

Review

## Copy number variations and cancer

Adam Shlien<sup>\*†</sup> and David Malkin<sup>\*†‡</sup>

Addresses: <sup>\*</sup>Departments of Genetics and Genome Biology and Division of Hematology/Oncology, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada, M5G 1X8. <sup>†</sup>Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada, M5G 2M9. <sup>‡</sup>Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada, M5G 2M9.

Correspondence: David Malkin. Email: david.malkin@sickkids.ca

Published: 16 June 2009

*Genome Medicine* 2009, **1**:62 (doi:10.1186/gm62)

The electronic version of this article is the complete one and can be found online at <http://genomemedicine.com/content/1/6/62>

© 2009 BioMed Central Ltd

### Abstract

DNA copy number variations (CNVs) are an important component of genetic variation, affecting a greater fraction of the genome than single nucleotide polymorphisms (SNPs). The advent of high-resolution SNP arrays has made it possible to identify CNVs. Characterization of widespread constitutional (germline) CNVs has provided insight into their role in susceptibility to a wide spectrum of diseases, and somatic CNVs can be used to identify regions of the genome involved in disease phenotypes. The role of CNVs as risk factors for cancer is currently underappreciated. However, the genomic instability and structural dynamism that characterize cancer cells would seem to make this form of genetic variation particularly intriguing to study in cancer. Here, we provide a detailed overview of the current understanding of the CNVs that arise in the human genome and explore the emerging literature that reveals associations of both constitutional and somatic CNVs with a wide variety of human cancers.

### Copy number variations: dynamic genomes

Our genomes are not the stable places we once thought they were. Recent genome-wide studies have shed light on copy number variations (CNVs), an unexpectedly frequent, dynamic and complex form of genetic diversity, and have quickly overturned the idea of a single diploid human 'reference genome'. Although the characterization of the extent and location of these regions in healthy genomes is far from complete, many groups, including ours, are actively trying to determine the clinical impact of CNVs in patient populations.

CNVs are structurally variant regions in which copy number differences have been observed between two or more genomes [1]. Defined as being larger than 1 kilobase (kb) in size, CNVs can involve gains or losses of genomic DNA that are either microscopic or submicroscopic and are, therefore, not necessarily visible by standard G-banding karyotyping. Until recently, only a few copy-number-variable loci had been identified, such as duplications at the  $\alpha 7$ -nicotinic

receptor gene (*CHRNA7*) at 15q13-15 [2] and variation at the major histocompatibility complex locus [3]. In 2004, significant advances in DNA array technology enabled the discovery of many CNVs, revealing a novel and pervasive form of inter-individual genomic variation [4,5]. These pioneering genome-scale efforts used two different platforms to find 76 CNVs in 20 individuals [5] and 255 CNVs in 55 individuals [4], some of which were common to both studies, suggesting possible hotspot regions of CNVs in the human genome. Even this was soon found to be an under-representation of the number of CNVs; follow-up studies have since ascertained many thousands of CNV regions in hundreds of healthy individuals. In fact, the recent increase in scientific interest in CNVs, combined with improvements in microarray fabrication (higher density at lower cost) and the development of new informatics techniques, have led to the ascertainment of approximately 21,000 CNVs, or around 6,500 unique CNV loci, in the five short years since this form of genetic variation was first revealed (these figures come

from the March 2009 update of the Database of Genomic Variants (DGV) [4]). CNVs are now thought to cover at least 10% of the human genome. Furthermore, next-generation sequencing technologies will soon be used to sequence thousands of genomes along with their CNVs.

### CNVs and disease: mutable genomes

The CNV map for the human genome is being continuously refined and has already pinpointed the location, copy number, gene content, frequency and approximate breakpoints of numerous CNVs in the healthy population. These structural variants can alter transcription of genes by altering dosage or by disrupting proximal or distant regulatory regions, as has been shown globally in the healthy human [6], mouse [7] and rat genomes [8]. It is, however, the specific disease-associated CNV loci that have been particularly scrutinized and that therefore provide the most detailed examples of how CNVs can alter cellular function. We will highlight three insights in particular from the literature: that pathogenic CNVs often contain multiple genes, that the effect of a pathogenic CNV is not limited to the gene(s) it contains, and that pathogenic CNVs can have reciprocal deletions/duplications.

### The number of genes in pathogenic CNVs

Genomic rearrangements give rise to a variety of diseases classified as 'genomic disorders' [9]. Because they involve large regions, it is common for genomic disorders to include many deleted or duplicated genes, unlike traditional mutations that affect a single coding-region change of one gene. These genes can be either fully encompassed or partially overlapped by the pathogenic CNV. Deletions of 22q11.2 are associated with DiGeorge/velocardiofacial syndrome and include the catechol-O-methyltransferase gene, the T box transcription factor 1 gene and others [10]. Similarly, the autosomal dominant Prader-Willi syndrome (15q11-q13 deletion) involves many genes [11], and the Williams-Beuren syndrome (7q11.23 deletion) involves 28 genes [12]. As microarray resolution increases, genomic disorders will certainly be found that are caused by small CNVs involving only a single gene, or even a portion of one gene.

### The source of the effect of a pathogenic CNV

Usually, the genes contained in the pathogenic CNV are candidates for association with the clinical phenotype under study. However, research on genomic disorders has shown that some genes within a CNV may not be necessary, or may not be sufficient, to cause the observed disease. For example, a recurrent 3.7 Mb microdeletion is responsible for 70% of cases of Smith-Magenis syndrome (SMS) [13], a neuro-behavioral disorder involving sleep disturbance, craniofacial and skeletal anomalies, intellectual disability and distinctive behavioral traits. Although the size of the deletions observed varies, the identification of a common 'critical region' (1.5 Mb) in SMS patients led to the conclusion that the retinoic acid induced 1 (*RAI1*) gene alone is responsible for

most SMS features. Indeed, *RAI1* point mutations have been seen in patients without deletions with similar phenotypes, thus confirming that this gene (of the 13 in the critical region) is necessary to cause SMS. Patients with additional genes deleted have a variable and more severe phenotype. In contrast, in Williams-Beuren syndrome, not only the aneuploid genes but also genes far outside the deleted region have reduced expression and are thought to contribute to the phenotype [14]. Such long-range influence of CNVs on distant gene expression is proposed to be caused by positional effects [15].

