
REVIEWS
AND THEORETICAL ARTICLES

The Rate of Human Germline Mutations—Variable Factor of Evolution and Diseases

N. Y. Uspenskaya^{a,*}, S. B. Akopov^a, E. V. Snezhkov^a, and E. D. Sverdlov^a

^a*Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, 117997 Russia*

**e-mail: natusp@gmail.com*

Received May 17, 2018; revised May 22, 2018; accepted July 4, 2018

Abstract—The emergence of genetic diseases and evolutionary processes are associated with the flow of genetic information from one generation to another, in which genetic information carried by gametes can be changed by the appearance of *de novo* germline mutations. The rate of germline mutations determines the rate of evolution and the incidence of heritable disorders. Despite the great theoretical and practical importance, the problem of establishing mutation rates and their dependence on different factors remains scarcely studied, and the mutation rate values obtained by different methods vary considerably. The review discusses different ways of estimating the rate of these mutations and makes an attempt to explain the reasons for discrepancies in the data obtained. Three levels of the mutation formation are considered: (1) mutations that are formed during the development of a given individual during gametogenesis (basic mutations); (2) mutations transmitted to offspring and determining differences in the genomes of consecutive generations (parents and offspring), which include basic mutations and possible changes resulting from complex processes of sperm transfer to oocyte, fertilization, and subsequent events that lead to only one viable offspring of hundreds of millions spermatozoa and oocytes; (3) mutations which are formed at level 2, fixed in evolution, and determine evolutionary processes and differences between genomes, in particular, of hominoids, hominids, and hominins.

Keywords: *de novo* mutations, genome sequencing, divergence, dating, substitutions

DOI: 10.1134/S1022795419050144

GERMLINE *de novo* MUTATIONS

The emergence of genetic diseases and evolutionary processes are associated with the flow of genetic information from one generation to another, carried out by germ cells, spermatozoa and oocytes. The formation of gametes, starting from fertilization of the egg with the formation of zygotes to the formation of mature gametes, occurs through the development of germ cells, including numerous acts of replication and recombination. At the same time, genetic information can be changed by the *de novo* emergence of different mutations, from point mutations, associated with the errors in DNA replication, to large deletions and insertions that can arise during the recombination process. Mutations can also arise under the influence of exogenous factors. The rate at which mutations arise is called the germline mutation rate and it is central to understanding the patterns of emergence of genomic differences between individuals and populations, as well as for evolutionary genetics. This rate determines the rate of evolution and the frequency of occurrence of heritable disorders.

The first attempts to determine the mutation rate in humans were made before the DNA structure was established, on the basis of the frequency of patholog-

ical, phenotypically identifiable mutations (see [1]). These estimates were involuntarily limited to mutations that caused phenotypic differences, for example, at dominant disease loci. These include the classic works of John Haldane (<http://vikent.ru/author/1052/Haldane>). Further estimates were based on phylogenetic comparisons of species with divergence times obtained from the paleontological and archaeological dating of fossils. In recent times, the developments in genome sequencing technology made it possible to estimate the mutation rate on the basis of the *de novo* mutation counting by comparing the genomes of the closest relatives, i.e., in parent–offspring trio or larger pedigrees.

Despite the great theoretical and practical importance, the problem of establishing the mutation rates and their dependence on different factors remains poorly studied and the mutation rates obtained by different methods vary considerably. For instance, rate estimates on the basis of direct genome sequencing of relative human trios give a value of 0.5×10^{-9} substitutions per bp per year, which is 2 times lower than the estimates obtained by phylogenetic methods.

The famous researcher of genome instability M. Lynch [2] wrote: “Despite its central importance to matters of health and phenotypic evolution, factors

affecting the rate and spectrum of mutations spontaneously arising in the human genome remain scarcely studied, and the figures reported by different authors vary considerably. How frequently do germline and somatic mutations arise and to what extent does this differ between the sexes? What is the relative frequency of different forms of mutations, for example, insertions, duplications, and deletions, especially among mutations having phenotypic effects? How does the mutational spectrum in humans compare with that in other species? And, most importantly, what are the consequences of genetic mutations for the long-term well-being of our species?”

This review focuses on the rates of germline mutations, i.e., mutations arising in the germ cells, from zygote up to mature gametes [3–5]. Different ways of estimating the rates of these mutations will be discussed, and the attempts to explain the reasons for the discrepancy in the data obtained will be made. The problem of somatic mutations was discussed in detail in our recent reviews [6, 7].

THREE DIFFERENT LEVELS OF ESTIMATING GERMLINE MUTATION RATE

In this review, the following definitions of the applied terms will be used. The mutation rate is the number of mutations acquired by the genome per unit of time. This unit can be absolute or conditional. For example, a year, million years, cell division, or generation.

It is noteworthy that these units of time are usually different for somatic and germ cells. For somatic cells, this is usually the number of mutations per base pair per cell division, while for germ cells, the number of mutations per base pair or the whole genome per *generation* is more often used. In population biology and evolutionary biology, generation time is the average time between two consecutive generations in a population. In human populations, the generation time is usually assumed to be 22–32 years. The generation time can also be defined as the time required for the birth of an individual, entering puberty, and reproduction. In evolution, a universal unit of the number of mutations per base pair per million years is also used.

The spectrum of mutations is understood as the ratios between different types of mutations: base substitutions, deletions, insertions, recombinations, changes in the repeat element copy number, etc.

In accordance with the method used in determining the mutation rate, three different concepts of germline mutations are used and compared with each other in the literature:

(1) Mutations that, as with somatic cells, are formed during the development of a given individual during gametogenesis and determine the differences (mosaicism) between its individual gametes (sperma-

tozoa or oocytes). This type of mutations is defined as a basic one.

(2) Mutations transmitted to offspring and determining differences in the genomes of consecutive generations (parents and offspring). They result from not only basic mutations identified in (1) but also from additional complex processes of sperm transfer to oocyte, fertilization, and subsequent events, which lead to only one viable offspring among hundreds of millions of spermatozoa and oocytes. In this case, a successful gamete may contain sets of mutations that either were selected for some traits or were simply stochastic mutation sets. Some of them can cause genetic diseases.

(3) Mutations that are transmitted to offspring, as in (2), but are also *fixed in evolution* and determine the evolutionary processes and differences between the genomes, in particular, of hominoids, hominids, and hominins.

In the first two cases, mutation rates are determined in a given generation or in two or three consecutive generations, while type 3 mutations depend on the processes occurring over many generations. It is quite possible that the rates and spectra of mutations varied from generation to generation, and the values that are obtained on the basis of phylogenetic approach represent a certain value averaged over the rates in different generations, which must differ from the first two. Intuitively, the rates of these mutations cannot be higher, but most likely should be lower than those determined in the first two cases. However, as will be seen later, actual experimental estimates of the mutation rates in the first two cases are lower, which requires an explanation and raises many debates.

