Genetic Findings in Autism:
Toward a Biological Understanding

Daniel B. Campbell, Ph.D.
Assistant Professor
Department of Psychiatry and the Behavioral Sciences
Zilkha Neurogenetic Institute
Center for Genomic Psychiatry
Keck School of Medicine
University of Southern California
Los Angeles, CA USA
Email: dbcampbe@usc.edu

The Help Group Summit
October 26, 2013
Outline

• Introduction to Non-Coding RNAs
• Overview of Autism Genetics
• Genome-Wide Association Study results
  – Point to non-coding RNAs
• Exome Sequencing results
  – Point to a transcription factor that regulates … non-coding RNAs
• A New Type of Pharmacology Targets Non-Coding RNAs
The Central Dogma
In the Human Brain, 62% of All Long RNAs are Non-Coding

Figure 3. The relationship between biological complexity and genome composition. In this plot, the 73 organisms with a previously defined number of distinct cell types (e.g., relative biological complexity, see Table S1; ref. 35) are shown as pairs of data points, with one depicting total protein-coding sequence bases (red) and one total non-protein-coding bases (blue) which cumulatively give the total genome size (x-axis). Non-protein-coding sequence increases exponentially with the number of distinct cell types, while protein-coding sequence is asymptotic. Note that the intersection of the protein-coding and non-protein-coding data sets occurs among simple multicellular organisms.
Translating dosage compensation to trisomy 21


Down's syndrome is a common disorder with enormous medical and social costs, caused by trisomy for chromosome 21. We tested the concept that gene imbalance across an extra chromosome can be de facto corrected by manipulating a single gene, XIST (the X-inactivation gene). Using genome editing with zinc finger nucleases, we inserted a large, inducible XIST transgene into the DYRK1A locus on chromosome 21, in Down’s syndrome pluripotent stem cells. The XIST non-coding RNA coats chromosome 21 and triggers stable heterochromatin modifications, chromosome-wide transcriptional silencing and DNA methylation to form a 'chromosome 21 Barr body'. This provides a model to study human chromosome inactivation and creates a system to investigate genomic expression changes and cellular pathologies of trisomy 21, free from genetic and epigenetic noise. Notably, deficits in proliferation and neural rosette formation are rapidly reversed upon silencing one chromosome 21. Successful trisomy silencing in vitro also surmounts the major first step towards potential development of 'chromosome therapy'.

Figure 3 | XIST induces long-range silencing in targeted iPSC cells. a, RNA
Outline

• Introduction to Non-Coding RNAs
• Overview of Autism Genetics
• Genome-Wide Association Study results
  – Point to non-coding RNAs
• Exome Sequencing results
  – Point to a transcription factor that regulates … non-coding RNAs
• A New Type of Pharmacology Targets Non-Coding RNAs
The Contributions to Autism?

• Rare Genetic Variants
  – Whole Exome Sequencing
  – Copy Number Variations

• Common Genetic Variants
  – Genome-Wide Association
  – Candidate Gene Association

• Environmental Factors

• The debate continues …
Global Contribution of Types of Genetic Variation to ASD

- Tagged by Common Variation
- De Novo CNVs
- Nonsense de Novo SNVs
- Missense de Novo SNVs
- 2-Hit LoF Rare Transmitted

Unexplored Rare Genetic Influences/Epistasis/Environment

<table>
<thead>
<tr>
<th>Locus</th>
<th>Cytoband</th>
<th>Pinto et al. [20**] # of events in cases/controls</th>
<th>Sanders et al. [38*] # of events in cases/controls</th>
<th>Combined # of events in cases/controls</th>
<th>P-value (cases vs. controls)</th>
<th>Frequency in 2120 ASD cases (males)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNV-16p11.2</td>
<td>16p11.2</td>
<td>4/996; 3/1287</td>
<td>14/1124; 0/872</td>
<td>16/2120; 3/2159</td>
<td>0.001</td>
<td>0.8%</td>
</tr>
<tr>
<td>PTCHD1/PTCHD1AS</td>
<td>Xp22.11</td>
<td>7/839; 0/383 males</td>
<td>3/968; 0/403 males</td>
<td>10/1807; 0/786 males</td>
<td>0.038</td>
<td>0.5% (0.6%)</td>
</tr>
<tr>
<td>NRXN1</td>
<td>2p16.3</td>
<td>6/996; 0/1287</td>
<td>3/1124; 1/872</td>
<td>9/2120; 1/2159</td>
<td>0.011</td>
<td>0.4%</td>
</tr>
<tr>
<td>CNV-7q11.23</td>
<td>7q11.23</td>
<td>0/996; 0/1287</td>
<td>4/1124; 0/872</td>
<td>4/2120; 0/2159</td>
<td>0.06</td>
<td>0.2%</td>
</tr>
<tr>
<td>CNV-22q11.2</td>
<td>22q11.2</td>
<td>3/996; 1/1287</td>
<td>1/1124; 0/872</td>
<td>4/2120; 1/2159</td>
<td>0.214</td>
<td>0.2%</td>
</tr>
<tr>
<td>CNV-1q21.1</td>
<td>1q21.1</td>
<td>1/996; 3/1287</td>
<td>3/1124; 0/872</td>
<td>4/2120; 3/2159</td>
<td>0.723</td>
<td>0.2%</td>
</tr>
<tr>
<td>CNV-15q13.3</td>
<td>15q13.3</td>
<td>2/996; 0/1287</td>
<td>3/1124; 0/872</td>
<td>5/2120; 0/2159</td>
<td>0.030</td>
<td>0.2%</td>
</tr>
<tr>
<td>CNV-15q11-q13</td>
<td>15q11-q13</td>
<td>1/996; 0/1287</td>
<td>1/1124; 0/872</td>
<td>2/2120; 0/2159</td>
<td>0.245</td>
<td>0.1%</td>
</tr>
<tr>
<td>SHANK2</td>
<td>11q13.3</td>
<td>2/996; 0/1287</td>
<td>0/1124; 0/872</td>
<td>2/2120; 0/2159</td>
<td>0.245</td>
<td>0.1%</td>
</tr>
<tr>
<td>SHANK3</td>
<td>22q13.33</td>
<td>1/996; 0/1287</td>
<td>0/1124; 0/872</td>
<td>1/2120; 0/2159</td>
<td>0.495</td>
<td>0.05%</td>
</tr>
<tr>
<td>NLGN3</td>
<td>Xq13.1</td>
<td>0/839; 0/383 males</td>
<td>1/968; 0/403 males</td>
<td>1/1807; 0/786 males</td>
<td>1</td>
<td>0.05% (0.06%)</td>
</tr>
<tr>
<td>NLGN4X</td>
<td>Xp22.3</td>
<td>0/839; 0/383 males</td>
<td>1/968; 0/403 males</td>
<td>1/1807; 0/786 males</td>
<td>1</td>
<td>0.05% (0.06%)</td>
</tr>
</tbody>
</table>

