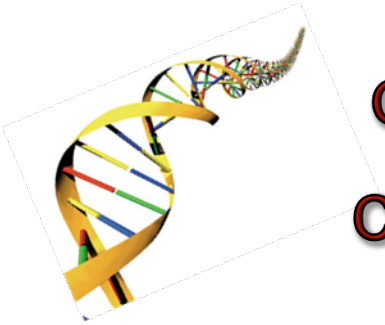


GENETICS AND GENOMICS IN CLINICAL RESEARCH COURSE

Copy Number Variations (CNVs)

October 1st 2013

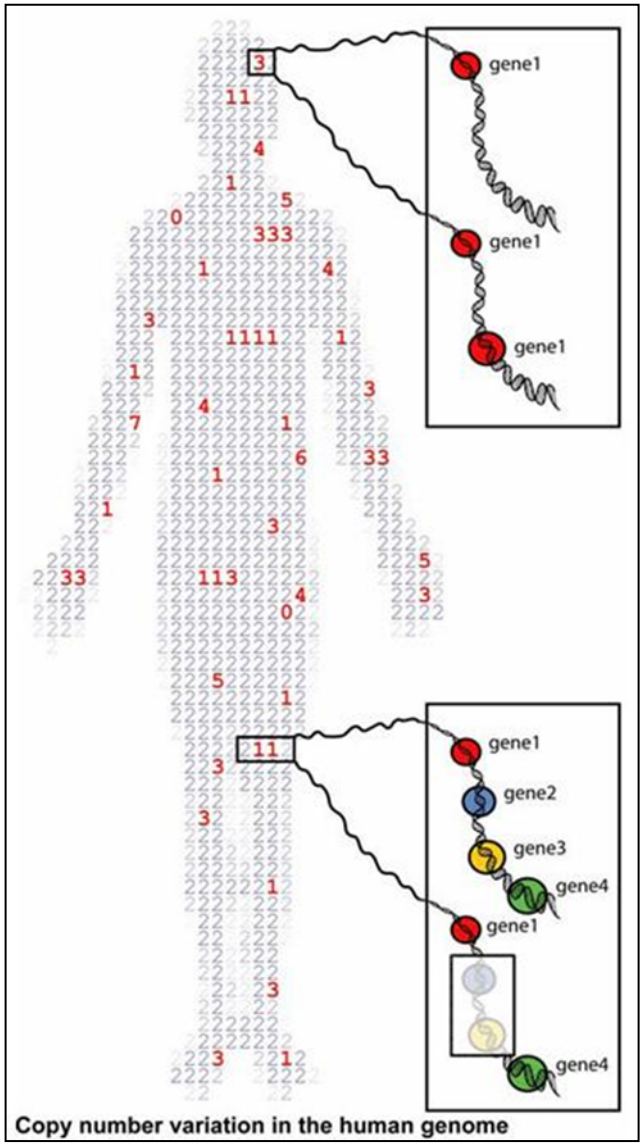
Fady M. Mikhail, MD, PhD

Associate Professor

Department of Genetics

Copy number variations (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 variation in the human genome

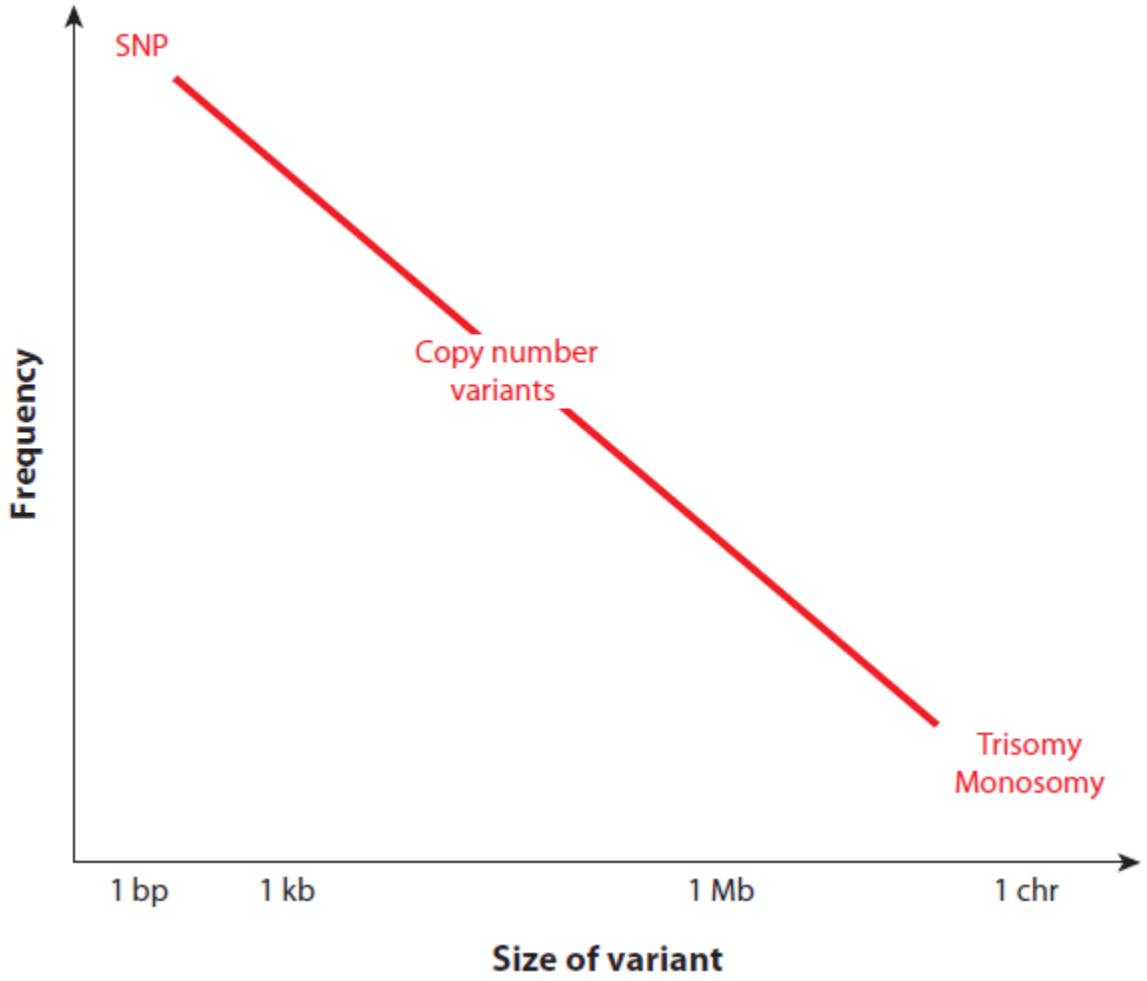
Copy number variations (CNVs) (cont'd)

