gli assiomi ma non deve abusarne, deve evitare l'introduzione di assiomi poco espressivi che servono solo a costruire dei giochi inutili e spesso poco divertenti, cercare con grande umiltà gli assiomi più "saggi", deve ascoltare con attenzione domande, osservazioni, critiche di altri matematici e anche di studiosi di discipline più o meno lontane dalla matematica, considerarle non attacchi da cui difendere la propria teoria, ma aiuti preziosi nel difficile cammino verso una maggiore saggezza.

Usando la parola saggezza (o sapienza) intendo attribuire a questo termine tutta la ricchezza di significati che esso ha negli antichi libri sapienziali, negli scritti dei migliori filosofi, nelle migliori tradizioni culturali dei diversi popoli, e quindi penso che gli assiomi "più saggi" sono quelli che servono ad accrescere la bellezza, la ricchezza espressiva, la capacità comunicativa, il valore culturale, l'utilità della matematica.

Credo che per fare della buona ricerca matematica e specialmente per scegliere bene gli assiomi fondamentali giovano uno spirito di indipendenza e una grande disponibilità al dialogo; a queste due virtù sono vicine molte altre "virtù sapienziali" come il rispetto per la dignità, la libertà, l'originalità di ogni persona, la comprensione e l'amicizia tra quanti, in ogni tempo e in ogni paese, hanno lavorato e lavorano nel campo delle scienze, delle arti, della tecnica, con quel sentimento che gli antichi chiamarono "filosofia", cioè "amore della sapienza".

Ho parlato di quelli che *hanno lavorato* con amore della sapienza perché credo che il matematico debba amare la tradizione, ricordare con ammirazione e gratitudine i grandi matematici, scienziati, filosofi del passato, apprezzare l'inesauribile fecondità delle loro idee più belle.

Al rispetto della tradizione il matematico deve unire la disponibilità all'innovazione, ricordare l'ammonimento di Shakespeare "vi sono più cose tra cielo e terra di quante ne sogni la tua filosofia" a cui vorrei aggiungere l'osservazione che vi sono più "cose" nella mente e nel cuore di ogni uomo di quante lui stesso non immagina.

Il matematico "saggio" guarda con ammirazione la bellezza, l'armonia, l'ordine del creato, che Galileo considerava un libro scritto con caratteri matematici, e considera pure con ammirazione la ragione umana che riesce a decifrare, almeno in parte, questo libro meraviglioso e a riflettere sulla propria stessa natura, scoprendo con lo studio della logica alcune regole fondamentali del ragionamento umano.

D'altra parte chi studia la matematica con serietà e passione sincera ha molte occasioni di riconoscere i propri eventuali errori, il limitato interesse di molti risultati da lui stesso conseguiti, la grande importanza di molti problemi che non riesce a risolvere e qualche volta nemmeno ad impostare chiaramente, ha in sostanza molte occasioni di esercitare quelle virtù dell'umiltà e dell'onestà intellettuale che mi sembrano componenti importanti di ciò che uno dei più antichi libri sapienziali della Bibbia, il Libro dei Proverbi, chiama «*timor Domini principium sapientiae*».

Per applicare l'ammonimento di Shakespeare ai fondamenti della matematica, il matematico deve essere disposto all'innovazione, deve cercare sempre di allargare l'orizzonte della sua ricerca in modo che comprenda nuove "cose" non contemplate dalle precedenti teorie.

Dovendo citare una innovazione con cui mi sembra che la "teoria base" allarga l'"orizzonte matematico" osserverei che il modo di trattare le collezioni mi sembra innovativo rispetto alle usuali presentazioni del concetto di insieme.

Ricordo a questo proposito che nella "teoria base" gli insiemi sono particolari collezioni, che la collezione vuota $\emptyset$ e la collezione dei numeri naturali N sono insiemi, mentre non sono insiemi la collezione universale V, la collezione di tutte le collezioni *Coll* e la collezione di tutti gli insiemi *Ins*. In termini un po' grossolani, si potrebbe dire che non sono insiemi le collezioni "troppo grandi", non solo infinite ma "infinitamente più grandi di ogni insieme infinito". Non vi sono invece differenze qualitative tra oggetti che appartengono alla collezione V e oggetti che appartengono a qualche insieme: tutti gli oggetti presi in considerazione dalla nostra teoria appartengono alla collezione V, per ogni oggetto x vi è un "insieme singolare" (o "singoletto") al quale appartiene solo l'oggetto x, questo insieme viene usualmente indicato con il simbolo $\{x\}$. Evidentemente i singoletti sono i più "piccoli" tra tutti gli insiemi, con esclusione dell'insieme vuoto, che è l'insieme minimo.

Avendo parlato di innovazione a proposito delle collezioni, debbo aggiungere per onestà intellettuale che la mia conoscenza della letteratura matematica e la mia abilità nelle ricerche bibliografiche sono notevolmente inferiori a quelle di molti miei colleghi e riesco a lavorare in matematica soprattutto perché molti colleghi e amici mi hanno sempre fornito informazioni bibliografiche interessanti; perciò, mentre mi sento relativamente sicuro nei giudizi sull'interesse di un'idea matematica, sono meno sicuro nell'affermare il suo carattere di assoluta novità.

Volendo unire innovazione "saggia" e rispetto della tradizione, un posto d'onore è stato riservato nella "teoria base" all'aritmetica che, dai tempi di Pitagora, Diofanto, Fermat ai giorni nostri è sempre stata la "regina della matematica". Per tale motivo, subito dopo l'insieme dei numeri naturali N, introduciamo la più semplice operazione aritmetica, cioè l'operazione N *succ* che ad ogni numero naturale associa il numero naturale successivo.

Con l'operazione semplicissima N *succ* introduciamo la qualità *Qops*, cioè la proprietà di essere un'*operazione semplice* e alcune semplici notazioni; precisamente se f è un'operazione semplice che può essere eseguita sull'oggetto x indichiamo con

$$fx$$

il risultato che si ottiene eseguendo l'operazione f sull'oggetto x, per esempio

$$N\,succ\,0 = 1, \quad N\,succ\,1 = 2, \quad N\,succ\,2 = 3, \quad N\,succ\,3 = 4, \quad \ldots, \text{ ecc.} \ldots$$

Dopo l'operazione N *succ* si introducono le altre operazioni aritmetiche per cui, non volendo appesantire il discorso con molte formule, rinvio alla nota [4].

Dopo l'introduzione dell'aritmetica la "teoria base" può essere sviluppata seguendo due strade: nella prima si dà la precedenza alle relazioni, alle collezioni, agli insiemi, nella seconda si dà la precedenza alle qualità, alle operazioni, alle proposizioni, ai predicati; in sostanza, la prima strada tende verso un'ampia estensione dell'usuale teoria degli insiemi, la seconda verso un'ancora più ampia estensione del calcolo delle proposizioni e del calcolo dei predicati. Nel seguito di questa esposizione accennerò soprattutto a questa "seconda strada", sia perché è ancora la meno esplorata, sia perché lungo di essa si incontrano molti problemi a mio avviso interessanti per la matematica, la logica, l'informatica, la filosofia, l'epistemologia, ecc. In sostanza, si tratta di valutare le possibili risposte alla domanda: "Come si possono costruire teorie matematiche che comprendono almeno in linea di principio la matematica e la logica tradizionali e abbiano un alto grado di 'autodescrizione'?" Uso il termine "autodescrizione" di una teoria per dire che tra i suoi oggetti vi è la teoria stessa; esempi di autodescrizione si possono avere pensando a un vocabolario che, fra le altre parole, comprende la parola "vocabolario", oppure a una grammatica, scritta seguendo le regole grammaticali e sintattiche.

Per descrivere almeno i primissimi passi lungo quella che abbiamo chiamato la seconda strada è necessario introdurre qualche semplicissima notazione riguardante i risultati che si ottengono mediante l'esecuzione successiva di più operazioni semplici. Cominciamo perciò con la considerazione di due operazioni semplici f, g e supponiamo che vi siano tre oggetti x, y, z, tali che $fx = y$, $gy = z$. In questo caso useremo la notazione

$$z = g(fx)$$

per indicare che z è ottenuto eseguendo prima l'operazione f sull'elemento x e poi eseguendo l'operazione g sul risultato così ottenuto. Nell'uso di questa notazione non si esclude il caso $f = g$, per esempio avremo

$$N\,succ\,(N\,succ\,0) = N\,succ\,1 = 2, \qquad N\,succ\,(N\,succ\,2) = N\,succ\,3 = 4, \ldots$$

$$N\,succ\,(N\,succ\,7) = N\,succ\,8 = 9, \ldots$$

Può accadere che, eseguendo una operazione α su un certo oggetto x, si ottiene un'altra operazione β; per tener conto di questa eventualità, stabiliamo che: se α e β sono due operazioni e x, y, z sono tre oggetti tali che $\beta = \alpha x$, $z = \beta y$, useremo la notazione

$$z = (\alpha x)y;$$

analogamente, se α, β, γ sono tre operazioni ed x, y, z, t quattro elementi tali che $\beta = \alpha x$, $\gamma = \beta y$, $t = \gamma z$, useremo la notazione

$$t = ((\alpha x)y)z.$$

Analogamente si potrebbe trattare il caso di un numero maggiore di operazioni, seguendo sempre la convenzione che le operazioni che si trovano dentro una parentesi devono essere eseguite prima di quelle che si trovano fuori. Stabilite queste notazioni elementari, possiamo introdurre due operazioni semplici importanti nella teoria dei predicati, l'operazione *Predord* che ci dà l'*ordine dei predicati* e l'operazione *Gpred*, operazione *generatrice di predicati*, che associa ai vari oggetti finora considerati i predicati che ne descrivono il comportamento. All'operazione *Predord* imponiamo due assiomi:

1. Affinché p sia una proposizione, cioè goda della proprietà *Qprop*, occorre e basta che p sia un predicato, cioè goda della proprietà *Qpred* e che l'ordine di p sia zero, cioè $Predord\ p = 0$. In altri termini, identifichiamo proposizioni e predicati di ordine zero.

2. Se n è un numero naturale, p è un predicato, l'ordine di p è il numero naturale successore di n, cioè

$$Predord\ p = N\ succn,$$

allora p è un'operazione semplice e, per ogni scelta dell'oggetto x, px è un predicato di ordine n.

In sostanza, questi due assiomi ci dicono che dai predicati di ordine più alto si possono ottenere in innumerevoli modi predicati di ordine più basso, ed infine ottenere delle proposizioni.

L'altra operazione *Gpred* associa a ogni oggetto x un predicato, il cui ordine caratterizza in sostanza la "complessità strutturale" di x e pertanto verrà chiamato anche col termine "arietà", suggerito dalle parole unitario, binario, ternario, quaternario, ecc. Volendo esprimere questo fatto con un assioma, si può enunciare il seguente assioma:

3. Esiste una operazione semplice *Ar* tale che per ogni oggetto x

$$Ar\ x = Predord(Gpred\ x).$$

Per quanto riguarda l'arietà dei vari oggetti finora introdotti, notiamo che:

Se *Qnum* x allora $Ar\ x = 0$, cioè tutti i numeri hanno arietà 0.
Se *Qprop* p allora $Ar\ p = 0$.
Se *Qqual* q allora $Ar\ q = 1$.
Se *Qcoll* C allora $Ar\ C = 1$.
Qops f se e solo se *Qop* f, $Ar\ f = 2$.

Accanto alle operazioni semplici si possono considerare delle operazioni complesse, binarie, ternarie, ecc., e per esse, detta *Qopb* la proprietà di essere un'operazione binaria, *Qopt* la proprietà di essere un'operazione ternaria, ecc. avremo:

Qopb g se e solo se *Qop* g, *Ar* $g = 3$,
Qopt h se e solo se *Qop* h, *Ar* $h = 4$, ecc.

Analogo discorso si può fare per le relazioni distinguendo le relazioni binarie che godono della qualità *Qrelb*, le relazioni ternarie che godono della qualità *Qrelt*, le operazioni quaternarie che godono della qualità *Qrelq*, ecc. per le quali avremo:

Qrelb α se e solo se *Qrel* α, *Ar* $\alpha = 2$,
Qrelt β se e solo se *Qrel* β, *Ar* $\beta = 3$,
Qrelq γ se e solo se *Qrel* γ, *Ar* $\gamma = 4$, ecc.

Questi assiomi mostrano che l'operazione *Gpred* è un'operazione essenziale per lo sviluppo della "seconda strada" che può essere seguita nella costruzione della teoria base dei fondamenti della matematica e il suo ruolo è meglio precisato dagli ulteriori assiomi seguenti:

4. Data una qualità q affinché un oggetto x goda della proprietà q occorre e basta che (*Gpred* q)x goda della qualità *Qver*.

5. Data una operazione semplice f affinché si abbia $fx = y$ occorre e basta che ((*Gpred* f)x)y goda della qualità *Qver*.

Per ricordare facilmente gli assiomi 4 e 5, conviene introdurre dopo le notazioni con parentesi le notazioni con virgolette e scrivere:

"qx" in luogo di (*Gpred* q)x, "$fx = y$" in luogo di ((*Gpred* f)x)y.

Analoghi assiomi valgono per le collezioni, le operazioni più complesse (binarie, ternarie, ecc.), le relazioni ed anche per esse si possono introdurre opportune notazioni con virgolette. Dopo aver introdotto in questo modo predicati e proposizioni, si possono introdurre le principali operazioni del calcolo dei predicati, per esempio negazione, disgiunzione, interpretazione delle formule, ecc. ed anzi tutte queste operazioni risultano immerse in un quadro molto più ampio dell'usuale. In tale quadro continua a valere il principio di non contraddizione, cioè una proposizione non può godere contemporaneamente della proprietà *Qver* e della proprietà *Qfals*, ma, avendo largheggiato con le operazioni che "producono proposizioni", non possiamo escludere che vi siano proposizioni che non godono di nessuna di queste due proprietà e quindi dobbiamo aggiungere alla qualità *Qprop* la qualità *QVF*, cioè la proprietà di essere una proposizione che gode di *Qver* oppure gode di *Qfals*.

Ciò in sostanza equivale a riconoscere che alla parola proposizione diamo una grande ampiezza di significati, intendiamo comprendere sia le proposizioni "sensate, chiare, univoche" sulle quali è possibile esprimere un giudizio sicuro di verità o di falsità, sia altre proposizioni più ambigue, oscure, meno sensate, sulle quali è difficile o impossibile esprimere un tale giudizio. Un esempio di tale impossibilità è dato dalle "proposizioni autoneganti" cioè dalle proposizioni p che verificano la condizione:

$$p = \text{“}Qfals\ p\text{”}$$

Le proposizioni autoneganti erano *sostanzialmente* conosciute fin dall'antichità attraverso l'antinomia del mentitore che io considero "madre di tutte le antinomie". Mentre è facile ammettere che vi siano alcune proposizioni che godono della proprietà *QVF* ed altre che non godono di tale proprietà, sembra difficile (forse impossibile) tracciare una buona "linea di demarcazione" tra le prime e le seconde. Forse proprio per questo lo studio della qualità *QVF* mi sembra molto interessante e probabilmente esso si collega a molti problemi di logica, informatica, psicologia, filosofia, linguistica, ecc., alla stessa riflessione sui concetti di intelligenza umana e di "intelligenza artificiale".

BIBLIOGRAFIA

[1] Autori Vari: *La qualità QVF e la teoria base dei fondamenti della matematica*, Preprint Scuola Normale Superiore di Pisa, 1994, lavoro in corso di redazione.

[2] M. Clavelli, E. De Giorgi, M. Forti, V.M. Tortorelli: *A self-reference oriented theory for the foundations of mathematics*, Contributions en l'honneur de J.L. Lions, Parigi, 1986.

[3] E. De Giorgi, M. Forti: *Una teoria quadro per i fondamenti della matematica*, Atti Accademia Nazionale Lincei, Rendiconti Classe Scienze Fisiche, Matematiche, Naturali, serie 8, volume 79 (1985), pp. 55-67.

[4] E. De Giorgi, M. Forti, G. Lenzi: *Una Proposta di Teorie Base dei Fondamenti della Matematica*, Atti Accademia Lincei, Rendiconti Classe Scienze Fisiche, Matematiche, Naturali, serie 9, volume 5 (1994), pp. 11-22.

[5] M. Forti, F. Honsell: *Models of Self-Descriptive set Theories*, in: Partial Differential Equations, Essays in honour of Ennio De Giorgi, editori F. Colombini et alii Birkhäuser, 1989, pp. 473-518.

[6] G. Lenzi: *Estensioni contraddittorie della teoria ampia*, Atti Accademia Nazionale Lincei, Rendiconti Classe Scienze Fisiche, Matematiche, Naturali, serie 8, volume 83 (1989), pp. 13-28.

[7] G. Lenzi, V.M. Tortorelli: *Introducing Predicates into a basic Theory for the Foundations of Matematics*, Preprint Scuola Normale Superiore di Pisa, luglio 1989.

MODELLING ANALYSIS AND NUMERICAL SIMULATION

LUIS A. CAFFARELLI *

When asked to make some remarks about fundamental principles of mathematics I took it in a generalized sense, in an ample sense, and I thought it would be of interest for some of you to make some remarks on some issues that I feel have arisen lately in the mathematical community about the relations between what one could call reality, modelling and computing simulation.

I thought it was a side issue, but I was surprised to hear a few references to it on the first day here. One of them was by Prof. Pullman when he mentioned the somewhat unfortunate remark of Einstein to the effect that reality better conformed to his model. So I would like to make a little review of what is the difference between reality, a model and a large-scale computing simulation. So let me start by describing how a model is made. I'll talk about continuum dynamics, but when making a model, say, about market economy or mathematical financing, maybe because of historical reasons or because of historical behaviour, the way the model is made is more or less the same: The scientist first looks at the phenomenon he wants to describe — for instance, compressible or incompressible flow in some configuration — around a body, inside a pipe, etc., and he makes a list of all the factors, all the variables that may describe it: velocity, pressure, density, external forces, if you have flame propagation, the chemical composition of the flame, and so on. Then, one writes relations, i.e. equations that relate these quantities, for instance, conservation laws that will say that mass is preserved, momentum is preserved, energy is preserved, which are usually based on general principles of continuum mechanics, and these are written as differential equations or just simple equations.

---

* Pontifical Academician.

This is the first set of equations, and the second family of equations is what in continuum mechanics is usually called "constitutive relations", for instance, relations between stress and strain in elasticity, or velocity and pressure. These are usually based on experimental behaviour.

In other words, the scientist makes some experiments that are often relatively simple, measures pressure along a column, a strain and strain relation for simple deformations of the material; on the basis of this experimental data, he draws some curves and deduces some more or less approximate laws of behaviour.

Once the scientist has written to his own satisfaction every possible element that will influence his phenomena, then he has to start making simplifications. For instance, you don't need relativity theory for most of your daily life, Newtonian physics is enough. We can assume that the process is isothermal, incompressible, or a flow is irrotational, and that means reducing and simplifying the equations, so from a large number of equations we come back to a few. In short, we have a hierarchy of models according to how precise we want to be in the description of our problem.

