Germline copy number variations and cancer predisposition

Ana Cristina Victorino Krepischi1,2, Peter Lees Pearson3 & Carla Rosenberg*3

1CIPe, A.C. Camargo Hospital, São Paulo, Brazil
2National Institute of Science & Technology in Oncogenomics, São Paulo, Brazil
3Department of Genetics & Evolutionary Biology, Biosciences Institute, University of São Paulo, São Paulo, Brazil
*Author for correspondence: Tel.: +55 11 3091 7573 n carla.rosenberg@uol.com.br

We present an overview of the role of germline copy number variations (CNVs) in cancer predisposition. CNVs represent a significant source of genetic diversity, although the mechanisms by which they influence cancer susceptibility still remain largely unknown. Approximately 100 highly penetrant germline mutant genes are now known to cause cancer predisposition inherited in a Mendelian fashion; in this review, we show that nearly half of these genes have also been observed as rare CNVs associated with cancer. However, these highly penetrant alleles seem to account for less than 5% of all familial cancers. We surmise that most of the genetic risk of cancer in the general population must largely involve genes of low or moderate penetrance. In the last 5 years, studies have demonstrated that although common low penetrant CNVs are modest contributors to cancer individually, their combined impact on cancer predisposition must be taken into account in estimating cancer risk.

Since the start of human chromosome analysis over half a century ago, chromosome alterations have indicated the location of genes contributing to cancer. In one of the earliest specific cancer rearrangements detected, individuals with retinoblastoma presented cytogenetically visible deletions encompassing the band 13q14 (review in [1]), allowing the \( RB1 \) gene to be mapped to this region. In the 21st century, genomic arrays became available, permitting the detection of gains and losses of chromosomal segments much smaller than those observable under the light microscope (>4–10 Mb), and the mapping of cryptic chromosome alterations has now become efficient and precise. Genomic arrays have revealed a surprisingly large variation in the number of copies of many genomic sequences between normal individuals [2–4], to the extent that an estimated ~53% of the human genome exhibits copy number differences between ‘normal’ individuals (so called ‘DNA copy number variations’ [CNVs]). A direct relationship between CNVs and expression of the genes encompassed by such CNVs has been extensively documented both in germlines and tumors [5–7]. Several CNVs have already been associated with both complex and common disorders, including cancer (review in [8]). A map of CNVs reported in controls identified 49 cancer-related genes involving copy number changes [9], most of which were probably of low penetrance, and giving only a small contribution to the cancer penetrance. There are now approximately 100 genes known to cause Mendelian-inherited cancer syndromes when mutated [10], and a recent review indicates that >30% of them have also been observed as deleterious CNVs in cancer patients [11]. In addition to already known cancer genes, new candidates genes and genomic regions have also been identified by whole-genome CNV profiling in high-risk cancer individuals [12–15]. Since gains and losses of full genes often escape detection by conventional sequencing, techniques that measure DNA sequence copy number, such as multiplex ligation-dependent probe amplification and real-time PCR, should also be added to routine clinical diagnosis [16]. Next-generation sequencing is a powerful approach that measures the depth of coverage of reads aligned to the human reference genome. The depth of coverage corresponds to the number of copies of a given chromosome region, and we can expect that much of what has been performed to date by genomic arrays will gradually be replaced by these new sequencing technologies [17].

While only a small proportion of genes overlapped by CNVs represent highly penetrant cancer susceptibility loci, such as \( RB1 \) or \( TP53 \), the vast majority are believed to contribute to less extreme variation, including cancers determined by low-impact point mutations or deletions [10]. Such low-penetrant variants are much harder to detect, not only because each confers a very modest increase in cancer risk (often less than twice the population frequency), but also because many of the CNVs involve genes that mediate interaction with the environment, and their pathogenic effect will only become evident in the presence or absence of some environmental factors. There is also growing evidence that CNVs may not be static, as lateral genetic transfer is likely. In future, the genetic diversity in the germline and epigenetic effects are likely to become important factors in cancer predisposition and therapy.