- 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

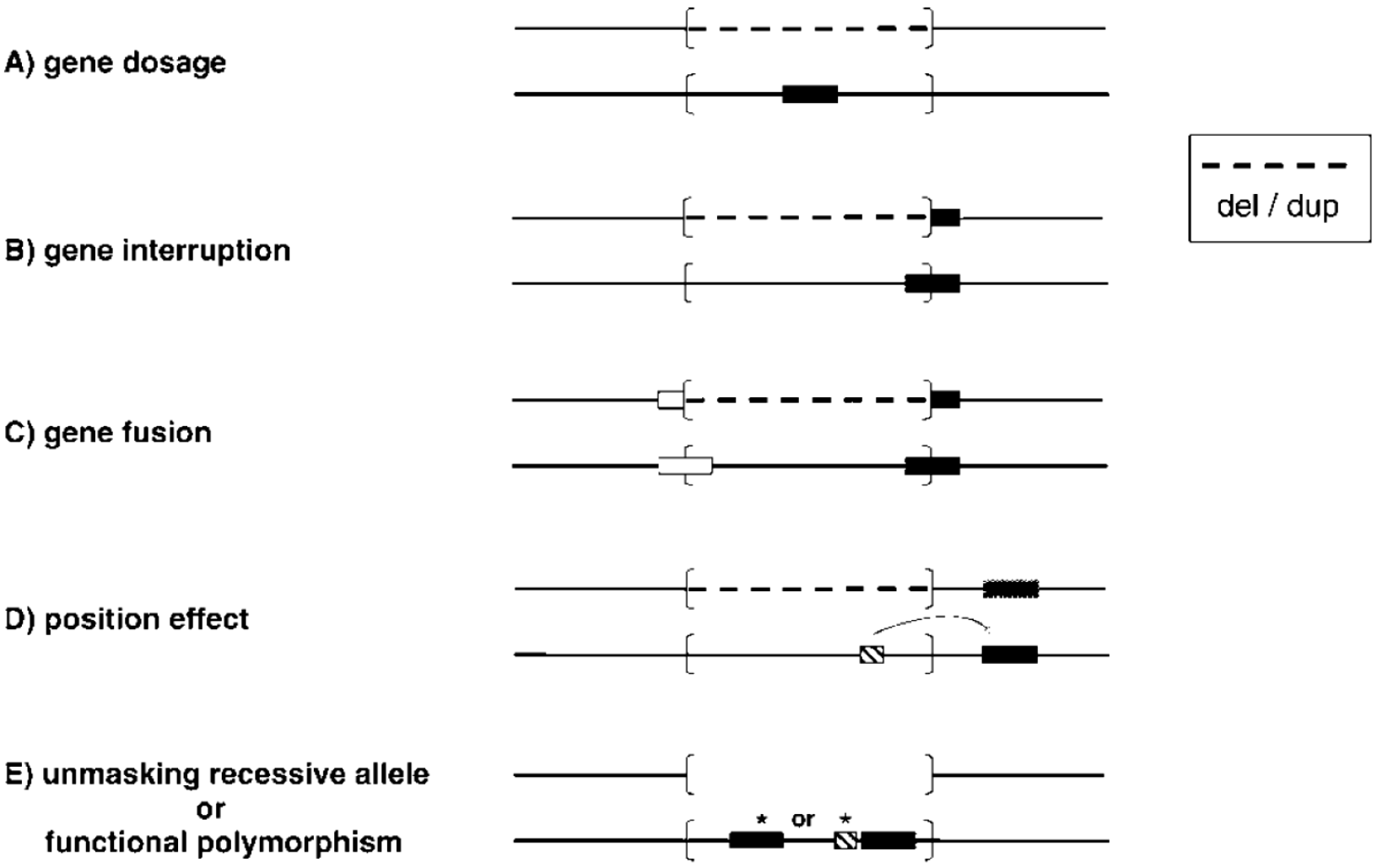
- A recent estimate of the proportion of the human genome that is structurally variant (i.e. benign CNVs) is in the order of ~5-10%

- The majority (>95%) of benign CNVs in humans are <100 kb in size

Can CNVs cause disease?

- Most CNVs are benign variants that will not directly cause disease
- CNVs that affect critical developmental genes can cause disease
- Recent reviews have listed 17 conditions of the nervous system alone – including Parkinson's Disease and Alzheimer's Disease – that can result from copy number variation
- Genes that are involved in the immune system and in brain development and activity – two functions that have evolved rapidly in humans – tend to be enriched in CNVs

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]

