

The design of life: part 3— an introduction to variation-inducing genetic elements

Peer Terborg

The inheritance of traits is determined by genes: long stretches of DNA that are passed down from generation to generation. Usually, genes consist of a coding part and a non-coding regulatory part. The coding part of the gene determines the functional output, whereas the non-coding portion contains switches and units that determine when, where and how much of the functional output should be generated. Point-mutations in the coding part are predominantly neutral or slightly detrimental genetic noise that accumulates in the genome, whereas point-mutations in the regulatory part of DNA units can induce variation with respect to the amount of output. Previously, in part 2, I argued that created kinds were frontloaded with baranomes: that is, pluripotent genomes with an ability to induce variation from within. The output of (morpho)genetic algorithms present in the baranome can readily be modulated by variation-inducing genetic elements (VIGEs). VIGEs are frontloaded genetic elements normally referred to as endogenous retroviruses, insertion sequences, LINEs, SINEs, micro-satellites, transposons, insertion sequences, and the like. In the present report, these transposable and repetitive DNA sequences are redefined as VIGEs, which solves the RNA virus paradox. The (morpho)genetic algorithms were designed in such way that VIGEs easily integrated into it and became a part of it, hence making the program explicit.

In order to fight off invading bugs and parasites, higher organisms have an elaborate mechanism that induces variation in immunological defence systems. One particular type of immune cells (B cells) produces defence proteins known as immunoglobulins. Immunoglobulins are very sticky; they bind to intruders as biological tags and mark them as ‘alien’. Other cells of the immune system then recognize the intruder, and a destruction cascade is activated. To have a tag available for every possible alien intruder, millions of B cells have their own highly specific gene for immunoglobulin production. In the genome there is only limited storage space for biological information, so how can there be millions of genes? Well, there aren’t. Immunoglobulin genes are assembled from several pre-existing DNA sequences that can be independently put together. The part of the immunoglobulin that does the ‘alien’ recognition contains several domains which are each highly variable. Every single B cell forms a unique immunoglobulin gene by picking from several short pre-existing DNA sequences. We also observe that later generations of immunoglobulins are more specific than the earlier generations, in the sense that they bind more tightly to invading microorganisms. Binding affinity to an invader is equivalent to recognition of that invader. And the better the immune system is able to recognize an intruder, the better it is able to clear it. The increased specificity is due to somatic mutations deliberately introduced in the genes of the immunoglobulins. A mechanism to rapidly induce mutations in immunoglobulin genes is present in the B cell genome. This mechanism ensures that the recognition pattern specified by the genes becomes increasingly specific

for the intruder. This ability to recognize and defeat all potential microorganisms is characteristic of the immune systems of higher organisms, including humans. The genomes contain all the necessary biological information required to induce variation from within. A flexible genome is required to effectively ward off diseases and parasitic infections. B cells don’t wait for mutations to happen; they generate the necessary mutations themselves.

Darwin revisited

Previously, in part 2,¹ I argued that organisms are equipped with flexible, highly adaptable, pluripotent, multipurpose genomes. Organisms are able to conquer the world through *adaptive radiation of baranomes*. But how do baranomes unleash information? Do organisms have to wait for selectable mutations to occur in order to rapidly invade and occupy novel ecological niches? Or were the baranomes of created kinds equipped with mechanisms to rapidly induce mutations, similar to the variation generated by B cells? Let’s turn to Darwin’s *The Origin of Species*, where we will find some clues. Darwin wrote quite extensively on variation, and in particular on the variation of feather patterns in pigeons:

‘Some facts in regard to the colouring of pigeons well deserve consideration. The rock-pigeon is of a slaty-blue, and has a white rump (the Indian sub-species, *C. intermedia* of Strickland, having it bluish); the tail has a terminal dark bar, with the bases of the outer feathers externally edged with white; the wings have two black bars; some

semi-domestic breeds and some apparently truly wild breeds have, besides the two black bars, the wings chequered with black. These several marks do not occur together in any other species of the whole family. Now, in every one of the domestic breeds, taking thoroughly well-bred birds, all the above marks, even to the white edging of the outer tail-feathers, sometimes concur perfectly developed. Moreover, when two birds belonging to two distinct breeds are crossed, neither of which is blue or has any of the above specified marks, the mongrel offspring are very apt suddenly to acquire these characters; for instance, I crossed some uniformly white fantails with some uniformly black barb, and they produced mottled brown and black birds; these I again crossed together, and one grandchild of the pure white fantail and pure black barb was of as beautiful a blue colour, with the white rump, double black wing-bar, and barred and white-edged tail-feathers, as any wild rock pigeon! We can understand these facts, on the well-known principle of reversion to the ancestral characters, if all the domestic breeds have descended from the rock-pigeon.²

