Chromosome microarrays in human reproduction

Evica Rajcan-Separovic*

Department of Pathology and Laboratory Medicine (Cytogenetics), University of British Columbia, Children’s and Women’s Health Centre of BC and Child and Family Research Institute, 950 West 28th Avenue, Vancouver, BC, Canada V5Z 4H4

*Correspondence address. Tel: +604-875-3121; Fax: +604-875-3601; E-mail: eseparovic@cw.bc.ca

Submitted on February 1, 2012; resubmitted on April 13, 2012; accepted on April 30, 2012

TABLE OF CONTENTS

- Introduction
- Methods
- Types of arrays and chromosome abnormalities detected by CMA
- CMA in PGD and PGS
- CMA in prenatal diagnosis
- CMA and miscarriage
- CMA and reproductive disorders
- Conclusion and future directions

BACKGROUND: Chromosome microarray (CMA) testing allows automatic and easy identification of large chromosomal abnormalities detectable by conventional cytogenetics as well as the detection of submicroscopic chromosomal imbalances.

METHODS: A PubMed search was performed in order to review the current use of CMA testing in the field of human reproduction. Articles discussing the use of CMA in the preimplantation setting, ongoing pregnancies, miscarriages and patients with reproductive disorders were considered.

RESULTS: A high rate of concordance between conventional methods of detecting chromosomal abnormalities [e.g. fluorescence in situ hybridization (FISH), karyotyping] and CMA was reported in the prenatal setting with CMA providing more comprehensive and detailed results as it investigates the whole genome at higher resolution. In preimplantation genetic screening, CMA is replacing FISH and the selection of embryos based on CMA has already resulted in live births. For ongoing pregnancies and miscarriages, CMA eliminates tissue culture failures and artifacts and allows a quick turnaround time. The detection of submicroscopic imbalances [or copy number variants (CNVs)] is beneficial when the imbalance has a clear clinical consequence but is challenging for previously undescribed imbalances, particularly for ongoing pregnancies. Recurrent CNVs have been documented in patients with reproductive disorders; however, the application of CMA in this field is still limited.

CONCLUSIONS: CMA enhances reproductive medicine as it facilitates better understanding of the genetic aspects of human development and reproduction and more informed patient management. Further clinical validation of CMA in the prenatal setting, creation of practice guidelines and catalogs of newly discovered submicroscopic imbalances with clinical outcomes are areas that will require attention in the future.

Key words: array comparative genomic hybridization / single nucleotide polymorphism array / PGD / PGS / miscarriage / prenatal diagnosis
Introduction

Numerical and structural chromosomal abnormalities are a common cause of human disease including reproductive failure. They are traditionally identified by karyotyping, which has low resolution and limited ability to detect gains and losses of chromosomal material (<4–10 Mb in size (typically the size of one chromosomal band). The development of chromosome microarrays (CMAs), which represent arrays of small DNA segments sampled at various densities across the whole genome (whole-genome arrays) or part of it (targeted arrays) and spotted on a glass slide revolutionized chromosome analysis as it allowed the detection of chromosome abnormalities at a higher resolution and in an automated, faster and more accurate way (Albertson and Pinkel, 2003).

The development of CMAs was spurred by the reports that this technique can detect causative submicroscopic chromosomal imbalances in ~10–15% of patients with developmental delay, multiple congenital abnormalities or autism, thus increasing the diagnostic yield over karyotype analysis (Vissers et al., 2003; de Vries et al., 2005; Shinawi and Cheung, 2008; Schaaf et al., 2011). Submicroscopic chromosomal imbalances (gains and losses) are referred to as copy number variants (CNVs) and those detected in affected individuals are cataloged in databases, such as Decipher (http://decipher.sanger.ac.uk/), Ecaruca (http://www.ecaruca.net/) and ISCA (https://www.iscasonsortium.org/), while CNVs detected in apparently healthy controls are reported in the Database of Genomic Variants (DGV, http://projects.tcg.cavariation/, lafrate et al., 2004). The overwhelming presence of CNVs in controls, encompassing in total ~18% of the genome (Pinto et al., 2007), requires the careful interpretation of CNVs detected in affected individuals. Guidelines addressing clinical and technical aspects of array application, including interpretation, in the post-natal setting are now available (Shaffer et al., 2007; Kearney et al., 2011). Based on extensive literature review and experience (Miller et al., 2010), it was recently proposed that arrays should replace karyotyping for patients with developmental delay, autism or congenital abnormalities.

In the field of human reproduction, the first application of CMAs was for the detection of chromosomal abnormalities in miscarriages (Schaef er et al., 2004) and fetuses with morphological abnormalities (Le Caignec et al., 2005). CMA testing in the prenatal setting has since increased dramatically, and a recommendation was made recently that CMA analysis replaces fluorescence in situ hybridization (FISH) in preimplantation genetic screening (PGS) as it provides a more comprehensive view of the genome (Harper et al., 2010). For ongoing pregnancies, the use of CMA is recommended as an adjunct tool to karyotyping in morphologically abnormal pregnancies with normal karyotype as well as in cases of fetal demise with congenital anomalies when karyotype cannot be obtained (ACOG, 2009). Targeted CMA analysis is recommended in these cases, as it screens chromosomal regions of known clinical relevance and minimizes uncovering of CNVs of uncertain relevance. Finally, the CMA analysis provides new opportunities to find genetic causes of miscarriage and reproductive disorders.

This review is organized to provide a brief background on types of arrays and illustrate their application, using specific examples, to screen for chromosome abnormalities in preimplantation embryos, ongoing pregnancies and miscarriages, and patients with reproductive disorders.

Methods

In order to obtain information on the use of CMAs in human reproduction, PubMed was searched using relevant terms and their combinations, e.g. PGD, PGS, array CGH, SNP array, microarray, reproduction, embryo, prenatal diagnosis, miscarriage.