*a CNV-16p11.2: encompasses 700 kb and 30 genes, 10 deletions and 8 duplications observed in ASD cases, one deletion and two duplications observed in controls; PTCHD1/PTCHD1AS: CNV region involves ~1 Mb region at Xp22.11, 1 deletion of PTCHD1 and 9 deletions affecting upstream PTCHD1AS non-coding RNA gene; NRXN1: 8 deletions and 1 duplication observed in cases, one deletion observed in controls; CNV-7q11.23: encompasses ~1.4 Mb and 22 genes, all four cases represent apparently reciprocal duplications of 7q11.23 region typically deleted in Williams-Beuren syndrome; CNV-22q11.2: encompasses ~2.5 Mb and 56 genes, two deletions and two duplications were observed in ASD cases, one duplication observed in controls; CNV-1q21.1: encompasses ~1.5 Mb and 14 genes, four duplications were observed in ASD cases; CNV-15q13.3: encompasses ~1.5 Mb and 6 genes, four deletions and one duplication were observed in ASD cases; CNV-15q11-q13: encompasses ~5 Mb and 12 genes, two duplications were observed in ASD cases; SHANK2: two deletions observed in cases; SHANK3: one duplication observed in cases; NLGN3: one deletion observed in cases; NLGN4X: one duplication observed in cases; for NRXN1, SHANK2, SHANK3, NLGN3 and NLGN4X CNVs are only counted in cases and controls if they affect one or more exons.

b Fisher’s exact 2-sided P-value.
3 Autism Exome Sequencing Papers: April 2012

**LETTER**

*De novo mutations revealed by whole-exome sequencing are strongly associated with autism*

Stephan J. Sanders1, Michael T. Murtha1, Abba R. Gupta2*, John D. Murdoch4*, Melanie J. Rausboeuf4*, A. Jeremy Willsey4*, A. Gulhan Ercan-Sencicek4, Nicholas M. Dilullo4, Noelmp N. Parrishale4, Jason L. Stein4, Michael F. Walker4, Gordon T. Ober1, Nicola A. Teran5, Yuewen Song6, Paul El-Fishawy7, Ryan C. Murtha1, Murim Choi8, John D. Overton9, Robert D. Bjornson9, Nicholas J. Cavigro4, Kyle A. Meyn9, Kaya Bilgic9, Shrikant M. Mane9, Nenad Sestan9, Richard P. Lifton9, Murat Gincel9, Kathryn Roeder9, Daniel H. Geschwind10, Bernie Devlin10 & Matthew W. State1

**LETTER**

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

Brian J. O’Roak1, Laura Vives1, Santhosh Gorinraj1, Emre Karcak1, Nikolas Krum1, Bradley P. Cox2, Role Levy2, Arthur Ko3, Choli Lee3, Joshua D. Smith1, Emily H. Turner1, Ian B. Stanaway1, Benjamin Vernot1, Maika Malig2, Carl Baker2, Beau Reilly3, Joshua M. Akey4, Elihvan Borenstein5,6, Mark J. Rieder1, Deborah A. Nickerson1, Raphael Bernier2, Jay Shendure1 & Evan E. Eichler1,2

**LETTER**

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

Benjamin M. New1,2,3,6, Yan Xu1,2,3,4, Li Liu3,4, Avi Maayan1, Katrin E. Samocha1,2, Antko Saas1, Chiao-Feng Lin5, Christine Stevens2, Li-Xuan Wang1, Vahidmir Makaniy6,9, Faz Polak2,3, Scanggl Jung2,4, Jared Maguire2, Emily L. Crawford2,4, Nicholas G. Campbell1,6, Evan T. Geller1, Otto Villalba1, Chad Schaffer1, Han Liu3, Ya Zhou3, Guang Chen3,4, Jiyun Lian4,4, Ruth Dannenfeldt1, Omar Zaitoun6, zuszana ivanyosi4, tamsam wasim6, Yona Azimza6, Jeffrey T. Li6, treeswai somu6, yingying wu6, Lora Lewis6, Yi-Han9, Benjamin F. Wright1,4, Elaine Linn1, Elizabeth Rosin2, Andrew Kirby2,2, Jason Flannick1,5, Michael Schork1,2, Khalid Shaikh2,3, Tim Fennell2,3, Kisha Garmire2, Eric Rank2, Ryan Poplin2, Stacey Gabriel2, Mark DeFraino7, Jack R. Wimbish1, Bradford E. Soome1, Shawn E. Levy3, Carolin Rekiszczak2, Shamil Suryanarayanan2, Eric Boerwinkle1,2, Joseph D. Buxbaum1,2,8,9, Edwin H. Cook Jr1,2, Bernie Devlin1,2, Richard A. Gibbs2, Kathryn Roeder4, Gerhard D. Schellenberg5, James S. Sulis1,6 & Mark J. Daly1,2
April Exome: Major Findings

Table 1 | Distribution of SNVs between probands and siblings

<table>
<thead>
<tr>
<th>Category</th>
<th>Total number of SNVs*</th>
<th>SNVs per subject</th>
<th>Per base SNV rate (x10^-8)</th>
<th>( P )</th>
<th>Odds ratio (95% CI)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pro ( N = 200 )</td>
<td>Sib ( N = 200 )</td>
<td>Pro ( N = 200 )</td>
<td>Sib</td>
<td></td>
</tr>
<tr>
<td>De novo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>154</td>
<td>125 §</td>
<td>0.77</td>
<td>0.63</td>
<td>1.58</td>
</tr>
<tr>
<td>Silent</td>
<td>29</td>
<td>39</td>
<td>0.15</td>
<td>0.20</td>
<td>0.29</td>
</tr>
<tr>
<td>All non-synonymous</td>
<td>125</td>
<td>87</td>
<td>0.63</td>
<td>0.44</td>
<td>1.29</td>
</tr>
<tr>
<td>Missense</td>
<td>110</td>
<td>82</td>
<td>0.55</td>
<td>0.41</td>
<td>1.13</td>
</tr>
<tr>
<td>Nonsense/splice site</td>
<td>15</td>
<td>5</td>
<td>0.08</td>
<td>0.03</td>
<td>0.16</td>
</tr>
<tr>
<td>All genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain-expressed genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>137</td>
<td>96</td>
<td>0.69</td>
<td>0.48</td>
<td>1.41</td>
</tr>
<tr>
<td>Silent</td>
<td>23</td>
<td>30</td>
<td>0.12</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>All non-synonymous</td>
<td>114</td>
<td>67</td>
<td>0.57</td>
<td>0.34</td>
<td>1.18</td>
</tr>
<tr>
<td>Missense</td>
<td>101</td>
<td>64</td>
<td>0.51</td>
<td>0.32</td>
<td>1.04</td>
</tr>
<tr>
<td>Nonsense/splice site</td>
<td>13</td>
<td>3</td>
<td>0.07</td>
<td>0.02</td>
<td>0.14</td>
</tr>
</tbody>
</table>


