FORESTRY IDEAS, 2024, vol. 30, No 2 (68): 272–287

Molecular markers vs. phylogeographic memory dilemma: an example of length-dependent mutation processes

Review paper

Krassimir D. Naydenov[1,3], Michel K. Naydenov[2], Alexander Alexandrov[3], Todor Gurov[4*], Georgi Hinkov[3], Sofiya Ivanovska[4], Veselka Gyuleva[3], Silvi-Maria Gurova[4], Makech Balakumar[5], and Biljana Nikolic[6]

[1]Ministry of National Defence – CAF, 25 Grand-Bernier, St-Jean-sur-Richelieu, QC, J0J1R0, Canada. E-mails: KrassimirDN@gmail.com, KrassimirDN@hotmail.com
[2]UN Secretariat – UNOHR | DMSPC/ GSPD/ STMS/ ODS, Headquarters, 405 E 42nd St., New York, NY 10017, USA. E-mails: michel.naydenov@un.org, michel92kn@gmail.com
[3]Forest Research Institute, Bulgarian Academy of Sciences, 132 Kliment Ochridskii Blvd., Sofia 1756, Bulgaria. E-mails: Alexandrov_38@abv.bg, gjuleva33v@abv.bg, georgihi@abv.bg
[4]Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 25-A, 1113 Sofia, Bulgaria. E-mails: todor.gurov@iict.bas.bg*, sofia@parallel.bas.bg, smgurova@parallel.bas.bg
[5]Departament of Botany, G-Venkataswamy Naidu College, Kovilpatti-628502, India. E-mail: bala.Makeshkumar@gvncollege.edu.in
[6]Institute of Forestry, Kneza Višeslava 3, 11000 Belgrade, Serbia. E-mail: smikitis@gmail.com

§ In Memory of Irena Naydenova

Received: 14 August 2024 / Accepted: 17 October 2024 / Available online: 25 October 2024

Abstract

In this research, we focused on the following question: What element can be used for phylogeography and phylogenetic studies for all species using results of the genome-wide analysis (GWA) of numerous model organisms (e.g., *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*)? We especially paid attention to the average mutation rate variation in the DNA fragment size across alleles from diverse loci (i.e., related to length-dependent mutation processes). Over the past 30 years, all molecular markers have been developed from the fundamentally different mechanisms of four mutations: single-nucleotide polymorphisms (SNPs), simple sequence repeats (SSRs), insertions–deletions (INDELs), and chromosome aberrations. The challenge is to accurately fix the molecular clock for size-variant DNA marker data to uncover the phylogeographic histories of many plants and animals. In this paper, we provided a lot of examples from a *Homo sapience* to a forest tree from *Pinus* genus, illustrating the importance of differentiating between molecular markers, their origins, and the locus molecular size to better calibrate the molecular clock for back-in-time simulation analyses.

Key words: genome analysis, length-dependent mutation processes, molecular clock, phylogeography.

Introduction

Over the past 20–25 years, genome-wide analysis (GWA) has enabled the collection of high-quality genome assemblies of diverse species (Rhie et al. 2021). The millions of sequences of nucleotide data obtained with overall good, effective depth and high resolution have been object of intensive investigations related to the principal, fundamentally different mechanisms underlying four mutations: single-nucleotide polymorphisms (SNPs), simple sequence repeats (SSRs), insertions–deletions (INDELs), and chromosome aberrations. This has, to some extent, changed the definition, models, meaning, and mutation significance for species evolution. These molecular markers have been used for numerous practical applications, such as forensics (Ruitberg et al. 2001), paternity testing (Thomson et al. 1999), personal medicine (Sturzeneker et al. 2000), and determination of the time of divergence between populations (Goldstein and Clark 1995, Stumpf and Goldstein 2001). The initial traits of genomic data of scientific and practical interest are the density, frequency, allele size, and mutation rate of molecular markers. DNA sequence changes are often studied using their chromosome position and length-dependent frequency. The average mutation rate per base pair (μ_{bp}) is observed to be reciprocally proportional to the genome size (*G*) in the base pair (*bp*), as reported by Drake (1991), Drake et al. (1998), Lynch (2010), Sanjuan et al. (2010), and Sung et al. (2012). In addition, the between-species genomic differences in all mechanisms underlying principal mutations have been confirmed (Kruglyak et al. 1998, Dieringer and Schlötterer 2003). The results show the following differences and similarities: a significant difference in the molecular marker density and mutation rate between species; and a significant observable similarity related to length-dependent mutation processes among species. Often, genomic studies focus on SNPs and from two different mutation mechanisms, microsatellites and short INDELs. For SNPs, due to the nature of DNA change, length-dependent mutation processes are not directly applicable. We decided to focus on the similarities between genomic sequence data with free access. Next, we present the results obtained by different researchers.

Short tandem repeats (STRs) or SSRs, known as microsatellites, are repeats of 1–6-nucleotide motifs. These could be *mono*-, *di*-, *tri*-, *tetra*-, *penta*-, and *hexa*-nucleotides. A microsatellite can be *perfect* (or uninterrupted), *imperfect* (or interrupted), or *isolated* (Calabrese and Durrett 2003). Strong evidence suggests that the mutation rates of SSR/STR loci depend on the length of a repeat unit (Weber and Wong 1993; Chakraborty et al. 1997; Schug et al. 1997; Kruglyak et al. 1998, 2000; Sibly et al. 2001, 2003; Calabrese and Durrett 2003; Dieringer and Schlötterer 2003; Lai and Sun 2003; Whittaker et al. 2003; Verbiest et al. 2023). This dependency has been confirmed in different organisms, such as humans, mice, the fruit fly, and yeast. The length-dependent density of SSRs is often related to point mutations, which break up long repeats; long microsatellite repeats are the result of mutation robustness at the genomic level (Kruglyak et al. 1998). Additional features of this DNA change include: (1) the number of repeat motifs is more abundant in *noncoding* DNA (i.e., non-genic region) than in *coding* DNA (i.e., genic region), (2) the microsatellite density differs between chromosome pairs, (3) between intergenic versus intronic sequences (Toth et al.

