ON A BREAK WITH THE X: THE ROLE OF REPAIR OF DOUBLE-STRANDED DNA BREAKS IN X-LINKED DISEASE

Laura De Caris, Lucia Cecceroni and Hemanth Tummala
University of Abertay Dundee, School of Contemporary Sciences, Scotland, UK
Correspondence to: Hemanth Tummala
E-mail: H.Tummala@abertay.ac.uk

ABSTRACT

The problem of managing free reactive DNA ends in eukaryotic cells has resulted in the development of a number of mechanisms in order to ensure that free ends are rendered non-reactive, or that the double-strand DNA breaks generating the free ends are promptly repaired; or that if the cell requires the introduction of double-strand breaks in its DNA at some point in its life cycle, the generated free DNA ends would be subject to strictest management. Different mechanisms for dealing with double-strand breaks in DNA exist, contributing to the maintenance of the exquisite balance between entering an evolutionary dead end resulting from extremely faithful DNA repair and genome instability resulting from relaxed control of repair. Recombination-based mechanisms related to the mechanism for repair of double-strand breaks and template slippage during replication are believed to be able to account for multiplication of parts of the genome, generating regions of homology which may serve as sites for mispairing during meiotic and mitotic recombination, and resulting, in turn, in translocations, deletions and inversions. This is especially true for the X chromosome, which engages in homologous recombination with the partnering X chromosome in female meiosis, but will participate in only limited partial genomic exchange with the Y chromosome in male meiosis, providing ample opportunities for mispairing and non-allelic recombination. Several X-linked monogeneous diseases and even some chromosome diseases such as variants of Turner syndrome have been found to be related to non-allelic homologous recombination during male meiosis, resulting in various rearrangements of the X chromosome. Apparently, the mechanisms which protect the mammalian genomes from stagnation or, alternatively, from hypermutability, may not always perform adequately. The genome rearrangement disorders may be unfortunate manifestations of the ongoing process of reorganization of genetic architecture in the eukaryotic genomes, which is inherent to evolution.

Double-strand DNA breaks – the breakup that hurts both ways

Double-strand DNA breaks (DSB) in DNA occur when the phosphodiester backbone of both the strands of the same DNA helix is broken, the breakpoints being in proximity of each other so that the broken ends may become physically separated. This may happen for a variety of reasons – physiological and pathological alike – and indeed occurs, albeit not very frequently, throughout the life of the cell. DSB are considered be a particularly genotoxic type of damage, as they not only cause serious alteration of the DNA structure, but also generate free (unprotected) DNA ends. The latter are reactive and would readily interact with other DNA ends or intact DNA, resulting in genomic fusions, deletions, insertions, duplications, allelic conversion, loss of heterozygocity, in other words, overall genomic instability which in the short term may lead to the breakdown of eukaryotic chromosomes into smaller fragments, and in the long term may result in inherited disease or disease of somatic origin, such as cancer (7, 72, 87). As a rule, eukaryotic genomes are not tolerant to the presence of

DSB, and a burden of several double-strand breaks per genome is known to cause rerouting of the cell programme of a non-cancerous cell to the apoptotic pathway.

The problem of safekeeping and managing free reactive DNA ends in eukaryotic cells has brought the development of a number of mechanisms in order to ensure that free ends are protected and therefore non-reactive (as is the case with the physical ends of the chromosomes which are protected by the telomere complex); or that the double-strand breaks are promptly repaired (for which several mechanisms exist); or that when the cell really requires the presence of DSB at some point in its life cycle, the control over the cell programme and the management of the generated free ends would be very rigorous. The general rule is that unprotected DNA ends are sought for and processed where possible, or in most cases the cell is sacrificed. This 'ultimate control' policy, however, has some inherent issues, which may produce different outcomes.

Shaking it and actually breaking it – events that lead to double-strand breaks

DSB may occur in the genome for a variety of reasons, endogenous and exogenous in their origin, and physiological and pathological in their nature. Certain clastogenic agents, such as ionizing radiation (short-wavelength UV included)

and radiomimetic chemicals increase the rate of occurrence of double-strand breaks in DNA (70). The presence of Z-DNA has been shown to increase the risk for occurrence of double-strand breaks (42). Meiotic and mitotic recombination, apparently the only mechanisms for generation of genetic variability that Nature deemed necessary to specifically designate, operate by creating and subsequently repairing double-stranded breaks. Shortening of telomeres of somatic cells as they move towards ageing creates potential targets for the cellular repair machinery. Finally, some types of cells explicitly require generation of DSB at some point of their life cycle. In any case, double-strand breaks in DNA are usually meticulously tracked and dealt with, as the potential for generating lasting damage is estimated to be much larger for DSB than for any other type of damage.

Unprotected chromosome ends
The telomeric complex at the ends of eukaryotic chromosomes serves multiple purposes, besides the obvious role in keeping the changes in the genome content to the bare minimum. Among these purposes specifically prominent is keeping the reactive DNA ends away from other DNA and preventing the cellular machinery for repair of DSB from being activated by the presence of natural chromosome ends. The length of the telomeres in the chromosomes of normal somatic cells is usually a good indicator of how close the cell is to replicative senescence (30), that is, through how many more telomere shortenings the cell can safely go without risking triggering of genome instability and, ultimately, cancerous growth. It has been shown, however, that instead of launching the process of ageing at the first appearance of shortened telomeres in a cell, these tend to accumulate in the cell for a while before the onset of actual senescence (39, 46), making the feat of assessing the median length of telomeres in a cell population and correlating it to any indices pertaining to the status of the cell population, the tissue or the organism very difficult indeed (5, 60, 75). Recently, it has been determined that at least five dysfunctional telomeres are required in a human cell in order to trigger p53-dependent senescence (40). Apparently, the risk for genomic instability due to telomere attrition is actually tolerated by the genome to a certain point before launching any of the mechanisms eliminating the cell containing chromosomes with shortened telomeres from the population.

