Genetics and Genomics in Clinical Research Course

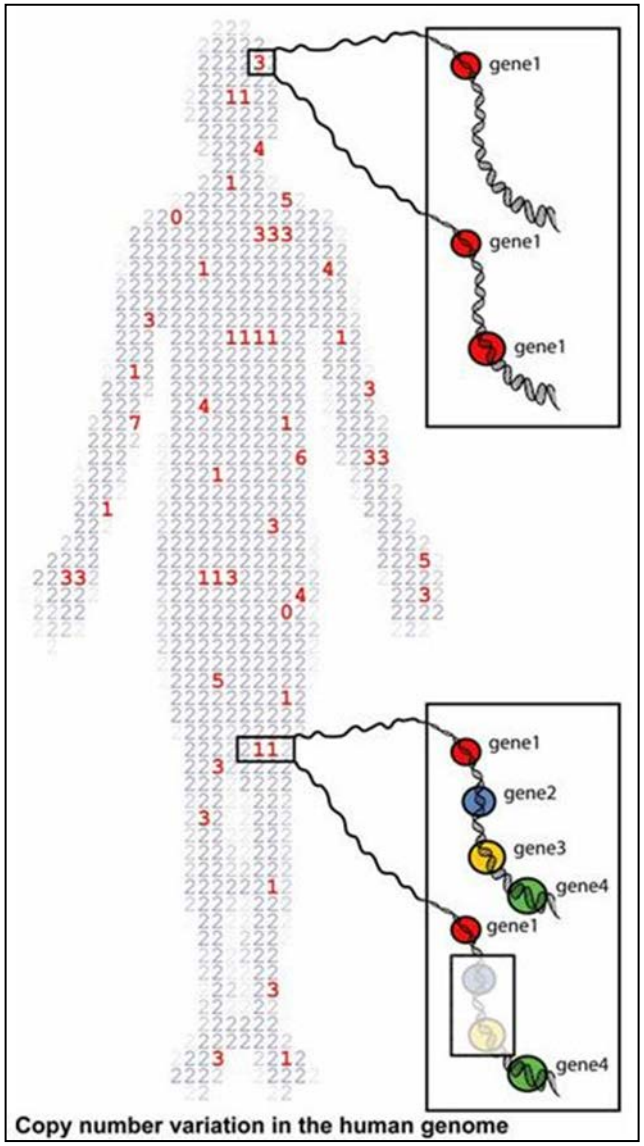
Copy number variants (CNVs) and chromosomal microarray technologies

October 13th 2015

Fady M. Mikhail, MD, PhD
Associate Professor
Department of Genetics

Copy number variants (CNVs)

- Stretches of genomic DNA present in more than or less than two copies that can range in size from kilobases (kb) to megabases (Mb)
- Cannot be identified by G-banded chromosome analysis, but can be identified by Cytogenomic array methodologies and whole genome sequencing
- Can be germline or somatic
- Can be inherited or sporadic (*de novo).* Large *de novo* CNVs are more likely to be disease causative



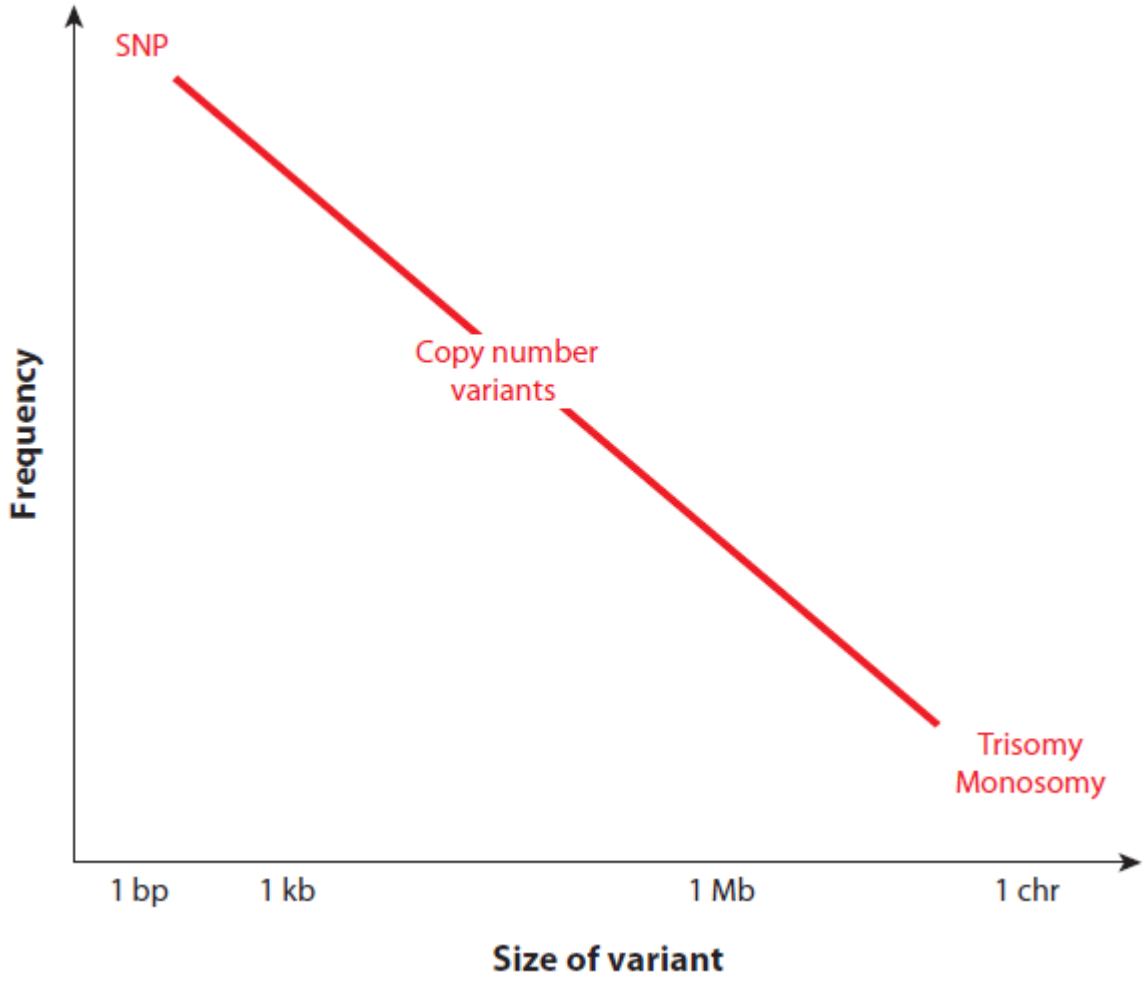
Copy number variants (CNVs) (cont'd)

- Recent studies have indicated that CNVs are widespread in the human genome and are a significant source of human genetic variation accounting for population diversity and human disease. Between any two individuals the number of base-pair differences due to CNVs is >100-fold higher compared with SNPs

- The phenotypic effects of CNVs are sometimes unclear and depend on whether they span dosage-sensitive genes or regulatory sequences

- In a clinical setting, CNVs have been categorized into five groups (according ACMG practice guidelines):
  1. Benign
  2. Variant of unknown significance (VOUS) - most likely benign
  3. VOUS - uncertain significance
  4. VOUS - most likely pathogenic
  5. Pathogenic

Size and frequency of major categories of genetic variants



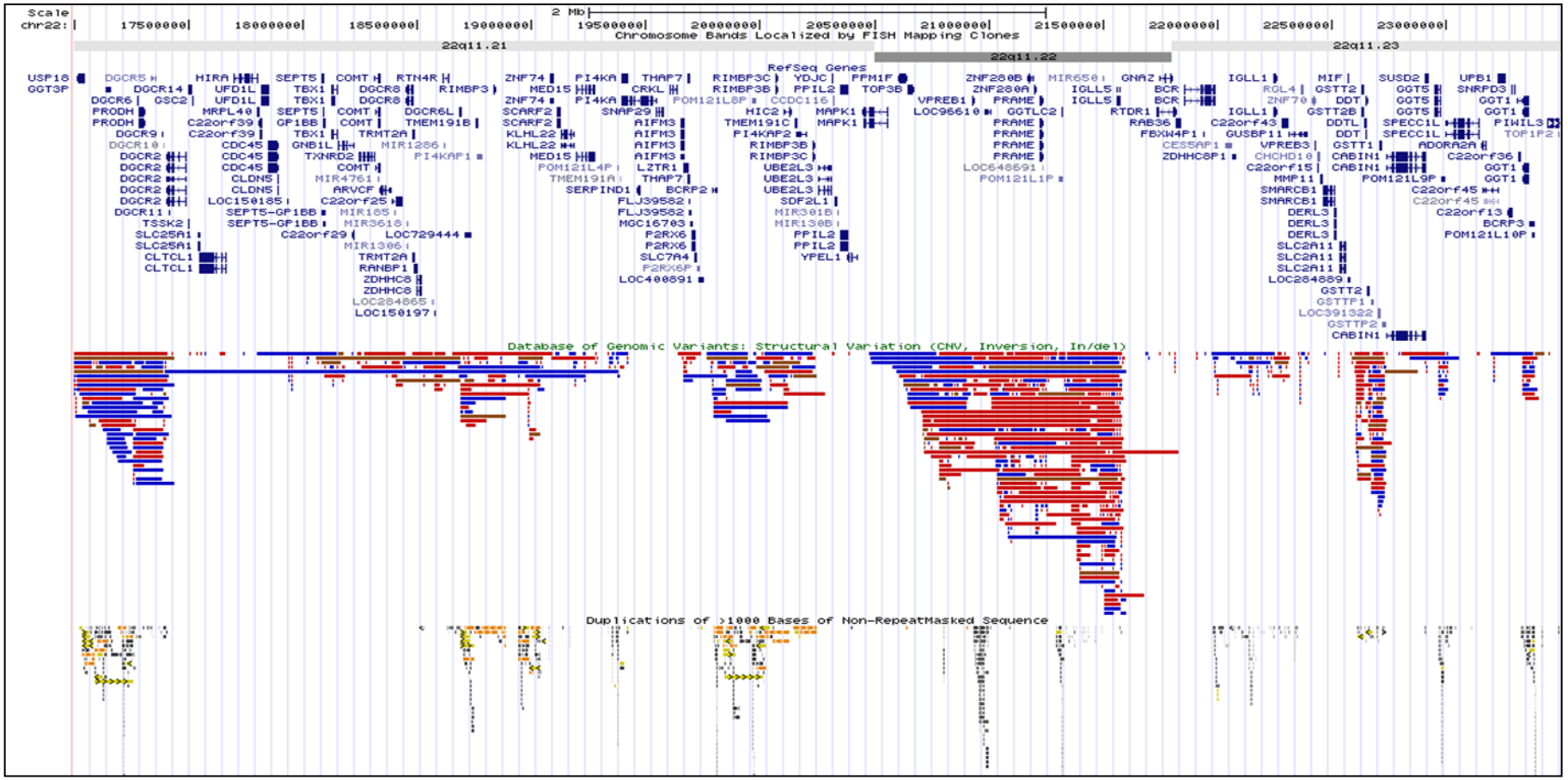
Girirajan S et al. Annu Rev Genet 2011;45:203-26

