

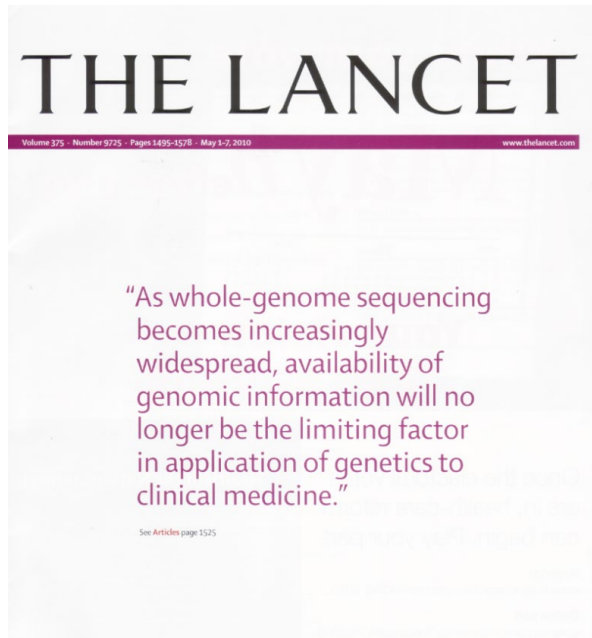
Update from Ashley Lab

- Publication update
- Major allele human reference
- Exome sequencing and HCM
- Whole genome sequencing for clinical medicine
- The intermediate step here now: targeted capture chip for clinical sequencing
- Whole genome sequencing for molecular autopsy
- Personalized medicine clinical trial



Clinical assessment incorporating a personal genome

Euan A Ashley, Atul J Butte, Matthew T Wheeler, Rong Chen, Teri E Klein, Frederick E Dewey, Joel T Dudley, Kelly E Ormond, Aleksandra Pavlovic, Alexander A Morgan, Dmitry Pushkarev, Norma F Neff, Louanne Hudgins, Li Gong, Laura M Hodges, Dorit S Berlin, Caroline F Thorn, Katrin Sangkuhl, Joan M Hebert, Mark Woon, Hersh Sagreiya, Ryan Whaley, Joshua W Knowles, Michael F Chou, Joseph V Thakuria, Abraham M Rosenbaum, Alexander Wait Zaranek, George M Church, Henry T Greely, Stephen R Quake, Russ B Altman



SCIENCE WATCH



HOME | ABOUT THOMSON REUTERS | PRESS ROOM | CONTACT US | Site Search SEARCH

ScienceWatch.com Data & Rankings New Hot Papers

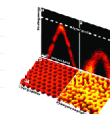
FOLLOW US SHARE

NEW HOT PAPERS

MAIN MENU

SCIENCEWATCH HOME
INSIDE THIS MONTH...
INTERVIEWS
SPECIAL TOPICS
ANALYSES
DATA & RANKINGS
Sci-Bytes
Fast Breaking Papers
New Hot Papers
Emerging Research Fronts
Fast Moving Fronts
Research Front Maps
Current Classics
Top Topics
Rising Stars
New Entrants
Country Profiles
ABOUT SCIENCE WATCH

FOLLOW US ON
twitter
SHARE



Every two months, *Essential Science Indicators*SM from Thomson Reuters lists a new crop of what it calls hot papers in science. Hot papers are selected by virtue of being cited among the top one-tenth of one percent (0.1%) in a current bimonthly period. Papers are selected in each of 22 fields of science and must be published within the last two years. Because the hot papers are updated every two months, new papers are added with every update. and *ScienceWatch.com* tracks these new additions.

Here, *ScienceWatch.com* highlights the most-cited of these new entries, one from each field, which are, in addition, not more than one year old. Since new hot papers are very recent scientific contributions that are receiving recognition during a current period, they may signal important new trends in research and serve as leading indicators of scientific advancement. Full information, including bimonthly citation histories, for these papers can be found in *Essential Science Indicators*. See also, *Understanding Core Data: Hot Papers Methodology*.

Some papers have comments sent in by the author(s) of the paper which may include images and descriptions of their work.

SELECT 2011 2010 2009 2008

Choose the year/month you would like to view from the tabs above.

EMAIL | BACK TO TOP

CITE THIS PAGE: <http://sciencewatch.com/dhtml/>

ScienceWatch.com Data & Rankings New Hot Papers

NEW HOT PAPERS

New Hot Papers Home
2011 Menu

MORE PAPERS

Fast Breaking Papers
Current Classics

CURRENT SPECIAL TOPIC



Special Topic of:
Metabolic Syndrome
VIEW >

FEATURED IMAGE

A figure from a New Hot Paper commentary titled: "Liv Homelær Discusses Hydrogen Functionalization of Graphene."



Personalis EX2045.2

Phased whole genome genetic risk in a family quartet using a major allele reference sequence

Dewey FE¹, Chen R², Cordero SP³, Ormond KE⁴, Caleshu C^{1,4}, Karczewski KJ^{3,4}, Carrillo MW⁴, Wheeler MT¹, Dudley JT^{2,3}, Byrnes JK⁴, Comejo OE⁴, Knowles JW¹, Woon M⁴, Sangkuhl K⁴, Gong L⁴, Thorn CF⁴, Hebert JM⁴, Capriotti E⁴, David SP⁴, Pavlovic A¹, West A⁵, West JS⁶, Thakuria JV⁷, Ball MP⁸, Zaranek AW⁸, Rehm HL⁹, Church GM⁸, Bustamante CD⁴, Snyder MP⁴, Altman RB⁴, Klein TE⁴, Butte AJ², Ashley EA¹

¹Center for Inherited Cardiovascular Disease, Division of Cardiovascular Medicine, Stanford University, Stanford, CA

²Division of Systems Medicine, Stanford University School of Medicine, Stanford, CA

³Biomedical Informatics Graduate Training Program, Stanford University School of Medicine, Stanford, CA

⁴Department of Genetics, Stanford University School of Medicine, Stanford, CA

⁵The Harker School, San Jose, CA

⁶Personalis, Palo Alto, CA

⁷Division of Genetics, Massachusetts General Hospital, Boston, MA

⁸Department of Genetics, Harvard Medical School, Boston, MA

⁹Department of Pathology, Harvard Medical School, Boston, MA

Running title: Genetic risk in a family quartet

Please address correspondence to:

Euan A. Ashley
Stanford University School of Medicine
Falk CVRB
300 Pasteur Drive
Stanford, CA, 94305
650 498 4900 (p)
650 723 8392 (f)
euana@stanford.edu



Invited revision – Plos Genetics

PLOS GENETICS
a peer-reviewed open-access journal published by the Public Library of Science

Login | Create Account | Feedback

Search articles... GO Advanced Search

Browse RSS

Home Browse Articles About For Readers For Authors and Reviewers Journals Help PLoS.org

<Previous Issue | Archive |

Table of Contents | June 2011

ABOUT THIS IMAGE
Maize tassels in full blossom. Millions of pollen grains are ready to fertilize the ovules of ears of maize. This is easy to control because male flowers on the top (tassel) are seen) and female flowers on the side (ears) are physically separated, allowing either self- or hybrid crosses, which has advanced breeding in agriculture and made maize the highest yielding grain. To further enable the improvement of grain quality, [Mickelson et al.](#) studied the architecture, expression, and epigenetic regulation of a medium-size gene family, the alpha-zymes, in two maize haplotypes, B73 and B65553. Image Credit: [Mikael Mickelson](#) (Wageningen Institute of Microbiology)

[View Image](#)

Interview

Vive La Différence: An Interview with Catherine Dulac
Jane Gitchev

Top
Interview
Viewpoints
Perspective
Review
Research Articles



Personalis EX2045.3

Paper introduces:

- the development and application of ethnicity-specific major allele reference sequences
- the definition of meiotic crossover sites to a median resolution of <1000 base pairs
- the development of a computational algorithm to achieve genome-wide long-range haplotype phasing
- the application of this phasing to HLA typing and clinical interpretation of genomes from a family quartet using internally developed databases of all publicly accessible genotype-phenotype association data
- the demonstration and quantification of discordance of inherited disease risk within a nuclear family
- the successful prediction of the only manifest disease phenotype in this family.



WHAT'S NEW?

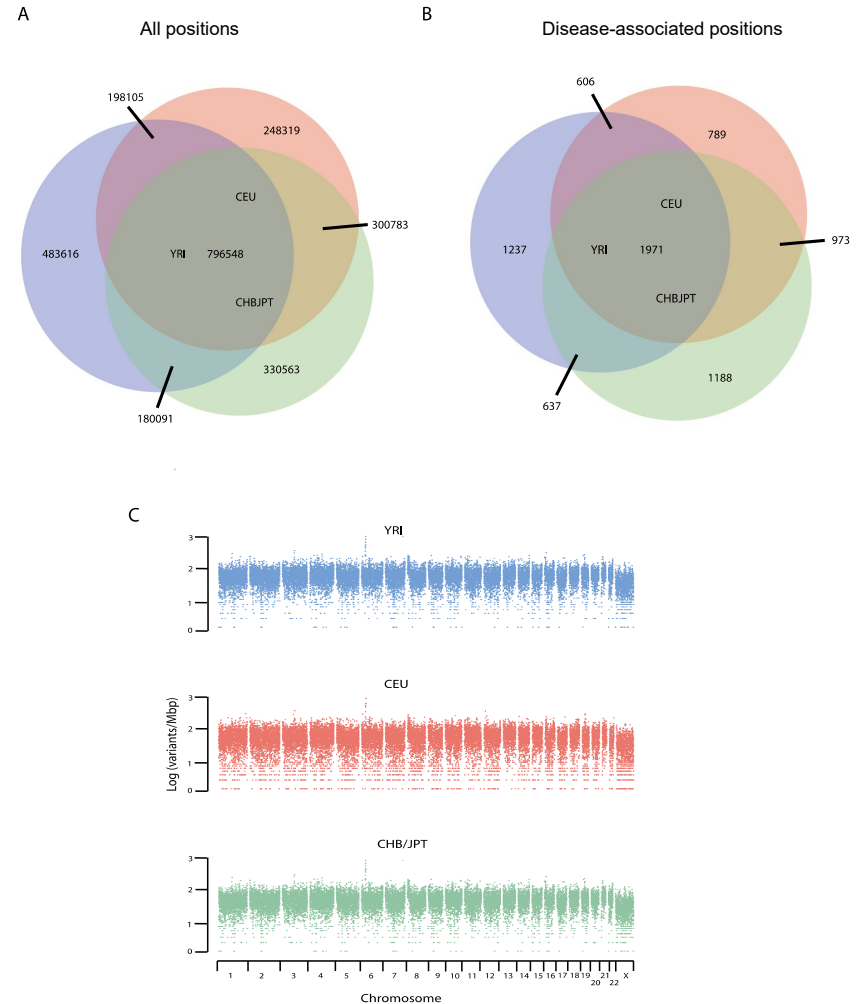


Personalis EX2045.5

Variant identification with a major allele human reference sequence



Towards an (improved) human reference sequence



Disease risk alleles on the human reference sequence

Table 2: Rare disease-associated minor alleles on the reference genome

	Minor alleles on the reference genome (hg19)	Rare minor alleles	
		Allele freq <5%	Allele freq <1%
All populations	1971	23	3
CEU	4337	131	32
CHB/JPT	4593	553	319
YRI	4450	466	267

Rong Chen





Disease risk alleles on the human reference sequence

Rong Chen



Table 3: Very-rare alleles on the NCBI human reference genome with highly-validated disease risk associations