Now, for mathematics, before large-scale computations, the value of this model was relatively restricted. One will look at the asymptotic behaviour, find limiting equations when some parameters like viscosity will become very small, compute the effective constant of materials, etc.

About twenty years ago, reasonably large-scale computation started. From my colleagues at Courant and from what I remember, twenty years ago people would compute, for instance, plane, compressible potential flow around a wing profile which meant one or two equations in a grid of 30 by 30 nodes, with a hundred time step. This was a simple model of two mathematical equations trying to describe what happened around a wing profile.

Now, for instance, one of my colleagues at Courant, Charles Peskin, is doing a tridimensional model of a beating heart. Let me try to give you an idea of the magnitude of that computation.

For the flow he uses the standard classical equation of incompressible fluid, the heart; the structure of the heart is a set of points in some sense attached to each other as a set of fibers: that means points attached to each other by some imaginary elastic fiber whose elasticity changes periodically.

He reproduces something that looks very much like a heart beat, and to do that, he is using a grid of 200 by 200 points and five thousand time steps, eight equations in mixed Eulerian Lagrangian coordinates.

Now most models are based on relatively simple experiments at reasonable states. They are based on small perturbations around a well-understood state, around rest or some well understood plane flow. On top of that, when doing numerical simulations there is discretization. That means numerical viscosity and many other effects that may change the final results dramatically.

So, what I wanted to warn you about is that what one should really do at this point is understand better mathematically what large deformations are which are not around a simple state. These are the problems that appear as soon as you do want to look at more sophisticated models, for instance, in non-linear material science, non-linear plasticity, etc. We have to understand changes in Lagrangian coordinates, not just by looking at velocity fields, but by actually following the deformation of the domain, and this appears in non-linear, material science. There is already some work in progress but still a lot has to be done to know to which degree computations agree to modelling and modelling to reality.

II.

ARTIFICIAL INTELLIGENCE

CONTROLE ET INTELLIGENCE ARTIFICIELLE *

J.L. LIONS **

La Société de Téléologie

En Novembre-Décembre 1944, il y a donc 50 ans, John Von Neumann et Norbert Wiener envisagent la création d'une "Société de Téléologie" pour "l'étude des buts de l'action chez les êtres vivants et dans les constructions mécaniques et électriques qui peuvent les imiter". Dès 1943, Norbert Wiener a recommandé à Von Neumann l'étude du travail, qui vient de paraître, de W.S. Mc Culloch et W. Pitts (*A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bull. Math. Biophys. 5, 1943, p. 115-133). Von Neumann a, comme toujours avec lui, immédiatement compris l'importance de ce travail qui présente les premiers exemples de modèles de réseaux neuronaux conçus pour effectuer des opérations de logique. En fait, introduisant la redondance, où plusieurs neurons sont affectés pour un travail qui pourrait être accompli par un seul d'entre eux, Von Neumann montrera, une dizaine d'années plus tard (Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components, dans *Automata Studies*, édité par C.E. Shannon et J. Mc Carthy, Princeton University Press, 1956) comment concevoir des réscaux de Mc Culloch et Pitts pour effectuer des calculs mathématiques avec une grande fiabilité.

Néanmoins, lorsque le 24 Janvier 1945 Von Neumann reçoit une lettre de Wiener exprimant le souhait que la Société de Téléologie réfléchisse "au problème du passage de la machine à calculer à la machine à contrôler", il trouve l'idée intéressante certes, mais prématurée. Il s'en expliquera longuement et retournera sur ce sujet dans une lettre détaillée à Wiener le 29 Novembre 1946 (source: W. Aspray, *The Origins of John*

* Read by André Lichnerowicz.

** Académicien Pontifical. Collège de France, Paris.

Von Neumann's Theory of Automata, dans *The Legacy of John Von Neumann*; Amer. Math. Soc., Vol 50, 1988, p. 289-309). Il a deux types d'objections:

a) le caractère trop général des contributions de A. Turing et de Mc Cullogh et Pitts "these authors have demonstrated in absolute and hopeless generality ..." (tout en notant d'ailleurs le caractère très élémentaire des réseaux de Mc Culloch et Pitts, simple point de départ de très nombreux travaux récents et en cours sur les réseaux neuronaux);

b) l'absence d'outils adéquats d'observation et de mesure des activités cérébrales.

La machine à contrôler

Mais à quoi pense vraiment N. Wiener lorsqu'il propose une réflexion et une étude sur "la machine à contrôler"? Il publie ses idées là-dessus en 1948. La terminologie a changé. La "Téléologie" a été remplacée par le nom, désormais classique, de "Cybernétique" (*Cybernetics: or Control and Communication in the Animal and the Machine*. Technology Press, Cambridge, 1948), Etymologiquement "l'art de gouverner", plus précisément l'art de gouverner les "systèmes" par l'étude des mécanismes de communication et de contrôle chez les êtres vivants et leur réalisation dans des machines.[1] Si l'entreprise est possible voilà les machines à contrôler que Wiener évoque dans sa lettre à Von Neumann.

Elles auront donc vocation à remplacer, en totalité ou en partie, l'intervention humaine dans le "gouvernement" des systèmes, plus précisément pour la conduite de processus physiques, pour la gestion économique, pour la prise de décision en situation imprévue. Cela serait, dans une certaine mesure, la réalisation du mythe de Dedale, considéré (cf. J.L. Lions, Daedalon Gold Medal Lecture, Athènes, 1992) comme:

(i) précurseur de la Robotique (les forgerons mécaniciens de Crète ou de Rhodes parlent de "vie dédalique");

(ii) maître de la complexité (lorsque, par exemple, il résout le problème de complexité intellectuelle maximum, la Laryrinthe);

(iii) héros de l'intelligence (selon la formule de F. Frontisi-Ducroux, *Dédale, Myrthololgie de l'artisan en Grèce ancienne*, Paris, Maspers, 1975).

[1] Le grec "Kybernetis" devient "gubernator" en latin, d'où le nom "governor" (donc "l'art de gouverner") donné en 1868 par Maxwell à un mécanisme de *feedback* régulant la vitesse d'une machine à vapeur.

John Von Neumann et le Contrôle du Climat

L'idée de Wiener de la "machine à contôler" n'est pas poursuivie dans les dernières années 40, ce qui n'empêche pas Wiener et Von Neumann de travailler ensemble, dans un groupe très pluridisciplinaire, les Conférences MACY, consacrées aux Mécanismes de *Feedbacks* en Biologie et en Sciences Sociales, appelées ensuite "Les Conférences Cybernétiques". Von Neumann y participe activement. Son intérêt pour les aspects "contrôle" de systèmes se manifeste particulièrement lorsque, en 1955, dans un article publié dans Fortune "*Can we Survive Technology*" il évoque les possibilités de contrôle du climat. C'est l'époque où il dit qu'à son avis "climate is easier to control than to predict" (Dvoretsky, communication personnelle). Il y a donc, dès la fin des années 40, accord sur l'importance fondamentale du contrôle des systèmes — l'un des chapitres de l'Intelligence Artificielle — dans la terminologie qui apparaît quelques années après.

Les points de vue sur les méthodes possibles pour aborder le contrôle des systèmes sont très variés, comme le montrent les très riches discussions de Von Neumann, Wiener et leurs collègues des Conférences Cybernétiques.

S'appuyant sur une série d'exemples, la suite de l'exposé va présenter quelques-unes des évolutions de ces questions.

Le temps du contrôle et de l'intelligence

"Gouverner" un système, ou encore le contrôler, c'est faire en sorte (si possible) qu'il se comporte selon nos souhaits. Le temps dont on dispose pour cela joue un rôle tout à fait fondamental, pour deux raisons principales:

(i) l'échelle de temps "intrinsèque" du système;[2]

(ii) les horizons temporels des "souhaits" que le contrôle a pour but de faire réaliser.

L'intervention du contrôle, sa réalisation, ou encore les réactions "intelligentes" de la machine à contrôler humaine, dépendent de ces échelles de temps. Donnons quelques exemples.

[2] C'est en fait beaucoup plus compliqué que cela, puisqu'un système peut avoir *plusieurs* échelles de temps selon ses composantes, liées entre elles par les "feedbacks".

Exemple 1

Considérons le système {cycliste + bicyclette}.

Si l'objectif à réaliser, le souhait, est d'éviter la chute et l'accident, il y a une seule échelle de temps réel d'une chute, à éviter. Il semble extrêmement peu probable que quelqu'un ait jamais appris le contrôle du système par la machine, constituée par la tête et les jambes du cycliste, autrement que par apprentissage — et non par usage d'un modèle mathématique! — La "machine" apprend à connaître les relations entre les entrées (position, effort, ...) et les sorties (vitesse, stabilité, ...).

Supposons maintenant que le souhait, l'objectif, soit la production d'une nouvelle bicyclette — objectif à réaliser pour une échelle de temps tout à fait différente que dans la situation précédente. On passe de la seconde à l'année, ou plus. Les méthodes vont ici changer; il est tout à fait raisonnable que des modèles mathématiques soient utilisés pour la conception. Par exemple, des modèles d'aérodynamique. La "machine à contrôler" a changé: le cycliste débutant est remplacé par un professionnel qui pourra produire ses efforts en tenant compte de plusieurs échelles de temps, selon, par exemple, la distance à parcourir.

Exemple 2

Une sonde spatiale destinée à observer plusieurs planètes doit suivre une trajectoire calculée au préalable à partir des lois de la Mécanique céleste. Mais la trajectoire réelle, même en l'absence de tout incident, différera toujours quelque peu de la trajectoire optimale (dite nominale) préalablement calculée: aléas dans la traversée de l'atmosphère, dans les poussées des moteurs, ... La machine à contrôler consiste ici, en simplifiant, en l'ordinateur (embarqué *et* au sol) et en les moteurs (qui mettent en oeuvre les actions de contrôle ou de commande). Il faut, pour décider de l'action, connaître la "distance" entre la trajectoire réelle et la trajectoire nominale, donc disposer de mesures fiables, en un temps cohérent avec celui dont on dispose pour la prise de décision.

Le temps de transmission des télémesures devient significatif, compte tenu des distances. C'est l'une des difficultés de la Robotique lunaire ou martienne lorsqu'elle doit être coordonnée avec les bases terrestres.

Dans le premier exemple, la "machine à contrôler" est "humaine", dans le deuxième, elle est "automatique". Elle n'utilise pas de modèle mathématique dans le premier cas (sauf si l'on modifie les objectifs), elle repose en revanche sur des modèles mathématiques dans le deuxième cas.

Voyons maintenant un exemple "mixte", où les deux formes "d'intelligence" interviennent en coopération.

Exemple 3

La mission lunaire Apollo a nécessité, comme dans l'Exemple 2, des calculs préalables de trajectographie et l'usage de contrôles "temps réel" basés sur des filtres de R. Kalman (qui utilisent essentiellement, à la différence des filtres antérieurement introduits par N. Wiener, "l'équation d'état" du système à contrôler, c'est-à-dire un modèle mathématique du système) et les équations de Riccati (introduites en 1724 en réponse à divers problèmes posés par Leibniz). [On pourra consulter J.L. Lions, *Le temps du contrôle*, *La Vie des Sciences*, et la *Bibliographie* de ce travail]. Mais bien entendu, les astronautes sont à bord et ont leur rôle (essentiel!), non seulement pour l'exploration lunaire, mais aussi dans l'alunissage et dans l'atterrissage.

Il s'agit là, toujours dans le cadre du contrôle, de la coopération entre deux "machines à contrôler". L'une est basée sur les modèles mathématiques pour l'analyse et sur les logiciels. L'autre s'appuie sur une expérience longue (dans le temps) d'ingénieurs et de pilotes hors normes et sur un apprentissage spatial, fruits de simulations, qui ne sont elles-mêmes possibles que grâce aux modèles mathématiques, aux logiciels correspondants et à la visualisation des résultats.

Ce type de coopération entre diverses formes d'intelligences et des méthodes variées de contrôles est mis en oeuvre régulièrement, dans des situations de très grande complexité dans les vols du Shuttle, dans des situations souvent très complexes mais qui tendent à devenir routinières dans le pilotage des avions de transport.

Donnons maintenant un exemple où le temps "intrinsèque" du système est plus grand mais où d'autres facteurs de complexité apparaissent.

Exemple 4

Les possibles inondations des docks de Londres, sous l'effet conjugué de hautes marées et de conditions météorologiques particulières, sont contrôlées à l'embouchure de la Tamise, par un système de fermetures modulables.

L'objectif est simple et sans ambiguité. La mise en oeuvre des variables de contrôle (le "degré" d'ouverture ou de fermeture des diverses portes) est également simple (ce qui, bien sûr, ne veut pas dire que la conception et la construction de ces remaquables installations aient été chose simple!).

La décision est basée sur les résultats numériques obtenus à partir des modèles mathématiques des marées et de la météorologie. Une situation un peu semblable se rencontre en Hollande, pour la gestion des polders et, également, dans l'usage de l'énergie hydraulique dans le système global de production d'électricité.[3] Tout cela conduit, assez logiquement, à une interrogation sur les contrôles "naturels". Ainsi l'équilibre écologique naturel de la forêt Amazonienne, observé avant des interventions humaines, a-t-il donné lieu à des interprétations par analogie avec les réseaux neuronaux?[4]

Peut-être est-il utile de signaler, à l'échelle globale de la Planète Terre, la théorie de Gaia proposée par Lovelock ("The age of Gaia", Norton and Company, N.Y. 1988), faisant de la Terre entière un système "intelligent". Cela pourrait être interprété comme une action intelligente de type contrôle effectuée par les végétaux sur les êtres vivants pour assurer la stabilité du système.[5]

Il y a peut être d'ailleurs là — *indépendamment* de la validité éventuelle de Gaia — une source d'inspiration pour des réflexions sur "l'intelligence". On peut d'ailleurs noter l'usage récent d'algorithmes d'optimisation dits "génétiques" ou "écologiques" basés sur des analogies, réelles ou supposées, avec les équilibres "naturels" (Pour des résultats précis dans ce sens, en liaison avec les systèmes dynamiques, cf. Y. Rabinovic, A. Sinclair et A. Wigderson, *Quadratic Dynamical Systems, Proc. of the 33th IEEE Symposium on Foundations of Computer Sciences*, 1992, p. 304-313).

Exemple 5

Revenons encore sur le rôle du temps dans les processus de contrôle et de décision à propos d'un exemple de pollution dans l'océan suite à un accident maritime, d'un pétrolier par exemple. Première échelle de temps: celle, très courte, du sauvetage des vies humaines. Deuxième échelle de temps: celle des prévisions météorologiques locales, temps court ou moyen, de l'heure à quelques jours au plus. Troisième échelle: le temps, long cette fois, des conséquences écologiques à moyen et long terme. Dans chacune des étapes, et tout particulièrement pour les deuxième et troisième étapes, les modèles mathématiques fournissent les plus probables scénarios et contribuent aux décisions. Nouvel exemple

---

[3] Consulter l'article déjà cité, *Le temps de Contrôle*, pour les références sur l'exemple hollandais et sur la gestion de l'énergie hydraulique par E.D.F. (Electricité de France).

[4] A. Labeyrie, *Communication personnelle*.

[5] L'objection, souvent faite, selon laquelle les ordres de grandeur des énergies en cause sont trop différents pour la plausibilité de Gaia, n'est pas absolument évidente si l'on se réfère au fait que les systèmes instables ou chaotiques ou turbulents sont, d'une certaine façon, *plus faciles* à contrôler que les systèmes très stables.

de coopération homme-machine, à tous les niveaux. Dans l'exemple présent et dans le précédent, les calculs pour les échelles de temps longues, supposent que l'on dispose de moyens de calcul très puissants et nécessitent un certain temps de calcul et de simulation. Ils ne peuvent être conduits que "off-line", c'est-à-dire en dehors du processus décisionnel immédiat et de contrôle, donc que les principes, les modèles et les programmes aient été étudiés analysés et écrits *avant* tout début d'intervention.

Face à une situation nouvelle, ou imprévue, ils sont, aussi rapidement que possible, actualisés et mis à jour; ils sont utilisés ensuite dans les processus dimensionnels. On rencontre d'ailleurs, à ce propos, une autre tendance:[6] les très grandes simulations (météorologie, climatologie, aéronautique, ...) sont des processus qui ont besoin, pour être contrôlés, de systèmes homme-machine, rencontre entre l'I.A. et les calculs à très haute performance.

Quelques remarques finales

Le demi-siècle écoulé montre qu'en ces matières, il faut associer progrès scientifique et moyens technologiques nouveaux.

Dans ce cadre, et en restant dans le domaine du contrôle — du "gouvernement" au sens où l'entendait Wiener — il faut ajouter deux importantes évolution en cours:

(i) les nouvelles possibilités que la technologie offre pour la mise en oeuvre des contrôles, tels les matériaux à mémoire (matériaux dits "intelligents"), ou l'usage possible des lasers dans des interventions à échelle moléculaire, etc. ...

(ii) le parallélisme — l'usage coordonné et coopératif de plusieurs ordinateurs qui revient à l'organisation de la coopération de nombreuses "intelligences".

L'Intelligence Artificielle peut être abordée sous l'angle du contrôle des systèmes. C'est la démarche que nous avons, très superficiellement, suivie dans cet exposé. Bien entendu, cela ne donne accès qu'à quelques éléments de ce sujet immense et complexe. Mais ce point de vue conduit naturellement à mettre l'accent sur les indispensables coopérations entre les diverses formes d'intelligence et à nombre croissant d'acteurs humains et de "machines à contrôler" et à s'interroger sur les relations entre le temps et l'intelligence.

---

[6] Signalée par P. Caseau, *La modélisation numérique: où en serons-nous dans dix ans?* Note interne EDF, 1991.

INTELLECTUAL PROSTHESES. THEORY AND PRACTICE

ANTONIO M. BATTRO *

Nec manus, nisi intellectus, sibi permissu, multam valent; instruments et auxilibus res perficitur.
Francis Bacon, *Novum Organum.*

I would like to quote Sir Francis Bacon's *Novum Organum* to introduce the concept of intellectual prosthesis: *nec manus, nisi intellectus, sibi permissu, multam valent; instrumentis et auxilibus res perficitur.* Bacon's *instrumentis et auxilibus* are our "prosthetic devices". The psychologist Jerome Bruner rightly translated "neither the hand nor the mind alone, left to itself, would amount to much". And he asked: "What are these prosthetic devices that perfect them?" (Bruner, 1986).

The answer may be: the computer. Since the beginning of computation and robotics a great effort has been made to develop "friendly" human/machine interfaces: i.e. voice synthesizers, speech recognition devices, electro-mechanical switches, scanners, multimedia tools, virtual reality, tele-computing, etc. The remarkable growth in the power and speed of computers, and of the corresponding software, stimulates new uses of these machines in most human activities. In this paper I shall deal with the application of computers in the field of neuro-cognitive sciences, in particular in the education and rehabilitation of the disabled person.

* Battro & Denham, Buenos Aires, Argentina.