- Each paper finds a few genes with mutations in 2 affected individuals and 0 controls
- In a total of 584 families, the same gene was found to have a *de novo* mutation in no more than 2 (<0.4%) families
  - 4 genes: *SCN2A, CHD8, NTNG1, KATNAL2*
- Each paper concludes that there is no gene that is causal for autism, and that several hundred genes will contribute to risk
The 4th Autism Exome Sequencing Paper

### Table 2. Summary of De Novo Single Nucleotide Variants (SNVs) in 343 SSC Families

<table>
<thead>
<tr>
<th>SNV Effect</th>
<th>40x (High) Coverage</th>
<th>All Loci</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proband</td>
<td>Sibling</td>
<td>Proband</td>
<td>Sibling</td>
<td>Proband F (29)</td>
<td>Proband M (314)</td>
<td>Sibling F (182)</td>
<td>Sibling M (161)</td>
<td>Both</td>
</tr>
<tr>
<td>Splice site</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Nonsense</td>
<td>15</td>
<td>7</td>
<td>19</td>
<td>9</td>
<td>3</td>
<td>16</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Missense</td>
<td>125</td>
<td>121</td>
<td>207</td>
<td>207</td>
<td>19</td>
<td>188</td>
<td>116</td>
<td>91</td>
<td>3</td>
</tr>
<tr>
<td>Synonymous</td>
<td>53</td>
<td>42</td>
<td>79</td>
<td>69</td>
<td>8</td>
<td>71</td>
<td>43</td>
<td>26</td>
<td>4</td>
</tr>
<tr>
<td>Promoter</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>UTR</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Intron</td>
<td>34</td>
<td>35</td>
<td>59</td>
<td>64</td>
<td>5</td>
<td>54</td>
<td>38</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>Intergenic</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>236</td>
<td>216</td>
<td>380</td>
<td>364</td>
<td>36</td>
<td>344</td>
<td>209</td>
<td>155</td>
<td>10</td>
</tr>
</tbody>
</table>

De novo SNVs were tabulated according to affected status, gender, and type of mutation. Data under “40× coverage” indicate variants in the subset of the exome target region in which all members of a given family were covered by at least 40 sequence reads. The power to detect de novo variants in children from this well-covered portion of the target is very high, and we found no bias in coverage between affected and unaffected children. No significant difference was seen for missense mutations (125 in probands to 121 in unaffected siblings), but larger ratios of nonsense (15:7) and splice site (4:1) mutations were observed in probands relative to unaffected siblings. When we expanded our set of calls to include every variant that passed our thresholds (under “all loci”; see Experimental Procedures), similar ratios were observed. Probands and unaffected siblings are further subdivided based on gender: “proband F” indicates an affected female; “proband M” an affected male; “sibling F” an unaffected female; and “sibling M” an unaffected male. In parentheses, we indicate the number of children with the corresponding affected status and gender. The “both” column shows de novo SNVs that were shared by both siblings from the same family.
Autism Exome Sequencing: June 2012

• 967 families (quads) exome sequenced
• Still no gene with *de novo* LGD mutations in more than 2 (0.2%) families
“Mutations” that Cause Loss of Protein Function are Shockingly Common

A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding Genes

Daniel G. MacArthur,1,2* Suganthi Balasubramanian,3,4 Adam Frankish,1 Ni Huang,1 James Morris,1 Klaudia Walter,1 Luke Jostins,1 Lukas Habegger,3,4 Joseph K. Pickrell,5 Stephen B. Montgomery,6,7 Cornelis A. Albers,1,8 Zhengdong D. Zhang,9 Donald F. Conrad,10 Gerton Lunter,11 Hancheng Zheng,12 Qasim Ayub,1 Mark A. DePristo,13 Eric Banks,13 Min Hu,1 Robert E. Handsaker,13,14 Jeffrey A. Rosenfeld,15 Menachem Fromer,13 Mike Jin,3 Xinmeng Jasmine Mu,3,4 Ekta Khurana,3,4 Kai Ye,16 Mike Kay,1 Gary Ian Saunders,1 Marie-Marthe Suner,1 Toby Hunt,1 If H. A. Barnes,1 Clara Amid,1,17 Denise R. Carvalho-Silva,1 Alexandra H. Bignell,1 Catherine Snow,1 Bryndis Yngvadottir,1 Suzannah Bumpstead,1 David N. Cooper,18 Yali Xue,1 Irene Gallego Romero,1,15 1000 Genomes Project Consortium, Jun Wang,12 Yingrui Li,12 Richard A. Gibbs,19 Steven A. McCarroll,13,14 Emmanouil T. Dermitzakis,7 Jonathan K. Pritchard,5,20 Jeffrey C. Barrett,1 Jennifer Harrow,1 Matthew E. Hurles,1 Mark B. Gerstein,3,4,21† Chris Tyler-Smith1†

Genome-sequencing studies indicate that all humans carry many genetic variants predicted to cause loss of function (LoF) of protein-coding genes, suggesting unexpected redundancy in the human genome. Here we apply stringent filters to 2951 putative LoF variants obtained from 185 human genomes to determine their true prevalence and properties. We estimate that human genomes typically contain ~100 genuine LoF variants with ~20 genes completely inactivated. We identify rare and likely deleterious LoF alleles, including 26 known and 21 predicted severe disease-causing variants, as well as common LoF variants in nonessential genes. We describe functional and evolutionary differences between LoF-tolerant and recessive disease genes and a method for using these differences to prioritize candidate genes found in clinical sequencing studies.

Science. 2012.
All “Mutations” are Shockingly Common

<table>
<thead>
<tr>
<th>Variant type</th>
<th>Mean number of variants (± sd) in African Americans</th>
<th>Mean number of variants (± sd) in European Americans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel variants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>303 (± 32)</td>
<td>192 (± 21)</td>
</tr>
<tr>
<td>Nonsense</td>
<td>5 (± 2)</td>
<td>5 (± 2)</td>
</tr>
<tr>
<td>Synonymous</td>
<td>209 (± 26)</td>
<td>109 (± 16)</td>
</tr>
<tr>
<td>Splice</td>
<td>2 (± 1)</td>
<td>2 (± 1)</td>
</tr>
<tr>
<td>Total</td>
<td>520 (± 53)</td>
<td>307 (± 33)</td>
</tr>
<tr>
<td>Non-novel variants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>10,828 (± 342)</td>
<td>9,319 (± 233)</td>
</tr>
<tr>
<td>Nonsense</td>
<td>98 (± 8)</td>
<td>89 (± 6)</td>
</tr>
<tr>
<td>Synonymous</td>
<td>12,567 (± 416)</td>
<td>10,536 (± 280)</td>
</tr>
<tr>
<td>Splice</td>
<td>36 (± 4)</td>
<td>32 (± 3)</td>
</tr>
<tr>
<td>Total</td>
<td>23,529 (± 751)</td>
<td>19,976 (± 505)</td>
</tr>
<tr>
<td>Total variants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missense</td>
<td>11,131 (± 364)</td>
<td>9,511 (± 244)</td>
</tr>
<tr>
<td>Nonsense</td>
<td>103 (± 8)</td>
<td>93 (± 6)</td>
</tr>
<tr>
<td>Synonymous</td>
<td>12,776 (± 434)</td>
<td>10,645 (± 286)</td>
</tr>
<tr>
<td>Splice</td>
<td>38 (± 5)</td>
<td>34 (± 4)</td>
</tr>
<tr>
<td>Total</td>
<td>24,049 (± 791)</td>
<td>20,283 (± 523)</td>
</tr>
</tbody>
</table>