Types of arrays and chromosome abnormalities detected by CMA

Array CGH

This technology was initially developed using the principle of chromosomal comparative genome hybridization (CGH). In array CGH, differentially labeled control or reference DNA (typically labeled red) and patient or test DNA (typically labeled green) are co-hybridized to small segments of human DNA arrayed on a glass slide (for review see Shaffer and Bejjani, 2004; Shinawi and Cheung, 2008). The distinction between gain, loss or a balanced status is based on the green-to-red fluorescence ratio for each DNA segment arrayed on the slide (Fig. 1A). Using bioinformatics tools, the green-to-red fluorescence ratio for each DNA segment is mapped to the chromosome, resulting in an array profile.

Sampling larger human DNA segments (100–200 kb in size) incorporated into bacterial artificial chromosome (BAC) clones results in BAC arrays, while smaller DNA segments (~60 nucleotides) constitute oligonucleotide (oligo) arrays. The DNA segments can be sampled throughout the genome and the frequency of their sampling is one of the parameters that determines the resolution of the array (i.e. the more densely they cover the region or the whole genome, the higher is the resolution of the array, see Fig. 2). BAC arrays typically have a lower resolution of 1 Mb (i.e. gains and losses of 1 Mb or ~1/10th of a chromosome band can be identified), while oligonucleotide arrays are available at different resolutions. For example, commercially available oligonucleotide arrays with 105 000 or 244 000 oligonucleotides provide an overall median probe spacing (resolution) of ~0.02 Mb and ~0.01 Mb, respectively.

Single nucleotide polymorphism (SNP) arrays

SNP arrays consist of small oligonucleotides (~20–60 bp DNA) containing a base known to show sequence variability (polymorphism) in humans (LaFramboise, 2009; Nowak et al., 2009; Schaaf et al., 2011). Most SNPs are biallelic, and occur in either of the two forms (usually labeled A and B). Oligonucleotides representing the two SNP variants (alleles) are spotted on the array. In one form of SNP array analysis, test DNA, digested, amplified and labeled with a fluorescent dye, is applied to the array and hybridized to the oligonucleotides based on sequence homology. Copy number changes are established by comparing the intensity of the fluorescence for each oligonucleotide variant on the array and the signal intensities for the same oligonucleotide in a large number of reference controls. Since the sequence at each SNP in the test sample can be established, SNP arrays also allow genotyping, i.e. determining the proportion of each allelic variant combination (AA, AB and BB) at the polymorphic position for each of the oligonucleotides (Fig. 1B).
Types of chromosome abnormalities detected by CMA

Large scale chromosomal changes

Both array CGH and SNP arrays allow easy identification of most types of large-scale numerical and segmental chromosomal gains and losses. In addition, as detailed in sections below, both arrays identify low-level mosaicism (10–30%) and eliminate cell culture artifacts as the analysis is DNA based.

Array CGH cannot detect balanced chromosome rearrangements and ploidy changes; however, improvement in ploidy detection is possible when reference DNA with 47 XXY chromosome complement is used (Ballif et al., 2006). Unique application of array CGH was demonstrated by Gabriel et al. (2011) who studied the chromosome composition of polar bodies in comparison with the control DNA and determined that gains or losses of chromatids, rather than whole chromosomes, were more prevalent. This finding confirmed the previous observations (Angell, 1997, Fragouli et al., 2011) of the presence of extra or missing chromatids in oocytes resulting from metaphase I (MI) errors.

In addition to detecting structural and numerical chromosome imbalances, SNP arrays can provide genotype information which facilitates the detection of ploidy changes, uniparental disomy (UPD, both copies of the chromosome derive from one parent) and parent and stage (meiotic versus mitotic) origin of aneuploidy; for the latter application, the possibility of misclassification cannot be excluded as meiotic errors without recombination can be called a mitotic error (Bisignano et al., 2011, 2012; Handyside, 2011). SNP arrays have been shown to simultaneously detect aneuploidy and transmission of a chromosome segment carrying a single gene mutation (Handyside et al., 2010). The possibility that SNP genotyping may be used to distinguish cells containing a balanced chromosome translocation from cells with a normal karyotype has been indicated (Handyside et al., 2010) but not yet tested in practice. This distinction was possible when the parental balanced translocation resulted in a detectable microdeletion at the translocation breakpoint (Treff et al., 2011).

Submicroscopic chromosomal imbalances

Detection of cytogenetically invisible chromosomal gains and losses (or CNVs) is the frequent reason for array testing using both platforms when chromosomes are normal. Interpreting small CNVs is one of the challenges of CMAs and algorithms for CNV classification and interpretation have been published for both post-natal (Koolen et al., 2009; Miller et al., 2010; Kearney et al., 2011) and prenatal cases (D’Amours et al., 2011; Fiorentino et al., 2011a). Generally speaking, CNVs are classified as likely benign, likely pathogenic (disease causing) and variants of unknown significance (VOUS). CNVs that overlap with known syndromic areas, that are larger in size, de novo, contain biologically relevant genes and are deletions are typically considered pathogenic (Miller et al., 2010; Kearney et al., 2011). DGV, which catalogs CNVs detected in healthy controls, aids interpretation, as CNVs detected in controls and reported in more than one independent control cohort, are considered benign (Koolen et al., 2009; Qiao et al., 2010). When referring to DGV, caution should be exercised...
Preimplantation genetic screening (PGS) refers to the selection of euploid embryos in patients undergoing IVF because of impaired fertility or increased risk for aneuploidy (Harper and Harton, 2010; Harper and Sengupta, 2011).

