Reprinted from *American Biotechnology Laboratory* March/April 2010

Technical Article

by Douglas Hurd

Delving Deep Into the Genetic Basis of Disease

Cytogenetics, the study of the chromosomes, has been an active area of biological study since the 19th century. Cytogeneticists are interested in detecting chromosomal amplifications, deletions, and other structural rearrangements (such as translocations and inversions), since there are a large number of constitutional genetic disorders that are caused by such rearrangements. In addition, chromosomal abnormalities have been implicated in certain cancers.

Traditional cytogenetic techniques such as karyotyping and fluorescence in situ hybridization (FISH) are being increasingly supplemented by array comparative genomic hybridization (array CGH or aCGH). Array CGH is also beginning to be used by molecular geneticists. Furthermore, amplification and deletion of genetic material, known as copy number variants (CNVs), are a major source of genetic variation and have been implicated in complex multifactorial diseases. This article focuses on the role of array CGH in cytogenetics, molecular genetics, and elucidating genetic mechanisms in complex diseases, and highlights the benefits of using commercial arrays and services such as CytoSure* and Genefficiency from **Oxford Gene Technology** (**OGT**, Oxford, U.K.).

Traditional cytogenetic techniques

Karyotyping

Karyotyping involves the careful staining of metaphase chromosomes (most commonly) using one of a number of different techniques[1] to produce distinctive banding patterns on each chromosome. The bands are usually formed due to the different packing densities of the heterochromatin and euchromatin regions. Micrograph images of the stained chromosomes are then arranged in numerical order, in their diploid pairs (for somatic cells), with the short (p) arms uppermost (see *Figure 1* for a typical diploid karyogram). The introduction of a range of fluorophores has enabled spectral karyotyping (SKY),[2] where each chromosome is identified by a unique color (using a combination of fluorophores tagged to chromosome-specific probes).

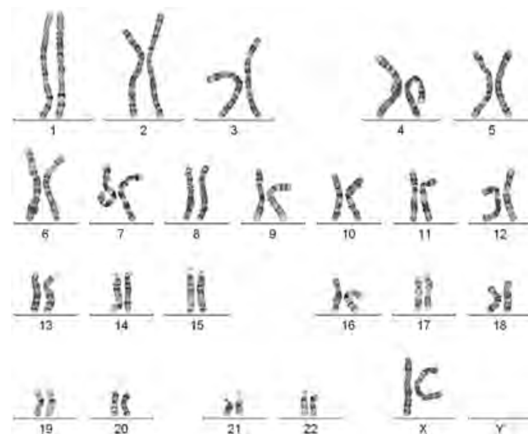

Figure 1 Typical example of a classic human female karyogram, showing a full complement of 44 autosomal and 2 sex chromosomes in their diploid pairs. (Image courtesy of the West Midlands Regional Genetics Laboratory [WMRGL], Birmingham, U.K.)

Karyotyping has been an important technology for cytogeneticists because it provides complete views of an individual's chromosomal complement, enabling the identification of any gross genetic differences, for example, trisomies (e.g., Down syndrome), and structural abnormalities, such as translocations, inversions, deletions, or duplications. Karyotyping, however, requires cells to be grown in culture and is also limited by its low resolution; typically only aberrations of over 5 Mb can be reliably detected.

Fluorescence in situ hybridization

FISH and its derivations, such as multiplex FISH (M-FISH), are powerful cytogenetic techniques that use fluorescently labeled DNA probes to identify specific sequences of interest within an interphase or metaphase spread.[3,4] See *Figure 2* for a typical example of M-FISH. FISH techniques have an advantage over karyotyping in that they provide higher resolution and are capable of detecting microdeletions, as well as balanced and unbalanced translocations not detectable by karyotyping. Importantly, FISH techniques can be performed on uncultured cells and thus are classed as "rapid tests," meaning that copy number changes can be detected within 1 day of receipt of a sample into a laboratory, compared to the 3–7 days for karyotyping (due to the need to culture the sample cells). The limited number of spectrally distinct fluorophores available, though, restricts the number of probes that can be used to "interrogate" the DNA simultaneously.

