LINEAGEN

GOLDEN HELIX
Accelerating the Quest for Significance™

THE UNIVERSITY OF UTAH

The Children's Hospital of Philadelphia®

New Study Identifies High-Risk Variants Associated with Autism Spectrum Disorders

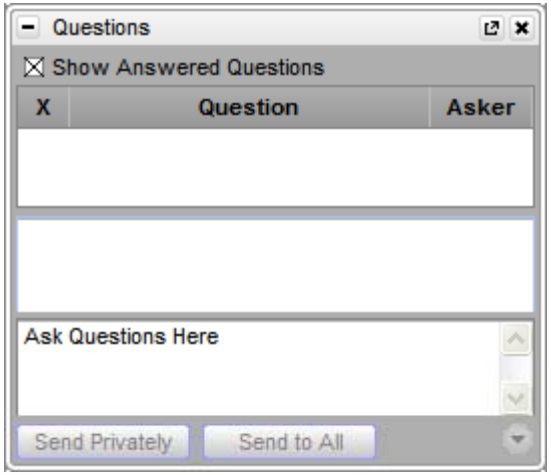
Twenty-four new variants discovered, each conferring more than a 2-fold risk of developing ASD

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Q&A

During the presentation, ask questions

Use the Questions pane in your GoToWebinar window



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Speakers & Agenda

					
Dr. Christophe Lambert	**Dr. Michael Paul**	**Dr. Hakon Hakonarson**	**Dr. Mark Leppert**	**Dr. Bryce Christensen**	**Dr. Charles Hensel**
CEO at Golden Helix	*President and Chief Executive Officer at Lineagen*	*Director of the Center for Applied Genomics at the Children's Hospital of Philadelphia*	*Professor of Human Genetics at the University of Utah and Chief Science Advisor at Lineagen*	*Director of Services and Statistical Geneticist at Golden Helix*	*Senior Research Manager at Lineagen*
1. Introduction of presenters and agenda	**2. Background on Lineagen**	**3. The science behind Autism Spectrum Disorders (ASDs)**	**4. Family-based genetics of ASDs**	**5. The analytic process**	**6. Study results**

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Dr. Michael Paul

2. Background on Lineagen

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Lineagen, Inc.

Private-Stage Molecular Diagnostics Company

- **Customized** clinical testing to accelerate and enhance the diagnostic evaluation of ASD and other neurological disorders
- **2007** – Incorporated with venture capital backing from Sanderling Ventures and Signal Peak Ventures (previously vSpring Capital)
- **2007/2009** – completed pedigree-based CNV and next-generation sequence variant discovery programs in ASD and MS using Golden Helix as genetic data and predictive analytics partner

Lineagen, Inc. | Salt Lake City

- **2010** – Launched commercial genetic testing and counseling business for individuals with ASD and other disorders of childhood development
- **2011/2012** – Sold **more than 1400 tests** and experienced **98%** year-over-year revenue growth first two years on the market
- **2013** – Launched customized genetic test that incorporates validated genetic variants from recent PLOS ONE publication (Matsunami et al., 2013)

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Prevalence of ASD continues to rise dramatically and genes are a significant contributor to etiology

- Between **10-15%** of children are thought to have a developmental disability
- Autism is the fastest-growing developmental disability, with historic annual growth rates reaching **10 – 17%**
- Prevalence of autism has been recently revised by the CDC to about **1:88** children, up from **1:150** in 2002
- Genes are one of the only **scientifically** validated factors shown to be causative for Autism

Significant economic burden with $35 billion being spent annually by society
$3.2 million in lifetime cost of care for an individual with autism

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Genetic testing in ASD is recommended and significantly influences clinical management

American College of Medical Genetics Recommended Practice Guidelines
September 2010

Chromosomal microarray testing influences medical management
Genetics in Medicine
September 2011

Chromosomal microarray (CMA) testing for copy number variations (CNV) is recommended for individuals with:

A. Autism spectrum disorders (ASD)
B. Apparently non-syndromic developmental delay (DD)/intellectual disability (ID)
C. Multiple anomalies

- Avoidance of additional testing
- Improved access to treatment services
- Medical screening recommendations (**to perform** appropriate screening or **to stop** previously recommended screening)
- Recurrence risk counseling
- Referral to medical specialists
- Clarify a clinical diagnosis with a genetic diagnosis

CMA Genetic Testing is covered by many private insurance companies

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Early intervention leads to significant improvement in cognition and life-time achievement

Randomized, controlled clinical trials show benefit of treatment in children as young as 18 months

PEDIATRICS

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Randomized, Controlled Trial of an Intervention for Toddlers With Autism: The Early Start Denver Model

Geraldine Dawson, Sally Rogers, Jeffrey Munson, Milani Smith, Jamie Winter, Jessica Greenson, Amy Donaldson and Jennifer Varley

Pediatrics 2010;125;e17-e23; originally published online Nov 30, 2009;
DOI: 10.1542/peds.2009-0958

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://www.pediatrics.org/cgi/content/full/125/1/e17

PEDIATRICS is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. PEDIATRICS is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 2010 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 0031-4005. Online ISSN: 1098-4275.

American Academy of Pediatrics
DEDICATED TO THE HEALTH OF ALL CHILDREN™

Downloaded from www.pediatrics.org by guest on November 18, 2010

- Early Intensive Behavioral Intervention
  - Sustained IQ gains of 20 points
  - Normal education placement

Treatment is so effective that 32 States legislatively mandate that private insurance companies pay for ASD treatment

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Partnered with world-leading academic institutions to translate discoveries into patient care

- Genetic Discoveries Licensed From The Children's Hospital of Philadelphia (CHOP) and University of Utah (Utah)
- CHOP genetic variants may account for up to **15%** of autism cases
  - Published in high-impact peer-reviewed publications, including **Nature**[1]
  - Named as one of Time Magazine **Top 10 Medical Breakthroughs** in 2009[2]
- Over **2000** novel genetic variants discovered using the Utah Population Database resource
  - More disease-causing genes, and more successful commercial genetic tests, have been discovered in Utah than in any other place world-wide
- ***Most comprehensive set of proprietary genetic markers associated with ASD***

MYRIAD

FAMILION
Genetic Tests for Inherited Cardiac Syndromes

Sources: [1]Nature, Glessner et al, 2009, Nature, Wang et al, 2009, Bucan et al, PLoS Genetics, 2009
[2]http://www.time.com/time/specials/packages/article/0,28804,1945379_1944376_1944378,00.html)

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Completed one of largest genetic validation studies in ASD to confirm clinical relevance of discoveries

- **9,000** subject autism genetic validation study performed in collaboration with The Children's Hospital of Philadelphia, University of Utah, and Golden Helix that validated novel genetic variants in ASD with **OR > 2**
- Study validated **24 novel** CNV genetic markers that were not previously identified in literature and **31** previously reported markers
- Aggregate sensitivity estimates from proprietary markers is approximately **5.6%**
- Represents an immediate **two-fold** increase in ASD-sensitivity over other chromosomal microarray tests
- Publication **does not include** additional genetic variants discovered by CHOP, which may account for an additional increases in test sensitivity once further validation studies have been completed

OPEN ACCESS Freely available online

PLOS ONE

Identification of Rare Recurrent Copy Number Variants in High-Risk Autism Families and Their Prevalence in a Large ASD Population

Nori Matsunami[1☯], Dexter Hadley[2☯], Charles H. Hensel[3☯], G. Bryce Christensen[4], Cecilia Kim[2], Edward Frackelton[2], Kelly Thomas[2], Renata Pellegrino da Silva[2], Jeff Stevens[1], Lisa Baird[1], Brith Otterud[1], Karen Ho[3], Tena Varvil[1], Tami Leppert[1], Christophe G. Lambert[4], Mark Leppert[1], Hakon Hakonarson[2,5]*

1 Department of Human Genetics, University of Utah, Salt Lake City, Utah, United States of America, 2 Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States of America, 3 Lineagen, Inc., Salt Lake City, Utah, United States of America, 4 Golden Helix, Inc., Bozeman, Montana, United States of America, 5 Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, United States of America

individuals with ASD and three have only been observed once. Finally, we confirmed the association of 31 of 185 published ASD-associated CNVs in our dataset with odds ratios greater than 2.0, suggesting they may be of clinical relevance in the evaluation of children with ASDs. Taken together, these data provide strong support for the existence and application of high-impact CNVs in the clinical genetic evaluation of children with ASD.

Citation: Matsunami N, Hadley D, Hensel CH, Christensen GB, Kim C, et al. (2013) Identification of Rare Recurrent Copy Number Variants in High-Risk Autism Families and Their Prevalence in a Large ASD Population. PLoS ONE 8(1): e52239. doi:10.1371/journal.pone.0052239

Editor: Michael Edward Zwick, Emory University School Of Medicine, United States of America

Received September 6, 2012; **Accepted** November 9, 2012; **Published** January 14, 2013

Copyright: © 2013 Matsunami et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: All Utah subjects were ascertained and DNA collected with support from R01 MH 06359 from the National Institute of Mental Health and U19HD035476 from the National Institute of Child Health and Human Development. DNA was processed with support from GCRC M01-RR025764 from the National Center for Research Resources. The Autism Genetic Resource Exchange is a program of Autism Speaks and is supported, in part, by grant 1U24MH081810 from the National Institute of Mental Health to Clara M. Lajonchere (PI). Dr. Hakonarson is additionally supported by the Margaret Q. Landenberger Foundation. Additional funding for this study was provided by Lineagen, Inc. Scientific input into study design, data analysis, and preparation of the manuscript were provided by two authors who are Lineagen employees (CHH, KH). The remaining funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Bryce Christensen and Christophe Lambert are paid employees of Golden Helix Inc., which derives commercial revenue from the SNP & Variation Suite software used for data analysis for this publication. Nori Matsunami, Charles Hensel and Mark Leppert have stock options in Lineagen, Inc. These affiliations do not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: hakonarson@email.chop.edu

☯ These authors contributed equally to this work.

Introduction

Twin studies [1–3], (reviewed in [4]), family studies [5–7], and reports of chromosomal aberrations in individuals with ASD (reviewed in [8]) all have strongly suggested a role for genes in the development of ASD. Although the magnitude of the genetic effect observed in ASD varies from study to study, it is clear that genetics plays a significant role.

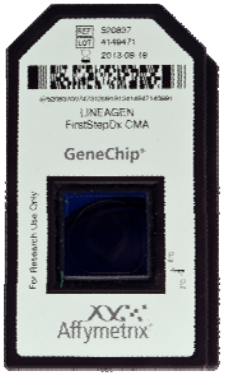
While a number of genes associated with ASD susceptibility have been observed in multiple studies, variants in a single gene cannot explain more than a small percentage of cases. Indeed, recent estimates suggest that there may be nearly 400 genes or chromosomal regions involved in ASD predisposition [9–12].

In the past few years, a number of studies have identified both *de novo* and inherited structural variants, including CNVs, that are associated with ASD [13–23]. *De novo* CNVs may explain at least some of the "missing heritability" of ASD as understood to date. While it is clear that CNVs play an important role in susceptibility to ASD, it is also clear that the genetic penetrance of many of these CNVs is less than 100%. Although many of the duplications or

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

FirstStep^{Dx} PLUS – delivers the most clinically informative results for patients

- The **most comprehensive** whole genome array clinically available
  - In partnership with Affymetrix, customized the CytoScan microarray with a unique probe design that allows for detection of novel validated genetic variants
  - Yields a **> 2x** increase in detection of Autism-related genetic variants over competitive tests
- Increased coverage of other developmental delay genetic alterations not readily detectable by competitive array platforms



Affymetrix | Powered by Affymetrix™

- On a single platform, FirstStep^{Dx} allows for maximum detection of genetic variants associated with ASD and other disorders of childhood development

The most clinically-actionable information per test result

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

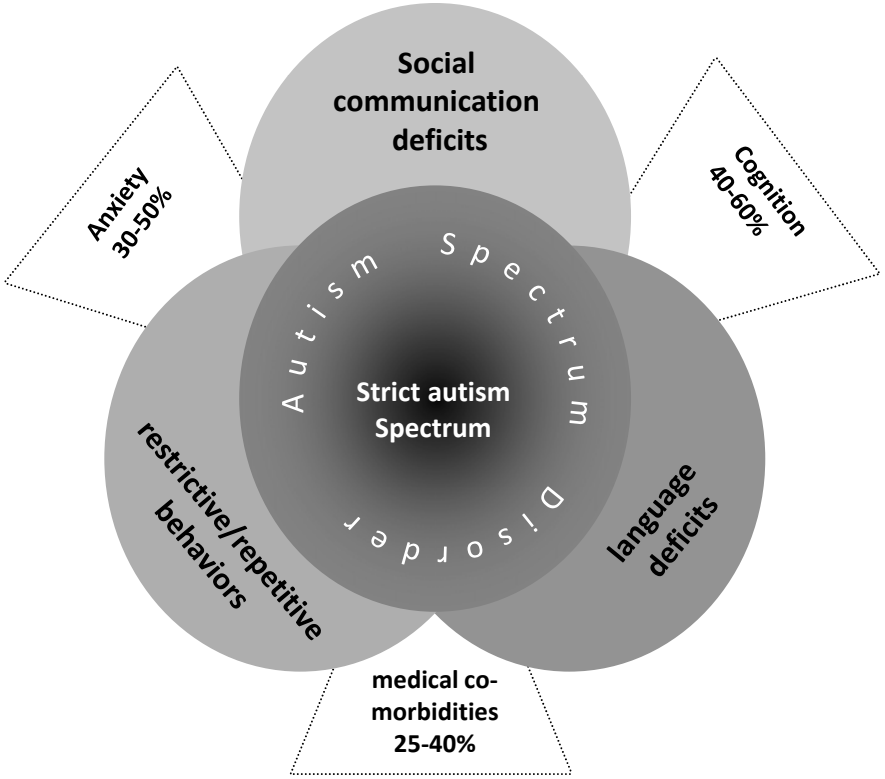
Dr. Hakon Hakonarson

3. Background on Autism Spectrum Disorders (ASDs)

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Autism Spectrum Disorders

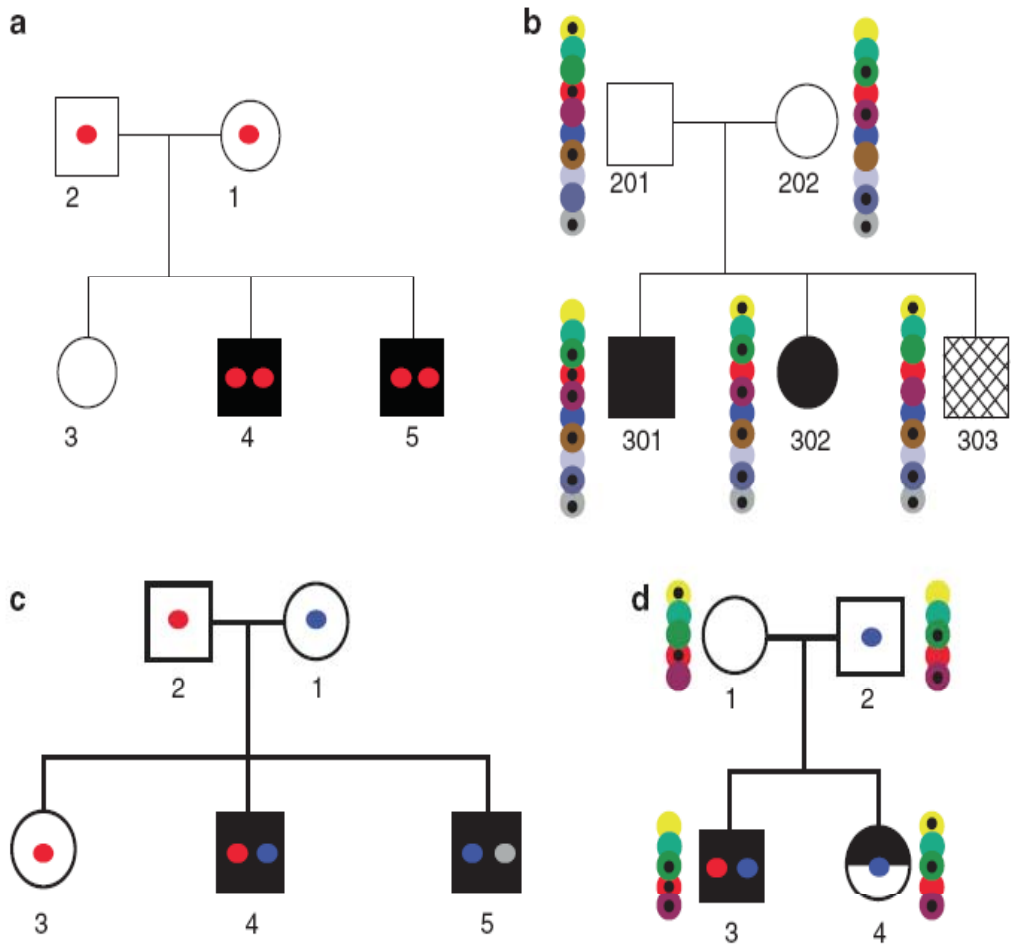
A heterogeneous 'spectrum' disorder involving deficits in 3 domains of function



