Karl Valentin

All has been said already, but not yet by everyone.

3.8 billion years of RNA

RNA.world@uni-muenster.de

Salzburg, 2014

J. Mol. Biol. (1968) **38**, 367–379

The Origin of the Genetic Code

F. H. C. Crick

Medical Research Council
Laboratory of Molecular Biology
Hills Road, Cambridge, England

(*Received 21 August 1968*)

The general features of the genetic code are described. It is considered that originally only a few amino acids were coded, but that most of the possible codons were fairly soon brought into use. In subsequent steps additional amino acids were substituted when they were able to confer a selective advantage, until eventually the code became frozen in its present form.

J. Mol. Biol. (1968) **38**, 381–393

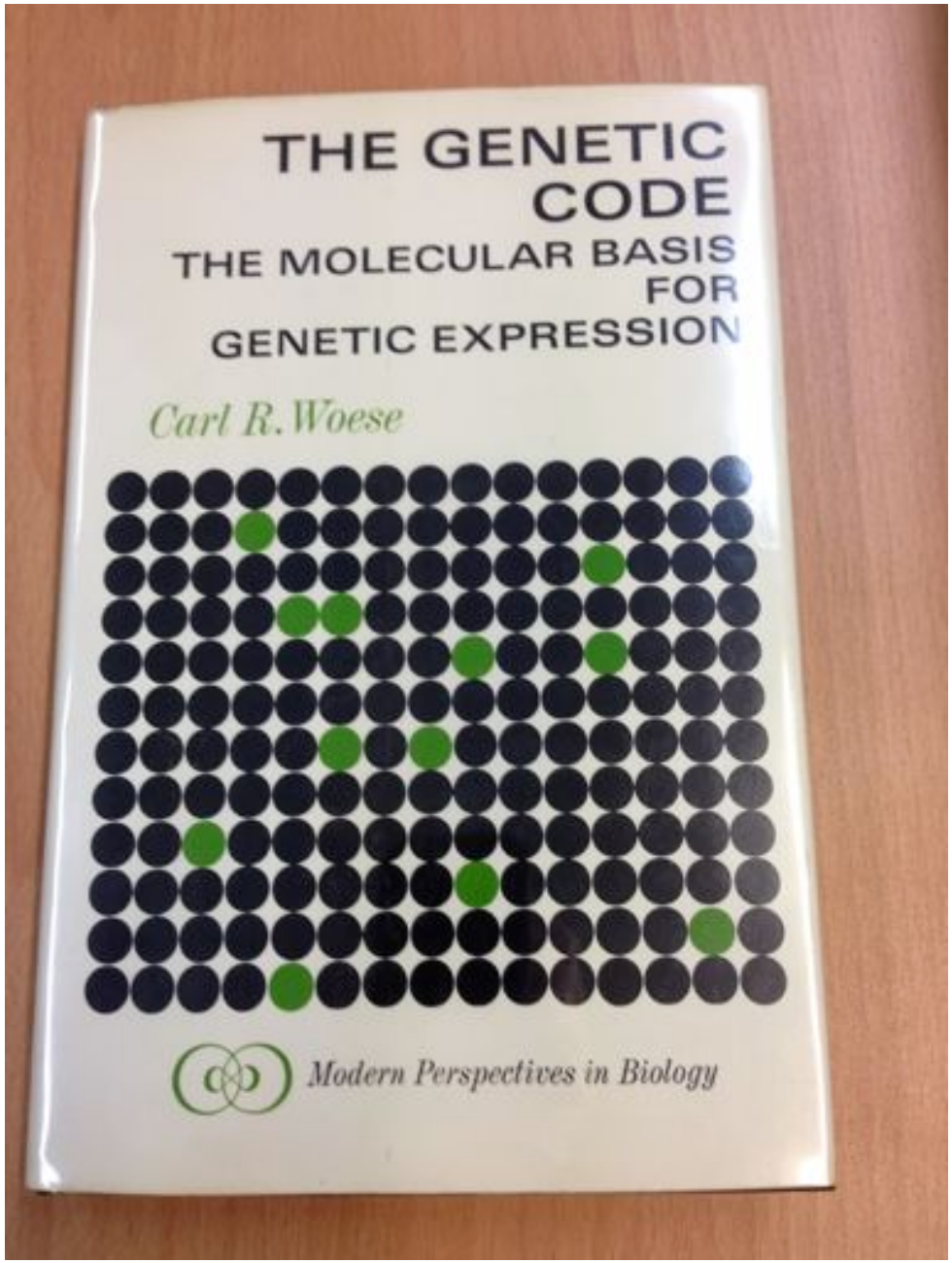
Evolution of the Genetic Apparatus

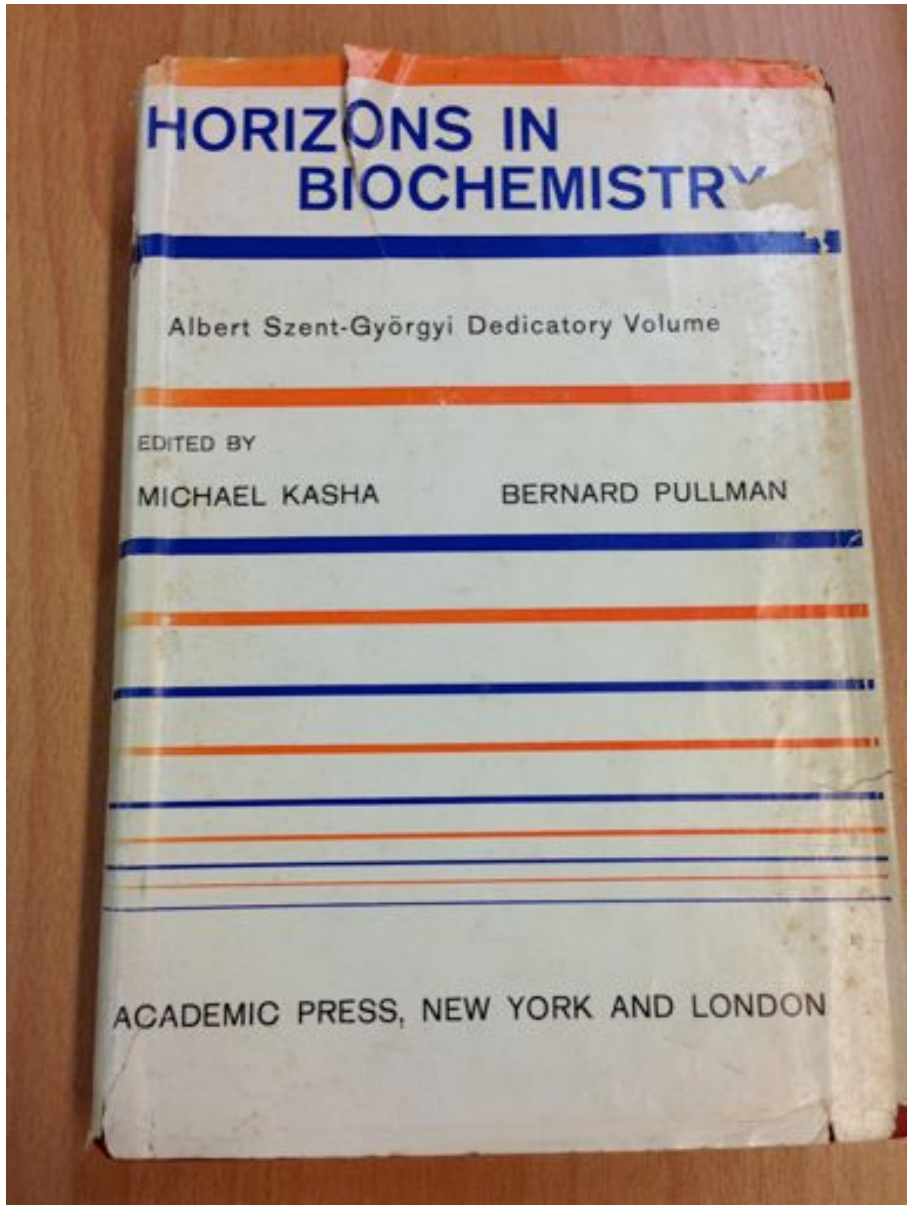
L. E. Orgel

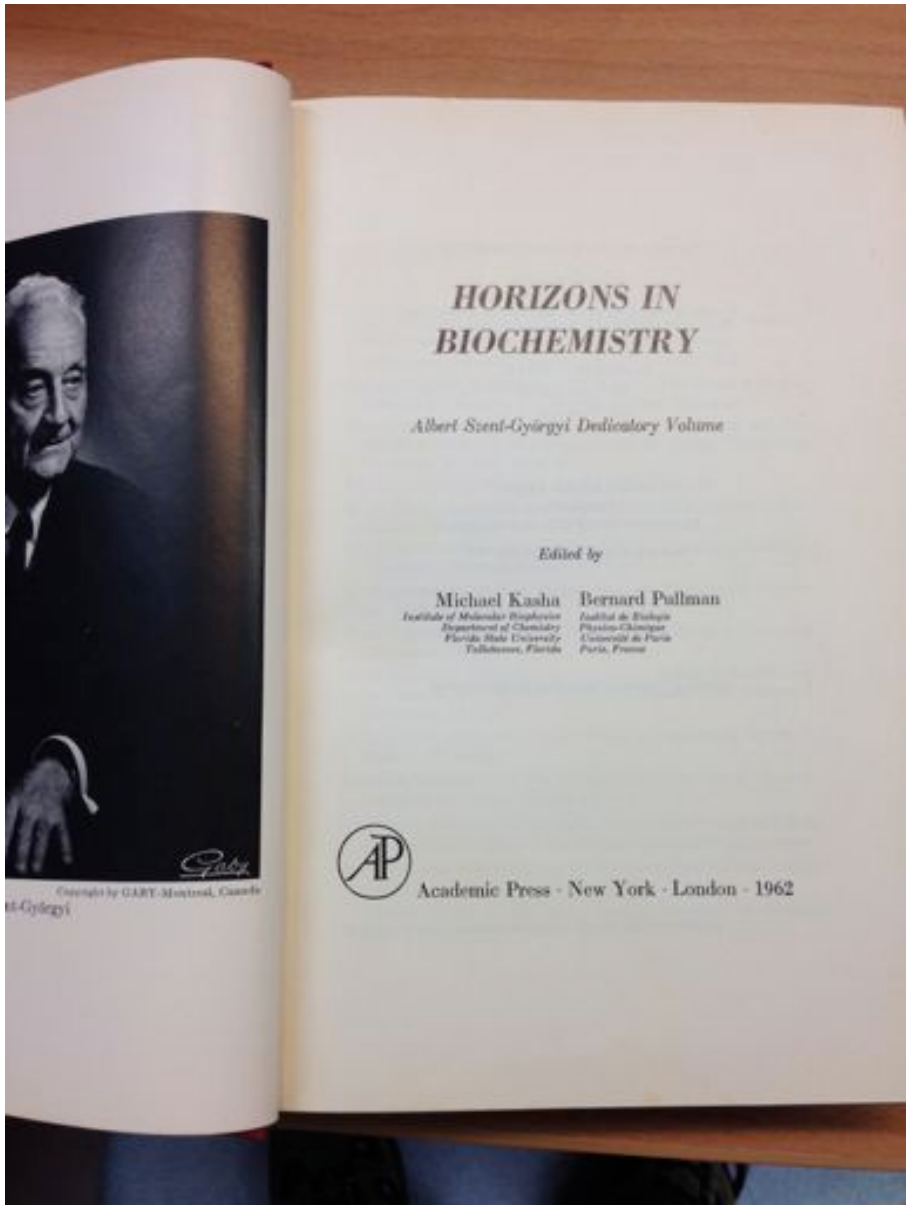
The Salk Institute for Biological Studies
San Diego, California 92112, U.S.A.

(*Received 21 August 1968*)

It is argued that the evolution of the genetic apparatus must have required the abiotic formation of macromolecules capable of residue-by-residue replication. This suggests that polynucleotides were present even in the most primitive ancestors of contemporary organisms. Models which explain the evolution of the association between polynucleotide and polypeptide sequences are discussed.


THE GENETIC CODE
THE MOLECULAR BASIS FOR GENETIC EXPRESSION
Carl R. Woese
Modern Perspectives in Biology





HORIZONS IN BIOCHEMISTRY

Albert Szent-Györgyi Dedicatory Volume

Edited by

Michael Kasha
Institute of Molecular Biophysics
Department of Chemistry
Florida State University
Tallahassee, Florida

Bernard Pullman
Institut de Biologie
Physico-Chimique
Université de Paris
Paris, France

Academic Press · New York · London · 1962

On the Problems of Evolution and Biochemical Information Transfer

ALEXANDER RICH

Department of Biology,
Massachusetts Institute of Technology,
Cambridge, Massachusetts

I. Introduction

Perhaps the most striking characteristic in the development of biochemical understanding during the past decade has been the discovery that nucleic acids play the central role in the transmission of molecular information. Although a great deal of data had been compiled on pathways of intermediary metabolism by the early 1950's, relatively little was known about those reactions which gov-

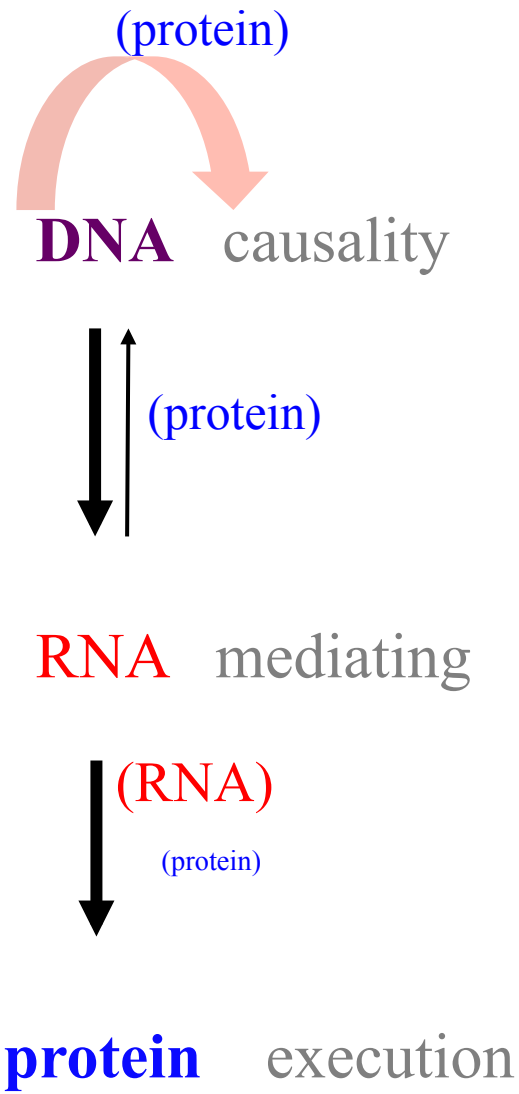
Here is the greatest weakness in a theory of this type, since it does not really explain this most difficult step, the evolution of nucleic acid-controlled protein synthesis. However, another type of theory can be described as the basis for the origin of life which places the nucleic acids in a more central role.

V. Polynucleotides as the Origin of Living Systems

Theories of the biochemical origin of life on this planet were seriously developed about 20 to 30 years ago. At that time, it was quite clear that proteins were the most characteristic molecules in living systems and that their specific catalytic properties were essential to the functioning of biochemical systems. Accordingly much attention was devoted to ways of developing primitive protein molecules through nonliving agents. However, the almost explosive development of our understanding about the role of the nucleic acids during the past decade has made it imperative for us to reformulate theories about the origin of life in order to place the nucleic acids in proper perspective. As discussed above, the sequence of amino acids in proteins is, in a sense, a derivative of the sequence information encoded in the order of nucleotides in some part of the nucleic acids. Accordingly, it may be more reasonable to consider a theory of the origin of life in which the nucleic acids were developed as the primary agents.

Here we imagine a large number of nucleotide monomers among other molecules which are floating freely in a primitive sea, having been created by chemical reactions promoted by the products of ultraviolet discharges. We postulate that these

Central Dogma of Molecular Biology

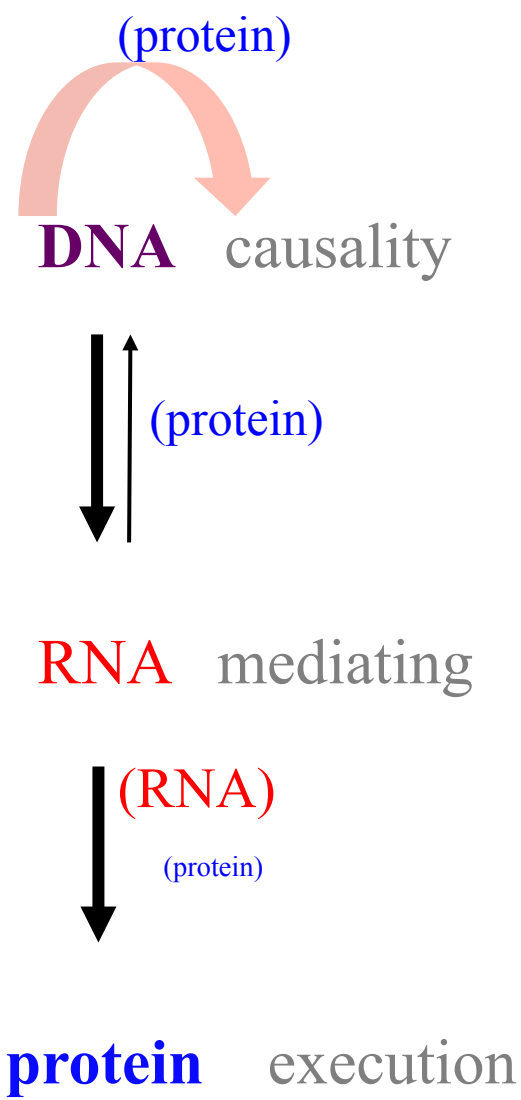


X. Why Are There Two Nucleic Acids?

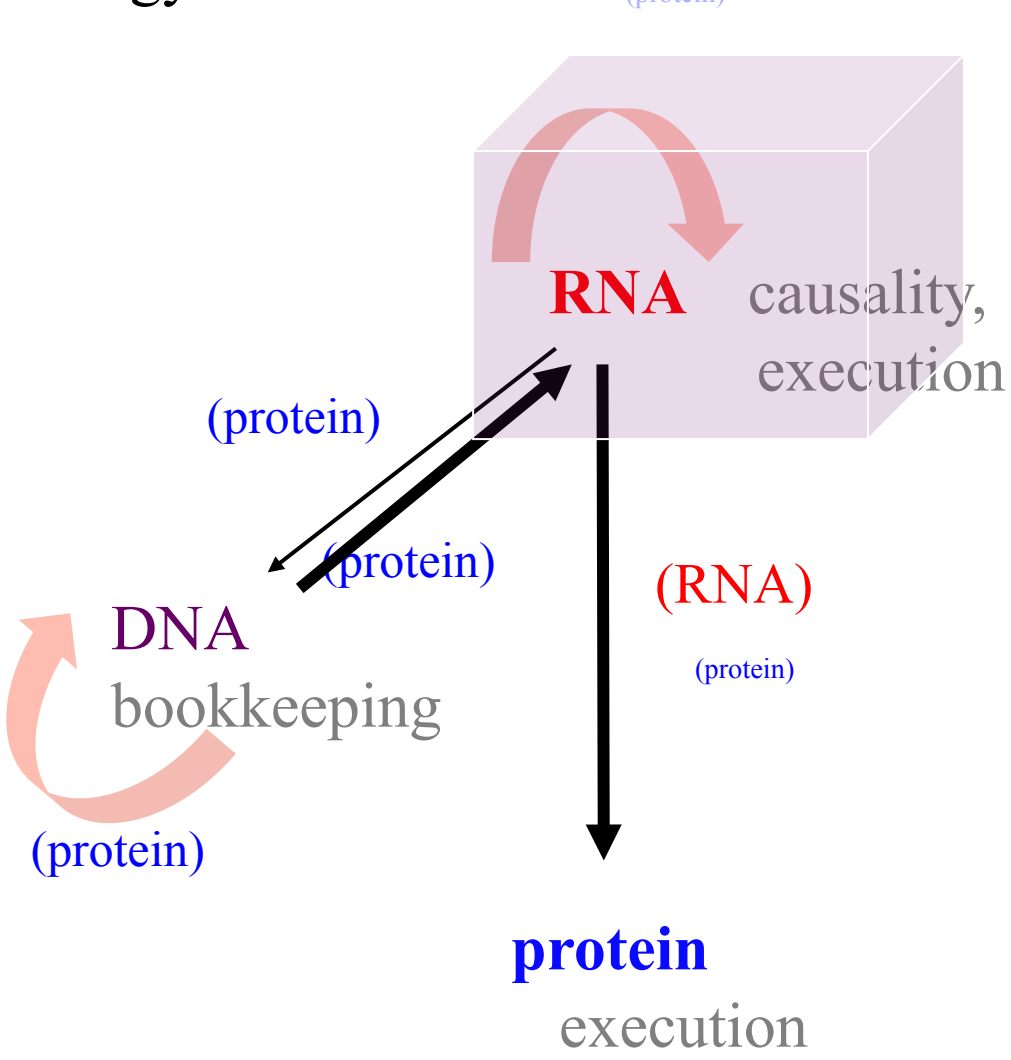
It is quite remarkable that contemporary biochemical systems have two nucleic acids, DNA and RNA, which differ only by a systematic hydroxyl group and an occasional methyl group. Despite the great chemical similarities, the molecules nevertheless have quite different functions in the cell. DNA acts as the major carrier of genetic information, while the RNA molecule is used to convert this genetic information into actual protein molecules. Because of the close chemical similarities, we are tempted to ask whether they could have originated historically from a common stem nucleic acid molecule which then specialized in the course of evolution to produce the two different classes of nucleic acids which we see today. To discuss this further we should note that the RNA molecule is also able to carry genetic information, as, for example, in the RNA-containing viruses. Thus, it may be reasonable to speculate that the

