Predictability in Science: Accuracy and Limitations
Pontifical Academy of Sciences, Acta 19, Vatican City 2008
www.pas.va/content/dam/accademia/pdf/acta19/acta19-vicunia.pdf

PREDICTING THE MINIMAL SUSTAINABLE GENOME

RAFAEL VICUÑA

Predictability is the capacity to foretell a given situation or event based on observation and reasoning. This ability is fundamental to the progress of science, since it relates to the processes of deduction and induction. Predictability is in itself a test of scientific understanding, since the power of scientific method shows its strength when predictions made by investigators coincide with observable events in the laboratory or in nature.

Recent advances in functional genomics and computational biology have led to considerable progress in our comprehension of gene structure and expression. One of the most defying issues that geneticists are presently addressing is the so-called minimal genome concept, in other words, the attempt to define the minimum number of genes that are necessary to sustain a free-living cellular organism. It is expected that the lowest number of genes will be identified under the most favorable conditions imaginable, that is, in the presence of a full complement of nutrients and in the absence of environmental stress. Until the size of such minimal genome is experimentally confirmed, this task epitomizes an exercise of predictability.

The ability to correctly define the minimal gene set goes to the heart of our understanding of cellular life. In addition, the fulfillment of this goal should provide some insight into the earliest stages of biological evolution, since it is assumed that simpler free living cells, with genomes much smaller than those of extant microbes, must have proliferated at the onset of life on earth. On the other hand, if a minimal genome proves to be something that we can observe in nature or obtain in the laboratory, some may be tempted to believe that the challenge of synthesizing simple forms of cells may be plausible in a not so distant future.

The Minimal Cell

Closely related to the minimal genome problem is that of the minimal living cell, defined as a theoretical entity having the smallest number of

components and functions that are necessary to be considered alive.[1,2,3] Although it might not be straightforward to reach an agreement on the meaning of 'alive' in this context, this minimal cell would need to exhibit the properties of self-maintenance or metabolic activity, self-reproduction and Darwinian evolution.[3] A pioneering effort to describe a minimal cell was achieved by Morowitz.[4] Based on enzymatic reactions that he considered essential, Morowitz arrived at the conclusion that such minimal cell should be about one tenth the size of Mycoplasma, a very small and well described bacterial pathogen. However, it is most likely that primitive cells were even simpler than the one predicted by Morowitz and that complexity gradually built up as a result of billions of years of evolution of new metabolic pathways and defense mechanisms.

In a recent publication, Szostak *et al.*[5] have envisaged how an early protocell might have looked like. The nucleic acid could have been a double stranded RNA molecule, with one strand possessing replicase activity and the complementary strand serving as its template. Physical confinement of this genome was necessary to facilitate preferential replication of those RNA molecules incorporating advantageous mutations, thus allowing Darwinian evolution. Self-assembling and self-replicating vesicles composed of amphipathic lipids must have been responsible for encapsulating the RNA molecules. At this point, coupling between membrane synthesis and replication of the genome would be required to improve survival of the entire entity. This condition could have been met by a second gene encoding a ribozyme catalyzing the synthesis of phospholipids necessary to build up the membrane.

Another proposal of a minimal cell, in a closer agreement with contemporary life, came as a result of the so-called E-CELL project.[6] The proposed

[1] Luisi, P.L., Oberholtzer, T. and Lazcano, A., 'The notion of a minimal cell: A general discourse and some guidelines for an experimental approach', *Helv. Chim*. Acta 85, 1759-1777, 2002.

[2] Islas, S., Becerra, A., Luisi, P.L. and Lazcano, A., 'Comparative genomics and the gene complement of a minimal cell', *Origins of Life and Evolution in the Biosphere* 34, 243-256, 2004

[3] Luisi, P.L., Ferri, F. and Stano, P., 'Approaches to semi-synthetic minimal cells: a review', *Naturwissenschaften* 93, 1-13, 2006.

[4] Morowitz, H.J., 'Biological self-replicating systems', *Progr. Theor. Biol*. 1, 35-58, 1967.

[5] Szostak, J.W., Bartel, D.P. and Luisi, P.L., 'Synthesizing life', *Nature* 409, 387-390, 2001.