- 0.9-1.0% prevalence
- ~15-20% of sibs have an ASD
- Subset of cases have genetic abnormality (rare single-gene disorders, chr. rearrangements)
- Multiple CNVs have been identified as risk factors
- Few common GWAs hits

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Genetic models in ASDs



- Different genetic models for common and rare variants in ASDs
- Penetrance is incomplete in most instances

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Diagnosis of ASDs – Domains of impairment

Domain	Autism	Asperger	PDD-NOS[b]	ASD
Social communication	Required	Required	Required	
Language	Required	–	Variable	
Repetitive and/or restrictive behaviors	Required	Required	Variable	
Sensory abnormalities[c]	>90%	80%	Variable	94%
Developmental regression[d]	15–40%	?	?	15–40%
Motor signs[e]	60–80%	60%	60%	60–80%
Gross motor delay	10%	?	?	5–10%
Sleep disturbance	55%	5–10%	40%	50%
Gastrointestinal disturbance[f]	45%	4%	50%	4–50%
Epilepsy[g]	10–60%	0–5%	5–40%	6–60%
Comorbid psychiatric diagnosis[h]	70%	60%	>25%	25–70%

Ann Rev Med, 2009

ADR-I: *Autism* Diagnostic Interview-Revised

ADOS: *Autism* Diagnostic Observation Schedule

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

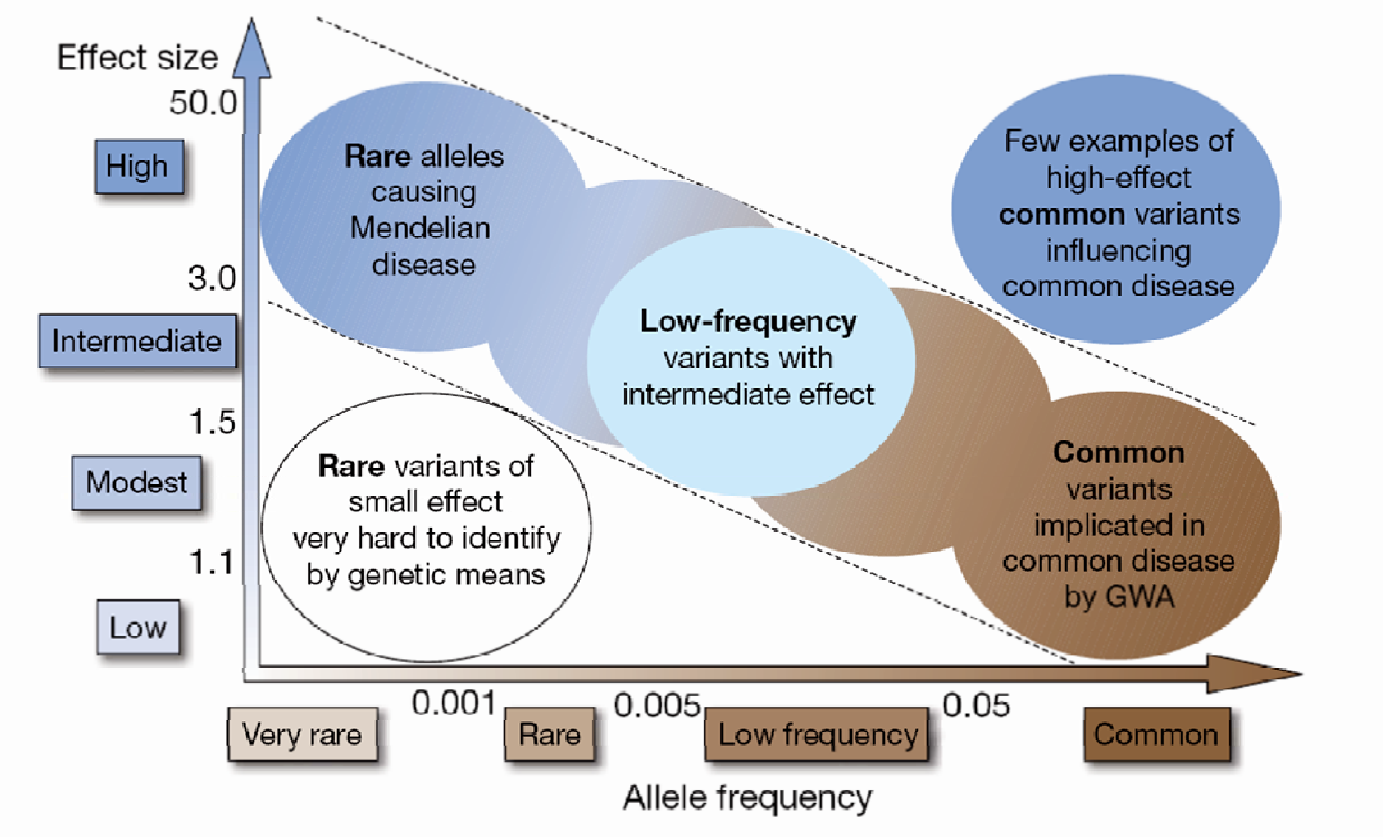
ASD-related syndromes

ASD-related syndrome	Associated gene(s)	Proportion with ASD	Proportion ASD with syndrome	References
1q21 Duplication	Many	50%	~1%?	[91, 128]
3p Deletion / duplication	*CNTN4*	<50%	~1%	[51, 61, 110]
15q Duplication (maternal)	Many (including *UBE3A, GABRB3, SNRPN,* and *SNURF*)	High	~1%	[41]
15q13 Deletion	Many (including *CHRNA7*)	<50%	Unknown	[15, 118]
16p11 Deletion	Many (including *SEZ6L2*)	High	~1%	[78, 79, 90, 144]
22q11 Deletion (aka VCFS / DiGeorge)	Many (including *TBX1* and *COMT*)	15–50%	<1%	[52, 139]
22q13 Deletion	*SHANK3*	High	~1%	[48, 89, 95]
Angelman (15q11-13)	Maternal *UBE3A*	40–80%	<1%	[22, 102]
Beckwith Weidemann (11p15)	*IGF2* and *CDKN1C*	~7%	Unknown	[73]
Cortical dysplasia focal epilepsy (7q35-36)	*CNTNAP2*	70%	Negligible	[68, 125]
Cowden/BRRS (10q23)	PTEN	20%	>10% with macrocephaly	[101, 135]
Down (trisomy chr.21)	Many	6–15%	Unknown	[86]
Fragile X (Xq27)	*FMR1*	25% of males 6% of females	1–2%	[64]
Potocki-Lupski (17p11)	Many (including *RAI1*)	~90%	Unknown	[106]
Smith–Lemli–Optiz (11q13)	*DHCR7*	50%	Negligible	[129]
Prader–Willi (15q11-13)	Paternal deletions	20–25%	Unknown	[45]
Rett (Xq26)	*MECP2*	N/A	~0.5%	[5]
Timothy (12p13)	*CACNA1C*	60–80%	Negligible	[120]
Tuberous sclerosis (9q34 and 16p13)	*TSC1, TSC2*	20%	~1%	[10]

- Multiple syndromes have ASD characteristics
- Fragile-X is the most common cause of autism (1-2%)
  - Molecular mechanism been well established

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

The genetic landscape in complex disease



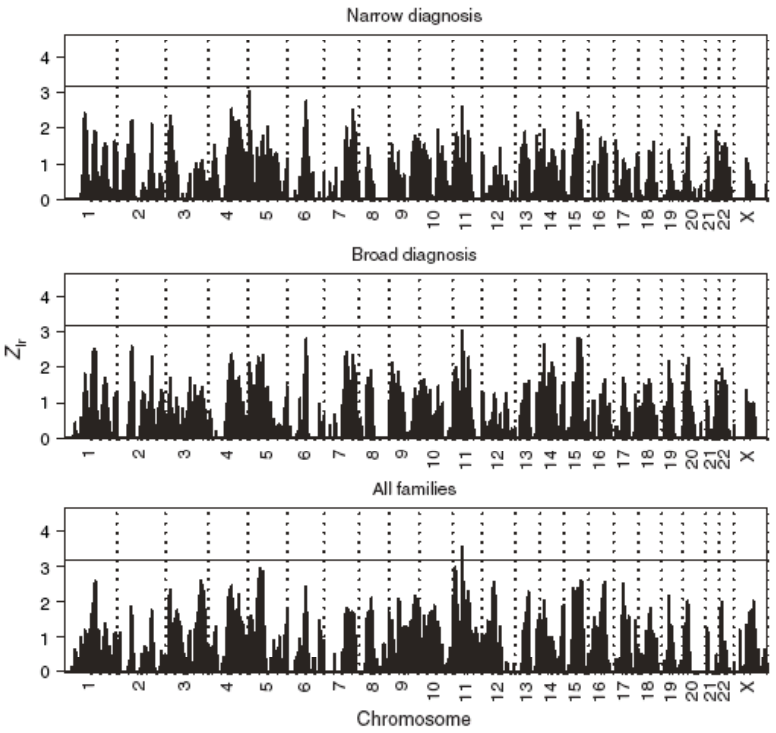
"Rare variants could be the primary drivers of common diseases."

- Nat Rev Genet. 2010

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

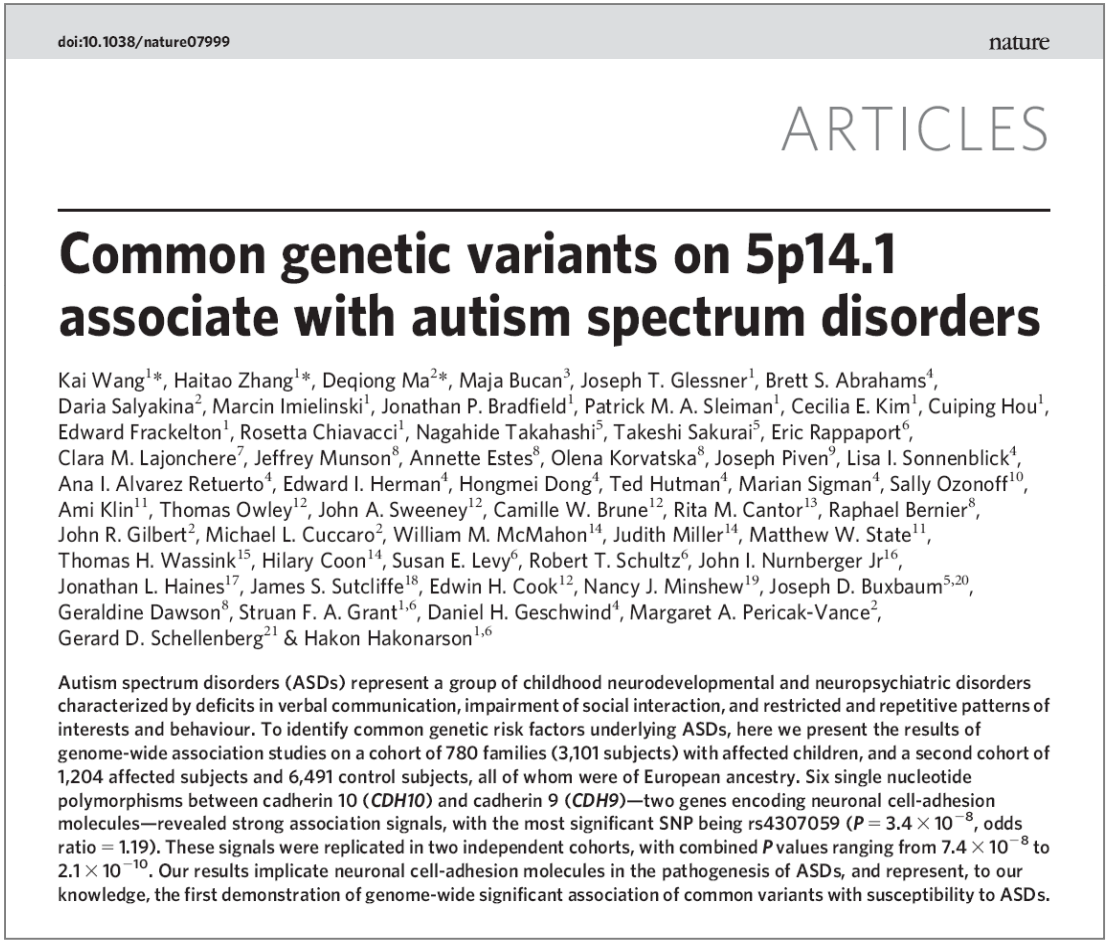
A GWAS odyssey in autism since 2007

- Autism Genome Project Consortium (Nat Gen 2007)
  - 10K SNP arrays
  - suggest 11p12-13 and neurexins
  - detect microdeletions and duplications in ASD families
- Sebat et al (Science 2007)
  - array CGH
  - Describe sub-microscopic *de novo* CNVs
- Weiss et al (NEJM 2008)
  - Affy 5.0
  - 16p11.12 micro-CNV is a high penetrance risk factor.
- Marshall et al (AJHG, 2008)
  - 427 cases and 500 controls
  - 277 unbalanced CNVs in 44% of ASD families not present in 500 controls (27 were de novo)



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

First common variant in ASDs



doi:10.1038/nature07999 nature

ARTICLES

Common genetic variants on 5p14.1 associate with autism spectrum disorders

Kai Wang[1]*, Haitao Zhang[1]*, Deqiong Ma[2]*, Maja Bucan[3], Joseph T. Glessner[1], Brett S. Abrahams[4], Daria Salyakina[2], Marcin Imielinski[1], Jonathan P. Bradfield[1], Patrick M. A. Sleiman[1], Cecilia E. Kim[1], Cuiping Hou[1], Edward Frackelton[1], Rosetta Chiavacci[1], Nagahide Takahashi[5], Takeshi Sakurai[5], Eric Rappaport[6], Clara M. Lajonchere[7], Jeffrey Munson[8], Annette Estes[8], Olena Korvatska[8], Joseph Piven[9], Lisa I. Sonnenblick[4], Ana I. Alvarez Retuerto[4], Edward I. Herman[4], Hongmei Dong[4], Ted Hutman[4], Marian Sigman[4], Sally Ozonoff[10], Ami Klin[11], Thomas Owley[12], John A. Sweeney[12], Camille W. Brune[12], Rita M. Cantor[13], Raphael Bernier[8], John R. Gilbert[2], Michael L. Cuccaro[2], William M. McMahon[14], Judith Miller[14], Matthew W. State[11], Thomas H. Wassink[15], Hilary Coon[14], Susan E. Levy[6], Robert T. Schultz[6], John I. Nurnberger Jr[16], Jonathan L. Haines[17], James S. Sutcliffe[18], Edwin H. Cook[12], Nancy J. Minshew[19], Joseph D. Buxbaum[5,20], Geraldine Dawson[8], Struan F. A. Grant[1,6], Daniel H. Geschwind[4], Margaret A. Pericak-Vance[2], Gerard D. Schellenberg[21] & Hakon Hakonarson[1,6]

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 7.4×10^{-8} to 2.1×10^{-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.