The table lists the mean number (± standard deviation (sd)) of novel and non-novel coding single nucleotide variants from 100 sampled African Americans and 100 European Americans. Non-novel variants refer to those found in dbSNP131 or in 200 other control samples.
Multiplex Targeted Sequencing Identifies Recurrently Mutated Genes in Autism Spectrum Disorders

Exome sequencing studies of autism spectrum disorders (ASD) have identified many de novo mutations, but few recurrently disrupted genes. We therefore developed a modified molecular inversion probe method enabling ultra-low-cost candidate gene resequencing in very large cohorts. To demonstrate the power of this approach, we captured and sequenced 44 candidate genes in 2,446 ASD probands. We discovered 27 de novo events in 16 genes, 59% of which are predicted to truncate proteins or disrupt splicing. We estimate that recurrent disruptive mutations in six genes—CHD8, DYRK1A, GRIN2B, TBR1, PTEN, and TBL1XR1—may contribute to 1% of sporadic ASD. Our data support associations between specific genes and reciprocal subphenotypes (CHD8-macrocephaly, DYRK1A-microcephaly) and replicate the importance of a β-catenin/chromatin remodeling network to ASD etiology.
CHD8: 11 de novo LoF mutations

Table 1. Six genes with recurrent de novo mutations. Abbreviations: M-male, F-female, Mut-mutation type, fs-framework indel, ns-nonsense, sp-splice-site, aa-amino acid deletion, ms-missense, HGVS-Human Genome Variation Society nomenclature; NVIQ-nonverbal intellectual quotient.

<table>
<thead>
<tr>
<th>Proband</th>
<th>Sex</th>
<th>Gene</th>
<th>Mut</th>
<th>Assay†</th>
<th>HGVS</th>
<th>NVIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>12991p1</td>
<td>M</td>
<td>CHD8*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Glu2103ArgX3</td>
<td>67</td>
</tr>
<tr>
<td>12752p1</td>
<td>F</td>
<td>CHD8*</td>
<td>fs</td>
<td>EX</td>
<td>p.Leu2120ProX13</td>
<td>93</td>
</tr>
<tr>
<td>14233p1</td>
<td>M</td>
<td>CHD8*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Asn237LysX2</td>
<td>19</td>
</tr>
<tr>
<td>14406p1</td>
<td>M</td>
<td>CHD8*</td>
<td>as</td>
<td>MIP</td>
<td>p.His249del</td>
<td>98</td>
</tr>
<tr>
<td>12099p1</td>
<td>M</td>
<td>DYRK1A*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Ile485LysX2</td>
<td>55</td>
</tr>
<tr>
<td>13552p1</td>
<td>M</td>
<td>DYRK1A*</td>
<td>fs</td>
<td>MIP</td>
<td>p.His498AsnX94</td>
<td>66</td>
</tr>
<tr>
<td>11691p1</td>
<td>M</td>
<td>GRIN2B^</td>
<td>fs</td>
<td>MIP</td>
<td>p.Ser340GlnX25</td>
<td>62</td>
</tr>
<tr>
<td>13932p1</td>
<td>M</td>
<td>GRIN2B^</td>
<td>ms</td>
<td>MIP</td>
<td>p.Cys450Tyr</td>
<td>55</td>
</tr>
<tr>
<td>12547p1</td>
<td>M</td>
<td>GRIN2B^</td>
<td>ms</td>
<td>MIP</td>
<td>p.Trp559X</td>
<td>65</td>
</tr>
<tr>
<td>12681p1</td>
<td>F</td>
<td>GRIN2B^</td>
<td>sp</td>
<td>EX</td>
<td>c.2172-2A&gt;G</td>
<td>65</td>
</tr>
<tr>
<td>14433p1</td>
<td>M</td>
<td>PTEN</td>
<td>ms</td>
<td>MIP</td>
<td>p.Thr131Ile</td>
<td>50</td>
</tr>
<tr>
<td>14641p1</td>
<td>M</td>
<td>PTEN</td>
<td>fs</td>
<td>MIP</td>
<td>p.Cys136MetX44</td>
<td>33</td>
</tr>
<tr>
<td>11390p1</td>
<td>F</td>
<td>PTEN</td>
<td>ms</td>
<td>EX</td>
<td>p.Thr167Asn</td>
<td>77</td>
</tr>
<tr>
<td>12335p1</td>
<td>F</td>
<td>TBL1XR1*</td>
<td>ms</td>
<td>EX</td>
<td>p.Glu283Pro</td>
<td>47</td>
</tr>
<tr>
<td>14612p1</td>
<td>M</td>
<td>TBL1XR1*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Ile397SerX19</td>
<td>41</td>
</tr>
<tr>
<td>11480p1</td>
<td>M</td>
<td>TBR1↑</td>
<td>fs</td>
<td>MIP</td>
<td>p.Ala136ProX80</td>
<td>41</td>
</tr>
<tr>
<td>13814p1</td>
<td>M</td>
<td>TBR1↑</td>
<td>ms</td>
<td>MIP</td>
<td>p.Lys228Glu</td>
<td>78</td>
</tr>
<tr>
<td>13796p1</td>
<td>F</td>
<td>TBR1↑</td>
<td>fs</td>
<td>MIP</td>
<td>p.Ser351X</td>
<td>63</td>
</tr>
</tbody>
</table>

*Part of 49-member connected component reported in (3). †Part of expanded 74-member connected component. ‡Primary assay that identified the variant. ††Proband was exome sequenced by cited study and variant was not reported or reported. †Variant reported in MIP screen from (3).