[6] Tomita, M., K. Hashimoto, K., Takahashi, K, Shimizu, T.S., Matsuzaki, Y., Miyoshi, F. Saito, K., Tanida, S., Yugi, K., Venter, J.C. and Hutchison III, C.A., 'E-CELL: software environment for whole-cell simulation', *Bioinformatics* 15, 72-84, 1999.

autonomous cell in this case required a total of 105 protein-coding genes. These were involved in very rudimentary metabolic pathways for glucose metabolism and phospholipid biosynthesis, in gene transcription and in protein synthesis. Accordingly, this minimal cell should be able to maintain metabolic homeostasis, but could not reproduce or evolve since it could not synthesize DNA with the proposed genome.

Theoretical Approaches to the Minimal Genome

However, although highly provocative, this hypothetical minimal cell is pure speculation. The simplest living cells today possess a few hundred DNA genes encoding a corresponding number of proteins. Most of these proteins are catalysts of numerous reactions taking place in a highly confined and organized environment. There are also some genes in the cell that encode various types of RNAs. Table 1 illustrates some small genomes found in bacteria, most of which are endosymbionts and obligate parasites. For comparative purposes, the genome sizes of *Escherichia coli* and *Streptomyces coelicolor* are also shown.

The hyperthermophile *Nanoarchaeum equitans* possesses the smallest genome that has been sequenced and analyzed.[7] It is an obligate symbiont that grows in co-culture with the crenarchaeon *Ignicoccus* and therefore, as opposed to *Mycoplasma* and *Haemophilus*, it cannot grow in the laboratory as a free-living microorganism. In contrast to other microbial genomes that are undergoing reductive evolution, *N. equitans* has an unusually high gene density, with little non-coding DNA and very few pseudogenes. In addition to its 556 protein coding genes, it has 17 genes encoding ribosomal RNA. Adaptation to an obligatory parasitic life is evidenced by the lack of most genes involved in *de novo* synthesis of aminoacids, nucleotides, cofactors and lipids, as well as genes of the glycolytic/gluconeogenesis pathways, the pentose shunt and the Krebs cycle. However, in contrast to other parasites, it contains most of the enzymatic machinery required for DNA repair and the complete machinery for DNA replication, transcription and translation. The lack of genes encoding transfer RNAs and the large amount of coding capacity devoted to surface proteins that interact with the host seem to leave little room for further reductions of the genome.

[7] Das, S., Paul, S., Bag, S.K. and Dutta, C., 'Analysis of *Nanoarchaeum equitans* genome and proteome composition: indications for hyperthermophilic and parasitic adaptation', *BMC genomics* 7, 186, 2006 (Epub).

TABLE 1. GENOME SIZE OF SOME PROKARYOTES

Microorganism	Genome size (kb)	Protein coding genes
Mycoplasma genitalium	580	482
Nanoarchaeum equitans	491	556
Buchnera aphidicola BBp	616	545
Blochmannia floridanus	706	625
Chlamydia tracomatis	1,000	895
Rickettsia prowazekii	1,100	834
Aquifex aeolicus	1,591	1,553
Haemophilus influenzae	1,830	1,703
Escherichia coli	4,640	4,288
Streptomyces coelicolor	8,500	7,825

Some other very small genomes found in nature are those of bacteria living in endosymbiosis with insects. These include *Blochmannia floridanus*, *Wigglesworthia glossinidia* and *Buchnera aphidicola*, the endosymbiotic bacteria of carpenter ants, tsetse flies and aphids, respectively. In these obligate, metabolic interdependent insect-bacterial relationships, the bacterial cells are contained in specialized host cells called bacteriocytes and the infection is vertically transmitted via eggs and young embryos. These endosymbionts diverged from their free-living relatives approximately 70 (*Blochmannia*) to 250 (*Buchnera*) million years ago. Since then, they have experienced a genome minimization process that started with a rapid decline and gradually slowed down until reaching the sizes found in present day genomes, which are 20-25% of the original ones. There are species of the genus *Buchnera* possessing genomes as small as 450 kb, much smaller than that of *Mycoplasma genitalium*.[8] However, they have not been sequenced and annotated to date. On the other hand, the hyperthermophile *Aquifex aeolicus* possesses the smallest sequenced genome of an autotrophic microorganism.

[8] Gil, R., Sabater-Muñoz, B., Latorre, A., Silva, F.J. and Moya, A., 'Extreme genome reduction in *Buchnera* spp.: Toward the minimal genome needed for symbiotic life', *Proc. Natl. Acad. Sci. USA* 99, 4454-4458, 2002.