Genomic rearrangements versus base pair alterations

	Genomic rearrangements (including CNVs)	Base pair (bp) alterations
Size	Thousands to millions of bp	Small scale gene mutations (e.g. point mutations)
Gene content	One to several genes	One gene
Molecular mechanism	• Mechanisms mediated or stimulated by genomic architecture OR • Exogenous factors (e.g. ionizing radiation)	• Errors of DNA replication and/or repair OR • Exogenous factors (e.g. chemical mutagens)
Locus-specific mutation rate (μ)	CNVs: 1.7×10^{-6} - 1.2×10^{-4}	Single-nucleotide changes: 1.8 - 2.5×10^{-8}
Method of detection	• G-banded chromosomes • FISH • Cytogenomic arrays	• DNA sequencing • Other molecular techniques

Benign CNVs

- A recent estimate of the proportion of the human genome that is structurally variant (i.e. benign CNVs) is in the order of ~5-10%
- The majority (>95%) of benign CNVs in humans are <100 kb in size
- The Database of Genomic Variants (DGV) is an important resource that catalogues benign CNVs reported in healthy controls and is continuously updated with new data from peer reviewed research studies



Can CNVs cause disease?

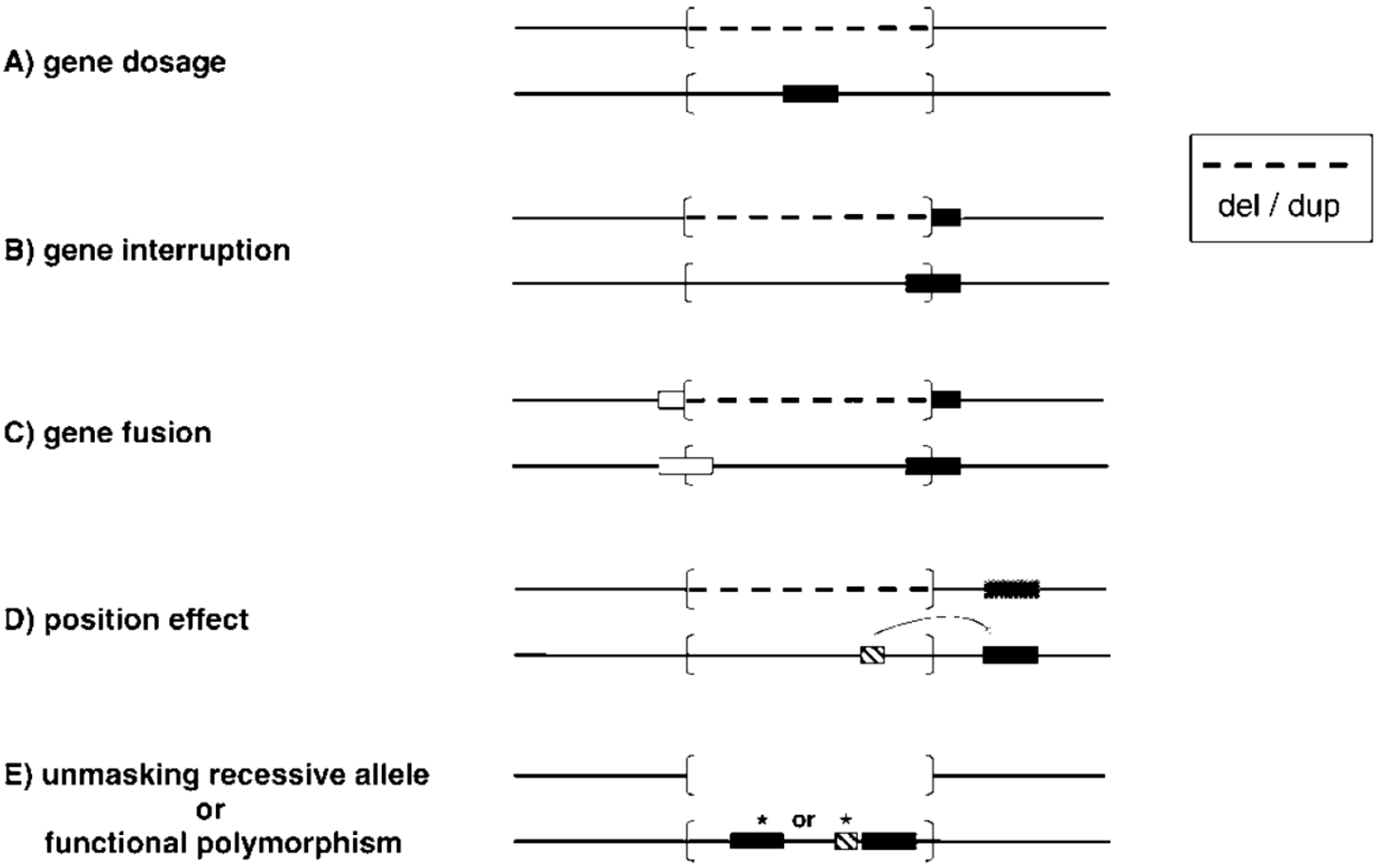
- Most CNVs are benign variants that will not directly cause disease

- However, benign CNVs can encompass genes, especially those of the immune and environmental response pathways, which suggests that they likely play an important role in local adaptive selection in human populations

- CNVs that encompass critical developmental genes can cause disease. Some of these encompass multiple contiguous genes, including dosage-sensitive genes, each contributing to the phenotype independently. Others encompass a single gene or just few genes

- Clinically, pathogenic CNVs are observed in approximately 20% of patients with neurodevelopmental problems. Because of their high diagnostic yield, Cytogenomic arrays were recommended in 2010 by the ACMG as the preferred first-tier clinical diagnostic test for individuals with developmental delay (DD), intellectual disability (ID) or multiple congenital anomalies (MCAs)

Molecular mechanisms by which genomic rearrangements can convey phenotypes



Lupski JR, Stankiewicz P. PLoS Genet 2005;1:e49

Interpretation of the clinical significance of CNVs

Table 1. Assessment of Pathogenicity of a CNV[a]

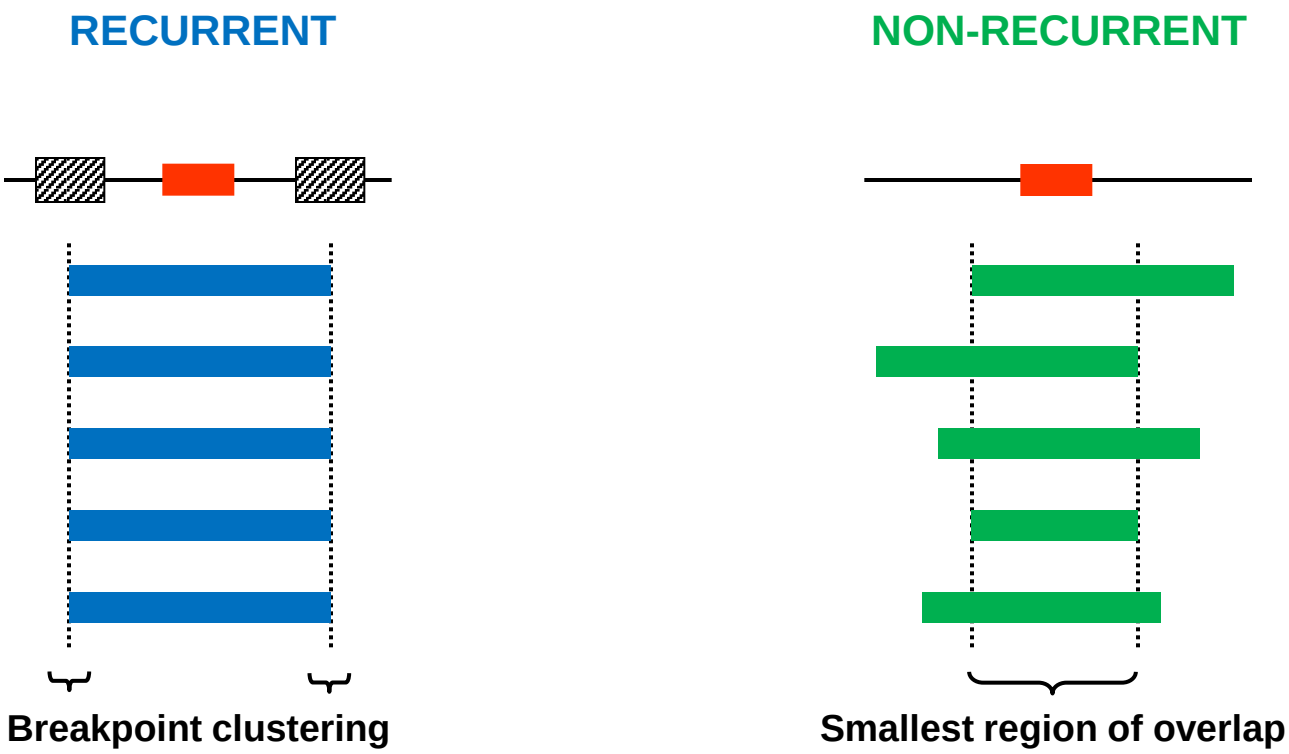
Primary Criteria		Indicates CNV Is Probably Pathogenic	Indicates CNV Is Probably Benign
1.	a. Identical CNV inherited from a healthy parent[b]		✓
	b. Expanded or altered CNV inherited from a parent	✓	
	c. Identical CNV inherited from an affected parent	✓	
2.	a. Similar to a CNV in a healthy relative		✓
	b. Similar to a CNV in an affected relative	✓	
3.	CNV is completely contained within genomic imbalance defined by a high-resolution technology in a CNV database of healthy individuals		✓
4.	CNV overlaps a genomic imbalance defined by a high-resolution technology in a CNV database for patients with ID/DD, ASD, or MCA	✓	
5.	CNV overlaps genomic coordinates for a known genomic-imbalance syndrome (i.e., previously published or well-recognized deletion or duplication syndrome)	✓	
6.	CNV contains morbid OMIM genes[c]	✓	
7.	a. CNV is gene rich	✓	
	b. CNV is gene poor		✓

