The theory of evolution in the light of 21st century’s science

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Österreich
und
The Santa Fe Institute, Santa Fe, New Mexico, USA

Conference on Evolutionism and Religion

Florence, 19.-21.11.2009
Web-Page for further information:

http://www.tbi.univie.ac.at/~pks
Populations adapt to their environments through multiplication, variation, and selection – Darwin's natural selection.

All forms of (terrestrial) life descend from one common ancestor – phylogeny and the tree of life.
1. Darwin's natural selection

2. The tree of life

3. From evolution *in vitro* to biotechnology

4. Genotypes with multiple functions

5. How complex is biology?
1. Darwin’s natural selection

2. The tree of life

3. From evolution *in vitro* to biotechnology

4. Genotypes with multiple functions

5. How complex is biology?
Genotype, Genome

Collection of genes

Developmental program

Unfolding of the genotype

Highly specific environmental conditions

Phenotype

Evolution explains the origin of species and their interactions
Three necessary conditions for Darwinian evolution are:

1. Multiplication,
2. Variation, and
3. Selection.

*Variation* through mutation and recombination operates on the *genotype* whereas the *phenotype* is the target of *selection*.

One important property of the Darwinian scenario is that *variations* in the form of mutations or recombination events occur *uncorrelated* with their effects on the selection process.
Two variants with a mean progeny of ten or eleven descendants

\[ s = \frac{f_2 - f_1}{f_1} = 0.1 \]
Selection of advantageous mutants in populations of $N = 10000$ individuals

$N_1(0) = 9999, N_2(0) = 1; \quad s = 0.1, 0.02, 0.01$
1. Darwin's natural selection

2. The tree of life

3. From evolution *in vitro* to biotechnology

4. Genotypes with multiple functions

5. How complex is biology?
Genotype, Genome

Quantitative biology

‘the new biology is the chemistry of living matter’

Biochemistry
molecular biology
structural biology
molecular evolution
molecular genetics
systems biology
bioinformatics
epigenetics

Highly specific environmental conditions

Unfolding of the genotype

Phenotype

evolution of RNA molecules, ribozymes and splicing, the idea of an RNA world, selection of RNA molecules, RNA editing, the ribosome is a ribozyme, small RNAs and RNA switches.

The exciting RNA story

Molecular evolution
Linus Pauling and Emile Zuckerkandl

Hemoglobin sequence
Gerhard Braunitzer

Manfred Eigen

Max Perutz

John Kendrew

James D. Watson und Francis H.C. Crick
The geometry of the double helix is compatible only with the base pairs:

**AT, TA, CG, and GC**
The structure of DNA suggests a mechanism for reproduction
The logics of DNA replication
The molecular mechanism of mutation

point mutation

Adenine  Thymine  Guanine  Cytosine
Molecular phylogeny
Motoo Kimuras population genetics of neutral evolution.


What is neutrality?

Selective neutrality =
= several genotypes having the same fitness.

Several genotypes \Rightarrow \text{ one phenotype}
The molecular clock of evolution

Results from molecular evolution:

• The molecular machineries of all present day cells are very similar and provide a strong hint that all life on Earth descended from one common ancestor (called „last universal common ancestor“, LUCA).

• Comparison of DNA sequences from present day organisms allows for a reconstruction of phylogenetic trees, which are (almost) identical with those derived from morphological comparison of species and the paleontologic record of fossils.
1. Darwin's natural selection
2. The tree of life
3. From evolution *in vitro* to biotechnology
4. Genotypes with multiple functions
5. How complex is biology?
Three necessary conditions for Darwinian evolution are:

1. Multiplication,
2. Variation, and
3. Selection.

Variation through mutation and recombination operates on the genotype whereas the phenotype is the target of selection.

One important property of the Darwinian scenario is that variations in the form of mutations or recombination events occur uncorrelated with their effects on the selection process.

