Germline mutation: *de novo* mutation in reproductive lineage cells

Kunihiko Sakumi*

Division of Neurofunctional Genomics, Medical Institute of Bioregulation, Kyushu University,
3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

(Received 14 November 2018, accepted 8 January 2019; J-STAGE Advance published date: 21 March 2019)

Next-generation sequencing (NGS) has been used to determine the reference sequences of model organisms. This allows us to identify mutations by the chromosome number and sequence position where the base sequence has been altered, independent of any phenotypic alteration. Because the re-sequencing method by NGS covers all of the genome, it enables detection of the small number of spontaneous *de novo* germline mutations that occur in the reproductive lineage. The spontaneous mutation rate varies depending on the environment; for example, it increases when 8-oxoguanine accumulates. If the mutation rate (per replication) is greater than 1/genome size ($2n$), at least one mutation would generally occur in each cell division on average, producing cells with a different genome from the parent cell. Organisms with larger genomes and more divisions by cells in the reproductive lineage are expected to show higher mutation rates per generation, if the mutation rate per replication is constant among species. The accumulation of mutations that arose in the genome of ancestor cells has resulted in heterogeneity and diversity among extant species. In this sense, the ability to produce mutations in cells of the reproductive lineage can be considered as a key feature of organisms, even if mutations also present an unavoidable risk.

Key words: germline mutation, mutator mouse, 8-oxoguanine, reproductive lineage cell

INTRODUCTION

What is a mutation? In *DNA Repair and Mutagenesis* (2nd ed.), Friedberg et al. (2006) defined a mutation as being “a heritable change in the sequence of an organism’s genome; the full complement of an organism’s genetic material is referred to as its genome”. In the 20th century, mutation was considered to involve an alteration of a gene that causes a change of a heritable trait, such as color or virus resistance (Luria and Delbruck, 1943; Gluecksohn-Waelsch, 1979). Projects to sequence the whole genome of organisms and next-generation sequencing (NGS) have enabled the establishment of reference sequences of organisms such as human and mouse (IHGSC, 2001; Venter et al., 2001; MGSC, 2002). This allows us to identify any sequence by the chromosome number and base sequence position. Mutations can also be described by the chromosome number and position, along with the particular altered base sequence.

We usually use the term *mutation* for a variety of pur-

poses with divergent meanings in various fields of the life sciences. For classifying types of mutation, we use terms such as base substitution, insertion, deletion, amplification, inversion and translocation. To differentiate their functional effects, we use terms such as loss-of-function, gain-of-function and dominant negative mutations. Terms such as silent, missense, nonsense and truncating mutations are also used to describe the effects of mutations in protein-coding genes. Moreover, in genomic analysis, the terms synonymous, non-synonymous, intronic and intergenic mutations are used. We also sometimes compare spontaneous mutations and induced mutations, as well as germline mutations and somatic cell mutations. These criteria are summarized in the textbook (Friedberg et al., 2006). In this minireview, I use the definition that a mutation is a genetically transmittable change in the sequence of the genome passed from mother cell to daughter cell and from parent to child.

REPRODUCTIVE CELLS AND GERMLINE MUTATIONS IN MAMMALS

In Japanese, we use the term “*seishoku*” to describe reproduction in terms of both individual reproduction in

Edited by Kenji Ichiyanagi

* Corresponding author. E-mail: sakumi@bioreg.kyushu-u.ac.jp

DOI: <http://doi.org/10.1266/ggs.18-00055>

the population, and that in the germline, namely the cell lineage specific for gamete production, which is distinct from somatic cells in the body. Generally, a process in which individuals (i.e., parents) produce a new individual (i.e., offspring) is called reproduction, which is one of the most important features of organisms. The cells of the reproductive lineage of each organism are immortal and responsible for the continuity of life.

There are two types of reproduction: sexual reproduction and asexual reproduction. In sexual reproduction, meiosis and the following conjugation (mating) process are required, in addition to mitosis. In contrast, in asexual reproduction, only mitosis is used and there is no conjugation process. Higher plants such as angiosperms can use both sexual reproduction (seeding) and asexual reproduction (vegetative reproduction). In *Escherichia coli* and other bacteria, each cell behaves as a reproductive cell, through asexual reproduction. All of the mutations occurring in a parent cell are continuously transmitted to subsequent generations. The mutations occurring in each bacterium thus behave as reproductive cell mutations. The occurrence of spontaneous mutation in reproductive cells produces genetic polymorphism, which provides the raw material for evolution (Williams, 2014).

In nature, the reproductive ability of mammals is maintained by specific cells, sperm and egg (gametes). In contrast to the case in unicellular organisms, the reproductive cell mutations of mammals are the cumulative

total of all mutations that have occurred during the series of cell divisions in the reproductive lineage starting from a fertilized egg. Only those mutations that have occurred and accumulated in reproductive lineage cells, from a fertilized egg to gametes, are transmitted to the next generation.