Events in the cellular life cycle which generate free DNA ends
Double-strand breaks in DNA may occur as a result of biological processes in the cell – not necessarily pathological ones. The prime example for this is the differentiation of T-cells and B-cells of the immune system, which need to introduce multiple DSB in their DNA in order to carry out the requisite genomic rearrangements for V(D)J recombination. Failure to do so usually results in severe immunodeficiency states. The process of genomic rearrangements in immune cells is, however, under strict control. Basically, any cell that has not received a designated survival signal would proceed to apoptosis by default (57, 80). This is rather severe selection, and indeed only about 2% of all immature double-positive thymocytes survive the rearrangement of their T-cell receptors.

Some types of differentiated cells which are exposed to high levels of oxidative damage (and, therefore, tend to experience significant amount of double-strand breaks as a consequence of the impact of the generated reactive oxygen species), such as neurons, tend to carry out preferential repair of transcribed genes at the expense of the global genome repair (19, 68). This presumably serves a dual purpose. On the one hand, since those cells do not generally divide, they are unlikely to replicate their damaged DNA, and are therefore in no need to repair genome regions that are not currently under transcription. On the other hand, this may also ensure that all DSB in critically important regions of the genome (currently transcribed genes) are repaired in a timely fashion, so that the need to eliminate cells with damaged DNA from a cell population that is supposed to last for a long time and is not likely to be replenished would arise infrequently. Indeed, the current theories of brain ageing place a strong emphasis on the link between accumulation of DNA damage and neurodegeneration. An example for that is ataxia-telangiectasia (A-T), a genetic disease related to a defect in the DSB-activated protein ATM (Ataxia Telangiestesia Mutated), which is a serine/threonine protein kinase that is recruited and activated by DNA double-strand breaks (43). In A-T patients, the cells of the thymus (where T-cells mature and carry out their genomic rearrangements) and brain cells (exposed to oxidative stress) would promptly die by apoptosis triggered by physiological events instead of repairing the DNA damage, causing the characteristic phenotype of immune deficiency and neurological lesions (100).

Meiotic recombination and sister chromatid exchange
Recombination during meiosis produces genetic diversity in the resulting gametes, thus creating unique allelic combinations to be tested and, if found favourable, propagated further. Recombination in prophase I of meiosis is fully legitimate between regions which are both allelic and homologous between the DNA sequences participating in the recombination. Non-allelic homologous recombination (NAHR), that is, recombination between regions sharing homology but belonging to different loci) occurs relatively frequently too, though the effects are generally mitigated by the bulk of noncoding DNA in the genome. It is currently believed that non-allelic homologous recombination and other recombination events related to replication (such as break-induced replication, see below) are responsible for a significant proportion of inherited genomic disease and cancer (23, 44).

Exchange of genetic material also occurs between sister chromatids during mitosis, though this is usually 'silent' in genetic aspect, as sister chromatids are (presumably) identical and the event is relatively infrequent, amounting to several exchange events per chromosome pair per cell cycle (95). Excessive sister chromatid exchange, caused by molecular defects in certain DNA repair proteins produces chromosomal

instability and a disease phenotype, namely, Bloom syndrome (91). Homologous recombination is believed to be the main mechanism in sister chromatid exchange, providing precise repair of double-strand breaks in actively proliferating cells and suppressing tumorigenesis (63).

Break up to make up – strategies cells use to process double-strand breaks in DNA

Apart from the specific mechanisms managing natural chromosome ends, several mechanisms for dealing with DSB exist. Each of these has its advantages as well as its limitations, contributing to the exquisite balance between genomic stagnation resulting from extremely faithful DNA repair and hypermutability resulting from relaxed control of repair.

Non-homologous end joining (NHEJ)

Basically, broken ends may be rejoined by ligation after making sure that a free 3'–OH and a 5'–phosphate are present, for which some nucleotides at the damage site may be excised and subsequently resynthesized. This does not require the presence of a homologous template, and is therefore known as non-homologous end joining (NHEJ). NHEJ is critically important in V(D)J recombination in immune cells, allowing the B- and T-cells to deploy their huge antigen-recognizing potential by introducing complex genomic rearrangements into the genes coding for the T-cell receptor and the antibody loci, respectively. An important part of the process is adding of non-template nucleotides, increasing the diversity of the resulting molecules (24). Defects in genes coding for products participating in the NHEJ pathway (DNA ligase IV and NHEJ1) in humans usually produces a phenotype of severe immune deficiency with radiosensitivity (12, 69). NHEJ is a generally error-prone repair pathway. On the other hand, it is associated with loss or rewriting of the genetic information from the site surrounding the DSB (52). To this adds the error rate of the polymerases participating in the resynthesis of DNA, which generally lack proofreading properties. However, since the eukaryotic genomes contain a significant amount of noncoding DNA, NHEJ is a viable option for repairing double-strand breaks, as it is highly efficient. Its error-proneness introduces a certain amount of genetic variability, too, though it can cause incorrect joining of ends producing potentially deleterious genomic rearrangements (3).

Homologous recombination (HR)

When there are areas of homology between the sequences containing broken DNA ends, more than one mechanism may be engaged to ensure that repair is carried out without loss of genetic information or that the sequence of the region in the vicinity of the DSB is recovered with minimal losses or alterations. Homologous recombination may be subdivided into several sub-pathways – generally, double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA), and break-induced replication (BIR or break-copy replication) (55). Generally, all three mechanisms require the presence of a homologous template DNA molecule and a single-stranded 3'–OH end to initiate the invasion and pair to the homologous template. DSBR may generate non-crossover or crossover products between the template DNA molecule and the molecule undergoing repair. SDSA generally results in non-crossover products, leaving the template strand in its initial state. BIR is carried out when the double-strand breaks occur in DNA molecules currently undergoing replication. It generates entirely non-reciprocal crossovers, as a replication fork at the site of the free-end invasion may proceed with subsequent DNA synthesis up the end of the replicated region – presumably, to the end of the chromosome (33, 58). All three mechanisms may result in gene conversion, i.e. substituting one part of the DNA sequence for another. When there is multiple allelism in the loci undergoing repair, this may result in loss of heterozygocity. In SDSA, however, the region of the conversion is limited on both ends of the initial double-strand break, while both DSBR and BIR could presumably cause substitution of all DNA sequence after the break site.

