Biosemiotics
DOI 10.1007/s12304-008-9018-0

The Viral Origins of Telomeres and Telomerases and their Important Role in Eukaryogenesis and Genome Maintenance

Guenther Witzany

Received: 27 November 2007 / Revised: 3 January 2008 / Accepted: 15 January 2008

Abstract Whereas telomeres protect terminal ends of linear chromosomes, telomerases identify natural chromosome ends, which differ from broken DNA and replicate telomeres. Although telomeres play a crucial role in the linear chromosome organization of eukaryotic cells, their molecular syntax most probably descended from an ancient retroviral competence. This indicates an early retroviral colonization of large double-stranded DNA viruses, which are putative ancestors of the eukaryotic nucleus. This contribution demonstrates an advantage of the biosemiotic approach towards our evolutionary understanding of telomeres, telomerases, other reverse transcriptases and mobile elements. Their role in genetic/genomic content organization and maintenance is no longer viewed as an object of randomly derived alterations (mutations) but as a highly sophisticated hierarchy of regulatory networks organized and coordinated by natural genome-editing competences of viruses.

Keywords Telomeres · Telomerases · Eukaryotic nucleus · Persistent viruses

Introduction

Biosemiotics investigates rule-governed, sign-mediated interactions both within and among cells, tissues, organs and organisms. It also investigates genetic sequences as codes/texts that are coherent with the laws of physics and chemistry but, in addition, follow a complementary mix of combinatorial (syntactic), context-sensitive (pragmatic), content-specific (semantic) rules (Witzany 2000, 2007). In this respect,

G. Witzany (✉)
Telos–Philosophische Praxis, Vogelsangstrasse 18c, 5111 Bürmoos, Austria
e-mail: witzany@sbg.at

the roles of telomeres and telomerases in evolution, structure and content arrangement of genomes are of particular interest. This involves deciphering the relationships between the 'molecular syntax' (Eigen and Winkler 1975) of telomere repeats and their meaning, i.e. their function in the genomic content. This requires their evolutionary roots to be examined. The telomere replication process by telomerase is the most important feature here because it is processed by a very ancient competence (Nosek et al. 2006), i.e. reverse transcriptase with a great variety of functions in most key processes of living nature (Eickbush 1997).

Upon close examination the specific characteristics of telomeres reveals certain features common to all genomes that possess telomeres: telomeres are highly conserved, non-mobile, repetitive DNA sequences. Telomeres are nucleoprotein structures that protect the ends of chromosomes from erosion, degradation, colonization, or adhering chromosome ends (Blasco 2007). They are necessary only in linear chromosomes, not in circular ones. Telomere repeats are building nodes. These nodes stabilize telomeres and *do* not consist of linear DNA. These nodes also prevent recognition as DNA damage, which would induce a DNA repair pathway. Intact nodes serve as a signal for the cell that it is fit for further replication. Telomeres are thought to be the forerunners of centromeres which probably derived from an ancestral telomere–telomere fusion (Ijdo et al. 1991). Similar to telomeres, centromeres are highly conserved, non-mobile, repetitive DNA sequences. They interact with spindle microtubules and are therefore crucial for distributing chromosomes to offspring cells (Villasante et al. 2007). They also encode small RNAs which are responsible for heterochromatin formation (Couzin 2002; Grewal and Elgin 2007).

Linear chromosomes of eukaryotes have the so-called end replication problem: DNA polymerases which replicate leading strands of double-stranded DNA only in the 5′ to 3′ direction are unable to replicate lagging strands, i.e. in the 3′ to 5′ direction. For leading strand replication, DNA polymerases add polynucleotides to an RNA primer. These RNA strands are later replaced by DNA. At the terminal end of the chromosome, the RNA primer cannot be replaced completely by DNA, so it cannot code for proteins or further replications. When the last RNA is added, DNA polymerase and DNA ligase transform the RNA of the primer to DNA. This process requires the presence of another DNA strand in front of the RNA primer. The end replication problem is the lack of another DNA strand in front of the last attached RNA primer. That RNA is degraded by enzymes. Thus, a section of telomeres would be lost during each replication cycle to replicate a completely lagging strand, another technique is necessary. A reverse transcriptase known as telomerase uses its integrated subunit, an inherent RNA template, to replicate the overhanging RNA primer. This allows the terminal end of the lagging strand to be fully completed without loss (Haoudi and Mason 2000).

Telomere function needs a certain length of base pairs. If this length is not available due to continued end replication problems or damage, then chromosome ends are unprotected (Du and Traktman 1996). This has prompted the suggestion that continuous telomere shortening is a main reason for cell ageing. However, recent research has documented that this is the case only in rare situations, not in general (Laun et al. 2007).

Differences in the Molecular Syntax of Telomere Sequences

From the biosemiotic perspective it would be of interest to determine whether the telomere sequences differ between various organisms, species and kingdoms. The lack of a difference would indicate that the telomere repeat function depends on strict sequence order whereas differences would indicate that the specific function of telomere repeats is of primary importance, not the sequence order that encodes this function.

Interestingly, the molecular syntax of telomere repeats differs in those organisms in which it has been identified. This indicates that no unique molecular syntax is necessary to guarantee the function that telomere repeats have to fulfil. Rather, the same important function can be coded by different nucleic acid sequences. For instance, we find: TTAGGG in vertebrates, humans, mice, *Xenopus*, filamentous fungi, *Neurospora crassa*, the slime moulds *Physarum* and *Didymium*: TTGGGG in *Tetrahymena* and *Glaucoma*; TTGGG(T/G) in *Paramecium*; TTTTGGGG in *Oxytricha, Stylonychia* and *Euplotes*; TTAGGG(T/C) in the apicomplexan protozoan *Plasmodium*; TTTAGGG in *Arabidopsis thaliana*; TTTTAGGG in green algae *Chlamydomonas*; TTAGG in the insect *Bombyx mori*; TTAGGC in the roundworm *Ascaris lumbricoides*; TTAC(A)(C)G(1–8) in the fission yeast *Schizosaccharomyces pombe*; TGTGGGTGTGGTG (from RNA template) in *Saccharomyces cerevisiae*; GGGGTCTGGGTGCTG in *Candida glabrata*; and GGTGTACGGATGTCTAA CTTCTT in *Candida albicans*.