METHODOLOGICAL APPROACHES TO DETERMINING GERMLINE MUTATION RATES

Germline mutations have been the subject of quantitative research for a long time, beginning in the 1920s (for a review see [8]). However, the efforts of the past century have produced only a fragmented picture. Study of the rate of spontaneous mutations did not reach such a level of accuracy where it can be considered a quantitative science. The authors of the above review ask: “Why, 100 years after determining the electron mass and the speed of light, are we still in the process of trying to make accurate measurements of this fundamental parameter?” They list six reasons for this lag.

First of all, it is because the per-nucleotide mutation rate is extremely low. To measure such low rates, extremely careful and laborious methods are required. Second, mutation is a *random* process, and there is no way to know where and when a mutation will occur. Third, there are several types of mutations, from point gene mutations to chromosomal and genomic, and

their rates need to be measured separately. Fourth, the mutation rate may depend on the *genetic background*, especially on the context around the site in question or its methylation. This background is different in different individuals. Fifth, the mutation rate can be influenced dramatically by many mutagenic environmental and physiological factors, including temperature, chemicals (the compilation of hazardous substances, <http://www.evol.nw.ru/~spirov/hazard/hazard.html>, contains about 700 known mutagens), radiation, age, sex, etc. Finally, the mutation rate itself is subjected to selection and evolution. Moreover, some authors suggest that organisms can increase the mutation rate in times of stress. Therefore, there is very large natural variability, which leads to considerable scattering in the mutation rate values obtained by different authors and using different methods.

In the case of type 1 mutations, which, as with somatic cells, are formed during the development an individual in the course of DNA replication at gametogenesis and determine the differences (mosaicism) between its individual gametes, the germline mutation rate can be determined by sequencing individual spermatozoa (oocytes). In the case of type 2 mutations, transmitted to offspring and determining differences in the genomes of consecutive generations (mainly parents and their offspring), the mutation rate can be determined by analyzing the frequency of occurrence of genetic diseases in the population, as well as by comparative analysis of the parental and offspring genomes. Finally, in the case of type 3 mutations, which are transmitted to offspring, fixed in evolution, and determine the evolutionary processes, in particular, the differences between the genomes of hominoids, hominids, and hominins, the rate is determined by comparing the whole genome sequencing data of modern humans and their evolutionary relatives that originated from a common ancestor, determining the number of differences between them (divergence) and assigning the number of differences to the time or the number of generations separating modern humans (or their evolutionary relatives) from a common ancestor. The latter can be determined independently by paleontological and archaeological dating, including radiocarbon and other types of radioactive isotopes. The problems that arise from this will be discussed below.

It is not surprising that the different history behind the origin of mutations detected by different methods leads to different values of the mutation rate obtained by these methods.

We will consider these methods in somewhat more detail. In this case, we will focus mainly on mutations caused by replication errors leading to base substitutions. This is the most frequent source of errors [9]. Other types of mutational changes are less understood and, accordingly, less attention will be paid to them.

INTRAOrganISMAL (TYPE 1) GERMLINE MUTATION RATE

Single-cell DNA sequencing [10–12] provides a new approach to studying the mechanisms that lead to germline alterations. In one of the first studies in this field, individual sperm cells were isolated, which revealed an average of 22.8 recombination events and 25–36 *de novo* mutations in each sperm cell [13].

The mutation rate is roughly estimated to be 10^{-8} per generation. This is not completely consistent with the mutation rate determined by evaluating *de novo* mutations in sperm by comparing the genome sequences of relatives, which will be discussed below. For instance, in [4], the overall germline mutation rate was reported to be, on average, 64 new mutations per child, 78% of which, on average, 50, were of paternal origin, i.e., represented the number of mutations per spermatozoon. This is higher than 25–36 *de novo* mutations reported in [13]. Unfortunately, these comparisons are hardly legitimate, given a low accuracy of determination.

The calculations performed in [13] showed that 7% of single spermatozoa had aneuploid genomes [13]. In accordance with this study, another group of researchers used the whole-genome sequencing to determine haplotypes in single sperm cells of one individual, which revealed an average of 25.3 cases of recombination per cell [14]. This is also slightly lower than the number of recombinations determined by other methods to be 36 recombination events per generation [15]. In another study, 99 sperm sequences in an Asian individual were determined, and aneuploidy was reported in 4% of cells and 26 recombination events per single sperm cell [16].

While most germline studies focus on sperm cells, in a recent study, fertilized oocytes were analyzed [17]. Oocytes from eight individual females were analyzed and 43 recombination events per oocyte were detected, with the recombination rate 1.6 times higher than in sperm. This study also reported a much higher rate of aneuploidy in oocytes (17.6%) compared to sperm (4–7%). Taken together, these studies showed high genomic diversity that occurred in germline cells prior to the transfer of genetic material to offspring [18].

THE RATE OF TYPE 2 GERMLINE MUTATIONS, DETERMINED BY THE POPULATION FREQUENCIES OF HERITABLE DISEASES

A certain (limited) pattern of the point mutation rate estimates can be obtained from databases of point mutations at the loci that are known to confer monogenic disorders with manifested phenotypic effects. In the case of autosomal dominant and X-linked disorders, the affected individuals can easily be identified as *de novo* mutants. Evaluation of the rate of human mutations on the basis of the frequency of monogenic

(Mendelian) diseases, as mentioned above, has long history dating back to John Haldane.

Historically, human germline mutation rate was calculated by analyzing data on the frequency of new cases of monogenic hereditary diseases [8, 19, 20].

The calculations were based on the hypothesis that severe Mendelian disease arose from the presence in a population of mutant alleles of a given locus. Autosomal dominant Mendelian diseases caused by single highly penetrant and deleterious alleles may be easier to study than others. The frequency of occurrence of this disease (we denote it f) in a population is determined by the balance between the rate μ of the appearance of affected individuals in a population with each new generation and the rate of their removal from the population (negative selection), which takes place because the patients are less adapted to survival in the environment. In the balanced state, the rate of appearance of patients is equal to the rate of their removal, so that the frequency f remains approximately constant (if the conditions in which the population exists do not change considerably). In the first approximation, $f = \mu/s$ (where μ is the rate at which mutations lead to the appearance of a given disease, and s is the coefficient of selection against the phenotype corresponding to the disease, which in the first approximation represents the proportion of mutant alleles that are not transmitted and therefore are lost as a result of selection). If the dominating disease is not deadly, patients can reproduce, but produce less than the “normal” number of offspring; i.e., their fitness is reduced. This mutant allele is lost from the gene pool through selection. The frequency of mutant alleles responsible for causing a disease in a population is the balance between the loss of mutant alleles through selection and the appearance of mutant alleles as a result of new mutations. When a genetic disease restricts reproduction so much that $s = 1$, it is called a lethal genetic disease. The mutation rate per base pair per generation can then be estimated by the disease frequency f , taking into account estimates of the s values and the mutational target size (i.e., the number of sites at which mutations would lead to the disease). One disease can be caused by a number of mutations in different parts of the gene responsible for it. This increases the rate of disease occurrence.