Darwin argues—and correctly so—that all domestic pigeon breeds have descended from the rock-pigeon. He even knew, as demonstrated above, how to breed the rock-pigeon from several distinct pigeon races following a breeding pattern. Darwin describes a *breeding algorithm* for pigeons, to obtain the ancestor to all pigeons! But does he also describe an algorithm for breeding turkeys from pigeons? No. Darwin doesn't know such an algorithm. If he had found an algorithm for breeding ducks or magpies from pigeon genomes, he would have had solid evidence in favour of his proposal *On The Origin of Species Through the Preservation of Favoured Races*. His breeding experiments led him to discover 'the principle of reversion to ancestral characters', but contrary to common Darwinian wisdom, it is also *the falsifying observation* to his proposal for the *origin of species*. The observation that pigeons bring forth pigeons, and nothing else but pigeons, is not exactly the evidence needed to argue for the common descent of all birds. On the contrary! Darwin's breeding experiments demonstrated that a pigeon is a pigeon is a pigeon. Characteristics and traits within single species of pigeons may vary tremendously, but he always started and ended with pigeons. Breeding experiments have always shown, without exception, that novel and distinct bird species do not arrive through artificial selection. Even Darwin argues that there is no doubt that all varieties of ducks and rabbits have descended from the common wild duck and rabbit.³ From the variation Darwin observed in wild and domesticated populations, it does not follow that rabbits and ducks have some hypothetical common ancestor in a fuzzy distant past. Darwin observed

inborn, innate variation that already existed in the genomes of the pigeons and it only had to be activated or expressed.

From the excerpt above, we may even get an impression of how it works. A genetic algorithm for making feathers (a feather program) is part of the pigeon's genome and is present in every single cell. The feather program is present in billions of pigeon cells, but it is NOT active in all those cells. Feathers are only formed when the program is activated. The feather program is silent in cells where it should normally not operate. Activation of the feather program in the wrong cells may often be incompatible with life, but sometimes it may produce pigeons with (reversed) feathers on the feet. The program may be derepressed or activated through a mechanism that operates in the pigeon's genome. Whether feathers appear on the feet or on the head, and whether they appear normal or reversed is merely a matter of activation and regulation of the feather program. But Darwin didn't know about silent genomic programs or how they could become active. He didn't know about gene regulation and molecular switches. Darwin did not know anything about genes and genomes.

Analogous variation

The idea that Darwin had been working on for over two decades prior to the publication of *Origin*, his *idée fixe*, was how organic change (i.e. variation) present in populations might explain how novel species came into being. Unchanging, stable species is not what Darwin had in mind. He pondered the riddles of variation; he thought about laws and principles associated with the process of variation and believed he could disclose them by the study of the formation of new breeds. Drawing from what he knew about pigeon breeding and equine varieties, Darwin describes some of his ideas about the 'laws of variation' in chapter five of *Origin*:

'Distinct species present analogous variations; and a variety of one species often assumes some of the characters of an allied species, or reverts to some of the characters of an early progenitor. These propositions will be most readily understood by looking to our domestic races. The most distinct breeds of pigeons, in countries most widely apart, present sub-varieties with reversed feathers on the head and feathers on the feet, characters not possessed by the aboriginal rock-pigeon; these then are analogous variations in two or more distinct races.'⁴

Darwin describes that the exact same traits can appear in distinct breeds of pigeons and—importantly—these traits appeared *independently* in 'countries most widely apart'. If several breeds arrive with the same characteristics independently, it is unlikely they do so because of chance. Rather, the pigeon genomes may activate or derepress the same feather program independently. The effect is that distinct breeds 'in countries most widely apart'

acquire the same characteristics. Over and over the same traits appear in separated populations of organisms as the result of mutations ‘from within’. Animal breeders like exuberant patterns and rarities; that is exactly what they are looking for to select. Aberrant traits that are normally under stringent negative selection, as might be the case for the pigeon’s reversed feathers, may readily become visible as soon as the selective pressure is relieved; that is, when organisms are reared and fed in the protective environment of captivity. Darwin called the phenomenon of independent acquisition of the same traits *analogous variation*. It is a common phenomenon well known to breeders, and Darwin easily found more examples of analogous variation:

‘The frequent presence of fourteen or even sixteen tail-feathers in the pouter, may be considered as a variation representing the normal structure of another race, the fantail. I presume that no one will doubt that all such analogous variations are due to the several races of the pigeon having inherited from a common parent the same constitution and tendency to variation, when acted on by similar unknown influences. In the vegetable kingdom we have a case of analogous variation, in the enlarged stems, or roots as commonly called, of the Swedish turnip and Ruta бага [*sic*] plants which several botanists rank as varieties produced by cultivation from a common parent: if this be not so, the case will then be one of analogous variation in two so-called distinct species; and to these a third may be added, namely, the common turnip. According to the ordinary view of each species having been independently created, we should have to attribute this similarity in the enlarged stems of these three plants, not to the *vera causa* of community of descent, and a consequent tendency to vary in a like manner, but to three separate yet closely related acts of creation.’⁷⁵

Analogous variation originates in the genome. Through rearrangement and/or transposition of DNA elements, previously silent (cryptic) traits can be activated. The underlying molecular mechanism can’t be merely random; if it were, then Darwin, and other breeders, would not have observed the expression of the same traits independently of each other. A more contemporary translation of analogous variation would be *non-random* (or: *non-stochastic*) *variation*, and it implies some sort of mechanism.