Telomeres act as immune functions against genomic agents with high recombination or degradation competences, i.e. viral genetic parasites, and seems to function similar to an RNAi system. RNAi protects the genome against genomic parasites, i.e. viruses, by silencing genomic transcripts of exogenous infective RNA viruses or endogenous transposons or retroposons (Fire et al. 1998; Fire 2005; Couzin 2002). In addition, telomeres serve as recognition sequences, primer functions and genetic/genomic raw material for sequence generation (genome duplication, RNA template).

In *Drosophila* and some plants, telomere elongation during replication does not occur by telomerase but through recombination facilitated by the non-LTR retroposons HetA and TART (Nakamura and Cech 1998; Fajkus et al. 2005; Blasco 2007). They transport their *gag* protein into the nucleus to produce more copies of the chromosome ends (Rashkova et al. 2002). These retroposons, which fulfil the same function of telomere elongation as telomerase, are regulated by the same epigenetic regulations that govern mobile element activity, including RNAi (Savitsky et al. 2006; Slotkin and Martienssen 2007).

Telomere Replication in Most Cases by Telomerase, a Reverse Transcriptase

In most cases, except that described above, telomeres are replicated by telomerase, a reverse transcriptase. This indicates that the function of telomeres in the eukaryotic replication cycle is very ancient (Curcio and Belfort 2007). Some authors have suggested that reverse transcriptases derived from RNA-dependent RNA polymerases which themselves derived from an ancient RNA world (Boeke 2003).

Telomerase is a ribonucleoprotein enzyme that is an assembly of telomerase RNA and telomerase reverse transcriptase (Jady et al. 2004). Telomerase is clearly related to mobile elements, especially to the non-LTR retroposons (Eickbush 1999).

Reverse Transcriptases and Mobile Elements

Mobile elements in the genome may be transposons that integrate directly into a host genome, or retroposons that integrate via an RNA intermediate, reverse transcriptase. Copying from RNA into DNA generally involves reverse transcriptases. Mobile elements are important for genotype processing, with far-reaching consequences for phenotype expression during its various developmental stages. Recent research has demonstrated that overlapping epigenetic marking in eukaryotic cells is an important evolutionary feature to silence the expression of mobility of these mobile elements (Slotkin and Martienssen 2007). Mobile elements can silence single genes as well as larger chromosomal regions and, therefore, play an important role in the evolution of diversity. They share their competence to recombine, rearrange and insert into genomic content with other retroelements (Coffin et al. 1997). They influence neighbouring genes through alternative splicing and are active agents as enhancers and promotors or act by polyadenylation patterns (Slotkin and Martienssen 2007).

Reverse transcriptases play key roles in mobile elements like transposons and retroposons. One type of retroposon has direct repeats at its ends (LTR), others do not (non-LTRs). Interestingly, the number of retroposons increases with every transposition (transposition duplication) so that they can expand genomes: LINE-1 is 20% of the human genome (Maita et al. 2004). In contrast, transposons contain a code for the transposase protein. This enzyme identifies the terminal inverted repeats which flank mobile elements, excises them and *integrates itself instead* of those excised. The gap at the donor site is repaired in a cut-and-paste transposition or filled up with a copy of the transposon by a gap repair technique (Slotkin and Martienssen 2007). Transposons can also *integrate themselves* in phages and plasmids, and are transferred with them into other cells (Frost et al. 2005). This is evidence for a self/non-self differentiation competence.

In contrast to non-mobile telomeres and centromeres, mobile sequences such as transposons and retroposons (Volff 2006) and non-coding repetitive elements such as LTRs, SINEs and LINEs enable far-reaching DNA rearrangement and reorganization (Shapiro 2002; Sternberg 2002; Shapiro and Sternberg 2005). Together, they play a decisive role in the evolution of new genomic structures (Shabalina and Spiridonov 2004; Shapiro and Sternberg 2005; Sternberg and Shapiro 2005). Interestingly, the non-coding DNA also contains the regulations of transcription, promoter, enhancer and suppressor (Bird et al. 2006). The repetitive sequences are highly species specific and are more suitable for determining species than the coding sequences (Villarreal 2005).

Major Roles of Reverse Transcriptases in Natural Genome Editing

In addition, reverse transcriptases play key roles in altering genomic structures and, therefore, in evolutionary processes facilitated by natural genome editing (Witzany 2006). Reverse transcriptases are used to generate (a) copies of mRNAs which they

need for integration into a genome and (b) copies of non-mRNAs such as small nucleolar RNAs, one of the largest classes of non-coding RNAs (Zemann et al. 2006) which, like DNA copies, are SINEs. SINEs can initiate new genes which code for small RNAs with regulatory competences on existing genes.

One further key feature of reverse transcriptases is that they are a primer for retroposons such as LTRs (copia, gypsy, Ty1, IAPs, HERVs). Non-LTRs (Het-A/TART, SINEs, LINEs) act like telomerases in several arthropods and plants. Moreover, reverse transcriptases are encoded and used by open reading frames (ORF), ORF1 (an RNA-binding and shuttling protein), ORF2 (endonuclease, reverse transcriptase activities), as well as ALUs (manipulation of LINE-1 function for mobilization), group II self-splicing introns and snoRNAs (type 1–3 retroposons), all of which act as important regulatory functions (Yang et al. 1999; Batzer and Deininger 2002; Tomlinson et al. 2006; Weber 2006; Matera et al. 2007).

Reverse transcriptases are also found in retroviruses of mammals and birds, in the hepadnavirus of mammals and birds, and the caulimovirus of plants, in LTR retroposons of animals, plants, fungi and protozoa, in non-LTR retroposons of animals, plants, fungi and in protozoa, group II introns of bacteria, fungi, plant mitochondria, chloroplasts and plastids, in mitochondrial plasmids of *Neurospora* mitochondria, and in multiple single-stranded DNAs (Villarreal 2005).

RNA-dependent RNA polymerases together with reverse transcriptases replicate positive-strand RNA viruses, double-stranded RNA viruses, negative-strand RNA viruses and retroviruses (Koonin et al. 2006). RNA-dependent RNA polymerases are involved in the coupling of heterochromatin for the production of siRNAs (Sugiyama et al. 2005). The RNAi system is competent in post-transcriptional gene silencing and is, therefore, a crucial instrument in keeping the balance between the need for expression and the need for silencing (Grewal and Elgin 2007). SiRNAs therefore act similar to endogenously encoded microRNAs (Doench et al. 2003).

