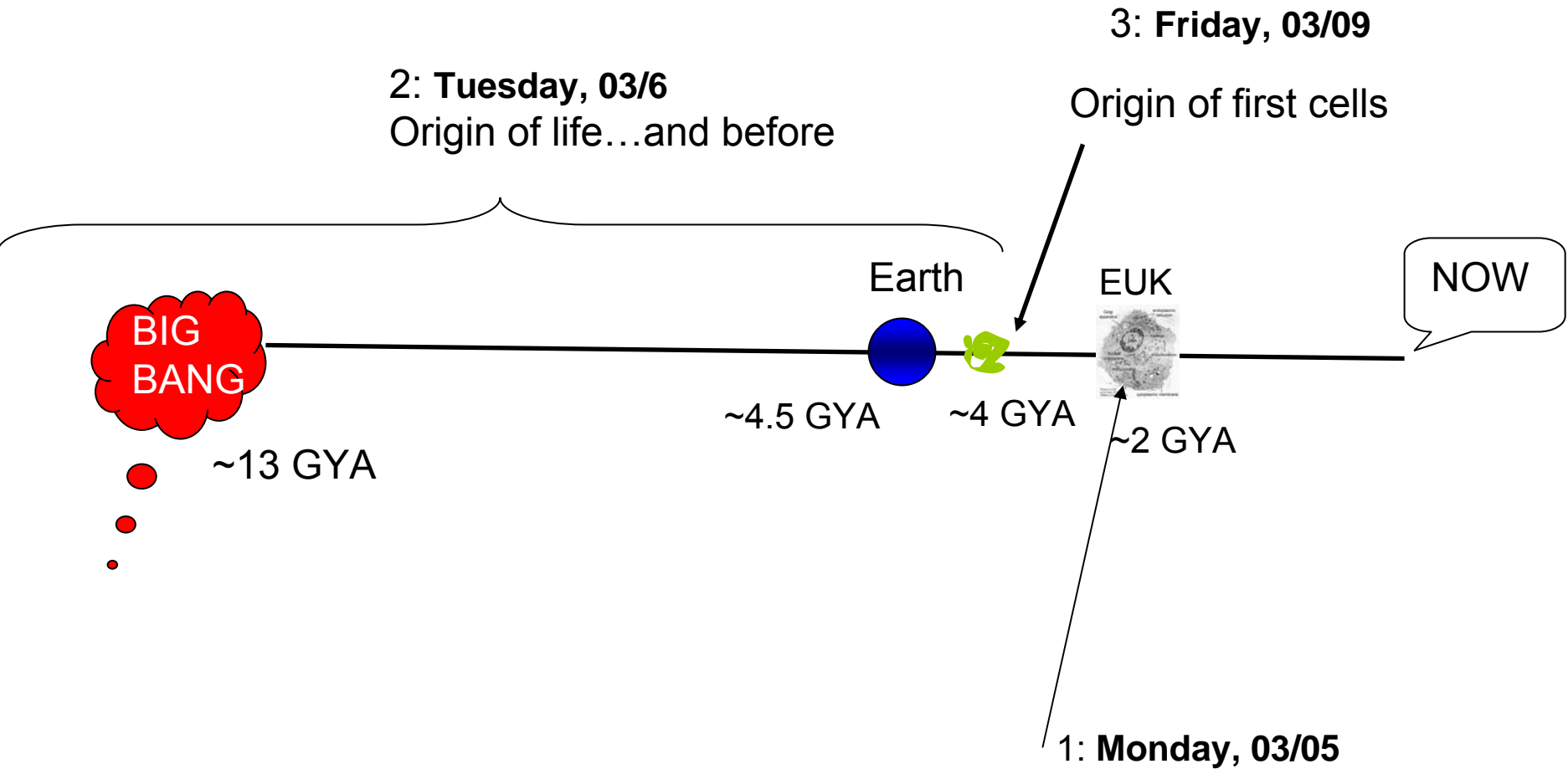


A three-part origin saga



On the origin of viruses and cells: the ancient virus world as the cradle of cellular life

Eugene V. Koonin

National Center for Biotechnology Information, NIH,
Bethesda, MD, USA

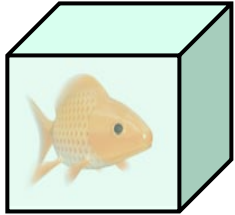
KITP, Santa Barbara, March 9, 2007

- Comparative genomics and evolution of viruses:
the ancient Virus World
- The nature of the Last Universal Cellular Ancestor (LUCA)
and the origin of cells
- An attempt on a grand synthesis: the viral model of the early
evolution of life

“The reputedly intractable problem of the origin of viruses has long been neglected. In the modern literature, ‘virus evolution’ has come to refer to studies more akin to population genetics, such as the worldwide scrutiny of new polymorphisms appearing daily in the H5N1 avian flu virus [1], than to the fundamental question of where viruses come from. This is now rapidly changing, as a result of the coincidence of bold new ideas (and the revival of old ones), the unexpected spectacular features of some recently isolated viruses [2,3], as well as the steady increase in the numbers of genomic sequences for ‘regular’ viruses and cellular organisms, which enhances the power of comparative genomics [4].”

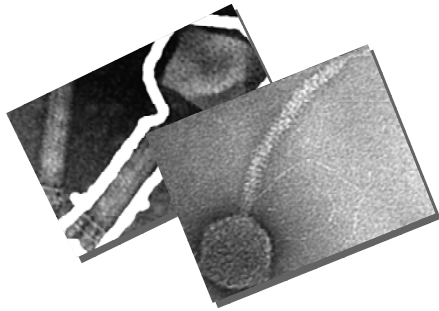
J.-M. Claverie. 2006. Viruses take center stage in cellular evolution.
Genome Biol. 7: 110

*The first paradox of virus evolution:
The metagenomes of viruses and cellular organisms
have comparable complexities*



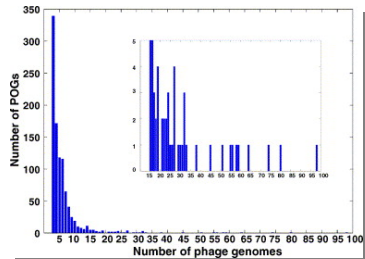
1 cm³ of seawater contains 10⁶-10⁹ virus particles;
~10 per each cell

Suttle, C.A. (2005) *Nature* **437**:356



Millions of distinct bacteriophage
species in the water, soil, and gut

Edwards and Rohwer (2005) *Nat. Rev. Microbiol.* **3**:504

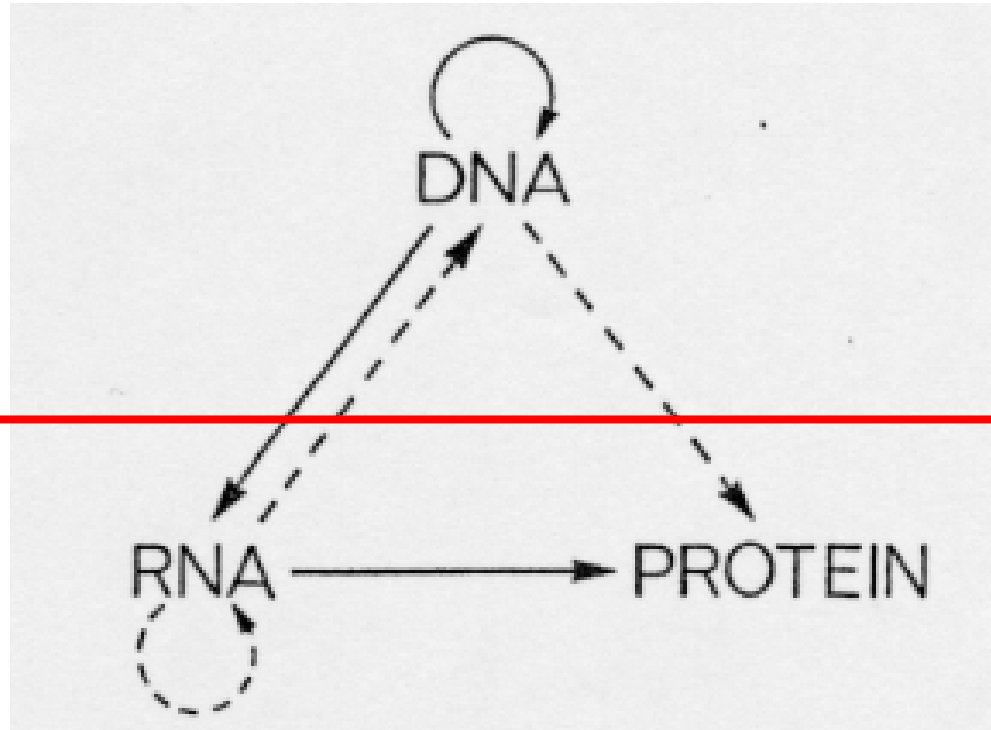


~ 90% of phage genes have no homologs in bacterial
genomes with large proportion of unique genes

Liu *et al.* (2005) *Virus Res.* **117**:68

The second paradox of virus evolution:
the diversity of viral genetic cycles
as opposed
to the single genetic cycle of cellular organisms

Francis Crick's classification of information transfer pathways, 1970



Transfers relevant for origin of life

A general transfer is one which can occur in all cells. The obvious cases are

DNA --> DNA

DNA --> RNA

RNA --> Protein

A special transfer is one which does not occur in most cells, but may occur *in special circumstances*.