- First common gene variant identified and replicated in ASDs
- T risk-allele next to CHD10 and CHD9 is present in 65% of children with autism
- CHD10 expressed in frontal lobe of brain, synaptic function/connectivity
- Neuronal cell adhesion molecules enriched is ASD

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

CHD10 locus

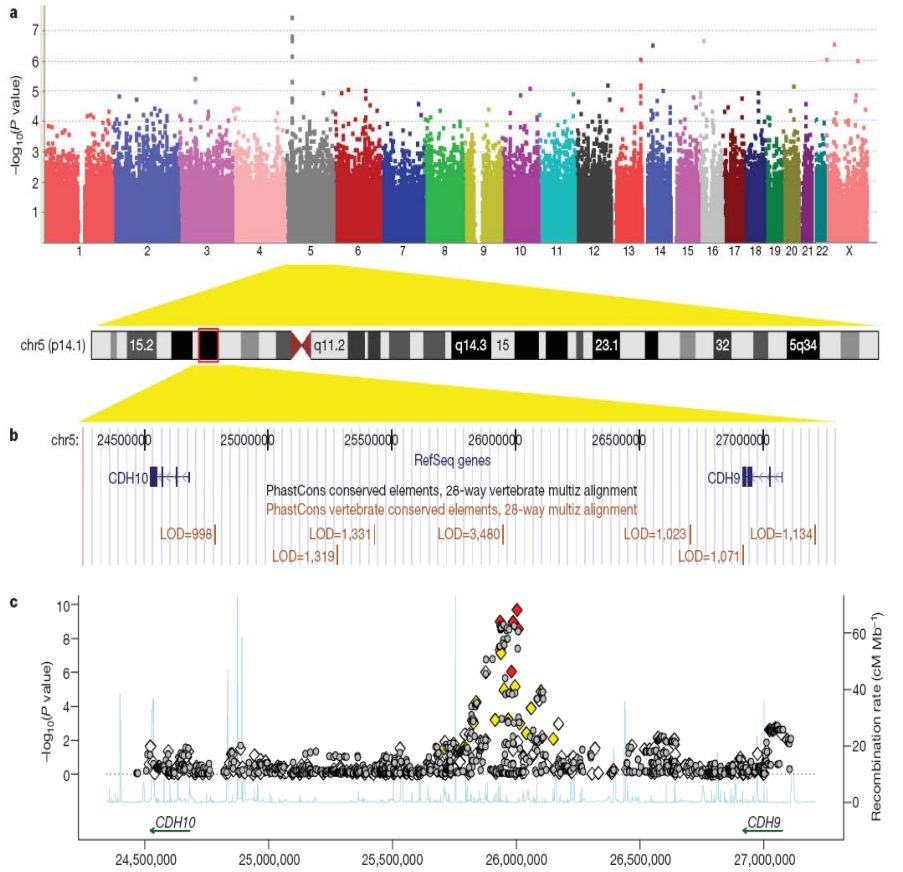


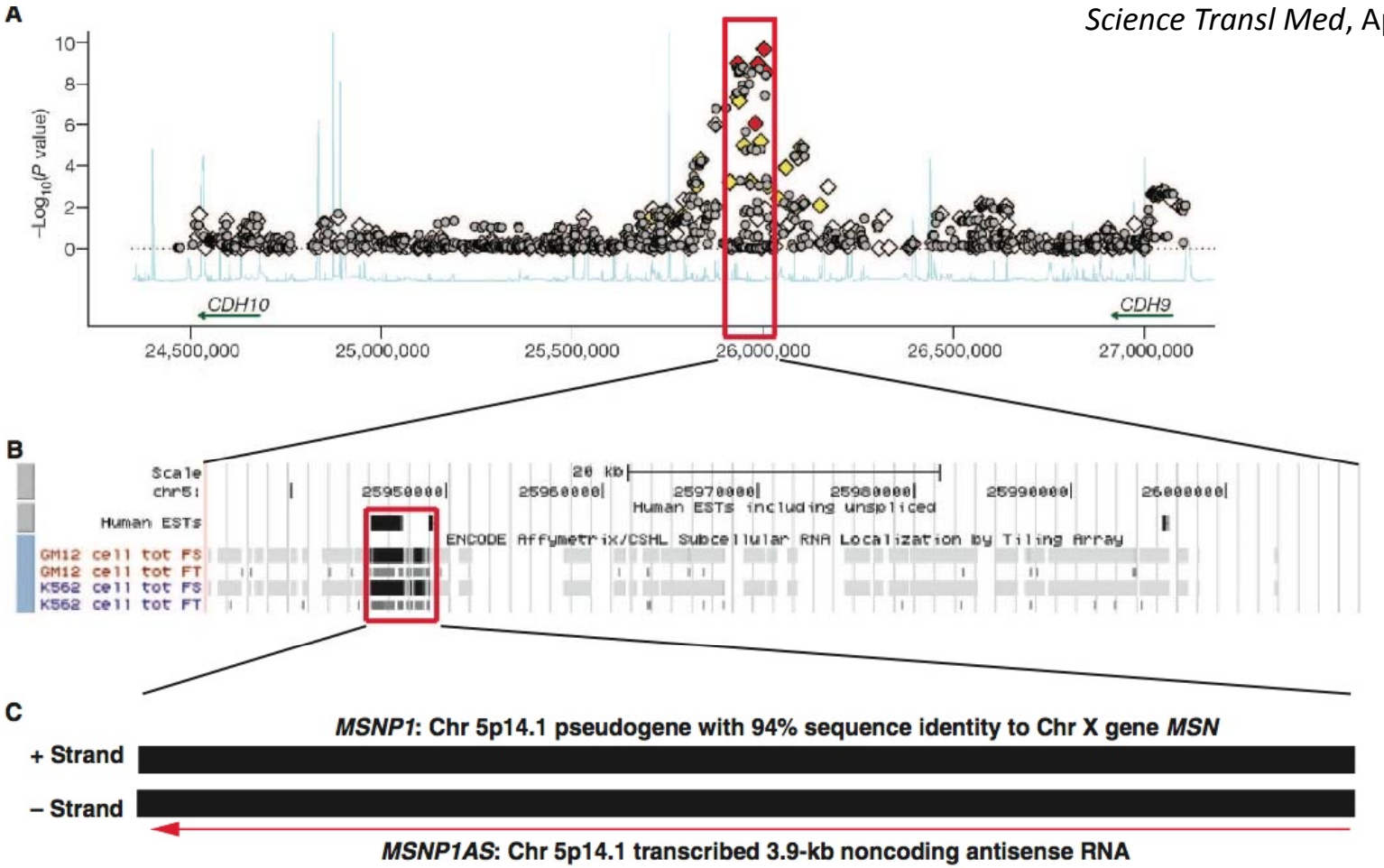
Figure 1 | Genome-wide association results for the 5p14.1 region. **a**, A Manhattan plot showing the $-\log_{10}(P$ values) of SNPs from the combined association analysis of the AGRE and ACC cohorts. **b**, The 5p14.1 genomic region is displayed in UCSC Genome Browser, with conserved genomic elements in the PhastCons track. **c**, Both genotyped (diamonds) and imputed (grey circles) SNPs are plotted with their combined *P* values in all four cohorts. Genotyped SNPs were coloured on the basis of their correlation with rs4307059 (red: $r^2 \geq 0.5$; yellow: $0.2 \leq r^2 < 0.5$; white: $r^2 < 0.2$). Estimated recombination rates from HapMap data are plotted to reflect the local linkage disequilibrium structure.

- Strong association in the intergenic region on chr 5p14.1
- Association replicated in several independent cohorts

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Autism locus on 5p14

Science Transl Med, April 2012

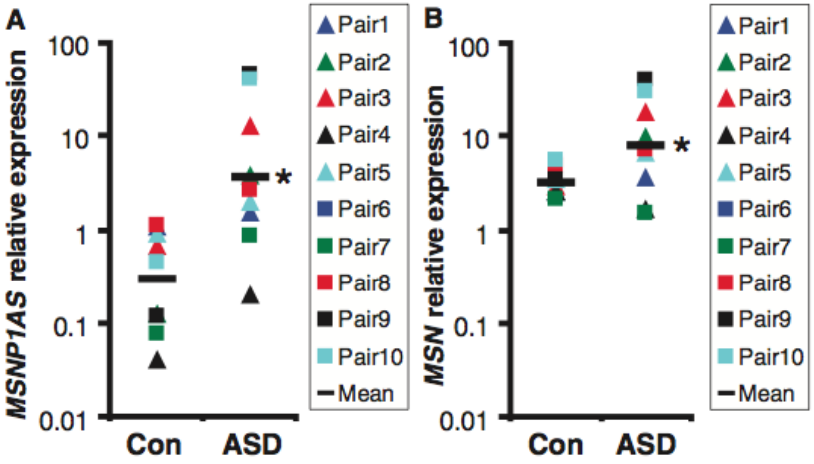


LD block containing core signal, with a few conserved elements

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Autism 5p14 locus

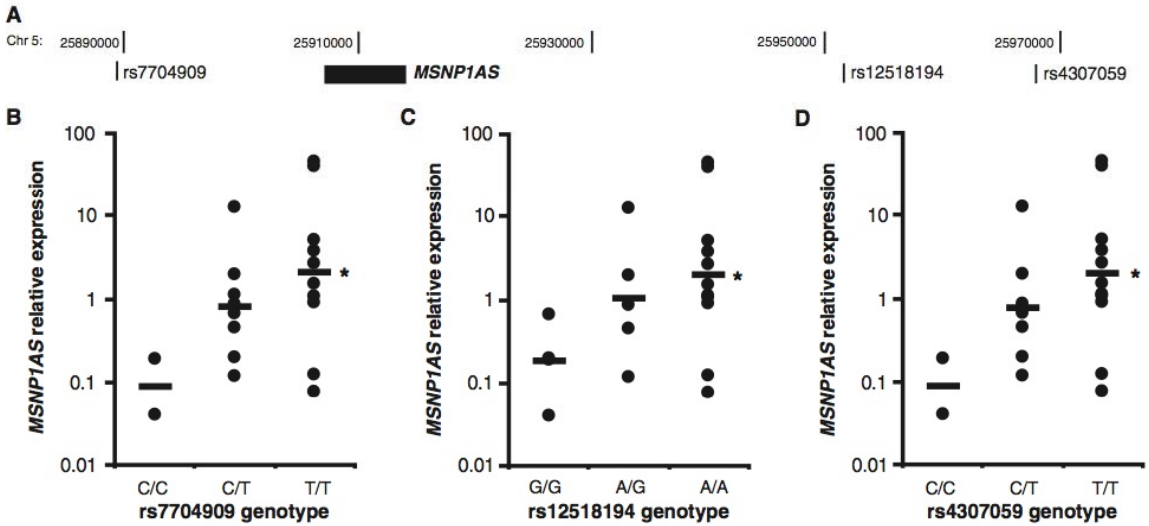
Science Transl Med, April 2012



Expression of MSNP1AS in brain 12.7-fold higher in ASD vs control

MSN 2.4-fold higher in ASD brain vs control

MSNP1AS expression correlated with ASD associated genotype in postmortem brain



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Moesin in ASD

- Moesin key regulator of neuronal architecture
- Knockdown with antisense RNA neurons results in:
  - growth cone collapse
  - suppressed neurite formation
  - 10-fold reduction in neurite advancement rate
  - suppression of glutamate-induced increase in active presynaptic boutons
  - suppression of estrogen-induced increase in the formation of dendritic spines
- Decreased moesin at critical developmental stages could:
  - contribute to altered short and long-range connectivity in the brains of individuals with ASD
  - early brain overgrowth and later reduction in brain size beginning at 2 to 3 years in ASD

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Novel Autism Variants

LETTER

doi:10.1038/nature10945

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

Stephan J. Sanders[1], Michael T. Murtha[1], Abha R. Gupta[2]*, John D. Murdoch[1]*, Melanie J. Raubeson[1]*, A. Jeremy Willsey[1]*, A. Gulhan Ercan-Sencicek[1]*, Nicholas M. DiLullo[1]*, Neelroop N. Parikshak[3], Jason L. Stein[3], Michael F. Walker[1], Gordon T. Ober[1], Nicole A. Teran[1], Youeun Song[1], Paul El-Fishawy[1], Ryan C. Murtha[1], Murim Choi[4], John D. Overton[4], Robert D. Bjornson[5], Nicholas J. Carriero[5], Kyle A. Meyer[6], Kaya Bilguvar[7], Shrikant M. Mane[8], Nenad Šestan[6], Richard P. Lifton[4], Murat Günel[7], Kathryn Roeder[9], Daniel H. Geschwind[3], Bernie Devlin[10] & Matthew W. State[1]

Multiple studies have confirmed the contribution of rare *de novo* copy number variations to the risk for autism spectrum disorders[1–3]. But whereas *de novo* single nucleotide variants have been identified in affected individuals[4], their contribution to risk has yet to be clarified. Specifically, the frequency and distribution of these mutations have not been well characterized in matched unaffected controls, and such data are vital to the interpretation of *de novo* coding mutations observed in probands. Here we show, using whole-exome sequencing of 928 individuals, including 200 phenotypically discordant sibling pairs, that highly disruptive (nonsense and splice-site) *de novo* mutations in brain-expressed genes are associated with autism spectrum disorders and carry large effects. On the basis of mutation rates in unaffected individuals, we demonstrate that multiple independent *de novo* single nucleotide variants

systematic bias in variant detection between affected and unaffected siblings through comparisons of silent *de novo*, non-coding *de novo*, and novel transmitted variants (Fig. 1a; Supplementary Figs 1–5; Supplementary Information).

Among 200 quartets (Table 1), 125 non-synonymous *de novo* SNVs were present in probands and 87 in siblings: 15 of these were nonsense (10 in probands; 5 in siblings) and 5 altered a canonical splice site (5 in probands; 0 in siblings). There were 2 instances in which *de novo* SNVs were present in the same gene in two unrelated probands; one of these involved two independent nonsense variants (Table 2). Overall, the total number of non-synonymous *de novo* SNVs was significantly greater in probands compared to their unaffected siblings ($P = 0.01$, two-tailed binomial exact test; Fig. 1a; Table 1) as was the odds ratio (OR) of non-synonymous to silent mutations in probands versus

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Novel Autism Variants

LETTER

doi:10.1038/nature10945

LETTER

doi:10.1038/nature10989

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

Brian J. O'Roak[1], Laura Vives[1], Santhosh Girirajan[1], Emre Karakoc[1], Niklas Krumm[1], Bradley P. Coe[1], Roie Levy[1], Arthur Ko[1], Choli Lee[1], Joshua D. Smith[1], Emily H. Turner[1], Ian B. Stanaway[1], Benjamin Vernot[1], Maika Malig[1], Carl Baker[1], Beau Reilly[2], Joshua M. Akey[1], Elhanan Borenstein[1,3,4], Mark J. Rieder[1], Deborah A. Nickerson[1], Raphael Bernier[2], Jay Shendure[1] & Evan E. Eichler[1,5]

It is well established that autism spectrum disorders (ASD) have a strong genetic component; however, for at least 70% of cases, the underlying genetic cause is unknown[1]. Under the hypothesis that *de novo* mutations underlie a substantial fraction of the risk for developing ASD in families with no previous history of ASD or related phenotypes—so-called sporadic or simplex families[2,3]—we sequenced all coding regions of the genome (the exome) for

per generation, in close agreement with our previous observations[4], yet in general, higher than previous studies, indicating increased sensitivity (Supplementary Table 2 and Supplementary Table 4)[7]. We also observed complex classes of *de novo* mutation including: five cases of multiple mutations in close proximity; two events consistent with paternal germline mosaicism (that is, where both siblings contained a *de novo* event observed in neither parent); and nine events