1. Human/machine interfaces

The idea of using the computer as a cognitive tool for the disabled was introduced by Seymour Papert at MIT in 1978. He coined the name of "information prosthetics" at a time when few could imagine any possible humanitarian use of computers (Papert and Weir, 1978). This technological and social breakthrough opened new ways to improve the quality of life of physically and mentally handicapped people by employing computers (Valente, 1979). In this report I shall deal with "intellectual prostheses" as a subset of the more general "informational prostheses". Indeed the latter include "physical prostheses" as well (cochlear implants for the deaf, limb computer stimulation for quadriplegic persons, array of electrodes in the visual cortex for the blind, etc.). The "intellectual prosthesis" instead, does not imply any physical and direct contact between the nervous system and the computer. It is a pure "functional" prosthesis not a "physical" one. However, it can be expected that both "intellectual" and "physical" informational prostheses will interact in the future in ways that we hardly imagine today. A good introduction into this subject can be found in the science-fiction book *The Turing Option*, written by Marvin Minsky, one of the founders of the field of Artificial Intelligence, in collaboration with Harry Harrison (Harrison and Minsky, 1993).

Fig. 1 shows a communication network among humans H, machines M and environments E. This prism offers a global view about the specific links or interfaces: M/M, H/H, M/H, H/E, E/E. When one of these links is perturbed or abolished, some new path in the network can be opened by means of specific technologies. I shall summarize now our clinical and educational experience in this field.

a) *H/H discommunications.* When "face to face" communication is impaired, as with deaf persons, the computers may bridge the gap. In 1982

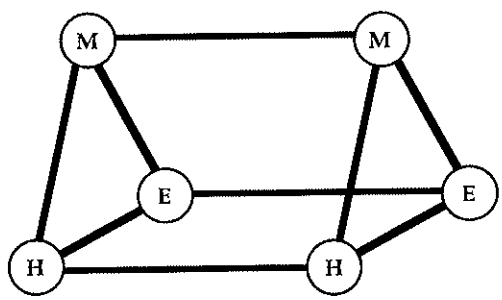


Fig. 1. Links and interfaces between humans H, machines M and environments E.

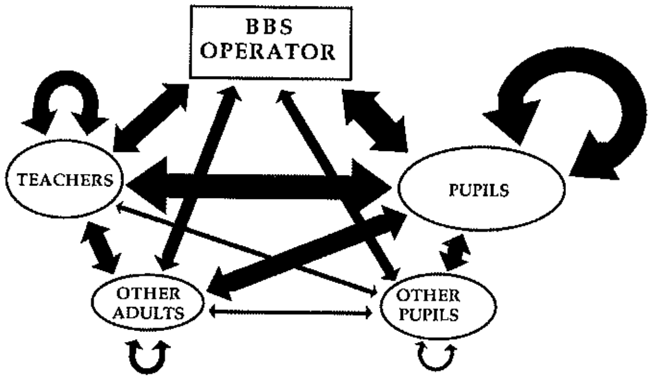


Fig. 2. Computer networking in a school for the deaf. The relative number of communications and users is represented by the size of the arrows and nodes.

we first introduced in Argentina computer networking in a primary school for deaf children (Battro and Denham, 1989). A BBS server allowed written communication (E-mail, conference, etc.) among pupils, teachers and other users. After the first 10 months of training, the number of electronic messages produced by the deaf children in this school increased from 30 to 600. Fig. 2 shows a picture of the relative significance of computer links and users in the network at a given time. We have since successfully experimented with radio-frequency communication (via radio-packet technology and radio low-orbit satellites) with the same deaf children. Radio-packet technology has the important advantage of free communication. Interesting experiences with radio-packet and the deaf are now in practice in Brazil (Fagundes, 1994). For a deaf person, the use of a telephone or a radio equipment is a formidable cognitive and cultural acquisition that became available for the first time in the eighties thanks to improved computer technology. In a sense, this "informational prosthesis" solves what we may call the *Graham Bell paradox*. Bell was a dedicated teacher of the deaf but is remembered as the genius who invented the telephone; unfortunately, the telephone excluded the deaf user for almost a century! The new communication path thus established is H/M/M/H, via telephone lines, direct networking or radio coupled with computers. We should mention that a good example of a telephone/computer interface is given by *Minitel* in France, a public system that completely changed communication among disabled users on a very large social scale.

Also for communication among deaf people or users with speech impairments and a common audience, a tele-computing link H/M/M/H can be established with the help of new informational devices that transform written text into voice and vice versa. This path "bypasses" the communication obstacles at the audio-speech H/H interface.

b) *M/H* obstacles can be classified essentially in two groups, *motor*: for example a motor disabled person cannot easily type on the computer keyboard, and *sensory*: a blind person cannot read the messages on the computer screen. A great number of devices are now available on the market that overcome most of the M/H barriers of sensory or motor origin. Recently, a remarkable catalogue in CD-ROM with a detailed list of more than 17,000 assistive products, recordings of 50 different speech synthesizers and special software for the disabled has been published. In this particular case, a blind person can even navigate through the CD-ROM and "hear" what he is actually browsing (Borden *et al.*, 1993).

M/H motor interfaces are being developed very rapidly. Among the most used in rehabilitation and special education are switches that replace the keyboard, touch screen devices, very large or miniature keyboards, "unicorns" that allow users to type with their head, devices that control the computer with gentle head movements or eye movements alone, voice recognition equipment that empowers speech to program and operate a computer. *M/H sensory interfaces* include voice synthesizers that transform computer characters on the screen into artificial speech and produce reasonable auditory information added to the standard visual format of written texts. Vision impaired people as well as young children during their learning process of reading and writing are frequent users of this friendly interface that can be connected to any modern personal computer.

c) *M/M links*. The old personal computers or PCs, have been transformed nowadays into truly *inter-personal computers*. An isolated computer will become a rarity in the near future, every machine will be interconnected, and so millions of people. The M/M link of Fig. 1 includes channels of communication via modem, telephone lines, satellites, optic fibers, radio-frequency, etc. A direct consequence of this link is the building of a "24 hour-society" on planet earth. This fact sheds a light of hope for the education and work of our brothers and sisters who are unable to move, speak, see, or learn at the proper pace required by modern living.

d) *H/E links*. When a man opens the door of his house, takes his car out of the garage, travels to his work, controls a huge machine with the help of levers and buttons, he is changing his environment, he is acting upon artificial or natural objects and systems. A common journey is a very complex path in the human environment. Unfortunately, many physical objects become obstacles to those people unable to use their hands, walk, hear, see, or who are mentally disabled to process the complexity of our everyday world. Computers may help to overcome these barriers. Robots can do the work of humans, houses can be transformed into "intelligent homes", navigational devices can help our movements. In our educational experience even disabled children may take advantage of the use of robots as instructional tools. For instance "Lego-Logo" interfaces provide micro-environments of great cognitive potential for them (Papert, 1993). Handicapped adults can move into intelligent houses with enough automation and control for independent living. This field is growing at an increasing speed and many kinds of *H/E interfaces* are now available or under study.

2. Amplification of cognitive performances

When using computers with the handicapped, we may perceive several dramatic changes in mental activities and specific behaviors. Essentially the subject becomes more motivated, emotionally satisfied and shows significant improvement in his cognitive performance. I will give two examples now.

Learning to write. Mentally handicapped persons can dramatically improve their writing abilities with the use of computers. Fig. 3 shows the learning curves of four mentally retarded subjects using a simple word-processor. After a latency period, that varies between 2 and 40 hours, the subject can manage to write 20 words (mostly as a copy of some printed text). Then a most remarkable "explosion" in the quality and quantity of writing takes place until a plateau is reached. The four subjects have different pathologies: S1. Age 18: mild mental retardation, S2. Age 20: borderline, spina bifida, quadriplegia, S3. Age 17: Down's syndrome (trisomy 21) and S4. Age 16: autism with visceral and somatic malformations. All show a similar acceleration of their written output (a similar slope of the learning curves) after a specific delay. It must be emphasized that standard special education and rehabilitation practice might never show a comparable development in writing abilities. A significant learning acceleration is the rule when computers and word processors are used as intellectual prostheses for reading and writing. This implies a fairly long time of training, a condition that is not always fulfilled in common practice.

A simple explanation of the permanent delays in the process of

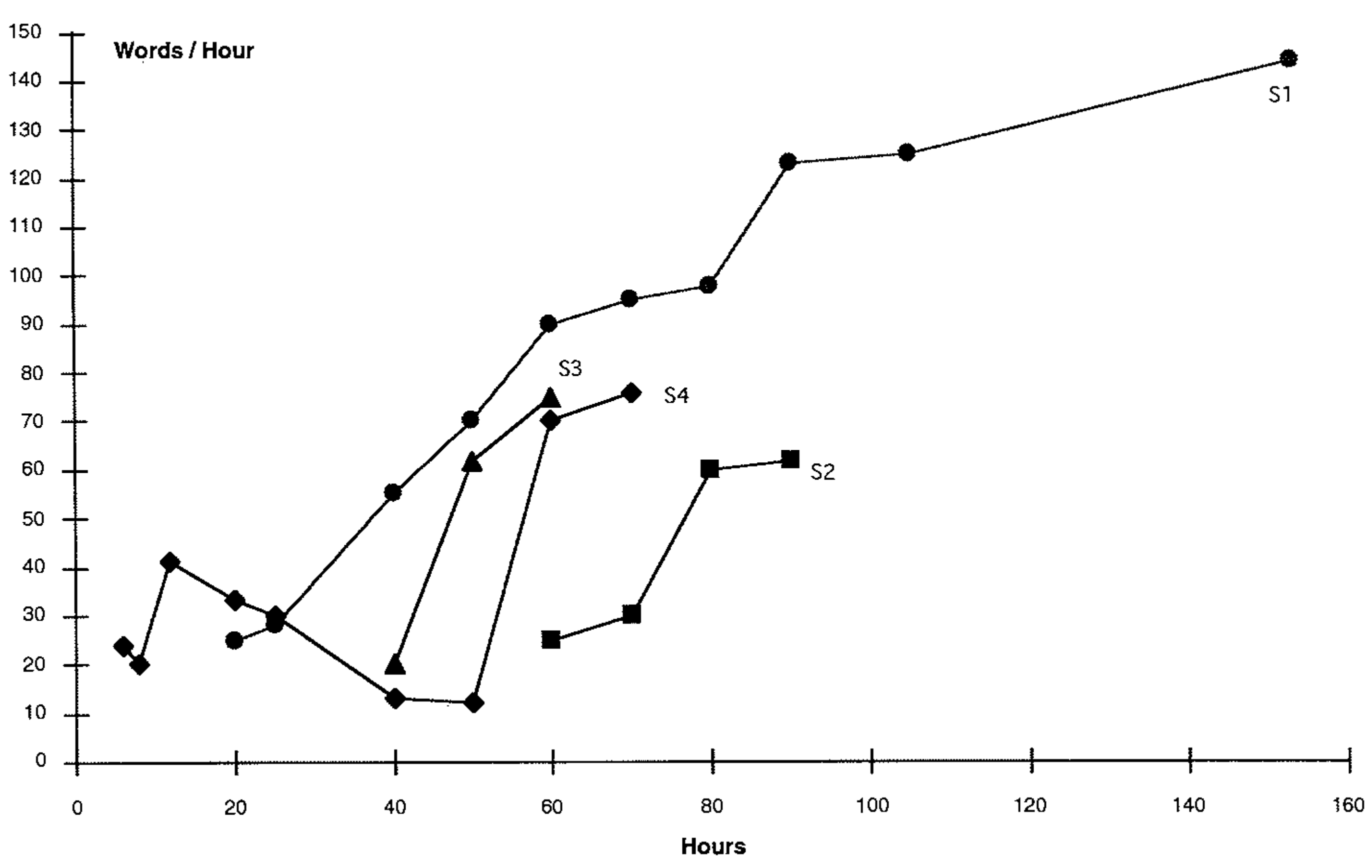


Fig. 3. Learning curves using computers to write, by four mentally retarded adolescents.

acquisition of new abilities by the mentally handicapped persons is that the brain needs a long period of activation to trigger a new way of doing things. For instance, standard learning offers poor intellectual feedback because of the very short attention span of the mentally retarded. Therefore, no long-term memory can be firmly established and learning of any intellectual valuable skill is impaired. Without the help of computers it seems that the training never reaches the threshold that opens new cognitive paths. But this remains to be proven with the help of the new non-invasive techniques like PET, nuclear magnetic resonance and the like.

The phenomenon of cognitive amplification becomes evident also during tele-communication via computers in the rehabilitation of aphasic persons. It is necessary to create new habits of electronic mail and tele-conference between the user and the instructor. A great deal of patience and time is needed but the results are encouraging. This was the case with an aphasic patient with a permanent deficit in speech and very limited motor abilities with his right hand. His handwriting was severely impaired although he has managed to become a good painter with his left hand. With the prosthetic help of the computer, he has shown a significant improvement of the length of the written messages (measured by the number of words per sentence). After 25 hours of training he improved from a poor average of 4 words per sentence to a maximum of 22. Also the number of one-word messages were reduced from 18 to 8. His impaired linguistic cortical networks were somehow "bypassed" with the help of the intellectual prosthesis at work. We assume that new cognitive paths became active in his brain.

Fifteen years of clinical observation of hundreds of disabled persons have shown us that perhaps a logarithmic scale should be used in order to measure the cognitive changes produced with the help of an intellectual prosthesis. As a matter of fact, many psychophysical laws are intrinsically logarithmic, and our consistent observation in clinical practice could be related to deep changes in brain activation that remain undetectable even with the most advanced technologies. The first changes are usually produced after 10 hours of practice (Figure 3 depicts the considerable delays in reaching a meager 20 word per hour level!). After 100 hours a plateau can be attained. But a new scaling factor is triggered and a new explosion in the written output can be observed after 1,000 hours of computer work, when many disabled students can attain a reliable rate (for those who think that this is too long, please recall the 50,000 hours that a chess master needs to work on chess boards to excel in international competitions). Usually the first 100 hours of computer practice are quite difficult and cannot be compressed a into few weeks. It takes almost 10 months for a mentally retarded student to assimilate the first 100 hours of computer work but the training can be accelerated after this first stage (mostly by home-work). This time-consuming and expensive training imply a firm chain of solidarity and care in order to foster higher levels of performance. But the results are certainly worth the effort. Several of our mentally retarded students have been socially and economically integrated after two years work with computers and became good copyists and reliable data-entering clerks. Automatic translation, dictionaries and glossaries improve the polish of the written text, error correction and final editing. This acquired ability engages the disabled person in a truly competitive intellectual activity, instead of being limited to do simple manual tasks that are underpaid or overprotected in special workshops. The worst possible attitude is to disengage a mentally disabled person from intellectual work! The intellectual prosthesis, on the contrary, enables the subject to close the gap with his own cognitive capabilities and also with the intellectual world.

Some *threshold* needs to be reached after a training period of computer practice. The clinical difficulty is to keep the subject active during the long and boring latency period, that can be long enough to inhibit further learning — because of the lack of satisfactory feedback — or even discourage the instructor because of continuing poor results. We may interpret the accelerated segment of the written performance and the correlated decrease of errors as a behavioral sign of the opening in the brain of new cognitive paths for writing. Some day, it can be safely predicted that new non-invasive techniques for brain imaging should add reliable data at the neuronal or even molecular level.

Learning to draw. The second example is related to the prosthetical applications of standard *Computer Assisted Design* (CAD) to students or professionals unable to use their hands to draw. The computer input can be given by speech using a voice recognition device with a microphone coupled with the CAD system. The remarkable feature in this case is that the informational system totally bypasses the area for motor control of arm, hand and fingers. The computer instructions to make a drawing can be vocally given and by this interface it has been possible, for instance, to train a quadriplegic architect to return to his professional practice after multiple sclerosis. The learning began with the voice control of a "screen turtle" using Papert's Logo language (Papert, 1981), first with simple instructions like *Forward* and *Right* to draw two dimensional pictures, then with the help of Reggini's few but powerful 3D Logo instructions (like *Pitch*, *Roll* and *Veer*) that can draw very complex figures in space (Reggini, 1985). Once this elementary stage with voice-controlled Logo was reached, a systematic training with voice-controlled CAD produced remarkable professional results in a few months (Battro, 1990).

3. Neuronal and cognitive paths

It is interesting to analyze from the neuro-cognitive point of view the differences between the prosthetic practice with speech recognition devices and the standard CAD drawings by instructions typed on the keyboard. It must be said that a good deal of training is always required for an architect or painter to bypass, by way of CAD, the common gestures of drawing and sketching. The difference between both situations, typed or vocally produced instructions for computer drawing, remains in the cognitive paths having been activated in the brain. It is known that many different areas are involved in speaking: Broca and Wernicke areas, some bilateral areas of the motor cortex, the supplementary motor cortex, the insular cortex and even the middle cerebellum are active during a simple speech utterance (Posner and Raichle, 1994). Other cortical areas however are involved when moving the hand, as in the act of drawing. The computer used as an intellectual prosthesis works as a bridge to pass from one mode (hand-drawing) to the other ("speech-drawing"). Fig 4 shows a simplified representation of this functional "cortical switch".

It is certain that the talents of draftsmen and painters are related to the cortical processing of spatial images (mostly in the right hemisphere) and to the motor control (by the left hemisphere) of our (right) hand (the inverse for pure left-handers). Therefore, many "agencies" in both hemispheres in Minsky's sense must be activated to produce the most simple drawing (Minsky, 1989). How many of them should then be active to produce the detailed layout and computerized images of a building! The finding reported here of an architect who makes high quality 2D and 3D technical drawings "with his speech" encourages us to explore the detail of cognitive specific paths of the brain when intellectual prostheses are in use.

"Drawing by hand" is a typical "analogical" ability. "Drawing by speech", instead, is digital, in the sense that uttered sentences are chains of discrete elements like words, syllables or phonemes. Space cognition includes space perception, space memory, space images and 2D and 3D representations. When drawing by hand, the cognitive paths must be opened to continuous motor control. This is not the case when the artist is drawing by speech with the aid of a computer, because of the digital nature of speech. There is perhaps a kind of a "navigational device" in these cases too, but of a different nature. For instance, it would be difficult if not impossible for an aphasic artist to make a computer drawing by computer using only speech instructions because of his specific linguistic impairment. On the contrary, the analogical cognitive path used to make a drawing by hand may remain intact. It is recognized that right-handed draftsmen can

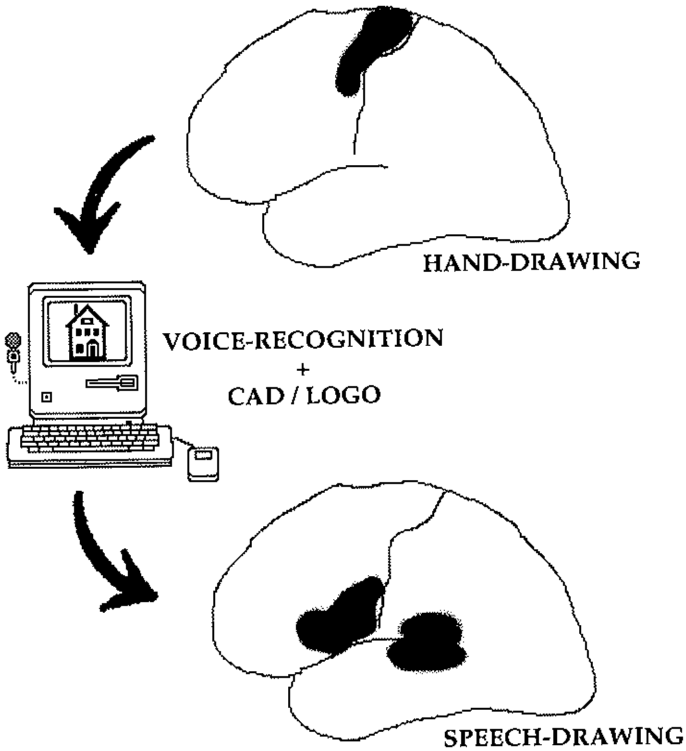


Fig. 4. The intellectual prosthesis can switch the brain from one cognitive mode (hand-drawing) to the other (speech-drawing)

produce the same pattern as left-handed ones, i.e. from the quality of the drawing it is quite impossible to distinguish between brain dominances. Moreover, clinical practice shows that many right-handed hemiplegic aphasics can perfectly relearn to use the left hand to draw and paint, even with severe (written or spoken) language impairments.