RNA --> RNA

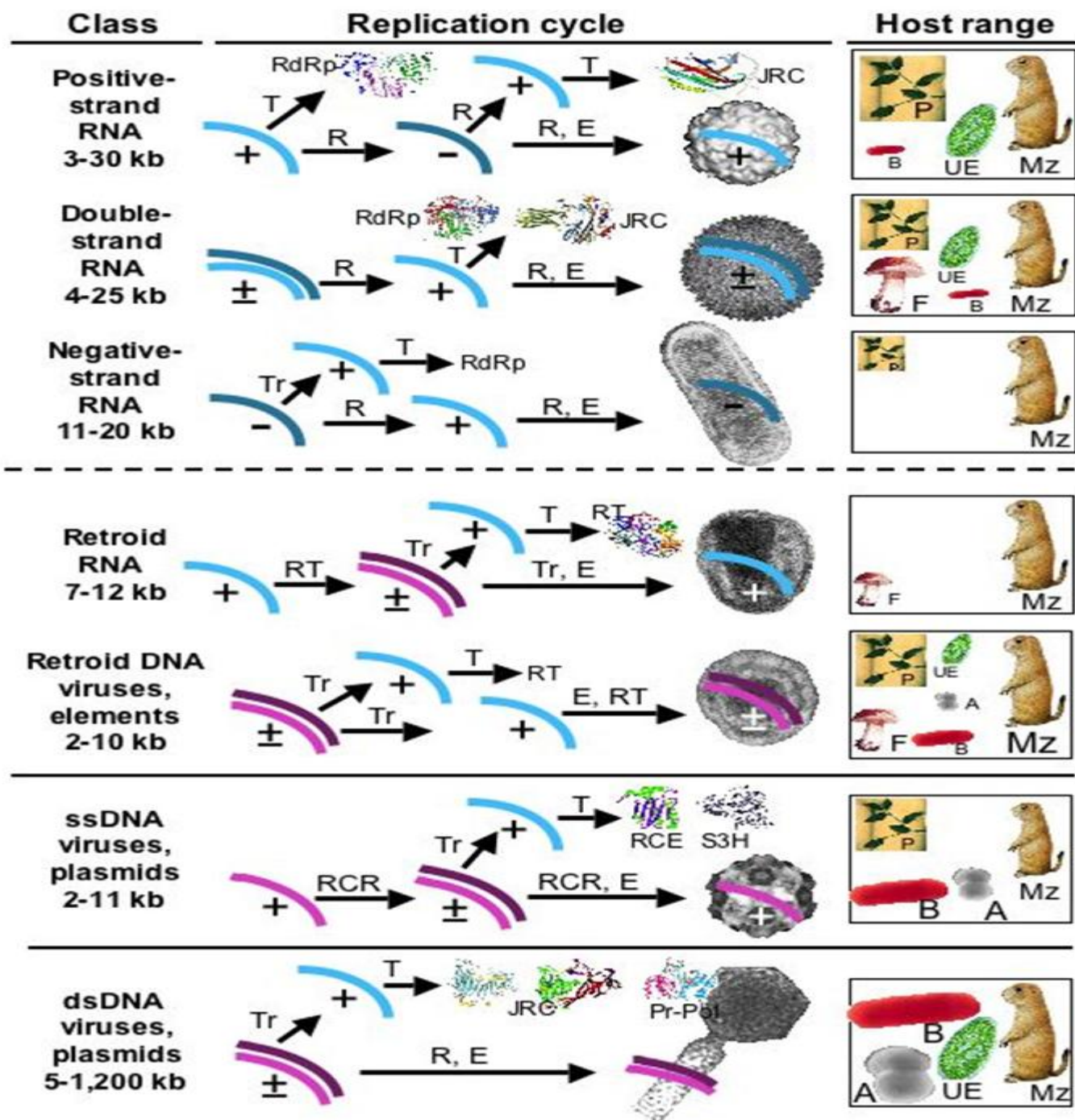
RNA --> DNA

DNA --> Protein

It [classification] was intended to apply only to present-day organisms, and not to events in the remote past, such as the origin of life or the origin of the code

Crick FHC, in Symp. Soc. Exp. Biol. The Biological Replication of Macromolecules, XII, 138 (1958).

Crick FHC, Central Dogma of Molecular Biology, Nature, vol. 227, pp. 561-563 (August 8, 1970)



Natural history of viral genes: a one-page summary of *viral comparative genomics*

I. Genes with readily detectable homologs from cellular life forms:

1. Genes with closely related homologs from cellular organisms (typically, the host of the given virus) present in a narrow group of viruses
2. Genes that are conserved within a virus lineage or even several lineages and have moderately close cellular homologs

Origin: relatively recent (1) or ancient (2) acquisition from host

II. Virus-specific genes

3. ORFans, i.e., genes without detectable homologs except, possibly, in closely related viruses
4. Virus-specific genes that are conserved within a virus lineage

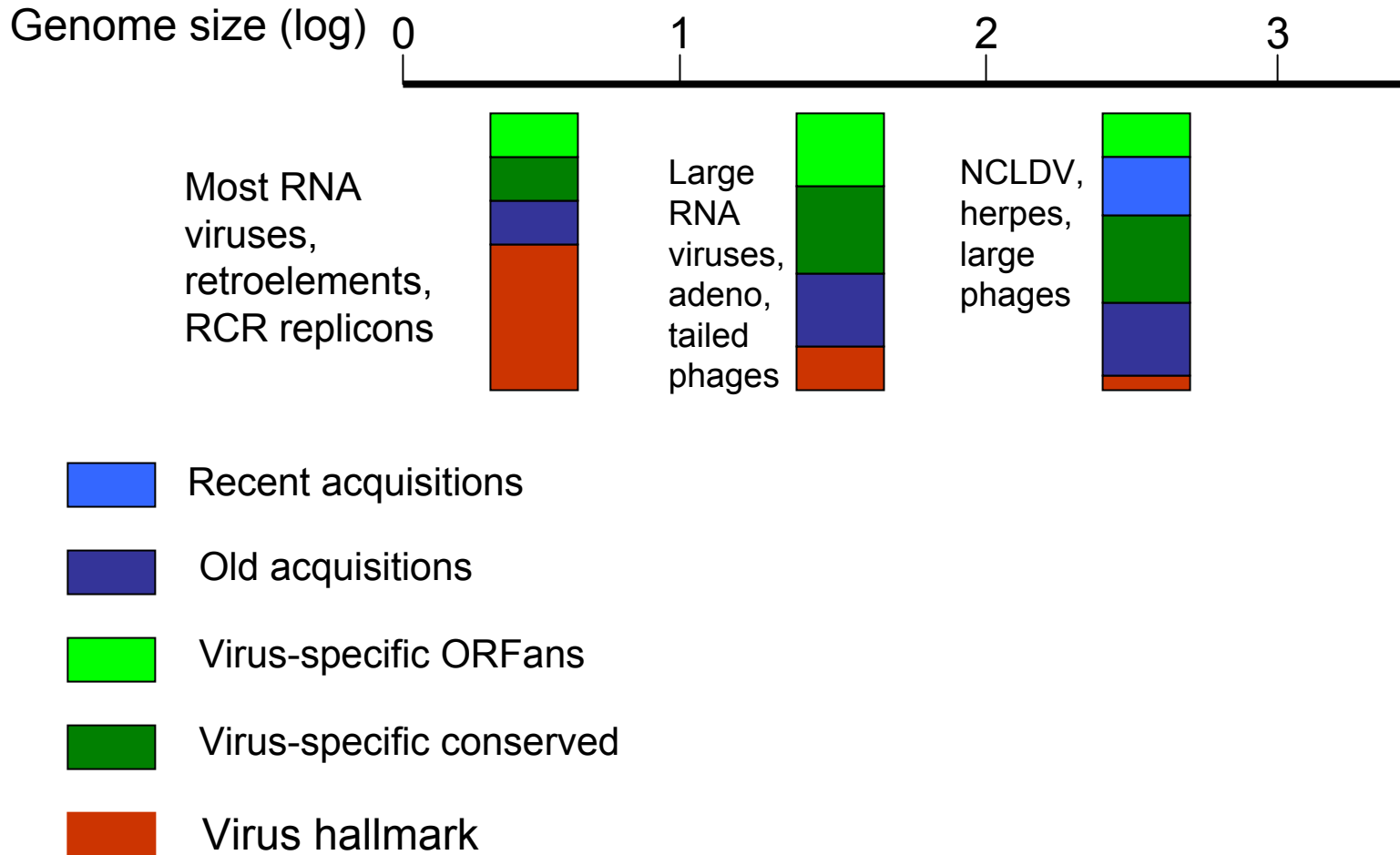
??

Acquisition from host but with rapid divergence from ancestor once within viral genomes?

III. Viral hallmark genes

5. Genes shared by many diverse virus lineages, with only very distant homologs in cellular organisms

Contributions of different classes of viral genes to the genomes of different classes of viruses: strong dependence on genome size



Natural history of viral genes: Viral Hallmark Genes

Shared by many diverse groups of viruses

Strong support for monophyly of all viral members of the respective gene families

Only distant homologs in cellular organisms

Play major roles in genome replication, packaging and assembly

Can be viewed as distinguishing characters of the 'virus state'

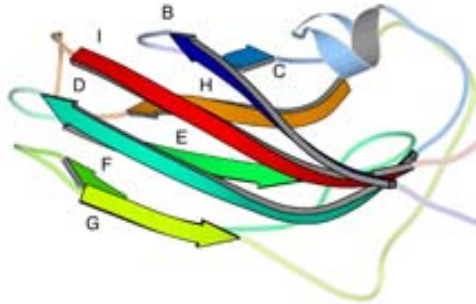
The hallmark viral genes/proteins

Proteins *crucial for virus replication and structure* that are shared by a broad variety of virus lineages but are only represented by very distant homologs (except for obvious acquisitions from viruses) in cellular life forms

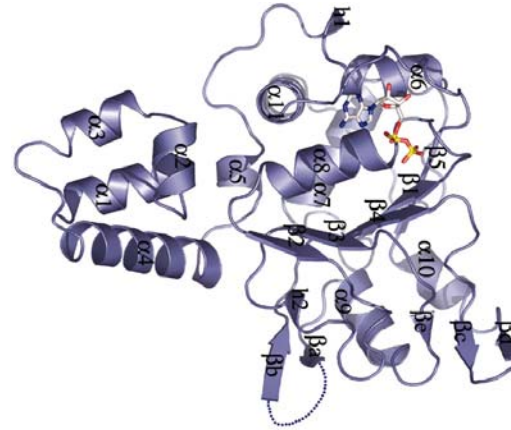
- *RNA-dependent RNA polymerase – most RNA viruses – eukaryotes and bacteria*
- *Reverse transcriptase – retroid viruses/retroelements/group II introns – eukaryotes and bacteria (very few instances in archaea)*
- **Superfamily III helicase – numerous groups of RNA, small DNA, and large DNA viruses – eukaryotes and bacteria**
- Distinct family of archaeo-eukaryotic primases – larger DNA viruses and plasmids - eukaryotes and archaea
- Rolling circle replication initiator nuclease – a variety of small DNA viruses, plasmids, and transposons – bacteria, archaea, and eukaryotes
- Distinct (A32-like) family of DNA-pumping ATPase involved in packaging – various large and small DNA viruses – bacteria and eukaryotes
- **Icosahedral capsid protein (jelly-roll domain) – most icosahedral viruses, from smallest RNA viruses to mimivirus**