Disease/trait	dbSNP	symbol	Type	Risk	P value	Odds Ratio	Pop
Blood pressure	653178	ATXN2	intron	C	3.00E-18		ALL, CHB/JPT, YRI
Celiac disease	653178	ATXN2	intron	C	7.15E-21	1.20	ALL, CHB/JPT, YRI
Celiac disease	3184504	SH2B3	missense	T	1.23E-12	1.21	ALL, CHB/JPT, YRI
Coronary artery disease	3184504	SH2B3	missense	T	4.23E-11	1.14	CHB/JPT, YRI
End-Stage renal disease	2032487	MYH9	intron	C	1.98E-15	2.61	CHB/JPT
End-Stage renal disease	4821480	MYH9	intron	G	8.48E-17	2.29	CHB/JPT
End-Stage renal disease	4821481	MYH9	intron	C	2.28E-16	2.62	CHB/JPT
End-Stage renal disease	5756152	MYH9	intron	A	5.84E-11	2.82	ALL, CHB/JPT
Focal segmental glomerulosclerosis	4821481	MYH9	intron	C	1.28E-09		CHB/JPT
Ige levels	1295685	IL13	UTR-3	A	2.00E-07	1.61	YRI
Rheumatoid arthritis	1217413		intergenic	G	4.00E-08	1.38	YRI
Rheumatoid arthritis	2476601	PTPN22	missense	A	2.30E-98	1.96	ALL, CHB/JPT, YRI
Systemic lupus erythematosus	2476601	PTPN22	missense	A	3.4E-12	1.57	ALL, CHB/JPT, YRI
Type 1 diabetes	2476601	PTPN22	missense	A	2.11E-87	2.25	ALL, CHB/JPT, YRI
Type 1 diabetes	3184504	SH2B3	missense	T	5.62E-31	1.24	ALL, CHB/JPT, YRI
Type 2 diabetes	5219	KCNJ11	missense	T	6.7E-11	1.26	YRI
Type 2 diabetes	2970847	PPARGC1A	cds-synon	T	1.59e-18	2.10	YRI
Vitiligo	2476601	PTPN22	missense	A	6.79E-09	1.82	ALL, CHB/JPT, YRI



Personalis EX2045.9

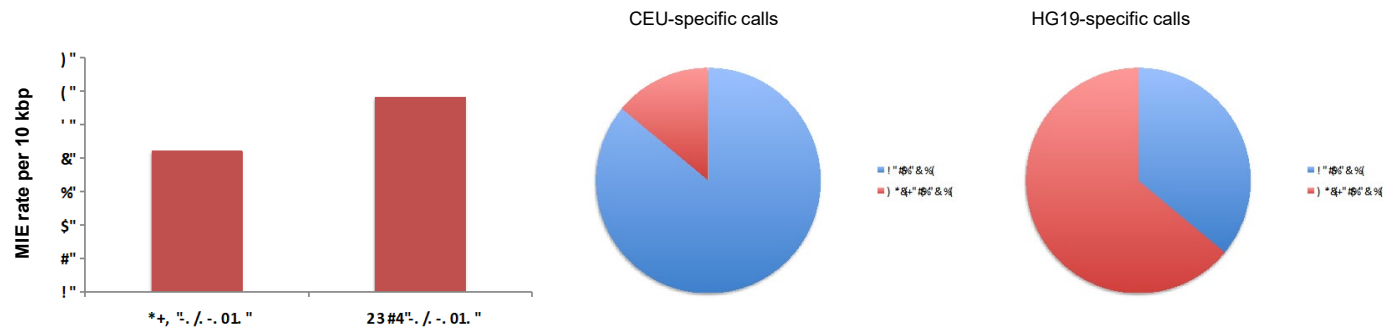
Alignment to a major allele reference sequence

	HG19 reference	CEU reference
Bases aligned – chr 22	13,757,281	13,758,089
Bases aligned – chr 6	26,186,555	26,189,514
Reads aligned – chr 22	3,740,839	3,741,223
Reads aligned – chr6	4,503,067	4,503,444

One lane of data from the Illumina GA II was aligned to each reference sequence using BWA



Variant identification using a major allele reference sequence



Hg19 -> CEU	PG21	PG20	PG22	PG23
Risk allele	23	11	14	21
Protective allele	46	29	25	44
No change	11	63	68	14
Missing risk allele	48	25	21	49
Unmatched allele [#]	2	2	2	2

- CEU reference resulted in ~40% lower error rate at disease-associated loci
- 85% of CEU-specific calls validated with chip
- 36% of HG19-specific calls validated with chip