One thing is certain, before the introduction of computers no human mind has ever produced a drawing by speech instructions alone! (although in ancient Peru the Inca architects were called "the men that give orders aloud"). We cannot imagine Michelangelo making the plans of San Pietro by telling some assistants how to produce a detailed layout step by step, or painting the Sistina shouting to his aides to "put a red there" or "draw a line here". Nowadays, however, this exploit is technically possible. When a human operator uses his speech in order to draw, he activates a new neuronal network on his brain that was never before engaged in the act of drawing. Only a computer can enable this brain-switching. *This switching to a new cognitive path is the essence of any intellectual prosthesis.* In other words, intellectual prosthesis help the brain to perform some cognitive tasks that were normally processed by a quite different area of the cortex. Neuronal networks can be substituted or bypassed by new cognitive paths driven by computer instructions. This fact has very important consequences in clinical practice and special education.

Some day non-invasive techniques should show the different paths involved in cognitive growth too. In our experiments, four classical Piagetian tests were submitted to right-handed and left-handed children in a blind situation with only one hand at a time to manipulate the objects of each test (clay, sticks, marbles, etc.) (Battro, 1981). We could show that for "concrete operations", like class-inclusion or conservation of substance and for "formal operations", like probabilities, the right-handed children performed at higher Piagetian level when using their left-hand! The inverse phenomenon occurred with spatial concrete operations as in the conservation of length of a rod. The neuronal paths taken by both types of operational thinking were perfectly crossed:

a) For conservation of substance, class inclusion, probabilities:
Left-hand → right hemisphere → left (logical) hemisphere

b) For conservation of length:
Right-hand → left hemisphere → right (spatial) hemisphere.

This simple mapping of cognitive paths may inspire new experiments in neuro-cognition when the subject (young or old) is involved with computer prosthetic activities. The availability of non-invasive techniques now in place in many institutions will certainly help to understand the incredible plasticity of our brain. The computer permits the neuro-imaging of cognitive paths at the same time it provides the tools for the opening of new ones when used as intellectual prosthesis. This impressive technological feat enhances our moral and scientific responsibility. We have already the means to improve the quality of life of the disabled persons in many fields. We will need a surplus of wisdom to do it correctly and certainly a definite ethical involvement with those who need all our solidarity and love.

REFERENCES

Battro, A.M., 1981, «Hemispheric lateralization in the development of spatial and logical reasoning in left and right-handed children», *Archives de Psychologie*, Genève, 49, 83-90.

— 1986, *Computación y aprendizaje especial. Aplicaciones del lenguaje Logo en el tratamiento de niños discapacitados* (Buenos Aires: El Ateneo).

— Denham, P.J., 1989, *Discomunicaciones. Computación y niños sordos* (Buenos Aires: Fundación Navarro Viola, El Ateneo; 1994, Padova: Piccin).

— 1991, «Logo, talents et handicaps», *Logo et apprentissages*, eds., J.L. Gurtner, J. Retschitzki (Neuchâtel: Delachaux et Niestlé, 167-173).

Bruner, J., 1986, *Actual Minds, Possible Words* (Cambridge, MA: Harvard University Press) p. 72.

Fagundes, L., 1994. «Em busca de novos recursos para ajudar o desenvolvimento cognitivo de crianças surdas», *Revista em Aberto*. (Brasilia: Inep).

Minsky, M., 1987, *The Society of Mind* (New York: Simon & Schuster).

— Harrison, H., 1993, *The Turing Option* (New York: Warner Books).

Papert, S., Weir, S., 1978, «Information prosthetics for the handicapped», *Artificial Intelligence Memo* 496 (Cambridge, MA: MIT).

— 1993, *The Children's Machine. Rethinking School in the Age of the Computer* (New York: Basic Books).

Posner, M.I. and Raichle, M.E., 1994, *Images of Mind* (New York: Scientific American Library).

Reggini, H.C., 1985. *Ideas y formas. Explorando el espacio con Logo* (Buenos Aires: Ediciones Galápago).

Valente, J., 1983. *Creating a computer-based learning environment for physically handicapped children. Unpublished doctoral dissertation* (Department of Mechanical Engineering, Cambridge, MA: MIT).

Borden, P. A., Fatherly, S., Ford, K., and Vanderheiden, G.C., eds., 1993-1994, *Trace Resource-book. Assistive Technologies for Communication, Control and Computer Access* (Madison: Waisman Center, University of Wisconsin-Madison) also: *CO-NET CD. Cooperative Database Distribution Network for Assistive Technology* (Trace R&D Center, S-151 Waisman Center, 1500 Highland Avenue, Madison, WI 53705).

FUNCTIONAL BRAIN ACTIVATION STUDIES IN MAN WITH ADVANCED NEUROIMAGING TECHNIQUES: THE EXPERIMENTAL BASIS FOR COMPUTATIONAL THEORIES

BALÁZS GULYÁS *

INTRODUCTION

A most common question that lay people usually ask a neuroscientist is the following: "Is it true that we only use 5-10% of our brain?".

In fact, it is a most appropriate scientific question to ask to what extent the central nervous system in general, and the cerebral cortex in particular, is engaged during various brain operations. This question is even more appropriate these days when various mathematical models and computational theories of artificial intelligence aim at the interpretation and/or re-creation of a "hardware-independent software" of brain operations, derived from the working of biological nervous systems.

The main body of our knowledge of the functional organisation of the brain comes predominantly from experimental studies in animals and, to a lesser extent, from clinical obervations on neurological patients with well-circumscribed anatomical lesions resulting in functional deficits. Experimental studies in animals in recent decades used a versatile battery of techniques to explore the functioning brain. Of these methodologies, the ablation technique and single unit recording techniques contributed most significantly to our understanding of the brain. In ablation studies a part of the brain of the animal is eliminated and the locus of the anatomical deficit is correlated with the resulting functional loss in brain functions. In single cell recording techniques, the activities of single neurons are recorded with

* Laboratory for Brain Research and PET, Nobel Institute of Neurophysiology, Stockholm, Sweden.

intra- or extracellularly placed microelectrodes during perceptual or motor activities of the animal, and the recorded activity pattern is correlated with the characteristics of the perceptual or motor function. These techniques have provided us with a basic understanding of the functional specialisation in the cerebral cortex. Due to ethical limitations, however, these techniques cannot be used routinely in humans.

During the past years the advance of functional neuroimaging techniques has made the exploration of the working human brain possible. Whereas certain neuroimaging techniques, such as magnetic resonance imaging (MRI), provide us with high spatial resolution anatomical images of the human brain, other techniques, for instance positron emission tomography (PET), yield us information about the metabolic changes during different brain functions. The complementary use of these two techniques, in turn, made it possible for us to co-localise metabolic changes with anatomic structures in the functioning human brain.

In the present paper our aim is: (1) to overview data obtained in some of our earlier experiments with functional neuroimaging techniques during brain activation, and, on the basis of that, (2) to come up with realistic estimates of some quantitative features of cortical fields (their size and location, the number of neurons in them, etc.) involved in different perceptual and cognitive brain operations.

Methodological considerations

Brain activation

The term "brain activation" covers those experimental conditions wherein the human brain is engaged by sensory stimulation, motor activation, or cognitive tasks (Roland, 1993). On the one hand, the experimental tasks can be elicited externally, i.e. the stimuli or the instructions are externally generated (stimulus –> brain [–> response]) and the brain's "executive activities" are the object of our scientific exploration. On the other hand, the stimuli can be "self-generated" by the brain, e.g. the brain itself internally generates stimulus-conditions (e.g. by recalling from visual memory images or by generating words or imagined movements). In that case, both the generation of tasks and their execution by the brain are under our scrutiny.

Under experimental conditions, the psychophysical performance levels of the experimental subjects (or in cases for which human psychophysics does not have proper measurement techniques, such as silent thinking, the "behavioural signatures" of the subjects; Kosslyn and Ochsner, 1994), as well as a number of characteristic physiological parameters (e.g. EEG, eye movement frequencies, etc.) are strictly controlled and correlated with both the stimulus parameters and the results of the imaging techniques (Gulyás and Roland, 1994 a,b,c).

The most widely used stimulus paradigms are usually based upon the principle of the *subtraction technique* and, therefore, are differential: two stimulation conditions (an 'experimental condition' or 'test condition' and a 'reference condition'; an experimental task versus a reference task) are paired with each other; the experimental condition differs only in one feature from the reference condition. The brain regions related to the processing and analysis of that specific feature are in the focus of our exploration. Since with the exception of one given stimulus feature the tasks are identical, by subtracting the images belonging to the reference condition from those belonging to the experimental condition we create subtraction images (or difference images) wherein only those brain regions are highlighted which are activated by the specific stimulus feature we are interested in. A less frequently used reference state is the rest condition. In this case the stimulation conditions may differ from the rest condition in several features, and therefore the resulting subtraction images may contain brain regions involved in a relatively large number of neuronal operations.

Neuroimaging techniques

Our experiments are based upon a complementary imaging with two imaging techniques: magnetic resonance (MR) scans provide us with anatomically precise images of the individual brains, whereas positron emission tomography (PET) images provide us with images of regional cerebral blood flow (rCBF) changes during the various brain activation conditions. As changes in regional cerebral blood flow closely correspond to changes in regional cerebral metabolism, in rCBF we have a faithful indicator of regional cerebral metabolism.

A high resolution MR scan was made of each brain. The head fixation in the MR and PET scans were identical (Bergström *et al.*, 1981) so that the corresponding MR and PET images were superimposable. The PET scanning was made by a Scanditronix PC-384 (Litton *et al.*, 1984) or a Scanditronix PC2048-15B positron emission tomograph (Litton *et al.*, 1980; Evans *et al.*, 1991). The spatial resolution of the former camera was 7.8 mm, the latter one was 5 mm in-plane, numbers of slices and interslice distances were 8 and 13.5 mm, and 15 and 6.75 mm, respectively. In the earlier experiments $^{11}CH_3F$, in the latter experiments ^{15}O-butanol, a freely diffusible flow tracer, was used to measure rCBF. (The Scanditronix 384 camera and $^{11}CH_3F$ were used in the visual learning, recognition, and recall of coloured patterns experiments [Roland and Gulyás, 1994b], whereas in all the other experiments the other camera and tracer were used). During the experiments EEG, EOG (saccadic eye movements), arterial radiotracer concentrations, PaO_2 and $PaCO_2$ levels were continuously monitored. The response latencies and performance levels of the subjects were measured on-line. Differences in rCBF between the tasks resulting from differences in $PaCO_2$ were corrected (Olesen, 1971). Other aspects of the method, including the adequacy of resolution parameters to the present studies, were described earlier (Roland *et al.*, 1987, 1993; Gulyás and Roland, 1994 a,b,c; Roland and Gulyás 1994 a,b).

Image analysis

The MR and PET images were transferred into a computerised brain atlas system (CBA) (Bohm *et al.*, 1983, 1986; Greitz *et al.*, 1991; Seitz *et al.*, 1991) and transformed into standard size and shape (Figure 1). The contours of the CBA were adjusted to the MR tomograms of the individual brains. Thereafter the transformation parameters were used to transform the individual PET images into anatomically standardised PET images. Individual subtraction images (ΔIMA 'test-reference', e.g. ΔIMA_1, ΔIMA_2, etc.) were created (Figure 1), which were then averaged across the whole subject population to provide averaged subtraction (Ave-IMA = $\sum_{n=1}^{N} \Delta IMA/n$; where n = number of subjects) images, as well as variance images (Var-IMA) and descriptive Student's-t images ($\sum_{n=1}^{N} \Delta IMA \sqrt{variance}/\sqrt{n}$) (Figure 2). Volume elements (voxels) in the image (volume of a voxel: 44 mm^3 having t values ≥ 2.26 were considered to be clustered if they were attached by side, edge or corner. On the basis of an analysis of false positive clusters with high t-values, it was decided to reject the hypothesis that all clusters of size 8 and above belong to the distribution of false positives (Roland *et al.*, 1993). The probability of finding *one* false positive cluster of size 8 and above in the whole brain was ≤ 0.5, whereas the probability of finding *two* or *three* false positive clusters is ≤ 0.13 and ≤ 0.008, respectively. (Raising the acceptable cluster size to 9 or more in order to eliminate the false positives could lead to false negatives, i.e. to the omission of positively identified activated fields). The descriptive t-image was thresholded to comprise only clusters of voxel size 8 or more with voxel values t ≥ 2.26. The remaning voxel values were set to zero. The resulting image is called a cluster image (Clu-IMA; Figure 2). In a cluster image, all clusters of size 8

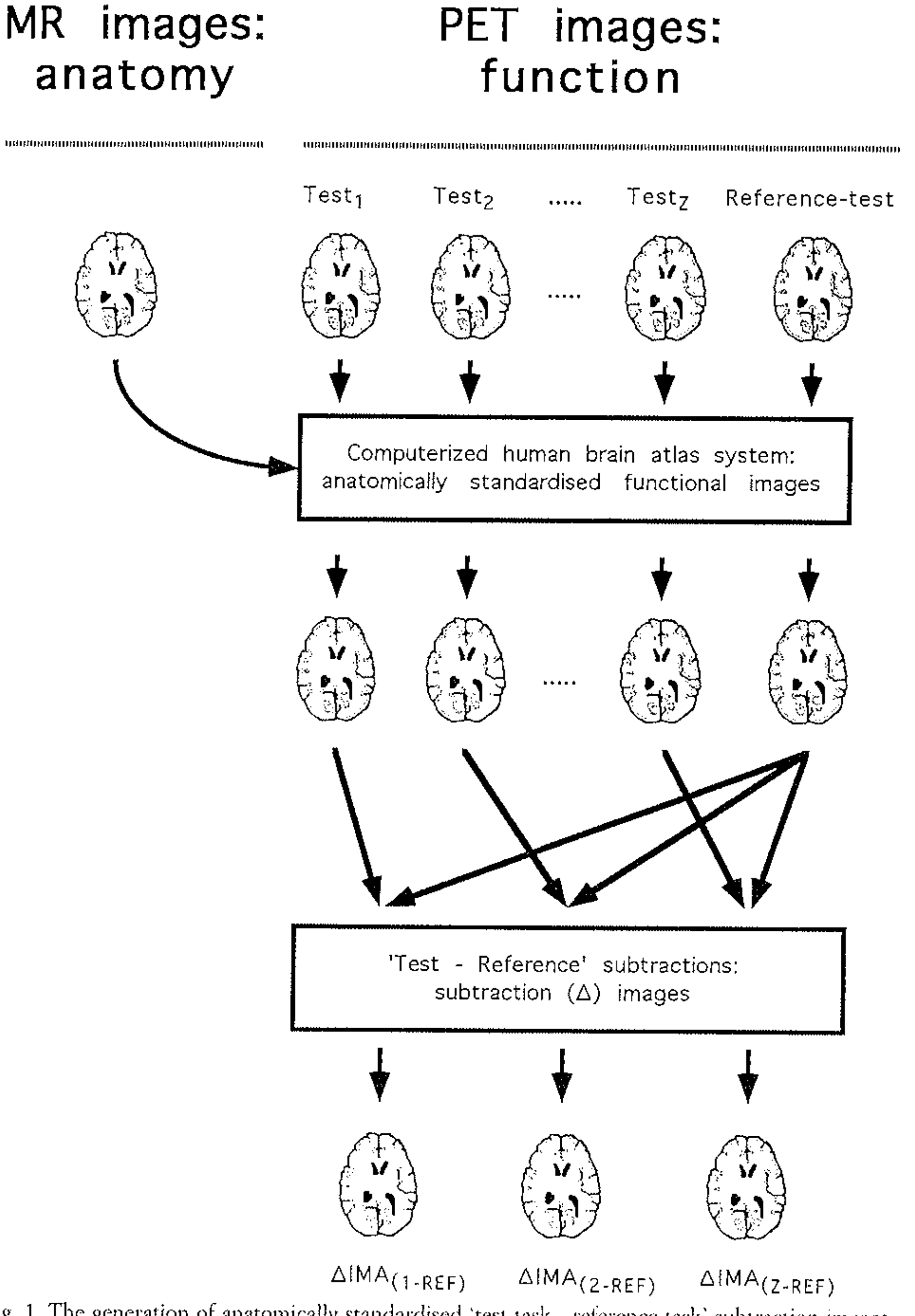


Fig. 1. The generation of anatomically standardised 'test task - reference task' subtraction images. For details, see section on Image analysis.