2000, Morgante et al. 2002), and (4) between sex chromosomes and autosomes (Bachtrog et al. 2000), and (5) the microsatellite frequency decreases from short to large repeat motifs (Calabrese and Durrett 2003, Dieringer and Schlötterer 2003). Another particularity of microsatellites is that mutation robustness depends on the nucleotide repeat motif composition (Schlotterer and Tautz 1992, Chakraborty et al. 1997, Bachtrog et al. 2000). Furthermore, SSR *interrupted* alleles mutate less than *perfect* alleles at the same fragment/allele size (Brinkmann et al. 1998). In human paternity testing (pedigree analysis), Brinkmann et al. (1998) and Makova et al. (2024) also demonstrated statistically significant sex differences in mutations on the X and Y chromosomes for nine loci. Similar results have been reported for humans and other animals by Weber and Wong (1993), Ellegran and Fridolfsson (1997), and McVean (2000). Regarding locus evolution, Garza et al. (1995) showed that microsatellites have a '*target length*', that is, short microsatellites have a mutation bias-up, whereas longer ones have a mutation bias-down. This finding was later supported by Wierdl et al. (1997), Harr and Schlötterer (2000), and Bhargava and Fuentes (2010). Additionally, for genome species evolution, *in silico* analysis have shown that microsatellites of a minimum threshed size exist for slippage mutations (Rose and Falush 1998). Some of the more popular microsatellite mutation mechanisms include: (1) the stepwise mutation model (SMM; Ohta and Kimura 1973); (2) the continuous-time Markov chain, which incorporates length-dependent slippage events, also known as proportional slippage (PS; Kruglyak et al. 1998); (3) the constant exponential model (ConExp; Xu et al. 2000); and (4) slippage occurs only when the length exceeds some threshold K (PSwK; Calabrese et al. 2001). In addition, Calabrese and Durrett (2003) proposed four additional mutation models: (5) piecewise linear bias (PLBias), (6) the asymmetric quadratic model (AsyQuad), (7) the asymmetric linear model (AsyLin), and (8) the multinomial model. It is widely accepted that the SSR/STR mutation mechanism is complex (Lai and Sun 2003).

The second more popular mutation mechanism studied in genomic data is INDELs. This DNA change consists of specific insertions or deletions of genome bases, from 1 to 10,000 nucleotides. Short INDELs include single bases, *monometric* bases, and multi-bp expansions of 2–15 bp repeat units (i.e., *dimetric*, *trimetric*, *tetrametric*, *pentrametric*, *hexametric*, *heptametric*, *octametric, nonametric*, and *decametric*), as shown by Mills et al. (2006). INDELs containing random DNA sequences of more than 50 bases are classified as structural variants (SVs; Mahmoud et al. 2019, Ebert et al. 2021). Short INDELs can be compared to point mutations and microsatellites. The three mutations SNPs, SSRs/STRs, and INDELs result from fundamentally different mechanisms (Buettner et al. 1999, Hill et al. 2003, Yan et al. 2014). Generally, genomic data indicate that for all species, long INDELs (SVs) are less abundant than short INDELs, microsatellites, and SNPs. Studies on genetic variation in *Homo sapiens*, *Drosophila melanogaster*, and *Caenorhabditis elegans* have shown that SVs and chromosome aberrations are formed by strong length-dependent mutation processes, similar to the aforementioned microsatellites (Berger et al. 2001, Wicks et al. 2001, Mills et al. 2006, Rhie et al. 2023, Hallast et al. 2023).

A database from a large geographic region comprises more than 6000 gene

disorders and chromosome aberrations in humans. Known also as abnormalities, on average, a frequency of 7.4 per 10,000 births is responsible for the large numbers of phenotypic and congenital anomalies (Wellesley et al. 2012). The frequency of gene disorders is significantly lower compared to INDELs, SSRs, and SNPs. Here, the mutation load of gene disorders is higher than the mutation changes for INDELs, SSRs, and SNPs. The high mutation load in this example of gene disorders is the result of a large DNA fragment size with affected coding sequences. For example, Zuberi (2013) reported the following chromosome aberrations related to length-dependent mutation processes: partial terminal zone deletion chromosome (1p36), whose frequency is 2 per 10,000 live births, followed by partial chromosome arm deletion with a frequency of 1.86 per 10,000 live births; a duplication rate with a frequency of 0.70 per 10,000 live births; a marker chromosome with a frequency of 0.43 per 10,000 live births; Wolf-Hirschhorn syndrome (4p-) with a frequency of 0.20 per 10,000 live births; and, if the mutation is in a high-density coding DNA region, a frequency of 0.03 per 10,000 live births, as observed for Miller-Dieker syndrome (del 17p13.3). In this example, a gradual decline of surviving frequency is observed for partial chromosome aberration, which results from a combination of coding sequencing density and chromosome arm fragment size DNA. The mutation load's lethal effect in this example is proportional to the affected chromosome fragment size and its coding DNA density.

Another example of length dependent mutation process are *aneuploidy* and *polyploidy*. Aneuploidy is an abnormal number of chromosomes in a cell, that is, cells have one or more chromosomes missing or present in surplus. A good example of aneuploidy (a karyology phenomenon in humans) is the trisomy of chromosome 13 (Patau syndrome), with a frequency of 1 per 10,000–20,000 live births, followed by chromosome 18 (Edwards syndrome), with a frequency of 1 per 3600–8500 live births, and chromosome 21 (Down syndrome), with a frequency of 1 per 800 live births (Greene and Benjamin 2023). In aneuploidy example here, the total extra chromosome size gradually decreases from chromosomes N13 through N18 up to N21, while the lethal effect decreases. When extra complete sets of chromosomes are found, the condition is called polyploidy (high mutation charge). In humans, polyploidy is dangerous; embryo development is observed for a few days or weeks, but finally, pregnancy is terminated by abortion or chirurgical intervention to protect the woman from dangerous complications. The condition is also known as a *partial molar pregnancy* or a *partial hydatidiform mole*, with a frequency of 1 per 1000–1200 births at the zygote stage and zero chances of survival (Mahdieh and Rabbani 2013, De Franciscis et al. 2019, Al Ghadeer et al. 2022). The mutation load is high, with a lethal effect. Polyploidy is found in amphibians, such as salamanders, frogs, and leeches, and in invertebrates, such as clams. In more developed animals, only sterile *triploid* fish, nonsterile *tetraploid* Salmonidae fish, and only one *tetraploid* mammal, the plains viscacha rat (*Tympanoctomys barrerae*), have survived. In contrast to humans and other animals, plants support aneuploidy and polyploidy relatively better. For example, some successful cultivars in agriculture and forest practice are *triploid* (seedless watermelon, banana, and some *Populus* cultivars). In reality, the initial random mutation pattern is under constant selection

pressure from day 1 to reproductive age and death. The change in the genetic/genomic structure follows important steps: prezygotic, prenatal, natal, and classes of ages. For example, in support of constant selection pressure in the population, Korf et al. (2019) reported that in human gametes, abnormality rates of 4–5 % in sperm and 12–15 % in oocytes have been observed, in significant contrast with an abnormality rate of 0.07 % in new births. The postnatal survival rates in humans are inversely related to the affected chromosome fragment. Selection pressure is proportional to the DNA fragment size under mutation and explains part of mutation robustness at the population level.