Keywords
- cancer predisposition
- chromosomal alteration
- CNV = genomic microarray
- germline mutations
- next-generation sequencing

ISSN 1479-6694

Future Oncology (2012) 8(4), 441–450

10.2217/FON.12.34 © 2012 Ana CV Krepischi
of a particular environmental trigger. Several examples have been reported regarding the relationship between CNVs and environmental factors, such as the impact of copy number changes encompassing detoxification genes in response to drugs [18–21]. Another classical example concerns the relationship of CNV-related susceptibility to viral infection, such as HIV [22].

We present an updated review of the major cancer genes known to be included in germline CNVs (Table 1). These are rare in the population (typically much less than 1%) and highly penetrant. We also present an overview of the few studies reporting rare and common CNVs affecting genes or genomic segments not previously correlated with cancer that appear to affect cancer susceptibility. However, as is the case for most complex disorders, the data presented here are still the tip of the iceberg.

### Rare syndromes, rare CNVs

The clinical criteria for diagnosing most cancer syndromes are well-established, and there is no need to describe them here (review in [101]). However, the concept of rare variants in genetics is far from consensual; several articles refer to rare variants as those present in the population at frequencies <1%. In fact, the highly penetrant dominant variants that cause familial cancer exhibit frequencies between 1 in 1000 and 1 in 100,000, as exemplified by juvenile polyposis syndrome (1 in 100,000) and for familial adenomatous polyposis (FAP; 1 in 8300 to 1 in 14) [101]. The scientifically correct way to classify a variant, be it a CNV or any other genetic marker, is to consider those maintained in the population by recurrent mutation as rare. Although CNV mutation rates are not yet well defined, at least two studies suggest a mutation frequency varying between 1.2 × 10⁻² and 3 × 10⁻⁴ [23,24]. Furthermore, both studies conclude that the selection against new CNVs (at least in the context of studies on schizophrenia and autism) is extremely high and that such new CNV mutations will only last a few generations and eventually disappear.

The common guidelines for interpretation of the possible phenotypic impact of CNVs include comparing them with genomic imbalances recorded in healthy individuals [25]. The Database of Genomic Variants (DGV) [102] compiles 42 worldwide peer-reviewed studies on structural variations (deletions, duplications and inversions, mostly >1 Kb) in control samples. At present, nearly 16,000 CNV loci have been identified and, theoretically, those should be disregarded as potential disease variants. However, the DGV is not a curated database and is known to contain spurious data because:

- The size of the CNVs is known to be overestimated [26,27];
- The same populations, in particular involving HapMap individuals, have been analysed in several studies;
- There is a high false discovery rate.

These problems all lead to overestimating both the size and frequency of CNVs [28]. In particular, regarding cancer variants, there are additional issues to take into account, namely that cancer often manifests later in life and penetrance is not always complete. Therefore, it is reasonable to assume that cancer predisposition CNVs may be present in a normal population and will be reported in the DGV.

Because of the drawbacks in the use of the DGV as a ‘normal’ comparison, another criterion for evaluating causality of a chromosome variant in a medical setting is determining if it is de novo, and although noncausative de novo CNVs have been sporadically reported [29], the correspondence between relatively rare events such as a de novo CNV and de novo cancer may be regarded as providing reasonable grounds for suspecting an association between the two. This criterion should therefore be used with discretion and, wherever possible, other parameters should be used to support a role of the particular CNV in a pathogenic phenotype.

