Cold Spring Harbor Laboratory Meeting "Telomeres & Telomerases" 2-6 May 2007

TELOMERES IN EVOLUTION AND DEVELOPMENT FROM BIOSEMIOTIC PERSPECTIVE

Guenther Witzany
Telos – Philosophische Praxis
Vogelsangstrasse 18c, A-5111-Buermoos, Austria
Tel./Fax: ++43/6274 6805
Mail: witzany@sbg.at

Abstract:

Telomeres identify natural chromosome ends being different from broken DNA through differences in their "molecular syntax" (M.Eigen) which determines the functions of reverse transcriptase and its integrated RNA template, telomerase. Although telomeres play a crucial role in the linear chromosome organisation of eukaryotic cells, their molecular syntax descended from an ancient retroviral competence. This is an indicator for the early retroviral colonization of large double stranded DNA viruses, which are putative ancestors of the eukaryotic nucleus.

This talk will demonstrate certain advantages of the biosemiotic approach towards our evolutionary understanding of telomeres: focus on the genetic/genomic structures as language-like text which follows combinatorial (syntactic), context-sensitive (pragmatic) and content-specific (semantic) semiotic rules. Genetic/genomic organisation from the biosemiotic perspective is not seen any longer as an object of randomly derived alterations (mutations) but as functional innovation coherent with the broad variety of natural genome editing competences of viruses.

Introduction:

- For decades, non-coding, repetitive DNA sequences were interpreted as being ancient remnants of former evolutionary stages.
- It is now recognized that these non-coding regions of the genome are decisive for higher-order regulatory and constitutional functions of protein structural vocabulary.
- Repetitive Elements in eukaryotic genome organisation descended from RNA- and retroviral infection events in eukaryotic nucleus in a non-lytic but persistent status.
- Telomerase is a reverse transcriptase. Reverse transcriptases are key features in repetitive elements with higher-order regulatory functions.
- Telomeres are repetitive elements witch protect only linear DNA-chromosome-ends of eukaryotes against repair enzymes and genetic parasites.

Biosemiotics (Bios=Life; Semeion= Sign)

Biosemiotics investigates

a) communication processes within and among cells, tissues, organs, organisms as sign-mediated interactions
b) nucleotide sequence order as codes which follow syntactic, pragmatic and semantic rules

ad b) Biosemiotics investigates genetic sequences as texts which are coherent with laws of physics and chemics, but additionaly follow

- combinatorial (syntactic),
- context-specific (pragmatic)
- content-sensitive (semantic) rules.

Although this rules are likely conservative, in difference to the laws of physics and chemics they may be generated de novo, changed, combined and recombined.

1. Evolutionary Roles of Viruses

Precellular RNA gen-pool

RNA viruses, Retroviruses, DNA Viruses: Network of solely chemical and semiotic connected molecules. Several genes that are central for viral replication are missing from cellular genomes although phylogenetic analyses show that they are older than cellular elements.

Overlapping arrays of unrelated viruses enshure key functions in genome replication: capsid protein, helicase superfamily in all RNA,DNA viruses.

All RNA viruses share RNA dependend RNA polymerase and reverse transcriptase which indicates a RNA virus dependend function essential for eukaryotic replication, keeping in mind, that the eukaryotic nucleus is an large DNA virus.

Capsid proteins involved in jelly roll capsid protein may be a starting event in building true viruses. Alternative capsid proteins with helical capsid features might be a parallel development.

Several origins of viruses – several unrelated cellular key compenents

Membrane lipids, cell walls , as many other features are unrelated in bacteria and archaea. Complex colonization of unrelated viral descents into the large DNA virus which is the ancestor of the eukaryotic nucleus forced a digital/symbolic molecular grammar in the eukaryotic genome. Through this new grammatical competence it was possible to create diverse new features of eukaryotic cellular organization and coordination which the prokaryote world is lacking.