Many organisms have ORFs that code for proteins with sequences very similar to retroviral reverse transcriptases (Xiong and Eickbush 1990; Mesnard and Lebeurier 1991). RNA-dependent DNA polymerase (reverse transcriptase) has relations to RNA-dependent RNA polymerase. Rooting these lines of descent in RNA-dependent RNA polymerases yields two groups: (1) group 1 contains LTR retroposons, RNA viruses, DNA viruses; (2) group 2 contains non-LTR retroposons, bacterial and other organelle parts (Nakamura and Cech 1998).

The telomerase function is cell cycle regulated. It functions exclusively if its suppression is deleted. Once the telomerase function in telomere replication is fulfilled, a signal initiates its suppression again. A disturbed signalling process may lead to uncontrolled cell replication. Telomerase has to be transported to telomere repeats for its elongation during the S phase of the cell cycle. The delivery agents are Cajal bodies—small nucleolus-like organelles competent in (1) splicing, (2) ribosome production and (3) transcription (Platani et al. 2002; Jady et al. 2004). They are located in the periphery of nucleoli (Darzacq et al. 2002; Matera 2006). Cajal bodies move throughout the area of the nucleus and, for certain properties, they fuse with other Cajal bodies or associate with nucleoli (Tomlinson et al. 2006; Kiss et al. 2002). Telomerase trafficking is restricted to the S phase of the cell cycle, which avoids telomerase activity at non-telomeric sites of the chromosomes (Tomlinson et al. 2006).

Eukaryotic Key Features Absent in Prokaryotes

Because telomere repeats and telomerases are key features of eukaryotes, and not of prokaryotes, it may be concluded that eukaryotic telomeres and telomerases are interconnected with the evolution of the eukaryotic cells. Deciphering the evolutionary roots of telomeres and telomerases necessitates the main differences between eukaryotes and prokaryotes to be examined. The evolutionary agents of the eukaryotic nucleus may even point to the roots of telomeres and telomerases.

First, eukaryotic genomes share a great variety of repeat elements with higher-order regulatory functions. In contrast to prokaryotes, eukaryotic replication proteins have very different amino acid sequence compositions. In addition, eukaryotes share the control of DNA packaging and replication, whereas prokaryotes do not have chromatin proteins such as histones (Villarreal 2005).

The eukaryotic DNA replication starts in numerous (thousands) sites and is regulated by a complex cell cycle regulatory system. Eukaryotic replication control proteins do not resemble prokaryotic ones. A further difference between eukaryotes and prokaryotes is that daughter cells segregate by attachment to a microtubule system (spindles), not by attachment at the membrane. The highly conserved mitotic spindle system is not found in any prokaryote (Cottingham and Hoyt 1997).

Also, the eukaryotic nucleus possesses three classes of DNA-dependent RNA polymerases that do not resemble the polymerases of any prokaryote. A further crucial difference is that in eukaryotes the products of RNA polymerases must undergo post-transcriptional modifications (splicing) before they can function in the cytoplasm as mRNA, tRNA or rRNA. No prokaryote exhibits splicing of pre-mRNAs. To prevent mistranslation of mRNA or unspliced tRNA, the nucleus has to separate transcription/processing of mRNA from the cytoplasm transport of processed RNAs. This requires a nuclear membrane to segregate transcription, mRNA processing, transport and translation in the cytoplasm (Vale 2003). The nuclear membrane is distinct from the cell membrane and is dissolved after the S phase, but is restored at late anaphase/telophase. All complex modifications of mRNA and nuclear RNA seem to be acquired during the evolution of the eukaryotic nucleus; they are highly conserved in eukaryotes but absent in prokaryotes.

Only a very few prokaryotic genomes share some of the above-mentioned features. In the case of the spirochetes *Borrelia*, the genomes possess three types of telomeres, segmented genomes of linear and circular plasmids and extensive DNA rearrangements (Chaconas 2005; Tourand et al. 2006). This could indicate intensive infection by competing genetic parasites which are in balance as ‘addiction modules’ (see below) in a persistent status. This does not harm the host but is harmful to those organisms (even close relatives) that lack these persistent inhabitants.

A Viral Progenitor of the Eukaryotic Nucleus?

The eukaryotic cell most probably evolved by a symbiogenetic integration event of former free-living bacteria. This integration, however, cannot explain the progenitor of the eukaryotic nucleus because its key features could not derive from prokaryotes (Bell 2001, 2006). The eukaryotic nucleus resembles numerous key features,

proteins and RNAs described above which are not found in any prokaryote. Interestingly, these key features are present in certain prokaryote viruses (Villarreal 2005; Forterre 2006a, b). These viruses use linear chromosomes, telomere repeats, multiple membranes, histone-packaged chromosomes with marking effect for self/non-self identification and nuclear pores.

No single virus encompasses all of these key features, but every key feature of the eukaryotic nucleus is present in some large dsDNA viruses. This requires consideration of a process in which different viral competences have been integrated into a single dsDNA virus that is the progenitor of the eukaryotic nucleus. Alternatively, the large dsDNA virus functioned as a eukaryotic nucleus and later integrated different viral competences. (Competences are capabilities which may be used in a special situational context but need not necessarily be used). On examination of the key features of several candidates for this integration, the focus is primarily on prokaryotic, eukaryotic and archaeal phages.

Prokaryotic phages such as cyanophages have double-stranded DNA, DNA polymerases and RNA polymerases similar to eukaryotes. Eubacterial phages possess linear double-stranded DNA, telomeres, DNA polymerases, RNA polymerases, chromatin and internal membranes. Archaeal phages with linear double-stranded DNA have telomere repeats similar to eukaryotes. They also possess chromatin and an internal lipid tendency to non-lytic, persistent (and often mixed) infections (Villarreal 2005).