As shown in Fig. 1, the reproductive lineage cells are separated from nonreproductive cell lineage cells (somatic cells) during development, and form germ lineage cells. The mutations that occur in the nonreproductive lineage cells, after the stage of divergence from germ lineage cells, will be restricted to that individual. These are called somatic cell mutations and are not transmitted to the offspring. Reproductive lineage cell mutations can be defined as mutations occurring in the reproductive lineage from a fertilized egg to gametes. Such mutations are transmitted to the next generation, but do not exist in the parents. These are often called (*de novo*) germline mutations. In this review, I often use the term “germline mutation” to describe a mutation in a reproductive lineage cell.

Such mutations are classified into two groups: pre-germ-cell-stage mutations and germ-cell-stage mutations. The latter group occur in primordial germ cells (PGCs) and cells derived from them after the point of divergence from somatic cells, and are generally called germline mutations. In contrast, the former group of mutations occur before this branching point, so they can exist in both gametes and somatic cells at the same time. Therefore,

Fig. 1. Reproductive lineage cells and germline mutations. The reproductive cell lineage from fertilized egg to differentiated gamete constitutes a generation. Mutations occurring during the reproductive lineage are classified into two groups: pre-germ-cell-stage mutations and germ-cell-stage mutations. The latter occur in the primordial germ cells (pGCs) and their offspring after the point of divergence from somatic cells, generally being referred to as germline mutations. The former occur before this point of divergence, so such mutations can exist in both gametes and somatic cells at the same time.

in sexual reproduction of higher multicellular organisms such as mammals, to determine the biological significance of a mutation, it is important to determine when and where the mutation originated. Spontaneous *de novo* germline mutations in humans and mice are intensively discussed by Ohno (Ohno, 2018).

OCCURRENCE OF MUTATIONS BY DNA REPLICATION: A DNA LESION IS DIFFERENT FROM A MUTATION

The occurrence of a mutation usually requires a replication event. In the case of 8-oxoguanine (8-oxoG), for example, a reactive oxygen species (ROS) attacks a guanine base, resulting in the formation of an 8-oxoG:C pair in the DNA (Kasai and Nishimura, 1984; Kasai, 2016). The 8-oxoG can pair with adenine as well as cytosine, so an 8-oxoG:A pair forms in DNA via replication (Shibutani et al., 1991). During the next round of DNA replication, the adenine in the 8-oxoG:A pair makes a pairing with thymine to form a T:A base pair (Fig. 2A). Because T:A is a canonical base pair, it can be retained stably in the DNA. Consequently, through two rounds of DNA replication, 8-oxoG leads to a G:C to T:A transversion mutation (Cheng et al., 1992; Ohno et al., 2014). Similar to the case with 8-oxoG, most chemically modified (damaged) DNA is not actually mutated itself, but is rather a cause of mutation as a pre-mutagenic lesion. This can result in the formation of a noncanonical base pair, which may be recognized as an abnormal component by DNA repair systems (Friedberg et al., 2006).

The G:C to T:A transversion mutation caused by 8-oxoG requires three steps: (1) the formation of 8-oxoG

(DNA modification), (2) DNA replication of the region including the premutagenic lesion and formation of a noncanonical base pair (mutation intermediate), and (3) formation of a new (different from the original) canonical base pair (fixing of the mutation) in DNA as a mutated allele (Fig. 2B). Such premutagenic lesions are also produced by alkylation, oxidative deamination or tautomeric isomerization, in addition to oxidation (Friedberg et al., 2006). It should be noted that the formation of a DNA lesion (premutagenic lesion) is independent from the fixation of a mutation. During its resting stage, an oocyte may accumulate numerous oxidized guanine nucleotides in its DNA and nucleotide pool. However, these oxidized guanine nucleotides do not cause mutations until DNA replication is initiated after fertilization. Similarly, DNA lesions that accumulate in resting stem cells do not cause mutations until DNA replication has restarted.

WHEN AND WHERE DO GERMLINE MUTATIONS OCCUR?

In the adult body, most cells do not divide frequently. However, during human development, starting from a fertilized egg, more than 45 cell divisions are required to form a body composed of 3×10^{13} cells (Sender et al., 2016). Upon reaching adulthood, only limited numbers and kinds of cells divide to maintain the size and complexity of the body. In this sense, the cells in which somatic mutations can occur and the timing of this are restricted in the body.