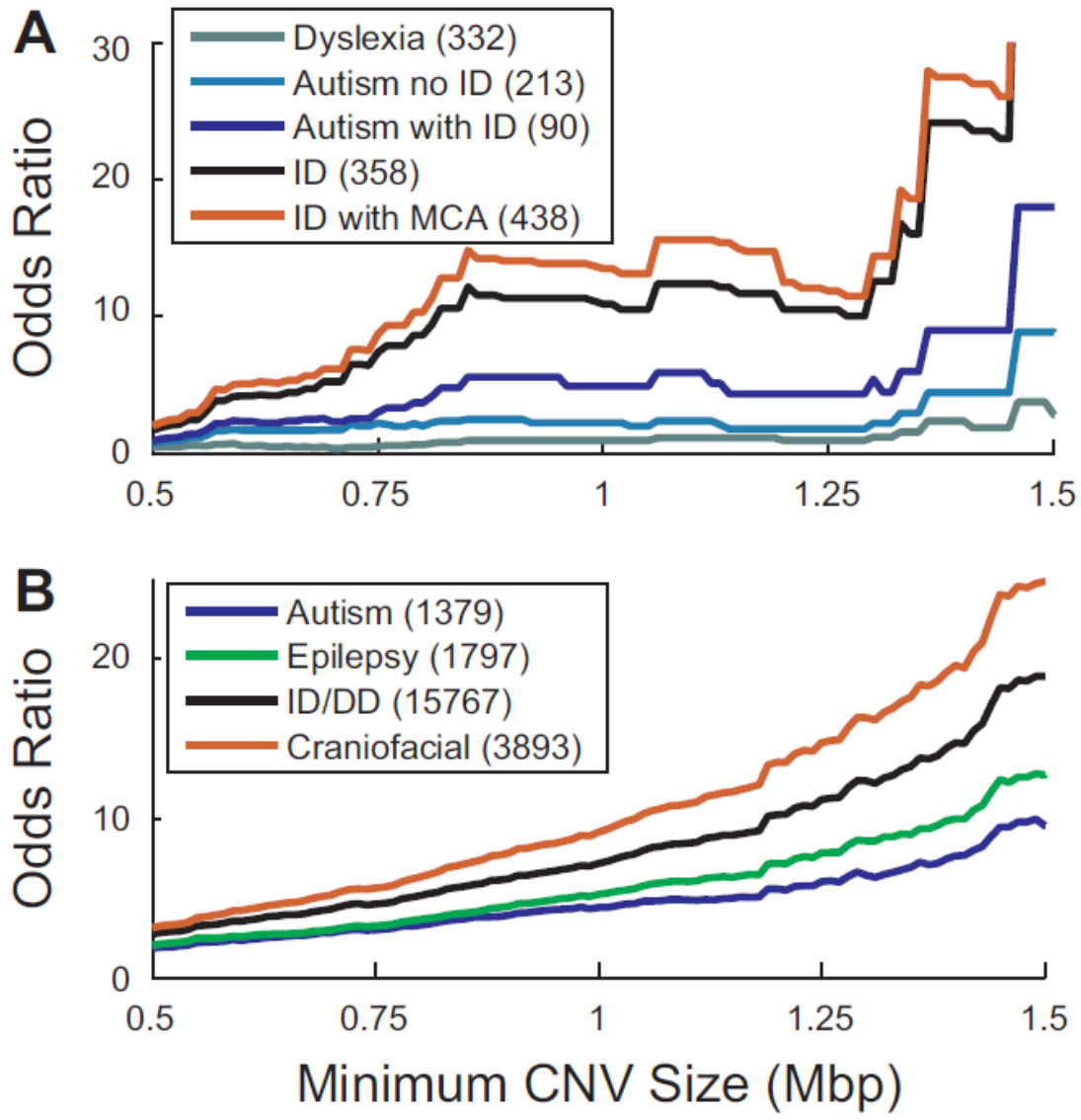
Primary Criteria		Indicates CNV Is Probably Pathogenic	Indicates CNV Is Probably Benign
1.	a. Identical CNV inherited from a healthy parent[b]		✓
	b. Expanded or altered CNV inherited from a parent	✓	
	c. Identical CNV inherited from an affected parent	✓	
2.	a. Similar to a CNV in a healthy relative		✓
	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		✓

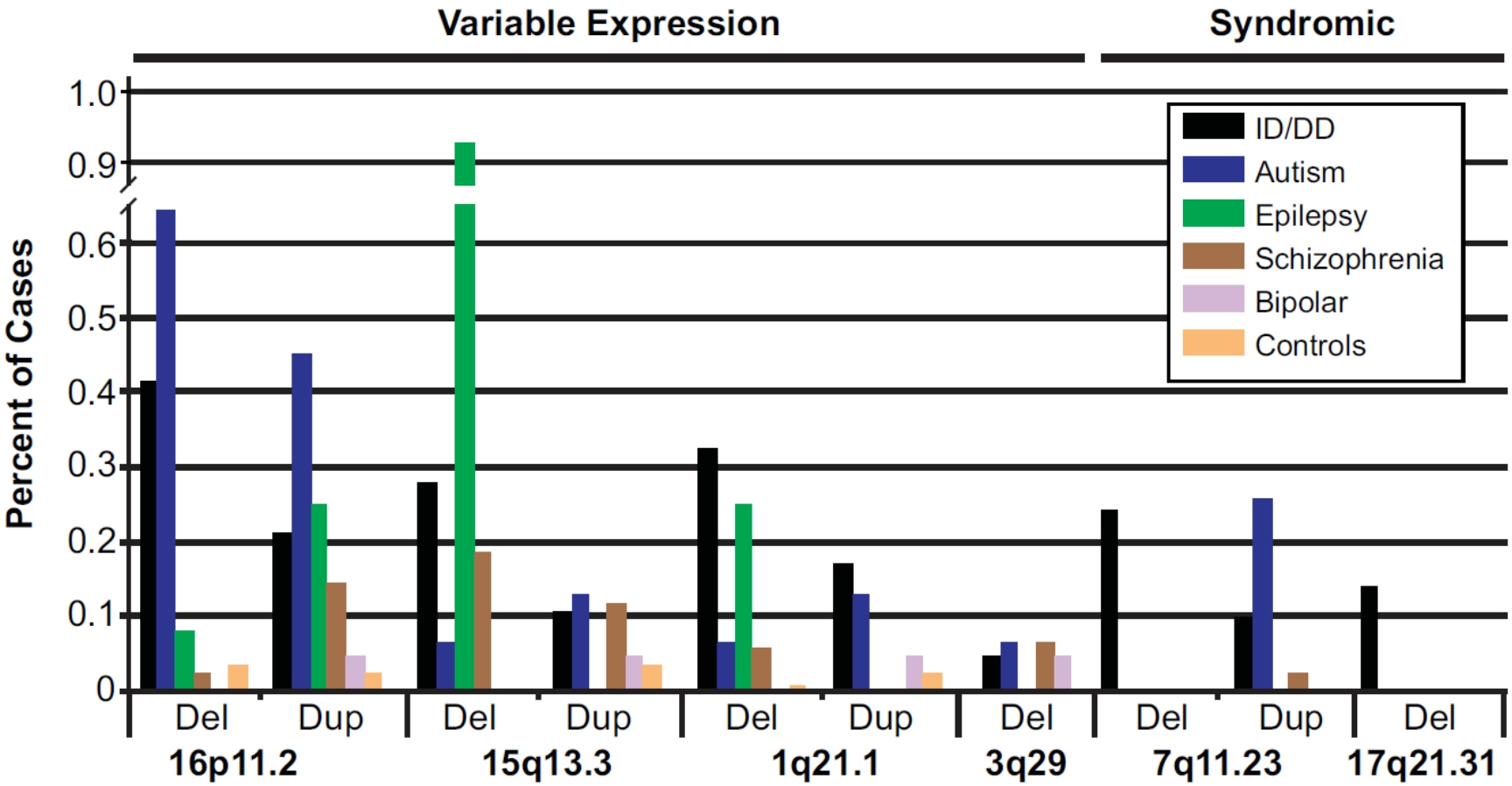
Miller DT et al. Am J Hum Genet 2010;86:749-64