When occurring at sites of microhomology (homologous sequences only several bp long), the mechanism of BIR may alter the genome in a quantitative as well as in a qualitative manner (microhomology-mediated BIR, MMBIR), introducing at breakpoint junctions DNA sequences derived from elsewhere in the genome (31, 104). On one hand, this would increase the number of copies in which a sequence is presented in the genome. On the other, it would produce discontinuity in genomic sequences, deleting certain regions, multiplicating others and translocating certain parts of the genome to other locations, placing them in the 'correct' or the 'reverse' orientation to the direction of transcription of the recipient locus.

Fork stalling and template switching (FoSTeS) is a mechanism closely related to MMBIR. Basically, it proposes that during DNA replication the replication fork can stall and switch templates using complementary template microhomology to anneal and prime DNA replication (47, 107). The involved forks can be close to each other or may be separated by considerable distances. The mechanism enables the joining of different sequences located in different positions in the genome and can result in complex non-reciprocal rearrangements, recurring as well as non-recurring ones.

Single-strand annealing

There is yet another mechanism for repair of DSB which require a homologous DNA sequence as a template, but not necessarily on a separate DNA molecule. This mechanism is known as the single-strand annealing (SSA) pathway for repair of double-strand breaks (34, 89). SSA uses repeated sequences within the same DNA molecule as sources of partial homology. SSA is error-prone, as is NHEJ, and bears significant mutagenic potential, not only because the DNA polymerase that performs the resynthesis of DNA is not very accurate, but also because of the fact that the DNA sequence between the repeats is invariably lost, as is one of the two participating repeats.

Breaking may be a new beginning – but of what, exactly?

Apparently, leaving double-strand breaks unrepaired is not an option, as this would result in undesired genetic instability. Adequateness of recombinational repair is, however, a different matter. Homologous recombination is the staple mechanism for generating potential for adaptation. It works to resolve double-strand breaks in a relatively precise manner, utilising the capacity of DNA to preserve its sequence information in duplicate. The greatest strength of homologous recombination is, however, the source of its greatest weakness, as sequences with only partial homology may serve as templates for homologous recombination. Such sequences are quite common in mammalian genomes. It has been proposed that the mechanisms of gene duplication coupled with subsequent divergence and exon shuffling, in other words the basic drives of molecular evolution of genomes, are results of the FoSTeS/ MMBIR mechanisms (107), and, that they operated on a 'trial and error' principle like all evolutionary mechanisms. As a result of multiplication and translocation of sequences, regions of more or less partial homology are densely interspersed within all mammalian chromosomes, within the same chromosome as well as between different chromosomes. For example, A domains of human Factor VIII (located on the X-chromosome, Xq28) are highly homologous to the A domains of Factor V (chromosome 1q) and ceruloplasmin (chromosome 3q) (38). Similarly, the C domains of Factor VIII are homologous to the C domains of Factor V; the human clotting proteins IX (Xq27), VII (chromosome 13q34), X (also 13q34), and protein C (chromosome 2q) all possess an N-terminal gamma-carboxyglutamic acid-containing domain followed by two EGF-like domains, etc. (97). In the human genome, copy number is considered a significant part of the overall genome variation, with the regions with variable copy number estimated to be up to 12% of the overall genome content (78). Apparently, variation in copy number is tolerated to an extent and may even have some hitherto unknown advantage to the carrier, but disorders related to abnormal gene dosage due to multiplication and rearrangement of genomic fragments are quite common.

Non-allelic homologous recombination (NAHR) occurs between DNA sequences which exist in more than one copy and exhibit high sequence homology, but are not polymorphic forms of the same allele. Especially noteworthy are the so-called low-copy repeats (LCR), which are spaced relatively far apart from each other (in the order of several hundreds of kilobases), in either orientation to one another (36, 96). Recombination between regions of homology may cause rearrangement of sizeable genome fragments – deletions of the intervening sequence, when the recombination occurs between straight repeats and inversions, when the breakpoint junctions are within inverted repeats. Unreciprocal exchanges of genome fragments between sequences sharing partial homology may result in duplication. Low-copy repeats may be of variable length, from several kb up to several hundred kb and usually share over 95% identity.

Homologous recombination is stimulated by several orders of magnitude after introduction of a DSB in a repeated sequence (63). Presumably, this ensures that the break is repaired in the safest possible way, using closely related sequences. When the genome contains more than one closely related sequence, however, errors are unavoidable.

Better leave it broken than get hurt trying to repair it – how repair machinery may actually go wrong when doing everything right

Balanced genomic rearrangements, even large ones, may be asymptomatic to their carriers and may only be discovered when the consequences of an unbalanced rearrangement manifest in the progeny. Many genes actually can perform very well with a single genomic copy only, too, provided that this copy is not subject to epigenetic activation (e.g. genomic imprinting) – though somatic inactivation of the surviving copy may trigger neoplastic growth. Autosomal genes generally can rely on homologous repair of DSB using the copy on the other chromosome as a template, though the resultant gene conversion may have undesired effects. The situation is radically different for the X chromosome, which can pair up and engage in homologous recombination in prophase I with the partnering X chromosome in females, but would have to negotiate a complicated genomic exchange with the Y chromosome in males.

The X chromosome is quite large, approximately 155 Mb in length (81) and containing between 1000 and 2000 genes and about 1000 pseudo genes. Less than 30 genes on the X chromosome are shared with the Y chromosome (59). During meiosis, the X and the Y chromosomes only pair within their pseudoautosomal regions (PAR), located in the terminal regions of the arms of these chromosomes. PARs are relatively short and may be not even be present, as deletion of parts or the whole of these regions is not very rare, the genomic fragments containing the missing genes sometimes being added to the X and Y chromosomes (29, 59). Interspersed repeats account for more than half of the X-chromosomal euchromatin, with LINE1 repeat elements covering about 30% of all X chromosome sequence. Intrachromosomal segmental duplications account for about 2.5% of the X chromosome, which is over than 10 times more than the rate of segmental duplication seen in any of the autosome pairs.