For each of these proteins, homologs with the same fold are found in cellular life forms but monophyly of viral proteins is beyond reasonable doubt

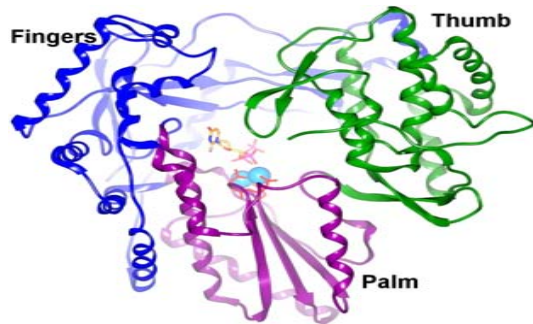
A few examples of viral hallmark proteins



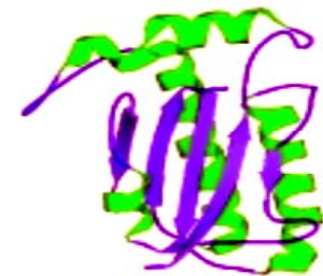
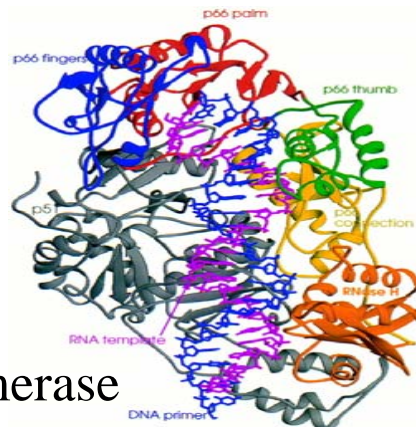
Jelly-roll capsid protein



Superfamily 3
helicase



RNA-dependent RNA polymerase
and Reverse transcriptase



Rolling circle
replication
initiation
endonuclease

Jelly-roll capsid proteins are found in an astonishing variety of RNA and DNA viruses



Bcl-A
bacterium
(anthrax)



C1-q
human
(TNF family)

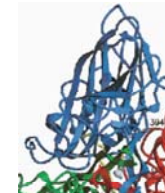
cellular jelly-roll proteins



BMV CP
plant
bromovirus



HRV-14 VP2
human
picornavirus



PaV CP
insect
nodavirus

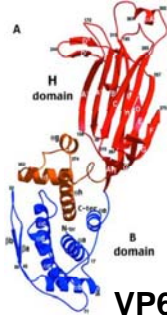


TNV CP
plant
carmovirus

ssRNA viruses

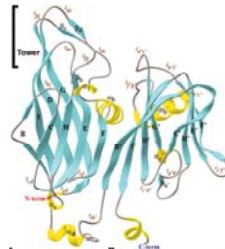


IBDV VP2
birnavirus

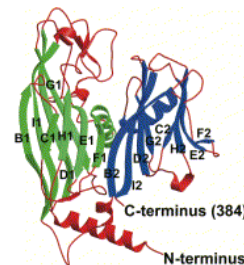


VP6
human rotavirus

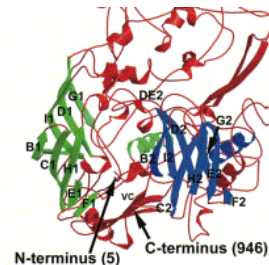
dsRNA viruses



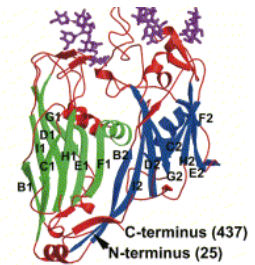
STIV major CP
archaeal virus



PRD1 P3
bacteriophage



Ad5 Hexon
human adenovirus



PBCV-1 Vp54
algal virus

dsDNA viruses (double-barrel jelly-roll)

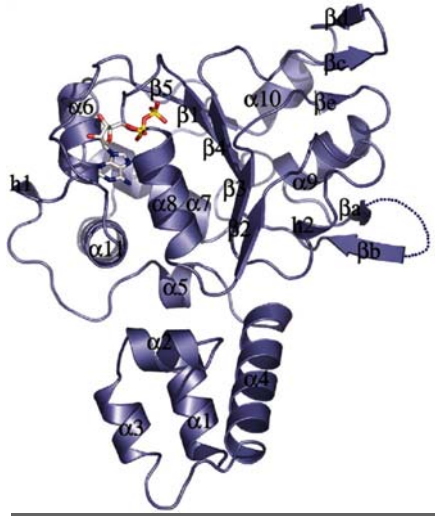
Superfamily 3 DNA helicases from diverse viruses are very similar