CNV burden across various neurodevelopmental phenotypes



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

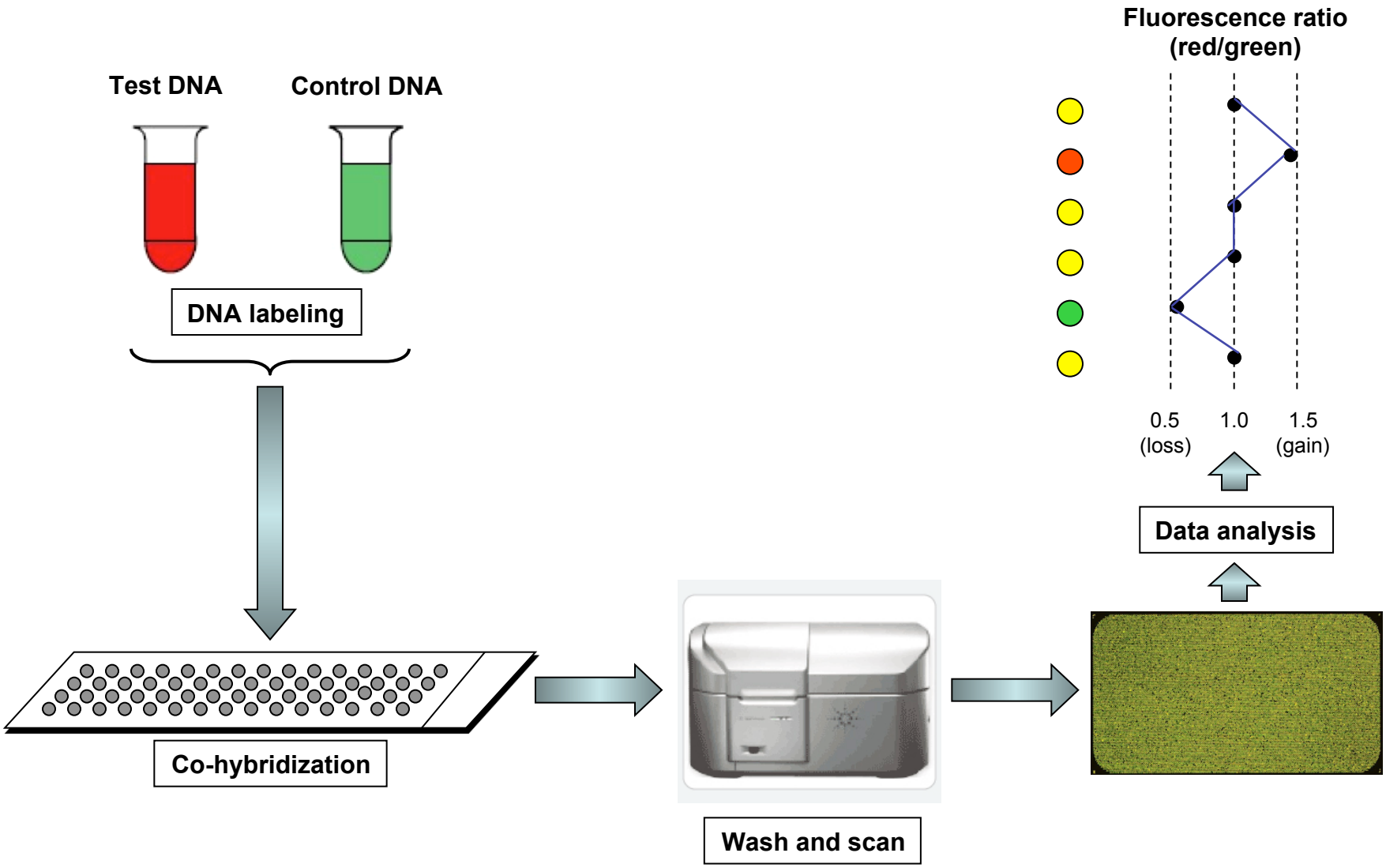
Variable expressivity of hotspot CNVs



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

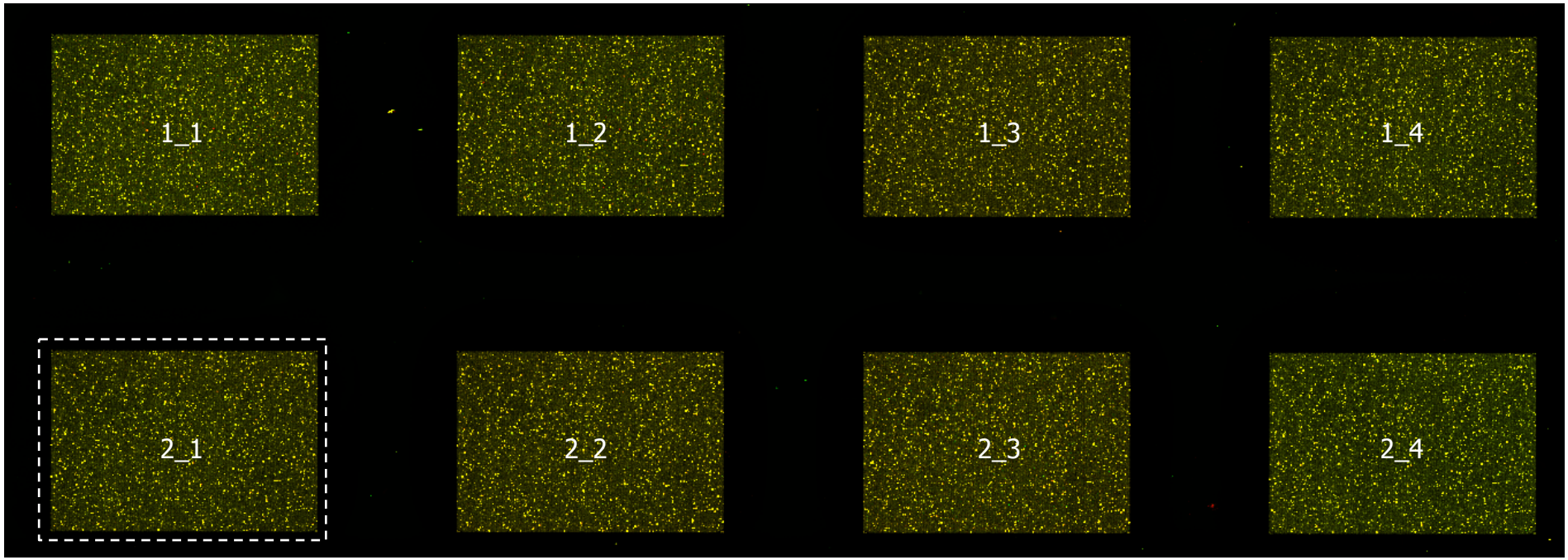
Array Comparative Genomic Hybridization (array CGH)



