

# Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits

JAMES R. LUPSKI (jlupski@bcm.tmc.edu)

Structural characteristics of the human genome predispose to rearrangements that result in human disease traits. The gene is still the key in mediating the phenotype, but is altered almost coincidentally by the rearrangement. I will refer to such conditions that result from genome architecture as genomic disorders. Genomic disorders are caused by an alteration of the genome that might lead to the complete loss or gain of a gene(s) sensitive to a dosage effect or, alternatively, might disrupt the structural integrity of a gene. This mechanism is in sharp contrast to the classical mechanism for genetic disease, by which an abnormal phenotype is primarily a result of point mutation. Genome alterations can occur through many mechanisms, one of which is homologous recombination during meiosis between region-specific, low-copy repeated sequences. These homologous recombination events result in a type of DNA rearrangement that is a function of the orientation of the repeated sequences that act as substrates for homologous recombination<sup>1</sup>. Recombination between direct repeats can lead to deletion and/or duplication of the genetic material between the repeats, while recombination between inverted repeats results in inversion of the intervening genomic sequence (Fig. 1).

## Types of genomic disorders

Genomic disorders are defined in this review as those that result from DNA rearrangements owing to homologous recombination involving region-specific, low-copy repeats. The repeats might represent: (1) gene deficient segments; (2) fragments of genes; (3) pseudogenes; (4) gene copies; (5) gene family members; or (6) repeat gene clusters. The type of the resulting genomic disorder is a function of the orientation of the repeat and the position of the dosage-sensitive gene(s), or specific exons of a gene, relative to the repeat. Three major types of genomic disorders can be delineated on the basis of genome architecture.

### *Tandemly repeated genes*

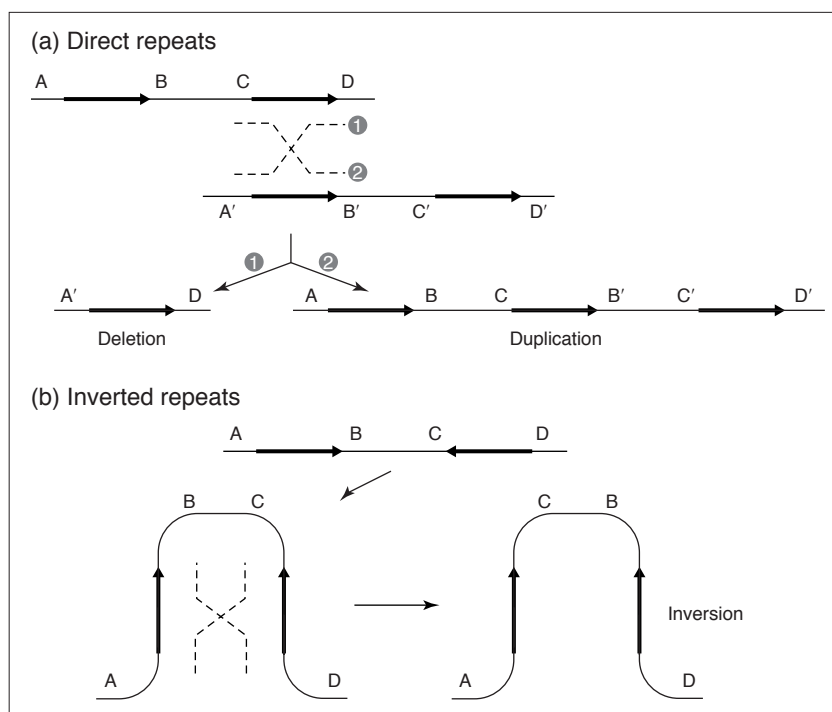
In this group of disorders the genes can be arranged in tandem and act as homologous recombination substrates, or the genes can have adjacent sequences that are repeated and are the preferred recombination substrates (Fig. 2a). Recombination between repeats leads to the loss of one gene copy, resulting in haploinsufficiency, or it can result in a recombinant hybrid gene with different properties. One of the first disorders recognized to be the result of a unique genomic arrangement was  $\alpha$ -thalassemia, which is caused by  $\alpha$ -globin gene deletions<sup>2-4</sup>. These deletions are the outcome of unequal crossing-over events between repeated segments (Z and X) of approximately 4 kb within the  $\alpha$ -globin locus<sup>4</sup>. The duplicated Z boxes are 3.7 kb apart and the X boxes are 4.2 kb apart. Misalignment and reciprocal crossover between the Z boxes or the X boxes at meiosis yields chromosomes with either one or three  $\alpha$ -globin genes. Individuals that inherit chromosomes with only one copy manifest the phenotype. Quantification of the amount of  $\alpha$ -globin mRNA correlates directly with the number of  $\alpha$ -globin genes present – a gene-dosage effect<sup>2</sup>.

