

Understanding life, constructing life

Nikita Tirjatkin

Understanding life and constructing life are two fundamental problems at the very base of bio(techno)logical research and development respectively. These problems are intricately interwoven reflecting that science evolves by continuous exchange of facts and ideas between research and development. Understanding life is the only way to become able to construct it. Constructing life is the best way to improve life understanding. However, despite of much effort, life continues to be notoriously difficult to understand and construct. Here, I show that the investigation of life from information processing perspective makes life understanding sufficient enough for successful life constructing.

Without sufficient life understanding, successful life constructing is impossible. Unfortunately, most of the empirical and theoretical investigation in bio(techno)logical research and development is largely analytical. Although recent few decades were marked by very active implementation of synthetic approach (Andrianantoandro E *et al* 2006, Barrett CL *et al* 2006, Drubin DA *et al* 2007, Heinemann M and Panke S 2006, Ideker T *et al* 2001, Kitano H 2002, Mesarovic MD *et al* 2004, Westerhoff HV and Palsson BO 2004, Wolkenhauer O 2001), understanding life (Anbar M 2001, Bedau MA 1998, Dix DE 2002, Emmeche C 1997, Harold FM 2001, Kornberg A 1991, Penzlin H 2009) and constructing life (Andrianantoandro E *et al* 2006, Bedau MA *et al* 2000, Chopra P and Kamma A 2006, Drubin DA *et al* 2007, Endy D 2005, Heinemann M and Panke S 2006, Kim KJ and Cho SB 2006, Szostak JW *et al* 2001) remain unsolved fundamental problems.

The investigation of life from information processing perspective allows recognition in living world of complete hierarchy of universal life patterns and many important specific life patterns (Tirjatkin N 2005a, 2005b, 2005c, 2007, 2008a, 2008b). Here, I claim that the familiarity with these patterns is a prerequisite for both sufficient life understanding and successful life constructing.

Patterns of information processing in living world

Subcellularly, the information processing involves two tightly coupled reactions: genome expression and genome replication. During genome expression, information is converted first from DNA into RNA (transcriptome) form by DNA transcription, then from RNA into polypeptide (proteome) form by RNA translation, and finally from polypeptide into metabolite (metabolome) form by catalysis.

It is important to note that the genome is a limited set of genes and each gene is usually expressed separately to be fully converted into the corresponding element of the cell structure or function. For each gene, its own sequence DNA transcription -> RNA translation -> catalysis can be determined. This directed sequence of chemical reactions builds the most fundamental unit of information processing in the living world. This life pattern can be called the gene expression network (abbreviated GEN). Additionally, in some GENs, the obligatory sequence of chemical reactions can be restricted or

extended. So, in many GENs, end products are polypeptides functioning always as substrate molecules and never as catalysts. In many other GENs, end products are RNAs that never become translated into polypeptides, but function always at the level of RNA as substrate molecules. On the other hand, in many GENs, products of DNA transcription or RNA translation undergo post-transcriptional or post-translational processing respectively.

Since the genes are usually associated in a genome, the corresponding GENs are organised in more complicated unit of information processing in the living world. This life pattern can be called the genome expression network (abbreviated GENome). This life pattern is roughly equal to the cell. Whereas gene and genome are notions that refer to how information is stored in the living world, GEN and GENome refer to how the gene and genome work. The GENome can be considered as a highly regular composition of interacting GENs. During information processing in particular GEN, it is just the job of other GENs to provide necessary elements for gene expression machinery. Collectively, GENs in GENome work to replicate the complete genome so that the life history of the single cell begins with one cell but ends with two. Generally, the cell life history begins at the point where two newly produced sister cells halve the matrix inherited from the mother cell and each starts a self-dependent life. What the newborn cell has to do is just what its mother done: it starts its own genome expression which results in genome replication and in division in two daughter cells. In particular cell, the GENome is suited to specific subset of sources of mass, impulse (momentum), and energy to produce their usable forms essential for the cell life.

Thus, subcellularly, all chemical reactions are organized highly regular: first into gene expression networks and then into genome expression networks.

Supercellularly (supracellularly), the information processing involves other two important reactions: genome multiplication and genome diversification. Mechanism of genome multiplication is always the same: the genome replication by genome expression. On the contrary, mechanisms of genome diversification differ greatly ranging from the spontaneous sequence mutation to the highly regulated sequence transfer.

Progressive genome replication by genome expression leads to much more complicated unit of information processing in the living world. This life pattern can be called the genome multiplication network. Progressive genome replication is usually associated with progressive cell propagation producing a cell progression: one cell -> two cells -> four cells -> eight cells -> and so on... The entire living world is the only one cell progression which arose from one single primordial cell. However, the genome diversification produces cell progressions each of which is specified by a particular individual genome and can be called individual cell progression. Respectively, the entire living world can be considered as a growing composition of an increasing number of individual cell progressions. Spatiotemporal organization of a particular individual cell progression mostly depends upon whether the cells divide symmetrically or asymmetrically, whether the asymmetric cell divisions occur occasionally or regularly, whether the asymmetric cell division is associated with symmetric or asymmetric kinetics of the cell propagation, whether the cells will be rather randomly dispersed in

space to become autonomous in behaviour or remain in an association to form cell colony (primary, secondary, etc.), whether the cell association grows continuously or is a steady state system, and so on.