Reversions

In the excerpt above, Darwin also describes what he calls *reversions*. By this term he meant traits that are present in ancestors, then disappear in

first generation offspring, and then reappear in subsequent generations. Darwin acknowledged that unknown laws of inheritance must exist, but still he talks about ‘the proportion of blood’. Reversions are easily explained as traits present on separate chromosomes, and the inheritance of such traits is best understood from Gregor Mendel’s inheritance laws. Through Mendel’s discovery of the genetic laws that underlie the inheritance of traits associated with chromosome segregation (a hallmark of sexual reproduction), Mendel gave us a quantum theory of inheritance. He found that traits are always inherited in well-defined and predictable proportions, and do not just come and go. Darwin’s ‘reversions’ are traits that reappear in later generations due to the inheritance of the same genes (alleles) from both parents.⁵ Darwin didn’t know about Mendel’s laws of inheritance, neither did he know about how variation is generated in genomes. What Darwin described in *Origin*, however, is that variation in offspring is a rule of biology. What Darwin described in isolated species (whether domesticated breeds or island-bound birds) was the result of a burst of abundant speciation resulting from multipurpose genomes. Variant breeds of pigeons are the phenotypes of a rearranged multipurpose pigeon genome. The Galápagos finches (with their distinct beaks and body sizes) are the phenotypes of a rearranged multipurpose finch genome. Where does the variation stem from in populations of Galápagos finches?

Darwin was well aware of the profound lack of knowledge on the origin of variation, and did not exclude mechanisms or laws to drive biological variation:

‘I have hitherto sometimes spoken as if the variations so common and multiform in organic beings under domestication, and in a lesser degree in those in a state of nature had been due to chance. This, of course, is a wholly incorrect expression, but it serves to acknowledge plainly our ignorance of the cause of each particular variation.’⁷⁶

Since Darwin’s days, almost all corners of the living cell have been explored and our biological knowledge has

bacteria	<i>IS</i> -elements
yeast (<i>S. cerevisiae</i>)	<i>Ty</i> -elements
plants (<i>Zea spp</i>)	<i>Bs</i> -elements, <i>Cin</i> -elements <i>Mu</i> -elements
insects (<i>D. Melanogaster</i>)	<i>Copia</i> , <i>gypsy</i> , <i>P</i> -elements
mammals	<i>ERVs</i> , <i>LINES</i> , <i>SINES</i> (e.g. <i>Alu</i> -elements)

Box 1. Common names of some well-known variation-inducing genetic elements (VIGEs) in prokaryotes (bacteria) and eukaryotes (yeast, plants, insects and mammals).

expanded greatly. Through a vast library of data generated by new research in biology, we now have the answers to many questions of a biological nature that had puzzled Darwin. We may also have the answer to ‘the cause of each particular variation’, although we may not be aware of it (yet). That is not because it is hidden between billions of other books and hard to find. No, it is because of the Darwinian paradigm. The mechanism(s) that drive biological variations have been elucidated but are not yet recognized as such.

One of the findings of the new biology was that the DNA of most (if not all) organisms contains jumping genetic elements. The mainstream opinion is that these elements are the remnants of ancient invasions of RNA viruses. RNA viruses are a class of viruses that use RNA molecule(s) for information storage. Some of them, such as influenza and HIV, pose an increasing threat to human health. Are virus invasions responsible for all the beautiful intricate complexity of organic beings? Is a virus a creator? Most likely it is not. Otherwise why would we pump billions of research dollars into research to fight off viruses? Could it be that mainstream science is mistaken?

The RNA virus paradox

Here is one good reason for believing that mainstream science is indeed mistaken: the RNA virus paradox. It has been proposed that these RNA viruses have a long evolutionary history, appearing with, or perhaps before, the first cellular life forms.⁷ Molecular genetic analyses have demonstrated that genomes, including those of humans and primates, are riddled with ‘endogenous retroviruses’ (ERVs), which are currently explained as the remnants of ancient RNA virus-invasions. RNA virus origin can be estimated using homologous genes found in both ERVs and modern RNA virus families. By using the best estimates for rates of evolutionary change (i.e. nucleotide substitution) and assuming an approximate molecular clock,^{8,9} the families of RNA viruses found today ‘could only have appeared very recently, probably not more than about 50,000 years ago’.¹⁰ These data imply that present-day RNA viruses may have originated much more recently than our own species. The implication of a recent origin of RNA viruses and the presence of genomic ERVs poses an apparent paradox that has to be resolved. I will argue, in order to resolve the paradox, we should abstain from the mainstream idea that ERVs are remnants of ancient RNA virus invasions.