The question is whether the very tiny genomes shown in Table 1 can be further reduced without affecting life's viability. Several laboratories have approached this problem, both theoretically and experimentally.

Some years ago, Mushegian and Koonin made a comparison between the genomes of two parasitic bacteria, namely *Mycoplasma genitalium,* a Gram positive human urogenital pathogen with 482 open reading frames (ORFs), and *Haemophilus influenzae,* a Gram negative human parasitic bacterium causing pneumonia and other diseases, with 1,700 ORFs.[9] The rationale followed by these investigators was that the genes that are conserved in these two bacteria are almost certainly essential for modern-type cellular function and therefore likely to approximate the minimal gene set. The authors found 240 *M. genitalium* genes have orthologues[10] in the genome of *H. influenzae*, although at first sight this resulting assemblage appeared insufficient since some key enzymes from intermediary metabolism were missing. The reason for this presumed incompleteness, which has also been observed by other authors in similar studies, is the so-called nonorthologous gene displacement (NOGD). This implies that unrelated or distantly related proteins are adapted to perform the same function in their respective cells. After identifying 22 nonorthologous displacements, the authors selected them from the genome of *M. genitalium* and they were added to the set of shared orthologues. Of the resulting 262 genes, 6 were eliminated from the minimal set because they were likely to be specific for parasitic bacteria. The conclusion was then reached that a hypothetical microorganism with 256 genes would be minimally equipped with the required cellular components, metabolic pathways, systems to copy and express its genome, a signal transduction apparatus and a few chaperones. It must be kept in mind, however, that genomes also contain some sequences encoding RNA species such as ribosomal RNA or transfer RNA, which are definitely essential for life. Therefore, these non-protein coding genes should be included in a minimal gene set. In the case of *M. genitalium*, the number of RNA coding genes is 43.

In a different study, comparative genomics by means of computational methods showed that a total of 462 protein-coding genes are shared among

[9] Mushegian, A.R. and Koonin, E.V., 'A minimal gene set for cellular life derived by comparison of complete bacterial genomes', *Proc. Natl. Acad. Sci. USA* 93, 10268-10273, 1996.

[10] Ortologues are homologous genes in different species that originate from the same ancestral gene in the last common ancestor of the species compared.

three aphid endosymbionts of the genus *Buchnera*.[11] They include genes to synthesize amino acids required by the host and genes necessary for cell division, replication, transcription and protein synthesis. When *B. floridanus* and *W. glossinidia* were added to this analysis, the number of conserved genes in the five endosymbionts decreased to 276. Of these, 156 are also conserved among host-dependent parasites such as *Rickettsia prowazekii*, *Chlamydia trachomatis* and *M. genitalium*. In addition to the 276 protein-coding genes, the five endosymbionts also share 36 RNA specifying genes,[12] which have to be considered in a minimal genome. On the other hand, the number of genes shared between the group of five insect endosymbionts and *Rickettsia*, *Chlamydia* and *Mycoplasma* are 220, 218 and 179, respectively. It is tempting to suggest this last figure as the basic subset of genes required for bacterial cell life. The rest of the genes shared by the five endosymbionts but absent in *M. genitalium* could be involved in endosymbiotic functions.

One of the most thorough studies on minimal genomes has been conducted by Gil and collaborators.[13] Although basically theoretical, this analysis includes data from experimental work and computational comparisons carried out by several groups. Therefore, this contribution by Gil *et al.* could be considered a typical exercise of prediction. These authors defined functions performed in any living cell and then listed the genes that would be necessary to support such functions. The resulting gene list was corrected to fill the gaps in biochemical pathways considered to be essential to maintain homeostasis in any living cell. The hypothetical core able to sustain a functional bacterial cell under ideal conditions turned out to contain 206 protein-coding genes. The number of genes and their respective functions were: 16 implicated in DNA replication and repair, 106 involved in gene transcription and protein synthesis, 15 related to protein folding and secretion, 56 participating in energetic and intermediary metabolism, 4 involved in transport, 1 in cell division and 8 in poorly characterized functions. The latter were identified as essential in *B. subtilis* and therefore added to the list. The authors cautiously acknowledged that there is no conceptual or

[11] Klasson, L. and Anderson, S.G.E., 'Evolution of minimal-gene-sets in host-dependent bacteria', *Trends in Microbiol*. 12, 37-43, 2004.