All conditions can be fulfilled not only by cellular organisms but also by nucleic acid molecules in suitable cell-free experimental assays.
Evolution of RNA molecules based on Qβ phage


Evolution in the test tube:

Application of serial transfer technique to evolution of RNA in the test tube

RNA sample

Stock solution: Qβ RNA-replicase, ATP, CTP, GTP and UTP, buffer
Decrease in mean fitness due to quasispecies formation

The increase in RNA production rate during a serial transfer experiment
Mutation and (correct) replication as parallel chemical reactions

M. Eigen. 1971. *Naturwissenschaften* 58:465,
Chemical kinetics of molecular evolution
Quasispecies

Driving virus populations through threshold

The error threshold in replication
Molecular evolution of viruses
Evolutionary design of RNA molecules


C. Tuerk, L. Gold, **SELEX - Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.** Science **249** (1990), 505-510


Y. Wang, R.R. Rando, *Specific binding of aminoglycoside antibiotics to RNA.* Chemistry & Biology **2** (1995), 281-290

An example of ‘artificial selection’ with RNA molecules or ‘breeding’ of biomolecules.
tobramycin

Formation of secondary structure of the tobramycin binding RNA aptamer with $K_D = 9 \text{ nM}$

The three-dimensional structure of the tobramycin aptamer complex

Application of molecular evolution to problems in biotechnology
Artificial evolution in biotechnology and pharmacology


Results from laboratory experiments in molecular evolution:

• Evolutionary optimization does not require cells and occurs in molecular systems too.

• *In vitro* evolution allows for production of molecules for predefined purposes and gave rise to a branch of biotechnology.

• Direct evidence that neutrality is a major factor for the success of evolution.
1. Darwin's natural selection
2. The tree of life
3. From evolution \textit{in vitro} to biotechnology
4. \textit{Genotypes with multiple functions}
5. How complex is biology?
What is conformational multiplicity?

Conformational multiplicity = several structures formed by one sequence.

One genotype $\Rightarrow$ several phenotypes
Extension of the notion of structure
An experimental RNA switch


Two ribozymes of chain lengths $n = 88$ nucleotides: An artificial ligase (A) and a natural cleavage ribozyme of hepatitis-$\delta$-virus (B)
The sequence at the **intersection**:

An RNA molecule which is 88 nucleotides long and can form both structures.
Two neutral walks through sequence space with conservation of structure and catalytic activity
The thiamine-pyrophosphate riboswitch

M. Mandal, B. Boese, J.E. Barrick, W.C. Winkler, R.R, Breaker.
1. Darwin’s natural selection
2. The tree of life
3. From evolution *in vitro* to biotechnology
4. Genotypes with multiple functions
5. How complex is biology?
Three-dimensional structure of the complex between the regulatory protein **cro-repressor** and the binding site on λ-phage **B-DNA**
A model genome with 12 genes

Sketch of a genetic and metabolic network
Biochemical Pathways

The reaction network of cellular metabolism published by Boehringer-Mannheim.
The citric acid or Krebs cycle (enlarged from previous slide).
The bacterial cell as an example for the simplest form of autonomous life

Escherichia coli genome:
4 million nucleotides
4460 genes

The structure of the bacterium *Escherichia coli*
**E. coli:**  Genome length \(4 \times 10^6\) nucleotides
   Number of cell types 1
   Number of genes 4,460

**Man:**  Genome length \(3 \times 10^9\) nucleotides
   Number of cell types 200
   Number of genes \(\approx 30,000\)

Complexity in biology
Evolution does not design with the eyes of an engineer, evolution works like a tinkerer.

François Jacob. The Possible and the Actual. Pantheon Books, New York, 1982, and

A model for the genome duplication in yeast 100 million years ago

A model for the genome duplication in yeast 100 million years ago

A model for the genome duplication in yeast 100 million years ago

A model for the genome duplication in yeast 100 million years ago

The difficulty to define the notion of “gene”.

Helen Pearson, Nature 441: 399-401, 2006

The idea of genes as beads on a DNA string is fast fading. Protein-coding sequences have no clear beginning or end and a DNA string is a key part of the information package, reports Helen Pearson.

Gene is not a typical four-letter word. It is not offensive. It is never dropped out of TV shows. And the meaning of most four-letter words is too clear, that of gene is not.

The most expert scientists become in molecular genetics, the less easy it is to be sure about what, if anything, a gene actually is.