Evolution and Development

- All key processes of combinant and recombinant RNA- and DNA- Processing derived from viral or retroviral competences which either compete or cooperate symbiotically
- Evolutionary novelty and development occured by competent genetic/genomic text-editing in all major steps and substeps. Chance mutations didn't play an important role

Biology: Symbiology

Competitive and symbiotic viral colonization strategies determined the history of bacteria evolution and development by colonizing any unicellular organism (phage evolution). I ml sea water contains 1 Million bacteria but 10 times from viral agents.

Lytic versus Persistent Viral Life-Strategies

- Acute viruses that exhibit lytic action induce disease and even death.
- Persistent life-style of viruses implies compatible interactions with the host, either by being integrated into the hosting genome or within the cell plasma, and act non-destructive during most life stages of the host.
- The persistent life-style allows the virus to transmit complex viral phenotypes to the hosting organism. Doing so enables the host to broaden evolutive potentials that may well lead to the formation of new species

Addiction Modules for Persistent Status

- Most of the genetic/genomic text editing competences which are inherent to cells, bacteria, protozoa, plants, animals, fungi are a complementary mix of former antagonistic viral features.
- We can identify them even today as toxin/antitoxin-, restriction/modification-, insertion/deletion – modules.
- As symbiotic neutralization and counterpart regulation they represent new phenotypic features which may consist up to 100 new genes.
- The feature of one competence is regulated exactly by the antagonist according to developmental stages in cell-cycle, replication, tissue growth or similar contexts.
- Is this suppressor-function out of balance the normally downregulated part may become lytic or disease causing.

Gene Functions of Eukaryotes acquired from Persistent Viruses

- Immunity (restriction and modification modules, toxic and antitoxic modules);
- Silencing functions/micro-RNAs, (methylation, suppression);
- Recognition functions (replicate expression, receptors, expression factors);
- Immune regulation (signal mediating, heredity, adaptation).

Endogenous Retroviral Competences

- Endogenious retroviral competences in the persistent status are often characterized by features which are expressed only in a strictly time window of the developmental process, such as e.g., axis formation, trophectoplast formation, s-phase of the cell cycle.
- In this highly spezialized contexts they are replicated through signaling which blocks the suppression of the replication process. After the function is fullfilled, a signal initiates suppressor function again.(see “Important Role of Addiction Modules)

Descent of Retroelements

All Retroelements with its (i) higher order regulatory functions, (ii) genetic creativity potential and (iii) innovation competence of new regulatory patterns and combinations descended from retroviruses which can be easely identified in their three essential parts *gag*, *pol* and *env*.

Examples of Viral Inventions

- RNA, DNA
- linear chromosomes
- replicase, polymerase, integrase
- Natural DNA repair techniques
- restriction / modification
- methylation/histone modification
- bilayer nuclear envelope
- eukaryotic nucleus
- division of transcription and translation
- nuclear pores
- tubulin-based chromosome duplication
- chitin, calcification
- innate immune system (MHC-Komplex, RNAi)
- adaptive immune system
- cartilage, bones
- skin, dermal glands for poison, mucus and milk
- larvae, egg, placenta, flowering plants
- viviparous mammals

Deep Grammar and Superficial Grammar

- Higher order regulations which are performed by non-coding RNAs and are inherent in all repeat elements like e.g. retroposons have a similar relationship to proteincoding sequences as (1) **deep grammar** and (2) **superficial grammar** of utterances.
- Through this two different levels it is possible to determine the proteincoding data-sets according to different needs into multiple protein meanings.
- In Eukaryotic Genome Evolution the step from analog sequence order to digital sequence order occurred. This symbiogenetically induced invention opened the possibility to use proteincoding data-set by various types of higher order regulation

Two independent replication processes

All eukaryotic proteins involved in DNA replication differ from those found in prokaryotes. Hence, nuclear properties of eukaryotes are completely different from those of prokaryotes. These differences include:

- use of linear chromosomes, with repetitive termination points and several origins for replication,
- transcription and translation separated by multiple membranes,
- existence of complex nuclear pore structures that actively mediate RNA translocation,
- a tubulin-system that enables separation of duplicated chromosomes.

