

ACCELERATED VARIABILITY OF HUMAN GENES AND TRANSPORTABLE ELEMENTS

GENESIS OF NETWORK

Boris B. Fuks, Alexandr B. Konstantinov

Corresponding author: Boris Fuks

Abstract

A human differs from all other species in that he acts according to a plan, or an idea he has chosen. The discovery of HAR (human accelerated region) showed that evolutionarily new regulatory regions play an important role in the functioning and development of the human brain. In *Homo sapiens*, conserved sequences in this area underwent numerous single nucleotide substitutions. In the five selected HARs, substitution rates were 26 times higher than those for chimpanzees showing 63 extremely fast-paced regions for *H. sapiens*. Human genes that regulate the development of the nervous system during evolution underwent positive selection mainly within their non-coding sequences. 92% of the detected HARs are located in intergenic regions and introns and, therefore, are regulatory sequences, such as enhancers. Only 2% of our genome consists of genes encoding a protein, and the remaining 98% encode regulatory elements that control gene expression in different tissues. Eukaryotic genomes contain thousands to millions of copies of transportable elements (TE). Authors believe that evolution is driven by the dynamics of TEs (transposons) and natural selection. Population studies have found thousands of individual TE insertions in the form of common genetic variants, i.e., TE polymorphisms. Active human TE families include Alu, L1, and SVA elements. These active families of human TE are retrotransposons. Analysis of human polyTE genotypes shows that patterns of TE polymorphism repeat the pattern of human evolution and migration over the past 60,000-100,000 years. They are involved in changes in human regulatory genes. The similarity of patterns allows one to see the effect of TE on regulatory structures that create the structure of the human body, using encoded structures. This conclusion is consistent with studies of intelligence genes, which are based on SNP associations with IQ, as well as with the foundations of a structural and functional network. High proportion of positive selection of genetic variants of our species for last 6 million years may explain the accelerated evolution of *H. sapiens*.

Keywords: HARs; transposones; enhancers; SNP; intelligence

Introduction

Conceptual aspects of the human genome variability

The contemporary paradigm of human evolution is based on the established fact of the two species, humans and chimpanzees, sharing a common predecessor [Pollard, K.S., Salama, S.R., Lambert, N et al., Pollard, K.S., Salama, S.R., King, B., et al. 2006 a b].

Evolution of *Homo sapiens* and chimpanzees has been developing in parallel. Nowadays, chimpanzees have 99% of the *H. sapiens* genes. For about 6 million years, the evolution of humankind has been accelerated relative to other species, including chimpanzee.

The evolutionary acceleration in humans appears to be related to the rate of substitutions in human genome. In the five described human accelerated regions (HARs) substitution rates were

found to be 26 times higher than those for chimpanzees [Pollard, K.S., Salama, S.R., King, B., et al., 2006 b].

Now, what is the sequence of modifications that the human genome underwent on its path to the modern *H. Sapiens* and why did they happen?

As questions are raised regarding the conceptual aspects of the variability of the human genome, one has to keep in mind two major aspects related to this issue. One of them is, traditionally, the usual variability of our species and, on the other hand, the features of its variability over the past 6 million years, which contribute to its successful evolutionary development. The second acceleration concept is discussed in this article.

The textbooks claim, common genetic variation can be caused by mutation (which can create entirely new alleles in a population), random mating, random fertilization, and recombination between homologous chromosomes during meiosis, which reshuffles alleles within an organism's offspring. There are three sources of genetic variation: mutation, gene flow, and sexual reproduction. A mutation is simply a change in the DNA. Mutations are not very common and are usually harmful to a population. Because of this, mutations are usually selected against through evolutionary processes. The human genome comprises about 3×10^9 base pairs of DNA, and the extent of human genetic variation is such that no genomes of two humans are identical. Between any two humans, the amount of genetic variation, the biochemical individuality, is about 1 percent. This means that about one base pair out of every 1,000 will be different between any two individuals.

The most common genetic differences in the human genome are single base-pair differences, the SNP polymorphisms. Other types of polymorphisms – for example, differences in copy number, insertions, deletions, duplications, and rearrangements – also occur, but much less frequently.

Next, we have to note mutations leading to human diseases. A classic example is the mutation for sickle hemoglobin, which in the heterozygous state provides a selective advantage in areas where malaria is endemic. This mutation is a point change in the hemoglobin gene: GAG codon changing to GTG [Alenzi F.Q., and Al Shaya, D.S. 2019].

A number of genetic variations [Jackson, M., Marks, L., May, G. H.W., 2018] cause health problems. Every human disease has a genetic component. In some diseases, such as Huntington disease, Tay-Sachs disease, and cystic fibrosis, this component is dominant. Tay Sachs disease for example is caused by mutations of the HEXA gene on chromosome 15. HEXA gene consists of approximately 35,000 nucleotide pairs and has 14 exons. 100 mutations of this gene have been described [Kaback, M.M., and Desnick, R.J., 2011]. In some other diseases, such as cancer, diabetes, and heart disease, the genetic component is more modest. In fact, we do not typically think of these diseases as “genetic diseases”, because we do not inherit the certainty of developing a disease, but rather a predisposition to developing it [Hubisz, M.J., and Pollard K.S., 2014]. Let us restrict ourselves to these examples, since the section of medical genetics is beyond the scope of this article.

Instead we will focus on the variability of human genes in relation to the problem of successful evolution of *H. sapiens* over the past 6 million years. In the pioneering work [Pollard, K.S., Salama, S.R., Lambert, N. et al., 206a] and in numerous subsequent works, the authors found specific human areas of genetic variability.

The discovery of HAR (human accelerated region) showed that evolutionarily new regulatory regions play an important role in the functioning of the human brain. Most of the

detected HARs were located near telomeres, in areas with high recombination rates. Conserved sequences in these regions underwent numerous single nucleotide substitutions [Bush, E.C, and Lahn, 2008]. Test statistics [Levchenko A., Kanapin A., Samsonova A.,et al.] at an F-adjusted value of $P=0.1$ showed 63 extremely fast-changing regions for *H. sapiens*. The majority, 76%, of HARs have developed due to positive selection [Chen, B., Zhang B., Xu, L., Li, Q., Jiang, F., et al 2017]. HARs provided a higher advantage to the organism in the struggle for survival. As shown in [1, 2] and [7], only ~3% of HARs were found in exons encoding a protein. 92% of the detected HARs are located in intergenic regions and introns and, therefore, can constitute regulatory sequences, such as enhancers. The most important question is why the rapid molecular evolution took place in the evolutionary branch of *H. sapiens*. Authors of [Bush, E.C., and Lahn, 2008] reported that many HARs were developmental gene regulatory elements and RNA genes, most of which evolved their uniquely human mutations through positive selection before divergence of archaic hominins and diversification of modern humans took place.