hypothetical stem or parent polynucleotide molecule was initially an RNA-like polymer which was able to convey genetic information as well as organize the amino acids into a specific sequence to make proteins. This implies that the RNA polynucleotide strand had the ability to replicate itself and produce a complement in a manner somewhat similar to that which is found in DNA. It is possible that an enzyme of this type may have been observed already. By this view, DNA may be regarded as a derivative molecule which has evolved in a form such that it only carries out part of the primitive nucleic acid function. It specialized in the molecular replicating cycle that is part of the mechanism for transmitting genetic information. DNA is metabolically less reactive than RNA, perhaps because of the absence of the hydroxyl group on carbon 2. The loss of this hydroxyl group may have made it impossible for the DNA molecule to have attached to it the amino acids which are used in protein synthesis. However, considerable selective advantage may be derived from the development of two different classes of nucleic acids, one of which is less active metabolically and specializes in self-replication. In a sense, this tends to preserve the primary copy of the genetic information. It will be of considerable interest to study the available simple life forms to see whether some of them may exist with only one type of nucleic acid rather than two types. It is possible that the RNA containing viruses may be regarded as present-day examples which may have degenerated evolutionarily from such a primitive life form.

Central Dogma of Molecular Biology



Revisited

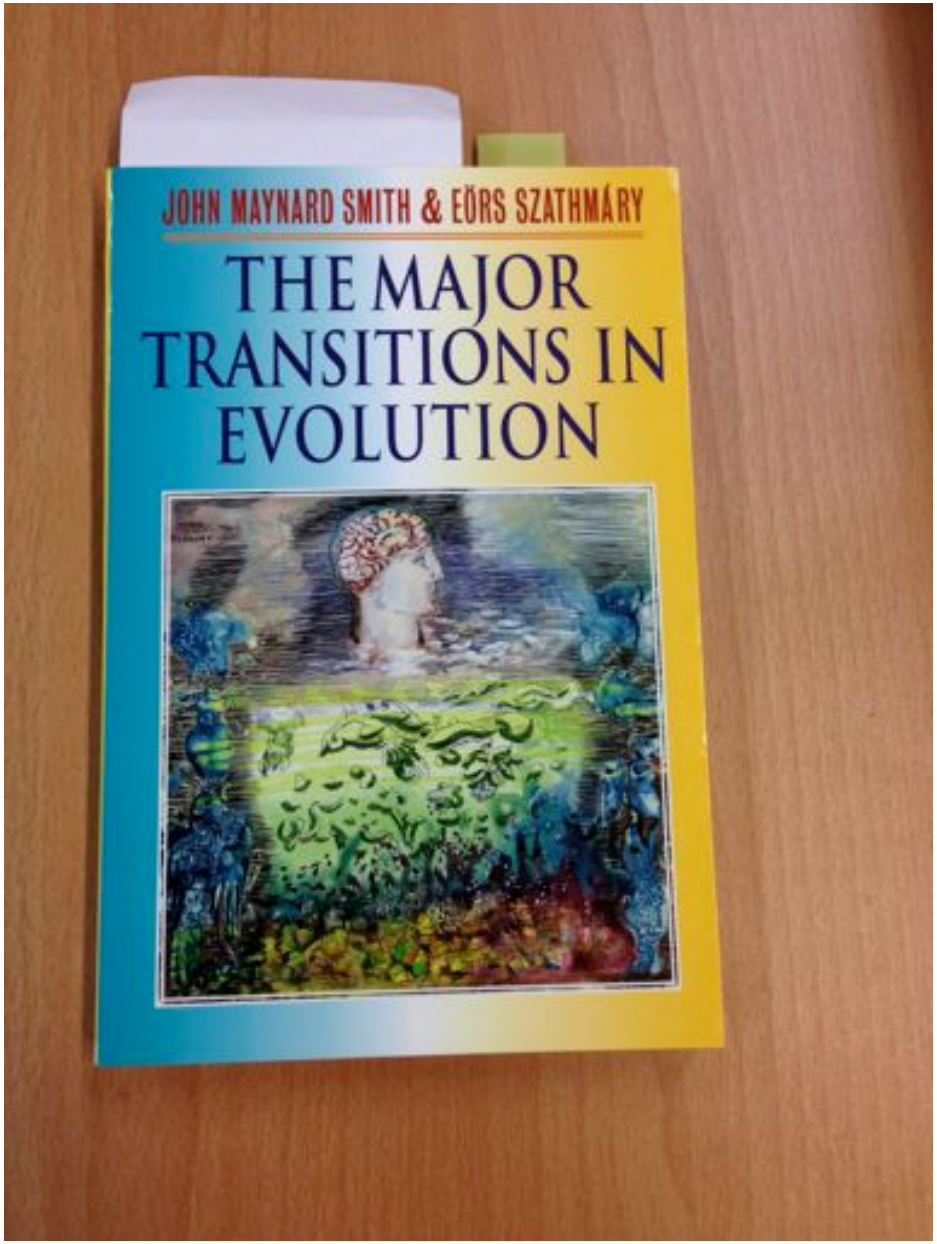


nucleotides could only define 16 amino acids. Since there are about 20 major amino acids, it is believed that at least three nucleotides are needed in this process. Recent genetic experiments by Crick and his collaborators (1961) strongly suggest that the coding ratio is three. However, it is quite likely that the ratio will be determined in the very near future using systems such as those described above with synthetic Messenger RNA. In this regard it should be pointed out that although it is commonly assumed that the coding ratio is the same for all amino acids, this is not necessary, since Nature may have developed a system in which amino acids fall into different classes utilizing a different number of nucleotides for coding the different classes. However, this probably is unlikely in view of the genetic experiments cited above (Crick *et al.*, 1961).

We have mentioned the possibility that Messenger RNA may be made *in vivo* as complementary copies of one or both strands of DNA. If both strands are active, then the DNA would produce two RNA strands which are complementary to each other. Only one of these might be active in protein synthesis, and the other strand might be a component of the control or regulatory system. However,

which may prove of utility in understanding and interpreting the results.

We have a considerable body of information about macroscopic evolution in terms of the changes, with time, of organisms and the development of new species. From this we see that the mechanism of evolution is such that slow refinements are usually made within a given organism in relation to its environment which lead toward increased efficiency of reproduction. Occasionally there are discontinuities which are opportunistically utilized in the exploitation of a new environment or of an altered physiological function. In general, the trend in macroscopic organic evolution is toward increased complexity. We can trace the development of increasingly subtle physiological functions and greater control leading to a progressive extension of the organism's ability to utilize the environment and maximize reproductive ability.


JOHN MAYNARD SMITH & EÖRS SZATHMÁRY
THE MAJOR TRANSITIONS IN EVOLUTION

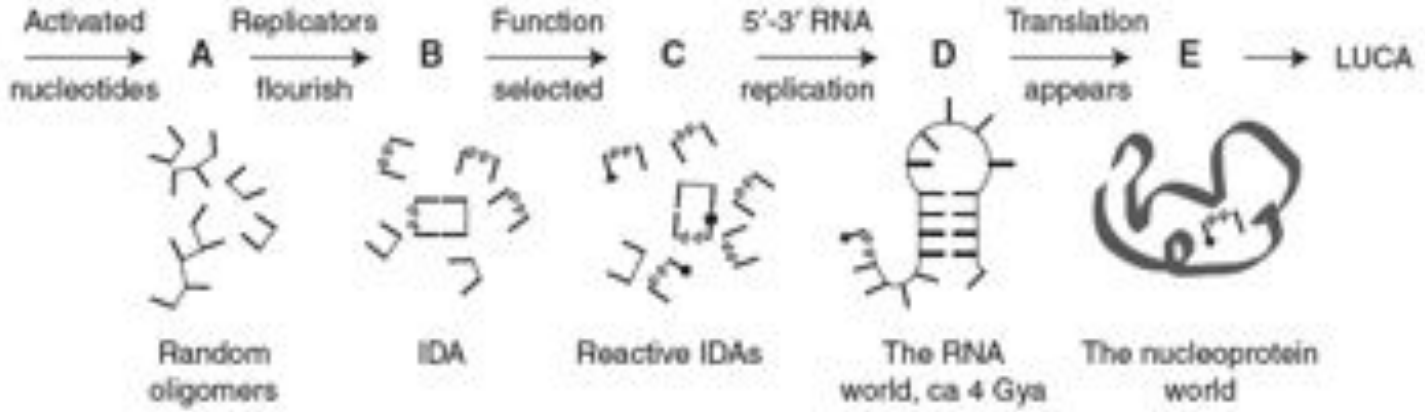


Figure 2. The IDA in context; the origin of life. (*A*) Activated nucleotides and compatible molecules oligomerize arbitrarily. (*B*) Replicators necessarily become abundant, by templating with minimal catalysis. (*C*) 5′-5′ replicators with reactive nucleotides are selected to participate in metabolism. (*D*) 5′-3′ RNA replicase creates an RNA world, ± 5′-5′ cofactor initiation and reactivity. (*E*) RNAs devise translation; 5′-5′ cofactors are adopted by peptide catalysts (ribbon). The pathway is initiated by its most complex event, geochemical creation of several activated nucleotide-like materials. Thus, while "simple" is a debatable evolutionary characterization, progress might be relatively simple once begun. After panel (*B*), all crucial transitions depend on somewhat similar selections for enhanced chemical proficiency.

Cite as *Cold Spring Harb Perspect Biol* 2011;3:a003590

Michael Yarus

IDA = Initial Darwinian Ancestor

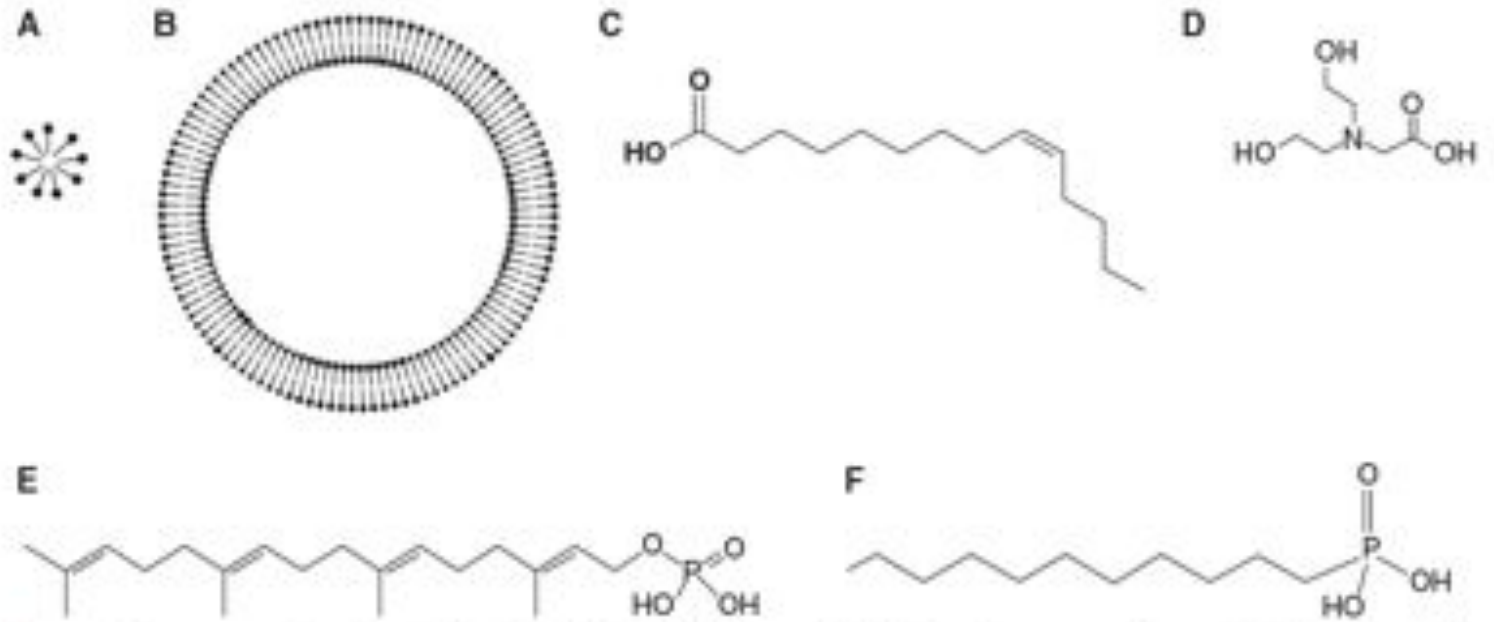


Figure 2. Structures of prebiotically plausible single chain amphiphiles and a commonly used buffer. (*A*) micelle; (*B*) vesicle; (*C*) myristoleic acid; (*D*) bicine; (*E*) geranylgeranyl phosphoric acid; (*F*) *n*-decylphosphonic acid.

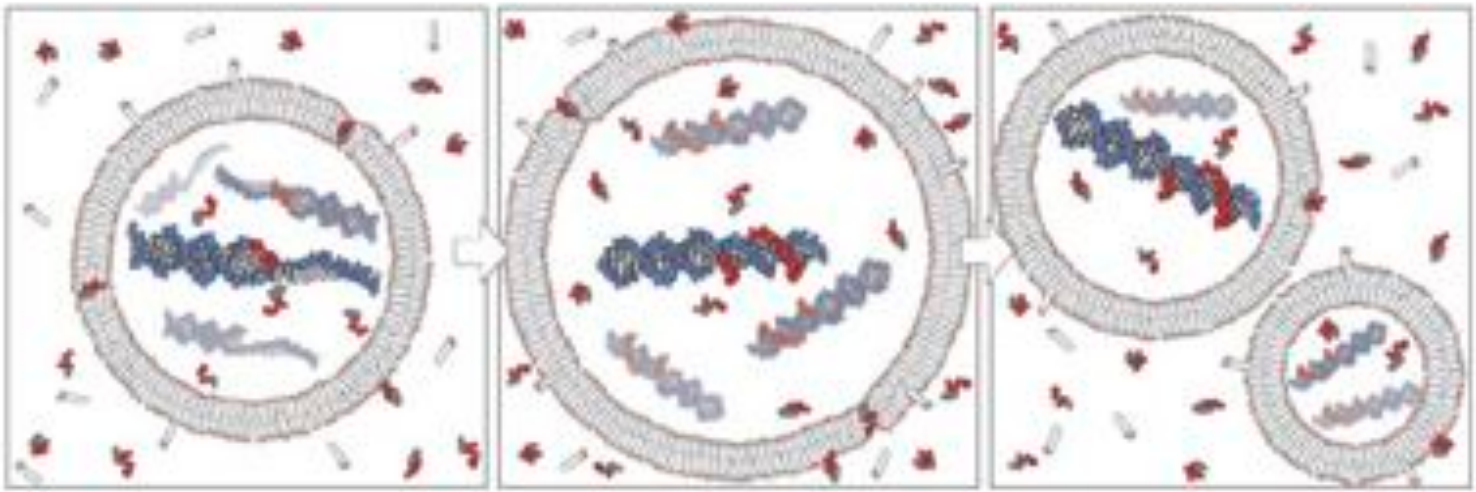


Figure 1. A simple protocell model based on a replicating vesicle for compartmentalization, and a replicating genome to encode heritable information. A complex environment provides lipids, nucleotides capable of equilibrating across the membrane bilayer, and sources of energy (*left*), which leads to subsequent replication of the genetic material and growth of the protocell (*middle*), and finally protocellular division through physical and chemical processes (*right*). (Reproduced from Mansy et al. 2008 and reprinted with permission from Nature Publishing ©2008.)

Orig Life Evol Biosph (2008) 38:329–341
DOI 10.1007/s11084-008-9131-8

PRIMITIVE MEMBRANES, SELF-ASSEMBLY

Stability of Model Membranes in Extreme Environments

Trishool Namani · David W. Deamer

Orig Life Evol Biosph (2008) 38:57–74
DOI 10.1007/s11084-007-9113-2

Lipid-assisted Synthesis of RNA-like Polymers from Mononucleotides

Sudha Rajamani · Alexander Vlassov · Seico Benner · Amy Coombs · Felix Olasagasti · David Deamer

Vol 454|3 July 2008 nature

NEWS & VIEWS

ORIGINS OF LIFE

How leaky were primitive cells?

David W. Deamer

If the first cells were simple vesicles, how did nutrients cross their membranes without help from transport proteins? A model of a primitive cell suggests that early membranes were surprisingly permeable.