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Novel Autism Variants

LETTER

doi:10.1038/nature10945

LETTER

doi:10.1038/nature10989

LETTER

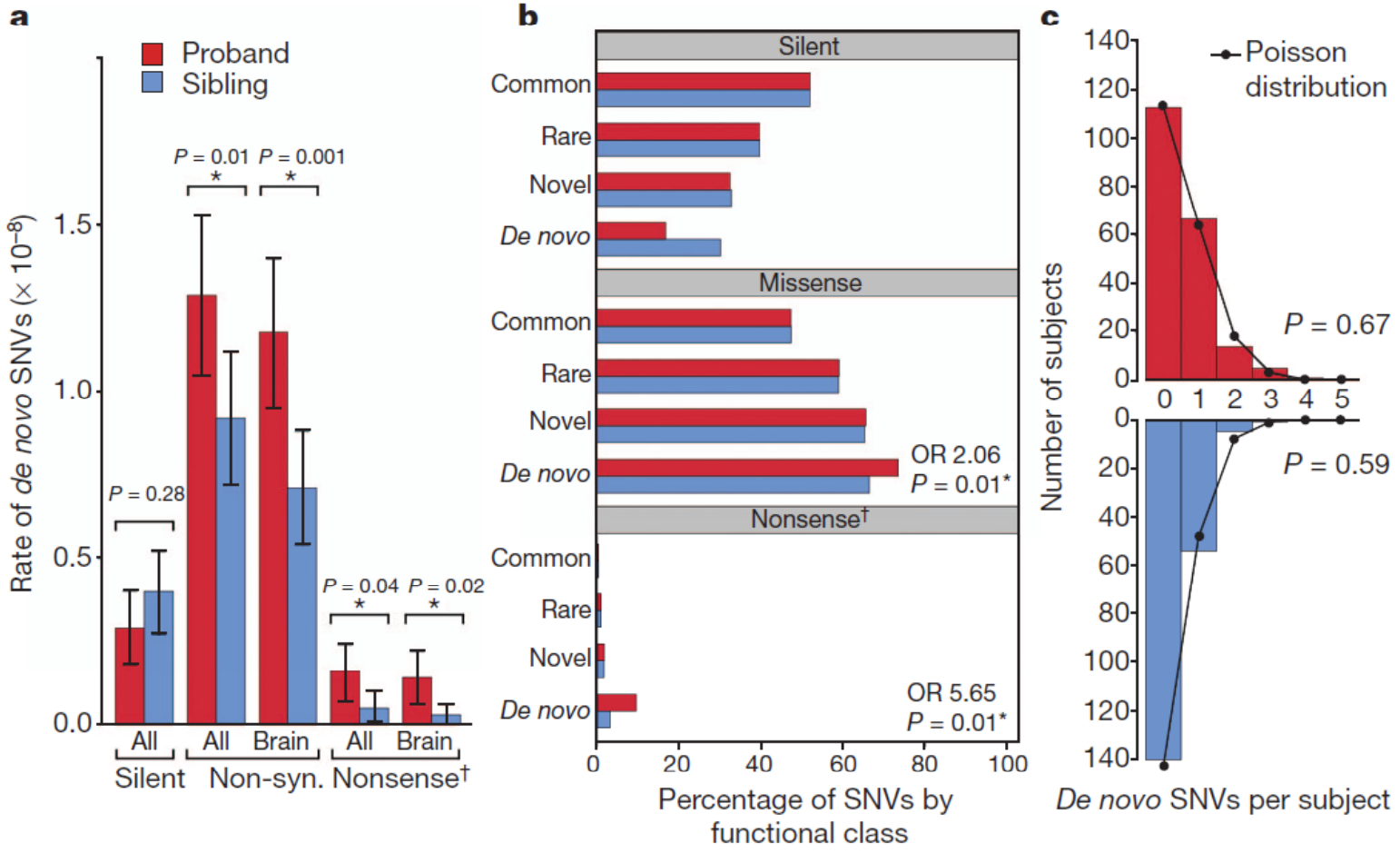
doi:10.1038/nature11011

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

Benjamin M. Neale[1,2], Yan Kou[3,4], Li Liu[5], Avi Ma'ayan[3], Kaitlin E. Samocha[1,2], Aniko Sabo[6], Chiao-Feng Lin[7], Christine Stevens[2], Li-San Wang[7], Vladimir Makarov[4,8], Paz Polak[2,9], Seungtai Yoon[4,8], Jared Maguire[2], Emily L. Crawford[10], Nicholas G. Campbell[10], Evan T. Geller[7], Otto Valladares[7], Chad Schafer[5], Han Liu[11], Tuo Zhao[11], Guiqing Cai[4,8], Jayon Lihm[4,8], Ruth Dannenfelser[3], Omar Jabado[12], Zuleyma Peralta[12], Uma Nagaswamy[6], Donna Muzny[6], Jeffrey G. Reid[6], Irene Newsham[6], Yuanqing Wu[6], Lora Lewis[6], Yi Han[6], Benjamin F. Voight[2,13], Elaine Lim[1,2], Elizabeth Rossin[1,2], Andrew Kirby[1,2], Jason Flannick[2],

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Collective finding from WES in ASDs



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

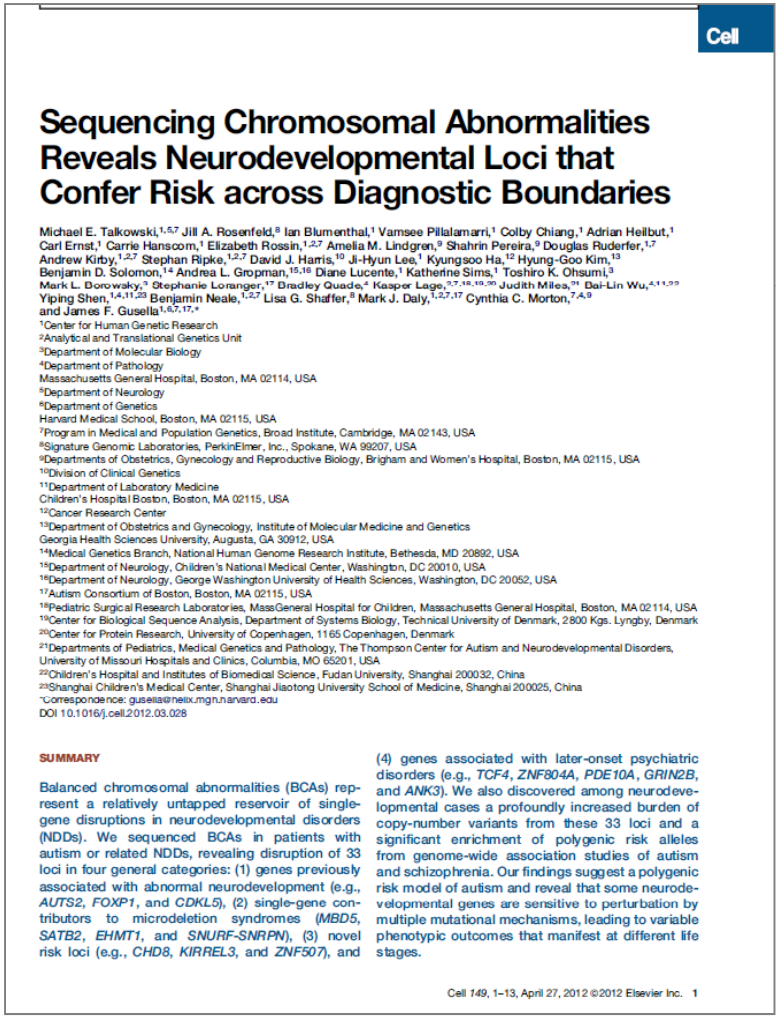
Loss of function mutations in probands

Gene symbol	Gene name	Mutation type
ADAM33	ADAM metallopeptidase domain 33	Nonsense
CSDE1	cold shock domain containing E1, RNA-binding	Nonsense
EPHB2	EPH receptor B2	Nonsense
FAM8A1	family with sequence similarity 8, member A1	Nonsense
FREM3	FRAS1 related extracellular matrix 3	Nonsense
MPHOSPH8	M-phase phosphoprotein 8	Nonsense
PPM1D	protein phosphatase, Mg^{2+}/Mn^{2+} dependent 1D	Nonsense
RAB2A	RAB2A, member RAS oncogene family	Nonsense
SCN2A	sodium channel, voltage-gated, type II, α subunit	Nonsense
SCN2A	sodium channel, voltage-gated, type II, α subunit	Nonsense
BTN1A1	butyrophilin, subfamily 1, member A1	Splice site
FCRL6	Fc receptor-like 6	Splice site
KATNAL2	katanin p60 subunit A-like 2	Splice site
NAPRT1	nicotinate phosphoribosyltransferase domain containing 1	Splice site
RNF38	ring finger protein 38	Splice site
SCP2	sterol carrier protein 2	Frameshift*
SHANK2	SH3 and multiple ankyrin repeat domains 2	Frameshift*

* Frameshift *de novo* variants are not included in any of the reported case-control comparisons

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Balanced chromosomal abnormalities in Autism



Cell

Sequencing Chromosomal Abnormalities Reveals Neurodevelopmental Loci that Confer Risk across Diagnostic Boundaries

Michael E. Talkowski,[1,5,7] Jill A. Rosenfeld,[8] Ian Blumenthal,[1] Vamsee Pillalamarri,[1] Colby Chiang,[1] Adrian Heilbut,[1] Carl Ernst,[1] Carrie Hanscom,[1] Elizabeth Rossin,[1,2,7] Amelia M. Lindgren,[9] Shahrin Pereira,[9] Douglas Ruderfer,[1,7] Andrew Kirby,[1,2,7] Stephan Ripke,[1,2,7] David J. Harris,[10] Ji-Hyun Lee,[1] Kyungsoo Ha,[12] Hyung-Goo Kim,[13] Benjamin D. Solomon,[14] Andrea L. Gropman,[15,16] Diane Lucente,[1] Katherine Sims,[1] Toshiro K. Ohsumi,[3] Mark L. Borowsky,[5] Stephanie Loranger,[17] Bradley Quade,[4] Kasper Lage,[5,7,18,19,20] Judith Miles,[21] Bai-Lin Wu,[4,11,22] Yiping Shen,[1,4,11,23] Benjamin Neale,[1,2,7] Lisa G. Shaffer,[8] Mark J. Daly,[1,2,7,17] Cynthia C. Morton,[7,4,9] and James F. Gusella[1,6,7,17,*]
[1]Center for Human Genetic Research
[2]Analytical and Translational Genetics Unit
[3]Department of Molecular Biology
[4]Department of Pathology
Massachusetts General Hospital, Boston, MA 02114, USA
[5]Department of Neurology
[6]Department of Genetics
Harvard Medical School, Boston, MA 02115, USA
[7]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA 02143, USA
[8]Signature Genomic Laboratories, PerkinElmer, Inc., Spokane, WA 99207, USA
[9]Departments of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Boston, MA 02115, USA
[10]Division of Clinical Genetics
[11]Department of Laboratory Medicine
Children's Hospital Boston, Boston, MA 02115, USA
[12]Cancer Research Center
[13]Department of Obstetrics and Gynecology, Institute of Molecular Medicine and Genetics
Georgia Health Sciences University, Augusta, GA 30912, USA
[14]Medical Genetics Branch, National Human Genome Research Institute, Bethesda, MD 20892, USA
[15]Department of Neurology, Children's National Medical Center, Washington, DC 20010, USA
[16]Department of Neurology, George Washington University of Health Sciences, Washington, DC 20052, USA
[17]Autism Consortium of Boston, Boston, MA 02115, USA
[18]Pediatric Surgical Research Laboratories, MassGeneral Hospital for Children, Massachusetts General Hospital, Boston, MA 02114, USA
[19]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[20]Center for Protein Research, University of Copenhagen, 1165 Copenhagen, Denmark
[21]Departments of Pediatrics, Medical Genetics and Pathology, The Thompson Center for Autism and Neurodevelopmental Disorders, University of Missouri Hospitals and Clinics, Columbia, MO 65201, USA
[22]Children's Hospital and Institutes of Biomedical Science, Fudan University, Shanghai 200032, China
[23]Shanghai Children's Medical Center, Shanghai Jiaotong University School of Medicine, Shanghai 200025, China
*Correspondence: gusella@helix.mgh.harvard.edu
DOI 10.1016/j.cell.2012.03.028

SUMMARY

Balanced chromosomal abnormalities (BCAs) represent a relatively untapped reservoir of single-gene disruptions in neurodevelopmental disorders (NDDs). We sequenced BCAs in patients with autism or related NDDs, revealing disruption of 33 loci in four general categories: (1) genes previously associated with abnormal neurodevelopment (e.g., *AUTS2*, *FOXP1*, and *CDKL5*), (2) single-gene contributors to microdeletion syndromes (*MBD5*, *SATB2*, *EHMT1*, and *SNURF-SNRPN*), (3) novel risk loci (e.g., *CHD8*, *KIRREL3*, and *ZNF507*), and (4) genes associated with later-onset psychiatric disorders (e.g., *TCF4*, *ZNF804A*, *PDE10A*, *GRIN2B*, and *ANK3*). We also discovered among neurodevelopmental cases a profoundly increased burden of copy-number variants from these 33 loci and a significant enrichment of polygenic risk alleles from genome-wide association studies of autism and schizophrenia. Our findings suggest a polygenic risk model of autism and reveal that some neurodevelopmental genes are sensitive to perturbation by multiple mutational mechanisms, leading to variable phenotypic outcomes that manifest at different life stages.

- Sequencing of patients with autism revealed disruption of 33 BCA loci:
  - (1) genes previously associated with abnormal neurodevelopment (e.g., AUTS2, FOXP1, and CDKL5);
  - (2) single-gene contributors to microdeletion syndromes (MBD5, SATB2, EHMT1, and SNURF-SNRPN),
  - (3) novel risk loci (e.g., CHD8, KIRREL3, and ZNF507), and
  - (4) genes associated with later-onset psychiatric disorders (e.g., TCF4, ZNF804A, PDE10A, GRIN2B, and ANK3).
- Neurodevelopmental cases have profoundly increased burden of CNVs
- Findings suggest a polygenic risk model of autism and reveal that some neurodevelop-mental genes are sensitive to perturbation by multiple mutational mechanisms, leading to variable phenotypic outcomes that manifest at different life stages.

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Genes disrupted in Autism

Table 1. Genes Disrupted by Chromosomal Rearrangements[a]

Cat	ID	Dx	ChrA	ChrB	Disrupted	Fisher's Exact p[b]	Function
1	DGAP201	ASD	7q11.22	7q36.3	*AUTS2*	5.6×10^{-4}	unknown
1 and 4	NDR27031	NDD	3q13.32	18q21.2	*TCF4*	6.2×10^{-4}	transcription factor
1	DGAP093	NDD	Xp22.13	19p13.3	*CDKL5*	7.2×10^{-2}	protein kinase
1	DGAP157	NDD	3p13	10q21.2	*FOXP1*	4.5×10^{-2}	transcription factor
1 and 4	NDR25941	ASD	12p13.1	12q21.31	*GRIN2B*	7.9×10^{-2}	glutamate receptor
1	DGAP189	NDD	11p13	12p12.1	*SOX5*	8.4×10^{-2}	transcription factor in embryonic development
2	DGAP232	ASD	9p11.2	15q11.2	*SNURF-SNRPN*	1.1×10^{-13}	genomic imprinting in angelman – pws region
2 and 4	DGAP155	ASD	9q34.3	11p11.2	*EHMT1*	3.3×10^{-7}	histone methyltransferase
2	DGAP142	ASD	2q23.1	22q13	*MBD5*	3.1×10^{-5}	methylation binding
2	DGAP211	ASD	2q33.1	6q16.3	*SATB2*	1.1×10^{-3}	transcriptional regulation and chromatin remodeling
3	DGAP148	NDD	Xp11.4	11q24.2	*KIRREL3*	1.6×10^{-4}	cell adhesion
3	DGAP154	NDD	Xq22	17p13.3	*SMG6*	5.9×10^{-4}	nonsense-mediated decay
3	NDR26867	ASD	3q25.31	14q11.2	*CHD8*	2.4×10^{-2}	chromatin remodeling
3	DGAP125	NDD	7q32.1	19q13.11	*ZNF507*	8.0×10^{-2}	zinc finger
3	DGAP132[c]	NDD	5q12.2	7q21.3	*PON3*	1.5×10^{-1}	lactonase
3	AC02-0053	ASD	6q16.1	9q21.13	*GNA14*	2.7×10^{-1}	g-protein signaling
3	DGAP131	NDD	1p22.3	5q33	*ZNHIT6*	2.7×10^{-1}	zinc finger protein
3	DGAP193	ASD	2p22.3	2q31.3	*SPAST*	2.7×10^{-1}	membrane trafficking
3 and 4	DGAP143	NDD	6q22.1	6q27	*PDE10A*	5.2×10^{-3}	phosphodiesterase
3 and 4	DGAP171	NDD	17p13.2	18p11.21	*C18orf1*	3.2×10^{-2}	unknown
3 and 4	DGAP180[c]	NDD	2q32	11q14	*ZNF804A*	4.7×10^{-2}	zinc finger protein

The following abbreviations are used: Cat, disruption category; Dx, diagnosis; ASD, autism spectrum disorder; NDD, other neurodevelopmental disorders; ChrA and ChrB = sequenced chromosomal sub-band containing the BCA. For the entire data set used to generate this table, see also Tables S1, S2, and S3.

[a]BCA-disrupted genes individually implicated by case-control CNV burden at uncorrected $p < 0.10$ or by a minimum of 3 CNVs in cases with none in controls are provided. See Table S1 and Supplemental Information for all subjects and phenotypes and Table S2 for CNV counts on all subjects.

[b]Fisher's exact test p value from comparison of CNV burden between NDD cases and controls.

[c]BCA inherited from similarly affected parent.