### Reciprocal deletions and duplications

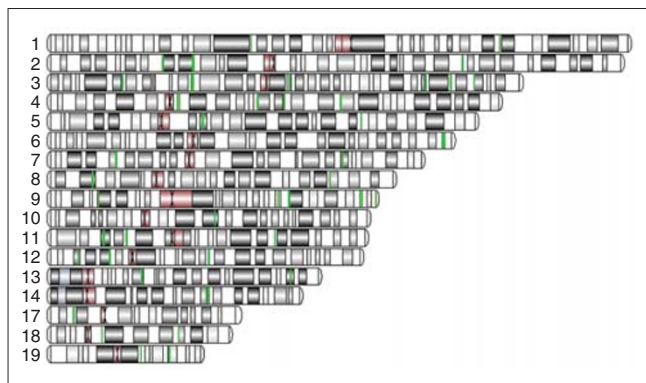
Recombination between highly homologous sequences (non-allelic homologous recombination) can generate deletions, duplications, inversions and translocations. The sequence architecture that allows one copy number change can also allow its reciprocal at the same locus. The reciprocal events usually cause different phenotypes and occur at different frequencies in the population and at different rates during meiosis [16].

### CNVs and cancer predisposition: first hits to the tumor genome

The goal of cancer genetics is to discover all variant alleles that predispose to neoplasms. To this end, single nucleotide polymorphisms (SNPs) have been the most widely studied form of genetic variation and, by using massive whole-genome studies (genome-wide association (GWA) studies), many common SNPs have been shown to be associated with cancer and other complex traits. However, the results of these efforts have not explained much of the heritability of disease [17]. This is perhaps because GWA studies have mostly ignored the inter-individual genetic variation provided by CNVs, which affect more than 10% of the human genome. CNVs, especially smaller variants, have been essentially hidden from view until recently; thus, only a handful of studies have found an association of CNVs with cancer. Once these CNVs have been identified, one can only assume that CNVs will explain a larger portion of the genetic basis of cancer. Once identified, common and rare CNVs should be considered separately, as they may have very different roles in cancer.

### Common cancer CNVs

As with SNPs, CNVs that are found frequently in the healthy population (common CNVs) are very likely to have a role in cancer etiology. In the only study published so far that begins to test the hypothesis that common CNVs are associated with malignancy, we [18] created a map of every known CNV whose locus coincides with that of *bona fide* cancer-related genes (as catalogued by [19]); we called these cancer CNVs. In an initial analysis [18], we examined 770 healthy genomes using the Affymetrix 500K array set, which has an average inter-probe distance of 5.8 kb. As CNVs are



**Figure 1**

Distribution of common cancer CNVs in the human genome. The chromosomes containing common cancer CNVs in the human genome are shown, with centromeric regions in red (using data from [19]) and Giemsa banding patterns in white, grey or black. Loci are in green if they were found to contain a cancer-related gene that is overlapped or encompassed by a CNV (as found by [18]).

generally thought to be depleted in gene regions [20], it was surprising to find 49 cancer genes that were directly encompassed or overlapped by a CNV in more than one person in a large reference population (Figure 1). In the top ten genes, cancer CNVs could be found in four or more people. In this analysis only CNVs directly overlapping a cancer gene were selected (either both breakpoints were inside the genomic interval containing the gene, both were outside the interval, or one breakpoint was inside while the other was outside). However, this is probably an underestimate of the actual number of common cancer CNVs, for two reasons. First, many smaller variants are missed at the resolution of this array: the mean size of CNVs found using the Affymetrix 500K array is 206 kb [20], whereas the CNVs found using the newer Affymetrix 6.0 platform with a median inter-marker distance of less than 700 bp are 5-15 times smaller [21]. Second, as discussed above, there are unquestionably additional, more distal CNVs that have a long-range effect on cancer gene transcription levels.

Validating the initial observation [18], many of these genes are also found in the DGV, a curated list of CNVs compiled from numerous publications [4]. Analysis of the DGV [22] shows that nearly 40% of cancer-related genes are interrupted by a CNV. This trend continues: even among the ten most recent CNV publications in the DGV (those published after February 2008), many important tumor suppressor genes and oncogenes can be found with diverse functions, including apoptosis, control of cell cycle checkpoints and DNA repair, and numerous translocation and fusion gene partners. An example of this is *Rad51L1*, a gene that is a member of the RAD51 family; this is essential for DNA repair by homologous recombination and has been shown by a GWA study to contain a SNP that is strongly associated with breast cancer [23].

