De novo mutations, genetic mosaicism and human disease

Mohiuddin Mohiuddin[1†]*, R. Frank Kooy[2] and Christopher E. Pearson[1,3]*

[1]Program of Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON, Canada, [2]Department of Medical Genetics, University of Antwerp, Edegem, Belgium, [3]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

Mosaicism—the existence of genetically distinct populations of cells in a particular organism—is an important cause of genetic disease. Mosaicism can appear as de novo DNA mutations, epigenetic alterations of DNA, and chromosomal abnormalities. Neurodevelopmental or neuropsychiatric diseases, including autism—often arise by de novo mutations that usually not present in either of the parents. De novo mutations might occur as early as in the parental germline, during embryonic, fetal development, and/or post-natally, through ageing and life. Mutation timing could lead to mutation burden of less than heterozygosity to approaching homozygosity. Developmental timing of somatic mutation attainment will affect the mutation load and distribution throughout the body. In this review, we discuss the timing of de novo mutations, spanning from mutations in the germ lineage (all ages), to post-zygotic, embryonic, fetal, and post-natal events, through aging to death. These factors can determine the tissue specific distribution and load of de novo mutations, which can affect disease. The disease threshold burden of somatic de novo mutations of a particular gene in any tissue will be important to define.

KEYWORDS

de novo mutation, mosaicism, timing of mutation, repeat instability, germline mutation, somatic mutation, autism spectrum disorder, genetic diseases

Mosaicism - From germline to somatic

The term "mosaic" refers to an intricate image or pattern created by craftsmen from small pieces of colored hard material, such as glass, gems, ornamental stones or other precious material. At a distance, the collective image seems as it would in a painting. Only on close assessment, the individual components become distinguishably different. In biology, mosaicism means the presence of genetically different cells within a tissue of single organism (Youssoufian and Pyeritz, 2002; Biesecker and Spinner, 2013). An individual who has developed from a single fertilized egg and has two or more populations of cells with different genotypes is mosaic. In terms of the whole organism, rise of the mosaic phenotype depends on tissue-to-tissue genetic variations that do not conform to Mendelian rules of inheritance. During the last several years, germline and somatic mosaicism have appeared as critical factors that contribute to phenotypic variability in almost every area of biology.
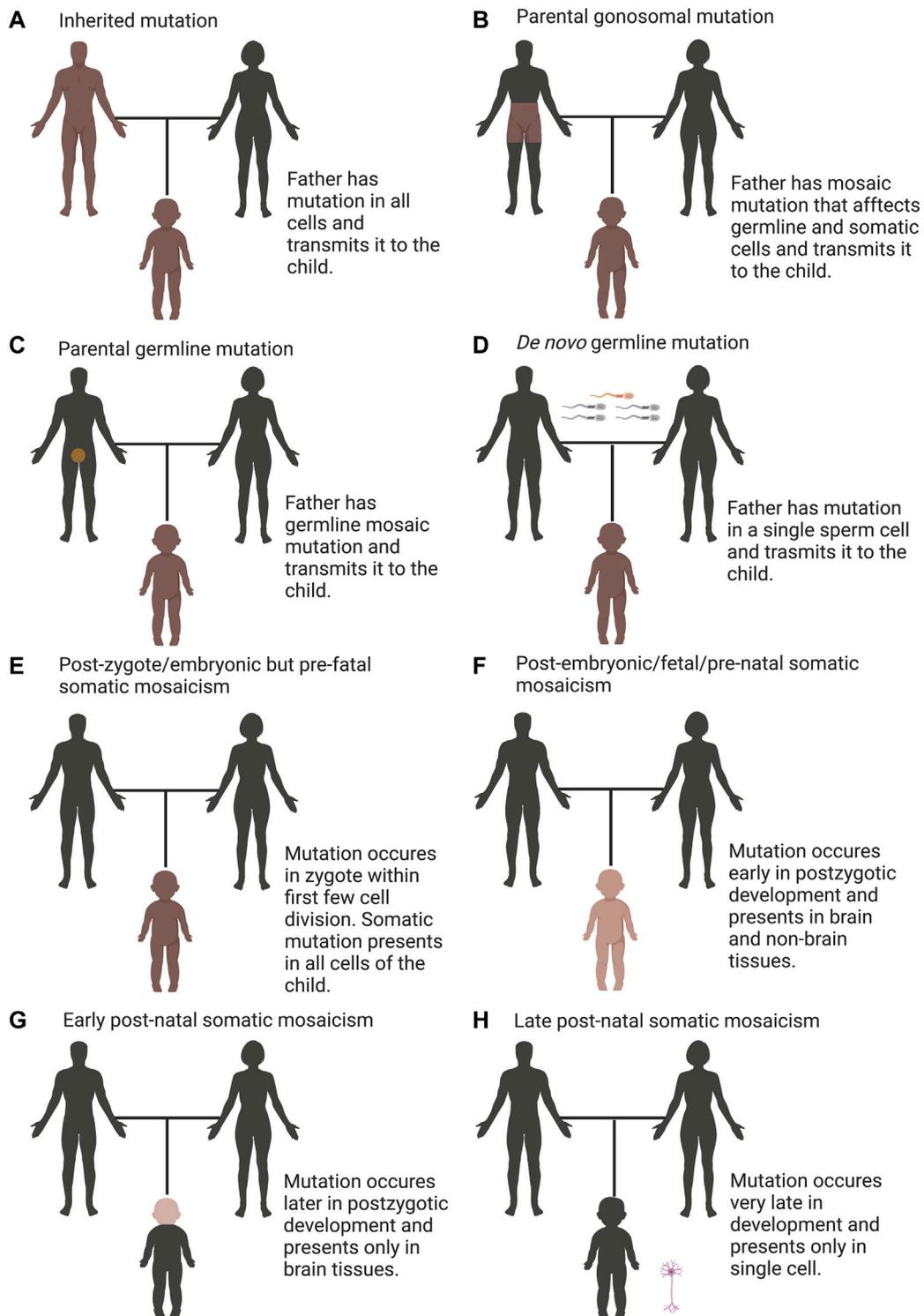
FIGURE 1

Overview of categories of mutations including inherited, *de novo*, and somatic variation. **(A)** Inherited mutations are constantly transmitted through the germline, which is detectable in all tissues of the child and their parent. **(B)** Parental gonosomal mutation is detectable in some tissues of the parents and in all tissues of the child, which is transmitted from a parent with mosaic mutation (combination of somatic and germline mosaicism). **(C)** Parental germline mosaicism is detectable in gametes of the parent and in all tissues of the child. **(D)** *De novo* germline mutation is detectable in all tissues of the child but not detectable in the parent. **(E)** post-zygote/embryonic but pre-fetal somatic mosaicism occurs in zygote within first few cell divisions such that the mutation presents in all cells that contribute to the embryo, is detectable in all tissues of the child but not

(Continued)

FIGURE 1 (Continued)
detectable in the parent. **(F)** post-embryonic/fetal/pre-natal somatic mosaicism, which is present in nonbrain and brain tissues, occurs early in post-zygotic development, is detectable in some tissues of the child but not detectable in the parent. **(G)** early post-natal somatic mosaicism, which is present only in the brain, occurs later in post-zygotic development, is detectable only in brain tissue of the child but not detectable in the parent. **(H)** late post-natal somatic mosaicism, which occurs very late in post-zygotic development, is detectable only in single cell of the child, which requires single-cell sequencing to detect but not detectable in the parent. In all panels, brown denotes the mutation and darker shades designate increasing degree of mosaicism.

Genetic variations that cause neurological, neuropsychiatric and neurodevelopmental diseases are generally considered as either inherited or *de novo* germline mutations. Inherited mutations are present in one or both parents and in all tissues of the affected individual (Figure 1A). Hence, the mutation can be conveniently assayed in any tissue of the affected offspring. Despite the presence of inherited mutations in essentially all cells, they may affect some tissues more than others, depending upon when and where the gene involved exerts its crucial roles.

The mutation rate of single-nucleotide variants in humans to be approximately 1×10^{-8} mutations per generation, giving rise to 45–60 *de novo* mutations (DNMs) per genome (Roach et al., 2010; Kong et al., 2012; Michaelson et al., 2012; Campbell and Eichler, 2013). Mounting evidence shows the importance of DNMs mutations in neuropsychiatric and pediatric disorders (Awadalla et al., 2010; O'Roak et al., 2011; O'Roak et al., 2012; Devlin et al., 2012; Michaelson et al., 2012; O' Roak et al., 2012; Sanders et al., 2012; Veltman and Brunner, 2012). DNMs are not detectable in either parent of an affected offspring. These mutations are typically existent in the sperm or egg of one parent and are not evident in blood taken of those parents, however, once transmitted to the embryo, the mutation can be detectable in any tissue of the affected individual (Figure 1D).

Any genetic variation in the genomes of germinal cells within an individual is termed as germline mosaicism. Germline mosaicism is a driving force behind evolution and a fundamental cause of genetic diseases. A human zygote receives 50% of its genome from the father through the sperm and the other 50% from the mother via the oocyte. In addition to the genomic material passed on from generation to generation, we are born with a little number of novel mutations—that arose either during the formation of the gametes or post-zygotically (Lynch, 2010; Roach et al., 2010). Each individual's genome carries approximately one *de novo* germline mutation in their protein-coding region of the genome that are not evident in their parents' somatic cells (Awadalla et al., 2010; O'Roak et al., 2011; Devlin et al., 2012; Sanders et al., 2012). Such mutations may manifest in individuals with neurodevelopmental and neuropsychiatric conditions such as autism spectrum disorder (ASD) (Awadalla et al., 2010; O'Roak et al., 2011; Devlin et al., 2012; Sanders et al., 2012). These *de novo* germline mutations are probably damaging, suggesting that many of these *de novo* germline mutations are associated with diseases. The contribution of *de novo* mutations to a disease will be enhanced if the mutations have a large negative effect on survival and reproduction (Veltman and Brunner, 2012). Neurodevelopmental or neuropsychiatric diseases, including autism—often arise by *de novo* mutations that are not evident in either of the parents.

De novo mutations can also occur post-zygotically during embryo development, through fetal development (beginning on the 9th week post-fertilization), and post-natally, through to death. These mutations lead to individuals who are mosaic, the presence of genetically distinct cells within a single organism (Youssoufian and Pyeritz, 2002; Biesecker and Spinner, 2013). These mutations are considered *de novo* because they are not evident in the parents or in the zygote of the affected individuals but are more precisely characterized as somatic mutations. Somatic mutations trigger antigenic variation, and have long been associated to cancer, with spans of studies and technological advances supporting and illuminating their role (Philp et al., 2001; Martincorena and Campbell, 2015). The latest progresses in DNA sequencing have shown mosaicism to be more prevalent than previously assumed in humans (Macosko and McCarroll, 2012; Biesecker and Spinner, 2013; Campbell et al., 2015; Forsberg et al., 2017), whether occurring in development or ageing, with the term 'somatic evolutionary genomics' used to explain the study of the accrual of somatic mutations in the body (Frank, 2010). Somatic mutations can indeed be used to reconstruct the developmental cell lineage in an organism (Frank, 2010; Woodworth et al., 2017; Coorens et al., 2021; Fasching et al., 2021; Park et al., 2021; Spencer Chapman et al., 2021). The proof of the involvement of somatic mutations in a wide range of neurodevelopmental and neuropsychiatric disorders is already robust (McConnell et al., 2017). A role of somatic mutations in neurodegenerative diseases has been constantly postulated (Pamphlett, 2004; Frank, 2010; Van Broeckhoven, 2010; Proukakis et al., 2013). Moreover, variable levels of mutations have been reported in somatic human tissues, including the brain, skin, and blood, and in patient-derived induced pluripotent stem cells (Baillie et al., 2011; Abyzov et al., 2012; Evrony et al., 2012; Macosko and McCarroll, 2012; O'Huallachain et al., 2012). It is well-known that germline mosaicism is difficult to evaluate, especially if a maternally originated mutation is suspected, because it is difficult to derive eggs from ovaries besides the very fact of it being a highly invasive procedure. Then, germline mosaicism is

suspected when two or more affected children are born to apparently unaffected parents (Gajecka, 2016). When a mosaicism in the germ line was identified because of the birth of more than one affected child, the mutation was also present in somatic cells in half of the cases (Zlotogora, 1998).

Types of mosaicism

Specific types of mosaicism describe which parts of the body harbor the variant cells and the potential for transmission to offspring. A new classification of genetic mosaicism depending on the distribution and pattern of the mosaicism, the affected tissue, the pathogenicity of the variant, the postzygotic mutational mechanism and the direction of the change (benign to pathogenic vs. pathogenic to benign) were reviewed elsewhere (Martínez-Glez et al., 2020). Apart from somatic mosaicism, which indicates to mutations that occur in somatic cells and are not evident in germ cells (Figures 1E–H), germline mosaicism (also known as gonadal mosaicism) (Figures 1C,D), which indicates to mutations that occur in germ cells and are not evident in somatic cells and gonosomal mosaicism (combination of somatic and germline mosaicism) where mutations evident in a subset of somatic and germ cells (Biesecker and Spinner, 2013; D'Gama and Walsh, 2018) (Figure 1B). Gonosomal mosaicism arises due to two possible mechanisms: the mutation arises in a germ cell that continues to divide. The other path to gonosomal mosacism is that the mutation arises very early in a somatic cell before the separation to germinal cells and hence evident in both germinal and somatic cells (Jansen et al., 1994; Zlotogora, 1998).

Although these categorizations are beneficial in an exceedingly practical sense, we concede that they cannot be decisively assigned due to the restrictions of tissue sampling. It is difficult or impossible to disentangle the differences. As an example, the detection of germline mosaic patient is usually based on the detection of a mutation in various germ cells (usually sperm) and on the lack of the mutation in skin fibroblasts and/or peripheral blood; however, the incidence of the mutation in other somatic cells cannot strictly be excluded.

Disease-associated somatic mutations can only be existent in the affected tissue but absent in other tissues, including the blood of the same patient (Rivière et al., 2012a; Lee et al., 2012; Poduri et al., 2012). The latter obscures the mutation from detection using routine diagnostic profiling. Considering the timing of mutation, somatic mosaicism can be classified as: i) post-zygote/embryonic but pre-fetal somatic mosaicism- where mutation arises within the zygote or during the primary mitotic divisions, such all cells that contribute to the embryo carry the mutation. This mutation is not detectable in the parent but can be identified in all tissues of the child (Figure 1E); ii) post-embryonic/fetal/pre-natal somatic mosaicism where mutation occurs early in development (following the blastula

stage of 64 cells is the beginning of the fetal stage), may be existent in a high percentage of cells across multiple tissues (brain and non-brain tissues) and is not detectable in the parent but identified relatively easily in some tissues of the child (Figure 1F); iii) early post-natal somatic mosaicism- where mutation arises late in development, may be evident in a proportion of cells in only one tissue (brain) and can be identified only if tested, which is not detectable in the parent (Figure 1G); and iv) late post-natal somatic mosaicism- where mutation occurs in a post-mitotic cell like a neuron, is not evident in the parent and can be identified only in that neuron of the child, which requires ddPCR or single-cell sequencing (D'Gama and Walsh, 2018) (Figure 1H). Subsequent mosaicism can arise during infancy, childhood, teen, young adult, adult, and aged adult hood.