There are now approximately 100 highly penetrant cancer-predisposing genes known to cause inherited cancer syndromes, characterized by evident strong familial history, frequently affecting multiple generations and usually obeying a Mendelian pattern of inheritance. The transmission of the majority of such cancers fits an autosomal dominant model, generally by segregation of a heterozygous point mutation or deletion of a tumor suppressor gene; tumorogenesis occurs when the remaining normal allele is somatically mutated (Knudson two-hit hypothesis – a dominantly inherited predisposition to cancer entails transmission of a germline mutation to the next generation, while tumorogenesis requires a second, somatic, mutation of the same locus). The relevance of the two-hit hypothesis to CNV involvement in tumorogenesis is that germline CNVs then represent the first hit and increase cancer predisposition. The second event is somatic and occurs in the tissues giving rise to the tumor, and can either be a CNV or a point mutation inactivating the second allele.
Table 1. Known cancer predisposition genes reported as rare constitutive copy number variations.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Locus</th>
<th>Syndrome</th>
<th>Cancer type</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>5q22.2</td>
<td>Adenomatous polyposis coli, Turcot syndrome</td>
<td>Colorectal, pancreatic, desmoid, hepatoblastoma, glioma, other CNS cancers</td>
<td>[47,16]</td>
</tr>
<tr>
<td>BMPR1A</td>
<td>10q22.3</td>
<td>Juvenile polyposis</td>
<td>Gastrointestinal polyps</td>
<td>[48]</td>
</tr>
<tr>
<td>BRCA1</td>
<td>17q21</td>
<td>Hereditary breast/ovarian cancer</td>
<td>Breast, ovarian</td>
<td>[49,50]</td>
</tr>
<tr>
<td>BRCA2</td>
<td>13q12.3</td>
<td>Hereditary breast/ovarian cancer</td>
<td>Breast, ovarian, pancreatic, leukemia (FANCb, FANCd1)</td>
<td>[51]</td>
</tr>
<tr>
<td>CDH1</td>
<td>16q22.1</td>
<td>Familial gastric carcinoma</td>
<td>Gastric, breast</td>
<td>[52]</td>
</tr>
<tr>
<td>CDKN1B</td>
<td>12p13.1</td>
<td>Multiple endocrine neoplasia type IV</td>
<td>Pituitary tumor, testicular tumor</td>
<td>[33]</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>9p21</td>
<td>Familial malignant melanoma</td>
<td>Melanoma, pancreatic</td>
<td>[53]</td>
</tr>
<tr>
<td>CHEK2</td>
<td>22q12.1</td>
<td>Familial breast cancer</td>
<td>Breast, prostate</td>
<td>[54,55]</td>
</tr>
<tr>
<td>CREBBP</td>
<td>16p13.3</td>
<td>Rubinstein–Taybi syndrome</td>
<td>Nervous system, brain, leukemia</td>
<td>[56]</td>
</tr>
<tr>
<td>CYLD</td>
<td>16q12.1</td>
<td>Brooke–Spiegler syndrome, familial cylindromatosis, multiple familial trichoepithelioma</td>
<td>Multiple skin appendage tumors</td>
<td>[57]</td>
</tr>
<tr>
<td>EPCAM</td>
<td>2p21</td>
<td>Lynch syndrome</td>
<td>Colorectal, endometrial</td>
<td>[58]</td>
</tr>
<tr>
<td>EXT1</td>
<td>8q24.11</td>
<td>Multiple exostoses type 1</td>
<td>Exostoses, osteosarcoma</td>
<td>[59]</td>
</tr>
<tr>
<td>EXT2</td>
<td>11p11.2</td>