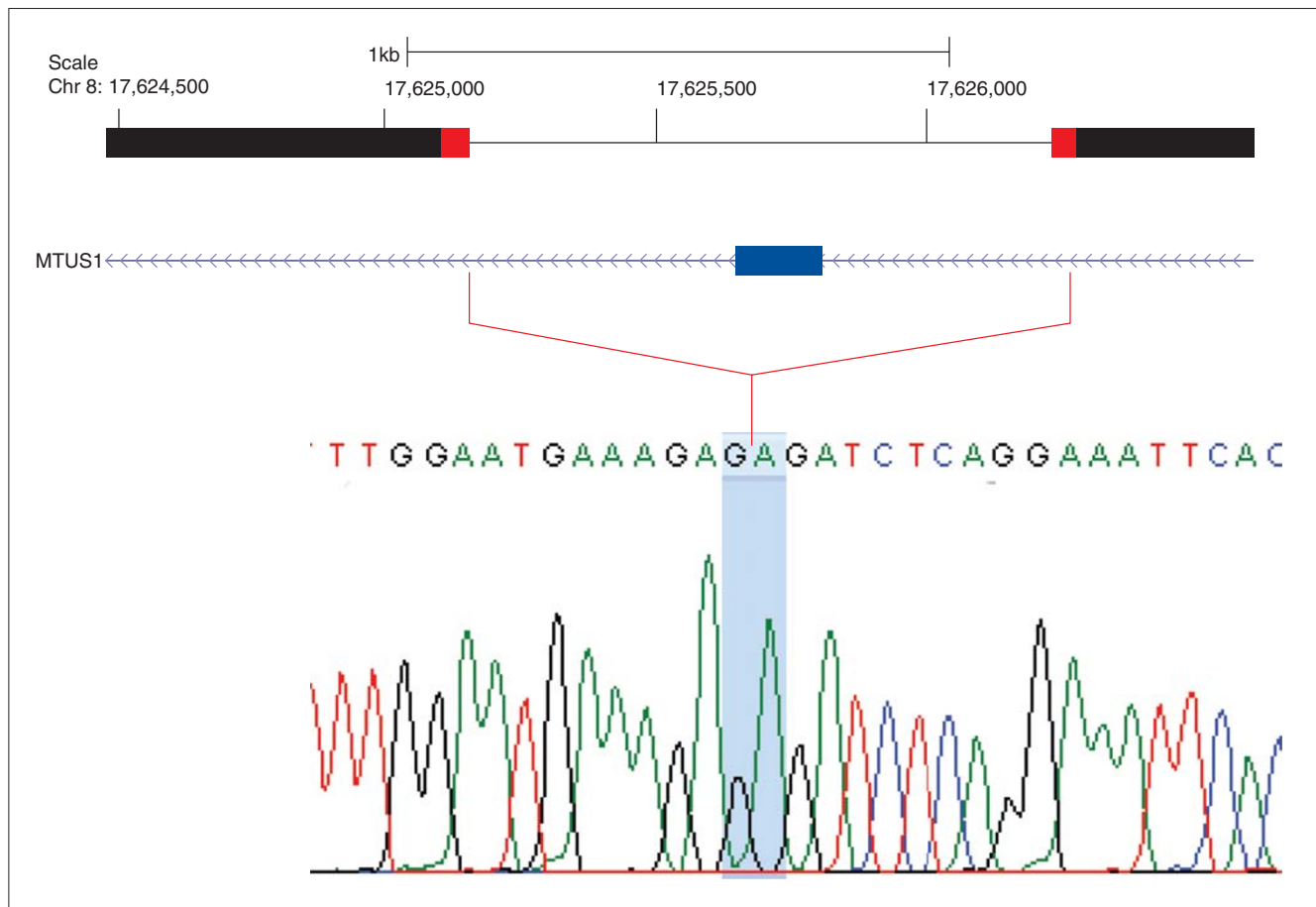
The challenge will be to determine which of these genes are dosage-sensitive and which tissues containing these common cancer CNVs will be susceptible to malignant transformation and growth. One approach is to characterize specific cancer CNVs in great detail, in terms of both population frequency and breakpoint sequence [24]. For example, in a pilot candidate-gene association study, we found a cancer CNV at the gene *MLLT4* (a *Ras* target encoding a protein that regulates cell-cell adhesion) that seems to be associated with the Li-Fraumeni cancer predisposition disorder (LFS); individuals affected with LFS harbor a germline heterozygous mutation of the *Tp53* tumor suppressor gene [18]. The frequency of this CNV is significantly increased in LFS ( $P = 0.006$ , Fisher's exact test): 3 of the 19 LFS probands (15.8%; observed/expected =  $3/0.4 = 7.5$ ) harbored the CNV duplication, whereas only 12 of 710 healthy individuals from the reference population (1.69%; observed/expected =  $12/14.6 = 0.82$ ) harbored the CNV.

A nice illustration of a focal CNV with phenotypic effect is given by the mitochondrial tumor suppressor gene (*Mtus1*); Frank *et al.* [25] found that a small deletion in *Mtus1* is associated with a decreased risk of familial and high-risk breast cancer. Using long-range PCR, we independently fine-mapped this common cancer CNV and genotyped it in a panel of healthy controls. Although it is only 1.1 kb in size, the deletion removes an entire exon of *Mtus1*. Direct sequencing reveals a 41 bp stretch of homology flanking the exon, which leads to this deletion by non-allelic homologous recombination (Figure 2).

These examples demonstrate hypothesis-driven approaches, which are restricted to genes for which there is an *a priori* association with cancer. Ultimately, it will be important to be able to discover and test every CNV in a genome for cancer susceptibility, but although this hypothesis-free approach is becoming technically tractable and more economical, such studies do have unique analytical challenges. As elaborated upon elsewhere [26,27] these challenges include: the unknown allele frequency and integer copy number of most CNVs, both within and among populations; the absence of sequence-level breakpoint information for most CNVs and the architectural complexity of some CNV regions, including smaller CNVs within larger ones [24].

#### Rare cancer CNVs

Common cancer SNPs - and by analogy common cancer CNVs - each confer only a minor increase in disease risk, but collectively they may cause a substantially elevated risk. In contrast, the mutations associated with hereditary cancer syndromes are frequently highly penetrant on their own and are usually inherited in an autosomal dominant manner. Unlike low-penetrance alleles, rare high-penetrance mutations will almost always co-segregate with the disease in families.



**Figure 2**  
 Cancer CNV breakpoint mapping. We mapped a 1.1 kb deletion in the mitochondrial tumor suppressor gene, *MTUS1*, to base-pair resolution. The affected portion of the gene is shown, including an exon (blue) that is deleted in the presence of the CNV. Two 41 bp repeats (with sequence AAATAAGAACCAAGTCCAAATACATCTTTGGAATGAAAGAG) were found at the breakpoints (red), while the sequence of the junction fragment is shown in the chromatogram.

There are over 200 cancer syndromes and although most arise infrequently, they account for 5-10% of all cancer cases [28]. These are caused by base-pair-sized germline mutations in many central tumor suppressor genes - such as *TP53*, *APC*, *BRCA1*, *BRCA2*, *PTEN*, and *RB1* - and (fewer) oncogenes, including *HRAS* and *RET*.

The role of large structural mutations in cancer syndromes has been less appreciated, probably because genomic deletions or duplications are not readily detected by PCR-based sequencing. New multiplexing methods, especially multiplex ligation-dependent probe amplification (MLPA) [29], allow targeted copy number assessment of single gene or exon changes. This has led to a recent upsurge in discoveries of patients and families with rare pathogenic CNVs that strongly predispose to cancer. Of the 70 germline cancer genes in the Cancer Genes Census [30], 28 have been reported to be mutated by genomic deletion or duplication (the genes and citations are shown in Table 1). We hypothesize

that many of the remaining gene mutations will be found to have a genomic equivalent and, perhaps more importantly, that predisposing CNVs will be found in other regions not usually associated with hereditary cancer. A recent report by Jackson *et al.* [31] describing five patients with rhabdoid predisposition syndrome and deletions at *SMARCB1* (22q11.2) highlights the benefits of a global approach to CNV detection: using SNP arrays to gain a broad perspective on the *SMARCB1* deletion and surrounding chromosomal landscape, it was found that the extent of two patients' deletions in fact extended past *SMARCB1*, impinging on neighboring genes, and explaining their clinical phenotype.