Exome sequencing and hypertrophic cardiomyopathy



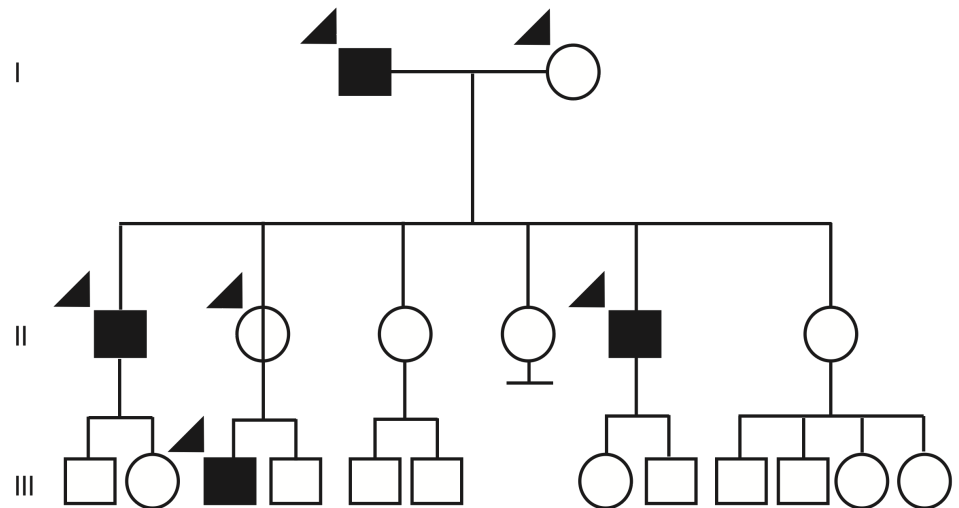
Study subjects

➤ Pedigree ascertained through 52 y/o proband

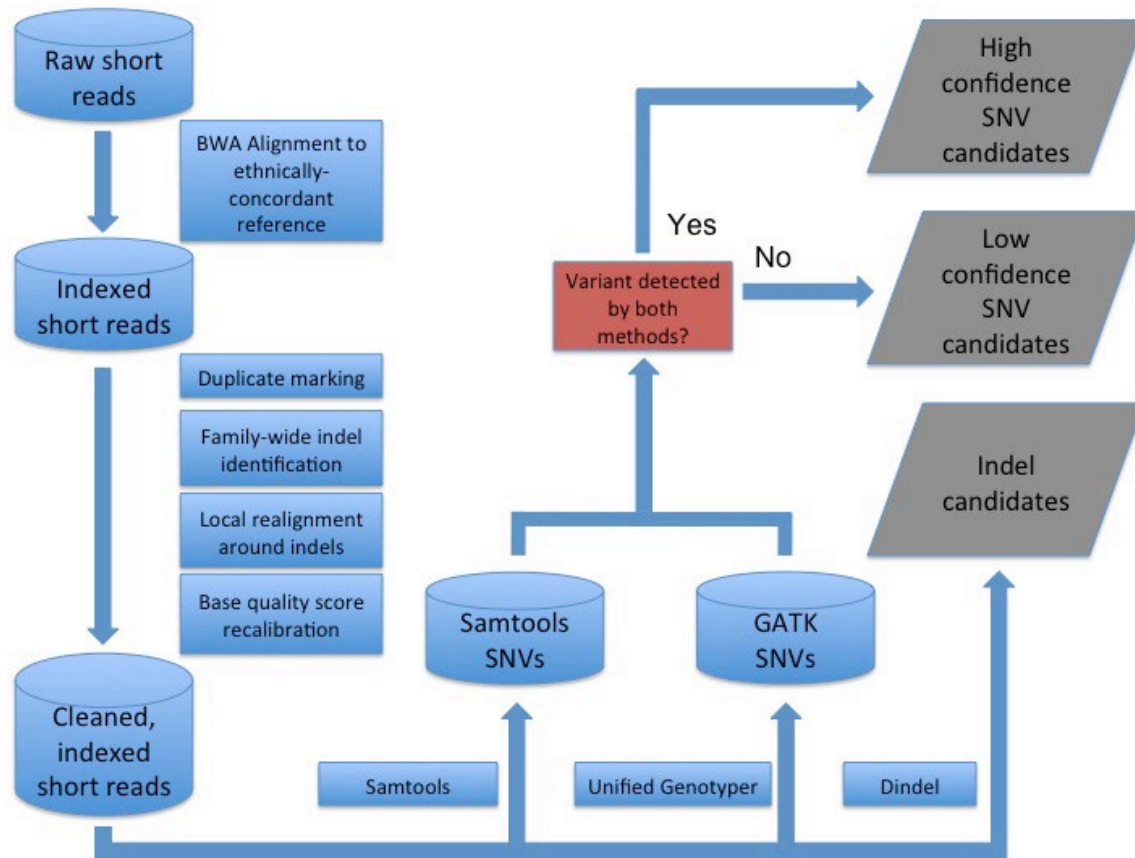
➤ 4 affected individuals, including one severely affected individual in generation III

➤ One obligate carrier in generation II

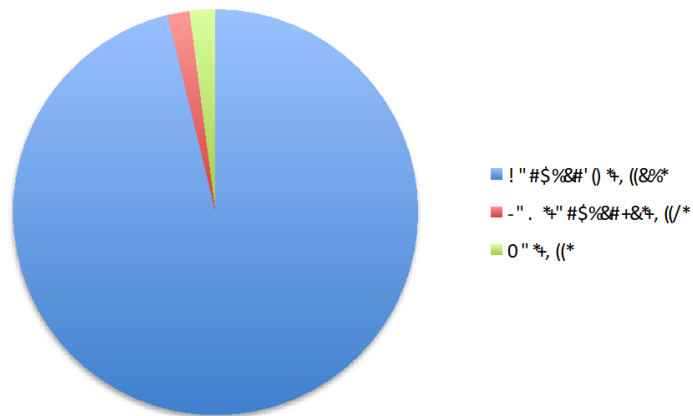
➤ One unaffected spouse in generation I



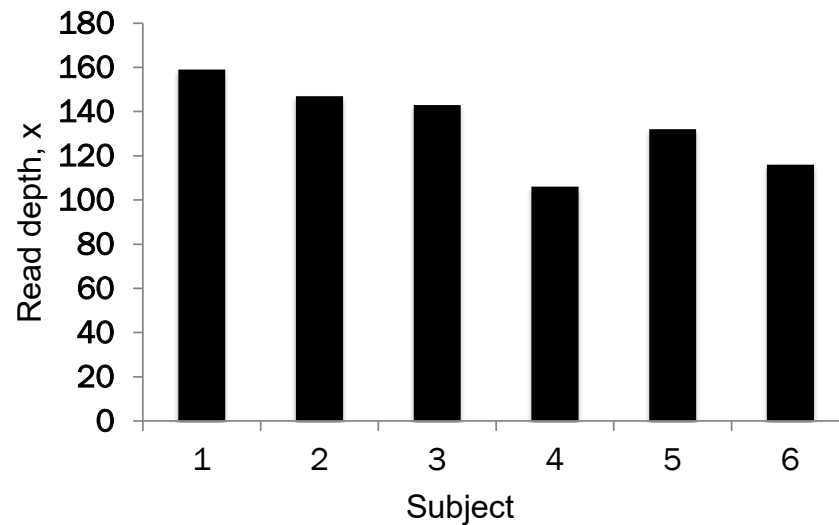
Alignment and variant identification



Sequencing coverage and depth



% of targeted positions covered



Total reads: 2.4B
Bases called: 73 M
Ti/Tv: 2.62

Variant statistics

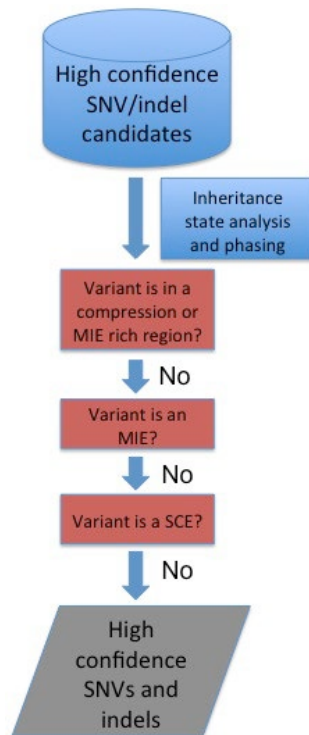
Variant type	Study subject					
	I.1*	I.2	II.1*	II.2 [†]	II.5*	III.3*
All snps	34,810	35,403	35,264	35,821	35,741	35,397
NS/MS/SS	6,006	6,170	5,995	6,154	6,078	6,243
Novel NS/MS/SS	297	305	289	310	281	299
All indels	2,659	2,728	2,723	2,737	2,732	2,647
Coding indels	235	239	233	222	238	237
Frameshift coding indels	112	112	111	105	118	113
Novel frameshift coding indels	97	99	96	92	103	95

* Affected subject. [†] Obligate carrier

No putative loss of function variants were seen in 29 genes previously associated with hypertrophic cardiomyopathy



Variant filters



➤ High confidence “damaging” SNVs/indels were filtered by:

➤ Novelty (not found in dbSNP 132 or 1KG): 415

➤ Cosegregation with phenotype: 38

➤ Variants passing these filters were annotated by:

➤ Polyphen2/SIFT

➤ GERP/Phastcons scores

➤ Association with known CV disease genes in OMIM



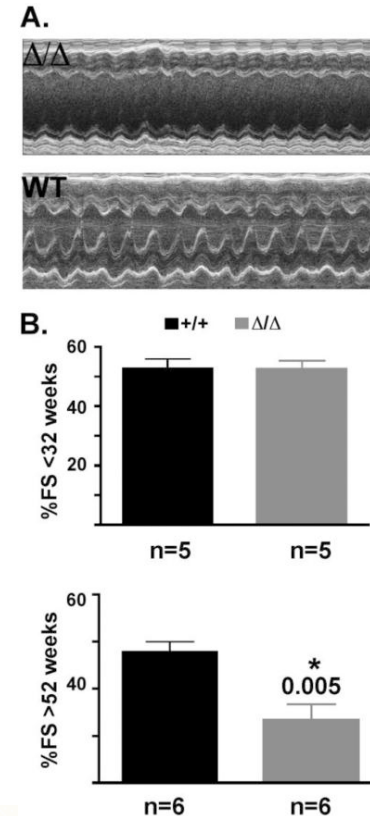
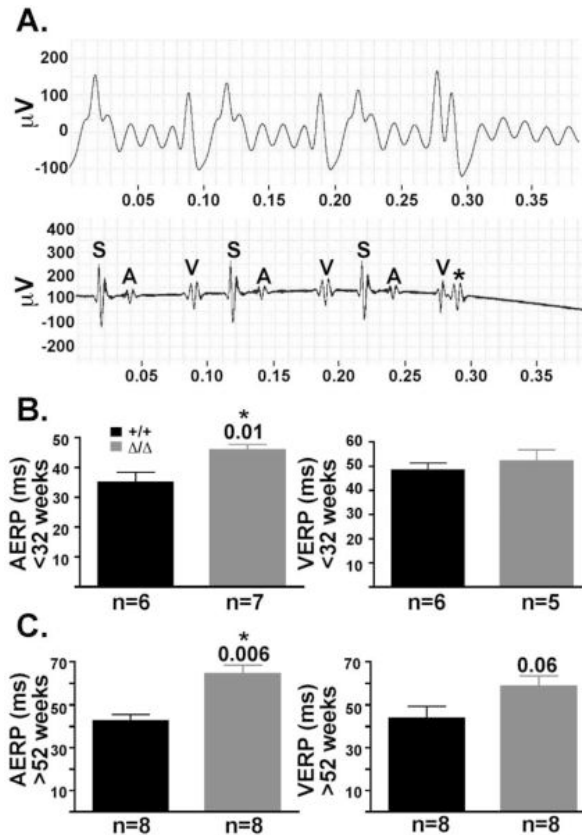
Top hits - SNVs

Variant	GERP score*	PhastCons score^
<i>SYNE1</i> p.P1585S	5.62	1
<i>ASTN1</i> p.M1095L	5.23	0.998
<i>KALRN</i> p.N1163T	5.10	0.99
*Range -11.6 to 5.82, with 5.82 being the most conserved		
^Range 0 to 1, with 1 being the most conserved		





Nesprin 1 modulates atrial and ventricular refractory times and contractility



J Mol Cell
Cardiol. 2010
Apr;48(4):600-8.
Epub 2009 Nov
24.



Top hits - indels

Variant type	Gene	Chromosome	Position	Exon
frameshift deletion	<i>PDE4DIP</i>	chr1	144917828	11
frameshift deletion	<i>KIAA0040</i>	chr1	175129948	5
frameshift deletion	<i>LGR6</i>	chr1	202183358	1
frameshift deletion	<i>ZNF83</i>	chr19	53116934	4
frameshift deletion	<i>ZNF28</i>	chr19	53304140	4



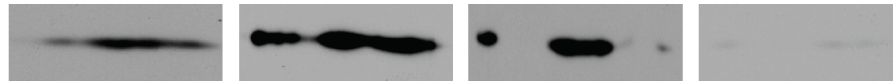


Myomegalin is an A-kinase associated protein necessary for adrenergically-mediated phosphorylation of MYBPC

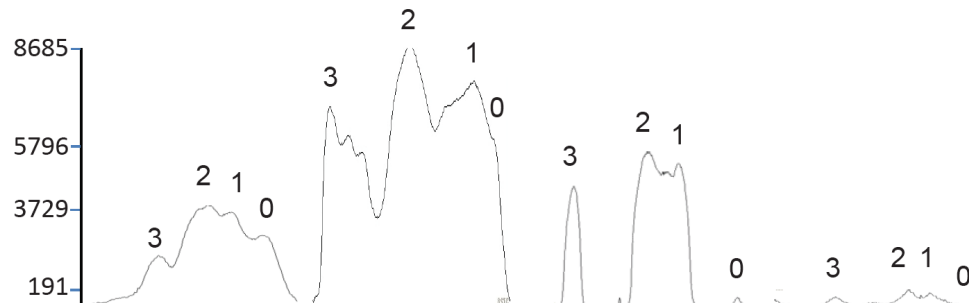
B)

	(i)	(ii)	(iii)	(iv)
MMGL siRNA:	-	-	+	+
CaCl ₂ , Isopro:	-	+	-	+

WB: JL-8



C)



BMC Cell Biol. 2011 May 10;12:18.