Other DNA viruses share similar features which are characteristic for the eukaryotic nucleus but are not found in prokaryotes. An example is the vaccinia virus (poxvirus) (Takemura 2001). These viruses have a membrane-bound segregation of transcription and translation, multiple membranes, and their DNA synthesis combines membrane loss and a cell cycle-dependent restoration as well as an actin/tubulin-bound transport system (Villarreal 2005; Van Lent and Schmitt-Keichinger 2006) and, interestingly, nuclear pores. Cytoplasmic DNA viruses (African swine fever virus) have chromatin and linear chromosomes with telomeres. Phyto DNA viruses have mRNA capping, introns and diverse DNA replication proteins. TTV (1–4) have linear double-stranded DNA genomes with a molecular basis for the evolution of eukaryotic chromatin; they also have capsids which integrate internal and external lipid proteins.

In addition, all these viruses have the competence for self and non-self identification. 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, the host genome or host-related transcripts (Villarreal 2005).

Evolutionary Roles of Viruses as Natural Genome Editors

To understand the evolutionary emergence of the eukaryotic nucleus with its key features such as telomeres and telomerases in the eukaryotic replication process, it could be useful to reconstruct the natural genome-editing competences of viruses (Witzany 2006). Recent research in microbiology, based on comparative genomics and phylogenetic analyses, has demonstrated that life must be viewed from the

perspective of the crucial role played by viruses (Forterre 2001, 2002, 2005, 2006a, b; Koonin 2006; Villarreal 2005; Tran et al. 2004).

This contradicts former concepts which focused on viruses in the framework of (1) escape theories, i.e. viruses are intact or deformed genetic parasites which escaped from cellular life, or considered that viruses (2) evolved from cellular ancestors or (3) that they are not living beings because they cannot live without cellular life. From these perspectives, viruses could not play crucial roles in the evolution of cellular life.

Interestingly, phylogenetic analyses do not support the former concept of RNA and DNA viruses descending from cellular life. These analyses also show that DNA viruses and RNA viruses most probably did not have a common ancestor but evolved independently. Viruses probably have to be placed at the very beginning of life, long before cellular life evolved (Villarreal 2005).

Pre-cellular Life

Recent research suggests thinking about the early stages of life as a pre-cellular RNA gene pool with RNA viruses, retroviruses and—by reverse transcriptase of single-stranded RNA viral genomes—also double-stranded DNA viruses (Leipe et al. 1999; Martin 2005; Koonin et al. 2006; Brosius 2003; Flavell 1995). Prior to cellular life forms, we can imagine networks of solely chemically connected molecules coherent to the molecular syntax of RNA and, later on, DNA. Several genes central to viral replication are missing from cellular genomes, and phylogenetic analyses show that they are older than cellular elements. Overlapping arrays of unrelated viruses ensure key functions in genome replication such as capsid proteins and the helicase superfamily found in all RNA and DNA viruses (Koonin 2006).

All RNA viruses share RNA-dependent RNA polymerase and reverse transcriptase. This indicates that an RNA virus-dependent function is essential for eukaryotic replication (Temin 1985).

Membrane lipids, cell walls and many other features are unrelated in bacteria and archaea. Complex colonization by (unrelated) viral descents into the large DNA virus, which is the ancestor of the eukaryotic nucleus, forced the emergence of a digital molecular syntax in the eukaryotic genome. The emergence of higher-order regulations on a given protein-coding data set is analogous to a limited repertoire of signs of an alphabet ready for use by the unlimited potential multiple regulatory combinations. The recent findings about multiple functions of non-coding RNAs, especially in plant development (Rodríguez-Alvarado and Roossinck 1997) but also their high number and abundant functions in nervous systems demonstrate this important role in understanding neuronal systems in general (St. Laurent and Wahlestedt 2007). This new grammatical competence made it possible to generate diverse and new, complex features of eukaryotic cellular organization and coordination. These features are lacking in the prokaryotic world and depend on competent agents that use this digital molecular syntax.

Persistent Viral Life Strategies are Beneficial for their Hosts

Acute viruses that exhibit lytic action induce disease and even death. In contrast, a persistent lifestyle of viruses implies compatible interactions with the host, either by

being integrated into the hosting genome (Gorinsek et al. 2004) or within the cell plasma. The result is non-destructive symbiosis during most life stages of the host. The persistent lifestyle allows the virus to transmit complex viral phenotypes to the hosting organism. This process enables the host to broaden its evolutive, adaptational potential and may promote the formation of new species (Villarreal 2005).

Persistent lifestyles of viruses are typically tissue specific, i.e. host tissues are colonized by different non-lytic viruses which integrate themselves into the host genomes and co-evolve with them. During host cell replication they function in a tissue-specific, replication cycle-dependent manner. Interestingly, micro-RNAs in eukaryotic cells have similar tissue-specific or developmental expression patterns (Mattick 2007). Micro-RNAs play important roles in Dicer- and Risc-mediated mRNA degradation or mRNA translation inhibition (Bartel 2004). This implies an RNAi immune function. Because micro-RNAs act on mRNAs, not on proteins, they are probably encoded by persistent nuclear DNA viruses (Cullen 2006).

Persistent Status Through Addiction Modules

The persistent status emerges through multiple colonization events into a host. This neutralizes former antagonistic and incompatible features of competing viral agents without harming the host (Ryan 2004, 2006, 2007). Most of the endogenous genetic/genomic inhabitants inherent to cells, bacteria, protozoa, plants, animals and fungi are a complementary mix of formerly antagonistic viral features. They can still be identified today as toxin/antitoxin, restriction/modification, insertion/deletion modules (Villarreal 2005; Gerdes 2000; Pandey and Gerdes 2005; Makarova et al. 2006). As symbiotic neutralization and counterpart regulation, they represent new phenotypic features. The feature of one competence is regulated exactly by the antagonist according to developmental stages in the cell cycle, replication, tissue growth, or similar contexts. Should this suppressor function become unbalanced, then the normally downregulated part may become lytic with even lethal consequences as documented for *Symbiodinium* and its major role in coral bleaching (Witzany 2007).

The gene functions of eukaryotes acquired from persistent viruses include immunity (restriction and modification modules, toxic and anti-toxic modules), silencing functions/micro-RNAs (methylation, suppression), recognition functions (replicate expression, receptors, expression factors) and immune regulation (signal mediating, heredity, adaptation) (Villarreal 2005).