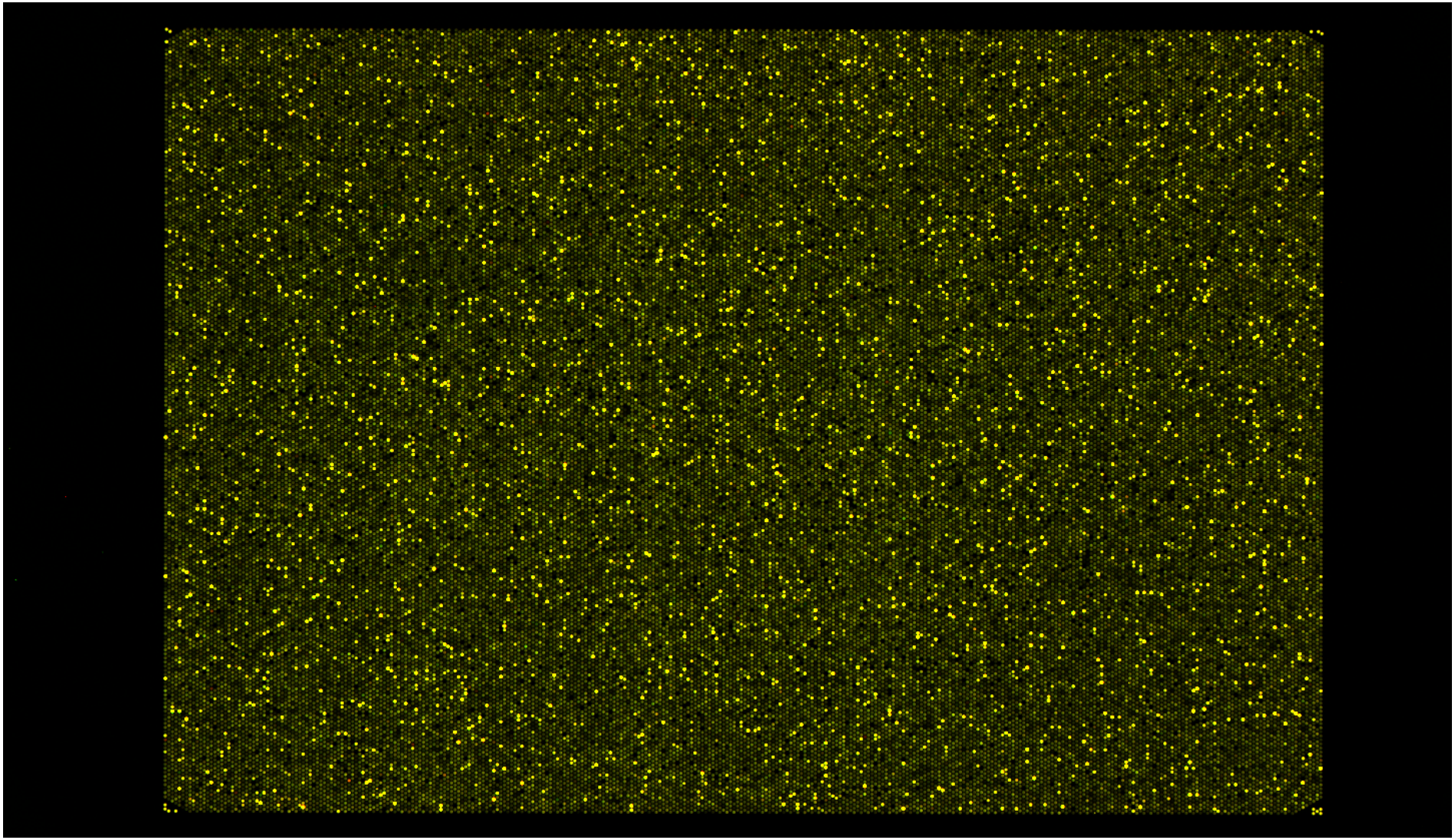
Cytogenomic array methodologies

Array CGH	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
No detection of UPD or consanguinity	Able to detect consanguinity and most UPD
Limited SNP addition possible recently	Detection of most known clinically relevant CNVs but not exon by exon