Different mutations form different alleles. Each allele causes a disease. And the rate of disease occurrence is equal to the sum of the mutant allele rates, each of which is approximately equal to the mutation rate per base pair if point mutations are concerned. These disease-incidence approaches not only make several assumptions about the coefficient of selection and the target size but also require total confidence that the observed phenotype is the one that is caused by given genotype, rather than a genocopy or phenocopy [21].

Our compatriot Alexey Kondrashov [22] was the first to use information on the mutant allele sequences causing monogenic genetic defects. He compared data on per locus mutation rates with sequences of nucleotide substitutions, deletions, insertions, and more complex events at nine loci causing autosomal dominant diseases and 12 loci causing X-linked diseases. The average estimate of the combined rate of all mutations was 1.8×10^{-8} per nucleotide per generation, and the coefficient of variation of this rate across the 20 loci was 0.53. It is clear from general considerations that this rate cannot be higher than the germline mutation rate (if we exclude additional mutations arising during the fertilization process), and the fact that it is still higher suggests once again the high degree of inaccuracy of all definitions of this type.

Single nucleotide substitutions occur about 25 times more frequently than all other mutations, while deletions are about 3 times more frequent than insertions, complex mutations are very rare, and in CpG rich areas the substitution rate is an order of magnitude higher.

Since the time of Kondrashov’s research, a considerable body of information on larger number of alleles has been accumulated, and more disorders have been investigated, which has made it possible to clarify preliminary rate estimates and their variations between different loci (for more information, see [23, 24]). However, the figures for the mutation rates remain about the same.

“DIRECT” METHODS FOR ASSESSING THE RATE OF TYPE 2 MUTATIONS AT TRANSMISSION FROM PARENTS TO OFFSPRING

The general approach, which has become possible only during the last few years, is to count the newly emerging mutations on the basis of deep sequencing of family members, especially of parents–offspring trios. This approach gives a direct estimate, but can be technically complicated, since it can produce errors inherent in modern high-throughput methods. In particular, random errors in sequencing and alignment errors can be difficult to distinguish from true *de novo* mutations. In addition, somatic mutations in offspring of the sequenced families cannot always be distinguished from new germline variants [25].

Unexpectedly, these estimates appeared to be noticeably lower (about 2 times) than those based on the species divergence comparison (see below). Comparison of family trios gave estimates of $1–1.8 \times 10^{-8}$ substitutions per base pair per generation, in contrast to $2–2.5 \times 10^{-8}$ obtained in interspecific comparisons [1, 4, 20, 26, 27]. Moreover, the rate varied considerably from family to family. This means that the

genome of an ordinary individual contains 40–100 single-nucleotide *de novo* mutations, with one or two of these in the coding sequence [28, 29].

Modern genomic approaches make it possible to determine additional characteristics of *de novo* mutations, for example, from which parent this mutation originated, and whether it occurred in the germline or postzygotically. We now know that most germline *de novo* mutations have paternal origins and that the number of *de novo* mutations in the offspring increases with more advanced paternal age at conception [20] (see also below).

Analysis of the genome-wide mutation rates in 78 Icelandic parents–offspring trios using whole-genome sequencing showed that, with an average paternal age at conception of 29.7 years, the average *de novo* mutation rate was 1.20×10^{-8} per bp per generation. These observations point to the importance of the father's age to the risk of such diseases as, in particular, schizophrenia and autism [29]. Each newborn receives 30–100 (on average, 60) new mutations. This study has been extended further [30]. The figures obtained are of some interest. They give an idea about the range of the mutation process. The genomes of 1548 Icelanders, their parents, and in 225 cases at least one child were sequenced. A total of 108778 *de novo* mutations were found, both single nucleotide polymorphisms and indels. The number of *de novo* mutants from mothers increases with age considerably slower than that from fathers. Interestingly, these age-related changes are unevenly distributed throughout the genome. A striking example is the 20 Mb region on chromosome 8p with the mutation frequency 50 times higher than the frequency in the rest of the genome. This feature was observed in chimpanzees and to a lesser extent in gorillas and was almost absent from orangutans. This demonstrates that sequence diversity in humans results from the evolving interaction between age, sex, mutation type, and genomic location [30].

A study that estimated the rate of *de novo* repetitive element copy number variants (CNV) in the germline showed that about 0.0065 (with a length of >500 kb) to 0.0123 (~30 kb) of new CNV could be expected to occur per genome per generation [31, 32]. Another study analyzing *de novo* structural changes gives the rate of 2.94 indels (1–20 bp long) and 0.16 structural variants (>20 bp) per genome per generation [33]. Although the number of *de novo* structural variants described is lower than the number of *de novo* point mutations, the number of bases per genome per generation that are involved in these structural changes is actually 50–100 times higher [32].

ESTIMATION OF THE RATE OF TYPE 3 GERMLINE MUTATIONS BY COMPARING THE GENOMES OF MODERN HUMANS AND THEIR EVOLUTIONARY RELATIVES, APES. MOLECULAR DATING IN THE EVOLUTION OF PRIMATES

Having accepted the existence of a global evolutionary clock (i.e., a clock that assumes the same mutation rate for all species or at least for large taxa, such as primates or hominoids, which later turned out to be wrong) and given an independent time estimate of when the compared species branched, it is possible to translate the number of substitutions between the two modern genomes into the yearly mutation rate. For example, an analysis of the fossil record points to 30 million years (My) as the time of split between *Homo* and rhesus macaques (Old World monkey, OWM). Using the nucleotide divergence between the two species of ~6.2% and the diploid genome size of 6×10^9 bp, it is possible to calculate the mutation rate of 10^{-9} per bp per year. Up to 2010, this unit was used as the substitution rate in the evolution of primates [27].