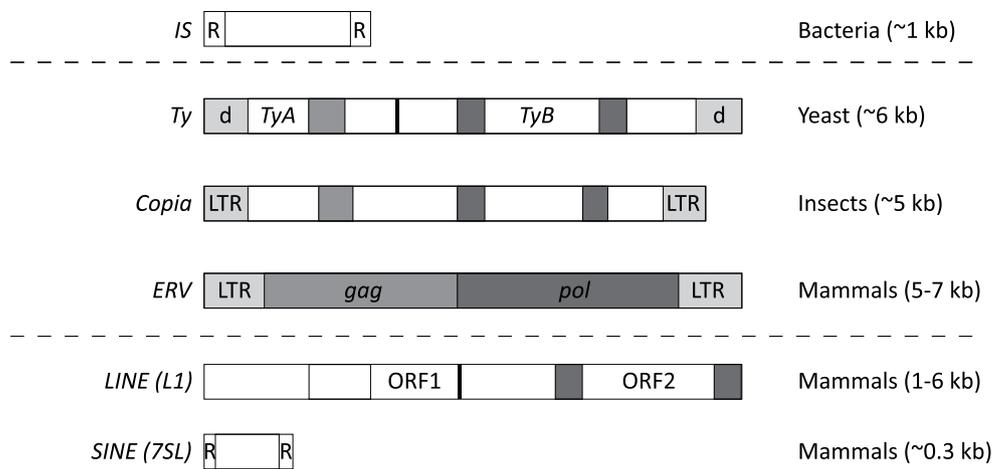


Figure 1. Variation-inducing genetic elements (MIGEs) are found throughout all biological domains, ranging from bacteria to mammals. In yeast, insects and mammals we observe similar designs. (Homologous sequences are indicated by the same colour).

Solving the RNA paradox can only be accomplished by asking questions. First, we have to ask ourselves, What do scientists mean when they refer to genetic elements as *endogenous retroviruses* (ERVs)? In addition, we have to ask, How do ERVs behave, and what—if any—are their functions? ERVs have been extensively studied in microorganisms, such as baker’s yeast (*Saccharomyces cerevisiae*) and the common gut bacterium *Escherichia coli*. Most of our knowledge on the mechanisms of transposition of ERVs comes from those two organisms. In yeast, the ERV known as *Ty* is flanked by long terminal repeats and specifies two genes, *gag* and *pol*, which are similar to genes found in free operating RNA viruses. This is the main argument why scientists believe RNA viruses and ERVs are evolutionarily closely related. The long terminal repeats enable the ERV to insert into the host’s DNA. The transposition and integration is a stringently regulated process and seems to be target or site-specific.^{11,12} During the transpositions of an ERV, the host’s RNA polymerase II makes an RNA template, which is polyadenylated to become messenger RNA. The *gag* and *pol* mRNAs are translated and cleaved into several individual proteins. The *gag* gene specifies a polyprotein that is cleaved into three proteins, which form a capsid-like structure surrounding the ERV’s RNA. We may ask here: why is a capsid involved? It should be noted that single stranded RNA molecules are very sticky nucleotide polymers and the capsid may prevent the ERV from sticking at wrong places. The capsid may also be required to direct the ERV to the right spots in the genome. The *pol* polyprotein is cleaved into four enzymes: protease, reverse transcriptase, RNase and integrase. Protease cleaves the polyproteins into the individual proteins and then the RNA and proteins are packed into a retrovirus-like particle. Reverse transcriptase forms a single-stranded DNA molecule from the ERV RNA template, whereas RNase removes the RNA. The DNA is then circularized and the complementary DNA

strand is synthesized to create a double-stranded, circular copy of the ERV, which is then integrated into a new site in the host's genomic DNA by integrase's activity. This intricate mechanism for transposition of ERVs seems to be irreducibly complex (and thus a sign of intelligent design) since all ERVs and RNA viruses use the same or similar genetic components.

Variation-inducing genetic elements (VIGEs).

What can the function, if any, of ERVs be? If we follow the mainstream opinion, ERVs integrated into the genomes a very long time ago as viral infections. Currently, ERVs are not particularly helpful. They merely hop around in the genome as selfish genetic elements that serve no function in particular. They are mainly upsetting the genome. Long ago, however, RNA viruses are alleged to have significantly contributed to evolution by helping to shape the genome.