Over the past 30 years, molecular geneticists' laboratories have developed numerous molecular markers from the three principal, fundamentally different mutation mechanisms (SNPs, SSRs, INDELs), which are specific DNA fragments associated with certain nucleotide patterns, fragment size and locations within the genome (Grover and Sharma 2016). These molecular markers include restriction fragment length polymorphisms, simple sequence length polymorphisms (SSLPs), amplified fragment length polymorphisms, random amplifications of polymorphic DNA (RAPDs), variable number tandem repeats (VNTRs), SSR microsatellite polymorphisms, SNPs, STRs, single-feature polymorphisms, diversity array technology, restriction-site-associated (RAD) DNA markers, allele-specific associated primers, inverse sequence-tagged repeats, inter-retrotransposon amplified polymorphisms (IRAPs), and sequence-tagged sites. More popular existing classifications are: (1) variations at the DNA level, such as nucleotide changes, including deletion, duplication, inversion, and/or insertion; (2) modes of inheritance, that is, dominant/recessive and codominant; (3) *allosomal* and *autosomal*; and (4) plastids and chromosomes. For example, the four principal mutation mechanisms in the human genome include ~10 million SNPs, ~0.5 million SSRs, ~0.5 million short INDELs (single, mono-metric and multi-bp nucleotide repeat expansions), and less than 0.02 million SVs (long INDELs). Chromosome aberrations, full or partial, have a low frequency. The densities of SNPs, SSRs, short INDELs, and SVs have been confirmed for numerous organisms (Mills et al. 2006, Galindo et al. 2009).

We focused on the question, what element can be used for phylogeography and phylogenetic studies for all species using WGA results? We especially focused on the average mutation rate variation in the DNA fragment size (bp) across alleles from diverse loci from SSRs/STRs, short INDELs, SVs, and chromosome aberrations. First, we analyzed how the length-dependent mutation processes are manifested in real genomic data. Second, we focused on where the length-dependent mutation processes could be used for molecular clock calibration for size-variant DNA data, such as SSRs, short INDELs, and SVs, in phylogenetic and phylogeographic studies. The mutation frequency is a well-known fundamental factor in evolution studies. However, the concept of molecular clock calibration for size-variant DNA is poorly understood. In this study, we proposed a simple model for the molecular clock mutation rate list of universal principles and discussed how it helps compare different molecular clock speeds from different molecular markers. The forest trees have been less studied at genome level generally do to their huge total genome size. They have low migration speed and distance compared to animals. That allows us to make back-time

phylogeographic simulation more interesting (Naydenov et al. 2016). More advanced analyses are from deferent species of genera *Pinus, Picea,* and *Populus* (Kovach et al. 2010, Nystedt et al. 2013, Gonzalez-Ibeas et al. 2016, Christenhusz et al. 2024). The obtained results have shown similar length-dependent mutation processes. We also presented a short comparative example of empirical data with three species: *P. nigra* Arn., *P. sylvestris* L.*,* and *P. halepensis* Mill. from the *Pinus* genus (Fig. 1).

Can length-dependent mutation processes in population genetics improve the precision of phylogenetic and phylogeographic hypotheses?

An overview of empirical data on DNA mutation frequency, specifically regarding length-dependent mutation processes in organisms, reveals that each organism's genome contains various mutations with differing frequencies. These mutations can be roughly categorized into different groups at the population level. This categorization could be used to create a more accurate molecular clock for each mutational group, significantly enhancing the accuracy of future predictions in phylogenetics and phylogeography. Below are some suggestions we believe could be valuable for future research.

Specific molecular markers and molecular clock calibration

Two principal factors are involved in computer data simulations for forward and backward molecular clock simulation for evolution, speciation, and phylogenetic and phylogeographic questions: (1) mutation frequency (Lynch 2010, Bergeron et al. 2023, Lynch et al. 2023) and (2) time mark molecular clock fixing. For more information, see Rogers and Harpending (1992), Sanderson (1997, 2003), Ho et al. (2015), O'Reilly et al. (2015), and Donoghue and Yang (2016).

Mutation frequency

Each of the molecular markers (SNPs, SSRs, and long INDELs) has specific strengths and weaknesses, so knowledge of these molecular markers is necessary before using them. For phylogenetic and phylogeographic studies, it is important to determine the mutation rate (μ) of the molecular marker used. The mutation rate depends on multiple factors, some of which are general, while others are specifically related to the species biology. The more important general factors are the molecular marker's size (bp; i.e., the molecular mass of the DNA fragments), the nucleotide fragment order as the *coding* or *noncoding* DNA mutation mechanism, and the origin of the molecular marker (chromosome, mitochondrion, ribosome, or chloroplast; Schug et al. 1997; Dieringer and Schlötterer 2003; Sibly et al. 2001, 2003; Verbiest et al. 2023). Based on the cell structure and morphology, we proposed four universal principles with the limit of one species and/or one genus:

(1) Long fragments (chromosome, locus, allele) have a proportionally lower mutation rate than short fragments (Chakraborty et al. 1997, Calabrese and Durrett 2003, Whittaker et al. 2003, Makova et al. 2024).

(2) *Coding* DNA fragments have a lower mutation rate than *noncoding* DNA fragments of the same size (Weber and Wong 1993; Kruglyak et al. 1998, 2000; Lai and Sun 2003).