The discrepancy between germline and somatic mosaicism is based on the finding of genetically different populations of cells in the germ-line and somatic tissues, respectively (Hall, 1988; Youssoufian, 1996). Mosaicism that arises during embryogenesis can be coexist as both germ-line and somatic mosaicism in the same individual. Coexistence depends on the specific cell affected and the timing of the development of the mosaicism-inducing event. Reckoning on various factors, like the gene involved, the location in the gene (Monroe et al., 2022; Zhang, 2022) and/or the degree and tissue location of mosaicism, the carrier of a somatic and germline mosaicism may be asymptomatic or may evident with various symptoms of the disease. In most cases, symptoms evident in the mosaicism carrier were milder than that present in the offspring (Zlotogora, 1998).

In germ cells, *de novo* mutations will almost always present as mosaicism if it occurs before meiosis, and, depending on the timing, may be evident in a very few to up to almost half of the germ cells. Human families are usually small. So, it is plausible that, in most of the cases in which a mutation is transmitted to more than one child, the mutation is evident in a significant number of germ cells and therefore must have occurred very early after a few divisions of the committed germ cells (Zlotogora, 1998). Whereas mutations during post-fertilization cell division and throughout life can contribute to somatic mutations, transmitted instability must involve germline mutations.

It has been anticipated that *de novo* germline mutations derive from errors in DNA replication during gametogenesis, specifically in sperm cells and their precursors. However, *de novo* mutations during spermatogenesis can also arise due to ineffective repair of spontaneous DNA lesions, as constant proliferation and short periods between cell divisions can translate into there being less time to repair these lesions (Goriely and Wilkie, 2012; Gao et al., 2016). Likewise, in oogenesis, ineffective repair mechanisms coupled to spontaneous DNA mutations might play a prominent role (Gao et al., 2016). Therefore, while the *de novo* mutation rate could be a reflection of the replication error rate and also the number of mitoses a cell has experienced, this number is additionally influenced by the quantity of time between

mitoses and therefore the proficiency of the DNA repair (Gao et al., 2016).

The origin of mosaicism

The origination of genetically distinct cells from a single zygote necessitates *de novo* post-zygotic mutational events (Veltman and Brunner, 2012) as the cause of mosaicism, and these mutations can result in sporadic disease. *De novo* mutational events can also arise pre-zygotically; where mosaic parent, usually unaffected, but the mutation might be inherited in the zygote and possibly in all cells of the developing offspring also to result in the creation of a *de novo* disease phenotype (Veltman and Brunner, 2012).

Mosaicism can also arise from either meiotic or mitotic circumstances. If a meiotic error has occurred and revised at the cleavage stage, it will cause mosaicism. The meiotic errors are usually more damaging because of the initial onset of the abnormality, which permits the mosaicism to dominate. Nevertheless, mitotic errors may be as extreme as meiotic errors depending on their time of occurrence. Mosaicism could be present through all embryonic tissues if a mitotic error occurs in the first one or two divisions. However, if the mitotic errors occur further along in development, the mosaicism could be tissue specific.

Genetic variation between individuals can occur due to chromosome segregation or DNA replication errors, leading to CNVs, chromosome aneuploidy, genomic rearrangements, repeat expansions, single-nucleotide variation, and microsatellite instabilities. These mutational processes can happen at any stage of development; in differentiating cells, stem cells, and in terminally differentiated somatic cells, leading to mosaicism (Forsberg et al., 2017), and have now been reported in normal brain (McConnell et al., 2017). DNA sequence characteristics (e.g., CpG dinucleotides) and genomic architectural features (e.g., inverted and direct repeats) can accelerate genome instability and susceptibility to mutation. Mosaicism can also occur during either repair of DNA damage (single-strand breaks (SSBs) and double-strand breaks (DSBs), or DNA replication (Lindahl and Wood, 1999; Hoeijmakers, 2001; McKinnon, 2009; Rulten and Caldecott, 2013). DNA damage may cause transient chemical lesions, but inaccurate repair of these damages cause somatic mutation (Vijg and Suh, 2013). Furthermore, DNA damages caused by exogenous sources, such as tobacco smoke and other carcinogens, may cause somatic mosaicism in targeted tissues. Accumulative exposures to exogenous mutagens, together with ongoing development of the organism, tissue regeneration, and cell proliferation and renewal, caused accumulation of mutations with age. Moreover, previous study proposed that nonallelic homologous recombination–predicted inversions, such as structural variations, are mosaic and seem to accumulate as the individual ages (Flores et al., 2007). L1 transposition

during embryogenesis can caused mosaicism (Kano et al., 2009). Several unbalanced translocations seem to originate post-zygotically, apparently appearing *de novo* during embryogenesis in a manner that is based on homologous interspersed transposable elements as substrates (Robberecht et al., 2013), and recombination-restarted replication forks contribute to other post-zygotic mutational events (Mizuno et al., 2013).

Cellular and DNA metabolism through human gametogenesis

While mutations during post-fertilization cell divisions and throughout life can contribute to somatic instability (Figure 2A), transmitted instability must encompass germline mutations (Figure 2B). Numerous genetic diseases are inherited primarily through either maternal or paternal transmissions. Each parent-of-origin mutation bias is driven by complex processes specific to sperm or oocyte development. Mosaic mutation could occur in proliferating, arrested meiotic cells or within the relatively dormant haploid germ cells (Figure 2). Knowledge about the timing of mutation events during germ-cell development in patients is critical to understanding particular important mutation and its transmission in affected families. It will also suggest which DNA metabolic process facilitates instability. During gametogenesis, germline mosaicism might arise due to highly specialized DNA metabolic activities that engage stage-specific expression of repair, replication or recombination factors, replication programs, expression profiles and epigenetic modifications. Previous studies demonstrated that many repair, replication or recombination genes show reformed or particular expression during gametogenesis (Drost and Lee, 1995; Kamel et al., 1997; Richardson et al., 2000; Baarends et al., 2001; Velasco-Miguel et al., 2003). Replication mediated mutation can be complex and could vary between different tissues and/or developmental stages. Mutation events must appear at some point between segregation of the primordial germ cells, development to puberty, or during the life-long post-pubertal spermatogonial stem-cell divisions (Figure 2). The large number of life-long spermatogonial cell divisions might push the paternal mutation bias. As mutations are evident in human pre-meiotic diploid cells, the early mutation event must involve DNA replication fork errors, genome maintenance repair or replication-associated repair. Additional mutations might involve post-meiotic genome maintenance or recombination events (McKinnon, 2009; Rulten and Caldecott, 2013; Vijg and Suh, 2013).

The genome of the haploid spermatids is re-packaged with protamine at the completion of spermatogonial proliferation and meiosis, thereby entering a comparatively inactive or 'sleeping' state, and ultimately becoming spermatozoa. After meiosis, an additional phase of expansion might happen and probably
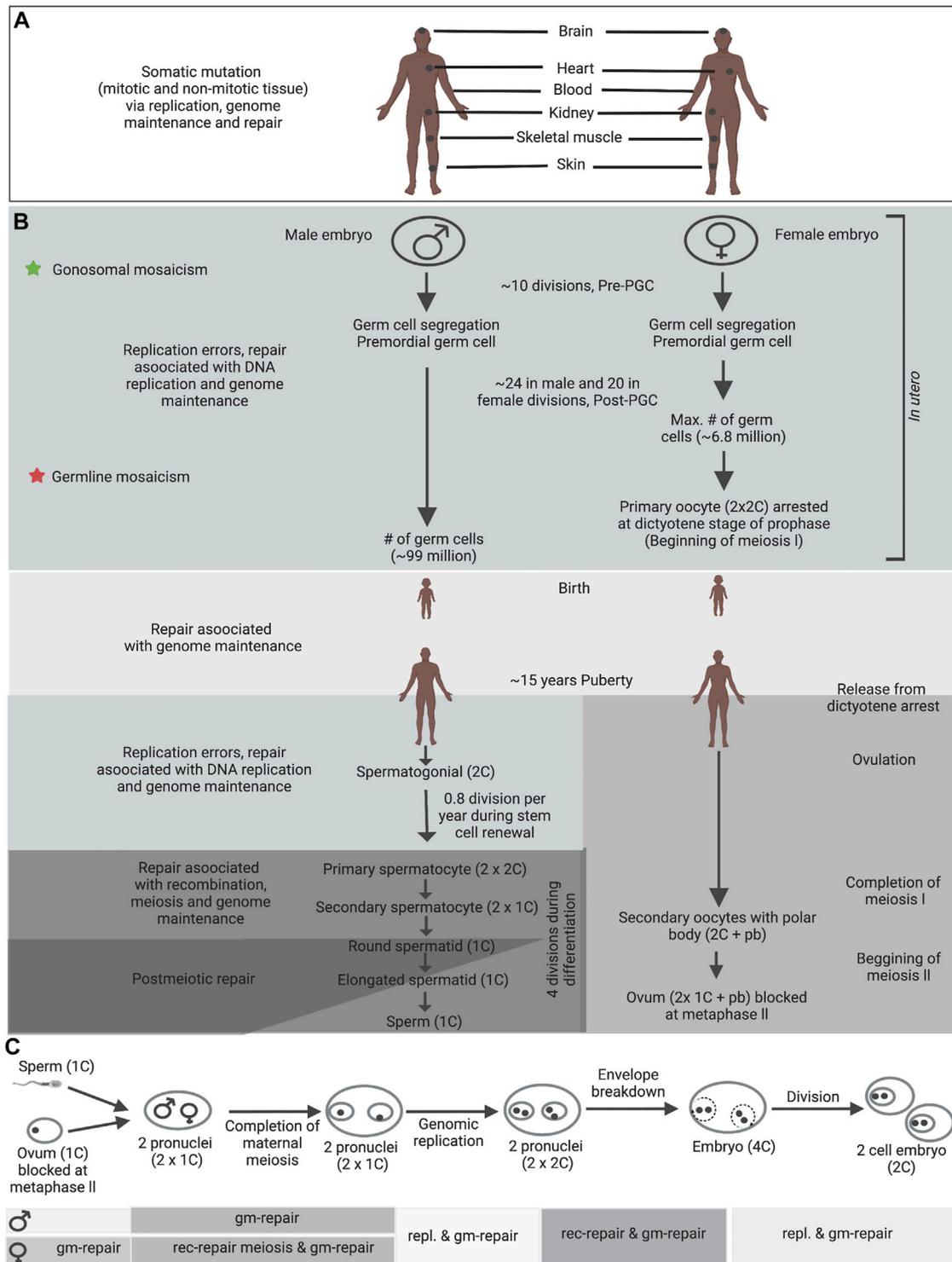
FIGURE 2
Germline and somatic DNA metabolic processes in human. **(A)** Somatic mosaicism in patients with a genetic disease can exhibit both pre-natal and post-natal tissue specific mosaicism. **(B)** Depending on the timing of mutations during embryonic development, different types of germline mosaicism can arise; star signs indicate different stages at which mutations can arise and the consequential types of mosaicism. Germline mosaic variants, which were apparent within the parents' blood, were possibly established before mesoderm tissue separation from PGCs within the parents (green stars). One potential explanation for mosaic mutations that are only shared by siblings—were not apparent in the parents' blood—is that the mutations arose after separation of PGCs from mesoderm in the mosaic parents (red stars). **(C)** Fertilization and also the steps leading to the
(Continued)

through to elongating spermatids. Spermatogonia and primordial germ cells undertake genome replication and are subject to DNA replication errors, genome maintenance repair, and repair related to replication (Figure 2). The mitotically inactive post-natal pre-pubertal stages undergo genome maintenance repair. Later stages are subject to genome maintenance repair, meiotic recombination and, ultimately, chromatin decondensation (coincident with dormancy) (Figure 2C).

Although the timing has not been thoroughly established, it is assumed that both male and female germ-cell segregation is occur between day 5–12 post-fertilization. Primordial germ cells (PGCs) emerge from the epiblast, which requires ~10 cell divisions after fertilization (Campbell et al., 2014b). Notably, initial few cell divisions of embryogenesis is relatively mutagenic, leading to germline and somatic mosaic mutations in mice (Lindsay et al., 2019) and humans (Huang et al., 2014; Acuna-Hidalgo et al., 2015; Rahbari et al., 2016a; Ju et al., 2017; Lim et al., 2017), as well as to mutations that are discordant between monozygotic twins (Dal et al., 2014; Jónsson et al., 2017). During first few cell divisions, male and female embryos seem to have similar mutation rate (according to models estimating the mutation rate during gametogenesis, approximately 0.2–0.6 mutations per haploid genome per cell division) (Rahbari et al., 2016a). After their specification, PGCs expand to form the complete population of primary oocytes and the pool of spermatogonial stem cells in female and male embryos, respectively (Campbell et al., 2014b; Rahbari et al., 2016a). According to computational modeling, the mutation rate during the expansion of PGCs to oogonia or spermatogonia is comparable in both sexes, with approximately 0.5–0.7 mutations per haploid genome per cell division (Rahbari et al., 2016a). However, spermatogonial stem cells undergo post-pubertal spermatogonial cell divisions throughout life (23 cell division per year during stem cell renewal), maintaining the spermatogonial stem cell pool while generating differentiated spermatogonial cells which produce sperm cells through an extra round of mitosis followed by meiosis (Crow, 2000). By disparity, oogenic meiosis initiates *in utero* and is halted in dictyotene for years, re-starting only minutes before ovulation, after which meiosis I is complete. Meiosis II then starts after the completion of meiosis I and is not completed until fertilization. Thus, oocytes undertake only one extra round of DNA replication in their evolution, after the completion of PGC

expansion in embryogenesis, to a mature ovum. On the other hand, spermatogonial cells can experience hundreds of rounds of cell division and DNA replication before their maturation to sperm cells. More than 75% of all germline *de novo* point mutations appear on the paternal allele (Francioli et al., 2015; Rahbari et al., 2016a). Both at the population level and within the same family, advanced paternal age at conception has been established as the major factor that linked to the increase in the number of *de novo* mutations in the offspring (Kong et al., 2012; Rahbari et al., 2016a; Goldmann et al., 2016). Male germ cells continue to divide throughout life, which allows the gradual accumulation of mutations due to DNA replication errors and repair defect of non-replicative DNA damage between cell divisions (Gao et al., 2016). Moreover, the proficiency of endogenous defense systems against radical oxygen species and of DNA repair mechanisms might also weakening with age (Paul and Robaire, 2013; Goriely, 2016). In contrast to post-meiotic male germ cells, DNA recombination and repair activities are linked with arrested and/or resting oocytes (Baarends et al., 2001; Nouspikel and Hanawalt, 2002).