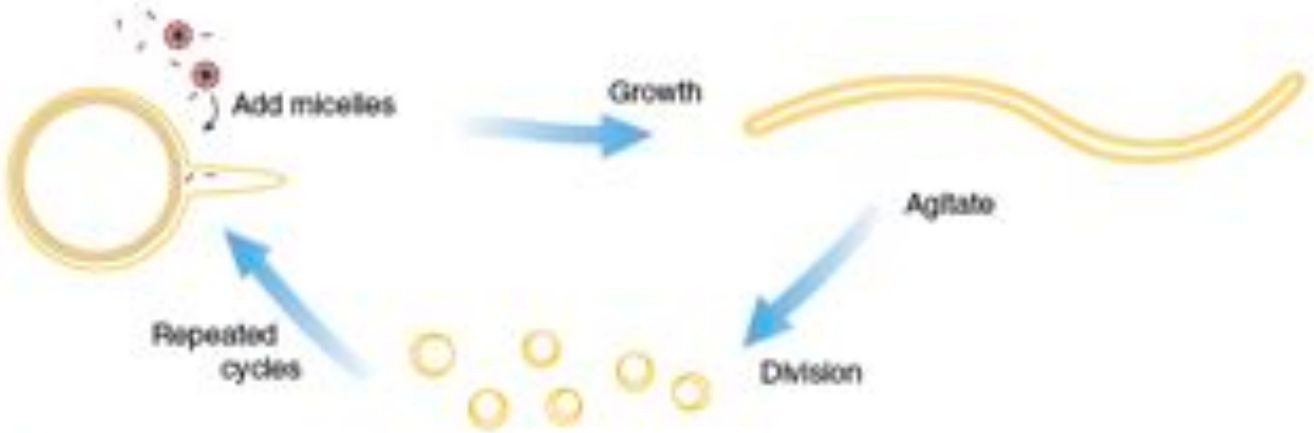
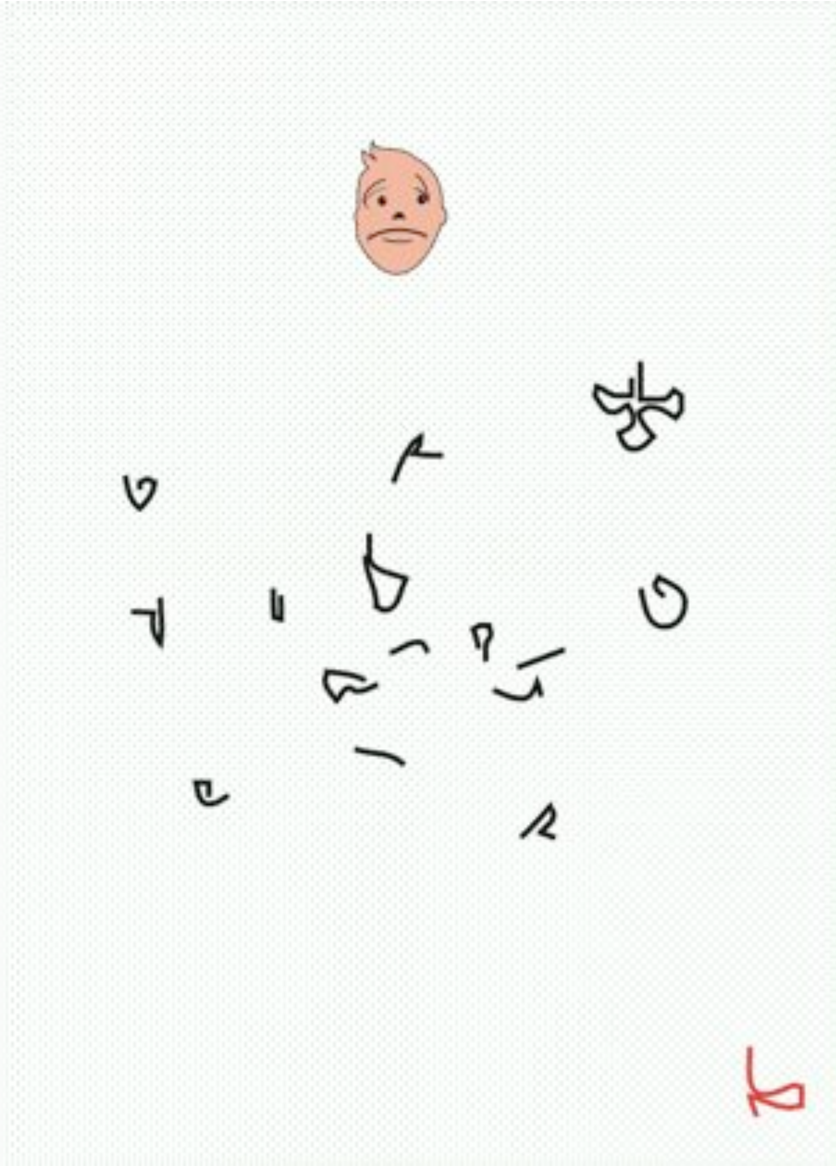
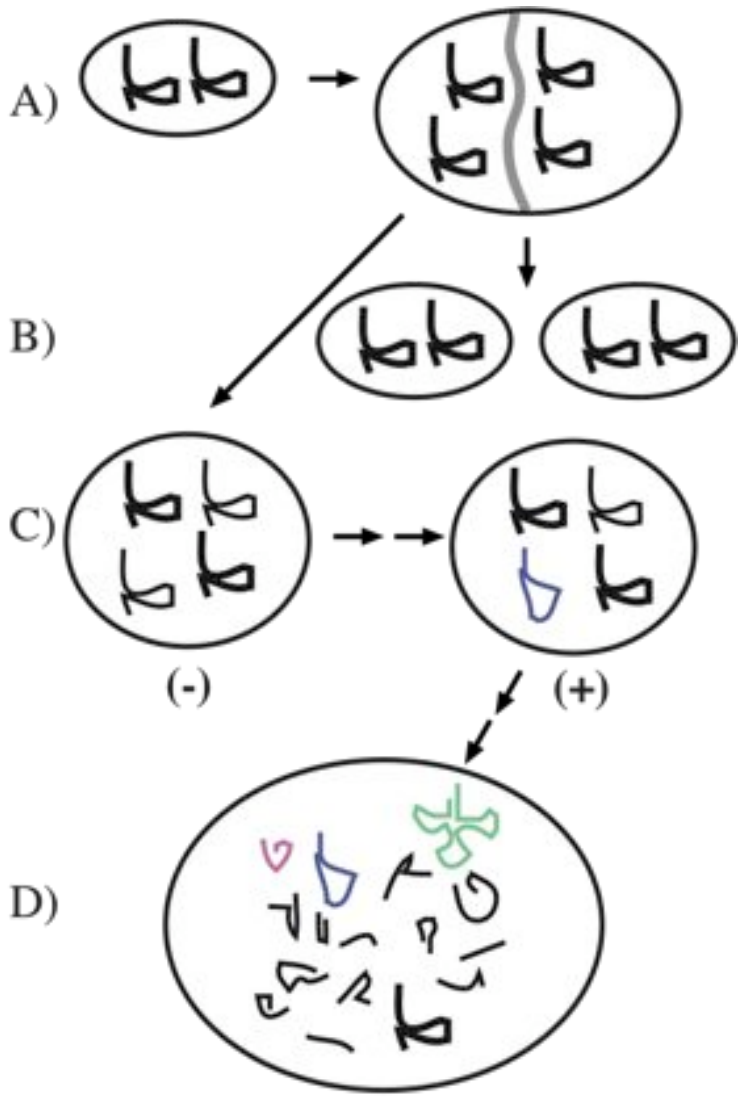


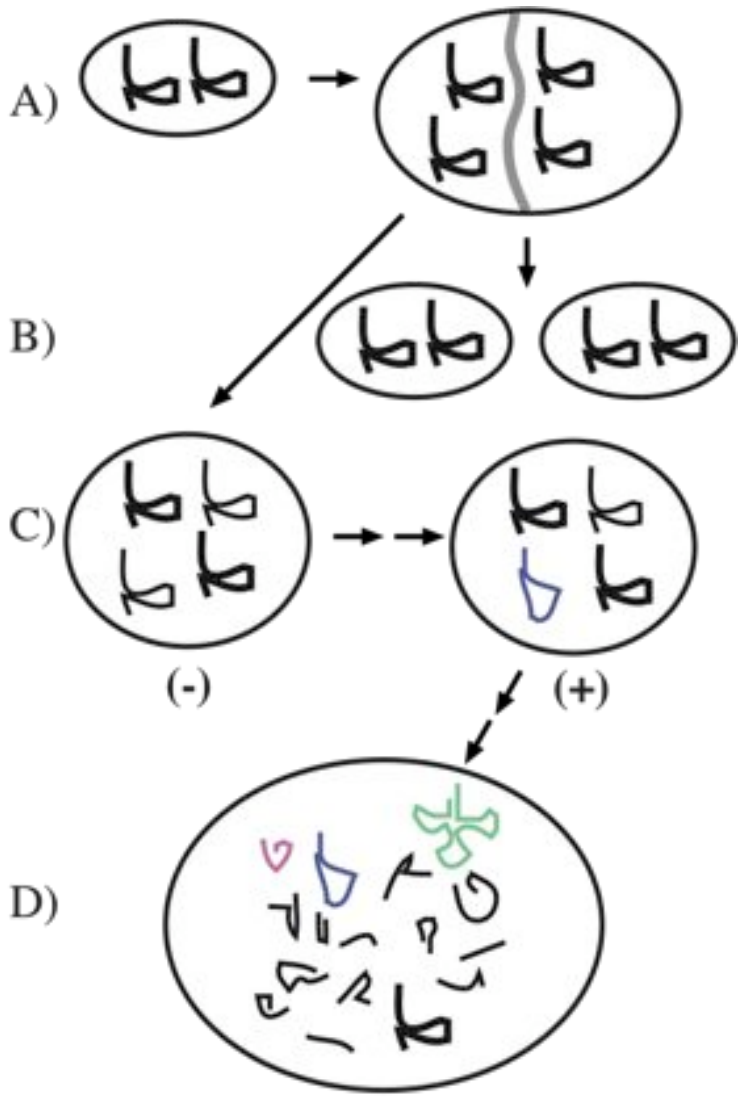
Figure 3. Schematic diagram of coupled vesicle growth and division. (Reproduced from Zhu and Szostak 2009a and reprinted with permission from the Journal of the American Chemical Society ©2009.)



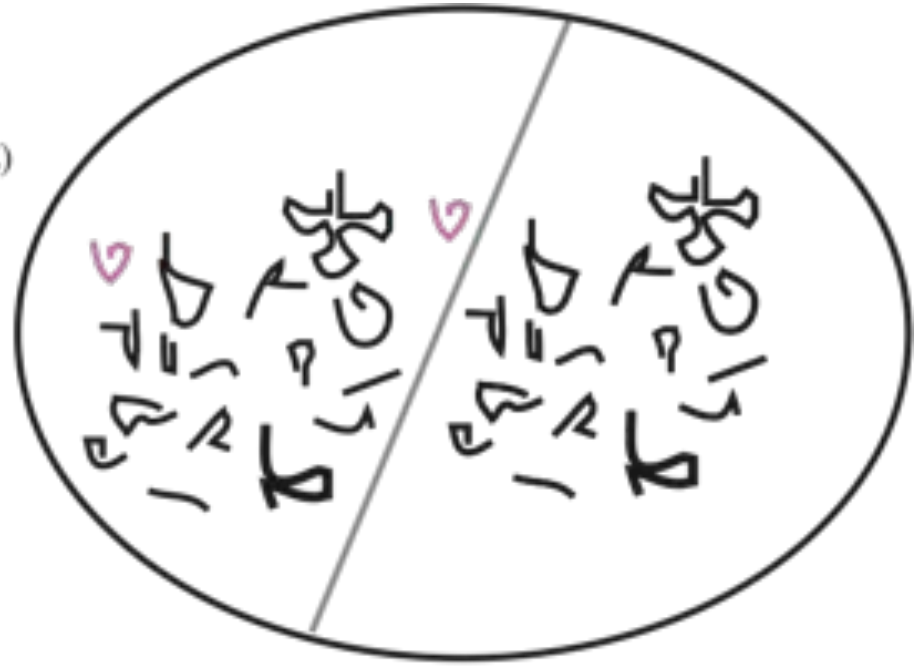

A)
B)
C)
(-)
(+)
D)

The Beginning of Life?

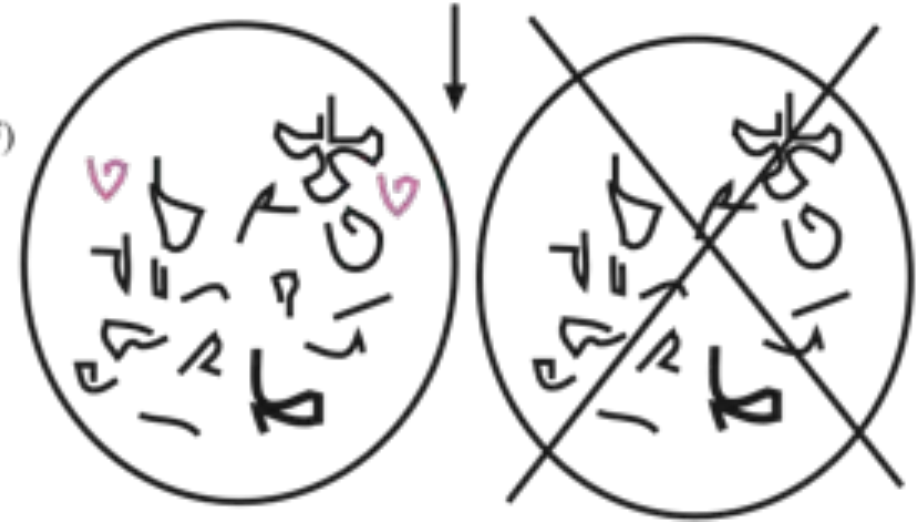
When - by chance - one of the [possibly] RNA polymers became able to replicate itself and later other RNA molecules.

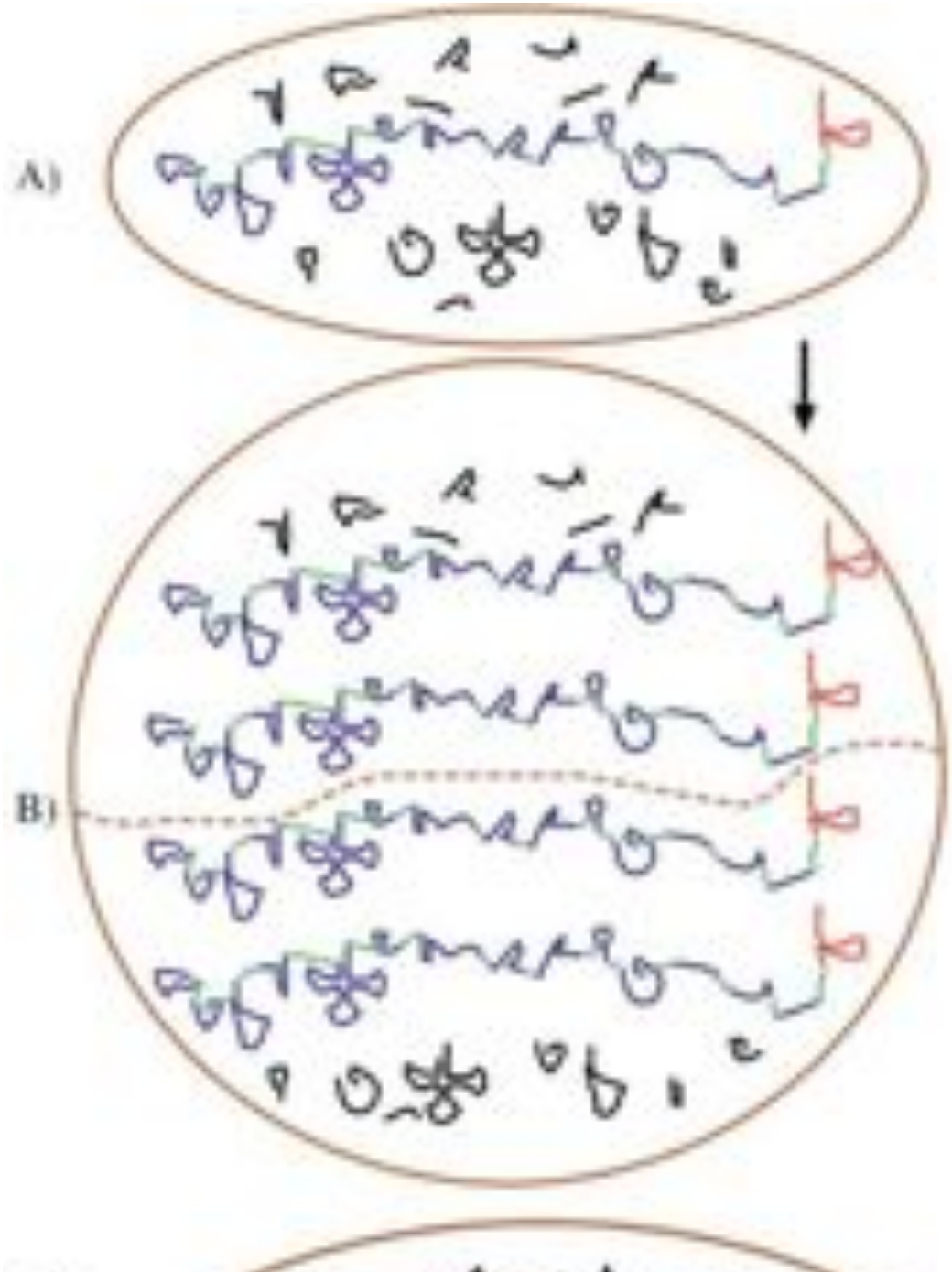

A)
B)
C)
(-)
(+)
D)

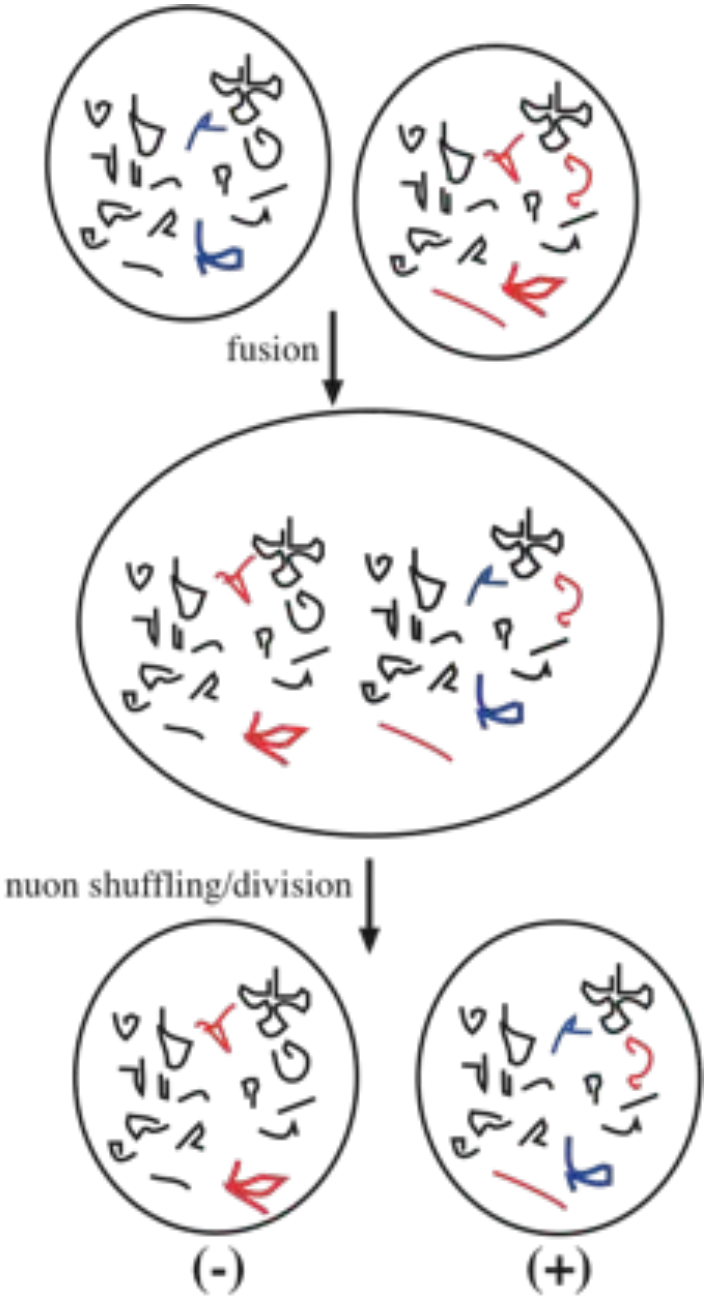
E)



F)




A)
B)


fusion
nuon shuffling/division
(-)
(+)

Journal of Structural and Functional Genomics **3**: 1–13, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*

Gene duplication and other evolutionary strategies: from the RNA world to the future

Jürgen Brosius
Institute of Experimental Pathology, Center for Molecular Biology of Inflammation, University of Münster, Von-Esmarch-Str. 56, D–48149 Münster, Germany
E-mail: RNAworld@uni-muenster.de; fax: +49 251 835 8512

Received 16.05.2002; Accepted in final form 29.06.2002

Key words: barriers of genetic exchange, Darwinian principles, evolutionary transitions, exaptation, group selection, Lamarckism, memes, mass duplication, neons, primordial sex, retroposition, RNA world, RNP world

What Evolution Is

- amplification
- modification
- selection

- Conflict (competition)
- cooperation
- transitions

Competition
Mutation bad
RT litters genome
Sloppy T, wasted RNA
Sloppy splicing, mRNA defective

Cooperation
Mutation good
RT provides raw material
Sloppy T, provides raw mat.
novel splice variants evolve

nature chemistry

ARTICLES

PUBLISHED ONLINE: 19 MAY 2013 | DOI: 10.1038/NCHEM.1650

Competition between model protocells driven by an encapsulated catalyst

Katarzyna Adamala[1,2] and Jack W. Szostak[1]*

The advent of Darwinian evolution required the emergence of molecular mechanisms for the heritable variation of fitness. One model for such a system involves competing protocell populations, each consisting of a replicating genetic polymer within a replicating vesicle. In this model, each genetic polymer imparts a selective advantage to its protocell by, for example, coding for a catalyst that generates a useful metabolite. Here, we report a partial model of such nascent evolutionary traits in a system that consists of fatty-acid vesicles containing a dipeptide catalyst, which catalyses the formation of a second dipeptide. The newly formed dipeptide binds to vesicle membranes, which imparts enhanced affinity for fatty acids and thus promotes vesicle growth. The catalysed dipeptide synthesis proceeds with higher efficiency in vesicles than in free solution, which further enhances fitness. Our observations suggest that, in a replicating protocell with an RNA genome, ribozyme-catalysed peptide synthesis might have been sufficient to initiate Darwinian evolution.