... and association with autism P<10^-8
... note: next best genes have 3 de novo LGD

Table 2

GWA studies in ASDf

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>N</th>
<th>SNP</th>
<th>Chr band</th>
<th>Position</th>
<th>Gene</th>
<th>MAF</th>
<th>OR</th>
<th>P discovery</th>
<th>P meta analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang et al.</td>
<td>AGRE&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3101 subjects; 1299 cases; 780 families</td>
<td>rs4307059</td>
<td>5p14.1</td>
<td>26,003,460</td>
<td>None</td>
<td>0.38</td>
<td>1.19</td>
<td>3.4 x 10^-8</td>
<td>2.1 x 10^-10</td>
</tr>
<tr>
<td>[83]</td>
<td>ACC&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1204 cases; 6491 controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAP&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>1390 subjects; 504 cases; 447 families</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CART&lt;sup&gt;b&lt;/sup&gt;</td>
<td>108 cases, 540 controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiss et al.</td>
<td>AGRE&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3000 subjects, 780 families</td>
<td>rs10513025</td>
<td>5p15.2</td>
<td>9,676,622</td>
<td>SEMA5A TAS2R1</td>
<td>0.041</td>
<td>0.55</td>
<td>1.7 x 10^-6</td>
<td>2.1 x 10^-7</td>
</tr>
<tr>
<td>[84]</td>
<td>NIMH&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1233 subjects, 341 families</td>
<td>rs10513026</td>
<td></td>
<td>9,677,106</td>
<td></td>
<td>0.04</td>
<td>0.53</td>
<td>4.5 x 10^-6</td>
<td>na&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Montreal&lt;sup&gt;b&lt;/sup&gt;</td>
<td>318 trios, 1755 trios</td>
<td>rs16883317</td>
<td></td>
<td>9,701,592</td>
<td></td>
<td>0.033</td>
<td>0.53</td>
<td>7.2 x 10^-5</td>
<td>na&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Anney et al.</td>
<td>AGP</td>
<td>1369 families, 1385 probands</td>
<td>rs4141463</td>
<td>20p12.1</td>
<td>14,695,221</td>
<td>MACROD2</td>
<td>0.43</td>
<td>0.56</td>
<td>2.1 x 10^-8</td>
<td>4.7 x 10^-8</td>
</tr>
<tr>
<td>[85]</td>
<td>AGRE</td>
<td>810 families</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Discovery.  
<sup>b</sup> Replication.  
<sup>c</sup> Included in Ref. [83] and reported in Ref. [100].  
<sup>d</sup> Imputed SNPs. Table is based on Ref. [88].
Outline

• Introduction to Non-Coding RNAs
• Overview of Autism Genetics
• Genome-Wide Association Study results
  – Point to non-coding RNAs
• Exome Sequencing results
  – Point to a transcription factor that regulates … non-coding RNAs
• A New Type of Pharmacology Targets Non-Coding RNAs
Common genetic variants on 5p14.1 associate with autism spectrum disorders


Autism spectrum disorders (ASDs) represent a group of childhood neurodevelopmental and neuropsychiatric disorders characterized by deficits in verbal communication, impairment of social interaction, and restricted and repetitive patterns of interests and behaviour. To identify common genetic risk factors underlying ASDs, here we present the results of genome-wide association studies on a cohort of 780 families (3,101 subjects) with affected children, and a second cohort of 1,204 affected subjects and 6,491 control subjects, all of whom were of European ancestry. Six single nucleotide polymorphisms between cadherin 10 (CDH10) and cadherin 9 (CDH9)—two genes encoding neuronal cell-adhesion molecules—revealed strong association signals, with the most significant SNP being rs4307059 ($P = 3.4 \times 10^{-8}$, odds ratio = 1.19). These signals were replicated in two independent cohorts, with combined $P$ values ranging from $7.4 \times 10^{-8}$ to $2.1 \times 10^{-10}$. Our results implicate neuronal cell-adhesion molecules in the pathogenesis of ASDs, and represent, to our knowledge, the first demonstration of genome-wide significant association of common variants with susceptibility to ASDs.
Genome Wide Association Study (GWAS) Revealed Association of Common Genetic Variants on Chromosome 5


Possible Interpretations:
1. The GWAS peak implicates the neighboring *CDH10* and *CDH9* genes in ASD.
rs4307059 Genotype Did Not Correlate with Expression of \textit{CDH9} or \textit{CDH10}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Examination of brain expression for \textit{CDH10} and \textit{CDH9}. a, The \textit{in situ} hybridization of \textit{CDH10} in human fetal brain sectioned in the sagittal plane. Medial and lateral sections from a representative sample are shown above corresponding cresyl-violet-stained marker slides. Orbitofrontal cortex (OFC) and frontal cortex (FC) are highlighted, with marked expression enrichment. b, The SNP genotypes of rs4307059 are not associated with \textit{CDH9} or \textit{CDH10} transcript levels in 93 cortical brain tissues.}
\end{figure}

Possible Interpretations:
1. The GWAS peak implicates the neighboring *CDH10* and *CDH9* genes in ASD.
2. The GWAS data indicate that no common variants contribute to ASD.
3. The GWAS peak indicates significant contribution of a functional, non-protein-coding genetic element to ASD risk.
A Long Non-Coding RNA is Expressed Directly Under the ASD GWAS Peak

Moesin Pseudogene 1 (MSNP1)
Northern Hybridization: The Long Non-Coding RNA is Complementary to *MSNP1*

**MSNP1AS** is the Long Non-Coding RNA Directly Under the ASD GWAS Peak

Postmortem Temporal Cortex: *MSNP1AS* Expression is Increased 12.7-Fold in ASD

MSNP1AS Expression is Correlated with ASD Risk Allele Genotypes

Neither *CDH9* nor *CDH10* Expression is Correlated with ASD Risk Allele Genotype

Despite 2.4-fold Increase in MSN RNA, Moesin Protein Levels are Unchanged

Over-Expression of *MSNP1AS* Causes a Decrease in Moesin Protein

Correlations Among MSN, MSNP1AS, and Moesin Protein in Postmortem Cortex

MSN is the major determinant of moesin protein levels

MSN and MSNP1AS appear to be co-regulated

MSNP1AS contributes to the regulation of moesin protein

Over-Expression of *MSNP1AS* Causes a Decrease in the Average Neurite Length

---

**Neurite Length**

- **Control**
- **MSN**
- **MSNP1AS + MSN**
- **MSNP1**
- **MSNP1AS**

**Construct Transfected**

- HEK 24 hr
- HEK 72 hr
- SK-N-SH 24 hr
- SK-N-SH 72 hr

---

**Preliminary Data**
**MSNP1AS Summary**

- **MSNP1AS** is the second anti-sense of a pseudogene demonstrated to regulate expression of a gene on a different chromosome
  - First was Oct4-pg5 regulation of Oct4 (Hawkins & Morris, 2010)

- Moesin is an X chromosome-encoded protein that acts:
  - (1) presynaptically to maintain axonal growth cones;
  - (2) postsynaptically to induce dendritic spine formation; and
  - (3) at the immune synapse (APCs-lymphocytes)
Table 2