The presence of rare cancer CNVs leads to many questions: do they differ from base-pair changes at the same locus? What is their penetrance? What are the mutational processes that give rise to them? Do they have reciprocal deletions/duplications? Do they have long-range effects on gene expression? These questions provide fertile ground for

**Table 1****Rare cancer CNVs at known cancer-predisposing genes**

Gene	Cancer syndrome	References on genomic deletions or duplications
<i>APC</i>	Adenomatous polyposis coli; Turcot syndrome	Hodgson et al. [36], Su et al. [37], Aretz et al. [38] and Charames et al. [39]
<i>BMPRI A</i>	Juvenile polyposis	Delnatte et al. [40]
<i>BRCA1</i>	Hereditary breast/ovarian cancer	Petrij-Bosch et al. [41] and Montagna et al. [42]
<i>BRCA2</i>	Hereditary breast/ovarian cancer	Casilli et al. [43]
<i>CDKN2A-p14ARF</i>	Familial malignant melanoma	Lesueur et al. [44]
<i>CDKN2A-p16(INK4a)</i>	Familial malignant melanoma	Lesueur et al. [44]
<i>CHEK2</i>	Familial breast cancer	Cybulski et al. [45,46]
<i>FANCA</i>	Fanconi anemia A	Levrin et al. [47]
<i>MADH4</i>	Juvenile polyposis	van Hattem et al. [48]
<i>MEN1</i>	Multiple endocrine neoplasia type I	Kishi et al. [49]
<i>MLH1</i>	Hereditary non-polyposis colorectal cancer, Turcot syndrome	Nystrom-Lahti et al. [50] Chan et al. [51]
<i>MSH2</i>	Hereditary non-polyposis colorectal cancer	Stella et al. [52]
<i>MSH6</i>	Hereditary non-polyposis colorectal cancer	Plaschke et al. [53]
<i>NF1</i>	Neurofibromatosis type I	Riva et al. [54], Bausch et al. [55]
<i>NF2</i>	Neurofibromatosis type 2	Tsilchorozidou et al. [56]
<i>PRKAR1A</i>	Carney complex	Horvath et al. [57]
<i>PTCH</i>	Nevoid basal cell carcinoma syndrome	Shimkets et al. [58]
<i>RBI</i>	Familial retinoblastoma	Bremner et al. [59]
<i>SDHB</i>	Familial paraganglioma	Cascon et al. [60]
<i>SDHC</i>	Familial paraganglioma	Baysal et al. [61]
<i>SDHD</i>	Familial paraganglioma	McWhinney et al. [62]
<i>SMARCB1</i>	Rhabdoid predisposition syndrome	Swensen et al. [63]
<i>STK11</i>	Peutz-Jeghers syndrome	Le Meur et al. [64]
<i>TP53</i>	Li-Fraumeni syndrome	Bougeard et al. [65,66]
<i>TSC1</i>	Tuberous sclerosis I	Kozlowski et al. [67]
<i>TSC2</i>	Tuberous sclerosis 2	Kozlowski et al. [67]
<i>VHL</i>	von Hippel-Lindau syndrome	Richards et al. [68]
<i>WT1</i>	Denys-Drash syndrome, Frasier syndrome, Familial Wilms tumor	Huff et al. [69]

The 28 genes of the 70 germline cancer genes in the Cancer Genes Census [30] that have been reported to be mutated by genomic deletion or duplication are shown.

future research. These studies may involve identifying novel CNVs in unexplained familial clusterings of cancer, or the use of *in vitro* models in which cancer CNVs are created to measure their effect on cellular proliferation, genomic instability and the other hallmarks of cancer [32].

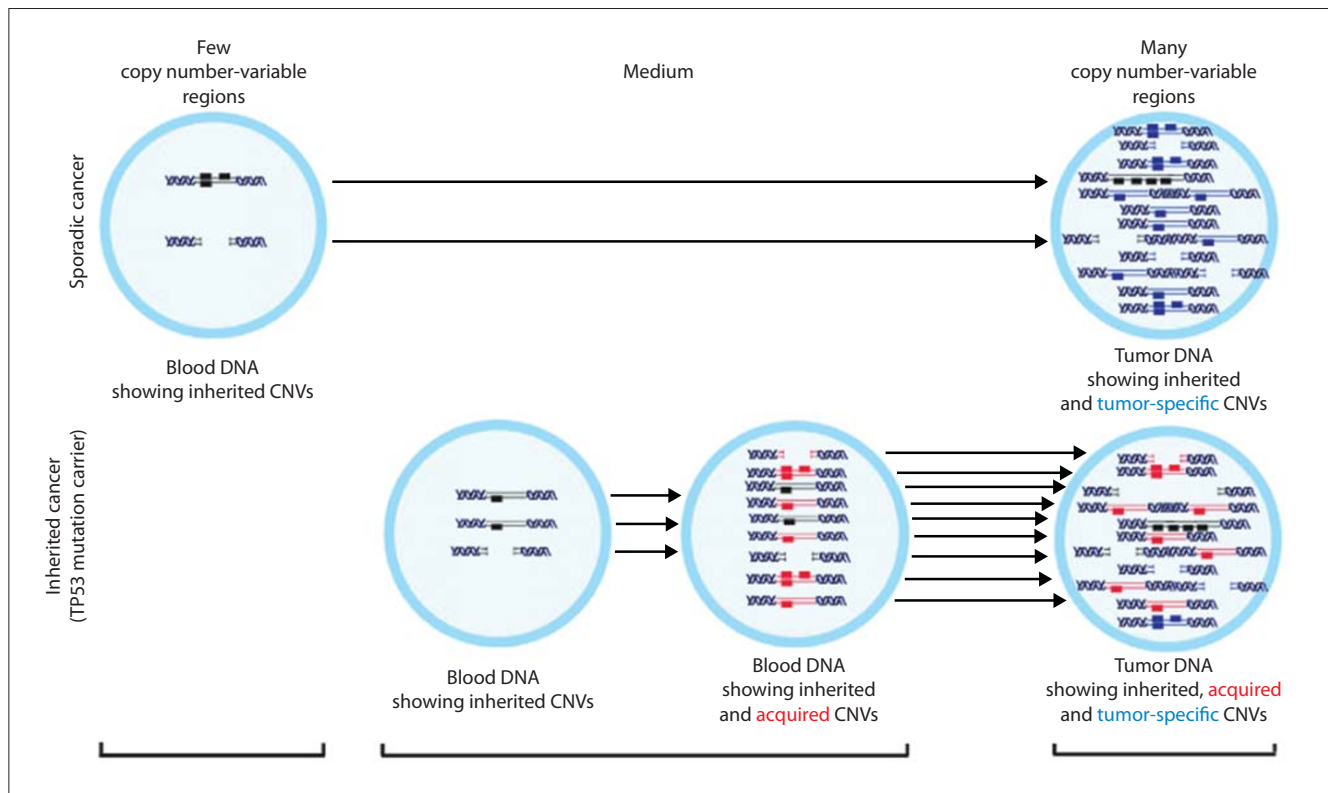
One potential model to explain the contribution of common and rare CNVs to cancer predisposition is shown in Figure 3. We propose that the number of copy-number-variable regions in healthy persons is maintained by efficient DNA repair, while CNVs are more abundant in cancer-prone