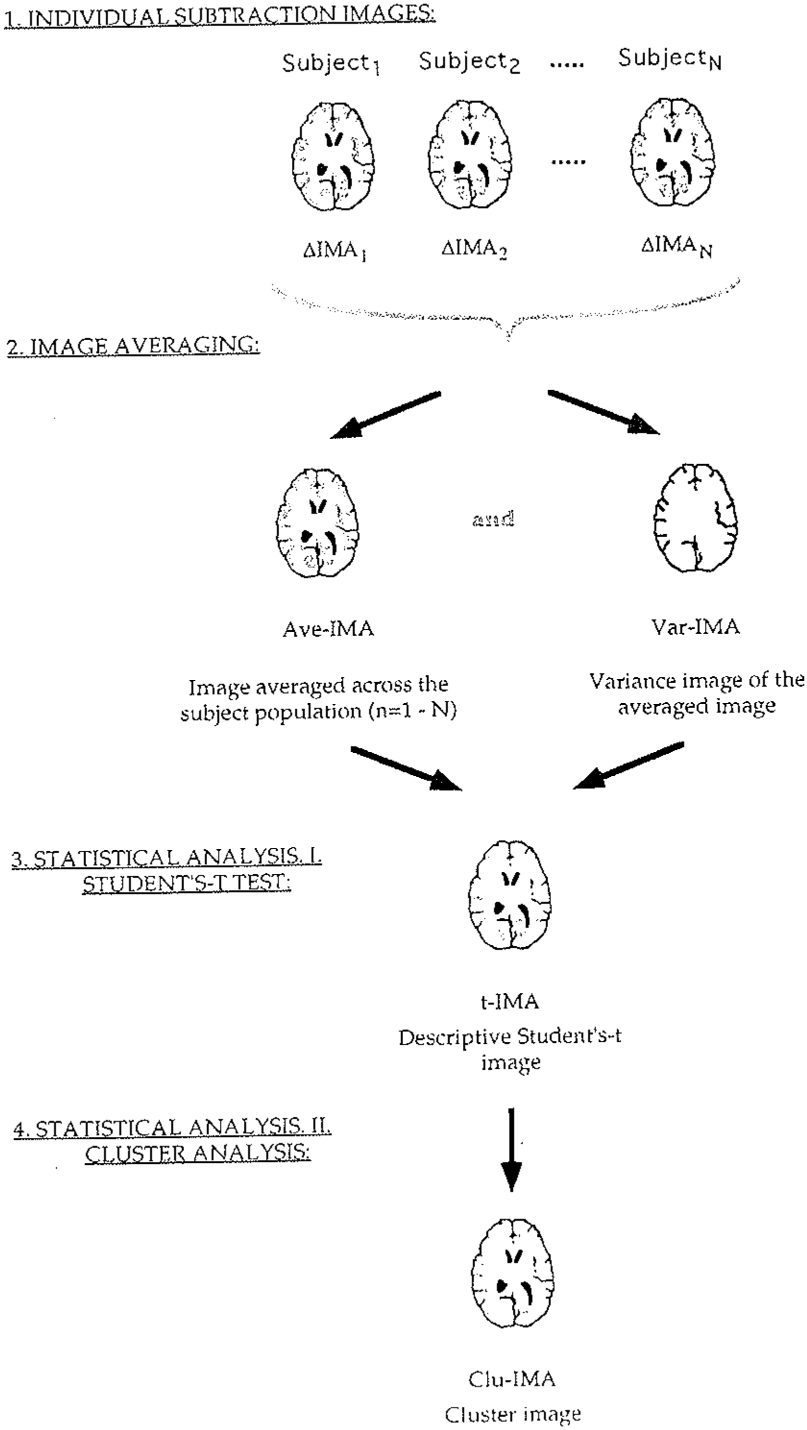


Fig. 2. Statistical analysis of functional images. For details, see section on Image analysis.

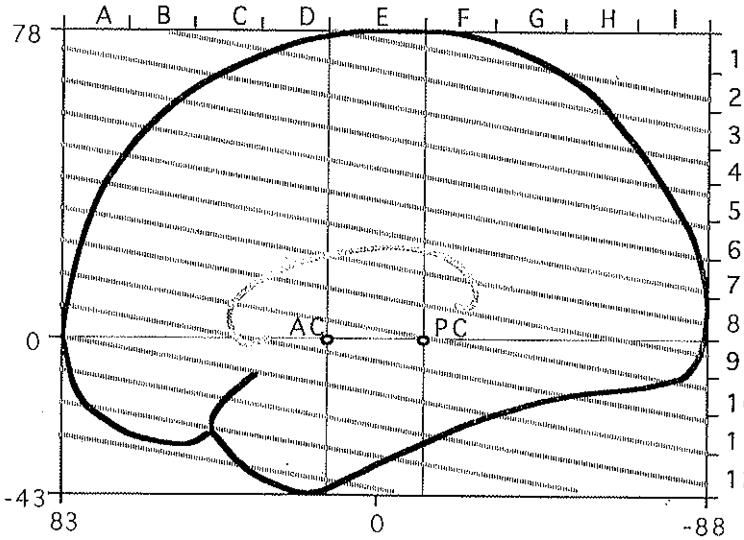


Fig. 3. The Talairach proportional stereotactic system (Talairach *et al.*, 1967) and the location of the horizontal MR and PET image slices used in the study. The brain contours are shown with solid lines, the corpus callosum is in grey. The dashed oblique lines correspond to the 15 horizontal image slices of the MR and PET scans. AC: anterior commissure, PC: posterior commissure. Lettering and numbering along the top and right sides represent the original Talairach denotions, figures along the left and bottom sides correspond to mm values.

and above are shown as regions of significantly changed rCBF. The size of cortical fields was determined on the basis of the cluster image. In order to express the data in a stereotactically standard image space, we use the Talairach-convention (Talairach *et al.*, 1967) and express the extension of the activated cortical fields in a "collapsed" sagittal plane of the Talairach stereotactic coordinate system (Figure 3).

Visual tasks used

The tasks to which we refer hereafter were selected from a larger series of experiments we have performed in the past years in order to map the functional organisation of the human brain. As the experiments were described in detail, only a brief description is given here. The data are collected from six experimental series. In each series at least ten young healthy volunteers participated.

In one experimental series (Gulyás and Roland, 1994 a,b), we explored the cortical fields involved in the discrimination of colour, form, and disparity. The subjects had to perform simple discrimination tasks related to the visual sub-modalities of colour, form, and disparity as well as a reference task. In another series (Gulyás and Roland, 1994c) the subjects had to discriminate in separate task-conditions spatial frequencies and orientations of gratings; whereas in the reference tasks, gratings had to be discriminated from random dot patterns. In another two series (Gulyás *et al.*, 1994 a,b) the task was form discrimination. The visual forms in the various tasks were, however, created from different visual cues including colour, form, disparity, and luminance. Each form discrimination task was paired with a reference task in which discrimination had to be made along the specific visual cues, i.e. colour, form, disparity, luminance, but not along the form component. In all the above tasks the test paradigms in the experimental and reference tasks differed only in one stimulus feature along which the discrimination had to be made.

In the remaining two tasks, on the other hand, the reference condition was rest and the experimental conditions were compared with the rest condition. In one series (Roland and Gulyás, 1994b) the subjects had to learn coloured geometrical patterns, then recall them from memory and, finally, recognise the learnt patterns from a series on unlearnt patterns of similar type. In another series (Gulyás *et al.*, 1992; Roland and Gulyás, 1994a), in the first task the subjects had to internally rehearse (with closed eyes and not saying anything loudly) the letters of the alphabet in due order followed by the first verse of the Hungarian national anthem. In the second and third tasks the subjects had to visualise internally, with closed eyes, the capital letters of the alphabet and that of the text of the Hungarian national anthem.

Stereological measurements

There is a relatively large variety in the literature regarding the size of the cerebral cortex and the number of cells in it. Since the first scientific determination of the cortical volume by Jensen (1875), the values obtained by various studies in the literature range between 310 and 711 cm^3 (Pakkenberg, 1966). Recent estimates using modern stereological techniques give average values usually above 500 cm^3 (Pakkenberg, 1966: 570 cm^3; Paul, 1971: 584 cm^3; Regeur and Pakkenberg, 1989: 549 cm^3; Badsberg Jensen and Pakkenberg, 1993: 486 cm^3). For our calculations, we have chosen Regeur and Pakkenberg's (1989) value (549 cm^3), since this study was specifically focusing on the measurement of cortical volume and used the most advanced techniques available at present.

As for the number of cortical cells in man, the variety of data is even larger in the literature. Since the first determination of cortical cell numbers (Donaldson, 1895), the numbers vary between 1.2×10^9 and 23.2×10^9 mm^{-3}. This latter value seems to be the most acceptable one, as this value was obtained with the use of advanced stereological techniques (e.g. the dichotom) with appropriate corrections (Badsberg Jensen and Pakkenberg, 1993) and other recent approaches give values close to the above figure (Pakkenberg, 1993). Therefore, in the further calculations, we use value of 23.2×10^9 mm^{-3}.

Cortical fields activated during different visual tasks

A great number of physiological studies in primates have indicated in the past that the various features of a complex visual scene, such as colour, form, motion, or disparity, in lower levels of the visual system are processed and analysed by anatomically and physiologically segregated "processing streams" or functional pathways, before an integrated and unified visual percept is generated in the brain (Van Essen and Maunsell, 1983; Livingstone and Hubel, 1987; De Yoe and Van Essen, 1988). The exploration of these functional pathways in primates has been a main target of visual physiology in the past decades. On the basis of these investigations, a wealth of evidence suggested that the functional organisation of the human visual system may be similar to that of other primates (Gulyás *et al.*, 1993).

Using functional neuroimaging techniques, we set out to explore *in the human brain* the cortical fields engaged by the analysis of fundamental visual cues or sub-modalities, including colour, form, orientation, spatial frequency, and binocular disparity (Gulyás and Roland, 1994 a,b,c,). The location of the activated cortical fields are shown in Figure 4 and the parametric description of the fields (number, size, maximum number of cells involved) is shown in Table 1. It is worth noting here that in this case the experimental (or test) conditions were compared with a reference condition which differed from the test condition in only one aspect, namely that it did not contain the stimulus feature (i.e., the visual cue) which was in the focus of our exploration (such as colour, form, disparity, etc.). Consequently, the fields shown in the subtraction images are those involved in the processing and analysis of the given visual cues.

A peculiar aspect of visual perception is that the same visual attribute, such as a visual form, can be generated on the basis of various visual cues, such as luminance, colour, motion, or disparity. With the purpose of exploring the extent to which different cortical functional pathways are

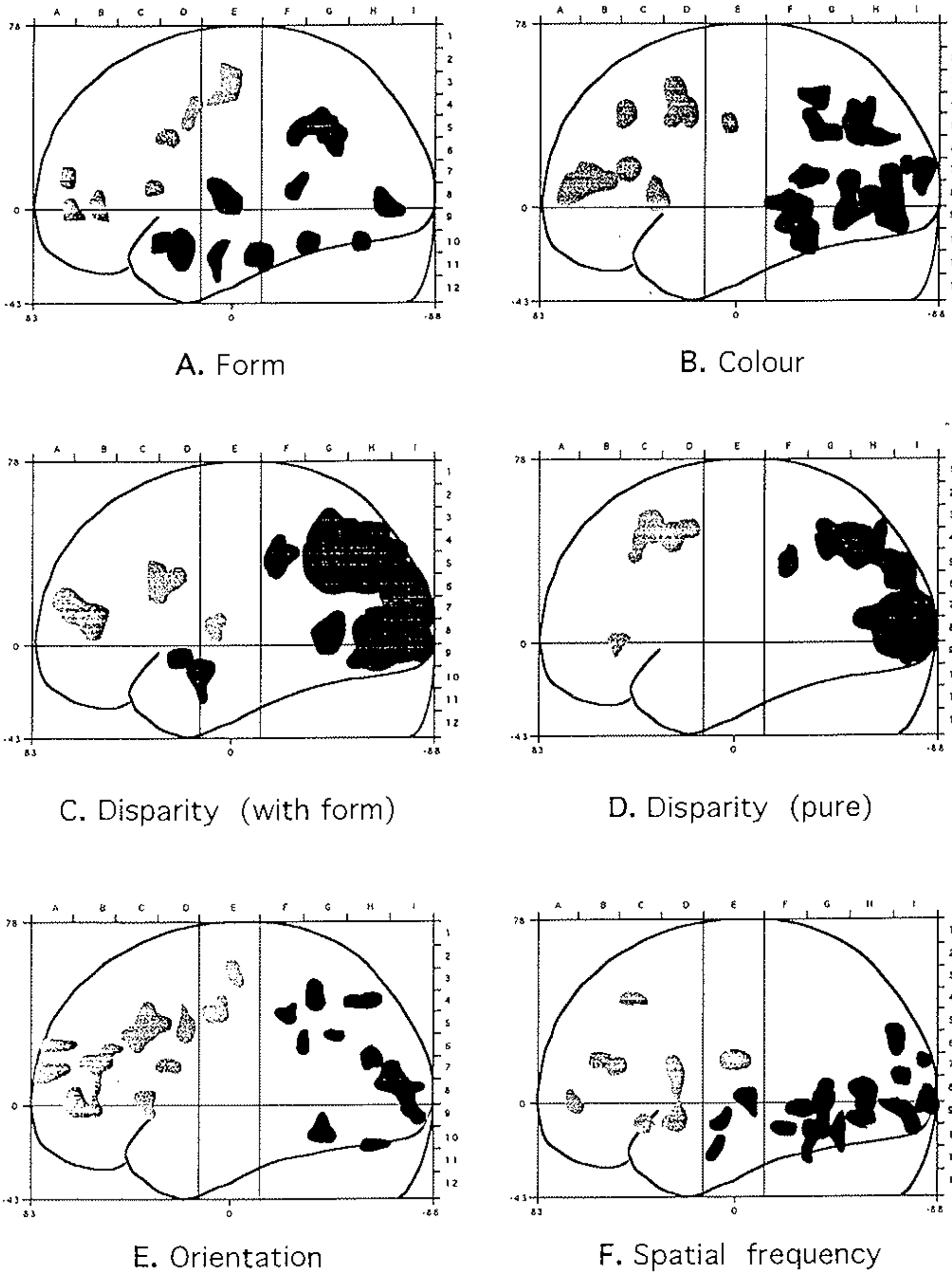


Fig. 4. Cortical fields activated during the visual discrimination experiments. The location and extension of the fields are expressed in the standard Talairach coordinate system, "collapsed" parasagittal plane. Apparent overlaps in the figures do not necessarily indicate field overlaps in reality. Fields constituting the "visual cortex" are in black, extravisual fields are in grey in the present and the following figures.

Table 1

Task	Total size of cortical fields (mm^3)	Number of cortical fields	Smallest and largest field size	% of cerebral cortex *	Proportion of visual cortex **	Number of cells ***	Reference
A. Subtraction type: Test - Control (Test = Control + one additional visual cue							
Discrimination of visual sub-modalities:							
Form discrimination	9,636	15	440 - 1,012	1.76	56%	0.41×10^9	Gulyás and Roland, 1994a
Orientation discrimination	18,738	25	352 - 1,980	3.41	41%	0.39×10^9	Gulyás and Roland, 1994c
Spatial frequency discrimination	9,152	19	352 - 924	1.67	76%	0.39×10^9	Gulyás and Roland, 1994c
Colour discrimination	18,620	27	440 - 1,716	3.46	54%	0.80×10^9	Gulyás and Roland, 1994a
Disparity (pure) discrimination	19,492	14	352 - 4,356	3.55	86%	0.82×10^9	Gulyás and Roland, 1994b
Disparity (with form) discrimination	32,604	21	440 - 6,248	5.94	76%	1.38×10^9	Gulyás and Roland, 1994a
Discrimination of a visual attribute based on different visual sub-modalities:							
Form from colour discrimination	5,280	12	352 - 528	0.96	50%	0.22×10^9	Gulyás *et al.*, 1994a
Form from motion discrimination	6,908	15	352 - 616	1.26	57%	0.29×10^9	Gulyás *et al.*, 1994a
Form from disparity discrimination	8,756	19	352 - 1,232	1.59	61%	0.37×10^9	Gulyás *et al.*, 1994b
Form from luminance discrimination	1,496	4	352 - 396	0.27	100%	0.06×10^9	Gulyás *et al.*, 1994b
B. Subtraction type: Test - Rest							
Direct visual perception:							
Visual learning of large coloured patterns	130,768	49	460 - 11,710	23.82	85%	5.53×10^9	Roland and Gulyás, 1994a
Recognition of large coloured patterns	95,455	43	460 - 4,730	17.38	80%	4.03×10^9	Roland and Gulyás, 1994a
Visual imagery:							
Visual recall of large coloured patterns	22,660	14	680 - 3,250	4.13	29%	0.96×10^9	Roland and Gulyás, 1994a
Visualization of letters of the alphabet	6,468	13	352 - 704	1.18	41%	0.27×10^9	Gulyás *et al.*, 1992
Visualization of letters of the anthem	5,632	11	352 - 616	1.02	47%	0.24×10^9	Gulyás *et al.*, 1992
Internal listing:							
Silent rehearsal of anthem or alphabet	1,672	5	352 - 440	0.30	32%	0.07×10^9	Gulyás *et al.*, 1992

* Total volume: 549.000 mm^3 (Regeur and Pakkenberg, 1989); ** As tentatively described by Roland and Gulyás, 1994b; *** Total number: 23.2×10^9 mm^{-3} (Badsberg Jensen and Pakkenberg, 1993).

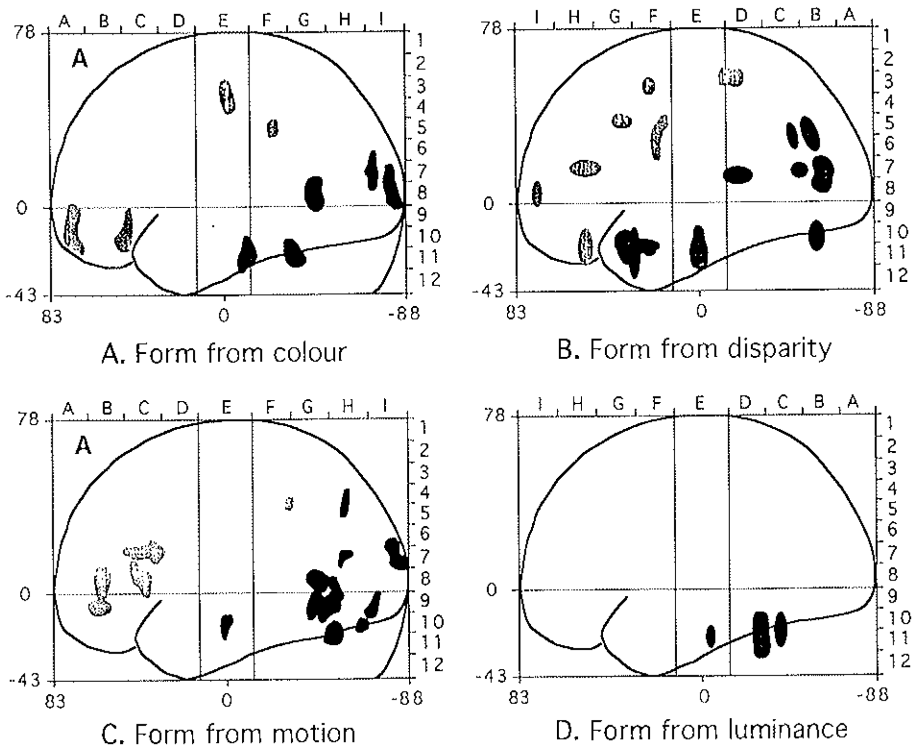


Fig. 5. Cortical fields activated in the form discrimination experiments.

involved in processing and analysing different types of information that yield the very same perceptual entity, visual form, we mapped the anatomical structures in the human brain participating in the discrimination of visual forms mediated either by motion, colour, disparity, and luminance cues (Gulyás *et al.*, 1994 a,b). Surprisingly, we could demonstrate that non-overlapping disparate networks of cortical fields are activated when the same visual forms are generated by different input cues. Not only the location of the fields involved were different, but their total volume differed, also, indicating that depending upon the input cue, the "computational load" of generating a visual form percept varies (Figure 5 and Table 1).

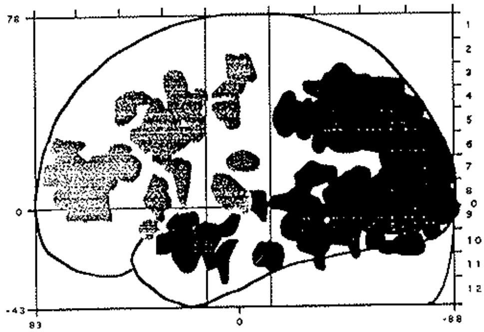
Naturally, in our everyday visual experience the human visual system is dealing with complex visual scenes which at the same time contain all or most types of the different visual cues. In order to explore the anatomical structures in the brain underlying comple high level visual perceptual processes, we mapped the cortical fields during the learning of complex coloured geometrical patterns and the recognising of the learnt patterns from a series of similar but unfamiliar patters (Roland and Gulyás, 1994b). The reference condition in these tests was rest. As shown in Figure 6B,C and Table 1, both visual learning and recognition processes mobilised a large part of the cerebral cortex.