Currently available data indicate that the substitution rate varies from species to species and, moreover, at different stages of species evolution [27, 34]. Obviously, changes determining the trajectory of the evolution of species and depending on differences in the life cycle characteristics (for example, age and size of individuals belonging to a given species in maturity; age at first reproduction; offspring number, size, and sex ratio) should also be found among the primates. A group of researchers [34] analyzed whole genomes from ten primate species, including Old World monkeys (OWM), New World monkeys (NWM), and great apes. These authors examined putatively neutral autosomal sites and concluded that the substitution rates in the evolutionary lineage leading from hominoid-NWM ancestor to NWM are about 64% higher than in lineages leading from a common ancestor to apes. In apes, in turn, the substitution rate in chimpanzees is ~2% higher, and in gorillas, it is 7% higher than in humans. Here, the so-called local molecular clocks are active.

However, not all types of mutations behave the same. In particular, transitions at CpG sites exhibit more regular (clocklike) behavior than other types, which is probably associated with their nonreplicative origin. (Recall that CpG is a sequence in which C is subjected to methylation and, being easily deaminated, turns into T. As a result, at these sites, C>T transitions occur more frequently.) Thus, among primates, not only the overall mutation rate but also the spectrum of mutations varies.

The authors think that events in the evolution of primates are most reliably dated using CpG transitions. By accepting this approach, they estimate the divergence time between humans and chimpanzees to be 12.1 My (compared to the traditionally taken about

6 My), and the separation of humans and gorillas to be 15.1 My (compared to the traditional 7–8 My) [34]. Time will tell who is closer to the truth.

As mentioned in the previous paragraph, recent direct comparisons of sequences of related genomes gave estimates of the rate about 2 times lower, about 0.5×10^{-9} per bp per year. The difference turns out to be considerable. This figure pushes the separation time between primate species very strongly back, and it becomes incompatible with paleontological data. For example, at this rate, the genetic divergence time between monkeys and Old World monkeys is estimated at 47 My, and therefore, the approximate speciation time is 40 My, while fossil dating gives an estimate of the species splitting of no more than 25–30 My [1]. Here we are faced with a discrepancy between the genetic divergence time and actual splitting of species [35].

There are a lot of papers trying to explain the discrepancies between methods of estimating mutation rate. A summary of the literature and a table of mutation rates obtained by different authors are given in [27]. One possible explanation is that, in the evolution of hominoids, there was a decrease in the mutation rate [1]. Another explanation is that a direct analysis of somatic mutations and the derivation of germline mutation rate on its basis can lead to numerous errors and underestimation of the number of mutations [36]. It was also pointed out that, in assessing the rate of hominoid mutations, it was necessary to take into account the life cycle characteristics.

An attempt at such accounting leads to noncontradictory results and estimates the time of splitting between humans and chimpanzees as 6.6 My [37]. This figure is consistent with traditional estimates. A detailed analysis of the situation was given recently in a good review [1]. The author, like many other researchers, suggests that over the past 15 million years, the mutation rate could have slowed down, which explains the above contradictions. Fossil analysis suggests that ancestral monkeys were smaller than modern ones, and small animals are prone to faster reproduction and, consequently, higher mutation rate. In particular, the author of this review notes that additional evidence supporting the lower level of germline mutations in modern humans as compared with their ancestors can be obtained by comparing the DNA of ancient human ancestors and the DNA of modern humans (see below).

Discussions on the rate of germline mutations are likely to continue for a long time. Let us mention very briefly approaches based on the mutation rate calibration by comparison with another process, the data on the rate of which are more accurate. This comparative process can be the microsatellite mutation rate or the rate of recombination. In the first case, the rate of substitutions per bp per generation of $1.4\text{--}2.3 \times 10^{-8}$ ($0.56\text{--}0.9 \times 10^{-9}$ per year) was obtained; in the second

case, 1.6×10^{-8} substitutions per bp per generation (0.64×10^{-9} per year) [26].

ESTIMATION OF THE RATE OF TYPE 3 GERMLINE MUTATIONS BY COMPARING THE GENOMES OF MODERN HUMANS AND DATED ANCIENT GENOMES OF THE SAME SPECIES

DNA from many ancient specimens (ancient DNA, aDNA), which were previously considered unsuitable for analysis because of extensive degradation, can now be successfully sequenced. High-throughput methods of next generation sequencing can dramatically increase the number of sequence reads and hope that this will help to identify the artifacts associated with contamination of ancient DNA, especially in human studies (see below). Currently, whole genomes of ancient anatomically modern humans, archaic hominins, ancient pathogens, and plants have been sequenced. They revealed important functional and genetic information [38].

A recent comparison of the genome sequences of modern humans and an Upper Paleolithic 45000-year-old human led to an estimate of the mutation rate of $0.44\text{--}0.63 \times 10^{-9}$ per bp per year or $1.3\text{--}1.8 \times 10^{-8}$ per bp per generation (under the assumption that the generation time was 29 years) [26].

An elegant method of “branch reduction” (Fig. 1) was used to study the history of populations, determine mutation rates, and establish time scales for phylogenetic trees in humans and other species. For the first time, on the full genomic scale, this method was used by Meyer et al. [39] to estimate the age of a bone from Denisovan individual. Its age is likely more than 50000 years old and, therefore, it could not be dated by the radiocarbon method. The authors compared the discrepancy between the Denisovan sequence and the genomes of modern humans. The result is shown in Fig. 1.

Comparison of the number of substitutions that presumably occurred between the human ancestral genome, chimpanzees, and the Denisovan genome and the number of substitutions between the ancestral and modern human genomes shows that the number of substitutions in the Denisovan genome is 1.16% lower (1.13–1.27% [25, 39, 40]; in Fig. 1, the corresponding branch is shorter). This, probably, reflects the lower evolutionary age of the Denisovan bone, which had less time for accumulating changes, compared to modern humans. Assuming that sequence divergence time between humans and chimpanzees is 6.5 My, it is possible to conditionally determine the bone age by reducing the Denisovan branch, approximately from 74000 to 82000 years, in accordance with the archaeological data. From these data, it is difficult to derive the mutation rate, since many sources of error can affect this estimate. The use of chimpanzee

