Copy number variations—is there a biological difference between submicroscopic and microscopically visible ones?

T Liehr*

Abstract

Introduction
Copy number variations are nowadays most often understood as submicroscopic gains or losses of chromosomal material, either connected with a disease or just one of the many possible genetic variants in man. However, besides such submicroscopic copy number variations, already decades ago chromosome analysis revealed the existence of cytogenetic visible copy number variations; first discovered were the chromosomal heteromorphisms, more recently euchromatic variants and unbalanced chromosome abnormalities without phenotypic consequences. Recently, for submicroscopic copy number variants a so-called two-hit model suggested. This suggests that the combination of per se harmless size variants may be led to clinical aberrations if they are present together in a patient.

A literature based comparison of cytogenetic visible copy number variations and MG-CNVs revealed that, biologically, besides size, these two groups of copy number variations do not really differ. Both may be present as heterochromatic (gene-poor to gene-free) and euchromatic copy number variation-variants. Identical regions have even been reported as cytogenetic visible copy number variation and a submicroscopic copy number variations independently. This study reviews the differences between submicroscopic and microscopically visible CNVs.

Conclusion
Yet cytogenetic visible copy number variations, especially chromosomal heteromorphisms are by far not considered enough in evaluation of routine cytogenetic analysis as well as in array-based comparative genomic hybridisation analysis.

Introduction
Since 1956 and the correct assessment of human chromosome numbers, cytogenetics plays a crucial role in pre- and postnatal, as well as tumour cytogenetics and research. Currently, the GTG-banding approach (G-bands by Trypsin using Giemsa) is still the gold-standard for all cytogenetic techniques and the preferred technique of molecular cytogenetics is fluorescence in situ hybridization (FISH). One additional major proceeding in molecular (cyto) genetics is nowadays the array-based variant of comparative genomic hybridisation (aCGH). All these mentioned approaches provide information on the human genome at different levels of resolutions (Figure 1). Public and patients expect from such analysis clear results in that sense that the tested person has (a) a normal or (b) an abnormal genetic condition. As in many cases nature is much more complex than layman—or even specialists—may expect. Besides the sheer impossibility per se to define what is ‘normal’, there is the problem of the so-called ‘copy number variant’ (CNV) regions in the human genome; CNVs can be larger or smaller in size in each clinically healthy individual tested, and any ‘norm’ can only be an artificial definition of a standard size.

In 2004 two key publications reported that no two clinically healthy individuals are alike and yet it is known that each person distinguishes from another by up to 0.5 megabasepair (Mb) of submicroscopic CNVs; the latter are abbreviated from now on as MG-CNVs, for ‘CNVs detectable by molecular genetics’. The description of MG-CNVs was celebrated in the years after 2004 as some kind of big surprise. However; it was an already early finding of cytogenetics that on a chromosomal level there are lots of cytogenetically visible copy number variations (CG-CNVs), and that there are no individuals which are really the same on a chromosomal level. The aim of this study was to review if there is a biological difference between submicroscopic and microscopically visible CNVs.

Discussion
The author has referenced some of his own studies in this review. The protocols of these studies have been approved by the relevant ethics committees related to the institution in which they were performed.

What are CG-CNVs and what are MG-CNVs ones?
Any kind of CNV may be found in otherwise normal as well as abnormal karyotypes/genomes. In general, CNVs are considered as having no (major) clinical impact on the...
Critical review

**Figure 1:** Current approaches for analysis of the human genome at different resolutions. (A) Banding cytogenetics has an approximate resolution of 10 Mb; arrowheads highlight heteromorphisms (CG-CNVs) on the short arms of the acrocentric chromosomes 13 and 22. (B) Molecular cytogenetics provides resolutions between ~10 Mb down to ~1 Mb; in interphase resolution may be higher than in metaphase due to differences in DNA packing. 1. A locus-specific probe for the Kallman syndrome (KAL1) specific region in Xp22.3 (red) and a centromeric probe (CEP X, green) are applied on a metaphase and an interphase nucleus of a healthy female. 2. The FISH-banding approach multicolour banding (MCB) was used to characterise the breakpoints (arrowheads) of an inversion in one chromosome 5 to 5q15 and 5q31. (C) Here the most important segment of an aCGH result using the platform and software of Agilent is depicted. A clinical case with different clinical abnormalities was studied. The resolution of aCGH depends on the amount of the probes spotted on the so-called ‘DNA chip’; even though the results are given down to basepair level by the software, the resolution in diagnostic cases is considered to be between ~10 and ~200 kb. Here detected imbalances are shown in three different resolutions. The lowest resolution (‘all chromosomes’) gives an overview on all imbalances along the entire genome: one can see ~10 small and one large imbalance. While the smaller ones are considered as MG-CNVs, the large imbalance on chromosome 13 of ~37,577 Mb in size is the disease-causing one. Two higher magnifications of the affected regions are shown (‘chromosome 13’ and ‘affected region’). The final result would read as arr[GRCh37/hg19] 13q14.11q31.1(44,819,113x2,44,844,282-82,421,008x1,82,474,913x2) and the MG-CNV are not detailed in the report. Figure kindly provided by Dr. Kristin Mrasek, Human Genetics, Jena, Germany. (D) Next generation sequencing (NGS) is at present used for targeted analysis, for example, to detect single basepair mutations in a tumour suppressor gene. NGS provides the highest possible resolution for human genome. Here a result for a heterozygous, disease causing mutation with in the BRCA1 gene was found and is highlighted by arrowheads. NB: NGS is not suited for the analysis of higher order repeat units, the latter being subject to CNVs. Figure kindly provided by M.Sc. Martin Voigt, Human Genetics, Jena, Germany.
phenotype of its carrier and thus, may be present in clinical healthy and diseased persons.

