

Array comparative genomic hybridisation (aCGH) delivers superior performance in the detection and quantification of CNVs

Nicole Sparkes, Head of High Throughput Services, Oxford Gene Technology



The importance of copy number

Copy number variations (CNVs) are generally defined as inter- or intra-chromosomal duplications or deletions of segments of DNA greater than 1kb in length¹. The vast majority are unbalanced, resulting in net gains or losses of genomic DNA.

The prevalence of CNVs in the healthy population suggests that these represent a significant proportion of total genomic variation, higher than that of single nucleotide polymorphisms (SNPs). It is estimated that CNVs may affect as much as 4-5% of the human genome². This includes inherited and *de novo* losses or gains in genomic sequence.

A proportion of these CNVs are found in coding or regulatory regions of the genome, where they can affect expression levels of genes. As a result, these CNVs have been implicated in the pathogenesis of a growing number of diseases, including autoimmune disease³, asthma⁴, schizophrenia⁵ and obesity^{6,7}. It has been estimated that the *de novo* rate of CNV mutation in each newborn is approximately 1 in 8 for deletions and 1 in 50 for duplications, which puts it between the rates for *de novo* point mutation and chromosomal aneuploidy⁸, and highlights CNVs as significant contributors to genomic and phenotypic variation.

Choice of platform for high throughput CNV analysis

Two divergent array-based methods have emerged for the detection and analysis of CNVs. The first uses existing SNP genotyping platforms, on the assumption that it should be possible to “tag” regions of CNV with SNPs, with the SNP acting as a marker for the actual CNV. It has been shown that there is little or no linkage disequilibrium between SNPs and common forms of CNVs^{2,9}. However, many

regions of the genome do not contain any SNPs, so these platforms developed to become “hybrid” genotyping arrays through the inclusion of non-polymorphic probes in regions of known copy number variation that did not contain SNPs.

The second approach is based on aCGH – comparative genomic hybridisation of reference and test genomic samples to long (60-mer) oligonucleotide probes. This uses dual-colour labelling of reference and test samples, followed by competitive hybridisation onto the probes on the array. Dual colour fluorescence detection allows calculation of the relative intensity of both dyes, highlighting differences in copy number between the samples. The long oligonucleotide probes are designed to cover the whole genome, and are designed to directly detect known and *de novo* CNVs.

A recent study published at BioMed Central (BMC) compared the various array platforms available for DNA copy number analysis – aCGH platforms from Agilent and Nimblegen and SNP-CNV platforms from Affymetrix and Illumina¹⁰. The Agilent aCGH platform (as used by Oxford Gene Technology Limited [OGT]) emerged as the best aCGH platform, and also showed clear benefits over SNP-CNV hybrid approaches, outperforming the others in:

- **Platform reproducibility:** detection of copy number changes relies on the ability to distinguish “true” change over background noise. Long oligonucleotide arrays exhibit signal to noise ratios an order of magnitude greater than SNP genotyping arrays¹¹, and in the BMC study, the Agilent platform showed superior signal to noise performance to all others, allowing superior sensitivity even when the number of probes was smaller. Agilent probes also showed the least variation between replicate probes, and greater specificity of detection in self-



self hybridisation tests.

- **Detection of whole chromosome gains or losses:** the Agilent aCGH platform performed best in the identification of whole-chromosome events, enabling quantification of the change as well as discrimination of different copy number states.
- **Detection of chromosome arm and subchromosomal gains or losses:** although all four platforms were able to detect these alterations, Agilent showed the greatest clarity of signal.
- **Detection of *de novo* CNVs:** unlike aCGH platforms, which, by their design, allow detection of known and *de novo* CNVs, SNP-CNV platforms are primarily designed to target known CNVs. This study did not examine the ability of either method to detect *de novo* CNVs, but concluded that the platform with dedicated probe coverage in the region, and the performance to allow detection, will perform best. The Agilent aCGH platform consistently showed the best performance throughout the study.
- **Analysis of tumour samples:** aCGH showed clear benefits in the detection of copy number aberrations in samples which were diluted to simulate 70% stromal contamination. Of the two aCGH platforms, Agilent emerged as the clear leader, robustly detecting the alteration at nearly the same level as the undiluted case.