*Molecular medicine began with Pauling's seminal work, which recognized sickle-cell anemia as a molecular disease, and with Ingram's demonstration of a specific chemical difference between the hemoglobins of normal and sickled human red blood cells. During the four decades that followed, investigations have focused on the gene – how mutations specifically alter DNA and how these changes affect the structure and expression of encoded proteins. Recently, however, the advances of the human genome project and the completion of total genome sequences for yeast and many bacterial species, have enabled investigators to view genetic information in the context of the entire genome. As a result, we recognize that the mechanisms for some genetic diseases are best understood at a genomic level. The evolution of the mammalian genome has resulted in the duplication of genes, gene segments and repeat gene clusters. This genome architecture provides substrates for homologous recombination between nonsyntenic regions of chromosomes. Such events can result in DNA rearrangements that cause disease.*

$\beta$ -Thalassemia and globin fusion genes can also be the result of unequal crossing over and homologous recombination between repeated genes in the  $\beta$ -globin locus<sup>4</sup>. The homology requirements for unequal crossing over in the  $\beta$ -gene cluster were examined by nucleotide sequence of the recombinant chromosomes. Each strand-exchange event occurred in an extensive region of uninterrupted identity between the parental genes. Crossovers between misaligned homologous genes occur consistently within regions with the largest available stretches of identity for a particular pair of mismatched genes. These observations support the hypothesis that sequence identity is a critical factor in efficient homologous recombination<sup>5,6</sup>.

Familial isolated growth-hormone deficiency is characterized by the complete absence of growth hormone owing to homozygous deletion of the gene encoding growth hormone (*GHI*). The deletion of *GHI* results from recombination between repeated segments at the *GHI* locus within the growth hormone gene cluster<sup>7,8</sup>. In one study, nine of ten patients had crossovers within the 594 bp segments that flank *GHI*; these segments are 99% identical and contain the longest perfect repeats, or stretches of sequence identity, found in the 2.24 kb segments of homology that flank *GHI* in the growth hormone gene cluster<sup>8</sup>.

## REVIEWS



**FIGURE 1.** Genomic rearrangements resulting from recombination between repeated sequences. Repeated sequences are depicted as black arrows with the orientation indicated by the direction of the arrowhead. Capital letters above or below the thin horizontal lines refer to the flanking unique sequence (e.g. A). Chromosome homologs are also shown (e.g. A'). Dashed lines refer to a recombination event with the results shown by numbers 1 and 2. (a) Recombination between direct repeats results in deletion and/or duplication. (b) Recombination between inverted repeats results in an inversion.

The cytochrome P450 enzyme debrisoquine 4-hydroxylase (*CYP2D6*) metabolizes many different classes of commonly used drugs and is, therefore, responsible for a common pharmacogenetic trait. The trait is expressed by the ability of an individual within a population to metabolize certain drugs in a rapid or poor manner. Therapeutic efficacy, as well as common side effects, of medications metabolized by the *CYP2D6* gene product can depend on an individual's genotype at this locus. *CYP2D6*, a member of the *CYPD* cluster of three tandemly arranged genes, is flanked by a 2.8 kb repeat. The breakpoints of the common poor-metabolizer-associated deletion allele occur within the repeat<sup>9</sup>.

Red-green color blindness is an extremely common trait with the incidence of red and green color vision variation being found in approximately 8% of Caucasian males. The red-green pigment gene complex maps to Xq28 and consists of a tandem array of one gene encoding red opsin and one or more genes encoding green opsin. The high degree of sequence similarity between these genes makes them prone to unequal crossing over, and the resulting deletions or duplications account for numerical polymorphisms at this locus<sup>10</sup>. Loss of pigment genes, or the creation of a hybrid gene from the fusion of a red and green opsin gene on a recombinant chromosome, lead to red-green color vision defects, which were among the first recognized X-linked traits in humans<sup>10,11</sup>.

Genes encoding aldosterone synthase and steroid 11 $\beta$ -hydroxylase are 95% identical and lie 45 kb apart on chromosome 8q. Glucocorticoid-remediable aldosteronism (GRA) is an autosomal dominant disorder that

is characterized by hypertension with variable hyperaldosteronism and by high levels of abnormal adrenal steroids, which are under the control of adrenocorticotrophic hormone and suppressible by glucocorticoids. GRA is caused by gene duplication arising from unequal crossing over that fuses the 5' regulatory region of the gene encoding 11 $\beta$ -hydroxylase (*11-OHase*) to the coding sequences of the gene encoding aldosterone synthase (*AldoS*). This results in inappropriate expression of aldosterone synthase in tissues or at times when it should not be expressed<sup>12</sup>.