CG-CNVs are here defined as gain or loss of genetic material, which leads to a different appearance of a chromosome visualised in banding cytogenetics; MG-CNVs are such, which are detectable in aCGH together with a normal karyotype and phenotype. For more comprehensive characterisation of a CG-CNV and/or an MG-CNV molecular cytogenetics may be applied.

In the literature, also clinically relevant changes in copy numbers, especially regions involved in microdeletion/microduplication syndromes\(^1\) are sometimes reported as MG-CNVs. In this review, the major focus is on clinically primarily inert MG-CNVs. It has to be admitted, that some potentially harmless MG-CNVs have already been linked in population studies to an enhanced or diminish susceptibility for special disorders or (age-related) diseases. However, mechanisms, therefore, are still not understood yet\(^1\).

**A closer look on CG-CNVs**

CG-CNVs comprise enlargement or reduction in the size of chromosomal subbands and/or appearance of additional band(s). Without knowledge about regions prone to develop CG-CNVs, and sometimes also if a cytogeneticist analyses GTG-banded chromosomes, such variants may easily be mixed up with clinically relevant gain or loss of chromosomal material\(^6\).

However, specific chromosomal regions in the human karyotype are prone to variations more than others, like the centromeric satellite DNA of any of the 24 human chromosomes, the band Yq11.2 or the short arms of all acrocentrics. All these regions (Figure 2) are considered as heterochromatin, which are extremely gene-poor. It is important to notice that cytogenetically visible heterochromatic blocks may vary in size (in their original chromosomal location), but also may be inserted at any place in the human genome\(^15\). Even mitochondrial DNA may be inserted and amplified in human chromosomes (Dr. M. Groß, Heidelberg, personal communication).

Overall, the majority of differences between individuals expressed as heterochromatic CG-CNVs are found in 1q12, 9q12, 13pter-q11, 14pter-q11.1, 15pter-q11.1, 16q11.2, 19p12-q12, 21pter-q11.1, 22pter-q11.1 and in male in Yq12; the cytogenetic approaches from the mid-1970s identified an average of four to five chromosomal heteromorphisms per person\(^6\).

Besides heterochromatic CG-CNVs, there are also euchromatic ones, which are more recent findings. They were reported as so-called euchromatic variants (i.e. large scale CG-CNVs) (Figure 2) and unbalanced chromosome abnormalities without phenotypic consequences\(^6,16\). These CG-CNVs can be due to gain or loss of copy numbers at the original location; moreover, such regions can be translocated to another chromosomal region due to rearrangements. It is suggested that any region not containing dose-dependent genes may be amplified or lost in the genome without clinical consequences, as also seen in euchromatic small supernumerary marker chromosomes\(^6,17\).

**A closer look on MG-CNVs**

A MG-CNV is a submicroscopic amplification or loss of DNA stretches. One of the first identified MG-CNV is a 28 kilobasepair (kb) deletion in 15q11.2\(^18\). MG-CNVs range from 1 kb to several Mb in size and are distributed along the entire human genome. As detected by aCGH, which for technical reasons only can identify...
euchromatic regions of the human genome, no submicroscopic heterochromatic CNVs were reported yet. However, there are hints from exceptional studies that such also exist, but as aCGH is (almost) blind for these regions yet, they cannot be analysed systematically at present. MG-CNVs which consist of low repetitive elements are summarised in Figure 3\textsuperscript{19}. Nonetheless, clinically relevance was found already for some of them\textsuperscript{19,20}.