Germ cell specification process occurs in mammals over the time of gastrulation. Being separated from nonreproductive cell lineage cells, the reproductive lineage cells are able to form germ lineage cells. The reproductive lineage originates from a fertilized egg located in the mother's body. For the preparation of the future body, cell propagation continues and an epiblast is created (De Felici, 2013; Behringer et al., 2014). Following invagination of the epiblast, several originated cells find a place in the wall of the yolk sac and successively migrate as primordial germ cells (PGCs) to the gonadal ridge. Somatic lineages are also differentiated around the same time, and also have their origins within the epiblast (Figure 3, adapted from Sakumi (2019). Non-reproductive lineage cell mutations, that arise after the stage of divergence from germ lineage cells, will be restricted to that individual. These are designated as somatic cell mutations and are unable to transmit to the offspring. On the other hand, mutations that occur in the reproductive lineage from a fertilized egg to gametes can be defined as reproductive lineage cell mutations. Such mutations can transmit to the next generation, but do not exist in the parents, are often called germline (*de novo*) mutations.

Reproductive lineage cell mutations are classified into two groups: pre-germ cell-stage mutations and germ-cell-stage mutations (Sakumi, 2019). The latter group arise during or
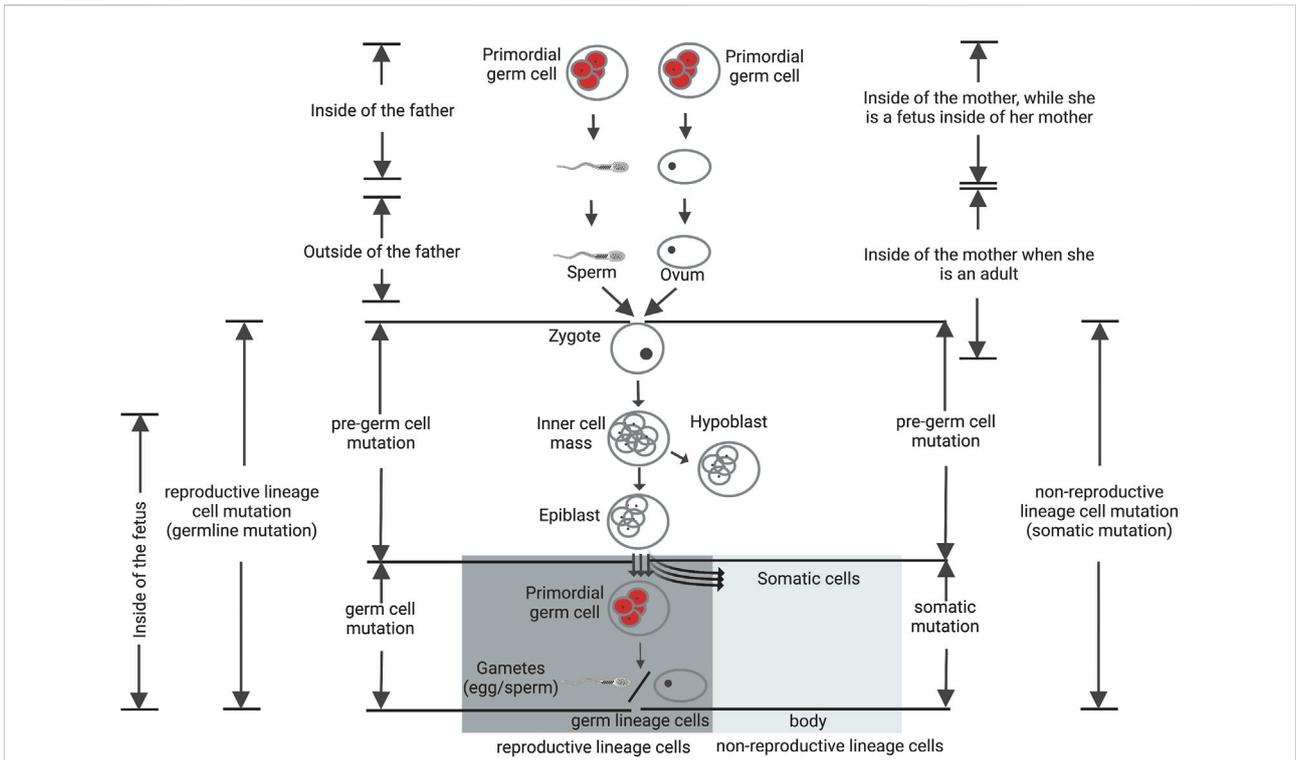
FIGURE 3
Reproductive and non-reproductive lineage cells and *de novo* mutations. Mutations that occur in the reproductive lineage from a fertilized egg to gametes can be defined as reproductive lineage cell mutations and categorized into two groups: pre-germ-cell-stage mutations (gonosomal mutation) and germ-cell-stage mutations. The germ-cell-stage mutations arise in the primordial germ cells (pGCs) and their offspring after the point of divergence from somatic cells, usually being designated as germline mutations. Gonosomal mutations arise before this point of divergence. These mutations can exist in both somatic cells and gametes at the same time. The mutations that arise in the non-reproductive lineage cells, after the stage of divergence from germ lineage cells, are designated as somatic cell mutations and are unable to transmit to the offspring. Figure adapted from Sakumi (2019).

after PGCs specification, may be absent from somatic cells but present in multiple resulting gametes (Biesecker and Spinner, 2013; Campbell et al., 2014b, 2015; Acuna-Hidalgo et al., 2015; Rahbari et al., 2016b; Tang et al., 2016; Jónsson et al., 2018). These variants can consequently be existing in more than one offspring as evident *de novo* mutations. In contrast, the former group of mutations arise before PGCs specification, but likely following zygote fertilization. Mutations that arise prior to PGCs specification can be exist in both somatic cells and gametes at the same time, characterize as "gonosomal" variants that possibly aroused early during the early post-zygotic development of second-generation individuals (Campbell et al., 2014b, 2014a, 2015; Besenbacher et al., 2015; Rahbari et al., 2016b; Jónsson et al., 2018). Overall, gonosomal mutations, that aroused during the early cell divisions of the offspring, rather than in a single parental gamete, consist of approximately 10% of candidate *de novo* germline mutations (Sasani et al., 2019). Therefore, to define the biological significance of *de novo* mutations in sexual reproduction of higher multicellular organisms such as

mammals, it is essential to define when and where these mutations originated.

Selfish spermatogonial selection

The fitness of the spermatogonia themselves can be affected by spermatogonial mutations. Some mutations in critical genes might lead to the eradication of the spermatogonial lineage, while other mutations can result in selective advantages and lead to clonal expansion, the outgrowth of the specific cell lineage. The later behaves selfishly and are therefore denoted as "selfish mutations" (Goriely et al., 2009). Due to the expansion of mutated spermatogonia that happens in the testes of healthy men during aging (Meng et al., 2000; Goriely et al., 2009; Giannoulatou et al., 2017), the quantity of mutated sperm raises continuously at higher rates than genome-wide paternal age effect. If these mutations are passed on to offspring, most of the cases they will trigger developmental disorders. The
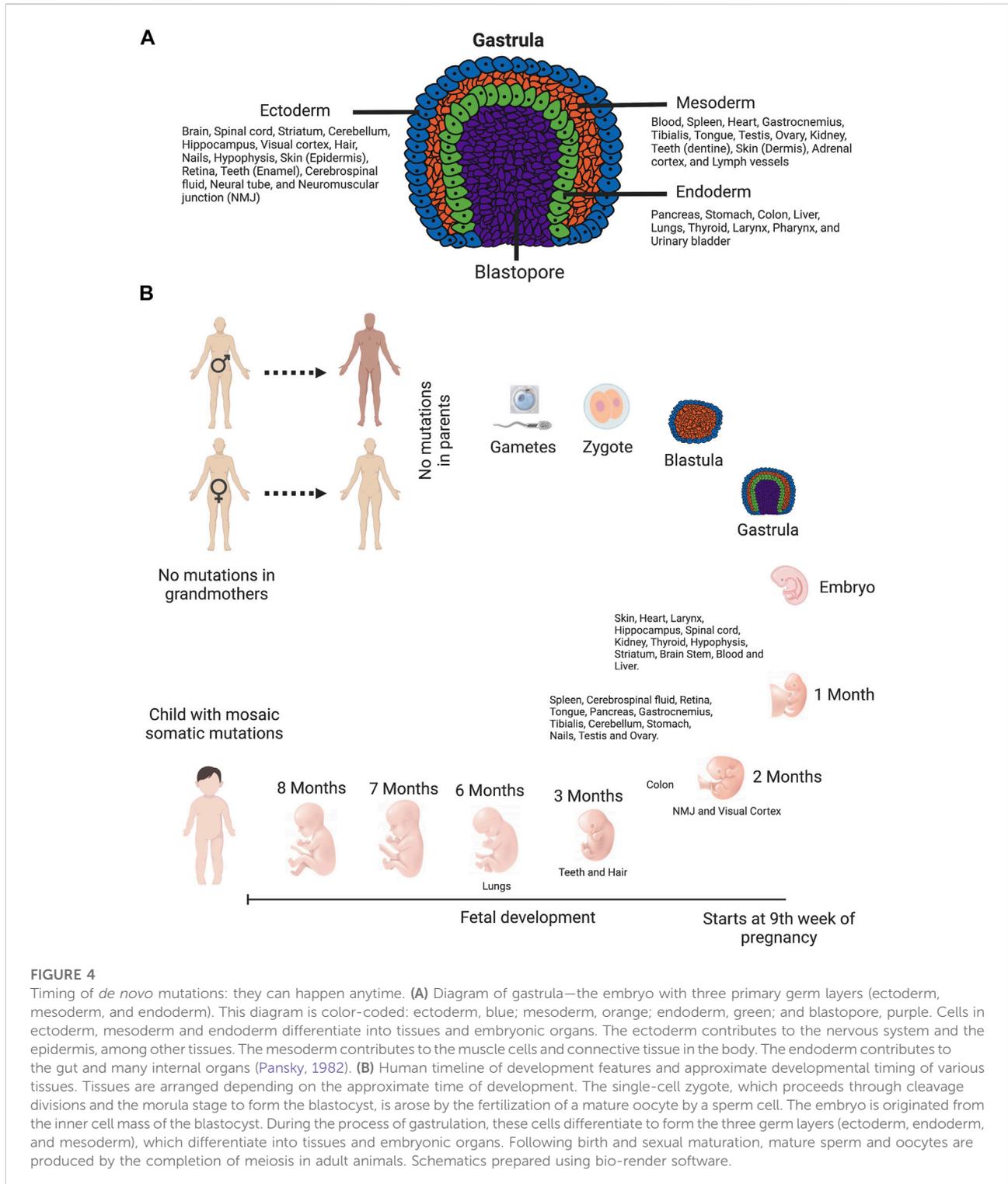
FIGURE 4
Timing of *de novo* mutations: they can happen anytime. **(A)** Diagram of gastrula—the embryo with three primary germ layers (ectoderm, mesoderm, and endoderm). This diagram is color-coded: ectoderm, blue; mesoderm, orange; endoderm, green; and blastopore, purple. Cells in ectoderm, mesoderm and endoderm differentiate into tissues and embryonic organs. The ectoderm contributes to the nervous system and the epidermis, among other tissues. The mesoderm contributes to the muscle cells and connective tissue in the body. The endoderm contributes to the gut and many internal organs (Pansky, 1982). **(B)** Human timeline of development features and approximate developmental timing of various tissues. Tissues are arranged depending on the approximate time of development. The single-cell zygote, which proceeds through cleavage divisions and the morula stage to form the blastocyst, is arose by the fertilization of a mature oocyte by a sperm cell. The embryo is originated from the inner cell mass of the blastocyst. During the process of gastrulation, these cells differentiate to form the three germ layers (ectoderm, endoderm, and mesoderm), which differentiate into tissues and embryonic organs. Following birth and sexual maturation, mature sperm and oocytes are produced by the completion of meiosis in adult animals. Schematics prepared using bio-render software.

frequency of these developmental disorders shows an exponential increase with paternal age at conception (Goriely et al., 2009). This is much higher than observed increase for other *de novo* mutations related disorders (Arnheim and Calabrese, 2009).

Timing of *de novo* mutations

During development through aging, mutation timing could lead to mosaicism levels fluctuating from less than heterozygous to approaching homozygous level (Mohiuddin et al., 2022). The

distribution of somatic mutation in the whole body is regulated by the developmental timing of mutations. Mutation loads between tissues could uncover timing of mutation and may shed light on disease etiology and clinical variability.

Mutation loads greater than heterozygosity (approaching homozygosity) support second-hit mutation theory, which produce mosaicism (Poduri et al., 2013). Moreover, varying levels of tissue specific mutation burden were evident between the post-mortem tissues, both above and below heterozygosity (Mohiuddin et al., 2022). Individuals with less than heterozygous mutations in at least one tissue argue against mutations having arisen in the transmitting germline. Moreover, the highly variable mutation loads between tissues, that can approach homozygous level, suggest extremely high rates of post-zygotic and probably post-natal somatic mutations.

The single-cell zygote, which proceeds through cleavage divisions and the morula stage to form the blastocyst, is arose by the fertilization of a mature oocyte by a sperm cell. The embryo is originated from the inner cell mass of the blastocyst. During the process of gastrulation, these cells differentiate to form the three germ layers (ectoderm, endoderm, and mesoderm), which differentiate into tissues and embryonic organs. The ectoderm contributes to the nervous system and the epidermis, among other tissues, whereas the mesoderm contributes to the muscle cells and connective tissue in the body, and the endoderm contributes to the gut and many internal organs (Figure 4A) (Pansky, 1982). Depending on the timing at which a *de novo* mutation occurs during embryonic development (Figure 4B), it could be organ specific or be evident at different levels in numerous tissues (Youssoufian and Pyeritz, 2002).

Parental influence on the rate and spectrum of human *de novo* germline mutations

The higher rate of mutation for the hemophilia-associated gene in men than in women was noted in 1947 by HALDANE (1947). Haldane's observation was decades later confirmed by whole genome sequencing that the male germ line is more mutagenic (Roach et al., 2010; Conrad et al., 2011; Kong et al., 2012; Michaelson et al., 2012; Venn et al., 2014). Mounting evidence have established the observation of the paternal age effect in humans (Francioli et al., 2015; Rahbari et al., 2016a; Goldmann et al., 2016, 2018; Wong et al., 2016; Jónsson et al., 2017). Each additional year in father's age at conception, on average, results in ~2 additional DNMs in the child (Kong et al., 2012). Consistently, the chance of dominant genetic disorders in the child rises with increasing paternal age (Goriely and Wilkie, 2012; Momand et al., 2013). The increasing number of cell divisions in the male germ line might be the cause of the paternal age effect (Penrose, 1955; Crow, 2000; Kong et al.,

2012). To confirm the continuous generation of sperm cells, the male germline undergoes very frequent genome replications. It has been predicted that in a 20-year-old male the male germ line has experienced 153 genome replications (10 very early embryonic divisions, plus 24 early embryonic divisions, plus 5 years of sperm production each with 23 divisions (Heller and Clermont, 1963)- assuming a sperm production onset at the age of 15 years, and finally 4 replications for spermatogenesis) with the number mounting to 613 genome replications in a 40-year-old male. The replication frequency increased by a factor of ~four between the parental ages of 20 and 40 years. However, these estimated numbers of cell divisions are not proportionate with the observations on DNM accrual (Goldmann et al., 2019). Several studies have reported this unsolved mismatch correlation between genome replications and the number of DNMs (Ségurel et al., 2014; Forster et al., 2015; Gao et al., 2016, 2019; Scally, 2016).