The detection of large-scale chromosomal abnormalities (numerical and structural) is the most common reason for PGD and PGS (Simpson, 2010). Until recently FISH, which uses DNA probes specific to a chromosomal region, was the main test for determining the chromosomal composition in early human development (i.e. in oocyte/zygote, single blastomeres of a Day-3 embryo or multiple cells from a 5 to 6 day blastocyst; Harper and Harton, 2010; Simpson, 2010). Screening for numerical abnormalities requires a set of chromosome-specific FISH probes (5–12 probes), while the presence of a known structural chromosome abnormality inherited from one of the parents requires FISH probes from the specific region of chromosome rearrangement. Extensive FISH analysis of blastomeres from preimplantation embryos obtained from couples at risk of aneuploidy/infertility showed that ~50% of embryos had aneuploidy and >50% of them were mosaic (Wells and Levy, 2003). The mosaic nature of the embryo and the limited number of chromosomes assayed by FISH are some of the possible reasons for the lack of improvement in the birth rate in randomized controlled trials (RCTs) of IVF with and without FISH screening for aneuploidy (Harper et al., 2010).

First reports of CMA testing of preimplantation embryos confirmed and extended the FISH observations of the high rate of chromosome abnormality. Array CGH and SNP array analysis of multiple blastomeres from 23 embryos detected chromosomal imbalances in 90% of embryos, with the majority showing mosaicism for aneuploidy (83%; Vanneste et al., 2009a). In ~50% of the mosaic embryos, the presence of normal cells was evident. In addition to chromosome number changes, this study detected UPD and segmental deletions and duplications in 9 and 70% of cases, respectively, underscoring the increased power of array technology in detecting a much broader scope of chromosomal abnormalities in comparison with FISH (Vanneste et al., 2009a,b). Most of the abnormalities in the embryos were mitotic in nature, and some aneuploides involved chromosomes not typically present in FISH screening panels. In contrast to blastocysts, the frequency of chromosomally abnormal oocytes was much lower (12.5%). The couples studied by Vanneste et al. (2009b) had normal fertility, suggesting that genomic instability is not necessarily associated with impaired fertility. Whether the mosaicism noted at preimplantation stages of human development is a normal biological event, or is influenced by in vitro conditions and manipulation, remains uncertain (Munne et al., 1997).

To overcome the problems associated with mosaicism and limitations of FISH, the European Society of Human Reproduction and Embryology PGD Consortium Steering Committee has recently recommended that in future RCTs stages of development other than blastomeres (i.e. first and second polar bodies or later stage developing embryos, such as blastocysts) are examined and that array-based technologies are used for the assessment of the whole genome to determine if this will result in an improved delivery rate after PGS (Harper et al., 2010).

Efforts to validate and implement whole-genome arrays for PGS into clinical practice are ongoing. Investigations usually involve a comparison of the chromosome abnormality detection rate for the chosen array technology (CGH based or SNP based) with more traditional

**Figure 2** CMAs improve banding resolution. The middle band of chromosome 15 (15q21 representing cc 10 Mb) is seen as a single dark band at low resolution G banding (~400 bands) (A). It resolves to two dark and one middle light band at ~550 band level resolution (B). In (C) the resolution is further improved by sampling 10 equally distributed DNA segments from 15q21, increasing the resolution 10 times (1 Mb instead of initially 10 Mb). In (D) an even higher number of DNA segments are sampled from 15q21, further increasing the resolution and allowing smaller changes in copy number to be detected (e.g. if 500 DNA segments are sampled from the same 10 Mb region, the resolution is 0.02 Mb). When the sampling of DNA segments is extrapolated to the whole genome, whole-genome arrays of 1 Mb or 0.02 Mb resolution are obtained in this example.

because many CNVs have not been confirmed using independent methods (thus the requirement that they are reported in more than one cohort), and the reproductive history of the controls is largely unknown or not reported. This is of importance when CNVs that are associated with reproduction-related pathologies (e.g. miscarriage, infertility) are identified. Unique, previously not reported, CNVs transmitted to the patient from one of the phenotypically normal parents are, in general, considered benign. However, there are many reasons why these CNVs can also represent VOUS (Kearney et al., 2011) as it is rarely possible to completely exclude that their integral genes show variable expressivity between parent and the patient, differ in allelic expression as a result of imprinting (i.e. the disorder manifests only when inherited from a particular sex), or have a mutation in the remaining copy of the gene in the affected individual.

**CMA in PGD and PGS**

**Overview of PGD and PGS**

Preimplantation genetic diagnosis (PGD) refers to diagnosis of genetic disorders in gametes or early stages of human embryo development.

**Figure 2** CMAs improve banding resolution. The middle band of chromosome 15 (15q21 representing cc 10 Mb) is seen as a single dark band at low resolution G banding (~400 bands) (A). It resolves to two dark and one middle light band at ~550 band level resolution (B). In (C) the resolution is further improved by sampling 10 equally distributed DNA segments from 15q21, increasing the resolution 10 times (1 Mb instead of initially 10 Mb). In (D) an even higher number of DNA segments are sampled from 15q21, further increasing the resolution and allowing smaller changes in copy number to be detected (e.g. if 500 DNA segments are sampled from the same 10 Mb region, the resolution is 0.02 Mb). When the sampling of DNA segments is extrapolated to the whole genome, whole-genome arrays of 1 Mb or 0.02 Mb resolution are obtained in this example.
methods of detecting chromosomal abnormalities such as karyotype, FISH, metaphase chromosome CGH or PCR, in single cells from cell lines with known abnormality or embryos from couples undergoing IVF. Amplification of DNA from single cells is a requirement for CMA testing in PGS and its efficiency is particularly important for SNP arrays in order to minimize biased allele amplification. The application of CMAs for detecting a variety of chromosomal defects in PGS and examples of studies validating array CGH and SNP array testing in cells from different stages of embryo development as well as their clinical outcomes are provided below.

