Array comparative genomic hybridization (array CGH)

for diagnosis of the genetic causes of learning disability

Learning disability

Learning disability (LD) may be defined as a significant impairment of cognitive and adaptive functions, with onset before the age of 18 years. It is a relatively common condition, affecting 1-3% of individuals worldwide. It is generally estimated that there is a primarily genetic cause in around half of all patients with severe LD and around 15% of patients with mild LD.

Why diagnose?

Whilst a genetic diagnosis does not necessarily offer a 'cure', wherever possible diagnosing the underlying cause of LD is important both from a patient and a clinical perspective. A diagnosis may be helpful for understanding aetiology, giving a prognosis, guiding treatment and providing appropriate advice and support to families. A diagnosis is helpful for the parents to better understand the future needs of the child, to access relevant services and support groups, and to make informed decisions about further children. Both groups also benefit from ending the uncertain, lengthy and often costly 'diagnostic odyssey'.

Genetic variation

Recent research has uncovered several sources of variation between human genomes, which may cause disease in certain cases:

- *Microscopically visible changes* of more than 5 million bases, including entire chromosome duplications (*e.g.* Down's syndrome). These are present in less than 1% of the population, and are almost always pathogenic.
- *Copy number variations (CNVs)* are common changes in the amount of a region of the genome, including duplications and deletions of around a thousand to five million bases. Everyone has around a hundred CNVs in their genome, some of which may predispose individuals to disease but most of which are benign.
- *Single nucleotide polymorphisms (SNPs)* are common changes to single base code in the DNA. Everyone has around 3 million SNPs in their genome, some of which may predispose individuals to disease but most of which are benign. [Rare single base changes are called mutations and may cause serious inherited diseases, such as cystic fibrosis.]

Recent advances in genomics have widened understanding of genetic factors and particularly genomic imbalance (CNVs) in the aetiology of LD. Whether a CNV is benign or pathogenic depends largely upon its size and genetic content.

Genetic tests

Clinical assessment of children with LD typically involves examination by a paediatrician and investigations including biochemical, haematological and genetic (cytogenetic or chromosome) tests. There are several different techniques available for diagnosing the genetic causes of LD, which focus principally on finding pathogenic CNVs. Established testing procedures involve karyotype analysis, in which the whole genome of the individual is visualised in order to identify large copy number changes. This can be followed by targeted techniques such as fluorescent *in situ* hybridisation (FISH) or multiplex ligation-dependent probe amplification (MLPA) to identify specific smaller chromosomal deletions associated with specific syndromes.