Table 1. Assessment of Pathogenicity of a CNV[a]

General Findings[d]		Indicates CNV Is Probably Pathogenic	Indicates CNV Is Probably Benign
1.	a. CNV is a deletion	✓	
	b. CNV is a homozygous deletion	✓	
2.	a. CNV is a duplication (no known dosage-sensitive genes)		✓
	b. CNV is an amplification (greater than 1 copy gain)	✓	
3.	CNV is devoid of known regulatory elements		✓

Miller DT et al. Am J Hum Genet 2010;86:749-64

CNVs associated with genomic disorders can be



- Gene

- Segmental duplication , also called low copy repeat (LCR)

Three major mechanisms have been proposed for the generation of CNVs

1. Non-Allelic Homologous Recombination (NAHR) } **RECURRENT CNVs**
2. Non-Homologous End-Joining (NHEJ) } **NON-RECURRENT CNVs**
3. Fork Stalling and Template Switching (FoSTeS)

Microdeletion/microduplication syndromes

- Aka: contiguous gene syndromes, segmental aneusomy syndromes, genomic disorders

- A group of clinically recognizable disorders characterized by a deletion or a duplication of a chromosomal segment spanning multiple dosage-sensitive genes, each contributing to the phenotype independently

- Clinically, each syndrome is characterized by a specific and complex phenotype, which was recognized in most cases as a genetic syndrome before knowledge of their cytogenetic etiology

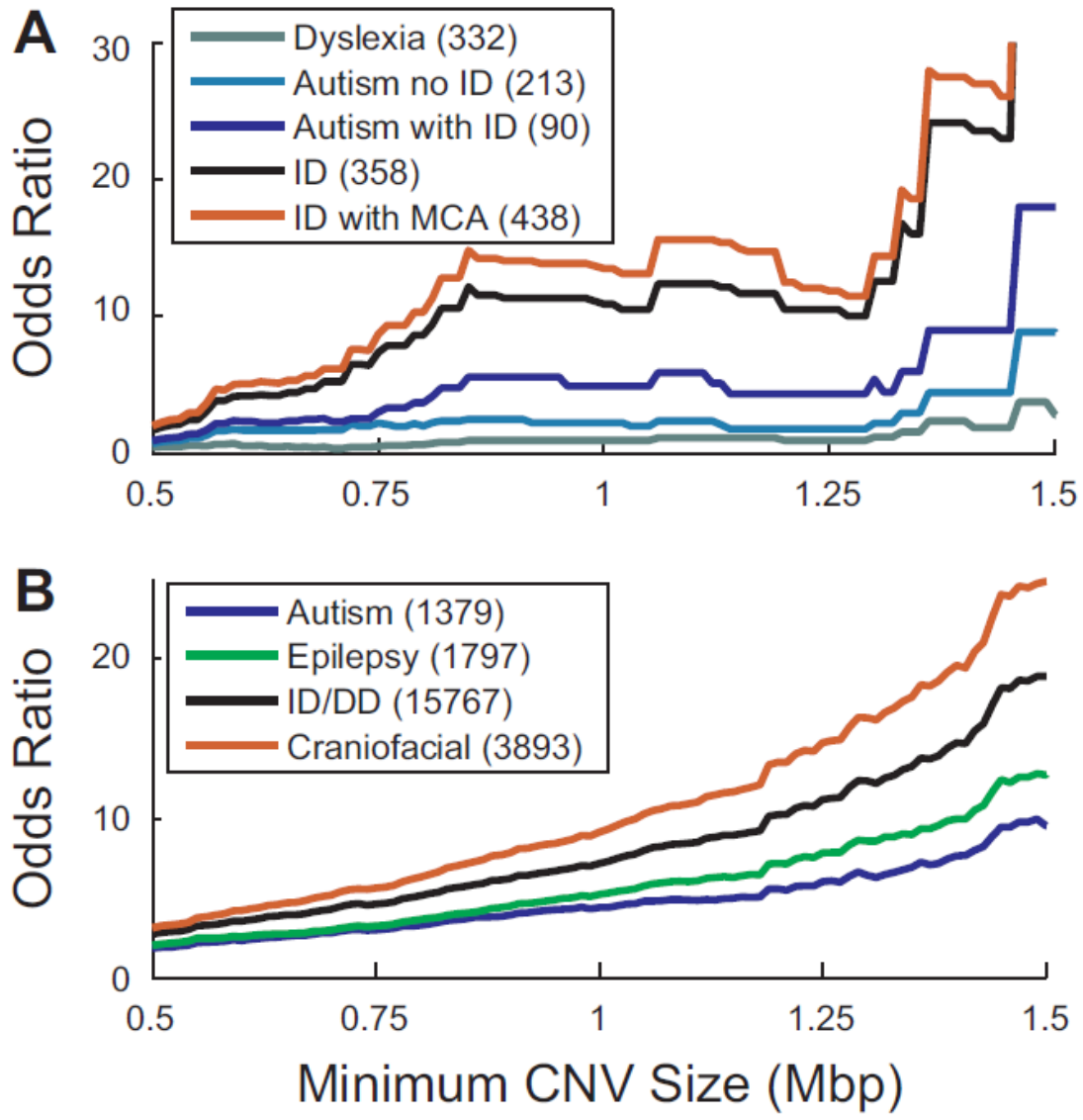
- May be due to deletions on the X chromosome in males, with resulting structural and functional nullisomy

- May encompass ‘imprinted’ gene(s) and therefore result in different phenotypes depending on whether the deletion or duplication involves the paternal or maternal homologues

DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources)

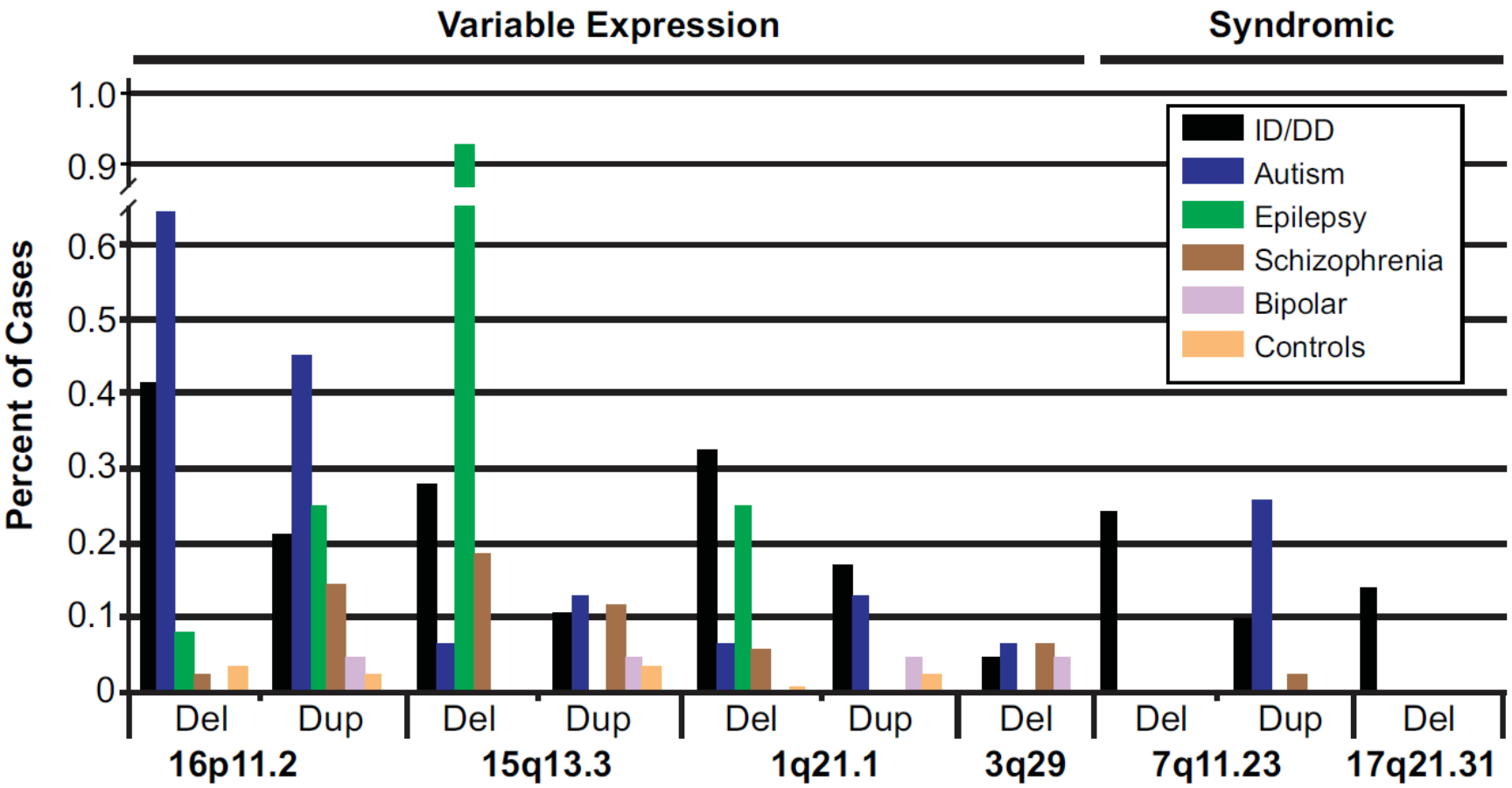
(70 DECIPHER microdeletion – microduplication syndromes reported to date)

