Psychiatria Danubina, 2009; Vol. 21, No. 4, pp 562–569 *Conference paper*

HUMAN GENOME VARIATION IN HEALTH AND IN NEUROPSYCHIATRIC DISORDERS

Alena Buretić-Tomljanović[1] & Draško Tomljanović[2]

[1]Department of Biology and Medical Genetics, School of Medicine, University of Rijeka, Croatia
[2]Private psychiatric practice, Rijeka, Croatia

SUMMARY

Objectives: *Variation in the human genome may explain genetic contributions to complex traits and common diseases.*

Findings: *Until recently, single nucleotide polymorphisms were thought to be the most prevalent form of interindividual genetic variation. However, structural genomic rearrangements such as deletions, duplications, and inversions lead to variation in gene copy number and contribute even more to genomic diversity. Other sources of genomic variation include noncoding genes, pseudogenes, and mobile genetic elements (transposons).*

Conclusions: *Genome dynamics, including changes in gene number and position as well as epigenetic modifications of coding and noncoding sequences, can affect regulation of gene expression and may contribute to the variability of complex phenotypes.*

Key words: *human genome variation - single nucleotide polymorphisms - copy number variation - noncoding genes - pseudogenes - transposable elements - neuropsychiatric disorders*

* * * * *

INTRODUCTION

It has long been recognized that both genetic and environmental factors contribute to neuropsychiatric conditions. Twin studies, and most notably twin studies of schizophrenia, clearly reveal the genetic etiology of neuropsychiatric diseases. However, finding the genes that are responsible has proved challenging. Extensive association analyses have been performed on hundreds of single nucleotide polymorphisms (SNPs) in 14 candidate genes for schizophrenia, but no single SNP carries a substantial genetic risk for the disease (Sanders et al. 2008). In addition to SNPs, structural genomic variants or copy number variation (CNV) have been identified recently as a source of genomic variability that is associated with traits that impact health and disease (Sharp et al. 2005, Freeman et al. 2006, Lee & Lupski 2006, Redon et al. 2006). CNV includes deletions, duplications, and disruptions of dosage-sensitive genes. The contribution of CNV to variability in the human genome is greater than previously suspected, even exceeding the variability due to SNPs.

Some genes affected by CNV may contribute to neuropsychiatric disease susceptibility. The complex nature of neuropsychiatric disorders, along with the stable genetic background, indicates the significant role that variability in the human genome plays in their etiology. The human genome is a highly dynamic structure that is affected by environmental factors. Our growing knowledge about genomic structure, dynamics, and function may provide insights into the interplay of genetic and environmental components of complex traits and common diseases. In the human genome, structure and function are shaped by a variety of noncoding sequences that include introns, noncoding genes, mobile genetic elements, and pseudogenes. The aim of this report is to review the dynamic organization of the human genome and the implication of genomic variation for neuropsychiatric disorders.

PROTEIN-CODING AND NONCODING GENES

The human genome consists of coding and noncoding regions. Interestingly, exons coding for

proteins comprise only ~1.5% of the genome. Until recently, most of the noncoding regions were regarded as "junk sequences" without any particular function. However, most (60–70%) of the genome is transcribed (the so-called "transcriptome"); however, only a small portion of the transcribed RNA is translated into protein. Approximately 27–30% of the transcriptome plays a part in protein-coding genetic expression but does not code for protein (e.g., introns that are excised after transcription); the rest of the transcriptome represents noncoding genetic expression. About 98% of genomic transcription is noncoding (genome.wellcome.ac.uk). Noncoding genes are transcribed into small microRNA (miRNA) molecules that help regulate protein-coding gene expression. As of 2008, 677 miRNAs had been identified in the human genome (www.microrna.org). Each miRNA regulates a set of genes, or, alternatively, multiple miRNAs can work cooperatively to regulate a particular messenger RNA (mRNA) from a single gene. Thus, tissues or cell lines from normal or diseased mammalian tissue can have different miRNA expression profiles.