individuals because of germline defects in these processes. Although tumors are known to have increased somatic CNV and instability, our model suggests these alterations arise much earlier in cancer-predisposed individuals.

### CNVs and tumor genomes

So far we have focused here on CNVs and cancer predisposition, but similar high-resolution approaches have also driven recent studies on acquired (somatic) copy number alterations (CNAs) in tumor DNA.





**Figure 3**  
 Proposed model for CNVs in tumorigenesis. A model of copy-number-variable DNA regions in patients with sporadic (top) or inherited (bottom) cancer. We propose that healthy people maintain a similar low number of CNVs in their genomes (left; black blocks indicate inherited CNVs), whereas those at risk of developing early onset cancer have an excess of CNVs and a greater overall genomic burden of copy-number-variable DNA (middle; red blocks indicate somatically acquired CNVs). As a tumor grows, it acquires more copy-number-variable regions, including tumor-specific regions (blue). Reproduced with permission from [18], copyright (2008) National Academy of Sciences, USA.

**Copy number alterations**

Genome-scale analyses have found many formerly invisible CNAs. In an analysis of 371 lung adenocarcinoma samples using a 250,000 probe array, Weir *et al.* [33] identified seven recurrent homozygous deletions and 24 recurrent amplifications. The most significant amplification, at 14q13.3 and containing the novel oncogene *NKX2-1*, had not been found in previous studies; because of insufficient resolution and sample size, the target gene it contained had not been identified. Using an even denser array, Mullighan *et al.* [34] profiled the DNA copy number changes of 242 pediatric acute lymphoblastic leukemia (ALL) patients, including 192 with B-progenitor leukemia (B-ALL) and 50 with T-lineage leukemia (T-ALL). Global differences between the subtypes' genomes and recurrent abnormalities at specific loci were identified. An average of six CNAs were found per leukemia genome, but significant differences in the number of CNAs were found within the B-ALL group and between the B-ALL and T-ALL subtypes. Intriguingly, in 30% of B-ALL patients, the authors [34] detected deletions of *PAX5*, a transcription factor that is expressed during early stages of B-cell development. Using CNA analysis to

pinpoint critical genes can also help to plan subsequent sequencing efforts. For example, having identified deletions at *PAX5*, the authors [34] found that an additional 14 patients had point mutations in the same gene.

**Using CNAs to define the key pathways of a tumor**

In glioblastoma, CNA information, mRNA expression levels and methylation changes have been measured and nucleotide mutational analyses have been carried out [35]. Integrative analysis has shown that over 70% of tumors carry alterations in the retinoblastoma, p53 and receptor tyrosine kinase pathways. Although cancer is driven primarily by alterations of the genome, this study [35] and others have shown that CNA profiles can be combined with other high-throughput data to create insights that are 'greater than the sum of their parts'.

**Conclusions and perspectives**

The study of cancer and CNVs is in its infancy but is maturing quickly. In considering the effect of this form of genetic variation on cancer predisposition, cancer gene

expression and tumor genome profiling, there is much to learn from past studies on genomic disorders. Denser microarrays, next-generation sequencing and integrative informatics analyses are around the corner and promise to uncover new CNVs and CNAs.

There are, therefore, many exciting questions to be addressed: what role do CNVs have in cancer predisposition and how can we use this newly discovered form of genetic variation to identify those most at risk? Which cancer-related genes are affected by CNVs and, of these changes, which are both necessary and sufficient to cause neoplastic growth? Can incipient cancer cells use these constitutional deletions and duplications to induce or accelerate tumorigenesis and tumor proliferation? As these questions are resolved, the potential value of cancer CNVs as novel biomarkers of cancer susceptibility and initiation, and of cancer progression and metastases, will become apparent. Whether cancer CNVs offer insight into genes that might be targets for novel drug development remains to be determined.

### Abbreviations

CNA, copy number alteration; CNV, copy number variation; DGV, Database of Genomic Variants; GWA, genome-wide association; LFS, Li-Fraumeni syndrome; SMS, Smith-Magenis syndrome; SNP, single nucleotide polymorphism.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

AS and DM contributed equally to the development of and writing of this manuscript and to approving the final draft.

### Acknowledgements

The authors' research was supported by a grant from the Canadian Cancer Society, SickKids Foundation and Canada Foundation for Innovation. AS is supported by the Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes for Health Research.