Other genomic disorders resulting from unequal crossing over of physically linked repeated genes, or a gene and a related pseudogene, affect a significant number of patients with: (1) 21-hydroxylase deficiency owing to recombination between *CYP21* genes<sup>13</sup>; (2) Bartter syndrome type III, subsequent to recombination between the related genes encoding chloride channels, *CLCNKB* and *CLCNKA* (Ref. 14); and (3) Gaucher disease owing to recombination between the gene for acid  $\beta$ -glucosidase and a nearby pseudogene<sup>15</sup>.

### *Tandem repeats separated from genes*

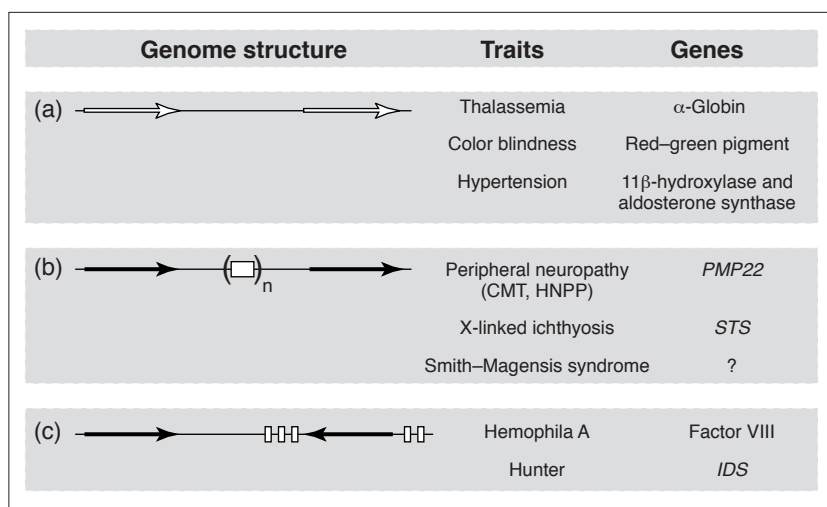
Genomic disorders are also associated with repeats that are physically distinct and located some distance away from the gene locus involved in the phenotype (Fig. 2b). The repeated sequence flanks a genomic region that might contain one or many genes, any number of which is dosage sensitive and predominantly responsible for the phenotype. In this case, the gene responsible for the phenotype is affected by virtue of a dosage alteration, but it is not involved as a substrate for the recombination resulting in the DNA rearrangement. Recombination between repeats leads either to deletion of one copy of the dosage-sensitive gene or to a duplication of that gene.

One example of a special case of this type of genomic disorder is steroid-sulfatase (STS) deficiency caused by deletions in Xp22.3. This is considered a special case because, in the strict sense, a gene dosage effect does not apply to phenotypes that manifest in a hemizygous state. STS deficiency is associated with the skin disorder X-linked ichthyosis. About 90% of patients with STS deficiency have their entire *STS* gene deleted<sup>16</sup>. Because the prevalence of STS deficiency is one in 2000–5000 males, about 0.01% of individuals in the population have an X chromosome with the *STS* locus deleted<sup>16</sup>. In the majority of deletion patients the breakpoints lie within a low-copy repeat called S232 (Refs 16, 17). The S232 repeats are located 1.9 Mb apart and consist of 5 kb of unique sequence in addition to two elements composed of a variable number of tandem repeats<sup>18</sup>.

## REVIEWS

Perhaps one of the most extensively characterized genomic disorders caused by recombination between tandem repeats is the peripheral neuropathy Charcot-Marie-Tooth disease type 1A (CMT1A), which is associated with a 1.5 Mb tandem duplication in 17p12 (Refs 19–21). This duplication arises from unequal crossing over and homologous recombination between 24 kb flanking repeats termed CMT1A-REP (Ref. 22). The reciprocal recombination product involving CMT1A-REP results in a 1.5 Mb deletion that is associated with a clinically distinct peripheral neuropathy – hereditary neuropathy with liability to pressure palsies (HNPP)<sup>23</sup>. CMT1A and HNPP result from an altered copy number of the dosage-sensitive myelin gene *PMP22*, which is located 0.5 Mb from the proximal or centromeric copy of CMT1A-REP and 1.0 Mb from the distal or telomeric CMT1A-REP. The two CMT1A-REP copies share 24 011 bp of 98.7% sequence identity<sup>24</sup>. The CMT1A-REP repeat appears to have been duplicated during primate genome evolution because humans and chimpanzees have two copies, whereas gorillas and other lower primates have only one copy<sup>24,25</sup>. Thus, the evolution of the mammalian genome itself might create structural features that leave particular regions of the human genome susceptible to genomic disorders.