(3) Molecular markers with a plastid origin generally have a lower mutation rate

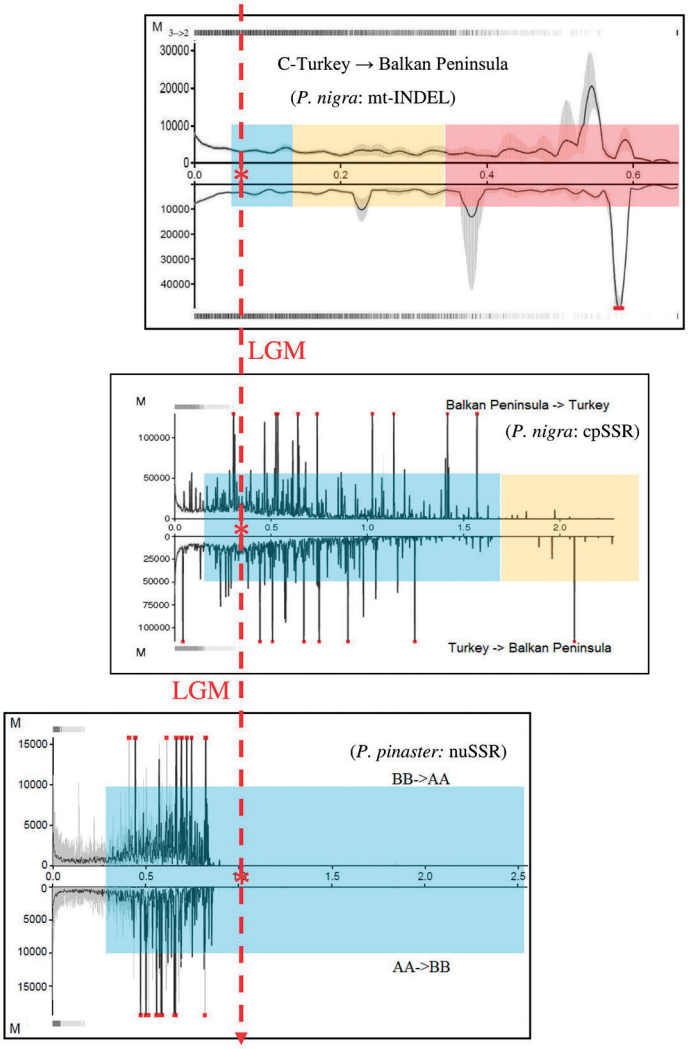

Fig. 1. An example of the phylogenetic memory of three of the most frequently used molecular markers from phylogeographic studies on pines.

Note: The horizontal axis for each chart is time by mutation rate. The time units are different for the different molecular markers, and the average time span for each migration wave shows a different number of generations for each chart. The results are taken from Naydenov et al. (2014, 2017, 2023); YBP – years before present; MY – millions of years before present. Figure 1 also shows a comparison of the asymmetric migration patterns of the mitochondrial INDELs, chloroplast SSRs, and nuclear SSRs for two Mediterranean pines (*P. nigra* and *P. pinaster* Aiton). The red dashed line connecting the red asterisks is the time mark for the Last Glacial Maximum (LGM ≈ 27,000 YBP). The horizontal, colored bar in the middle of each chart represents the geological time grid: white – Holocene epoch (up to 0.0117 MY), light blue – Pleistocene epoch (between 0.0117 and 1.806 MY), yellow – Pliocene epoch (between 1.806 and 5.332 MY), and orange – Miocene epoch (≥ 5.332 MY). The pine seed radiation molecular markers, INDEL (top) and

nuSSR (bottom), show a low average migration intensity of approximately 10,000 individuals per migration wave, but the pollen radiation marker, cpSSR (middle), shows a high average migration intensity of approximately 50,000 individuals per migration wave. The three molecular markers used show different phylogenetic memories: that of mt-INDEL is the longest, followed by cpSSR, and that of nuSSR is the shortest.

than chromosomes of the same size due to the plastids' double membranes in plants and animals (Toth et al. 2000, Morgante et al. 2002). The principal plastids are mitochondria and chloroplasts for plants and mitochondria and ribosomes for animals. In plant and animal genetics, cell plastids are different organelles in the cytoplasm and transfer genetic material, which is known as cytoplasmic inheritance. Some of the genetic material is strictly maternal or paternal, but species with mixed sex exist, too (Reboud and Zeyl 1994, Xu 2005, Munasinghe and Agren 2023, Sakamoto and Takami 2023).

(4) Among species from different genera up to a high taxonomic level, specific fluctuations in the mutation rate of the same molecular marker are different (Drake 1991, Drake et al. 1998, Lynch 2010, Sanjuan et al. 2010, Sun and Lind 2022, Lynch et al. 2023). These results are due to differences in the genome size and probably karyological specificity (chromosome number fluctuations and size difference). Generally, for the same origins and DNA fragment size, species with large genome sizes have lower mutation rates than species with small genomes. The molecular marker ranking is important for computer simulations for the aforementioned markers, for high-performing wide genome analysis (WGA) developed in the past 20 years, and for future developments.

Time mark molecular clock fixing

For time mark molecular clock calibration, it is necessary to match the aforementioned universal principles with a time mark, such as climatic fluctuations, tectonic history maps, or paleo-records (Rogers and Harpending 1992; Sanderson 1997, 2003; Ho et al. 2015; O'Reilly et al. 2015; Donoghue and Yang 2016). Time mark molecular clock fixing is closely related to the molecular marker ranking. For example, it is impossible for molecular markers with a high molecular mass difference (bp difference) and/or different origins to refer to the same historical period. In this context, postglacial migration routes with molecular markers with a high mutation rate difference (i.e., the mutation class, order, and origin [MCOO] profile difference) cannot be described. It is also important to mention that paleo-records must be used with precaution (Schoonmaker and Foster 1991, Carrion and Scott 1999, Charles et al. 2003, Rohde and Muller 2005), as false-positive paleo-botanical artifacts could be pollen (Naydenov et al. 2014, 2017, 2023).

Technically, molecular clock theory is well developed, and robust computer programs, such as LAMARC (Kuhner 2006) and MIGRATE (Beerli 1998), exist for phylogeographic simulations. These programs offer numerous hypothetical scenarios, often including population sizes, migration rates, growth rates, recombination rates, bottlenecks, and isolation-level fluctuations. Multiple rigorous publications involving paleo-map projects are available for tectonic history (Stampfli and Borel 2002; Stampfli et al. 2002a, 2002b; Blakey 2008). We used those developed by Dr. Ron Blakey from Colorado Plateau Geosystems (USA; Blakey 2008). For

molecular clock calibration for phylogenetic analysis, the BEAUti and BEAST 1.7 (Drummond et al. 2012) and CladeDate (Claramunt 2022) programs are used more often.

Example: Forest trees and molecular markers

Over the past 30 years, many studies have discussed molecular-based phylogeography using numerous molecular markers and the postglacial colonization routes of forest trees through the temperate zones to the boreal and tundra regions (Jaramillo-Correa et al. 2004, Godbout et al. 2005, Naydenov et al. 2007). The main question is whether all molecular markers (SNPs, SSRs, and large INDELs) can describe postglacial migration. The simple answer is no. In this paper, we showed that, first, the mutation rate significantly differs between molecular markers by locus/allele length-dependent mutation processes and that, second, the DNA mutation process in plastids versus chromosomes differs. For example, in the *Pinus* genus, the mitochondrial genome is transmitted maternally, while that of chloroplasts is transmitted paternally, that is, by pollen (Neale and Sederoff 1989). In addition, based on our studies on the *Pinus* genus, we considered maternally inherited molecular markers to be more conservative than paternal ones of the same size (Naydenov et al. 2014, 2017, 2023). Various interacting factors directly impact the mutation rate μ. We abbreviated the molecular marker ranking as MCOO. In short, the molecular mutation–rated difference carries a different phylogenetic memory; that is, for different contrasting ranges of molecular markers, it is impossible to describe the same historical migration pattern and phylogeographic structures.