Moreover, in the case of germline mutations, the timing and site at which such mutations can occur are highly controlled. The reproductive lineage starts from

Fig. 2. 8-OxoG as a pre-mutagenic lesion. (A) 8-OxoG causes a G:C to T:A transversion mutation. An 8-oxoG:C pair is formed by the oxidation of G in DNA. Because 8-oxoG can pair with adenine as well as cytosine, the 8-oxoG:A pair forms via replication. In the next replication, the adenine in the 8-oxoG:A pair makes a pairing with thymine to form a T:A base pair. Because this T:A base pair is a canonical base pair, it is retained stably in the DNA. (B) Mutation is thus completed by two rounds of replication via a noncanonical base pair.

a fertilized egg located in the mother's body. Cell propagation continues and an epiblast is formed for the preparation of the future body (De Felici, 2013; Behringer et al., 2014). In mouse, the primordial germ cells (pGCs) that are derived and separated from the epiblast continue to divide for the formation of gametes. During oogenesis, the pGCs, which first appear around 6.25 days post-coitum (dpc), increase in number to about 2.5×10^4 cells by 13.5 dpc, and they enter meiosis (Saitou and Yamaji, 2012). This ceases at prophase of the first division, and the cells enter the resting stage. Therefore, the period in which germline mutations can become fixed by replication is limited to between fertilization (0.5 dpc) and the end of the propagation of pGCs (13.5 dpc). At the time when germline mutation occurs in pGCs, the mother's body holding the pGCs is present in the pregnant grandmother's uterus. A maternal germline mutation observed in a child will thus have occurred within the grandmother's body (Fig. 3).

In spermatogenesis, the pGCs that increase in number in early embryogenesis also enter the mitotic resting stage (Saitou and Yamaji, 2012). However, the spermatogonia restart their cell division around five days after birth, and maintain this division to supply mature sperm continuously for a long period. Therefore, the number of DNA replications for the preparation of fertilization-ready gametes is greater in spermatogenesis than in oogenesis. Thus, there should be more paternal mutations than maternal ones.

Indeed, according to a human study, 75% of germline mutations are of paternal origin, and this proportion increases with age (Goldmann et al., 2016). Specifically, it was reported that the number of cell divisions associated with oogenesis is about 22 at the embryonic stage, while in spermatogenesis, there are 23 cell divisions/year in addition to 30 cell divisions during embryogenesis (Goriely, 2016). According to a calculation by the author, paternal mutations are increased by aging at a

Fig. 3. A maternal germline mutation observed in a child occurred in the grandmother's body. When a germline mutation occurs in pGCs in the mother, the mother's body is present in the pregnant grandmother's uterus. Red color indicates the cells carrying the mutation.

rate of one mutation/year. In contrast, in the case of oogenesis, maternal mutations increase with age at a rate of 0.25 mutations/year (Goldmann et al., 2016). It is likely that during the resting stage, DNA and nucleotides in oocytes continuously suffer damage due to oxidative radicals and other chemical/physical attacks. The lesions that result from this and accumulate during aging cause mutations when DNA replication restarts after fertilization. Besides the mutations associated with DNA replication, other mutations occurring without DNA replication, such as in nonhomologous end joining, may also affect the total mutation rate associated with reproduction.

Particular mutations that lead to cell death are by definition not transmitted to progeny cells. In addition, mutations that cause embryonic lethality or infant lethality (i.e., death before reaching reproductive age) are not transmitted to offspring as germline mutations. Thus, some germline mutations are not counted by standard methods for calculating the mutation rate. In this sense, the observable rate of germline mutations is expected to be lower than the rate of somatic cell mutations.

It has been reported that a higher speed of DNA replication results in a higher mutation rate (Maya-Mendoza et al., 2018). Considering this, early embryogenesis should be a mutation-prone stage because it involves a very high rate of replication (De Felici, 2013; Behringer et al., 2014). However, we cannot rule out the existence of germline-specific systems for avoiding mutations in reproductive lineage cells. The mutation rate of reproductive lineage cells is probably not constant among the differentiation stages; even within each cell type, it may fluctuate from one cell to the next. It remains unclear when and how many reproductive cell mutations occur during development.

STRATEGY FOR DETECTING GERMLINE MUTATIONS

We can determine the nucleotide sequence of model animals by performing resequencing, and detect any changes from the reference sequence as sequence variations. By defining mutations according to their genome location, we can characterize and evaluate the mutations across the genome independently of bias, such as would occur if analyzing the phenotypic change of individuals.

To evaluate the occurrence of mutations, we have to compare sequences between parent cells and progeny cells. However, it remains difficult to determine the whole genome sequence from a single cell. Practically, 100 ng to 1 μ g of DNA is required for NGS analysis, which is equivalent to the genomic DNA of $1.75 \times 10^4 - 10^5$ cells (human female, 2n).

The genome sequence of a parent cell is usually estimated based on the sequence data of progeny cells. As

shown in Fig. 4, for example, after 25 cell divisions, each G_{25} cell has its own *de novo* mutations (DNMs), in addition to the original sequence variations (SVs) of the G_0 cell. Because the loci of DNMs produced in each cell are random and not identical among the cell population, the signal of each DNM is hidden within the signals of the DNA of 10^4 G_{25} cells. Thus, we can simulate the G_0 DNA sequence by using the DNA sequence data from a mixture of G_{25} cells. Similarly, each G_{25} cell (indicated by blue or red in Fig. 4) contains DNMs that have accumulated over 25 cell divisions. Similarly, after clonal expansion for 25 more cell divisions, the $SV_{G_{25}}$ can be estimated using the DNA sequences of G_{50} cells. By comparing SVs between G_0 ($\Sigma SV_{G_{25}}$) and G_{25} ($\Sigma SV_{G_{50}}$), we can determine the mutations that occurred between G_0 and G_{25} .