Another important feature of human genome variability is transposed elements –TE. Eukaryotic genomes contain thousands to millions of copies of transportable elements. But in our species TE have their own characteristics that are important. Authors of [Rishishwar, L., Tellez Villa,C.E., and Jordan, I.K. 2015] believe that evolution is driven by the dynamics of TEs (transposones) and natural selection. Most TE-derived sequences in the genome are remnants of ancient insertion events that are no longer capable of transposition [Sniekers, S., Stringer, S., Watanabe, K. et al. 2017]. Active human TE families include Alu, L1, and SVA elements. Alu elements are short disseminated nuclear elements (SINE) derived from 7SL RNA, L1 is a family of long disseminated nuclear elements (LINE), and SVA elements are complex TEs consisting of a sequence of endogenous human retrovirus and repeats of Alu sequences. These three active families of human TE are retrotransposones, which are transposed by reverse transcription of intermediate RNA. An analysis by the author of the entire genome of the set of human polyTE genotypes shows that the patterns of TE polymorphism repeat the pattern of human evolution and migration over the past 60,000-100,000 years.

Extensive research is currently underway around the so-called intelligence genes. They are based on associations with SNP and IQ: GWAS, Genome-wide association. A meta-analysis of the genomic association in 78,308 people revealed new loci and genes that affect human intelligence [12]. There, in [12], a meta-analysis of studies of genomic associations for cognitive ability identified 190 new loci that contain 939 new genes related to neurogenesis, neuronal differentiation, and synaptic structure.

A paper [13], on the role of soft sweeps, which is discussed in detail below, may have some relevance to the above analysis. So far there is no data to discuss what genetic variability refers to network neurobiology.

In conclusion, a few general aspects need to be addressed. The variability of genes in evolution in general and in the evolution of our species is not completely random. There are always spatial limitations. Crossingover in meiosis is an example. The exchange of regions of homologous chromosomes begins with their dissection in promoters and GC-enriched regions. Typically, these regions are located in recombination hot spots – areas consisting of approximately 1000-2000 base pairs and having a high recombination frequency. Therefore, the occurrence of mutations in crossingover is also spatially limited. Spatial limitations of gene variability in evolution can have a serious consequence, namely, an increase in the likelihood of positive selection of the arising variants. This takes place in HARs e.g. in regions of human acceleration. Authors of [Kern, A.D., and Schrider, D.R. 2018] believe that positive selection can often occur through “soft sweeps” using mutations already present in the population. A soft sweep occurs when a previously neutral mutation becomes beneficial due to an environmental

change. Publication [Kern, A.D., and Schrider, D.R., 2018] provided evidence that soft sweeps are widespread and account for the vast majority of recent human adaptations.

Hard sweep: the classical selective sweep model in which a new advantageous mutation arises, and spreads quickly to fixation due to natural selection. The dominant paradigm in modern evolutionary genetics is the neutral theory, which argues that the vast majority of molecular changes are the result of genetic drift – a stochastic process, and positive selection plays a minor role. Negative selection has been proposed as a classic model to explain differences in levels of genetic diversity in the genome.

In [Pääbo, S., 2015] both theoretical and empirical evidence have accumulated sufficiently as to suggest that soft sweeps might be the more frequent mode of adaptation in many natural populations. Soft sweep selection is obviously a positive selection in evolution and consequently positive selection in *H. sapiens* evolution. A relatively high proportion of the positive selection of genetic variants may explain the accelerated evolution of our species.

Predecessors of Modern Human

Results of analyses of the genomes of Neanderthals and Denisovans [Pääbo, S. 2015] show that the closest evolutionary relatives of present-day humans, suggest that our ancestors were part of a web of now-extinct populations linked by limited, but intermittent or sometimes perhaps even persistent, gene flow. The development of modern man was carried out in stages. Some important steps are: African man - Neanderthal and Denisovan man-modern man. It took about 6 million years. Human-chimpanzee average DNA sequence divergence happened about 6 million years ago [Goodman, M. 1999]. The Denisovan-present-day human divergence took place approximately 800,000 years ago [Reich, D., Green, R.E., Kircher, M., Krause, J., 2010]. Relatively short Denisovan branch allows the bone to be tentatively dated to between 74,000 and 82,000 years before present.

Neanderthals [18], the closest evolutionary relatives of present-day humans, lived in large parts of Europe and Western Asia before disappearing 30,000 years ago. Draft sequence of the Neanderthal genome is composed of more than 4 billion nucleotides from three individuals. Morphological features typical of Neanderthals first appear in the European fossil record about 400,000 years ago. Gradually, more characteristic forms of Neanderthals developed. During the later part of their history, Neanderthals lived in Europe and West Asia as far in the east as Southern Siberia and as far in the south as the Middle East. At this time, Neanderthals apparently came into contact with anatomically modern people in the Middle East, at least 80,000 years ago, and then in Europe and Asia [Green, R.E., Krause, J., Briggs, A.W. et al. 2010].

Limited size of this article allows to focus only on Denisovan and Neandertal gene flow. 6.0% of the genomes of present-day Papuans derive from Denisovans. Denisovans share more alleles with Papuans than with mainland Eurasians. Denisovans have contributed to the genomes of present-day Melanesians, Australian Aborigines, and other South-East Asian islanders. More allele sharing between Denisovan and populations from southern China, such as the Dai, than with populations from northern China, such as the Han. Papuans share more alleles with the Denisovan genome on the autosomes than on the X chromosome. Excess archaic material is more closely related to Neandertals than to Denisovans. The proportion of Neandertal ancestry in Europe is 24% lower than in eastern Asia and South America. One possible explanation is that there were at least two independent Neandertal gene flow events into modern humans. How it was that modern humans came to expand dramatically in population size as well as cultural complexity while archaic humans eventually dwindled in numbers and became physically extinct?

The accelerated development of *H. sapiens* over the past 6 million years has been discussed in the extensive literature on this subject. Several almost non-overlapping lines of research lead to an explanation of the development of intelligence. It seems pertinent to name these lines of research and compare them to evaluate possible new steps towards the main goal. A serious argument explaining the accelerated adaptive evolution of *H. sapiens* is demography [Hawks, J., Wang, E.T., Cochran, G.M., et al. 2007]. A significant increase in population size happened in the last stages: “Demographic change may be the major driver of new adaptive evolution, but the detailed pattern must involve 50,000 years or more”. This is the last evolutionary stage: “Human populations have increased vastly in numbers during the past 50,000 years or more”. The authors note “(iii) an incredibly large number of adaptive substitutions and (iv) almost 100 times the observed number of high-frequency units of non- equilibrium coupling”. They point to the development of culture as the main achievement. Obviously, intelligence is a part of a broad concept of culture, but the genetics of this basic property of *H. sapiens* has remained little studied to this day.