Rick Young, a geneticist at the Whitehead Institute in Cambridge, Massachusetts, says that when he first started teaching as a young professor two decades ago, it took him about two hours to teach fresh-faced undergraduates what a gene was and the nuts and bolts of how it worked. Today, he and his colleagues need three months of lectures to convey the concept of the gene, and that is not because the students are any less bright. “It takes a whole semester to teach this stuff to talented undergraduates,” Young says. “It used to be we could give a one-off definition and now it’s much more complicated.”

In classical genetics, a gene was an abstract concept — a unit of inheritance that ferried a characteristic from parent to child. As biochemistry came into its own, those characteristics were associated with enzymes or proteins, one for each gene. And with the advent of molecular biology, genes became real, physical things — sequences of DNA which when converted into strands of so-called messenger RNA could be used as the basis for building the associated protein piece by piece. The great coiled DNA molecules of the chromosomes were seen as long strings on which gene sequences sat like beads on a string.

This picture is still the working model for many scientists. But those at the forefront of genetic research see it as increasingly flimsy — a crude approximation that, at best, hides fascinating new complexities and, at worst, blinds us to useful new paths of enquiry.

Information, it seems, is parcelled out along chromosomes in a much more complex way than was originally supposed. DNA molecules are not just passive conduits through which the gene’s message flows into the world but active regulators of cellular processes. In some cases, RNA may pass information across generations — directly from the plant to the seed of its next generation. Newly emerging RNA machines have been observed to recruit proteins that act as RNA activators, interacting with DNA sequences to regulate gene expression.

This is a fundamental change in the way we think about genes. The idea that a gene is a single, linear block of DNA that codes for a single protein, with no regulatory mechanisms or overlapping sequences, is no longer sufficient. Instead, we must consider a much more complex reality, where genes can overlap, interact with other genes, and be regulated by RNA molecules.

The challenge for scientists is to understand how these overlapping sequences and regulatory mechanisms work together to control gene expression. This is a difficult task, as the complexity of gene regulation is not well understood.

Still, there are some promising developments in this area. For example, recent studies have shown that RNA molecules can be used to control gene expression in a process known as RNA interference. This process involves the production of small RNA molecules that can bind to specific sequences of DNA and prevent the expression of those genes.

As we continue to learn more about the complexity of gene regulation, we will undoubtedly discover new and exciting ways to control gene expression and potentially treat a wide range of diseases.

References:

---

Spook of DNA (above) fills harbours with new protein-coding genes often overlapping the next.
ENCODE stands for ENCYclopedia Of DNA Elements.

Coworkers

Peter Stadler, Bärbel M. Stadler, Universität Leipzig, GE

Paul E. Phillipson, University of Colorado at Boulder, CO

Heinz Engl, Philipp Kügler, James Lu, Stefan Müller, RICAM Linz, AT

Jord Nagel, Kees Pleij, Universiteit Leiden, NL

Walter Fontana, Harvard Medical School, MA

Martin Nowak, Harvard University, MA

Christian Reidys, Nankai University, Tien Tsin, China

Christian Forst, Los Alamos National Laboratory, NM

Thomas Wiehe, Ulrike Göbel, Walter Grüner, Stefan Kopp, Jaqueline Weber, Institut für Molekulare Biotechnologie, Jena, GE

Ivo L. Hofacker, Christoph Flamm, Andreas Svrček-Seiler, Universität Wien, AT

Kurt Grünberger, Michael Kospach, Andreas Wernitznig, Stefanie Widder, Stefan Wuchty, Jan Cupal, Stefan Bernhart, Lukas Endler, Ulrike Langhammer, Rainer Machine, Ulrike Mückstein, Erich Bornberg-Bauer, Universität Wien, AT
Acknowledgement of support

Fonds zur Förderung der wissenschaftlichen Forschung (FWF)
Projects No. 09942, 10578, 11065, 13093
13887, and 14898

Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF)
Project No. Mat05

Jubiläumsfonds der Österreichischen Nationalbank
Project No. Nat-7813

European Commission: Contracts No. 98-0189, 12835 (NEST)

Austrian Genome Research Program – GEN-AU: Bioinformatics
Network (BIN)

Österreichische Akademie der Wissenschaften

Siemens AG, Austria

Universität Wien and the Santa Fe Institute
Thank you for your attention!
Web-Page for further information:

http://www.tbi.univie.ac.at/~pks