For molecular clock calibration for forest tree phylogeography and phylogenetic analysis, the time mark we must assume is that species have a large genome size and reproductive age. Conifers have a genome that is more than three times that of humans, and the reproductive age of conifers starts from 20 years and reaches its maximum effect at 50–60 years for *P. nigra* and *P. sylvestris* (Dobrinov et al. 1982, Alexandrov et al. 1988). Based on the mutation rate, the time marks to fix the molecular clock could be well-known global climate change records and tectonic history events. For example, in some places in the world, the tectonic history is more dynamic and has been well studied, such as the Mediterranean basin, which allows researchers to try using different time marks to fix the molecular clock back in time. For coniferous species, the most interesting geological period is between 0.01 and 8 million years (Fig. 1). For the *Pinus* genus, we proposed that mitochondrial SNPs (mtSNPs) and nuclear SSRs (nuSSRs) should be used for the Pleistocene epoch, chloroplast SSRs (cpSSRs) for the Pliocene epoch, and mitochondrial INDELs (mtINDELs) for the late Miocene epoch (Burban and Petit 2003; Naydenov et al. 2014, 2017, 2023). For places where it is difficult to find a sufficiently credible time mark, it is possible to use a time mark comparable to those proposed here for coniferous species with similar genome sizes. In the future, an open access data set must be created with the mutation rates for each molecular marker by species and genus.

Conclusion

The mutation rate ranking, and the principles cited before reflecting that different

genome elements have different speeds of evolution and phylogeographic memory, which is important to consider for phylogeographic research. Phylogenetic and phylogeographic memories are valuable for research in various fields, including plants, animals, and humans. For example, understanding the phylogeography of forest trees helps identify historical mutation events and their connections to population migration dynamics, climate change, and the conservation and management of biodiversity. In agriculture and livestock, these insights can determine the origin and period of domestication and guide conservation strategies for important genetic lines, which are crucial for a sustainable economy. In personalized medicine for humans, phylogeographic memory is important for identifying diseases, understanding resistance and parasite vectors, tracing evolutionary timelines, and analyzing their relationships with geographic distribution and social factors. This knowledge will aid in supporting future health programs and developing new medicines.

Acknowledgements

We would like to thank Irena M. Naydenova for her technical assistance. The work of authors (T. Gurov and S. Ivanovska) was carried out using the infrastructure purchased under the National Roadmap for RI, financially coordinated by the MES of the Republic of Bulgaria (Grant No. DO1-325/01.12.2023). The work of author S.-M. Gurova was partly supported by the National Geoformation Centre (part of the National Roadmap of RIs) under grant No. D01-321/30.11.2023.

Declarations

Conflict of interest: None declared.

References

Al Ghadeer H.A., Al Kishi N., Algurini K.H., Albesher A.B., AlGhadeer M.R., Alsalman A.A., Bubshait A.A., Alkishi B.M. 2022. Partial Molar Pregnancy with Normal Karyotype. Cureus 14(10), e30934. doi: 10.7759/cureus.30934

Alexandrov A., Rafailov G., Nedelin G., Tsanov K., Bogdanov B., Spasov S. 1988. Coniferous forest in Bulgaria. Zemizdat, Sofia, Bulgaria. 152 p. (in Bulgarian).

Bachtrog D., Agis M., Imhof M., Schlötterer C. 2000. Microsatellite variability differs between dinucleotide repeat motifs – evidence from *Drosophila melanogaster*. Molecular Biology and Evolution 17(9): 1277–1285. https://doi.org/10.1093/oxfordjournals.molbev.a026411

Beerli P. 1998. Estimation of migration rates and population sizes in geographically structured populations. NATO ASI Series a Life Sciences 306: 39–54.

Berger J., Suzuki T., Senti K.-A., Stubbs J., Schaffner G., Dickson B.J. 2001. Genetic mapping with SNP markers in *Drosophila*. Nature Genetics 29: 475–481. https://doi.org/10.1038/ng773

Bergeron L.A., Besenbacher S., Zheng J., Li P., Bertelsen M.F., Quintard B., Hoffman J.I., Li Z., Leger J.St., Shao C., Stiller J., Gilbert M.T.P., Schierup M.H., Zhang G. 2023. Evolution of the germline mutation rate across vertebrates. Nature 615: 285–291. https://doi.org/10.1038/s41586-023-05752-y

Bhargava A., Fuentes F.F. 2010. Mutational dynamics of microsatellites. Molecular Biotechnology 44(3): 250–266. doi: 10.1007/s12033-009-9230-4

Blakey R.C. 2008. Gondwana paleogeography from assembly to breakup – A 500 m.y.

odyssey. Overview. In: Fielding C.R., Frank T.D., and Isbell J.L. (Eds), Resolving the Late Paleozoic Ice Age in Time and Space. Geological Society of America. Special Paper 441: 1–28.

Brinkmann B., Klintschar M., Neuhuber F., Hühne J., Rolf B. 1998. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. American Journal of Human Genetics 62(6): 1408–1415. doi: 10.1086/301869

Buettner V.L., Hill K.A., Halangoda A., Sommer S.S. 1999. Tandem-base mutations occur in mouse liver and adipose tissue preferentially as G:C to T:A transversions and accumulate with age. Environmental and Molecular Mutagenesis 33(4): 320–324. doi: 10.1002/(sici)1098-2280(1999)33:4<320::aid-em9>3.0.co;2-s

Burban C., Petit R.J. 2003. Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. Molecular Ecology 12(6): 1487–1495. doi: 10.1046/j.1365-294x.2003.01817.x

Calabrese P., Durrett R. 2003. Dinucleotide Repeats in the Drosophila and Human Genomes Have Complex, Length-Dependent Mutation Processes. Molecular Biology and Evolution 20(5): 715–725. https://doi.org/10.1093/molbev/msg084

Calabrese P.P., Durrett R.T., Aquadro C.F. 2001. Dynamics of microsatellite divergence and proportional slippage/point mutation models. Genetics 159(2): 839–852. https://doi.org/10.1093/genetics/159.2.839

Carrion J.S., Scott L. 1999. The challenge of pollen analysis in paleoenvironmental studies of hominid beds: the record from Sterkfontein caves. Journal of Human Evolution 36(4): 401–408. https://doi.org/10.1006/jhev.1998.0276

Chakraborty R., Kimmel M., Stivers D.N., Deka R. 1997. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. PNAS 94(3): 1041–1046. https://doi.org/10.1073/pnas.94.3.1041