<td>Multiple exostoses type 2</td>
<td>Exostoses, osteosarcoma</td>
<td>[59]</td>
</tr>
<tr>
<td>FANCA</td>
<td>16q24.3</td>
<td>Fanconi anemia A</td>
<td>Acute myeloid leukemia</td>
<td>[60]</td>
</tr>
<tr>
<td>FH</td>
<td>1q42.1</td>
<td>Hereditary leiomyomatosis and renal cell cancer</td>
<td>Leiomyomatosis, renal</td>
<td>[61]</td>
</tr>
<tr>
<td>FLCN</td>
<td>17p11.2</td>
<td>Birt–Hogg–Dubé syndrome</td>
<td>Renal cell carcinoma</td>
<td>[33]</td>
</tr>
<tr>
<td>GPC3</td>
<td>Xq26</td>
<td>Simpson–Golabi–Behmel syndrome</td>
<td>Wilms‘ tumors</td>
<td>[33]</td>
</tr>
<tr>
<td>HRPT2</td>
<td>1q31.2</td>
<td>Hyperparathyroidism–jaw tumor syndrome</td>
<td>Parathyroid carcinoma, renal cell carcinoma</td>
<td>[62]</td>
</tr>
<tr>
<td>JAG1</td>
<td>20p12</td>
<td>Alagille syndrome</td>
<td>Hepatocellular carcinoma, papillary thyroid carcinoma</td>
<td>[63,64]</td>
</tr>
<tr>
<td>MADH4</td>
<td>18q21.1</td>
<td>Juvenile polyposis</td>
<td>Gastrointestinal polyps</td>
<td>[65]</td>
</tr>
<tr>
<td>MEN1</td>
<td>11q13</td>
<td>Multiple endocrine neoplasia type 1</td>
<td>Parathyroid adenoma, pituitary adenoma, pancreatic islet cell, carcinoid</td>
<td>[66]</td>
</tr>
<tr>
<td>MSH2</td>
<td>2p21</td>
<td>Lynch syndrome</td>
<td>Colorectal, endometrial, ovarian</td>
<td>[47]</td>
</tr>
<tr>
<td>MSH6</td>
<td>2p16</td>
<td>Lynch syndrome</td>
<td>Colorectal, endometrial, ovarian</td>
<td>[67]</td>
</tr>
<tr>
<td>NF1</td>
<td>17q11.2</td>
<td>Neurofibromatosis type 1</td>
<td>Neurofibroma, glioma</td>
<td>[68]</td>
</tr>
<tr>
<td>NF2</td>
<td>22q12.2</td>
<td>Neurofibromatosis type 2</td>
<td>Meningioma, acoustic neuroma</td>
<td>[69]</td>
</tr>
<tr>
<td>NSD1</td>
<td>5q35.3</td>
<td>Sotos syndrome</td>
<td>Increased risk of benign or malignant tumors, including neuroblastoma and gastric carcinoma</td>
<td>[33]</td>
</tr>
<tr>
<td>PMS2</td>
<td>7p22</td>
<td>Lynch syndrome; Turcot syndrome</td>
<td>Colorectal, endometrial, ovarian, medulloblastoma, glioma</td>
<td>[70]</td>
</tr>
<tr>
<td>PRKAR1A</td>
<td>17q24.2</td>
<td>Carney complex</td>
<td>Myxoma, endocrine, papillary thyroid</td>
<td>[71]</td>
</tr>
<tr>
<td>PTC1</td>
<td>9q22.3</td>
<td>Gorlin syndrome</td>
<td>Skin basal cell, medulloblastoma</td>
<td>[72]</td>
</tr>
<tr>
<td>PTEN</td>
<td>10q23.31</td>
<td>Cowden disease; Lhermitte–Duclos syndrome</td>
<td>Breast cancer, leukemia, renal cell adenocarcinoma, neuroendocrine carcinoma, Merkel cell carcinoma</td>
<td>[73]</td>
</tr>
<tr>
<td>RB1</td>
<td>13q14.2</td>
<td>Familial retinoblastoma</td>
<td>Retinoblastoma, sarcoma, breast, small cell lung</td>
<td>[1]</td>
</tr>
<tr>
<td>RUNX1</td>
<td>21q22.12</td>
<td>Familial platelet disorder</td>
<td>Acute myeloid leukemia</td>
<td>[74]</td>
</tr>
<tr>
<td>SDHB</td>
<td>1p36.13</td>
<td>Familial paraganglioma</td>
<td>Paraganglioma, pheochromocytoma</td>
<td>[75]</td>
</tr>
<tr>
<td>SDHC</td>
<td>1q21</td>
<td>Familial paraganglioma</td>
<td>Paraganglioma, pheochromocytoma</td>
<td>[76]</td>
</tr>
</tbody>
</table>