The entire living world is the only one cell progression which arose from one single primordial cell and has 3 or 4 billions years of uninterrupted history. It represents the most complicated unit of information processing in the living world. This life pattern can be called the genome diversification network or the general cell progression. The present-day biosphere is merely a tiny slice from it, a visible top of iceberg in ocean of time. The ancient part of this gigantic life pattern leaves very scarce traces.

Thus, supercellularly, chemical reactions are organized highly regular too: first into genome multiplication networks and then into the genome diversification network.

Understanding terrestrial life

Explicit recognition in the living world of complete hierarchy of universal life patterns and many important specific life patterns (Tirjatkin N 2005a, 2005b, 2005c, 2007, 2008a, 2008b) makes the understanding of life complexity and life diversity on the Earth amazingly easy.

The living world – an extremely complex network composed of huge numbers of different chemical reactions continuously creating an enormously complex matrix composed of bewildering numbers of different chemicals involved in these reactions – becomes at once comprehensible as soon as one becomes familiar with how such basic chemical reactions as DNA transcription, RNA translation, and catalysis arrange in strong hierarchy of life patterns shown in the table:

Table. Complete hierarchy of universal life patterns

Level	Life pattern...	... is roughly equal to:
4	Genome diversification network	General cell progression (living world or biosphere)
3	Genome multiplication network	Individual cell progression
2	Genome expression network (GENome)	Cell
1	Gene expression network (GEN)	

The general cell progression occupies the apex of the hierarchy. Most likely, it is unique and merits its own name (for example, Zoe). Other three life patterns in this hierarchy are doubtlessly universal. Their innumerable variations underlie the life diversity at corresponding levels.

Therefore, it is reasonable to have at least three taxonomic hierarchies (or genealogies): first for GENs, second for cells (GENomes), and third for individual cell progressions. It is important to note that, in addition to abstract taxonomic genealogies, such individual living things as GENs, cells (GENomes), and individual cell progressions also arrange in concrete genealogy and that this concrete genealogy is just the general

cell progression. According to the nature of universal life patterns in hierarchy, this concrete genealogy is one single genealogy of individual cell progressions at lower resolution, one single genealogy of cells (GENomes) at middle resolution, but a multiple N_1 -fold genealogy of GENs at higher resolution where N_1 is a number of genes in genome (or GENs in GENome) of the primordial cell from which the general cell progression arose. Both abstract and concrete genealogies may be presented in form of a tree-like drawing (dendrogram). Some individual cell progressions may produce cell associations in form of true trees.

If combined with taxonomic hierarchies embracing diversity of corresponding life patterns, the complete hierarchy of universal life patterns provides basic reference frame for secure orientation within the living world. Additionally, this basic reference frame is suited very well for ordering of innumerable specific life patterns: they either disclose specific reciprocal relations between some GENs within any cells (GENomes), between some cells within any individual cell progressions, or between some individual cell progressions within the general cell progression or disclose specific reciprocal relations between some universal life patterns and environment. In addition to conventional specific life patterns, many important specific patterns of information processing in living world have been recognized (Tirjatkin N 2005a, 2005b, 2005c, 2007, 2008a, 2008b).

Constructing life *in silico*

Explicit recognition in the living world of complete hierarchy of universal life patterns and many important specific life patterns (Tirjatkin N 2005a, 2005b, 2005c, 2007, 2008a, 2008b) makes life understanding sufficient enough for successful life constructing.

There are large numbers of life definitions (for review see, for example, Popa R 2004). However, no one can be used for life constructing. It is important to note that the non-living and living are undistinguishable from the perspective of mass, impulse, and energy processing. The difference becomes apparent only from the perspective of information processing: life originates by coupling of genome replication to genome expression and develops by continuous genome multiplication and genome diversification. This statement defines exactly what life is from information processing perspective. However, it is a minimal definition of life. By contrast, maximal definition of life must contain an extension pointing out that, after origin of the Life on the Earth, history of the Earth is inseparable from the history of the living world: it is obvious that the geosystems and biosystems coevolve and the geoecosystems and bioecosystems emerge at the interface of this coevolution (Tirjatkin N 2008a).

The proposed minimal definition reveals life as a special case of information processing and suggests that the simplest way to construct life is the programming.

Using knowledge on patterns of information processing in the living world, I have already obtained first positive results by means of the simplest C++ programs (www.nikita-tirjatkin.de). Each program must be compiled to produce an executable code which then conducts the corresponding experiment for *in silico* life construction

every time it runs. By contrast to known applications, these programs explicitly construct in the memory of computer genome-containing objects with genome-based information processing: gene expression networks, genome expression networks, etc. For simplicity, *in silico* life (InSiLi) is constructed by instantiation of uncomplicated object types. However, all these object types can be easily replaced by more realistic or – why not! – by more unrealistic object types depending on research and/or development aims.