Comparative genomic hybridization

It can be more efficient to identify such chromosomal aberrations by direct comparison of an unknown sample with a known reference, i.e., CGH. DNA from a sample is labeled with one fluorophore and DNA from a reference sample is labeled with another fluorophore. The DNAs are then hybridized to a normal metaphase spread. CGH has been combined with microarray technology to allow array CGH, which has revolutionized cytogenetics.

*CytoSure products are for research use only. The product is provided under an agreement between **Agilent Technologies, Inc.** (Santa Clara, CA) and **OGT**. The manufacture, use, sale, or import of this product may be subject to one or more U.S. patents, pending applications, and corresponding international equivalents, owned by **Agilent Technologies, Inc.** The purchaser has the nontransferable right to use and consume the product for research use only and not for diagnostics procedures. It is not intended for use, and should not be used, for the diagnosis, prevention, monitoring, treatment, or alleviation of any disease or condition, or for the investigation of any physiological process, in any identifiable human, or for any other medical purpose.

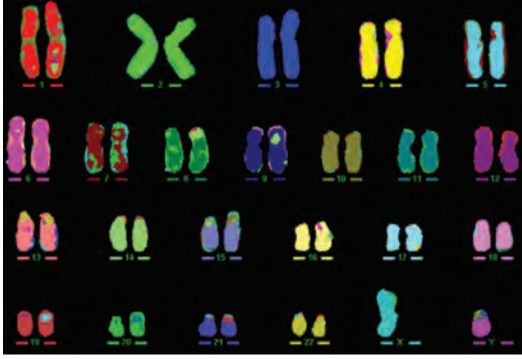

Figure 2 *Multiplex FISH karyogram.*

Overview of array CGH

Microarrays are typically a glass slide with discrete spots consisting of individual DNA probes. In aCGH, the DNA probes on the array can be bacterial artificial chromosomes (BACs), or, as is becoming more commonplace, oligonucleotides. The DNA probes can be deposited presynthesized to produce so-called spotted arrays. However, the use of oligonucleotide arrays in which the oligos are synthesized on the array surface is becoming the preferred method. The oligos can be synthesized onto the specially coated glass surface using either lithography methods or by depositing the phosphoramidites using a sophisticated ink jet printer (**OGT**). This latter method, called in situ ink jet printing, produces extremely high coupling efficiencies, permitting high-quality long oligonucleotides to be fabricated on the arrays. With in situ ink jet printing, 60-mer oligos are typically used.

An overview of a typical array CGH process is shown in *Figure 3*. In general, 1 μg of the sample DNA and a reference DNA are fragmented, and then the sample DNA is labeled with Cy3 and the reference is labeled with Cy5 (Cy3 and Cy5 are trademarks of **GE Healthcare**, Piscataway, NJ). The labeled DNAs are then cohybridized to the probes on the array. The results of the hybridization can then be analyzed using an array scanner and analysis software, which detect the ratio of Cy3 to Cy5 signal intensity at each probe. Most probes will have approximately the same amounts of each fluorophore, showing that the sample and reference DNA have equal copies of that specific DNA fragment. Where there are differences, i.e., more fluorescence from one dye than the other, then there are unequal quantities of that fragment between the sample and reference DNA, indicating a potential CNV. Figure 3 shows a schematic of this process.