Adapted with permission from [11].
Much less frequently, affected families present recessive patterns of inheritance, and their cancer is usually initiated by defects in genes acting in DNA repair or genome stability, including Bloom syndrome, ataxia-telangiectasia and xeroderma pigmentosum. By contrast to the dominant syndromes, in which loss of function of the second allele only occurs in the tumor, patients affected by recessive cancer syndromes carry biallelic mutations in the germline and often have other clinical signs in addition to cancer. We are unaware of germline CNVs covering cancer genes with a recessive pattern of inheritance, but CNVs have already been implicated in other recessive genetic disorders, involving homozygous deletions encompassing a known gene implicated in deafness [30], and deletions unmasking pathogenic point mutations in the homologous chromosome [31,32].

**Known cancer genes and rare CNVs**

Although point mutation is the most common mechanism reported in germlines causing cancer predisposition, a review from 2010 claims that approximately 30% of known cancer genes have already been described as being fully deleted [11]. However, the detection of deletions and duplications is steadily increasing with wider application of genomic microarrays, and Table 1 shows that deletions have now been observed in >40% of familial cancer genes. Table 1 compiles those cancer genes already described as involved in pathogenic CNVs, although a few of these CNVs have only been observed in patients referred for investigation because of mental deficiency or congenital abnormalities [33], rather than for cancer. While in many cases deletions and point mutations in these genes cause similar phenotypes, it may not always be the case; for example, for over two decades cases of cytogenetically visible deletions, including the APC gene, have been reported. In most of these patients, regardless of ascertainment by cancer or mental impairment, very large numbers of adenomatous polyps of the colon and rectum were present at the time of cytogenetic diagnosis, which eventually progressed to colorectal carcinoma, a phenotype referred to as familial adenomatous polyposis (FAP) [34]. Although the phenotype caused by different APC point mutations can be variable, a distinction between point mutations and deletions of the entire APC has been suggested [35]; the heterozygous deletion of the full APC gene is associated with an attenuated form of FAP, characterized by the occurrence of less than 100 colonic adenomas (rather than hundreds to thousands in classical FAP) and a later onset of colorectal cancer (age >40 years) [36]. Another example is seen in mutations in the TTP53 gene causing Li–Fraumeni syndrome, a usually highly penetrant disorder that predisposes to a variety of early-onset tumors. Because none of the ten reported mentally impaired patients carrying germline 17p13 deletions fully encompassing TP53 [37] have so far manifested cancer, it has been suggested that full deletions may not cause cancer predisposition, in contrast to point mutations. However, most of these patients are still fairly young and may yet develop cancer. It has been postulated that the defective proteins resulting from point mutation, both for the APC and TP53 genes, may interfere with the protein product of the normal allele and cause a more severe disease [35,37]. In both cases, a dominant-negative effect or gain-of-function would be apparent, leading to a more severe phenotype for point mutations than for CNV deletions.

### Table 1. Known cancer predisposition genes reported as rare constitutive copy number variations (cont.)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Locus</th>
<th>Syndrome</th>
<th>Cancer type</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDHD</td>
<td>11q23</td>
<td>Familial paraganglioma</td>
<td>Paraganglioma, pheochromocytoma</td>
<td>[76]</td>
</tr>
<tr>
<td>SMAD4</td>
<td>18q21.2</td>
<td>Juvenile polyposis syndrome</td>
<td>Colon, stomach, small bowel and pancreas</td>
<td>[65]</td>
</tr>
<tr>
<td>SMARCB1</td>
<td>22q11</td>
<td>Rhabdoid tumor predisposition syndrome-1</td>
<td>Schwannomas, malignant rhabdoid</td>
<td>[77]</td>
</tr>
<tr>
<td>STK11</td>
<td>19p13.3</td>
<td>Peutz–Jeghers syndrome</td>
<td>Jejunal hamartoma, ovarian, testicular, pancreatic</td>
<td>[78]</td>
</tr>
<tr>
<td>TP53</td>
<td>17p13.1</td>
<td>Li–Fraumeni syndrome</td>
<td>Breast, sarcoma, adrenocortical carcinoma, glioma, multiple other tumor types</td>
<td>[79]</td>
</tr>
<tr>
<td>TSC1</td>
<td>9q34</td>
<td>Tuberous sclerosis 1</td>
<td>Hamartoma, renal cell</td>
<td>[80]</td>
</tr>
<tr>
<td>TSC2</td>
<td>16p13.3</td>
<td>Tuberous sclerosis 2</td>
<td>Hamartoma, renal cell</td>
<td>[80]</td>
</tr>
<tr>
<td>VHL</td>
<td>3p25.3</td>
<td>von Hippel–Lindau syndrome</td>
<td>Renal, hemangioma, pheochromocytoma</td>
<td>[81]</td>
</tr>
<tr>
<td>WT1</td>
<td>11p13</td>
<td>Denys–Drash syndrome, Frasier syndrome, Familial Wilms’ tumor</td>
<td>Wilms’ tumor</td>
<td>[82]</td>
</tr>
</tbody>
</table>

Adapted with permission from [11].
New candidate genes for cancer susceptibility: studies of cancer predisposition cohorts

If we consider the current state of mapping and identifying genes for Mendelian genetic diseases in general, of the estimated 7024 human diseases present in the McKusick catalogue OMIM [103], 47% still have no gene mutation attributed to them. Clearly, mutations for genes that segregate dominantly, such as \textit{RBI} and \textit{TP53}, are easier to identify than recessive mutations that do not cause an abnormal phenotype in the heterozygous state. In the majority of patients with histories suggestive of hereditary cancer in which the genetic factors for cancer predisposition remain elusive, rare CNVs can indicate a chromosome region related to their cancer susceptibility.