It is worth noting that when we expressed the cortical fields active during the discrimination of various visual sub-modalities (Figure 4) in one stereotactical frame, the resulting image (Figure 6A) shows remarkable similarities, especially as far as the "visual cortex" is concerned, with the images displaying the cortical activation patterns during visual learning and recognition (Figure 6B,C). This has been expected, in fact, as the learning and recognition of complex visual patterns requires the recruitment of those cortical areas which participate in the processing and analysis of individual visual sub-modalities or cues, including colour, orientation, spatial frequency, form, etc.

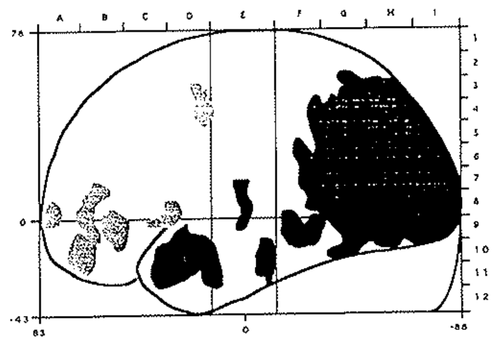
A higher form of conscious visual image generation is visual imagery, during which visual representations are generated by will in the absence of direct visual input. Visual imagery involves different operations, including the recall of images from visual memories, the maintainance of the recalled visual images, and various operations on the imagined images, such as changes in perspective, zooming, rotations, etc. (Roland and Gulyás, 1994a). We have investigated the cortical structures in man participating in visual imagery processes, including the recall of large coloured visual patterns (those which the subject had to "learn" in an earlier test; Roland and Gulyás, 1994b), the recall of the letters of the alphabet and those of the Hungarian national anthem (Gulyás *et al.*, 1992; Roland and Gulyás, 1994a). In comparison with the visual recall tasks, the cortical structures engaged by the silent internal rehearsal of the alphabet and the national anthem were also investigated (Gulyás *et al.*, 1992). The cortical areas involved in the imagery and rehearsal tasks are displayed in Figure 7 and the parameters of the activated fields are shown in Table 1. In the latter tasks, just as in the learning and recognition tasks, the reference condition was a rest condition.

THE CONCERTED ACTION OF CORTICAL FIELDS

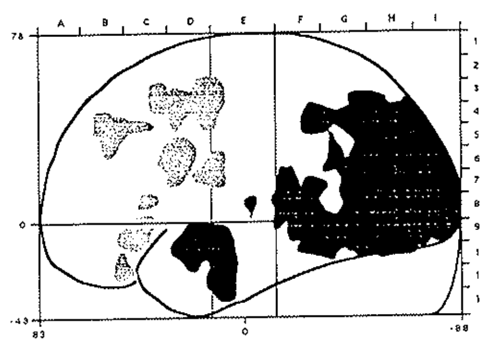
As shown in Table 1, the human brain uses a relatively large number of cortical fields even in seemingly simple perceptual or cognitive tasks. In the 'test - control' subtraction types, the number of fields ranged between 4 and 27, with a median value of 17. The size of the cerebral cortex involved



A. Visual perception



B. Visual learning



C. Visual recognition

Fig. 6. A. The summary picture of the fields activated during the visual discrimination experiments of colour, form, disparity, spatial frequency, and orientation. B. Cortical fields activated during visual learning of complex coloured geometrical patterns. C. Cortical fields activated by the recognition of learnt patterns.

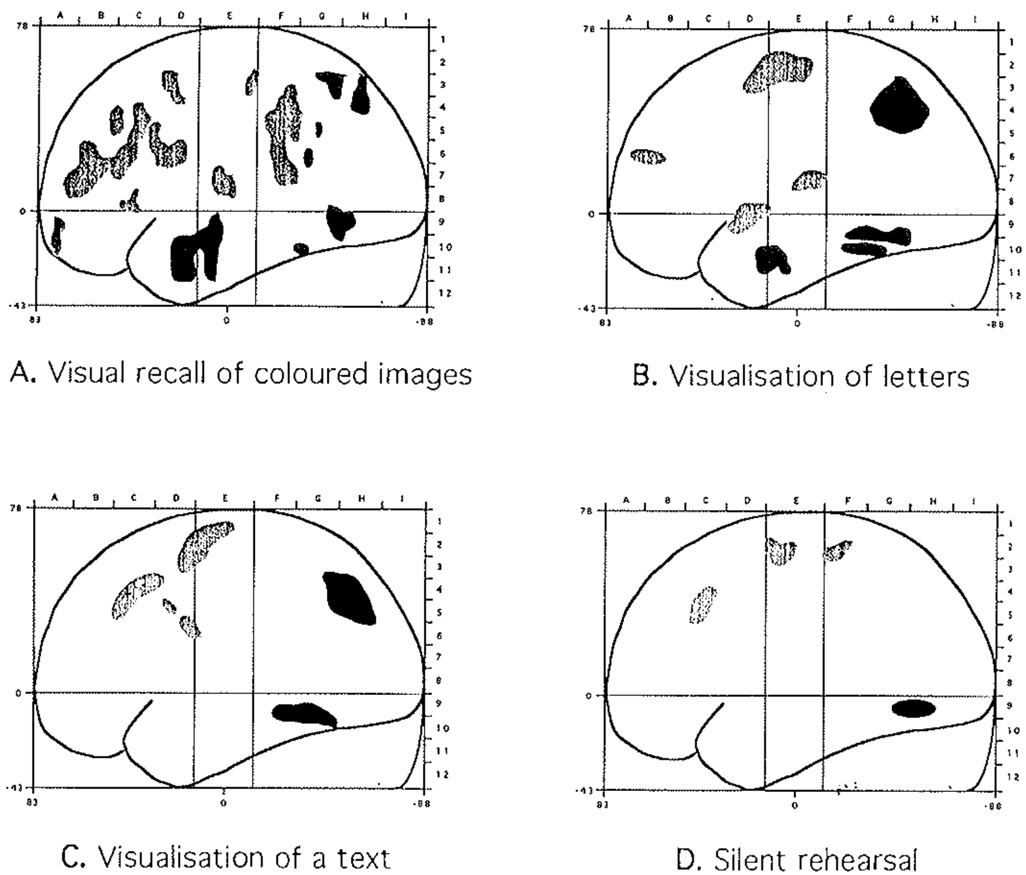


Fig. 7. Cortical fields activated by visual imagery and silent rehearsal.

in the tasks also varied remarkably, between 0.27% and 5.94%, with a median value of 1.72%. This twenty-fold variation of the cortical fields involved in the various tasks indicates that the brain is capable to recruit a large number and variety of cortical fields depending upon the actual task to be performed. The number of cortical cells involved in the fields varied between 6×10^7 and 1.38×10^9, with 3.9×10^8 as median value. In the 'test - rest' subtraction types, the number of fields varied between 4 and 49, corresponding to between 0.30% and 23.83% of total cortical volume involved in the tasks, and between 7×10^7 and 5.53×10^9 maximum number of cells in the fields.

Since in the 'test - control' conditions the test task and control task differ in only one stimulus (or task) feature, in this case the cortical fields represent those regions in the cortex which participate in the processing and analysis of one specific stimulus feature. In the 'test - rest' conditions, however, the cortical fields present after the subtraction, involve all those cortical regions which become active during the experimental task as compared to the rest state. It may be surprising in the light of this reasoning that some of the tasks in the 'test - rest' condition activate fewer fields and, consequently, smaller cortical volume, than most tasks in the 'test - control' condition (cfr. Table 1). It is not surprising, however, in the light of those studies which indicate that one can hardly avoid free-floating thinking activities in awake resting situations (Singer, 1993), and in this respect our silent rehearsal or visualisation tasks may only slightly differ from uncontrolled streams of thoughts during rest. For this reason, several authors have in the past years questioned the appropriateness of rest as a control state in brain activation experiments (Roland, 1993).

Whereas the cortical fields' volume in the above tasks varied between relatively large values (352 - 11,710 mm^3), their distribution indicated that the larger fields may, in fact, be "composite fields", i.e. the aggregates of a number of smaller fields which with the present methodology cannot be separated into smaller "constituent fields". The highest number of fields were found between 400 and 500 mm^3. Such a cortical field involves about 17-21 x 10^6 cells and it may represent a functional unit of the neocortex, as suggested by Roland (1985, 1993).

It is important to realise in the present context, that even relatively simple perceptual or cognitive operations require the concerted action of a relatively large number of cortical fields. The activated fields form a cortical network and the networks underlying various perceptual or cognitive tasks may be formed from different cortical fields, depending upon the nature of the task. As demonstrated with the discrimination tasks of single visual submodalities (Figure 4 and Figure 6A) and the complex visual tasks using a multitude of sub-modalities (Figure 6B,C), the activated visual fields in large parts overlap. Following this reasoning, it can be predicted that using a sufficiently large variety of perceptual, motor or cognitive tasks, activated fields will be present in the whole cerebral cortex, i.e., we indeed need our whole brain to be able to perform a large variety of tasks, even if the total size of activated fields remain relatively limited in one given task.

Caveats: sources of uncertainties

The above data on the number of cells represent the maximum number of neurons involved in those cortical fields activated during the different visual tasks. The figures present in Table 1 can, in fact, be *both* underestimations and overestimations of the number of neurons participating in a given visual task.

On the one hand, the above figures can be *understimations*. Since in all these experiments we used the subtraction technique, the above figures stand for only those cortical fields which participate in the processing and analysis of information related to some stimulus features involved in the task (e.g. in the case of the perception of a visual scene) or one single stimulus feature (e.g. in the case of colour discrimination task) which "remain(s)" after the subtraction. Let the state of comparison be either the rest condition or a given test condition, in both cases the brain is active and a large number of cells are functioning. In this respect the cell numbers shown in Table 1 are only referring to the maximum number of cells in the cortical fields engaged by the specific components of the tasks.

On the other hand, the figures can be *overestimated*, since they represent the maximum number of cells in the active cortical fields. It is reasonable to believe that only a proportion of the cells are active in a cortical region under physiological circumstances. However, with the present methodology we cannot calculate the exact proportion of the cells active in a cortical field, neither can we estimate the intensity (e.g. in form of action potential frequency) of their activation.

Conclusion

Advanced neuroimaging studies can provide us with reasonable estimates on the physiological and anatomical characteristics of cortical fields active during cognitive, perceptual, or motor activities of the brain. In our data, obtained in visual perceptual and imagery tasks, the total volume of cerebral cortical fields active during the various tasks varied between 0.27% and 23.82% of cortical grey matter. Whereas these volumes correspond to between 0.06×10^9 and 5.53×10^9 maximal number of cells, it is reasonable to believe that the actual number of cells participating in the processing and analysis of stimulus features investigated in the present experiments is below the total number of cells involved in the fields. However, since in the present experiments the subtraction technique was used, i.e. a rest or a reference condition was subtracted from the experimental condition, in order to estimate the total amount of neuronal elements active in brain operations, one should also take into consideration the number of those cells active in rest or reference states. In order to be able to take into consideration these facts as well, further studies are required which would provide us with more precise and careful estimates on the actual number of neurons active during brain operations.

Acknowledgements

I am thankful to Per E. Roland, my sparring partner, for our everlasting discussions about these issues. This work has been supported by grants from the Human Frontier Program Organization, the Wenner-Gren Foundation, and the Karolinska Institute.

REFERENCES

Badsberg Jensen, G. and Pakkenberg, B., Do alcoholics drink their neurons away? *Lancet*, 342 (1993): 1201-1204.

Bergström, M., J. Boethius, L. Eriksson, T. Greitz, T. Ribbe, and L. Widén, Head fixation device for reproducible positron alignement in transmission CT and positron emission tomography, *J. Comput. Assist. Tomogr.*, 5 (1981): 136-141.

Bohm, C., Gretz, T., Blomqvist, G., Farde, L., Forsgren, P.O., Kingsley, D., Sjögren, I., Wiesel, F.-A., and Wik, G., Applications of a computerized adjustable brain atlas in positron emission tomography, *Acta Radiol.*, Suppl., 369 (1986): 449-452.

Bohm, C., Greitz, T., Kingsley, D., Berggren, B.M., and Olsson, L., Adjustable computerized stereotaxic brain atlas for transmission and emission tomography, *Am. J. Neurorad.*, 4 (1983): 731-733.

De Yoe, E.A. and Van Essen, D.C., Concurrent processing streams in monkey visual cortex, *Trends Neurosci.*, 11 (1988): 219-226.

Donaldson, H.H., *The Growth of the Brain*, Chicago, 1895.

Evans, A.C., Thompson, C.J., Marrett, S., Meyer, E., Mazza, S., Performance evaluation of the PC-2048, A new 15-slice encoded-crystal PET scanner for neurological studies, *IEEE Trans. Med. Imaging*, 10 (1991): 90-98.

Greitz, T., Bohm, C., Holte, S., and Eriksson, L., A computerized brain atlas, construction, anatomical content, and some applications, *J. Comp. Assist. Tomogr.*, 15 (1991): 26-38.

Gulyás B., Heywood, C.A., Popplewell, D.B., Cowey, A., and Roland, P.E., Visual form discrimination from colour or motion cues: Functional anatomy by positron emission tomography, *Proc. Natl. Acad. Sci. US*, 91 (1994a): 9965-9969.

Gulyás, B., Ottoson, D., and Roland, P.E. (eds.), *Functional Organization of the Human Visual Cortex*, Oxford: Pergamon Press, 1993, p. 391.

Gulyás, B. and Roland, P.E., Binocular disparity detection in human visual cortex: Functional anatomy by positron emission tomography, *Proc. Natl. Acad. Sci. US*, 91 (1994a), 1239-1243.

Gulyás, B., and Roland, P.E., Cortical fields activated by form, colour, and binocular disparity discrimination: Functional anatomy by positron emission tomography, *Eur. J. Neurosci.*, 6 (1994b): in press.

Gulyás, B., and Roland, P.E., *Visual cortical fields participating in spatial frequency and orientation discrimination: functional anatomy by positron emission tomography*, Human Brain Mapping (194c) submitted.

Gulyás, B., Persson, H., Stone-Elander, S., and Eriksson, L., Neuroimaging studies in visual imagery of reading with positron emission tomography (PET), nuclear magnetic resonance (MR) imaging, and brain electric activity mapping (BEAM), *Eur. J. Neurosci.*, Suppl. 5 (1992): 22.

Gulyás, B., Roland, P.E., Heywood, C.A., Popplewell, D.B., and Cowey, A., Visual form discrimination from luminance or disparity cues: Functional anatomy by positron emission tomography, *Neuro. Report.*, 5 (1994b): in press.

Jensen, J., Untersuchungen über die Beziehung zwischen Grosshirn und Geistesstörung an 6 Gehirnen geisteskranker Personen, *Arch. Psychiatr. Nervenkr.*, 5 (1875): 587-757.

Kosslyn, S.M. and Ochsner, K.N., In searcb of occipital activation during visual mental imagery, *Trends Neurosci.*, 17 (1994): 290-292.

Litton, J.E., Bergström, M., Eriksson, L., Bohm, C., Blomqvist, G., and Kesselberg, M., Performance study of the PC-384 positron camera system for emission tomography of the brain, *J. Comp. Ass. Tomogr.*, 8 (1984): 74-87.

Litton, J.E., Holte, S., Eriksson, L., Evaluation of the Karolinska new positron camera system; the Scanditronix PC2048-15B, *IEEE\Trans. Nucl. Sci.*, 37 (1990): 743-748.

Livingstone, M.S. and Hubel, D.M., Psychophysical evidence for separate channels for the perception of form, color, movement, and depth, *J. Neurosci.*, 3 (1987): 3416-3468.

Olesen, J., Paulson, O., and Lassen, N.A., Regional cerebral blood flow in man determined by initial slope of the clearance of intraarterially injected 133Xe, *Stroke*, 2 (1971): 519-540.

Pakkenberg, B., Total nerve cell number in neocortex in chronic schizophrenics and controls estimated using optical disectors, *Biol. Psychiatry*, 34 (1993): 768-772.

Pakkenberg, H.J., The number of cells in the cerebral cortex of man, *Comp. Neurol.*, 128 (1996): 17-20.

Paul, F., Biometrische Analyse der Frischvolumina der Grosshirnrinde und des Prosencephalon von 31 menschlichen, adulten Gehirnen, *Z. Anat. Entwickl. Gesch.*, 133 (1971): 325-368.

Regeur, L., and Pakkenberg, B., Optimizing sampling design for volume measurements of components of human brain using a stereological method, *J. Microscopy* 155 (1989): 113-121.

Roland, P.E., The cortical field activation hypotheses. In: Sokoloff, L. (ed.), *Brain Imaging and Brain Function*, New York: Raven Press, 1985, pp. 87-106.

Roland, P.E., *Brain Activation*, New York: Wiley, 1993.

Roland, P.E., Eriksson, L., Stone-Elander, S., and Widén, L., Does mental activity change the oxidative metabolism of the brain?, *J. Neurosci.* 7 (1987): 2373-2389.

Roland, P.E. and Gulyás, B., Visual imagery and visual representation, *Trends Neurosci.* 17 (1994a): 281-287.

Roland, P.E. and Gulyás, B., Visual memory, visual imagery, and visual recognition of large field patterns by the human brain. Functional anatomy by positron emission tomography, *Cerebr. Cortex*, 4 (1994b): in press.

Roland, P.E., Levin, B., Kawashima, R., and Åkerman, S., Three-dimensional analysis of clustered voxels in ^{15}O-butanol brain activation images, *Human Brain Mapping*, 1 (1993): 3-19.

Seitz, E.J., Bohm, C., Greitz, T., Roland, P.E., Eriksson, L., Blomqvist, G., Rosenqvist, G., and Nordell, B., Accuracy and precision of the computerized brain atlas programme (CBA) for localization and quantification in positron emission tomography, *J. Cer. Blood Flow Metab.*, 10 (1990): 443-457.

Singer, J.L., Experimental studies of ongoing conscious experience. In: Bock, G.R. and March, J. (eds.), *Experimental and Theoretical Studies of Consciousness*, Chichester: John Wiley and Sons, 1993, pp. 100-122.

Talairach, J., Szikla, G., Tournoux, P., Prossalentis, A., Bordas-Ferrer, M., Covello, L., Iacob, M., and Mempel, E. (1967), *Atlas d'Anatomie Stéréotaxique du Télencéphale*, Paris, Masson & Cie., p. 326.

Van Essen, D.C. and Maunsell, J.H.R., Hierarchical organization and functional streams in the visual corte, *Trends Neorusci.*, 6 (1983): 370-375.