Charles H.W., Osterloff P.L., Mohiuddin U. 2003. Fragments of the Earliest Land Plants. Nature 42(6955): 282–285. https://doi.org/10.1038/nature01884

Christenhusz M.J.M., Bastiaanse H.; Royal Botanic Gardens Kew Genome Acquisition Lab; Plant Genome Sizing collective; Darwin Tree of Life Barcoding collective; Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team; Wellcome Sanger Institute Scientific Operations: Sequencing Operations; Wellcome Sanger Institute Tree of Life Core Informatics team; Tree of Life Core Informatics collective; Darwin Tree of Life Consortium. 2024. The genome sequence of black poplar, *Populus nigra* subsp. *betulifolia* L., 1753 (Salicaceae). Wellcome Open Research 9, 228. 16 p. https://doi.org/10.12688/wellcomeopenres.21300.1

Claramunt S. 2022. CladeDate: Calibration information generator for divergence time estimation. Methods in Ecology and Evolution 13(11): 2331–2338. https://doi.org/10.1111/2041-210X.13977

De Franciscis P., Schiattarella A., Labriola D., Tammaro C., Messalli E.M., La Mantia E., Montella M., Torella M. 2019. A partial molar pregnancy associated with a fetus with intrauterine growth restriction delivered at 31 weeks: a case report. Journal of Medical Case Reports 13, 204. 5 p. https://doi.org/ 10.1186/s13256-019-2150-4

Dieringer D., Schlötterer C. 2003. Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. Genome Research 13(10): 2242–2251. http://www.genome.org/cgi/doi/10.1101/gr.1416703

Dobrinov I., Doykov G., Gagov V. 1982. Forest genetic pool in Bulgaria. Zemizdat, Sofia, Bulgaria. 258 p. (in Bulgarian).

Donoghue P.C., Yang Z. 2016. The evolution of methods for establishing evolutionary timescales. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 371(1699), 20160020. https://doi.org/10.1098/rstb.2016.0020

Drake J.W. 1991. A constant rate of spontaneous mutation in DNA based microbes.

PNAS 88(16): 7160–7164. https://doi.org/10.1073/pnas.88.16.7160

Drake J.W., Charlesworth B., Charlesworth D., Crow J.F. 1998. Rates of spontaneous mutation. Genetics 148(4): 1667–1686. https://doi.org/10.1093/genetics/148.4.1667

Drummond A.J., Suchard M.A., Xie D., Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29(8): 1969–1973. https://doi.org/10.1093/molbev/mss075

Ebert P., Audano P.A., Zhu Q., Rodriguez-Martin B., Porubsky D. et al. 2021. Haplotype-resolved diverse human genomes and integrated analysis of structural variation. Science 372, 6537. 13 p. https://doi.org/10.1126/science.abf7117

Ellegran H., Fridolfsson A.-K. 1997. Male driven evolution of DNA sequences in birds. Nature Genetics 17: 182–184. https://doi.org/10.1038/ng1097-182

Galindo C.L., McIver L.J., McCormick J.F., Skinner M.A., Xie Y., Gelhausen R.A., Ng K., Kumar N.M., Garner H.R. 2009. Global Microsatellite Content Distinguishes Humans, Primates, Animals, and Plants. Molecular Biology and Evolution 26(12): 2809–2819. https://doi.org/10.1093/molbev/msp192

Garza J.C., Slatkin M., Freimer N.B. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology and Evolution 12(4): 594–603. https://doi.org/10.1093/oxfordjournals.molbev.a040239

Godbout J., Jaramillo-Correa J.P., Beaulieu J., Bousquet J. 2005. A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. Molecular Ecology 14(11): 3497–3512. https://doi.org/10.1111/j.1365-294X.2005.02674.x

Goldstein D.B., Clark A.G. 1995. Microsatellite variation in North American populations of *Drosophila melanogaster*. Nucleic Acids Research 23(19): 3882–3886. https://doi.org/10.1093/nar/23.19.3882

Gonzalez-Ibeas D., Martinez-Garcia P.J., Famula R.A., Delfino-Mix A., Stevens K.A., Loopstra C.A., Langley C.H., Neale D.B. 2016. Assessing the Gene Content of the Megagenome: Sugar Pine (*Pinus lambertiana*). G3 (Bethesda) 6(12): 3787–3802. https://doi.org/10.1534/g3.116.032805

Greene E., Benjamin R. 2023. Pediatric disorders of stature. In: Halpern-Felsher B. (Ed.), Encyclopedia of Child and Adolescent Health 1: 336–351. https://doi.org/10.1016/B978-0-12-818872-9.00132-1

Grover A., Sharma P.C. 2016. Development and use of molecular markers: past and present. Critical Reviews in Biotechnology 36(2): 290–302. https://doi.org/10.3109/07388551.2014.959891

Hallast P., Ebert P., Loftus M., Yilmaz F., Audano P.A. et al. 2023. Assembly of 43 human Y chromosomes reveals extensive complexity and variation. Nature 621: 355–364. https://doi.org/10.1038/s41586-023-06425-6

Harr B., Schlötterer C. 2000. Long microsatellite alleles in *Drosophila melanogaster* have a downward mutation bias and short persistence times, which cause their genome-wide underrepresentation. Genetics 155(3): 1213–1220. https://doi.org/10.1093/genetics/155.3.1213

Hill K.A., Wang J., Farwell K.D., Sommer S.S. 2003. Spontaneous tandem-base mutations (TBM) show dramatic tissue, age, pattern, and spectrum specificity. Mutation Research 534(1–2): 173–186. https://doi.org/10.1016/S1383-5718(02)00277-2

Ho S.Y., Tong K.J., Foster C.S., Ritchie A.M., Lo N., Crispal M.D. 2015. Biogeographic calibrations for the molecular clock. Biology Letters 11(9), 20150194. https://doi.org/10.1098/rsbl.2015.0194

Jaramillo-Correa J.P., Beaulieu J., Bousquet J. 2004. Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. Molecular Ecology 13(9): 2735–2747. https://doi.org/10.1111/j.1365-294X.2004.02258.x

Korf B.R., Pyeritz R.E., Grody W.W. 2019.