Considering recent observations of DNM accumulation and single-cell transcriptome analyses of the spermatogenic epithelium, Goldmann et al. (2019) suggest that estimated rate of spermatogonium replication may have been constructed on wrong assumptions. In the genome-replication model, the estimate of 23 divisions per year is based on the observations by in the 1960s Heller and Clermont (1963). They showed that one division cycle lasts 16 days, which would conclude to 23 divisions per year, assuming that there are no interruptions or resting phases between the divisions. However, it has recently been observed by several group that the nature of spermatogonial self-renewal is driven by stochastic cell transitions, including cell stages without replicative activity (Hara et al., 2014; Krieger and Simons, 2015; Sharma et al., 2019). Recent studies on single cell transcriptome supported such nonhierarchical model of spermatogonial self-renewal with stochastic, continuous oscillation between cell stages (Guo et al., 2017, 2018; Green et al., 2018; Hermann et al., 2018; Wang et al., 2018). A spermatogonium division model has been proposed and its parameters have been estimated by a fit to DNM accumulation data (Scally, 2016). According to this newly proposed model, the yearly increase of spermatogonium divisions would be only 0.8 instead of 23 divisions per year. Consequently, the number of germline divisions for sperm of 20-year-olds and 60-year-olds would be 42 and 74, respectively. This represents an increase by a factor of 1.76, which would be closer to the increase in DNM numbers. However, recent study has reported a speculative sign of a maternal age effect on the quantity of mutations on the paternal genome, which in favor of the hypothesis that mother age at conception effects mutagenesis in early embryonic development of her child (Gao et al., 2019).

The number of DNMs in offspring rises with both paternal and maternal age, and that some genome regions show enhancement for maternally derived DNMs (Goldmann et al., 2016). The number of DNMs from mothers rises by 0.37 per year

of age, which is one-fourth of the 1.51 per year from fathers (Francioli et al., 2015; Rahbari et al., 2016a; Jónsson et al., 2017). It has been reported that this ratio of paternal to maternal is independent of the age of the parents (Gao et al., 2019), as maternal DNMs also accumulate during aging, although the rate of accumulation is lesser than for the paternal allele. Several recent studies have confirmed the maternal age effect on DNMs (Goldmann et al., 2016; Wong et al., 2016; Jónsson et al., 2017; Gao et al., 2019). Due to the production of oocytes ceases prenatally, and no more genome replications occur, the mechanisms that trigger maternal aging-associated DNMs must fundamentally differ from those underlying the paternal aging-associated DNMs. On this foundation, recently detected maternal age effect (Goldmann et al., 2016; Wong et al., 2016; Jónsson et al., 2017; Sasani et al., 2019) has been interpreted as reflecting the accrual of DNA lesions or damage induced mutations in (primary) oocytes during the lengthy meiotic arrest phase (Ségurel et al., 2014; Gao et al., 2016; Goriely, 2016; Jónsson et al., 2017). However, other explanations for a maternal age effect cannot be excluded (Wong et al., 2016). For example, it could be an effect on post-zygotic mutations (Wong et al., 2016). Recently a positive association between the number of DNMs on paternal chromosomes and maternal age was reported which supports the hypothesis that the age of a mother at conception affects the post-zygotic mutation rate in the developing embryo (Gao et al., 2019).

Recent evidence suggest that mutations can accumulate at a high rate without cell division (Lodato et al., 2018; Abascal et al., 2021). By characterizing the mutational landscape of post-mitotic neurons and polyclonal smooth muscle, Abascal et al., confirm that neurons accumulate somatic mutations at a constant rate throughout life without cell division, with similar rates to mitotically active tissues (Abascal et al., 2021).

Mutational signatures underlying specific mutational processes

A pattern of mutations that is specific to a mutational process occurring in a cell, tissue, or organism is defined as a "mutational signature" (Alexandrov et al., 2013). Mutations arising from different processes, such as DNA replication errors, failure to repair DNA damage, or exposure to mutagens, showed distinct mutational patterns. The various underlying biology of female and male gametogenesis ends up in differences in mutational signatures between maternally and paternally transmitted DNMs, respectively. Nucleotide type (Alexandrov et al., 2013; Aggarwala and Voight, 2016), sequence context (Alexandrov et al., 2013; Aggarwala and Voight, 2016), replication timing (Koren et al., 2012), functional constraints (Gudbjartsson et al., 2015; Aggarwala and Voight, 2016), apolipoprotein B messenger RNA-editing enzyme catalytic (APOBEC) polypeptide activity (Chan and Gordenin, 2015), and epigenetics (Polak et al., 2015;

Rahbari et al., 2016a) have also been reported to affect the mutational landscape. The parent-of-origin specific mutation signatures that become more prominent with increased parental age, indicating to different mutational mechanisms in oogenesis and spermatogenesis (Goldmann et al., 2016). Consistent with biological difference, the mutational spectra of paternal and maternal DNMs are also different. The mutational spectrum, as opposed to the mutational frequency, was affected more by maternal than by paternal age (Jónsson et al., 2017). Paternal DNMs are enriched for C>T mutations and maternal DNMs are enriched for C>G mutations at CpG sites (Jónsson et al., 2017; Goldmann et al., 2018).The natures of *de novo* mutation from mothers change substantially with age, with a 0.33% increase in C>G *de novo* mutations per year. Surprisingly, these age-related changes are not evenly distributed through the genome (Jónsson et al., 2017). Some genome regions on chromosome 8, 9, and 16 are highly enriched for maternal DNMs (up to eight times compared with paternal DNMs) (Goldmann et al., 2016, 2018; Jónsson et al., 2017). The relative frequency of C>G mutations in these regions is drastically larger than among paternal DNMs, and even larger than in the remaining maternal DNMs. In contrast to female germline CpG sites, male germline CpG sites in cells are more favorably methylated, making the sites more prone to C>T transitions and spontaneous deamination (Kobayashi et al., 2013; Gao et al., 2019).

Mutational clusters and hotspots

DNMs arise throughout the human genome, but intermittently several mutations can appear at a nearer distance than usual by casual distribution (Michaelson et al., 2012). Mutational clusters, the occurrence of DNMs in an individual on the same DNA molecule at a closer distance than usual, with numerous DNMs within short reciprocal distances, usually several kilobases ranging from 10 to 100 kb (Michaelson et al., 2012; Chan and Gordenin, 2015; Francioli et al., 2015). These mutations characterize a distinctive subset of 2–3% of DNMs rising due to a single mutational event, as the frequency of clusters is too high to be justified by the incidental co-occurrence of numerous independent sequential DNMs (Schrider et al., 2011; Michaelson et al., 2012; Terekhanova et al., 2013; Francioli et al., 2015; Besenbacher et al., 2016; Goldmann et al., 2016; Yuen et al., 2016).

Clustered DNMs are also unique in terms of their parental age effects and parent-of-origin characteristics. Clustered DNMs are equally rich on maternal and paternal alleles, which is very unlikely to other DNMs. In addition, the quantity of these mutations are strongly associated with maternal than with paternal age, increases quicker with the mother's age than with the father's (Jónsson et al., 2017; Goldmann et al., 2018). Moreover, mutational clusters display a distinctive mutational

spectrum with a significant enhancement of C>G mutations at non-CpG sites (Francioli et al., 2015; Besenbacher et al., 2016; Goldmann et al., 2016, 2018), which is also significantly different between the parental alleles, with the distances between the DNMs of a maternal cluster are greater than for paternal clusters, and the maternal allele being more inclined to C>G mutations than the paternal. In the perspective of cancer, where it is known as "kataegis", it has been proposed that clusters comprising C > G transversions could be linked to the formation of single stranded DNA in diverse cellular processes, such as dysfunctional replication forks and double-strand breaks (DSBs) (Sakofsky et al., 2014). Chromosomal rearrangements could also be another origin for some of these clusters. Previous studies show that the mutation rate for SNVs is prominent and SNVs can cluster in vicinity to the breakpoints of de novo CNVs (Neumann et al., 2010; Carvalho et al., 2013). Previous study on yeast supports the finding that DSB-induced replication is a source of mutation clusters (Sakofsky et al., 2014).

Unlike mutation clusters, that occur within one individual, mutational hotspots are considered overlapping loci that are mutated more frequently than usual in the population. Mutation hotspots in coding sequence has identified by recent research based on whole genome sequencing (WGS) datasets and modeling (Michaelson et al., 2012). The presence of mutational hotspots has also been established in a larger study that exhibited specific bins of 1 Mb within the human genome with elevated mutation rates (Goldmann et al., 2016).

Male germline mosaicism to predict the recurrence risk of genetic diseases

Recently, Breuss et al. (2020), (2021) sequenced fathers' sperm DNA to define paternal germline mosaicism that could improve risk predictions for families with children with autism. Prezygotic de novo mutation in an unaffected parent may exist in a mosaic fashion. Mosaic de novo mutation might be inherited in the zygote and possibly in all cells of the developing offspring, which may cause a de novo disease phenotype. Because de novo prezygotic mutations arise before fertilization in the parental germline, their existence is closely connected to the biology of germ cells. A DNM has been assumed to arise in a single sperm or egg during meiosis, and in such cases, the hypothetical possibility of recurrence of this mutation would be remarkably low in the subsequent children. However, the probability of germline mosaicism uplifts risk of recurrence of the same mutation in families higher than in the general population. The developmental timing and cell lineage affected, merged with the phenotypic consequences of the mutation, eventually regulate the tissue distribution of mosaicism and also the patterns of disease reoccurrence within families.

Breuss et al. (2020), (2021) measure the allele frequency of de novo single nucleotide variants (dSNVs) in sperm relative to that in blood and discovered that a large majority of the mutations are evident in sperm only or are enriched in sperm comparative to their abundance in blood. Depending on their findings, they conclude that germline mosaicism in fathers can be identified better by sequencing of sperm than does sequencing of blood (Breuss et al., 2020; Morrow, 2020). Only somatic cells are routinely assessed in genetic analyses, and a clue about possible germline mosaicism will be obtained if somatic mosaicism is detected, or if more affected siblings are born with an analogous de novo variant. The incidence of parental germline mosaicism has been reported in many disorders now. Recently, in case of routine genetic screening for germline mutation in fathers of patients with apparently de novo mutations, it has been suggested that analysis of sperm could more often be performed (Pacot and Pasmant, 2019).

The identification of a de novo mutation as the cause of disease in a patient has numerous implications for the patient and his family. In terms of family planning, the identification of a de novo mutation as the cause of disease in a child can be positive news with respect to recurrence risk. There is a low risk of recurrence of the mutation in an additional child for the parents with a post-zygotic mutation, anticipated as being the same as the population risk. Breuss et al. (2020) proposed a genetic testing whereby paternal sperm can directly be assessed for pathogenic variants which is previously transmitted and these variants may thereby be categorized as low risk for recurrence or high risk for recurrence. The authors further propose that in case of family planning, men may have their sperm DNA sequenced as a method of discerning high-risk mosaic mutations. Undoubtedly, this process may additionally be appropriate to sperm donor clinics.

Somatic repeat instability

Expansion of gene-specific tandem repeat DNAs, like (CAG)n, (CGG)n, and (GGGGCC)n underlies over fifty neurological, neuromuscular and neurodegenerative diseases, including Huntington disease (HD), myotonic dystrophy (DM1), spinocerebellar ataxias (SCAs), fragile X syndrome (FXS), schizophrenia, amyotrophic lateral sclerosis (ALS) and autism (ASD) (Castel et al., 2010; DeJesus-Hernandez et al., 2011; Song et al., 2018; Course et al., 2020; Trost et al., 2020). The key mechanisms by which expansion of repeat DNA causes disease are R-loop formation followed by the activation of the DNA damage response and repeat-induced transcriptional gene silencing, repeat associated RNA translation of toxic repeat peptides, and RNA-facilitated gain of function through gelation and sequestration of RNA- binding proteins, gain of function of canonically translated repeat-harboring proteins. The severity

of these mechanisms is amplified by somatic repeat instability, a process that differentially regulates repeat length across the cells and tissues of individuals over their lifespan (La Spada and Taylor, 2010; Nussbacher et al., 2019; Khristich and Mirkin, 2020), which impacts both disease age of onset and tissue specificity of pathogenic features. Several studies on somatic instability and age at disease onset in DM1 patients with variant repeats (Braida et al., 2010; Pešović et al., 2018) and with pure repeats (Morales et al., 2020) illustrate that variation in age at disease onset is modulated by variation in somatic instability.

DNA mismatch repair pathways (MMR) resolve unusual DNA–DNA and DNA–RNA structures formed during transcription. MMR can intensify intergenerational repeat instability and permit additional expansion in somatic cells (Castel et al., 2010; Reddy et al., 2014; Khristich and Mirkin, 2020; Malik et al., 2021). Most repeat expansion diseases displayed somatic instability, which creates variation in toxicity and repeat size across tissues in the same patient. Remarkably, somatic expansions are detected in terminally differentiated cells such as neurons and myofibers, makes it DNA replication independent process (Pearson et al., 1997; Gonitel et al., 2008; Keogh et al., 2017; Neil et al., 2021). The detailed mechanisms of somatic instability in repeat expansion diseases were recently reviewed (Khristich and Mirkin, 2020). Recent studies supported a role for somatic instability in disease pathogenesis implicating FAN1 and MMR proteins as modifiers of the age of onset in Huntington disease (Lee et al., 2015, 2017; Deshmukh A. et al., 2021b, Deshmukh et al., 2021 A. L.a; Porro et al., 2021) and spinocerebellar ataxias (SCAs) (Bettencourt et al., 2016), and alteration of MMR is sufficient to overturn somatic instability and can moderate toxicity in Huntington disease mice (Kovalenko et al., 2012; Pinto et al., 2013). In DM1, expansion-biased repeat instability in somatic cells is continuous throughout an individual's life (Monckton et al., 1995; Wong et al., 1995), and is reflected to contribute directly to the progressive nature of the disease (Morales et al., 2012). The differential rates of somatic instability cause differences in the pathogenicity of repeat expansions across tissues. Somatic instability of the CTG repeat generates alleles in brain and skeletal muscle with repeat tracts considerably longer than in leukocytes (Ashizawa et al., 1993; Thornton et al., 1994), possibly aggravating RNA toxicity in those tissues.