### References

1. Feuk L, Carson AR, Scherer SW: **Structural variation in the human genome.** *Nat Rev Genet* 2006, **7**:85-97.
2. Gault J, Hopkins J, Berger R, Drebing C, Logel J, Walton C, Short M, Vianzon R, Olincy A, Ross RG, Adler LE, Freedman R, Leonard S: **Comparison of polymorphisms in the alpha7 nicotinic receptor gene and its partial duplication in schizophrenic and control subjects.** *Am J Med Genet B Neuropsychiatr Genet* 2003, **123B**:39-49.
3. Traherne JA: **Human MHC architecture and evolution: implications for disease association studies.** *Int J Immunogenet* 2008, **35**:179-192.
4. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C: **Detection of large-scale variation in the human genome.** *Nat Genet* 2004, **36**:949-951.
5. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M: **Large-scale copy number polymorphism in the human genome.** *Science* 2004, **305**:525-528.
6. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavare S, Deloukas P, Hurler ME, Dermitzakis ET: **Relative impact of nucleotide and copy number variation on gene expression phenotypes.** *Science* 2007, **315**:848-853.
7. Henrichsen CN, Vinckenbosch N, Zollner S, Chaignat E, Pradervand S, Schutz F, Ruedi M, Kaessmann H, Reymond A: **Segmental copy number variation shapes tissue transcriptomes.** *Nat Genet* 2009, **41**:424-429.
8. Guryev V, Saar K, Adamovic T, Verheul M, van Heesch SA, Cook S, Pravenec M, Aitman T, Jacob H, Shull JD, Hubner N, Cuppen E: **Distribution and functional impact of DNA copy number variation in the rat.** *Nat Genet* 2008, **40**:538-545.
9. Inoue K, Lupski JR: **Molecular mechanisms for genomic disorders.** *Annu Rev Genomics Hum Genet* 2002, **3**:199-242.
10. Prasad SE, Howley S, Murphy KC: **Candidate genes and the behavioral phenotype in 22q11.2 deletion syndrome.** *Dev Disabil Res Rev* 2008, **14**:26-34.
11. Bittel DC, Butler MG: **Prader-Willi syndrome: clinical genetics, cytogenetics and molecular biology.** *Expert Rev Mol Med* 2005, **7**:1-20.
12. Meyer-Lindenberg A, Mervis CB, Berman KF: **Neural mechanisms in Williams syndrome: a unique window to genetic influences on cognition and behaviour.** *Nat Rev Neurosci* 2006, **7**:380-393.
13. Elsea SH, Girirajan S: **Smith-Magenis syndrome.** *Eur J Hum Genet* 2008, **16**:412-421.
14. Merla G, Howald C, Henrichsen CN, Lyle R, Wyss C, Zabot MT, Antonarakis SE, Reymond A: **Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes.** *Am J Hum Genet* 2006, **79**:332-341.
15. Kleinjan DA, van Heyningen V: **Long-range control of gene expression: emerging mechanisms and disruption in disease.** *Am J Hum Genet* 2005, **76**:8-32.
16. Turner DJ, Miretti M, Rajan D, Fiegler H, Carter NP, Blayney ML, Beck S, Hurler ME: **Germline rates of de novo meiotic deletions and duplications causing several genomic disorders.** *Nat Genet* 2008, **40**:90-95.
17. Maher B: **Personal genomes: The case of the missing heritability.** *Nature* 2008, **456**:18-21.
18. Shlien A, Tabori U, Marshall CR, Pienkowska M, Feuk L, Novokmet A, Nanda S, Druker H, Scherer SW, Malkin D: **Excessive genomic DNA copy number variation in the Li-Fraumeni cancer predisposition syndrome.** *Proc Natl Acad Sci USA* 2008, **105**:11264-11269.
19. Higgins ME, Claremont M, Major JE, Sander C, Lash AE: **Cancer-Genes: a gene selection resource for cancer genome projects.** *Nucleic Acids Res* 2007, **35**:D721-726.
20. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, et al.: **Global variation in copy number in the human genome.** *Nature* 2006, **444**:444-454.
21. McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, de Bakker PI, Maller JB, Kirby A, Elliott AL, Parkin M, Hubbell E, Webster T, Mei R, Veitch J, Collins PJ, Handsaker R, Lincoln S, Nizzari M, Blume J, Jones KW, Rava R, Daly MJ, Gabriel SB, Altshuler D: **Integrated detection and population-genetic analysis of SNPs and copy number variation.** *Nat Genet* 2008, **40**:1166-1174.
22. Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, Clements J, Menzies A, Teague JW, Futreal PA, Stratton MR: **The Catalogue of Somatic Mutations in Cancer (COSMIC).** *Curr Protoc Hum Genet* 2008, **Chapter 10**:Unit 10.11.
23. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, Hankinson SE, Hutchinson A, Wang Z, Yu K, Chatterjee N, Garcia-Closas M, Gonzalez-Bosquet J, Prokunina-Olsson L, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Diver R, Prentice R, Jackson R, Kooperberg C, Chlebowski R, Lissowska J, et al.: **A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1).** *Nat Genet* 2009, **41**:579-584.
24. Perry GH, Ben-Dor A, Tsalenko A, Sampas N, Rodriguez-Revena L, Tran CW, Scheffer A, Steinfeld I, Tsang P, Yamada NA, Park HS, Kim JJ, Seo JS, Yakhini Z, Laderman S, Bruhn L, Lee C: **The fine-scale and**