Human Accelerated Regions. HARs

The evolution of *Homo sapiens* and chimpanzees took place in parallel for 6 million years.

Nowadays, chimpanzees have 99% of the *Homo Sapiens* genes. For 6 million years, the evolution of humankind has been accelerated. One of the arguments is the pioneering work of K. Pollard and subsequent authors: Timothy R. Dreszer, et al. (2007), Pollard KS, et al. (2006a), Pollard KS, et al. (2006b), Pollard KS, et al. (2010), Hubisz MJ, et al. (2010), Rosenbloom KR, Siepel A. (2010), Franchini LF, Pollard KS. (2015). In her first work, Pollard, KS stated the wide biological significance of her work: “we scanned the entire human genome and revealed the most extreme examples of recent rapid molecular evolution”.

An in-depth review of this group of papers was made recently by Anastasia Levchenko, et al. (2018). The authors begin with the trivial statement “the human brain appeared in evolution, probably as a result of natural selection among intellectually developed representatives of the genus *Homo*. Only *Homo Sapiens* survived, presenting an exceptionally developed terminal brain, in particular the prefrontal cortex, which enabled us to invent more complex instruments and plan ahead”. But how did this exceptionally developed terminal brain form? How this happened over a relatively short period of time? The discovery of HAR (human accelerated region) showed that evolutionarily new regulatory regions play an important role in the functioning of the human brain. In *Homo Sapiens*, conserved sequences in this region underwent numerous single nucleotide substitutions (Hubisz, and Pollard, 2014).

Pollard et al. (2006a) examined a wide range of data for 202 HARs (in subsequent studies, this number increased significantly). *In the five selected HARs, substitution rates were 26 times higher than those for chimpanzees*. Most of the detected HARs were located near telomeres, in areas with high recombination rates. Besides, these were areas with a high content of GC, two bases, connected to each other through three hydrogen bonds, thus creating strong pairs of nucleotides.

And the new options presented in the HAR also often followed this pattern, since they mainly consisted of a transition from weak AT pairs to strong GC pairs. Test statistics (Bush and Lan, 2008) at an F-adjusted value of $P = 0.1$ *showed 63 extremely fast-changing regions for Homo Sapiens*. Contrary to the results of Pollard et al. (2006a), these 63 HARs do not have a significant number of AT replacements with GC, although they are more common near telomeres. Human genes that regulate the development of the nervous system during evolution underwent positive selection mainly within their non-coding sequences (Haygood et al. 2007).

Taken together, scientific evidence shows that two non-coding regulatory regions are modified by single nucleotide substitutions and deletions, and new coding sequences created by segmental duplication are important for human evolution and, apparently, form the human brain. Bird et al. (2007), who originally studied HAR, found that non-coding accelerated sequences are statistically more likely to be found in recent segmental duplicates. Almost the same changes were found among Neanderthals and Denisovans. The authors of the review cited by us earlier believe that the majority, 76%, of HARs are due to positive selection. The following possibility should also be considered. At some stage, HAR appeared, providing a higher advantage for this organism in the struggle for survival.

Previously, we called such a hypothetical gene, a gene leader (B. Berel, 2010). The appearance in the genome of new HARs that contribute to the survival of the leader gene increased the body's advantage in the struggle for survival, which was expressed in accelerating the evolution of "human properties." The emergence of new HARs, for example, out of 202 found by C. Pollard, should have been accompanied by an increasingly rapid positive selection. The important question is, what genes changed during the evolution of Homo sapiens. As shown (Pollard et al. 2006a and Hubisz and Pollard 2014), only ~3% HAR is found in exons encoding a protein. 92% of the detected HARs are located in intergenic regions and introns and, therefore, can be regulatory sequences, such as enhancers. Approximately 30% of these sequences can be enhancers (Capra et al. 2013). Another 62% HAR may constitute another type of regulatory sequence, such as insulators or silencers (Hubisz and Pollard 2014). An initial analysis (Pollard et al. 2006a) showed that HARs are found mainly in the vicinity of genes encoding transcription factors, DNA-binding proteins and regulators of nucleic acid metabolism. It is point variants – SNP polymorphism. Point variants can occur spontaneously during DNA replication. Based on the data on DNA replication, it cannot be excluded that spontaneous replicative SNPs take place, and the acceleration effect is produced by accelerated positive selection of HAR genes. It is a minimum assumption.

In the brain of adults, in situ hybridization revealed HAR1A expression in the frontal lobe and hippocampus, while quantitative PCR indicated high levels of expression in the cerebellum, diencephalon, and frontal lobe (Pollard et al. 2006b). All 18 human-specific substitutions in HAR1 originate from weak ATs turning into strong GC pairs (Pollard et al. 2006b) This strengthens RNA helices in a three-dimensional structure, possibly by creating a stable clover-like construct instead of an elongated and unstable chimpanzee hairpin (Benjaminov et al. 2008). This can determine the different functions of such RNAs (Griffiths-Jones et al. 2005). Both RNAs in HAR1 are long non-coding RNAs. They are involved in many intracellular processes, from transcriptional and post-transcriptional regulation to epigenetic regulation and maintenance of the chromatin structure, often acting in complexes with proteins (Wilusz et al. 2009; Chen 2016; Engreitz et al. 2016 Kim and Shiehattar 2016). In particular, lncRNAs play an important role in the development of the human cerebral cortex (Ng et al. 2013; Lipovich et al. 2014; Hart and Goff 2016). In this sense, the human version of lncRNA HAR1A can form, for example, another transcript in Cajal-Retzius human neurons, which are crucial for the proper development of the six layers of the human cerebral cortex.

It was found that the haplotype of the UDPGC formed by SNP rs6011613, rs2427496, rs6122371, rs750696, rs750697, and rs2427498 completely covers the region including the HAR1A gene [Tolosa et al. 2008]. Another studied HAR-associated gene is AUTS2. Half of the region containing the first four exons of the gene is the region most distinguished between humans and Neanderthals [Green et al., 2010]. The differences are 293 consecutive SNPs that vary in humans but have only allelic ancestors in Neanderthals; all but one SNP do not alter the content of amino acids; therefore, they can play a regulatory role. The protein coding gene

contains up to 19 coding exons depending on the isoform, as indicated in GENCODE v.24 [Harrow et al. 2012]. In this case, the protein has a length of 1259 amino acids. The function of this protein is still mysterious.