In the last few years, few whole-genome studies profiling germline CNVs have been conducted in cohorts of cancer predisposition patients negative for mutations in the major genes related to their specific cancer. Table 2 shows (in addition to studies in sporadic cases, see next section) the few whole-genome germline CNV screens published so far on high-risk cancer patients or families. Although all screens identified sporadic individuals carrying rare CNVs, none exhibited consistent recurrence of given CNVs, nor have the genes encompassed by them been reported as mutated in the germline of cancer patients. The strongest indication that these rare CNVs are related to cancer predisposition is that many of the affected genes have been previously implicated in the tumors themselves. Because of the limited recurrence data obtained so far, it is impossible to state explicitly which of the variants contribute to cancer predisposition. Although not proof, it is reasonable to assume that a variant required for tumor development being carried in the germline is an event that is likely to increase the probability of tumor occurrence and decrease the time necessary for tumor manifestation. An interesting example is provided by the work of Yang et al., which reports a rare 4q13 duplication detected in a melanoma-prone family; although the duplication was observed in only a single family, it segregates with the melanoma in the three affected patients [15]. Furthermore, the duplicated region contains ten genes, including \textit{CXCL1} and \textit{IL-8}, that have been shown to stimulate melanoma growth \textit{in vitro} and \textit{in vivo}. Another CNV, a 9p21.3 germline microdeletion encompassing the \textit{KIA1797} and \textit{MIR491} genes, was recently reported in two independent cancer studies, concentrating on colorectal [14] and breast cancer [12], respectively. This finding of a similar rare CNV in independent cohorts of cancer patients supports a pathogenic role in cancer predisposition.

Table 2. Literature review of whole-genome studies associating germline copy number variations with cancer susceptibility.

<table>
<thead>
<tr>
<th>Type of cancer cohort</th>
<th>Number of unrelated patients</th>
<th>Relevant CNVs (n)</th>
<th>Germline CNV details</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familial pancreatic cancer</td>
<td>57</td>
<td>56</td>
<td>Nonrecurrent rare CNVs (not present in the DGV and 607 study controls)</td>
<td>[13]</td>
</tr>
<tr>
<td>Familial and early-onset colorectal cancer</td>
<td>41</td>
<td>7</td>
<td>Nonrecurrent rare CNVs (not present in the DGV and 1600 controls from in-house database)</td>
<td>[14]</td>
</tr>
<tr>
<td>Familial and early-onset breast cancer</td>
<td>68</td>
<td>26</td>
<td>Nonrecurrent rare CNVs (not present in the DGV, 100 study controls and 158 from in-house database)</td>
<td>[12]</td>
</tr>
<tr>
<td>Familial melanoma</td>
<td>30</td>
<td>1</td>
<td>Nonrecurrent rare 4q13 duplication (three affected individuals from the same family; CXC genes)</td>
<td>[15]</td>
</tr>
<tr>
<td>Aggressive prostate cancer</td>
<td>498</td>
<td>1</td>
<td>Recurrent common 2p24.3 deletion (no known gene; 12.63% in patients in comparison to 8.28% in 494 study controls)</td>
<td>[83]</td>
</tr>
<tr>
<td>Neuroendocrine tumors of the ileum</td>
<td>226</td>
<td>4</td>
<td>Recurrent 18q22.1 deletion (no known genes; 6.19% in patients in comparison to 2.06% in 97 study controls)</td>
<td>[84]</td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>846</td>
<td>1</td>
<td>Recurrent common CNV at 1q21.1 (gene \textit{NBPF23}; odds ratio 2.49 [803 study controls])</td>
<td>[44]</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>386</td>
<td>6</td>
<td>Recurrent common CNV at 1p36.33 (no known genes; odds ratio 17.0 [687 study controls])</td>
<td>[45]</td>
</tr>
<tr>
<td>Nasopharyngeal carcinoma</td>
<td>278</td>
<td>8</td>
<td>Recurrent common 6p21.33 deletion (MICA and HCP5 genes) Gender-specific association (male)</td>
<td>[46]</td>
</tr>
</tbody>
</table>