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

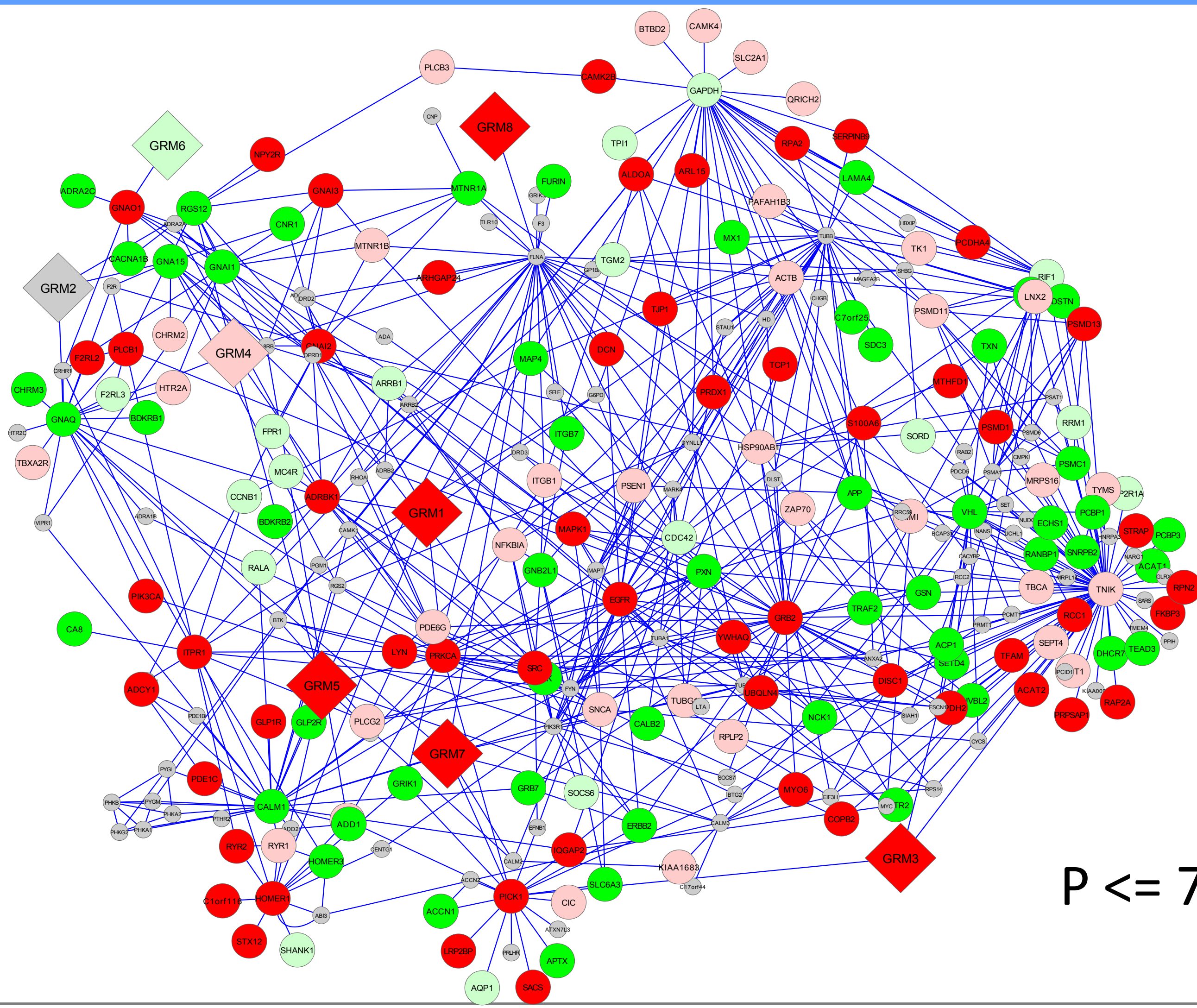
Golden Helix

The Children's Hospital of Philadelphia

The University of Utah

Lineagen

mGluR also significant in idiopathic ASD



P <= 7.64E-12

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Mutations

ARTICLE

doi:10.1038/nature10658

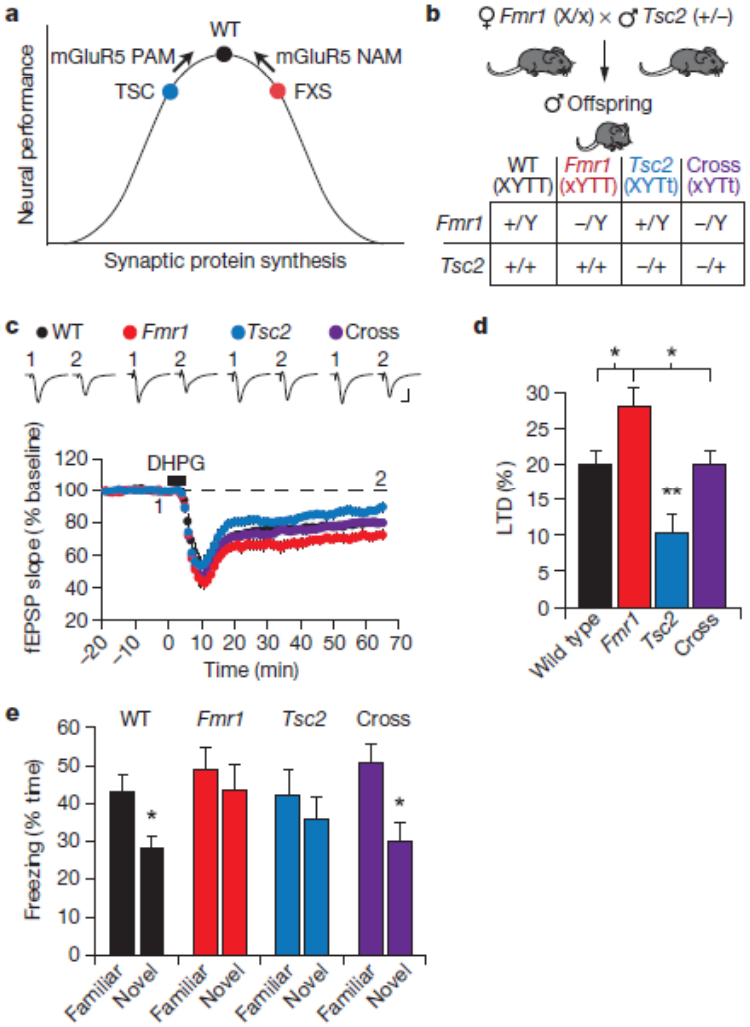
Mutations causing syndromic autism define an axis of synaptic pathophysiology

Benjamin D. Auerbach[1], Emily K. Osterweil[1] & Mark F. Bear[1]

Tuberous sclerosis complex and fragile X syndrome are genetic diseases characterized by intellectual disability and autism. Because both syndromes are caused by mutations in genes that regulate protein synthesis in neurons, it has been hypothesized that excessive protein synthesis is one core pathophysiological mechanism of intellectual disability and autism. Using electrophysiological and biochemical assays of neuronal protein synthesis in the hippocampus of *Tsc2*$^{+/-}$ and *Fmr1*$^{-/y}$ mice, here we show that synaptic dysfunction caused by these mutations actually falls at opposite ends of a physiological spectrum. Synaptic, biochemical and cognitive defects in these mutants are corrected by treatments that modulate metabotropic glutamate receptor 5 in opposite directions, and deficits in the mutants disappear when the mice are bred to carry both mutations. Thus, normal synaptic plasticity and cognition occur within an optimal range of metabotropic glutamate-receptor-mediated protein synthesis, and deviations in either direction can lead to shared behavioural impairments.

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Genetic cross of Tsc2 (+/-) and Fmr1 (-/y) mice



- Genetic cross of **Tsc2 (+/-) and Fmr1 (-/y)** mice rescues synaptic and behavioral impairments present in both single mutants
- The data suggest that optimal synaptic function requires a narrow and tightly regulated level of synaptic protein synthesis and that deviations in either direction can impair function

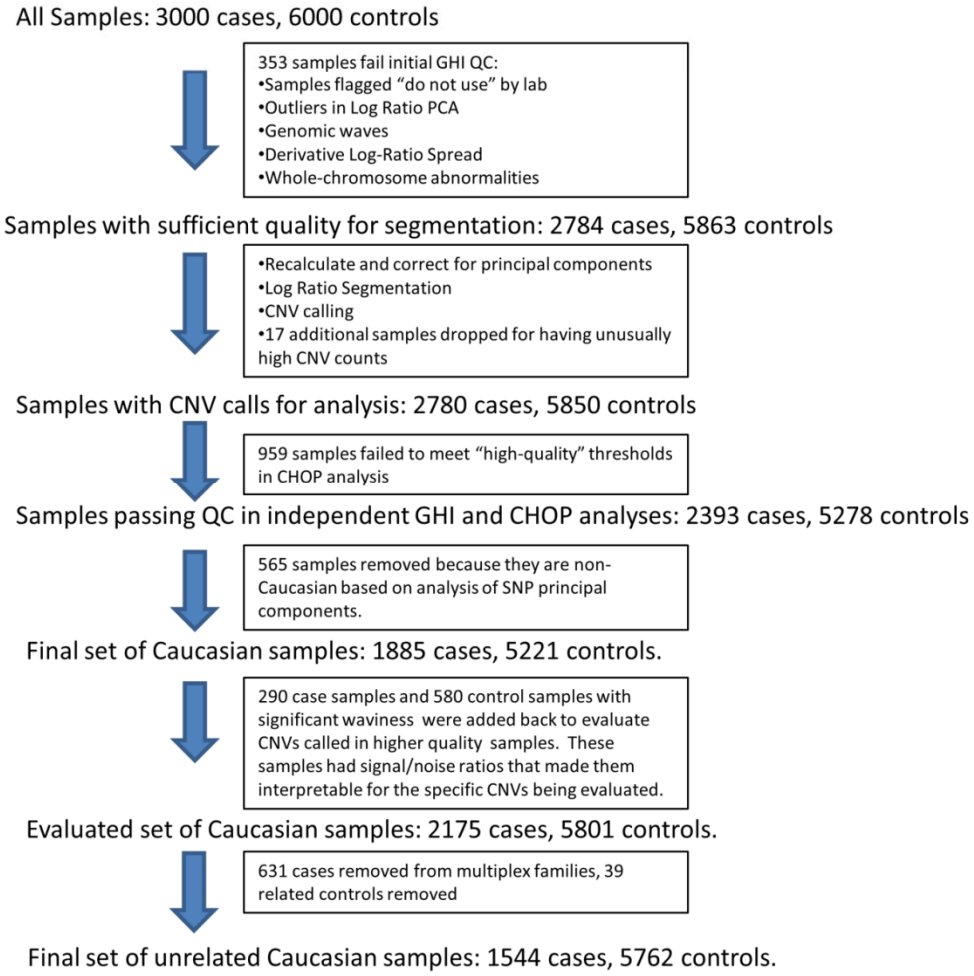
© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Summary of current knowledge

- Both common and rare variants predispose to ASDs
- Biological validation of the statistical signal at 5p14 locus
- Enrichment of nonsense and missense mutations in ASDs
- Balanced chromosomal abnormalities predispose to ASDs
- mGluR gene networks are important risk factors for ASDs
- mGluR5 loss or gain leads to neurodevelopmental phenotype spectrum in animal models – restored with Rx

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

What the current study addresses



- Still, only a small proportion of autism heritability and causality is explained
- We designed a well powered study to discover new high-impact variants and validate existing variants (>2-fold risk) in autism

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Dr. Mark Leppert

4. Family-based genetics of ASDs

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Utah Autism family discoveries

- This study was initiated six years ago and is funded by Lineagen Inc.
  - The purpose of the study was to identify causative variants in multiplex Utah families with autism spectrum disorder (ASD)
  - Identify high-impact variants
- Nine multigenerational families with a maximum of 9 affected individuals were identified
- 55 ASD individuals comprised the discovery cohort.
  - CNVs were identified utilizing the Affymetrix genome-wide human SNP array 6.0
- 153 putative CNVs were identified by the Golden Helix SVS program.
  - These CNVs were absent from Utah control samples.
  - These CNVs therefore were considered to be good candidateASD risk CNVs.
  - This set of 153 included 131 novel CNVs and 22 CNVs present in the Autism Chromosomal Rearrangement Database. Thirty-two autism specific CNVs were detected in multiple (2 or more) autism subjects, and 121 CNVs were detected in only one person among the 55 autism subjects. Of these 153 CNVs, 112 were copy number losses (deletions) and 41 were copy number gains (duplications). The average size of the CNVs was 91 kb.

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Replication in the general ASD population

- We designed a custom Illumina iSelect array containing probes covering all of the 153 Utah CNVs. Also included in this custom array were 185 autism associated CNVs culled from the literature, as well as 2,800 putative functional SNVs detected by next generation sequencing of genes in regions of haplotype sharing among the high-risk ASD families. The SNVs allowed us to identify 25 additional CNVs.
- We then carried out a large CNV replication study from an independent (non-Utah) population of 3,000 ASD cases and 6,000 typically developing controls (9,000 individuals total) using our custom Illumina iSelect array.
- We used two independent CNV calling algorithms, CNAM from Golden Helix and PennCNV from CHOP to evaluate these CNVs in our case/control study.

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

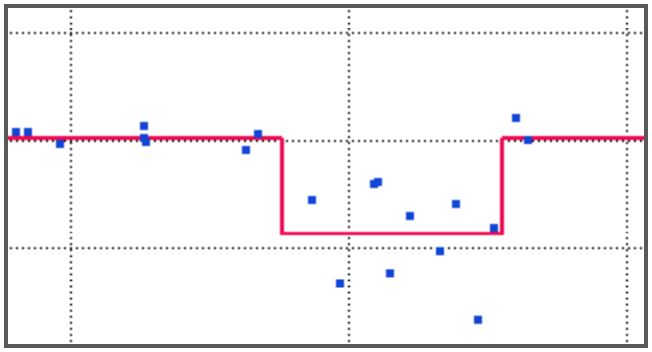
Dr. Bryce Christensen

5. The analytic process

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Laying the foundation

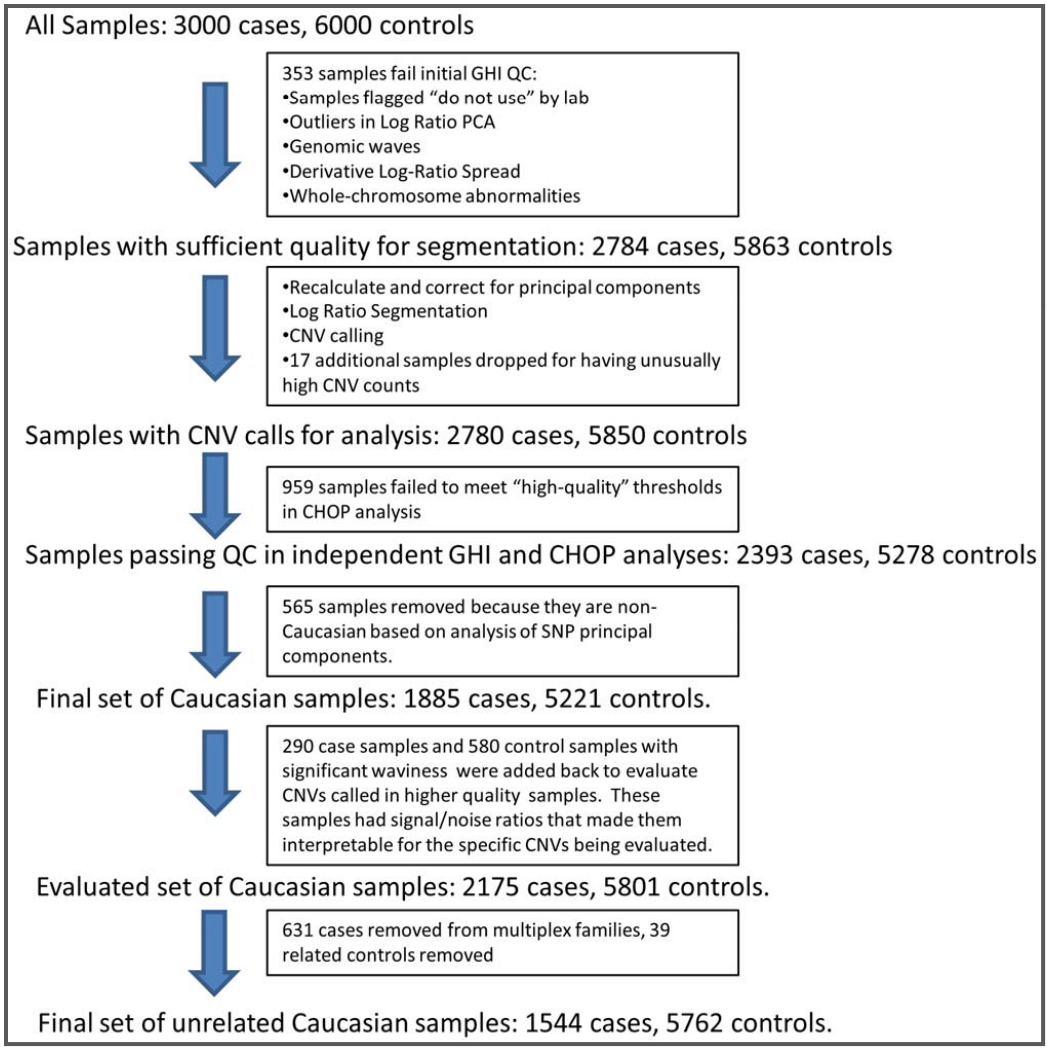
- Study analyzed targeted content from several sources:
  - CNVs from University of Utah/Lineagen WG autism analysis
    - Golden Helix assisted with the original CNV identification
  - DNA sequence variants identified in linkage regions in Utah families
  - CNVs found in previous autism research at CHOP
  - Autism CNVs identified through literature review
- Custom Illumina iSelect chip designed to assay targeted content
  - Designed chip with about 10 markers in each targeted CNV, plus about 5 flanking markers on either side.
  - Golden Helix assisted in probe design for chip
  - Final chip included about 8600 markers