3 – Nature and Frequency of Genetic Disease In: Pyeritz R.E., Korf B.R., Grody W.W. (Eds), Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics (Seventh edition), Academic Press: 47–51. https://doi.org/10.1016/B978-0-12-812537-3.00003-2

Kovach A., Wegrzyn J.L., Parra G., Holt C., Bruening G.E., Loopstra C.A., Hartigan J., Yandell M., Langley C.H., Korf I., Neale D.B. 2010. The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. BMC Genomics 11, 420. 14 p. https://doi.org/10.1186/1471-2164-11-420

Kruglyak S., Durrett R., Schug M.D., Aquadro C.F. 2000. Distribution and Abundance of Microsatellites in the Yeast Genome Can Be Explained by a Balance Between Slippage Events and Point Mutations, Molecular Biology and Evolution 17(8): 1210–1219. https://doi.org/10.1093/oxfordjournals.molbev.a026404

Kruglyak S., Durrett R.T., Schug M.D., Aquadro C.F. 1998. Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. PNAS 95(18): 10774–10778. https://doi.org/10.1073/pnas.95.18.10774

Kuhner M.K. 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22(6): 768–770. https://doi.org/10.1093/bioinformatics/btk051

Lai Y., Sun F. 2003. Microsatellite mutations during the polymerase chain reaction: mean field approximations and their applications. Journal of Theoretical Biology 224(1): 127–137. https://doi.org/10.1016/S0022-5193(03)00155-3

Lynch M. 2010. Evolution of the mutation rate. Trends in Genetics 26(8): 345–352. https://doi.org/10.1016/j.tig.2010.05.003

Lynch M., Ali F., Lin T., Wang Y., Ni J., Long H. 2023. The divergence of mutation rates and spectra across the Tree of Life. EMBO reports 24, 10. https://doi.org/10.15252/embr.202357561

Mahdieh N., Rabbani B. 2013. An overview of mutation detection methods in genetic disorders. Iranian Journal of Pediatrics 23(4): 375–388.

Mahmoud M., Gobet N., Cruz-Dávalos D.I., Mounier N., Dessimoz C., Sedlazeck F.J. 2019. Structural variant calling: the long and threshold of it. Genome Biology 20, 246. 14 p. https://doi.org/10.1186/s13059-019-1828-7

Makova K.D., Pickett B.D., Harris R.S., Hartley G.A., Cechova M. et al. 2024. The complete sequence and comparative analysis of ape sex chromosomes. Nature 630: 401–411. https://doi.org/10.1038/s41586-024-07473-2

McVean G. 2000. Evolutionary genetics: What is driving male mutation? Current Biology 10(22): R834–R835. https://doi.org/10.1016/S0960-9822(00)00787-9

Mills R.E., Luttig C.T., Larkins C.E., Beauchamp A., Tsui C., Pittard W.S., Devine S.E. 2006. An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Research 16(9): 1182–1190. http://www.genome.org/cgi/doi/10.1101/gr.4565806

Morgante M., Hanafey M., Powell W. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genetics 30(2): 194–200. https://doi.org/10.1038/ng822

Munasinghe M., Agren J.A. 2023. When and why are mitochondria paternally inherited? Current Opinion in Genetics & Development 80, 102053. 8 p. https://doi.org/10.1016/j.gde.2023.102053

Naydenov K.D., Alexandrov A., Matevski V., Vasilevski K., Naydenov M.K., Gyuleva V., Carcaillet C., Wahid N., Kamary S. 2014. Range-wide genetic structure of maritime pine predates the last glacial maximum: evidence from nuclear DNA. Hereditas 151(1): 1–13. https://doi.org/10.1111/j.1601-5223.2013.00027.x

Naydenov K.D., Naydenov M., Alexandrov A., Gurov T., Gyuleva V. et al. 2023. Speciation and historical migration pattern interaction: examples from *P. nigra* and *P. sylvestris* phylogeography. Europian Journal of Forest Research 142: 1–26. https://

doi.org/ 10.1007/ s10342-022-01513-0

Naydenov K.D., Naydenov M.K., Alexandrov A., Vasilevski K., Gyuleva V. et al. 2016. Ancient split of major genetic lineages of European Black Pine: evidence from chloroplast DNA. Tree Genetics and Genomes 12, 68. https://doi.org/10.1007/s11295-016-1022-y

Naydenov K.D., Naydenov M.K., Alexandrov A., Vasilevski K., Hinkov G. et al. 2017. Ancient genetic bottleneck and Plio-Pleistocene climatic changes imprinted the phylobiogeography of European Black Pine populations. European Journal of Forest Research 136(5–6): 767–786. https://doi.org/10.1007/s10342-017-1069-9

Naydenov K.D., Senneville S., Beaulieu J., Tremblay F., Bousquet J. 2007. Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. BMC Evolutionary Biology 7, 233. 12 p. https://doi.org/10.1186/1471-2148-7-233

Neale D.B., Sederoff R.R. 1989. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. Theoretical and Applied Genetics 77(2): 212–216. https://doi.org/10.1007/BF00266189

Nystedt B., Street N., Wetterbom A., Zuccolo A., Lin Y.-C. et al. 2013. The Norway spruce genome sequence and conifer genome evolution. Nature 497: 579–584. https://doi.org/10.1038/nature12211

Ohta T., Kimura M. 1973. A model of mutation appropriate to calculate the number of electrophoretically detectable alleles in a genetic population. Genetical Research 22(2): 201–204. doi: 10.1017/s0016672300012994

O'Reilly J.E., Dos Reis M., Donoghue P.C.J. 2015. Dating Tips for Divergence-Time Estimation. Trends in Genetics 31(11): 637–650. https://doi.org/10.1016/j.tig.2015.08.001

Reboud X., Zeyl C. 1994. Organelle inheritance in plants. Heredity 72: 132–140. https://doi.org/10.1038/hdy.1994.19

Rhie A., McCarthy S.A., Fedrigo O., Damas J., Formenti G. et al. 2021. Towards complete and error-free genome assemblies of all vertebrate species. Nature 592: 737–746. https://doi.org/10.1038/s41586-021-03451-0

Rhie A., Nurk S., Cechova M., Hoyt S.J., Taylor D.J. et al. 2023. The complete sequence of a human Y chromosome. Nature 621: 344–354. https://doi.org/10.1038/s41586-023-06457-y

Rogers A.R., Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution 9(3): 552–569. https://doi.org/10.1093/oxfordjournals.molbev.a040727

Rohde R.A., Muller R.A. 2005. Cycles in fossil diversity. Nature 434: 208–210. https://doi.org/10.1038/nature03339

Rose O., Falush D. 1998. A threshold size for microsatellite expansion. Molecular Biology and Evolution 15(5): 613–615. https://doi.org/10.1093/oxfordjournals.molbev.a025964

Ruitberg C.M., Reeder D.J., Butler J.M. 2001. STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acids Research 29(1): 320–322. https://doi.org/10.1093/nar/29.1.320

Sakamoto W., Takami T. 2023. Plastid inheritance revisited: emerging role of organelle DNA degradation in angiosperms. Plant and Cell Physiology 65(4): 484–492. https://doi.org/10.1093/pcp/pcad104