The added strength of the Agilent array lies in the ease of design and manufacture which, in contrast to expensive photolithography techniques, makes customisation of array content affordable and routine. Agilent employs ink-jet *in situ* printer technology (IJSS), under license from OGT, and OGT in turn exploits this strength to offer complete flexibility in

Genefficiency™ high throughput CNV services. Researchers can choose between catalogue arrays or they can customise arrays to achieve maximum resolution over desired regions of the genome.

OGT and Agilent

OGT was recently named as Agilent's first high throughput microarray service provider. OGT combines the superior performance of Agilent aCGH arrays with the OGT Oligome™ database of more than 22 million dedicated CNV probes, allowing the user to completely customise their CNV analysis to either specific regions of the genome, or a whole genome study. Design of probes across the region of interest can enable high resolution detection of breakpoints.

In addition, OGT provides a team of highly experienced project scientists, and the capacity to analyse more than 1,000 samples per week in a fully automated workflow, to provide high throughput, high quality results rapidly.

Conclusion

The Agilent aCGH platform delivers superior performance in the detection and quantification of CNVs. OGT has developed a high throughput processing and analysis service in collaboration with Agilent, and has since been named as Agilent's first certified high throughput service provider, allowing the screening of more than 1,000 samples per week. The combined benefits of Agilent and OGT for aCGH include:

- Superior sensitivity in the detection of *de novo* CNVs
- More accurate quantification of copy number changes
- Better sensitivity in the detection of CNVs in tumour samples



- Flexibility to completely customise your CNV study, allowing very high resolution analysis and more comprehensive genome coverage
- Capacity to run >1,000 samples per week, with >30 QC checks on each sample
- Batch processing of feature extracted files to detect regions with copy number changes using the CBS algorithm¹²
- Delivery of a database file with all pre-processed results loaded, including a dedicated software package that enables population analysis based on CNV calls

References

1. Gökçümen, O. and Lee, C. (2009) Copy number variants (CNVs) in primate species using array-based comparative genomic hybridisation. *Methods* 49, 18-25
2. Conrad, D.F. *et al* (2009) Origins and functional impact of copy number variation in the human genome. *Nature advance online publication* Oct 7, doi:10.1038/nature08516
3. Fanciulli, M. *et al* (2007) FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. *Nature Genetics* 39, 721-723
4. Brasch-Andersen, C. *et al* (2004) Possible gene dosage effect of glutathione-S-transferases on atopic asthma: using real-time PCR for quantification of GSTM1 and GSTT1 gene copy numbers. *Human Mutation* 24, 208-214
5. Walsh, T. *et al* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008, 539-543
6. Walters, R.G. *et al* (2010) A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* 463, 671-675
7. Bochukova, E.G. *et al* (2010) Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature advance publication online* doi:10.1038/nature08689
8. Shen, Y. and Wu, B.-L. (2009) Microarray-based genomic DNA profiling technologies in clinical molecular diagnostics. *Clinical Chemistry* 55, 659-669
9. Kato, M. *et al* (2010) Population-genetic nature of copy number variations in the human genome. *Human Molecular Genetics* 19, 761-773
10. Curtis, C. *et al* (2009) The pitfalls of platform comparison: DNA copy number array technologies assessed. *BMC Genomics* 10, 588-610
11. Ionita-Laza, I. *et al* (2009) Genetic Association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics* 93, 22-26
12. Ohlsen, A.B. *et al* (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 5, 557-572



Genomic Services from OGT: Combining industry leading platforms, expert people, unparalleled process power and performance — to rapidly deliver high quality genomic data to you.
We call this Genefficiency™