*Figure 1. G-banded Karyotype of normal male (46,XY)*
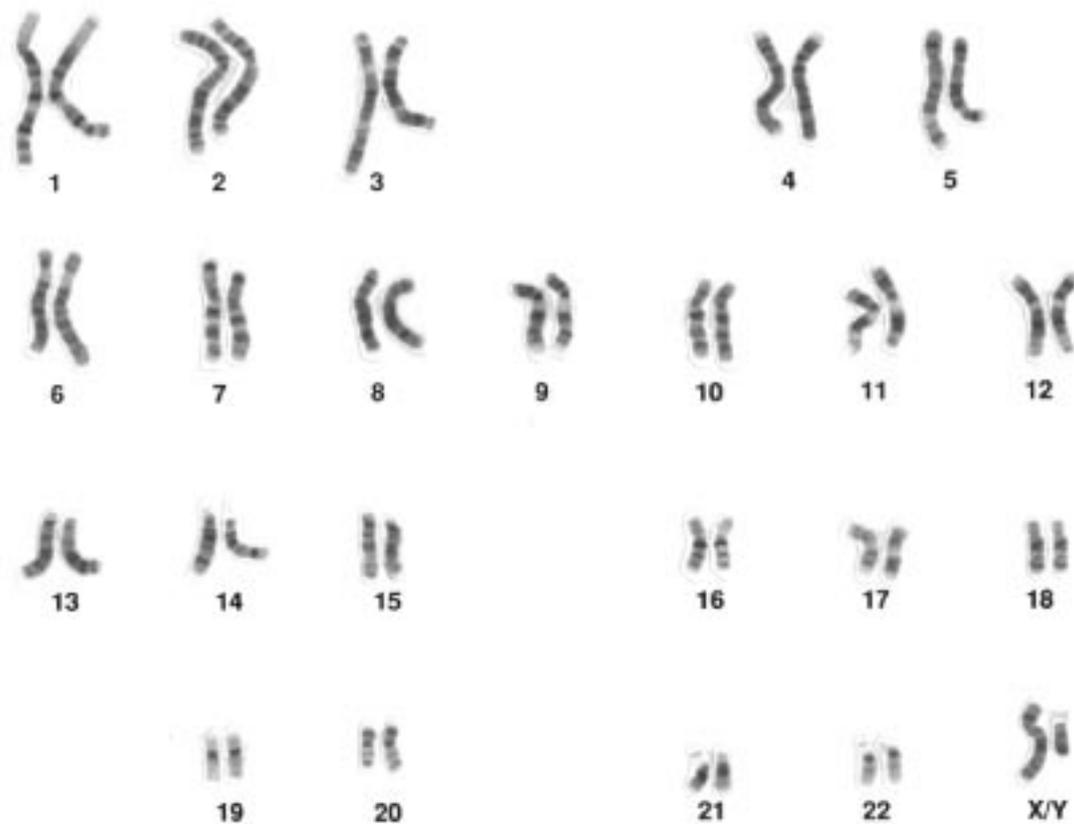
Array comparative genomic hybridisation (array CGH) is a relatively new technique which is at present increasingly being used in patients in whom other cytogenetic tests are negative. Array CGH compares the patient's genome against a reference genome (normal control or standard) and identifies differences between the two genomes and hence locates regions of genomic imbalance (CNVs) in the patient. To aid analysis, the whole genome is fragmented into many small regions and the array is arranged so the exact location of each fragment within the whole genome can be identified. From this, the gene content of any imbalance can be established and the genes can be evaluated against the patient's phenotype.

Figure 2. Principles of array-CGH.

To determine how copy numbers differ from a reference (control) sample:

- *The sample and reference DNA are labeled with different coloured fluorescent probes (green and red).*

- *The two samples are applied to immobilized DNA on the array, and complementary sequences bind.*

- *Where there is no change in sequence copy number in the test sample, there will be equal binding of test and reference sample DNA, and equal amounts of green and red fluorescence will produce a net combined emission colour (yellow).*

- *For sequences where there has been a duplication in the test sample, there will be more green than red fluorescence and an overall green emission; conversely, deletions will result in a reduced level of green fluorescence relative to the red fluorescence from the reference sample, and a net emission of red light.*

Sample DNA   **Reference DNA**

Hybridize to array

Ratio

Position on chromosome
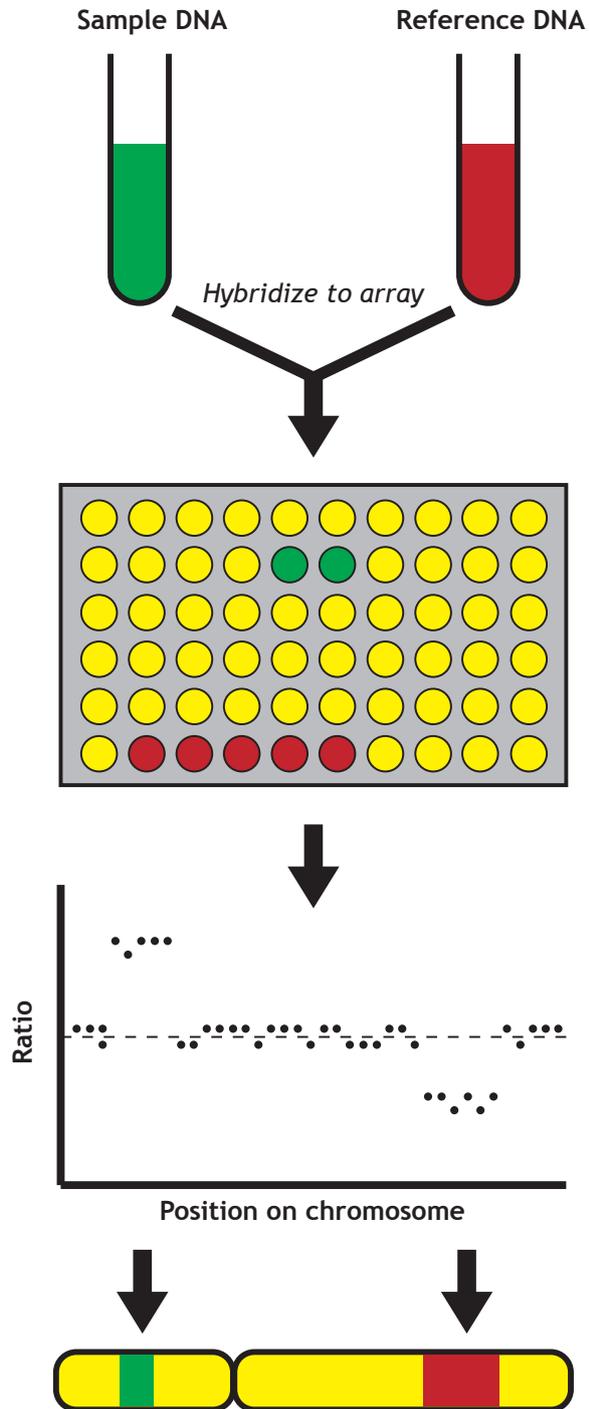
Different array platforms use different levels of analysis; some examine DNA spaced at large intervals along the genome and are described as low resolution (*e.g.* BAC arrays), whilst others examine more DNA spaced close together and hence are described as high resolution (*e.g.* oligo or SNP arrays). Importantly, the larger the genomic spacing, the lower the resolution of the array, and the greater the likelihood that CNVs may fall in the spaces and be missed.