Apparently, recombination has been (and currently is) at work at the X chromosome, occasionally producing disease phenotypes. Since a significant number of genes are specific to the X-chromosome only, it is not surprising that these phenotypes become manifest principally in males, with females being asymptomatic carriers. What is highly significant, however, is the fact that they actually occur predominantly during male meiosis, presumably in an X chromosome that cannot pair full-length with the Y and therefore becomes prone to non-allelic homologous recombination involving its various repeated elements. Several allegedly monogeneous diseases and even chromosome diseases such as some variants

of Turner syndrome have been found to be, in fact, genome rearrangement disorders, related to genomic reorganization on the X chromosome. These rearrangements affect not only the gene which produces the general phenotype but also the adjoining genomic sequences, the latter being of considerable length in some cases. Replication-based mechanisms such as break-induced replication are likely to account for most genomic rearrangements of the duplication/multiplication type (22). Subsequent non-allelic homologous recombination between the generated regions of homology results, in turn, in translocations, deletions and inversions, the direct result being a high male/female ratio of germline mutations in the X chromosome (96, 102). The idea that diseases traditionally associated in the public mind with female transmission actually may result from a mutation originating in males apparently still presents quite a challenge for genetic counseling for X-linked disease of the affected families (6, 32, 73, 77). The specificities of the genomic rearrangements resulting in some of these common conditions are presented below.

Rearrangements in the Factor VIII and Factor IX genes

Hemophilia A is a severe disorder of blood clotting, resulting from mutations in the Factor VIII gene (Xq28). The latter contains 16 exons and spans 186 kb, taking about 0.1% of the length of the X-chromosome. The mutation spectrum of haemophilia A is fairly heterogeneous, except from a gross molecular rearrangement which affects 40-50% of the severe cases and approximately 20-25% of all cases of haemophilia A. This molecular defect is essentially an inversion-translocation of the genome region containing exons 1-22 and the adjoining noncoding sequence, which places the inverted region at a 300-400 kb distance from the distal part of the gene (45, 66). The rearrangement has been traced to recombination between repeated regions approximately 10 kb in length sharing a very high degree of homology (over 99.9%) which are represented in one intragenic and two or (rarely) more extragenic copies located on the same chromosome at a distance of about 300 and 400 kb, respectively, from the Factor VIII locus (51). The relatively large scale of the mutation caused considerable difficulties to genetic analysis (35, 53, 74), which actually resulted in obtaining very detailed data on the nature of the mutation and the underlying mechanism. The prevalence of the inversion is very similar in all studied populations, as it results from a random event (non-allelic homologous recombination within the X-chromosome is believed to be a stochastic event). There is slight difference between the prevalence of rare inversion types (which involve the intragenic and the third, the fourth, etc. extragenic copy) in different populations, which is related to how common the individuals carrying more than two extragenic copies in the studied population are. Virtually all mothers of boys with inversion-type haemophilia A are found to be carriers for the inversion, which led to the conclusion that the causative mutation originated predominantly in male germ cells (83), therefore causing healthy males to have carrier daughters which, in turn, give birth to affected boys. With support from linkage analysis of the X chromosome, it was proposed that mispairing resulting in incorrect recombination occurs during male meiosis, when the single X chromosome is only partially paired with the Y, creating ample opportunity for close contact of homologous regions elsewhere on the chromosome (10, 83). The phenomenon of 'mutation type-dependent sex ratio of mutation frequencies', as the authors called it, does not extend only to gross gene rearrangements. Another relatively common (5% of severe cases) defect causing haemophilia A, inversion of intron 1, has been proven to originate exclusively in the male germline too (1). Partial or complete factor VIII gene deletions are found in about 5% of the families with severe hemophilia A, some of which include recombinations of repeated sequences of the LINE1 and Alu type (82, 103). Family analysis shows that, similarly to the inversion with breakpoints in the 10 kb repeat in intron 22, the index patient carrying the rearrangement in Factor VIII gene is usually the first case in the family, born to a carrier mother (50), and that most (though not all) rearrangements occur in the male germline (10, 64).

Homologous recombination between repeats producing deletions or inversions in the Factor IX gene (a close neighbour to Factor VIII on the long arm of the X chromosome) have been found to be the most common causative mutation in Haemophilia B as well, and the male/female ratio of occurrence of mutations has been found to be close to 4 (41). In 2007, studying a patient with a contiguous gene syndrome including haemophilia A, two adjacent deletion/insertion events were found that were reportedly explained best by break-induced replication mechanism (93).

Rearrangements in the dystrophin gene

Duchenne/Becker muscular dystrophy is another common X-linked disease in which the affected males are often the first and only case in the family, born to asymptomatic carrier females and exhibiting distinctive male-to-female germline ratio of occurrence of some type of mutations (26). The gene is located on Xp21 and is one of the largest genes in the human genome, containing 79 exons and spanning about 2 300 kb of genomic DNA (1.5% of all the X-chromosome) (101). Genomic rearrangements, mainly deletions, but also duplications and inversions with flanking deleted sequences are responsible for over 60% of the cases of DMD/BMD (71). It has been found that the sites of rearrangements often contained short regions of homology with other sites on the X chromosome, which presumably self-paired during male meiosis (28, 62, 94). In 2009, a recombination-related mechanism related to presence of repeated elements was proposed to explain the apparent genomic instability of the regions between exons 8 and 13 and between exons 45 and 52 which constituted the two major mutational hotspots in the dystrophin gene (71). Recently, in a larger study, comprised of DMD cases as well as other X-linked disorders, it was found that regions of microhomology (2-10 bp) were observed at breakpoint junctions in about 60% of the study cases which, according to the authors, were remnants of template slippage during break-induced replication (4).

Turner syndrome with other than 45(XO) karyotype

Turner syndrome results from complete or partial monosomy for the X chromosome. The typical 45, XO karyotype is seen in about half of the patients. The most frequent (almost 20%) variant karyotype, however, presents with an isodicentric X chromosome with two q-arms, so that the cell is partially monosomic for the p arms (37). It has been found that the breakpoint regions in many isodicentric X chromosomes included large inverted repeats composed of clusters of repeated sequences, and that most isodicentric X chromosomes result from non-allelic homologous recombination between palindromic sequences (44, 92). The role of regions of microhomology and FoSTeS in generation of dicentric X chromosomes has also been discussed (44).