De novo mutation contributes to diseases

A lot has been found about the cause of *de novo* mutations, but how it contributes to disease is not well understood. There is no question that identifying the mechanism & the mutation load between tissues, specifically affected tissues, is going to help to understand how this contributes. For example, autism spectrum disorder (ASD) - a known and common form of neurodevelopmental intellectual disabilities (ID) described by a combination of deficits in communication and social interaction together with repetitive and restrictive behaviors (Kanner, 1968; Baio et al., 2018), is associated with *de novo* mutations. *ADNP* is one of the most frequently mutated genes in blood DNA through targeted molecular inversion probe sequencing studies and multiple recent whole-exome sequencing in ASD/ID cohorts (McRae et al., 2017; Stessman et al., 2017; Satterstrom et al., 2020). Nothing is known about how *de novo* mutations contribute to the extreme clinical variability of autism, a knowledge that would be beneficial in clinical assessment, diagnosis, and management.

De novo mutations are the most extreme form of rare genetic variation, have been subjected to less stringent evolutionary selection, more deleterious, on average, than inherited variation (Crow, 2000; Eyre-Walker and Keightley, 2007), which makes these mutations crucial candidates for causing genetic diseases that arise sporadically. A list of sporadic genetic disorders that are caused by *de novo* germline mutations are listed in the Table 1. Cancer is known to have somatic *de novo* mutations, which is covered elsewhere [for coverage of this topic, see reviews (Ding et al., 2010; Meyerson et al., 2010)]. Disorders caused by genetic mosaicism where reviewed elsewhere recently (Moog et al., 2020).

Clinical consequences

The clinical consequences of mosaicism depend upon which chromosome is involved, and where and when a mutation arises. The identification of mosaicism is imperative in establishing a disease diagnosis, evaluating recurrence risk, and counseling. Mutation detection can be hampered by the presence of mosaicism in a patient. Both somatic and germline mosaicisms in humans have several clinical implications. Harmful *de novo* point mutations and indels that affect essential genes in development have been recognized as a prominent source of both common and rare genetic disorders (Vissers et al., 2010; O'Roak et al., 2011; De Ligt et al., 2012; Rauch et al., 2012; Allen et al., 2013; Iossifov et al., 2014; Chong et al., 2015). Somatic expansion of the *HTT* CAG repeat sequence is the rate-determining mechanism of HD (Lee et al., 2019). It is challenging to forecast the clinical consequences of mosaicism, as the patterns and distribution of abnormal cells can diverge widely depending on the timing of the mutation events. Depending on various factors, such as the degree of mosaicism and/or the gene involved, the clinical outcome of the mosaicism may be different in different disorders. For example, mosaicism of a specific chromosome may affect organ development, while mosaicism of other specific chromosome may affect

TABLE 1 List of genetic disorders caused by *de novo* mutations.

Disorders	Characteristics
Autism spectrum disorder (ASD) (O'Roak et al., 2011; O'Roak et al., 2012; Devlin et al., 2012; O' Roak et al., 2012; Sanders et al., 2012; Iossifov et al., 2014)	Difficulties in social communications and interactions. Restricted, repetitive patterns of behavior, activities, or interests and sensory problems
Intellectual disability (ID) (De Vries et al., 2005; Vissers et al., 2010)	Certain limitations in mental abilities that affect cognitive functioning, communication skills, social and self-care skills
Schizophrenia (Girard et al., 2011; Xu et al., 2011)	Psychosis, apathy and withdrawal, and cognitive impairment, which cause problems in social and occupational functioning, and self-care
Down syndrome (Allen, 1974)	Child is born with an extra copy of their 21st chromosome—thus its other name is trisomy 21. This causes mental and physical developmental delays and disabilities
Schinzel–Giedion syndrome (Hoischen et al., 2010)	Skeletal dysplasia, congenital hydronephrosis, and severe developmental retardation
Kabuki syndrome (Ng et al., 2010)	Skeletal abnormalities, distinctive facial features and intellectual disabilities (ID)
Bohring–Opitz syndrome (Hoischen et al., 2011)	Multiple malformations, failure to thrive, facial anomalies and severe intellectual disabilities (ID)
Proteus syndrome (Lindhurst et al., 2011)	Mosaic or patchy overgrowth and hyperplasia of various organs and tissues
CHARGE syndrome (Vissers et al., 2004)	Arises during early fetal development and affects various organ systems, such as heart, eyes and ears
KBG syndrome (Sirmaci et al., 2011)	Facial dysmorphisms, macrodontia, skeletal anomalies and developmental delay
Baraitser–Winter syndrome (Rivière et al., 2012b)	Intellectual disability (ID) that ranges from mild to profound and typical craniofacial features
Microdeletion syndrome (Hassfurther et al., 2015; Kogan et al., 2015)	Mild to moderate intellectual disabilities, autism spectrum disorders, learning delays, or normal intelligence, epilepsy and mental illness
Coffin–Siris syndrome (Santen et al., 2012; Tsurusaki et al., 2012)	Abnormal head and facial (craniofacial) area, resulting in a coarse facial appearance
Adrenoleukodystrophy (Wang et al., 2011)	Failure of adrenal glands, progressive brain damage and, eventually, death
Crouzon syndrome (Toriello and Meck, 2008)	Premature fusion of the skull bones (craniosynostosis), exophthalmos, and midface hypoplasia
Multiple Endocrine Neoplasia Type 2 (Toriello and Meck, 2008)	Prevalence of medullary thyroid carcinoma and risk of developing other specific tumors that affect additional glands of the endocrine system
Charcot–Marie–Tooth disease type 1a (Pentao et al., 1992; Berg et al., 2010)	A rare genetic neurological disorder. It affects the peripheral nerves
Achondroplasia (Toriello and Meck, 2008)	Limited range of motion at the elbows, dwarfism, small fingers, large head size (macrocephaly), and normal intelligence
Apert Syndrome (Toriello and Meck, 2008)	Premature closing of cranial sutures. Certain fingers and toes fused or webbed
Duchenne muscular dystrophy (Helderman-van Den Enden et al., 2009)	Skeletal muscle weakness and degeneration
PIK3CA-related overgrowth spectrum (PROS) (Canaud et al., 2021)	Severe functional impairment, pain, vascular & neurological complications, seizures, and developmental delay, etc.
Paroxysmal nocturnal hemoglobinuria 1 (PNH1) (Li and Williams, 2013)	Hemoglobinuria, abdominal pain, smooth muscle dystonias, fatigue, and thrombosis
X-linked alpha-thalassemia mental retardation (Li and Williams, 2013)	Sometimes associated with myelodysplastic syndrome, with cases often associated with somatic mutations
Neurofibromatosis 1 (NF1) (Li and Williams, 2013)	Cafe-au-lait spots, Lisch nodules in the eye, and fibromatous tumors of the skin
Cardiac myocyte (Li and Williams, 2013)	Affect electrical communication and associate with a large minority of atrial fibrillation cases
Alport syndrome (Li and Williams, 2013)	X-linked dominant disorder characterized by kidney disease, hearing loss, and eye abnormalities
Lissencephaly, or smooth brain (Li and Williams, 2013)	Lethal in males, but milder forms have been associated with somatic mosaics in two patients with predominantly posterior subcortical band heterotopia
Autoimmune lymphoproliferative syndrome (ALPS) (Li and Williams, 2013)	A disease of benign lymphoproliferation, elevated immunoglobulins, plasma IL-10 and FAS-L, and accumulation of double-negative T cells

muscular development. In addition to the chromosome, the prevalence or sheer numbers of the mosaic cell line are also matters. For an individual with a few mosaic cells, the effect of abnormal cells is masked by the number of normal cells. The opposite also true, for an individual with a large number of mosaic cells, the normal cells will be masked by the mosaic cells.

For unaffected parents who have an affected child and are planning a pregnancy, the recurrence risk may relate to the occurrence of new mutations at a particular gene or locus, the nature of mutational mechanism, the severity of the phenotype conferred by mosaicism, or the age and sex of the mosaic parent. For parents with germline mosaicism, the risk for a recurrence of another child with the disease is high. Mosaicism is also vital for

disease mechanism. As an example, the protein kinase AKT1 with somatic mutations are correlated with Proteus syndrome, whereas mosaicism for post-zygotic mutations in genes for three essential components of the phosphatidylinositol 3-kinase (PI3K)–AKT signaling pathway that increase signaling can trigger a variety of related megalencephaly syndromes (Poduri et al., 2013). Compared to the same mutation present in a constitutional state, mosaic mutations can result in a less severe phenotype (Wallis et al., 1990), suggesting that more prevalent mosaicism is prone to have a more severe phenotype.

Most of the somatic mutations can be neutral and there has been speculation of how they could even be beneficial in some cases. For example, they may be contributing to functional diversity of neurons (Newman et al., 2017). Somatic mosaicism in neurons is generated mostly by *de novo* insertions of long interspersed nuclear elements (LINE retrotransposons) while undertaking the last neural progenitor divisions. As a population, neurons with low or high transposable elements insertions may express diverse subcategories of ion channels or neuronal adhesion molecules and thus have different firing properties. The origination of such a diverse neuronal population initiates the stochastic input required for network formation (Newman et al., 2017).

Tools and techniques to detect mosaicism

Detection of mosaicism relies on applying subtle genotyping techniques that can detect low-level mosaicism in a more tedious fashion. Fluorescence *in situ* hybridization (FISH) have been used to analyze mosaic embryos. Molecular cytogenetic techniques (e.g., single-nucleotide polymorphism [SNP] array, array comparative genome hybridization [aCGH], chromosomal microarray analyses (CMA), quantitative polymerase chain reaction [qPCR], high-resolution next-generation sequencing (NGS) are better tools than FISH, because they can provide information on the copy number of all 24 types of chromosome (Munné and Wells, 2017). Mosaicism for primary trisomies in prenatal samples can be detected by quantitative fluorescent PCR (QF-PCR) and karyotype analysis (Donaghue et al., 2005). Non-invasive prenatal testing (NIPT) is used to detect confined placental mosaicism (CPM) (Eggenhuizen et al., 2021). Targeted strategies (e.g., droplet digital PCR) is a very powerful and efficient tool for sensitive detection and quantification of mosaicism (Mohiuddin et al., 2022).

Conclusion and future perspectives

The identification of mosaic disorders is mounting day by day, which can be uniformly distributed throughout an

organism, tissue-specific or segmental, and the germ line, somatic tissues or both can be affected. It can occur at numerous stages of development or adult life and can be triggered by mutation from a variant genotype to a normal genotype or vice versa.

The element of genetic counselling that addresses recurrence risks for patients with mosaic disorders is challenging. This is because genetic counselling necessitates analysis of sample cells within a given tissue and also because of the fact that mosaicism may be tissue-specific or tissue-limited. It is undoubtedly challenging to diagnose of tissue-limited mosaicism when the affected tissue is not skin or blood, tissues that are most frequently analyzed in clinical laboratories. Powerful genomic tools like NGS or targeted strategies (e.g., droplet digital PCR) might be useful for sensitive detection and quantification of mutations. These techniques will develop standard tools for the superior assessment of recurrence risks in families for which a genetic disease is triggered by *de novo* disease gene mutation event.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abascal, F., Harvey, L. M. R., Mitchell, E., Lawson, A. R. J., Lensing, S. V., Ellis, P., et al. (2021). Somatic mutation landscapes at single-molecule resolution. *Nature* 593, 405–410. doi:10.1038/s41586-021-03477-4

Abyzov, A., Mariani, J., Palejev, D., Zhang, Y., Haney, M. S., Tomasini, L., et al. (2012). Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* 492, 438–442. doi:10.1038/nature11629

Acuna-Hidalgo, R., Bo, T., Kwint, M. P., Van De Vorst, M., Pinelli, M., Veltman, J. A., et al. (2015). Post-zygotic point mutations are an underrecognized source of de Novo genomic variation. *Am. J. Hum. Genet.* 97, 67–74. doi:10.1016/j.ajhg.2015.05.008

Aggarwala, V., and Voight, B. F. (2016). An expanded sequence context model broadly explains variability in polymorphism levels across the human genome. *Nat. Genet.* 48, 349–355. doi:10.1038/ng.3511

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., et al. (2013). Signatures of mutational processes in human cancer. *Nature* 500, 415–421. doi:10.1038/nature12477

Allen, A. S., Berkovic, S. F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E. E., et al. (2013). De novo mutations in epileptic encephalopathies. *Nature* 501, 217–221. doi:10.1038/nature12439

Allen, G. (1974). Aetiology of Down's syndrome inferred by Waardenburg in 1932. *Nature* 250, 436–437. doi:10.1038/250436a0

Arnheim, N., and Calabrese, P. (2009). Understanding what determines the frequency and pattern of human germline mutations. *Nat. Rev. Genet.* 10, 478–488. doi:10.1038/nrg2529

Ashizawa, T., Dubel, J. R., and Harati, Y. (1993). Somatic instability of ctg repeat in myotonic dystrophy. *Neurology* 43, 2674–2678. doi:10.1212/wnl.43.12.2674

Awadalla, P., Gauthier, J., Myers, R. A., Casals, F., Hamdan, F. F., Griffing, A. R., et al. (2010). Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. *Am. J. Hum. Genet.* 87, 316–324. doi:10.1016/j.ajhg.2010.07.019

Baarends, W. M., van der Laan, R., and Grootegoed, J. A. (2001). DNA repair mechanisms and gametogenesis. *Reproduction* 121, 31–39. doi:10.1530/rep.0.1210031

Baillie, J. K., Barnett, M. W., Upton, K. R., Gerhardt, D. J., Richmond, T. A., De Sapio, F., et al. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479, 534–537. doi:10.1038/nature10531

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., et al. (2018). Prevalence of autism spectrum disorder among children aged 8 Years - autism and developmental disabilities monitoring network, 11 Sites, United States, 2014. *MMWR. Surveill. Summ.* 67, 1–23. doi:10.15585/mmwr.ss6706a1

Behringer, R., Gertsenstein, M., Vintersen Nagy, K., and Nagy, A. (2014). *Manipulating the mouse embryo: A laboratory manual.* Fourth Edition. Cold Harb: Lab. Press.