Array CGH
A high concordance rate between FISH and array CGH applied to blastomeres was reported by Gutierrez-Mateo et al. (2011) with only 1.9% of embryos misdiagnosed by array CGH in comparison with a 12-probe FISH panel. On the other hand, 13% of embryos identified as abnormal by array CGH were called normal by FISH, confirming that array CGH has an increased power in detecting abnormalities. When array CGH and FISH were used to study multiple cells from blastocysts, a high concordance rate (95%) in identifying embryos that are chromosomally abnormal was reported (Fragnoli et al., 2010). A similar concordance rate (94%) was detected when the chromosomal constitution in both polar bodies and the zygote was compared (Geraedts et al., 2011). The main reason for discrepancy between FISH and array CGH in studies using blastomeres and blastocysts was mosaicism and technical artifacts, while the majority of discordant findings between the polar bodies and zygote were related to aneuploidy in the polar body and euploidy in the zygote. This was postulated to be caused by compensation of aneuploidy by the sperm or from anaphase lagging (Geraedts et al., 2011).

There was a 100% concordance between array CGH and PCR in the detection of unbalanced chromosomal rearrangements in blastomeres of embryos from carriers of balanced reciprocal and Robertsonian translocations (Fiorentino et al., 2011b), with array CGH also detecting aneuploidy for a chromosome not involved in a translocation in 66.8% embryos. The loss of a small 2.5 Mb terminal chromosomal segment caused by a translocation was the smallest unbalanced fragment detected in preimplantation embryos using whole-genome BAC array with an effective 0.5 Mb resolution at the whole genome and 0.25 Mb in the telomeric and pericentromeric regions (Fiorentino et al., 2011b). Segmental chromosome imbalances, other than the ones suspected based on parental chromosome rearrangement, have rarely been reported in PGD studies using array CGH (Traversa et al., 2011), although they were very common in studies by Vanneste et al. (2009a). Some of the possible explanations of the discrepancy include the use of different array platforms and algorithms focusing on detection of larger imbalances or the fact that the Vanneste et al. (2009a) study used both SNP and BAC arrays, which increased confidence in the detection of segmental chromosome imbalances.

One of the frequently cited drawbacks of array CGH is the limited confidence in the detection of segmental chromosome imbalances. The comparison of SNP array analysis (based on the quantification of signal intensity) and FISH showed a higher rate of interpretable results (96% versus 83%) and the reduction in intercell variability (31 versus 100%) for SNP arrays in randomized blastomeres (Treff et al., 2010a). The latter finding suggested that FISH technology may overestimate the contribution of mitotic errors to the origin of aneuploidy. Using a similar SNP approach, the accuracy of detecting numeric chromosome errors was determined to be 99.2% in single cells from cell lines with a known karyotype and concurrence between multiple blastomeres from 16 embryos was noted (Treff et al., 2010b). The minimum deletion/duplication size that can be identified with SNP arrays has not been explored, although a 2.4 Mb microdeletion for Alagille syndrome (occurring at the breakpoint of a parental balanced translocation) was detected (Treff et al., 2011). FISH, SNP array and PCR were in 100% concordance for detecting this microdeletion and, in addition, SNP arrays detected aneuploidy in ∏40% embryos.

SNP arrays
In PGS, the SNP analysis is performed by quantification of signal intensity for alleles from amplified sample DNA in comparison with normal controls (Treff et al., 2010a,b, 2011) or by genotype and linkage analysis which includes the embryo, parents (Johnson et al., 2010) and, in some cases, additional family members (Handyside et al., 2010; Handyside, 2011).

The comparison of SNP array analysis (based on the quantification of signal intensity) and FISH showed a higher rate of interpretable results (96% versus 83%) and the reduction in intercell variability (31 versus 100%) for SNP arrays in randomized blastomeres (Treff et al., 2010a). UPD was very rare (0.16% per chromosome) and always associated with aneuploidy (Rabinowitz et al., 2012). Most of the monosomies were found to be mitotic, while trisomies were predominantly of maternal meiotic origin. The presence of a maternal meiotic trisomy was found to be positively predictive of a fully aneuploid embryo (which would not be selected for transfer; Johnson et al., 2010). Upd was very rare (0.16% per chromosome) and always associated with aneuploidy (Rabinowitz et al., 2012). Using this parent support approach, haploidy and triploidy were detected (Johnson et al., 2010; Rabinowitz et al., 2012); however, their false negative or detection rate is unknown. Segmental imbalances were observed in ∏15% of blastomeres and 38.5% of embryos (Johnson et al., 2010; Rabinowitz et al., 2012).

Simultaneous detection of aneuploidy and a single gene disorder (cystic fibrosis-CF) was achieved by Handyside et al. (2010) based on genotyping of the embryo, parents and a different embryo from the same couple. The genotypes were used to develop a linkage-based karyomap (Handyside et al., 2010), which allowed the determination of the CF carrier status in DNA extracted from 3 to 10 cells from five embryos. In addition to simultaneously detecting single-gene disorders and aneuploidy (and its origin), the benefit of karyomapping is the possibility to detect balanced rearrangements.
It is conceivable that both methodologies could help resolve the debate regarding the benefit of both methodologies. Side-by-side comparison of the types of abnormalities detectable by the two arrays (examples in Table I) and diagnostic yield for different array testing modalities in a larger number of embryos could help resolve the debate regarding the benefits and pitfalls of the array CGH versus SNP-based methods in detecting chromosomal abnormalities in PGS (Bisignano et al., 2011, 2012, Handyside, 2011). It is conceivable that both methodologies will continue to exist and be used in practice in parallel, as is the case for CMA application to prenatal and post-natal constitutional disorders.