The human genome contains about 20,500 protein-coding genes (Clamp et al. 2007), although this number is still an estimate. Protein-coding genes have remained relatively stable throughout evolution, as evidenced by the similarity of the proteomes of different species. In contrast, noncoding DNA has been dynamic throughout evolution. Several lines of evidence suggest that noncoding RNAs (ncRNAs) affect the mechanisms used to regulate the expression of protein-coding genes (Mattick & Makunin 2006, Gingeras 2007), but also heterochromatic silencing, transposable element silencing, X chromosome inactivation, genomic imprinting, and other epigenetic phenomena (Zaratiegui et al. 2007). Thus, sophisticated RNA-based gene regulation mechanisms, i.e. small regulatory RNAs, have emerged and assisted in the evolution of complex multicellular organisms.

Small regulatory RNAs or ncRNAs may be involved in all aspects of protein-coding gene expression in eukaryotes. ncRNAs include a variety of molecules: small nucleolar RNAs (snoRNA) that are involved in rRNA maturation in the nucleolus; miRNAs that target mRNA molecules and suppress their translation into protein; small interfering RNAs (siRNAs) involved in the process of RNA interference (RNAi), which is associated with gene silencing; antisense RNAs that arise from transcription of the noncoding DNA strand concordantly with the transcription of the coding strand; and pseudogene RNAs, some of which are transcribed and may have a biological function.

Small regulatory RNAs represent a network of intracellular signals that regulate gene expression during normal physiology and development and modulate developmental timing, cellular proliferation, asymmetric gene expression (the development of the left and right sides of the body and left-right neuronal patterning), neuronal cell fate, apoptosis, brain morphogenesis, embryonic stem cell proliferation, lipid metabolism, and so forth. Given the multiple roles of ncRNA, RNA species may be critical in determining many if not most complex traits. Small RNAs represent an as-yet unexplored arena of genetic variability, both in humans and in other species. They may contribute to inter-individual variability and to complex disease susceptibility, including both malignant and neurologic diseases (Mattick & Makunin 2006).

INTRONS

Protein-coding exons of eukaryotic genes are separated by noncoding sequences called introns that span approximately 30% of the human genome. After protein-coding genes are transcribed, introns are precisely excised in the nucleus and the exons are combined to form mature mRNA molecules. This process is called RNA splicing. A complex of small nuclear RNAs and proteins, termed small nuclear ribonucleoprotein particles (snRNPs), are essential for splicing. Spliceosomal introns in eukaryotic genes probably evolved from group II introns, found in bacteria and mitochondria, which have the ability to self-splice (Martin & Koonin 2006). During evolution, some spliceosomal introns have acquired a number of functions and have become an important source of noncoding RNAs. It was thought previously that introns were degraded after excision, but they are actually cut into smaller RNA molecules (miRNAs and siRNAs) that have regulatory functions. In addition to being derived from introns, small regulatory RNAs may also be encoded in intergenic DNA, pseudogenes, or mobile genetic elements.

MOBILE GENETIC ELEMENTS

Most plant and animal genomes allow expression and transposition of several families of mobile genetic elements or transposons. Transposons are abundant in the genomes of placental mammals: In humans, for example, transposons make up approximately 40% of the genome. Mobile genetic elements are very polymorphic in humans and are useful for determining genetic distances between populations and for analyzing population structure (Muotri et al. 2007).

There are two main classes of retrotransposons, which are a class of RNA transposons that use reverse transcriptase for transposition: long-interspersed nuclear (LINE-1 or L1) elements and short-interspersed nuclear elements (SINEs). L1 elements are closely related to self-spliced group II introns. L1 elements represent 20% of the human genome, encoding reverse transcriptase, endonuclease, and the RNA II promotor sequence; this gives them autonomy for transposition. The most abundant SINEs are the *Alu* elements. *Alu*s contain sequences encoding small RNAs and require L1 for transposition. SINEs modulate the key basic genetic mechanisms of post-transcription regulation, including alternative splicing, RNA editing, and translation. SINEs can act as promotors or enhancers to regulate gene expression. *Alu* elements, for instance, are found in the 3'-UTR (untranscribed regions) of protein-coding genes and also interact with miRNAs. Because they contain RNA polymerase II promotor sequences, L1 elements can affect epigenetic modifications, such as methylation status, at the insertion site and change the pattern of the expression of the neighboring gene(s).