ARTIFICIAL INTELLIGENCE AND COMPUTER VISION

SHIMON ULLMAN *

In this paper, I would like to describe some aspects of computer vision — which is the attempt to build "seeing machines". The main goal of the paper, however, is not to describe computer vision in detail, but to use it as an example to illustrate a number of general points about Artificial Intelligence and its relationship to human cognition.

1. The goals of Artificial Intelligence

Artificial Intelligence (A.I.) in general is an attempt to create intelligent machines, or to produce human-like intelligent behavior in artificial devices. This includes, for example, machines that will be able to interact with their environment, to see, to understand spoken language, to move around, and to manipulate objects. Attempts are being made to construct systems that will be able to solve complex problems and assist human experts in a variety of areas, such as medical diagnosis, engineering design, the interpretation of signals from various sources, the maintenance and monitoring of complex equipment, the management of large projects, and so on. This kind of systems are also expected to learn from their own experience and improve their behavior with practice without being explicitly taught or programmed.

This overall research endeavor is motivated by two rather distinct objectives. Some people are interested in A.I. as an area in the frontier of computer science. Computer power is becoming increasingly accessible, with computers becoming faster, more powerful, and cheaper, every year. At the same time, computers are still being used primarily for fairly routine

* The Weizmann Institute of Science, Israel.

jobs, such as computing payrolls, tedious numerical calculations, tabulating data and so on. In many respects they are still clumsy, inflexible, and difficult to interact with. There is, therefore, an obvious challenge to make computer-based systems more intelligent, responsive, and easier for humans to interact with.

1.1 *Simple machines and complex behavior*

A fundamental question that comes up in attempting to use standard digital computers to perform intelligent tasks is whether there are reasons to believe that current computers will be able in principle to perform intelligent functions that characterize the human brain. When we examine how digital computers operate, we find that their operations are composed of long sequences of quite simple elementary operations. The computer typically stores numbers composed of 0's and 1's, and then performs simple operations such as adding, comparing, or moving such numbers around. In contrast with such simple manipulation of binary digits stand the complex functions of the human brain, including the use of language, sight, solving complex problems, and learning from past experience. The question, then, is whether one can hope to reproduce some of these complexities by machines that are based on simple, limited operations.

The answer is that there is at least some basis for such hope, coming from theoretical work on the foundations of computer science by people such as Turing, Church, Klenee, Post, and others during the 30's and 40's. Roughly speaking, Turing has claimed that any computation that can be performed by a machine of any kind, either simple or complex, can already be performed by a very simple conceptual device called a Turing Machine. This means that even the most complicated computation can in principle be performed by a sequence of simple operations of the type performed by computers. This gave rise to the hope that, with the appropriate programming, it should be possible to produce intelligent behavior by simple computers. At least, it should be possible to demonstrate such behavior in principle, although practical use may of course require more powerful and sophisticated devices.

In this area of practical, working AI systems, there have been a number of interesting and sometimes impressive developments. Systems were built to troubleshoot large telephone networks at AT&T (the ACE system), to design computer configurations at DEC (the XCON system), to analyze geological data and locate, for example, a new area containing molybdenum (PROSPECTOR at SRI), to diagnose certain lung diseases (PUFF at Stanford), or to perform symbolic mathematical computations (MACSYMA at MIT). JPL (Jet Propulsion Laboratory) is using a system (SkiCat) to analyze space science data that reduces the analysis of data from the Palomar Total Sky Survey from about two years to only a few weeks. NASA is using AI systems to monitor and control space shuttle operations at the Johnson Space Center, and another system (GPSS, for ground-processing scheduling system) to manage the thousands of people and other resources needed to prepare the space shuttle for its next flight. Computer systems are also doing well in the game of chess, often considered a prime example for intelligent activity. The best programs have reached the level of Chess Masters, and recently even the world champion lost for the first time to a chess playing program.

1.2 *AI and natural cognition*

Practical considerations are a significant driving force behind research in Artificial Intelligence. There is, however, a second and more fundamental reason for this research. For many scientists in the field, the main goal is to develop general theories of thinking processes and intelligent behavior, that will apply not only to artificial systems, but also to the human brain, and human cognition.

For this goal, the construction of intelligent systems is viewed as an experimental vehicle for studying and testing out general principles. An analogy is sometimes drawn in this regard to a simpler and better understood domain — the role of constructing flying machines in understanding the phenomena of flight. In studying these phenomena, one may be motivated primarily by the attempt to construct flying machines, or by the attempt to understand the principles of natural flight performed by birds. Although airplanes and birds are quite different, they also have a lot in common. The main principles, such as the generation of lift over the wing profile, are quite similar, and the same branch of science, aerodynamics, can deal with both classes. Similarly, the hope in Artificial Intelligence is to develop basic scientific principles for understanding thinking and intelligent behavior, that will apply to both natural and artificial intelligence.

As it turned out, there are indeed useful connections between the study of artificial intelligence on the one hand and some functions of the human brain on the other. Perhaps the best example is provided by the field of computer vision that I would like to consider next.

2. Computer vision

In the area of computer vision, the goal is to build in some sense "seeing machines". This goal is defined functionally: there is no attempt to reproduce in any way something related to our own perceptual experience when we see the world. Rather, the goal is to develop systems that can use visual information to perform certain tasks. One can connect an electronic camera to a computer to obtain a rudimentary "eye-brain" system. The goal is then to program the computer to interpret the incoming image and use it to locate objects, determine their shape, recognize them, and so on.

One may wonder perhaps whether vision should be included in the study of intelligence. Subjectively, vision appears to require little thinking or mental effort. It seems that all we have to do in order to see is to open our eyes and look around. It took a long time to appreciate that this is a misleading perception, and that vision is in fact a highly complicated activity that is extremely difficult to reproduce in artificial devices.

As in other areas of AI, research in computer vision is in part practical, and in part related to studies of the human brain. On the practical side, machines with some vision capabilities can have a large number of useful applications, for example in robotic applications, industrial inspection and quality control, cartography, mail sorting, analyzing aerial photographs for monitoring crops, forests, and other resources, and developing visual aids for the blind.

In terms of the brain, it turned out that in primates a large portion of the brain is devoted to visual perception. In the macaque monkey, for example, about 60% of the cortex is involved with the analysis of visual information. Not surprisingly, therefore, the study of vision is one of the main research efforts in the study of the human brain. To appreciate the difficulties involved in vision, it is useful to remember that the process starts with a raw image, which is nothing more than a large collection of measurements of light intensities at different points in the visual field. For the computer, it is a large array of numbers as in Figure 1.* From this, the system needs to determine what are the objects in the scene, where they are located, what are their shapes and colors, and so on. For the human visual system too, the information delivered from the eye to the brain is similar to this array of independent measurements, and the brain performs all the rest of the process.

Let me turn now to a few examples to illustrate some of the complexities of the task, and some of the problems that need to be solved. I will

* The figure accompanying the text have not been reproduced for technical reasons.

only give a small number of illustrative examples, and then use them to make more general points related to AI and computer vision.

The first example has do to with the problem of separating the scenes into objects and delineating their boundaries. In Figure 2 one can see two shapes — a square and a triangle. The complete shapes are easy to perceive, but one may notice that they do not really exist; they are in fact constructed by the visual system from fragmentary evidence. The visual system interprets the colinear arrangement of the line terminations as evidence for occluding objects and then fills in the complete bounding contours. This process does not take place at a conscious level: the visual system "makes its own mind" about the plausible interpretation of the scene.

The problem of finding object boundaries and segmenting the scene into objects is also one of the main problems in computer vision. In fact, a boundary extraction process is the first processing stage in almost all computer vision systems. The process is not straightforward, and it must include the capability to infer the presence of boundaries from fragmentary and indirect evidence.

The next figure depicts a number of objects that are shown only in part. The image is difficult to interpret, but the interpretation becomes easier when the occluder that hides parts of the figure is shown explicitly. We can now see that the figure contains a number of instances of the letter "B". However, even though we now know what the figures are, when we look back to inspect the original image the grouping of the fragments into objects is still difficult. We see again evidence for sophisticated processes operating in the visual system in an automatic, unconscious manner, and deciding for us, in this case, what goes with what, and which parts form together a complete object. And again, such grouping and segmentation processes need to be incorporated in computer vision systems to allow them to deal successfully with images of natural scenes. In this area we are still way behind the natural system. The problems of segmentation and grouping proved to be quite complicated, and considerable effort is currently invested in trying to understand them better.

The final example along this line shows a part of a scene that is initially difficult to interpret. After a while, the figure of a nurse, standing on the right, becomes clear. The visual system is able to combine correctly the crude black and white regions in the image, discover the figure and separate it from the background. This kind of a process is a remarkable achievement that we are trying to understand and simulate in computer vision systems. Figure 5 shows an example from my own work on image segmentation. Figure 5a shows the input: a noisy contour image of a car in front of a textured background. Figure 5c shows the output of the computation, the car was selected automatically as the object of interest in the image. Figure 5b depicts an intermediate stage in the processing of the image. During this computation some parts of the image become highlighted, others become less salient and drop into the background. The figure-ground separation is obtained in this computation without using any high-level knowledge concerning specific objects such as cars, but based on geometrical considerations, including the length and smoothness of the contours.

Let me turn now to a second problem: the perception of three-dimensional shape of objects in the image. The image itself is of course flat, but a number of cues allow the visual system to extract from it three-dimensional shape. One cue for shape is the variations in brightness across surfaces, as we can see in Figure 6. The problem of extracting shape from shaded images has been studied extensively in computer vision. It turned out to be a complicated problem that even now we can solve with only partial success. We do not have an access to the methods used by our visual system to solve this problem, and therefore, in computer vision, solutions to the problem are sought by using mathematical analysis and computer simulations. The results of such studies also help us to understand the basic principles that are involved in extracting shape from shading, and the insights we gain are also useful in attempting to understand how the task is performed by the visual system.

In deriving shape from shading it is useful to know the direction from which the scene was illuminated. It turns out that the visual system tends to assume, in the lack of evidence to the contrary, the existence of a single light source, illuminating the scene from above. We are not aware of making such assumptions, but we can see the result in Figure 7. The figure contains two identical copies of the same image, but one is rotated by 180 degrees. We perceive one image as a volcano, and the other as a crater. These are reasonable interpretations under the assumption of illumination from above. A different assumption about the illumination will result in a different, three-dimensional interpretation. An important aspect of computer vision is to understand these kinds of assumptions and to use them whenever appropriate in artificial systems as well.

Shading provides one source of information concerning three dimensional shape. A second important source is stereoscopic vision, or seeing with two eyes. The two images seen by the two eyes are slightly different because they are taken from slightly different viewing directions. The brain can extract these small differences and use them to obtain detailed and precise, three-dimensional information. It is impressive that the visual system can perform this function even when the images contain no apparent structure, as in Figure 8. When these images are shown separately, the left one to the left eye only and the right to the right eye, we see a clear, three-dimensional shape emerging from the image.

The problem of stereoscopic vision has been a major focus of research in computer vision. We now have systems that perform this task almost as well as human observers. Such a system can take a pair of images as an input and produce automatically a three-dimensional map of the scene as the output. This capability has various practical applications. One is in cartography, to produce three-dimensional maps from aerial photographs. Figure 9 shows an example of a pair of images of a ravine together with a depth map that was generated from them by a computer system. The figure shows only a rough rendition of the computed depth, but the system was in fact able to generate an accurate map of the terrain from the two images. Another application is in the field of robotics. Figure 10 shows a robotic arm trying to grasp an object. Tasks that involve finding objects, grasping them, and putting them together lie beyond the capability of current industrial robots, because most existing robots do not have sight. They usually perform blindly a simple repetitive movement, tasks such as welding or spray painting. In the future, robots are likely to perform sophisticated tasks that cannot be accomplished today, such as the assembly of complex pieces of machinery. Such robots will probably be equipped with vision capabilities, including stereo vision for analyzing three-dimensional shape.

The final example is more futuristic: it shows an autonomous vehicle. Instead of a driver it has a pair of cameras in front, and a few tons of computer equipment in the back. Perhaps one day it will be possible for a blind person to tell his car: "take me to the office" and the car will comply; however, this day is certainly not close at hand.

3. Conclusions

We have seen only a small number of examples, but they can serve to illustrate a number of general observations. It is instructive to observe that although we have systems that can solve complex problems or play chess at a world class level, in terms of visual capabilities we cannot yet reach the level of a five-year old. One of the main general lessons we have learned from research in AI and computer vision is the enormous hidden complexities of some of the tasks that for us, as humans, are natural and effortless. It took a long time to appreciate the difficulties in the processes of vision because the brain performs them with great efficiency by extensive and highly specialized mechanisms that operate in an autonomous manner, without the need to invest conscious mental effort. The attempts to perform visual tasks artificially is instrumental in unraveling the fundamental problems underlying visual perception and for proposing possible directions that can help us not only in constructing seeing machines, but also in understanding the processes that take place in our own brains.

SCIENTIFIC PAPERS

on:

PONTIFICAL ACADEMY OF SCIENCES' ACTIVITY CONCERNING THE ANALYSIS OF THE IMPACT OF SCIENCE ON ECONOMIC AND SOCIAL LIFE

RESOURCES AND POPULATION

BERNARDO M. COLOMBO * and ANTONIO GOLINI **

Introduction

I would like to start this final session.

There are two talks. There will be one by Prof. Colombo and one by Mr. Ramel.

Now, the first one is on the general problem of population and resources, and we've circulated the English translation to the members. We'll in fact have two speakers on the subject, Prof. Colombo and Prof. Golini, who also took part in our work, both in the Conference and then in preparing the report. I think it's interesting to have both of them contribute, because both of them were actually present at the Cairo Conference: Prof. Colombo in the Vatican Delegation and Prof. Golini in the Italian Delegation. So, in fact, I think the Academy had a *longa manus* on what happened at this Conference.

Report on the Study Week

As to the report you've received, it is, as I said in the closing session, an overview of what was said and debated in the study week on resources and population.

The study week concluded with four distinct reports.

I was rapporteur of the first one on population and these reports were a little two technical, short and disjointed to give a right idea of the conclusions of the study week.

So the Academy asked, also in view of the time needed to have the

* Università degli Studi di Padova, Dipartimento di Scienze Statistiche, Italia.
** Università La Sapienza, Italia e Istituto Ricerche sulla Popolazione del C.N.R.

papers published, to accelerate the presentation of the conclusion of the study week to the higher Authorities, and asked a group of persons — none of them was an Academician, all of them had participated in the study week — to make this overview.

As I said, we met several times with President Marini Bettòlo; we had almost finished just before he fell ill. That prolonged the presentation of the report. It was ready at the beginning of last summer, last year; it was taken to Dr. Perutz who was one of the inspirers and organizers of the study week, and Dr. Perutz found that something was missing on the medical aspects which were treated by several persons during the study week, and he added an Appendix. That is what you have: it is not a statement of the Academy. It is not a political statement, it is just a conclusion reached by persons who looked at facts, and tried to interpret what the situation was, what was going on for the future.

If you wish to have some other clarification on the report, I am here. We are here, Golini worked with us.

Speaking of Cairo, there is a nice opportunity to be two persons here, because in Cairo, as in Bucharest before, things were treated in three different places. One, a plenary session, where the heads of the Delegations spoke, and which voted at the end by consensus on the text which was prepared. In another site, called the Main Committee; Golini worked as a member of the Italian Delegation in the previous meetings of the European Union.

The meeting in Cairo was the fifth initiative of the United Nations. The first two, in Rome in 1963, in Belgrade in 1965, were a sort of general, scientific international meeting, no conclusions, no voting. What I remember of Rome is that each nation of the Soviet block and other communist countries took the floor. In everything they said, they had only one target: to praise their regime. All said the same things, and all in Russian. It was really boring.

In Belgrade we had some very interesting — four — background papers for presentation and discussion. I remember one on population and development, demographic and economic growth. The Chairman was the Vice-President of the World Bank. Many of those who took the floor started their talk by saying: "I am a pessimist", "I am an optimist", this was the level of science at that moment. At the end the Chairman said "I heard many theories, few are facts". There was a background paper by Simon Kuznets, who later on, just a few years later, got the Nobel prize for Economics, which is a classic, on the relation of population, the dynamics of population and economic growth.

Bucharest was totally different. It was a political undertaking. So, there was a text to be taken into consideration, amended, voted, accepted. This text was prepared with a large and very good initiative of the Population Division of the United States Secretariat. The merit goes to Léon Tabar, a French demographer of the Institut National d'Etudes Démographiques, who was the Director of the Population Division. There were several meetings of various kinds. The text was presented to the delegates, this time they were delegates with the right to vote; it was presented by the Secretary of the United Nations. It was put to our attention together with a long list of papers, of study papers, which was not the case in Cairo. There was an amendment, the most spectacular amendment I saw in a political meeting by the Argentinean Delegation, which turned the text upside-down. Just before Bucharest there was a movement about a new international economic order. This was the spirit taken by the Argentinean amendment. Then there was much discussion of this aspect and very little time left for concluding and approving. The conclusions at the end were on family planning, which passed easily and briefly, except for one aspect: that is, quantitative targets. At that time the American Delegation insisted on having quantitative targets in the programme and the Developing Countries would not accept these prescriptions. Then this part was abandoned and everything went on smoothly.

In Mexico City, ten years later, things were prepared both by the Population Division and the United Nations Fund for Population Activities, at the same level. The climate changed completely because there, in Mexico City, developing countries asked for help in family planning programmes and particularly the American Delegation insisted on market, free enterprise, economic development and so on.

In Cairo the work was done, and the Secretary of the Conference was the Secretary of the United Nations Population Fund; now, this is its name, which is a part of the United Nations Development Programme just for population. The Population Division and the Population Commission of the United Nations were separated; everything was done by the United Nations Population Fund, all the preparation. There were six meetings on specific themes — Prof. Golini went to Colombia for a meeting on migration —, four meetings for regional discussions, but very little of these meetings came into the text, which was prepared certainly in the United States and mainly by American-speaking people, not necessarily American, but with the idiomatic choice of Americans. This was the text they presented, which created some language problems. They presented it in Cairo. It was a text not supported, as the one in Bucharest was, for example, by studies of the United Nations Organizations. In Bucharest the FAO presented two studies on the number of people that the earth could sustain. Here in Cairo there was no such study of this kind. There were four or five pages read in the plenary session not by the Director of FAO but by a Vice-Director of FAO. They were the organizations that WHO, the UNESCO and so on downgraded. Everything was there and was treated by the United Nations Population Fund.

What was in the title, a new title, a World Conference on Population and Development (the former Conferences were on Population), turned out to be a Conference on Population forgetting about Development.