Endogenous Retroviral Competences are a Persistent Symbiotic Lifestyle

Endogenous retroviral competences in the persistent status are often characterized by features expressed only in the strict time window of a developmental process, such as axis formation, trophectoplast formation, or S phase of the cell cycle. In these highly specialized contexts they are replicated through signalling, which blocks the suppression of the replication process. After the function is fulfilled, a signal once again initiates suppressor function. Retroelements—with their (1) higher-order regulatory functions, (2) capability for genetic creativity and (3) innovation competence of new regulatory patterns and combinations—are descended from retroviruses which can be easily identified by their three essential parts *gag*, *pol* and *env* (Rashkova et al.

2002; Weiss 2006; Tang et al. 1999). Most endogenous retroviruses have been degraded into formerly connected domains, but they can still be recognized by one of these three genes (Gao et al. 2003; Sfakianos and Hunter 2003; Ryan 2004; Gabus et al. 2006). The *gag* gene encodes structural proteins, *pol* encodes enzymes such as reverse transcriptase and integrase functions, and *env* encodes envelope proteins.

'Elements', 'Entities', 'Parasites'—Agents of Natural Genome Editing

Recent research shows extensive dynamic DNA remodelling by small RNAs and micro-RNAs, which are competent in a great variety of DNA arrangements, rearrangements and recombinations (Shapiro 2002; Vaughn and Martienssen 2005; Mattick 2001, 2006). Some authors refer to agents of genomic creativity (Ryan 2006), mobile or regulatory elements (Eickbush 1999; Brosius 1999) or entities (Daubin and Ochman 2004), while others refer to transposable elements (Slotkin and Martienssen 2007), non-coding RNA populations (Mattick 2007) and still others to mobile DNA species or genetic parasites (Nakamura and Cech 1998; Villarreal 2005). Together, these agents enable complex organisms to integrate several temporal steps and a great variety of coordinated signalling processes in eukaryotic cell replication, fix them in a conserved DNA storage medium and, if necessary, resolve conservation, change, rearrange or newly construct the whole genomic content and sequence order (Shapiro 2006).

The DNA information storage medium is and has to be edited. I predict a future discussion on how to refer to these editing agents, for example as interactions of more or less chemical molecules or as 'non-random genetic change operators' (Shapiro 2007, personal communication)

From a biosemiotic perspective—which investigates combinatorial (syntactic), content-specific (semantic) and contextual (pragmatic) rules of natural genome editing and genetic text processing—it is important to note that there can be no editing without a subject that edits, i.e. an editor or a swarm of editors (Vetsigian et al. 2006). For example, the spliceosome works as an integrated network of several small nuclear RNAs and their associated proteins on the primary RNA-transcript into the pre-mRNA (Vaughn and Martienssen 2005).

Life could not function without the key agents of DNA replication, namely mRNA, tRNA and rRNA. Not only rRNA, but also tRNA and the processing of the primary transcript into the pre-mRNA and the mature mRNA, are clearly descended from retroelements (Rao et al. 1989; Maizels and Weiner 1993; Maizels et al. 1999; Flavell 1995; Eickbush and Eickbush 2007)

It is now possible to appreciate how sophisticatedly the competent, subject-like operators act in the case of endogenous retroviruses, which reached a persistent and non-lytic lifestyle. We also know that all related retroelements share a common genome-editing competence like transposable 'elements'. Nonetheless, it remains difficult to reconstruct how all these DNA-encoded RNA agents reached persistent status in hundreds, thousands and tens of thousands of elements. We only know that they act in a precisely coordinated manner which would be impossible without competent signalling. This includes a strict competence for self/non-self identification, which is a major asset of RNAs in general and of small nucleolar RNAs in particular (Filipowicz 2000).

Persistent endogenous agents competent in both natural genetic engineering and natural genome editing apparently prefer a special kind of habitat characterized as non-coding DNA sectors. They use a syntax mainly consisting of repeats. They colonized analogous DNA genomes by inserting their sites between coding elements; then they use these coding elements for different needs. This developed to the point that, in the human genome, only 3% of coding regions remained. The remaining 97% serves as a habitat for persistent viral operators that orchestrate a highly sophisticated division of labour. From these genomic locations they can actively regulate close-coding sequences. Of special interest is the highly sophisticated production of mRNA with its cut-and-paste process in which non-coding elements, i.e. introns, are spliced out; the remaining exons which code for proteins are combined into a coherent protein-coding content ready for translation.

As opposed to persistent endogenous agents of natural genome editing in eukaryotes, we find persistent exogenous agents in prokaryotes that are competent in natural genome editing in the prokaryotic gene pool. This process has long been visualized as horizontal gene transfer and is now recognized as occurring by plasmids, phages and transposons, all with viral ancestors (Frost et al. 2005).

It is difficult to perceive mere molecules or molecule buildings as being 'competent' to process the sophisticated DNA language. It is less difficult to think of viruses being these subject-like agents.

Deep Grammar and Superficial Grammar of Eukaryotic Genome Content

Higher-order regulations which are performed by agents inherent in non-coding RNAs and in most repeat elements such as subtelomeric repeats and all the other retroelements have a similar relationship to protein-coding sequences as operators competent in using a (1) deep grammar with which they determine (2) the superficial grammar of sequence content. Through these two different levels it is possible to determine the protein-coding data sets, according to different needs, into "multiple protein meanings" (Ast 2005). Eukaryotic genome evolution involved the step from a continuous coding sequence order to an interrupted sequence order. Interestingly, the former is characteristic for circular prokaryotic genomes, the latter for linear genomes.

Biosemiotically, this symbiogenetically induced innovation of multiple-invaded coding data sets by retroelements opened up the possibility of using protein-coding data sets according to various types of higher-order regulation. The protein-coding data sets are the structural vocabulary, the non-protein-coding 'underworld' (Mattick 2006) of RNAs is the text-editing operators. This involved a massive invasion by non-coding introns (viruses) into the genomic habitats of protein-coding data sets (Rogozin et al. 2005; Mattick 2007). Thus, the molecular syntax of protein-coding data sets could be used for different requirements in different contexts (pragmatics) to serve for different genetic content arrangements (semantics). This could explain:

- that in evolutionary history certain genotypes from one species are transferred and integrated into the genomic content of other species to yield a new role or a new phenotypic feature in another context. This occurred with telomeres in the

linear chromosomes of ancient double-stranded DNA viruses (poxvirus, vaccinia virus, archaeal phages: AFV-1, SIRV-1, TTV 1–4), where they had other functions than in the eukaryotic genomic content (Villarreal 2005);
- the close coherence of protein-coding data sets between humans and chimpanzees (99%), keeping in mind that the percentage of protein coding in humans and chimpanzees is only 3%, whereas the percentage of non-coding DNA with higher-order regulatory functions is 97%, which determines different expression patterns (Witzany 2006);
- that specific cellular functions are encoded in a weakly conserved manner at the sequence level, in contrast to their preserved domains, for example the genes of nuclear pores (Bapteste et al. 2005);
- that telomeres themselves are not typical sites for colonization events, in contrast to sites very close to these telomeres. This is similar to the phylogenetically related centromeres. Because telomeres and centromeres themselves are relatively free of inverted repeats or retroelements, this could indicate an ancient immune (RNAi) function that protects both from massive invasions by genetic parasites.