Retrotransposons have expanded until there are more than one million copies in the human genome. These elements have probably been crucial in the evolution and shaping of mammalian genomes (Deininger & Batzer 1999, Biémont & Vieira 2006, Muotri et al. 2007). The evolutionary rate of retrotransposition (10^{-3}–10^{-5}) is higher than for nucleotide substitution (10^{-8}–10^{-9}). During evolution, intrachromosomal or interchromosomal illegitimate (non-allelic) homologous recombination between transposons has resulted in insertional mutagenesis, deletions, and gene rearrangements. Although not all retroelements are capable of moving, L1 elements are very active in germline cells and embryonic stem cells. Importantly, genome rearrangements caused by retrotransposition events in these cells are transmitted to the next generation. Keeping in mind the extensive redistribution of epigenetic markers during gametogenesis and in early stages of embryonic development, retrotransposition activity might be expected. Retrotransposition of L1 seems to preferentially involve genomic regions surrounding neuronal genes that are active during early phases of neuronal differentiation. Random insertions and gene rearrangements might change gene expression patterns and influence neuronal fate. In this instance, the neurons become genetic mosaics in terms of L1 content. The results of Coufal et al. (2009) strongly support the existence of somatic mosaicism in the hippocampus and in several regions of the human brain. By developing a quantitative multiplex polymerase chain reaction, this group determined that there was a higher copy number of endogenous L1s in human brain tissue compared to heart or liver tissue samples. It is still unclear whether the genetic mosaicism of the retrotransposons in the neuronal network has functional consequences, possibly influencing complex traits such as behavior and contributing to the interindividual variability of cognitive and other psychological features. Mosaicism could also contribute to genetic predisposition to disease.

Somatic mosaicism due to retrotransposition could have important consequences. Retrotransposons are usually methylated and, as such, are probably inactivated in somatic cells. If somatic retrotransposition occurs, individuals would have a variety of random genomic rearrangements in their somatic tissues (somatic mosaicism). Consequently, the mechanisms of regulation of gene expression would vary and be specific to each individual cell.

The epigenetic mechanisms of retrotransposon regulation are poorly understood. It is thought that environmental factors such as nutrition and chemicals affect methylation enzymes, thus modulating normal physiological processes (Biemont & Vieira 2006). Retrotransposons are good candidates for being the sites at which complex genome/environmental interactions take place.

Transposons, and L1 elements in particular, have also been crucial in the genesis of a class of pseudogenes comprising processed or retrotransposed pseudogenes that were first copied from mRNA and then incorporated into the chromosome. In summary, during evolution, transposons

have been integrated into a complex regulatory network that affects gene expression regulation.

PSEUDOGENES

By definition, pseudogenes are genes that have lost their function. With an average length of 830 bp, pseudogenes cover 0.39% of the human genome. They arise through different mechanisms, for example, via gene duplication events, mutations or, in most instances, through retrotransposition of copies of random genes by L1 retroelements acting as mediators. These processed pseudogenes lack introns. The number of pseudogenes in the human genome is estimated to be 8,000–20,000. A biological function has been proposed for one mouse pseudogene, *Makorin1,* which appears to regulate the mRNA stability of its homologous protein-coding gene (Hirotsune et al. 2003). Subsequently, a genome-wide survey for pseudogenes with a biological function revealed the first examples of conserved pseudogenes that are common to human and mouse. The identified pseudogenes belonged to the poorly characterized *Ataxin* gene family, which includes a number of genes related to neurodegenerative disorders (Svensson et al. 2006). More recently, noncoding siRNA that regulates transcripts in mouse oocytes has been found to be derived from a pseudogene (Watanabe et al. 2008). The involvement of pseudogenes in the complex network of gene expression regulation, along with their evolutionary conservation, strongly suggests that they have important biological functions. Pseudogenes mostly arise from genes expressed in embryonic stem cells and germ cells, implying that they contribute to genome dynamics and variability.

