Clinical Application of Microarray-Based Molecular Cytogenetics: An Emerging New Era of Genomic Medicine

Marilyn M. Li, MD and Hans C. Andersson, MD

Geneticists have long recognized the role of genomic imbalances (eg, deletions or duplications of chromosomal material) in the pathogenesis of human disorders. Numerous methods have been developed to detect genomic alterations since the discovery of the correct chromosome number in human cells in 1956. In 1959, Lejeune et al[1] discovered that an extra copy of chromosome 21 (trisomy 21) caused Down syndrome, the first evidence linking genomic imbalances with human disease. Soon after, new clinical syndromes were delineated on the basis of the identification of multiple patients with the same cytogenetic abnormality, such as trisomy 13 in Patau syndrome and trisomy 18 in Edwards syndrome. The identification of the Philadelphia chromosome, which was later showed to be caused by a translocation between chromosomes 9 and 22, and its association with chronic myelocytic leukemia in 1960 marked the beginning of cancer cytogenetics.[2] The invention of chromosome banding techniques in 1970 led to the discovery of numerous structural chromosome aberrations and their association with human diseases.[3] By optimizing culture conditions to arrest cellular division at prometaphase, high-resolution banding could detect chromosomal changes to a resolution of 3 to 5 Mb. The next breakthrough in cytogenetics was the development of fluorescent in situ hybridization (FISH) technology, which laid the foundation for molecular cytogenetics.[4] The technology not only allows the detection of small genomic alterations of 50 Kb to 100 Kb, but also permits the direct visualization of these alterations in uncultured cells. These features made FISH testing ideal not only in detecting microdeletion/microduplication syndromes, but also for prenatal aneuploidy screens, where a fast turnaround time is highly desirable, and for cancer genetics studies, where metaphase chromosomes may not be obtainable. Although FISH allows the detection of genomic imbalances with great accuracy, it can only probe specific sequences that are known and suspected to be associated with known syndromes.

Microarray-based technology, developed in the last decade, affords the capacity to examine the whole human genome on a single chip with a resolution as high as a few hundred base pairs, a process also known as microarray-based cytogenetics.[5] This resolution is at least 10-fold greater than the best prometaphase chromosome analysis, heretofore the most sensitive whole-genome screen for genomic deletions and duplications.[6] Microarray technology represents the technical convergence of molecular genetics and cytogenetics and is rapidly revolutionizing modern cytogenetics. Submicroscopic chromosome copy number variations (CNVs), including 0, 1, or 3 copies, defined as deletions or duplications involving >1 Kb DNA, are detected in patients with mental retardation (MR), autism spectrum disorders (ASD), developmental delay (DD), and multiple congenital anomalies (MCA) of unknown causes. New syndromes begin to emerge on the basis of findings of similar genomic alterations. More than 5000 CNVs have been collected in the Toronto database (http://projects.tcag.ca/variation) so far, and many are associated with human diseases.[7] This article reviews the most commonly used microarray-based cytogenetics platforms, their strengths and limitations, and the implications for medical practice.

Microarray-Based Cytogenetic Technology

Two major groups of microarray-based platforms are currently used in clinical cytogenetics: microarray-based comparative genomic hybridization (aCGH), and single nucleotide polymorphism (SNP) genotyping-based arrays. aCGH, which directly measures genomic copy number differences between the patient DNA and a normal reference DNA, allows the construction of a high-resolution map of genome-wide copy number alterations. aCGH arrays contain thousands of bacterial artificial chromosome (BAC)/P1–derived artificial chromosome (PAC) clones or in situ–synthesized oligonucleotide probes. These probes may either be enriched for known genes or specific chromosomal regions for known syndromes, or distributed relatively evenly across the whole genome. SNP-based arrays probe thousands of SNPs and provide data about both

aCGH	Microarray-based comparative genomic hybridization
AML	Acute myeloid leukemia
ASD	Autism spectrum disorder
BAC	Bacterial artificial chromosome
CNV	Copy number variation
DD	Developmental delay
FISH	Fluorescent in situ hybridization
MCA	Multiple congenital anomalies
MR	Mental retardation
PAC	P1–derived artificial chromosome
PD	Parkinson disease
SNP	Single nucleotide polymorphism
UPD	Uniparental disomy

From the Hayward Genetics Center and the Department of Pediatrics, Tulane University School of Medicine, New Orleans, LA (M.L., H.A.)

The authors declare no conflicts of interest, real or perceived.

