Yale School of Medicine

# Autism Genetics

*Matthew W. State, MD, PhD*
*Donald J. Cohen Associate Professor*
*Of Child Psychiatry, Psychiatry and Genetics*
*Deputy Chairman for Research,*
*Department of Psychiatry*
*Yale University School of Medicine*
*New Haven CT,*
*Matthew.State @yale.edu*

# 2. Outline

- Why genetics?
- Why has it been so hard to find genes involved in ASD?
- What does the recent data show?
- What's next?

# 3. Why Genetics?

- Genetics as an entre into biology:
  - An understanding of molecular and cellular mechanisms promises to dramatically improve the opportunities to intervene in a rational way
  - Multiple concrete examples in common medical conditions: Alzheimer, hypercholesterolemia, cancer, hypertension
  - Fragile X, Tuberous Sclerosis, Neurofibromatosis
- Identifying population risks
- Clinical diagnoses
- Study natural history and response to treatment
- Helps clarify the role of the environment

# 4. The Heritability of ASD

- Heritability is the proportion of the variance of a phenotype/disorder explained by genetic variation
- Heritability is not inheritance. Disorder may have a genetic component without being inherited: new (*de novo*) mutation
- 8 studies since 1977 evaluated twins and ASD diagnosis
  - Prior heritability estimates 0.73-0.93
  - Hallmayer et al (*Arch Gen Psych* 2011) show the lowest overall heritability (~0.4; 95% CI 0.08 - 0.84)

# 5. The Heritability of ASD

- Every study has strengths and weaknesses:
  - Careful clinical assessment, large sample (202 pairs)
  - very low response rate (202/1156 total twins) which can influence estimates of dizygotic concordance
  - Large for a twin study, but small numbers for some analyses (<10 MZ female pairs)
  - Slightly higher concordance for identical twins, much higher concordance of non-identical twins (vs. Bailey et al, *Psychol Med* 1995)
- Adds to a body of knowledge; does not supplant it

# 6. Heritability

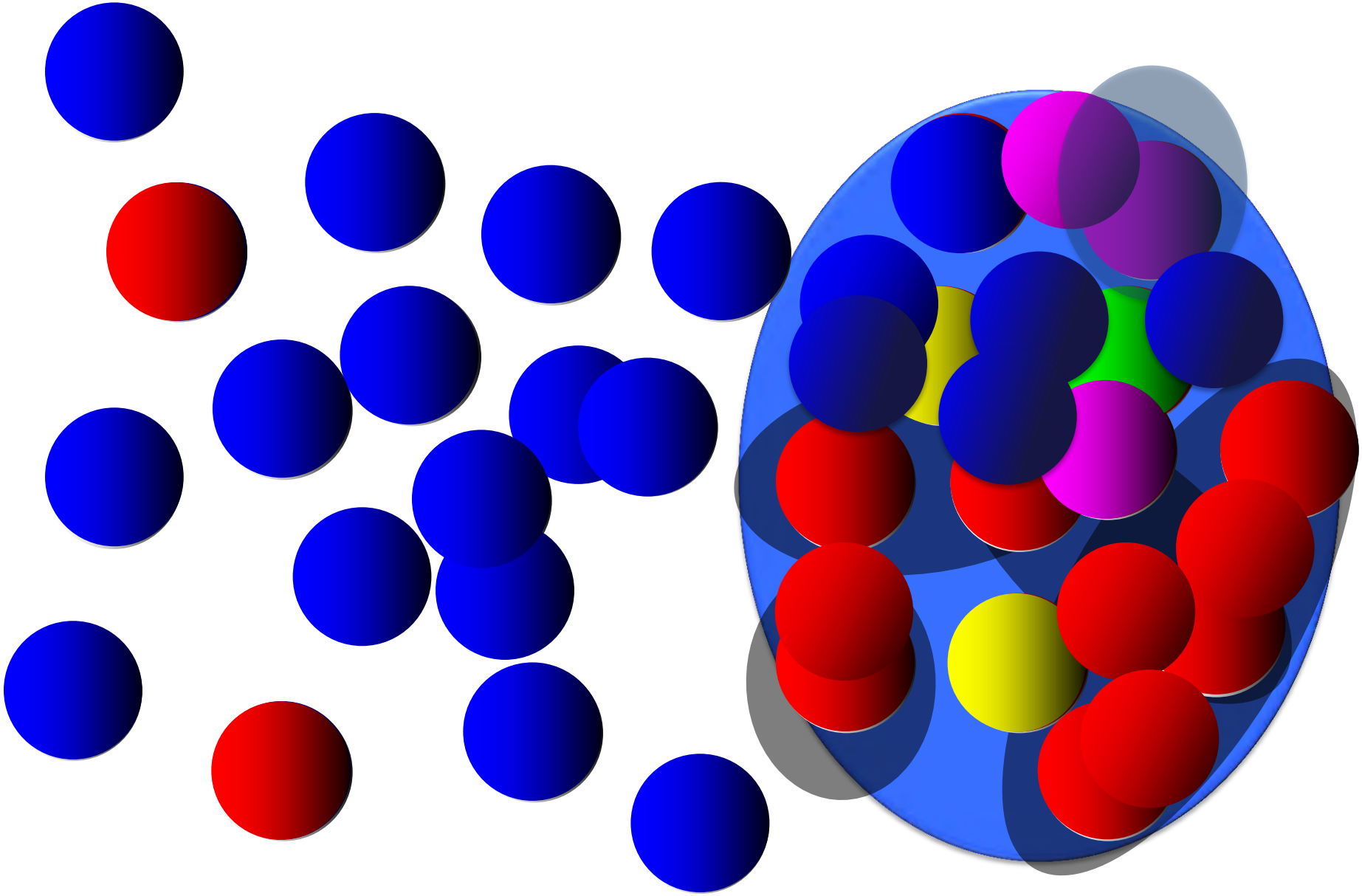
- If these estimates are replicated, it still does not tell us that we should study one and not the other
- Genes vs. environment is a false dichotomy
  - Hard to think of a common medical condition that does not involve the interplay of environment and genetic risks
  - The heritability of high blood pressure is ~0.3-0.5; the heritability of breast cancer is ~0.3;
  - Genetic studies in these and other conditions have led to critical insights into biology leading to new treatments, improved diagnosis and prognosis
  - The boundary between genes and environment is not distinct (e.g. epigenetics)