### CMA in prenatal testing

Large-scale chromosome abnormalities (aneuploidy and structural rearrangements) are detected in 2–7% of pregnancies undergoing amniocentesis (or in 9–35% if ultrasound abnormalities are present; D’Amours et al., 2011) and are the main reason for cytogenetic analysis of cultured cells from amniotic fluid or chorionic villus sampling (CVS) or pregnancy tissue. The benefits of CMA analysis of pregnancies include the elimination of tissue culture and tissue culture artifacts and the reduction of maternal contamination (Fiorentino et al., 2011a; Srebniak et al., 2011), as the analysis is performed on DNA extracted directly from the amniotic fluid and CVS cells. This allows a much faster turn-around time (~3 days (Fiorentino et al., 2011a; Srebniak et al., 2011)) which is critical in the management of ongoing pregnancies. Although PCR and interphase FISH-based testing for a selection of the most common aneuploidies or suspected specific chromosomal imbalances also provide the necessary quick turn-around time (Lewin et al., 2000; Schmidt et al., 2000; Boormans et al., 2010), they are limited to a selection of chromosomal regions. The CMA-based detection of mosaicism at a low level (~10%; Fiorentino et al., 2011a), elucidation of the origin of unidentifiable chromosomal rearrangements (marker chromosomes) as well as gains or losses at the breakpoints of apparently balanced de novo chromosomal rearrangement and UPD when SNP array is used (Faas et al., 2010) all increase the spectrum of identifiable prenatal chromosomal abnormalities (Van den Veyver et al., 2009) and facilitate more informed genetic counseling.

Detection of submicroscopic imbalances in karyotypically normal pregnancies further adds to the power of CMA testing, but also introduces challenges in interpretation. The distinction of CNVs as benign and pathogenic in an ongoing pregnancy is of importance as uncertain findings can cause anxiety in couples. In general, targeted arrays, which test for a selection of chromosomal regions with known clinical consequences, detect a lower percentage of CNVs but with more predictable clinical outcome than whole-genome arrays, which sample a wider selection of probes genome wide and can detect both imbalances with an established clinical consequence as well as those that are of unknown relevance to the pregnancy (ACOG, 2009; Faas et al., 2010). However, in one comparative study, whole-genome arrays did not increase the detection of CNVs with uncertain clinical relevance in comparison with targeted arrays (Coppinger et al., 2009).

The prevalence of pathogenic chromosomal changes detected using CMAs in karyotypically normal pregnancies is very broad (~1–16%; Faas et al., 2010; Valduga et al., 2010; D’Amours et al., 2011; Fiorentino et al., 2011a; Hillman et al., 2011; Srebniak et al., 2011) and depends on the type of array used (targeted or whole genome), resolution of the whole-genome array, reason for referral and CNV size cut-off (if used for CNV selection). In a systematic review and meta-analysis which considered 798 prenatal cases from 10 publications (out of 135 publications with the topic of array CGH use in pregnancy), Hillman et al. (2011) established the presence of CMA-detected chromosomal imbalances of pathogenic or

<table>
<thead>
<tr>
<th>Detection of</th>
<th>Array CGH</th>
<th>SNP array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aneuploidy</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Unbalanced structural abnormalities (large)</td>
<td>Yes, the detection of expected segment ~2 Mb in size reported</td>
<td>Yes, the detection of expected and unexpected segments ~2 Mb in size</td>
</tr>
<tr>
<td>Unbalanced structural abnormalities (small)</td>
<td>No, may be improved if 47,XXY used as reference</td>
<td>Possible, but the detection rate is unknown</td>
</tr>
<tr>
<td>Polyploidy</td>
<td>No</td>
<td>Yes (parental analysis required)</td>
</tr>
<tr>
<td>Balanced rearrangements (translocations, inversions)</td>
<td>No</td>
<td>No, but theoretical possibility exists (karyomapping)</td>
</tr>
<tr>
<td>Origin of aneuploidy: paternal versus maternal; meiotic (MI/ML) versus mitotic</td>
<td>No</td>
<td>Yes (parental analysis required)</td>
</tr>
<tr>
<td>Chromatid versus chromosome aneuploidy</td>
<td>Yes</td>
<td>Not tested</td>
</tr>
<tr>
<td>Single gene defect (based on haplotype)</td>
<td>No</td>
<td>Yes (analysis of additional family members required)</td>
</tr>
<tr>
<td>Mosaicism level (i.e. the prevalence of abnormal cell line if DNA is extracted from multiple cells)</td>
<td>Unknown, but chromosome abnormality not detected when present in ~30% cells</td>
<td>Unknown</td>
</tr>
<tr>
<td>Artifacts caused by amplification of DNA from single cell</td>
<td>Artifacts exist but small</td>
<td>Allele drop-out leads to reduction in usable information</td>
</tr>
</tbody>
</table>

| UPD | No | Yes |

| MI, metaphase I; MII, metaphase II. |
potentially pathogenic relevance in $3.6\%$ (95% confidence interval (CI), 1.5–8.5) of pregnancies with a normal karyotype, regardless of the clinical indication [advanced maternal age (AMA), anxiety, ultrasound abnormality, family history] and array type (targeted and whole-genome array CGH of different resolution); while this frequency was 5.2% (95% CI, 1.9–13.9) when the referral was solely for morphologic abnormality of the fetus. Two studies of morphologically abnormal fetuses with high-resolution whole-genome SNP arrays and a CNV size cut-off of 0.015 Mb for deletions and 0.02 Mb for duplications (Faas et al., 2010; Srebniak et al., 2011) confirmed a high incidence of pathogenic chromosomal imbalances (16 and 6.5%, respectively), undetected by karyotyping. This frequency was smaller (1%) in a recently studied cohort of pregnancies (referred predominantly for AMA and anxiety) using a whole-genome BAC array (Fiorentino et al., 2011a).