SINGLE NUCLEOTIDE POLYMORPHISMS

Since the completion of the Human Genome Project in 2001, single nucleotide polymorphisms (SNPs) have been considered the main source of human genome variation. SNPs account for most of the estimated 0.1% variability between the genomes of different individuals. SNPs include both single nucleotide substitutions and single base pair insertions or deletions. SNPs can occur in protein-coding regions, in introns, or in intergenic regions. SNPs occur as frequently as every 100-300 bases, and more than 10 million SNPs may be present in the human genome. Numerous association studies have been conducted to detect genes associated with a predisposition to a variety of common diseases, including neurologic and psychiatric diseases. It was expected that SNP variability would be the basis for pharmacogenomics and personalized medicine, i.e. that SNPs would be identified that were associated with individuals' responses to drugs. Genome-wide association studies investigating more than 300,000 polymorphic markers spread evenly across the genome have detected genomic regions and genes involved in neuronal function and development; however, the findings of associated alleles could not generally be replicated in other populations. This was most likely due to the modest effects of the particular SNPs, difficulties in detecting rare alleles, interactions between different genes, and also, in the case of neuropsychiatric disorders, the lack of measurable physiological parameters (Baum et al. 2008, Adeyemo & Rotimi 2009, Hardy & Singleton 2009, Ng et al. 2009, Ollila et al. 2009).

STRUCTURAL GENOMIC REARRANGEMENTS AND COPY NUMBER VARIATION

In addition to qualitative differences between individuals in the form of SNPs, structural genomic rearrangements yielding individual-specific genomic architecture leads to quantitative variation in the human genome and is an additional source of genomic variability.

In the broadest sense, structural genomic rearrangements include all genomic variations that are not SNPs: deletions, inversions, translocations, and duplications. These rearrangements can result in copy number differences in chromosomal region(s) or gene(s), which is termed copy number variation (CNV). CNV refers to a segment of DNA that exists in different copy numbers in the genomes of different individuals. The following has become clear: CNVs are widespread and common in the human genome; approximately 6–19% of each chromosome is affected by CNVs; CNVs may be inherited or originate *de novo*; and their distribution across the genome is not random.

CNVs appear mostly in unstable genomic regions that are capable of rapid evolution of new genes and gene variants. CNVs were originally detected during efforts to map human SNPs

(International HapMap Project). Unexpectedly, some SNPs were missing from the genomes of some individuals. Large-scale submicroscopic deletions (>100 kb) are the most common form of CNV. CNV affects gene dosage, shows interpopulation differences, and contributes substantially to phenotypic variability in both health and disease states (McCarroll & Altshuler 2007, Cook & Scherer 2008, Wain et al. 2009). A well known example of CNV is the variable copy number, ranging from 1 to 14 copies, of the CCL3L1 gene and the corresponding effect of copy number on resistance to HIV infection. The higher the CCL3L1 copy number, the better the resistance.

CNVs are affected by environmental factors. One example is the copy number of the AMY1 gene, which codes for amylase (Perry et al. 2007). The copy number of the AMY1 gene is associated with expression of the amylase protein in saliva. Higher copy numbers of this gene are found in populations that consume more starch in their diet.

The effect of CNVs on gene dosage has substantial implications for Mendelian traits and disorders. An individual inherits two copies of nearly all genes in the human genome, one copy from each parent. CNVs are an exception to this rule. CNVs can change the biallelic state at a genetic locus into a monoallelic or triallelic state (Lupski, 2008). When this happens, CNVs can disrupt the SNP evaluation (which is biallelic) and give false results in a linkage analysis.

CNVs occur in genomic (or chromosomal) regions that show instability and are prone to rearrangements. These regions are enriched in repeated DNA fragments or segmental duplications (segments of DNA with near-identical sequences). The existence of duplications across the genome implies errors during replication, illegitimate non-allelic homologous recombination, or non-disjunction events during cell division that are the basis for the genesis of the structural rearrangements and aneuploidy. Several studies have suggested that there is constitutional aneuploidy in normal mouse and human brains (Rehen et al. 2005, Kingsbury et al. 2005, Iourov et al. 2008, Arendt et al. 2009). Therefore, genomic architecture could be the crucial factor for predisposition to genomic instability. Aneuploidy of chromosomes 17 and 21, for instance, has been found in 4–19% of human brain cells (both non-neuronal cells and functionally active postmitotic neurons), and sex-chromosome aneuploidy has been observed in 0.2% of functionally active neurons. Since the aneuploidy status has not yet been determined for most autosomes, the extent of constitutional aneuploidy in the normal human brains could be even greater. Although the functional significance of the intermixed population of euploid and aneuploid neurons is unknown, CNV modulates the gene dosage of numerous genes. This is likely to have a profound impact on cellular physiology, possibly increasing or decreasing disease susceptibility. It is interesting that genomic regions affected by CNVs are involved in detecting environmental sensory signals and contain genes crucial for molecular-environmental interactions such as sensory perception, cell adhesion, chemical stimuli processing, neurophysiologic processing, detoxication processes, immune response and disease resistance, and inflammation. CNVs also involve genes that contribute to interindividual variation in drug response. CYP2D6, for instance, has a highly polymorphic copy number that modulates enzyme expression (Wain et al. 2009).