- complex architecture of human copy-number variation.** *Am J Hum Genet* 2008, **82**:685-695.
25. Frank B, Bermejo JL, Hemminki K, Sutter C, Wappenschmidt B, Meindl A, Kiechle-Bahat M, Bugert P, Schmutzler RK, Bartram CR, Burwinkel B: **Copy number variant in the candidate tumor suppressor gene MTUS1 and familial breast cancer risk.** *Carcinogenesis* 2007, **28**:1442-1445.
  26. McCarroll SA, Altshuler DM: **Copy-number variation and association studies of human disease.** *Nat Genet* 2007, **39**:S37-S42.
  27. Ionita-Laza I, Rogers AJ, Lange C, Raby BA, Lee C: **Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis.** *Genomics* 2009, **93**:22-26.
  28. Nagy R, Sweet K, Eng C: **Highly penetrant hereditary cancer syndromes.** *Oncogene* 2004, **23**:6445-6470.
  29. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G: **Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification.** *Nucleic Acids Res* 2002, **30**:e57.
  30. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR: **A census of human cancer genes.** *Nat Rev Cancer* 2004, **4**:177-183.
  31. Jackson EM, Shaikh TH, Gururangan S, Jones MC, Malkin D, Nikkel SM, Zuppan CW, Wainwright LM, Zhang F, Biegel JA: **High-density single nucleotide polymorphism array analysis in patients with germline deletions of 22q11.2 and malignant rhabdoid tumor.** *Hum Genet* 2007, **122**:117-127.
  32. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **7**:57-70.
  33. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhi R, Lin WM, Province MA, Kraja A, Johnson LA, Shah K, Sato M, Thomas RK, Barletta JA, Borecki IB, Broderick S, Chang AC, Chiang DY, Chirieac LR, Cho J, Fujii Y, Gazdar AF, Giordano T, Greulich H, Hanna M, Johnson BE, Kris MG, Lash A, Lin L, Lindeman N, et al.: **Characterizing the cancer genome in lung adenocarcinoma.** *Nature* 2007, **450**:893-898.
  34. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, Girtman K, Mathew S, Ma J, Pounds SB, Su X, Pui CH, Relling MV, Evans WE, Shurtleff SA, Downing JR: **Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia.** *Nature* 2007, **446**:758-764.
  35. Cancer Genome Atlas Research Network: **Comprehensive genomic characterization defines human glioblastoma genes and core pathways.** *Nature* 2008, **455**:1061-1068.
  36. Hodgson SV, Coonar AS, Hanson PJ, Cottrell S, Scriven PN, Jones T, Hawley PR, Wilkinson ML: **Two cases of 5q deletions in patients with familial adenomatous polyposis: possible link with Caroli's disease.** *J Med Genet* 1993, **30**:369-375.
  37. Su LK, Steinbach G, Sawyer JC, Hindi M, Ward PA, Lynch PM: **Genomic rearrangements of the APC tumor-suppressor gene in familial adenomatous polyposis.** *Hum Genet* 2000, **106**:101-107.
  38. Aretz S, Stienen D, Uhlhaas S, Pagenstecher C, Mangold E, Caspari R, Propping P, Friedl W: **Large submicroscopic genomic APC deletions are a common cause of typical familial adenomatous polyposis.** *J Med Genet* 2005, **42**:185-192.
  39. Charames GS, Ramyar L, Mitri A, Berk T, Cheng H, Jung J, Bocangel P, Chodirker B, Greenberg C, Spriggs E, Bapat B: **A large novel deletion in the APC promoter region causes gene silencing and leads to classical familial adenomatous polyposis in a Manitoba Mennonite kindred.** *Hum Genet* 2008, **124**:535-541.
  40. Delnatte C, Sanlaville D, Mougnot JF, Vermeesch JR, Houdayer C, Blois MC, Genevieve D, Goulet O, Fryns JP, Jaubert F, Vekemans M, Lyonnet S, Romana S, Eng C, Stoppa-Lyonnet D: **Contiguous gene deletion within chromosome arm 10q is associated with juvenile polyposis of infancy, reflecting cooperation between the BMPRIA and PTEN tumor-suppressor genes.** *Am J Hum Genet* 2006, **78**:1066-1074.
  41. Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drusedau M, Hogervorst FB, Hageman S, Arts PJ, Ligtenberg MJ, Meijers-Heijboer H, Klijn JG, Vasen HF, Cornelisse CJ, van 't Veer LJ, Bakker E, van Ommen GJ, Devilee P: **BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients.** *Nat Genet* 1997, **17**:341-345.
  42. Montagna M, Dalla Palma M, Menin C, Agata S, De Nicolo A, Chieco-Bianchi L, D'Andrea E: **Genomic rearrangements account for more than one-third of the BRCA1 mutations in northern Italian breast/ovarian cancer families.** *Hum Mol Genet* 2003, **12**:1055-1061.
  43. Casilli F, Tournier I, Sinilnikova OM, Coulet F, Soubrier F, Houdayer C, Hardouin A, Berthet P, Sobol H, Bourdon V, Muller D, Fricker JP, Capoulade-Metay C, Chompret A, Nogues C, Mazoyer S, Chappuis P, Mailet P, Philippe C, Lortholary A, Gesta P, Bezieau S, Toulas C, Gladieff L, Maugard CM, Provencher DM, Dugast C, Delvincourt C, Nguyen TD, Faivre L, et al.: **The contribution of germline rearrangements to the spectrum of BRCA2 mutations.** *J Med Genet* 2006, **43**:e49.
  44. Lesueur F, de Lichy M, Barrois M, Durand G, Bombled J, Avril MF, Chompret A, Boitier F, Lenoir GM, Bressac-de Paillerets B, Baccard M, Bachollet B, Berthet P, Bonadona V, Bonnetblanc JM, Caron O, Chevrant-Breton J, Cuny JF, Dalle S, Delaunay M, Demange L, De Quatrebarbes J, Dore JF, Frenay M, Fricker JP, Gauthier-Villars M, Gesta P, Giraud S, Gorry P, Grange F, et al.: **The contribution of large genomic deletions at the CDKN2A locus to the burden of familial melanoma.** *Br J Cancer* 2008, **99**:364-370.
  45. Cybulski C, Wokolorczyk D, Huzarski T, Byrski T, Gronwald J, Gorski B, Debniak T, Masojc B, Jakubowska A, Gliniewicz B, Sikorski A, Stawicka M, Godlewski D, Kwias Z, Antczak A, Krajka K, Lauer W, Sosnowski M, Sikorska-Radek P, Bar K, Klijer R, Zdrojowy R, Malkiewicz B, Borkowski A, Borkowski T, Szwiec M, Narod SA, Lubinski J: **A large germline deletion in the Chek2 kinase gene is associated with an increased risk of prostate cancer.** *J Med Genet* 2006, **43**:863-866.
  46. Cybulski C, Wokolorczyk D, Huzarski T, Byrski T, Gronwald J, Gorski B, Debniak T, Masojc B, Jakubowska A, van de Wetering T, Narod SA, Lubinski J: **A deletion in CHEK2 of 5,395 bp predisposes to breast cancer in Poland.** *Breast Cancer Res Treat* 2007, **102**:119-122.
  47. Levrin O, Diotti R, Pujara K, Batish SD, Hanenberg H, Auerbach AD: **Spectrum of sequence variations in the FANCA gene: an International Fanconi Anemia Registry (IFAR) study.** *Hum Mutat* 2005, **25**:142-149.
  48. van Hattem WA, Brosens LA, de Leng WW, Morsink FH, Lens S, Carvalho R, Giardiello FM, Offerhaus GJ: **Large genomic deletions of SMAD4, BMPRIA and PTEN in juvenile polyposis.** *Gut* 2008, **57**:623-627.
  49. Kishi M, Tsukada T, Shimizu S, Futami H, Ito Y, Kanbe M, Obara T, Yamaguchi K: **A large germline deletion of the MEN1 gene in a family with multiple endocrine neoplasia type 1.** *Jpn J Cancer Res* 1998, **89**:1-5.
  50. Nystrom-Lahti M, Kristo P, Nicolaidis NC, Chang SY, Aaltonen LA, Moisio AL, Jarvinen HJ, Mecklin JP, Kinzler KW, Vogelstein B, et al.: **Founding mutations and Alu-mediated recombination in hereditary colon cancer.** *Nat Med* 1995, **1**:1203-1206.
  51. Chan TL, Yuen ST, Ho JW, Chan AS, Kwan K, Chung LP, Lam PW, Tse CW, Leung SY: **A novel germline 1.8-kb deletion of hMLH1 mimicking alternative splicing: a founder mutation in the Chinese population.** *Oncogene* 2001, **20**:2976-2981.
  52. Stella A, Surdo NC, Lastella P, Barana D, Oliani C, Tibiletti MG, Viel A, Natale C, Piepoli A, Marra G, Guanti G: **Germline novel MSH2 deletions and a founder MSH2 deletion associated with anticipation effects in HNPCC.** *Clin Genet* 2007, **71**:130-139.
  53. Plaschke J, Ruschoff J, Schackert HK: **Genomic rearrangements of hMSH6 contribute to the genetic predisposition in suspected hereditary non-polyposis colorectal cancer syndrome.** *J Med Genet* 2003, **40**:597-600.
  54. Riva P, Corrado L, Natacci F, Castorina P, Wu BL, Schneider GH, Clementi M, Tenconi R, Korf BR, Larizza L: **NFI microdeletion syndrome: refined FISH characterization of sporadic and familial deletions with locus-specific probes.** *Am J Hum Genet* 2000, **66**:100-109.
  55. Bausch B, Borozdin W, Neumann HP: **Clinical and genetic characteristics of patients with neurofibromatosis type 1 and pheochromocytoma.** *N Engl J Med* 2006, **354**:2729-2731.
  56. Tsilchorozidou T, Menko FH, Laloo F, Kidd A, De Silva R, Thomas H, Smith P, Malcolmson A, Dore J, Madan K, Brown A, Yovos JG, Tsaligopoulos M, Vogiatzis N, Baser ME, Wallace AJ, Evans DG: **Constitutional rearrangements of chromosome 22 as a cause of neurofibromatosis 2.** *J Med Genet* 2004, **41**:529-534.
  57. Horvath A, Bossis I, Giatzakis C, Levine E, Weinberg F, Meoli E, Robinson-White A, Siegel J, Soni P, Groussin L, Matyakhina L, Verma S, Remmers E, Nesterova M, Carney JA, Bertherat J, Stratakis CA: **Large deletions of the PRKARIA gene in Carney complex.** *Clin Cancer Res* 2008, **14**:388-395.
  58. Shimkets R, Gailani MR, Siu VM, Yang-Feng T, Pressman CL, Levanat S, Goldstein A, Dean M, Bale AE: **Molecular analysis of chromosome**