Conclusion

The acquisition of telomere repeats in eukaryotes was a key event in eukaryotic nucleus evolution. The eukaryotic nucleus most probably evolved from a large DNA virus. The changing structure of the eukaryotic genome, however, with its coding and non-coding sections and its typical repetitive (higher-order regulatory) elements, indicates high rates of persistent, non-lytic viral infections. In contrast to most of these mobile, higher-order regulatory agents, telomere repeats (as well as centromeres) attained a non-mobile status.

Telomerase, from the biosemiotic perspective, is a natural genetic engineering tool with different functions in different contexts. Whereas in the RNA virus life cycle reverse transcriptase is used for replication functions, it serves as an acquired tool for complete replication of chromosomal ends in linear eukaryotic genomes. In eukaryotes, telomerases and other reverse transcriptases act as endogenous viral competences.

In these symbiogenetic infection events, the eukaryotic host acquired a higher-order regulated genomic syntax. This is the precondition for multiple protein meanings from the same genetic data set through post-transcriptional modifications such as alternative splicing pathways. The transformation of the continuous (prokaryotic) molecular syntax into a eukaryotic molecular syntax invaded by a great diversity of natural genome-editing agents is, therefore, a major step in the evolution of multicellular complexity.

Acknowledgements This work was first presented at the Cold Spring Harbor Laboratory Meeting on 'Telomeres and Telomerases', 3–6 May 2007. I would like to thank Cold Spring Harbor Laboratory for the invitation and participation support.

References

Ast, G. (2005). The alternative genome. *Scientific American*, *292*, 58–65.

Bapteste, E., Charlebois, R. L., MacLeod, D., & Brochier, C. (2005). The two tempos of nuclear pore complex evolution: highly adapting proteins in an ancient frozen structure. *Genome Biology*, *6*, R85.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, *116*, 281–297.

Batzer, M. A., & Deininger, D. L. (2002). ALU repeats and human genomic diversity. *Nature Reviews Genetics*, *3*, 370–380.

Bell, P. J. L. (2001). Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus. *Journal of Molecular Evolution*, *53*, 251–256.

Bell, P. J. L. (2006). Sex and the eukaryotic cell cycle is consistent with a viral ancestry for the eukaryotic nucleus. *Journal of Theoretical Biology*, *243*(1), 54–63. doi:10.1016/j.jtbi.2006.05.015.

Bird, C. P., Stranger, B. E., & Dermitzakis, E. T. (2006). Functional variation and evolution of non-coding DNA. *Current Opinion in Genetics & Development*, *16*, 559–564.

Blasco, M. (2007). The epigenetic regulation of mammalian telomeres. *Nature Reviews*, *8*, 299–309.

Boeke, J. D. (2003). The unusual phylogenetic distribution of retrotransposons: a hypothesis. *Genome Research*, *13*, 1975–1983.

Brosius, J. (1999). RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene*, *238*, 115–134.

Brosius, J. (2003). The contribution of RNAs and retroposition to evolutionary novelties. *Genetica*, *118*, 99–115.

Chaconas, G. (2005). Hairpin telomeres and genome plasticity in Borrelia: all mixed up in the end. *Molecular Microbiology*, *58*, 625–635.

Coffin, J. M., Hughes, A. H., & Varmus, H. E. (1997). *Retroviruses*. New York: Cold Spring Harbor Laboratory Press.

Cottingham, F. R., & Hoyt, M. A. (1997). Mitotic spindle positioning in saccharomyces cerevisiae is accomplished by antagonistically acting microtubule motor proteins. *Journal of Cell Biology*, *138*, 1041–1053.

Couzin, J. (2002). Small RNAs make big splash. *Science*, *298*, 2296–2297.

Cullen, B. R. (2006). Viruses and microRNAs. *Nature Genetics*, *38*, S25–S30.

Curcio, M. J., & Belfort, M. (2007). The beginning of the end: links between ancient retroelements and modern telomerases. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 9107–9108.

Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., & Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. *EMBO Journal*, *21*, 2746–2756.

Daubin, V., & Ochman, H. (2004). Start-up entities in the origin of new genes. *Current Opinion in Genetics & Development*, *14*, 616–619.

Doench, J. G., Christian, P., Petersen, C. P., & Sharp, P. A. (2003). siRNAs can function as miRNAs. *Genes & Development*, *17*, 438–442.

Du, S., & Traktman, P. (1996). Vaccinia virus DNA replication: two hundred base pairs of telomeric sequence confer optimal replication efficiency on minichromosome templates. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 9693–9698.

Eickbush, T. (1999). Mobile introns: retrohoming by complete reverse splicing. *Current Biology*, *9*, 11–14.

Eickbush, T. H. (1997). Telomerase and retrotransposons: which came first. *Science*, *277*, 911–912.

Eickbush, T. H., & Eickbush, D. G. (2007). Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics*, *175*, 477–485.

Eigen, M., & Winkler, R. (1975). *Das spiel—naturgesetze steuern den zufall*. München: Piper.

Fajkus, J., Sykorova, E., & Leitch, A. R. (2005). Telomeres in evolution and evolution of telomeres. *Chromosome Research*, *13*, 469–479.

Filipowicz, W. (2000). Imprinted expression of small nucleolar RNAs in brain: time for RNomics. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 14035–14037.

Fire, A. (2005). Nucleic acid structure and intracellular immunity: some recent ideas from the world of RNAi. *Quarterly Reviews of Biophysics*, *38*, 303–309.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*, 806–811.

Flavell, A. J. (1995). Retroelements, reverse transcriptase and evolution. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, *110*, 3–15.