It's hard to imagine this story to be true, and not only because of the RNA virus paradox. Modern viruses usually do not integrate into the DNA of the germ line-cells; that is, the genes of an RNA virus don't usually become a part of the heritable material of the infected host. If we obey the uniformitarian principle, we are allowed to argue: 'What currently doesn't happen didn't happen a long time ago, either'. To answer the question raised above, we must start finding out more about some biological characteristics of a less complicated jumping genetic element, the so-called *insertion-sequence* (IS) element. IS elements are DNA transposons abundantly present in the genomes of bacteria. IS elements share an important characteristic with ERVs: transposition. Genome shuffling takes place in bacteria so frequently that we can hardly speak of a specific gene order. The shuffling of pre-existing genetic elements may unleash cryptic information instantly as the result of position effects. Shuffling seems to be an important mechanism to generate variation. But what is the mechanism for genome shuffling? The answer to this question comes unexpectedly from evolutionary experiments, in which genetic diversity ('evolutionary change') was determined between reproducing populations of *E. coli*. During the breeding experiment, which ran for two decades, it was observed that the number and location of IS ('insertion sequence') elements dramatically changed in evolving populations, whereas point mutations were not abundant.¹³ After 10,000 generations of bacteria, the genomic changes were mostly due to duplication and transposition of IS elements. A straightforward conclusion would thus be that jumping genetic elements, such as the IS elements, were designed to deliberately generate variation—variation that might be useful to the organism. In 2004, Lenski, one of the co-authors of the studies, demonstrated that the IS elements indeed generate fitness-increasing mutations.¹⁴ In *E. coli* bacteria IS elements activate cryptic—or silent—catabolic operons: a set of genetic programs for food digestion. It has been reported that IS element transposition overcomes

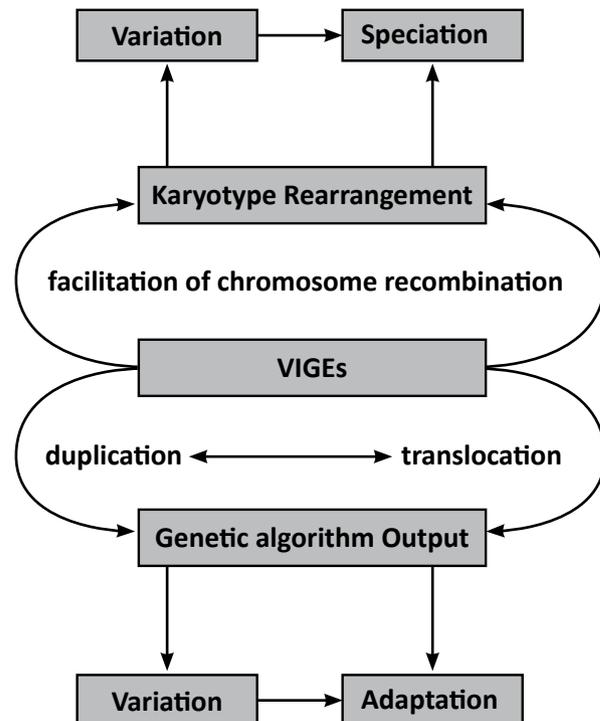


Figure 2. Schematic view of the central role VIGEs may play to generate variation, adaptations and speciation events. Lower part: VIGEs may directly modulate the output of (morpho)genetic algorithms due to position effects. Upper part: VIGEs that are located on different chromosomes may be the result of speciation events, because their homologous sequences facilitate chromosomal translocations and other major karyotype rearrangements.

reproductive stress situations by activating cryptic operons, so that the organism can switch to another source of food. IS elements do so in a regulated manner, transposing at a higher rate in starving cells than in growing cells. In at least one case, IS elements activated a cryptic operon during starvation only if the substrate for that operon was present in the environment.¹⁵

It is clear that in Lenski's experiments, IS elements did not evolve over night. Rather, the IS elements reside in the genome of the original strain. During the two decades of breeding, the IS elements duplicated and jumped from location to location. There was ample opportunity to shuffle genes and regulatory sequences, and plenty of time for the IS elements to integrate into genes or to simply redirect regulatory patterns of gene expression. Microorganisms may thus induce variation simply through shuffling the order of genes and put old genes in new contexts: variation through position effects that can be inherited and propagated in time. It's hardly an exaggeration to state that jumping genetic elements specified by the bacterium's genome generated the new phenotypes.

Transposition of IS elements is mostly characterized by local hopping, meaning that novel insertions are usually in the proximity of the previous insertion and may be a more-or-less random phenomenon; the site of integration isn't sequence dependent. Bacteria have a restricted set of genes and they divide almost indefinitely. Therefore, sequence-dependent insertion and stringent regulation of transposition may not be required for IS-induced reshuffling of bacterial genomes; in a population of billions of microorganisms all possible chromosomal rearrangements may occur due to stochastic processes. In 'higher' organisms the order of genes in the chromosomes is more important, but there is no reason to exclude jumping genetic elements as a factor affecting the expression of genetic programs through position effects. Transposable elements may therefore be a class of variation-inducing genetic elements (VIGEs) in 'higher' organisms. Indeed, ERVs, LINES and SINES resemble IS elements in bacteria in that they are able to transpose. In fact, these elements may be responsible for a large part of the variability observed in higher organisms and may even be responsible for adaptive phenotypes. The genomic transposition of VIGEs is not just a random process. As observed for *Ty* elements in yeast, integration of all VIGEs may originally have been designed as site or sequence specific. It should be noted that VIGEs might qualify as redundant genetic elements, of which the control over translocation may have deteriorated over time.