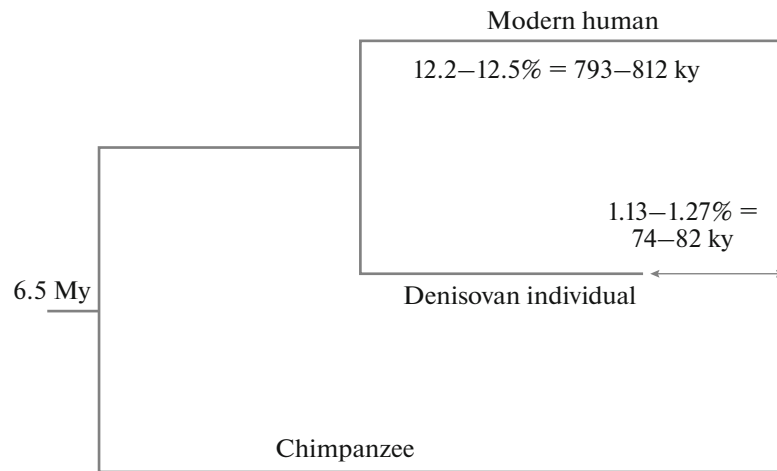


Fig. 1. The average divergence of sequences and the difference in the length of the branches between a Denisovan individual and 11 modern humans. Divergence here is presented as the proportion (%) of the branch leading from human to the human–chimpanzee common ancestor and is recalculated in tens of thousands of years, taking the divergence time between humans and chimpanzees of 6.5 My (cited from [39]).

and human divergence time for calibration is based on the mutation rate estimates, which, as we have seen, are inaccurate. Comparing million-year intervals with ten-thousand-year intervals can lead to large errors in estimating the lifetime of the Denisovan individual. And the interval of tens of thousands of years is just the time period covering the age of the vast majority of ancient human specimens sequenced so far [41].

Because of this, an ideologically similar approach, but using the “recombination clock,” was recently proposed.

ESTIMATION OF AGE AND GENERATION TIME BY GENETIC METHOD USING RECOMBINATIONS IN FOSSIL AND MODERN DNA

Priya Moorjani et al. [41] developed a genetic approach for dating ancient genomes by comparing the number of genetic recombination events accumulated in the genome of modern non-Africans in the time that has passed since the insertion of Neanderthal DNA into DNA of the ancient human ancestor (Fig. 2). This is a more or less accurately dated event from the relatively recent past. The incorporation of Neanderthal DNA into the genome of the modern human ancestor occurred as a result of interbreeding between two populations 37000–86000 years ago. These insertions constitute from 1 to 4% of the modern human genome in Eurasia (but not in Africa). Since the vast majority of ancient specimens sequenced to date have been found in Eurasia (with estimated age values of ~2000–45000 years), the authors used them to determine the number of recombinations from the moment of “spread” of Neanderthal DNA into the genome of modern humans and their extinct relatives, Denisovan individuals [41]. Under the assumption of an approxi-

mately constant recombination rate per generation, the number of recombinations that occurred since the initial insertion of Neanderthal DNA can be used as a molecular clock to determine (1) the number of generations that have passed since the insertion of Neanderthal DNA into the genomes of ancient extinct humans until the death of the ancient owner this genome and (2) the number of generations that have passed between the time of death of the ancient humans and our time, the time of existence of modern genomes.

The idea of the method is the same that was reported in [39] in the method of “branch reduction” described above. The ancient genome of extinct population passed fewer generations in its evolution after radiation from the common ancestor compared to the genomes of extant relatives. This means that it underwent fewer recombination events. The more recently the ancient ancestor died, the smaller the difference in the number of recombinations compared to modern humans. Since recombination rates are more constant in evolution than substitution rates, the accumulated number of recombination events provides a molecular clock that makes it possible to estimate the number of generations separating the lifetime of an ancient relative and the lifetime of a contemporary. The differences in recombination numbers can be translated to absolute time in years using the generation time estimate.

The authors tested their method by comparing the dates obtained on the basis of the recombination rate of five sequenced genomes of ancient human specimens from North America, Europe, and Siberia with radiocarbon dates and obtained consistent age estimates between 12000 and 45000 years ago. Using the correlation between radiocarbon dating, measured in years, and the dates of Neanderthal DNA insertion

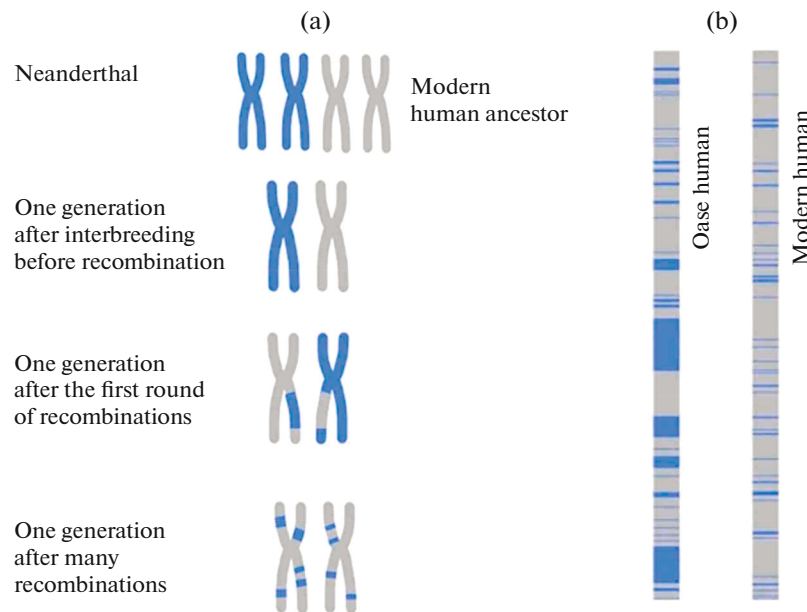


Fig. 2. Fossil dating on the basis of the recombinant fragment length and number distributions (see text) (a). Comparison of chromosome 6 from 40000-year-old fossil from Pesteră-cu-Oase in Romania with the chromosome of a modern human (b). The blue bands represent segments of Neanderthal DNA from past interbreeding. Oase's segments are longer because he had a Neanderthal ancestor just 4–6 generations before he lived. Estimates are based on the number of recombination events (cited from [15]).

into modern human genome, measured in generations, the authors calculated that, in this historical interval, the generation time was approximately 26–30 years, suggesting that the generation time did not change considerably over the past 45000 years. According to the authors, this estimate is consistent with the estimate for modern Western Eurasians [41].

The idea is good, but the difficulties lie in the fossil dating, which cannot always be done with confidence. Even if the fossil ages can be determined accurately, it is often difficult to determine their correct phylogenetic positions on the evolutionary tree. Moreover, a series of data [42] indicate that the time separating us from the lowest common ancestor determined by the gene divergence is always longer, and sometimes much longer, than the divergence time of the species themselves, i.e., the time when they stopped contact [26]. This means that paleontological and archaeological calibrations based on the determination of the time of separate existence of species (populations) cannot always be directly applied to genetic differences [26, 35]. Therefore, for a number of other reasons that we are not able to discuss in the framework of this review, there remains uncertainty regarding the divergence time between humans, chimpanzees, and other hominoids.

In conclusion, it should be emphasized once again that all measurements are extremely approximate, and conclusions that are made on a comparative analysis of the obtained figures must be accepted respectfully, but skeptically.