Berg, I. L., Neumann, R., Lam, K. W. G., Sarbajna, S., Odenthal-Hesse, L., May, C. A., et al. (2010). PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat. Genet.* 42, 859–863. doi:10.1038/ng.658

Besenbacher, S., Liu, S., Izarzugaza, J. M. G., Grove, J., Belling, K., Bork-Jensen, J., et al. (2015). Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. *Nat. Commun.* 6, 5969. doi:10.1038/ncomms6969

Besenbacher, S., Sulem, P., Helgason, A., Helgason, H., Kristjansson, H., Jonasdottir, A., et al. (2016). Multi-nucleotide de novo Mutations in Humans. *PLoS Genet.* 12, e1006315. doi:10.1371/journal.pgen.1006315

Bettencourt, C., Hensman-Moss, D., Flower, M., Wiethoff, S., Brice, A., Goizet, C., et al. (2016). DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann. Neurol.* 79 (6), 983–990. doi:10.1002/ana.24656

Biesecker, L. G., and Spinner, N. B. (2013). A genomic view of mosaicism and human disease. *Nat. Rev. Genet.* 14, 307–320. doi:10.1038/nrg3424

Braida, C., Stefanatos, R. K. A., Adam, B., Mahajan, N., Smeets, H. J. M., Niel, F., et al. (2010). Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum. Mol. Genet.* 19, 1399–1412. doi:10.1093/hmg/ddq015

Breuss, M. W., Antaki, D., George, R. D., Kleiber, M., James, K. N., Ball, L. L., et al. (2020). Autism risk in offspring can be assessed through quantification of male sperm mosaicism. *Nat. Med.* 26, 143–150. doi:10.1038/s41591-019-0711-0

Breuss, M. W., Yang, X., and Gleeson, J. G. (2021). Sperm mosaicism: Implications for genomic diversity and disease. *Trends Genet.* 37, 890–902. doi:10.1016/j.tig.2021.05.007

Campbell, C. D., and Eichler, E. E. (2013). Properties and rates of germline mutations in humans. *Trends Genet.* 29, 575–584. doi:10.1016/j.tig.2013.04.005

Campbell, I. M., Shaw, C. A., Stankiewicz, P., and Lupski, J. R. (2015). Somatic mosaicism: Implications for disease and transmission genetics. *Trends Genet.* 31, 382–392. doi:10.1016/j.tig.2015.03.013

Campbell, I. M., Stewart, J. R., James, R. A., Lupski, J. R., Stankiewicz, P., Olofsson, P., et al. (2014a). Parent of origin, mosaicism, and recurrence risk: Probabilistic modeling explains the broken symmetry of transmission genetics. *Am. J. Hum. Genet.* 95, 345–359. doi:10.1016/j.ajhg.2014.08.010

Campbell, I. M., Yuan, B., Robberecht, C., Pfundt, R., Szafranski, P., McEntagart, M. E., et al. (2014b). Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *Am. J. Hum. Genet.* 95, 173–182. doi:10.1016/j.ajhg.2014.07.003

Canaud, G., Hammill, A. M., Adams, D., Vikkula, M., and Keppler-Noreuil, K. M. (2021). A review of mechanisms of disease across PIK3CA-related disorders with vascular manifestations. *Orphanet J. Rare Dis.* 16, 306. doi:10.1186/s13023-021-01929-8

Carvalho, C. M. B., Pehlivan, D., Ramocki, M. B., Fang, P., Alleva, B., Franco, L. M., et al. (2013). Replicative mechanisms for CNV formation are error prone. *Nat. Genet.* 45, 1319–1326. doi:10.1038/ng.2768

Castel, A. L., Cleary, J. D., and Pearson, C. E. (2010). Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. Cell Biol.* 11 (3), 165–170. doi:10.1038/nrm2854

Chan, K., and Gordenin, D. A. (2015). Clusters of multiple mutations: Incidence and molecular mechanisms. *Annu. Rev. Genet.* 49, 243–267. doi:10.1146/annurev-genet-112414-054714

Chong, J. X., Buckingham, K. J., Jhangiani, S. N., Boehm, C., Sobreira, N., Smith, J. D., et al. (2015). The genetic basis of mendelian phenotypes: Discoveries, challenges, and opportunities. *Am. J. Hum. Genet.* 97, 199–215. doi:10.1016/j.ajhg.2015.06.009

Conrad, D. F., Keebler, J. E. M., Depristo, M. A., Lindsay, S. J., Zhang, Y., Casals, F., et al. (2011). Variation in genome-wide mutation rates within and between human families. *Nat. Genet.* 43, 712–714. doi:10.1038/ng.862

Coorens, T. H. H., Moore, L., Robinson, P. S., Sanghvi, R., Christopher, J., Hewinson, J., et al. (2021). Extensive phylogenies of human development inferred from somatic mutations. *Nature* 597, 387–392. doi:10.1038/s41586-021-03790-y

Course, M. M., Gudsnuk, K., Smukowski, S. N., Winston, K., Desai, N., Ross, J. P., et al. (2020). Evolution of a human-specific tandem repeat associated with ALS. *Am. J. Hum. Genet.* 107, 445–460. doi:10.1016/j.ajhg.2020.07.004

Crow, J. F. (2000). The origins, patterns and implications of human spontaneous mutation. *Nat. Rev. Genet.* 1, 40–47. doi:10.1038/35049558

Dal, G. M., Ergüner, B., Sağiroğlu, M. S., Yüksel, B., Onat, O. E., Alkan, C., et al. (2014). Early postzygotic mutations contribute to de novo variation in a healthy monozygotic twin pair. *J. Med. Genet.* 51, 455–459. doi:10.1136/jmedgenet-2013-102197

De Felici, M. (2013). Origin, migration, and proliferation of human primordial germ cells. *Oogenesis*, 19–37. doi:10.1007/978-0-85729-826-3_2

De Ligt, J., Willemsen, M. H., Van Bon, B. W. M., Kleefstra, T., Yntema, H. G., Kroes, T., et al. (2012). Diagnostic exome sequencing in persons with severe

intellectual disability. *N. Engl. J. Med.* 367, 1921–1929. doi:10.1056/NEJMoa1206524

De Vries, B. B. A., Pfundt, R., Leisink, M., Koolen, D. A., Vissers, L. E. L. M., Janssen, I. M., et al. (2005). Diagnostic genome profiling in mental retardation. *Am. J. Hum. Genet.* 77, 606–616. doi:10.1086/491719

DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256. doi:10.1016/j.neuron.2011.09.011

Deshmukh, A. L., Caron, M. C., Mohiuddin, M., Lanni, S., Panigrahi, G. B., Khan, M., et al. (2021a). FAN1 exo- not endo-nuclease pausing on disease-associated slipped-DNA repeats: A mechanism of repeat instability. *Cell Rep.* 37, 110078. doi:10.1016/j.celrep.2021.110078

Deshmukh, A., Porro, A., Mohiuddin, M., Lanni, S., Panigrahi, G., Caron, M., et al. (2021b). FAN1, a DNA repair nuclease, as a modifier of repeat expansion disorders. *J. Huntingt. Dis.* 10 (1), 95–122. doi:10.3233/JHD-200448

Devlin, B., Boone, B. E., Levy, S. E., Lihm, J., Buxbaum, J. D., Wu, Y., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485, 242–245. doi:10.1038/nature11011

D'Gama, A. M., and Walsh, C. A. (2018). Somatic mosaicism and neurodevelopmental disease. *Nat. Neurosci.* 21, 1504–1514. doi:10.1038/s41593-018-0257-3

Ding, L., Wendl, M. C., Koboldt, D. C., and Mardis, E. R. (2010). Analysis of next-generation genomic data in cancer: Accomplishments and challenges. *Hum. Mol. Genet.* 19, R188–R196. doi:10.1093/hmg/ddq391

Donaghue, C., Mann, K., Docherty, Z., and Ogilvie, C. M. (2005). Detection of mosaicism for primary trisomies in prenatal samples by QF-PCR and karyotype analysis. *Prenat. Diagn.* 25, 65–72. doi:10.1002/pd.1086

Drost, J. B., and Lee, W. R. (1995). Biological basis of germline mutation: Comparisons of spontaneous germline mutation rates among drosophila, mouse, and human. *Environ. Mol. Mutagen.* 25, 48–64. doi:10.1002/em.2850250609

Eggenhuizen, G. M., Go, A., Koster, M. P. H., Baart, E. B., and Galjaard, R. J. (2021). Confined placental mosaicism and the association with pregnancy outcome and fetal growth: A review of the literature. *Hum. Reprod. Update* 27, 885–903. doi:10.1093/humupd/dmab009

Evrony, G. D., Cai, X., Lee, E., Hills, L. B., Elhosary, P. C., Lehmann, H. S., et al. (2012). Single-neuron sequencing analysis of l1 retrotransposition and somatic mutation in the human brain. *Cell* 151, 483–496. doi:10.1016/j.cell.2012.09.035

Eyre-Walker, A., and Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* 8, 610–618. doi:10.1038/nrg2146

Fasching, L., Jang, Y., Tomasi, S., Schreiner, J., Tomasini, L., Brady, M. V., et al. (2021). Early developmental asymmetries in cell lineage trees in living individuals. *Science* 80, 1245–1248. doi:10.1126/science.abe0981

Flores, M., Morales, L., Gonzaga-Jauregui, C., Domínguez-Vidaña, R., Zepeda, C., Yañez, O., et al. (2007). Recurrent DNA inversion rearrangements in the human genome. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6099–6106. doi:10.1073/pnas.0701631104

Forsberg, L. A., Gisselsson, D., and Dumanski, J. P. (2017). Mosaicism in health and disease-clones picking up speed. *Nat. Rev. Genet.* 18, 128–142. doi:10.1038/nrg.2016.145

Forster, P., Hohoff, C., Dunkelmann, B., Schürenkamp, M., Pfeiffer, H., Neuhuber, F., et al. (2015). Elevated germline mutation rate in teenage fathers. *Proc. Biol. Sci.* 282, 20142898. doi:10.1098/rspb.2014.2898

Francioli, L. C., Polak, P. P., Koren, A., Menelaou, A., Chun, S., Renkens, I., et al. (2015). Genome-wide patterns and properties of de novo mutations in humans. *Nat. Genet.* 47, 822–826. doi:10.1038/ng.3292

Frank, S. A. (2010). Evolution in health and medicine sackler colloquium: Somatic evolutionary genomics: Mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* 107, 1725–1730. doi:10.1073/pnas.0909343106

Gajecka, M. (2016). Unrevealed mosaicism in the next-generation sequencing era. *Mol. Genet. Genomics* 291, 513–530. doi:10.1007/s00438-015-1130-7

Gao, Z., Moorjani, P., Sasani, T. A., Pedersen, B. S., Quinlan, A. R., Jorde, L. B., et al. (2019). Overlooked roles of DNA damage and maternal age in generating human germline mutations. *Proc. Natl. Acad. Sci. U. S. A.* 116, 9491–9500. doi:10.1073/pnas.1901259116

Gao, Z., Wyman, M. J., Sella, G., and Przeworski, M. (2016). Interpreting the dependence of mutation rates on age and time. *PLoS Biol.* 14, e1002355. doi:10.1371/journal.pbio.1002355

Gianaroli, L. (2000). Preimplantation genetic diagnosis: Polar body and embryo biopsy. *Hum. Reprod.* 15, 69–75. doi:10.1093/humrep/15.suppl_4.69

Giannoulatou, E., Maher, G. J., Ding, Z., Gillis, A. J. M., Dorssers, L. C. J., Hoischen, A., et al. (2017). Whole-genome sequencing of spermatocytic tumors provides insights into the mutational processes operating in the male germline. *PLoS One* 12, e0178169. doi:10.1371/journal.pone.0178169

Girard, S. L., Gauthier, J., Noreau, A., Xiong, L., Zhou, S., Jouan, L., et al. (2011). Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat. Genet.* 43, 860–863. doi:10.1038/ng.886

Goldmann, J. M., Seplyarskiy, V. B., Wong, W. S. W., Vilboux, T., Neerincx, P. B., Bodian, D. L., et al. (2018). Germline de novo mutation clusters arise during oocyte aging in genomic regions with high double-strand-break incidence. *Nat. Genet.* 50, 487–492. doi:10.1038/s41588-018-0071-6

Goldmann, J. M., Veltman, J. A., and Gilissen, C. (2019). De novo mutations reflect development and aging of the human germline. *Trends Genet.* 35, 828–839. doi:10.1016/j.tig.2019.08.005

Goldmann, J. M., Wong, W. S. W., Pinelli, M., Farrah, T., Bodian, D., Stittrich, A. B., et al. (2016). Parent-of-origin-specific signatures of de novo mutations. *Nat. Genet.* 48, 935–939. doi:10.1038/ng.3597

Gonitel, R., Moffitt, H., Sathasivam, K., Woodman, B., Detloff, P. J., Faull, R. L. M., et al. (2008). DNA instability in postmitotic neurons. *Proc. Natl. Acad. Sci. U. S. A.* 105, 3467–3472. doi:10.1073/pnas.0800048105

Goriely, A. (2016). Decoding germline de novo point mutations. *Nat. Genet.* 48, 823–824. doi:10.1038/ng.3629

Goriely, A., Hansen, R. M. S., Taylor, I. B., Olesen, I. A., Jacobsen, G. K., McGowan, S. J., et al. (2009). Activating mutations in FGFR3 and HRAS reveal a shared genetic origin for congenital disorders and testicular tumors. *Nat. Genet.* 41, 1247–1252. doi:10.1038/ng.470

Goriely, A., and Wilkie, A. O. M. (2012). Paternal age effect mutations and selfish spermatogonial selection: Causes and consequences for human disease. *Am. J. Hum. Genet.* 90, 175–200. doi:10.1016/j.ajhg.2011.12.017

Green, C. D., Ma, Q., Manske, G. L., Shami, A. N., Zheng, X., Marini, S., et al. (2018). A comprehensive roadmap of murine spermatogenesis defined by single-cell RNA-seq. *Dev. Cell* 46, 651–667. doi:10.1016/j.devcel.2018.07.025

Gudbjartsson, D. F., Helgason, H., Gudjonsson, S. A., Zink, F., Oddson, A., Gylfason, A., et al. (2015). Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* 47, 435–444. doi:10.1038/ng.3247

Guo, J., Grow, E. J., Mlcochova, H., Maher, G. J., Lindskog, C., Nie, X., et al. (2018). The adult human testis transcriptional cell atlas. *Cell Res.* 28, 1141–1157. doi:10.1038/s41422-018-0099-2

Guo, J., Grow, E. J., Yi, C., Mlcochova, H., Maher, G. J., Lindskog, C., et al. (2017). Chromatin and single-cell RNA-seq profiling reveal dynamic signaling and metabolic transitions during human spermatogonial stem cell development. *Cell Stem Cell* 21, 533–546. doi:10.1016/j.stem.2017.09.003

Haldane, J. B. (1947). The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. *Ann. Eugen.* 13, 262–271. doi:10.1111/j.1469-1809.1946.tb02367.x

Hall, J. G. (1988). Review and hypotheses: Somatic mosaicism: Observations related to clinical genetics. *Am. J. Hum. Genet.* 43, 355–363.