Other rearrangements with clinical significance in the X chromosome

Pelizaeus-Merzbacher disease (PMD) is an X-linked demyelinating disorder caused by mutation in the gene encoding proteolipid protein-1 (*PLP1*) in humans (Xq22). The major cause for PMD are duplications (about 60% of all *PLP1* mutations), which occur predominantly in the male germline (61) and with breakpoints in low-copy repeats (48, 49, 79, 107).

Similar is the situation with the *MECP2* gene (Xq28), coding for a chromatin-associated protein which binds methylated CpGs, activating or repressing transcription of target sequences. Defects in the *MECP2* gene may produce the phenotype of Rett disorder in humans (2). Duplications including the *MECP2* gene are commonly seen in genomic rearrangements identified in neurodevelopmentally delayed males (14). These rearrangements are supposedly generated by the FoSTeS mechanism (9, 14). Both in PLP1 (Xq22) and in MECP2 (Xq28) loci a complicated rearrangement has been found, namely, a duplication-inverted triplication-duplication mediated by inverted repeats (13).

Mucopolysaccharidosis II is often a product of massive (20-30 kb) inversions or deletions of the IDS locus on Xq, involving the functional gene copy and the respective pseudogene (27). This is believed to be a NAHR-mediated event, occurring predominantly during male meiosis (27, 90).

Fixing it the wrong way – aberrant repair of free DNA ends and double-strand breaks in cancer cells

Apparently, the diversity of mechanisms Nature has provided to make sure that all free DNA ends are joined to one another or protected may not always work in the right direction. This does not extend to meiosis only, but may involve somatic cells as well, especially cancer cells. For example, alternative telomere lengthening (ATL) is a mechanism for telomere elongation other than the conventional TERT-catalyzed synthesis of telomeric repeats using the RNA template provided by TERC. Alternative elongation of telomeric DNA is carried out using as a template telomeric DNA from other telomeres (15, 34). Utilization of ATL has been demonstrated so far in cancer cells only, but makes up for an impressive 10-15% of all cancers (25).

Cancer cells may or may not express high levels of TERT, but they often possess fusion-prone free chromosome ends resulting from replication stress. These reactive free ends may become joined by the NHEJ mechanism, forming a dicentric chromosome which will promptly break at the next anaphase in cellular division, when the two centromeres are pulled apart, producing more free ends, thereby closing the vicious cycle (56, 65).

The prevalence of null genotypes for glutathione S-transferases such as GSTT1 and GSTM1 is quite common in Caucasian and Asian populations, amounting to an average frequency above 50% (11). The GSTM1 and GSTM2 genes are likely to have been produced by gene duplication, as they have very high level of sequence identity, with the GSTM1 null allele probably resulting from non-allelic homologous recombination (106). Carriership of null alleles of these genes is associated with increased level of DNA damage and elevated baseline level of sister chromatid exchanges increasing the probability of somatic recombination, which has been linked to an increase of the risk of various cancers (67, 105).

Recently, it has been proposed that in some types of tumours, especially bone cancer (76, 98), the genomic instability occurs not in the random fashion, which could be expected according to the gradual damage accumulation model of carcinogenesis, but in a single catastrophic event, termed 'chromothripsis', affecting localized regions on one or several chromosomes. Briefly, according to the authors, the affected region is literally shattered into fragments, some (but not all) of which are subsequently brought together again by the cellular machinery for DNA. Since it is not possible to determine the 'right' order of the genomic fragments, these are urgently patched together using the error-prone mechanism of NHEJ into a more or less random order, with all possible consequences of such a 'cut-and-paste' activity (22). Under such circumstances, it is likely that at least one, and, possibly, more molecular events typical of cancerous transformation would occur, therefore, cells who have undergone chromotripsis are much more likely to have acquired exceptional proliferation capacity and to have become capable of ignoring the apoptotic signals as a result.

Re(tying) the knot – how the deleterious results of double-strand breaks could be repaired by introducing yet more double-strand breaks

Modern research and medicine already possess various tools to assess capacity for DNA repair and, sometimes, to intervene where this is possible (8, 16, 17, 18, 20, 54, 86, 99). *In vitro* assays for the rate of repair of double-strand breaks have been developed in order to be used in testing for cancer susceptibility and radiosensitivity (8, 85, 88). Recently, it has been proposed that molecular events causing disruptions of the reading frame of a gene could be corrected by introducing double-strand breaks in the desired DNA sequences using designated

nucleases, then employing the error-prone mechanism of NHEJ to join the free ends. Mutations disrupting the reading frame of a gene, even by deleting or inserting only a single nucleotide, often have more deleterious effects than mutations eliminating whole exons of the same gene, provided that in the second case the remaining parts of the sequence splice in-frame. Thus, by using NHEJ on pre-cut sequences containing mutations disrupting the reading frame, it could be expected that micro-insertions or micro-deletions would be generated at the joining site, which may effectively restore the reading frame though at the expense of several improperly added nucleotides or, alternatively, risking a loss of one or more amino acid residues. As of now this has been demonstrated *in vitro* with out-of-frame mutated dystrophin gene (21, 84). It is also believed that specific nuclease activities coupled with subsequent NHEJ could be used for targeting sequences containing a nonsense mutation and deleting them, or introducing more than one DSB in sites flanking an inserted or translocated sequence, causing its excision (21, 49).

Conclusions

The role of DNA repair mechanisms in maintaining the balance between genome stability and genome variability has been intensively studied for the last decades. It has only recently been proposed that the mechanisms which protect the genomes from stagnation or, alternatively, from hypermutability, may not always perform adequately and that some allegedly monogeneous diseases are, in fact, manifestations of the ongoing process of rearrangement of genetic architecture in the eukaryotic genomes, which is inherent to evolution. Continued research in the field of genome dynamics and rearrangement-based disorders is likely to bring new opportunities for combating human disease, inherited as well as conditions related to somatic mutagenesis.

Acknowledgements

REFERENCES

1. **Acquila M., Pasino M., Santoro C., Lanza T., Molinari A.C., Bottini F., Bicocchi M.P.** (2003) Haemophilia, **9**(6), 717-720.

2. **Amir R.E., Van den Veyver I.B., Wan M., Tran C.Q., Francke U., Zoghbi H.Y.** (1999) Nature Genet., **23**, 185-188.