CNV burden across various neurodevelopmental phenotypes



Coe BP et al. Am J Med Genet C Semin Med Genet 2012;160C:118-29

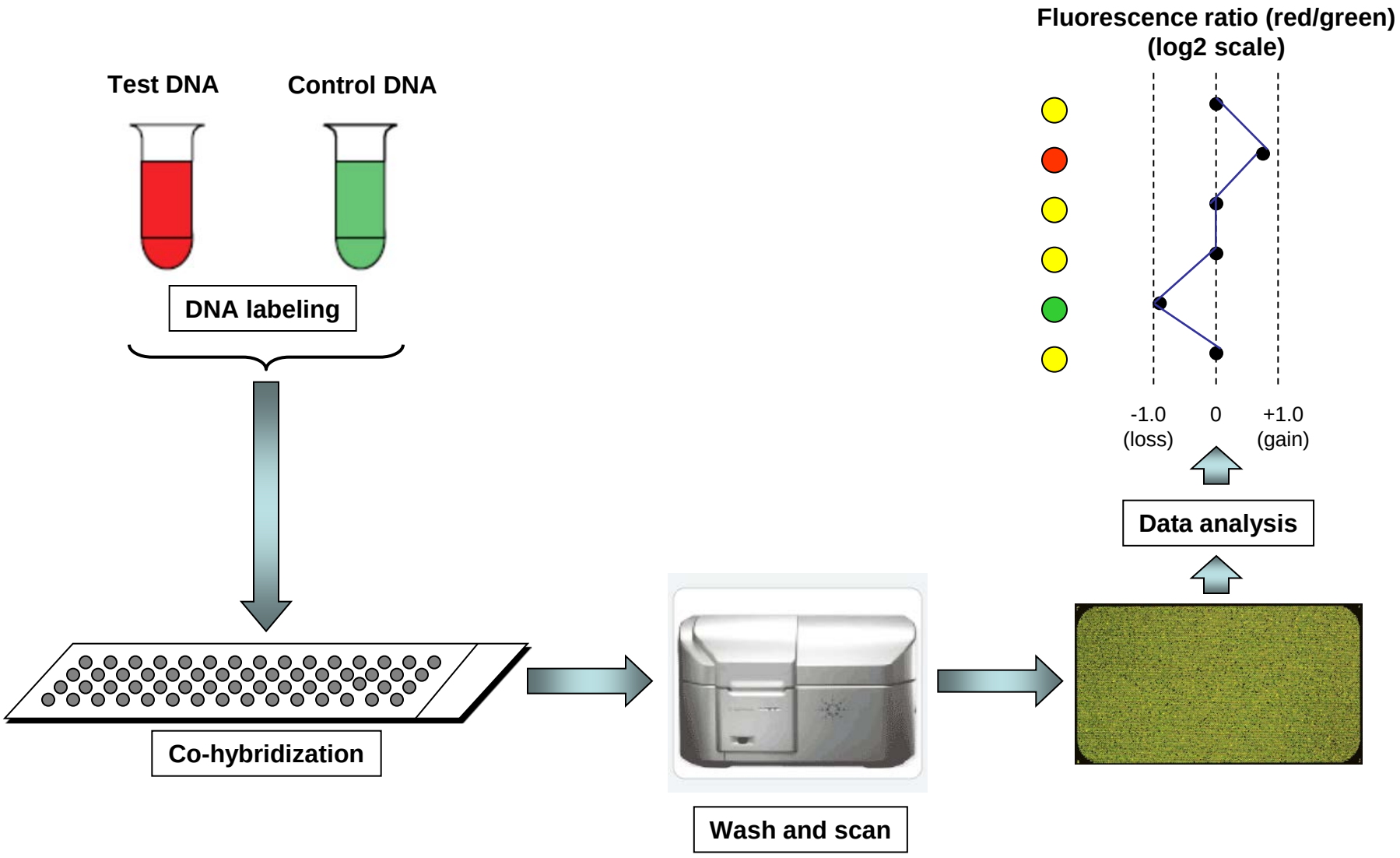
Variable expressivity of hotspot CNVs



The frequency of CNV deletions and reciprocal duplications for six genomic hotspots associated with neurological disease are shown (ID/DD, autism, epilepsy, schizophrenia, and bipolar disorders).

Coe BP et al. Am J Med Genet C Semin Med Genet 2012;160C:118-29

Array Comparative Genomic Hybridization (array CGH)

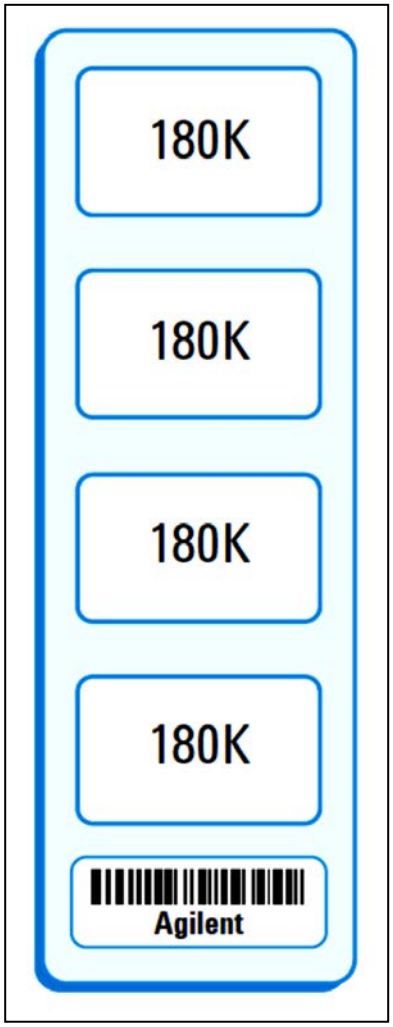


Cytogenomic array methodologies

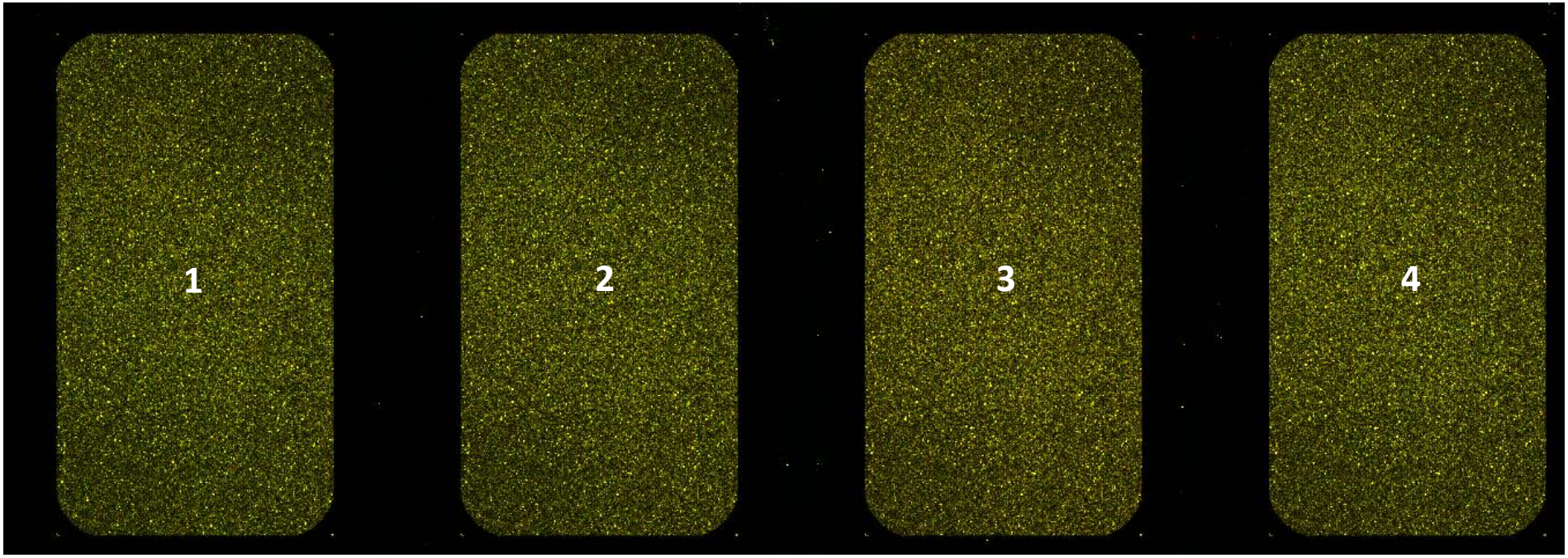
aCGH+SNP arrays	SNP arrays
Single-sequence oligonucleotides of ~60 bp	Two 20–60 bp oligonucleotides of different sequence
Two labeled DNAs (patient and control) per hybridization	Only patient DNA labeled and hybridized
Resolution down to size of oligonucleotides; exon by exon	Resolution limited by SNP distribution
Able to detect consanguinity, most uniparental disomy (UPD), and copy-neutral loss of heterozygosity (cnLOH)	Able to detect consanguinity, most UPD, and cnLOH at higher resolution

The **Agilent 4x180k aCGH+SNP array** platform adds SNP detection to the robust aCGH platform, allowing the rapid and reliable identification of both copy number changes and copy-neutral loss of heterozygosity (cnLOH)