A

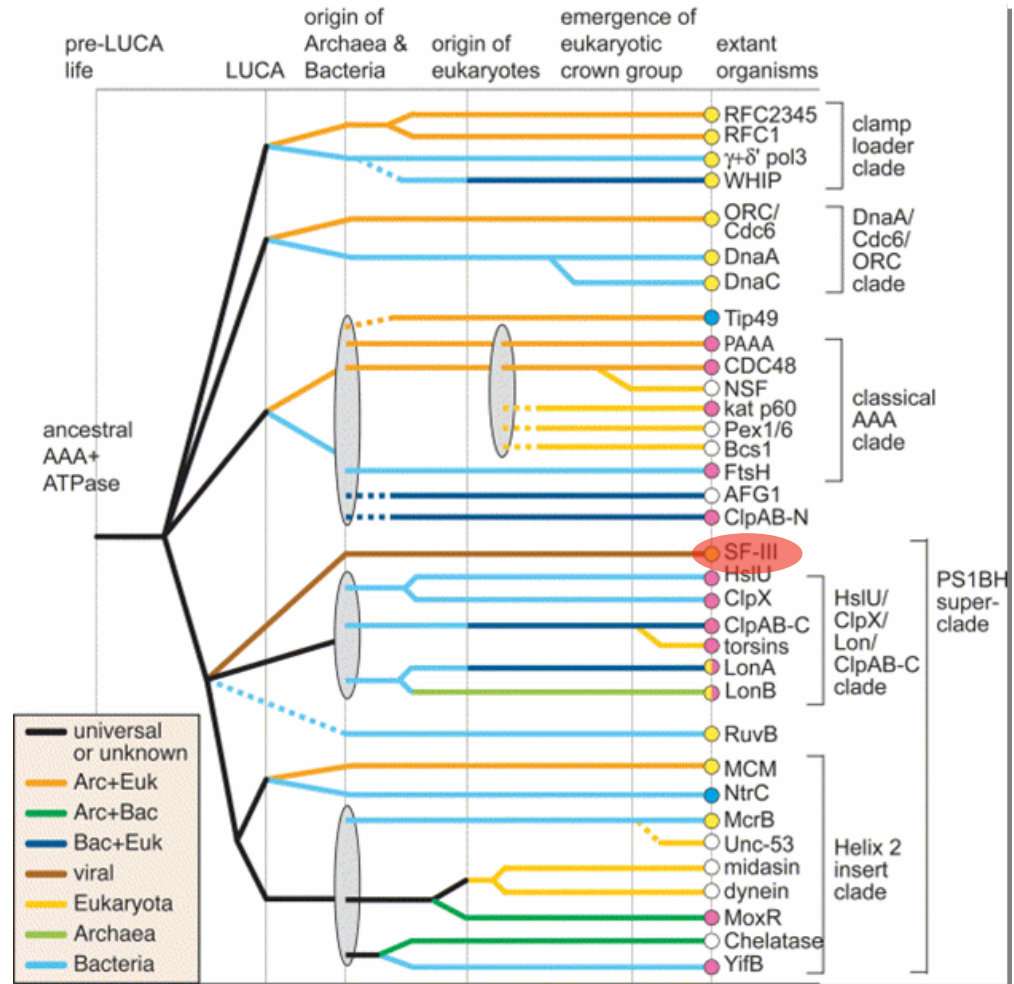
		D5- family specific motifs				Walker A		
MSV089_MSUV_9631496	432	TYGTPNPIIGGKGGIPFQKQ-NIPIYIGIEAKKYIKMYPHPIESY	3	EEMSEEE	6	NYNMLVTV	5	PENHDKR---KIPFENISS-SLQD---ISKRTITFLVQPTKSGKTTIKELL-AALFQNLFLTIP--IIDY
AMV087_AMV_9964401	431	TIICPNFYICQDNGVYDQK-SKPTTQENAKKYIRLAIYIKIMD	3	KEMSDDEE	6	MNMLLKSL	5	PKSNMPK---IVVPTNLSG-VLHY---CHKSVITFLVQPTSGKSTIKYLL-RQLFPMFLFPPFPQW
MC0948_MCV1_9629026	407	STDVYFERLQFADQVQIDAD-QSPFQADAKDFMCTVSTQVPL	5	ARDLALP	3	ELEMSLSD	2	FRSPQAMRNALTRKVLAS-CLCG---ATKPCIPFPFSTFATSGSTSKLL-QSAGHGLFLEGTQITLD
FPV058_FPV_9634728	409	STDTYFERLQFADQVQIDAD-QSPFQADAKDFMCTVSTQVPL	5	GIYFSDI	3	ELMSLSD	2	PKTKENFENRELTRQQLSS-CLMG---VTRQCIIPFPFSTFATSGSTSKLL-KSVGHHMFLSTGTQITLIE
D5R_VV_137602	405	STDTYFERLQFADQVQIDAD-QSPFQADAKDFMCTVSTQVPL	3	KFVDSFP	3	ELMMLDMD	2	PLTDENKRELTRKVLAS-CLCG---ATKPCIPFPFSTFATSGSTSKLL-KSAGLDFVSTGTQITLD
C962R_ASPV_9628179	543	FLDTNPRGLQVGNQVLSREV-IPAKLIDHFRHPIHTYTHICY	1	PNPHEHP	3	LLLALSD	2	FELDAR---LWIMFVLSF-AIFRG---LKEALLMLFQGGCCHRTFLARLY-AMFLDHWYAKKINLSLT
orf6_FLDV_13358409	505	KIDSNPYLMAFQNVYVYQV-KLFRQGRLED-YCCKSLTINMY	5	QLLCSNP	9	RTLFLRQ	2	FDIELR---VPIRGLAS-AFVGG---NSKICLFWTGGGCHRTITVTLM-EGMDFPVAIKKIPKGMVITG
CIV184R_CIV_2738452	576	KMDQNKQLIATFNQVYVLSL-PTFRQGLRPEY-YITQOMT-IPY	2	TLTMINP	3	KMLAFPK	2	FDRELV---EYFMKACE-WYVGG---NRKICLFWTGGGCHRTITVTLM-EGMDFPVAIKKIPKGMVITG
A456L_PBCV_9632024	282	PLMKSGRQVYFQNVYIAPF-DRFHCFTCKEPLSDSIVASK	2	EGMFINP	10	PHLDSIMI	2	QMEEEV---QHWLALLORILYPAVQVSDKQVIFPKGLAATQKSTILKVIKIPFTVVDGILLSDNVER
ORF109_ESV_13242580	230	DLANAKRYVSPNDIYDATE-DTFFTYDQD---IDELVLSCK	7	DVVFSTP	14	PLFSDSLQ	2	RMQDQ---LWQYTVGRLFYVSELSDKQVIFPKGLAATQKSTILKVIKIPFTVVDGILLSDNVER
gp9a_phic31_9631173	365	KFDANARLSPANVYVLRP-GKGRADKQDQ---MLTVLPL	6	DMQAAP	8	RMEQFLQ	2	PNMALD---VQVNSLVD-YGITQW-TSEQCFAVLMKGMKMGKSTVFTETL-TQVFRITKTIPTTATFD
C.RepA_S1_11137542	463	QLLEPLHIAFQNTLSTWGIKIWYDCKKRTYVAFVYHIV	7	EKNFNR	11	RVCFFALE	3	QWVDQ---MILYKTIIG-YTLYPK-YIPKAILLVNMGSGKSTFLNLL-LKILGNVNVTAUSKLRIM
C.Primase_LcbA2_6599326	425	KFDKIKILANTINGVYVLSL-GEKHIDHVVK---MPSKRTGVWY	7	TDVDSV	3	QPLQFDP	2	QMEEEV---IDYIQAIG-YSLTGS-TSEQCFALVLMKGMKMGKSTVFTETL-TQVFRITKTIPTTATFD
C.Primase_phi-105_7444375	427	KLDGKYLKPCNMGVYDQK-DGLPHDRDL---LPTKISVWAX	6	EKDACP	5	MESIFDR	3	PNYEV---IFPIKAIG-YSLTGS-TTQVQVFLVGMGSGKSTFLNLL-LKILGNVNVTAUSKLRIM
XP2505_XF_11362498	450	QLDTPMLANCAQNVYVLRP-GTKKAHFPED-YMKRVPVWY	7	PDAAAP	5	LERITCE	2	AQQPL---SPTQRWFG-YCATGS-VRHKLAVYVGMGSGKSTFLNLL-LKILGNVNVTAUSKLRIM
C.Primase_Ec_13359760	406	QSNPARQLIFQNVYVLRP-GLFSPHDKQK---MLRSLCKVY	7	TOVDGE	11	WLDRAAG	2	NPEK---RDIIILAL-PMVLANR-YDQGLFVTVGGGSGKSTFLAEL-TSLAGRDNAT--ATIME
p143_AcNPV_293982	907							DYQKALVYMGSGKSTFLAEL-TSLAGRDNAT--ATIME
E1_NPV08_137643	421							DYQKALVYMGSGKSTFLAEL-TSLAGRDNAT--ATIME
TALA_SV40_1351194	416							DYQKALVYMGSGKSTFLAEL-TSLAGRDNAT--ATIME
NS1_AAV_4160147	320							DYQKALVYMGSGKSTFLAEL-TSLAGRDNAT--ATIME
AL1_MstV_9719358	219							DYQKALVYMGSGKSTFLAEL-TSLAGRDNAT--ATIME
F2C_Polio_9627037	1246							DYQKALVYMGSGKSTFLAEL-TSLAGRDNAT--ATIME
consensus/85%		p.s....bh.h.Nghh-h.p..b.....c.....h.b.....p.....h.....hbb.h.s.h.h.....b.h.bb.G.s.gKoe...l...hh.....						

		Walker B	Motif C						
MSV089_MSUV_9631496	12	LGRKNGKLVCFASESN	1	CDKFDSTQIKMTERI---ISIKVANS---MEGDQENYATIIDTNIFP	30	SIDITDAILKRLAVINIDHSY	30	NPNFNRKIKNGDF--SLAFNHLKRWVRYKYL	716
AMV087_AMV_9964401	10	LQVVEKRLVSPASEGD	3	NEVFLNKNIKQYTEQY--ILGRDLNK---SKCVQNTLTQIDILNFKP	24	NPSSVDPALVKRLAVININETQ	24	STFD-DKILNNEF--TLQPFYELKRWKRYKHK	710
MC0948_MCV1_9629026	9	LAMHGLKRAVPCSELP	6	AKKIRADNVKQLTEPC--LVGRPCYS---NRINDRQATIIDTNVYK	18	VFKCVNDALGRVGLVRFPTH	18	AALD-RKIQSHYF--RFAFLLELLSNYQRHHA	681
FPV058_FPV_9634728	9	IAMBGLKRSVPCSELP	6	SKKIRSDNKKLTEPC--VVGRSCYS---MKINDRQATIIDTNVYK	19	VFKCVNDALGRVGLVRFPTH	19	ESLD-SKIQSHYF--RFAFLLELLSNYQRHHA	680
D5R_VV_137602	9	IAMBGLKRSVPCSELP	6	SKKIRSDNKKLTEPC--VIQRPCFS---MKINDRQATIIDTNVYK	22	VFKCVNDALGRVGLVRFPTH	22	EGLD-GKIQSHYF--RFAFLLELLSNYQRHHA	681
pC962R_ASPV_9628179	12	PNRKGGRQVYFPEST	1	SEVLTSTRKLVNPGD-VIARISEAQ---KQSPQENATVIAAIVNYP	23	IIDTIDHGTQULHRYRKYKVT	23	MPDCCQAPPFSL--VYFWELDKRYGQIKK	815
orf6_FLDV_13358409	12	VRTGGVGRMAVSEFP	1	DERIHAGIKLQNGEDF-FWAKLECYTGDQTKLIIIPMKLIVLGNL	21	IIDYADQAVRQVYVYFPEST	21	DLKFSKCLKILIE--PLVYTLIYGLANORLN	787
CIV184R_CIV_2738452	12	TRAGQVGRMAVSEFP	1	DETVMAGVYDQVGGIDVYARDEIQGQKQIDIDFPFKIYICMPTI	20	NIRNPNQATVRIKIVYFPEST	20	DTSPCEKRETRIEGLKAPVLYKLVQVPIKQKRD	850
A456L_PBCV_9632024	3	ISAFNKICYLVAPEIK	1	DLAIWQAEFQSHYSESD-VQVNVKIK---KAFABEHSVYPALAQHVEP	6	QAMDNGGSIQRNVVYFPEST	6	SNQD-MKLEKRESVQ--IVANAVGAVRYANGL	491
ORF109_ESV_13242580	3	ISTAYNKTIYVYVPELK	1	DFAMDQAEFQSHYVQSTY-LSPVWVNG---SPITQVMTTQVIAQGHESA	3	KMKDKSGISIKRIYVYFPEST	3	KVNP-GEMHIDQTELPIIPIKSLAVYRQVQ	491
gp9a_phic31_9631173	10	LAAERGRSIVASSEGE	1	GKFMSEAVLQVYQDKR-VIARIFLRQ---EFFTPAFTPLDLAATNHL	6	KPKSQDEGLKQVYVYFPEST	6	YDLD-RKLEKRESVQ--IVANAVGAVRYANGL	613
C.RepA_S1_11137542	6	PTEYHKLANINSELP	1	FKVYNTGLVRLTQGEDY-ILIKRKLRF---DMYFYINAKLINAATNHL	6	EYKQYQYQPRQVYVYFPEST	6	NPFDKTKRESIEG-IITVSYVAFSRVLAQKQL	678
C.Primase_LcbA2_6599326	10	IARLKGARLVTASEFN	1	GVRLDGLVRLTQGEDY-VIARIFLRQ---SEFYKPEFKLMLAATNHL	6	IIRGTDQYHQRLVYVYFPEST	6	KRLA-YKLEKRESVQ--IVANAVGAVRYANGL	723
C.Primase_phi-105_7444375	10	IARLKGARLVTASEFN	1	GVRLDGLVRLTQGEDY-VIARIFLRQ---SEFYKPEFKLMLAATNHL	6	IYVSGDESHKRLVYVYFPEST	6	KQLP-CKLAEKRFQ--ILKAVGACCLMQRKGL	680
XP2505_XF_11362498	9	IADVYGRMVTYNETS	1	QGLLSESLVQLTQGEDY-MSAKFLRQ---EYFETFEFKVYVYFPEST	6	VIKGGQVYVYVYVYFPEST	6	HKIT-ERLAEKRFQ--ILKAVGACCLMQRKGL	711
C.Primase_Ec_13359760	6	RAALYFSLTILYDQE	1	KWGDGAGKELTQGEDY-VYVDFRYQ---NAYSTHIFAVLIVVSRNFM	6	RFTDRQVYVYVYVYFPEST	6	TQLK-ERLAEKRFQ--IVVYQMLQRSPDHPKAR	653
p143_AcNPV_293982	8	YKSL-NSCGLYTHLQK	1	CSRSYFKKASDQSEK-DQKRFYQ---GLLKYANVYKLVVSRNFM	8		8		1031
E1_NPV08_137643	3	LQFADCKIALDQVY	4	LVMYF-FRIKGLDQVY-WLSCYKY---APMQLKFP--LLTSHNML	4		4		541
TALA_SV40_1351194	3	LGVAIQDFVAVVDFVY	12	QQINLDRHRYLQDQVYKHLKSHL---NKRTQIFPP-GIVTSHNEP	4		4		545
NS1_AAV_4160147	3	LGVAIQDFVAVVDFVY	12	QQINLDRHRYLQDQVYKHLKSHL---NKRTQIFPP-GIVTSHNEP	4		4		445
AL1_MstV_9719358	2	YHSDA---IYHIVDIP	3	CPC-NKGLVCCQKRPVMPKRYGK-KRVQKSKP-TIELAMDE	4		4		335
F2C_Polio_9627037	3	FQGVQQGVYVYVYVY	4	GAINKLCQNVSTVYFIPMASLE---EKGILPSTNYPVYVYVYVY	4		4		1373
consensus/85%		h.h.....h.h.c.....p...hbphs.....h.p.b.....p.....hh.ssh.psh.hh.h...h.....bb.....h.b.b.h...p...							

