Array comparative genomic hybridisation: a new tool in the diagnostic genetic armoury

Renate Marquis-Nicholson, Salim Aftimos, Ian Hayes, Alice George, Donald R Love

Abstract

The traditional understanding of genetic disease, that, with the exception of aneuploidy, it is due primarily to single base pair changes or small deletions and duplications has been challenged over the last decade. This challenge has been spearheaded by increasing evidence of the frequency and significance of larger genomic rearrangements. It now appears that a substantial proportion of Mendelian conditions are caused by deletions and duplications that involve the copy number of one or more contiguous genes. It is becoming apparent too that de novo chromosomal events are much more frequent than spontaneous point mutations and that chromosomal rearrangement is likely to account for the vast majority of sporadic disease.

Background

The extent of copy number variation between apparently healthy individuals and the importance of this variation in the development of disease were not appreciated until early this decade. Much of this discovery has been precipitated by a novel molecular cytogenetic technique, array comparative genomic hybridisation (aCGH). This technique is paving the way for a revolution not only in disease gene discovery and our understanding of the molecular basis of complex disease, but also in clinical diagnostics.

While the use of aCGH as a research tool has been examined extensively in the literature, the principal aim of this review is to discuss its place in diagnostics, with an emphasis on its advantages and limitations when compared to more conventional cytogenetic techniques.

Introduction

It was initially thought that the completion of the Human Genome Project would provide a standardised reference template of the entire genetic code.\(^1\) It was estimated that the genomes of healthy individuals were 99.9% identical, with differences in the remaining 0.1% resulting in phenotypic variation.\(^1\)

Studies published in the early 2000s, such as the landmark studies of Sebat et al\(^2\) and Iafrate et al\(^3\), established that there is significant variation between apparently normal individuals. This variation lies not in the DNA sequence itself but in the number of copies an individual has of each particular DNA sequence.\(^1\)\(^-\)\(^4\)

Copy number variants (CNVs) have been defined as chromosomal segments of more than 1kb (kilobase) in length whose copy number varies, as a result of deletion or duplication, between individuals in the population.\(^1\)
The majority of CNVs appear to be benign and have no clinical significance, but others have been linked to single gene disorders with high penetrance (such as Charcot-Marie-Tooth disease type 1A), single gene disorders with incomplete penetrance (hereditary neuropathy with liability to pressure palsies) and multiple gene disorders with high penetrance (Williams syndrome).\textsuperscript{5,6}

The term genomic disorders is used to refer to diseases that are caused by abnormal dosage or dysregulation of one or more genes resulting from rearrangement of the human genome.\textsuperscript{5,6}

The role that CNVs play in the development of disease is not yet completely understood. Several modes of action have been suggested: a dosage sensitive gene may be contained within the CNV region; a hemizygous deletion within a gene region may unmask a recessive mutation on an homologous chromosome; there may be a positional effect on a gene due to disruption of regulatory elements that are located within a CNV; or a gene may be directly disrupted by the breakpoint of a CNV.\textsuperscript{7}

The cytogenetic techniques conventionally used to detect chromosomal abnormalities include the Giemsa-banded karyotype and fluorescence \textit{in situ} hybridisation (FISH). Standard karyotyping (550 band), however, is limited by the resolution of the banding and is therefore only able to accurately identify deletions or duplications of more than \textasciitilde5Mb (Megabases),\textsuperscript{8} while FISH can only examine a certain number of loci at once and requires a clinical suspicion of the genes involved.\textsuperscript{9}

Recent advances in array-based technology allow the simultaneous screening of the whole genome at a high degree of resolution, detecting unsuspected genomic changes that are too small to be resolved by karyotyping or FISH, and too large to be observed by DNA sequencing.

\textbf{Array comparative genomic hybridisation (aCGH)}

Comparative genomic hybridisation (CGH) was first developed as a genome-wide method to detect copy number changes of >10Mb in solid tumours.\textsuperscript{10} It was initially performed on a spread of metaphase chromosomes, but as the technique evolved it was recognised that greater resolution could be achieved if many small probes were employed instead—indeed the resolution was limited only by the number and length of probes.\textsuperscript{11,12}

Array CGH entails differential labelling of a test and a control sample followed by co-hybridisation of these samples to an array of DNA segments of known sequence. The consequent fluorescence ratio is measured and the results interpreted to determine if there are differences in the copy number between the two samples.