The main results of germline mutation rate estimates are summarized in Table 1.

PARENTAL AGE, ESPECIALLY PATERNAL AGE, INFLUENCES THE PROBABILITY OF GERMLINE MUTATIONS

Sperm of older fathers undergoes more rounds of cell division and, therefore, accumulates more mutations. On the other hand, mothers transmit fewer mutations (about 0.25 per year), since female oocytes are formed mainly simultaneously, before her birth. The mutation rate also depends on factors such as the onset of puberty, the age of reproduction, and the sperm production rate. These life cycle characteristics vary in living primates and are likely to also differ between extinct species of human ancestors [15].

It was estimated that the male germline experiences 160 genome replications in a 20-year-old man, increasing to 610 genome duplications in a 40-year-old man [4].

This is consistent with what was described in the study [29]: the total number of *de novo* mutations in a child strongly correlated with the father's age at conception, increasing by ~1 or 2 *de novo* mutations for each additional year of paternal age, while mothers made a small contribution (~10 *de novo* mutations), which is considerably less dependent on their age. Similar results were obtained in a number of other studies [4, 43]. Today, a consensus on the rate of germline mutations and, consequently, the risk of genetic

Table 1. The main results of the analysis of germline mutation rates in human genome

Mutation type (see the text)	Test method	Rate of substitutions (recombinations) per bp (recombinations per genome) per generation	Reference
1	Sperm sequencing	10^{-8}	[13, 18]
1	Oocyte sequencing	43 recombinations	[17]
2	Direct sequencing of relative genomes*	$1-1.8 \times 10^{-8}$	[1, 4, 20, 25–29, 32]
2	Analysis of heritable disease population frequencies and sequencing	1.8×10^{-8}	[22]
3	Sequencing of human–monkey genomes	$2-2.5 \times 10^{-8}$	[27]
3	Comparison of the substitution rates with the rates of recombinations or microsatellite mutations	$1.4-2.3 \times 10^{-8}$	[26]
3	Comparison of modern human genomes with dated ancient genomes of the same species	$1.3-1.8 \times 10^{-8}$	[26]

* A detailed summary and a table of mutation rates obtained by different authors by this method are given in [27].

diseases is associated with the fact that the main contributor is the error-prone process of DNA replication, the number of cycles of which increases with the paternal age [44].

In a recent large-scale study [45], data on 816 family trios were used and 36 441 *de novo* mutations were identified (on average, ~45 mutations per child). This is a lower value than previously described. The authors confirmed some previous results, including 80% of mutations that arise during spermatogenesis and the effect of paternal age, i.e., the appearance of approximately one new mutation with each year of increasing paternal age. In earlier studies, it was not possible to detect a small, but important, age-related effect of *de novo* maternal mutations; this paper shows that mothers introduce approximately one mutation for every four additional years of maternal age. It was also found that some genome regions were enriched in maternal mutations.

The mechanisms that lead to an increase in genetic risks for the offspring of older fathers are becoming clearer with improved genomics technologies. In addition to mutations associated with DNA replication, there are many other mutation variants. In this case, base substitutions, insertions–deletions, repeat expansions, and chromosomal changes follow different rules [46].

For example, there is now convincing genetic evidence that, aside from an approximately linear increase in the number of paternally originating point mutations with age [4, 29], there is a steeper rise with age for mutations conferring gain-of-function properties, for example, in the *FGFR2*, *FGFR3*, *HRAS*, *PTPN11*, and *RET* genes. All of these genes are involved in the growth factor receptor–RAS signaling which is active in spermatogonia and is commonly dysregulated in cancer. It is suggested that these muta-

tions provide selective advantages for the proliferation and/or survival of mutant spermatogonia [47].

CONCLUSION. THERE IS A DISBALANCE BETWEEN COMPLEX SOFTWARE ALGORITHMS FOR MOLECULAR CLOCK ANALYSIS AND SCARCE DATA OF GENETIC, PALEONTOLOGICAL, AND ARCHAEOLOGICAL INVESTIGATIONS

Many researchers have noted a huge imbalance between complex software algorithms used to analyze the molecular clock, on one hand, and poor experimental data, on the other hand.

This imbalance does not allow for a rational explanation of the discrepancies between the data on the germline mutation rates obtained by different methods.

Recent studies provide many explanations for why these estimates differ from each other (see, for example, [21, 48, 49]). Possible drawbacks include the following: (a) small sample sizes in terms of the number of individuals used to obtain data; (b) inaccurate estimates of the number of false-positive and false-negative results, which are very likely, since the sequences obtained from different people and having different levels of these errors are compared; (c) analysis of only mutations arising in one generation, which leads to incomplete detection of postzygotic mutations [4]; (d) insufficient consideration of correlation with paternal age and a number of others.

Interestingly, the above list does not include huge errors inherent in the methods used to obtain data for analysis. For example, sequencing of ancient DNA carries many sources of errors that were recently analyzed [50–52]. These include not a high quality of ancient DNA, which, naturally, degraded at different rates for different samples, depending on temperature,

humidity, pH, etc. A very important role is played by contamination of this DNA with modern admixtures of microbiological nature. Specimens obtained from different places and, consequently, stored in different conditions are inevitably compared. The detected differences may be associated not with evolutionary processes, but with processes occurring during the time when the specimen was at the burial place. Caldararo [50] in a very critical analysis of the paper [53] wrote: “Although the authors of the report on the two Denisovan sequences interpret a large number of differences between them as evidence of genetic diversity and, consequently, the long-term presence of Denisovans at this location, these differences can be the result of contamination and degradation.”

The paper [50] deserves a much more detailed analysis than what we can give in the short conclusion to this short review, but together with other similar articles, it sows serious doubts not only about the reliability of the data obtained as a result of sequencing of ancient DNA, but, moreover, in the rationality of their interpretation. Unfortunately, here, despite many warnings (see from quite recently [52]), there is a situation that William Faulkner expressed in his novel *The Sound and the Fury*: “All at once they started talking—hotly, interrupting each other, passionately, turning the unreal into possible, then into probable, then into the indisputable fact, as people always get it when they put desires into words.”

There are no doubts that experimental sequencing data obtained with modern specimens are much less error prone, but their interpretation seems to require serious improvement. The dependence of the mutation rates on the life cycle characteristics should be more carefully taken into account. It cannot be excluded that, for the same species, the mutation rate could be different depending on habitat, where environmental conditions and food sources changed. It seems likely that, in the evolution of hominoids, there was a decrease in the mutation rates on approaching modern populations. Finally, a direct analysis of somatic mutations and the derivation of germline mutation rates from it can lead to numerous errors and underestimation of the number of mutations [36]. These errors can be systematic, since in the process of soma development, different selection processes could have occurred, discarding a considerable part of the mutations that were neutral for the conditions of the germline development, and *vice versa*. Is it possible to correctly consider all these sources of variability?