\*Studies of familial cancer or high-risk cancer patients. CNV: Copy number variation; DGV: Database of Genomic Variants.
A CNV can also act as a modifier factor for cancer risk. One of the first examples was reported by Frank et al. [38]. Interestingly, they detected a protective effect in breast cancer progression associated with deletion of the exon 4 of the MTUS1 gene (odds ratio: 0.41; p = 0.003). It is important to note that for genes with lower penetrance, much larger cohorts need to be investigated for their effect to be detectable. In this case, this modest effect was detected in a case–control study involving 732 controls and 593 breast cancer patients. Further studies are needed to gather enough data to determine the precise role of rare CNVs in familial and populational high-risk cancer individuals.

A different approach consists of investigating CNV profiles in cancer predisposition cohorts of individuals with known mutations in a cancer predisposition gene. The first study investigated Li–Fraumeni families [39], and reported a significantly higher number of CNVs among TP53 mutation carriers compared with controls; several of the detected CNVs affected cancer-related genes that the authors suggested could act as risk modifiers. More recently, a similar study performed on a cohort of BRCA1-associated ovarian cancer individuals compared germline CNVs between patients and controls [40]; the number of CNV deletions per genome was found to be significantly higher in the BRCA1 group, including 31 previously unreported CNV regions. In cases with a previously detected mutation in a well-known cancer gene, it is difficult to determine if the increased number of CNVs is a consequence of chromosome instability/deficient apoptosis, or whether these CNVs are directly involved in oncogenesis along with the recognized mutation. These two hypotheses are not mutually exclusive; for example such patients may exhibit an increased number of CNVs as a result of the initial cancer mutation, but part of these CNVs may also contribute to cancer, and their accumulation may result in more severe or earlier-onset cancer manifestation.

Studies of CNVs related to cancer predisposition in sporadic cancer patients
Cancer syndromes are infrequent and only explain a minor part of the familial clustering of the common cancers. Despite the fact that a syndrome is regarded as a group of clinically recognizable features, several of the so-called cancer syndromes do not include distinctive clinical signs other than cancer. The fact that we have already identified so many highly penetrant genes in cancer families is entirely due to the fact that Mendelian inheritance is easy to detect through segregation. Many instances occur where familial transmission is not detected as such, because only single affected individuals are present and cannot be distinguished from sporadic cases; this partly explains why a large proportion of familial cancers go unrecognized despite the fact that genetic predisposing factors determine approximately 30% of total cancer risk in the population. Increased familial risk of cancer in the general population must largely involve combinations of genes of low or moderate penetrance, and the highly penetrant genes must play a relatively infrequent role in familial predisposition.

Until recently, efforts to identify cancer predisposition genes with low penetrance have proven singularly unrewarding. However, in the past few years, developments in this area have sped up and several large cohort studies have led to the identification of recurrent common CNVs associated with cancer manifestation. One such CNV was a variant at 20p13 associated with aggressiveness of prostate cancer and detected in a case–case study comparing 448 aggressive and 500 nonaggressive prostate cancer cases, a finding replicated in an additional 2895 aggressive and 3094 nonaggressive cases [41]. CNVs covering detoxifying genes such as GSTT1 and GSTM1 have been extensively studied regarding their contribution to disease susceptibility and drug response [20,42]. However, most studies were inconclusive on the contribution of CNVs involving detoxifying genes to cancer susceptibility, with the exceptions of prostate and bladder cancer [43], and leukemia [20]. Other common CNVs have also been associated with neuroblastoma [44], hepatocellular carcinoma [45] and nasopharyngeal carcinoma (Table 2) [46].