Journal of Theoretical Biology 267 (2010) 605–613

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Selfishness versus functional cooperation in a stochastic protocell model

Elias Zintzaras [a,b,c], Mauro Santos [a,d,*], Eörs Szathmáry [a,e,f]

[a] Collegium Budapest, Institute for Advanced Study, Szentháromság u. 2, H-1014 Budapest, Hungary
[b] Department of Biomathematics, University of Thessaly School of Medicine, 2 Panepistimiou Str, Biopolis, Larissa 41100, Greece
[c] The Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Tufts University School of Medicine, 800 Washington Str, Boston, MA 02111, USA
[d] Departament de Genètica i de Microbiologia, Grup de Biologia Evolutiva (GBE), Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[e] Institute of Biology, Eötvös University, 1/c Pázmány Péter sétány, H-1117 Budapest, Hungary
[f] Parmenides Center for the Study of Thinking, Kirchplatz 1, D-82409 Munich/Pullach, Germany

ARTICLE INFO

Article history:
Received 29 March 2010
Received in revised form
30 August 2010
Accepted 5 September 2010
Available online 17 September 2010

Keywords:
Artificial cells
Functional complexity
Monte Carlo methods
Qβ replicase
Ribozymes

ABSTRACT

How to design an "evolvable" artificial system capable to increase in complexity? Although Darwin's theory of evolution by natural selection obviously offers a firm foundation, little hope of success seems to be expected from the explanatory adequacy of modern evolutionary theory, which does a good job at explaining what has already happened but remains practically helpless at predicting what will occur. However, the study of the major transitions in evolution clearly suggests that increases in complexity have occurred on those occasions when the conflicting interests between competing individuals were partly subjugated. This immediately raises the issue about "levels of selection" in evolutionary biology, and the idea that multi-level selection scenarios are required for complexity to emerge. After analyzing the dynamical behaviour of competing replicators within compartments, we show here that a proliferation of differentiated catalysts and/or improvement of catalytic efficiency of ribozymes can potentially evolve in properly designed artificial cells where the strong internal competition between the different species of replicators is somewhat prevented (i.e., by choosing them with equal probability). Experimental evolution in these systems will likely stand as beautiful examples of artificial adaptive systems, and will provide new insights to understand possible evolutionary paths to the evolution of metabolic complexity.

© 2010 Elsevier Ltd. All rights reserved.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 239 (2006) 247–256
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Coexistence and error propagation in pre-biotic vesicle models: A group selection approach

José F. Fontanari[a,*], Mauro Santos[b,c], Eörs Szathmáry[c,d]

[a]*Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-970 São Carlos SP, Brazil*
[b]*Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain*
[c]*Collegium Budapest, Institute for Advanced Study, Szentháromság u. 2, H-1014 Budapest, Hungary*
[d]*Department of Plant Taxonomy and Ecology, Eötvös University, 2 Ludovika tér, H-1083 Budapest, Hungary*

Received 26 April 2005; received in revised form 29 June 2005; accepted 23 July 2005
Available online 21 October 2005

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2006) **361**, 1761–1776
doi:10.1098/rstb.2006.1912
Published online 7 September 2006

The origin of replicators and reproducers

Eörs Szathmáry[1,2,*]

[1]*Collegium Budapest (Institute for Advanced Study), 2 Szentháromság utca, 1014 Budapest, Hungary*
[2]*Department of Plant Taxonomy and Ecology, Institute of Biology, Eötvös University, 1/c Pázmány Péter sétany 1117 Budapest, Hungary*

Replicators are fundamental to the origin of life and evolvability. Their survival depends on the accuracy of replication and the efficiency of growth relative to spontaneous decay. Infrabiological systems are built of two coupled autocatalytic systems, in contrast to minimal living systems that must comprise at least a metabolic subsystem, a hereditary subsystem and a boundary, serving respective functions. Some scenarios prefer to unite all these functions into one primordial system, as illustrated in the lipid world scenario, which is considered as a didactic example in detail. Experimentally produced chemical replicators grow parabolically owing to product inhibition. A selection consequence is survival of everybody. The chromatographized replicator model predicts that such replicators spreading on surfaces can be selected for higher replication rate because double strands are washed away slower than single strands from the surface. Analysis of real ribozymes suggests that the error threshold of replication is less severe by about one order of magnitude than thought previously. Surface-bound dynamics is predicted to play a crucial role also for exponential replicators: unlinked genes belonging to the same genome do not displace each other by competition, and efficient and accurate replicases can spread. The most efficient form of such useful population structure is encapsulation by reproducing vesicles. The stochastic corrector model shows how such a bag of genes can survive, and what the role of chromosome formation and intragenic recombination could be. Prebiotic and early evolution cannot be understood without the models of dynamics.

Keywords: replicator; origin of life; ribozyme; autocatalysis; compartmentation; error threshold

ARTICLE

doi:10.1038/nature11549

Spontaneous network formation among cooperative RNA replicators

Nilesh Vaidya[1], Michael L. Manapat[2], Irene A. Chen[3]†, Ramon Xulvi-Brunet[3], Eric J. Hayden[4] & Niles Lehman[1]

The origins of life on Earth required the establishment of self-replicating chemical systems capable of maintaining and evolving biological information. In an RNA world, single self-replicating RNAs would have faced the extreme challenge of possessing a mutation rate low enough both to sustain their own information and to compete successfully against molecular parasites with limited evolvability. Thus theoretical analyses suggest that networks of interacting molecules were more likely to develop and sustain life-like behaviour. Here we show that mixtures of RNA fragments that self-assemble into self-replicating ribozymes spontaneously form cooperative catalytic cycles and networks. We find that a specific three-membered network has highly cooperative growth dynamics. When such cooperative networks are competed directly against selfish autocatalytic cycles, the former grow faster, indicating an intrinsic ability of RNA populations to evolve greater complexity through cooperation. We can observe the evolvability of networks through *in vitro* selection. Our experiments highlight the advantages of cooperative behaviour even at the molecular stages of nascent life.

© 2012 Macmillan Publishers Limited. All rights reserved

The cooperative gene

The origin of life on Earth remains one of the great unsolved mysteries. A new study suggests that cooperation among molecules could have contributed to the transition from inanimate chemistry to biology. SEE ARTICLE P.72

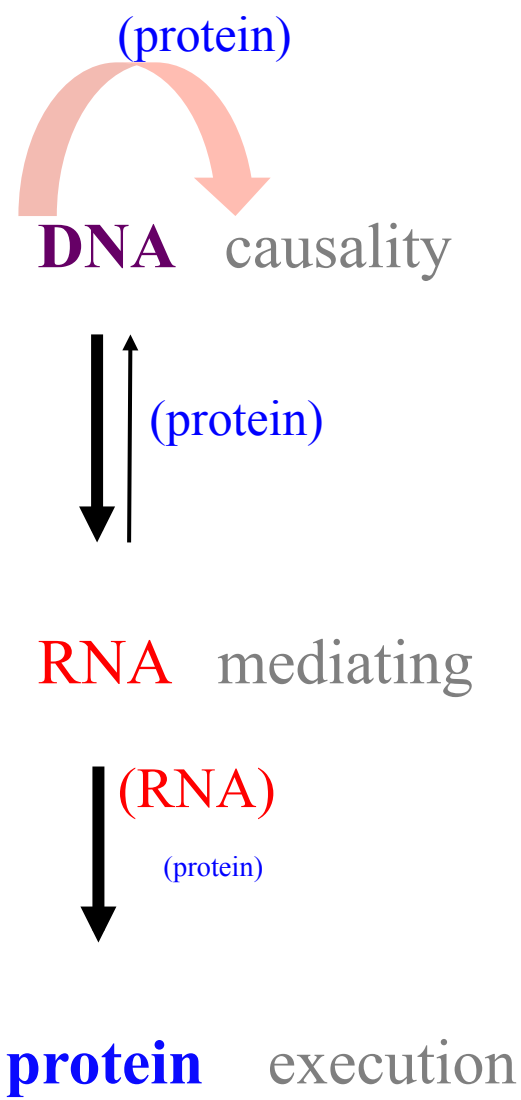
James Attwater *and* **Philipp Holliger** *are at the MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK.*
e-mail: ph1@mrc-lmb.cam.ac.uk

X. Why Are There Two Nucleic Acids?

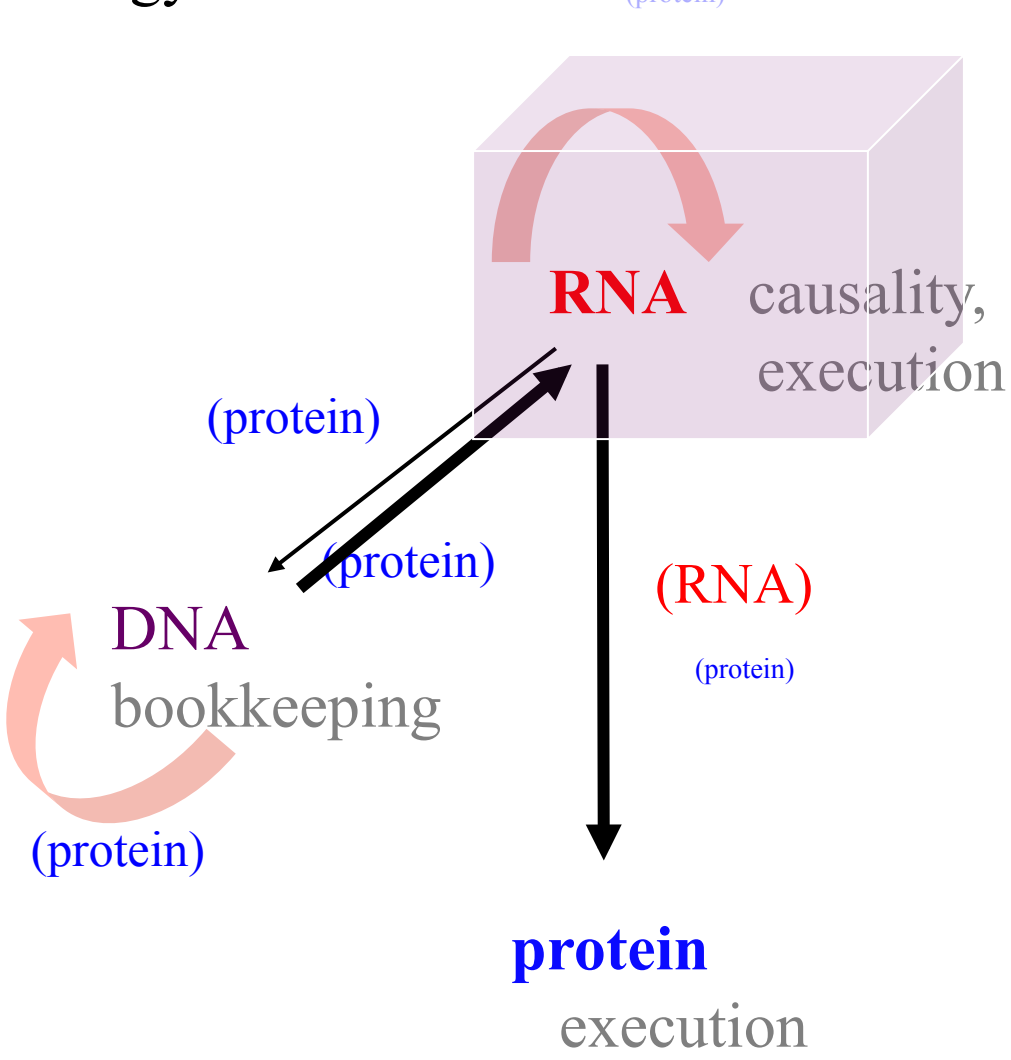
It is quite remarkable that contemporary biochemical systems have two nucleic acids, DNA and RNA, which differ only by a systematic hydroxyl group and an occasional methyl group. Despite the great chemical similarities, the molecules nevertheless have quite different functions in the cell. DNA acts as the major carrier of genetic information, while the RNA molecule is used to convert this genetic information into actual protein molecules. Because of the close chemical similarities, we are tempted to ask whether they could have originated historically from a common stem nucleic acid molecule which then specialized in the course of evolution to produce the two different classes of nucleic acids which we see today. To discuss this further we should note that the RNA molecule is also able to carry genetic information, as, for example, in the RNA-containing viruses. Thus, it may be reasonable to speculate that the

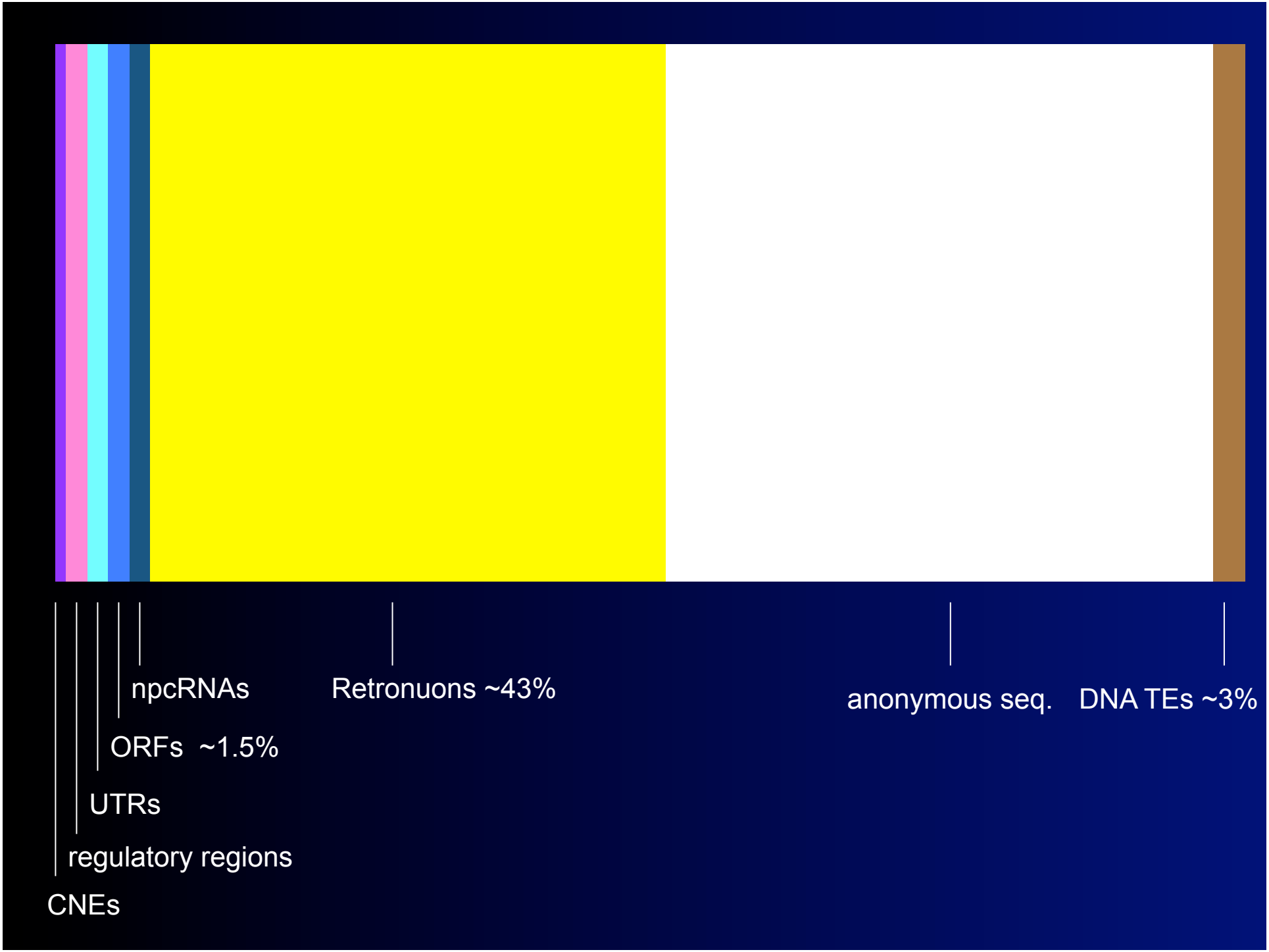
hypothetical stem or parent polynucleotide molecule was initially an RNA-like polymer which was able to convey genetic information as well as organize the amino acids into a specific sequence to make proteins. This implies that the RNA polynucleotide strand had the ability to replicate itself and produce a complement in a manner somewhat similar to that which is found in DNA. It is possible that an enzyme of this type may have been observed already. By this view, DNA may be regarded as a derivative molecule which has evolved in a form such that it only carries out part of the primitive nucleic acid function. It specialized in the molecular replicating cycle that is part of the mechanism for transmitting genetic information. DNA is metabolically less reactive than RNA, perhaps because of the absence of the hydroxyl group on carbon 2. The loss of this hydroxyl group may have made it impossible for the DNA molecule to have attached to it the amino acids which are used in protein synthesis. However, considerable selective advantage may be derived from the development of two different classes of nucleic acids, one of which is less active metabolically and specializes in self-replication. In a sense, this tends to preserve the primary copy of the genetic information. It will be of considerable interest to study the available simple life forms to see whether some of them may exist with only one type of nucleic acid rather than two types. It is possible that the RNA containing viruses may be regarded as present-day examples which may have degenerated evolutionarily from such a primitive life form.