The expression of AUTS2 in humans is localized in various parts of the developing brain, including the prefrontal cortex [Zhang et al. 2011]. A neuronal protein is described that contains the PAS 3 domain (NPAS3), which belongs to the main family of PAS helix-loop-helix PAS genes with 14 HAR in its genomic sequence [Kamm, Pisciotano, et al. 2013]. In humans, NPAS3 is expressed in the developing neocortex, hippocampus, and cerebellum [Gould and Kamnasaran, 2011]. 11 out of 14 HARs act as enhancers in the central nervous system, as demonstrated by analysis of an expression reporter containing HAR, promoter, and GFP in transgenic zebrafish [Kamm, Pisciotano, et al., 2013]. Another HAR-associated gene expressed in the brain is Frizzled 8 (FZD8), whose enhancer contains HAR. It was called HARE5 [Boyd et al. 2015]. Studying differences in enhancer activity between HARE5 and its ortholog in chimpanzees Boyd et al. (2015) demonstrated that human HARE5 induces higher LacZ expression starting on day 10 of the embryo. Fluorescent proteins, in turn, showed expression 30 times higher in a human enhancer. Also, the results of Boyd et al. (2015) point to a significant acceleration of the cell cycle of neural progenitors and increased brain size in mice with human HARE5-containing constructs. The basis for the study of human-specific single nucleotide substitutions, deletions and segmental duplications is the determination of the genetic basis of human-specific morphology, of which the human brain, in particular, the cerebral cortex, is the most remarkable. The list of detected variants could be incomplete: potential genomic conversions and inversions throughout the genome. The most important question is why in the evolutionary branch of *Homo sapiens*, rapid molecular evolution took place. Already in her first work, K. Pollard put forward several assumptions on this issue: mutations in the recombination process, gene conversion, errors in DNA repair, and even spontaneous mutations.

Transposed Elements (Transposones)

Only about 2% of our genome consists of genes encoding proteins, while the remaining 98%, once considered junk DNA, encode epigenetic regulatory elements that control gene expression in different tissues and over time from conception to death (Simona Gaudi et al. 2016). For tens of millions of years, evolution has virtually preserved proteins and genes encoding these proteins, but it has been continuously reforming the system of regulatory genes. Using the same source materials – essentially the same set of coding genes and their proteins – regulatory genes have built up a whole variety of creatures.

Eukaryotic genomes contain thousands to millions of copies of transportable elements (TE). [Bing Chen Bo Zhang et al. (2017)] believe that evolution is driven by the dynamics of TEs (transposons) and natural selection. With time, the dynamics of transposons "found" more and more positive solutions that were preserved by natural selection. How did this dynamic change regulatory genes and how did it change the phenotype accordingly? Undoubtedly, TEs somewhat damaged both coding regions and negative selection has obviously rejected such organisms. Evolutionary development has radically changed the shape and size of creatures. It is interesting that sometimes, for example, in mammals, the changes in the shape and size were not accompanied by similar radical changes in the structure of internal organs. Genome variability in phylogenesis is a fundamental phenomenon in the nature of living things. Combined with natural selection, it created humans and their mind. Genetics stands on the eve of an in-depth attempt to directly answer the question of how a certain change in TE causes a corresponding change in the phenotype.

Population studies of the sequences of the human genome have revealed thousands of individual TE insertions that stand out as common genetic variants, i.e., TE polymorphisms [Wang L et al., 2017]. These relatively recent insertions of TE are a source of natural human genetic variation. Researchers are already starting to use population genomic data sets to assess the phenotypic effects of TE polymorphisms on humans, such as specific regulatory effects of genes, as well as coordinated changes in networks of regulatory genes. We face, in particular, the question of what changes have occurred from the transferred elements in the genome and how they contribute to the rapid evolution of *Homo sapiens* from the moment of bifurcation, about 6 million years ago, to the present. Obviously, this variability was not totally stochastic: conditions arose in the genome and in the cell in general that made the changes non-random. In a more general form, the question can be posed as follows: is there such a structural and functional evolution of the genome and its environment that makes its variability not entirely random? It appears natural to consider how the variability of non-coding and coding structures of the genome is carried out at certain periods of the phylogenetic process. At any time, intragenomic and intracellular mechanisms carry out variability according to their own rules, while external causes only activate or inhibit these mechanisms. Sometimes these external factors are almost invisible as, for example, with alternative splicing.

Some mechanisms and events of variability are not available for today's researcher. It is especially true for early ontogenesis. Ontogenesis "repeats" phylogenesis. However, most of the especially early "pages" of phylogenesis are lost to today's ontogenesis. Accessible to the researcher and an important stage in genome variation is the dynamics of TEs transportable units discovered by Barbara McClintock back in 1946. Most TE-derived sequences in the genome are remnants of ancient insertion events that are no longer capable of transposition (Lavanya Rishishwar, 2015). Active human TE families include Alu, L1, and SVA elements. Alu elements are short disseminated nuclear elements (SINE) derived from 7SL RNA, L1 is a family of long disseminated nuclear elements (LINE), and SVA elements are complex TEs consisting of a sequence of endogenous human retrovirus and repeats of Alu sequences. These three active families of human TE are retrotransposons, which are transposed by reverse transcription of intermediate RNA. L1s are autonomous retrotransposons that encode the enzymatic mechanism necessary for catalysis of their own retrotransposition while Alu and SVA elements are transposed using the L1 mechanism. The authors cited above dealt with a set of genomes that yielded 16,192 TE polymorphisms as part of the 1000 Genomes project, genotyped among 2,504 individuals from 26 global populations. It has become possible to systematically evaluate the utility of polyTE loci for the human genetic population. An analysis by the authors of the entire genome of the set of human polyTE genotypes shows that the patterns of TE polymorphism repeat the pattern of human evolution and migration over the past 60,000-100,000 years.

A detailed analysis of the problem was undertaken by Savannah J. Klein and Rachel J. O'Neill (2018). LINE1s, the only remaining mobile family of LINE1 in humans, accounts for ~17-20% of the human genome (Lander et al. 2001). Alu, the active and mobile SINE family in humans, accounts for a smaller portion of the human genome (~11%) in terms of nucleotides, but is more numerous in terms of copy numbers than LINE1, because of their 20-fold smaller element size (Cordaux and Batzer 2009). Unlike LINE1 and Alu, SVA makes up only ~0.2% of the human genome [Cordaux and Batzer 2009; Wang et al. 2005]. Interactions between the TE and the host genome are considered in the context of the Red Queen hypothesis (Van Valen, 1973), where TE and the host genomes experience antagonistic co-evolution (McLaughlin and Malik 2017). Inhibition of TE at the transcription level includes epigenetic modifications of DNA and/or chromatin, which can alter the access to DNA of the protein necessary for transcription and, therefore, regulate the transcriptional activity of TE. While epigenetic modifications are inherited, the TE sequence itself has not been changed in any way and, thus, it

can retain its ability to be mobilized by transcription in the event of a change in epigenetic modifications. There are chromatin modifications that can lead to the repression of TE transcription. These include histone tail modifications, DNA methylation, and changes in chromatin packaging and condensation [Slotkin and Martienssen 2007]. By 2016, there were 124 documented LINE1-mediated insertions that led to genetic disease [Hancks and Kazazian 2016], with LINE1-mediated retrotransposition events accounting for approximately one out of every 250 pathogenic human mutations [Wimmer et al. 2011].