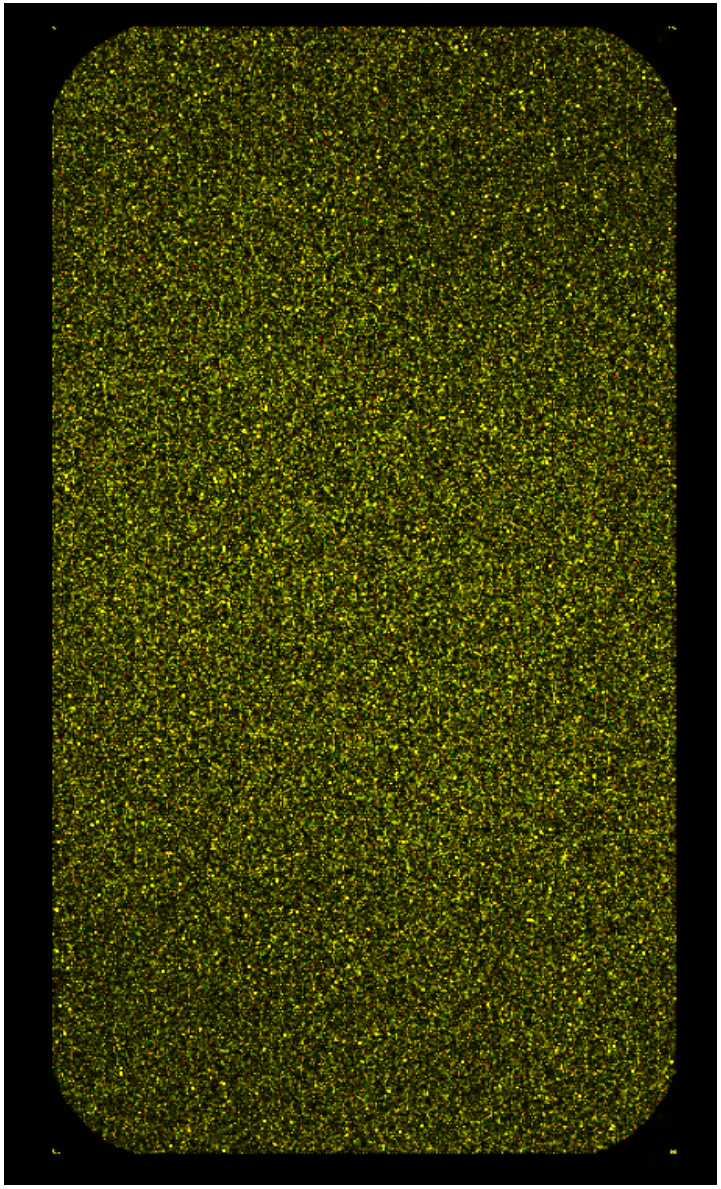
	Agilent 4x180k aCGH+SNP array
Number of copy number change probes	~120,000
Median copy number change probe spacing	~25 kb
Copy number change resolution	~50 kb
Number of SNP probes	~60,000
Copy-neutral LOH resolution	~5 Mb



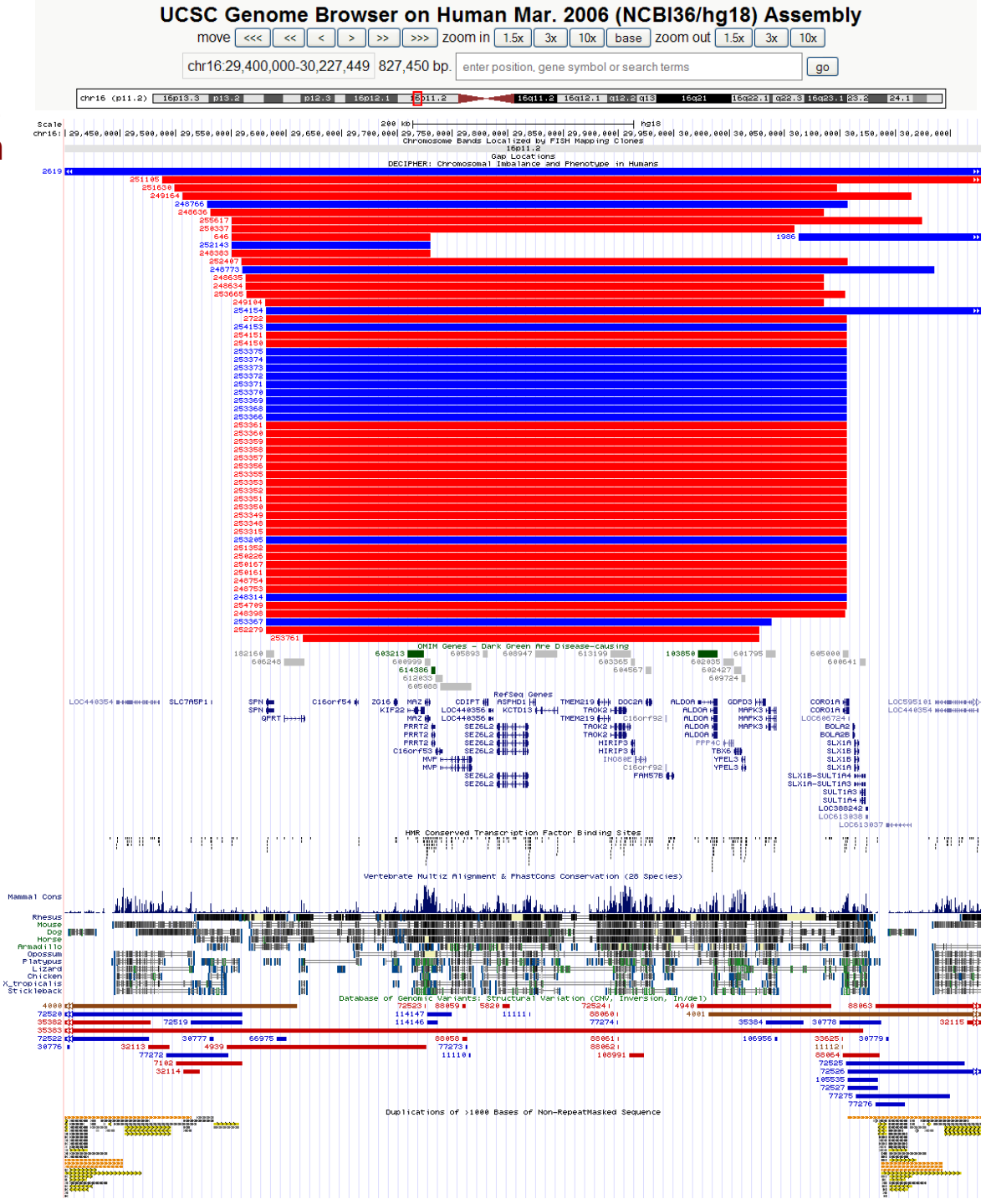
Agilent 4x180k aCGH+SNP array



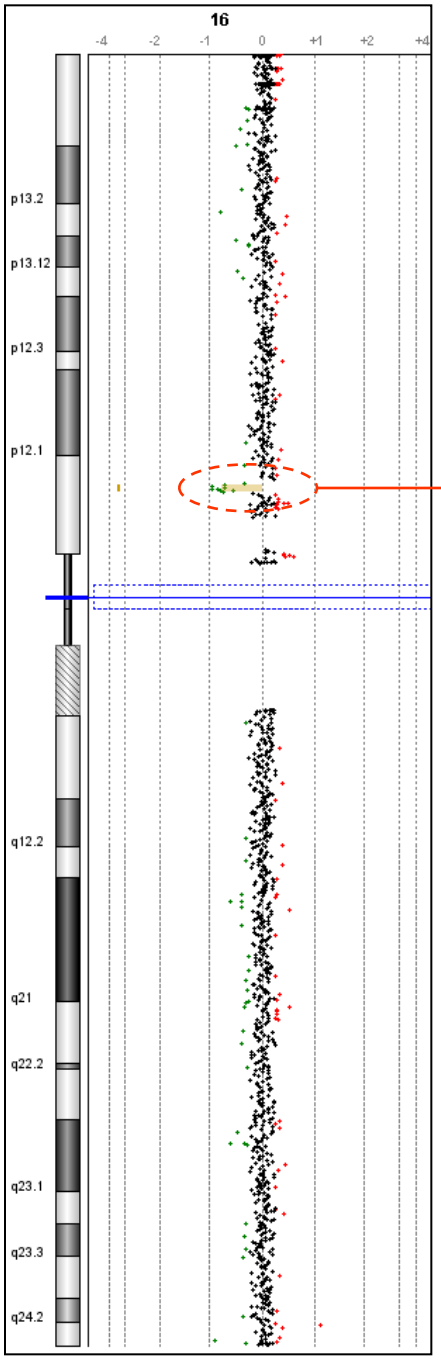
Agilent 4x180k aCGH+SNP array
sub-array 2

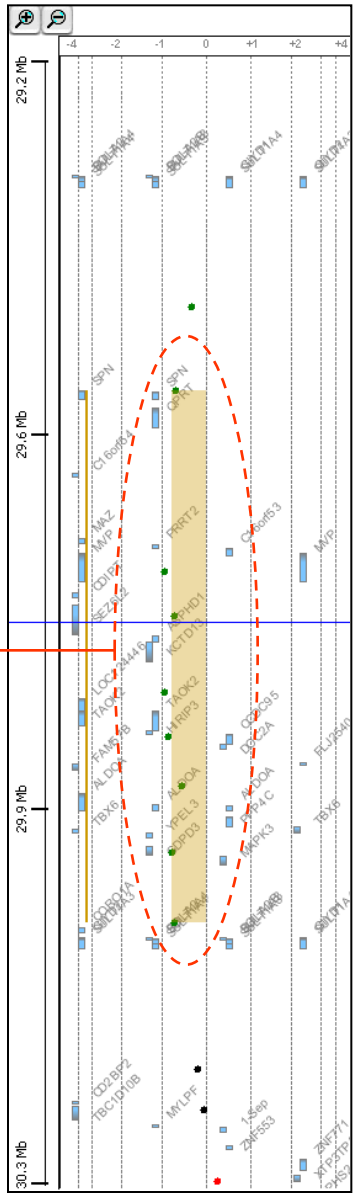


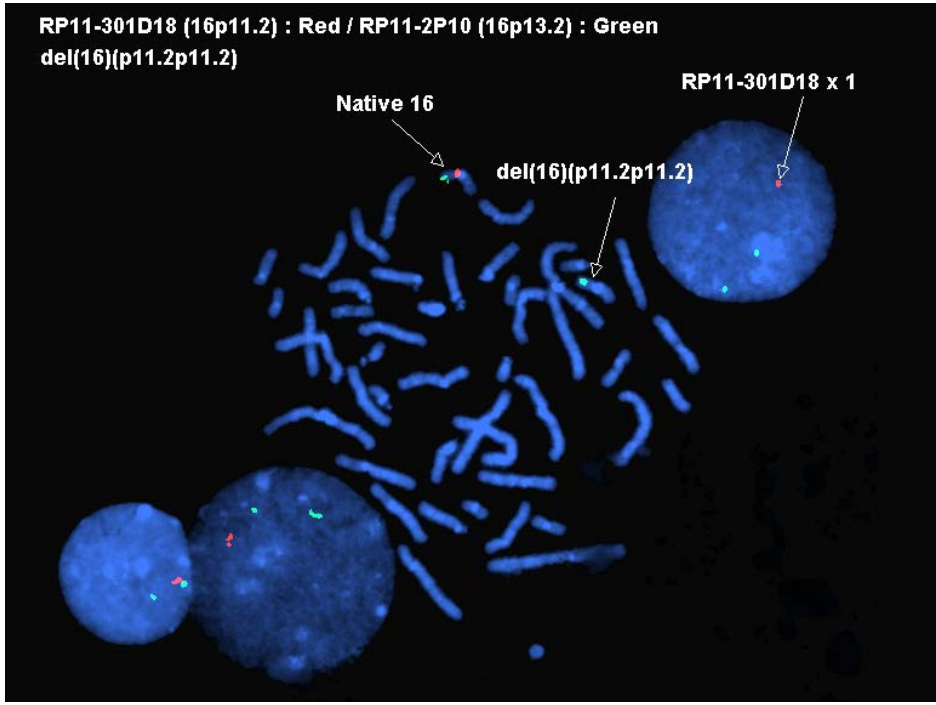
Constitutional recurrent CNVs at the 16p11.2 region