The resources generated for the Human Genome Project had already allowed a library of cloned DNA fragments to be produced which accurately spanned the entire genome.\textsuperscript{13} Probes derived from this library were selected to interrogate a particular genomic region of interest, then fixed to a glass or silica slide and the first microarray was born.\textsuperscript{13}

Probes employed since then have ranged from large-insert clones of 40–200kb, to small-insert clones (1.5–4.5kb), cDNA clones of 0.5–2kb, genomic PCR products of
100bp–1.5kb and oligonucleotide probes of 25–80bp.\textsuperscript{13} The first arrays were constructed by manually or robotically spotting the desired probes onto the slide.

Current practice involves either spotting or the \textit{in situ} synthesis of oligonucleotides using piezoelectric printing or photolithography, respectively.\textsuperscript{13} It is now possible to have more than 2 million probes on a single array slide (25mm × 76mm).\textsuperscript{14} A range of microarray formats is available, covering a spectrum from genome-wide to highly targeted arrays. Traditional genome-wide arrays employ evenly spaced probes across the entire genome.\textsuperscript{15} The problem with the use of genome-wide arrays in a clinical setting has been that casting the net so widely inevitably led to the detection of CNVs of unknown clinical significance.\textsuperscript{15} Targeted arrays, on the other hand, have had denser probe coverage, and therefore higher resolution, in regions of particular interest and hence could better delineate the breakpoints of genomic rearrangements.\textsuperscript{15} With ongoing refinement of technology, the distinction between formats is increasingly less clear-cut.

Most arrays currently available for clinical use employ a combination of the two methods: probe coverage across the entire genome, together with coverage of a higher density in chromosomal regions known to be important in pathogenesis (such as the telomeres) and in clinically important genes.\textsuperscript{15} In this way a single microarray is able to harness the power of both targeted and genome-wide approaches simultaneously. The conventional workflow of an aCGH approach is described in Box 1. Although variant formats exist that offer the same outcome, comparative analysis of these formats is beyond the scope of this review.

\textbf{Array CGH in the clinical setting}

The high resolution and broad surveying capacity of aCGH are perfectly suited to the research laboratory, where novel findings are key. For the diagnostician, however, the optimism engendered by the technique is necessarily tempered by the need to put the new wealth of information into a clinical context. The identification of a microdeletion as the cause of a child’s developmental delay, for example, is of obvious interest, but of greater importance is how this result can be translated into an improvement in the clinical management of that child.

\textbf{Box 1. Array CGH—the method}

Equal amounts of test and control DNA are labelled with different fluorescent dyes and competitively hybridised to the microarray slide. This slide contains probes of known chromosomal location in the genome.\textsuperscript{9} The excess (unbound) DNA is washed off, the array is scanned and the image files are exported to a computer.\textsuperscript{9}
Specially designed software extracts the signal intensities at each probe location and obtains a log ratio. If the ratio of the two fluorescence intensities is equal then the copy number of the particular genomic sequence is assumed to be the same in the test and control DNA. If the fluorescent signal from the test DNA is more intense in a particular area then a copy number gain (for example a duplication or a trisomy) can be inferred and vice versa for a copy number loss (deletion or monosomy).

Any genomic imbalances that are identified are validated by other cytogenetic and molecular methods such as FISH analysis, quantitative PCR methods, customized multiplex ligation-dependent probe amplification (MLPA) assays or microarray formats with higher resolution in the area concerned. Testing of parental samples is sometimes also required to determine whether the abnormality is inherited or de novo.