The frequency of CNVs of uncertain relevance also varies between the studies; their presence was reported in 0.4–2.7% of chromosomally abnormal fetuses from 12 to 6%, respectively (Tyreman et al., 2011). Reduced the percentage of CNVs of unknown relevance in morphologic samples for determining if the CNV is de novo (Zuffardi et al., 2011). The availability of parental tests for determining if the CNV is de novo or familial in origin was not performed in any of the above studies. Chromosomal abnormalities are typically explained ultrasound abnormalities. All the couples agreed not to be informed about findings of uncertain clinical relevance. A wide-spread use of CMAs in prenatal diagnosis is inevitable in the future and will yield new information to help standardize clinical application and determine its role in relation to conventional karyotyping (Bui et al., 2011).

CMA and miscarriage

Miscarriage, defined as spontaneous loss of an intrauterine pregnancy before 20 weeks, is a common reproductive problem as it occurs in 15% of clinically recognized pregnancies. Cytogenetic abnormalities represent the major etiology for reproductive failure (Hassold and Jacobs, 1984; Stephenson et al., 2002). About half of the miscarriages are caused by numerical chromosomal abnormalities, while structural aberrations and mosaicism occur less frequently (2 and 10%, respectively; Bruyere et al., 2002). Chromosomal abnormalities are typically detected by cytogenetic analysis of cultured miscarriage tissue, and therefore the main benefit of using CMAs is to eliminate or minimize problems associated with tissue culture, such as tissue culture failure and contamination by maternal cells (Lomax et al., 2000; Menasha et al., 2005). In a comparative study, CMAs showed a reduced failure rate (owing to degraded DNA) in comparison with cytogenetic analysis (Menten et al., 2009; Robberecht et al., 2009). In addition, CMAs offers new opportunities for improved detection of mosaicism (Fig. 3) and subtle chromosomal imbalances in miscarriages, neither of which is easily detectable by karyotyping because of the small number of cells analyzed and low banding resolution for chromosomes obtained from tissues.

Most of the CMA studies of miscarriages were based on BAC arrays which contained a small targeted selection of BAC probes from defined chromosomal areas (Schaeffer et al., 2004) or a genome-wide selection (Benkhalifa et al., 2005; Shimokawa et al., 2006; Menten et al., 2009; Robberecht et al., 2009). Nevertheless, the studies showed that CMAs can be used to identify maternal cell contamination in 33% of miscarriages (Menten et al., 2009), detect large-scale chromosomal abnormalities in 57% of miscarriages that failed to grow in culture (Benkhalifa et al., 2005), mosaicism for an abnormality affecting 10–30% cells (Schaeffer et al., 2004; Robberecht et al., 2009) and submicroscopic imbalances (CNVs) in 5% of chromosomally normal miscarriages (Shimokawa et al., 2006). A higher resolution oligonucleotide array (244 k Agilent with 0.01 Mb resolution) documented submicroscopic abnormalities in 13% of chromosomally normal miscarriages (Zhang et al., 2009). Confirmation of the CMA-detected CNVs by FISH was performed in only a few of the above studies (Schaeffer et al., 2004; Shimokawa et al., 2006; Robberecht et al., 2009), and, except in rare instances (Shimokawa et al., 2006; Zhang et al., 2009), their exact breakpoints, gene content and/or overlap with CNVs in controls remained unknown, thus making the interpretation of their clinical impact challenging. A follow-up of the confirmed CNVs to determine if they are de novo or parental in origin was not performed in any of the above studies.

Investigations of submicroscopic chromosomal imbalances in miscarriages, which included CNV origin were recently performed using higher resolution arrays (Warren et al., 2009; Rajcan-Separovic et al., 2010a,b). Confirmed de novo CNVs were rare (~10%) and small (~250 kb or less) and did not contain obvious candidate genes for miscarriage. This is in contrast to CNV findings in post-natal cases with normal karyotypes and developmental abnormalities where the median de novo CNV size is 10 times bigger (2.5 Mb; Koolen et al., 2009) and the majority of pathogenic CNVs (>75%) are >1 Mb (de Vries et al., 2005). In chromosomally normal sporadic and recurrent miscarriages, the vast majority of unique (previously unreported) confirmed CNVs were of familial origin (Rajcan-Separovic et al., 2010a,b).
CNVs of familial origin could have a detrimental effect on embryo/placenta development but not on the carrier parent owing to imprinting, recessive mutation on the remaining allele and variable expressivity. For example, familial CNVs can uncover genes suspected or known to be imprinted in placenta, i.e. differentially expressed depending on the parental origin. Figure 4 shows an example of a CNV disrupting a tissue inhibitor of metalloproteinases 2 (TIMP2) gene, transmitted from a woman to her multiple miscarriages, most of which had placental abnormalities (Rajcan-Separovic et al., 2010a). TIMP2 is of interest as it is suspected to be maternally expressed in the placenta in the first trimester and has a role in modulating invasion of the trophoblast into maternal endometrium, as well as in vascular remodeling and angiogenesis of the placenta. Finally, familial CNVs present in the miscarriage and parent could impact the pregnancy if their integral genes play a role in the stress response and if the embryo/fetus was exposed to more stress (because of a less favorable microenvironment) during prenatal development than the parent with the same CNV (Rajcan-Separovic et al., 2010b).

CMA analysis of miscarriages and couples faces several challenges. One of the challenges is the lack of CNV data in appropriate controls. A database of CNVs in individuals with established fertility (reproductive DGV) is therefore desired, as well as a database of CNVs detected in miscarriages and couples with recurrent pregnancy loss. An additional challenge in CMA analysis of miscarriages using the array CGH is the detection of polyploidy, occurring in 10–20% miscarriages (Bruyere et al., 2002; Stephenson et al., 2002). However, polyploidy detection can be improved if CMA analysis is combined with flow cytometry, microsatellite genotyping or interphase FISH (Menten et al., 2009; Robberecht et al., 2009; Zhang et al., 2009) or if a 47, XXY reference DNA is used in the array CGH (Robberecht et al., 2009). SNP arrays would be expected to detect polyploidy but their use to study miscarriages has not yet been reported. Finally, the detection of balanced chromosome rearrangements is not possible by CMAs; however, this type of chromosomal abnormality is rare in miscarriages (2%) (Bruyere et al., 2002) and if the rearrangement is truly balanced, it is less likely to be causative of the miscarriage.