Personalis EX2045.21

Follow-up work

- Functional assay of myomegalin indel and nesprin-1 SNV
- IBD analysis to identify possible modifier loci



Whole genome sequencing for clinical diagnosis



Complex diseases and *de novo* mutations

- *de novo* mutation is often overlooked in complex diseases

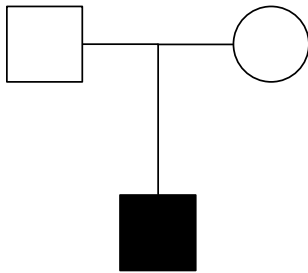
Increased exonic *de novo* mutation rate in individuals with schizophrenia

Simon L Girard¹, Julie Gauthier¹, Anne Noreau¹, Lan Xiong¹, Sirui Zhou¹, Loubna Jouan¹, Alexandre Dionne-Laporte¹, Dan Spiegelman¹, Edouard Henrion¹, Ousmane Diallo¹, Pascale Thibodeau¹, Isabelle Bachand², Jessie Y J Bao³, Amy Hin Yan Tong³, Chi-Ho Lin³, Bruno Millet^{4,5}, Nematollah Jaafari^{4,6}, Ridha Joobier⁷, Patrick A Dion^{1,8}, Si Lok³, Marie-Odile Krebs^{4,9,11} & Guy A Rouleau^{1,2,10,11}

Strong Association of De Novo Copy Number Mutations with Autism

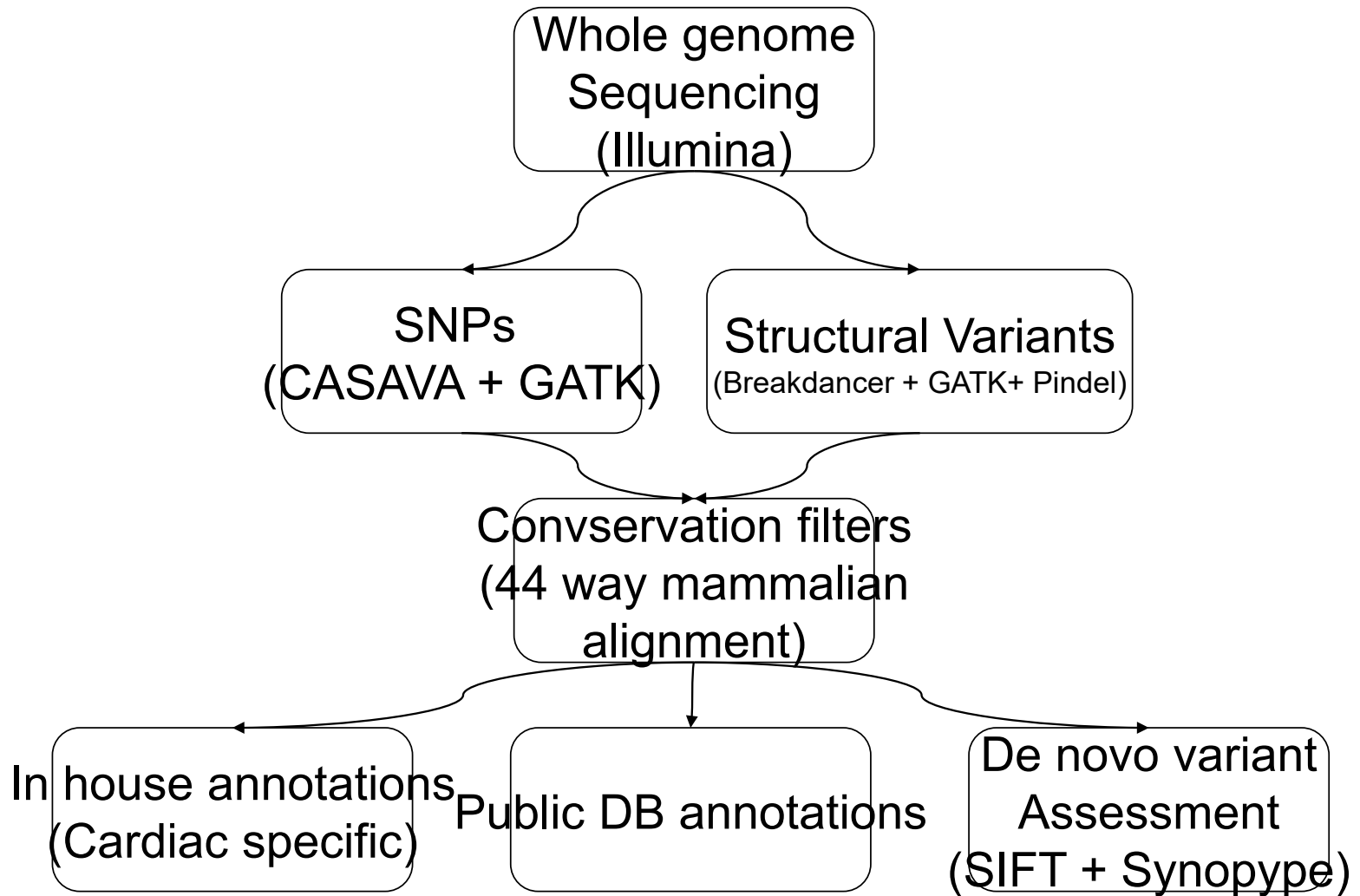
Jonathan Sebat,^{1*} B. Lakshmi,¹ Dheeraj Malhotra,^{1*} Jennifer Troge,^{1*} Christa Lese-Martin,² Tom Walsh,³ Boris Yamrom,¹ Seungtae Yoon,¹ Alex Krasnitz,¹ Jude Kendall,¹ Anthony Leotta,¹ Deepa Pai,¹ Ray Zhang,¹ Yoon-Ha Lee,¹ James Hicks,¹ Sarah J. Spence,⁴ Annette T. Lee,⁵ Kaija Puura,⁶ Terho Lehtimäki,⁷ David Ledbetter,² Peter K. Gregersen,⁵ Joel Bregman,⁸ James S. Sutcliffe,⁹ Vaidehi Jobanputra,¹⁰ Wendy Chung,¹⁰ Dorothy Warburton,¹⁰ Mary-Claire King,³ David Skuse,¹¹ Daniel H. Geschwind,¹² T. Conrad Gilliam,¹³ Kenny Ye,¹⁴ Michael Wigler^{1†}

Hunting for *de novo* mutations using WGS



- **Hypertrophic Cardiomyopathy** manifestation at young age
- No family history
- Parents have no related phenotype
- Genotyping for common associated variants gave negative results





SNPs

Coordinates	Symbol	Description
chr15,51692419,C/T	GLDN	Both variations on gliomedin are scored as damaging. It is expressed in the heart, but not in cardiac myocytes.
chr15,51689763,G/A	GLDN	
chr16,29922353,G/C	KCTD2	These are heart-expressed potassium channel tetramerization domains.
chr17,73043403,G/T	KCTD13	
chr2,179634421,T/G	TTN	Mutations in titin are causal of hypertrophic cardiomyopathies. Note that there is a discrepancy between the protein and DNA sequences annotated in the UCSC browser in this region.