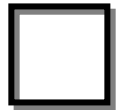
Helicase superfamily 3 is a monophyletic, deep-branching viral clade within a major class of P-loop NTPases



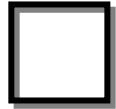
Present in a vast variety of viruses: from the smallest (+) RNA viruses (picornaviruses) to the largest DNA viruses (e.g., mimivirus)



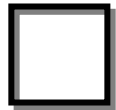
What is the evolutionary explanation for the striking features of Viral Hallmark Genes?



Artifact of the rapid sequence divergence of viral genes



Heritage of the last universal common ancestor of viruses



Spread by horizontal gene transfer



Origin from a pre-cellular phase of life's evolution: A primordial gene pool

Shared by many diverse groups of viruses

Strong support for monophyly of all viral members of the respective gene families

Only distant homologs in cellular organisms

Play major roles in genome replication, packaging and assembly

Can be viewed as distinguishing characters of the **V**irus state

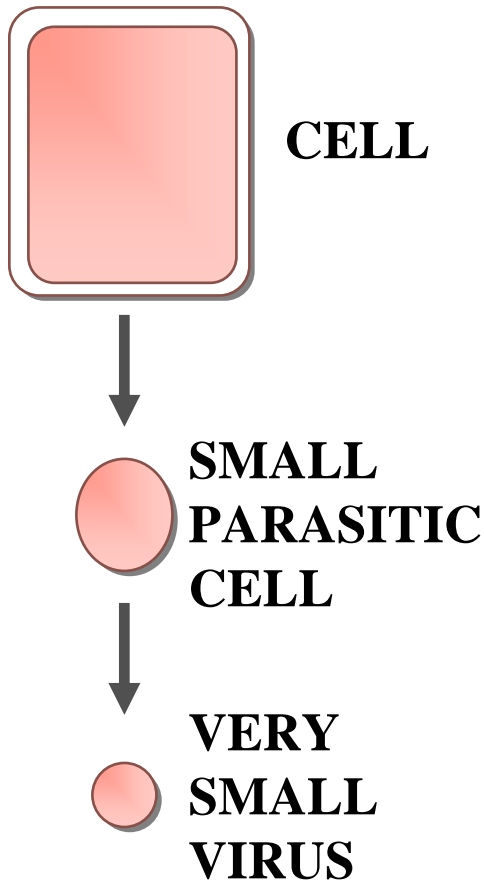
The origin and evolutionary history of hallmark viral genes - the primordial gene pool hypothesis:

The hallmark genes AND, by implication, the major lineages of modern viruses descend directly from a primordial gene pool

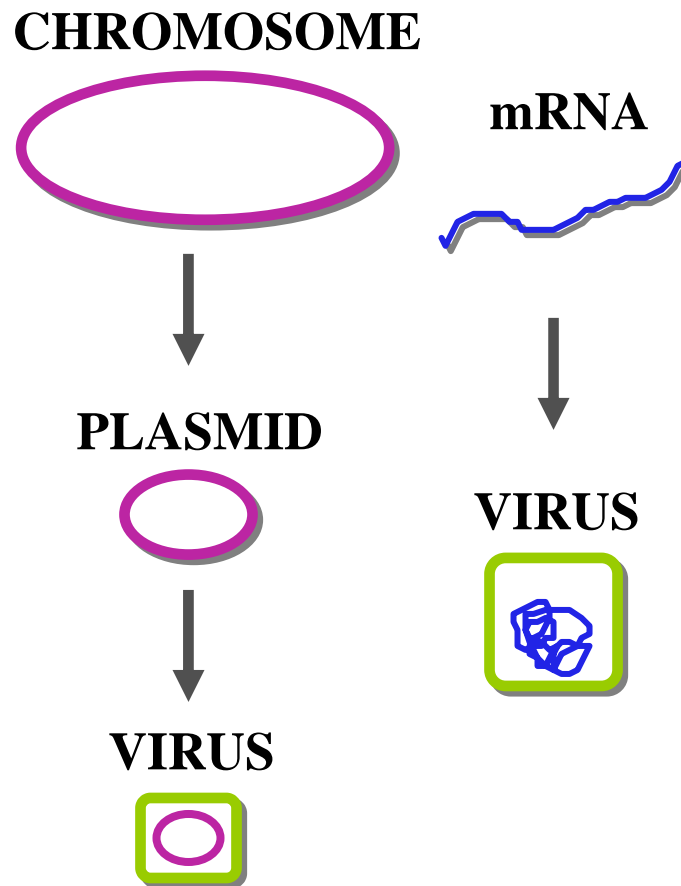
A crucial corollary: If viruses come directly from a primordial gene pool, then, origin of viruses is *inextricably linked to the origin of cells*