Given that these polymorphic TE inserts with “functional implications”, in terms of gene expression profiles, are found in a healthy population, TE inserts are not harmful, but should also lead to regulatory changes and gene expression options that can be chosen for the human genome evolution [Wang et al. 2017b]. An initial analysis of human polyTEs within and between populations showed that polymorphic TE inserts are detected at a low frequency. Elimination of polyTEs by selection purification is consistent with the fact that polyTEs are significantly underrepresented in gene and exon regions. However, some polyTEs rise to high allele frequencies, and many also exhibit high levels of geographic differentiation, consistent with what was observed for SNPs. This differentiation suggests the possibility that polyTE inserts may influence population-specific phenotypes, formed as a result of selection. To verify this requirement, an additional analysis of the effect of selection on TE polymorphisms is required, as well as the relationship between polymorphic TE and common adaptive phenotypes. Polymorphic TE loci are used as markers in genetic studies of human populations, and it is known that they provide many advantages for such studies.

The genetic diversity represented by TE polymorphisms reflects the known patterns of human evolution and defines subsets of polyTE loci that can be used as informative markers of ancestors. It appears that we are talking about complex organized processes with a pronounced non-random factor over a relatively short evolutionary period. Nature has been creating a human for about 6 million years. During this time, the nonrandom factor in genome variability also had to play a significant role. Data has been recently presented on the positive selection of TE inclusions in the human genome [Lavanya Rishishwar et al, 2018]. The authors sought to evaluate the effect of TE inserts on people's fitness and on the role that natural selection plays in shaping patterns of TE polymorphisms in populations. A population-based genomic study of the effects of natural selection on human genetic variation as a result of recent TE activity was presented. Several families of human TE – primarily Alu, L1, and SVA – continue to actively transpose, thereby causing insertion polymorphism among individual genomes. The authors found a number of positively selected (adaptive) human TE inserts and concluded that they were involved in changes in human regulatory genes. Seven cases of polymorphic TE inserts were identified, which appeared to increase in frequency in specific human populations as a result of positive selection. Five of the seven selected TE inserts correspond to tissue-specific enhancers, and two cases correspond to loci of quantitative expression traits that are associated with inter-individual gene regulatory differences. It was the first report of recent local adaptation associated with polymorphic human TEs. The above data points to the effect of TE on regulatory structures that create the structure of the human body, using encoded structures. It could be reasonable that so far, there was no possibility to establish a direct connection between these two events. Nevertheless, the question arises whether this mechanism of transfers and insertions is the main architect of the evolving human body. Favorable changes in the structure of organisms and the development of intelligence depend on the acceleration of evolution.

This mechanism uses the genes encoding proteins as tools. And besides, natural selection constantly corrects the system, adapting organisms to the environment.

The same reasoning applies to changes in shape and size that occurred as a result of mutations in regulatory structures during the meiotic crossover. A question may arise how much random is the process of changing regulatory structures in the genome during the insertion of pieces of DNA (TEs) in the genome and during the meiotic crossover. The partnership between TEs and the genome revolves between the selfish spread of TE and the "taming" of TE to fulfill a critical function of chromosome inheritance. The ongoing «conflict» between TE and the host genome finds a balance that allows the continued existence and development of TEs. Positive selection selects positive adaptive changes. Favorable changes in the structure of organisms and, in the end, the development of intelligence depends on accelerated evolution. We tried to describe a molecular machine that works on the genome itself and which is encoded in the genome. This auto-variability of the heredity apparatus may add to the understanding of the functions of double-stranded DNA.

Enhancers and SNP

We need to make an assumption. The mutation (s), which separated the future man from a common predecessor with chimpanzees, approximately 6 million years ago, was crucial. It could change the function of the brain so much that it significantly increased the ability of organisms to survive. This mutation should have a fundamental impact on the evolution of this branch of the evolutionary tree. In 2010 we (Berel B.) assumed the term gene-leader as a likely material factor in human evolution. It could contribute to accelerating positive selection of new HARs genes. According to S. Shastry (2002), most SNPs are located not in protein-coding regions, but in other places of the genome; therefore, their distribution is not a subject to selection pressure. This conclusion seems inaccurate. SNP-modified HARs are primarily related to enhancers and therefore are in physical contact with the promoters of the coding genes and therefore are under selection pressure. HAR variability is usually associated with SNP. TEs and HARs, but not just them, created a set of designs over 6 million years of evolution that formed the basis of intelligence, consciousness, thinking. A direct relationship between gene changes and these higher functions has not yet been traced. From [Boyd et al. (2015)], we learn that mice with human HARE5-containing constructs had a significant acceleration of the cell cycle of neural progenitors and increased brain size. Then, accelerated proliferation could result in a faster DNA replication. And this should lead to an increase in the appearance of SNPs and the emergence of new HARs i.e. to accelerated evolution. It has been shown that the SNP index to some extent depends on the time of DNA replication [Amnon Koren et al., 2014].

We can look at the genome as a container of the algorithm in which molecular machines are described along with their structure, place, and sequence of putting into operation. They will create new molecular machines based on their own structure and additional information obtained from DNA. 98.5% of the genome DNA does not code for protein genes. Cloning experiments show that the algorithm is stored in the cell nucleus at a late stage of ontogenesis, i.e., permanently. The interaction of the algorithm with the cytoplasm of the egg only leads to a normal reading of the algorithm and the development of a normal adult organism. The emergence and improvement of the algorithm in evolution is a quite mysterious event. According to K. Pollard, the formation of HARs is mainly about point variants of SNP polymorphism. Accelerated regions (HARs) are conservative genomic loci with increased divergence in humans. Using sequencing of chromatin interaction, massively parallel reporter assays (MPRA) and transgenic mice, Ryan N. Doan (2016) identified disease-related biallelic mutations of HARs in the active enhancers for CUX1, PTBP2, GPC4, CDKL5 and other genes involved in neural function. Levchenko A. et al. (2018) listed HARs associated genes involved in the evolution and function of the brain: AUTS2, NPAS3, FZD8, CUX1, PTBP2, and GPC4. We could not find a match between these genes and 52 intelligence genes in the article by [Sniekers S. (2017)] and

[Junkiart-Czarnecka A., Haus O. (2016)] and we did not compare with a set of 939 genes in the article by Jeanne E. Savage et al. (2018). SNP has a special role in human evolution and disease. People differ by 0.1% DNA. Most of the differences in 0.1% of the DNA are SNPs. Studies of intelligence genes are based on associations between SNP and IQ [Suzanne Sniekers et al. (2017)].