3. **Andrade P., Martín M.J., Juárez R., López de Saro F., Blanco L.** (2009) Proc. Natl. Acad. Sci. U.S.A., **106**(38), 16203-16208.

4. **Ankala A., Kohn J.N., Hegde A., Meka A., Ephrem C.L., Askree S.H., Bhide S., Hegde M.R.** (2012) Genome Res., **22**(1), 25-34.

5. **Arkus N.** (2005) J. Theor. Biol., **235**(1), 13-32.

6. **Atanassov V., Vazharova R.** (2011) Biotechnol. Biotech. Eq., **25**(4), 2566-2571.

7. **Bardeesy N., DePinho R.A.** (2002) Nat. Rev., **2**, 897-909.

8. **Baria K., Warren C., Roberts S.A., West C.M., Scott D.** (2001) Br. J. Cancer., **84**(7), 892-896.

9. **Bauters M., Van Esch H., Friez M.J., Boespflug-Tanguy O., Zenker M., Vianna-Morgante A.M., Rosenberg C., Ignatius J., Raynaud M., Hollanders K., Govaerts K., Vandenreijt K., Niel F., Blanc P., Stevenson R.E., Fryns J.P., Marynen P., Schwartz C.E., Froyen G.** (2008) Genome Res., **18**(6), 847-858.

10. **Becker J., Schwaab R., Moller-Taube A.** (1996) Am. J. Hum. Genet., **58**, 657-670.

11. **Board P., Coggan M., Johnston P., Ross V., Suzuki T., Webb G.** (1990) Pharm. Ther., **48**, 357-336.

12. **Buck D., Malivert L., de Chasseval R., Barraud A., Fondaneche M.-C., Sanal O., Plebani A., Stephan J.-L., Hufnagel M., le Deist F., Fischer A., Durandy A., de Villartay J.-P., Revy P.** (2006) Cell, **124**, 287-299.

13. **Carvalho C.M., Ramocki M.B., Pehlivan D., Franco L.M., Gonzaga-Jauregui C., Fang P., McCall A., Pivnick E.K., Hines-Dowell S., Seaver L.H., Friehling L., Lee S., Smith R., Del Gaudio D., Withers M., Liu P., Cheung S.W., Belmont J.W., Zoghbi H.Y., Hastings P.J., Lupski J.R.** (2011) Nat. Genet., **43**(11), 1074-1081.

14. **Carvalho C.M., Zhang F., Liu P., Patel A., Sahoo T., Bacino C.A., Shaw C., Peacock S., Pursley A., Tavyev Y.J., Ramocki M.B., Nawara M., Obersztyn E., Vianna-Morgante A.M., Stankiewicz P., Zoghbi H.Y., Cheung S.W., Lupski J.R.** (2009) Hum. Mol. Genet., **18**(12), 2188-2203.

15. **Cesare A.J., Reddel R.R.** (2010) Nat. Rev. Genet., **11**(5), 319-330.

16. **Cafardi J.A., Elmets S.A**. (2008) Expert Opin. Biol. Ther., **8**(6), 829-838.

17. **Chakarov S.A., Petkova R.D., Russev G.Ch.** (2008) Biotechnol. Biotech. Eq., **22**(4), 1011-1012.

18. **Chakarov S., Roeva I., Russev G.** (2011) Biotechnol. Biotech. Eq., **25**(3), 2505-2507.

19. **Chakarov S., Russev G.** (2010) Biotechnol. Biotech. Eq., **24**(2), 1804-1806.

20. **Chakarov St., Stoilov P., Alexandrov A., Russev G.** (1997) Eur. J. Biochem., **248**, 669-675.

21. **Chapdelaine P., Pichavant C., Rousseau J., Pâques F., Tremblay J.P.** (2010) Gene Ther., **17**(7), 846-58.

22. **Chen J.M., Ferec C., Cooper D.N.** (2012) Mutat. Res., **750**(1), 52-59.

23. **Colnaghi R., Carpenter G., Volker M., O'Driscoll M.** (2011) Semin. Cell Dev. Biol., **22**(8), 875-885.

24. **Domínguez O., Ruiz J.F., Laín de Lera T., García-Díaz M., González M.A., Kirchhoff T., Martínez -A. C., Bernad A., Blanco L.** (2000) EMBO J., **19**(7), 1731-1742.

25. **Durant S.T**. (2012) J. Cancer, **3**, 67-82.

26. **van Essen A.J., Abbs S., Baiget M., Bakker E., Boileau C., van Broeckhoven C., Bushby K., Clarke A., Claustres M., Covone A.E., et al.** (1992) Hum. Genet., **88**(3), 249-257.

27. **Froissart R., Maire I., Millat G., Cudry S., Birot A.-M., Bonnet V., Bouton O., Bozon D.** (1998) Clin. Genet. **53**, 362-368.

28. **Giacalone J.P., Francke U.** (1992) Am. J. Hum. Genet., **50**(4), 725-741.

29. **Graves J.A., Wakefield M.J., Toder R.** (1998) Hum. Mol. Genet., **7**(13), 1991-1996.

30. **Harley C.B., Vaziri H., Counter C.M., Allsopp R.C.** (1992) Exp. Gerontol., **27**, 375-382.

31. **Hastings P.J., Ira G., Lupski J.R.** (2009) PLoS Genet., **5**(1), e1000327.

32. **Helderman-van den Enden A.T., van den Bergen J.C., Breuning M.H., Verschuuren J.J., Tibben A., Bakker E., Ginjaar H.B.** (2011) Clin. Genet., **79**(3), 236-242.

33. **Helleday T., Lo J., Van Gent D.C., Engelward B.P.** (2007) DNA Repair (Amst.), **6**(7), 923-935.

34. **Henson J.D., Neumann A.A., Yeager T.R., Reddel R.R.** (2002) Oncogene, **21**(4), 598-610.

35. **Higuchi M., Kazazian H.H. Jr., Kasch L., Warren T.C., McGinniss M.J., Phillips J.A. 3rd, Kasper C., Janco R., Antonarakis S.E.** (1991) Proc. Natl. Acad. Sci. U.S.A., **88**(16), 7405-7409.