Hara, K., Nakagawa, T., Enomoto, H., Suzuki, M., Yamamoto, M., Simons, B. D., et al. (2014). Mouse spermatogenic stem cells continually interconvert between equipotent singly isolated and syncytial states. *Cell Stem Cell* 14, 658–672. doi:10.1016/j.stem.2014.01.019

Hassfurther, A., Komini, E., Fischer, J., and Leipoldt, M. (2015). Clinical and genetic heterogeneity of the 15q13.3 microdeletion syndrome. *Mol. Syndromol.* 6, 222–228. doi:10.1159/000443343

Helderman-van Den Enden, A. T. J. M., De Jong, R., Den Dunnen, J. T., Houwing-Duistermaat, J. J., Kneppers, A. L. J., Ginjaar, H. B., et al. (2009). Recurrence risk due to germ line mosaicism: Duchenne and Becker muscular dystrophy. *Clin. Genet.* 75, 465–472. doi:10.1111/j.1399-0004.2009.01173.x

Heller, C. G., and Clermont, Y. (1963). Spermatogenesis in man: An estimate of its duration. *Science* 80, 184–186. doi:10.1126/science.140.3563.184

Hermann, B. P., Cheng, K., Singh, A., Roa-De La Cruz, L., Mutoji, K. N., Chen, I. C., et al. (2018). The mammalian spermatogenesis single-cell transcriptome, from spermatogonial stem cells to spermatids. *Cell Rep.* 25, 1650–1667. doi:10.1016/j.celrep.2018.10.026

Hoeijmakers, J. H. J. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* 411, 366–374. doi:10.1038/35077232

Hoischen, A., Van Bon, B. W. M., Gilissen, C., Arts, P., Van Lier, B., Steehouwer, M., et al. (2010). De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat. Genet.* 42, 483–485. doi:10.1038/ng.581

Hoischen, A., Van Bon, B. W. M., Rodríguez-Santiago, B., Gilissen, C., Vissers, L. E. L. M., De Vries, P., et al. (2011). De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat. Genet.* 43, 729–731. doi:10.1038/ng.868

Huang, A. Y., Xu, X., Ye, A. Y., Wu, Q., Yan, L., Zhao, B., et al. (2014). Postzygotic single-nucleotide mosaicisms in whole-genome sequences of clinically unremarkable individuals. *Cell Res.* 24, 1311–1327. doi:10.1038/cr.2014.131

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515, 216–221. doi:10.1038/nature13908

Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., et al. (1994). Gonosomal mosaicism in myotonic dystrophy patients: Involvement of mitotic events in (CTG)(n) repeat variation and selection against extreme expansion in sperm. *Am. J. Hum. Genet.* 54, 575–585.

Jónsson, H., Sulem, P., Arnadottir, G. A., Pálsson, G., Eggertsson, H. P., Kristmundsdottir, S., et al. (2018). Multiple transmissions of de novo mutations in families. *Nat. Genet.* 50, 1674–1680. doi:10.1038/s41588-018-0259-9

Jónsson, H., Sulem, P., Kehr, B., Kristmundsdottir, S., Zink, F., Hjartarson, E., et al. (2017). Parental influence on human germline de novo mutations in 1, 548 trios from Iceland. *Nature* 549, 519–522. doi:10.1038/nature24018

Ju, Y. S., Martincorena, I., Gerstung, M., Petljak, M., Alexandrov, L. B., Rahbari, R., et al. (2017). Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. *Nature* 543, 714–718. doi:10.1038/nature21703

Kamel, D., Mackey, Z. B., Sjöblom, T., Walter, C. A., McCarrey, J. R., Uitto, L., et al. (1997). Role of deoxyribonucleic acid polymerase epsilon in spermatogenesis in mice. *Biol. Reprod.* 57, 1367–1374. doi:10.1095/biolreprod57.6.1367

Kanner, L. (1968). Autistic disturbances of affective contact. *Acta Paedopsychiatr.* 35, 100–136.

Kano, H., Godoy, I., Courtney, C., Vetter, M. R., Gerton, G. L., Ostertag, E. M., et al. (2009). L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev.* 23, 1303–1312. doi:10.1101/gad.1803909

Keogh, N., Chan, K. Y., Li, G. M., and Lahue, R. S. (2017). MutSβ abundance and Msh3 ATP hydrolysis activity are important drivers of CTG•CAG repeat expansions. *Nucleic Acids Res.* 45 (17), 10068–10078. doi:10.1093/nar/gkx650

Khristich, A. N., and Mirkin, S. M. (2020). On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* 295 (13), 4134–4170. doi:10.1074/jbc.REV119.007678

Kobayashi, H., Sakurai, T., Miura, F., Imai, M., Mochiduki, K., Yanagisawa, E., et al. (2013). High-resolution DNA methylome analysis of primordial germ cells identifies gender-specific reprogramming in mice. *Genome Res.* 23, 616–627. doi:10.1101/gr.148023.112

Kogan, J. H., Gross, A. K., Featherstone, R. E., Shin, R., Chen, Q., Heusner, C. L., et al. (2015). Mouse model of chromosome 15q13.3 microdeletion syndrome demonstrates features related to autism spectrum disorder. *J. Neurosci.* 35, 16282–16294. doi:10.1523/JNEUROSCI.3967-14.2015

Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., et al. (2012). Rate of de novo mutations and the importance of father-s age to disease risk. *Nature* 488, 471–475. doi:10.1038/nature11396

Koren, A., Polak, P., Nemesh, J., Michaelson, J. J., Sebat, J., Sunyaev, S. R., et al. (2012). Differential relationship of DNA replication timing to different forms of human mutation and variation. *Am. J. Hum. Genet.* 91, 1033–1040. doi:10.1016/j.ajhg.2012.10.018

Kovalenko, M., Dragileva, E., Claire, St.J., Gillis, T., Guide, J. R., New, J., et al. (2012). Msh2 acts in medium-spiny striatal neurons as an enhancer of CAG instability and mutant huntingtin phenotypes in huntington's disease knock-in mice. *PLoS One* 7, e44273. doi:10.1371/journal.pone.0044273

Krieger, T., and Simons, B. D. (2015). Dynamic stem cell heterogeneity. *Development* 142, 1396–1406. doi:10.1242/dev.101063

La Spada, A. R., and Taylor, J. P. (2010). Repeat expansion disease: Progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.* 11, 247–258. doi:10.1038/nrg2748

Lee, J. H., Huynh, M., Silhavy, J. L., Kim, S., Dixon-Salazar, T., Heiberg, A., et al. (2012). De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. *Nat. Genet.* 44, 941–945. doi:10.1038/ng.2329

Lee, J. M., Chao, M. J., Harold, D., Elneel, K. A., Gillis, T., Holmans, P., et al. (2017). A modifier of Huntington's disease onset at the MLH1 locus. *Hum. Mol. Genet.* 26, 3859–3867. doi:10.1093/hmg/ddx286

Lee, J. M., Correia, K., Loupe, J., Kim, K. H., Barker, D., Hong, E. P., et al. (2019). CAG repeat not polyglutamine length determines timing of huntington's disease onset. *Cell* 178, 887–900. doi:10.1016/j.cell.2019.06.036

Lee, J. M., Wheeler, V. C., Chao, M. J., Vonsattel, J. P. G., Pinto, R. M., Lucente, D., et al. (2015). Identification of genetic factors that modify clinical onset of huntington's disease. *Cell* 162 (3), 516–526. doi:10.1016/j.cell.2015.07.003

Li, C., and Williams, S. M. (2013). Human somatic variation: It's not just for cancer anymore. *Curr. Genet. Med. Rep.* 1, 212–218. doi:10.1007/s40142-013-0029-z

Lim, E. T., Uddin, M., De Rubeis, S., Chan, Y., Kamumbu, A. S., Zhang, X., et al. (2017). Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat. Neurosci.* 20, 1217–1224. doi:10.1038/nn.4598

Lindahl, T., and Wood, R. D. (1999). Quality control by DNA repair. *Science* 80, 1897–1905. doi:10.1126/science.286.5446.1897

Lindhurst, M. J., Sapp, J. C., Teer, J. K., Johnston, J. J., Finn, E. M., Peters, K., et al. (2011). A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N. Engl. J. Med.* 365, 611–619. doi:10.1056/nejmoa1104017

Lindsay, S. J., Rahbari, R., Kaplanis, J., Keane, T., and Hurles, M. E. (2019). Similarities and differences in patterns of germline mutation between mice and humans. *Nat. Commun.* 10, 4053. doi:10.1038/s41467-019-12023-w

Lodato, M. A., Rodin, R. E., Bohrson, C. L., Coulter, M. E., Barton, A. R., Kwon, M., et al. (2018). Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* 80, 555–559. doi:10.1126/science.aao4426

Lynch, M. (2010). Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl. Acad. Sci. U. S. A.* 107, 961–968. doi:10.1073/pnas.0912629107

Macosko, E. Z., and McCarroll, S. A. (2012). Exploring the variation within. *Nat. Genet.* 44, 614–616. doi:10.1038/ng.2311

Malik, I., Kelley, C. P., Wang, E. T., and Todd, P. K. (2021). Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat. Rev. Mol. Cell Biol.* 22, 589–607. doi:10.1038/s41580-021-00382-6

Martincorena, I., and Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science* 80, 1483–1489. doi:10.1126/science.aab4082

Martínez-Glez, V., Tenorio, J., Nevado, J., Gordo, G., Rodríguez-Laguna, L., Feito, M., et al. (2020). A six-attribute classification of genetic mosaicism. *Genet. Med.* 22, 1743–1757. doi:10.1038/s41436-020-0877-3

McConnell, M. J., Moran, J. V., Abyzov, A., Akbarian, S., Bae, T., Cortes-Ciriano, I., et al. (2017). Intersection of diverse neuronal genomes and neuropsychiatric disease: The Brain Somatic Mosaicism Network. *Science*, 80, eaal1641. doi:10.1126/science.aal1641

McKinnon, P. J. (2009). DNA repair deficiency and neurological disease. *Nat. Rev. Neurosci.* 10, 100–112. doi:10.1038/nrn2559

McRae, J. F., Clayton, S., Fitzgerald, T. W., Kaplanis, J., Prigmore, E., Rajan, D., et al. (2017). Prevalence and architecture of de novo mutations in developmental disorders. *Nature* 542, 433–438. doi:10.1038/nature21062

Meng, X., Lindahl, M., Hyvönen, M. E., Parvinen, M., De Rooij, D. G., Hess, M. W., et al. (2000). Regulation of cell fate decision of undifferentiated spermatogonia by GDNF. *Science* 80, 1489–1493. doi:10.1126/science.287.5457.1489

Meyerson, M., Gabriel, S., and Getz, G. (2010). Advances in understanding cancer genomes through second-generation sequencing. *Nat. Rev. Genet.* 11, 685–696. doi:10.1038/nrg2841

Michaelson, J. J., Shi, Y., Gujral, M., Zheng, H., Malhotra, D., Jin, X., et al. (2012). Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* 151, 1431–1442. doi:10.1016/j.cell.2012.11.019

Mizuno, K., Miyabe, I., Schalbetter, S. A., Carr, A. M., and Murray, J. M. (2013). Recombination-restarted replication makes inverted chromosome fusions at inverted repeats. *Nature* 493, 246–249. doi:10.1038/nature11676

Mohiuddin, M., Marusic, Z., Anicic, M., Anke, V. D., and Cappuyns, E. (2022). *De novo, post-zygotic, inter-tissue mosaicism of cell autonomous ADNP mutations in autistic individuals: Restricted environmental contribution.*

Momand, J. R., Xu, G., and Walter, C. A. (2013). The paternal age effect: A multifaceted phenomenon. *Biol. Reprod.* 88, 108. doi:10.1095/biolreprod.112.103440

Monckton, D. G., Wong, L. J. C., Ashizawa, T., and Caskey, C. T. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: Small pool PCR analyses. *Hum. Mol. Genet.* 4, 1–8. doi:10.1093/hmg/4.1.1

Monroe, J. G., Srikant, T., Carbonell-Bejerano, P., Becker, C., Lensink, M., Exposito-Alonso, M., et al. (2022). Mutation bias reflects natural selection in *Arabidopsis thaliana. Nature* 602, 101–105. doi:10.1038/s41586-021-04269-6

Moog, U., Felbor, U., Has, C., and Zirn, B. (2020). Disorders caused by genetic mosaicism. *Dtsch. Arztebl. Int.* 117, 119–125. doi:10.3238/arztebl.2020.0119

Morales, F., Couto, J. M., Higham, C. F., Hogg, G., Cuenca, P., Braida, C., et al. (2012). Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. *Hum. Mol. Genet.* 21, 3558–3567. doi:10.1093/hmg/dds185

Morales, F., Vásquez, M., Corrales, E., Vindas-Smith, R., Santamaría-Ulloa, C., Zhang, B., et al. (2020). Longitudinal increases in somatic mosaicism of the expanded CTG repeat in myotonic dystrophy type 1 are associated with variation in age-at-onset. *Hum. Mol. Genet.* 29, 2496–2507. doi:10.1093/hmg/ddaa123

Morrow, E. M. (2020). Paternal sperm DNA mosaicism and recurrence risk of autism in families. *Nat. Med.* 26, 26–28. doi:10.1038/s41591-019-0732-8

Munné, S., and Wells, D. (2017). Detection of mosaicism at blastocyst stage with the use of high-resolution next-generation sequencing. *Fertil. Steril.* 107, 1085–1091. doi:10.1016/j.fertnstert.2017.03.024

Neil, A. J., Hisey, J. A., Quasem, I., McGinty, R. J., Hitczenko, M., Khristich, A. N., et al. (2021). Replication-independent instability of Friedreich's ataxia GAA repeats during chronological aging. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2013080118. doi:10.1073/pnas.2013080118

Neumann, R., Lawson, V. E., and Jeffreys, A. J. (2010). Dynamics and processes of copy number instability in human γ-globin genes. *Proc. Natl. Acad. Sci. U. S. A.* 107, 8304–8309. doi:10.1073/pnas.1003634107

Newman, A. G., Bessa, P., Tarabykin, V., and Singh, P. B. (2017). Activity-DEP endent transposition. *EMBO Rep.* 18, 346–348. doi:10.15252/embr.201643797

Ng, S. B., Bigham, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I., et al. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* 42, 790–793. doi:10.1038/ng.646

Nouspikel, T., and Hanawalt, P. C. (2002). DNA repair in terminally differentiated cells, *DNA Repair (Amst)*, 1. doi:10.1016/S1568-7864(01)00005-2

Nussbacher, J. K., Tabet, R., Yeo, G. W., and Lagier-Tourenne, C. (2019). Disruption of RNA metabolism in neurological diseases and emerging therapeutic interventions. *Neuron* 102, 294–320. doi:10.1016/j.neuron.2019.03.014

O'Huallachain, M., Karczewski, K. J., Weissman, S. M., Urban, A. E., and Snyder, M. P. (2012). Extensive genetic variation in somatic human tissues. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18018–18023. doi:10.1073/pnas.1213736109

O'Roak, B. J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J. J., Girirajan, S., et al. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat. Genet.* 43, 585–589. doi:10.1038/ng.835

O'Roak, B. J., Vives, L., Fu, W., Egertson, J. D., Stanaway, I. B., Phelps, I. G., et al. (2012). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338, 1619–1622. doi:10.1126/science.1227764

Pacot, L., and Pasmant, E. (2019). Should we genotype the sperm of fathers from patients with "de novo" mutations? *Eur. J. Endocrinol.* 182, C1–C23. doi:10.1530/eje-19-0759