Conclusion
CNV analyses have disclosed copy number changes in bona fide cancer genes and also have indicated new candidate genes for cancer predisposition. Because most of the rare CNVs identified so far are nonrecurrent, it is impossible to determine at this juncture which of these variants contribute, possibly with high penetrance, to cancer predisposition. Nevertheless, these rare CNVs point to regions that require further investigation and putative new cancer determining genes. Taken together, data emerging from CNV analyses on hereditary cancer families or high-risk cohorts of patients disclose a background of genetic heterogeneity, in which an extremely large number of different rare CNVs individually confer a penetrant cancer predisposition. It must be acknowledged that this is no surprise, since it is now estimated...
that CNVs globally make an approximately 20% contribution across the whole spectrum of all genetic diseases, of which familial cancer is just one component. However, the human genome, with its approximately 21,000 genes, is much smaller than originally believed and the necessity to catalog the many CNVs contributing to familial cancer predisposition, as argued in this review, may not be as formidable a task as it is at present.

**Future perspective**

Whole-genome next-generation sequencing is revolutionizing the mapping of classical Mendelian genes and now permits detection of specific mutations on extremely small sample sizes, which, only 1 year ago, would have been considered ‘mission impossible’. We foresee that next-generation sequencing will soon replace Sanger sequencing and genomic arrays to identify both sequence and copy number variants in the human genome. The challenge will be dealing with such large amounts of information and correlating variants to specific phenotypes, particularly those involved in cancer susceptibility. Given that CNVs make a major contribution to the burden of all genetic diseases, we can expect that many new CNVs related to cancer susceptibility will be disclosed in the near future. Eventually, we expect that personal genome sequencing will become a tool of medical practice and differential diagnosis, and lead to familial approaches in treating cancer.

**Financial & competing interests disclosure**

This work was supported by grants from the Brazilian National Institute of Science and Technology in Oncogenomics (FAPESP 2008/57887-9; CNPq 573589/08-9) and FAPESP (2009/00898-1). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

**Executive summary**

**Background**

- Only a small proportion of genes overlapped by copy number variations (CNVs) represent highly penetrant cancer susceptibility loci.
- The vast majority of CNVs are supposed to have low to moderate penetrance and contribute modestly to human diseases.

**Rare syndromes, rare CNVs**

- The higher the penetrance of a given pathogenic variant, the lower its frequency in the population.
- CNVs never or hardly reported in databases of normal populations, such as the Database of Genomic Variants, are considered rare and potentially hazardous.

**Known cancer genes & rare CNVs**

- Nearly half the approximately 100 highly penetrant cancer predisposition genes are already documented as CNVs related to cancer.
- Full deletions of cancer predisposition genes do not necessarily result in the same phenotype as point mutations.

**New candidate genes for cancer susceptibility: studies of cancer predisposition cohorts**

- Complementary to the finding of rare CNVs affecting *bona fide* cancer genes, in the last few years germline CNV whole-genome profiling of cancer families and high-risk patients has revealed new candidate genes for cancer susceptibility.

**Studies of CNVs related to cancer predisposition in sporadic cancer patients**

- Some studies investigating large cohorts have shown that common CNVs may play a role in sporadic cancer. Such CNVs are generally associated with very modest increases in cancer risk.

**Conclusion**

- CNV analyses have disclosed new candidate genes for cancer risk. More studies are required to determine which of those rare and common variants have (minor or major) impact in cancer predisposition.

**References**

Papers of special note have been highlighted as:
- of interest
- of considerable interest

5. Mileyko Y, Joh RI, Weitz JS. Small-scale copy number variation and large-scale changes in...
Disclosure relating penetrance of a disease

- Discussion relating penetrance of a disease with its frequency in the population.
- Extensive review of copy number variations (CNVs) in cancer predisposition, mainly related to known cancer genes already reported as copy number variable in cancer patients.

- Discusses the frequency and potential impact of imbalances of selected cancer predisposition genes in patients referred for array-comparative genomic hybridization for developmental delay, behavioral abnormalities and birth defects.
- Jin G, Sun J, Liu W et al. Genome-wide copy-number variation analysis identifies...


**Outstanding genome-wide association study of CNVs in the childhood cancer neuroblastoma. They identified a novel member of the NBPF genes, which were found to exhibit expression level strictly correlated with CNV state in neuroblastoma cells.**


**Websites**