VIGEs in humans

Mobile genetic elements make up a considerable part of the eukaryotic genome and have the ability to integrate into the genome at a new site within their cell of origin. Mobile genetic elements of several classes make up more than one third of the human genome.

Human *endogenous retroviruses* (ERVs) are, as with yeast ERVs, first transcribed into RNA molecules as if they were genuine coding genes. Each RNA is then transformed into a double stranded RNA-DNA hybrid through the action of *reverse transcriptase*, an enzyme specified by the retrotransposon itself. The hybrid molecule is then inserted back into the genome at an entirely different location. The result of this copy-paste mechanism is two identical copies at different locations in the genome. More than 300,000 sequences that classify as ERVs have been found in the human genome, which is about 8% of the entire human DNA.¹⁶

Long terminal repeats retrotransposons (LTR retrotransposons) are transcribed into RNA and then reverse transcribed into a RNA-DNA hybrid and reinserted into the genome. LTRs and retroviruses are very similar in structure. Both contain *gag* and *pol* genes (figure 1), which encode a viral particle coat (GAG), reverse transcriptase (RT), ribonuclease H (RH) and integrase (IN). These genes provide proteins for the conversion of RNA into complementary DNA and facilitate insertion into the genome. Examples of LTR retrotransposons are human endogenous retroviruses (HERVs). Unlike RNA

retroviruses, LTR retrotransposons lack envelope proteins that facilitate movements between cells.

Non-LTR retrotransposons, such as long interspersed elements (LINES), are long stretches (4,000–6,000 nucleotides) of reverse transcribed RNA molecules. LINES have two open reading frames: one encoding an endonuclease and reverse transcriptase, the other a nucleic acid binding protein (figure 1). There are approximately 900,000 LINES in the human genome, i.e. about 21% of the entire human DNA. LINES are found in the human genome in very high copy numbers (up to 250,000).¹⁷

Short interspersed elements (SINES) constitute another class of VIGEs that may use an RNA intermediate for transposition. SINES do not specify their own reverse transcriptase and therefore they are *retroposons* by definition. They may be mobilized for transposition by using the enzymatic activity of LINES. About one million SINES make up another 11% of the human genome. They are found in all higher organisms, including plants, insects and mammals. The most common SINES in humans are *Alu elements*. Alu elements are usually around 300 nucleotides long, and are made up of repeating units of only three nucleotides. Some Alu elements secondarily acquired the genes necessary to hop around in the genome, probably through recombination with LINES. Others simply duplicate or delete by means of unequal crossovers during cell divisions. More than one million copies of Alu elements, often interspersed with each other, are found in the human genome, mostly in the non-coding sections. Many Alu-like elements, however, have been found in the introns of genes; others have been observed between genes in the part responsible for gene regulation and still others are located within the coding part of genes. In this way SINES affect the expression of genes and induce variation. Alu elements are often mediators of unequal homologous recombinations and duplications.¹⁸

Repetitive triplet sequences (RTSs) present in the coding regions of proteins are a class of VIGEs that cannot actively transpose. RTSs are usually found as an intrinsic part of the coding region of proteins. For instance, RTSs can be formed by a tract of glycine (GGC), proline (CCG), or alanine (GCC). Usually RTSs form a loop in the messenger (m)RNA that provides a docking site for chaperone molecules or proteins involved in the mRNA translation. RTSs may increase or decrease in length through slippery DNA polymerases during DNA replication.

Conclusions and outlook

Now that we have redefined ERVs as a specific class of VIGEs, which were present in the genomes from the day they were created, it is not difficult to see how RNA viruses came into being. RNA viruses have emerged from VIGEs. ERVs, LINES and SINES are the genetic ancestors of RNA viruses. Darwinists are wrong in promoting ERVs as remnants of invasions of RNA viruses; it is the other way around. In my opinion, this view is supported by several recent observations. RNA viruses contain

functional genetic elements that help them to reproduce like a molecular parasite. Usually, an RNA virus contains only a handful of genes. Human Immunodeficiency virus (HIV), the agent that causes AIDS, contains only eight or nine genes. Where did these genes come from? An RNA world? From space? The most parsimonious answer is: the RNA viruses got their genes from their hosts.