Sanderson M. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution 14(12): 1218–1231. https://doi.org/10.1093/oxfordjournals.molbev.a025731

Sanderson M.J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19(2): 301–302. https://doi.org/10.1093/bioinformatics/19.2.301

Sanjuan R., Nebot M.R., Chirico N., Mansky L.M., Belshaw R. 2010. Viral mutation

rates. Journal of Virology 84(19): 9733–9748. https://doi.org/10.1128/jvi.00694-10

Schlotterer C., Tautz D. 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Research 20(2): 211–215. https://doi.org/10.1093/nar/20.2.211

Schoonmaker P.K., Foster D.R. 1991. Some Implications of Paleoecology for Contemporary Ecology. Botanical Review 57(3): 204–245. https://www.jstor.org/stable/4354170

Schug M.D., MacKay T.F., Aquadro C.F. 1997. Low mutation rates of microsatellite loci in *Drosophila melanogaster*. Nature Genetics 15: 99–102. https://doi.org/10.1038/ng0197-99

Sibly R.M., Meade A., Boxall N., Wilkinson M.J., Corne D.W., Whittaker J.C. 2003. The structure of interrupted human AC microsatel-lites. Molecular Biology and Evolution 20(3): 453–459. https://doi.org/10.1093/molbev/msg056

Sibly R.M., Whittaker J.C., Talbot M. 2001. A maximum-likelihood approach to fitting equilibrium models of microsatellite evolution. Molecular Biology and Evolution 18(3): 413–417. https://doi.org/10.1093/oxfordjournals.molbev.a003817

Stampfli G.M., Borel G.D. 2002. A plate tectonic model for the Paleozoic and Mesozoic constrained by dynamic plate boundaries and restored synthetic oceanic isochrons. Earth and Planetary Science Letters 196(1–2): 17–33. https://doi.org/10.1016/S0012-821X(01)00588-X

Stampfli G.M., Mosar J., Favre P. et al. 2002a. The paleotectonic atlas of the PeriTethyan Domain (electronic resource). European Geophysical Society, CD-ROM.

Stampfli G.M., von Raumer J.F., Borel G.D. 2002b. Palaeozoic evolution of pre-Variscan terranes: From Gondwana to the Variscan collision. In: Martinez-Catalanl R., Hatcher R.D., Arenas R., Diaz Garcia F. (Eds), Variscan Appalachian dynamics: the building of the Upper Paleozoic basement. Geological Society of America, Special Paper 364: 263–280.

Stumpf M.P., Goldstein D.B. 2001. Genealogical and evolutionary inference with the human Y chromosome. Science 291(5509): 1738–1742. doi: 10.1126/science.291.5509.1738

Sturzeneker R., Bevilacqua R.A.U., Haddad LA, Simpson A.J.G., Pena S.D.J. 2000. Microsatellite instability in tumors as a model to study the process of microsatellite mutations. Human Molecular Genetics 9(3): 347–352. https://doi.org/10.1093/hmg/9.3.347

Sun T.A., Lind P.A. 2022. Distribution of mutation rates challenges evolutionary predictability. bioRxiv 12, 518962. 45 p. https://doi.org/10.1101/2022.12.12.518962

Sung W., Ackerman M.S., Miller S.F., Lynch M. 2012. Drift-barrier hypothesis, and mutation-rate evolution. PNAS 109(45): 18488–18492. https://doi.org/10.1073/pnas.1216223109

Thomson J.A., Pilotti V., Stevens P., Ayres K.L., Debenham P.G. 1999. Validation of short tandem repeat analysis for the investigation of cases of disputed paternity. Forensic Science International 100(1–2): 1–16. https://doi.org/10.1016/S0379-0738(98)00199-6

Toth G., Gaspari Z., Jurka J. 2000. Microsatellites in different eukaryotic genomes: Survey and analysis. Genome Research 10(7): 967–981. doi: 10.1101/gr.10.7.967

Verbiest M., Maksimov M., Jin Y., Anisimova M., Gymrek M., Sonay T.B. 2023. Mutation and selection processes regulating short tandem repeats give rise to genetic and phenotypic diversity across species. Journal of Evolutionary Biology 36(2): 321–336. doi: 10.1111/jeb.14106

Weber J.L., Wong C. 1993. Mutation of human short tandem repeats. Human Molecular Genetics 2(8): 1123–1128. https://doi.org/10.1093/hmg/2.8.1123

Wellesley D., Dolk H., Boyd P.A., Greenlees R., Haeusler M., Nelen V., Garne E., Khoshnood B., Doray B., Rissmann A., Mullaney C., Calzolari E., Bakker M., Salvador J., Addor M.C., Draper E., Rankin J., Tucker D. 2012. Rare chromosome abnormalities, prevalence and prenatal diagnosis rates from population-based congenital anomaly regis-

ters in Europe. European Journal of Human Genetics 20(5): 521–526. https://doi.org/10.1038/ejhg.2011.246

Whittaker J.C., Harbord R.M., Boxall N., Mackay I., Dawson G., Sibly R.M. 2003. Likelihood-based estimation of microsatellite mutation rates. Genetics 164(2): 781–787. https://doi.org/10.1093/genetics/164.2.781

Wicks S.R., Yeh R.T., Gish W.R., Waterston R.H., Plasterk R.H.A. 2001. Rapid gene mapping in *Caenorhabditis elegans* using a high-density polymorphism map. Nature Genetics 28: 160–164. https://doi.org/10.1038/88878

Wierdl M., Dominska M., Petes T.D. 1997. Microsatellite instability in yeast: dependence on the length of the microsatellite. Genetics 146(3): 769–779. https://doi.org/10.1093/genetics/146.3.769

Xu J. 2005. The inheritance of organelle genes and genomes: patterns and mechanisms. Genome 48(6): 951–958. https://doi.org/10.1139/g05-082

Xu X., Peng M., Fang Z., Xu X. 2000. The direction of microsatellite mutation is dependent upon allele length. Nature Genetics 24(4): 396–399. https://doi.org/10.1038/74238

Yan Y., Yi G., Sun C., Qu L., Yang N. 2014. Genome-Wide Characterization of Insertion and Deletion Variation in Chicken Using Next Generation Sequencing. PLoS ONE 9(8), e104652. 11 p. https://doi.org/10.1371/journal.pone.0104652

Zuberi S.M. 2013. Chapter 57 – Chromosome disorders associated with epilepsy. In: Dulac O., Lassonde M., Sarnat H.B. (Eds), Handbook of Clinical Neurology. Elsevier 111: 543–548, https://doi.org/10.1016/B978-0-444-52891-9.00057-9