MG-CNVs are determined as variable copy numbers when compared to a reference genome and may include deletions and duplications of genomic loci\textsuperscript{21}. They may encompass as much as 12\% of the human genome\textsuperscript{22}. Most of them are considered as benign MG-CNVs and are usually inherited from a parent\textsuperscript{23}. When determined as de novo, genomic imbalances are considered more likely pathological\textsuperscript{24}. It is also known that each human being carries about thousand MG-CNVs ranging from only a few hundred base pairs to over 1 Mb\textsuperscript{25}. The major determinant for the clinical impact of a CNV seems to be if dosage sensitive genes are present in the corresponding DNA stretch\textsuperscript{17,26}. However, 'it was recently observed that more than one (submicroscopic) CNV (larger than 500 kb) can contribute to severe developmental delay and often is responsible for phenotypic variability associated with genomic disorders'; this phenomenon is called 'two-hit' model\textsuperscript{9}.

CG-CNVs and MG-CNVs—a comparison

CG-CNVs and MG-CNVs have in common 8 of 10 of the features listed in Table 1. The theories how they may be formed are identical—unequal crossing-over events or breakage–fusion–bridge model are considered. Both may be either inherited or de novo formed. Both are size variants of DNA stretches and may be visualised by molecular cytogenetics. Both may be euchromatic, but also heterochromatic, as outlined before. They only are different in two points: (i) MG-CNVs may be associated with disease, CG-CNVs may not, and (ii) CG-CNVs are microscopically visible, MG-CNVs are submicroscopic. Both points are matters of definition: for (i) CG-CNVs considered only as such if they are not associated with clinical signs and symptoms; for (ii) MG-CNVs are submicroscopic just due to limitations of optical microscopy\textsuperscript{6}.

Thus, overall there is no biological reason to distinguish MG-CNVs

Figure 3: MG-CNVs which consist of low repetitive elements are highlighted according to Warburton et al.\textsuperscript{19}.

<table>
<thead>
<tr>
<th>Features</th>
<th>Type of CNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formation by unequal crossing-over events</td>
<td>+</td>
</tr>
<tr>
<td>Formation by breakage–fusion–bridge model</td>
<td>+</td>
</tr>
<tr>
<td>May be inherited</td>
<td>+</td>
</tr>
<tr>
<td>May be de novo</td>
<td>+</td>
</tr>
<tr>
<td>Are size variants of DNA stretches</td>
<td>+</td>
</tr>
<tr>
<td>May be visualised by molecular cytogenetics</td>
<td>+</td>
</tr>
<tr>
<td>May be euchromatic</td>
<td>+</td>
</tr>
<tr>
<td>May be heterochromatic</td>
<td>+</td>
</tr>
<tr>
<td>May be associated with disease</td>
<td>(−)*</td>
</tr>
<tr>
<td>Are microscopically visible</td>
<td>+</td>
</tr>
</tbody>
</table>

*Due to definition of CG-CNVs.
and CG-CNVs, i.e., besides size there is no real difference between them. Nothing could underline this fact better than identical CNV regions reported ‘independently’ by cytogenetic approaches on the one side and by molecular techniques on the other side; examples are CNVs in 8q21.227 (Figure 4) or 15q11.228.

Every human being provides besides a unique DNA-primary sequence also an individual combination of CG-CNV and MG-CNV. Monozygote twins are considered as exception from that rule, however, distinct MG-CNVs were already found there.28 In genetic analyses, irrespective which of the presently available approaches are applied, the geneticist has always to be aware of the fact that one never obtains the whole picture of a genome (Figure 1). (Molecular) cytogenetics provides a genome wide view with low resolution; however, also heterochromatic CG-CNVs are visible there. Molecular genetic analysis provide a high-resolution of the analysed genome, but low- and massive-copy repeats are neither ascertainable by aCGH nor by next generation sequencing approaches29 (Figure 1).

**Conclusion**

The future has to show if we start to consider CG-CNV and MG-CNV as the same biological phenomenon or keep the artificial borders of thinking here. If these obstacles are driven down the way may also be free to check the phenomenon ‘two-hit’ model for patients with multiple CG-CNVs or with CG-CNVs and MG-CNVs and syndrome or disease. Also, then it might be necessary to list all detected MG-CNVs in aCGH reports and all CG-CNVs, especially heterochromatin variants, in cytogenetic reports.

**References**


![Figure 4: FISH result in a case with a CG-CNV in the region 8q21.2; chromosome 8 with arrowhead shows an enlargement of the corresponding sub-band in GTG-banding and the BAC-probe RP11-96G1 gives an ~3 fold stronger signal on chromosome 8 with arrowhead than on the homologous one. In database of genomic variants, the corresponding region at position 86,898,422-86,955,420 (NCBI36/hg18) has a collection of 60 normal cases with gain and 150 cases with loss of copy numbers according to aCGH. Obviously this region can be expressed as MG-CNV or CG-CNV. A centromeric probe for chromosomes 8 (cep 8) in red was also applied in this FISH-experiment as a positive control.](image-url)
Critical review