DNA variability during replication, in particular, the spontaneous appearance of SNP, is an important property of DNA as a molecule. Not only negative natural selection preserves positive organisms – but positive selection also does exist, and it could be implemented through the properties of a DNA molecule.

Network neurobiology and positive selection

Extensive research is currently underway around the so-called intelligence genes. They are based on associations with SNP and IQ: GWAS, Genome-wide association. Despite the warning by Chabris C.F et al. (2012), according to which the majority of registered genetic associations with general intelligence could be false positive, the SNP-IQ-correlation-based works have been still accepted and published in 2017. A meta-analysis of the genomic association in 78,308 people revealed new loci and genes that affect human intelligence [Suzanne Sniekers, et al. (2017)]. In a paper by Jeanne E. Savage, et al. (2018), a meta-analysis of studies of genomic associations for cognitive ability identified 190 new loci that contain 939 new genes related to neurogenesis, neuronal differentiation, and synaptic structure. The authors discussed their functional connections with intelligence. Both above-mentioned articles were co-authored by Philip R. Jansen and Danielle Posthuma. The authors believed that the collective picture of interactions, communication, or information transfer forms a circuit (network) that provides behavior [Danielle S. Bassett et al. 2017]. Particular attention was also paid to functional magnetic resonance imaging (fMRI), a technique that can capture changes in inter-regional interactions in the brain. Network science is a branch of the science of complex systems that mathematically codifies systems whose functions can be described to a first approximation by patterns of interaction between components. The use of networked scientific tools to solve neurobiological hypotheses has become "network neurobiology."

Two networks are considered: a structural and a functional network. A structural network is a network in which neural units (cells, columns, regions) are connected by anatomical connections (synapses, axon projections, white matter tracts).

On the contrary, a functional network is a network in which the same neural units, the nodes, are connected to each other according to estimates of the statistical similarity of their activity. Functional near-infrared spectroscopy (fNIRS) has shown that the functional network tracks cognitive load, notably activating in the middle phase of training, for example. A measure of the structure of the global network is the average length of the shortest path, defined as the shortest path between any pair of nodes, averaged over all pairs of nodes. A community structure refers to the presence of groups of nodes, where nodes in a group are more tightly connected to other nodes in their group than to nodes in other groups. The core-periphery structure refers to the presence of a tightly connected core and rarely connected peripherals. In reality, intelligence – thinking – apparently occurs within this network. A network in the brain of modern humans is a product of genes in nerve cells of the brain.

We sought to track the evolution of these genes in two ways: the occurrence of HARs and the dynamics of transposons. [Junkiart-Czarnecka A., Haus O. (2016)] report: intelligence as the ability to reason, think abstractly, and effectively adapt to the environment has been the subject of research in the field of psychology, neurobiology, and genetics over the past twenty years.

Genetic testing of twins, conducted since the 20th century, showed the heritability of intelligence and therefore confirmed the influence of the genetic factor on cognitive processes. Studies of the genetic background of intelligence focus on dopaminergic (DRD2, DRD4, COMT, SLC6A3, DAT1, CCKAR) genes and genes of the adrenergic system (ADRB2, CHRM2), as well as neurotrophin (BDNF) and oxidative stress (LTF, PRNP) genes. Undoubtedly, the evolution of the totality of genes has led to the creation of such an extraordinary function of thinking. The roots of this function and structure go back to the deep past. They are probably associated with the emergence and evolution of sensory organs, especially vision, i.e., with the appearance of the gene of the protein opsin and the molecule of rhodopsin. An obvious fragment of the network – the vasopressin and oxytocin receptor genes – and hormones themselves suggests the possibility of using such structures in the network when thinking. Indeed, social activity cannot occur outside of thinking. Social cognition, memory, and behavior are controlled by these genes [B. Fuks, 2018]. Genes whose evolution was considered above probably participate in the activities of the functional network.

Can the genetic basis of a functional network be related to the accelerated evolution of enhancers? Their number has been selectively increasing during evolution, in particular in the form of HAR genes and TE inserts. That is, the DNA sequences are programmed for the transcription of the coding gene located close to enhancer on the same chromosome. Variations of these enhancer genes (in particular, SNP polymorphism) can lead to changes in their properties, which have not yet been explored. A change in the function of the enhancer gene is possible due to the appearance of new SNPs in it. Changes in enhancer genes may be one of the reasons for the high tone of the functional network. [Capra et al. (2013)] used transgenic LacZ mice to study the enhancer activity of five non-coding nHARs with a constant expression pattern and differences between chimpanzees and humans. These enhancers caused significantly different expression patterns in the central nervous system and limbs of the two species. Could the increasing variability of enhancers alter the activity of coding genes so that this leads to the creation of a functional network? Some important factors beyond the scope of this article are also worthy of evaluation: human-specific segmented duplicates, deletions, and other evolutionarily new sequence variations.

The classical paradigm in modern evolutionary genetics is the neutral theory, which argues that the vast majority of molecular changes are the result of genetic drift, a stochastic process, and positive selection plays a minor role. Genetic drift is a change in the genetic structure of a population caused by random causes and not leading to genotypic adaptation to the environment. Negative selection has been proposed as a classic null model to explain differences in levels of genetic diversity in the genome. *The question that researchers are currently asking is not whether a positive selection is present, but how frequent and strong is the positive selection?* A number of different approaches have been proposed for quantifying the frequency and strength of positive selection using population-based genetic (and genomic) approaches. [Daniel R. Schrider and Andrew D. Kern (2017)] believe that positive selection can often occur through “soft sweeps” using mutations already present in the population (formed in the genome). A soft sweep occurs when a previously neutral mutation that was present in a population becomes beneficial due to an environmental change. Daniel R. Schrider and Andrew D. Kern provided evidence that soft sweeps are widespread and account for the vast majority of recent human adaptations. Positive selection is the process by which new advantageous genetic variants sweep a population. Positive selection, also known as the Darwinian selection, is the main mechanism that Darwin envisioned as giving rise to evolution. Positive selection is a powerful mechanism that not only accelerates evolution but makes it adaptive. It remains to be seen whether HAR genes and TE inserts can be the tools for soft scans in human evolution.