Potential benefits

Array CGH is more sensitive than standard karyotyping or FISH and is therefore more likely to enable a definitive diagnosis to be made, a diagnosis which carries with it a probable clinical course and long-term prognosis. This increased diagnostic certainty has several significant benefits for the clinical team, the affected individual, and their family:

- For previously recognised conditions, making the diagnosis at a molecular level allows more accurate advice to be given when counselling family members with regard to risk of recurrence and possible pattern of inheritance. It also enables prenatal and carrier testing to be performed if appropriate.
- For chromosomal abnormalities that are not reported in the literature, analysis of the genes affected can reveal potential complications and indicate the most probable clinical course. Physicians involved in the care of the patient are then able to anticipate these complications and undertake appropriate surveillance. Mutations in the CEP290, NPHP1 and RPGIP1L genes, for instance, result in a cerebello-renal phenotype which includes renal failure in the first or second decade of life. If alterations of these genes are identified early then anticipatory measures can be undertaken.
- If there is no efficacious therapy for the particular disorder, a molecular diagnosis will allow the affected individual to be placed in an appropriate clinical trial (if available) and aid in new therapeutic developments.
• Children are able to be proactively placed in appropriate programmes involving, for example, occupational therapy and early speech and language therapy, as opposed to watching and waiting until particular functional deficits become obvious.  

• Affected individuals and their families appreciate being given concrete information. The guilt which frequently surrounds the birth of a child with multiple congenital anomalies, for instance, can be assuaged when parents learn that these anomalies are not because they ‘did something wrong’ when the baby was in utero. For many people even an unfavourable diagnosis is preferable to the uncertainty of the unknown. Local or international support groups can be sought, bolstering emotional well-being and strengthening support networks.

There are significant benefits for laboratory staff also. The diagnosis in question is reached more quickly when aCGH is employed. There is no need to culture cells, greatly reducing sample preparation time. Much of the process is able to be automated, simplifying the workflow regime. Furthermore, the quantity of DNA that is required is much smaller than that needed for karyotyping or FISH analysis, and as cells are not going through the culturing process, the patient’s tissue sample does not need to be of high quality.

Potential limitations

As mentioned above, the high resolution of aCGH is at once the attraction of the technique and also one of the main clinical limitations at this time. The high prevalence of CNVs in the healthy population means that not all variants discovered in affected individuals can be viewed as causative of the presenting phenotype.

If the particular variation in question is found in an unaffected parent it can usually be assumed to be benign (although this assumption is complicated by low penetrance and variable phenotypic expression); likewise if the variation is described in the literature as occurring in the normal population. If the variation is found to be de novo then it is more likely to be pathogenic.

Pathogenicity is supported if there are reports of similar variations in the literature and if the genes involved in the alteration are consistent with the phenotypic abnormalities that are evident clinically.

Large international databases such as the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER) (http://www.sanger.ac.uk/Software/analysis/decipher/database.shtml), the European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA) (http://www.ECARUCA.net) and the Database of Genomic Variants (http://projects.tcag.ca/variation; Toronto database) have been established to allow the assembly of data from multiple sources and to elucidate the connections between genotype and phenotype.

It is hoped that as these databases become more comprehensive, diagnostic laboratories around the world can be increasingly confident when evaluating the clinical significance of their CNV results.
There have been concerns expressed, too, about the cost of integrating microarrays into routine diagnostic practice. The largest outlay is the initial cost of the scanner, followed by that of the microarray slides themselves. A cost-analysis study performed in the United Kingdom by Wordsworth et al\(^\text{18}\) formally examined the possibility of the NHS funding aCGH as a first-line diagnostic test for idiopathic developmental delay.

Their analysis found the technique to be entirely feasible in this context,\(^\text{18}\) resulting in less money being spent in total per diagnosis. Despite the test being more expensive than karyotyping, multi-telomere FISH or multi-telomere MLPA, the process is more automated meaning that staff costs are lower and fewer follow-up or confirmatory tests need to be done to reach a diagnosis.\(^\text{18}\)

Diagnoses made earlier also avoid additional diagnostic tests, including neuroimaging or invasive procedures such as muscle biopsy, and a negative result on aCGH allows follow-up testing to be limited.\(^\text{18}\)

**Specific clinical scenarios**

Since its inception, aCGH has been used extensively in the research of a diverse group of disorders, including (but not limited to) idiopathic and syndromic developmental delay, congenital anomalies, various neurological and neuropsychological disorders and a range of malignancies,\(^\text{9}\) and is now finding its way into diagnostic laboratories.