Forterre, P. (2001). Genomics and early cellular evolution. The origin of the DNA world. *Comptes rendus de l'Académie des sciences. Série 3, Sciences de la vie*, *324*, 1067–1076.
Forterre, P. (2002). The origin of DNA genomes and DNA replication proteins. *Current Opinion in Microbiology*, *5*, 525–532.
Forterre, P. (2005). The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie*, *87*, 793–803.
Forterre, P. (2006a). The origin of viruses and their possible roles in major evolutionary transitions. *Virus Research*, *117*, 5–16.
Forterre, P. (2006b). Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 3669–3674.
Frost, L. S., Laplae, R., Summers, A. O., & Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology*, *3*, 722–732.
Gabus, C., Ivanyi-Nagy, R., Depollier, J., Bucheton, A., Pelisson, A., & Darlix, J. L. (2006). Characterization of a nucleocapsid-like region and of two distinct primer tRNA binding sites in the endogenous retrovirus Gypsy. *Nucleic Acids Research*, *34*, 5764–5777.
Gao, X., Havecker, E. R., Baranov, P. V., Atkins, J. F., & Voytas, D. F. (2003). Translational recoding signals between gag and pol in diverse LTR retrotransposons. *RNA*, *9*, 1422–1430.
Gerdes, K. (2000). Toxin–antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *Journal of Bacteriology*, *182*, 561–572.
Gorinsek, B., Gubensek, F., & Kordis, D. (2004). Evolutionary genomics of chromovirus in eukaryotes. *Molecular Biology and Evolution*, *21*, 781–798.
Grewal, S. I. S., & Elgin, S. C. R. (2007). Transcription and RNA interference in the formation of heterochromatin. *Nature*, *447*, 399–406.
Haoudi, A., & Mason, J. M. (2000). Reverse transcriptase can stabilize or destabilize the genome. *Genome*, *43*, 949–956.
Ijdo, J. W., Baldini, A., Ward, D. C., Reeders, S. T., & Wells, R. A. (1991). Origin of human chromosome 2: an ancestral telomere–telomere fusion. *Proceedings of the National Academy of Sciences of the United States of America*, *88*, 9051–9055.
Jady, B. E., Bertrand, E., & Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *The Journal of Cell Biology*, *164*, 647–652.
Kiss, A. M., Jady, B. E., Darzaq, X., Verheggen, C., Bertrand, E., & Kiss, T. (2002). A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic Acids Research*, *30*, 4643–4649.
Koonin, E. V. (2006). Temporal order of evolution of DNA replication system inferred by comparison of cellular and viral DNA polymerases. *Biology Direct*, *1*, 39. doi:10.1186/1745-6150-1-39.
Koonin, E. V., Senkevich, T. G., & Dolja, V. V. (2006). The ancient virus world and evolution of cells. *Biology Direct*, *1*, 29.
Laun, P., Bruschi, C. V., Dickinson, J. R., Rinnerthaler, M., Heeren, G., Schwimbersky, R., et al. (2007). Yeast mother cell-specific ageing, genetic (in)stability, and the somatic mutation theory of ageing. *Nucleic Acids Research*, *35*(22), 7514–7526. doi:10.1093/nar/gkm919.
Leipe, D. D., Aravind, L., & Koonin, E. V. (1999). Did DNA replication evolve twice independently. *Nucleic Acids Research*, *27*, 3389–3401.
Maita, N., Anzai, T., Aoyagi, H., Mizuno, H., & Fujiwara, H. (2004). Crystal structure of the endonuclease domain encoded by the telomere-specific long interspersed nuclear element, TRAS1. *Journal of Biological Chemistry*, *279*, 41067–41076.
Maizels, A., & Weiner, A. M. (1993). The genomic tag hypothesis: modern viruses as molecular fossils of ancient strategies for genomic replication. In R. F. Gesteland, & J. F. Atkins (Eds.), *The RNA world* (pp. 577–602, 2nd ed.). Cold Spring Harbor NY: Cold Spring Harbor Laboratory Press.
Maizels, N., Weiner, A. M., Yue, D., & Shi, P. (1999). New evidence for the genomic tag hypothesis: archaeal CCA-adding enzymes and tRNA substrates. *Biological Bulletin*, *196*, 331–334.
Makarova, K. S., Grishin, N. V., & Koonin, E. V. (2006). The HicAB cassette, a putative novel, RNA-targeting toxin–antitoxin system in archaea and bacteria. *Bioinformatics*, *22*, 2581–2584.
Martin, W. (2005). Archaebacteria (Archaea) and the origin of the eukaryotic nucleus. *Current Opinion in Microbiology*, *8*, 630–637.
Matera, A. G. (2006). *Drosophila* Cajal bodies: accessories not included. *The Journal of Cell Biology*, *172*, 791–793.
Matera, A. G., Terns, R. M., & Terns, M. P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nature Reviews Molecular Cell Biology*, *8*, 209–220.