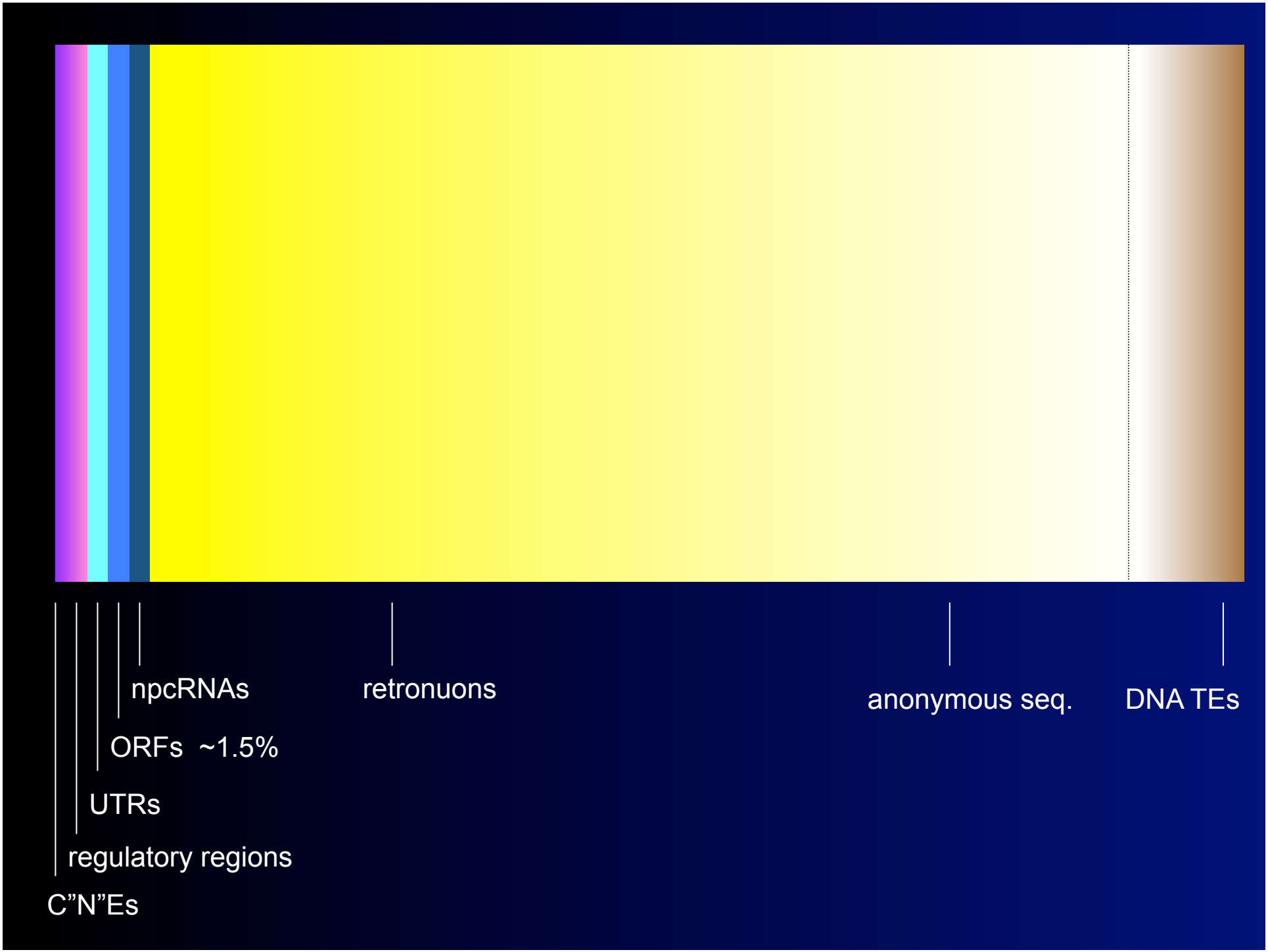
Central Dogma of Molecular Biology



Revisited




CNEs
regulatory regions
UTRs
ORFs ~1.5%
npcRNAs
Retronuons ~43%
anonymous seq.
DNA TEs ~3%


npcRNAs
retronuons
anonymous seq.
DNA TEs
ORFs ~1.5%
UTRs
regulatory regions
C”N”Es

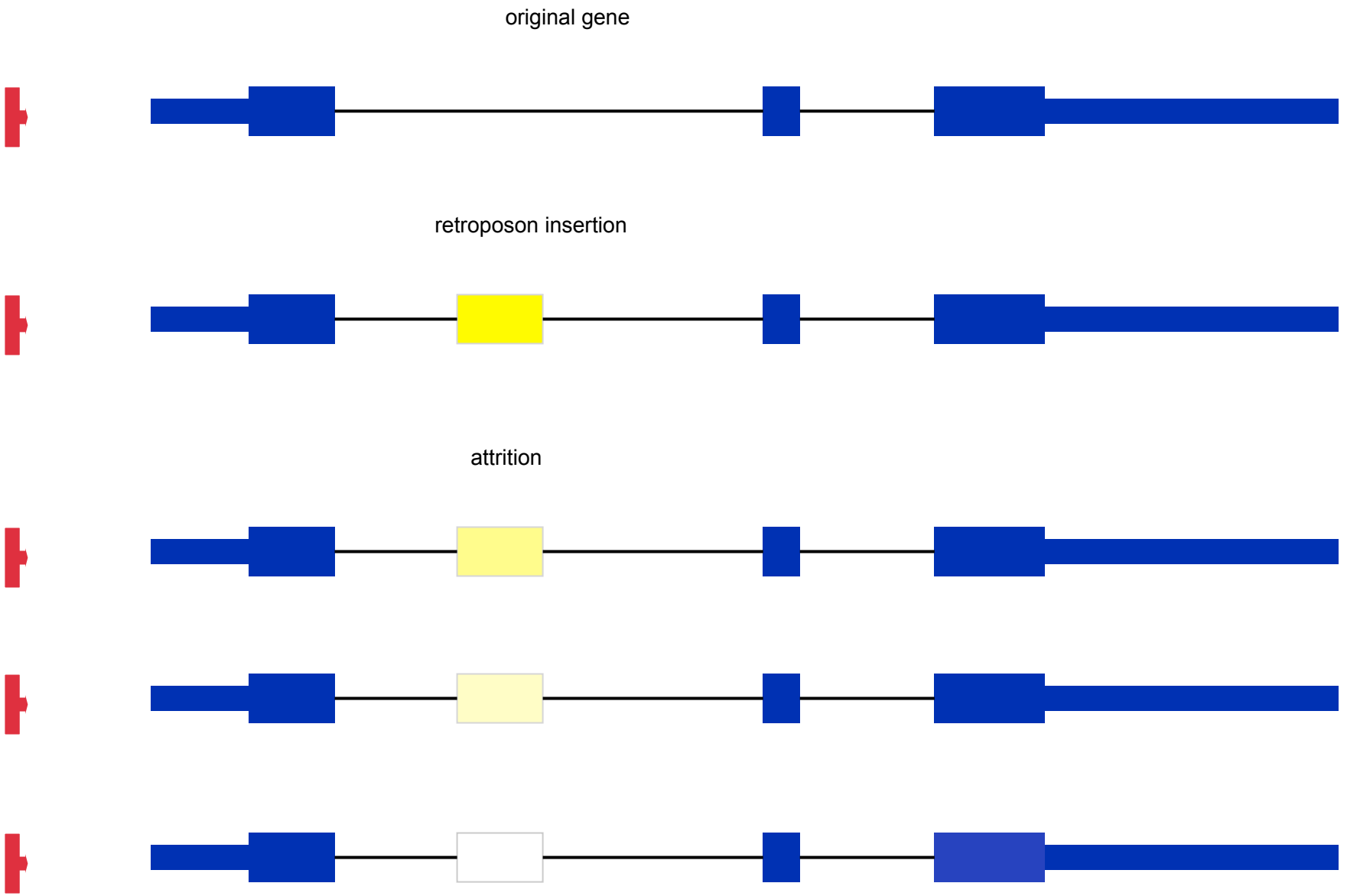
Retroposition yields

1) novel protein encoding genes or novel protein domains
2) regulatory elements

   (e.g., promoters, enhancers, silencers, poly(A) addition signals, splice sites → novel protein domains)

3) non-protein coding RNA genes (npcRNAs)

   (e.g., BC1 RNA, BC200 RNA, transcribed Ψ-genes)


original gene
retroposon insertion
attrition

exonization - at any stage:

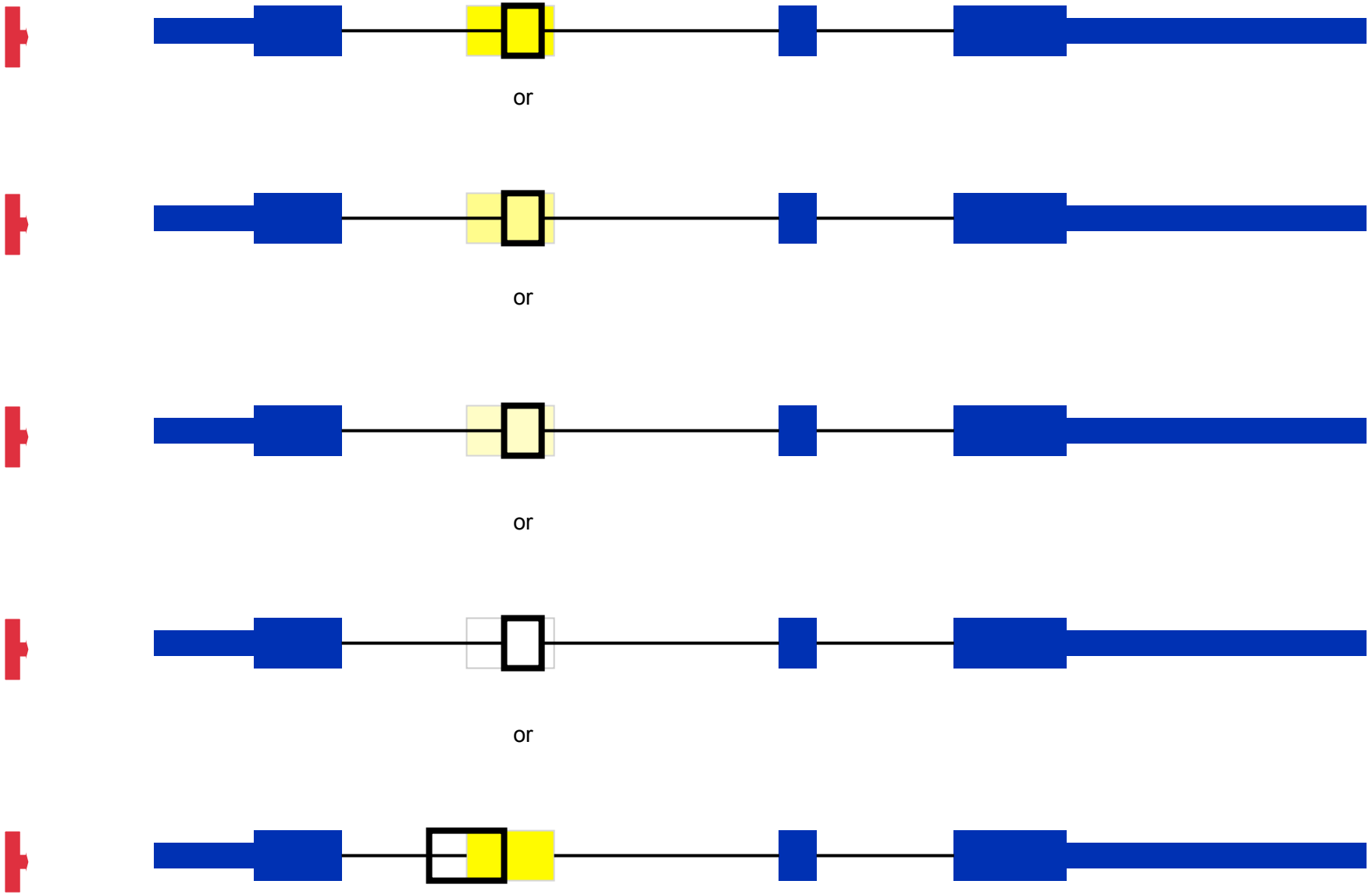
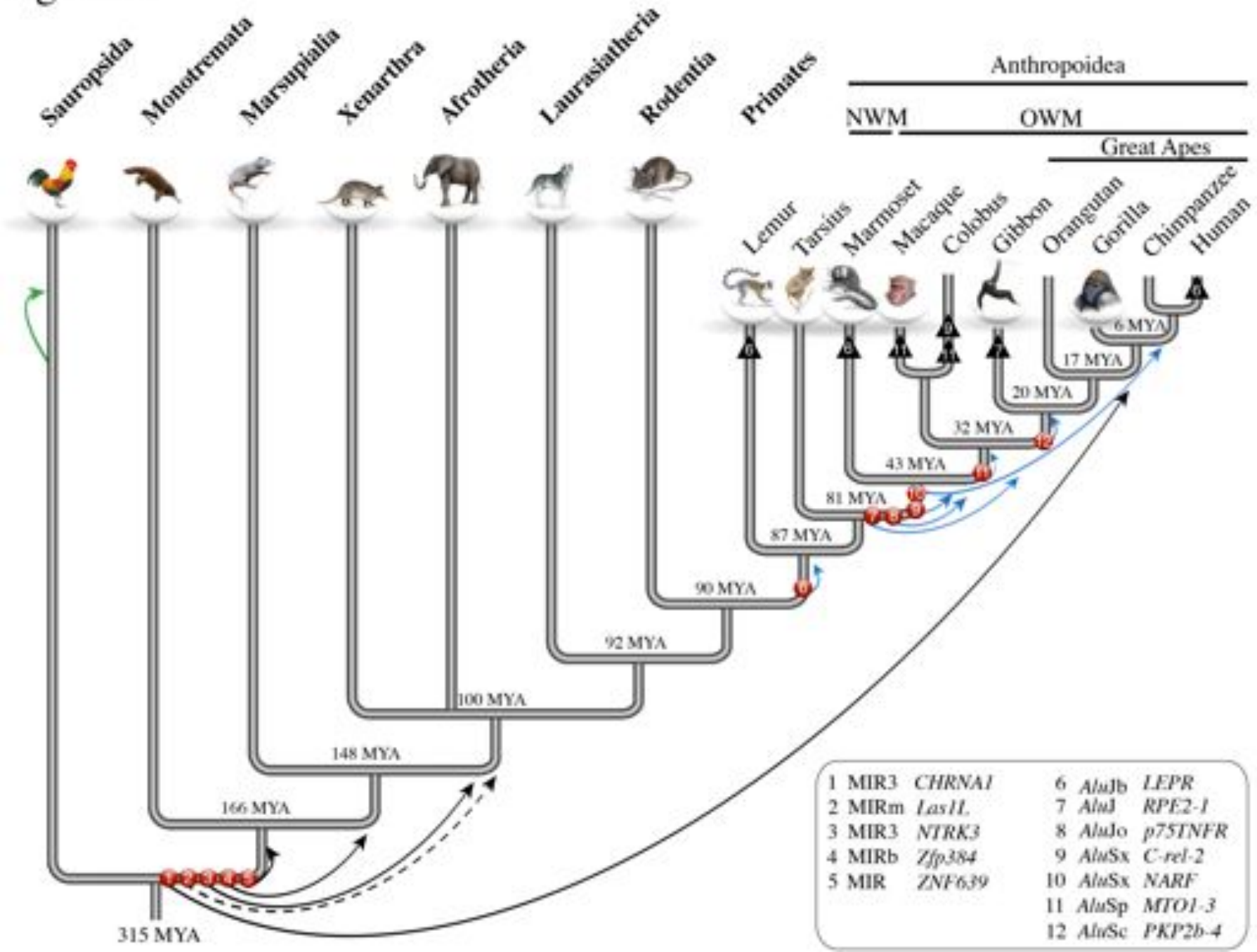
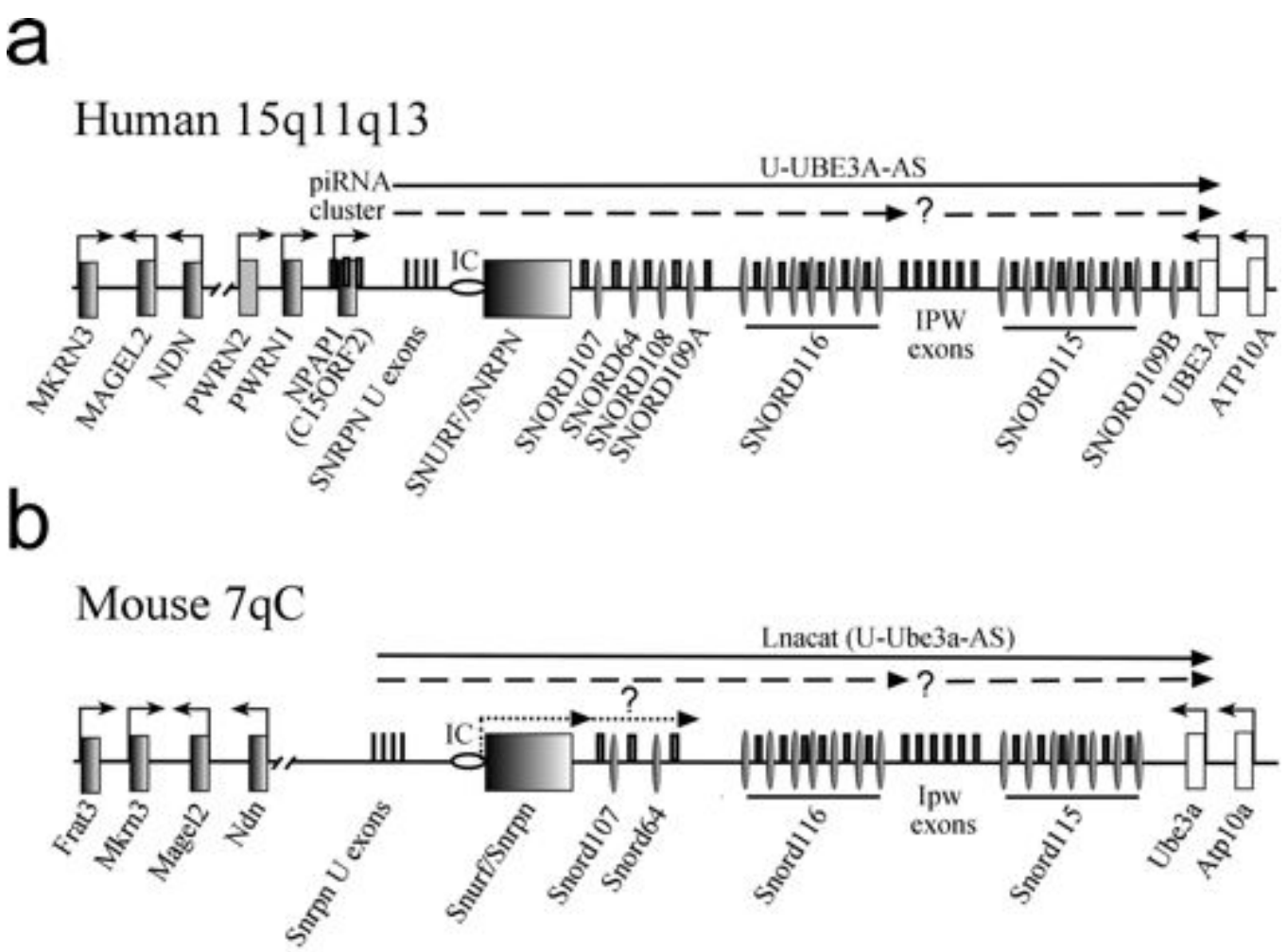
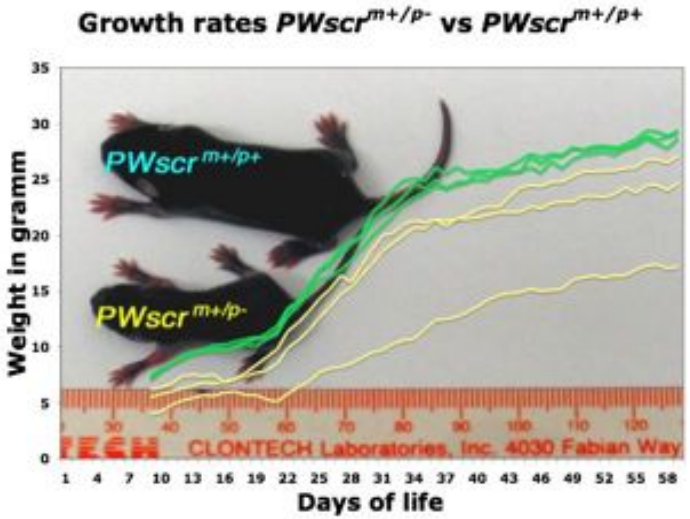
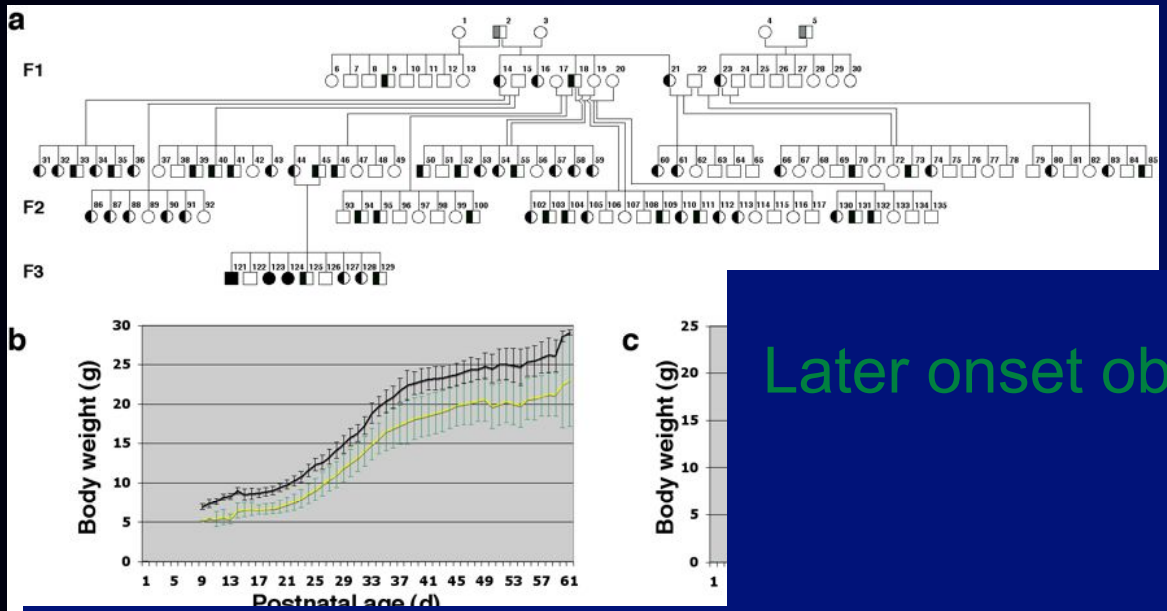