[12] Gil, R., Silva, F.J., *et al.*, 'The genome sequence of *Blochmannia floridanus*: Comparative analysis of reduced genomes', *Proc. Natl. Acad. Sci. USA* 100, 9388-9393, 2003.

[13] Gil, R., Silva, F.J., Peretó, J. and Moya, A., 'Determination of the core of a minimal bacterial gene set', *Microbiol. Molec. Biol. Rev*. 68, 518-537, 2004.

experimental support for the existence of one particular type of minimal cell. In a sense, there is no single type of minimal metabolism. Therefore, they proposed that their conclusions should be regarded as provisional.

Genome Reduction in the Laboratory

Several experimental approaches permit identification of non-essential genes. These include site-directed gene knockout, global transposon mutagenesis, plasmid insertion mutagenesis and use of antisense RNA.

In a pioneering work,[14] Itaya inserted an antibiotic resistance gene cassette at seventy nine randomly selected chromasomal loci in *Bacillus subtilis*, most of them rare restriction enzyme sites. Only six of the insertions affected bacterial growth in rich medium. Although it is likely that functional redundancy may lead to an underestimation of essential genes, the author also tested multiple (7-, 12- and 33) fold mutations among identified dispensable loci. These highly mutated strains retained the ability to form colonies. The indispensable DNA size was calculated by satistical analysis to be in the range of 318-562 kb. Considering that bacterial open reading frames encompass about 1 kb, this minimal genome would harbor between 300 and 500 genes.

Transposons are segments of DNA that can move from one location in a genome to another, often disrupting gene function by insertion. Once the location of enough transposon insertions are defined by DNA sequencing, the researcher is able to deduce with some certainty that regions in which transposon insertions are not observed are likely to be essential for viability. In other words, transposition not affecting cell viability allow the identification of non-essential genes. Global transposon mutagenesis has been used to identify non-essential genes in *M. genitalium*. In a first attempt,[15] 685 insertion events led to the conclusion that 265 to 350 of the 482 protein-coding genes are absolutely essential for growth under laboratory conditions. Of these, 100 are of unknown function. To confirm that the insertions had been indeed disruptive, only those found after the first three codons and before the 3'-most 20% of the coding sequence were considered in this work.

[14] Itaya, M., 'An estimation of minimal genome size required for life', *FEBS Lett*. 362, 257-260, 1995.

[15] Hutchinson, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O. and Venter, J.C., 'Global transposon mutagenesis and a minimal Mycoplasma genome', *Science* 286, 2165-2169, 1999.

In a more recent report, the same group expands this work by proving gene dispensability after isolation and characterization of pure colonies, a precaution that they did not originally take.[16] This step is necessary because there is always the possibility that other cells in the same pool of mutants may supply a gene product. In this refined study, the authors found that 100 instead of 120 genes are nonessential. None of the genes suspected to be essential for growth (DNA replication, glycolisis, cytoskeleton, etc.) were disrupted, including the 43 RNA-coding genes. This new study also showed that in spite of its small genome, *M. genitalium* possesses some enzymatic redundancy, a property that would mask the requirement of a gene disrupted by insertion mutagenesis. Taking this into consideration, 5 genes were added to the minimal set, thus ending up with a total of 387 essential protein-coding genes. This number is higher than the same group's initial prediction and surprisingly larger than those predicted in the theoretical calculations described above. Another unexpected characteristic of this essential gene set is that it includes 110 proteins of unknown function. On the other hand, it is likely that several of the dispensable genes are involved in the maintenance of *M. genitalium* in the human urogenital tract, its natural habitat.

Genome-scale transposon mutagenesis has also been conducted with the bacterium *H. influenzae*. The number of putative essential genes in this case turned out to be 670, also a much higher figure than the one deduced by theoretical comparison with *Mycoplasma*.[17] This implies that about 40% of the genome is essential under the conditions tested. Some of the genes identified have proven dispensable in other bacteria, whereas as many as 259 genes lack a defined functional role. It is possible that transposon mutagenesis overestimates the size of the minimal gene set by misclassification of non-essential genes that slow down growth without arresting it. Moreover, in this particular study, the authors included in the minimal set genes that had been the target of single insertions (191), following the rationale that some essential genes encode non-essential domains. On the other hand, computation can underestimate the set because it takes into account only those genes that have remained similar enough during evolution to be considered as canonical orthologues.