Competing concepts of the origin of viruses

Cell degeneration

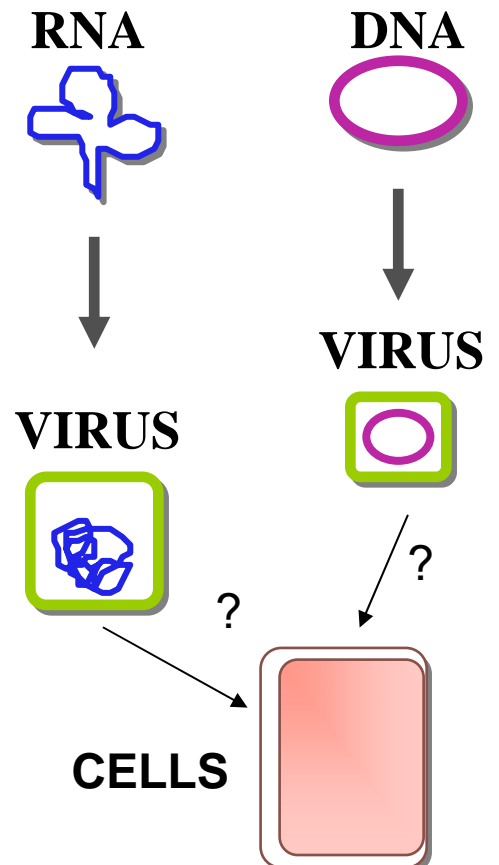


Escaped genes



Primordial genetic systems

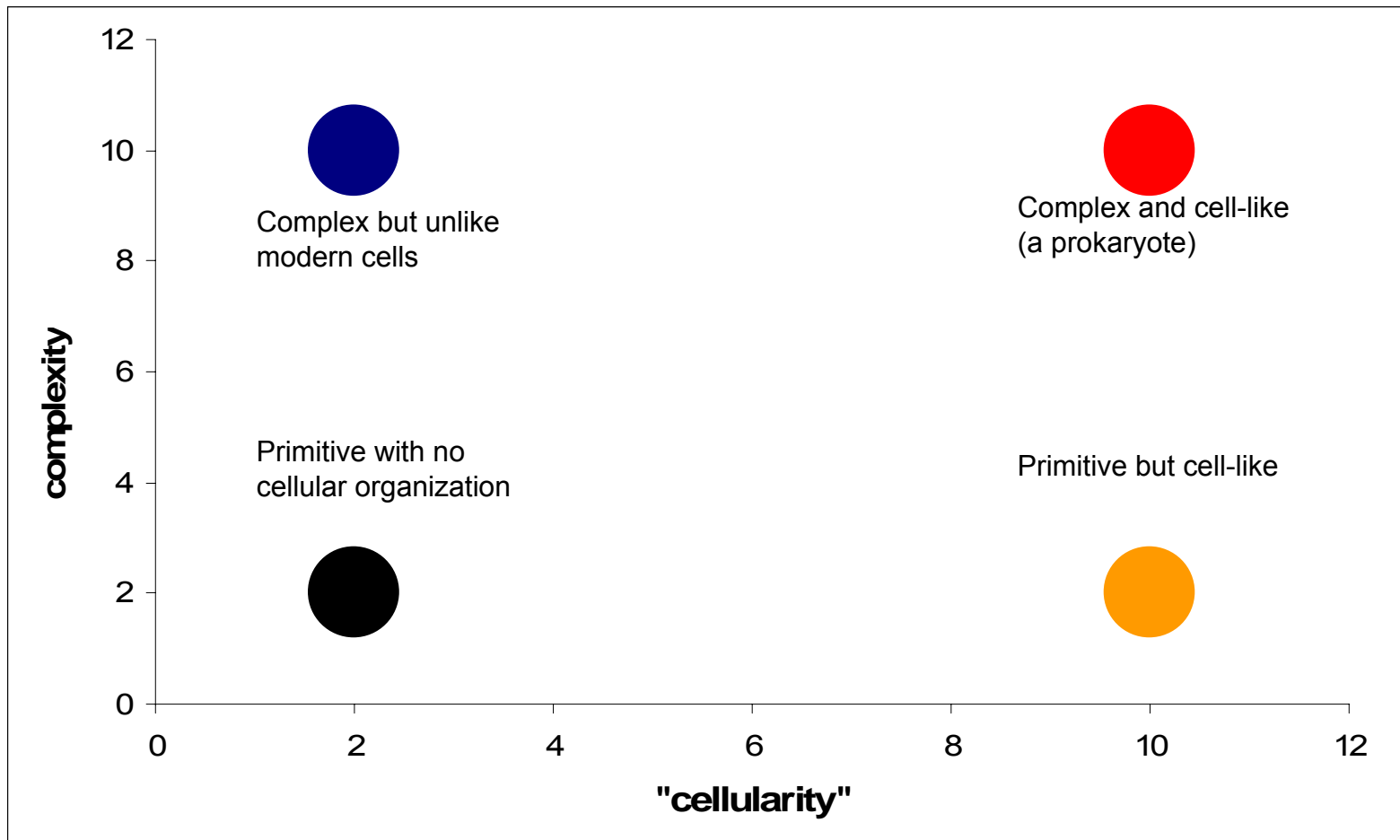
PRE-CELLULAR LIFE FORMS



Last Universal Common Ancestor: **LUCA**

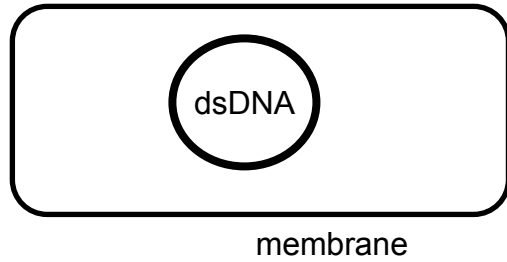
- The “uniformitarian” view (Hutton-Lyell-Darwin):
LUCA must have been like modern prokaryotes, perhaps, somewhat simpler
- The radical view: LUCA was dramatically different from any extant cell...at least in some respects

LUCA: complexity *versus* "cellularity"

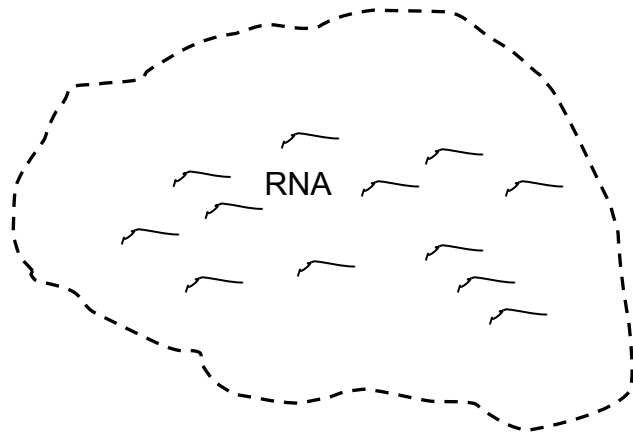
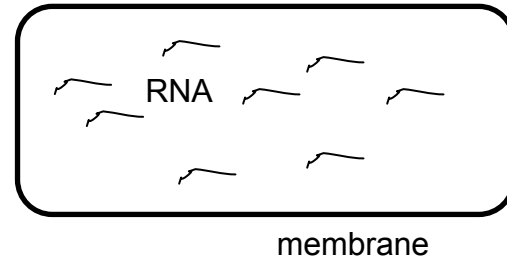


Different versions of LUCA

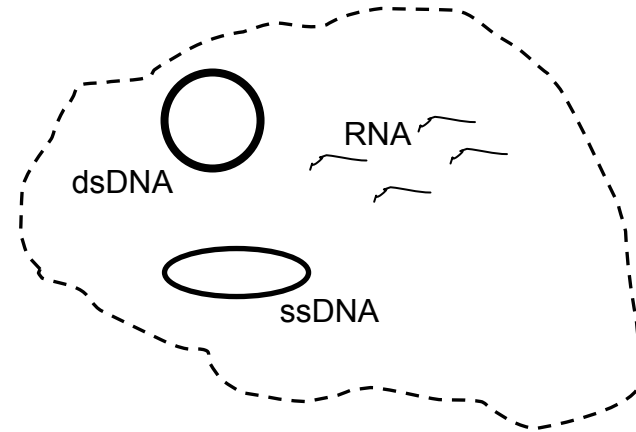
Modern-type cell



RNA cell



Pre-cellular stage, mixture of RNA segments



Pre-cellular stage, mixture of diverse genetic elements

What does comparative genomics have to say about LUCA?

At least 2 major problems

Apparently unrelated in archaea and bacteria:

- **membrane chemistry/biogenesis**
- **core DNA replication systems**

A radical ...but parsimonious solution: LUCA had neither

- modern-type membranes
- modern-type DNA replication

[Leipe DD](#), [Aravind L](#), [Koonin EV](#).

Did DNA replication evolve twice independently?

Nucleic Acids Res. 1999 Sep 1;27(17):3389-401

[Martin W](#), [Russell MJ](#).

On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells.

Philos Trans R Soc Lond B Biol Sci. 2003 Jan 29;358(1429):59-83

[Koonin EV](#), [Martin W](#). **On the origin of genomes and cells within inorganic compartments.**

Trends Genet. 2005 Dec;21(12):647-54

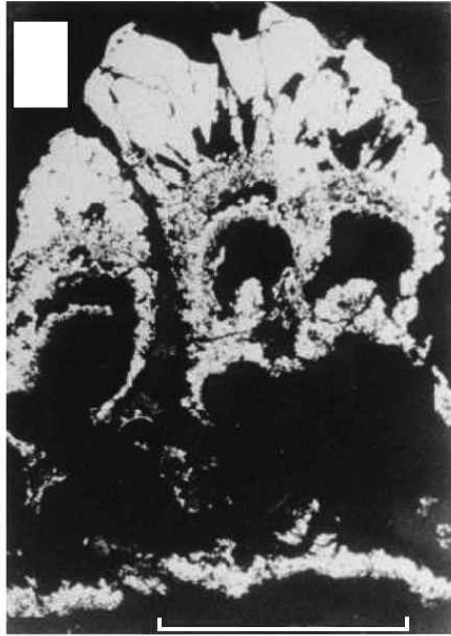
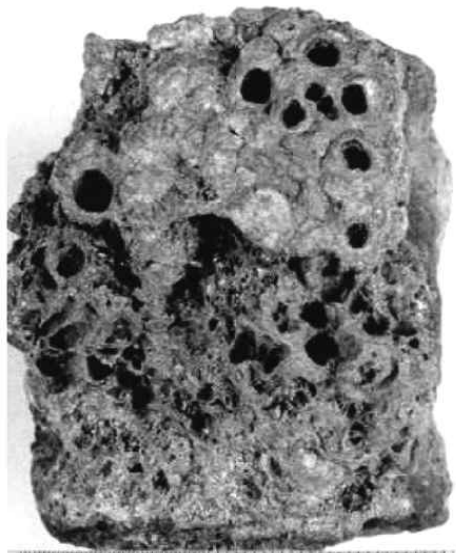
A membrane-less, non-cellular but spatially confined LUCA: a hypothesis to explain the disparity of membranes in bacteria and archaea

[Martin W](#), [Russell MJ](#).

On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells.

Philos Trans R Soc Lond B Biol Sci. 2003 Jan 29;358(1429):59-83

- We propose that life evolved in structured iron monosulphide precipitates in a seepage site hydrothermal mound at a redox, pH and temperature gradient between sulphide-rich hydrothermal fluid and iron(II)-containing waters of the Hadean ocean floor. The naturally arising, three-dimensional compartmentation observed within fossilized seepage-site metal sulphide precipitates indicates that these inorganic compartments were the precursors of cell walls and membranes found in free-living prokaryotes.



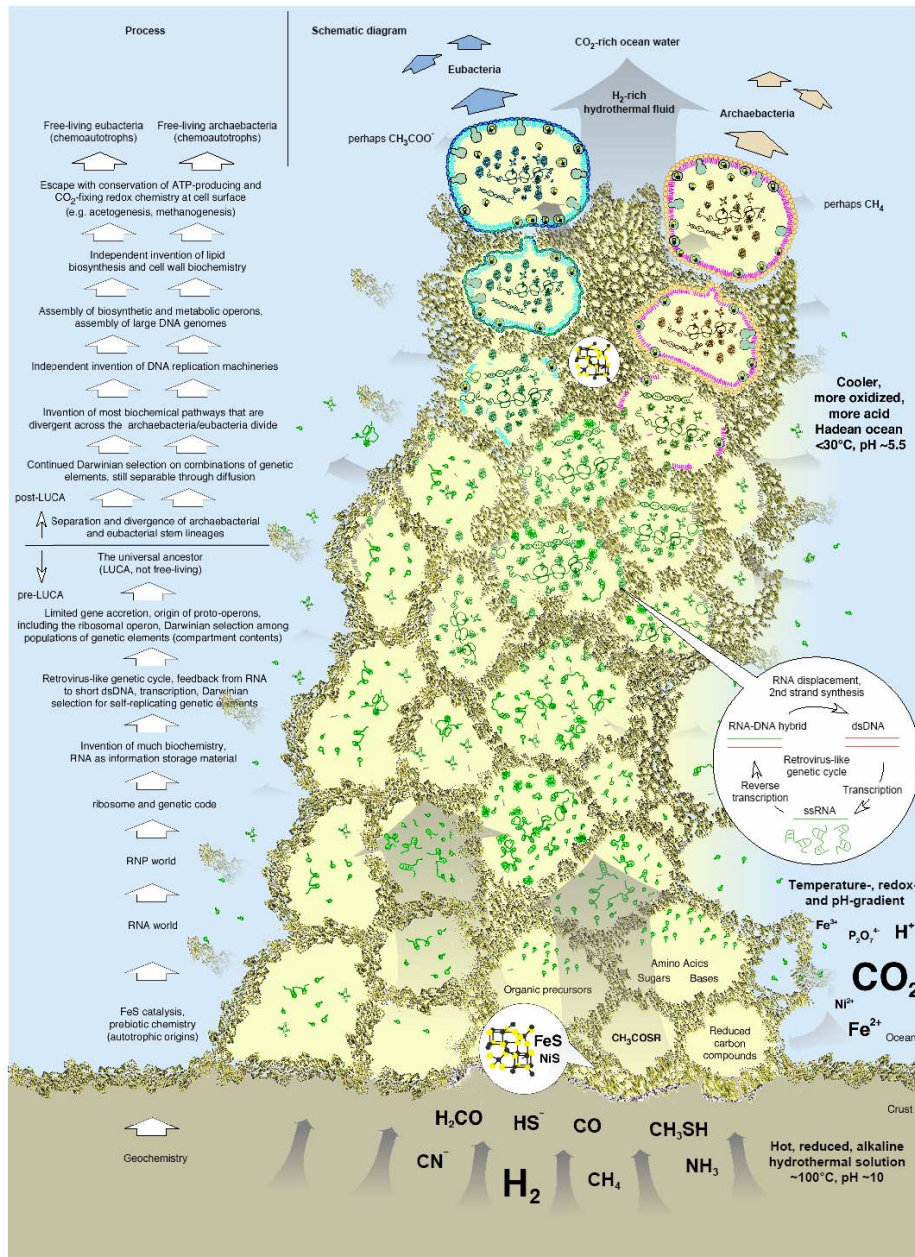
The crucible of Life?