**Prenatal screening**

Referrals for prenatal genetic testing essentially fall into one of a series of categories: positive maternal serum screen, advanced maternal age, abnormalities detected on ultrasound scanning, or a significant family history.\(^\text{21}\) In New Zealand this subset of expectant mothers are offered a diagnostic test for foetal chromosome aneuploidy and foetal trisomy 13, 18 or 21 via karyotype and FISH testing.\(^\text{22}\)

Further targeted testing is performed if there is a specific concern, such as a characteristic abnormality on ultrasound scan suggestive of a particular syndrome, a family history of a certain disorder (such as cystic fibrosis or sickle cell anaemia), or if parents or other family members are known carriers of balanced chromosome rearrangements.\(^\text{22}\)

The detection rate of clinically significant abnormalities when aCGH is applied to samples routinely referred for prenatal testing is approximately 7%, more than three times that of conventional analysis with karyotyping and FISH.\(^\text{23}\)

Array CGH can not only screen for all the copy number imbalances detected by standard G-banded karyotyping together with aneuploidy FISH, but is also useful when conventional analysis is difficult due to small sample size, poor growth of cultures or limited numbers of cells in mitosis.\(^\text{23}\)

The identification of CNVs of unknown clinical significance is particularly important in the prenatal setting, especially when consideration is being given to the termination of pregnancy. The detection of variants of unknown clinical significance is reduced by using a targeted microarray format.\(^\text{23}\)
The Medical Genetics Laboratory at Baylor College of Medicine, for example, has developed an array for prenatal testing which, in conjunction with a standard karyotype, is recommended for patients who are at an increased risk of genetic disorders, or for whom there are concerns about foetal genetic abnormalities.\(^\text{24}\) This array tests specifically for more than 150 genetic disorders, containing probes for almost all known microdeletion/duplication syndromes as well as having increased coverage in the subtelomeric and pericentromeric regions.\(^\text{24}\)

Van den Veyver et al\(^\text{21}\) recently performed karyotyping, FISH and aCGH analysis (using a custom-designed targeted array) on 300 prenatal samples. Additional significant information was provided by aCGH in 2.3% of cases (7/300), including 2 cases in which the chromosomal disorder would not have been found if only standard karyotyping or aneuploidy FISH had been performed.\(^\text{21}\)

The authors put this into striking perspective by noting that the ‘risk’ of missing the diagnosis if aCGH had not been used is equivalent to 1/150.\(^\text{21}\) This risk is comparable to the term pregnancy risk for Trisomy 21 of a 38-year-old woman, or to the total term risk of all common aneuploidies of a 36-year-old woman.\(^\text{21}\)

In order to ensure that adequate pre and post-test counselling are provided, further large studies are required to clarify:

- The appropriate resolution and format of the array;\(^\text{23}\)
- The ratio of the detection of clinically significant abnormalities to that of abnormalities of uncertain clinical significance;\(^\text{23}\) and
- How identifying the additional abnormalities will alter prenatal and immediate postnatal management.\(^\text{23}\)

Until these studies have been performed it is likely that aCGH will remain an adjunct to, rather than replacement for, standard karyotyping and aneuploidy FISH in the prenatal context.\(^\text{16,23}\)

**Developmental delay (with and without dysmorphic features)**

Global developmental delay is a common clinical problem—with an estimated prevalence of 1–3%.\(^\text{18}\) The underlying cause is unclear in more than 50% of cases but the condition is known to be extremely heterogeneous in aetiology, with multiple chromosomal anomalies implicated.\(^\text{9}^\) A wide range of syndromes include developmental delay amongst their phenotypic characteristics, and with the aid of aCGH the list of microdeletion and duplication syndromes is consistently increasing.\(^\text{9,25}\) A large number of children also have idiopathic developmental delay without dysmorphic features or congenital anomalies.\(^\text{9}\)

The current investigation of patients with developmental delay/intellectual disability includes a karyotype and a subtelomeric FISH screen, as abnormalities in the subtelomeric and pericentromeric regions are implicated in these conditions.\(^\text{26}\) Targeted FISH can also be used to look for particular syndromes if these are suggested by the clinical presentation.\(^\text{27}\)
One of the major benefits of employing aCGH in the setting of a heterogeneous condition such as developmental delay/intellectual disability is that the referring clinician does not need to have any suspicion of the particular chromosomal abnormality involved. It is also undoubtedly useful in the further investigation of abnormal conventional cytogenetic results in order to:

- Delineate the extent of the deletion or duplication, and more accurately identify the genes involved,
- In the case of an apparently balanced rearrangement, to exclude cryptic deletions or duplications at the breakpoints.