Perspectives

InSiLi strongly supports the claim that the familiarity with patterns of information processing in the living world is a prerequisite for both sufficient life understanding and successful life constructing. Now, we understand life good enough to construct it at least *in silico*.

Although extremely simple, InSiLi can be anticipated as starting point for broad family of sophisticated research and development tools. Bridging the gap between life understanding and life constructing, these tools would put bio(techno)logical research and development in a new perspective. On the one hand, they might contribute to the improvement of life understanding until we become able to construct life *ex silico* too. On the other hand, they might contribute to the improvement of life quality itself. Indeed, knowledge on patterns of information processing provides real opportunity for development of a variety of *in silico*, *in vitro*, and/or *in vivo* models being suited for simulation of differential genome expression during normal or pathologic ontogenesis and histogenesis to discover relevant pathways of disease pathogenesis and reveal appropriate targets and agents for disease diagnosis and therapeutic interventions.

Being extremely simple, InSiLi is comprehensible for the broadest scientific and non-scientific audience suggesting that the life understanding and life constructing are much easier than expected. Using knowledge on patterns of information processing in the living world, everyone can construct her or his own version of *in silico* life and so contribute to the solution of the most exciting fundamental problems of bio(techno)logical research and development.

References

- Anbar M 2001. What is life? *IEEE Eng. Med. Biol. Mag.* **20**, 15-19.
- Andrianantoandro E *et al* 2006. Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* DOI: 10.1038/msb4100073.
- Barrett CL *et al* 2006. Systems biology as a foundation for genome-scale synthetic biology. *Curr. Opin. Biotechnol.* **17**, 488-492.
- Bedau MA 1998. Four puzzles about life. *Artif. Life* **4**, 125-140.
- Bedau MA *et al* 2000. Open problems in artificial life. *Artif. Life* **6**, 363-376.
- Chopra P and Kamma A 2006. Engineering life through synthetic biology. *In Silico Biol.* **6**, 401-410.
- Dix DE 2002. What is life? Prerequisites for a definition. *Yale J. Biol. Med.* **75**, 313-321.
- Drubin DA *et al* 2007. Designing biological systems. *Genes Dev.* **21**, 242-254.
- Emmeche C 1997. Defining life, explaining emergence. www.nbi.dk/~emmeche.
- Endy D 2005. Foundations for engineering biology. *Nature* **438**, 449-453.
- Harold FM 2001. Postscript to Schrödinger: So what is life? *ASM News* **67**, 611-616.
- Heinemann M and Panke S 2006. Synthetic biology – putting engineering into biology. *Bioinformatics* **22**, 2790-2799.

- Ideker T *et al* 2001. A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* **2**, 343-372.
- Kim KJ and Cho SB 2006. A comprehensive overview of the applications of artificial life. *Artif. Life* **12**, 153-182.
- Kitano H 2002. Systems biology: a brief overview. *Science* **295**, 1662-1664.
- Kornberg A 1991. Understanding life as chemistry. *Clin. Chem.* **37**, 1895-1899.
- Mesarovic MD *et al* 2004. Search for organising principles: understanding in systems biology. *Systems Biol.* **1**, 19-27.
- Penzlin H 2009. The riddle of “life,” a biologist’s critical view. *Naturwissenschaften*, **96**, 1-23.
- Popa R 2004. *Between necessity and probability: Searching for the definition and origin of life* (Springer, New-York).
- Szostak JW *et al* 2001. Synthesizing life. *Nature* **409**, 387-390.
- Tirjatkin N 2005a. Subcellular patterns of information processing. www.nikita-tirjatkin.de 14 July 2005.
- Tirjatkin N 2005b. Supercellular patterns of information processing. www.nikita-tirjatkin.de 14 July 2005.
- Tirjatkin N 2005c. Diversity of individual cell progressions in biosphere. www.nikita-tirjatkin.de 14 July 2005.
- Tirjatkin N 2007. Diversity of asymmetric cell progressions in Mammalia. www.nikita-tirjatkin.de 7 March 2007.
- Tirjatkin N 2008a. Complete hierarchy of universal life patterns. www.nikita-tirjatkin.de 15 May 2008.
- Tirjatkin N 2008b. Patterns of information processing in living world. www.nikita-tirjatkin.de 15 May 2008.
- Westerhoff HV and Palsson BO 2004. The evolution of molecular biology into systems biology. *Nature Biotechnol.* **22**, 1249-1252.
- Wolkenhauer O 2001. Systems biology: The reincarnation of systems theory applied in biology? *Brief. Bioinform.* **2**, 258-270.

Information to article

Written **14 March 2009**.

Revised **19 July 2010**.

Published online at www.nikita-tirjatkin.de **19 July 2010**.

© 2010 Nikita Tirjatkin

This is an Open Access Article. Unrestricted use, reproduction, and distribution are permitted in all media for any purpose, provided the original work is properly cited.