36. **Holkers M., de Vries A.A.F., Gonçalves A.V.F.M.** (2012) Nucleic Acids Res., **40**(5), 1984-1999.

37. **Hook E.B., Warburton D.** (1983) Hum. Genet., **64**, 24-27.

38. **Kane W.H., Davie E.W.** (1986) Proc. Natl. Acad. Sci. U.S.A., **83**(18), 6800-6804.

39. **Karlseder J., Smogorzewska A., de Lange T.** (2002) Science, **295**(5564), 2446-2449.

40. **Kaul Z., Cesare A.J, Huschtscha L.I., Neumann A.A., Reddel R.R.** (2011) EMBO Rep., **13**(1), 52-59.

41. **Ketterling R.P., Vielhaber E., Bottema C.D., Schaid D.J., Cohen M.P., Sexauer C.L., Sommer S.S.** (1993) Am. J. Hum. Genet., **52**(1), 152-166.

42. **Kha D.T., Wang G., Natrajan N., Harrison L., Vasquez K.M.** (2010) J. Mol. Biol., **398**(4), 471-480.

43. **Khalil S.H., TummalaH., Oluwaseun O.A., Zhelev N.** (2011) Curr. Opin. Biotech., **22S**, S115-S116.

44. **Koumbaris G., Hatzisevastou-Loukidou H., Alexandrou A., Ioannides M., Christodoulou C., Fitzgerald T., Rajan D., Clayton S., Kitsiou-Tzeli S., Vermeesch J.R., Skordis N., Antoniou P., Kurg A., Georgiou I., Carter N.P., Patsalis P.C.** (2011) Hum. Mol. Genet., **20**(10), 1925-1936.

45. **Lakich D., Kazazian H. Jr., Antonarakis S., Gitschier J.** (1993) Nature Genet., **5**, 236-241.

46. **Lansdorp P.M.** (2000) Mech. Ageing Dev., **118**, 23-34.

47. **Lee J.A., Carvalho C.M., Lupski J.R.** (2007) Cell, **131**(7), 1235-1247.

48. **Lee J.A., Inoue K., Cheung S.W., Shaw C.A., Stankiewicz P., Lupski J.R.** (2006) Hum. Mol. Genet., **15**(14), 2250-2265.

49. **Lee H.J., Kim E., Kim J.S.** (2010) Genome Res., **20**(1), 81-89.

50. **Leuer M., Oldenburg J., Lavergne J.M., Ludwig M., Fregin A., Eigel A., Ljung R., Goodeve A., Peake I., Olek K.** (2001) Am. J. Hum. Genet., **69**(1), 75-87.

51. **Levinson B., Kenwrick S. Lakich D., Hammonds G. Jr., Gitschier J.** (1990) Genomics, **7**, 1-11.

52. **Lieber M.R., Ma Y., Pannicke U., Schwarz K.** (2004) DNA Repair (Amst), **3**, 817-826.

53. **Liu Q., Nozari G., Sommer S.S.** (1998) Blood, **92**, 1458-1459.

54. **Liu J., Sneeden J., Heyer W.D.** (2011) Methods Mol. Biol., **745**, 363-383.

55. **Llorente B., Smith C.E., Symington L.S.** (2008) Cell Cycle, **7**(7), 859-864.

56. **Lo A.W., Sabatier L., Fouladi B., Pottier G., Ricoul M., Murnane J.P.** (2002) Neoplasia, **4**(6), 531-538.

57. **Malissen B., Malissen M.** (1996) Curr. Opin. Immunol., **8**(3), 383-393.

58. **Malkova A., Ivanov E.L., Haber J.E.** (1996) Proc. Natl. Acad. Sci. U.S.A., **93**, 7131-7136.

59. **Mangs H.A., Morris B.J.** (2007) Curr. Genomics, **8**(2), 129-136.

60. **Martens U.M., Chavez E.A., Poon S.S., Schmoor C., Lansdorp P.M.** (2000) Exp. Cell Res., **256**(1), 291-299.

61. **Mimault C., Giraud G., Courtois V., Cailloux F., Boire J.Y., Dastugue B., Boespflug-Tanguy O., Clinical European Network on Brain Dysmyelinating Disease** (1999) Am. J. Hum. Genet., **65**, 360-369.

62. **Miyazaki D., Yoshida K., Fukushima K., Nakamura A., Suzuki K., Sato T., Takeda S., Ikeda S**. (2009) J. Hum. Genet., **54**(2), 127-130.

63. **Moynahan M.E., Jasin M.** (2010) Nat. Rev. Mol. Cell Biol., **11**(3), 196-207.

64. **Mühle C., Zenker M., Chuzhanova N., Schneider H.** (2007) Hum. Mutat., **28**(10), 1045.

65. **Murnane J.P.** (2012) Mutat. Res., **730**(1-2), 28-36.

66. **Naylor J.A., Brinke A., Hassock S., Green P.M., Giannelli F.** (1993) Hum. Mol. Genet., **2**, 1773-1778.

67. **Norppa H.** (2001) Int. J. Hyg. Environ. Health, **204**(1), 31-38.

68. **Nouspikel T., Hanawalt P.C.** (2002) DNA Repair (Amst), **1**(1), 59-75.

69. **O'Driscoll M., Cerosaletti K.M., Girard P.-M., Dai Y., Stumm M., Kysela B., Hirsch B., Gennery A., Palmer S. E., Seidel J., Gatti R. A., Varon R., Oettinger M. A., Neitzel H., Jeggo P. A., Concannon P.** (2001) Molec. Cell, **8**, 1175-1185.

70. **Ohnishi T., Mori E., Takahashi A.** (2009) Mutat. Res., **669**(1-2), 8-12.

71. **Oshima J., Magner D.B., Lee J.A., Breman A.M., Schmitt E.S., White L.D., Crowe C.A., Merrill M., Jayakar P.,**

Rajadhyaksha A., Eng C.M., del Gaudio D. (2009) Hum. Genet., **126**(3), 411-423.

72. **Petkova R., Chakarov S., Ganev V.** (2007) Biotechnol. Biotech. Eq., **21**(4), 385-392.

73. **Petkova R., Chakarov S., Ganev V.** (2011) In: Hemophilia (Batorova A., Ed.), InTech Open Access Publisher, Rijeka, Croatia.