The Rous sarcoma virus (RSV), which has the ability to cause tumours, has only 4 genes: *gag*, *pol*, *env* and *src*. In addition, the virus is flanked by a set of repeat sequences that facilitate integration and promote replication. *Gag*, *pol* and *env* are genes commonly present in ERVs. The *src* gene of RSV is a modified host-derived *src* gene that normally functions as a tyrosine kinase—a molecular regulator that can be switched on and off in order to control cell proliferation. In the virus, the regulator has been reduced to an on-switch only that induces uncontrolled cell proliferation. The *src* gene is not necessary for the survival of RSV, and RSV particles can be isolated that have only the *gag*, *pol* and *env* genes. These have perfectly normal life cycles, but do not cause tumours in their host. It is clear the virus picked up the *src* gene from the host. Why wouldn't the whole vector be derived from the host? VIGEs may easily pick up genes or parts thereof as the result of an accidental polymerase II read-through. This will increase the genetic content of the VIGE because the gene located next to the VIGE will also be incorporated. An improper excision of VIGEs may also include extra genetic information. Imagine for instance HERV-K, a well-known human-specific endogenous retrovirus, transposing itself to a location in the genome where it sits next to the *src* gene. If in the next round of transposition a part of the *src* gene was accidentally added to the genes of HERV-K, it has already transformed into a fully formed RSV (see figure 3). It can be demonstrated that most RNA viruses are built of genetic information directly related to that of their hosts.

The outer membranes of influenza viruses, for instance, are built of hemagglutinin and neuraminidase molecules. Neuraminidase is a protein that can also be found in the genomes of higher 'host' organisms, where it serves the function to modify glycopeptides and oligosaccharides. In humans, neuraminidase deficiency leads to neurodegenerative lysosomal storage disorders: sialidosis and galactosialidosis.¹⁹ Even so-called orphan

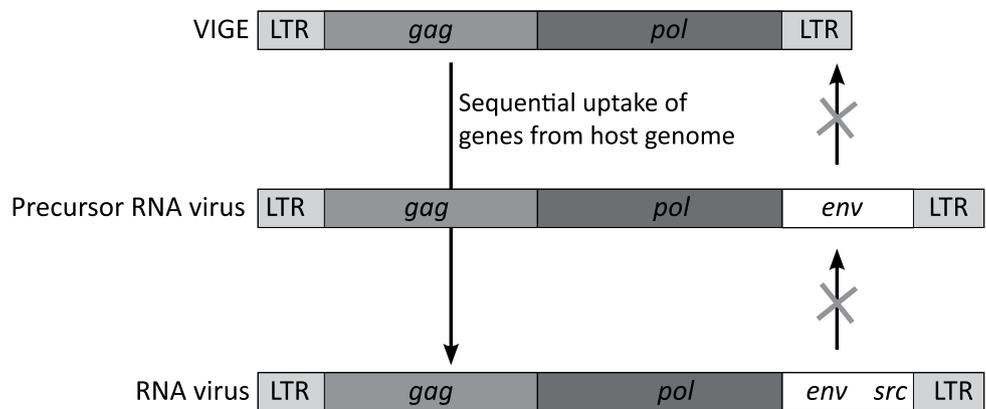


Figure 3. RNA viruses originate from VIGEs through the uptake of host genes. In the controlled and regulated context of the host DNA, genes and VIGEs are harmless. A combination of a few genes integrated in VIGEs may start an uncontrolled replication of VIGEs. In this way, VIGEs may take up genes that serve to form the virus envelope (to wrap up the RNA molecule derived from the VIGE) and genes that enable them to leave and re-enter host cells. Once VIGEs become full-blown shuttle vectors between hosts, they act as virulent, devastating and uncontrolled replicators. Hence, harmless VIGEs may degenerated into molecular parasites in a similar way normally harmless cells turn into tumors once they lose the power to control cell replication. VIGEs are on the basis of RNA viruses, not the other way around. The scheme outlined here shows how the Rous sarcoma virus (RSV) may have formed from a VIGE that integrated the *env* gene and part of the *src* gene (a proto-oncogene; for details see text).

genes, genes that are only found in viruses, can usually be found in the host genomes. Where? In VIGEs!

To become a shuttle-vector between organisms, all that is required is to have the right tools to penetrate and evade the defenses of the host cell. HIV, for instance, acquired part of the gene of the host's defence system (the gp120 core) that binds to the human beta-chemokine receptor CCR5.²⁰

These observations make it plausible that all RNA viruses have their origin in the genomes of living cells through recombination of host's DNA elements (genes, promoters, enhancers). Every now and then such an 'unfortunate' recombination produces a molecular replicator: it is the birth of a new virus. Once the virus escapes the genome and acquires a way to re-enter cells, it has become a fully formed infectious agent. It has long been known that bacteria use genes acquired from bacteriophages—i.e. bacterial viruses that insert their DNA temporarily or even permanently into the genome of their host—to gain reproductive advantage in a particular environment. Indeed, work reaching back decades has shown that prophage (the integrated virus) genes are responsible for producing the primary toxins associated with diseases such as diphtheria, scarlet fever, food poisoning, botulism and cholera. Diseases are secondary entropy-facilitated phenomena. Virologists usually explain the evolution of viruses as recombination: that is, a mixing of pre-existing viruses, a reshuffling and recombination of genes.²¹ In bacteria, viruses may therefore be recombined from plasmids carrying survival genes and/or transposable genetic elements, such as IS elements.