For example, trio analysis involves comparing the single-nucleotide variations (SNVs) of offspring with those of their parents, and selecting SNVs only found in the offspring as DNMs. According to this strategy, the mutation rate in humans has been estimated to be about 1.2×10^{-8} mutations/bp/generation (Kong et al., 2012). In contrast to the case in human pedigrees, we can systematically mate genetically manipulated animals to prepare the desired pedigree in the laboratory. In such a case, mutations are expected to accumulate through the generations. By increasing the number of generations, more mutations can be detected in the offspring. Uchimura et al. (2015) used this strategy to determine the mutation rate of *Pole1* mutant mice.

These methods appear extremely promising in theory, but are associated with certain problems when applied in practice. For example, to determine the mutations that have occurred during a generation, we have to compare the genomes of fertilized eggs between parents and offspring. However, given the difficulty in determining the whole genome sequence from a single cell, we usually use DNA prepared from tissues such as liver, blood or tail. Because these tissues are composed of cells derived from a limited number of stem cells, some somatic mutations may be concentrated in them and observed as noise, especially in the case of mutator mice.

At present, no gold-standard detection pipeline using NGS has been established for DNMs. Thus, confirmation by wet methods such as Sanger sequencing may be required to ascertain whether an identified mutation is true or false. Several parameters, such as sequencing depth and alternate/reference read ratio, could improve the efficiency of detecting DNMs (Masumura et al., 2016a, 2016b). Common standards regarding use of the same parameters and thresholds for the confirmation of mutations in addition to a better DNM detection pipeline would be helpful to compare mutation data among different studies.

Fig. 4. Strategy to characterize *de novo* mutations. For standard NGS sequence analysis, 100 ng of DNA equivalent to 1.75×10^4 progeny cells is required. To estimate the sequence of a G_0 cell, we have to use the offspring cell DNA expanded from the G_0 cell. During 25 cell divisions, each G_{25} cell has its own *de novo* mutations (DNMs) in addition to the original sequence variations (SVs) of the G_0 cell. However, the loci of new DNMs produced in each cell are random and differ from one cell to the next. The signal of each DNM is hidden in the DNA signals of 10^4 G_{25} cells. Each G_{25} cell (indicated by blue or red) contains an independent set of DNMs accumulated during 25 cell divisions. After clonal expansion for 25 more cell divisions, the $SV_{G_{25}}$ can be reconstructed using the DNA sequence of G_{50} cells. By comparing the SVs between G_0 ($\sum SV_{G_{25}}$) and G_{25} ($\sum SV_{G_{50}}$), we can identify the mutations that occurred between G_0 and G_{25} .

SPONTANEOUS MUTATIONS AND MUTATOR GENES

In aqueous conditions, DNA gradually degrades via deamination and depurination (Frederico et al., 1990; Lindahl, 1993). Such DNA lesions are one of the causes of spontaneous mutation. In contrast to induced mutations caused by exogenous factors such as radiation or exposure to chemicals, spontaneous mutations are caused by endogenous factors such as aberrant DNA polymerase or DNA repair genes, or the production of DNA lesions via oxidation or deamination under physiological conditions in cells. It is often difficult to distinguish endogenous from exogenous factors because ROS, a group of DNA-modifying molecules, are generated by intracellular mitochondrial respiration, and also as a result of attack by exogenous mutagens (Kawamura and Kobayashi, 2018).

Mutator phenotypes are considered as resulting from mutations that increase the spontaneous mutation rate relative to the wild type. Twenty-five mutator genes were summarized in a review of *E. coli* mutators (Table 1) (Horst et al., 1999). These genes encode DNA polymerases, mismatch repair proteins, lesion-specific DNA glycosylases and nucleoside triphosphatases, among others. In yeast, it has been shown that mutants of mis-

match and other DNA repair genes possess a mutator phenotype (Serero et al., 2014). In the case of plants, it is known that transposons are responsible for strong mutator activity (Lisch, 2002). It is also reported that mismatch repair deficiency increases microsatellite instability in plants (Spampinato et al., 2009).