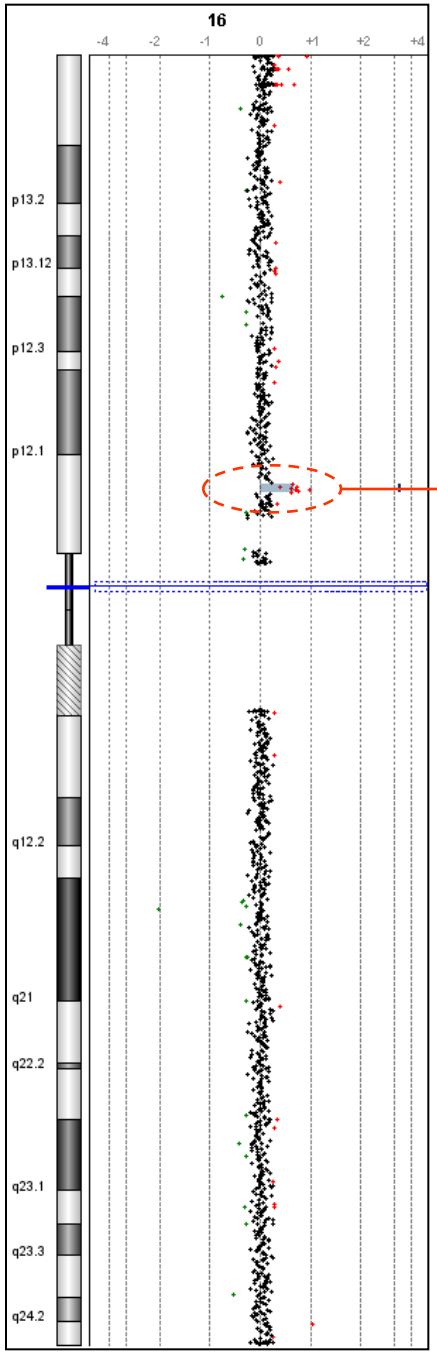
Constitutional 16p11.2 microdeletion

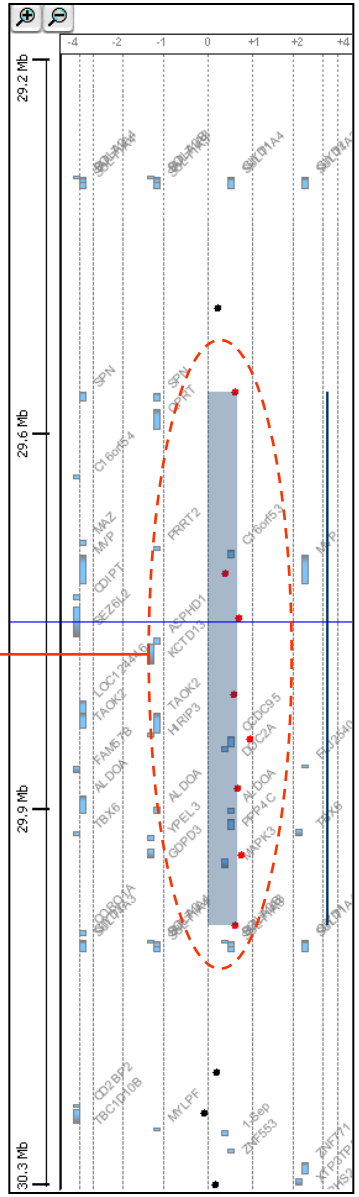


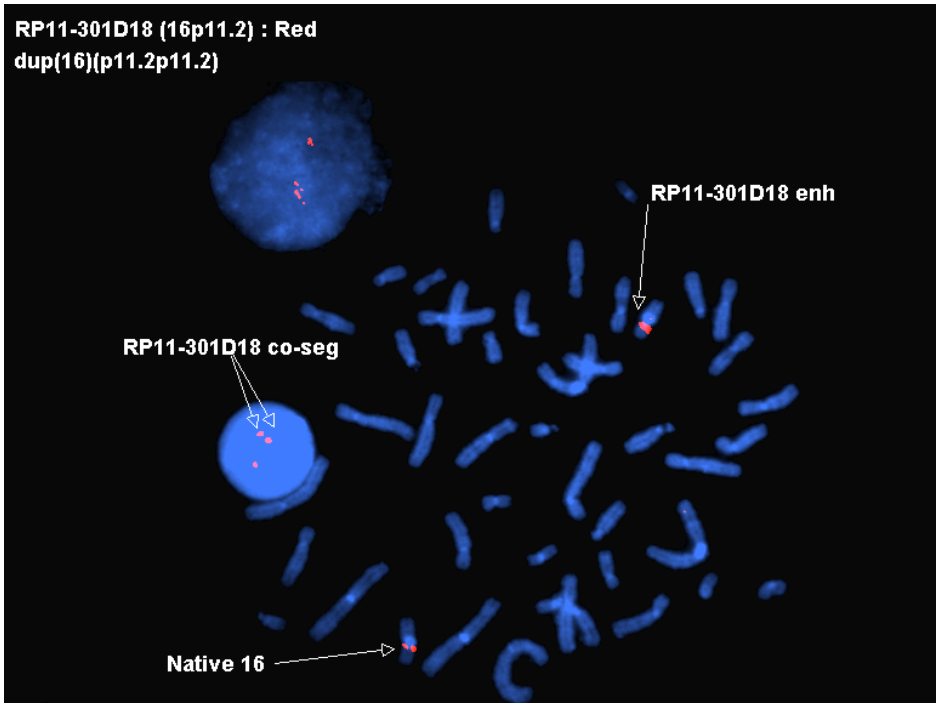




Constitutional 16p11.2 microduplication







CNV Databases

- Database of Genomic Variants: http://projects.tcag.ca
- USCS Genome Browser: http://www.genome.ucsc.edu/cgi-bin/hgGateway
- Ensembl Database: http://useast.ensembl.org/Homo_sapiens/Info/Index
- NCBI Map Viewer: http://www.ncbi.nlm.nih.gov/projects/mapview/
- DECIPHER Database: http://decipher.sanger.ac.uk/
- ISCA Consortium: https://www.iscaconsortium.org/

Conclusions

- CNVs are widespread in the human genome and are a significant source of human genetic variation accounting for population diversity and human disease

- High-resolution Cytogenomic array is a powerful and efficient method (in both clinical and research settings) for detecting constitutional pathogenic CNVs in patients with DD, ID, ASD, and MCAs

- Clinical high-resolution Cytogenomic array has proven to have an ~20% overall detection rate of pathogenic CNVs in these patients

- A specific genetic diagnosis in these cases facilitates comprehensive medical care and accurate recurrence risk counseling for the family

Use of Cytogenomic arrays in studying cancers

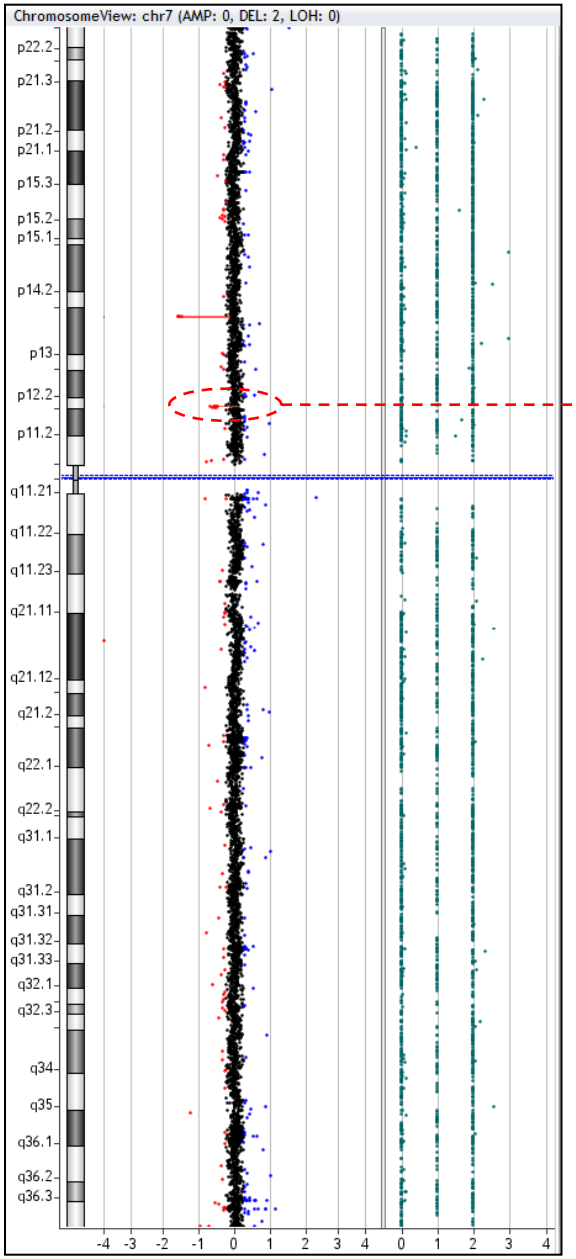
- Cytogenetic analyses, including conventional karyotyping and targeted FISH analyses, of malignancies are routinely performed to detect recurrent chromosomal abnormalities that have diagnostic, prognostic, and therapeutic implications