Despite the high potential to improve genetic analysis, the reported use of CMAs for studying miscarriages is still limited in comparison with its widespread use to test patients with developmental delay/congenital anomalies, preimplantation embryos and ongoing pregnancies. One of the possible reasons for the discrepancy is that chromosome analysis of miscarriages is not routinely performed in many centers and is limited to selected cases (e.g. those with morphological abnormalities). This is because the efforts to obtain miscarriage chromosomes by time-consuming culture and karyotype analysis are not considered justified, as many miscarriages are caused by age-related aneuploidy. As CMAs offer the opportunity to enhance, automate and speed up the array analysis, and is particularly effective in instances when the main goal is to detect aneuploidy and larger structural changes, it can facilitate a more widespread analysis of the genomic causes of miscarriage in the future. The interpretation of smaller, unique, miscarriage CNVs will require parental analysis and more extensive genetic counseling for a larger number of miscarriages (≏30–50%; Rajcan-Separovic et al., 2010a,b) in comparison with parental follow-up required for miscarriages that carry a visible structural chromosomal abnormality (≏6%; Stephenson and Kutteh, 2007). Nevertheless, CMA analysis will expand possibilities to support clinical management of the couples experiencing miscarriages and to develop
new knowledge of the factors and causes of early failure of human development. A summary of the benefits and challenges of CMA analysis in miscarriages is presented in Table II.

CMA and reproductive disorders

It has been well established that some reproductive disorders, such as infertility and congenital abnormalities of the reproductive organs e.g. Mullerian duct hypoplasia, premature ovarian failure (POF) and ovarian dysgenesis (OD), have a genetic etiology, either related to chromosomal abnormalities or single-gene changes (Biason-Lauber et al., 2004; Bashamboo et al., 2010; Ledig et al., 2010b; Harton and Tempest, 2012). However, in many cases their cause remains unknown.

CMA analysis of patients with reproductive disorders is rarely reported, and in many of the earlier studies the relevance of detected CNVs was difficult to establish because of the lack of precise breakpoint assignment for CNVs, comparison of patients’ CNVs with DGV, and unavailability of parental samples in adult cases which would help determine the significance of the CNV. Repeated and more standardized analysis using the available guidelines for analysis and interpretation of CNVs in post-natal developmental disorders (Miller et al., 2010; Kearney et al., 2011) may be helpful to gain insight into the role of CNVs in the development and function of the reproductive system. Because there are no adequate controls for which detailed information on reproductive history is available, the exclusion of CNVs reported in DGV as causative for the patient’s reproductive problems should be performed with caution. CNVs reported in DGV but containing genes known to affect sexual development should be considered as potential candidates.

POF and OD

Ledig et al. (2010b) used a 0.02 Mb resolution oligonucleotide array to study 74 patients with POF and OD, and in ~50% of cases reported rare CNVs not, or infrequently, listed in the DGV. The rare CNVs were predominantly small (>90% were <400 kb) and confirmed by quantitative PCR. Genes contained in these CNVs were not previously reported to be associated with POF and consisted, among others, of genes involved in meiosis, DNA repair and folliculogenesis, or in male fertility in homologs of model organisms. An earlier study of 90 women with POF using a ~1 Mb BAC array revealed eight CNVs which were considered relevant based on the frequency difference.
between controls and the study population (Aboura et al., 2009). None of the CNVs from the two studies were recurrent.

Mayer-Rokitansky-Kuster-Hauser (MRKH) syndrome

MRKH is characterized by the absence of the uterus and the upper part of the vagina in women with a normal karyotype. It can be isolated (Mullerian aplasia) or associated with other defects (syndromic), such as renal, skeletal and hearing problems (Bernardini et al., 2009). Recently, pathogenic CNVs were detected in 14% of MRKH cases (Nik-Zainal et al., 2011) and the recurrence of deletion of 16p11.2 (Nik-Zainal et al., 2011), 22q11.2 (Morcel et al., 2011; Nik-Zainal et al., 2011), 1q21.1 regions which includes the TAR syndrome critical region (Cheroki et al., 2008; Ledig et al., 2010c) and 17q12 (≏31.5–33.3 Mb; Fig. 1; Cheroki et al., 2008; Bernardini et al., 2009; Ledig et al., 2010c; Nik-Zainal, et al., 2011) were noted. Gene LHX1 from 17q12 was implicated based on the expression of its mouse homologue lim1 in the epithelium of the developing Mullerian duct (Cheroki et al., 2008) and the fact that female lim1 null mice lack uterus and oviducts, with ovaries unaffected (Kobayashi et al., 2004). Screening for mutation of this gene in additional patients with MRKH identified a putative missense mutation (Ledig et al., 2010c).