[16] Glass, J.I., Assad-García, N., Alperovich, N., Yooseph, S., Lewis, M.R., Maruf, M., Hutchinson, C.A., Smith, H.O. and Venter, J.C., 'Essential genes of a minimal bacterium', *Proc. Natl. Acad. Sci*. USA 103, 425-430, 2006.

[17] Akerley, B.J., Rubin, E.J., Novick, V.L., Amaya, K., Judson, N. and Mekalanos, J.J., 'A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*', *Proc. Natl. Acad. Sci. USA* 99, 966-971, 2002.

Another study utilizing a similar experimental technique known as plasmid insertion mutagenesis was conducted with the Gram-positive bacterium *Bacillus subtilis*. Unexpectedly, only 271 genes out of the 4,100 comprising the genome appeared essential when they are individually inactivated.[18] However, this approach does not detect essential functions encoded by redundant genes because a single gene is inactivated in each mutant strain. Therefore, the minimal set mentioned before may be an underestimation. The great majority of the 271 genes encode functions related with information processing, cell envelope, shape, division and energetics. Surprisingly, the authors found that genes encoding glycolytic enzymes are required for growth even though the experiments were conducted in a rich medium that contains numerous compounds that could provide energy under aerobic conditions. This suggests that these enzymes may have unknown functions in the cell. About 50% of the genes of this minimal set are found in all bacteria, even in those with the smallest genomes.

As mentioned previously, interference of gene expression with antisense RNA constitutes another experimental approach to identify essential genes. In this case, gene function is inhibited by formation of a duplex RNA structure between the target mRNA and the antisense RNA introduced to the cell. Essential genes are identified after conditionally expressing random genomic fragments representing the entire genome linked to an inducible promoter, and then screening for those cloned fragments whose expression blocks growth. DNA sequencing and BLAST analysis against the annotated genome identify the genes targeted by antisense RNA. Forsyth *et al.* followed this strategy with *Staphylococcus aureus*, the most frequent causative agent of nosocomial infections in humans.[19] These authors found that out of the 2,595 protein coding genes in this microorganism, 658 are essential for growth. Homology comparison showed that 168 of these genes are found in the genome of *M. genitalium*. Interestingly, a similar approach followed by another group with the same bacterium showed that antisense inactivation of only 150 genes led to lethal or growth inhibitory effects.[20] The reason for this discrepancy is not known.

[18] Kobayashi, K., Ehrlich, S.D., *et al.*, 'Essential *Bacillus subtilis* genes', *Proc. Natl. Acad. Sci. USA* 100, 4678-4683, 2003.

[19] Forsyth, R.A., Haselbeck, R.J. *et al.*, 'A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*', *Mol. Microbiol*. 43, 1387-1400, 2002.

[20] Ji, Y., Zhang, B., Van Horn, S.F., Warren, P., Woodnutt, G., Burnham, M.K.R. and Rosenberg, M., 'Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA', *Science* 293, 2266-2269, 2001.

Comparison of Both Approaches and Concluding Remarks

Both theoretical and practical approaches for calculating minimal gene set numbers have some pitfalls.[21] First, computational methods rely on our ability to identify actual orthologues in species that are distantly related, as well as on filling some essential functions corresponding to NOGD. In spite of the general belief that we have identified the essential functions in a cell, it is expected that some cases of NOGD may remain hidden. In addition, different solutions may have evolved for each function in different organisms, giving rise to vast combinatorial possibilities. For these reasons, a minimal gene-set derived by computational comparison of genomes is probably an underestimate.

On the other hand, for technical reasons, the mutagenizing protocols described above are able to target only about 50% the genes, although this is considered statistically sufficient for reliable extrapolations. The knockout experiments also tend to score as essential genes that slow down microbial growth but do not abolish it, leading to a potential overestimation of the minimal gene-set. This possibility is partially compensated by the fact that knocking out of individual genes does not unveil the synthetic lethal genes. That is, those for which only simultaneous mutation results in a lethal phenotype.

In spite of these drawbacks, the number of essential genes reached by the computational and experimental approaches are in the same range. Moreover, the protein functions encoded in both minimal sets are strikingly similar, with a notorious enrichment of genes involved in the processing of genetic information, a rather low number of genes encoding metabolic enzymes and very few genes of unknown function.