The homologous recombination events between flanking CMT1A-REPs are associated with a hotspot for crossovers, where >70% of all events occur<sup>25–29</sup>. Nucleotide sequence analysis of the strand-exchange region in HNPP deletion<sup>30</sup> and in CMT1A duplication<sup>31</sup> individuals reveals that the crossover occurs in long stretches (>400 bp) of identity within the highly homologous CMT1A-REPs. The length of these stretches of identity can be equal to or greater than minimal efficient processing segments (MEPS) for homologous recombination in human meiosis<sup>30</sup>. Moreover, analysis of the recombinant CMT1A-REPs revealed evidence of gene-conversion events, which are the hallmarks of the double-strand breaks that frequently initiate homologous recombination<sup>30</sup>. We have proposed that a *mariner*-like element, termed MITE, in conjunction with a *trans* acting transposase, could be responsible for initiating double-strand breaks at the CMT1A locus<sup>29</sup>. The recombination hotspots associated with the CMT1A duplication and HNPP deletion appear to overlap with the longest stretches of sequence identity between CMT1A-REPs that are nearest to the MITE element<sup>29,30</sup>. Other investigators have proposed that different short sequence motifs might be involved in initiating the recombination process<sup>31</sup>. Intriguingly, there seems to be a sex predilection for the *de novo*



**FIGURE 2.** Genome structural features and example genomic disorders. (a) Tandemly repeated genes. The features of the genome are shown (genome structure) with genes indicated as open arrows. Examples of disease traits that might be caused by this genomic architecture are given with the genes affected by the rearrangement. (b) Tandem repeats separated from genes. A dosage-sensitive gene (open horizontal rectangle) or genes ( $n > 1$ ) is flanked by a repeat (black arrows) in tandem orientation. If recombination occurs between mis-aligned flanking repeats (unequal crossing over) the genes located between the repeats can be deleted or duplicated. Examples of diseases that can occur because of this genomic architecture, and be the result of gene haploinsufficiency (deletion) or increased gene dosage (duplication) effects, are given with the particular gene involved [? Refers to dosage-sensitive gene(s) unknown]. (c) Inverted repeats with one copy of the repeat located within a gene. A multi-exon gene with one copy of a repeat (black arrow) located between exons (white rectangles) but an additional copy of the repeat is located either upstream or downstream from the gene. This genomic architecture can result in an inversion that disrupts the structural integrity of the gene.

CMT1A duplication, with the overwhelming majority of duplications resulting from unequal crossing over during male gametogenesis<sup>32,33</sup>. However, a few HNPP deletions have been documented as occurring during female gametogenesis and appear to result from intra-chromosomal exchange events<sup>33</sup>.

Smith–Magenis syndrome (SMS), associated with del(17)(p11.2), is a multiple congenital anomalies, mental-retardation syndrome that appears to be a contiguous-gene syndrome<sup>34–36</sup>. Molecular studies of SMS-deletion patients have revealed an approximately 5 Mb commonly deleted region in the majority of patients<sup>36–38</sup>. This region is flanked by a repeat gene cluster, SMS-REP, which appears to be >200 000 bp in length<sup>39</sup>. Although there is a third copy of SMS-REP located within the common deletion region, the copies flanking this region and located furthest apart appear to be the preferred substrates for homologous recombination<sup>39</sup>. The breakpoints in SMS patients with the common deletion occur within the flanking SMS-REP. Over 90% of patients with the common deletion have a novel junction fragment of the same apparent size, observed using pulsed-field gel electrophoresis (PFGE) with an SMS-REP probe, suggesting a precise recombination event<sup>39</sup>.

Recently, some patients with mild retardation and minor dysmorphic features have been shown to harbor a chromosomal duplication [dup(17)(p11.2)] of the genomic region that is deleted in patients with SMS. Interestingly, these patients with duplications of the SMS region also have a patient-specific novel junction

## REVIEWS

**TABLE 1. Physical features of regions associated with genomic disorders**

Trait	Rearrangement type	Distance between repeats (kb)	Repeat length (bp)
Color blindness	DEL	0	39 000
$\alpha$ -Thalassemia	DEL	3.7 or 4.2	4000
Growth hormone deficiency	DEL	6.7	2200
Debrisoquine sensitivity	DEL	9.3	2800
Hunter mucopolysaccharidosis	INV	20	3000
Glucocorticoid-remediable aldosteronism	DUP	45	10 000
Hemophilia A	INV	500	9500
CMT1A/HNPP	DUP/DEL	1500	24 011
X-linked ichthyosis	DEL	1900	20 000
Williams syndrome	DEL	~2000	>30 000
Smith-Magenis syndrome/dup(17)(p11.2)	DEL/DUP	~5000	>200 000

Abbreviations: DEL, deletion; DUP, duplication; INV, inversion.

fragment when PFGE-separated genomic DNA is hybridized with an SMS-REP probe, suggesting that these duplications represent reciprocal recombination events involving SMS-REP (Ref. 39 and J.R. Lupski *et al.*, unpublished).