# 7. The Heritability of ASD

- Different types of knowledge may be gleaned:
  - Genetic studies offer more direct path to elaborating molecular and cellular mechanisms, pharmacological intervention, and ability to study epigenetics;
  - environmental studies identify potentially modifiable risks.
- The two together will ultimately answer the key questions about risk, trajectory and intervention

# 8. Why has it been so hard to find ASD genes?

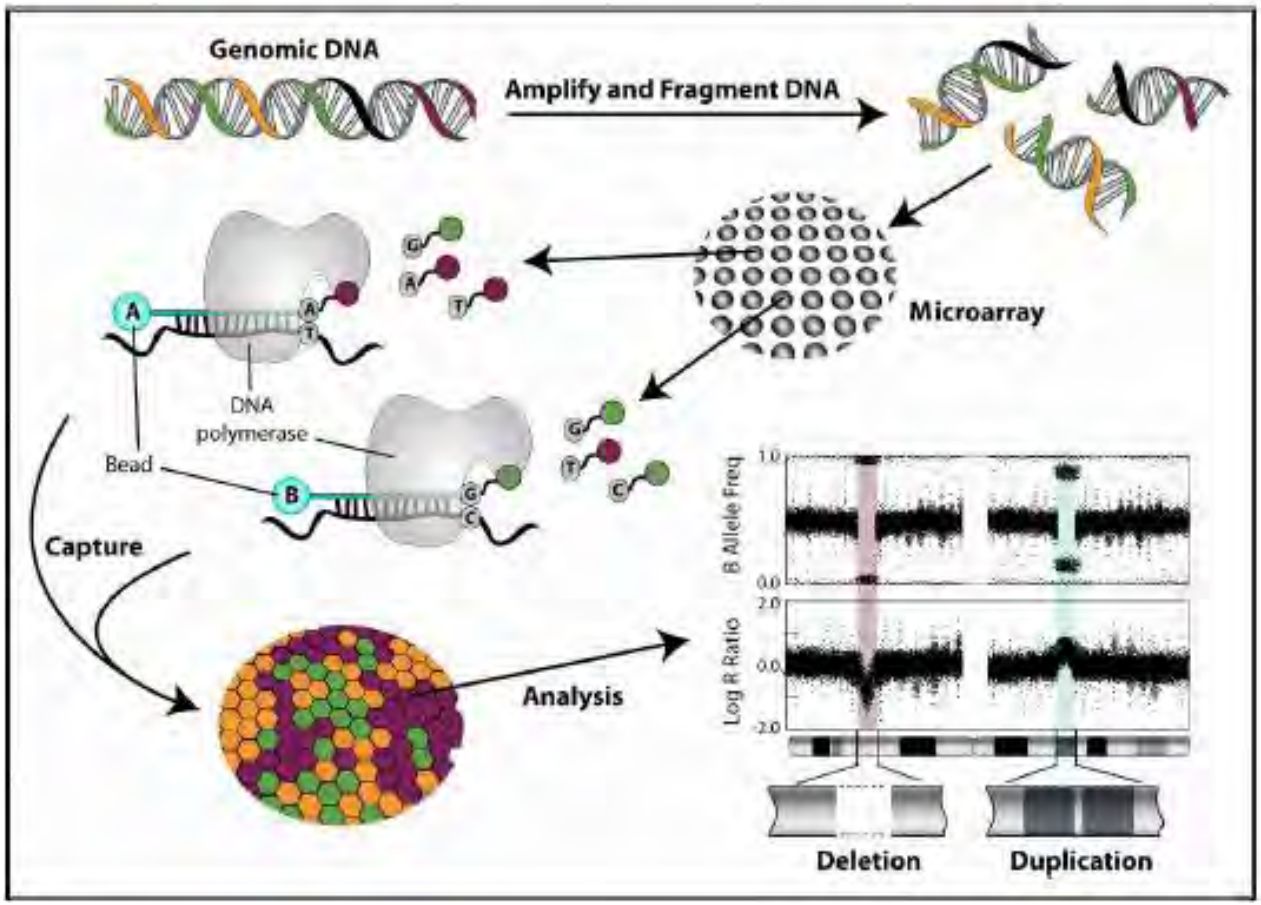
- Highly genetically and phenotypically heterogeneous disorder
- Until very recently, we have had a very limited ability to search through the genome
- Not a single gene; not a 1:1 relationship between genetic risk and outcome.