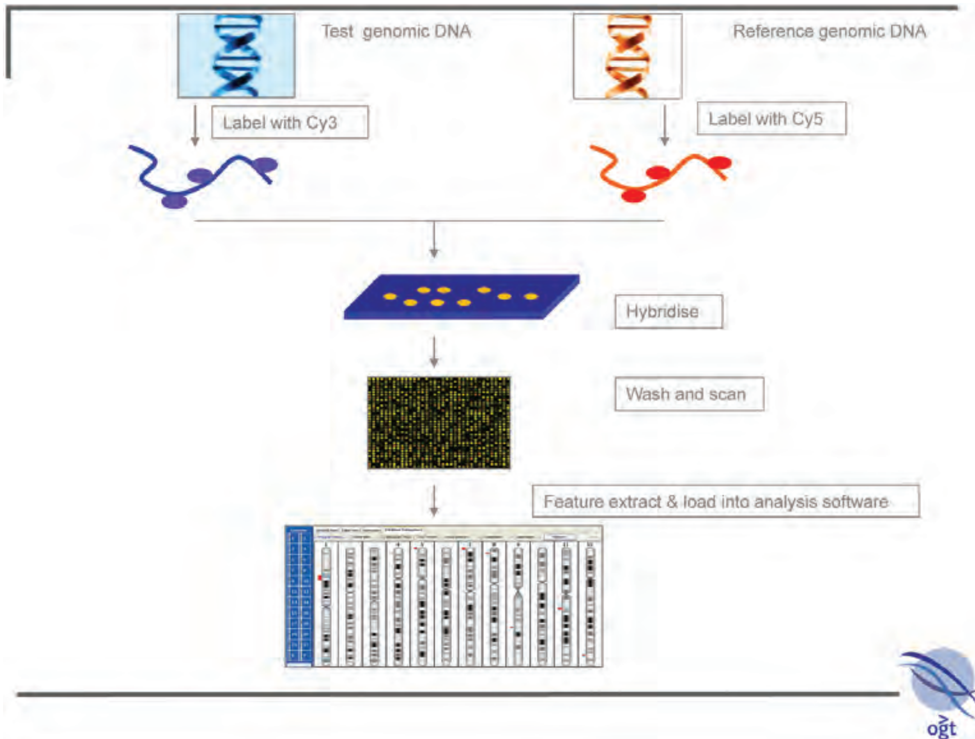

Figure 3 *Typical aCGH work flow. Genomic DNA and reference DNA are labeled with two different fluorophores. These are then combined and hybridized to a microarray containing 105,000 60-mer oligonucleotide probes. After washing and scanning, the results are plotted and the sites of genomic DNA deletions and amplification are determined.*

The primary advantage of array CGH is the substantial increase in overall resolution. Depending on the type of array used, aberrations of a few tens of kilobases can be detected, which is a dramatic improvement over karyotyping. Use of array CGH has resulted in the characterization of a number of new syndromes.

Array CGH, unlike FISH and karyotyping, however, cannot detect copy number neutral events such as unbalanced translocations and inversions. This suggests that there will remain a role for FISH and karyotyping in the future.

Applications of array CGH

1. *Cytogenetic microarrays.* Array CGH is becoming an increasingly common investigative technique for cytogeneticists to understand the genetic mechanism of constitutional diseases and some cancers. The design of specialized oligonucleotide cytogenetic arrays incorporates two aspects. The first is the actual design of the oligonucleotide probes on the arrays. Care needs to be taken to ensure that the oligonucleotide probes hybridize to unique sequences in the genome, that the melting temperature (Tm) is within a certain range to allow efficient hybridization, and that the G/C content is optimal for specific hybridization. Companies such as **OGT** have many years of experience in finding the optimal probes on the array.

The second aspect to take into consideration is the actual distribution of the probes on the genome. Experts such as the International Standard Cytogenomic Array (ISCA) Consortium (http://isca.genetics.emory.edu) are establishing clear aCGH format guidelines for cytogenetic microarrays. Typical cytogenetic arrays have a high number of probes spaced approximately evenly throughout the genome, and then increased density of probes at regions of the genome of interest to cytogeneticists. Companies such as **OGT** have vast experience in designing