Overall, the mutation rate of CNV in humans is high, with CNV arising in about 1/10,000 meiotic divisions (St Clair 2009). While the majority of CNVs are rare and found in small number of individuals, others occur frequently. Cytogenetic studies show that large deletions and duplications can be present in healthy, phenotypically normal people with no developmental delays or highly penetrant diseases. A number of microdeletion events have been associated with severe disorders such as Prader-Willi and Angelman syndromes (15q11-13), Smith-Magenis syndrome (17p11.2), and others. These observations led to the hypothesis that CNVs that do not cause early, highly penetrant disorders, might play a role in the genesis of later-onset genomic disorders or common diseases (Freeman et al. 2006).

HUMAN GENOME VARIATION AND NEUROPSYCHIATRIC DISORDERS

CNVs affecting gene dosage have been implicated in many clinically distinct entities, including neuropsychiatric disorders (Lee & Lupski, 2006). For example, germline duplication of the amyloid precursor protein locus (APP) on chromosome 21q21 has been associated with the risk of developing autosomal-dominant early-onset

Alzheimer's disease (AD). Individuals with Down syndrome are also at higher risk for developing neuropathologic features of early-onset AD. Regarding Parkinson's disease, variable copy number of the alpha-synuclein (SNCA) promoter on chromosome 4q21 is associated with an increased risk for the disease and for disease severity (Ross et al. 2008). Rearrangements of the 22q chromosomal region have been associated with both cognitive deficits and schizophrenia (Bassett et al. 2008). Uniparental disomy (UPD) for chromosome 4 has been repeatedly associated with mood disorders. CNVs in chromosomal regions 1q21, 15q11.2, 15q13.3, 16p11.2, 22q12, and at the *Neurexin 1* locus at 2p16.3 (ranging between 400 kb and 1.6 Mb) markedly increase the risk of mental retardation, autism, schizophrenia, bipolar disorder, and attention-deficit hyperactive disorder (ADHD) (St Clair et al. 2009, Singh et al. 2009). The frequency of rare CNVs (>100 kb) in schizophrenia and related disorders is consistently higher in affected individuals than in control healthy subjects (Stefansson et al. 2008; The International Schizophrenia Consortium 2008), and CNVs could potentially be used to distinguish familial and sporadic cases of schizophrenia. Specifically, rare CNVs that originate *de novo* are found more frequently in sporadic cases of schizophrenia. *De novo* CNVs together with those inherited confer a higher risk for developing schizophrenia or related disorders. Somatic mosaicism for CNVs might explain monozygotic (MZ) twins who are discordant for schizophrenia. Bruder et al. (2008) showed that 19 MZ twin pairs differed in terms of the CNVs in their genomes. These results suggest that CNVs are continuously acquired during one's lifetime, so that each individual is a mosaic. In the last two years, intense effort has gone into investigating the role of CNVs in schizophrenia. These studies show that CNVs that cause schizophrenia are multiple, rare, and individually specific, meaning each individual has a different set of affected genes. This genetic heterogeneity is consistent with the well known clinical heterogeneity of schizophrenia patients.

Retroviruses have also been implicated in the pathogenesis of multiple sclerosis, schizophrenia, and other complex diseases. The frontal cortex of patients with schizophrenia and bipolar disorder (Yolken et al. 2001), as well as the brain tissue and mononuclear blood cells of patients with multiple sclerosis (Mameli et al. 2007), show an increased transcription rate of several human endogenous retroviruses, a class of retrotransposons, compared to healthy controls. Retroelements may act as alternative promotors (Okamura & Nakai 2008), mediating the transcription rate of genes located downstream of the site of retrotransposition. Alternatively, endogenous retroviruses may be activated by hormones during fetal development, or by cytokines generated during acute infections. Recently, a role for *Alu* repeats has been proposed in the genesis of recurrent 22q11 microdeletions; these microdeletion confer a 30% greater risk of developing psychosis (Uddin et al. 2006). *Alu* repeats were found in close proximity to all known deletion breakpoints, suggesting their involvement in the intrachromosomal homologous recombination events, and therefore in the genesis of this particular CNV.