Iron sulfides (FeS – ancient or Fe_2S_3 -modern); NiS
3D compartments, not just 2D surfaces



Martin, Russell, 2003,
Phil. Trans. Roy. Soc.

Early evolution from monomers to first cells



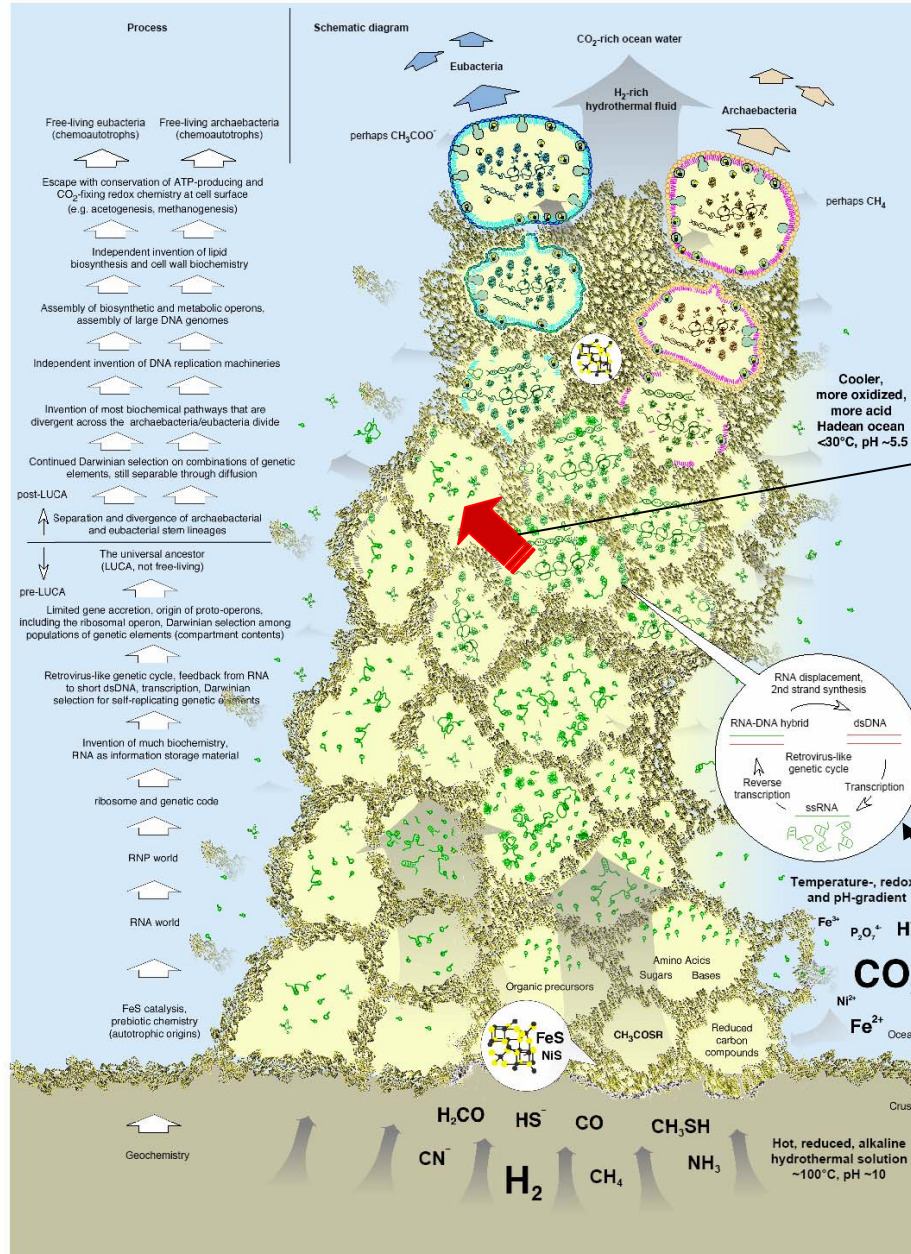
Russell-Martin chimney

- The Martin-Russell chimney – a system of hydrothermal compartments - is a **perfect home** for a multipartite, virus-like genome of LUCA postulated by Leipe et al.

[Koonin EV](#), [Martin W](#). On the origin of genomes and cells within inorganic compartments.

Trends Genet. 2005 Dec;21(12):647-54

Early evolution from monomers to first cells



Russell-Martin chimney

Transfer of genetic content between compartments (infection)

LUCA

Salient features of early evolution: first ensembles of genetic elements – LUCA - archaeal and bacterial cells

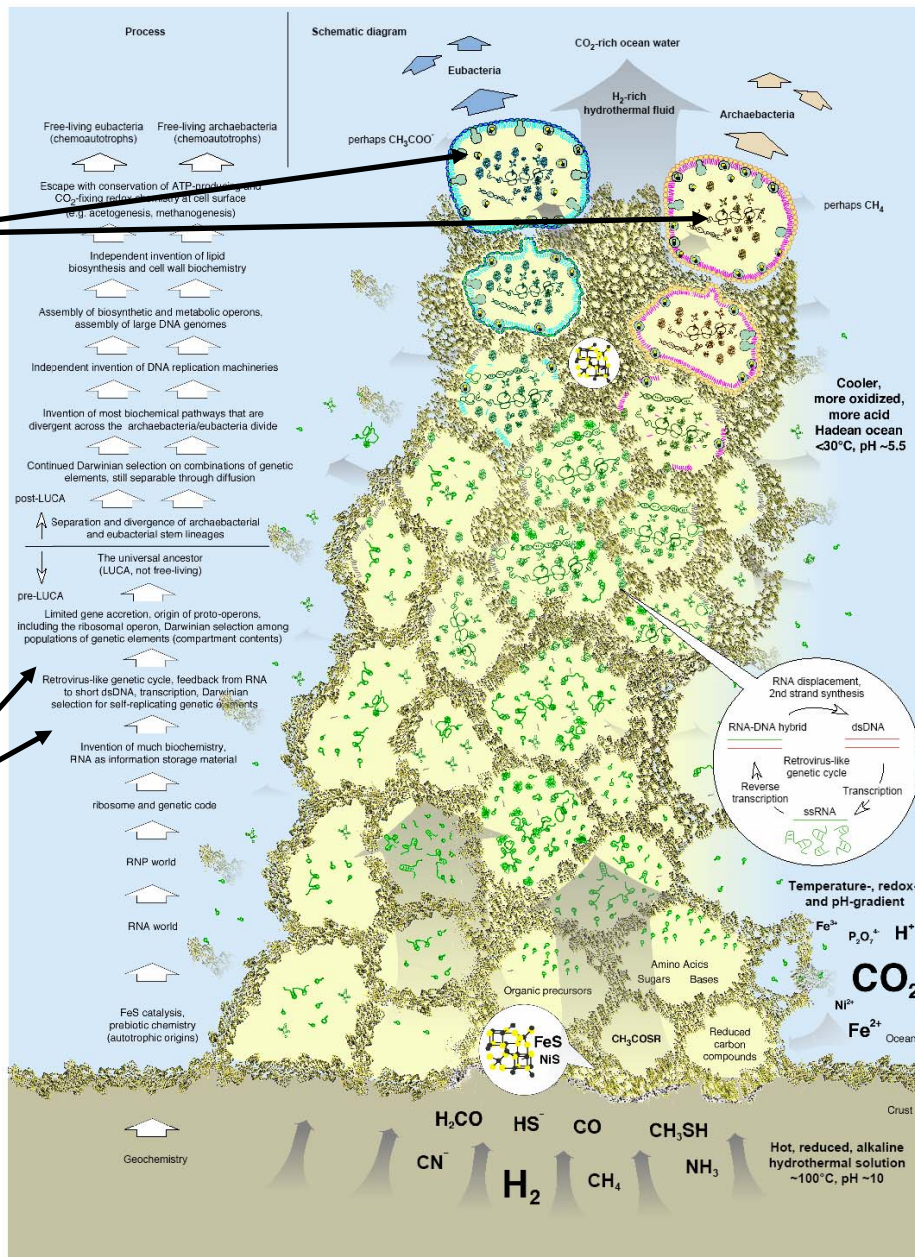
- LUCA: community of genetic elements inhabiting the Russell-Martin mound – more a *state* than an organism - **LUCAS**
- RNA replication+retrovirus-like cycle
- **Selfish cooperatives: ensembles of virus-like genetic elements** coding not only for replication but also for accessory functions (including translation, precursor synthesis etc)
- Advantage of physical connection between functionally coupled genes – origin of (super)operons from selfish cooperators – driving force for subsequent transition to DNA genomes
- No cell division but **effective selection for compartment contents** via “infection” of neighboring compartments – “horizontal transfer” is an intrinsic feature of LUCA’s lifestyle
- Independent invention of at least two machineries for membrane biogenesis and DNA replication
- Probably, many escape attempts but only two successful ones: archaea and bacteria

Viruses coming out of the chimney...