Structural Variants

Coordinates	Symbol	Description
chr1,236899900,frameshift deletion:1	ACTN2	Mutations in actinin alpha 2 have been previously related with sigmoidal and dilated cardiomyopathies
chr5,1800464931, frameshift deletion:2	FLT4	
		Encodes a tyrosine kinase, receptor for endothelial growth factors, which have been associated with hypertrophy

SNVs

Coordinates	Symbol	Description
chr15,51692419,C/T	GLDN	Both variations on gliomedin are scored as damaging. It is expressed in the heart, but not in cardiac myocytes.
chr15,51689763,G/A	GLDN	
chr16,29922353,G/C	KCTD2	These are heart-expressed potassium channel tetramerization domains.
chr17,73043403,G/T	KCTD13	
chr2,179634421,T/G	TTN	Mutations in titin are causal of hypertrophic cardiomyopathies. Note that there is a discrepancy between the protein and DNA sequences annotated in the UCSC browser in this region.

Structural Variants

Coordinates	Symbol	Description
chr1,236899900,frameshift deletion:1	ACTN2	Mutations in actinin alpha 2 have been previously related with sigmoidal and dilated cardiomyopathies
chr5,1800464931, frameshift deletion:2	FLT4	Encodes a tyrosine kinase, receptor for endothelial growth factors, which have been associated with hypertrophy

Targeted high-throughput sequencing in familial cardiomyopathy



Prior expectation – CV disease genes

Gene symbol	Gene name	Curated cellular location	OMIM disease
MYBPC3	myosin binding protein C, cardiac	sarcomere	Cardiomyopathy familial hypertrophic type 4 (CMH4) [MIM:115197]
MYH11	myosin, heavy chain 11, smooth muscle	sarcomere	Aortic aneurysm familial thoracic type 4 (AAT4) [MIM:132900]
MYH6	myosin, heavy chain 6, cardiac muscle, alpha	sarcomere	Atrial septal defect type 3 (ASD3) [MIM:160710]//Cardiomyopathy familial hypertrophic (CMH) [MIM:192600]
MYH7	myosin, heavy chain 7, cardiac muscle, beta	sarcomere	Cardiomyopathy dilated type 1S (CMD1S) [MIM:160760]//Cardiomyopathy familial hypertrophic type 1 (CMH1) [MIM:192600]
MYL2	myosin, light chain 2, regulatory, cardiac, slow	sarcomere	Cardiomyopathy familial hypertrophic type 10 (CMH10) [MIM:608758]//Cardiomyopathy hypertrophic with mid-left ventricular chamber type 2 (MVC2) [MIM:608758]
MYL3	myosin, light chain 3, alkali; ventricular, skeletal, slow	sarcomere	Cardiomyopathy familial hypertrophic type 8 (CMH8) [MIM:608751]//Cardiomyopathy hypertrophic with mid-left ventricular chamber type 1 (MVC1) [MIM:608751]
MYLK2	myosin light chain kinase 2	sarcomere	Cardiomyopathy familial hypertrophic (CMH) [MIM:192600]
MYO6	myosin VI	sarcomere	Deafness, sensineuronal, with hypertrophic cardiomyopathy [MIM:600970]
MYOZ2	myozenin 2	sarcomere	**

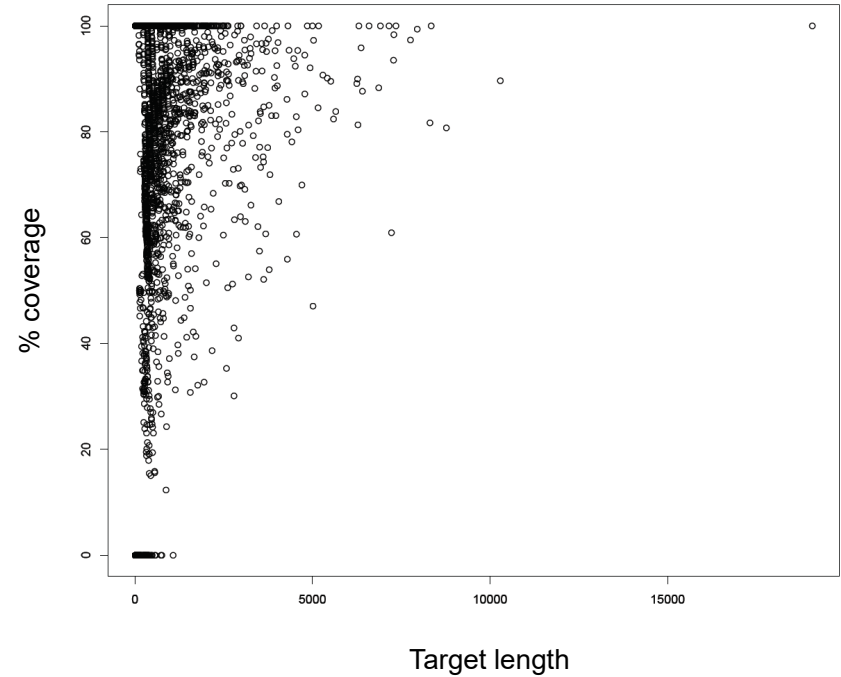
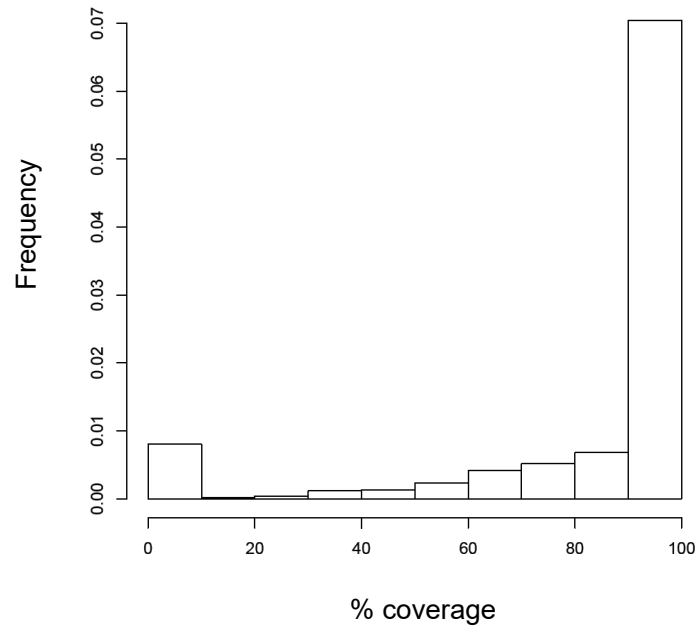
Prior expectation – CV disease genes