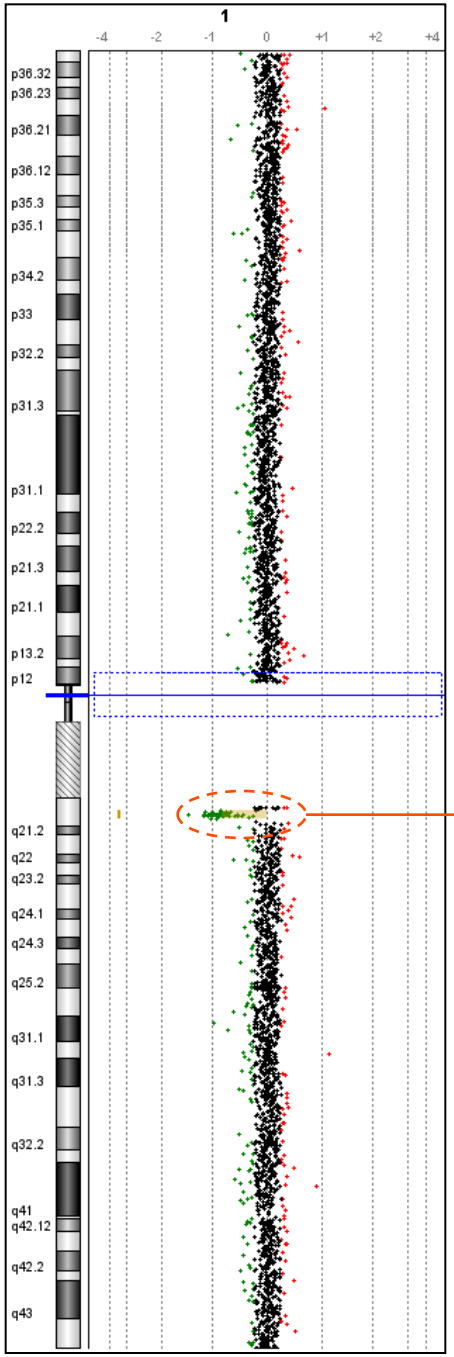
Agilent 8x60k array

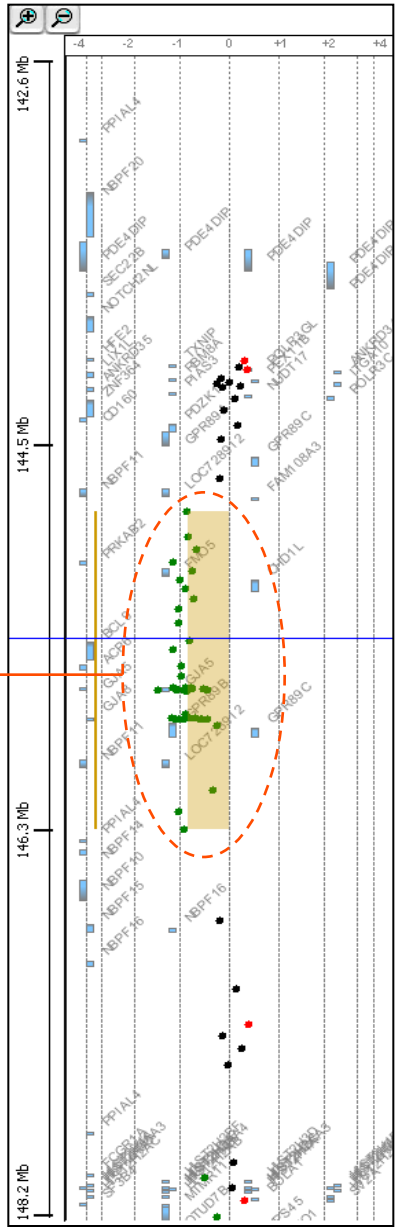


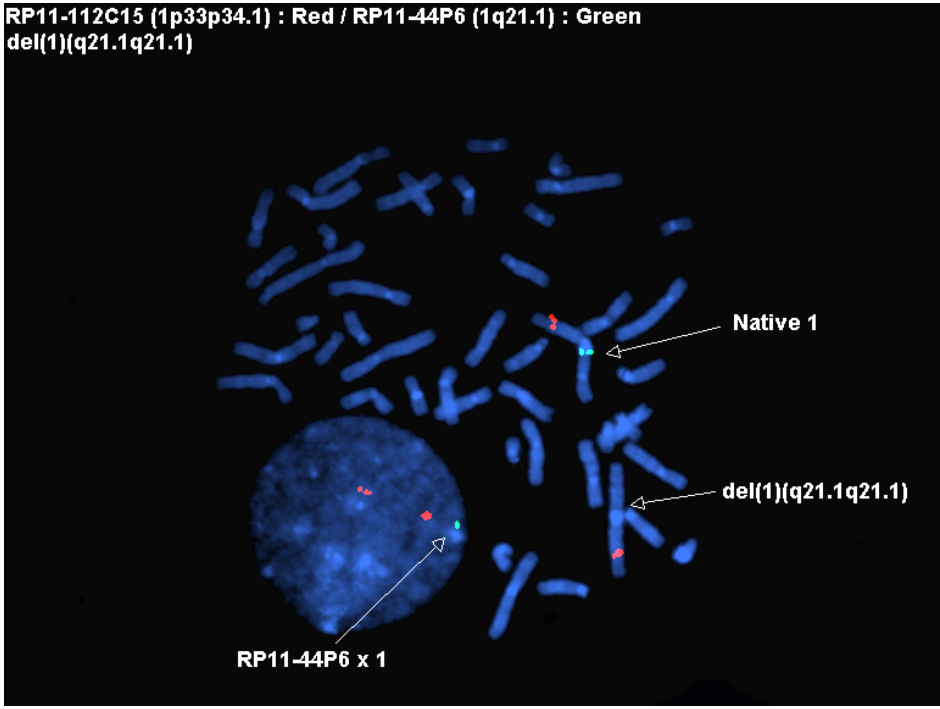
Agilent 8x60k array – subarray 2_1



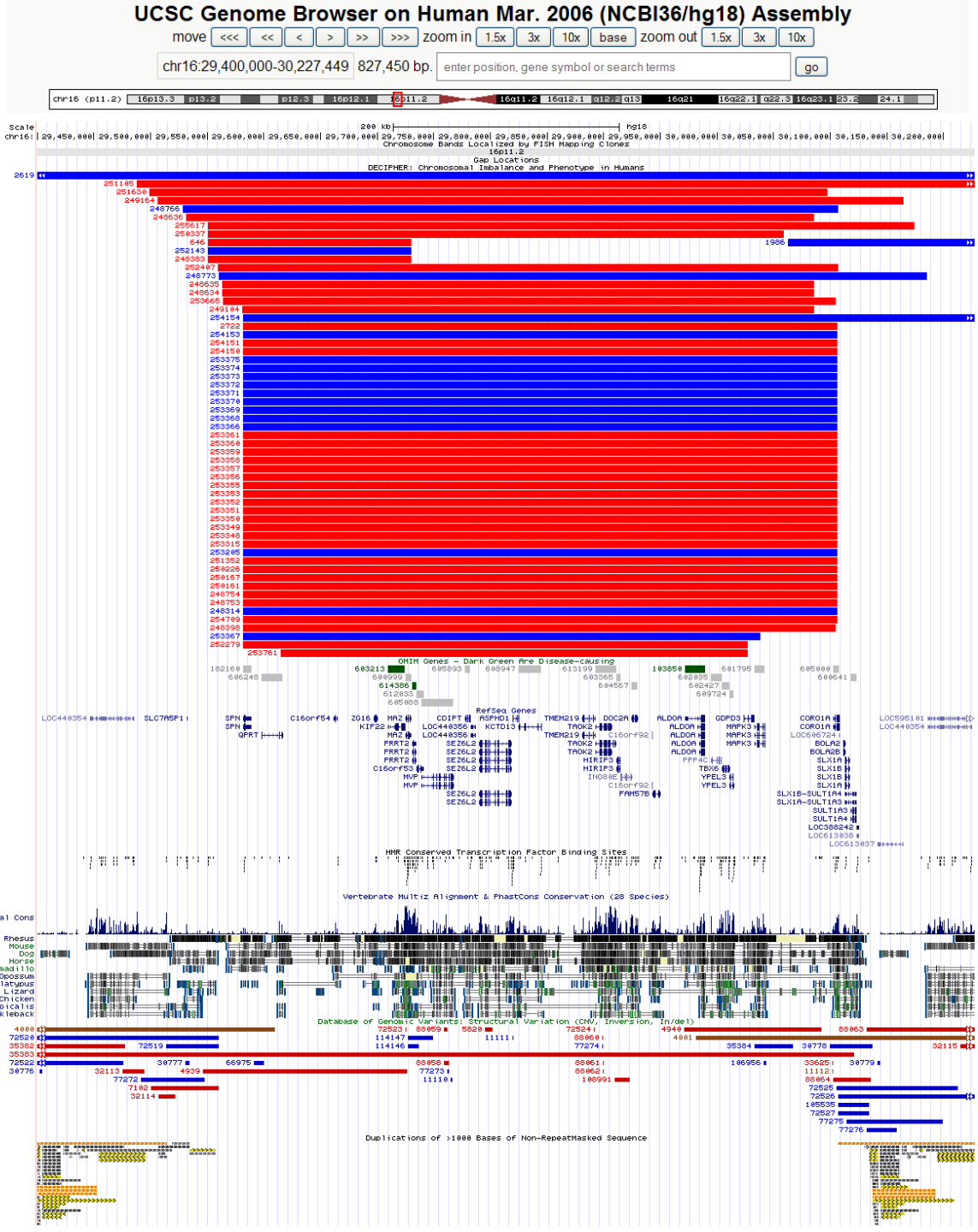
Distal 1q21.1 microdeletion



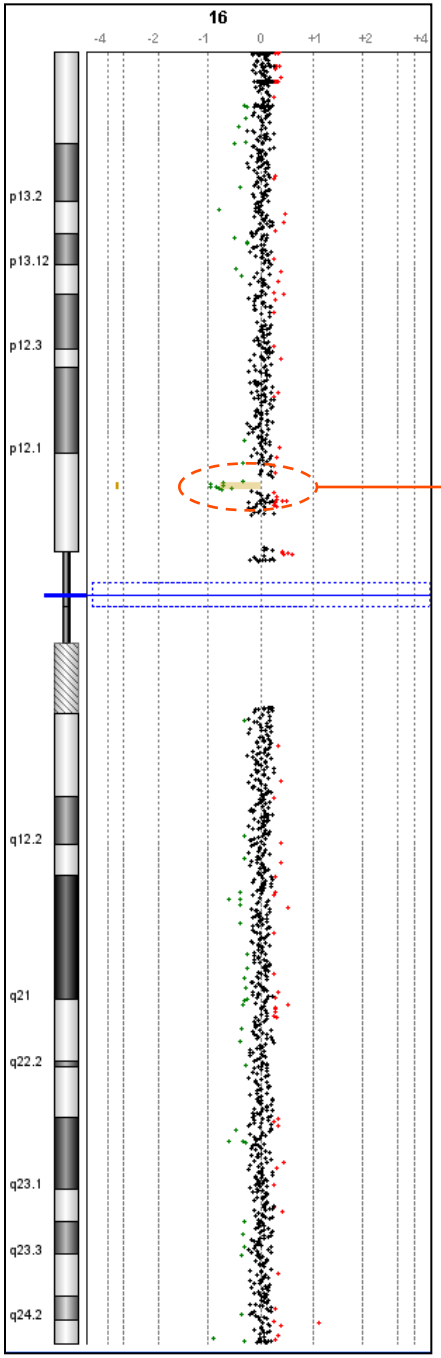


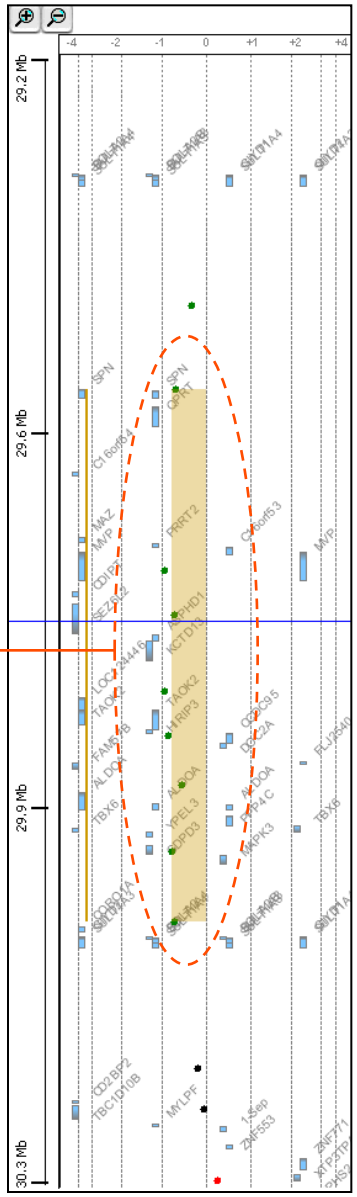