2. Eukaryotic Key Features Not Present in Prokaryotes

- Eukaryotes have linear DNA-chromosomes with telomere repeats on the ends
- Eukaryotic genomes share great variety of repeat elements with higher order regulatory functions
- Eukaryotic replication proteins have very different amino acid sequence compositions from prokaryotes.
- Eukaryotes share the control of DNA packaging and replication. Prokaryotes doesn't have chromatin proteins like histones.
- Eukaryotic DNA replication starts in numerous (thousands) of sites and is regulated by a complex cell cycle regulatory system
- Replication control proteins in eukaryotes doesn't have similarity to prokaryotic ones
- Daughter cells segregate by attachment to microtubule system (spindles) not by attachment at the membrane
- Eukaryotic nucleus posses three classes of DNA dependend RNA polymerases that lacks similarity to polymerases of any prokaryote
- In Eukaryotes the products of RNA polymerases must undergo posttranscriptional modifications (splicing) before they can function in the cytoplasm as mRNA, tRNA, rRNA. In no prokaryote splicing of pre-mRNAs is found.
- To prevent mistranslation of mRNA or unspliced tRNA the nucleus has to separate transcription/processing of RNA from the cytoplasm transport of processed RNAs. RNA processing evolved after eukaryot nucleus, because in no prokaryote splicing of pre-mRNAs is found. Therefore a nuclear membrane is needed to allow the evolution of introns within coding reagions.
- Introns allow splicing:
    - Group I introns(self-splicing) are mobile elements which code for DNA transposase
    - Group II introns code for reverse transcriptase
    - Small RNAs recognize the splice junctions and splice RNA after capping
- All 3 intron types (group I, II, small RNAs) are not existent in prokaryotes (but in viruses of prokaryotes)
- All complex modifications of mRNA und nuclear RNA are acquired during Evolution of the Eukaryotic Nucleus and are highly conserved in Eukaryotes but absent from prokaryotes
- The nuclear membrane is distinct from plasma membrane and is dissolved after S-phase and reformed at late anaphase/telophase.
- Highly conserved mitotic spindle system is not found in any prokaryote.

3. A Viral Progenitor of the Eukaryotic Nucleus

The eukaryotic cell evolved by a symbiogenetic integration of former free living bacteria. The eukaryotic nucleus doesn't have a prokaryotic progenitor. He resembles a lot of key features, proteins and RNAs which are not found in any prokaryote, but in some prokaryote viruses: These Viruses use

- linear chromosomes
- telomere repeats
- multiple membranes
- histone packaged chromosomes /with marking effect for self – non-self identification)
- nuclear pores

No single virus resembles all of these key features, but every key feature of the eukaryotic nucleus is present in a large dsDNA-Virus:

A) Prokaryotic Phages

• Cyanophages	dsDNA, DNA+RNA Polymerase similar to Eukaryotes
• Eubacterial Phages	linear dsDNA, Telomeres, DNA+RNA Polymerases, Chromatin, internal membranes
• Archaeal Phages	linear dsDNA, Telomere repeats similar to Eukaryotes but very dissimilar to Prokaryotes; Chromatin, internal lipid tendency to non-lytic, persistent and mixed infections. ASF-1: Eukaryotic like TATA sequences

B) Further Candidates

• Vaccinia Virus (Poxvirus)	membrane-bound division of transcription and translation; multiple membranes; DNA synthesis combines membrane loss and restoration (cell-cycle dependend); Actin/Tubulin bound transport system, nuclear pores
• Cytoplasmic DNA-Virus (ASFV)	Chromatin, linear chromosomes with telomeres
• Phycto DNA- Virus	mRNA capping, Introns, diverse DNA replication proteins
• TTV-1/4	linear dsDNA genomes with molecular basis for the evolution of eukaryotic chromatin. Capsid which integrates internal and external lipid proteins

Viral self and non-self identification competence

All Viruses mark their genomes, RNAs and Proteins by different kinds of chemical modifications e.g. methylation. This marking allows the differentiation between self and non-self. Non-self may be other viruses ort he host genome or host related transcripts.