Figure 2


Sauropsida
Monotremata
Marsupialia
Xenarthra
Afrotheria
Laurasiatheria
Rodentia
Primates
Anthropoidea
NWM
OWM
Great Apes
Lemur
Tarsius
Marmoset
Macaque
Colobus
Gibbon
Orangutan
Gorilla
Chimpanzee
Human
6 MYA
17 MYA
20 MYA
32 MYA
43 MYA
81 MYA
87 MYA
90 MYA
92 MYA
100 MYA
148 MYA
166 MYA
315 MYA
1 MIR3 CHRNA1
2 MIRm Las1L
3 MIR3 NTRK3
4 MIRb Zfp384
5 MIR ZNF639
6 AluJb LEPR
7 AluJ RPE2-1
8 AluJo p75TNFR
9 AluSx C-rel-2
10 AluSx NARF
11 AluSp MTO1-3
12 AluSc PKP2b-4


a
Human 15q11q13
piRNA cluster
U-UBE3A-AS
?
IC
MKRN3
MAGEL2
NDN
PWRN2
PWRN1
NPAP1 (C15ORF2)
SNRPN U exons
SNURF/SNRPN
SNORD107
SNORD64
SNORD108
SNORD109A
SNORD116
IPW exons
SNORD115
SNORD109B
UBE3A
ATP10A
b
Mouse 7qC
Lnacat (U-Ube3a-AS)
?
?
IC
Frat3
Mkrn3
Magel2
Ndn
Snrpn U exons
Snurf/Snrpn
Snord107
Snord64
Snord116
Ipw exons
Snord115
Ube3a
Atp10a


Growth rates PWscr m+/p- vs PWscr m+/p+
Weight in gramm
0
5
10
15
20
25
30
35
PWscr m+/p+
PWscr m+/p-
CLONTECH Laboratories, Inc. 4030 Fabian Way
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58
Days of life



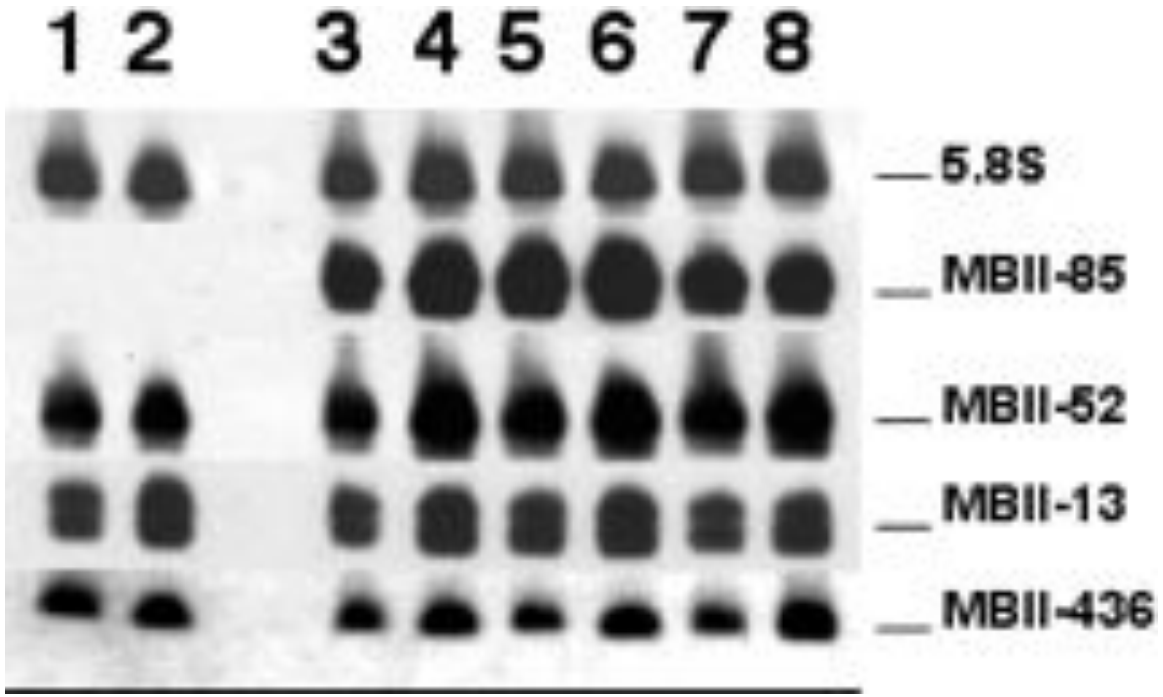
Infertility: **no**

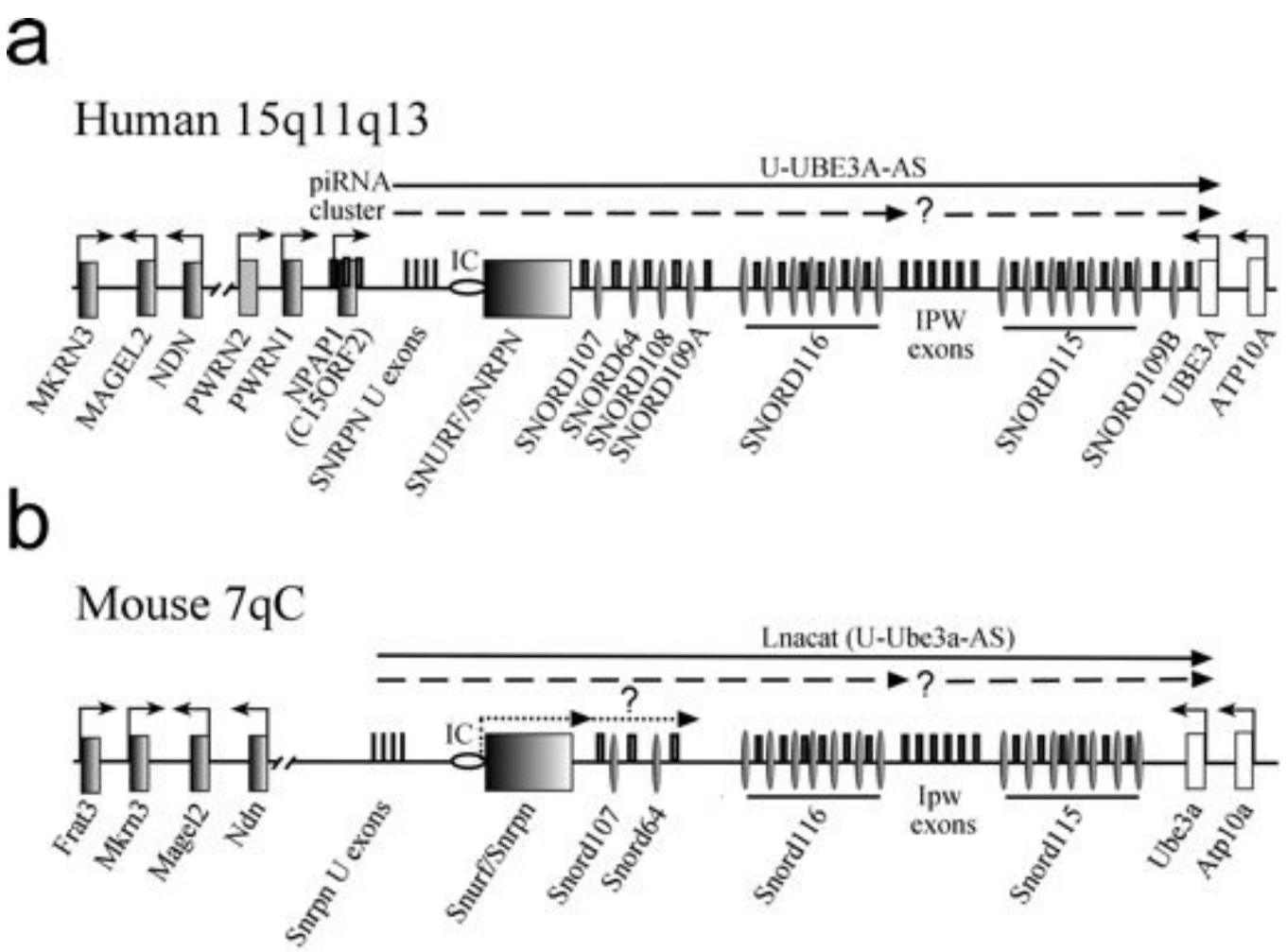
Later onset obesity: **no**

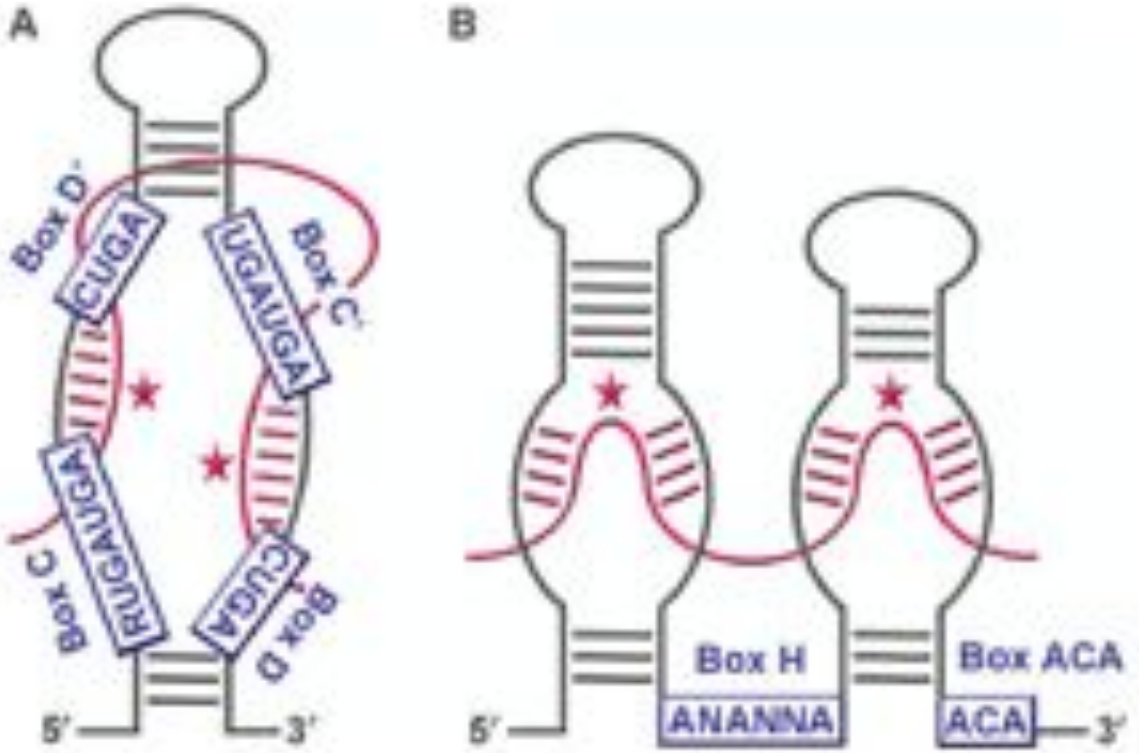
d

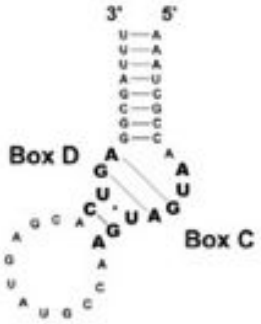
f

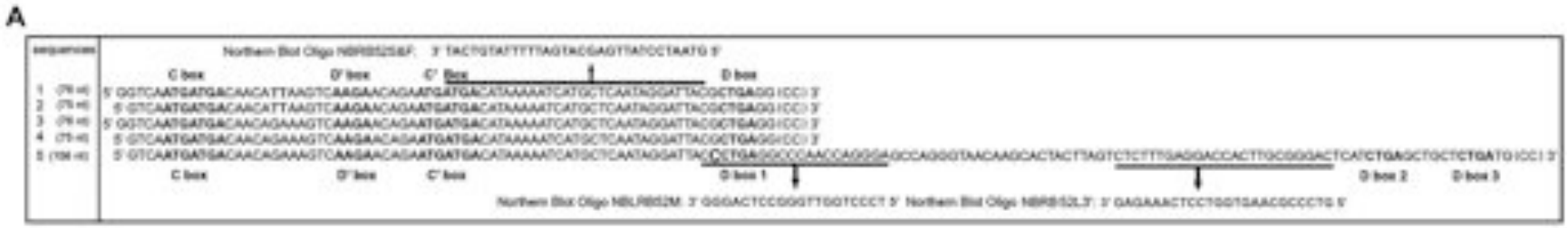
Skryabin, B.V., Gubar, L., Seeger, B., Pfeiffer, J., Handel, S. Robeck, T., Karpova, E., Rozhdestvensky, T.S., Brosius, J. (2007) Deletion of MBII-85 snoRNA gene cluster in mice results in postnatal growth retardation. PLoS Genet. 3, 2529-2539.

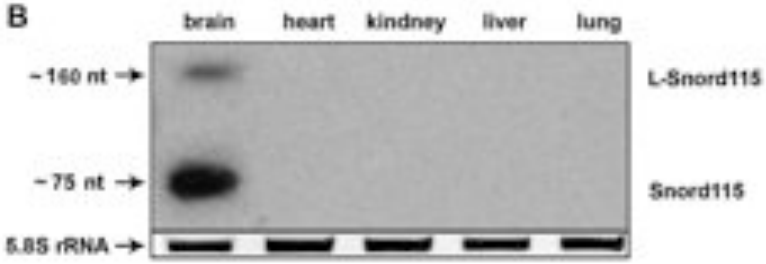

1 2 3 4 5 6 7 8
5.8S
MBII-85
MBII-52
MBII-13
MBII-436

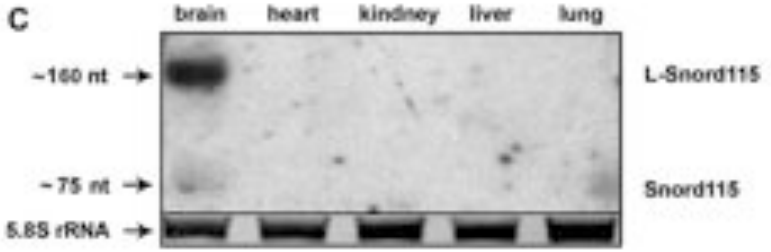

a
Human 15q11q13
piRNA cluster
U-UBE3A-AS
?
IC
MKRN3
MAGEL2
NDN
PWRN2
PWRN1
NPAP1 (C15ORF2)
SNRPN U exons
SNURF/SNRPN
SNORD107
SNORD64
SNORD108
SNORD109A
SNORD116
IPW exons
SNORD115
SNORD109B
UBE3A
ATP10A
b
Mouse 7qC
Lnacat (U-Ube3a-AS)
?
?
IC
Frat3
Mkrn3
Magel2
Ndn
Snrpn U exons
Snurf/Snrpn
Snord107
Snord64
Snord116
Ipw exons
Snord115
Ube3a
Atp10a

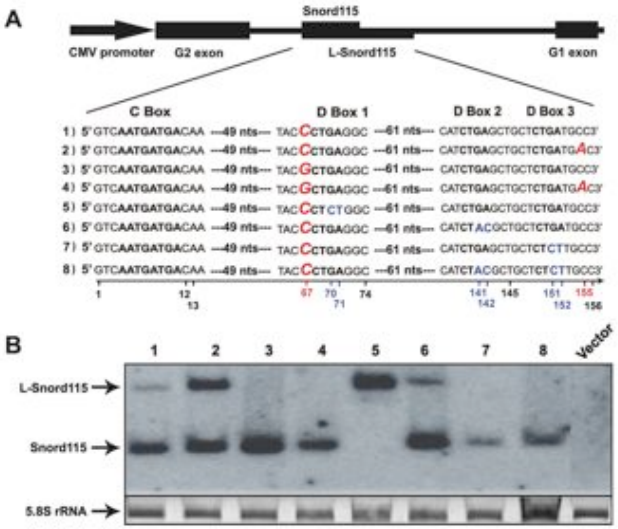

A
Box D'
CUGA
UGAUGA
Box C'
RUGAUGA
Box C
CUGA
Box D
5'
3'
B
Box H
ANANNA
Box ACA
ACA
5'
3'