References

- Alenzi F.Q., and AlShaya, D.S. Biochemical and molecular analysis of the beta-globin gene on Saudi sickle cell anemia, *Saudi J. Biol. Sci.*, 2019, vol. 26, no. 7, pp.1377–1384.
- Beniaminov, A., Westhof E., Krol, A. Distinctive structures between chimpanzee and human in a brain non-coding RNA, *RNA*, 2008. Vol.14(7), pp. 1270-1275.
- Berel, B., Revelations of the Bible and the molecular genetics of altruism, *Retro*, St. Petersburg 2010.
- Boyd, J.L., Skove, S.L., Rouanet, J.P., Pilaz, L.J., Bepler, T., Gordân, R. Human - chimpanzee differences in a FZD8 enhancer alter cell-cycle dynamics in the developing neocortex. *Curr Biol*, 2015, vol. 25(6), pp.772-779.
- Bush, E.C., Lahn, B.T. A genome-wide screen for noncoding elements important in primate evolution. *BMC Evol Biol*. 2008, Vol. 8, pp.17.
- Bassett, Danielle S., Marcelo, G., Mattar, A. Network Neuroscience of Human Learning: Potential to Inform Quantitative Theories of Brain and Behavior, *Trends Cogn Sci*. 2017, vol. 21(4), pp. 250-264.
- Capra, J.A., Erwin, G.D., McKinsey, G., Rubenstein, J.L., Pollard, K.S. Many human accelerated regions are developmental enhancers. *Philos Tran R Soc Lond B Biol Sci*. 2013, vol. 368(1632), pp, 00-25.
- Cordaux, R., Batzer, M.A. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet*. 2009, vol. 10, pp.691-703.
- Chabris, C.F., Hebert, B.M., Benjamin, D.J., Beauchamp, J., Cesarini, D. et al. Most reported genetic associations with general intelligence are probably false positives. *Psychol Sci*. 2012, vol. 23, no 11, pp.1314-23.
- Chen, B., Zhang B., Xu, L., Li, Q., Jiang, F., Yang, P., Xu, Y., and Kang, L. Transposable element-mediated balancing selection at Hsp90 underlies embryo developmental variation, *Mol. Biol. Evol.*, 2017, vol. 34, no. 5, pp. 1127–1139. doi: 10.1093/molbev/msx062
- Chen, L.L. Linking long noncoding RNA localization and function. *Trends Biochem Sci*. 2016, vol. 41, no 9, pp.761-772.
- Doan, Ryan N., Byoung-II Bae, Cubelos, Beatriz, Chang, Cindy, et al. The Homozygosity Mapping Consortium for Autism, 2016, Marta Nieto, and Christopher A. Walsh.
- Dreszer, Timothy R., Gregory D. Wall, David Haussler and Katherine S. Pollard. Biased clustered substitutions in the human genome: The footprints of male-driven biased gene conversion. *Genome Res*. (2007), vol.17, no.10, pp. 1420-1430.
- Engreitz, J.M., Ollikainen, N, Guttman, M., Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat. Rev. Mol. Cell Biol*. 2016, vol.17, no12, pp.756-770.
- Franchini, L.F., Pollard, K.S. Can a few non-coding mutations make a human brain? *Bioessays*. 2015, vol. 37, no10, pp.1054-61.

- Fuks, Boris. Genetic and Epigenetic Programming of Human Social Psychology And Behavior. *International Journal of Current Science and Technology*. 2018, Vol.6, Issue 2(A), pp. 551-558.
- Gaudi, S., Guffanti, G., Fallon, J., Macciardi, F.J. Epigenetic mechanisms and associated brain circuits in the regulation of positive emotions: A role for transposable elements. *J. Comp. Neurol.* 2016, vol. 15, no 524(15), pp.2944-54.
- Goodman, M. The genomic record of Humankind's evolutionary roots, *Am.J.Hum. Genet*, 1999, vol. 64, no. 1, pp. 31-39. doi: 10.1086/302218
- Green, R. E., Krause, J., Briggs, A.W. et al. A draft sequence of the Neandertal genome, *Science*, 2010, vol. 3287, no. 5979, pp. 710-722. doi:10.1126/science.1188021
- Griffiths-Jones S. et al. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* 2005, vol. 33 (Database issue), pp. D121-D124.
- Hancks, D.C., Kazazian, H.H. Jr Roles for retrotransposon insertions in human disease. *Mob DNA*, 2016, vol. 7, pp.9.
- Harrow, J. et al. GENCODE: the reference human genome annotation for. The GENCODE Project. *Genome Res.* 2012, vol. 22 no 9, pp.1760-1774.
- Hubisz, M.J., Pollard, K.S. (. Exploring the genesis and functions of Human Accelerated Regions sheds light on their role in human evolution. *Curr. Opin Genet Dev.* 2014, vol. 29, pp.15-21.
- Hart R.P., Goff L.A. Long noncoding RNAs: central to nervous system development. *Int J Dev Neurosci.* 2016, vol.55, pp.109-116.
- Haygood R, Fedrigo O, Hanson B, Yokoyama KD, Wray GA Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. *Nat Genet.*, 2007, vol. 39(9) pp.1140-1144.
- Hawks John, Eric T. Wang, Gregory M. Cochran, Henry C. Harpending, and Robert K. Moyzis). Recent acceleration of human adaptive evolution. *Proc Natl Acad Sci USA*, 2007, Dec 26, vol. 104, (52), pp. 20753-20758.
- Jackson, M., Marks, L., May, G. H.W., and Wilson, J.B. The genetic basis of disease, *Essays Biochem.*, 2018, vol. 62, no 5, pp. 643–723.
- Junkiart-Czarnecka A, Haus O Genetical background of intelligence. *Postepy Hig Med Dosw (Online)*, 2016, Jun 13, vol. 70(0), pp.590-8.
- Kaback, MM, and Desnick, RJ. Hexosaminidase A deficiency. [Updated 2011 Aug 11]. In. *GeneReviews®* [Internet], Adam, M.P., Ardinger H.H., Pagon, R.A. et al., Eds,1999, Seattle (WA): University of Washington, Seattle; 1993-2020. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1218/>
- Kamm, G.B., Pisciotto, F, Kligler R., Franchini, L.F. The developmental brain gene NPAS3 contains the largest number of accelerated regulatory sequences in the human genome. *Mol Biol Evol.*, 2013b, vol. 30(50), pp.1088-1102.
- Kern, A.D., and Schrider, D.R. diploS/HIC: An updated approach to classifying selective sweeps, *G3 (Bethesda)* , 2018, vol. 8, no 6 pp. 1959–1970. <https://doi.org/10.1534/g3.118.200262>