They took the stand of the pessimist. That hurt, but, apart from that, some ideological positions were introduced in the text, and these were the most important matters of debate: for instance, a tendency to obtain an internationally recognized right to abortion, or access to family planning, to contraceptives by adolescents excluding the parents, or family in any form, and not a family founded on the covenant of a marriage. There are certainly families of many forms, but the origin would be a covenant for a true family. Also on the reunification of families of immigrants a debate grew up on these problems. Certainly Golini can tell you much more than I about this, because he was in the Main Committee, where the text was negotiated. I was a little surprised, being outside, to see that a big obstacle was by-passed changing 'fertility regulation' into 'regulation of fertility'. Fertily regulation was a working definition of the WHO including abortion. Most of the Delegations who took a position in the plenary session excluded abortion except for the life of the mother or serious health problems of the mother. They could not accept that. 'Regulation of fertility' did not repeat the text of the WHO and then each one was free to give it the meaning he preferred. I don't know how many hours they discussed in formal and informal ways about this small piece of language. Not being English-speaking mother tongue, I wonder whether there is any difference between the two expressions. The Delegations made a division of efforts between the Delegates. I speak now as a member of the Holy See Delegation. I was given the task of following the plenary session. Certainly we had discussion also outside, but the possibility of a scientific contribution was nil, because the meeting was not a scientific meeting, it was a political meeting. I'll show you the kind of frustration I felt as a man who normally wants to discuss, to make up his mind by thinking.

There were figures presented first about maternal deaths due to induced abortion, and then down rates of maternal deaths in general, not specifically for induced abortion. In an official document informing about the preparation of the Conference, there was the figure — International Conference on Population and Development issued at least 95 forms — of 250,000 maternal deaths. Dr. Savick, in opening the presentation of the

Delegates' statements, spoke of 50-60,000 maternal deaths as, she said, the World Bank says. I thought that the World Bank was interested in the government debts — D-E-B-T-S — not in maternal 'deaths'. In any case, that was the authority called in cause by Dr. Savick. The Head of the U.K. Delegation spoke of 200,000, the same as the Head of Denmark's Delegation. The Director of the WHO spoke of 65,000; in fact, he said, there are 500,000 total maternal deaths, of which 13% are supposed to be from induced abortion.

Now, let's go to the rates of maternal deaths. An article in the Population Development Review made a critical evaluation of data presented in the Human Development Report of the United Nations and the World Development Report of the World Bank in the 1993 edition of this annual publication. I will take two countries: the United Nations give a rate of 600 for Pakistan, and the World Bank for the same estimate, 250. In Mexico it is the reverse. What are the sources from which these data come? For Pakistan, community studies, hospital studies, government estimates, other estimates; so you see, the community studies for Pakistan give a rate of 82 to 200; hospital studies 171 to 5,000; government estimates, other estimates of various kinds. Mexico has no sources of this kind, government estimates or other sources. The author of the article concludes that the figure of 500,000 is presented in a book by the WHO. The author of the article, Parker Molding, says what is abundantly clear is that reliable data on maternal mortality do not exist for most developing countries, for Zambia as for Argentina, for instance, and he concludes: "Even when they appear in prestigious secondary sources by WHO publications, data should be presumed guilty until found innocent". Well, the Director of WHO gives figures on this kind of information.

I have another example. In a publication of the United Nations Population Division, it is written that in the eighties there were between 220,000 and 800,000 illegal abortions. Just after the REPCO in April, Golini wrote in an article, among other things, that illegal, induced abortion has almost disappeared in Italy. No scientifical authority of any weight could have given this information, but then if the Secretary General of the United Nations cannot trust one of its Divisions, whom does he trust when he speaks? This part on maternal death and so on was in the preamble of the text presented to the Conference. In April, finally, it was taken out, perhaps set in another place. In the plenary session I saw the moving of all the facts, of many facts and problems, about a lot of countries, including Armenia, 300,000 refugees, Liberia, 150,000, Zaïre, 2 million refugees, all the countries of the former Soviet Union. This contrasted with the pride of Islam: "We are one and a half billion people. In 2005 we will be almost two billion people".

CHEMICAL HAZARDS IN DEVELOPING COUNTRIES

CLAES RAMEL *

THE COLLABORATION BETWEEN THE PONTIFICAL ACADEMY OF SCIENCES AND THE ROYAL SWEDISH ACADEMY OF SCIENCES

In 1988 the President of the Pontifical Academy, Professor Carlos Chagas, visited the Swedish Academy of Sciences and suggested a collaboration between the two academies on the urgent problem of preserving tropical forests. As a Brazilian he no doubt had direct contacts with the deterioration of the world's richest forest areas and the rapid loss of an essential part of the global biodiversity. Considering the fact that the major part of tropical forests are located in Catholic countries, it also seemed of great importance to have a statement by the Pope, which could serve as a guidance to these countries. It was therefore decided to arrange a conference in the Vatican on the topic "Tropical Forests and the Conservation of Space" and this meeting took place 14-18 May 1990.

The preservation of biodiversity in tropical forests as well as elsewhere, does not only involve scientific and economic but also ethical problems. Particularly in the Christian world. Man has been given a position above all other living creatures and Nature is subordinate to him. In my introductory talk at the meeting I emphasized that there are in fact also compelling biological reasons to look upon our own species as qualitatively different from other species particularly from one evolutionary point of view. Biological evolution is based on the transfer of information from one generation to the next. For all plant and animal species, except for Man, the source of information and evolution is provided by alteration of genetic information through mutations. But that is a slow and impractical process, where practically all genetic changes constitute useless mistakes, which have to be eliminated by natural selection. Man is the only species which has

* Royal Swedish Academy of Sciences.

managed to by-pass that cumbersome and time consuming evolutionary process by an entirely different kind of evolution — cultural evolution, where acquired experience is transmitted orally or by writing from one generation to the next. This cultural evolution is infinitely more efficient and selective and it no doubt constitutes the major reason for the formidable success of Man as a biological species.

But if Man exhibits a hegemony over the living world it does not relieve him from an ethical responsibility for other species. The role of Man as a steward for the living world was emphasized by the Pope in his speech at the conference, where he appealed to the world to protect the valuable ecosystems of tropical forests. This statement by the Holy Father was greeted with great satisfaction by participants in the conference.

Our conference on tropical forests was organized under the general heading "Man and His Environment", and this heading reflects the intentions for the collaboration between the Pontifical Academy of Sciences and the Royal Swedish Academy of Sciences, which was established through a letter of agreement in 1990. Beside the scientific problems, this area entails other ethical aspects — one important one being the responsibility of developed countries towards developing countries in environmental matters. Together with the former President of the Pontifical Academy of Science, Professor G.B. Marini-Bettòlo, we decided to bring up another critical problem in a joint conference, namely "Chemical Hazards in Developing Countries", which is an extremely important area, where industrialized countries have a heavy responsibility for many environmental problems.

The conference took place from 21-23 October 1993 at the Pontifical Academy in the Vatican, with 26 participants from 11 nations. With the time I have at my disposal. I can only give some glimpses of the material presented during the conference. I will divide my talk into four parts — industrial contamination, ecotoxicological aspects, pesticides and finally some words about national and international guidance and regulations.

Industrial contamination and waste handling

Thanks to Dr. Finkelman from Guatemala, the conference got a particularly comprehensive and interesting overview of the situation in Latin America with special emphasis on industrial problems. Developing countries have mainly produced basic goods which have been further processed by industrial countries and the large economic gain has been in the industrial and not in the developing countries. There is an understandable trend in developing countries to share the gain from processing through industrialization. In Latin America there has been a wave of industrialization since World War II, but because of a lack of economic foundation and infrastructure little attention has been payed to environmental consequences and the result has very often been fatal. The problems have been magnified by the population increase — since 1930 there has been a fourfold increase of the population in Latin America. Another severe problem has been the rapid urbanization and the formation of megacities with Mexico City as an example. With this urbanization has followed a sharp increase in traffic and air pollution. As an example of the consequences, it has been estimated that 2.3 million children suffer from respiratory problems in Latin America.

The toxic waste disposal is very badly regulated. In Mexico 450,000 tons of industrial waste are released per day, of which 14,500 have been considered hazardous. The largest problems are caused by small manufacturing plants. 5,748 tons of waste were generated by 39,000 small plants in Mexico City and only a fraction was properly handled. In industrial countries a legal protection against unregulated disposal of waste has been developed, but such regulations are largely lacking in developing countries, or existing laws and regulations are not implemented. Therefore, improper handling of industrial waste implies a major health threat in most developing countries. This applies also to domestic waste — less than 2% is subjected to any treatment and it is hardly astonishing that cholera epidemics break out periodically.

Many industries in Latin America cause pollution problems, but the petrochemical industries are particularly damaging. All oil-producing countries have serious problems with oil spills, not the least from off-shore drilling. Pipe lines for instance in Venezuela are corroded and leak oil constantly. But the very rapidly expanding petrochemical industries have also comprised large chemical enterprise based on petrochemical products in several Latin American countries. The chemical products include olefins, ethylene, propylene, butadiene and aromatics, benzene, xylene and styrene, which comprise 5% of the world's production. Some of these products are established carcinogens.

In Africa the pollution problems are no less severe than in Latin America, but industrial waste is not the dominating environmental problem for the simple reason that the population is mainly rural: 80% are small scale farmers. The predominant problem there is therefore the use of pesticides, which I will come back to.

Asian countries comprise all stages from highly industrialized countries, like Japan, to poor, overpopulated countries like Bangladesh. It is therefore impossible to give any short and generalized picture of chemical hazards in Asian countries. However, many of the problems met with in Latin America can be found here too. Increasing populations and formation of megacities have the same adverse effect as elsewhere in developing countries. The situation in China with its rapid economic development is of course of particular interest. In general the environmental problems in China remind us of the situation in the former Soviet Union. The centrally directed system in China has had similar disastrous effects in the former Soviet Union. One such problem is the fact that industrial and environmental problems have been handled by different administrative bodies. Among the health problems from chemicals, heavy metals seem to be particularly difficult. As a matter of fact, cases of the Minamata disease, that is methyl mercury intoxications, have been reported from China.

Ecotoxicology

The release of persistent chemicals in the environment often is only the beginning of a series of disastrous events, particularly because of bioaccumulation in various organisms. Classical examples of such bioaccumulation of chemicals are DDT, which accumulates from water to higher trophic levels in vertebrates by a factor of 1000 or more. Chlorinated compounds like hexachlorbenzene, dieldrin and lindane have been shown to accumulate in the same way in Lake Kariba in Zimbabwe as in temperate waters. But there are also differences between temperate and tropical zones. It was thus reported by Jensen that the most stable metabolite of DDT, DDE, in ampoules outdoors in Nigeria disappeared completely overnight, while Lindane and Toxaphene remained totally stable.

Heavy metals are particularly important from the point of view of ecotoxicology and human health. Perhaps most severe is the situation with mercury, partially due to the use of mercury in gold panning. By applying a layer of mercury in the bottom of the pan, the gold is amalgamated and will be retained by evaporating the mercury by heating. This method is old but it has been revived in many parts of the world, like Latin America, Africa, Russia, China, Vietnam, Indonesia, New Guinea and Philippines. The most affected area seems to be Amazonas and Dr. Moreira from Brazil elaborated on this problem. Of the mercury released in gold panning, 15-50% goes into the rivers, 65-83% as vapour. The airborne contamination from this source of mercury is most probably of global significance. In the rivers the mercury is converted into methyl mercury just as in temperate zones; that is, the compound which caused the Minamata and Iraque catastrophes. The methyl mercury is very stable and accumulates in fish and other aquatic organisms. Fish-eating people downstream evidently will be exposed to the mercury, but little seems to be known about the consequences on ecosystems and on human health.

The bioaccumulation of persistent chemicals makes analysis of suitable organisms the most efficient method to monitor the release of chemicals from industries and other human activities. Jernelöv reported on a recent international activity of this kind — the International Mussel Watch. Mussels function as a filter through which large quantities of water pass, while planktons and thereby chemical contaminants are retained. The purpose of this project has been to provide a general system for monitoring pollution in the sea and to be able to make comparisons between different parts of the world in that respect. The initial implementation of that project has been along the coasts of Latin America. Local hot spots of contamination have been identified and they will be followed up further.

Pesticides

Pesticides constitute the most severe chemical pollution problem in most developing countries and therefore a major part of the conference in the Vatican dealt with that problem. The world production of pesticides increases steadily and proportionally the share of developing countries increases even more. The pesticide problems of the developing countries are particularly severe because of the rapidly growing populations and the demand for higher food production. The tropical climate implies serious problems with agricultural pests; in Africa 50% of the crop is lost through pests. It has been estimated that there are 67,000 agricultural pests, of which 5% are considered serious. An additional and grave problem in this context is the development of pesticide resistance, which causes a demand for increased use of pesticides and for the development of new pesticides. Just cotton production in India requires about 30 pesticides today. A disastrous example of pesticide resistance is the fight against malaria, which seemed successful thirty years ago but must be considered a failure today. There are 270,000,000 cases of malaria per year now.

Another important aspect of pesticide problems, particularly obvious in developing countries, is poisoning from the use of pesticide in agriculture. There are at least 3 million cases of poisoning by pesticides per year, 99% of which occur in developing countries.

Many persistent and poisonous pesticides, which are banned in developed countries, are used in large amounts in developing countries. This applies, for instance, to chlorinated compounds. It was however also reported during the conference that the experience from developed countries concerning the persistency of pesticides, does not always apply to tropical countries. Thus as I mentioned before, DDT and DDD have turned out to be far less persistent in a tropical climate. The exchange of DDT for other pesticides in developing countries is therefore questionable in tropical climates.

The conference formulated at the end a series of recommendations, which can be summarized as follows:

Implementation of international rules and recommendations

— Agenda 21, Chapter 19
— WHO, FAO, UNEP, coordinated by IPCS — International Programme for Chemical Safety
— ILO on occupational Health
— PIC, Prior Informed Consent
— OECD guidelines for biotechnological products

Exchange of information between developed and developing countries

— Risk evaluation
— Judicious use of chemicals
— Safe disposal
— Administrative infra-structure and legislation

Responsibility of developed countries

— Restriction of marketing of hazardous chemicals
— Restriction of export for dumping of hazardous chemicals
— Application of PIC for material produced by genetic engineering

National education and information

— Training to promote safe and effective use in industry, agriculture and forestry
— Reduction of chemical accidents by training and establishing contingency plans

Regulatory control

— Collaboration between agencies and ministries
— Phasing out of hazardous chemicals
— Substitution of chemical technologies, when possible
— Developing regulatory services
— Cleaning up actions of contaminated sites

Research

— Developing reference National Laboratories
— Monitoring of exposure
— Epidemiological investigations

FROM THE EXPLORATION OF THE UNIVERSE TO THE EXPLORATION OF THE HUMAN BRAIN

MINORU ODA *

Keiji Tanaka *et al.* of RIKEN Frontier Research System have studied the activity of the monkey's brain, particularly visual activity, with an embedded probe in the cortex (K. Tanaka *et al.*, Nature 1992, vol. 360 No. 6402). It was concluded that the minimum spatial dimension of the brain activity, the column, is approximately 0.1 mm.

What about for MAN? Of course, we cannot embed a probe. But with some non-invasive method it was learnt that with the activity of the brain, the blood flow and/or de-oxidization of hemoglobin concentrate at a certain region. As for the non-invasive probing into the human brain, techniques of multiple scattering of light, PET, MRI and the measurement of very weak magnetic fields with SQUID have been developed, and the concentration of the blood flow has been identified to a scale of several mm. (e.g. M. Raichle, Scientific American, April 1994).

If the spatial resolution of the non-invasive probing is improved, the probing may result in a breakthrough in bridging high level activity of the brain or, say, the mind and physical observation. In the following, a proposed technique, which may be called "Fourier-Transform-Microscope" (FTM) is described. It is a composed structure of multiple modulation collimators to construct an image of the X-ray source produced in the brain.

One often sees the Moire's fringe pattern through overlapped lace curtains, grid plates, mesh screens etc. The modulation collimator was conceived as a device to determine the location, the size and the shape of celestial X-ray sources from how the sources are seen alternatively through the Moire's fringe projected on the sky. In an early period of X-ray astronomy, with this device the X-ray source was found to be as small as a

* Pontifical Academician. International Institute for Advanced Studies, Japan.

"star" and the precise location of a bright source, e.g. SCO X-1, led to the discovery of its optical counterpart.

The device usually consisted of two grids, separated by a certain distance. An X-ray detector placed behind the two-layered grids measures the integrated X-ray flux which is a Fourier component of the source transformed with the spatial wave-length determined by the spatial period of the grid.

In the late 1970s the modulation collimators were flown aboard the balloon at high attitudes and were used to produce the X-ray image of the Crab Nebula from spatial Fourier components at various position angles of the Nebula. Also, rotating modulation collimators had been utilized aboard X-ray astronomy satellites, "Hakucho" (launched in 1979) and "Tenma" (launched in 1983), and Solar X-ray astronomy satellite "Hinotori" (launched in 1980). The rotating modulation collimator produces Fourier components for a number of spatial wave numbers and position angles by which an X-ray image of the source may be reconstructed.

The Fourier-Transform-Telescope (FTT), which is sometimes referred to as the "ultimate modulation collimator", was realized as a Hard X-ray Telescope (HXT) on board the Solar X-ray satellite "Yohkoh" (launched in 1991), the total sensitive area of the X-ray detector being approximately 70 square cm. The HXT consists of 64 element modulation collimators of different spatial wave numbers and position angles which produce 64 Fourier components on the Fourier u-v plane. With the HXT, being aided by the Maximum Entropy Method (MEM), the X-ray images of typical Solar flares have been produced with 0.5 sec. of time resolution and approximately 5 arcsecond of spatial resolution for four energy bands of X-rays.

The concept of the FTT may be transferred to the Fourier-Transform-Microscope (FTM) to produce the image of microscopic X-ray or soft gamma ray sources with slight modification of the structure of the modulation collimator. Unlike the case of producing 2-dimensional X-ray images of astronomical objects at essentially infinite distances, here the sources are of 3-dimensional at near distances.

The difference may cause some difficulties. Kazuo Makishima of the University of Tokyo and his colleagues are undertaking laboratory simulations utilizing an optical system to investigate the situation.

In order to achieve the spatial resolution better than 100μ m or 10μ m, the advanced manufacturing techniques of the fine grids, like lithography, have to be accommodated and also the techniques of the alignment of the grids have to be devised.

The method of producing the X-ray or soft gamma-ray sources in the brain is represented in Table I (not given). First, as is utilized for the diagnosis of the heart, radioisotopes like ^{201}TI, ^{123}I, ^{133}Xe may be introduced into the blood vein. They radiate X-rays of energy of several tens to over 100 KeV which may be observed from outside with an essentially negligible amount of scattering through the body.

Secondly, a muon beam facility, which is under construction at Rutherford-Appleton-Laboratory in England under the collaboration between the Laboratory and RIKEN led by K. Nagamine, may be utilized. The region in the brain to be observed is irradiated with low energy minus-muons which replace the atomic lectons of oxygen atoms producing muonic oxygen atoms. The muonic atoms then generate hard KX-rays. The muonic oxygen atoms generate characteristic 133KeV KX-rays.

The third possibility is to utilize X-rays to produce fine shadows of blood veins, which contain iodine, at just above and below the KX-ray-edge of Iodine. The difference of the shadows enhances the fine image of the vein. As the X-ray emitting source, "SPring-8" i.e. a synchrotron radiation facility with 8Gev electron ring located in Hyogo-prefecture, Japan, may be used.

Question of dosage, when radioisotopes are introduced into blood and a part of the brain is irradiated by muon beams or X-ray beams, has to be carefully investigated in relation to the clarity of the image and the time of exposure, i.e. temporal resolution.