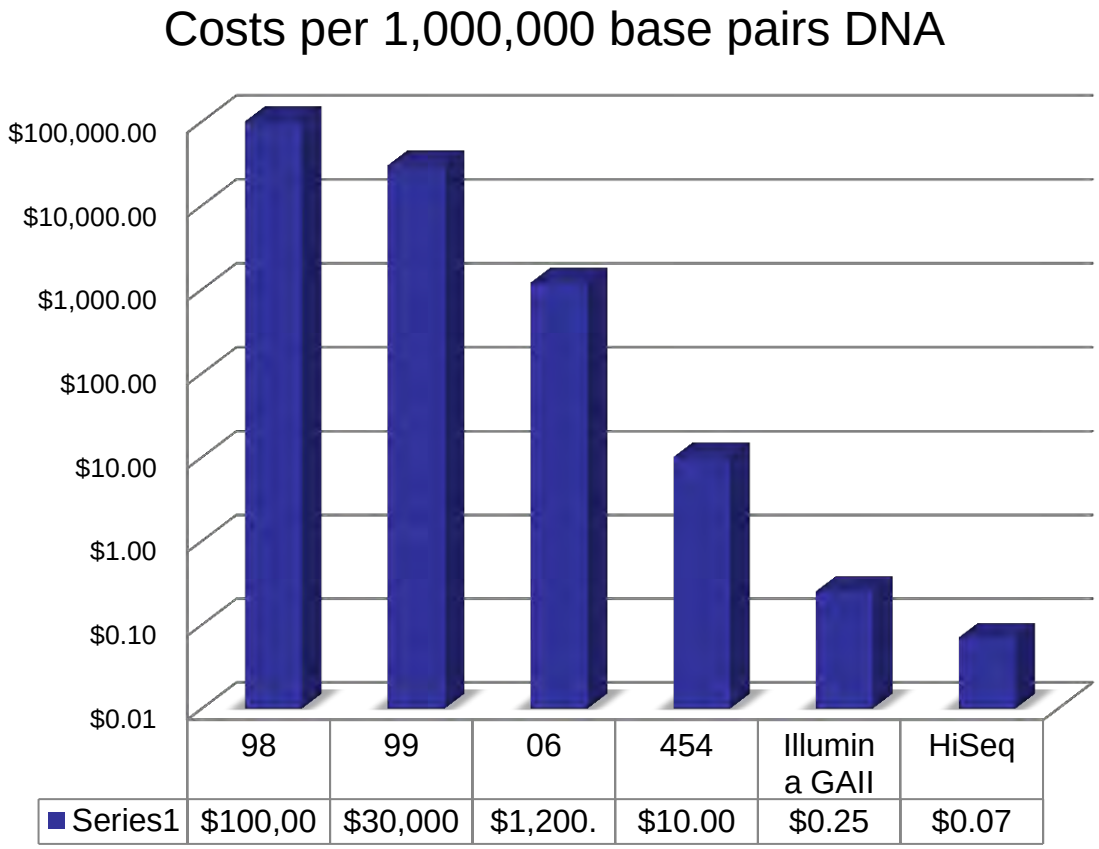
# 10. Genetic Variation

- Any two individuals are ~ 99% identical
- We are interested in the ~1% difference
- Variations are the basis of the genetic risk
- “Gene discovery” is “variation discovery”
  - Sequence: SNP (single nucleotide polymorphism) or SNV (single nucleotide variant)
  - Structural -- gains and losses of genetic material at submicroscopic resolution: CNV (Copy Number Variant)
  - Common and rare: boundary is ~1% of the population carrying the variation.