- 9q deletions in two Gorlin syndrome patients.** *Am J Hum Genet* 1996, **59**:417-422.
59. Bremner R, Du DC, Connolly-Wilson MJ, Bridge P, Ahmad KF, Mostachfi H, Rushlow D, Dunn JM, Gallie BL: **Deletion of RB exons 24 and 25 causes low-penetrance retinoblastoma.** *Am J Hum Genet* 1997, **61**:556-570.
60. Cascon A, Montero-Conde C, Ruiz-Llorente S, Mercadillo F, Leton R, Rodriguez-Antona C, Martinez-Delgado B, Delgado M, Diez A, Rovira A, Diaz JA, Robledo M: **Gross SDHB deletions in patients with paraganglioma detected by multiplex PCR: a possible hot spot?** *Genes Chromosomes Cancer* 2006, **45**:213-219.
61. Baysal BE, Willett-Brozick JE, Filho PA, Lawrence EC, Myers EN, Ferrell RE: **An Alu-mediated partial SDHC deletion causes familial and sporadic paraganglioma.** *J Med Genet* 2004, **41**:703-709.
62. McWhinney SR, Pilarski RT, Forrester SR, Schneider MC, Sarquis MM, Dias EP, Eng C: **Large germline deletions of mitochondrial complex II subunits SDHB and SDHD in hereditary paraganglioma.** *J Clin Endocrinol Metab* 2004, **89**:5694-5699.
63. Swensen JJ, Keyser J, Coffin CM, Biegel JA, Viskochil DH, Williams MS: **Familial occurrence of schwannomas and malignant rhabdoid tumour associated with a duplication in SMARCB1.** *J Med Genet* 2009, **46**:68-72.
64. Le Meur N, Martin C, Saugier-veber P, Joly G, Lemoine F, Moirrot H, Rossi A, Bachy B, Cabot A, Joly P, Frebourg T: **Complete germline deletion of the STK11 gene in a family with Peutz-Jeghers syndrome.** *Eur J Hum Genet* 2004, **12**:415-418.
65. Bougeard G, Brugieres L, Chompret A, Gesta P, Charbonnier F, Valent A, Martin C, Raux G, Feunteun J, Bressac-de Paillerets B, Frebourg T: **Screening for TP53 rearrangements in families with the Li-Fraumeni syndrome reveals a complete deletion of the TP53 gene.** *Oncogene* 2003, **22**:840-846.
66. Bougeard G, Sesboue R, Baert-Desurmont S, Vasseur S, Martin C, Tinat J, Brugieres L, Chompret A, de Paillerets BB, Stoppa-Lyonnet D, Bonaiti-Pellie C, Frebourg T: **Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families.** *J Med Genet* 2008, **45**:535-538.
67. Kozlowski P, Roberts P, Dabora S, Franz D, Bissler J, Northrup H, Au KS, Lazarus R, Domanska-Pakiela D, Kotulska K, Jozwiak S, Kwiatkowski DJ: **Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations.** *Hum Genet* 2007, **121**:389-400.
68. Richards FM, Phipps ME, Latif F, Yao M, Crossey PA, Foster K, Linehan WM, Affara NA, Lerman MI, Zbar B, et al.: **Mapping the Von Hippel-Lindau disease tumour suppressor gene: identification of germline deletions by pulsed field gel electrophoresis.** *Hum Mol Genet* 1993, **2**:879-882.
69. Huff V, Miwa H, Haber DA, Call KM, Housman D, Strong LC, Saunders GF: **Evidence for WT1 as a Wilms tumor (WT) gene: intragenic germinal deletion in bilateral WT.** *Am J Hum Genet* 1991, **48**:997-1003.