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Quality control

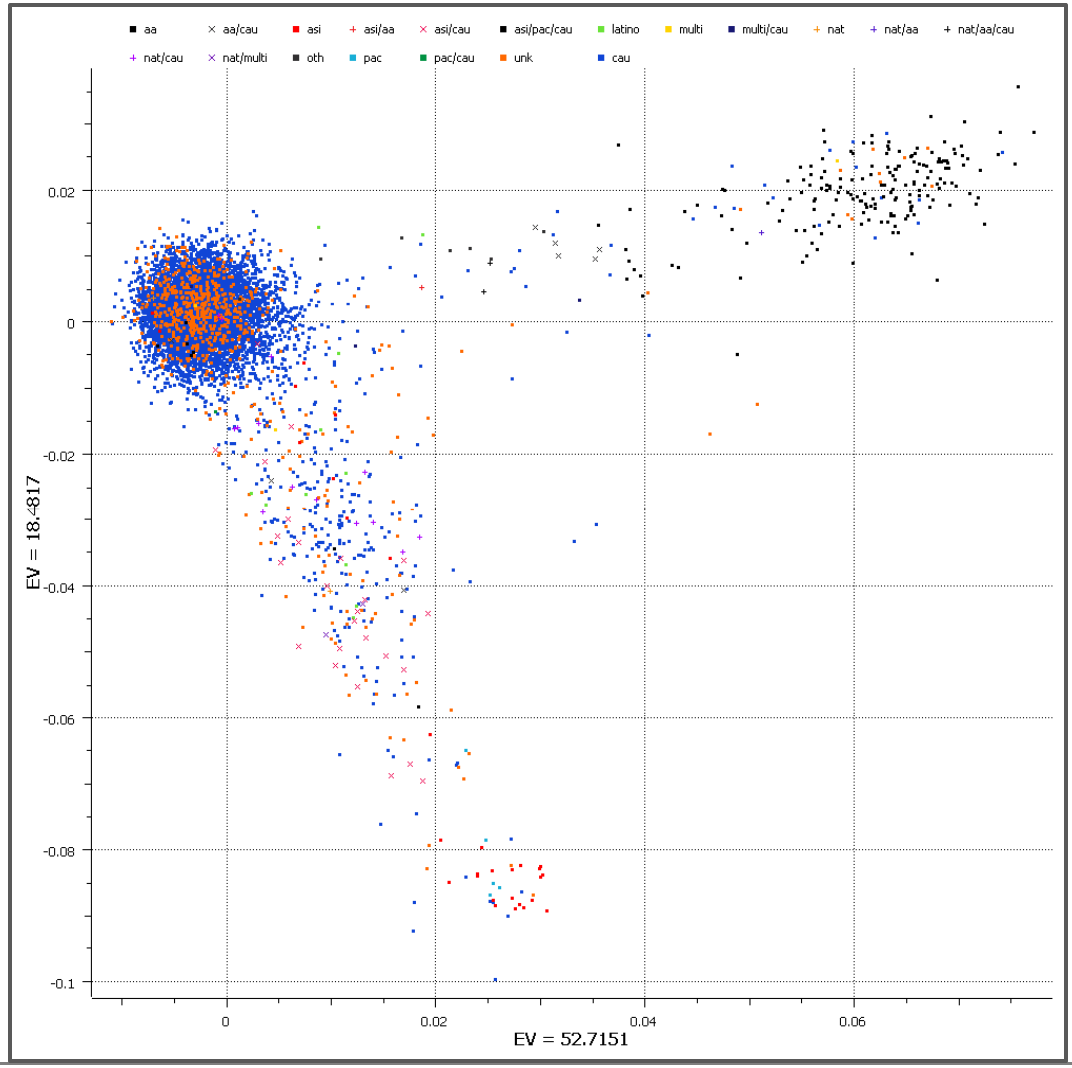
- 9000 subjects were genotyped
  - 3000 cases and 6000 controls
  - Subjects came from a variety of sources, including Utah and CHOP
- Used highest quality subjects for feature selection, included additional subjects for calculating associations and odds ratios



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

SNP principal components

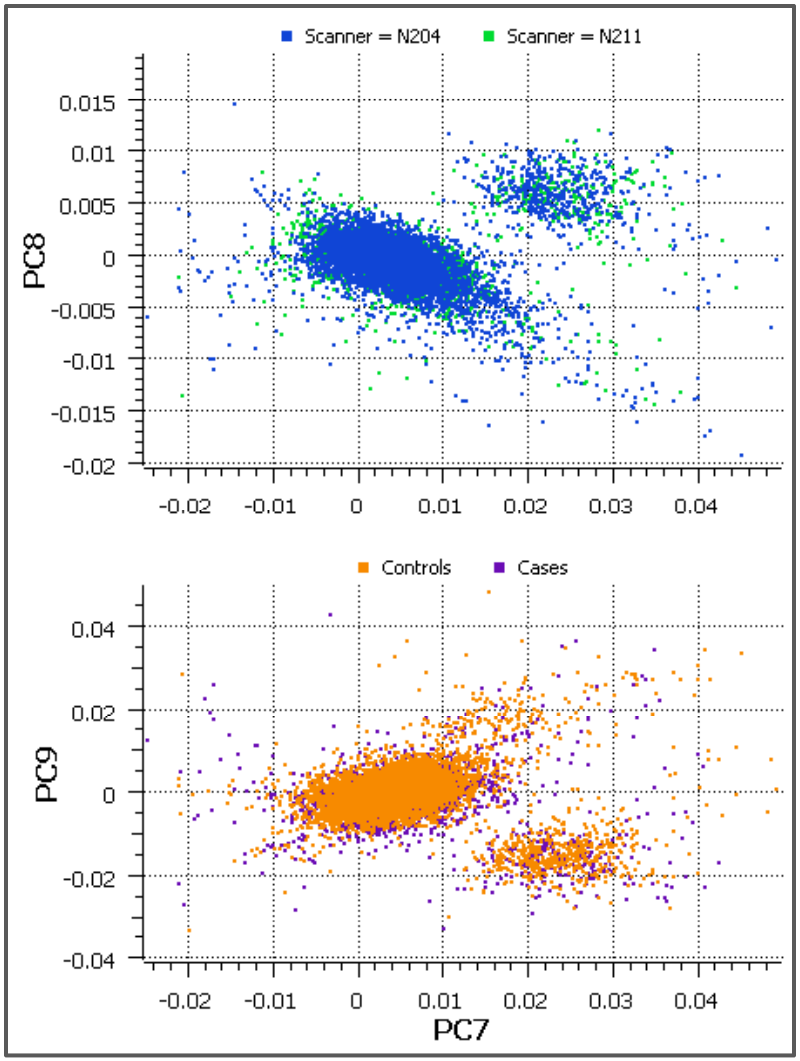
- Low number of polymorphic SNPs made ancestry estimation difficult.
  - Lesson: Include AIMs on custom chips!
- One very promising result in preliminary CNV analysis turned out to be a correlated with African ancestry.



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Log ratio principal components

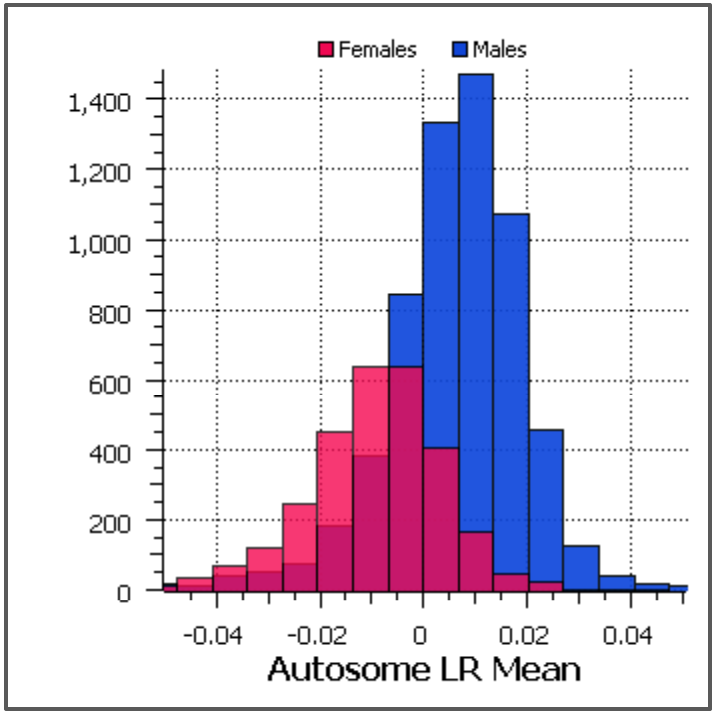
- Unusual systematic patterns were observed in the principal components of the signal intensity data used to generate CNV calls
- Some of these patterns were related to known experimental variables
- Others did not have an obvious cause
- Careful plate randomization scheme prevented any serious confounding

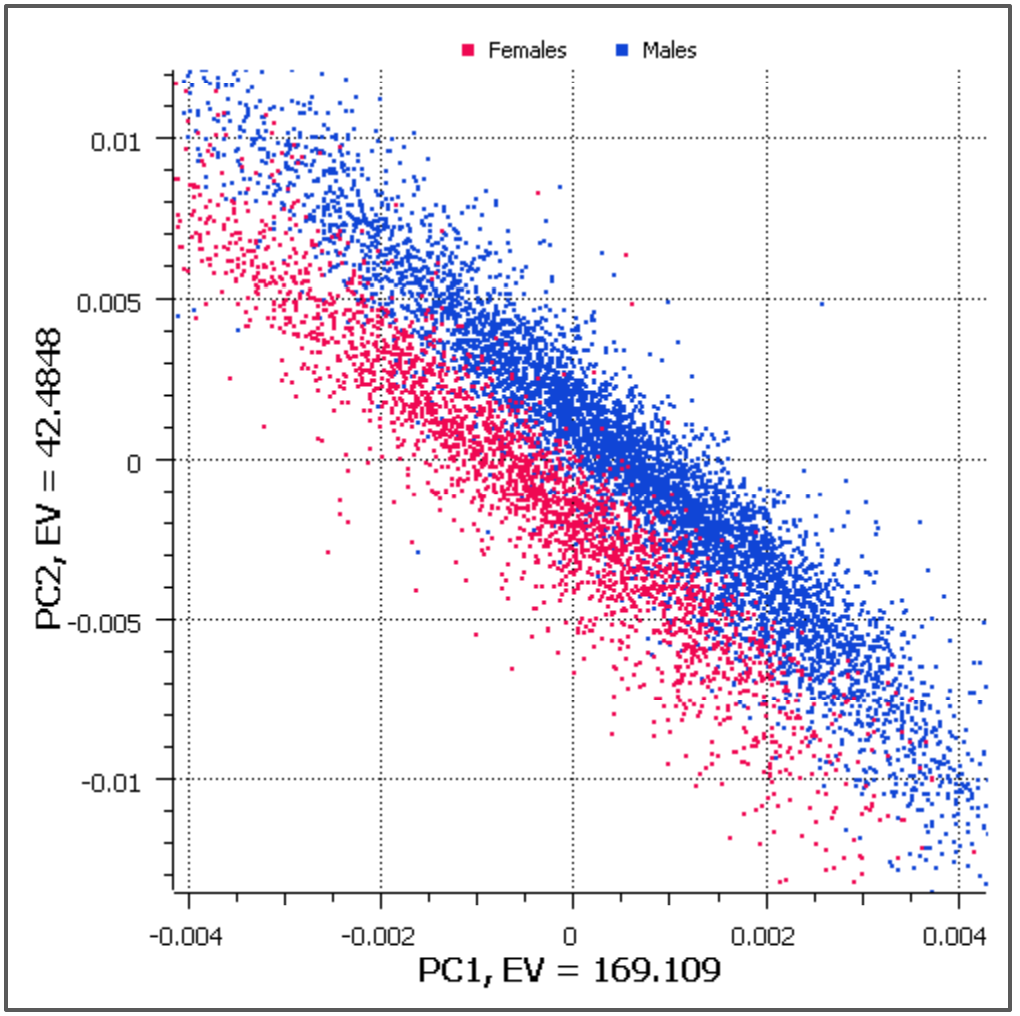


© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Gender bias in raw signal data

- This is a common issue for certain Illumina platforms
- Required special attention during analysis





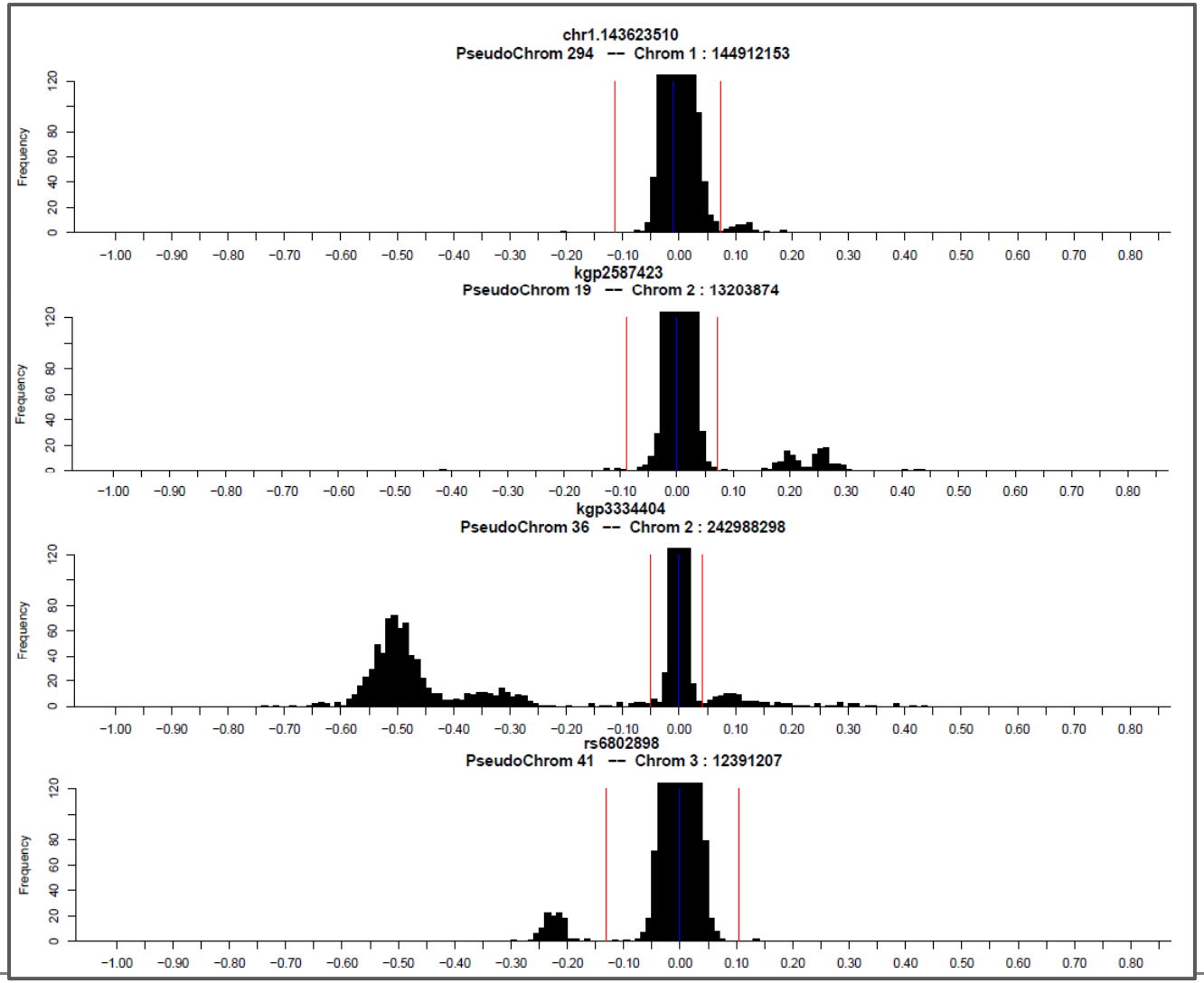
© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Parallel analyses