Self-Re-animation of Viruses

By mixing defective viral genomes it is possible to recombine a full function virus again. Their higly conserved repair and recombination capacity enbables them as the only one living beings which can reanimate as living being after their death

4. Telomerase and other Reverse Transcriptases

Reverse Transcriptases derived from RNA dependend RNA Polymerases

Reverse transcriptases play key roles in mobile elements like transposons and retroposons. Reverse transcriptases play key roles in altering genomic structures and therefore play an important role in evolutionary processes. Reverse transcriptases generate

- copies of mRNAs which they need for integration into a genome;
- copies of non mRNAs like snoRNAs which are as DNA copies SINEs
- SINEs can initiate new genes which code for small RNAs with regulatory competences on existing genes

Reverse Transcriptase is a Primer for Retroposons with Important Regulatory Functions

LTRs	*copia, gypsy, Ty1, IAPs, HERVs*
Non-LTRs	*act as telomeres in several arthropods: Het-A/TART*
SINEs	
LINEs	
ORF1	(RNA-binding protein)
ORF2	(endonuclease, reverse transcriptase activities)
ALUs	(manipulation of LINE 1 function for mobilization)
Group II self splicing introns	
snoRNAs	(Type 1-3 Retroposons)

Reverse Transcriptases are found also in

Other retroviruses	(mammals, birds)
Hepadnavirus	(mammals, birds)
Caulimovirus	(plants)
LTR-Retroposons	(animals, plants fungi, protazoa)
non-LTR Retroposons	(animals, plants fungi, protazoa)
group II introns	(bacteria, fungi, plant mitochondria, chloroplasts, plastids)
mitochondrial plasmids	(Neurospora mitochondria)
RTL Gene	(Chalimydomonas mitochondria)
multiple ssDNAs	

Telomerase Function has Alternatives:

Not only Telomerase reverse Transcriptase (TERT) replicate Telomere-repeats but:

- Protein Priming, terminal hairpins and recombination which allow complete replication of (i) viral linear DNA, (ii) bacterial plasmid genomes and (iii) linear mitochondrial genomes of certain eukaryotes.
- Homologous recombination between telomeric sequences instead of telomerases replicate telomeric repeats in some insects and plants.

Telomerase Function is Cell-Cycle Regulated

Telomerase functions exclusively if its suppression is deleted. Is the Telomerase function in Telomere-replication fulfilled a signal initiates its suppression again. If this signalling process is disturbed uncontrolled cell replication may occur. This indicates telomerase being part of an addiction module (see part (1): Addiction Modules).

Open reading frames (ORFs) code for reverse transcriptase

RNA dependent DNA polymerase (reverse transcriptase) has relations to RNA dependent RNA polymerase. Many organisms have ORFs which code for proteins that have very similar sequences as retroviral reverse transcriptases (Xion and Eickbush 1990). If we root these lines of descendend in RNA dependend RNA polymerases we find 2 groups:

1) group 1 contains: LTR retroposons, RNA Viruses, DNA Viruses
2) group 2 contains: non-LTR-retroposotons, bacterial and other organelle parts (Nakamura and Cech 1998)

Telomerase: A Natural Genetic Engineering Tool with Different Functions in Different Contexts

Whereas reverse transcriptase has been used in RNA-Virus life cycle for replication functions, it is as acquired tool (transported and integrated in a symbiogenetic event) for complete replication of chromosomal ends in linear eukaryotic genomes.
In Eukaryotes Telomerases are endogenous retroviral genome editing agents.

5. Telomere-Function is a Retroviral Competence

Telomeres are Repetitive Elements

Telomeres are highly conserved non-mobile repetitive DNA-sequences. We know various repetitive elements which are highly mobile elements all of retroviral origin

Telomeres are Nucleoprotein Structures

which protect the ends of chromosomes from erosion, degradation, colonization, or sticking together chromosome ends. They are necessary only in linear chromosomes not in circular

Telomere Repeats are Building Nodes

which stabilizes Telomeres and *are not of linear DNA*. These nodes also care for not being recognized as DNA damage which would induce a DNA repair mechanism. Is the node intact this serves as signal for the cell that she is fit for further replications.