- However, the genetic complexity of cancer cells requires higher resolution genome-wide analysis to enable the detection of small genomic changes

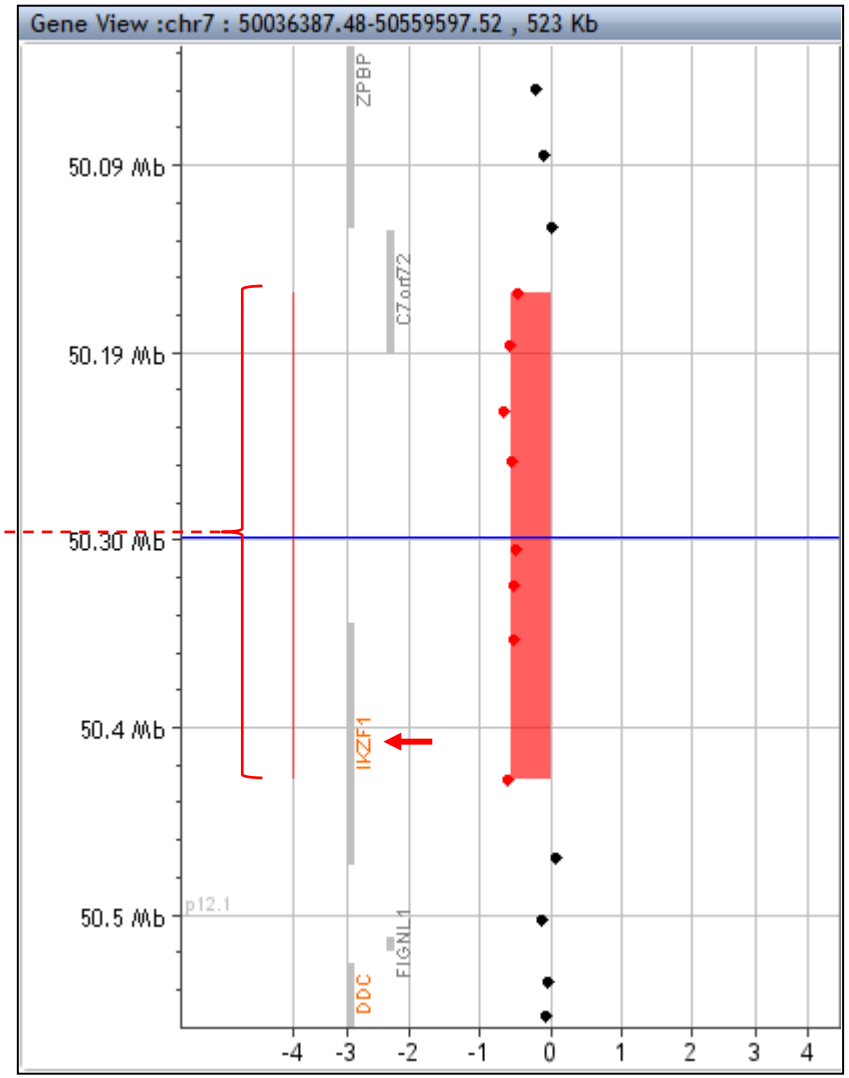
- The advent in Cytogenomic array methodologies, including aCGH and SNP arrays, have overcome many of the limitations of traditional cytogenetic techniques in picking up small clinically significant copy number changes in various malignancies

- The clinical utility of genome-wide Cytogenomic arrays in cancer diagnostics is growing rapidly. This technology is evolving into a diagnostic tool, to better identify high-risk patients and predict clinical outcomes

Somatic *IKZF1* gene (7p12.2) deletion in B-precursor ALL

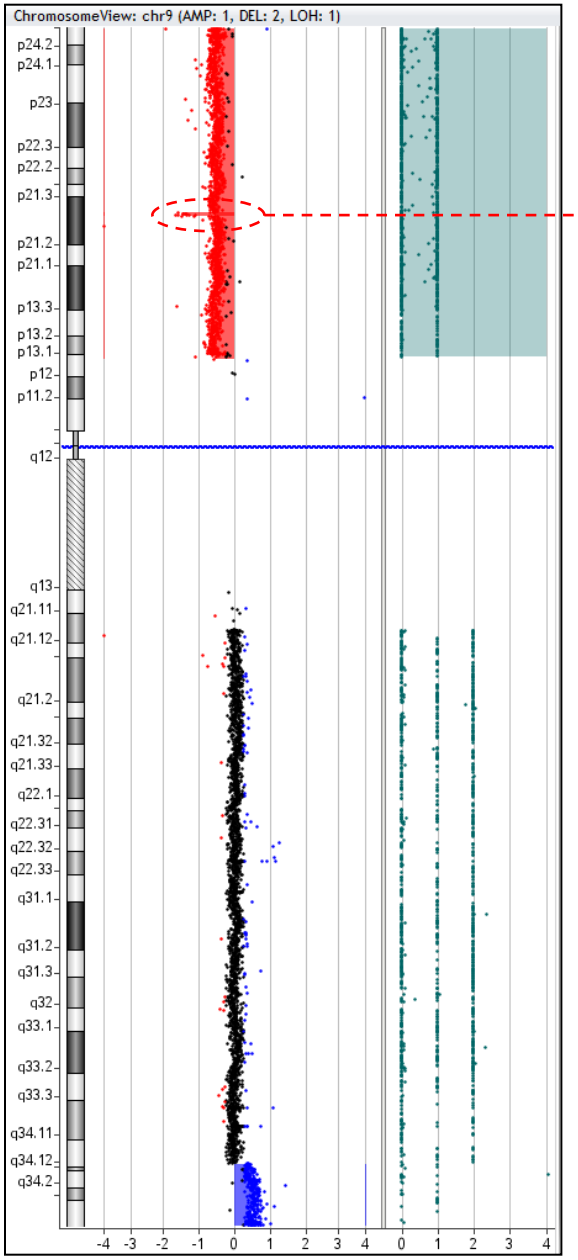


Chromosome 7

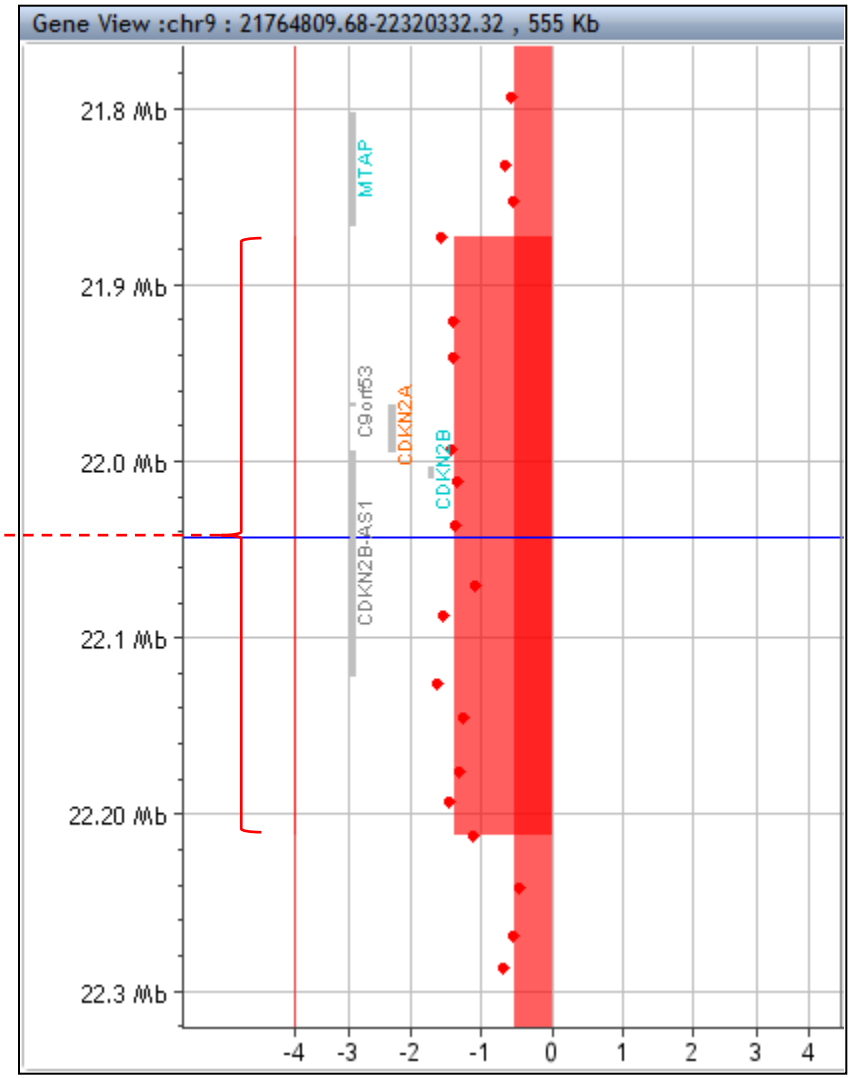


IKZF1 (7p12.2) deletion (~260 kb)
(a poor prognostic sign)

