Recent Progress in the genomics of autism spectrum disorders

Matthew W. State MD, PhD
Professor and Chairman, Department of Psychiatry

July 7, 2013
Autism Spectrum Disorders

• Fundamental impairment in reciprocal social interaction, language development and restricted interests/repetitive behaviors
• Onset in early childhood
• Limited treatment options; nothing for core social deficits
• Lack of understanding of basic pathophysiological mechanisms is a major obstacle
• Gene discovery can be a critical first step on the path to solving this
Treatment Targets

- Identify gene or region
- Confirm the link to human disease
- Elaborate expression and function
- Create and study model systems
- Clarify molecular and cellular mechanisms/pathways

- Cytogenetics
- Molecular cytogenetics
- Traditional linkage
- Copy Number Variation
- Next generation sequencing
Genetics 101

- Any two individuals are ~ 99% identical
- We are interested in the 1% difference
- These variations are the basis of the genetic contribution to risk
- “Gene discovery” is “variation discovery”
- Genetic variation can be common or rare in the population
  - common variation tends to have small effects and
  - rare variation tends to have big effects
- Genetic variation can involve the sequence of the DNA
  - Single Nucleotide Variants (SNVs; aka “point mutations”)
- Genetic variation can involve the structure of the DNA:
  - loses or gains = deletions or duplications.
  - Copy Number Variation (CNVs)
- Variation can be passed from generation to generation (transmitted) or new
  - Variation can occur in the parental germ-line/De novo in the child
Genetics of ASD

- Generally described as the most heritable NP disorder
- Few families with apparent Mendelian transmission
- Genetically complex, phenotypically heterogeneous group of disorders
- Lots of early emphasis on variation that is common in the genome (paralleling most early psychiatric genetics work)
- Candidate gene approaches; no clear results – similar to other areas of medicine
- Genome wide association studies (GWAS): powerful gene discovery approach in many common disorders -- no replicating loci in ASD ~N=3000 cases
Genetics of ASD

- Important but infrequent and sporadic findings of rare coding mutations in genes coding synaptic proteins (NLGN4X, SHANK2, SHANK3)
- Growing appreciation of the overlap of ASD with monogenic syndromes (Fragile X, NF)
- First hint of a systematic approach to gene discovery in early copy number variation studies
  - Increased burden of de novo variation in simplex families (Sebat et al Science 2007)
  - Modest increase in burden (amount in cases v controls) of CNVs (Pinto et al Nature 2010)
Simons Simplex Collection

Compare DNA in parents versus children

- Only one affected child
- At least one evaluated unaffected sibling
- Compare children in same family

De novo mutations

Mutation occurs in egg or sperm cell before fertilization, or immediately after fertilization.
Number of de novo CNVs  Number of genes found within de novo CNVs

Large risks: 5x-16x increase  N=\sim1000 matched pairs

Sanders et al. *Neuron* 2011
ASD (including 7q duplications)
7q deletions
NRXN1

1q21.1

7q11.23

16p11.2

22q11.2
ter

Scz

ASD

ID

WBS

ASD

ID

EP

Scz

ASD

ID

EP

Scz

BD

ASD

ID

Scz

ASD

ID

Scz

ASD

ID

15q11.2-13

22q11.2

ter

Willsey et al, unpublished
Costs per 1,000,000 base pairs DNA

<table>
<thead>
<tr>
<th>Year</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>98</td>
<td>$100,000.00</td>
</tr>
<tr>
<td>99</td>
<td>$30,000.00</td>
</tr>
<tr>
<td>06</td>
<td>$1,200.00</td>
</tr>
<tr>
<td>454</td>
<td>$10.00</td>
</tr>
<tr>
<td>Illumina GAI</td>
<td>$0.25</td>
</tr>
<tr>
<td>HiSeq</td>
<td>$0.07</td>
</tr>
</tbody>
</table>
De novo mutations revealed by whole-exome sequencing are strongly associated with autism


Patterns and rates of exonic de novo mutations in autism spectrum disorders


Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations

Brian J. O'Rourke, Laura D. Smith, Elhanan Borenstein, Neuron

De Novo Gene Disruptions in Children on the Autistic Spectrum


Stephan Sanders

School of Medicine
Gene chosen at random

21,000 genes

1,000 carry ASD risk

1 de novo loss of function (LoF)

N=142

2 de novo LoF

N=7

3 or more de novo LoF

N=4

KATNAL2
POGZ
CUL3
ANK2
KDM6B
DIP2A

SCN2A
GRIN2B
CHD8
DYRK1A

\[ p = 1.0 \]
\[ q = 0.95 \]

\[ p = 1.0 \]
\[ q = 0.40 \]

\[ p = 0.04 \]
\[ q = 0.007 \]

\[ p = 0.0003 \]
\[ q = 0.0003 \]
2 hit LoF consistent with 1,000 gene model

- Predicted genes, N = 2,650
- Locus Het. = 1,000 genes

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Number of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 hit de novo LoF</td>
<td>353</td>
</tr>
<tr>
<td>2 hit de novo LoF</td>
<td>38</td>
</tr>
<tr>
<td>3 hit de novo LoF</td>
<td>8</td>
</tr>
</tbody>
</table>