Pamphlett, R. (2004). Somatic mutation: A cause of sporadic neurodegenerative diseases? *Med. Hypotheses* 62, 679–682. doi:10.1016/j.mehy.2003.11.023

Pansky, B. (1982). *Review of MEDICAL EMBRYOLOGY.*

Park, S., Mali, N. M., Kim, R., Choi, J. W., Lee, J., Lim, J., et al. (2021). Clonal dynamics in early human embryogenesis inferred from somatic mutation. *Nature* 597, 393–397. doi:10.1038/s41586-021-03786-8

Paul, C., and Robaire, B. (2013). Ageing of the male germ line. *Nat. Rev. Urol.* 10, 227–234. doi:10.1038/nrurol.2013.18

Pearson, C. E., Ewel, A., Acharya, S., Fishel, R. A., and Sinden, R. R. (1997). Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum. Mol. Genet.* 6, 1117–1123. doi:10.1093/hmg/6.7.1117

Pearson, C. E. (2003). Slipping while sleeping? Trinucleotide repeat expansions in germ cells. *Trends Mol. Med.* 9 (11), 490–495. doi:10.1016/j.molmed.2003.09.006

Penrose, L. S. (1955). Parental age and mutation. *Lancet* 266, 312–313. doi:10.1016/S0140-6736(55)92305-9

Pentao, L., Wise, C. A., Chinault, A. C., Patel, P. I., and Lupski, J. R. (1992). Charcot-Marie-Tooth type 1A duplication appears to arise from recombination at repeat sequences flanking the 1.5 Mb monomer unit. *Nat. Genet.* 2, 292–300. doi:10.1038/ng1292-292

Pešović, J., Perić, S., Brkušanin, M., Brajušković, G., Rakoč Ević -Stojanović, V., and Savić-Pavić Ević, D. (2018). Repeat interruptions modify age at onset in myotonic dystrophy type 1 by stabilizing DMPK expansions in somatic cells. *Front. Genet.* 9, 601. doi:10.3389/fgene.2018.00601

Philp, A. J., Leet, C., Vincan, E., Rockman, S. P., Thomas, R. J. S., Phillips, W. A., et al. (2001). The phosphatidylinositol 3'-kinase p85alpha gene is an oncogene in human ovarian and colon tumors. *Cancer Res.* 61, 7426–7429. doi:10.1177/074873001129002178

Pinto, R. M., Dragileva, E., Kirby, A., Lloret, A., Lopez, E., Claire, St.J., et al. (2013). Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in huntington's disease mice: Genome-wide and candidate approaches. *PLoS Genet.* 9, e1003930. doi:10.1371/journal.pgen.1003930

Poduri, A., Evrony, G. D., Cai, X., Elhosary, P. C., Beroukhim, R., Lehtinen, M. K., et al. (2012). Somatic activation of AKT3 causes hemispheric developmental brain malformations. *Neuron* 74, 41–48. doi:10.1016/j.neuron.2012.03.010

Poduri, A., Evrony, G. D., Cai, X., and Walsh, C. A. (2013). Somatic mutation, genomic variation, and neurological disease. *Science* 341, 1237758. doi:10.1126/science.1237758

Polak, P., Karlic, R., Koren, A., Thurman, R., Sandstrom, R., Lawrence, M. S., et al. (2015). Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* 518, 360–364. doi:10.1038/nature14221

Porro, A., Mohiuddin, M., Zurfluh, C., Spegg, V., Dai, J., Iehl, F., et al. (2021). FAN1-MLH1 interaction affects repair of DNA interstrand cross-links and slipped-CAG/CTG repeats. *Sci. Adv.* 7, eabf7906. doi:10.1126/sciadv.abf7906

Proukakis, C., Houlden, H., and Schapira, A. H. (2013). Somatic alpha-synuclein mutations in Parkinson's disease: Hypothesis and preliminary data. *Mov. Disord.* 28, 705–712. doi:10.1002/mds.25502

Rahbari, R., Wuster, A., Lindsay, S. J., Hardwick, R. J., Alexandrov, L. B., Al Turki, S., et al. (2016a). Timing, rates and spectra of human germline mutation. *Nat. Genet.* 48, 126–133. doi:10.1038/ng.3469

Rahbari, R., Wuster, A., Lindsay, S. J., Hardwick, R. J., Alexandrov, L. B., Al Turki, S., et al. (2016b). Timing, rates and spectra of human germline mutation. *Nat. Genet.* 48, 126–133. doi:10.1038/ng.3469

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., et al. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: An exome sequencing study. *Lancet* 380, 1674–1682. doi:10.1016/S0140-6736(12)61480-9

Reddy, K., Schmidt, M. H. M., Geist, J. M., Thakkar, N. P., Panigrahi, G. B., Wang, Y. H., et al. (2014). Processing of double-R-loops in (CAG)●(CTG) and C9orf72 (GGGGCC)●(GGCCCC) repeats causes instability. *Nucleic Acids Res.* 42 (16), 10473–10487. doi:10.1093/nar/gku658

Richardson, L. L., Pedigo, C., and Ann Handel, M. (2000). Expression of deoxyribonucleic acid repair enzymes during spermatogenesis in mice. *Biol. Reprod.* 62, 789–796. doi:10.1095/biolreprod62.3.789

Rivière, J. B., Mirzaa, G. M., O'Roak, B. J., Beddaoui, M., Alcantara, D., Conway, R. L., et al. (2012a). De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat. Genet.* 44, 934–940. doi:10.1038/ng.2331

Rivière, J. B., Van Bon, B. W. M., Hoischen, A., Kholmanskikh, S. S., O'Roak, B. J., Gilissen, C., et al. (2012b). De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. *Nat. Genet.* 44, 440–444. doi:10.1038/ng.1091

Roach, J. C., Glusman, G., Smit, A. F. A., Huff, C. D., Hubley, R., Shannon, P. T., et al. (2010). Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 80, 636–639. doi:10.1126/science.1186802

Roak, O., Vernot, B. J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., et al. (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485, 246–250. doi:10.1038/nature10989

Robberecht, C., Voet, T., Esteki, M. Z., Nowakowska, B. A., and Vermeesch, J. R. (2013). Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. *Genome Res.* 23, 411–418. doi:10.1101/gr.145631.112

Rulten, S. L., and Caldecott, K. W. (2013). DNA strand break repair and neurodegeneration. *DNA Repair (Amst)* 12. doi:10.1016/j.dnarep.2013.04.008

Sakofsky, C. J., Roberts, S. A., Malc, E., Mieczkowski, P. A., Resnick, M. A., Gordenin, D. A., et al. (2014). Break-induced replication is a source of mutation clusters underlying kataegis. *Cell Rep.* 7, 1640–1648. doi:10.1016/j.celrep.2014.04.053

Sakumi, K. (2019). Germline mutation: De novo mutation in reproductive lineage cells. *Genes Genet. Syst.* 94, 3–12. doi:10.1266/ggs.18-00055

Sanders, S. J., Murtha, M. T., Gupta, A. R., Murdoch, J. D., Raubeson, M. J., Willsey, A. J., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485, 237–241. doi:10.1038/nature10945

Santen, G. W. E., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M., et al. (2012). Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* 44, 379–380. doi:10.1038/ng.2217

Sasani, T. A., Pedersen, B. S., Gao, Z., Baird, L., Przeworski, M., Jorde, L. B., et al. (2019). Large, three-generation human families reveal post-zygotic mosaicism and variability in germline mutation accumulation. *Elife* 8, e46922–e46924. doi:10.7554/eLife.46922

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J. Y., et al. (2020). Large-Scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* 180, 568–584. doi:10.1016/j.cell.2019.12.036

Scally, A. (2016). Mutation rates and the evolution of germline structure. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371, 20150137. doi:10.1098/rstb.2015.0137

Schrider, D. R., Hourmozdi, J. N., and Hahn, M. W. (2011). Pervasive multinucleotide mutational events in eukaryotes. *Curr. Biol.* 21, 1051–1054. doi:10.1016/j.cub.2011.05.013

Ségurel, L., Wyman, M. J., and Przeworski, M. (2014). Determinants of mutation rate variation in the human germline. *Annu. Rev. Genomics Hum. Genet.* 15, 47–70. doi:10.1146/annurev-genom-031714-125740

Sharma, S., Wistuba, J., Pock, T., Schlatt, S., and Neuhaus, N. (2019). Spermatogonial stem cells: Updates from specification to clinical relevance. *Hum. Reprod. Update* 25, 275–297. doi:10.1093/humupd/dmz006

Sirmaci, A., Spiliopoulos, M., Brancati, F., Powell, E., Duman, D., Abrams, A., et al. (2011). Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. *Am. J. Hum. Genet.* 89, 289–294. doi:10.1016/j.ajhg.2011.06.007

Song, J. H. T., Lowe, C. B., and Kingsley, D. M. (2018). Characterization of a human-specific tandem repeat associated with bipolar disorder and schizophrenia. *Am. J. Hum. Genet.* 103 (3), 421–430. doi:10.1016/j.ajhg.2018.07.011

Spencer Chapman, M., Ranzoni, A. M., Myers, B., Williams, N., Coorens, T. H. H., Mitchell, E., et al. (2021). Lineage tracing of human development through somatic mutations. *Nature* 595, 85–90. doi:10.1038/s41586-021-03548-6

Stessman, H. A. F., Xiong, B., Coe, B. P., Wang, T., Hoekzema, K., Fenckova, M., et al. (2017). Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat. Genet.* 49, 515–526. doi:10.1038/ng.3792

Tang, W. W. C., Kobayashi, T., Irie, N., Dietmann, S., and Surani, M. A. (2016). Specification and epigenetic programming of the human germ line. *Nat. Rev. Genet.* 17, 585–600. doi:10.1038/nrg.2016.88

Terekhanova, N. V., Bazykin, G. A., Neverov, A., Kondrashov, A. S., and Seplyarskiy, V. B. (2013). Prevalence of multinucleotide replacements in evolution of primates and Drosophila. *Mol. Biol. Evol.* 30, 1315–1325. doi:10.1093/molbev/mst036

Thornton, C. A., Johnson, K., and Moxley, R. T. (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann. Neurol.* 35, 104–107. doi:10.1002/ana.410350116

Toriello, H. V., and Meck, J. M. (2008). Statement on guidance for genetic counseling in advanced paternal age. *Genet. Med.* 10, 457–460. doi:10.1097/GIM.0b013e318176fabb

Trost, B., Engchuan, W., Nguyen, C. M., Thiruvahindrapuram, B., Dolzhenko, E., Backstrom, I., et al. (2020). Genome-wide detection of tandem DNA repeats that are expanded in autism. *Nature* 586 (7827), 80–86. doi:10.1038/s41586-020-2579-z

Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* 44, 376–378. doi:10.1038/ng.2219

Van Broeckhoven, C. (2010). The future of genetic research on neurodegeneration. *Nat. Med.* 16, 1215–1217. doi:10.1038/nm.2225

Velasco-Miguel, S., Richardson, J. A., Gerlach, V. L., Lai, W. C., Gao, T., Russell, L. D., et al. (2003). Constitutive and regulated expression of the mouse Dinb (Polκ) gene encoding DNA polymerase kappa, *DNA Repair (Amst)*, 2. doi:10.1016/S1568-7864(02)00189-1

Veltman, J. A., and Brunner, H. G. (2012). De novo mutations in human genetic disease. *Nat. Rev. Genet.* 13, 565–575. doi:10.1038/nrg3241

Venn, O., Turner, I., Mathieson, I., De Groot, N., Bontrop, R., and McVean, G. (2014). Nonhuman genetics. Strong male bias drives germline mutation in chimpanzees. *Science* 80, 1272–1275. doi:10.1126/science.344.6189.1272

Vijg, J., and Suh, Y. (2013). Genome instability and aging. *Annu. Rev. Physiol.* 75, 645–668. doi:10.1146/annurev-physiol-030212-183715

Vissers, L. E. L. M., De Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., De Vries, P., et al. (2010). A de novo paradigm for mental retardation. *Nat. Genet.* 42, 1109–1112. doi:10.1038/ng.712

Vissers, L. E. L. M., Van Ravenswaaij, C. M. A., Admiraal, R., Hurst, J. A., De Vries, B. B. A., Janssen, I. M., et al. (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat. Genet.* 36, 955–957. doi:10.1038/ng1407

Wallis, G. A., Starman, B. J., Zinn, A. B., and Byers, P. H. (1990). Variable expression of osteogenesis imperfecta in a nuclear family is explained by somatic mosaicism for a lethal point mutation in the α1(I) gene (COL1A1) of type I collagen in a parent. *Am. J. Hum. Genet.* 46, 1034–1040.

Wang, M., Liu, X., Chang, G., Chen, Y., An, G., Yan, L., et al. (2018). Single-cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. *Cell Stem Cell* 23, 599–614. doi:10.1016/j.stem.2018.08.007

Wang, Y., Busin, R., Reeves, C., Bezman, L., Raymond, G., Toomer, C. J., et al. (2011). X-Linked adrenoleukodystrophy: ABCD1 de novo mutations and mosaicism. *Mol. Genet. Metab.* 104, 160–166. doi:10.1016/j.ymgme.2011.05.016

Wong, L. J. C., Ashizawa, T., Monckton, D. G., Caskey, C. T., and Richards, C. S. (1995). Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.* 56, 114–122.

Wong, W. S. W., Solomon, B. D., Bodian, D. L., Kothiyal, P., Eley, G., Huddleston, K. C., et al. (2016). New observations on maternal age effect on germline de novo mutations. *Nat. Commun.* 7, 10486. doi:10.1038/ncomms10486

Woodworth, M. B., Girskis, K. M., and Walsh, C. A. (2017). Building a lineage from single cells: Genetic techniques for cell lineage tracking. *Nat. Rev. Genet.* 18, 230–244. doi:10.1038/nrg.2016.159

Xu, B., Roos, J. L., Dexheimer, P., Boone, B., Plummer, B., Levy, S., et al. (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat. Genet.* 43, 864–868. doi:10.1038/ng.902

Youssoufian, H. (1996). Natural gene therapy and the Darwinian legacy. *Nat. Genet.* 13, 255–256. doi:10.1038/ng0796-255

Youssoufian, H., and Pyeritz, R. E. (2002). Mechanisms and consequences of somatic mosaicism in humans. *Nat. Rev. Genet.* 3, 748–758. doi:10.1038/nrg906

Yuen, R. K. C., Merico, D., Cao, H., Pellecchia, G., Alipanahi, B., Thiruvahindrapuram, B., et al. (2016). Genome-wide characteristics of de novo mutations in autism. *NPJ Genom. Med.* 1, 160271–1602710. doi:10.1038/npjgenmed.2016.27

Zenzes, M. T. (2000). Smoking and reproduction: Gene damage to human gametes and embryos. *Hum. Reprod. Update* 6, 122–131. doi:10.1093/humupd/6.2.122

Zhang, J. (2022). *News & views Important genomic regions mutate less often.*

Zlotogora, J. (1998). Germ line mosaicism. *Hum. Genet.* 102, 381–386. doi:10.1007/s004390050708