Different Molecular Syntax of Telomere Sequences

Vertebrates Human, mouse, Xenopus	TTAGGG
Filamentous fungii Neurospora crassa	TTAGGG
Slime moulds Physarum, Didymium	TTAGGG
Tetrahymena, Glaucoma	TTGGGG
Paramecium	TTGGG(T/G)
Oxytricha, Stylonychia, Euplotes	TTTTGGGG
Apicomplexan protozoa Plasmodium	TTAGGG(T/C)
Arabidopsis thaliana	TTTAGGG
Green algae Chlamydomonas	TTTTAGGG
Insects Bombyx mori	TTAGG
Roundworms Ascaris lumbricoides	TTAGGC
Fission yeasts Schizosaccharomyces pombe	TTAC(A)(C)G(1-8)
Saccharomyces cerevisiae	TGTGGGTGTGGTG (from RNA template)
Candida glabrata	GGGGTCTGGGTGCTG
Candida albicans	GGTGTACGGATGTCTAACTTCTT
Candida tropicalis	GGTGTA[C/A]GGATGTCACGATCATT
Candida maltosa	GGTGTACGGATGCAGACTCGCTT
Candida guillermondii	GGTGTAC
Candida pseudotropicalis	GGTGTACGGATTTGATTAGTTATGT
Kluyveromyces lactis	GGTGTACGGATTTGATTAGGTATGT

Telomeres: Functions and evolution

- Telomeres act as immune functions against proteincoding enzymes with high recombination or degradation competences or viral genetic parasites and function similar to an RNAi system.
- Telomeres serve as recognition sequences, primer functions and genetic/genomic raw material for sequence generation (genome duplication, RNA template)

Telomere-function is a retroviral competence

In certain arthropods telomeres are built by non-LTR-retroposons HeT-A and TART which have retroviral origins. This may be an indicator, that telomere generation in common is a retroviral competence.

Conclusion:

- Natural genetic/genome-editing competences like telomerases and telomere-functions represent original skills of viruses.
- It seems likely that telomere repeats had a similar immune function in linear chromosomes of DNA viruses prior to the evolution of eukaryotes. Because repetitive elements are key features of Retroviruses, there could be an infection event of large DNA viruses by RNA viruses which attained a persistent status in the linear chromosome of the DNA virus. This protects linear chromosome ends against competing genetic parasites.
- The acquisition of the telomere repeats in eukaryotes has been a key event in eukaryotic nucleus evolution. The eukaryotic nucleus evolved from a large DNA virus. However, the digital structure of the eukaryotic genome, with its typical repetitive (higher-order regulatory) elements, indicates high rates of persistent, non-lytic retroviral infections.
- In these infection events the eukaryotic host acquired a discrete/digital genomic syntax which is the precondition for multiple protein meanings from the same genetic data-set through post-transcriptional modifications such as alternative splicing pathways.
- Therefore transforming of the analog (prokaryotic) molecular syntax in symbolic (eukaryotic) molecular syntax is a major step in evolution of multicellular complexity.

References

Haoudi A, Mason JM. (2000). Reverse transcriptase can stabilize or destabilize the genome. Genome 43: 949-956.

Nakamura TM, Cech TR. (1998). Reversing Time: Origin of Telomerase. Cell 92:587-590.

Ryan FP. (2007). Viruses as symbionts. Symbiosis 44:11-21.

Tomlinson RL, Ziegler TD., Supakorndej T, Terns RM, Terns MP. (2006). Cell Cycle-regulated Trafficking of Human Telomerase to Telomeres. Molecular Biology of the Cell 17: 955-965.

Villarreal LP. (2005). Viruses and the Evolution of Life. Washington, ASM Press.

Witzany G. (2006). Natural Genome Editing Competences of Viruses. Acta Biotheoretica 54:235-253.

Xiong Y., Eickbush TH. (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. The EMBO Journal 9(10):3353-3362.