Although not fully characterized at the molecular level, observations of other contiguous-gene syndromes suggest that they represent genomic disorders<sup>40-52</sup>. Low-copy repeats have been identified at the loci implicated in Williams syndrome (WS), Prader-Willi and Angelman syndrome, and DiGeorge/velo-cardio-facial syndrome. Furthermore, a common deletion is found in the majority of patients and unequal crossing over appears to be involved in the generation of the deletion. WS is a contiguous-gene syndrome associated with a deletion in 7q11.23. Recently, a genomic duplication has been identified flanking the commonly deleted interval<sup>53</sup>. The duplicated segment includes at least one transcribed gene and is >30 kb in length. An apparent novel deletion junction fragment of >3 Mb was detected with a cDNA probe that hybridizes to the flanking repeat in a hybrid cell line retaining the 7q11.23 deletion chromosome, as well as in four WS patient samples<sup>53</sup>. These results suggest that homologous recombination between the flanking repeat is responsible for generating the WS deletion.

### *Inverted repeats*

Recombination between inverted repeats, when one copy of the repeat is located within a gene, results in an inversion of the intervening genomic segment that disrupts the gene and functionally inactivates it (Fig. 1). A prototypical genomic disorder of this type is hemophilia A, in which 47% of cases in severely affected patients are caused by an inversion of a portion of the gene encoding factor VIII. This rearrangement is caused by repeat sequences located within and 5' to the gene<sup>54</sup> (Fig. 1). The repeat sequence is an intronless gene (gene A) present in intron 22 of the gene encoding factor VIII. Gene A is transcribed in the opposite direction of the gene encoding factor VIII, and two additional copies are present ~500 kb upstream of it. Recombination between the intronic copy<sup>54</sup> of gene A and either of the upstream copies leads to the inversion of exons 1-22 of the gene encoding factor VIII.

Interestingly, the recombination with the more distal copy occurs more frequently than recombination with the more proximal copy. The recombination was shown to occur within the region of homology that was estimated by chemical mismatch analysis to be approximately 99% identical<sup>55</sup>. The inversions in the gene encoding factor VIII causing severe hemophilia A were shown to originate almost exclusively in male germ cells<sup>56</sup>.

Inversion of the gene encoding iduronate-2-sulfatase (*IDS*), resulting from recombination with *IDS*-related sequences, is a common cause of Hunter syndrome<sup>57</sup>. A putative pseudogene (*IDS-2*) is located 20 kb distal to and in the opposite orientation to the *IDS* gene<sup>57,58</sup>. The *IDS-2* locus spans approximately 3 kb and shows greater than 88% identity with the *IDS* gene<sup>58,59</sup>. DNA sequence analysis of the junctions of the inversion showed that all recombination events take place within a 1 kb region where the sequence identity is greater than 98%<sup>59</sup>. The identification of regions with alternating *IDS* and *IDS-2* sequences present at one inversion junction represent a possible outcome of recombination events initiated by a double-strand break in intron 7 of the *IDS* gene<sup>59</sup>.

### **General features regarding mechanisms for genomic disorders**

It is apparent from the above discussion that there are several common features associated with the mechanisms that lead to genomic disorders.

(1) Significant regions of homology appear to be required for recombination. The homologous regions usually extend over thousands of base pairs. Short interspersed repetitive sequences, such as *Alu*, are not usually substrates, although recombination involving *Alu* can occur either by illegitimate or homologous recombination events, and it occasionally leads either to deletion or duplication through unequal crossing over<sup>60-64</sup>.

(2) Observations regarding the physical features of regions of the genome that are associated with genomic disorders reveal wide variations in the repeat length of the duplicated genome segments (Table 1). However, if genomic disorders are arranged according to increasing distance between repeats, the repeat length that is observed correlates positively with the distance between repeats (Table 1). Generally, the larger the distance between repeats, the greater the repeat length

## REVIEWS

potentially required for efficient recombination. This might reflect the fact that, for the unequal crossing-over and recombination to occur over large distances, a greater length of sequence similarity might be required to stabilize the recombination complex. Alternatively, when separated by greater distances, larger stretches of sequence similarity might be more likely to find each other and pair.

(3) The strand exchange or crossover appears to occur preferentially in a region of perfect identity located within a repeated sequence. The requirement for significant stretches of identity suggests that the mismatch-repair and recombination machinery recognizes sequence heterogeneities within similar regions and could break down recombination intermediates without sequence identity. Alternatively, a RecA-like protein might require and seek out identity within similar regions. These identity regions might reflect the MEPS required for homologous recombination<sup>65,66</sup>.