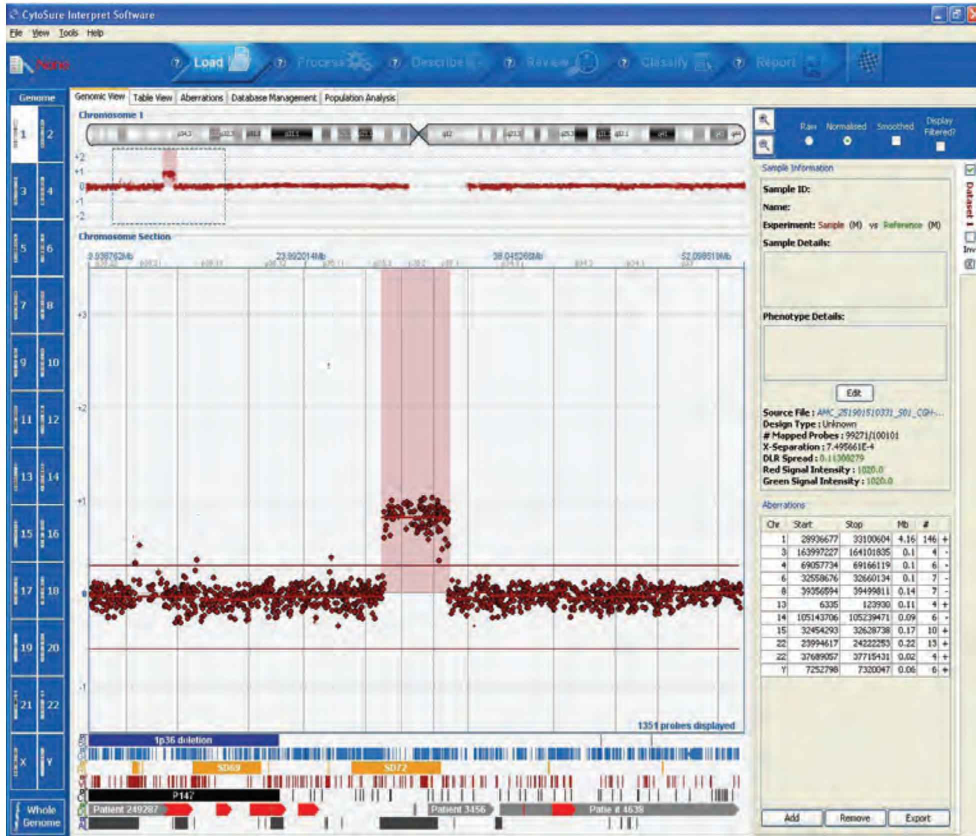

Figure 4 Automated aberration detection with CytoSure Interpret analysis software, showing clear detection of chromosomal abnormalities.

nucleotide arrays, and are working closely with the ISCA Consortium and other leading cytogeneticists to provide arrays to meet their requirements.

Array printing technologies are continually advancing, which allows greater probe densities on each array. This not only enables single arrays to contain more probes and therefore improve resolution, but also permits existing array probe sets to be printed numerous times per slide, permitting simultaneous screening of multiple samples, making each array more cost efficient.

An example of an array design used by cytogeneticists is commercialized by **OGT**. The arrays consists of long 60-mer oligonucleotide probes, which offer greater sensitivity and specificity.[5] The probes are distributed evenly throughout the genome as a backbone at a coverage of one probe every 40 kb, and there is extra probe coverage over 400 genes associated with developmental disorders (e.g., DiGeorge, Williams-Beuren, and hereditary neuropathy with liability to pressure palsies [HNPP]) using a 105,000-probe array.

2. *Microarrays for molecular geneticists.* Array CGH provides a high level of flexibility in resolution, since it is possible to assign the thousands of probes to cover a single gene or group of genes. This enables aCGH to be used for the detection of single-exon aberrations within individual genes. For example, commercial arrays are available that target only the dystrophin gene, using some 40,000 probes spaced at an average of just 10 bp within exons and 106 bp within introns. Arrays covering other genes of interest to molecular geneticists are becoming available.

Use of array CGH in understanding the role of CNVs in complex diseases

Chromosomal aberrations do not always cause disease; there are CNVs present in those with no apparent disease. However, CNVs are a major cause of variation in the genomes of different people, and it has been speculated that CNVs may be at least partially responsible for some multifactorial complex diseases.

Large population studies of CNVs are currently being carried out, and this is facilitated by the fact that aCGH can be automated, thus providing a high-throughput method of mapping CNVs. For large studies, services such as Genefficiency can provide significant amounts of highly accurate CNV data very quickly. A recent study conducted at the Wellcome Trust Sanger Institute (Cambridge, U.K.)[6] utilized the Genefficiency service and demonstrated the value in outsourcing aCGH for large-scale CNV studies.

Analysis of aCGH data

Data analysis of aCGH and interpretation of the results can be challenging. The first step is feature extraction of the files generated by the scanner. Specialized software is available that automates the process of finding each probe and quantitating the Cy3 and Cy5 signals within the probe areas. Following feature extraction, aCGH data are typically plotted as a $\log_2$ (Cy5 signal/Cy3 signal) ratio versus the genomic position.