If we want to know who we are, where we come from and when, and where we are going, we should invest in this problem the intellect and experience of different specialists in order to avoid the GIGO problem (garbage in, garbage out), which is a principle in computer science, implying that if the input data are incorrect, false results will be obtained even if the correct processing algorithm is used.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. This article does not contain any studies involving animals or human participants performed by any of the authors.

REFERENCES

1. Scally, A., Mutation rates and the evolution of germline structure, *Philos. Trans. R. Soc. Lond., B*, 2016, vol. 371, no. 1699. doi 10.1098/rstb.2015.0137
2. Lynch, M., Rate, molecular spectrum, and consequences of human mutation, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, vol. 107, no. 3, pp. 961–968. doi 10.1073/pnas.0912629107
3. Campbell, I.M., Yuan, B., Robberecht, C., et al., Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders, *Am. J. Hum. Genet.*, 2014, vol. 95, no. 2, pp. 173–182. doi 10.1016/j.ajhg.2014.07.003
4. Rahbari, R., Wuster, A., Lindsay, S.J., et al., Timing, rates and spectra of human germline mutation, *Nat. Genet.*, 2016, vol. 48, no. 2, pp. 126–133. doi 10.1038/ng.3469
5. Tang, W.W., Kobayashi, T., Irie, N., et al., Specification and epigenetic programming of the human germ line, *Nat. Rev. Genet.*, 2016, vol. 17, no. 10, pp. 585–600. doi 10.1038/nrg.2016.88
6. Alekseenko, I.V., Kuzmich, A.I., Pleshkan, V.V., et al., The cause of cancer mutations: improvable bad life or inevitable stochastic replication errors?, *Mol. Biol. (Moscow)*, 2016, vol. 50, no. 6, pp. 906–921. doi 10.7868/S0026898416060033
7. Sverdlov, E.D. and Mineev, K., Mutation rate in stem cells: an underestimated barrier on the way to therapy, *Trends Mol. Med.*, 2013, vol. 19, no. 5, pp. 273–280. doi 10.1016/j.molmed.2013.01.004
8. Kondrashov, F.A. and Kondrashov, A.S., Measurements of spontaneous rates of mutations in the recent past and the near future, *Philos. Trans. R. Soc. Lond., B*, 2010, vol. 365, no. 1544, pp. 1169–1176. doi 10.1098/rstb.2009.0286
9. Lynch, M., Evolution of the mutation rate, *Trends Genet.*, 2010, vol. 26, no. 8, pp. 345–352. doi 10.1016/j.tig.2010.05.003
10. Wang, J. and Song, Y., Single cell sequencing: a distinct new field, *Clin. Transl. Med.*, 2017, vol. 6, no. 1, p. 10. doi 10.1186/s40169-017-0139-4
11. Golov, A.K., Razin, S.V., and Gavrilov, A.A., Single-cell genome-wide studies give new insight into nongenetic cell-to-cell variability in animals, *Histochem. Cell Biol.*, 2016, vol. 146, no. 3, pp. 239–254. doi 10.1007/s00418-016-1466-z
12. Junker, J.P. and van Oudenaarden, A., Every cell is special: genome-wide studies add a new dimension to single-cell biology, *Cell*, 2014, vol. 157, no. 1, pp. 8–11. doi 10.1016/j.cell.2014.02.010
13. Wang, J., Fan, H.C., Behr, B., et al., Genome-wide single-cell analysis of recombination activity and *de novo* mutation rates in human sperm, *Cell*, 2012, vol. 150, no. 2, pp. 402–412. doi 10.1016/j.cell.2012.06.030