Arrays may also be targeted specifically at a particular region of the genome by enriching the number of DNA fragments examined in that region.

Figure 3. Genome-wide results from a high resolution microarray. CNVs in the test genome relative to the reference genome are highlighted across all the chromosomes (duplications in green and deletions in red).
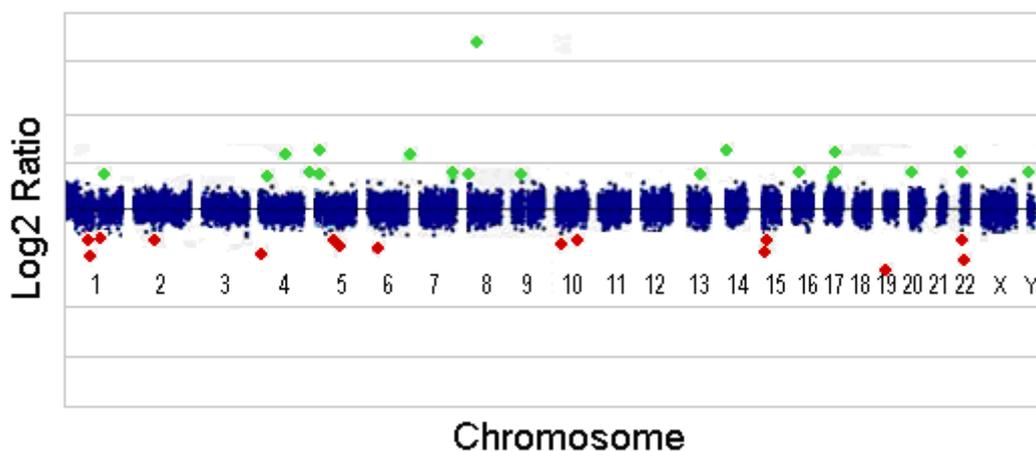
Karyotyping versus array-CGH

In principle, both karyotyping and arrays are genome-wide technologies which can be used to assess the presence of genomic imbalance such as CNVs. Although they may look like very different technologies, the primary difference between them is in the resolution, which is a measure of the level of magnification of the genome. A standard G-banded karyotype usually has a resolution of around 5 Mb (*i.e.* it can detect changes of greater than a five million basepairs). Modern arrays act like a more powerful microscope. Depending upon the particular array and how many DNA probes it uses, it is possible to detect changes greater than 1 Mb (one million basepairs) at low resolution or, changes as small as 10 kb (10 thousand basepairs) at high resolution.

Much smaller CNVs can be detected by using higher resolution technologies, which means that more pathogenic CNVs may be detected using modern arrays than through karyotyping. However, because CNVs are relatively common throughout the genome, numerous benign CNVs will also be detected, so careful interpretation and follow-up testing is needed.