Some of the mammalian ortholog mutants of *E. coli* mutator genes, such as *Pold1*, *Msh2* and *Mutyh*, have also been shown to possess a mutator phenotype (Fishel, 2001; Hirano et al., 2003; Uchimura et al., 2015). Here, I present the *Mth1*, *Ogg1*, *Mutyh* triple-KO (TOY-KO) mouse strain as an example of a mutator mouse strain. *Mth1*, *Ogg1* and *Mutyh* genes encode proteins that act in the avoidance of 8-oxoG-caused mutation (Ohno et al., 2014). MTH1 (*mutT* homolog 1, *NUDT1*) degrades 8-oxodGTP in the nucleotide pool to prevent its incorporation into DNA (Nakabeppu et al., 2006). OGG1 (8-oxoG DNA glycosylase) excises 8-oxoG from DNA, and MUTYH (*mutY* homolog, adenine DNA glycosylase) removes adenine that has been misincorporated opposite 8-oxoG in DNA (Nakabeppu et al., 2006). My colleagues and I as well as other groups have reported that mice deficient in these enzymes are prone to develop cancer, indicating a mutator phenotype in somatic cells (Tsuzuki et al., 2001; Sakumi et al., 2003; Sakamoto et al., 2007). *MUTYH* is

Table 1. Mutator genes in *Escherichia coli*

Gene	Gene product	Function	Specificity
<i>dnaQ (mutD)</i>	Epsilon subunit of DNA polymerase III, 3' → 5' exonuclease	Removes incorrectly paired nucleotides during replication (proofreading)	Mostly transversions
<i>polC (dnaE)</i>	Alpha subunit of DNA polymerase III	Correct nucleotide selection and proofreading	All base substitutions and frameshifts
<i>polA</i>	DNA polymerase I	General repair functions	Deletions, frameshifts
<i>mutA, mutC</i>	Mutated glycyl-tRNA	Missense suppression at aspartic acid codons in <i>dnaQ</i>	Similar to <i>mutD</i>
<i>mutT</i>	Nucleoside triphosphatase	Prevents mispairing of 8-oxoG with template A during replication	AT→CG
<i>dam</i>	DNA adenine methyltransferase	Methylation imparts strand specificity	GC→AT, AT→GC, frameshifts
<i>mutS</i>	DNA mismatch recognition	Binds DNA mismatches	GC→AT, AT→GC, frameshifts
<i>mutL</i>	68-kDa protein	Stimulates MutS, MutH and Vsr activity	GC→AT, AT→GC, frameshifts
<i>mutH</i>	Endonuclease	Nicks hemi-methylated GATC sequences	GC→AT, AT→GC, frameshifts
<i>uvrD</i>	DNA helicase II	Strand displacement	GC→AT, AT→GC, frameshifts
<i>mutY</i>	DNA glycosylase	Removes A from 8-oxoG–A or A–G mispairs	GC→TA
<i>mutM (fpg)</i>	DNA glycosylase	Removes 8-oxoG from 8-oxoG–C mispair	GC→TA
<i>miaA</i>	Transferase	Methylthio-isopentyladenosine tRNA modification	GC→TA
<i>sodA, sodB</i>	Superoxide dismutases	Removes superoxide radicals	
<i>oxyR</i>	Regulatory protein	Regulates hydrogen peroxide-inducible genes	AT→TA
<i>nth, nei</i>	Glycosylase and abasic-lyase activity	Removal of oxidized pyrimidine bases	GC→AT
<i>xthA, nfo</i>	Nucleases	5' abasic-endonuclease activity	AT→TA
<i>ung</i>	Uracil glycosylase	Removes U from U–G mispairs	GC→AT
<i>vsr</i>	Endonuclease	Cleaves adjacent to T–G mismatches	GC→AT
<i>ada, ogt</i>	Methyltransferases	Removes methyl groups from O ⁶ -methylguanine in DNA	GC→AT
<i>recA</i>	DNA-binding protein	Catalyzes strand pairing and recombination; co-protease exchange in general activity on LexA, UmuD, etc.	GC→TA, AT→TA
<i>recG</i>	DNA helicase	Branch migration of Holliday junctions	Frameshifts
<i>hns (bglY)</i>	DNA-binding protein	Histone-like protein involved in chromosome organization	Deletions
<i>topB (mutR)</i>	DNA topoisomerase III	Appears to decatenate chromosomes	Deletions between small repeats
<i>ssb</i>	Single-stranded DNA-binding protein	Protects single-stranded DNA	Point mutations, rearrangements

cited from Horst et al., 1999.

responsible for MUTYH-associated polyposis in humans (Venesio et al., 2012).

We prepared TOY-KO mice deficient for all three enzymes that are active in the avoidance of 8-oxoG-induced mutations (Ohno et al., 2014). TOY-KO mice accumulated 8-oxoG in the nuclear DNA of their gonadal cells, and exhibited heritable traits such as hydrocephalus and familial tumors. Whole-exome analysis also showed the increased presence of G-to-T mutations, occurring at a rate of 2×10^{-7} mutations/base/generation, in the

offspring of TOY-KO mice. Compared with the spontaneous base substitution mutation rate of wild-type mice (C57BL/6J) reported by Uchimura et al. (2015) (5.4×10^{-9} /mutations/base/generation), the mutation rate of TOY-KO mice is 37 times higher. In the wild-type mice, the contribution of G-to-T transversion mutation in the base substitution mutations was 9.3% (Uchimura et al., 2015), similar to the human population (9.0%) (Kong et al., 2012). Analysis of the mutation spectra observed in TOY-KO mice indicated a distinct feature in which 99%