Archaea and bacteria escape with their own complements of viruses

↑
Early virus world evolves

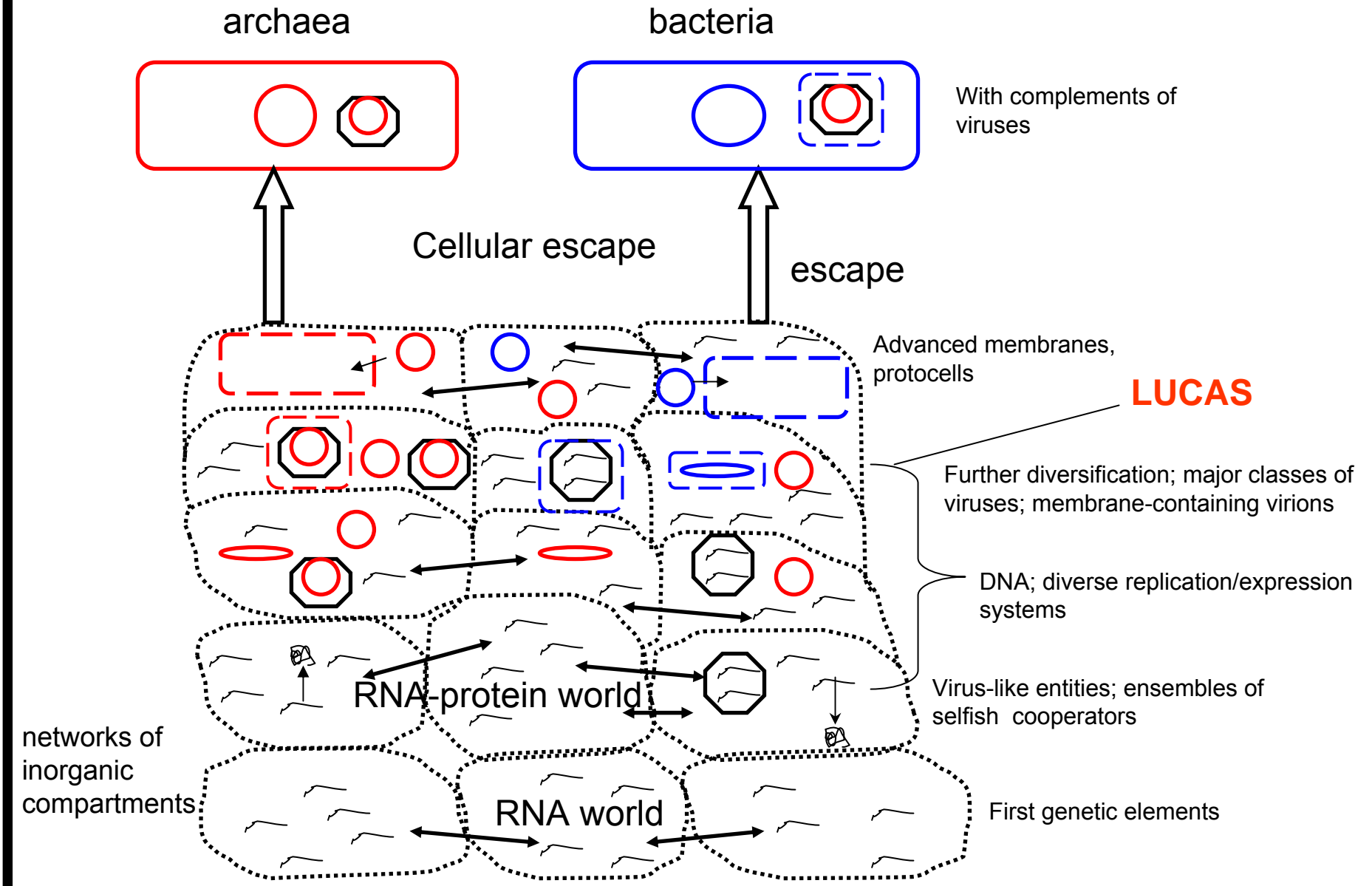
Persistence of purely selfish replicons parasitizing on selfish cooperatives



Russell-Martin chimney

Origin of cells and viruses: sketch of a synthesis

time



The Virus World

- There is a small pool of essential genes involved in viral replication and structure and shared by extremely diverse groups of viruses - **virus hallmark genes**
- ...and much larger gene pools shared by related groups of viruses, e.g., tailed phages
- All major classes of viruses emerged within the primordial gene pool (pre/post-LUCA/S)
- The Virus World has existed as a continuous gene flow since the pre-LUCA/S times
- Under this concept, **viruses and cells co-evolved, ancestors of viral and cellular genomes were indistinguishable at very early stages of evolution, and virus-like agents antedate fully-fledged cells**
- **Unification of the cellular and viral branches of life's evolution** at the earliest stages analogous to the unification of physical interactions in the first seconds after the Big Bang

Koonin, Dolja, 2006, Virus Res.

Koonin, Senkevich, Dolja, 2006, Biology Direct

On the origin of different types of cells

- Substantial genetic and metabolic complexity at a pre-cellular, ***virus-like*** stage – where LUCA/S (***not a prokaryote!***), probably, belongs
- Independent origin of two types of membrane organization, escape of two types of cells (archaea and bacteria) capturing two independently evolved DNA replication systems from the pre-cellular pool, probably, amidst many failed attempts
- Origin of eukaryotic cell through endosymbiotic collision of the two types of prokaryotic cellular organization; origin of the nucleus and some other major eukaryotic novelties as defense against the Invasion of Group II introns, the progenitors of spliceosomal introns

Koonin, Martin, Trends Genet. 2005 Dec;21(12):647-54

Martin, Koonin, Nature. 2006 Mar 2;440(7080):41-5.

Koonin, Biol Direct. 2006 Aug 14;1(1):22

Koonin, Senkevich, Dolja, 2006, Biology Direct

There is little new under the sun...

"...life may have remained in the virus stage for many millions of years before a suitable assemblage of elementary units was brought together in the first cell"

Haldane JBS: **The Origin of Life.**
Rationalist Annual 1928, **148**:3-10.

Acknowledgments

Bill Martin (U. Duesseldorf, Germany)

**Tania Senkevich
(NIAID, NIH)**

**Valerian Dolja
(Oregon State)**



Journal Welcome to *Biology Direct*

An online journal published by [BioMed Central](#)

Biology Direct considers original research articles, hypotheses and reviews in selected subject areas, and will eventually cover the full spectrum of biology. Subject areas already launched include Genomics, Bioinformatics and Systems Biology, and Immunology.

***Biology Direct* aims to provide authors and readers of research articles with a novel system of peer review.** This will include making the author responsible for obtaining reviewers' reports, via the journal's Editorial Board; making the peer review process open rather than anonymous; and publishing the reviewers' reports along with the articles, thus increasing both the responsibility and the reward of the referees and eliminating sources of abuse in the refereeing process. See '[about *Biology Direct*](#)' and the [Biology Direct instructions for authors](#) for further information.

[Submit your article](#) online

Research **Open Access**

Retrospective analysis of protein kinase C-beta (PKC-β) expression in lymphoid malignancies and its association with survival in diffuse large B-cell lymphomas

Shuyu Li, Mark Phong, Michael Lahn, Leslie Brail, Susan Sutton, Boris K Lin, Donald Thornton, Birong Liao

Biology Direct 2007, 2:8 (21 February 2007)

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#) [\[PubMed\]](#) [\[Related articles\]](#)

Research **Open Access**

Conserved intron positions in ancient protein modules

Albert DG de Roos

Biology Direct 2007, 2:7 (8 February 2007)

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#) [\[PubMed\]](#) [\[Related articles\]](#)

EDIT M
CLICK HE

Search

- [Biology](#)
- [BioMed](#)
- [Current Trials](#)
- [PubMed](#)
- [PubMed](#)

SUBM
MANU
CLICK HE

Sign
articl

Related BioM
gateways:
[Bioinformatic](#)
[Gateway](#)
[Systems Bio](#)



[research](#) | [comment](#) | [editorial](#) | [hypothesis](#) | [opinion](#) | [review](#) | [top 10 most accessed articles](#)

[edit my deta](#)

t for Top
articles of

Top 10 most accessed articles for [last 30 days](#) / [past year](#) / all time

1. Research [Open Access](#) [Highly accessed](#)

Accesses **Origin and evolution of the peroxisomal proteome**
 6120 Toni Gabaldón, Berend Snel, Frank van Zimmeren, Wieger Hemrika, Henk Tabak, Martijn A Huynen
 evaluated on *Biology Direct* 2006, 1:8 (23 March 2006)
[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#) [\[PubMed\]](#) [\[Related articles\]](#) [\[F1000 Biology\]](#) [\[Cited on BioMed Central\]](#)

2. Hypothesis [Open Access](#) [Highly accessed](#)

Accesses **The ancient Virus World and evolution of cells**
 6094 Eugene V Koonin, Tatiana G Senkevich, Valerian V Dolja
 evaluated on *Biology Direct* 2006, 1:29 (19 September 2006)
[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#) [\[PubMed\]](#) [\[Related articles\]](#) [\[F1000 Biology\]](#) [\[Cited on BioMed Central\]](#)

3. Research [Open Access](#) [Highly accessed](#)

Accesses **A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action**
 5899 Kira S Makarova, Nick V Grishin, Svetlana A Shabalina, Yuri I Wolf, Eugene V Koonin
 evaluated on *Biology Direct* 2006, 1:7 (16 March 2006)
[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#) [\[PubMed\]](#) [\[Related articles\]](#) [\[F1000 Biology\]](#) [\[Cited on BioMed Central\]](#)

4. Research [Open Access](#) [Highly accessed](#)

Accesses **Rooting the tree of life by transition analyses**
 4397 Thomas Cavalier-Smith
Biology Direct 2006, 1:19 (11 July 2006)
[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#) [\[PubMed\]](#) [\[Related articles\]](#) [\[Cited on BioMed Central\]](#)

Search

- [Biology](#)
- [BioMed](#)
- [Current Trials](#)
- [PubMed](#)
- [PubMed](#)