74. **Petkova R.D., Chakarov S.A, Kremensky I.M.** (2008) Biotechnol. Biotech. Eq., **22**(4), 1008-1010.

75. **Petkova R , Chicheva Z , Chakarov S.** (2011) Biotechnol. Biotech. Eq., **25**(4), 2576-2582.

76. **Rajapakse I., Scalzo D., Groudine M.** (2011) Nucleus, **2**(4), 249-252.

77. **Ranjan R., Biswas A., Meena A., Akhter M.S., Yadav B.K., Ahmed R.H., Saxena R.** (2008) Clin. Chim. Acta., **389**(1-2), 103-108.

78. **Redon R., Ishikawa S., Fitch K.R., Feuk L., Perry G.H., Andrews T.D., Fiegler H., Shapero M.H., Carson A.R., Chen W., Cho E.K., Dallaire S., Freeman J.L., González J.R., Gratacòs M., Huang J., Kalaitzopoulos D., Komura D., MacDonald J.R., Marshall C.R. et al.** (2006) Nature, **444**(7118), 444-454.

79. **Regis S., Biancheri R., Bertini E., Burlina A., Lualdi S., Bianco M.G., Devescovi R., Rossi A., Uziel G., Filocamo M.** (2008) Clin.Genet., **73**(3), 279-287.

80. **Robey E., Fowlkes B.J.** (1994) Ann. Rev. Immunol., **12**, 675-705.

81. **Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLay K., Muzny D., Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A., Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Wen G., Jones M.C., Hurles M.E., Andrews T.D. et al.** (2005) Nature, **434** (7031), 325-337.

82. **Rossetti L.C., Goodeve A., Larripa I.B., De Brasi C.D.** (2004) Hum. Mutat., **24**(5), 440.

83. **Rossiter J.P., Young M., Kimberland M.L., Hutter P., Ketterling R.P., Gitschier J., Horst J., Morris M.A., Schaid D.J., de Moerloose P. et al.** (1994) Hum. Mol. Genet, **3**, 1035-1039.

84. **Rousseau J., Chapdelaine P., Boisvert S., Almeida L.P., Corbeil J., Montpetit A., Tremblay J.P.** (2011) J. Gene Med., **13**(10), 522-537.

85. **Rübe C.E., Grudzenski S., Kühne M., Dong X., Rief N., Löbrich M., Rübe C.** (2008) Clin. Cancer Res., **14**(20), 6546-6555.

86. **Russev G., Anachkova A.** (2009) Biotechnol. Biotech. Eq., **23**(2), 1162-1169.

87. **Sahin E., Depinho R.A.** (2010) Nature, **464**, 520-528.

88. **Sakata K., Someya M., Matsumoto Y., Hareyama M.** (2007) Radiat. Med., **25**(9), 433-438.

89. **Salles D., Mencalha A.L., Ireno I.C., Wiesmüller L., Abdelhay E.** (2011) Carcinogenesis, **32**(1), 27-34.

90. **Sasaki M., Lange J., Keeney S**. (2010) Nat. Rev. Mol. Cell Biol., **11**(3), 182-195.

91. **Schroeder T.M., German J.B.** (1974) Humangenetik, **25**, 299-306.

92. **Scott S.A., Cohen N., Brandt T., Warburton P.E., Edelmann L.** (2010) Hum. Mol. Genet., **19**(17), 3383-3393.

93. **Sheen C.R., Jewell U.R., Morris C.M., Brennan S.O., Ferec C., George P.M., Smith M.P., Chen J.M.** (2007) Hum. Mutat., **28**(12), 1198-1206.

94. **Sironi M., Pozzoli U., Comi G.P., Riva S., Bordoni A., Bresolin N., Nag D.K.** (2006) The FASEB Journal, **20**, 1910-1912.

95. **Sjögren C., Ström L.** (2010) Exp. Cell Res., **316**(9), 1445-1453.

96. **Stankiewicz P., Lupski J.R.** (2002) Trends Genet., **18**(2), 74-82.

97. **Stenflo J.** (1999) Crit. Rev. Eukaryot. Gene Expr., **9**(1), 59-88.

98. **Stephens P.J., Greenman C.D., Fu B., Yang F., Bignell G.R., Mudie L.J., Pleasance E.D, Lau K.W., Beare D., Stebbings L.A., McLaren S., Lin M.L., McBride D.J., Varela I., Nik-Zainal S., Leroy C., Jia M., Menzies A., Butler A.P., Teague J.W., et al.** (2011) Cell, **144**, 27-40.

99. **Stoykova E., Sabotinov O., Ion R.-M., Alexandrova R.** (2008) Biotechnol. Biotech. Eq., **22**(2), 754-758.

100. **Swift M.** (1990) Immunodeficiency Reviews, **2**(1), 67-81.

101. **Tennyson C.N., Klamut H.J., Worton R.G.** (1995) Nature Genet., **9**, 184-190.

102. **Thomas G.H.** (1996) Am. J.Hum. Genet., **58**(6), 1364-1368.

103. **Van de Water N., Williams R., Ockelford P., Browett P.A.** (1998) Thromb. Haemost., **79**(5), 938-942.

104. **Vissers L.E., Bhatt S.S., Janssen I.M., Xia Z., Lalani S.R., Pfundt R., Derwinska K., de Vries B., Gilissen C., Hoischen A., Nesteruk M., Wisniowiecka-Kowalnik B., Smyk M., Brunner H.G., Cheung S.W., van Kessel A.G., Veltman J.A., Stankiewicz P.** (2009) Hum. Mol. Genet., **18**(19), 3579-3593.

105. **Vlaykova T., Gulubova M., Vlaykova D., Cirovski G., Yovchev Y., Dimov D., Chilingirov P.** (2009) Biotechnol. Biotech. Eq., **23**(1), 1084-1089.

106. **Xu S., Wang Y., Roe B., Pearson W.R.** (1998) J. Biol. Chem., **273**, 3517-3527.

107. **Zhang F., Khajavi M., Connolly A.M., Towne C.F., Batish S.D., Lupski J.R.** (2009) Nat. Genet., **41**(7), 849-853.