Fig. 5. Mechanism for preventing 8-oxoG-induced mutations in mammals. MTH1 degrades 8-oxodGTP in the nucleotide pool to prevent the incorporation of the oxidized nucleotide into DNA. OGG1 excises 8-oxoG from DNA, and MUTYH removes misincorporated adenine opposite 8-oxoG in DNA.

of the mutations were G-to-T transversions. Because G-to-T mutation had specifically increased in TOY-KO mice lacking the ability to avoid 8-oxoG-induced mutations (Fig. 5), we concluded that 8-oxoG is causative of spontaneous G-to-T mutation in mouse germ lineage cells (Ohno et al., 2014). By tracing the mutated alleles in a pedigree, we observed the fates of DNMs, in which the mutated alleles appeared, were transmitted, or disappeared, which are the typical fates of novel mutations in evolutionary processes.

In TOY-KO mice, the amount of 8-oxoG produced in cellular DNA does not differ from that in wild-type mice. Instead, the enzymes that act in preventing mutations caused by 8-oxoG are removed. This shows that 8-oxoG is endogenously produced and causes heritable spontaneous *de novo* G-to-T transversion mutations during a generation (the period from a fertilized egg to gametes) through the pre-germ and germ cell lineage.

In experimental animals, unlike in humans, it is easy to control genetic background, gene activity and mating to construct a desired pedigree. Mutator mutations increase the likelihood of detecting other spontaneous mutations and make it possible to clarify the mechanism by which each mutation has occurred. Mutator mouse strains are also useful as lines for accumulating mutations, in which various mutants spontaneously emerge as model animals with heritable traits.

CONCLUSION

The occurrence of spontaneous germline mutations,

alongside selection for survival, is important for evolution. However, it can also cause genetic diseases in individuals. If tumorigenesis is caused by an accumulation of mutations in cells, germline mutations accumulating over the generations would also increase the risk of tumors. Even in organisms that reproduce asexually, a significant number of mutations should occur and accumulate during a series of cell divisions (replications), if the mutation rate is greater than $1/\text{genome size } (2n)$. Therefore, mutations accompanying reproduction play an important role in genome diversification in populations.

Theoretically, if the mutation rate (per replication) exceeds $1/\text{genome size } (2n)$, at least one mutation will occur on average in each cell division, producing cells with a different genome from that in the parent cell. In humans, if the mutation rate is 1.67×10^{-10} mutations/bp/replication or more, more than one mutation should occur on average during each replication, and mutation should accompany every cell division. This indicates that each of the cells comprising a body should typically have a different genome from the other cells, forming a genetic mosaic. Organisms with a larger genome and more divisions among reproductive lineage cells are expected to show a higher mutation rate per generation, if the mutation rate per replication is constant among species. The concept that there is a risk of mutation accompanying each cell division also suggests that mutations would accumulate during the *in vitro* expansion of stem cells for medical applications, in addition to an increase of germline mutations occurring among older parents. To

develop and establish methods to control the mutation rate at a lower level could overcome these problems and improve the quality of health during the lifetime.

The amounts of 8-oxoG in tissue DNA and of ROS in cells vary in response to the oxygen concentration in the air and mitochondrial activity (Puente et al., 2014). Although the contribution of G-to-T transversion to substitution mutations is about 9% in *de novo* germline mutations in mammals, the 8-oxoG DNA repair enzymes MutM (FPG/OGG1) and MutY are conserved in a wide variety of species from cyanobacteria to humans (Aburatani et al., 1997; Radicella et al., 1997; Rosenquist et al., 1997; Ohtsubo et al., 1998; Nakamura et al., 2002; Dufresne et al., 2003; Hirano et al., 2003). It is likely that the oxidative environment accelerates the genetic diversity of species. The mutations caused by 8-oxoG should have been well controlled over evolutionary history in response to an increase in the atmospheric O₂ concentration.

The descent of reproductive lineage cells can be traced directly back to ancestral cells that are thought to have emerged a few billion years ago, a lineage that will continue into the future as long as life remains. In addition to mutations caused by 8-oxoG, there are numerous other causes and mechanisms by which mutations are produced, in manners dependent on or independent of DNA replication (Lieber, 2010; Mjelle et al., 2015). Various changes occurring in the genome over the course of evolution have resulted in heterogeneity and diversity among extant species. In this sense, the ability to produce mutations in reproductive lineage cells can be considered a key feature of organisms.

I thank Dr. M. Ohno for intensive discussions concerning germline mutations. This work was supported by JSPS KAKENHI Grant Numbers 15H04298 and 17K19913. I also thank Edanz (www.edanzediting.co.jp) for editing the English text of a draft of this manuscript.