GWA studies in ASDf

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>N</th>
<th>SNP</th>
<th>Chr band</th>
<th>Position</th>
<th>Gene</th>
<th>MAF</th>
<th>OR</th>
<th>P discovery</th>
<th>P meta or mega analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang et al. [83]</td>
<td>AGREa</td>
<td>3101 subjects; 1299 cases; 780 families</td>
<td>rs4307059</td>
<td>5p14.1</td>
<td>26,003,460</td>
<td>None</td>
<td>0.38</td>
<td>1.19</td>
<td>3.4 × 10^{-8}</td>
<td>2.1 × 10^{-10}</td>
</tr>
<tr>
<td></td>
<td>ACCa</td>
<td>1299 cases; 780 families</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAPb,c</td>
<td>1390 subjects; 504 cases; 447 families</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CARTb</td>
<td>108 cases, 540 controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiss et al. [84]</td>
<td>AGREa</td>
<td>3000 subjects, 780 families</td>
<td>rs10513025</td>
<td>5p15.2</td>
<td>9,676,622</td>
<td>SEMA5A</td>
<td>0.041</td>
<td>0.55</td>
<td>1.7 × 10^{-6}</td>
<td>2.1 × 10^{-7}</td>
</tr>
<tr>
<td></td>
<td>NIMHa</td>
<td>1233 subjects, 341 families</td>
<td>rs10513026</td>
<td></td>
<td>9,677,106</td>
<td>TAS2R1</td>
<td>0.04</td>
<td>0.53</td>
<td>4.5 × 10^{-6}</td>
<td>na^d</td>
</tr>
<tr>
<td></td>
<td>Montrealb</td>
<td>318 trios</td>
<td>rs16883317</td>
<td></td>
<td>9,701,592</td>
<td></td>
<td>0.038</td>
<td>0.53</td>
<td>7.2 × 10^{-5}</td>
<td>na^d</td>
</tr>
<tr>
<td></td>
<td>AGPb</td>
<td>1755 trios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anney et al. [85]</td>
<td>AGP</td>
<td>1369 families, 1385 probands</td>
<td>rs4141463</td>
<td>20p12.1</td>
<td>14,695,221</td>
<td>MACROD2</td>
<td>0.43</td>
<td>0.56</td>
<td>2.1 × 10^{-8}</td>
<td>4.7 × 10^{-8}</td>
</tr>
<tr>
<td></td>
<td>AGRE</td>
<td>810 families</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a Discovery.
^b Replication.
^c Included in Ref. [83] and reported in Ref. [100].
^d Imputed SNPs. Table is based on Ref. [88].
A genome-wide scan for common alleles affecting risk for autism

Richard Anney1, Lamberto Testa2, Dallia Pinto3, Regina Regan4, Judith Conroy5, 
Tiago R. Magalhaes4,5, Catarina Correia4,5, Brett S. Ahmed5, Nuala Sykes6, 
Alistair T. Pagamenta6, Ioana Almeida7, Elena Bacchiell8, Anthony J. Bailey1,1, 
Gillian Bailey1,1, Agatino Battaglia1,1, Tom Brereton1,1, Nadia Bolshakova1,1, Sven Bölte1,1, 
Patrick F. Bolton1,1, Thomas Bourgeron1,1, Sean Brennan1, Jessica Brian1, Andrew R. Carson1, 
Guilermo Casale9, Jilian Casey7, Su H. Chu9, Lynne Cochrane1,1, Cirianda Corsello1,1, 
Emily L. Crawford1,1, Andrew Cross1,1, Geraldine Dawson1,1,2,3,4, Marette de Jonge1,4, 
Richard Delorme1,5, Irene Dmoch2,3, Effie Dhulkis1,5, Frederico Duque9, Annette Estes5, 
Perry Farrar1,5, Bridget A. Fernandez1,5, Susan E. Falstein1,5, Éric Fombonne1,5, 
Christine M. Freitag1,5,6, John Gilbert1,5, Christopher Gillberg1,5, Joseph T. Gleesner1,5, 
Jeremy Goldberg1,5, Jonathan Green1,5, Stephen J. Gut1,5, Hakon Hakonarson1,5,1,1, 
Elisabeth A. Heron1,1, Matthew Hill1,3, Richard Holt1,5, Jennifer L. Howe1,5, Gillian Hughes1,1, 
Vanessa Hus1,5, Roberta Igloi1,5, Cecilia Kim1,5, Sabine M. Klack1,3,4,2, Alexander Kolevzon1,1, 
Olga Korvatska2,2, Vlad Kustanovich1,1, Clara M. Lajonchere1,1, Janine A. Lamb3, 
Magdalena Laskawiec1,5, Marion Leboyer1,5, Ann Le Coutre1,5, Bennett L. Leventhal1,4,5,6, 
Anath C. Lionel1,5, Xiao-Qing Liu1,5, Catherine Löst1,5, Linda Loitspeich1,1, Sabata C. Lund1,1, 
Elena Maestrini1,1, William Mahoney1,5, Carine Mantoulan1,1, Christian R. Marsh1,1, 
Helen McConachie1,1, Christopher J. McDougall1,5, Jane McGrath1,5, William M. McMahon1,1, 
Nadine M. Mehen1,5, Alison Merikangas1,5, Ohsuke Migita1,5, Nancy J. Minshew1,1,3,1,2, 
Ghazala M. Mirza1,5, Jeff Munson1,5, Stanley F. Nelson1,5,7, Carolyn Nosek1,5, Abdul Noor10, 
Gudrun Nygren1,5, Guiomar Oliveira1,5, Katerina Papenikou1,5, Jeremy R. Parr1,5, 
Barbara Parrini1,5, Tara Petron1,5, Andrew Pickles1,5, Joseph Piven1,5,6, David J. Posey1,5, 
Annamarie Pousta1,5, Fritz Pousta1,5, Aparna Prasad1,5, Jannis Ragoussi1,5, Katy Renshaw1,1, 
Jessica Rickaby1,5, Wendy Roberts1,5, Kathryn Roeder1,5, Bernadette Roge1,5, Michael L. Rutter1,5, 
Laura J. Bierut1,5, John P. Rice1,5, Jeff Salt1,5, Katherine Sansom1,5, Daiane Satô1,5, 
Ricardo Segurado1,5, Lili Seriman1,5, Naisha Shah1,5, Val C. Sheffield1,5, Lathe Soorya1,5, Inês Sousa1,5, 
Vera Stopf1,5, Christina Strawbridge1,5, Raffaella Tancredi1,5, Katherine Tansey1,5, 
Bhooma Thiruvahindrapu1,5, Ann P. Thompson1,5, Susanne Thomson1,5, Ana Tryfon1,5, 
John Tsiantis1,5, Herman Van Engeland1,5, John B. Vincent1,5, Fred Volkmar1,5, Simon Wallace1,5, 
Wai Wang1,5, Zhowhui Wang1,5, Thomas W. Wassef1,5,6, Kirsty Wing1,5, Kerstin Wittmeyer1,5, 
Shawn Wood1,5, Brian L. Yaspian1,5, Danielle Zurawiecki1,5, Lorraine Zwaigenbaum1,5, 
Catalina Belancur1,1, Joseph D. Bukowiecki1,1, Rita M. Cantor1,5,5, Edwin H. Cook1,1,1,2,3,1,2,3,4, 