Mattick, J. S. (2001). Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Reports*, *2*, 986–991.
Mattick, J. S. (2006). The underworld of RNA. *Nature Genetics*, *38*, 393.
Mattick, J. S. (2007). A new paradigm for developmental biology. *Journal of Experimental Biology*, *210*, 1526–1547.
Mesnard, J. M., & Lebeurier, G. (1991). How do viral reverse transcriptases recognize their RNA genome. *FEBS Letters*, *287*, 1–4.
Nakamura, T. M., & Cech, T. R. (1998). Reversing time: origin of telomerase. *Cell*, *92*, 587–590.
Nosek, J., Kosa, P., & Tomaska, L. (2006). On the origin of telomeres: a glimpse at the pre-telomerase world. *Bioessays*, *28*, 182–190.
Pandey, D. P., & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Research*, *33*, 966–976.
Platani, M., Goldberg, I., Lamond, A. I., & Swedlow, J. R. (2002). Cajal body dynamics and association with chromatin are ATP dependent. *Nature Cell Biology*, *4*, 502–508.
Rao, A. L. N., Dreher, T. W., Marsh, L. E., & Hall, T. C. (1989). Telomeric function of the tRNA-like structure of brome mosaic virus RNA. *Proceedings of the National Academy of Sciences of the United States of America*, *86*, 5335–5339.
Rashkova, S., Karam, S. E., Kellum, R., & Pardue, M. L. (2002). Gag proteins of the two *Drosophila* telomeric retrotransposons are targeted to chromosome ends. *The Journal of Cell Biology*, *159*, 397–402.
Rodríguez-Alvarado, G., & Roossinck, M. J. (1997). Structural analysis of a necrogenic strain of cucumber mosaic cucumovirus satellite RNA in planta. *Virology*, *236*, 155–166.
Rogozin, I. B., Sverdlov, A. V., Babenko, V. N., & Koonin, E. V. (2005). Analysis of evolution of exon–intron structure of eukaryotic genes. *Briefings in Bioinformatics*, *6*, 118–134.
Ryan, F. P. (2004). Human endogenous retroviruses in health and disease: a symbiotic perspective. *Journal of the Royal Society of Medicine*, *97*, 560–565.
Ryan, F. P. (2006). Genomic creativity and natural selection: a modern synthesis. *Biological Journal of the Linnean Society*, *88*, 655–672.
Ryan, F. P. (2007). Viruses as symbionts. *Symbiosis*, *44*, 11–21.
Savitsky, M., Kwon, D., Shpiz, S., Georgiev, P., Kalmykova, A., & Gvozdev, V. (2006). Telomere maintenance is under control of the RNAi-based mechanism in the *Drosophila* germline. *Genes & Development*, *20*, 345–354.
Sfakianos, J. N., & Hunter, E. (2003). M-PMV capsid transport is mediated by Env/Gag interactions at the pericentriolar recycling endosome. *Traffic*, *4*, 671–680.
Shabalina, S. A., & Spiridonov, N. A. (2004). The mammalian transcriptome and the function of non-coding DNA sequences. *Genome Biology*, *5*, 105e.
Shapiro, J. A. (2002). Genome organization and reorganization in evolution. *Annals of the New York Academy of Sciences*, *981*, 111–134.
Shapiro, J. A. (2006). Genome informatics: the role of DNA in cellular computations. *Biological Theory*, *1*, 288–301.
Shapiro, J. A., & Sternberg, R. (2005). Why repetitive DNA is essential to genome function. *Biological Reviews*, *80*, 1–24.
Slotkin, R. K., & Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics*, *8*, 272–285.
Sternberg, R. (2002). On the roles of repetitive DNA elements in the context of a unified genomic–epigenetic system. *Annals of the New York Academy of Sciences*, *981*, 154–188.
Sternberg, R., & Shapiro, J. A. (2005). How repeated retroelements format genome function. *Cytogenetic and Genome Research*, *110*, 108–116.
St. Laurent, G., & Wahlestedt, C. (2007). Noncoding RNAs: couplers of analog and digital information in nervous system function. *Trends in Neuroscience*, *30*(12), 612–621. doi:10.1016/j.tins.2007.10.002.
Sugiyama, T., Cam, H., Verdel, A., Moazed, D., & Grewal, S. I. S. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 151–157.
Takemura, M. (2001). Poxviruses and the origin of the eukaryotic nucleus. *Journal of Molecular Evolution*, *52*, 419–425.
Tang, Y., Winkler, U., Fredd, E. O., Torrey, T. A., Kim, W., Li, H., et al. (1999). Cellular motor protein KIF-4 associates with retroviral gag. *Journal of Virology*, *73*, 10508–10513.
Temin, H. M. (1985). Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons and retrotranscripts. *Molecular Biology and Evolution*, *2*, 455–468.

Tomlinson, R. L., Ziegler, T. D., Supakorndej, T., Terns, R. M., & Terns, M. P. (2006). Cell cycle-regulated trafficking of human telomerase to telomeres. *Molecular Biology of the Cell*, *17*, 955–965.
Tourand, Y., Bankhead, T., Wilson, S. L., Putteet-Driver, A. D., Barbour, A. G., Byram, R., et al. (2006). Differential telomere processing by borrelia telomere resolvases in vitro but not in vivo. *Journal of Bacteriology*, *188*, 7378–7386.
Tran, E., Brown, J., & Maxwell, E. S. (2004). Evolutionary origins of the RNA-guided nucleotide modification complexes: from the primitive translation apparatus. *Trends in Biochemical Sciences*, *29*, 343–350.
Vale, R. (2003). The molecular motor toolbox for intracellular transport. *Cell*, *112*, 467–480.
Van Lent, J. W. M., & Schmitt-Keichinger, C. (2006). Viral movement proteins induce tubule formation in plant and insect cells. In F. Baluska, D. Volmann, & P. Barlow (Eds.), *Cell–cell channels* (pp. 1–13). New York: Springer.
Vaughn, M. W., & Martienssen, R. (2005). It's a small RNA world, after all. *Science*, *309*, 1525–1526.
Vetsigian, K., Woese, C., & Goldenfeld, N. (2006). Collective evolution and the genetic code. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 10696–10701.
Villarreal, L. P. (2005). *Viruses and the evolution of life*. Washington: ASM.
Villasante, A., Abad, J. P., & Mendez-Lago, M. (2007). Centromeres were derived from telomeres during the evolution of the eukaryotic chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 10542–10547.
Volff, J. N. (2006). Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *BioEssays*, *28*, 913–922.
Weber, M. J. (2006). Mammalian small nucleolar RNAs are mobile genetic elements. *PLoS Genetics*, *2* (12), e205 (December).
Weiss, R. A. (2006). The discovery of endogenous retroviruses. *Retrovirology*, *3*, 67. doi:10.1186/1742-4690-3-67.
Witzany, G. (2000). *Life: the communicative structure*. Norderstedt: Libri Books on Demand.
Witzany, G. (2006). Natural genome-editing competences of viruses. *Acta Biotheoretica*, *54*, 235–253.
Witzany, G. (2007). *The logos of the bios 2. Bio-communication*. Helsinki: Umweb.
Xiong, Y., & Eickbush, T. H. (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. *The EMBO Journal*, *9*, 3353–3362.
Yang, J., Malik, H. S., & Eickbush, T. H. (1999). Identification of the endonuclease domain encoded by R2 and other site-specific, non-long terminal repeat retrotransposable elements. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 7847–7852.
Zemann, A., Beckke, A., Kiefmann, M., Brosius, J., & Schmitz, J. (2006). Evolution of small nucleolar RNAs in nematodes. *Nucleic Acids Research*, *34*, 2676–2685.