- CHOP Analysis
  - Used standard workflows with PennCNV software
  - Important that results could be replicated with standardized methods
- Golden Helix Analysis
  - Used Golden Helix SVS “CNAM” method with additional custom scripts
  - Manually reviewed intensity patterns at every locus to confirm correct thresholds for calling gains and losses
- Merged Results
  - Calculated P-values and odds ratios for all CNVs based on both individual and combined results
  - Primary focus was on CNVs called similarly by both methods
  - CNVs with highest odds ratios were selected for PCR validation

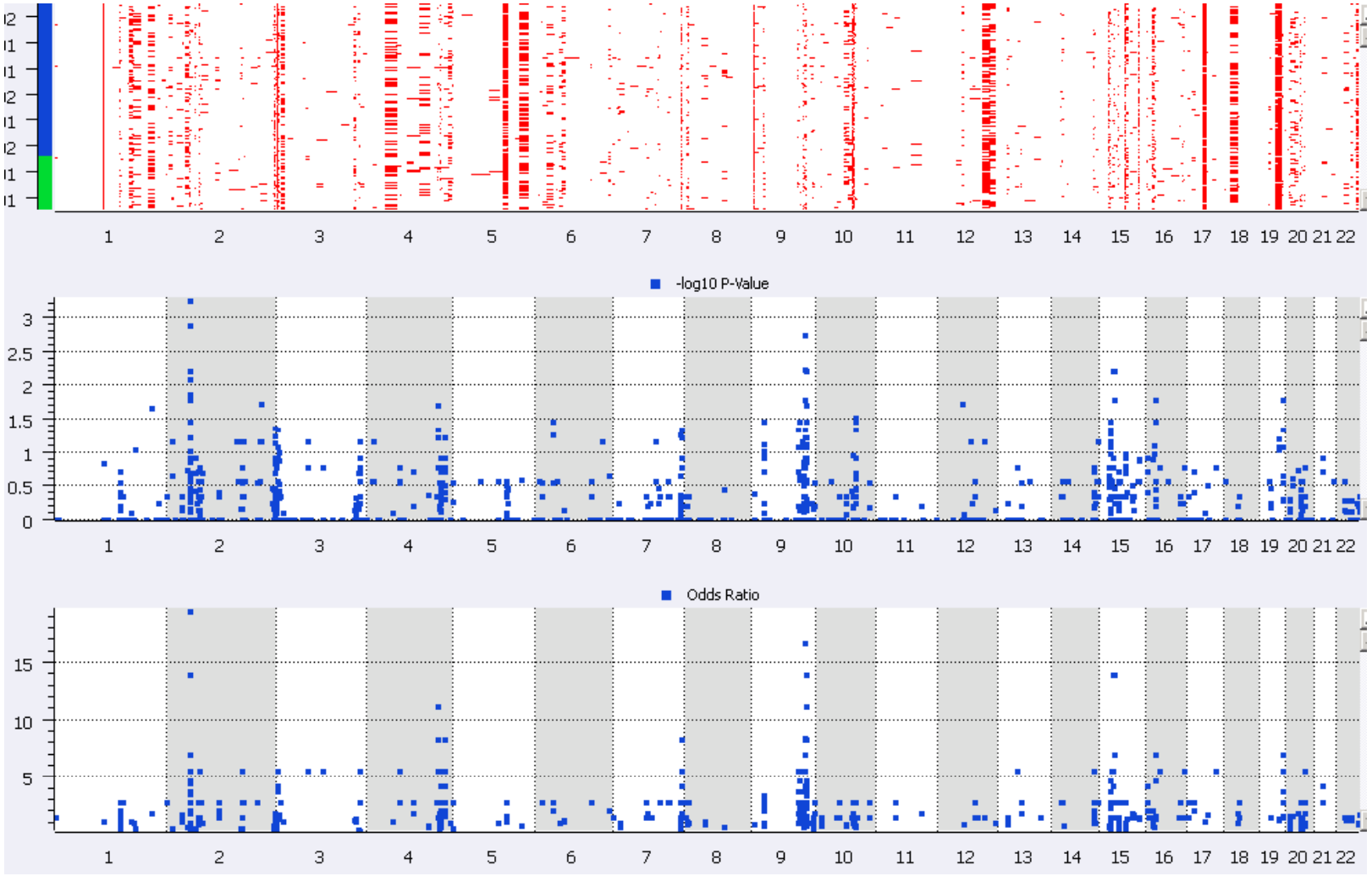
© 2013 Lineagen, Inc, University of Utah, Children’s Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Manual evaluation of CNVs



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Results…



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Dr. Charles Hensel

6. Study Results

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Diagnosing complex genetic disorders

Using symptoms alone misses the condition's etiology and often leads to improper diagnosis and treatment

We sought to develop for clinical use genetic variants that could aid in the genetic evaluation of children with ASDs

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Confirmatory study

Study Mandates

- Use large patient group to better understand the frequency of rare CNVs.
- Use large control group to eliminate rare CNVs also seen in controls.
- Establish clinical relevance of CNVs in a carefully defined population.
- Evaluate sequence variants from Utah families to identify ASD susceptibility genes relevant to the general ASD population.

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Results summary

- Selected for further analysis all CNVs with OR≥2
  - 88 CNVs met selection criteria
- Used TaqMan assays to confirm copy number changes by qPCR
  - >97% of individual CNVs called by both PennCNV and CNAM were confirmed by using TaqMan assays
  - CNVs called only by one of the two methods were confirmed <30% of the time
- Overall validation rates were similar for deletions and duplications

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Validation Results – Utah family variants

- Validated 15 out of 153 CNVs from Utah high risk ASD families
- Utah SNV probes on custom research array identified **11 novel CNVs**
  - **Suggests that both CNVs and SNVs within same gene can influence ASD etiology**
- 17 out of 2,800 SNVs were validated in at least 1 of the 2,175 cases and none of the 5,801 controls
  - One is in a gene previously observed to be disrupted by a translocation in a child with ASD
- ~550 SNVs were found in both controls as well as cases, thus not ASD "risk variants"
- While majority of remaining SNVs were not observed in Validation Study, they are still considered to be potential risk variants, pending further research
  - 75% of Utah SNVs were confirmed by a molecular lab test to be a "real" variant (not a sequence artifact)
  - These SNVs may comprise rare variants unique to the family/individuals in which they were identified
    - Next step is to sequence genes in a case-control study to identify other risk variants

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Validation results – Utah study

CNV Origin	CNV Region - Replication Cohort	CNV Type	OddsRatio (Unrelateds)	P Value (Unrelateds)	Cases (Unrelateds)	Controls (Unrelateds)	Gene/Region
Utah CNV	chr1:145703115-145736438	Dup	3.37	9.60E-03	9	10	CD160, PDZK1
Utah CNV	chr1:215854466-215861792	Del	2.12	5.02E-03	22	39	USH2A
Utah CNV	chr2:51266798-51339236	Del	14.96	8.26E-03	4	1	upstream of NRXN1
Utah CNV#	chr3:172591359-172604675	Dup	3.74	2.11E-01	1	1	downstream of SPATA16
Utah CNV#	chr4:189084240-189117031	Del	3.74	1.98E-01	2	2	downstream of TRIML1
Utah CNV#	chr6:7461346-7470321	Del	∞	2.11E-01	1	0	between RIOK1 and DSP
Utah CNV#	chr6:62426827-62472074	Dup	3.74	1.98E-01	2	2	KHDRBS2
Utah CNV	chr6:147577803-147684318	Del	∞	2.10E-01	1	0	STXBP5
Utah CNV#	chr7:6870635-6871412	Dup	7.47	1.15E-01	2	1	upstream of CCZ1B
Sequence SNP CNV#	chr7:93070811-93116320	Del	∞	4.46E-02	2	0	CALCR, MIR653, MIR489
Utah CNV#	chr9:28207468-28348133	Del	3.74	6.72E-02	4	4	LINGO2
Utah CNV#	chr9:28354180-28354967	Del	3.73	3.78E-01	1	1	LINGO2 (intron)
Utah CNV	chr10:83886963-83888343	Del	3.76	1.54E-02	7	7	NRG3 (intron)
Utah CNV#	chr10:92262627-92298079	Dup	7.47	1.15E-01	2	1	BC037970
Utah CNV#	chr12:102095178-102108946	Dup	7.47	1.15E-01	2	1	CHPT1
Utah CNV#	chr13:40089105-40090197	Del	∞	2.11E-01	1	0	LHFP (intron)
Sequence SNP CNV#	chr14:100705631-100828134	Dup	9.36	5.99E-03	5	2	SLC25A29, YY1, MIR345, SLC25A47, WARS
Sequence SNP CNV#	chr14:102018946-102026138	Dup	4.62	1.01E-14	60	50	DIO3AS, DIO3OS
Sequence SNP CNV#	chr14:102729881-102749930	Del	7.47	1.15E-01	2	1	MOK
Sequence SNP CNV#	chr14:102973910-102975572	Dup	3.82	8.29E-26	136	142	ANKRD9
Sequence SNP CNV*	chr15:25690465-28513763	Dup*	41.05	1.82E-08	11	1	ATP10A, GABRB3, GABRA5, GABRG3, HERC2
Sequence SNP CNV#	chr15:31092983-31369123	Del	∞	4.46E-02	2	0	FAN1, MTMR10, MIR211, TRPM1
Sequence SNP CNV#	chr15:31776648-31822910	Dup	4.40	6.91E-06	21	18	OTUD7A
Sequence SNP CNV#	chr20:32210931-32441302	Dup	2.72	3.16E-02	8	11	NECAB3, CBFA2T2, C20orf144, C20orf134, PXMP4, ZNF341, E2F1, CHMP4B

Total sample size of **unrelated** individuals: 1544 cases, 5762 controls

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

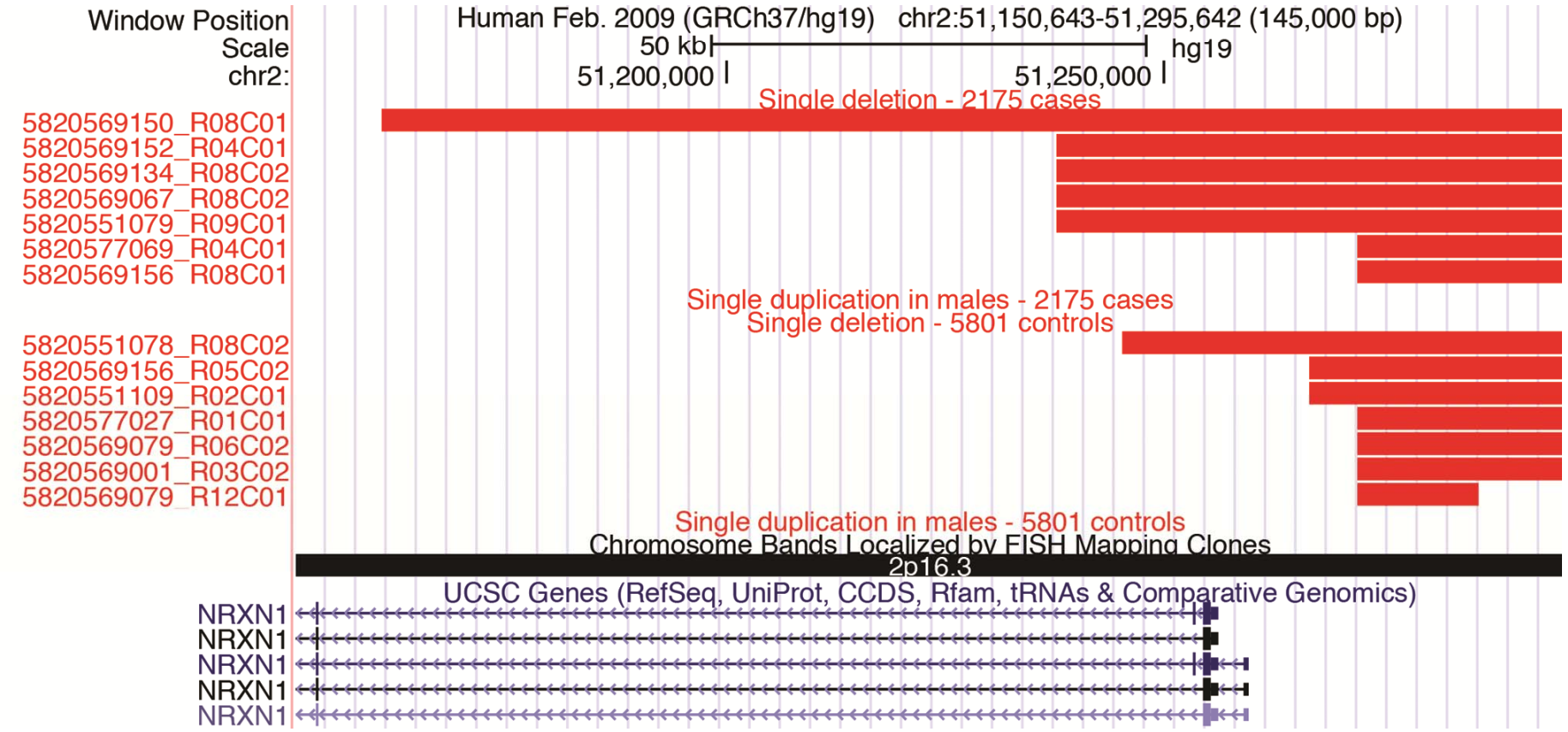
Validation Results – Utah study

TaqMan validated Utah and sequence SNP CNV regions of significance							
CNV Origin	CNV Region - Replication Cohort	CNV Type	OddsRatio (Unrelateds)	P Value (Unrelateds)	Cases (Unrelateds)	Controls (Unrelateds)	Gene/Region
Utah CNV	chr1:145703115-145736438	Dup	3.37	9.60E-03	9	10	CD160, PDZK1
Utah CNV	chr1:215854466-215861792	Del	2.12	5.02E-03	22	39	USH2A
Utah CNV	chr2:51266798-51339236	Del	14.96	8.26E-03	4	1	upstream of NRXN1
Utah CNV#	chr3:172591359-172604675	Dup	3.74	2.11E-01	1	1	downstream of SPATA16
Utah CNV#	chr4:189084240-189117031	Del	3.74	1.98E-01	2	2	downstream of TRIML1
Utah CNV#	chr6:7461346-7470321	Del	∞	2.11E-01	1	0	between RIOK1 and DSP
Utah CNV#	chr6:62426827-62472074	Dup	3.74	1.98E-01	2	2	KHDRBS2
Utah CNV	chr6:147577803-147684318	Del	∞	2.10E-01	1	0	STXBP5
Utah CNV#	chr7:6870635-6871412	Dup	7.47	1.15E-01	2	1	upstream of CCZ1B
Sequence SNP CNV#	chr7:93070811-93116320	Del	∞	4.46E-02	2	0	CALCR, MIR653, MIR489
Utah CNV#	chr9:28207468-28348133	Del	3.74	6.72E-02	4	4	LINGO2
Utah CNV#	chr9:28354180-28354967	Del	3.73	3.78E-01	1	1	LINGO2 (intron)
Utah CNV	chr10:83886963-83888343	Del	3.76	1.54E-02	7	7	NRG3 (intron)
Utah CNV#	chr10:92262627-92298079	Dup	7.47	1.15E-01	2	1	BC037970
Utah CNV#	chr12:102095178-102108946	Dup	7.47	1.15E-01	2	1	CHPT1
Utah CNV#	chr13:40089105-40090197	Del	∞	2.11E-01	1	0	LHFP (intron)
Sequence SNP CNV#	chr14:100705631-100828134	Dup	9.36	5.99E-03	5	2	SLC25A29, YY1, MIR345, SLC25A47, WARS
Sequence SNP CNV#	chr14:102018946-102026138	Dup	4.62	1.01E-14	60	50	DIO3AS, DIO3OS
Sequence SNP CNV#	chr14:102729881-102749930	Del	7.47	1.15E-01	2	1	MOK
Sequence SNP CNV#	chr14:102973910-102975572	Dup	3.82	8.29E-26	136	142	ANKRD9
Sequence SNP CNV*	chr15:25690465-28513763	Dup*	41.05	1.82E-08	11	1	ATP10A, GABRB3, GABRA5, GABRG3, HERC2
Sequence SNP CNV#	chr15:31092983-31369123	Del	∞	4.46E-02	2	0	FAN1, MTMR10, MIR211, TRPM1
Sequence SNP CNV#	chr15:31776648-31822910	Dup	4.40	6.91E-06	21	18	OTUD7A
Sequence SNP CNV#	chr20:32210931-32441302	Dup	2.72	3.16E-02	8	11	NECAB3, CBFA2T2, C20orf144, C20orf134, PXMP4, ZNF341, E2F1, CHMP4B

Total sample size of **unrelated** individuals: 1544 cases, 5762 controls

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

NRXN1 CNVs



Note that only 1 of the control CNVs extends into the NRXN1 coding region.