Although autism spectrum disorders (ASDs) have a substantial genetic basis, most of the known genetic risk has been traced to rare variants, principally copy number variants (CNVs). To identify common risk variation, the Autism Genome Project (AGP) Consortium genotyped 1558 rigorously defined ASD families for 1 million single-nucleotide polymorphisms (SNPs) and analyzed these SNP genotypes for association with ASD. In one of four primary association analyses, the association signal for marker rs4141463, located within MACROD2, crossed the genome-wide association significance threshold of \( P < 5 \times 10^{-8} \). When a smaller replication sample was analyzed, the risk allele at rs4141463 was again over-transmitted; yet, consistent with its effect size, the replication sample was much smaller; and, for the combined samples, the association signal barely fell below the \( P < 5 \times 10^{-8} \) threshold. Exploratory analyses of phenotypic subtypes yielded no significant associations after correction for multiple testing. They did, however, yield strong signals within several genes, KIAA0554, PLD5, POU6F2, ST8SIA2 and TAF1C.
An Uncharacterized Long Non-Coding RNA is Highly Expressed Near rs4141463

**RPS10P2AS**
Ribosomal Protein S10 Pseudogene 2, Anti-sense

UCSC Genome Browser
Expression: The Uncharacterized *RPS10P2AS* is Highly Expressed in Multiple Tissues
The Non-Coding RNA *RPS10P2AS* is Increased in Expression in Postmortem Autism Brain

*RPS10P2AS* 5-fold increase  
P=0.002

*MACROD2*  
No Sign Diff
Expression of the Non-Coding RNA RPS10P2AS is Correlated with the Autism-Associated rs4141463 C/C Genotype
**RPS10P2AS Summary**

- Under an autism GWAS peak
- Expression is increased in postmortem brains of individuals with autism
- Increased expression is correlated with the autism GWAS allele
- Function? …
Outline

• Introduction to Non-Coding RNAs
• Overview of Autism Genetics
• Genome-Wide Association Study results
  – Point to non-coding RNAs
• Exome Sequencing results
  – Point to a transcription factor that regulates … non-coding RNAs
• A New Type of Pharmacology Targets Non-Coding RNAs
Multiplex Targeted Sequencing Identifies Recurrently Mutated Genes in Autism Spectrum Disorders

Brian J. O’Roak,1 Laura Vives,1 Wenqing Fu,1 Jarrett D. Egertson,1 Ian B. Stanaway,1 Ian G. Phelps,2,3 Gemma Carvill,2,3 Akash Kumar,1 Choli Lee,1 Katy Ankenman,4 Jeff Munson,4 Joseph B. Hiatt,1 Emily H. Turner,1 Roie Levy,1 Diana R. O’Day,2 Niklas Krumm,1 Bradley P. Coe,1 Beth K. Martin,1 Elhanan Borenstein,1,5,6 Deborah A. Nickerson,1 Heather C. Mefford,2,3 Dan Doherty,2,3 Joshua M. Akey,1 Raphael Bernier,4 Evan E. Eichler,1,7,* Jay Shendure1,*

Exome sequencing studies of autism spectrum disorders (ASD) have identified many de novo mutations, but few recurrently disrupted genes. We therefore developed a modified molecular inversion probe method enabling ultra-low-cost candidate gene resequencing in very large cohorts. To demonstrate the power of this approach, we captured and sequenced 44 candidate genes in 2,446 ASD probands. We discovered 27 de novo events in 16 genes, 59% of which are predicted to truncate proteins or disrupt splicing. We estimate that recurrent disruptive mutations in six genes—CHD8, DYRK1A, GRIN2B, TBR1, PTEN, and TBL1XR1—may contribute to 1% of sporadic ASD. Our data support associations between specific genes and reciprocal subphenotypes (CHD8-macrocephaly, DYRK1A-microcephaly) and replicate the importance of a β-catenin/chromatin remodeling network to ASD etiology.
**CHD8: 11 de novo LoF mutations**


<table>
<thead>
<tr>
<th>Proband</th>
<th>Sex</th>
<th>Gene</th>
<th>Mut</th>
<th>Assay</th>
<th>HGVS</th>
<th>NVIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>12714-p1</td>
<td>M</td>
<td>CHD8*</td>
<td>ms</td>
<td>MIP</td>
<td>p.Ser62X</td>
<td>78</td>
</tr>
<tr>
<td>13986-p1</td>
<td>M</td>
<td>CHD8*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Tyr74*X</td>
<td>38</td>
</tr>
<tr>
<td>11654-p1</td>
<td>F</td>
<td>CHD8*</td>
<td>sp</td>
<td>MIP&lt;sup&gt;(d)&lt;/sup&gt;</td>
<td>c.3519-2A&gt;G</td>
<td>41</td>
</tr>
<tr>
<td>13044-p1</td>
<td>M</td>
<td>CHD8*</td>
<td>ns</td>
<td>EX</td>
<td>p.Gln1238X</td>
<td>34</td>
</tr>
<tr>
<td>14016-p1</td>
<td>M</td>
<td>CHD8*</td>
<td>ms</td>
<td>MIP</td>
<td>p.Arg1337X</td>
<td>92</td>
</tr>
<tr>
<td>12991-p1</td>
<td>M</td>
<td>CHD8*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Glu2103ArgX3</td>
<td>67</td>
</tr>
<tr>
<td>12752-p1</td>
<td>F</td>
<td>CHD8*</td>
<td>fs</td>
<td>EX</td>
<td>p.Leu2120ProX13</td>
<td>93</td>
</tr>
<tr>
<td>14233-p1</td>
<td>M</td>
<td>CHD8*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Asn2371LysX2</td>
<td>19</td>
</tr>
<tr>
<td>14406-p1</td>
<td>M</td>
<td>CHD8*</td>
<td>as</td>
<td>MIP</td>
<td>p.His2491del</td>
<td>98</td>
</tr>
<tr>
<td>12999-p1</td>
<td>M</td>
<td>DYRK1A*</td>
<td>fs</td>
<td>MIP&lt;sup&gt;(d)&lt;/sup&gt;</td>
<td>p.Ile484LysX2</td>
<td>55</td>
</tr>
<tr>
<td>13850-p1</td>
<td>F</td>
<td>DYRK1A*</td>
<td>sp</td>
<td>EX</td>
<td>c.1096+1G&gt;A</td>
<td>42</td>
</tr>
<tr>
<td>13552-p1</td>
<td>M</td>
<td>DYRK1A*</td>
<td>fs</td>
<td>MIP&lt;sup&gt;(d)&lt;/sup&gt;</td>
<td>p.Ala480ProX94</td>
<td>66</td>
</tr>
<tr>
<td>11691-p1</td>
<td>M</td>
<td>GRN2B&lt;sup&gt;†&lt;/sup&gt;</td>
<td>fs</td>
<td>MIP&lt;sup&gt;(d)&lt;/sup&gt;</td>
<td>p.Ser340IleX25</td>
<td>62</td>
</tr>
<tr>
<td>13932-p1</td>
<td>M</td>
<td>GRN2B&lt;sup&gt;†&lt;/sup&gt;</td>
<td>ms</td>
<td>MIP</td>
<td>p.Cys456Tyr</td>
<td>55</td>
</tr>
<tr>
<td>12547-p1</td>
<td>M</td>
<td>GRN2B&lt;sup&gt;†&lt;/sup&gt;</td>
<td>ms</td>
<td>MIP&lt;sup&gt;(d)&lt;/sup&gt;</td>
<td>p.Trp559X</td>
<td>65</td>
</tr>
<tr>
<td>12681-p1</td>
<td>F</td>
<td>GRN2B&lt;sup&gt;†&lt;/sup&gt;</td>
<td>sp</td>
<td>EX</td>
<td>c.2172-2A&gt;G</td>
<td>65</td>
</tr>
<tr>
<td>14433-p1</td>
<td>M</td>
<td>PTEN</td>
<td>ms</td>
<td>MIP</td>
<td>p.Thr131Ile</td>
<td>50</td>
</tr>
<tr>
<td>14611-p1</td>
<td>M</td>
<td>PTEN</td>
<td>fs</td>
<td>MIP</td>
<td>p.Cys136MetX44</td>
<td>33</td>
</tr>
<tr>
<td>11390-p1</td>
<td>F</td>
<td>PTEN</td>
<td>ms</td>
<td>EX</td>
<td>p.Thr167Asn</td>
<td>77</td>
</tr>
<tr>
<td>12335-p1</td>
<td>F</td>
<td>TBL1XR1*</td>
<td>ms</td>
<td>EX</td>
<td>p.Leu282Pro</td>
<td>47</td>
</tr>
<tr>
<td>14612-p1</td>
<td>M</td>
<td>TBL1XR1*</td>
<td>fs</td>
<td>MIP</td>
<td>p.Ile397SerX19</td>
<td>41</td>
</tr>
<tr>
<td>11480-p1</td>
<td>M</td>
<td>TBR1&lt;sup&gt;†&lt;/sup&gt;</td>
<td>fs</td>
<td>EX</td>
<td>p.Ala136ProX80</td>
<td>41</td>
</tr>
<tr>
<td>13814-p1</td>
<td>M</td>
<td>TBR1&lt;sup&gt;†&lt;/sup&gt;</td>
<td>ms</td>
<td>MIP</td>
<td>p.Lys228Glu</td>
<td>78</td>
</tr>
<tr>
<td>13796-p1</td>
<td>F</td>
<td>TBR1&lt;sup&gt;†&lt;/sup&gt;</td>
<td>fs</td>
<td>MIP&lt;sup&gt;(d)&lt;/sup&gt;</td>
<td>p.Ser351X</td>
<td>63</td>
</tr>
</tbody>
</table>