➤ Custom Agilent SureSelect in solution hybridization selected regions associated with 251 target genes implicated in myocardial disease and/or function:

- Exons +/- 100 base pairs
 - TFBS within 1.5 kbp
 - Enhancers/repressors within 1.5 Mbp
 - Conserved regions from MSA within 2kbp of coding regions and in introns
 - miRNA target sites
 - miRNA coding regions
- Targets 3.5 Mbp of genomic sequence



Predicted coverage



Study subjects and follow-up work

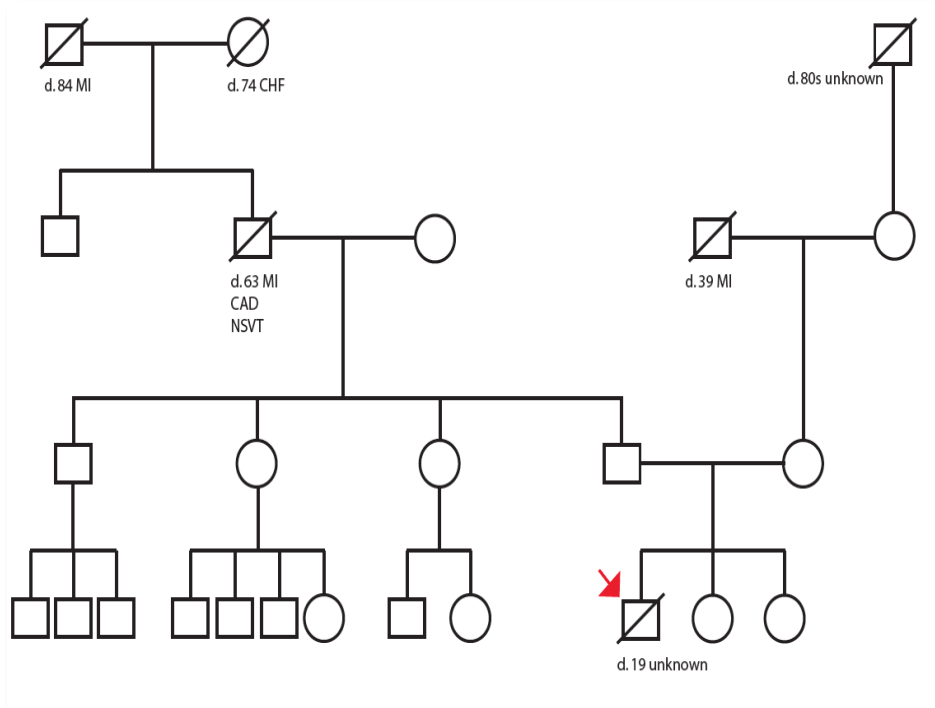
- 80 subjects with dilated and hypertrophic cardiomyopathy in whom clinical genetic testing has not identified a pathogenic mutation
 - 65 subjects have DNA already/15 subjects require DNA collection
- Plan for sequencing on Hi-Seq using indexed primers and sample pooling for maximum sequencing efficiency
- Ethnicity-specific variant identification as prior
- Targeted genotyping/re-sequencing/re-phenotyping contingent on results of discovery sequencing
- Potential clinical tool and tool for investigating modifier loci in other subjects



Molecular autopsy using whole genome sequencing



Pedigree



➤ DNA from paraffin-embedded formalin fixed myocardial tissue was sequenced on Helicos Heliscope





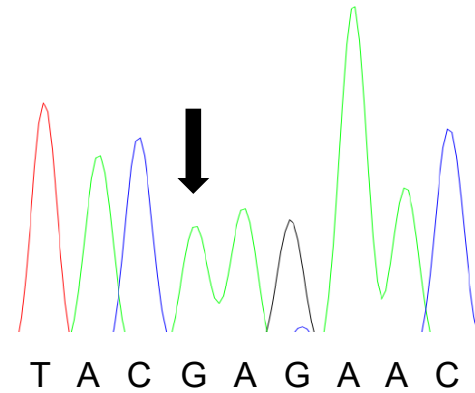
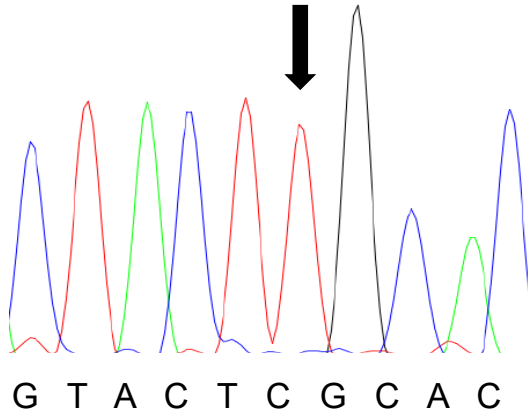
Compound heterozygous variants in the inward rectifying potassium channel Kir2.2



Sanger sequence confirmation

G T A C T T G C A C

T A C A A G A A C



Variant

c.1028C>T

c.1132G>A

Evidence for purifying selection at both nucleotide positions

Variant	GERP score*	PhastCons score^
<i>KCNJ12</i> c.C1028T	4.87	1
<i>KCNJ12</i> c.G1132A	4.84	1

*Range -11.6 to 5.82, with 5.82 being the most conserved

^Range 0 to 1, with 1 being the most conserved



Control chromosomes

- Neither *KCNJ12* variant was seen in 629 control subjects genotyped using low coverage sequencing in the 1000 genomes project
- Neither *KCNJ12* variant was seen in 60 full (high coverage) genomes released by Complete Genomics
- Overall, the variant was absent in 1400 control chromosomes
- However, both variants were present in 4/6 healthy control samples genotyped using a bacterial cloning strategy, one variant was seen in 27/30 separate controls, and the second was seen in 1/30 separate controls
- Both variants fall into a recently annotated pseudogene region, which may mask variation in short-read WGS



Follow-up work

- Sequence read alignment re-processed using GATK for local re-alignment and variant identification
- iPS cells currently are being made from the father and mother of the proband



PERSONALIZED MEDICINE CLINICAL TRIAL



Personalis EX2045.41

Utility of adding genotype risk data to management of coronary artery disease risk

- Well validated CAD SNPs
- Genotype data given to patient, not to doctor
- Follow up for hard end points
 - Lipid panel
- Follow up for behavior endpoints