# 11. Microarrays



# 12. Sequencing



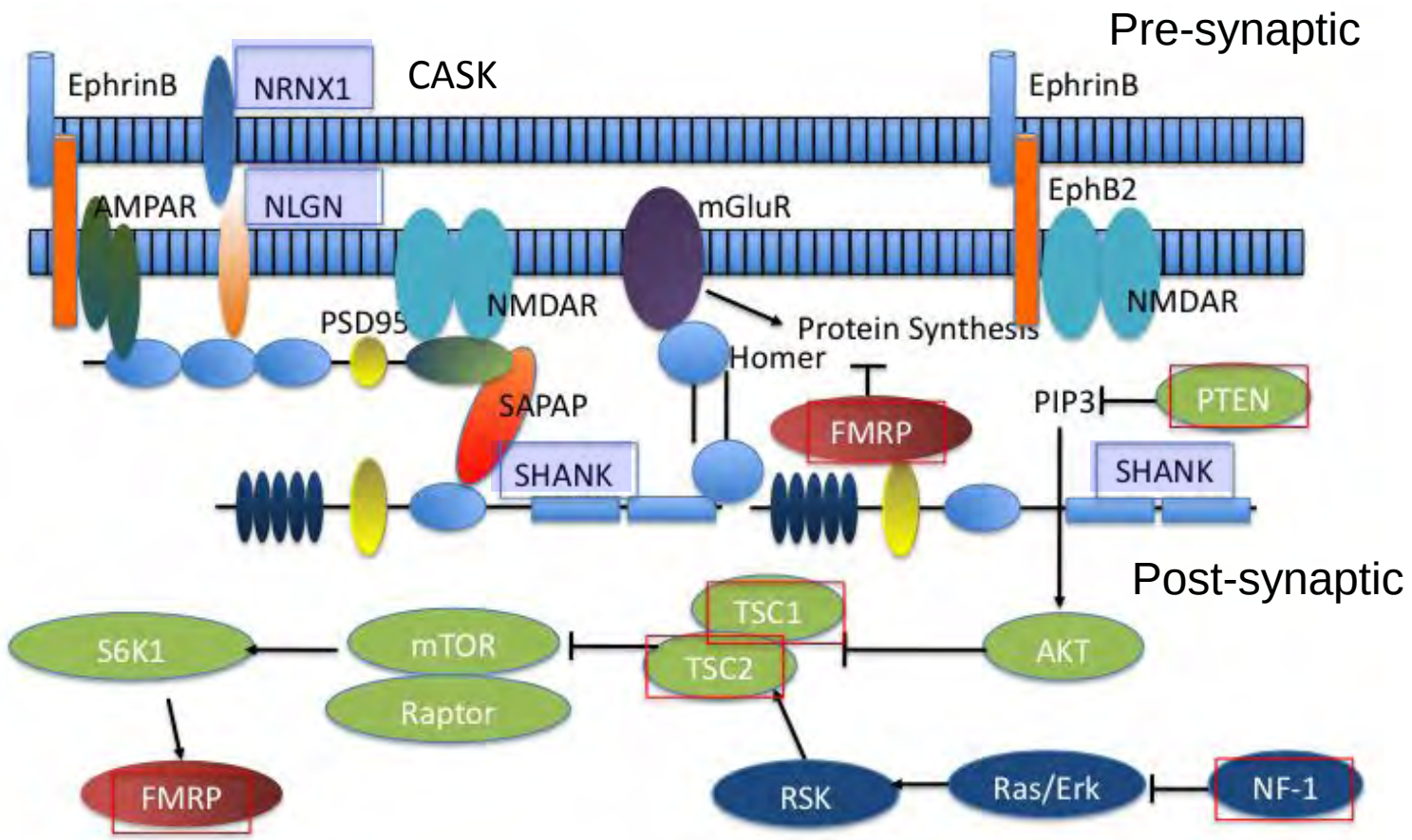

Cost of sequencing all the coding regions of the genome:
~$1500.00$/per individual

Costs of sequencing the entire human genome
~$4,000.00/per individual

# 13. What Does the Recent Data Show?

- Rare mutations contribute to ASD risk
- Contribution of common variants is an open question, but other conditions, including schizophrenia, suggest that larger studies will confirm alleles of small effect (Review: Devlin et al*, Brain Res* 2011)
- Rare mutations in “syndromic” and “idiopathic” ASD point to synaptic function; unlikely to be the only path, (Reviews: Toro et al, T*rends Genet* 2010; Bill and Geschwind, *Curr Opin Genet Devel*, 2009