CONCLUSION

The human genome is not a static repository of genetic information. On the contrary, the genome is dynamic, undergoing multiple and diverse structural rearrangements that are unique to each individual, individual tissue, or even individual cell. In most cases, structural rearrangements occur in unstable genomic regions that are prone to errors during DNA replication or during recombination events due to the presence of highly homologous duplicated sequences in these regions. Structural rearrangements can have a major impact on dosage-sensitive genes, such as those that control neurodevelopment and genetic interactions with environmental stimuli.

The dynamics of the human genome, the fluidity associated with the broad expansion of retrotransposons, and the evolution of the complex network of ncRNA molecules implicated in gene expression regulation may all play key roles in the evolution of the human genome and in the development of our extremely complex nervous system. Other sources of human genome variation, not presented in this review, may also contribute.

The dynamic nature of the human genome may explain interindividual diversity, interpopulation differences, and even susceptibility to common diseases, including neuropsychiatric conditions. New findings about genome dynamics and variation raise questions about clinical diagnostics, modes of inheritance, and genetic counselling for neuropsychiatric disorders; these

findings also hold promise for the development of novel treatments and advances in personalized medicine (Lee & Morton 2008).

REFERENCES

1. *Adeyemo A & Rotimi C: Genetic Variants Associated with Complex Human Diseases Show Wide Variation across Multiple Populations. Public Health Genomics 2009 (in press) (DOI: 10.1159/000218711).*
2. *Arendt T, Mosch B, Morawski M: Neuronal aneuploidy in health and disease: a cytomic approach to understand the molecular individuality of neurons. Int J Mol Sci 2009; 10:1609-27.*
3. *Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B et al: A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry 2008; 13:197-207.*
4. *Bassett AS, Marshall CR, Lionel AC, Chow EWC, Scherer SW: Copy number variations and risk for schizophrenia in 22q11.2 deletion syndrome. Hum Mol Genet 2008; 17:4045-53*
5. *Bruder C, Piotrowski A, Gijsbers A, Andersson R, Erickson S, Diaz de Stahl T et al: Phenotypically concordanr and discordant monozygotic tweens display different DNA copy number variation profiles. Am J Hum Genet 2008; 82:1-9.*
6. *Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF et al: Distinguishing protein-coding and noncoding genes in the human genome. Proc Natl Acad Sci U S A. 2007; 104:19428-33.*
7. *Cook EH Jr & Scherer SW: Copy-number variations associated with neuropsychiatric conditions. Nature 2008; 455:919-923.*
8. *Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M et al: L1 retrotransposition in human neural progenitor cells. Nature 2009; 460:1127-31.*
9. *Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM et al: Copy number variation: new insights in genome diversity. Genome Res 2006; 16:949-61.*
10. *genome.welcome.ac.uk [accessed August 25, 2009]*
11. *Gingeras TR: Origin of phenotypes: genes and transcripts. Genome Res 2007; 17:682-90.*
12. *Hardy JH & Singleton A: Genomewide Association Studies and Human Disease. N Engl J Med 2009; 360:1759-68*
13. *Hirotsune S, Yoshida N, Chen A, Garrett L, Sugiyama F, Takahashi S et al: An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 2003; 423:91-6.*
14. *Iourov IY, Vorsanova SG, Yurov YB: Molecular cytogenetics and cytogenomics of brain diseases. Curr Genomics 2008; 9.452-65.*
15. *Kingsbury MA, Friedman B, McConnell MJ, Rehen SK, Yang AH, Kaushal D et al: Aneuploid neurons are functionally active and integrated into brain circuitry. Proc Natl Acad Sci U S A 2005; 102:6143-7.*
16. *Lee JA, Lupski JR: Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron 2006; 52:103-21.*
17. *Lee C & Morton CC: Structural genomic variation snd personalized medicine. N Engl J Med 2008; 358:740-1.*
18. *Lupski JR: Incriminating genomic evidence. Nature 2008; 455:178-9.*
19. *Mameli G, Astone V, Arru G, Marconi S, Lovato L, Serra C et al: Brains and peripheral mononuclear cells of multiple sclerosis (MS) patients hyper-express MS-associated retrovirus/HERV-W endo-genous retrovirus, but not Human herpesvirus 6. J Gen Virol 2007; 88:264-7.*
20. *Martin W, Koonin EV: Introns and the origin of nucleus-cytosol compartmentalization. Nature 2006; 440:41-5.*
21. *Mattick JS, Makunin IV: Non-coding RNA. Hum Mol Genet 2006; 15 No 1:R17-29.*
22. *McCarroll SA & Altshuler DM: Copy-number variation and association studies of human disease. Nat Genet 2007; 38:S37-42.*
23. *Ng MYM, Levinson DF, Faraone SV, Suarez BK, DeLisi LE, Arinami T et al: Meta-analysis of 32 genome-wide linkage studies of schizophrenia. Mol Psychiatry 2009; 14:774-85.*
24. *Okamura K & Nakai K: Retrotransposition as a source of new promoters. Mol Biol Evol 2008; 25:1231-8.*
25. *Ollila HM, Soronen P, Silander K, Palo OM, Kieseppä T, Kaunisto MA et al: Findings from bipolar disorder genome-wide association studies replicate in a Finnish bipolar family-cohort. Mol Psychiatry 2009; 14:351-353.*
26. *Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R et al: Diet and the evolution of human amylase gene copy number variation. Nat Genet 2007; 39:1256-60.*
27. *Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD et al: Global variation in copy number in the human genome. Nature 2006; 444:444-54.*
28. *Rehen SK, Yung YC, McCreight MP, Kaushal D, Yang AH, Almeida BS: Constitutional aneuploidy in the normal human brain. J Neurosci 2005; 25:2176-80.*
29. *Ross OA, Braithwaite AT, Skipper LM, Kachergus J, Hulihan MM, Middleton FA et al: Genomic investigation of alpha-synuclein multiplication and parkinsonism. Ann Neurol 2008; 63:743-50.*