(4) Double-strand breaks might be the initiating event for recombination between repeats leading to genomic disorders. DNA sequence analysis of junctions in HNPP, CMT1A and Hunter syndrome patients have revealed interspersed patches of DNA sequence information from the two recombined repeats, suggesting gene-conversion events. These might result from the repair of heteroduplex DNA during the resolution of Holliday junctions.

(5) In some cases, the *de novo* rearrangements leading to genomic disorders display a parent-of-origin effect. It has been clearly documented, both in the overwhelming majority of CMT1A duplication cases and in the factor VIII inversion cases, that rearrangements occur on the paternally inherited chromosome. These data suggest that homologous recombination requirements for rearrangements resulting in genomic disorders sometimes differ for male and female gametogenesis.

(6) Not all copies of a given repeat are used equivalently as substrates for homologous recombination. As shown for the SMS deletion and the factor VIII gene inversion, the adjacent repeats are not necessarily the ones involved in the most frequent recombination events. Moreover, additional low-copy repeats have been identified in the SMS region on chromosome 17p11.2, but the flanking SMS-REPs appear to be the preferred recombination substrates. These observations suggest that other factors, perhaps higher-order structural features of chromosomes at the synaptonemal complex during meiosis, are required to align the homology substrates. Alternatively, proximity to the recombination initiation site may be important.

### Other genomic disorders

Genomic disorders are responsible for a number of common disease traits. Although exact prevalence estimates are not available for most genomic disorders, the relative frequency of determining that the genome rearrangement is responsible for a given trait can be quite significant. It has been determined that 13% of all cases of Hunter syndrome<sup>57</sup>, 20–25% of 21 hydroxylase deficiency<sup>67</sup>, 47% of severe hemophilia A (Ref. 54), 70% of all CMT1 (Refs 68, 69), 84% of HNPP (Ref. 69), and 90% of STS deficiency<sup>18</sup> are caused by specific DNA rearrangements. The frequency of the new mutation

event for genomic disorders can also be substantial, with an estimated mutation rate of  $10^{-4}$  for CMT1A (Ref. 20), SMS (Ref. 36) and Williams syndrome<sup>53</sup>.

In the cases of genome rearrangements caused by deletions, the reciprocal duplication events might be under-recognized. If other contiguous-gene-deletion syndromes, such as Williams, Prader-Willi and Angelman, and DiGeorge/velo-cardio-facial syndromes, are shown to result from a molecular mechanism similar to that of SMS, then the reciprocal duplication as seen for SMS might occur (Ref. 39). Such duplication patients might have different clinical findings and milder phenotypic features than those with deletions, because excess of genetic information is usually less detrimental to the organism than deficiency. Therefore, these cases could escape identification through under-ascertainment or be missed by routine cytogenetic analysis. Chromosomal duplications are frequent mutational events that have been documented across species<sup>70</sup>.

Other diseases commonly caused by DNA rearrangements might reflect unique genome structural features. These might include: the spinal muscular atrophies associated with repeated sequences in 5q13 (Refs 71–73); juvenile nephronophthisis (recessive medullary cystic kidney disease) associated with large homozygous deletions in 2q13, involving a 100 kb inverted duplication<sup>74</sup>; a duplicated *PLP* gene causing Pelizaeus-Merzbacher disease<sup>75</sup>; a regional duplication in Xq25–26 associated with X-linked recessive panhypopituitarism<sup>76</sup>; inversions around the emerin locus associated with an 11.3 kb inverted repeat in Xq28 (Refs 77, 78); and unique sequence arrangements at the locus in 4q35 associated with fascioscapulohumeral muscular dystrophy<sup>79</sup>. As additional large DNA sequence contigs become available through the efforts of the human genome project, the complexities of the human genome architecture will be revealed further. Subsequently, additional structural features will probably be uncovered and shown to be responsible for a variety of yet uncharacterized genomic disorders.

### Acknowledgements

I appreciate the critical reviews of my colleagues A. Beaudet, B. Bejjani, K-S. Chen, R. Gibbs, L. Potocki, L., Reiter, S. Rosenberg and L. Shaffer. Research in my laboratory has been generously supported by the National Institutes of Neurological Disorders and Stroke and the National Eye Institute, NIH, and the Muscular Dystrophy Association, as well as the Baylor Mental Retardation Research Center (HD24064), Child Health Research Center (HD94021), and the Texas Children's Hospital General Clinical Research Center (M01 RR-00188).