Picollo

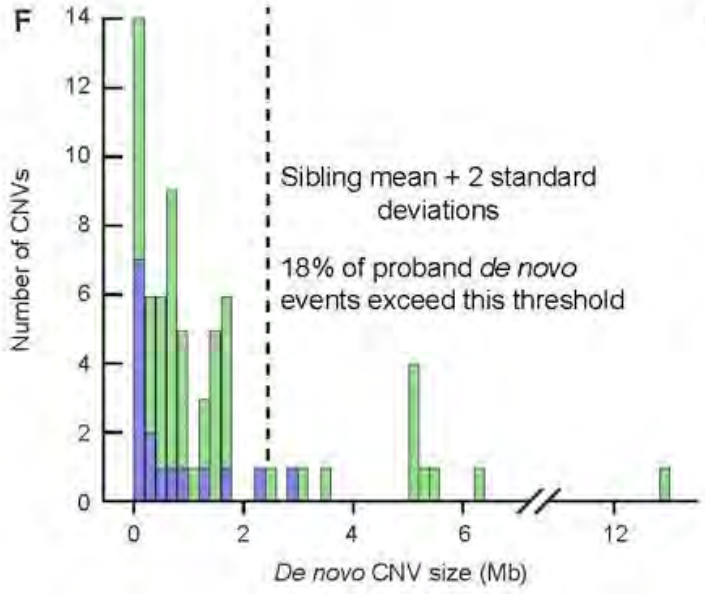


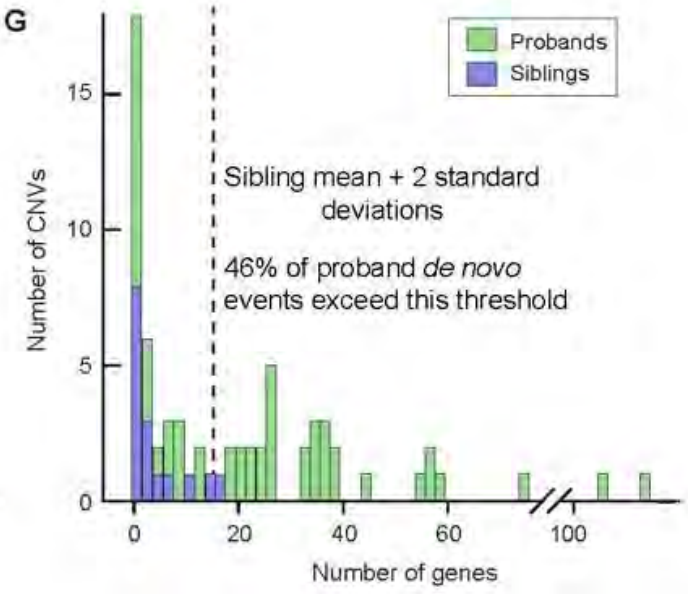

State MW, *Neuron* 2011

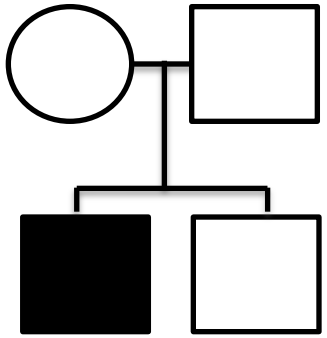
# Recent Data: Copy Number Variation

- Large *de novo* CNVs are present in about 6-10% of ASD and carry large, clearly identifiable, highly reproducible risks
- Particular risks for large multigenic *de novo* CNVs (OR= 5.6; Sanders et al *Neuron* 2011)
- Inherited CNVs also carry risk, but it is harder to discern than for regions that also have *de novo* events
- These *de novo* CNVs are large risk factors for ASD and are not strongly related to IQ
- Girls appear to be protected: CNVs need to be larger and more gene rich than in boys, do not show any relation to IQ (Sanders et al *Neuron* 2011: Levy et al *Neuron* 2011)
- Even strong genetic risk factors do not appear to be specific to ASD (Review: Sebat et al, *Trends Genet* 2009)