*Part of 49-member connected component reported in (3). †Part of expanded 74-member connected component. *Primary assay that identified the variant. †Proband was exome sequenced by cited study and variant was not reported or reported. Variant reported in MIP screen from (3).

... and association with autism P<10⁻⁸
... note: next best genes have 3 de novo LGD

CHD8 = Chromodomain Helicase DNA-Binding Protein 8

- Known to ...
  - *Chd8* knockout mouse is embryonic lethal before a brain appears (Nishiyama et al. 2004. *Mol Cell Biol.*)
  - CHD8 protein is known to interact with a handful of other proteins in cancer cells
    - Androgen Receptor (Menon et al. 2010. *Mol Endocrinology.*)
- Bottom line: CHD8 interacts with multiple proteins, but its function has not been studied in the brain or neurons
- We found: *CHD8* over-expression in human neuronal cell lines increased expression of the non-coding RNA *MSNP1AS*
Outline

• Introduction to Non-Coding RNAs
• Overview of Autism Genetics
• Genome-Wide Association Study results
  – Point to non-coding RNAs
• Exome Sequencing results
  – Point to a transcription factor that regulates … non-coding RNAs
• A New Type of Pharmacology Targets Non-Coding RNAs
Long Non-Coding RNA Targeting and Transcriptional De-Repression

Marc S. Weinberg¹,² and Kevin V. Morris¹,³

Our current understanding of the molecular events that functionally characterize cellular biology continues to be revised. Recent observations find that the vast majority of the human genome is transcribed and may be functionally relevant. Many of these previously unrecognized transcripts, both short and long non-coding RNAs, have been found to be active modulators of protein coding gene function. While such observations were in the past relegated to imprinted genes, it is now becoming apparent that several different genes in differentiated cells may be under some form of non-coding RNA based regulatory control. Emerging evidence suggests that some of these long non-coding RNAs are functional in controlling gene transcription by the targeted recruitment of epigenetic silencing complexes to homology-containing loci in the genome. Most notably when these repressor non-coding RNAs are targeted using small RNA-based inhibitors (such as with RNA interference), a derepression of the targeted gene can occur resulting in activation of gene expression. Knowledge of this emerging RNA based epigenetic regulatory network has implications not only in cellular evolution but also for the development of an entirely new area of pharmacology.
Proposed Mechanism of *MSNP1AS*
Long Non-Coding RNA

- **Chromosome X DNA**
  - Transcription
  - Chromosome X Moesin RNA
  - Translation
  - Moesin Protein

- **Chromosome 5 DNA**
  - Transcription
  - Chromosome 5 Non-Coding RNA
  - Antisense
  - Chromosome 5 Non-Coding RNA Binds Chromosome X Moesin RNA
  - Alters Translation of Moesin Protein
“Antisense to the Antisense”

Chromosome X DNA

Transcription

Chromosome X Moesin RNA

Translation

Moesin Protein

Chromosome 5 DNA

Transcription

Chromosome 5 Non-Coding RNA

Binds Chromosome X Moesin RNA

Alters Translation of Moesin Protein
Conclusions

Non-coding RNAs are abundant in human brain

Non-coding RNAs are the functional elements revealed by autism GWAS

Non-coding RNAs are treatment targets in HIV, cancer, diabetes
The Team

Campbell Lab (ZNI):
Nicole Grepo
Brent Wilkinson
Jessica DeWitt
Grace Kim
Christina Zdawczyk
Ani Misirian
Emily Holmes
Sarah Danehower
Kasey Rivas
Tara Kerin
Anita Ramanathan
Elisabeth Rutledge
Ranjita Raghavan
Young Kim
Giovanni Dandekar
Elizabeth Cortez-Toledo
Rachel Marshall

USC Collaborators:
Pat Levitt
Heather Volk
Gerry Coetzee
Jim Knowles
Carlos Pato
Kai Wang
Oleg Evgrafov
Marcelo Coba
Barbara Thompson
Wange Lu
Rob McConnell
Carrie Breton

Outside Collaborators:
Kevin Morris (Scripps)
Judy Van de Water (UC-Davis)
Paul Ashwood (UC-Davis)
Isaac Pessah (UC-Davis)
Irva Hertz-Picciotto (UC-Davis)
Susan Swedo (NIMH)
Audrey Thurm (NIMH)
Lisa Croen (Kaiser Permanente)

Funding:
NIMH
Autism Speaks

Resources:
Autism Tissue Program