 10.1016/j.jpeds.2009.04.001

Table I. Comparison of aCGH and SNP-based arrays

	aCGH	SNP-based arrays
Probes	BAC/PAC, oligos	oligos
SNP genotyping	No	Yes
SNP desert coverage	Yes	No
UPD detection	No	Yes

copy number and genotype; the latter can be used to study copy-neutral genomic alterations, such as uniparental disomy (UPD) seen in imprinting disorders. Although SNP-based arrays have the advantage of detecting UPD and consanguinity, they offer a poor representation of genomic regions with low SNP incidences (SNP deserts). Additionally, SNP-based platforms do not use intraexperimental control; rather, they compare patient data with a preestablished laboratory standard. A comparison of the 2 major microarray platforms is summarized in Table I. As microarray-based technologies continue to improve, many new platforms are being developed. These platforms will provide a combination of high probe density and optimal probe distribution across the genome, including SNP deserts and known chromosome regions that contain repetitive DNA sequences, and allow detections of both CNVs and loss of heterozygosity, including UPD. Better platforms that offer microarray-based cytogenetics with much higher resolution and significantly lower cost are likely to continue to emerge as the technology matures. Microarray-based whole genome sequencing may be just a few years away but is not currently ready for clinical use.

Clinical Application in Genetic Disorders

Detection of Subtle Genomic Imbalance in Patients with MR, ASD, DD, and MCAs

The prevalence of MR, ASD, and DD are reported to be 1% to 3%, 0.67%, and 3.7%, respectively,[8] for which a cause is unknown in as many as 60% to 70% of patients. With conventional cytogenetics, the diagnostic yield (ie, proportion of positive results) is about 3% to 4%; with subtelomere FISH, the yield is 5% to 7%.[9] The positive yield for clinically relevant CNVs with microarray-based cytogenetics is 15% to 20%.[10] The diagnostic yield for isolated MR and ASD may be slightly lower than 15%, but much higher than that of FISH testing. CNVs may occur within so-called genomic "hotspot" regions leading to recurrent microdeletion/microduplication syndromes such as DiGeorge/velocardiofacial syndrome (DGS/VCFS) and its reciprocal 22q11.2 duplication syndrome.[11,12] More often, CNVs are randomly distributed outside hotspot regions with higher incidences in the subtelomeric regions. The hotspot-associated CNVs, which have been postulated to be the result of nonallelic homologous recombination,[13] often present as simple deletions or duplications. Alternatively, CNVs outside hotspot regions often originate from nonhomologous end-joining,[14] some of which occur at the breakpoints of apparently balanced chromosomal translocations or inversions or as subtle unbalanced rearrangements of the subtelomere regions (Figure).[15] Because of the wide distribution and heterogeneity of CNVs in the human genome, whole-genome microarrays are the most useful method for detection of unpredictable, clinically-relevant genomic alterations. The diagnostic yield of whole-genome microarrays is largely dependent on the resolution (average inter-marker distance) of the arrays.

Delineation of Genotype-Phenotype Correlations of Known Syndromes. Phenotypic expression among patients with well-recognized microdeletion or microduplication syndromes varies considerably at least partially because of the size differences of the genomic alterations. FISH analysis, which is still the primary method in many cytogenetics laboratories for identifying deletions/duplications, does not delineate the specific size of the deletion or duplication. Microarray-based cytogenetic testing characterizes CNV size and genomic location, which facilitates genotype-phenotype correlations.

Phenotypic variability between individuals may be due to differences in the makeup of the rest of the genome other than CNV size. The phenotypes of patients with similar genomic alterations may range from apparently normal presentation to profound mental retardation. For example, in DGS/VCFS, intrafamilial phenotypic variation is a common phenomenon even when the deletions have been shown to be identical in different individuals.[16] These other genomic differences, which are often subtle and variable, can be characterized by use of aCGH- or SNP-based arrays, which can detect the allelic differences between individuals. Patients with the same genomic alteration but variable phenotypic expression are an important cohort for further study because they offer clues to the pathophysiologic study of syndromes with CNVs. As more information from other genomic variations is correlated with the effects of CNVs and other factors such as epigenetic and environmental factors, a clearer picture of the role of CNVs in the pathogenesis of genetic disorders will emerge.

Identification of Genes Responsible for Known Syndromes. Microdeletion syndromes may be the phenotypic effects of haploinsufficiency of single genes. Pertinent examples include the *UBE3A* gene in Angelman syndrome,[17] *RAI1* gene in Smith-Magenis syndrome,[18] and *NSD1* gene in Sotos syndrome.[19] Additionally, many monogenic diseases with MR and DD are due to genomic deletions. Relevant examples include a microdeletion at 11p13 where the *PAX6* gene resides in Aniridia type II, and a 7p21 deletion including the *TWIST* gene in Saethre-Chotzen syndrome.[20,21] Microdeletions/microduplications are estimated to comprise up to 15% of all disease-causing mutations underlying monogenic diseases.[22]

Microarray-based cytogenetics provides a powerful strategy for dosage-sensitive disease gene identification. A prime example of such application is the identification of the *CHD7* gene as the cause of CHARGE syndrome.[22] CHARGE syndrome is characterized by some combination of