REFERENCES

- Aburatani, H., Hippo, Y., Ishida, T., Takashima, R., Matsuba, C., Kodama, T., Takao, M., Yasui, A., Yamamoto, K., Asano, M., et al. (1997) Cloning and characterization of mammalian 8-hydroxyguanine-specific DNA glycosylase/apurinic, apyrimidinic lyase, a functional mutM homologue. *Cancer Res.* **57**, 2151–2156.
- Behringer, R., Gertsenstein, M., Nagy, K. V., and Nagy, A. (2014) *Manipulating the Mouse Embryo: A Laboratory Manual*, 4th ed. Cold Spring Harbor Laboratory Press, New York, USA.
- Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A. (1992) 8-hydroxyguanine, an abundant form of oxidative DNA damage, causes G→T and A→C substitutions. *J. Biol. Chem.* **267**, 166–172.
- De Felici, M. (2013) Origin, migration, and proliferation of human primordial germ cells. In: *Oogenesis*. (eds.: Coticchio, G., Albertini, D. F., and De Santis, L.), pp. 19–37. Springer, London, UK.
- Dufresne, A., Salanoubat, M., Partensky, F., Artiguenave, F., Axmann, I. M., Barbe, V., Duprat, S., Galperin, M. Y., Koonin, E. V., Le Gall, F., et al. (2003) Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome. *Proc. Natl. Acad. Sci. USA* **100**, 10020–10025.
- Fishel, R. (2001) The selection for mismatch repair defects in hereditary nonpolyposis colorectal cancer: revising the mutator hypothesis. *Cancer Res.* **61**, 7369–7374.
- Frederico, L. A., Kunkel, T. A., and Shaw, B. R. (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry* **29**, 2532–2537.
- Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T. (2006) *DNA Repair and Mutagenesis*, 2nd ed. ASM Press, Washington D. C., USA.
- Gluecksohn-Waelsch, S. (1979) Genetic control of morphogenetic and biochemical differentiation: lethal albino deletions in the mouse. *Cell* **16**, 225–237.
- Goldmann, J. M., Wong, W. S. W., Pinelli, M., Farrah, T., Bodian, D., Stittrich, A. B., Glusman, G., Vissers, L. E. L. M., Hoischen, A., Roach, J. C., et al. (2016) Parent-of-origin-specific signatures of *de novo* mutations. *Nat. Genet.* **48**, 935–939.
- Goriely, A. (2016) Decoding germline *de novo* point mutations. *Nat. Genet.* **48**, 823–824.
- Hirano, S., Tominaga, Y., Ichinoe, A., Ushijima, Y., Tsuchimoto, D., Honda-Ohnishi, Y., Ohtsubo, T., Sakumi, K., and Nakabeppu, Y. (2003) Mutator phenotype of MUTYH-null mouse embryonic stem cells. *J. Biol. Chem.* **278**, 38121–38124.
- Horst, J. -P., Wu, T., and Marinus, M. G. (1999) *Escherichia coli* mutator genes. *Trends Microbiol.* **7**, 29–36.
- International Human Genome Sequencing Consortium (IHGSC) (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.
- Kasai, H. (2016) What causes human cancer? Approaches from the chemistry of DNA damage. *Genes Environ.* **38**, 19.
- Kasai, H., and Nishimura, S. (1984) Hydroxylation of deoxyguanosine at the C-8 position by ascorbic acid and other reducing agents. *Nucleic Acids Res.* **12**, 2137–2145.
- Kawamura, K., Qi, F., and Kobayashi, J. (2018) Potential relationship between the biological effects of low-dose irradiation and mitochondrial ROS production. *J. Radiat. Res.* **59**, ii91–ii97.
- Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S. A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., et al. (2012) Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475.
- Lieber, M. R. (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715.
- Lisch, D. (2002) *Mutator* transposons. *Trends Plant Sci.* **7**, 498–504.
- Luria, S. E., and Delbruck, M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491–511.
- Masumura, K., Toyoda-Hokaiwado, N., Ukai, A., Gondo, Y., Honma, M., and Nohmi, T. (2016a) Estimation of the frequency of inherited germline mutations by whole exome sequencing in ethyl nitrosourea-treated and untreated *gpt* delta mice. *Genes Environ.* **38**, 10.