# Recent Data: Copy Number Variation



872 pairs

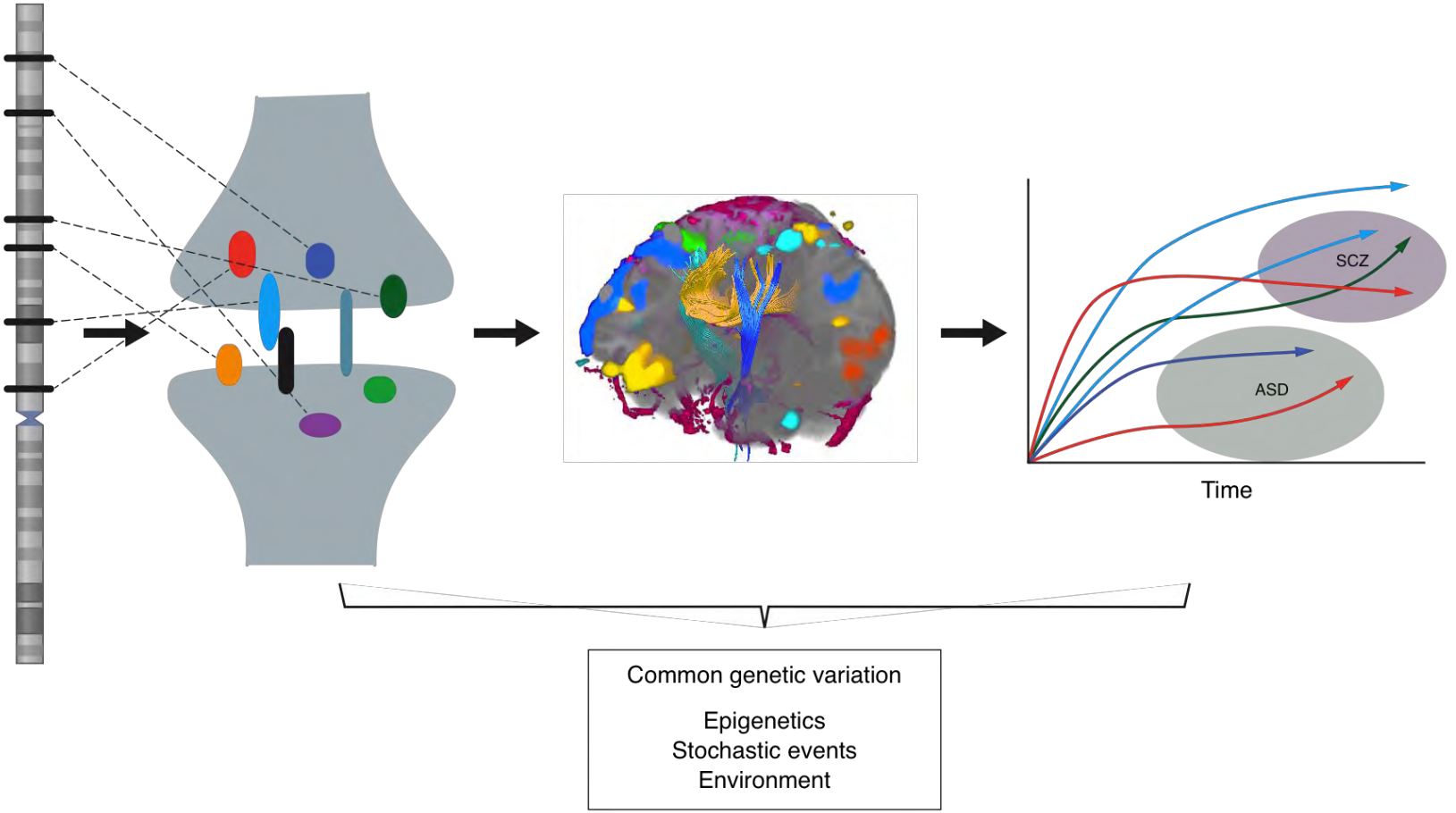
Sanders et al, *Neuron* 2011

# Recent Data: Copy Number Variation

- 16p11.2 deletions* and duplications*: clear and convincing replication. 1% of cases of “idiopathic” ASD (Weiss et al *NEJM* 2008; Kumar et al *Hum Mol Genet* 2008; Marshall et al. *Am J Hum Genet* 2008)
- 7q11.23 duplications strongly associated with ASD
- 7q11.23 is the Williams syndrome region, characterized by  a highly affiliative personality
- 4 regions had recurrent de novo and rare transmitted CNVs (found only in cases)
  - 1q21.1*, 15q13.3*, 16p13.2 *, 16q23.3 *(CDH13)*

- *De novo* events corresponding to previously identified regions
  - 22q11.2*, 15q11-13*, Neurexin 1*, PTCHD1-DDX53
- ~250-300 de novo CNV regions in the human genome carrying risk for ASD (Sanders et at *Neuron* 2011; Levy et al *Neuron* 2011)
- So many individually rare variations do not mean that each affected individual will need a different treatment

# Recent Data




State and Levitt, *Nat Neuroscience* in Press

# 20. What Next?

- Replication may not be as “exciting”, but reflects a critical milestone
- Common variants can provide important information on risk, course and gene x environment interactions:
  - False dichotomy #2:  common vs. rare mutations
- Rare CNVs: Still much more to discover
- Move to biology predicated on clear, reproducible and *large* risks: biochemistry, IPSCs, model systems, systems biology
- Next generation sequencing: the total dataset
- Genetic epidemiological studies to investigate gene x environment interactions; we are now able to think about this *vis a vis* both common and rare variation

# 21. What Next?

- Prospective cohort studies of individuals with the same high risk variants: really one risk multiple disorders?
- Tackle multiple variations contributing simultaneously
- Combine neuroimaging with genetic studies, based on clearly defined risk
- Clarify the role of epigenetics through genetic discoveries
- We can identify genetic risks in a significant minority of cases, but this is only the first step
- Move from a deeper understanding of mechanism to treatment and prevention

# 22. What Next?

- We *finally* have the tools in hand to address the major obstacles:
  - heterogeneity (sample sharing, advocacy groups; NIH)
  - limited ability to “look” at the genome (arrays and next generation sequencing)
  - and the lack of 1:1 relationship of genotype to phenotype (1 and 2 above)
- The pace is slower than anyone would like, but the progress is substantial  and rapidly accelerating