ASD genes identified vs. Number of families

- 100 genes
- 333 genes
- 667 genes
- 1000 genes
- Observed

Sanders et al Nature 2012
• When we started, the genetic architecture of ASD was largely speculation. We now know:
  – hundreds of CNVs perhaps 1000 genes
  – CNVs carry significant risk in ~5%-10% of cases
  – CNV risks for ASD are not specific
  – De novo SNVs in another(?) 15%
  – Increasing de novo SNV rate w paternal age
• Via the study of de novo mutation, there is a systematic path forward for gene discovery
• Clear association of SCN2A, CHD8, GRIN2B, DYRK1A
• How to manage the complexity: heterogeneity, phenotypic diversity and pleiotropy?
  – Pull on the thread and get all of biology
  – Can we determine when and where to look?
A Spatiotemporal human brain transcriptome

Periods 1 & 2
- FC
- PC
- TC
- OC
- HIP
- VF
- MGE
- LGE
- CGE
- DIE
- DTH
- URL

Periods 3-15
- OFC
- DFC
- VFC
- MFC
- M1C
- S1C
- IPC
- A1C
- STC
- ITC
- V1C
- HIP
- AMY
- STR
- MD
- CBC

15 periods of development

B

<table>
<thead>
<tr>
<th>Period</th>
<th>Description</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Embryonic</td>
<td>4-8 PCW</td>
</tr>
<tr>
<td>2</td>
<td>Early fetal</td>
<td>8-10 PCW</td>
</tr>
<tr>
<td>3</td>
<td>Early fetal</td>
<td>10-13 PCW</td>
</tr>
<tr>
<td>4</td>
<td>Early mid-fetal</td>
<td>13-16 PCW</td>
</tr>
<tr>
<td>5</td>
<td>Early mid-fetal</td>
<td>16-19 PCW</td>
</tr>
<tr>
<td>6</td>
<td>Late mid-fetal</td>
<td>19-24 PCW</td>
</tr>
<tr>
<td>7</td>
<td>Late fetal</td>
<td>24-38 PCW</td>
</tr>
<tr>
<td>8</td>
<td>Neonatal &amp; early infancy</td>
<td>0-6 M</td>
</tr>
<tr>
<td>9</td>
<td>Late infancy</td>
<td>6-12 M</td>
</tr>
<tr>
<td>10</td>
<td>Early childhood</td>
<td>1-6 Y</td>
</tr>
<tr>
<td>11</td>
<td>Middle and late childhood</td>
<td>6-12 Y</td>
</tr>
<tr>
<td>12</td>
<td>Adolescence</td>
<td>12-20 Y</td>
</tr>
<tr>
<td>13</td>
<td>Young adulthood</td>
<td>20-40 Y</td>
</tr>
<tr>
<td>14</td>
<td>Middle adulthood</td>
<td>40-60 Y</td>
</tr>
<tr>
<td>15</td>
<td>Late adulthood</td>
<td>60Y+</td>
</tr>
</tbody>
</table>

PCW, post-conceptional weeks; M, postnatal months; Y, postnatal years.

Graph showing gene expression over time:
- Embryonic
- Fetal
- Infancy
- Childhood
- Adolescence
- Adulthood

Gene expression (% of maximum)

Age (log scale)

- Cell proliferation
- Progenitors and immature neurons
- Synapse development
- Dendrite development
- Myelination

NCX
• Sea change in the genetics of ASD
• Systematic gene discovery can offer a foothold into biology
• Parallel advances in neurobiology and systems biology provide unprecedented traction
• The key to moving toward the development of novel and more effective treatments.
• Eric Morrow (Brown)  
  • Dilber Gamsiz  
• Jim Sutcliffe (Vanderbilt)  
  • Brian Yaspan  
  • Suzanne Thomson  
  • Sabata Lund  
• Ed Cook (UIC)  
  • Lea Davis  
  • Suma Jacob  
• Bernie Devlin (Pittsburgh)  
• Kathryn Roeder  
  • Lambertus Klei  
  • Su Chu  
• Donna Martin (UMich)  
• Dorothy Grice (Columbia)  
• Dan Geschwind (UCLA)  
  • Rui Luo  
  • Jenni Lowe  
  • Yuan Tian  
• Rita Cantor (UCLA)  
  • Ake Lu  
• Chris Walsh (Boston Childrens’ )  
  • Tim Yu  
• Matt State (Yale/UCSF)  
  • Stephan Sanders  
  • Michael Murtha  
  • Gulhan Erca  
  • Abha Gupta  
• Jay Tischfield and Andy Brooks (Rutgers)  
• David Ledbetter (Geisinger )  
• Christa Martin (Emory)  
  • Daniel Moreno-De Luca  
• Eric Fombonne (McGill)  
• Cathy Lord (Cornell)  
• Art Beaudet (Baylor)
State lab
Stephan Sanders
Michael T. Murtha
Abha R. Gupta
John D. Murdoch
Melanie J. Raubeson
A. Jeremy Willsey
A. Gulhan Ercan-Sencicek
Nicholas M. DiLullo
Michael F. Walker
Gordon T. Ober
Nicole A. Teran
Youeun Song
Paul El-Fishawy
Ryan Murtha

Yale Center Genome Analysis
Shrikant M. Mane
John D. Overton
Robert D. Bjornson
Nicholas J. Carrierio

Nenad Sestan
Kyle A. Meyer

Richard P. Lifton
Murim Choi

Murat Günel
Kaya Bilguvar

Bernie Devlin (Pittsburgh)
Kathryn Roeder (Carnegie Mellon)
Dan Geschwind (UCLA)
Neelroop N. Parikshak
Jason L. Stein
Simons Foundation
Overlook International Fund
National Institute of Mental Health

The families who participated in the
SFARI: Simons Foundation Autism Research Initiative (www.sfari.org)