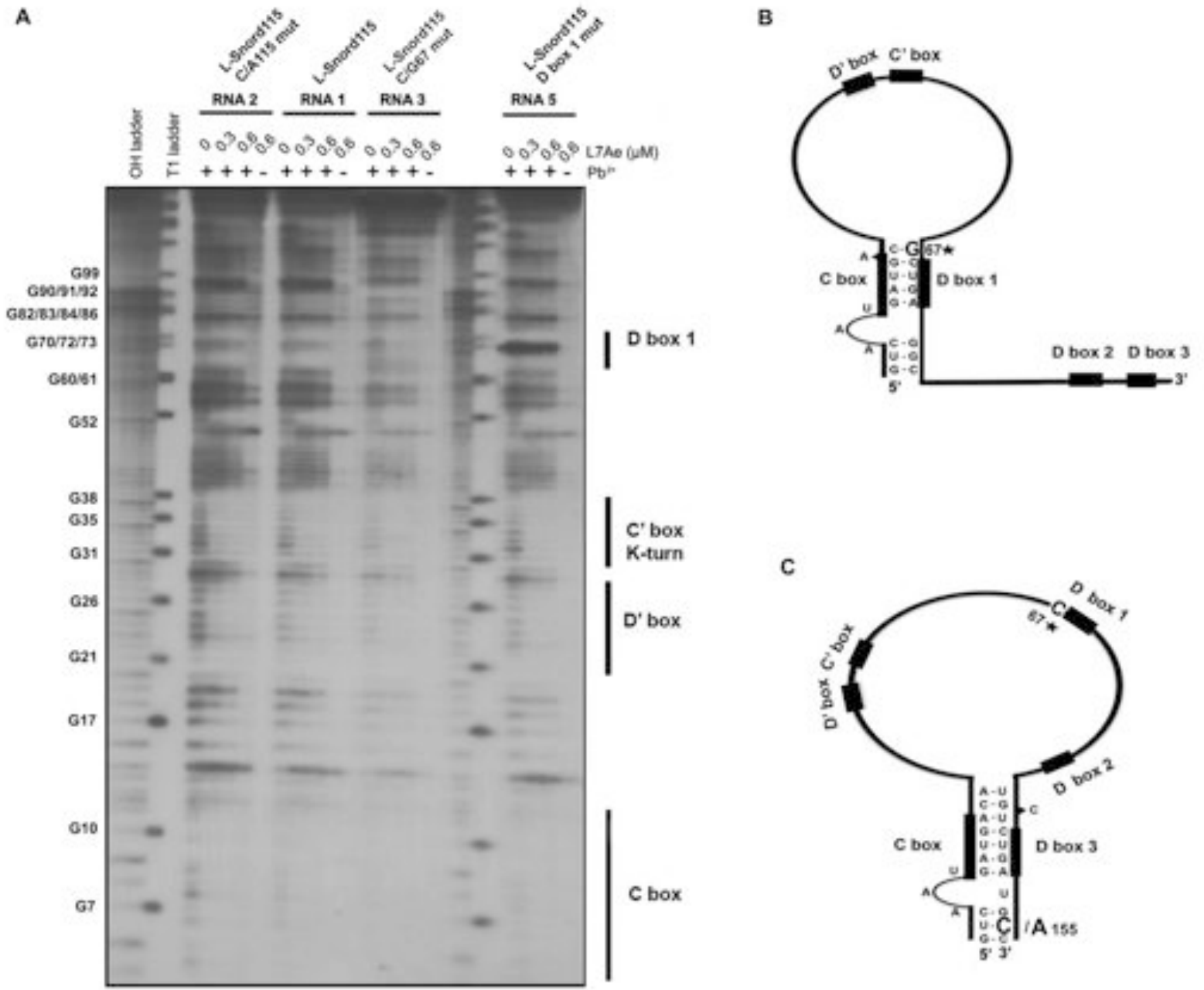

3'
5'
Box D
Box C

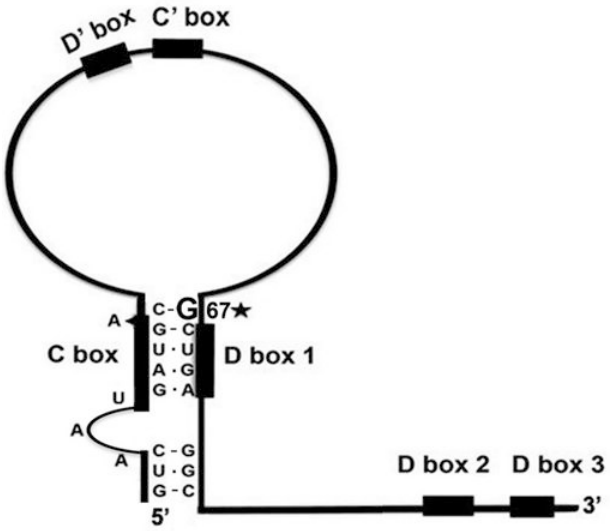

A
sequences
Northern Blot Oligo NBRB52SBF: 3' TACTGTATTTTTAGTACGAGTTATCCTAATG 5'
C box
D' box
C' box
Box
D box
1 (76 nt)
2 (75 nt)
3 (76 nt)
4 (75 nt)
5 (106 nt)
C box
D' box
C' box
D box 1
D box 2
D box 3
Northern Blot Oligo NBLRB52M: 3' GGGACTCCGGGTTGGTCCCT 5'
Northern Blot Oligo NBRB52L3': 3' GAGAAACTCCTGGTGAACGCCCTG 5'

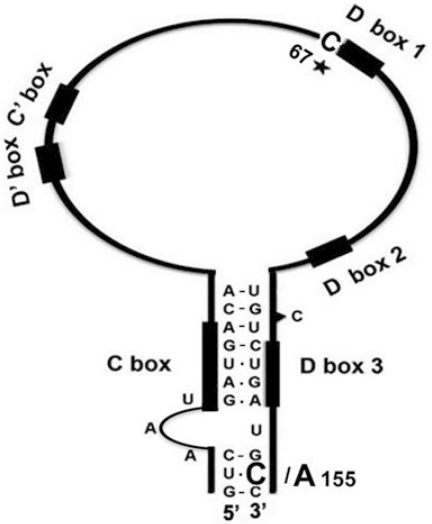

B
brain
heart
kindney
liver
lung
~160 nt
~75 nt
5.8S rRNA
L-Snord115
Snord115

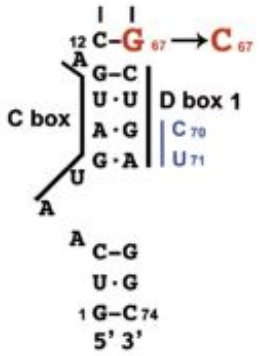

C
brain
heart
kindney
liver
lung
~160 nt
~75 nt
5.8S rRNA
L-Snord115
Snord115

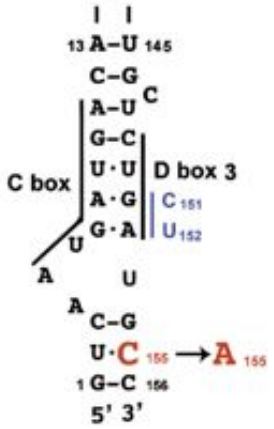

A
Snord115
CMV promoter
G2 exon
L-Snord115
G1 exon
C Box
D Box 1
D Box 2
D Box 3
1) 5'GTCAATGATGACAA ---49 nts--- TACCCTGAGGC ---61 nts--- CATCTGAGCTGCTCTGATGCC3'
2) 5'GTCAATGATGACAA ---49 nts--- TACCCTGAGGC ---61 nts--- CATCTGAGCTGCTCTGATGAC3'
3) 5'GTCAATGATGACAA ---49 nts--- TACGCTGAGGC ---61 nts--- CATCTGAGCTGCTCTGATGCC3'
4) 5'GTCAATGATGACAA ---49 nts--- TACGCTGAGGC ---61 nts--- CATCTGAGCTGCTCTGATGAC3'
5) 5'GTCAATGATGACAA ---49 nts--- TACCCTCTGGC ---61 nts--- CATCTGAGCTGCTCTGATGCC3'
6) 5'GTCAATGATGACAA ---49 nts--- TACCCTGAGGC ---61 nts--- CATCTACGCTGCTCTGATGCC3'
7) 5'GTCAATGATGACAA ---49 nts--- TACCCTGAGGC ---61 nts--- CATCTGAGCTGCTCTCTTGCC3'
8) 5'GTCAATGATGACAA ---49 nts--- TACCCTGAGGC ---61 nts--- CATCTACGCTGCTCTCTTGCC3'
1
12
13
67
70
71
74
141
142
145
151
152
156
B
1
2
3
4
5
6
7
8
Vector
L-Snord115
Snord115
5.8S rRNA
C
C box
D box 3
C 151
U 152
C 155
A 155
5' 3'
D
C box
D box 1
C 70
U 71
G 67
C 67
5' 3'


A
L-Snord115 C/A115 mut
L-Snord115
L-Snord115 C/G67 mut
L-Snord115 D box 1 mut
RNA 2
RNA 1
RNA 3
RNA 5
OH ladder
T1 ladder
0 0.3 0.6 0.6 L7Ae (μM)
+ + + - Pb2+
G99
G90/91/92
G82/83/84/86
G70/72/73
G60/61
G52
G38
G35
G31
G26
G21
G17
G10
G7
D box 1
C' box K-turn
D' box
C box
B
D' box
C' box
G67★
C box
D box 1
D box 2
D box 3
5'
3'
C
D box 1
C 67★
D' box
C' box
D box 2
C box
D box 3
/A 155
5' 3'

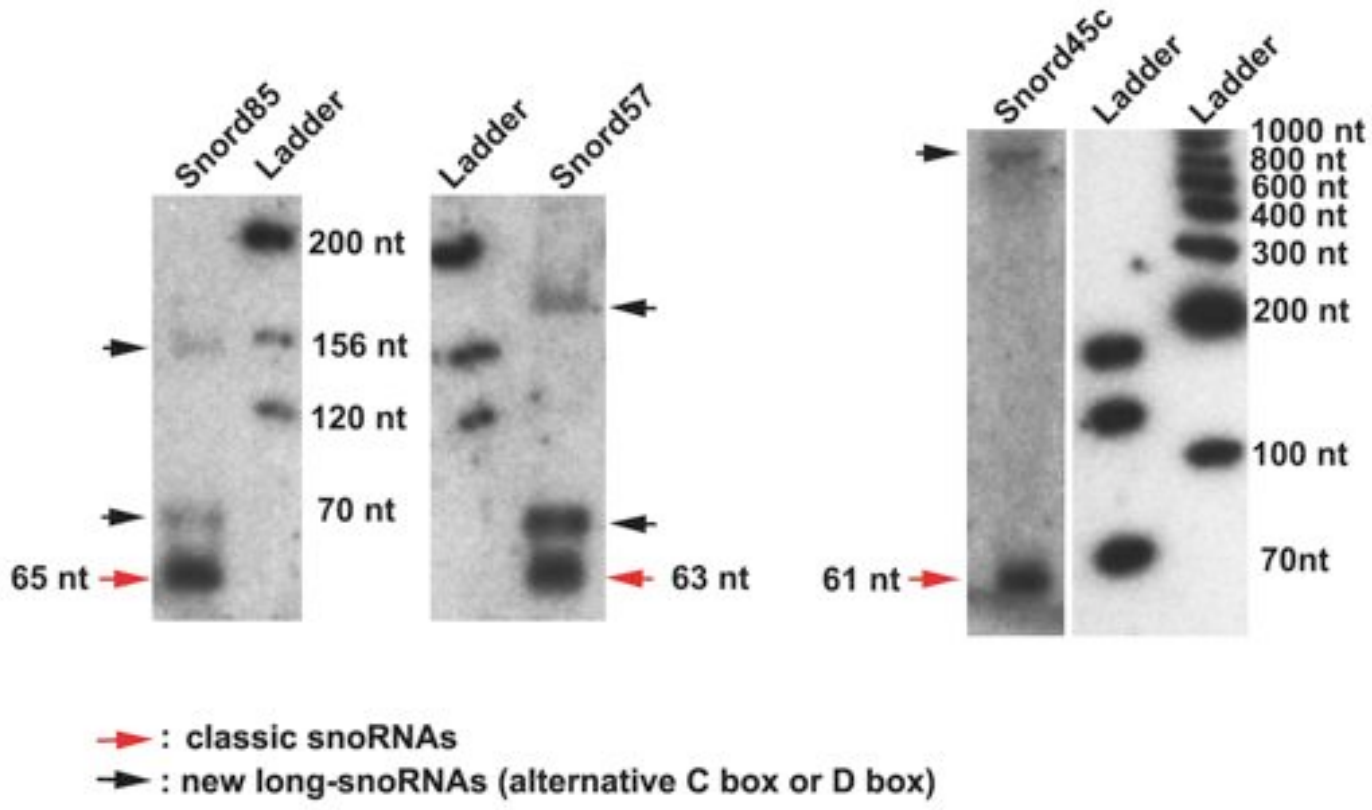

D' box
C' box
C box
D box 1
D box 2
D box 3
67★
5'
3'

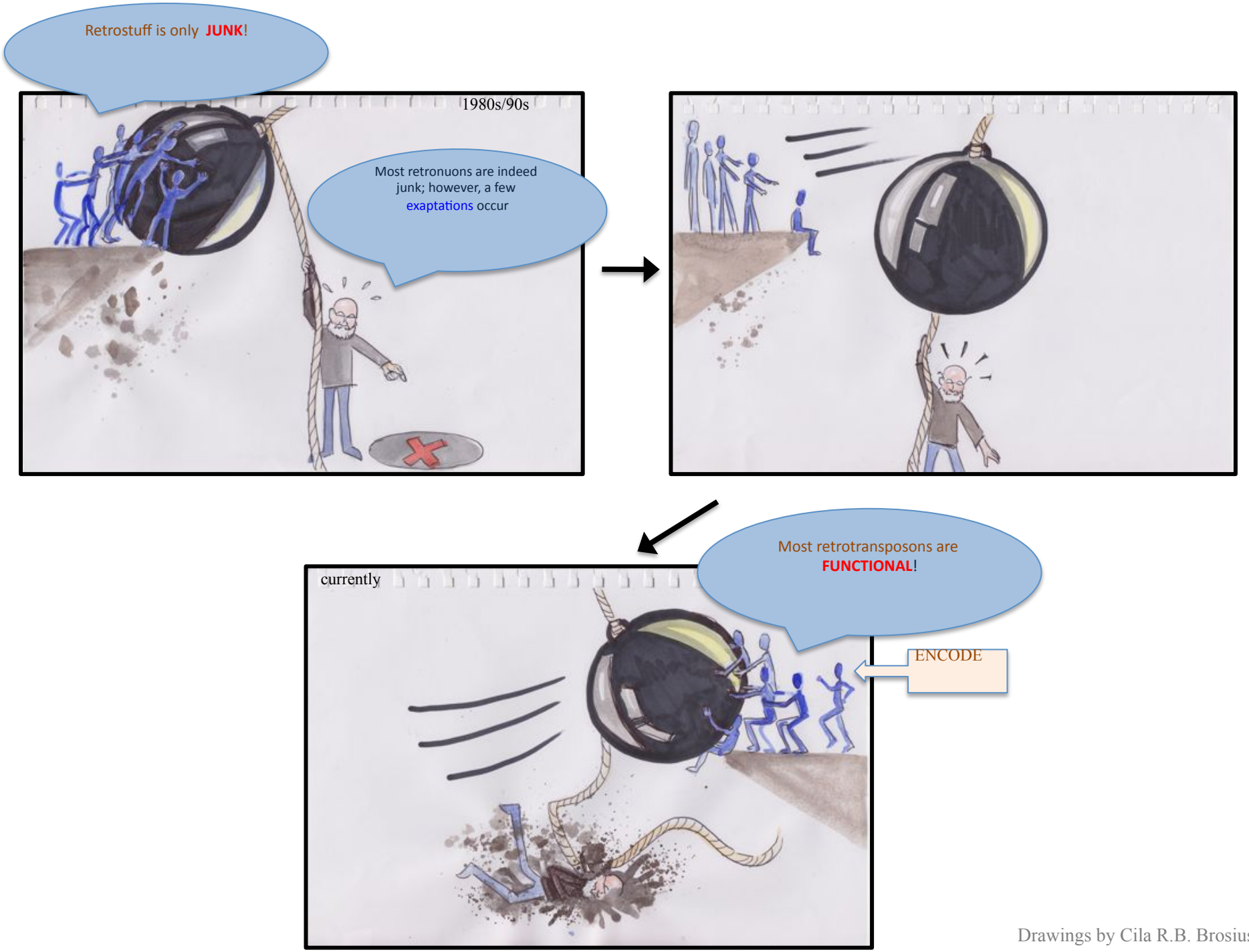

D box 1
C
67★
D' box
C' box
D box 2
C box
D box 3
/A 155
5' 3'


C box
D box 1
C 70
U 71
5' 3'

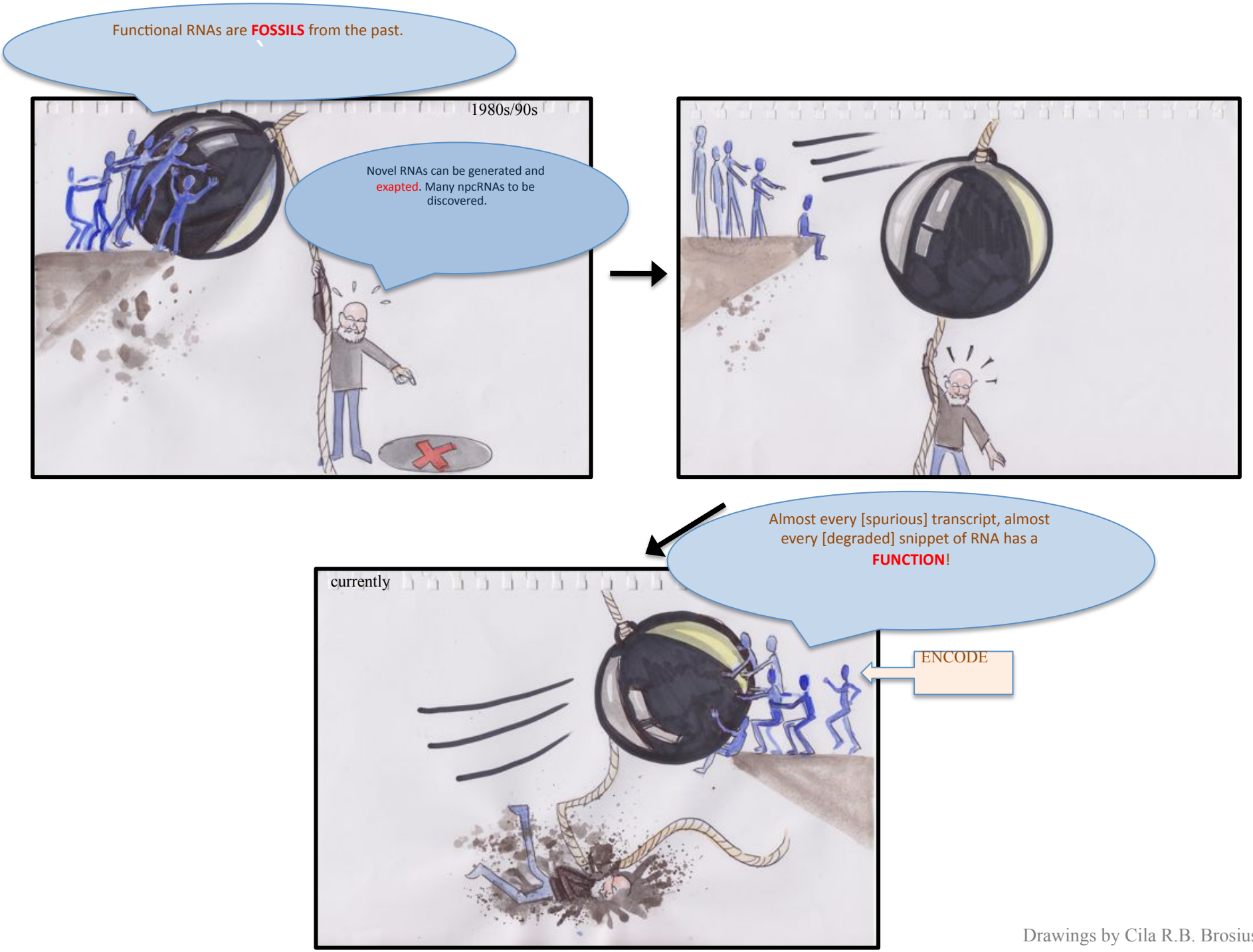

C box
D box 3
C 151
U 152
5' 3'

New long-snoRNAs





Drawings by Cila R.B. Brosius



Drawings by Cila R.B. Brosius

The Persistent Contributions of RNA to Eukaryotic Gen(om)e Architecture and Cellular Function

Jürgen Brosius

Institute of Experimental Pathology (ZMBE), University of Münster, D-48149 Münster, Germany

Correspondence: RNA.world@uni-muenster.de

Currently, the best scenario for earliest forms of life is based on RNA molecules as they have the proven ability to catalyze enzymatic reactions and harbor genetic information. Evolutionary principles valid today become apparent in such models already. Furthermore, many features of eukaryotic genome architecture might have their origins in an RNA or RNA/protein (RNP) world, including the onset of a further transition, when DNA replaced RNA as the genetic bookkeeper of the cell. Chromosome maintenance, splicing, and regulatory function via RNA may be deeply rooted in the RNA/RNP worlds. Mostly in eukaryotes, conversion from RNA to DNA is still ongoing, which greatly impacts the plasticity of extant genomes. Raw material for novel genes encoding protein or RNA, or parts of genes including regulatory elements that selection can act on, continues to enter the evolutionary lottery.