- Kim TK, Shiekhattar R Diverse regulatory interactions of long noncoding RNAs. *Curr Opin Genet Dev.*, 2016, vol. 36:73-82. Page 14.
- Klein Savannah J., Rachel J., O'Neill, Transposable elements: genome innovation, chromosome diversity, and centromere conflict. *Chromosome Res.*, 2018, vol. 26(1), pp. 5-23.
- Koren Amnon, Robert E. Handsaker, Nolan Kamitaki, Rosa Karlić, Sulagna Ghosh, Paz Polak, Kevin Eggan, and Steven A. McCarroll (2014). Genetic variation in human DNA replication timing. *Cell*; Nov 20; 159(5): 1015-1026.
- Lander ES et al. (2001). Initial sequencing and analysis of the human genome. *Nature*; 409:860-921.
- Lipovich L et al. (2014). Developmental changes in the transcriptome of human cerebral cortex tissue: long noncoding RNA transcripts. *Cereb Cortex* 24(6):1451-1459.
- Levchenko Anastasia, Alexander Kanapin, Anastasia Samsonova, and Raul R Gainetdinov (2018). Human Accelerated Regions and Other Human-Specific Sequence Variations in the Context of Evolution and Their Relevance for Brain Development. *Genome Biol Evol.* Jan; 10(1): 166-188.
- McLaughlin R.N. Jr, Malik H.S. (2017). Genetic conflicts: the usual suspects and beyond. *J Exp Biol.*; 220:6-17.
- Ng S.Y., Lin L., Soh BS, Stanton LW (2013). Long noncoding RNAs in development and disease of the central nervous system. *Trends Genet.* 29(8):461-468.
- Pääbo, S. The diverse origins of the human gene pool, *Nat. Rev. Genet.*, 2015, vol.16, no 6, pp. 313-314. <https://doi.org/10.1038/nrg3954>
- Pollard KS, , Salama, S.R., Lambert, N., Lambot, M.A., Coppens, S., Pedersen, J.S., Katzman, S., King, B., Onodera, C., Siepel, A., Kern, A.D., Dehay, C., Igel, H., Ares, M. Jr., Vanderhaeghen, P., and Haussler, D. 2006a. Forces shaping the fastest evolving regions in the human genome. *PLoSGenet.* 2(10).
- Pollard K.S., Salama, S.R., King, B., Kern, A.D., Dreszer, T., Katzman, S., Siepel, A., Pedersen, J.S., Bejerano, G., Baertsch, R., Rosenbloom, K.R., Kent, J., and Haussler, D. Forces shaping the fastest evolving regions in the human genome, 2006b. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* 443(7108): 167-172.
- Pollard K.S., Hubisz M.J., Rosenbloom K.R., Siepel A (2010). Detection of non-neutral substitution rates on mammalian phylogenies. *Genome Res.* 20(1): 110-121.
- Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., T. Maricic, Good, J.M., Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E.E., Marques Stoneking, M., Richards, M., Talamo, S., Shunkov, M.V. Derevianko, A.P., Hublin, J.-J., Kelso, J., Slatkin, M., and Pääbo, S. Genetic history of an archaic hominin group from Denisova Cave in Siberia, *Nature*, 2010, vol. 468, pp. 1053-1060. <https://doi.org/10.1038/nature09710>
- Rishishwar Lavanya, LuWang, JianrongWang, Soojin V.Yi, Joseph Lachance I.King Jorda (2018). Evidence for positive selection on recent human transposable element insertions. *Gene* Volume 675, 30 October, 69-79.

- Rishishwar Lavanya, Carlos E, Tellez Villa, and I. King Jordan (2015). Transposable element polymorphisms recapitulate human evolution. *Mob DNA.*; 6: 21.
- Savage, J.E., Jansen, P.R., Stringer, S. et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Genome-wide association intelligence, Nature Genet.*, 2018, vol. 50, pp. 912-919. <https://doi.org/10.1038/s41588-018-0152-6>
- Slotkin RK, Martienssen R, (2007). Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet.*; 8:272-285.
- Ryan N Doan, Byoung-II Bae, Beatriz Cubelos, Cindy Chang, Amer A. Hossain, Samira Al-Saad, Nahit M. Mukaddes, Ozgur Oner, Muna Al-Saffar, Soher Balkhy, Generoso G, Gascon. The Homozygosity Mapping Consortium for Autism. Marta Nieto, Christopher A. Walsh (2016). Mutations in Human Accelerated Regions (HARs), Disrupt Cognition and Social Behavior. *Cell.* Oct 6; 167(2): 341-354.
- Schrider Daniel R. and Andrew D. Kern (2017). Soft Sweeps Are the Dominant Mode of Adaptation in the Human Genome. *Molecular Biology and Evolution* 34(8), May 2017.
- Shastry S (2002). SNP alleles in human disease and evolution. *J Hum Genet.* 47 561-566.
- Sniekers, S., Stringer, S., Watanabe, K. et al. Erratum: Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet* (2017). 49, 1558, <https://doi.org/10.1038/ng1017-1558c>
- Sniekers Suzanne, Sven Stringer, Kyoko Watanabe, Philip R Jansen, Jonathan RI Coleman, Eva Krapohl, Erdogan Taskesen, Anke R Hammerschlag, Aysu Okbay, Jeanne E. Savage, Philip R. Jansen Danielle Posthuma (2018). Genome-wide association intelligence. *Nature Genetics* Vol. 50, 912-919.
- Tolosa A et al. (2008). Rapid evolving RNA gene HAR1A and schizophrenia. *Schizophr Res.* 99(1-3): 370-372.
- Van Valen L (1973). A new evolutionary law. *Evolutionary Theory*; 1:1-30.
- Wang L Norris ET, Jordan IK (2017). Human Retrotransposon Insertion Polymorphisms Are Associated with Health and Disease via Gene Regulatory Phenotypes. *Front Microbiol.* Aug 2; 8:1418.
- Wang L, Jordan I.K (2018). Transposable element activity, genome regulation and human health. *Curr Opin Genet Dev.* Apr 49:25-33.
- Wilusz JE, Sunwoo H, Spector DL (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* 23(13):1494-1504.
- Wimmer K, Callens T, Wernstedt A, Messiaen (2011). The NF1 gene contains hotspots for L1 endonuclease-dependent de novo insertion. *PLoS Genet.*; 7:e1002371.
- Zabaneh Delilah, Najaf Amin, Gerome Breen, David Cesarini, Christopher F Chabris, William G Iacono, M. Arfan Ikram, Magnus Johannesson, Philipp Koellinger, James J Lee, Patrik K E. Magnusson, Matt McGue, Mike B. Miller, William ER Ollier, Antony Payton, Neil Pendleton, Robert Plomin, Cornelius A Rietveld, Henning Tiemeier, Cornelia M van Duijn, and Danielle Posthuma (2017). Intelligence research should not be held back by its past. *Nat Genet.* Jul; 49(7): 1107-1112.

Zhang Bing Chen, Bo Lingling Xu, Qing Li Feng, Jiang Pengcheng Yang, Yanan XuLe Kang (2017). Transposable Element-Mediated Balancing Selection at Hsp90 Underlies Embryo Developmental, Molecular Biology, and Evolution, Vol. 34, Issue 5, 1 May, 1127-1139.

Zhang YE, Landback P, Vibranovski MD, Long M (2011). Accelerated recruitment of new brain development genes into the human genome. PLoS Biol. 9(10).