Are further reductions possible? The work by Gil *et al.* described above provides copious arguments supporting a minimal genome of 206 genes. However, the functions considered by these authors are those known in extant cells. It is quite obvious that primitive cells must have been far more simple. For example, the protein synthesizing machinery may have contained less ribosomal proteins, or perhaps, proteins (enzymes) were more commonly multifunctional and therefore less specific. Today it is widely accepted that increased cell complexity is due to gene duplication. Indepen-

[21] Koonin, E.V., 'Comparative genomics: minimal gene-sets and the last universal common ancestor', *Nature Rev. Microbiol*. 1, 127-136, 2003.

dent mutations in the resulting paralogue[22] genes gradually results in the addition of more functions in the cell. As the matter of fact, gene duplication together with horizontal gene transfer are considered the main sources of microbial diversity in the biosphere. Therefore, analysis of the genes comprising a minimal genome may give information with respect to the key functions in primitive cells, since the simultaneous presence of orthologues in the three primary kingdoms of life is a sign of a function already present in a common ancestor.

De Novo *Synthesis of a Minimal Genome*

Once a minimal genome suitable for a specific set of environmental conditions is defined with a certain degree of confidence, an inevitable question emerges: could a functional genome be assembled that would allow metabolic activity and replication? This 'bottom up' approach, as opposed to the 'top down' strategy of knocking out genes of an existing microbe, entails an enormous task. The means for synthesizing small pieces of DNA are presently available, but assembling an entire genome including the regulatory sequences represents a challenge of considerable magnitude. This is mainly due to the possibility of contamination with truncated species and the introduction of typographical mistakes into the code. Interestingly, construction of synthetic genomes has been accomplished with small viral chromosomes such as poliovirus (7,440 bases)[23] and bacteriophage ϕX174,[24] whose infectious genome possesses 5,386 bp. The latter involved the use of new methods that significantly improved the speed and accuracy of genomic synthesis.

Craig Venter and Hamilton Smith are determined to construct a microorganism with a synthetic minimal genome. These same investigators conducted the studies of transposon mutagenesis with *Mycoplasma* and the synthesis of ϕX174.[25] Their approach involves selection of genes comprising

[22] Paralogs are genes in the same cell that derive from a single ancestral gene.

[23] Cello, J., Paul, A.V. and Wimmer, E., 'Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template', *Science* 297, 1016-1018, 2002.

[24] Smith, H.O., Hutchinson, C.A., Pfannkoch, C. and Venter, J.C., 'Generating a synthetic genome by whole genome assembly: ϕX174 bacteriophage from synthetic oligonucleotides', *Proc. Natl. Acad. Sci. USA* 100, 15440-15445, 2003.

[25] Needless to say, Venter had a leading role in the sequencing of the human genome and has thereafter conducted interesting metagenomic studies.

the minimal genome, chemical synthesis of all genes, stitching in some undetermined fashion and introduction into a bacterial cell whose own genome has been irreversibly damaged or destroyed. But Venter's ambition doesn't stop there. His goal is to use this minimal microorganism as a foundation for building cells harboring additional genes that would enable them to consume pollutants from the environment or to produce hydrogen fuel at an industrial scale. In 2003, he predicted that he would accomplish the task in three years.[26] Obviously, this has not yet happened, or at least, it has not been reported. Being aware that this project raises some ethical concerns, Venter requested an ethics committee from Stanford University to weigh the risks of creating new life forms. The panel acknowledged that there is a large technological gap between defining a minimal gene set and actually 'creating life', ruling out moral obstacles to assemble a new microorganism. However, it recommended responsible use of the new technology, since it could pose threats to public health and safety due to possible environmental contamination or the development of biological weapons.

As mentioned above, predictability in science allows us to foretell a given situation or event based on observation and reasoning. Can we predict whether scientists will ever create microbial life in the laboratory? This seems to be an extremely difficult prediction to make. Some years ago, this goal would have been considered science fiction. However, with the rapid development of novel and highly sophisticated technologies, perhaps no one would be in a position to predict that life will never by created *de novo* in the laboratory.

[26] Zimmer, C., 'Tinker Taylor: Can Venter stitch together a genome from scratch?', *Science* 299, 1006-1007, 2003.