14. Kirkness, E.F., Grindberg, R.V., Yee-Greenbaum, J., et al., Sequencing of isolated sperm cells for direct haplotyping of a human genome, *Genome Res.*, 2013, vol. 23, no. 5, pp. 826–832. doi 10.1101/gr.144600.112
15. Alex, B. and Moorjani, P., DNA dating: how molecular clocks are refining human evolution's timeline, *Conversation*, 2017.
16. Lu, S., Zong, C., Fan, W., et al., Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing, *Science*, 2012, vol. 338, no. 6114, pp. 1627–1630. doi 10.1126/science.1229112
17. Hou, Y., Fan, W., Yan, L., et al., Genome analyses of single human oocytes, *Cell*, 2013, vol. 155, no. 7, pp. 1492–1506. doi 10.1016/j.cell.2013.11.040
18. Wang, Y. and Navin, N.E., Advances and applications of single-cell sequencing technologies, *Mol. Cell*, 2015, vol. 58, no. 4, pp. 598–609. doi 10.1016/j.molcel.2015.05.005
19. Keightley, P.D., Rates and fitness consequences of new mutations in humans, *Genetics*, 2012, vol. 190, no. 2, pp. 295–304. doi 10.1534/genetics.111.134668
20. Acuna-Hidalgo, R., Veltman, J.A., and Hoischen, A., New insights into the generation and role of *de novo* mutations in health and disease, *Genome Biol.*, 2016, vol. 17, no. 1, p. 241. doi 10.1186/s13059-016-1110-1
21. Segurel, L., Wyman, M.J., and Przeworski, M., Determinants of mutation rate variation in the human germline, *Annu. Rev. Genomics Hum. Genet.*, 2014, vol. 15, pp. 47–70. doi 10.1146/annurev-genom-031714-125740
22. Kondrashov, A.S., Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases, *Hum. Mutat.*, 2003, vol. 21, no. 1, pp. 12–27. doi 10.1002/humu.10147
23. de Ligt, J., Veltman, J.A., and Vissers, L.E., Point mutations as a source of *de novo* genetic disease, *Curr. Opin. Genet. Dev.*, 2013, vol. 23, no. 3, pp. 257–263. doi 10.1016/j.gde.2013.01.007
24. Smith, T., Ho, G., Christodoulou, J., et al., Extensive variation in the mutation rate between and within human genes associated with Mendelian disease, *Hum. Mutat.*, 2016, vol. 37, no. 5, pp. 488–494. doi 10.1002/humu.22967
25. Campbell, C.D. and Eichler, E.E., Properties and rates of germline mutations in humans, *Trends Genet.*, 2013, vol. 29, no. 10, pp. 575–584. doi 10.1016/j.tig.2013.04.005
26. Lipson, M., Loh, P.R., Sankararaman, S., et al., Calibrating the human mutation rate via ancestral recombination density in diploid genomes, *PLoS Genet.*, 2015, vol. 11, no. 11. e1005550. doi 10.1371/journal.pgen.1005550
27. Moorjani, P., Gao, Z., and Przeworski, M., Human germline mutation and the erratic evolutionary clock, *PLoS Biol.*, 2016, vol. 14, no. 10. e2000744. doi 10.1371/journal.pbio.2000744
28. Conrad, D.F., Keebler, J.E., DePristo, M.A., et al., Variation in genome-wide mutation rates within and between human families, *Nat. Genet.*, 2011, vol. 43, no. 7, pp. 712–714. doi 10.1038/ng.862
29. Kong, A., Frigge, M.L., Masson, G., et al., Rate of *de novo* mutations and the importance of father's age to disease risk, *Nature*, 2012, vol. 488, no. 7412, pp. 471–475. doi 10.1038/nature11396
30. Jonsson, H., Sulem, P., Kehr, B., et al., Parental influence on human germline *de novo* mutations in 1548 trios from Iceland, *Nature*, 2017, vol. 549, no. 7673, pp. 519–522. doi 10.1038/nature24018
31. Itsara, A., Wu, H., Smith, J.D., et al., *De novo* rates and selection of large copy number variation, *Genome Res.*, 2010, vol. 20, no. 11, pp. 1469–1481. doi 10.1101/gr.107680.110
32. Forsberg, L.A., Gisselsson, D., and Dumanski, J.P., Mosaicism in health and disease—clones picking up speed, *Nat. Rev. Genet.*, 2017, vol. 18, no. 2, pp. 128–142. doi 10.1038/nrg.2016.145
33. Kloosterman, W.P., Francioli, L.C., Hormozdiari, F., et al., Characteristics of *de novo* structural changes in the human genome, *Genome Res.*, 2015, vol. 25, no. 6, pp. 792–801. doi 10.1101/gr.185041.114
34. Moorjani, P., Amorim, C.E., Arndt, P.F., et al., Variation in the molecular clock of primates, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, vol. 113, no. 38, pp. 10607–10612. doi 10.1073/pnas.1600374113
35. Wilke, T., Schultheiß, R., and Albrecht, C., As time goes by: a simple fool's guide to molecular clock approaches in invertebrates, *Am. Malac. Bull.*, 2009, vol. 27, pp. 25–45. doi 10.4003/006.027.0203
36. Chen, C., Qi, H., Shen, Y., et al., Contrasting determinants of mutation rates in germline and soma, *Genetics*, 2017, vol. 207, no. 1, pp. 255–267. doi 10.1534/genetics.117.1114
37. Amster, G. and Sella, G., Life history effects on the molecular clock of autosomes and sex chromosomes, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, vol. 113, no. 6, pp. 1588–1593. doi 10.1073/pnas.1515798113
38. Der Sarkissian, C., Allentoft, M.E., Avila-Arcos, M.C., et al., Ancient genomics, *Philos. Trans. R. Soc. Lond., B*, 2015, vol. 370, no. 1660, p. 20130387. doi 10.1098/rstb.2013.0387
39. Meyer, M., Kircher, M., Gansauge, M.T., et al., A high-coverage genome sequence from an archaic Denisovan individual, *Science*, 2012, vol. 338, no. 6104, pp. 222–226. doi 10.1126/science.1224344
40. Shendure, J. and Akey, J.M., The origins, determinants, and consequences of human mutations, *Science*, 2015, vol. 349, no. 6255, pp. 1478–1483. doi 10.1126/science.aaa9119
41. Moorjani, P., Sankararaman, S., Fu, Q., et al., A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45000 years, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, vol. 113, no. 20, pp. 5652–5657. doi 10.1073/pnas.1514696113
42. Segurel, L., Thompson, E.E., Flutre, T., et al., The ABO blood group is a trans-species polymorphism in primates, *Proc. Natl. Acad. Sci. U.S.A.*, 2012, vol. 109, no. 45, pp. 18493–18498. doi 10.1073/pnas.1210603109
43. Francioli, L.C., Polak, P.P., Koren, A., et al., Genome-wide patterns and properties of *de novo* mutations in humans, *Nat. Genet.*, 2015, vol. 47, no. 7, pp. 822–826. doi 10.1038/ng.3292

44. Goriely, A., Decoding germline *de novo* point mutations, *Nat. Genet.*, 2016, vol. 48, no. 8, pp. 823–824. doi 10.1038/ng.3629
45. Goldmann, J.M., Wong, W.S., Pinelli, M., et al., Parent-of-origin-specific signatures of *de novo* mutations, *Nat. Genet.*, 2016, vol. 48, no. 8, pp. 935–939. doi 10.1038/ng.3597
46. Crow, J.F., The origins, patterns and implications of human spontaneous mutation, *Nat. Rev. Genet.*, 2000, vol. 1, no. 1, pp. 40–47. doi 10.1038/35049558
47. Maher, G.J., Rajpert-De Meyts, E., Goriely, A., et al., Cellular correlates of selfish spermatogonial selection, *Andrology*, 2016, vol. 4, no. 3, pp. 550–553. doi 10.1111/andr.12185
48. Narasimhan, V.M., Rahbari, R., Scally, A., et al., Estimating the human mutation rate from autozygous segments reveals population differences in human mutational processes, *Nat. Commun.*, 2017, vol. 8, no. 1, p. 303. doi 10.1038/s41467-017-00323-y
49. Scally, A. and Durbin, R., Revising the human mutation rate: implications for understanding human evolution, *Nat. Rev. Genet.*, 2012, vol. 13, no. 10, pp. 745–753. doi 10.1038/nrg3295
50. Caldararo, N., Denisovans, Melanesians, Europeans, and Neandertals: the confusion of DNA assumptions and the biological species concept, *J. Mol. Evol.*, 2016, vol. 83, nos. 1–2, pp. 78–87. doi 10.1007/s00239-016-9755-7
51. Allentoft, M.E., Collins, M., Harker, D., et al., The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils, *Proc. Biol. Sci.*, 2012, vol. 279, no. 1748, pp. 4724–4733. doi 10.1098/rspb.2012.1745
52. Llamas, B., Valverde, G., Fehren-Schmitz, L., et al., From the field to the laboratory: controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era, *Sci. Technol. Archaeol. Res.*, 2017, vol. 3, no. 1, pp. 1–14. doi 10.1080/20548923.2016.1258824
53. Vernot, B., Tucci, S., Kelso, J., et al., Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals, *Science*, 2016, vol. 352, no. 6282, pp. 235–239. doi 10.1126/science.aad9416

Translated by N. Maleeva