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Validation results – Utah study

TaqMan validated Utah and sequence SNP CNV regions of significance							
CNV Origin	CNV Region - Replication Cohort	CNV Type	OddsRatio (Unrelateds)	P Value (Unrelateds)	Cases (Unrelateds)	Controls (Unrelateds)	Gene/Region
Utah CNV	chr1:145703115-145736438	Dup	3.37	9.60E-03	9	10	CD160, PDZK1
Utah CNV	chr1:215854466-215861792	Del	2.12	5.02E-03	22	39	USH2A
Utah CNV	chr2:51266798-51339236	Del	14.96	8.26E-03	4	1	upstream of NRXN1
Utah CNV#	chr3:172591359-172604675	Dup	3.74	2.11E-01	1	1	downstream of SPATA16
Utah CNV#	chr4:189084240-189117031	Del	3.74	1.98E-01	2	2	downstream of TRIML1
Utah CNV#	chr6:7461346-7470321	Del	∞	2.11E-01	1	0	between RIOK1 and DSP
Utah CNV#	chr6:62426827-62472074	Dup	3.74	1.98E-01	2	2	KHDRBS2
Utah CNV	chr6:147577803-147684318	Del	∞	2.10E-01	1	0	STXBP5
Utah CNV#	chr7:6870635-6871412	Dup	7.47	1.15E-01	2	1	upstream of CCZ1B
Sequence SNP CNV#	chr7:93070811-93116320	Del	∞	4.46E-02	2	0	CALCR, MIR653, MIR489
Utah CNV#	chr9:28207468-28348133	Del	3.74	6.72E-02	4	4	LINGO2
Utah CNV#	chr9:28354180-28354967	Del	3.73	3.78E-01	1	1	LINGO2 (intron)
Utah CNV	chr10:83886963-83888343	Del	3.76	1.54E-02	7	7	NRG3 (intron)
Utah CNV#	chr10:92262627-92298079	Dup	7.47	1.15E-01	2	1	BC037970
Utah CNV#	chr12:102095178-102108946	Dup	7.47	1.15E-01	2	1	CHPT1
Utah CNV#	chr13:40089105-40090197	Del	∞	2.11E-01	1	0	LHFP (intron)
Sequence SNP CNV#	chr14:100705631-100828134	Dup	9.36	5.99E-03	5	2	SLC25A29, YY1, MIR345, SLC25A47, WARS
Sequence SNP CNV#	chr14:102018946-102026138	Dup	4.62	1.01E-14	60	50	DIO3AS, DIO3OS
Sequence SNP CNV#	chr14:102729881-102749930	Del	7.47	1.15E-01	2	1	MOK
Sequence SNP CNV#	chr14:102973910-102975572	Dup	3.82	8.29E-26	136	142	ANKRD9
Sequence SNP CNV*	chr15:25690465-28513763	Dup*	41.05	1.82E-08	11	1	ATP10A, GABRB3, GABRA5, GABRG3, HERC2
Sequence SNP CNV#	chr15:31092983-31369123	Del	∞	4.46E-02	2	0	FAN1, MTMR10, MIR211, TRPM1
Sequence SNP CNV#	chr15:31776648-31822910	Dup	4.40	6.91E-06	21	18	OTUD7A
Sequence SNP CNV#	chr20:32210931-32441302	Dup	2.72	3.16E-02	8	11	NECAB3, CBFA2T2, C20orf144, C20orf134, PXMP4, ZNF341, E2F1, CHMP4B

Total sample size of **unrelated** individuals: 1544 cases, 5762 controls

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Validation results – CHOP variants

- 11 out of 84 CNVs from CHOP were validated as being clinically relevant
  - All validated CNVs are recurrent (seen in more than 1 case)
  - Verified odds ratios >2 in unrelated (N=1544) ASD cases and controls (N=5762)
- 73 remaining CHOP CNVs did not reach clinical significance because they were seen in only 1 case and/or had odds ratios <2
- Validated clinical relevance of 16 of 101 CNVs from other publications
  - Included only unrelated cases in calculations to reduce inflated frequency estimates
  - Odds ratios >2 in unrelated cases

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Validation results – CHOP and Literature Studies

Cytoband	Region of Highest Significance	CNV Type	OddsRatio (Unrelateds)	P Value (Unrelateds)	Cases (Unrelateds)	Controls (Unrelateds)	Gene/Region
1q21.1	chr1:146656292-146707824	Dup	7.48	1.15E-01	2	1	FMO5
2p24.3	chr2:13203874-13209245	Del	∞	2.11E-01	1	0	upstream of LOC100506474
2p21	chr2:45489954-45492582	Dup	∞	4.46E-02	2	0	between UNQ6975 and SRBD1
2p16.3	chr2:51237767-51245359	Del	∞	1.99E-03	4	0	NRXN1
2p15	chr2:62230970-62367720	Dup	∞	2.11E-01	1	0	COMMD1
2q14.1	chr2:115133493-115140263	Del	7.47	1.15E-01	2	1	between LOC440900 and DPP10
3p26.3	chr3:1937796-1941004	Del	5.60	6.70E-02	3	2	between CNTN6 and CNTN4
3p14.1	chr3:67657429-68962928	Del	∞	2.11E-01	1	0	SUCLG2, FAM19A4, FAM19A1
4q13.3	chr4:73766964-73816870	Dup	∞	2.11E-01	1	0	COX18, ANKRD17
4q33	chr4:171366005-171471530	Del	∞	4.46E-02	2	0	between AADAT and HSP90AA6P
5q23.1	chr5:118527524-118589485	Dup	3.74	1.98E-01	2	2	DMXL1, TNFAIP8
6p21.2	chr6:39069291-39072241	Del	2.37	1.93E-02	12	19	SAYSD1
8q11.23	chr8:54855680-54912001	Dup	∞	2.11E-01	1	0	RGS20, TCEA1
10q11.22	chr10:49370090-49471091	Dup	3.77	1.96E-01	2	2	FRMPD2P1, FRMPD2
10q11.23	chr10:50884949-50943185	Dup	3.74	1.98E-01	2	2	OGDHL, C10orf53
12q13.13	chr12:53177144-53180552	Del	∞	4.46E-02	2	0	between KRT76 and KRT3
15q11.1	chr15:20192970-20197164	Dup	4.97	4.06E-02	4	3	downstream of HERC2P3
15q11.2	chr15:25099351-25102073	Del	3.75	1.13E-01	3	3	SNRPN
15q11.2	chr15:25099351-25102073	Dup	45.19	7.93E-08	12	1	SNRPN
15q11.2	chr15:25579767-25581658	Dup*	∞	3.86E-06	8	0	between SNORD109A and UBE3A
15q11.2	chr15:25582882-25662988	Dup*	30.08	2.82E-05	8	1	UBE3A
16p12.2	chr16:21958486-22172866	Dup	∞	4.47E-02	2	0	C16orf52, UQCRC2, PDZD9, VWA3A
16p11.2	chr16:29664753-30177298	Del	7.47	1.15E-01	2	1	DOC2A, ASPHD1, LOC440356, TBX6, LOC100271831, PRRT2 CDIPT, QPRT, YPEL3, PPP4C, MAPK3, SPN, MVP, FAM57B, ZG16, ALDOA, INO80E, SEZ6L2, TAOK2, KCTD13, MAZ, KIF22, GDPD3, C16orf92, C16orf53, TMEM219, C16orf54, HIRIP3
16q23.3	chr16:82423855-82445055	Dup	∞	4.46E-02	2	0	between MPHOSPH6 and CDH13
17p12	chr17:14132271-14133349	Dup	1.60	3.57E-01	3	7	between COX10 and CDRT15
17p12	chr17:14132271-15282708	Del	5.61	6.70E-02	3	2	PMP22, CDRT15, TEKT3, MGC12916, CDRT7, HS3ST3B1
17p12	chr17:14952999-15053648	Dup	3.74	1.98E-01	2	2	between CDRT7 and PMP22
17p12	chr17:15283960-15287134	Del	3.74	1.13E-01	3	3	between TEKT3 and FAM18B2-CDRT4
20p12.3	chr20:8162278-8313229	Dup	3.73	1.98E-01	2	2	PLCB1
Xp21.2	chrX:29944502-29987870	Dup	∞	4.47E-02	2	0	IL1RAPL1
Xq27.2	chrX:140329633-140348506	Del	7.48	2.06E-02	4	2	SPANXC
Xq28	chrX:148882559-148886166	Del	∞	4.46E-02	2	0	MAGEA8

Total sample size of **unrelated** individuals: 1544 cases, 5762 controls

Red color indicates CNVs validated by qPCR

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Developmental Disorders provide a key example of the problem

~70 Shared, Non-Specific Symptoms

>700 Classified Disorders of Childhood Development

Result:
Most kids get the wrong diagnosis or the right diagnosis too late

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

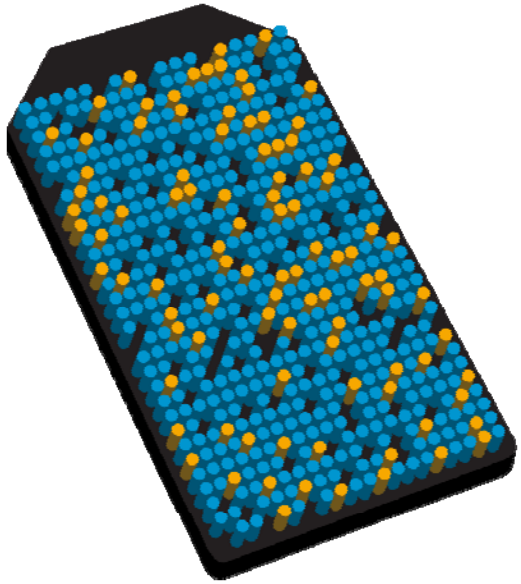
Rationale for FirstStep^{Dx} PLUS array design

Goal: the most comprehensive clinical microarray available for disorders of childhood development

- Best-in-class whole-genome coverage
- Dense coverage of validated proprietary ASD risk markers (CNVs, SNVs) from Lineagen family studies
- Dense coverage of literature ASD risk markers
- Dense coverage of genes/CNVs responsible for known developmental disorders
- Coverage of recurrent point mutations and small insertions/deletions in known ASD/DD/ID genes
- Coverage of genes/CNVs in additional pediatric conditions (ADHD, dyslexia, Tourette syndrome)

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Key FirstStep^{Dx} PLUS second generation array probe design elements

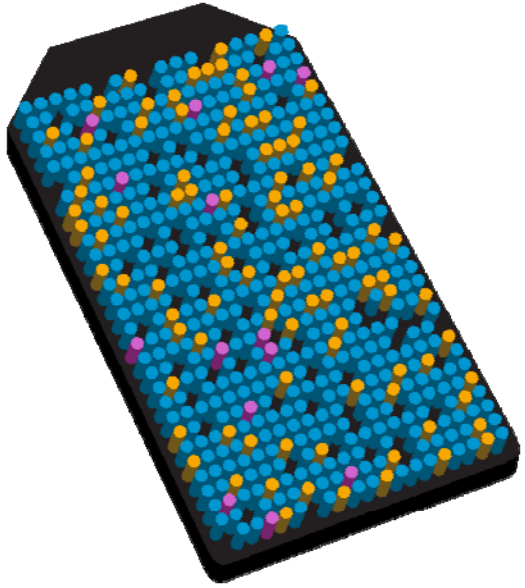


- Platform based on the Affymetrix CytoScan-HD array
- Added 83,443 probes to CytoScanHD base array
  - 2,779,993 total probes
- Coverage of Lineagen validation study CNVs and SNVs
- Coverage of known literature CNVs
- Doubles sensitivity of ASD-related genetic factors
  - **12% -14%** vs. **5% - 7%** base array sensitivity
- Additional increases in ASD-related sensitivity is "built into" FirstStep^{Dx} PLUS based on continuing validation studies with Utah and CHOP
  - All CHOP Study I, CHOP Study II, and Utah CNVs and SNVs were added to custom array
  - No need to manufacture a new array
  - Further increases in sensitivity possible with further validation studies using Lineagen proprietary array

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Key FirstStep^{Dx} PLUS second generation array probe design elements

- Additional probes cover DD-related alterations not readily detectable by generic CMA platforms
  - Recurrent small **Rett syndrome deletions**, usually detected by DNA sequencing
  - Recurrent point mutations in known ASD/DD genes e.g. **TSC1&2, MECP2**
- On a single platform, FirstStep^{Dx} PLUS allows for maximum detection of genetic variants associated with ASD and other disorders of childhood development



- **Key goal is limiting the need for follow-up genetic tests in normal clinical practice – such as single gene sequencing**

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

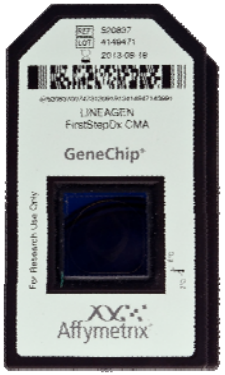
First and second generation technology comparison

	FirstStep^{Dx}	FirstStep^{Dx} PLUS
Number of DNA Probes	2,696,550	2,779,993 (CytoScan-HD + 83,443 custom probes)
Genome-Wide Resolution	1 kb	1 kb
Number of Target Regions	18,000 (RefSeq genes)	18,000+ and proprietary CNVs and SNVs
Target Region Resolution	0.6 kb	0.6 kb
ASD Sensitivity/Yield	**5-7%**	**12-14%**
Overall Sensitivity/Yield	**23%**	**26%**

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

FirstStep^{Dx} PLUS – delivers the most clinically informative results for patients

- The **most comprehensive** whole genome array clinically available
  - In partnership with Affymetrix, customized the CytoScan microarray with a unique probe design that allows for detection of novel validated genetic variants
  - Yields a **> 2x** increase in detection of Autism-related genetic variants over competitive tests
- Increased coverage of other developmental delay genetic alterations not readily detectable by competitive array platforms



Affymetrix | Powered by Affymetrix™

- On a single platform, FirstStep^{Dx} allows for maximum detection of genetic variants associated with ASD and other disorders of childhood development

The most clinically-actionable information per test result

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

LINEAGEN

GOLDEN HELIX
Accelerating the Quest for Significance™

THE UNIVERSITY OF UTAH

The Children's Hospital of Philadelphia®

New Study Identifies High-Risk Variants Associated with Autism Spectrum Disorders

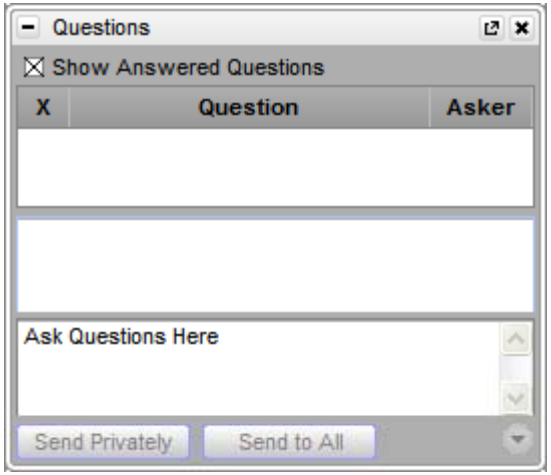
Twenty-four new variants discovered, each conferring more than a 2-fold risk of developing ASD

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Q&A

Do You Have Any Questions?

Use the Questions pane in your GoToWebinar window



© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.

Speakers & Agenda

					
Dr. Christophe Lambert	**Dr. Michael Paul**	**Dr. Hakon Hakonarson**	**Dr. Mark Leppert**	**Dr. Bryce Christensen**	**Dr. Charles Hensel**
CEO at Golden Helix	*President and Chief Executive Officer at Lineagen*	*Director of the Center for Applied Genomics at the Children's Hospital of Philadelphia*	*Professor of Human Genetics at the University of Utah and Chief Science Advisor at Lineagen*	*Director of Services and Statistical Geneticist at Golden Helix*	*Senior Research Manager at Lineagen*
1. Introduction of presenters and agenda	**2. Background on Lineagen**	**3. The science behind Autism Spectrum Disorders (ASDs)**	**4. Family-based genetics of ASDs**	**5. The analytic process**	**6. Study results**

© 2013 Lineagen, Inc, University of Utah, Children's Hospital of Philadelphia, and Golden Helix, Inc. All rights reserved.