30. Sanders AR, Duan J, Levinson DF, Shi J, He D, Hou C et al: No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. Am J Psychiatry 2008; 165:497-506.
31. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU et al: Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 2005; 77:78-88.
32. Singh SM, Castellani CA, O'Reilly RL: Copy number variation showers in schizophrenia: an emerging hypothesis. Mol Psychiatry 2009; 14:356-8
33. St Clair D: Copy number variation and schizophrenia. Sch Bull 2009; 35:9-12.
34. Stefansson H, Rujescu D, Cichon S, Pietiläinen OPH, Ingason A, Steinberg S et al: Large recurrent microdeletions associated with schizophrenia. Nature 2008; 455:232-6.
35. Svensson O, Arvestad L, Lagergren J: Genome-wide survey for biologically functional pseudogenes. PLoS Comput. Biol. 2006; 2: e46.
36. The International Schizophrenia Consortium: Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature 2008; 455:237-41.
37. Uddin RK, Zhang Y, Siu VM, Fan YS, O'Reilly RL, Rao J et al: Breakpoint associated with a novel 2.3 Mb deletion in the VCFS region of 22q11 and the role of Alu (SINE) in recurring microdeletions. BMC Med Genet 2006; 7:18
38. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y et al: Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 2008; 453: 539-43.
39. Wain LV, Armour JAL, Tobin MD: Genomic copy number variation, human health, and disease. Lancet 2009; doi10.1016/s0140-6736(09)60249-X
40. www.hapmap.org [accessed August 25, 2009]
41. www.microrna.org/microrna/releaseNotes.do [accessed August 25, 2009]
42. Yolken RH, Karlsson H, Bayer TA, Johnston-Wilson N, Yee F, Fuller Torrey E: Retroviruses, genes and schizophrenia. Clin Neurosci Res 2001; 1.164-9.
43. Zaratiegui M, Irvine DV, Martienssen RA: Noncoding RNAs and gene silencing. Cell 2007; 128:763-76.

Acknowledgements

This work was supported by grant No. 062-0982522-0369 from the Ministry of Science, Education, and Sports, Zagreb, Croatia.

Correspondence:

Alena Buretić-Tomljanović, PhD, Associate professor
Dept. of Biology and Medical Genetics, University of Rijeka
Braće Branchetta 22, Rijeka, Croatia
E-mail: alena medri.hr @pbsvi.hr