Somatic *CDKN2A* gene (9p21.3) deletion in B-precursor ALL

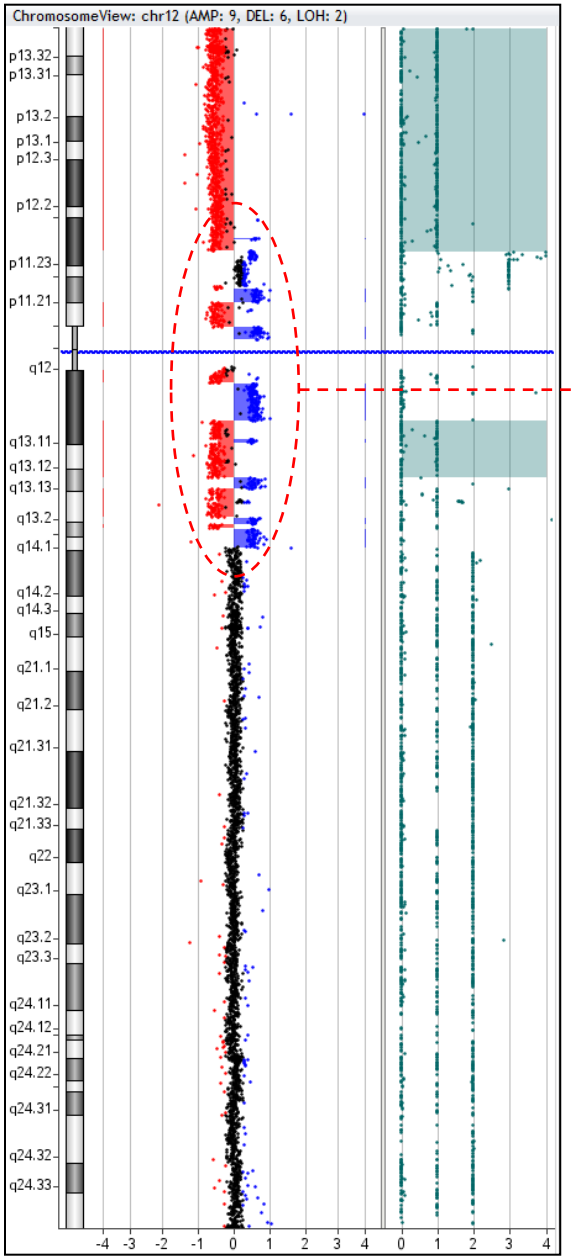


Chromosome 9

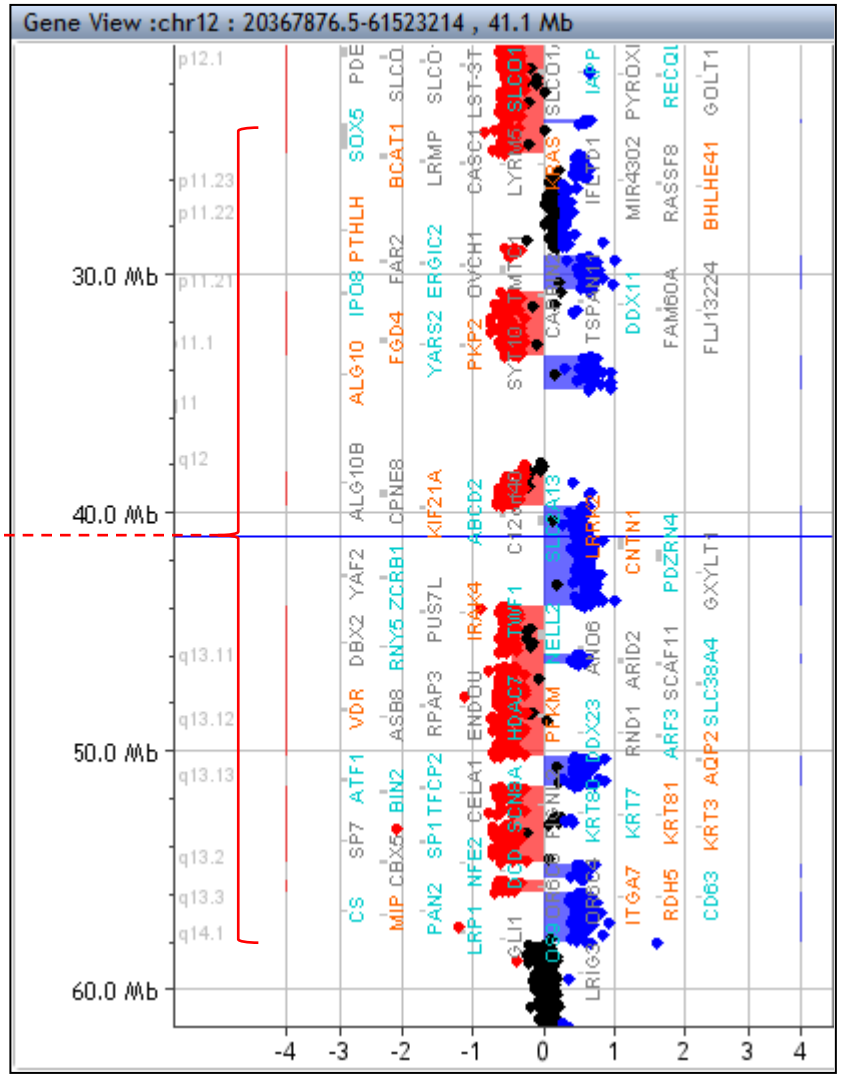


Nested homozygous 9p21.3 deletion spanning *CDKN2A* (~340 kb)

Somatic chromosome 12 chromothripsis in B-precursor ALL

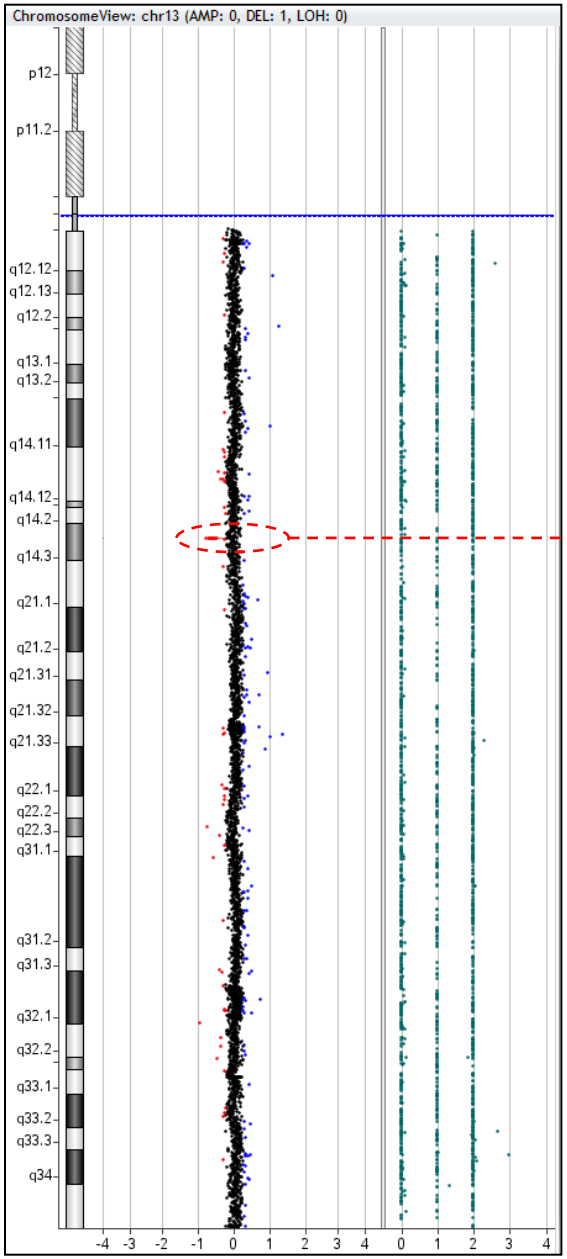


Chromosome 12

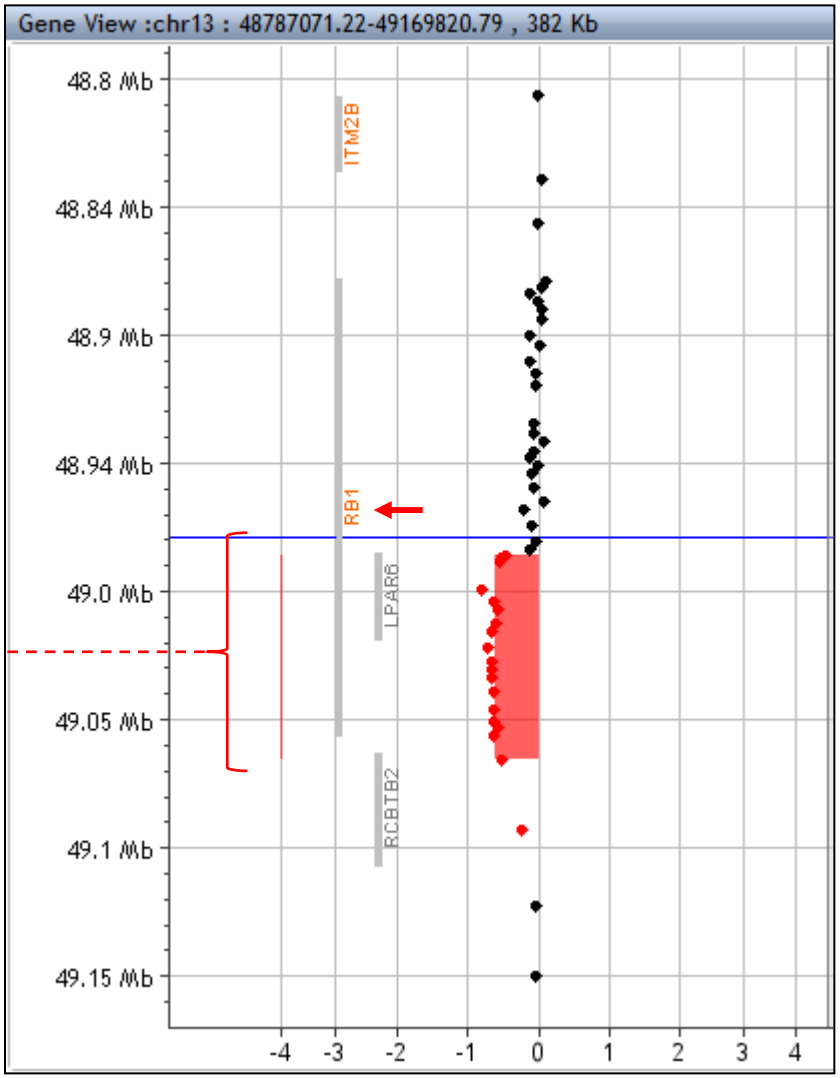


Chromosome 12 chromothripsis
(a poor prognostic sign)

Somatic *RB1* gene (13q14.2) deletion in B-precursor ALL



Chromosome 13



RB1 (13q14.2) deletion (~79 kb)
(a poor prognostic sign)

Thank you