Array CGH is also less labour intensive, offers higher throughput, and is more rapid and cheaper (per locus) than multiple targeted FISH tests.

The important questions to be considered, then, are how the diagnostic yield of aCGH compares to that of standard karyotyping and FISH, and whether the higher resolution of aCGH means that it should supplant these techniques in routine practice.

A recent retrospective analysis of 36,325 Dutch patients with developmental delay or intellectual disability, performed by Hochstenbach et al, aimed to answer these questions.

The diagnostic yield of aCGH as an initial test was estimated to be 19% based on an extensive literature review. This figure contrasts with an 8% abnormality rate detected by standard karyotyping and FISH. Critically, only 0.78% of those cases with an abnormal karyotype or FISH result involved balanced rearrangements which would not have been detected by aCGH. 0.48% of these were familial and 0.23% were de novo balanced rearrangements.

At this time the published guidelines for the use of aCGH in the investigation of developmental delay/intellectual disability recommend using arrays in the event of a normal karyotype. However, this situation is extremely dynamic. It is likely that as experience with aCGH in this clinical context grows, and with advances in streamlining the workflow and reduction in labour costs, aCGH will become the first-line diagnostic test. Karyotyping would then only be performed following a normal aCGH result in order to identify the <1% of patients with a balanced chromosomal rearrangement.

**Autistic spectrum disorder**

Autistic spectrum disorders (ASDs) are a heterogeneous group with a complex aetiology and a prevalence of approximately 1/166. They can be clustered into complex autism, when there are dysmorphic features or other factors suggestive of a chromosomal disorder/syndrome, and essential autism (without these features). Historically a genetic cause has been identified in only 5–10% of people affected by an ASD.

Gene mapping via linkage analysis has shown autism candidate gene loci on 20 different chromosomes, including the X chromosome. ASDs are also associated with several single gene disorders, including Rett syndrome, fragile X syndrome and tuberous sclerosis.
The higher resolution that aCGH allows has been used in several large studies, which show that de novo deletions and duplications play a significant role in the aetiology of autism, particularly complex autism.\textsuperscript{31,33} Two seminal studies detected copy number changes in 30\% of individuals with complex autism (Jacquemont et al)\textsuperscript{33} and 10\% of individuals with essential idiopathic autism (Sebat et al).\textsuperscript{31}

An Autism Chromosome Rearrangement Database has been established, and is a reference site delineating the breakpoints and other genomic features that have been described in publicly available literature.\textsuperscript{9}

**Cancer**

It has been well recognised that somatic chromosomal rearrangements and dosage alterations have a role to play in tumourigenesis.\textsuperscript{9} They affect gene expression and can disrupt normal growth control pathways by activating oncogenes or inactivating tumour suppressor genes.\textsuperscript{9}

Much research has been done to characterise the aberrations found in association with a wide range of cancers.\textsuperscript{34} It is hoped that ongoing high resolution microarray analysis will provide an ever more comprehensive understanding of the molecular mechanisms involved in the development of specific cancers.\textsuperscript{9,34} The goal is to develop a database of known genetic alterations responsible for each malignancy, along with the phenotype and clinical course.

Determination of each individual’s tumour profile at diagnosis would then enable not only classification of tumour type, but also prediction of prognosis and susceptibility to particular therapeutic agents—thereby improving the specificity of the therapy given and avoiding unnecessary treatments.\textsuperscript{34} Array CGH analysis in this instance will be complementary to existing diagnostic and prognostic markers, helping to build a more complete, complex and accurate picture for the clinician(s).