Discussion

Where did all the big, small and intermediate noses come from? Why are people tall, short, fat or slim? What makes morphogenetic programs explicit? The answer may be VIGEs. It may turn out that the created kinds were designed with baranomes that had an ability to induce variation from within. This radical view implies that the baranome of man may have been designed to contain only one morphogenetic algorithm for making a nose. But the program was implicit. The program was designed in such way that a VIGE easily integrated into it, becoming a part of it, hence making the program explicit. Most inheritable variation we observe within the human population may be due to VIGEs—Elements that affect morphogenetic and other programs of baranomes. It should be noted that a huge part of the genomic sequences are ‘redundant’ adaptors, spacers, duplicators, etc., which can be removed from the genome without major affects on reproductive success (fitness). In bacteria, VIGEs have been coined IS elements; in plants they are known as transposons; and in animals, they are called ERVs, LINEs, SINEs, and microsatellites. What these elements are particularly good at is inducing genomic variation. It is the copy number of VIGEs and their position in the genome that determine gene expression and the phenotype of the organism. Therefore, these transposable and repetitive elements should be renamed after their function: *variation-inducing genetic elements*. VIGEs explain the variations Darwin referred to as ‘*due to chance*’.

I will address the details of a few specific classes of VIGEs and argue why modern genomes are literally riddled with VIGEs in a future article. With the realization that RNA viruses have emerged from VIGEs the RNA paradox is solved. For many mainstream scientists this solution will be bothersome because VIGEs were frontloaded elements of the baranomes of created kinds and that implies a young age for their common ancestor and that all life is of recent origin.

References

1. Terborg, P., Evidence of the design of life: part 2—Baranomes, *J. Creation* **22**(3):68–76, 2008.
2. Darwin, C., *The Origin of Species by means of Natural Selection or The Preservation of Favoured Races in the Struggle for Life*, first published by John Murray, 1859. References from Penguin Classics, pp. 85–86, 1985.
3. Darwin, ref. 2, p. 80.
4. Darwin, ref. 2, p. 195.
5. Many recessive traits that come and go in offspring and inherit in a Mendelian fashion can be understood as inactivated redundant genes.
6. Darwin, ref. 2, p. 173.
7. Strauss, E.G., Strauss, J.H. and Levine A.J., Virus evolution; in: Fields, B.N., Knipe, D.M. and Howley, P.M. (Eds.), *Fundamental virology*, 3rd ed., Raven Press, New York, pp. 141–159, 1996.
8. Jenkins, G.M. *et al.*, Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis, *J. Mol. Evol.* **54**:152–161, 2002.
9. Sala, M. and Wain-Hobson, S., Are RNA viruses adapting or merely changing? *J. Mol. Evol.* **51**:12–20, 2000.
10. Holmes E.C., Molecular clocks and the puzzle of RNA virus origins, *J. Virology* **77**:3893–3897, 2003.
11. Barabaugh, P.J., Post-transcriptional regulation of transposition by Ty retrotransposons of *Saccharomyces cerevisia*, *J. Biol. Chem.* **270**:10361–10264, 1995.
12. Wilke, C.M., Maimer, E. and Adams, J., The population biology and evolutionary significance of Ty elements in *Saccharomyces cerevisiae*, *J. Genetics* **86**:155–173, 1992.
13. Papadopoulos, D. *et al.*, Genomic evolution during a 10,000-generation experiment with bacteria, *Proc. Natl Acad. Sci. USA*, **96**: 3807–3812, 1999.
14. Schneider, D. and Lenski, R.E., Dynamics of insertion sequence elements during experimental evolution of bacteria, *Res. Microbiol.* **155**:319–327, 2004.
15. Hall, B.G., Transposable elements as activators of cryptic genes in *E. coli*, *Genetica* **107**:181–187, 1999.
16. Belshaw, R. *et al.*, Long-term reinfection of the human genome by endogenous retroviruses, *Proc. Natl Acad. Sci. USA* **101**(14):4894–4899, 2004.
17. Pierce, B.A., *Genetics: A conceptual approach*, W.H. Freeman, New York, p. 311, 2005.
18. Lonngig, W.E. and Saedler, H., Chromosome rearrangements and transposable elements, *Annu. Rev. Genet.* **36**:389–410, 2002.
19. Neuraminidase deficiency, National Centre for Biotechnology Information, Online Mendelian Inheritance in Men, <www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=256550>, 5 December 2008.
20. Nolan, K.M., Jordan, A.P. and Hoxie, J.A., Effects of partial deletions within the HIV-1 V3 Loop on coreceptor tropism and sensitivity to entry inhibitors, *J. Virol.* **82**: 664–673, 2008.
21. Hamilton G., Virology: the genes weavers, *Nature* **441**:683–685, 2006.

Peer Terborg has an M.Sc. in Biology (Honours biochemistry and molecular genetics) and a Ph.D. in Medical Sciences from the University of Groningen, The Netherlands. He is currently working on the cellular and molecular aspects of pulmonary diseases, such as asthma and COPD, and is an

authority on the molecular biology of signal transduction and gene expression.