Everything has been said already, but not yet by everyone.

—Karl Valentin

Sturgeon's Revelation: *Ninety percent of science fiction is crud, but then, ninety percent of everything is crud.*

—Theodore Sturgeon

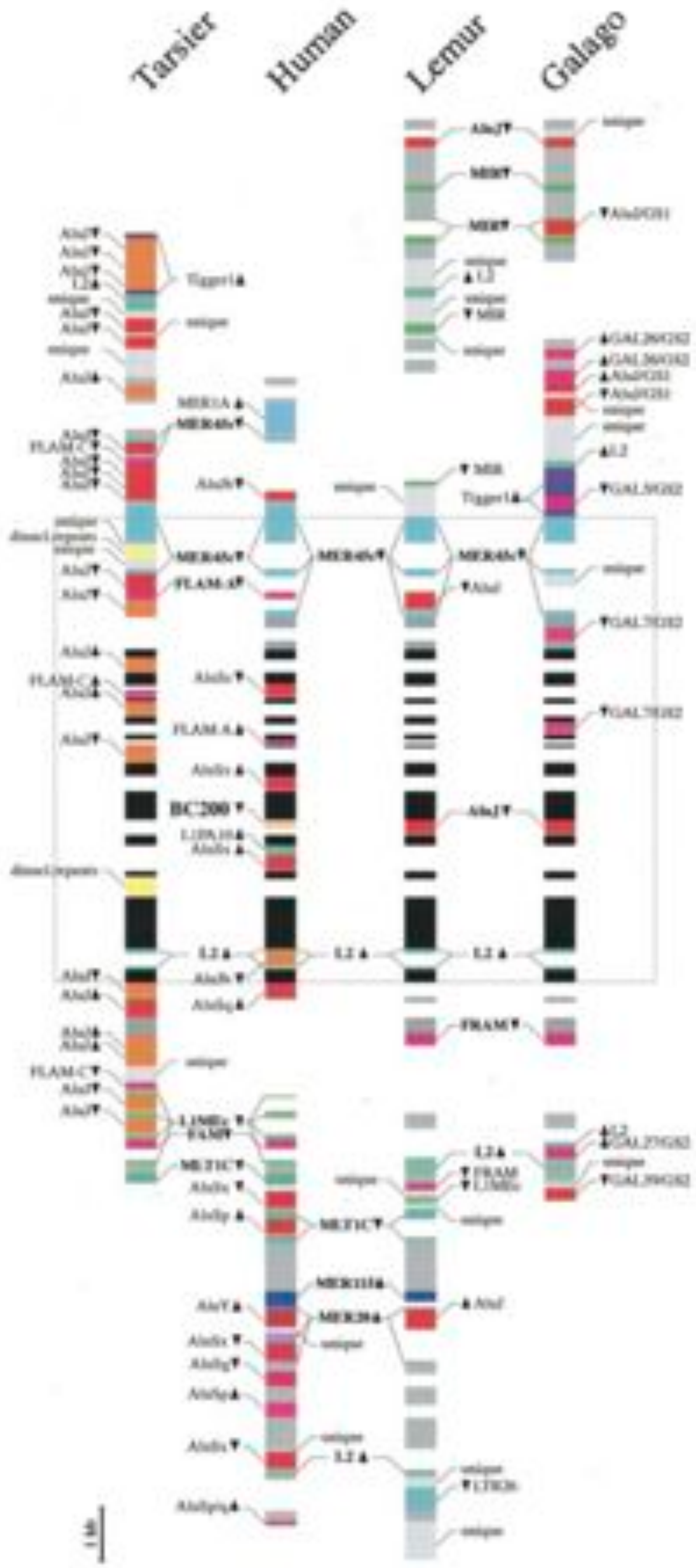
They think that intelligence is about noticing things that are relevant (detecting patterns); in a complex world, intelligence consists in ignoring things that are irrelevant (avoiding false patterns).

—Nassim Nicholas Taleb (Taleb 2010)

Of all extant cellular macromolecules, RNA is the most ancient, persisting as much as 4×10^9 years in our planet's life-forms. The ability to combine genotype with phenotype such as catalytic activity (Noller and Chaires 1972; Kruger et al. 1982; Guerrier-Takada et al. 1983; Noller et al. 1992) leveled a major hurdle in understanding the origin of life. The salient discoveries eliminated the virtually impossible prerequisite for two to three different classes of macromolecules to converge as an evolving unit. At the same time, RNA provides a required continuity in the path of evolution (Yarus 2011) during various genetic takeovers or evolutionary transitions (Cairns-Smith 1982; Szathmáry and Smith 1995). In a remarkably insightful article dating back half a century, Alex Rich foresaw much of what now is becoming mainstream, for example, that RNA was ancestral to protein and DNA (Rich 1962). This landmark publication received little attention over the years; even early proponents of an RNA world

Editors: Patrick J. Keeling and Eugene V. Koonin
Additional Perspectives on The Origin and Evolution of Eukaryotes available at www.cshperspectives.org

Copyright © 2014 Cold Spring Harbor Laboratory Press; all rights reserved.
Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a016089


Tarsier
Human
Lemur
Galago
BC200

although occasionally recruitment of extra DNA as novel gene module...

Transcriptional noise as a driver of gene evolution

Dmitrii Polev *

Department of Genetics and Breeding, St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg 199034, Russia

ARTICLE INFO

Article history:
Received 30 December 2010
Received in revised form
16 September 2011
Accepted 3 October 2011
Available online 10 October 2011

Keywords:
Tumor
Testis
Transcriptional noise
In-service evolution

ABSTRACT

As novel genes emerge in the evolution of species, pre-existing genes expand their expression patterns to diversify their functions and the expression patterns of gene duplicates diverge to pursue functional specialization. All these processes require genes to be expressed, however, the level and specificity of gene expression at the early stages of these processes are unclear. In this study, I propose that transcriptional noise is a mechanism to test genes for new functions, and I hypothesize the 'in-service' mechanism of gene evolution. In contrast to other hypotheses that suggest that there are specialized sites for gene evolution, such as tumors (Kozlov, 2010) or the testis (Kaessmann, 2010) this hypothesis proposes that emerging genes are expressed nonspecifically in many normal tissues, due to transcriptional noise. New genes are continuously 'tested' in various cells and under various conditions, thereby allowing the genes to evolve functions at the sites of their future work. The hypothesis of 'in-service' gene evolution also proposes that pre-existing genes are continuously tested under extrinsic conditions, due to transcriptional noise; this testing facilitates the emergence of alternative promoters and the diversification of the genes' expression patterns and functions.

© 2011 Elsevier Ltd. All rights reserved.

Opinion TRENDS in Genetics Vol.21 No.5 May 2005

Full text provided by www.sciencedirect.com

Waste not, want not – transcript excess in multicellular eukaryotes

Jürgen Brosius

Institute of Experimental Pathology, ZMBE, University of Münster, Von-Esmarch-Str. 56, Münster, Germany

There is growing evidence that mammalian genomes produce thousands of transcripts that do not encode proteins, and this RNA class might even rival the complexity of mRNAs. There is no doubt that a number of these non-protein-coding RNAs have important regulatory functions in the cell. However, do all transcripts have a function or are many of them products of fortuitous transcription with no function? The second scenario is mirrored by numerous alternative-splicing events that lead to truncated proteins. Nevertheless, analogous to 'superfluous' genomic DNA, aberrant transcripts or processing products embody evolutionary potential and provide novel RNAs that natural selection can act on.

Concluding remarks

Once more, economy recapitulates organismic evolution [40]. Like the music industry, which releases thousands of singles onto the market each year only a small percentage of which receive significant air play, the cell continuously churns out thousands of transcripts assembled from (parts of) retroposons, existing genes and/or the ever-changing, randomized genomic mass. Only a minority ever 'hit the charts' of purifying selection.

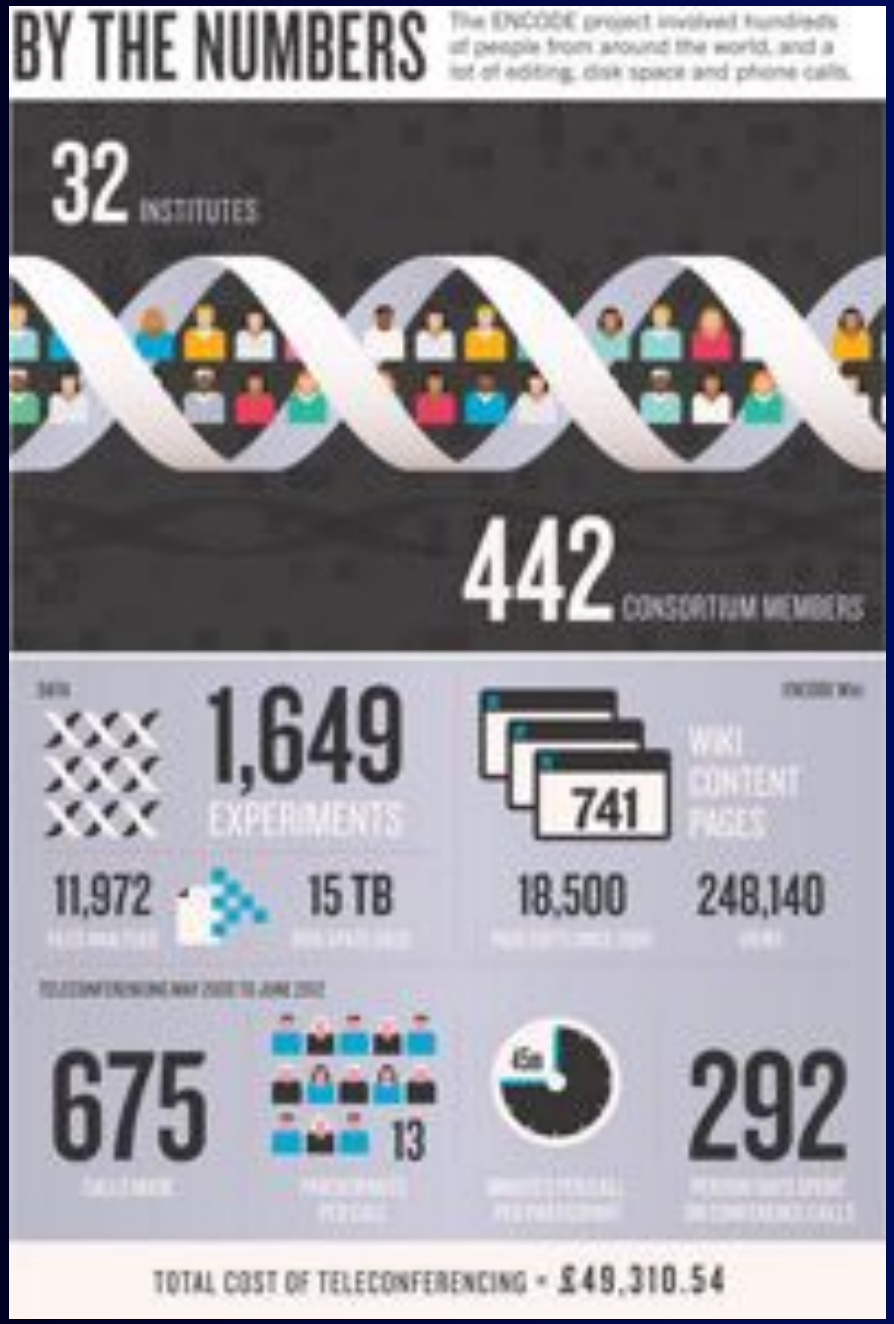
Darwin already recognized that nature can be wasteful and yet successful – 'Nor ought we to marvel if all the contrivances in nature be not, as far as we can judge, absolutely perfect; and if some of them be abhorrent to our ideas of fitness. We need not marvel at... the astonishing waste of pollen by our fir trees' [26]. He recognized the extravagance of nature, which is equally apparent in the cellular content and genomes of most multicellular organisms. The aforementioned examples of molecular waste are consistent with Darwin's natural selection and the concept of evolution by tinkering [27], illustrating the positive aspect of wasting resources.

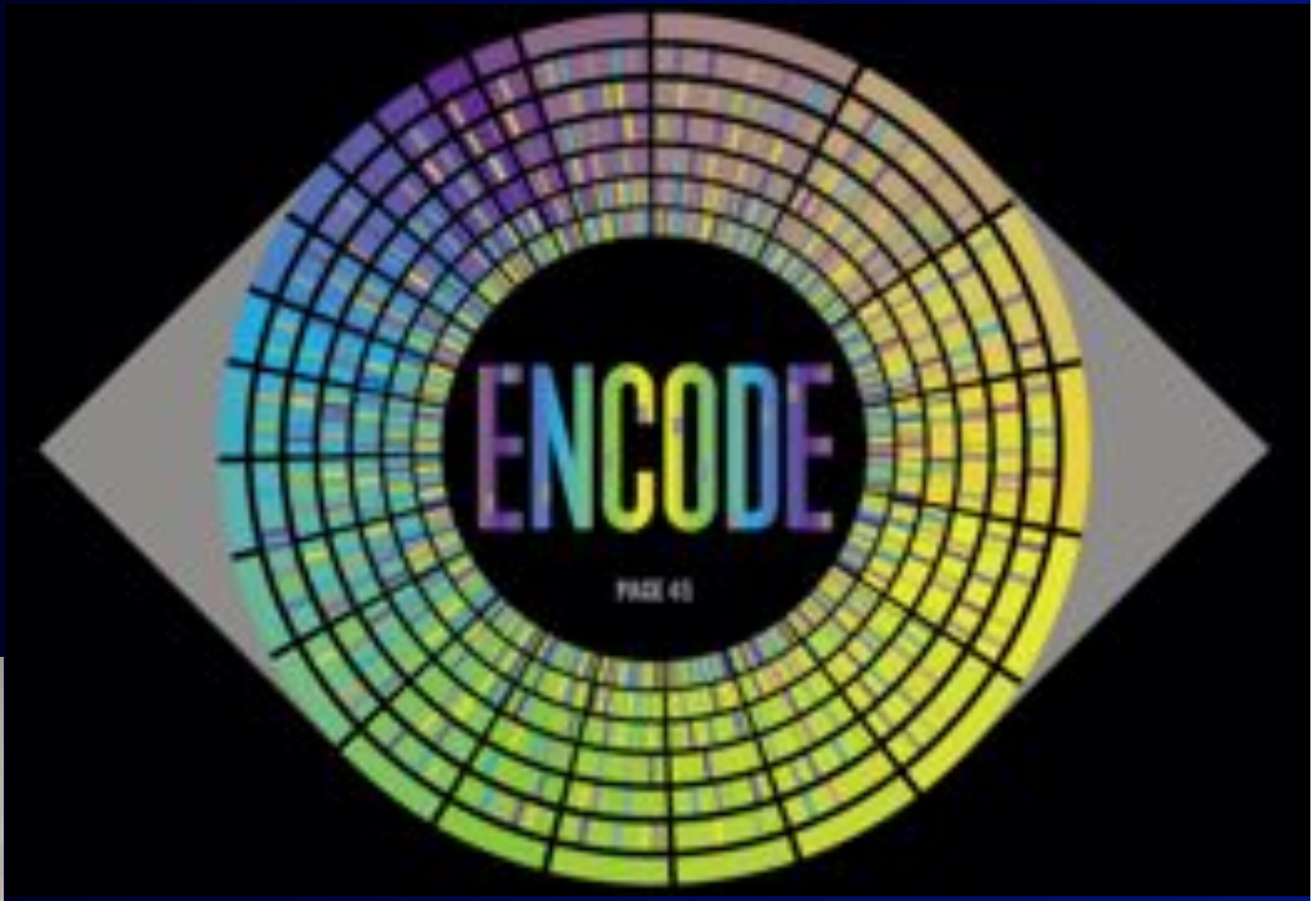
are two copies present. An accident of this type may seem to offer no selective advantage. However, during the process of mutation, alterations in these bases may occur in one of these copies but not in the second one which codes for the same protein. This means that a mutation which might be lethal for the organism would not be lethal in this duplex state. Accordingly an anomaly of this type allows for the development of a variety of new protein molecules which can, in a sense, explore the environment more readily than the original organism with only a single copy of its genetic information. Here the luxury of surplus, redundant information appears to provide a selective advantage in evolution. If this process goes on several times, eventually the organism ends up with many genetic copies of an original prototype protein molecule. These copies might then evolve along somewhat separate evolutionary lines and give rise to classes of molecules which, though similar, are in fact different in many ways. It is possible that this may be the origin of certain classes of related proteins such as myoglobin and the α- and β-chains of hemoglobin, in which there are many similarities in both amino acid sequence and secondary structure.

one country, although on the ordinary view supposed to have been created and specially adapted for that country, being beaten and supplanted by the naturalised productions from another land. Nor ought we to marvel if all the contrivances in nature be not, as far as we can judge, absolutely perfect, as in the case even of the human eye; or if some of them be abhorrent to our ideas of fitness. We need not marvel at the sting of the bee, when used against an enemy, causing the bee's own death; at drones being produced in such great numbers for one single act, and being then slaughtered by their sterile sisters; at the astonishing waste of pollen by our fir-trees; at the instinctive hatred of the queen-bee for her own fertile daughters; at ichneumonidæ feeding within the living bodies of caterpillars; or at other such cases. The wonder indeed is, on the theory of natural selection, that more cases of the want of absolute perfection have not been detected.

The complex and little known laws governing the production of varieties are the same, as far as we can judge, with the laws which have governed the production of distinct species. In both cases physical conditions seem to have produced some direct and definite effect, but how much we cannot say. Thus, when

The significance of RNA:
always a debate!


BY THE NUMBERS
The ENCODE project involved hundreds of people from around the world, and a lot of editing, disk space and phone calls.
32 INSTITUTES
442 CONSORTIUM MEMBERS
1,649 EXPERIMENTS
741
WIKI CONTENT PAGES
11,972
15 TB
18,500
248,140
675
13
292
TOTAL COST OF TELECONFERENCING = £49,310.54



US$: 280 000 000