### References

- 1 Weinstock, G.M. and Lupski, J.R. (1998) in *Bacterial Genomes: Physical Structure and Analysis* (de Bruijn, F.J., Lupski, J.R. and Weinstock, G.M., eds), pp. 112–118, Chapman & Hall
- 2 Higgs, D.R. *et al.* (1980) *Nature* 284, 632–635
- 3 Lauer, J., Shen, C.K.J. and Maniatis, T. (1980) *Cell* 20, 119–130
- 4 Weatherall, D.J., Clegg, J.B., Higgs, D.R. and Wood, W.G. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (7th edn) (Scriver, C.R., Beaudet, A.L., Sly, W.S. and Valle, D., eds), pp. 3417–3484, McGraw-Hill

## REVIEWS

- 5 Metzenberg, A.B., Wurzer, G., Huisman, T.H.J. and Smithies, O. (1991) *Genetics* 128, 143–161
- 6 Rayssiguier, C., Thaler, D.S. and Radman, M. (1989) *Nature* 342, 396–401
- 7 Vnencak-Jones, C.L., Phillips, J.A., Chen, E.Y. and Seeburg, P.H. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 5615–5619
- 8 Vnencak-Jones, C.L. and Phillips, J.A. (1990) *Science* 250, 1745–1748
- 9 Steen, V.M., Molven, A., Aarskog, N.K. and Gullbrandsen, A.K. (1995) *Hum. Mol. Genet.* 4, 2251–2257
- 10 Nathans, J. *et al.* (1986) *Science* 232, 203–210
- 11 Motulsky, A.G. and Deeb, S.S. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Vol. II) (Scriver, C.R., Beaudet, A.L., Sly, W.S. and Valle, D., eds), pp. 4275–4295, McGraw-Hill
- 12 Lifton, R.P. *et al.* (1992) *Nature* 355, 262–265
- 13 Donohoue, P.A., Jospé, N., Migeon, C.J. and Van Dop, C. (1989) *Genomics* 5, 397–406
- 14 Simon, D.B. *et al.* (1997) *Nat. Genet.* 17, 171–178
- 15 Zimran, A. *et al.* (1990) *J. Clin. Invest.* 85, 219–222
- 16 Yen, P.H. *et al.* (1990) *Cell* 61, 603–610
- 17 Ballabio, A. *et al.* (1990) *Genomics* 8, 263–270
- 18 Li, X.-M., Yen, P.H. and Shapiro, L. (1992) *Nucleic Acids Res.* 20, 1117–1122
- 19 Lupski, J.R. (1997) *Hosp. Pract.* 32, 83–122
- 20 Lupski, J.R. (1998) *Mol. Med.* 4, 3–11
- 21 Lupski, J.R. (1998) in *Scientific American Molecular Neurology* (Martin, J.B., ed.), pp. 239–256, Scientific American Inc.
- 22 Pentao, L. *et al.* (1992) *Nat. Genet.* 2, 292–300
- 23 Chance, P.F. *et al.* (1994) *Hum. Mol. Genet.* 3, 223–228
- 24 Reiter, L.T. *et al.* (1997) *Hum. Mol. Genet.* 6, 1595–1603
- 25 Kiyosawa, H. and Chance, P. (1996) *Hum. Mol. Genet.* 5, 745–753
- 26 Kiyosawa, H., Lensch, M.W. and Chance, P.F. (1995) *Hum. Mol. Genet.* 4, 2327–2334
- 27 Timmerman, V. *et al.* (1997) *J. Med. Genet.* 34, 43–49
- 28 Lopes, J. *et al.* (1996) *Am. J. Hum. Genet.* 58, 1223–1230
- 29 Reiter, L.T. *et al.* (1996) *Nat. Genet.* 12, 288–297
- 30 Reiter, L.T. *et al.* (1998) *Am. J. Hum. Genet.* 62, 1023–1033
- 31 Lopes, J. *et al.* (1998) *Hum. Mol. Genet.* 7, 141–148
- 32 Palau, F. (1993) *Hum. Mol. Genet.* 2, 2031–2035
- 33 Lopes, J. *et al.* (1997) *Nat. Genet.* 17, 136–137
- 34 Chen, K.-S., Potocki, L. and Lupski, J.R. (1996) *Mental Retard. Dev. Disab. Res. Rev.* 2, 122–129
- 35 Greenberg, F. *et al.* (1996) *Am. J. Med. Genet.* 62, 247–254
- 36 Greenberg, F. *et al.* (1991) *Am. J. Hum. Genet.* 49, 1207–1218
- 37 Guzzetta, V. *et al.* (1992) *Genomics* 13, 551–559
- 38 Juyal, R.C. *et al.* (1996) *Am. J. Hum. Genet.* 58, 998–1007
- 39 Chen, K.-S. *et al.* (1997) *Nat. Genet.* 17, 154–163
- 40 Dutly, F. and Schinzel, A. (1996) *Hum. Mol. Genet.* 5, 1893–1898
- 41 Urban, Z. *et al.* (1996) *Am. J. Hum. Genet.* 59, 958–962