- Masumura, K., Toyoda-Hokaiwado, N., Ukai, A., Gondo, Y., Honma, M., and Nohmi, T. (2016b) Dose-dependent *de novo* germline mutations detected by whole-exome sequencing in progeny of ENU-treated male *gpt* delta mice. *Mutat. Res.* **810**, 30–39.
- Maya-Mendoza, A., Moudry, P., Merchut-Maya, J. M., Lee, M., Strauss, R., and Bartek, J. (2018) High speed of fork progression induces DNA replication stress and genomic instability. *Nature* **559**, 279–284.
- Mjelle, R., Hegre, S. A., Aas, P. A., Slupphaug, G., Drabløs, F., Sætrom, P., and Krokan, H. E. (2015) Cell cycle regulation of human DNA repair and chromatin remodeling genes. *DNA Repair* **30**, 53–67.
- Mouse Genome Sequencing Consortium (MGSC) (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562.
- Nakabeppu, Y., Sakumi, K., Sakamoto, K., Tsuchimoto, D., Tsuzuki, T., and Nakatsu, Y. (2006) Mutagenesis and carcinogenesis caused by the oxidation of nucleic acids. *Biol. Chem.* **387**, 373–379.
- Nakamura, Y., Kaneko, T., Sato, S., Ikeuchi, M., Katoh, H., Sasamoto, S., Watanabe, A., Iriguchi, M., Kawashima, K., Kimura, T., et al. (2002) Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. *DNA Res.* **9**, 123–130.
- Ohno, M. (2018) Spontaneous *de novo* germline mutations in humans and mice: rates, spectra, causes and consequences. *Genes Genet. Syst.* ggs.18-00015.
- Ohno, M., Sakumi, K., Fukumura, R., Furuichi, M., Iwasaki, Y., Hokama, M., Ikemura, T., Tsuzuki, T., Gondo, Y., and Nakabeppu, Y. (2014) 8-oxoguanine causes spontaneous *de novo* germline mutations in mice. *Sci. Rep.* **4**, 4689.
- Ohtsubo, T., Matsuda, O., Iba, K., Terashima, I., Sekiguchi, M., and Nakabeppu, Y. (1998) Molecular cloning of AtMMH, an *Arabidopsis thaliana* ortholog of the *Escherichia coli* *mutM* gene, and analysis of functional domains of its product. *Mol. Gen. Genet.* **259**, 577–590.
- Puente, B. N., Kimura, W., Muralidhar, S. A., Moon, J., Amatruda, J. F., Phelps, K. L., Grinsfelder, D., Rothermel, B. A., Chen, R., Garcia, J. A., et al. (2014) The oxygen-rich postnatal environment induces cardiomyocyte cell-cycle arrest through DNA damage response. *Cell* **157**, 565–579.
- Radicella, J. P., Dherin, C., Desmaze, C., Fox, M. S., and Boiteux, S. (1997) Cloning and characterization of *hOGG1*, a human homolog of the *OGG1* gene of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **94**, 8010–8015.
- Rosenquist, T. A., Zharkov, D. O., and Grollman, A. P. (1997) Cloning and characterization of a mammalian 8-oxoguanine DNA glycosylase. *Proc. Natl. Acad. Sci. USA* **94**, 7429–7434.
- Saitou, M., and Yamaji, M. (2012) Primordial germ cells in mice. *Cold Spring Harb. Perspect. Biol.* **4**, a008375.
- Sakamoto, K., Tominaga, Y., Yamauchi, K., Nakatsu, Y., Sakumi, K., Yoshiyama, K., Egashira, A., Kura, S., Yao, T., Tsuneyoshi, M., et al. (2007) MUTYH-null mice are susceptible to spontaneous and oxidative stress-induced intestinal tumorigenesis. *Cancer Res.* **67**, 6599–6604.
- Sakumi, K., Tominaga, Y., Furuichi, M., Xu, P., Tsuzuki, T., Sekiguchi, M., and Nakabeppu, Y. (2003) *Ogg1* knockout-associated lung tumorigenesis and its suppression by *Mth1* gene disruption. *Cancer Res.* **63**, 902–905.
- Sender, R., Fuchs, S., and Milo, R. (2016) Revised estimate for the number of human and bacteria cells in the body. *PLoS Biol.* **14**, e1002533.
- Serero, A., Jubin, C., Loeillet, S., Legoux-Né, P., and Nicolas, A. G. (2014) Mutational landscape of yeast mutator strains. *Proc. Natl. Acad. Sci. USA* **111**, 1897–1902.
- Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* **349**, 431–434.
- Spampinato, C. P., Gomez, R. L., Galles, C., and Lario, L. D. (2009) From bacteria to plants: a compendium of mismatch repair assays. *Mutat. Res.* **682**, 110–128.
- Tsuzuki, T., Egashira, A., Igarashi, H., Iwakuma, T., Nakatsuru, Y., Tominaga, Y., Kawate, H., Nakao, K., Nakamura, K., Ide, F., et al. (2001) Spontaneous tumorigenesis in mice defective in the *MTH1* gene encoding 8-oxo-dGTPase. *Proc. Natl. Acad. Sci. USA* **98**, 11456–11461.
- Uchimura, A., Higuchi, M., Minakuchi, Y., Ohno, M., Toyoda, A., Fujiyama, A., Miura, I., Wakana, S., Nishino, J., and Yagi, T. (2015) Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Res.* **25**, 1125–1134.
- Venesio, T., Balsamo, A., D'Agostino, V. G., and Ranzani, G. N. (2012) *MUTYH*-associated polyposis (MAP), the syndrome implicating base excision repair in inherited predisposition to colorectal tumors. *Front. Oncol.* **2**, 83.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., et al. (2001) The sequence of the human genome. *Science* **291**, 1304–1351.
- Williams, A. B. (2014) Spontaneous mutation rates come into focus in *Escherichia coli*. *DNA Repair* **24**, 73–79.