Aberrations are apparent as data points that deviate from the Y = 0 axis, and algorithms to automate the calling of aberrations are now routinely used. Circular binary segmentation (CBS) divides up the data into segments of similar $\log_2$ ratio. A threshold can then be applied to "call" the segments.

The next step in the data analysis is for the cytogeneticist to determine which of the aberrations are likely to be pathogenic. A growing number of publicly available databases can help the investigator. For example, the Database of Genomic Variants (The Center for Applied Genomics, Toronto, Canada) is a collection of CNVs that have no apparent deleterious effect. The Decipher database (Sanger Institute) is a collection of aberrations that do cause disease. Various tools that provide information on the function of the genes are also available (e.g., Ensembl [EMBL-EBI, Cambridge, U.K./Sanger Institute], AceView [NCBI, NLM, Bethesda, MD], and OMIM® [Johns Hopkins, Baltimore, MD]); many laboratories also build their own databases of deleterious aberrations and benign CNVs.

Sophisticated software, such as CytoSure Interpret, integrates with these public and institutional databases to provide greater context for each array result (see *Figure 4*). As a result of this and other advanced tools, the investigator can perform very powerful interpretation of aCGH results more easily.

Conclusion

Array CGH is having a major impact on the work of cytogeneticists to elucidate the molecular mechanism of disease, and several new syndromes have recently been identified as a result of its adoption by cytogenetic laboratories. There is great potential for aCGH to be used to identify small, potentially single-exon, aberrations within individual genes. Array CGH is also being used for the study of complex genetic diseases.

Adoption of aCGH in these three areas is being facilitated by the increasing availability of economical commercial arrays and customized services. In addition, data analysis is becoming easier with the availability of specialized software. The combination of arrays and software assures researchers of specificity, accuracy, and repeatability, and provides increased confidence in results and clearer decision-making.

References

1. Gustashaw, K.M. Chromosome stains. In *The ACT Cytogenetics Laboratory Manual*, 2nd ed.; Barch, M.J., Ed.; The Association of Cytogenetic Technologists, Raven Press: New York, NY, 1991.
2. Schröck, E.; du Manoir, S.; Veldman, T.; Schoell, B.; Wienberg, J.; Ferguson-Smith, M.A.; Ning, Y.; Ledbetter, D.H.; Bar-Am, I.; Soenksen, D.; Garini, Y.; Ried, T. Multicolor spectral karyotyping of human chromosomes. *Science* **1996**, *273*(5274), 494.
3. Cheung, S.W.; Tishler, P.V.; Atkins, L.; Sengupta, S.K.; Modest, E.J.; Forget, B.G. Gene mapping by fluorescent in situ hybridization. *Cell Biol. Int. Rep.* **1977**, *1*(3), 255–62.
4. Speicher, M.R.; Gwyn Ballard, S.; Ward, D.C. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat. Genet.* **1996**, *12*(4), 368–75.
5. Curtis, C.; Lynch, A.G.; Dunning, M.J.; Spiteri, I.; Marioni, J.C.; Hadfield, J.; Chin, S.F.; Brenton, J.D.; Tavaré, S.; Caldas, C. The pitfalls of platform comparison: DNA copy number array technologies assessed. *BMC Genomics* **2009**, *10*, 588.
6. Conrad, D.F.; Pinto, D.; Redon, R.; Feuk, L.; Gokcumen, O.; Zhang, Y.; Aerts, J.; Andrews, T.D.; Barnes, C.; Campbell, P.; Fitzgerald, T.; Hu, M.; Ihm, C.H.; Kristiansson, K.; Macarthur, D.G.; Macdonald, J.R.; Onyiah, I.; Pang, A.W.; Robson, S.; Stirrups, K.; Valsesia, A.; Walter, K.; Wei, J.; The Wellcome Trust Case Control Consortium; Tyler-Smith, C.; Carter, N.P.; Lee, C.; Scherer, S.W.; Hurles, M.E. Origins and functional impact of copy number variation in the human genome. *Nature* **2009** Oct 7 [Epub ahead of print].

Dr. Hurd is Project Manager, ***Oxford Gene Technology****, Begbroke Science Park, Sandy Ln., Yarnton, Oxford OX5 1PF, U.K.; tel.: +44 1865 856800; fax: +44 1865 848684; e-mail: douglas.hurd@ogt.co.uk.*