Endometriosis

Endometriosis represents an ectopic transplantation and growth of endometrium, mainly through reflux menstruation (Guo et al., 2004; Zafrakas et al., 2008). It is therefore believed to represent benign dissemination and metastasis. The genomic composition of endometrial ovarian cysts and endometrial tissue was studied by two groups (Guo et al., 2004; Zafrakas et al., 2008) and the results were completely opposite: while no genomic changes were found in 10 endometrial ovarian cysts with the 1 Mb resolution BAC array (Zafrakas et al., 2008), all five endometrial samples studied by Guo et al. (2004) with the 3 Mb, lower resolution, array had a large number of genomic abnormalities, reminiscent of the genomic instability noted in cancer. There were many differences between the studies that could explain the observed discrepancy such as sample origin and purity (Guo et al. performed laser microdissection of the endometrial tissue to selectively collect epithelial cells from the endometrial glands and minimize stromal cell contamination), use of degenerate oligonucleotide primed-PCR to amplify DNA (Guo et al., 2004), type of array (earlier versus later generation of BAC arrays) and algorithms for identifying CNVs. CNV content of a normal endometrium was not reported in either study, and one of the studies (Guo et al., 2004) was performed at the time when the concept of DNA copy number variability was only starting to emerge and catalog of CNVs in controls was not yet available. Additional CMA studies of endometriosis are necessary to resolve the discrepancy.

XY gonadal dysgenesis (XY-GD)

XY-GD is a disorder manifested by failure of testicular development despite a normal male karyotype. Ledig et al. (2010a) studied 87 patients with syndromic and non-syndromic XY-GD and in 25% of the syndromic and 5.6% of non-syndromic cases detected CNVs containing genes previously associated with XY-GD (DMRT1, DAX1). A similar frequency of CNVs affecting genes known to have a role in non-syndromic XY-GD was reported by White et al. (2011). Previously unreported CNVs were also identified in both studies (≏in 20% of all cases; Ledig et al., 2010a) containing new putative candidate genes for sexual development. A review by Bashamboo et al. (2010) provides additional examples of submicroscopic imbalances detected by array CGH in XY-GD and other disorders of sex development, containing genes with a previously known dosage effect (SOX9, NROB1) as well as new candidate genes (e.g. genes from the recurrent 15q24 microdeletion region associated with micropenis/hypospadias).

Male infertility

Tutelmann et al. (2011) used a high-resolution array (244 k and 400 k) to determine the CNV content in 89 males with oligosperma and 37 males with azoosperma related to Sertoli-cell-only syndrome (SCOS) in comparison with 100 men with normo-sperma. They reported no difference in the number of CNVs or total amount of CNV-related DNA loss or gain between controls and men with oligo- and azoosperma, suggesting that the loss of germ cells in the patients is not due to increased genomic instability. Recurrent and non-recurrent CNVs were identified; sex chromosome CNVs were significantly over-represented in patients with SCOS. The CNVs listed as potentially causative were predominantly very small (<100 kb) and were not confirmed by an independent method, so their true number remains uncertain. The authors have cataloged CNVs detected in 100 normospermic controls to help alleviate the challenges in interpreting CNVs in patients with spermatogenic

### Table II Benefits and challenges of CMAs for studying miscarriages.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detects most types of large chromosomal imbalances, as with conventional cytogenetics</td>
<td>Neither array type (SNP and CGH based) can detect balanced rearrangements; array CGH is more limited for ploidy change detection than SNP array</td>
</tr>
<tr>
<td>Arrays are quick and DNA based; DNA from paraffin-embedded tissue can be used</td>
<td>DNA tends to be poor quality and degraded</td>
</tr>
<tr>
<td>Arrays are DNA based so the effect of tissue culture failure, artifacts or maternal contamination is minimized</td>
<td>If tissue culture is omitted, there are no chromosomes for FISH confirmation; CNV confirmation has to be DNA based</td>
</tr>
<tr>
<td>CMAs detect potentially pathogenic CNVs that cause or contribute to miscarriage</td>
<td>More parental investigations are needed to interpret CNVs, as in ≏30–50% miscarriages a previously unreported CNV is found. Therefore, more couples will have to be investigated to determine if they carry a miscarriage CNV (≏6% miscarriages show structural chromosomal abnormalities that require parental follow-up)</td>
</tr>
<tr>
<td>Improves understanding of genetic and biological factors implicated in early human development</td>
<td>More extensive genetic counseling necessary for uncertain findings</td>
</tr>
</tbody>
</table>
impairment in the future. Osborne et al. (2007) developed and validated a targeted array for screening for changes in AZF copy number known to be present in 5–10% of infertile men. This opened the path for using array CGH as an alternative to molecular detection of AZF deletions and duplications, allowing a wider range of abnormalities to be identified.

**Conclusion and future directions**

The advent of CMAs has revolutionized the field of medicine as many new and improved genetic diagnoses can be made. CMA testing allows a comprehensive detection of large-scale chromosomal abnormalities, which are a common cause of abnormal development and the main reason for genetic testing in a prenatal setting. In addition, CMAs uncover submicroscopic abnormalities, thereby facilitating diagnosis in an even higher number of cases. However, much of the ground work for CMA already performed for post-natal developmental disorders needs to be carried out in the field of human reproduction. Some examples include development and/or refinement of clinical guidelines for testing of embryos and pregnancies, catalog of CNVs in patients with reproductive disorders and fertile cohorts and determining couples’ attitudes regarding the information provided by CMA testing. Uncovering new genetic causes of impaired reproduction, followed by functional studies of the affected genes and screening for their recurrence in additional affected patients, will lead to an improved understanding of their role in reproduction and more improved clinical management.

**Acknowledgements**

The author thanks Dr Mary Stephenson and Dr Wendy Robinson for their comments on the manuscript.

**Authors’ roles**

E.R.S. reviewed the literature and wrote the manuscript.

**Funding**

E.R.S. is funded by the Michael Smith Foundation for Health Research Scholar award and the Canadian Institute of Health Research grant MOP106467.

**Conflict of interest**

None declared.

**References**


Coppinger J, Allman S, Lamb AN, Torcha BS, Bejani BA, Shaffer LG. Whole-genome microarray analysis in prenatal specimens identifies clinically significant chromosome alterations without increase in results of unclear significance compared to targeted microarray. Prenat Diagn 2009; 29:1156–1166.


Arrays in reproduction

567