There are already steps being taken in this direction. Chronic lymphocytic leukaemia is associated with a wide range of clinical outcomes, with disease progression ranging from indolent to rapid.\textsuperscript{35} There are, however, a range of characteristic copy number changes that are routinely used as prognostic markers.\textsuperscript{35} The deletion of 13q14 in the absence of other abnormalities, for example, is associated with a more favourable prognosis.\textsuperscript{9,35} Traditionally FISH has been used to detect these abnormalities and provide prognostic stratification.\textsuperscript{35}

A recent study by Patel et al\textsuperscript{35} showed that a custom-designed targeted microarray including all regions implicated in CLL was able to test for all abnormalities simultaneously and with a higher yield than conventional karyotyping and FISH (detecting 37 compared to 34 abnormalities). With B-cell clonal enrichment the sensitivity of aCGH was 100\% for cases with aberrations in at least 25\% of the cell population.\textsuperscript{35}

Array CGH has proven itself useful, too, in the identification of individuals with an increased susceptibility to developing malignant tumours. Adam et al\textsuperscript{36} have reported three cases in which genome wide testing using the commercially available EmArrayCyto6000 (an array with resolution equivalent to a 6000 band karyotype and targeted coverage of telomeric, centromeric and gene-rich regions) was performed on patients referred with developmental delay/dysmorphic features/congenital
anomalies. None of these patients had a clinically recognisable cancer predisposition syndrome and each had normal routine cytogenetic analysis.\textsuperscript{36}

Interestingly, aCGH revealed a different \textit{de novo} deletion in each patient, but with the loss of genes that are known to affect tumour susceptibility.\textsuperscript{36} The results of the copy number analysis on these patients had a direct bearing on their clinical management, with appropriate tumour surveillance protocols being initiated.\textsuperscript{36}

\textbf{Conclusions}

Array CGH is a powerful technique that is leading a revolution in the field of genomic medicine. The use of microarrays is enabling the elucidation of the chromosomal aetiology of known syndromes, the expansion of clinical phenotypes and the discovery of previously undefined syndromes.

Microarrays are less labour intensive than traditional testing platforms due to the automation of many of the processes. The great flexibility of their design and the recent major reduction in cost, coupled with the need for minimal laboratory space and personnel effort, make them a highly attractive platform for diagnostic use. As part of the arsenal of molecular cytogenetic techniques they are able to detect and define with great accuracy, and at a high-throughput, both microscopic and sub-microscopic deletions, duplications and rearrangements.

Over the next few years the targeted versus whole-genome debate will no longer be relevant as the two will be integrated, ensuring comprehensive coverage of all clinically relevant loci while excluding known areas of benign variation.

The interpretation of aCGH results, however, poses a significant challenge that should not be trivialised. As the degree of resolution used to analyse the genome increases, it becomes increasingly imperative to confidently filter the information to achieve outcomes that are understandable and clinically relevant.

It is important for diagnostic laboratories to be cautious in the adoption of novel technologies. This caution protects the welfare of patients by ensuring that a potential diagnostic test has well-established sensitivity, specificity and a suitable risk/benefit ratio.

Through research, array CGH has been shown to meet these criteria in a wide range of clinical settings. The time is ripe for the technique to take its rightful place as an integral part of diagnostic medicine.

\textbf{Competing interests:} None known.

\textbf{Author information:} Renate Marquis-Nicholson, Pathology Registrar (Trainee in Molecular Genetics)\textsuperscript{1}; Salim Aftimos and Ian Hayes, Clinical Geneticists\textsuperscript{2}; Alice George, Cytogeneticist\textsuperscript{1}; Donald R Love, Molecular Geneticist\textsuperscript{1,3}

1. LabPLUS, Auckland City Hospital, Auckland
2. Northern Regional Genetic Service, Auckland
3. School of Biological Sciences, The University of Auckland,

\textbf{Correspondence:} Dr Donald R Love, Diagnostic Genetics, LabPlus, Auckland City Hospital, PO Box 110031, Auckland, New Zealand. Fax: +64 (0)9 3074939; email: donaldl@adhb.govt.nz
References:


24. Baylor College of Medicine, Medical Genetics Laboratories [Internet]. Prenatal chromosomal microarray analysis – information for medical professionals. http://www.bcm.edu/geneticlabs/cma/index.html