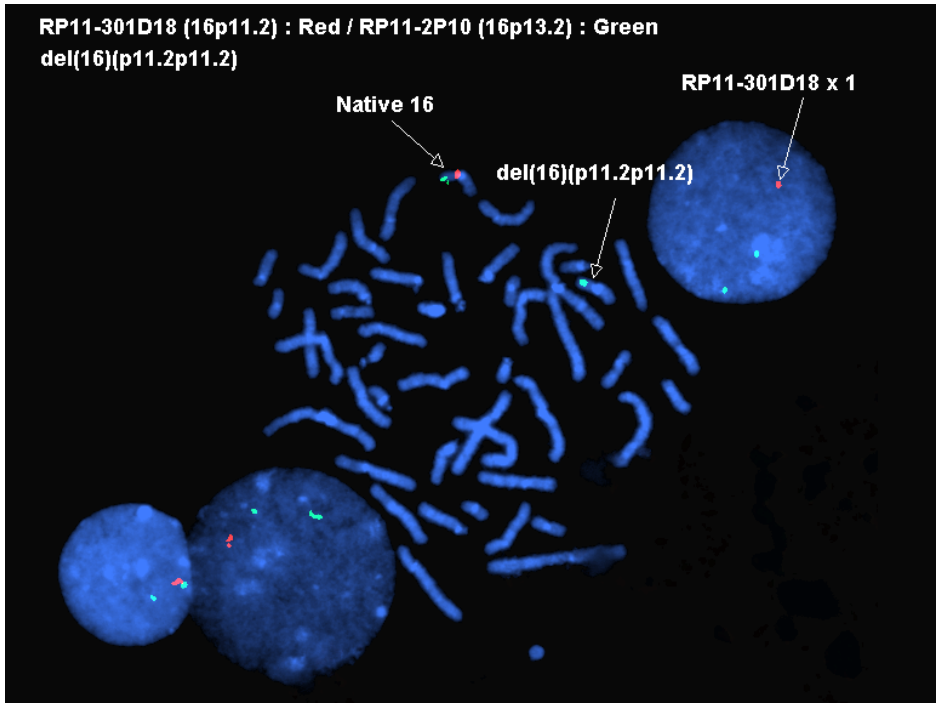
Recurrent CNVs at the 16p11.2 region



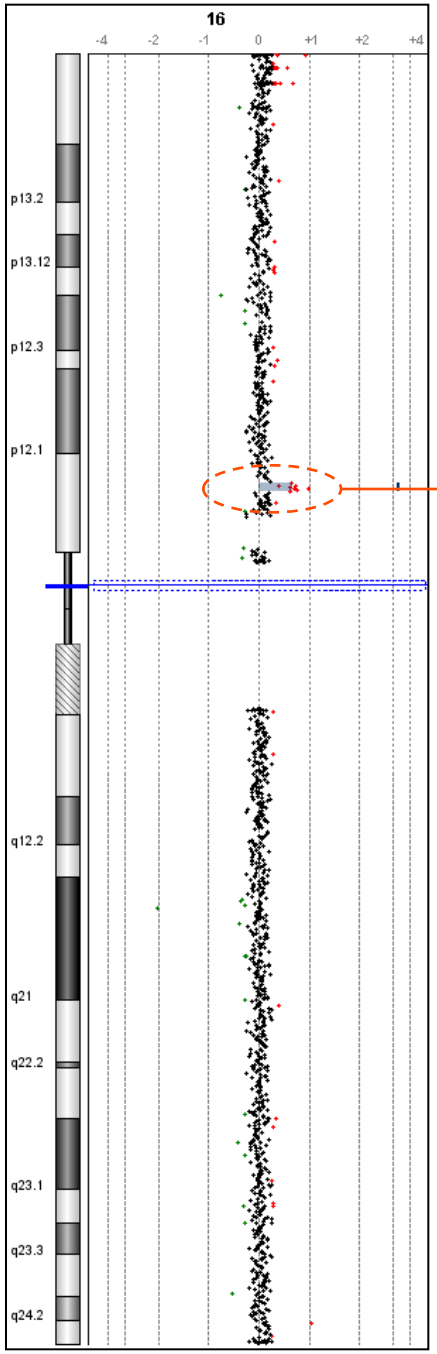
16p11.2 microdeletion

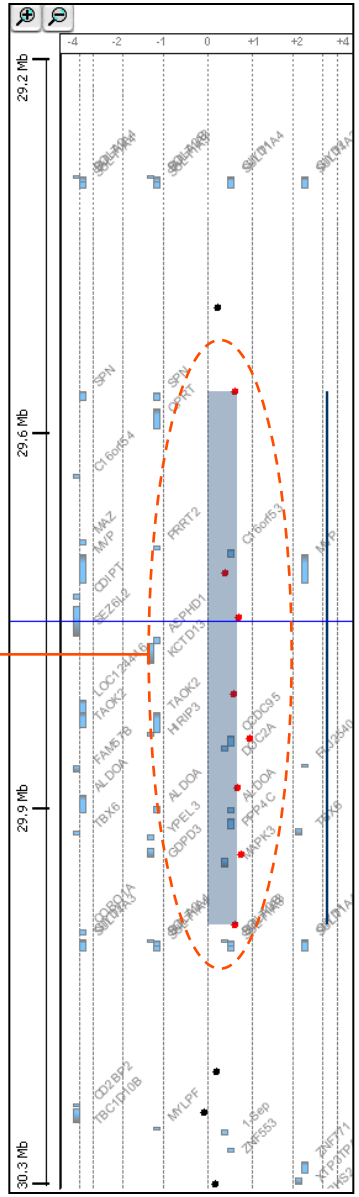


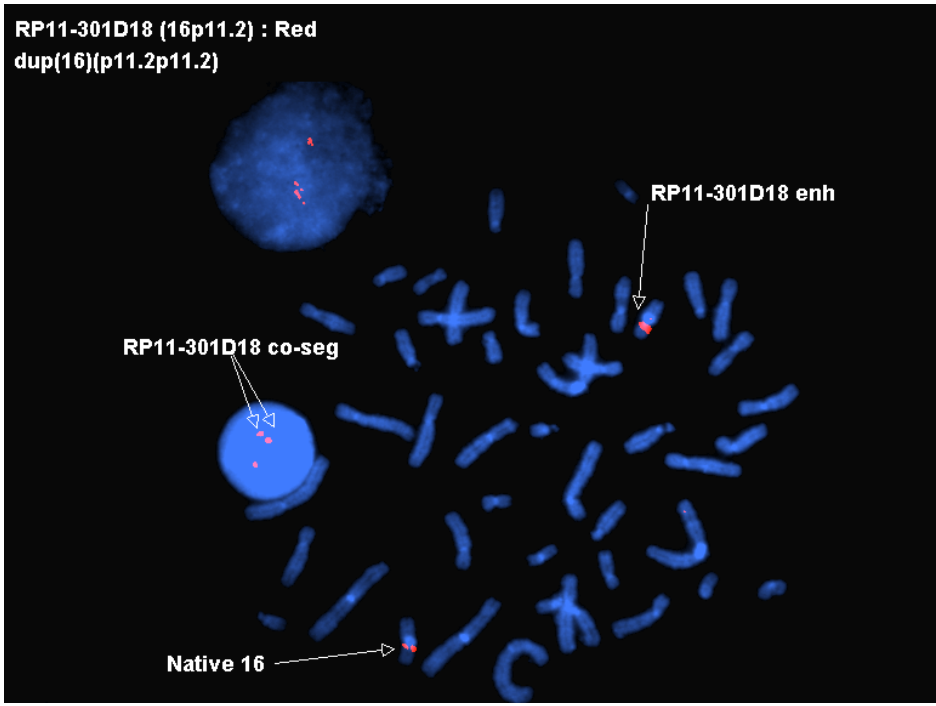




16p11.2 microduplication







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 pathogenic CNVs in patients with DD, ID, ASD, and MCAs
- Clinical high-resolution cytogenomic array has proven to have an ~15-20% overall detection rate of genomic rearrangements in these patients
- A specific genetic diagnosis in these cases facilitates comprehensive medical care and accurate recurrence risk counseling for the family