- 42 Perez-Jurado, L.A. *et al.* (1996) *Am. J. Hum. Genet.* 59, 781–792
- 43 Robinson, W.P. *et al.* (1996) *Genomics* 34, 17–23
- 44 Wu, Y. *et al.* (1998) *Am. J. Med. Genet.* 78, 82–89
- 45 Osborne, L.R. *et al.* (1997) *Genomics* 45, 400–404
- 46 Carrozzo, R. *et al.* (1997) *Am. J. Hum. Genet.* 61, 228–231
- 47 Christian, S.L. *et al.* (1995) *Am. J. Hum. Genet.* 57, 40–48
- 48 Christian, S.L. *et al.* (1997) *Am. J. Hum. Genet.* 61, A24
- 49 Morrow, B. *et al.* (1995) *Am. J. Hum. Genet.* 56, 1391–1403
- 50 Lindsay, E.A. *et al.* (1995) *Am. J. Med. Genet.* 56, 191–197
- 51 Halford, S. *et al.* (1993) *Hum. Mol. Genet.* 2, 191–196
- 52 Morrow, B.E. *et al.* (1997) *Am. J. Hum. Genet.* 61, A25
- 53 Perez-Jurado, L.A. *et al.* (1998) *Hum. Mol. Genet.* 7, 325–334
- 54 Lakich, D., Kazazian, J.H.H., Jr, Antonarakis, S.E. and Gitschier, J. (1993) *Nat. Genet.* 5, 236–241
- 55 Naylor, J.A. *et al.* (1995) *Hum. Mol. Genet.* 4, 1217–1224
- 56 Rossiter, J.P. *et al.* (1994) *Hum. Mol. Genet.* 3, 1035–1039
- 57 Bondeson, M.-L. *et al.* (1995) *Hum. Mol. Genet.* 4, 615–621
- 58 Timms, K.M. *et al.* (1997) *Hum. Mol. Genet.* 6, 479–486
- 59 Lagerstedt, K. *et al.* (1997) *Hum. Mol. Genet.* 6, 627–633
- 60 Rudiger, N.S. *et al.* (1991) *Clin. Genet.* 39, 451–462
- 61 Olds, R.J. *et al.* (1993) *Biochemistry* 32, 4216–4224
- 62 Kornreich, R., Bishop, D.F. and Desnick, R.J. (1990) *J. Biol. Chem.* 265, 9319–9326
- 63 Marcus, S. *et al.* (1993) *Hum. Genet.* 90, 477–482
- 64 Pousi, B. *et al.* (1994) *Am. J. Hum. Genet.* 55, 899–906
- 65 Shen, P. and Huang, H.V. (1986) *Genetics* 112, 441–457
- 66 Vulic, M., Dionisio, F., Taddei, F. and Radman, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 9763–9767
- 67 Donohoue, P.A., Parker, K. and Migeon, C.J. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Vol. II) (Scriver, C.R., Beaudet, A.L. and Valle, D., eds), pp. 2929–2966, McGraw-Hill
- 68 Wise, C.A. *et al.* (1993) *Am. J. Hum. Genet.* 53, 853–863
- 69 Nelis, E. *et al.* (1996) *Eur. J. Hum. Genet.* 4, 25–33
- 70 Lupski, J.R., Roth, J.R. and Weinstock, G.M. (1996) *Am. J. Hum. Genet.* 58, 21–27
- 71 Melki, J. *et al.* (1994) *Science* 264, 1474–1477
- 72 Lefebvre, S. *et al.* (1995) *Cell* 80, 155–165
- 73 Wang, C.H. *et al.* (1995) *Am. J. Hum. Genet.* 56, 202–209
- 74 Konrad, M. *et al.* (1996) *Hum. Mol. Genet.* 5, 367–371
- 75 Inoue, K. *et al.* (1996) *Am. J. Hum. Genet.* 59, 32–39
- 76 Lagerström-Fermer, M. *et al.* (1997) *Am. J. Hum. Genet.* 60, 910–916
- 77 Small, K., Iber, J. and Warren, S.T. (1997) *Nat. Genet.* 16, 96–99
- 78 Small, K. and Warren, S.T. (1998) *Hum. Mol. Genet.* 7, 135–139
- 79 Wijmenga, C. *et al.* (1992) *Nat. Genet.* 2, 26–30

**J.R. Lupski** is in the Department of Molecular and Human Genetics and the Department of Pediatrics, and Texas Children's Hospital, Baylor College of Medicine, One Baylor Plaza, Room 609E, Houston, TX 77030, USA.

### TBase has moved!

TBase, the database of transgenic animals and targeted mutations, has moved to the Jackson Laboratory, Maine, USA.

The new URL is:

<http://www.jax.org/tbase>

Be sure to adjust your bookmarks accordingly!

For questions or additional information please contact:

tbases@jax.org

**Anna V. Anagnostopoulos**

anna@jax.org

The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA.