Genome resources at the University of Washington, Seattle

Gail Jarvik MD, PhD
Professor and Head, Div. Medical Genetics
Motulsky Chair in Medicine,
Director, NWIGM,

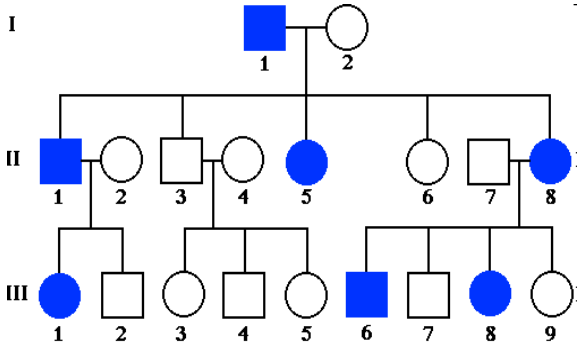
on behalf of UW, Seattle

Outline

- Northwest Genome Center (Nickerson, Rieder)
  - Mendelian Genomic Center (PIs Nickerson, Rieder, Shendure, Bamshad)
  - Seattleseqs; exome variant server
- eMERGE consortium (PIs Jarvik, Larson)
- CLIA sequencing
- Clinical sequencing study-NEXT Medicine (PI Jarvik; also Burke, Veenstra, Nickerson, Rieder, Fullerton, and others)
- Northwest Institute of Genetic Medicine

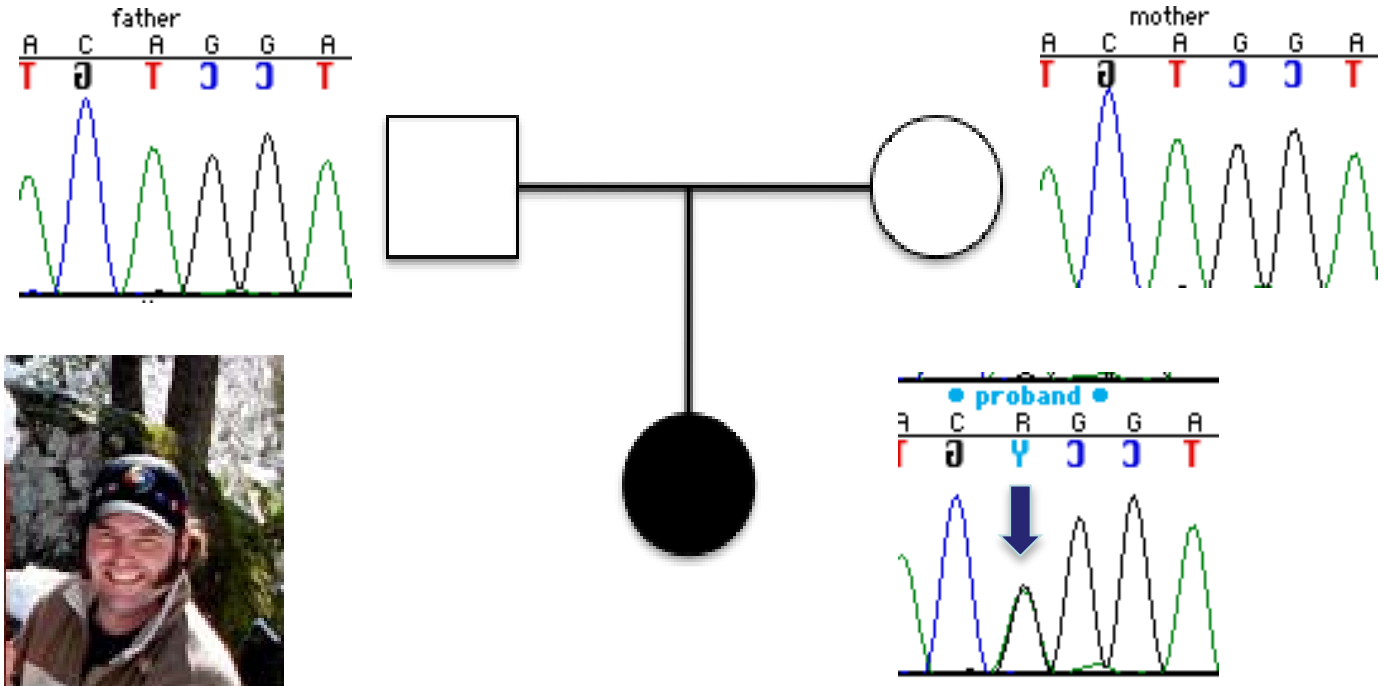
Next Generation Mendelian Genetics Center

- Successful Mendelian strategies
  - Group of unrelated patients with high locus homogeneity
  - Families, esp. recessive or linkage regions (Can have lod<3)
  - Parent-child trios with a *de novo* mutations
- Validation/Replication is crucial for mutations identified in single families
- PIs Nickerson, Rieder, Bamshad, and Shendure
- Accepting unknowns!



Autism Trio-based Exome Sequencing

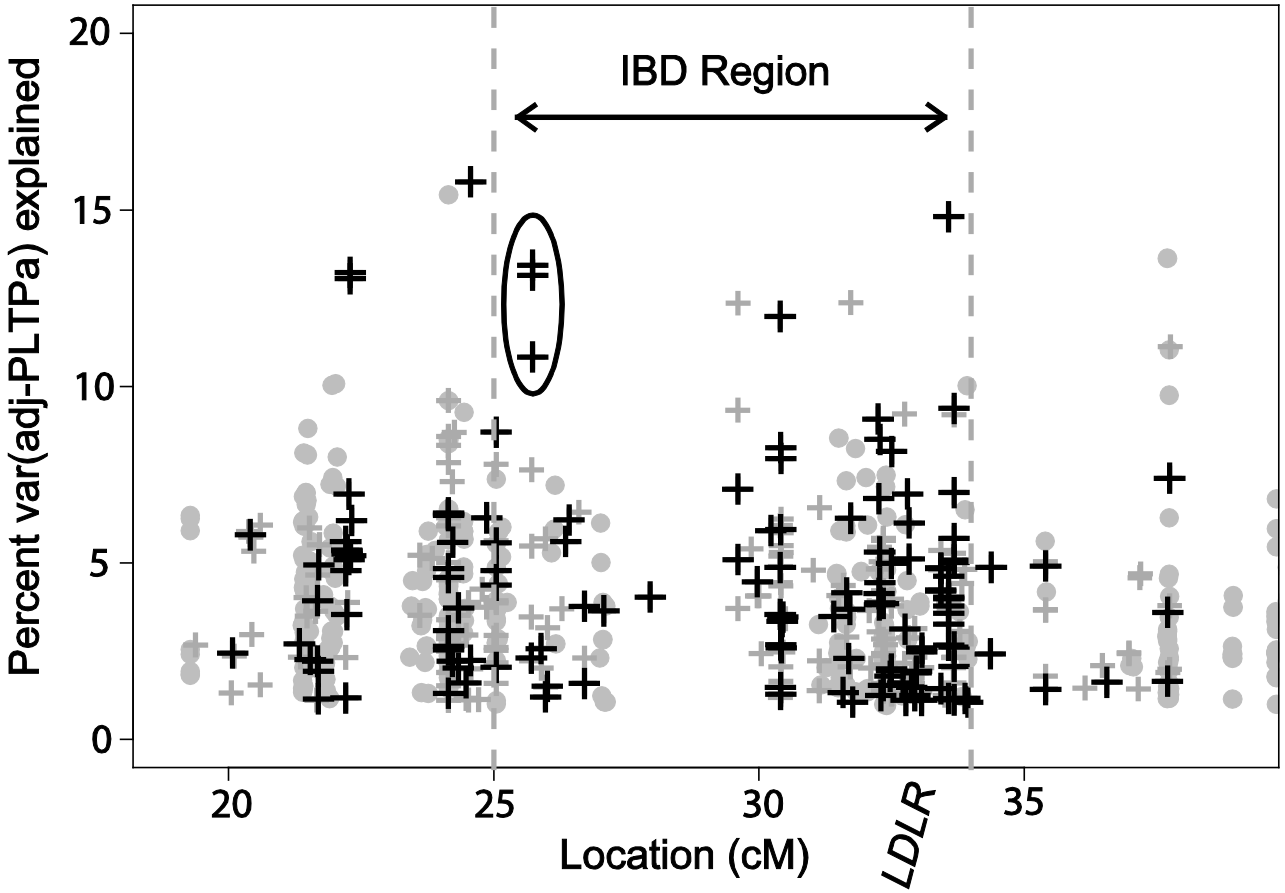
Simplified genetic model that focuses on **single families** & ***de novo*** mutations



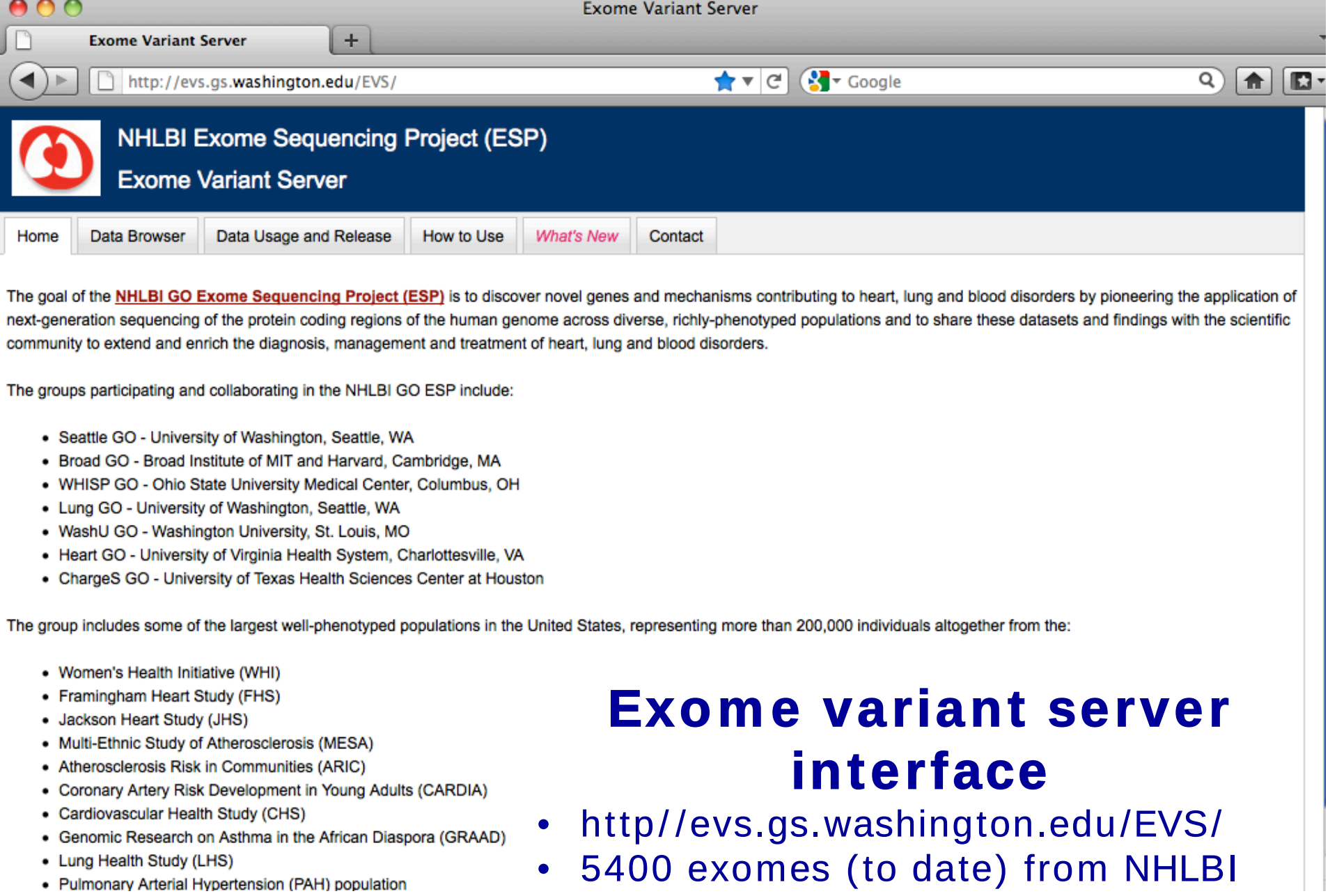
Brian O' Roak

O' Roak et al (2011) Nat. Genet. 43: 585-589. PMID:21572417.

Exomes solve a QTL:
LASS4 effects phospholipid transfer protein



Rosenthal et al. J Lipid Res. 2011 52(10):1837-46. PMID:21757428



Exome variant server interface

- http//evs.gs.washington.edu/EVS/
- 5400 exomes (to date) from NHLBI studies

Exome variant server LDLR query

http//evs.gs.washington.edu/EVS/

Gene Name: LDLR

Gene ID: 3949

Chromosome 19: 11200038 - 11244506 (+)

Population: EuropeanAmerican, AfricanAmerican

Variation Color Code:

splice or nonsense or frameshift

missense

coding-synonymous

coding

utr

Download Option:

File Format: Text

Zip Format: gzip

download

Add or Remove Columns (Description of Columns)

- [x] dbSNP rs ID
- [x] Alleles
- [x] EA Allele Count
- [x] AA Allele Count
- [x] Allele Count
- [x] Sample Read Depth
- [ ] MAF (%)
- [x] Genes
- [x] Gene Accession #
- [x] GVS Function
- [x] Amino Acid
- [x] Protein Position
- [ ] cDNA Position
- [ ] NCBI 37 Allele
- [ ] Chimp Allele
- [ ] Conservation (phastCons)
- [x] Conservation (GERP)
- [x] Grantham Score
- [ ] PolyPhen Prediction
- [x] Clinical Link
- [ ] Filter Status
- [ ] EA Genotype Count
- [ ] AA Genotype Count
- [ ] Genotype Count
- [x] Illumina HumanExome Chip

Sort SNPs by SNP Pos

reset

SNP Pos	rs ID	Alleles	EA Allele #	AA Allele #	All Allele #	Avg. Sample Read Depth	Genes	mRNA Accession #	GVS Function	Amino Acid	Protein Pos.	Conservation (GERP)	Grantham Score	Clinical Link	On Illumin HumanExor Chip
19:11200177	unknown	G/A	G=0/A=7018	G=1/A=3737	G=1/A=10755	26	LDLR	NM_001195803.1	utr-5	none	NA	-3.8	NA	unknown	no

eMERGE – www.gwas.net

- electronic MEdical Records and GEnomics Research Consortium
  - Cooperative Agreement of 7 Partner Institution
    - eMERGE1, Group Health/University of Washington, Marshfield, Mayo, Northwestern, and Vanderbilt
    - eMERGE2 added Geisinger and Mt. Sinai
    - NHGRI funded
  - to develop, disseminate, and apply approaches that combine DNA biorepositories with electronic medical record (EMR) systems for large-scale, high-throughput translational genetic and clinical genomic research
- Plan deployment of the pharmacogenetics research network (PGRN) sequencing array
- Strong bioethics component

CLIA sequencing at UW

- Peter Byers' Collagen arrays, all genes
- Mary Claire King/Tom Walsh
  - 29 gene cancer array; 7 colon (Coloseq)
    - Primers available
  - Many other research interests
  - Current trial of 29 gene sequencing for new invasive breast cancer cases
- Laboratory Medicine
  - King/Walsh Coloseq chip, fee for service
    - King lab supports variant classification
- Coming soon: Nickerson/Rieder CLIA exomes, genomes!

Clinical sequencing in cancer: Clinical, ethical, and technological studies

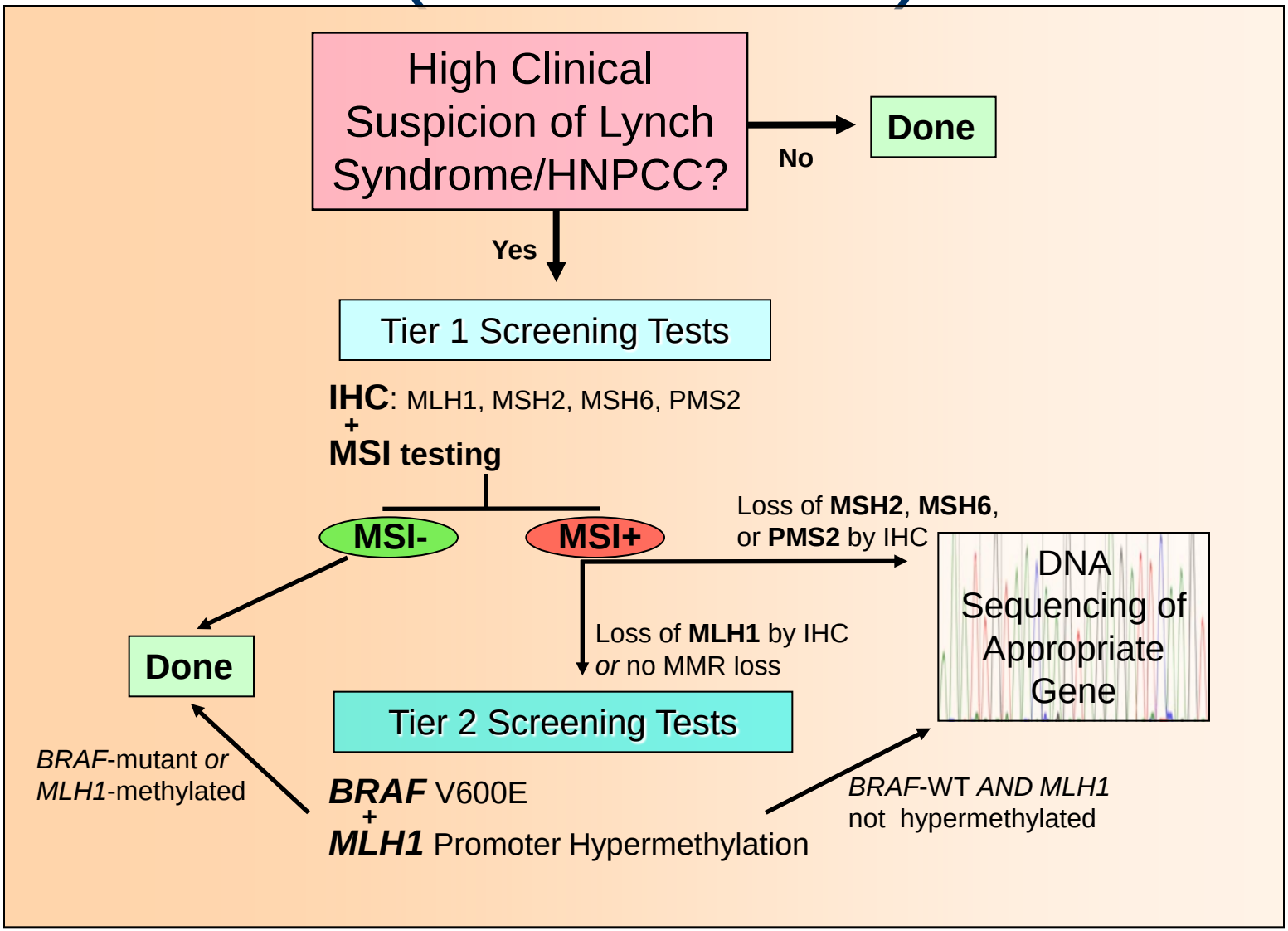
NEXT Medicine (New Exome Technology)

- Project 1: Clinical Genomics study (Jarvik (PI), Veenstra, Patrick, Regier, Heagerty
- Project 2: WXS (Nickerson, Reider)
  - Return of results process (Burke, Evans, Jarvik, Tarczy-Hornoch, et al)
- Project 3: Patient and clinician perspectives (Fullerton, Trinidad, Burke)
- Separate Return of Results RO1: Tabor

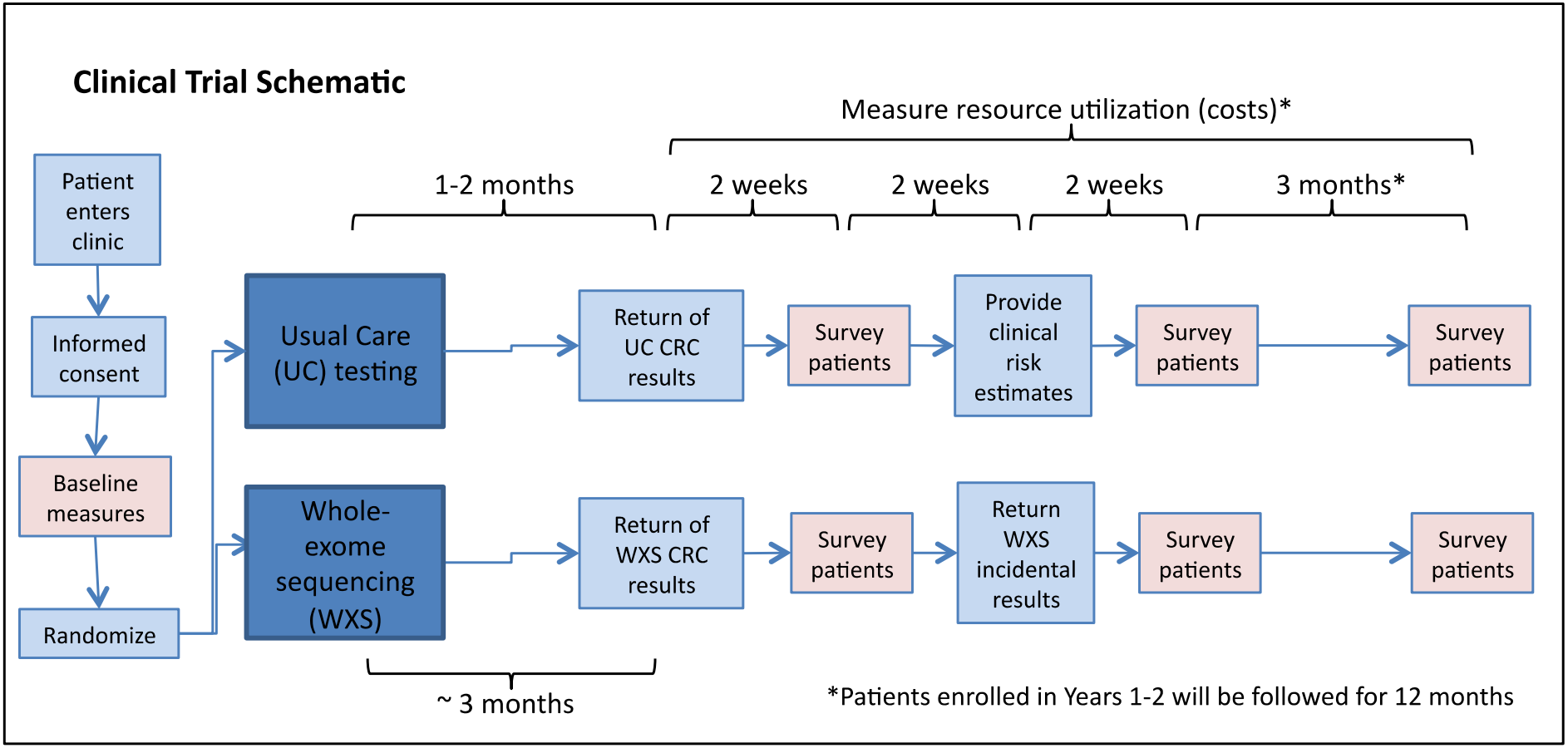
Study Rationale

- Familial CRCP is an ideal disorder to evaluate the utility of exomes for three reasons.
  - First, multiple genes are known to cause similar phenotypes.
  - Second, to arrive at a genetic diagnosis can be time consuming and expensive, requiring multiple clinical visits and tests as well as obtaining tumor samples for pathology studies.
  - Third, in as many as 50% of cases for which the clinician expects Lynch, the causative mutation is not identified
- Thus WXS may offer more efficient and effective approach to identifying genetic causes of CRC

Lynch Syndrome Screening (usual care)



Pritchard and Grady. *Gut* (2011) Slide courtesy of Pritchard



Note that
-Usual Care (UC) may involve multiple visits for MSI/IHC and serial gene tests
-WXS arm includes UC

RCT Study Design

- Comparative
  - Usual care vs. whole exome sequencing (UC vs. WXS plus UC)
- Randomized
  - Control for confounding factors
  - Blinded until return visit (patient and clinician)
- Primary outcome
  - Proportion of patients with a causative genetic mutation identified
  - N = 220
  - 86% power to detect a 20% increase (50->70%)
- Unsolved cases move to a discovery aim, families collected

Patient reported psychosocial and economic outcomes

Patient reported outcome psychosocial (PRO) measures

	Measure	# Items	Length
Symptoms			
Anxiety symptoms	OASIS-5	5	2 minutes
Depressive symptoms	PHQ-9	9	2 minutes
Perceptions			
Self-rated health	NCHS	1	<1 minute
Worry – genetic testing	IGT-AD, modified	16	3 minutes
Satisfaction – genetic testing	CAHPS	1	<1 minute
Decisional conflict	Gotay	3	1 minute

- Healthcare utilization followed by postcards of medical utilization
- Query regarding insurance changes, family members informed
- Also Discrete Choice Experiments (DCE) to value genetic services

Return of incidental exome findings: which?

- Clinical validity and utility (actionable)
- Committee of physicians (mainly medical geneticists) to "bin" results to be returned (Consortium work?)

PANEL MEMBER	INSTITUTION, ROLE	EXPERTISE
Wylie Burke MD PhD	UW, Co-Chair, Co-I	Medical genetics, internal medicine, bioethics
James P Evans MD PhD	UNC, Co-Chair	Medical genetics, genomics
Robin Bennett, MS, CGC	UW, Co-I	Genetic counselor, cancer genetics
Thomas Bird MD	VAMC Seattle	Neurogenetics, neurology
Peter Byers MD, PhD	UW, Co-I	Medical genetics, collagen/vascular, molecular lab
Frederick Chen MD	UW, consultant	Family medicine
William Grady, MD	UW,Co-I	Gastroenterology, Cancer
Fuki Hisama MD	UW, Co-I	Medical Genetics, Neurology
Gail Jarvik MD PhD	UW, PI	Medical genetics, genomics
Katherine Leppig MD	Group Health, consultant	Medical genetics, cytogenetics, eMERGE RORC
Jeff Murray, MD, PhD	Univ. Iowa	Medical genetics, pediatrics
Wendy Raskind, MD	UW, consultant	Medical Genetics, General Int. Med, cancer
Virginia Sybert, MD	UW, consultant	Medical & Dermatological Genetics, Turner syndrome
Benjamin Wilfond MD	UW/CHRC, consultant	Pediatrics, bioethics
EXPERT ADVISORS		
Mark Rieder	UW, Co-I	Genomics, pharmacogenomics
Debbie Nickerson	UW, Co-I	Genomics
S. Malia Fullerton	UW, Co-I	Bioethics, eMERGE RORC
Genetic counselor,	TBN	

NEXT Medicine Bioethics (Burke, Fullerton, Trinidad)

- **Characterize patients' and referring providers' attitudes and preferences regarding the return of exome sequencing results (focus groups).**
- **Explore patients' views and experiences of receiving genetic test findings generated from exome sequencing:**
  - **Elicit end-to-end first-person accounts from patients who receive both CRC and non-CRC risk information from exome sequencing, as well as the views of their referring providers.**
  - **Describe and compare the experiences of patients who receive CRC risk information via exome sequencing to those who receive the usual-care workup for CRC risk.**
  - **Describe and compare the views and experiences of patients who receive different types of exome sequence information (unrelated to CRC risk).**
- **Legal analysis of whether a requirement for CLIA compliance as a precondition to returning results from genomic research studies violates the First Amendment (Barbara Evans, JD, U Houston).**

Northwest Institute of Genetic Medicine

Gail Jarvik
Medical Genetics

UNIVERSITY OF WASHINGTON

Peter Tarczy-Hornoch
Biomedical Informatics,
Pediatrics, Computer Science

Children's
Hospital & Regional Medical Center
Seattle, Washington

Debbie Nickerson
Genome Sciences
NW Genome Center

UNIVERSITY OF WASHINGTON

Bruce Weir
Biostatistics

UNIVERSITY OF WASHINGTON

Mike Bamshad
Pediatrics

Children's
Hospital & Regional Medical Center
Seattle, Washington

Supported by
Eric Larson,
Group Health

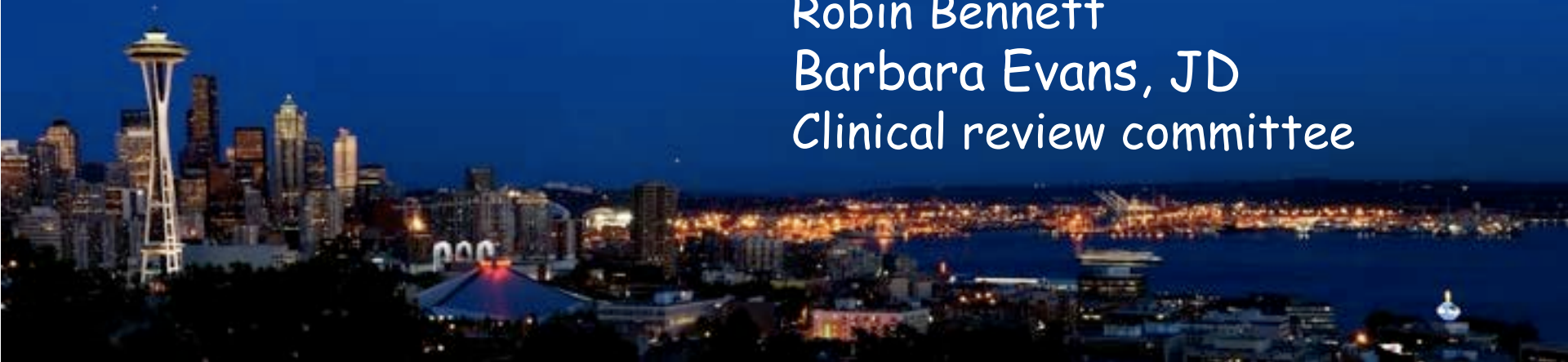
GroupHealth

Partial Acknowledgements

Debbie Nickerson
Mark Rieder
Jay Shendure
Mike Bamshad
Evan Eichler

NHLBI, NHGRI
Simons Foundation,
WA State Life Sciences
Discovery Fund,
NICHD, NIGMS, NIAID

Dave Veenstra
Wylie Burke
Malia Fullerton
Donald Patrick
Chris Nefcy
Peter Byers
Dean Regier
Fuki Hisama
Peter Tarczy-Hornoch
Brian Browning
Patrick Heagerty
Robin Bennett
Barbara Evans, JD
Clinical review committee



Enrollment

- In (first) clinic visit
- Subjects with CRCP where a single gene is not highly implicated
  - Exclude
    - Very likely APC (>100 polyps?)
    - Known mutation in family
    - Syndromic features suggest the diagnosis

Who and when enrolls

- GCs can enroll
- Fulltime (junior) GC to support study
- Martha can enroll

- Enroll and randomize at first visit

Clinically, what then

- Randomized to UC or WXS plus UC
- For both do your usual protocol (lets discuss)
- For WXS they have a blood test for exome exome
- Return to clinic for UC billed visits
- Each will have 1 extra, non-billed visit
  - Incidental Exome findings for WXS
  - Review of family risks for UC

Patient Outcomes

- Prior studies of CRC genetic testing report distress and anxiety scores within normal limits or moderately increased following disclosure of results
- Collins et al. reported an increase in cancer-specific distress in carriers at 2 weeks post-disclosure, followed by a return to baseline levels at 12 months that was stable 3 years later.
- Several studies have identified demographic and psychological factors (e.g., baseline mood disturbance, state anxiety, cancer worry, resilience, cognitive style, coping style) that are correlated with increased distress.
- Given the potential extensive scope of incidental findings from exome sequencing, these effects warrant further study.

Follow psychosocial and economic outcomes

- Healthcare-related resource utilization (HRU) will be collected using a patient survey implemented with a postcard [online?]return every month
- Patients will be asked about
  - use of medical services such as physician visits, hospitalization, prescription and non-prescription drug use, screening, ancillary care, and mental health services.
  - how many family members they have informed of their test results, and what actions their family members have taken to their knowledge – e.g., received genetic testing or CRC screening.
  - actual or intended changes to their health and life insurance policies.

Discrete Choice Experiments (DCEs)

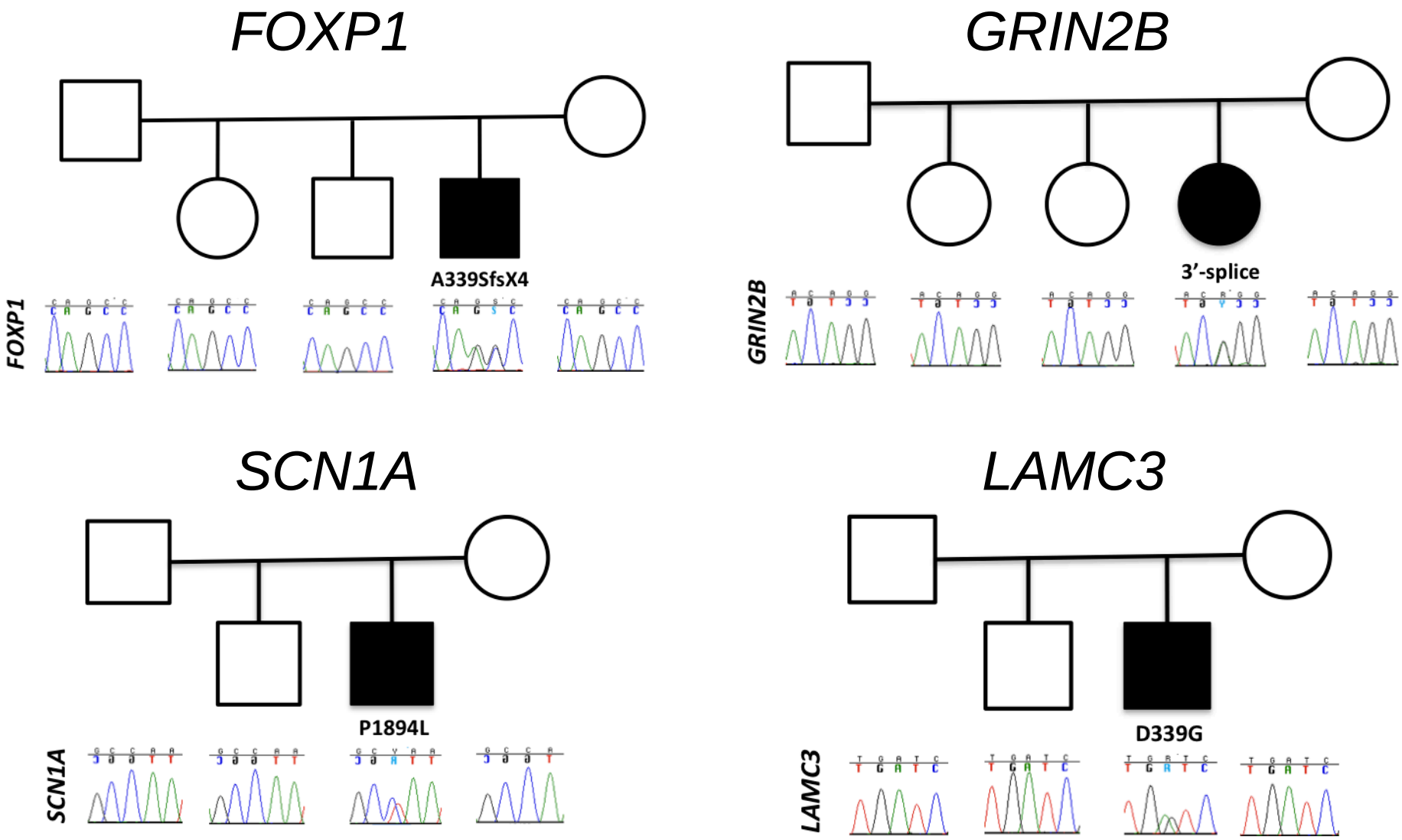
- DCEs assume
  - that health care ‘goods’ can be described by two or more attributes (e.g., probability of finding a genetic risk of CRCP; time waiting for results; cost of testing),
  - that each attribute is defined on a number of levels (e.g., 40% chance, 80% chance; 2 weeks, 8 weeks; $750, $2000)

How to find a needle in a haystack?

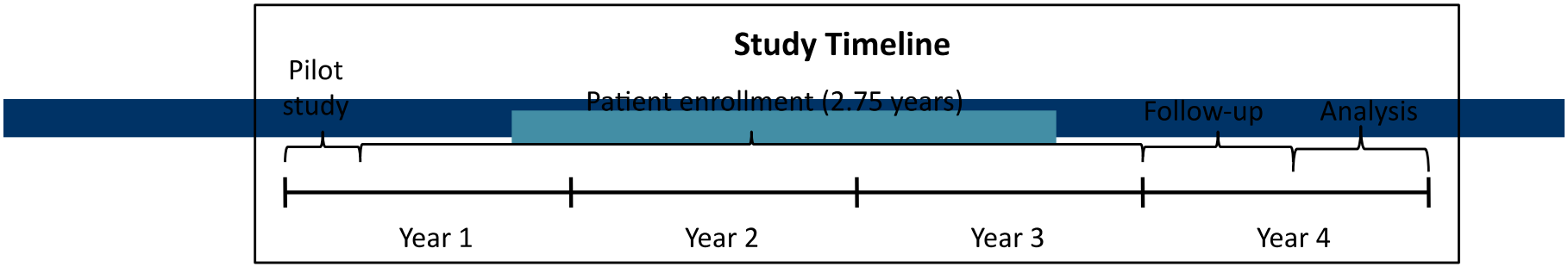
~1 *de novo* event expected per trio

16,000-20,000 exome variants

20 Pilot Trios



O'Roak et al. Nat Gen. May


Study Timeline
Pilot study
Patient enrollment (2.75 years)
Follow-up
Analysis
Year 1
Year 2
Year 3
Year 4

Exome sequencing is transforming Mendelian Genetic Analysis

Disorder	Mode	N	Strategy	Gene(s)	PMID
			Comparison of unrelated cases		
Kabuki	AD	10	10 cases / 10 kindreds	*MLL2*	20711175
Schinzel-Giedion	AD	4	4 cases / 4 kindreds	*SETBP1*	20436468
Hadju-Cheney	AD	3	3 cases/ 3 kindreds	*NOTCH2*	21378985
Fowler	AR	2	2 cases / 2 kindreds	*FLVCR2*	20518025
Sensenbrenner	AR	2	2 cases / 2 kindreds	*WDR35*	20817137
			Comparison of related cases		
Miller	AR	4	4 cases / 3 kindreds	*DHODH*	19915526
hyperphosphatasia-MR	AR	3	3 cases / 1 kindred	*PIGV*	20802478
hypolipidemia	AR	2	2 cases / 1 kindred	*ANGPTL3*	20942659
retinitis pigmentosa	AR	3	3 cases / 1 kindred	*DHDDS*	21295283
novel skeletal dysplasia	AR	4	2 cases + parents / 1 kindred	*POP1*	21455487
spinocerebellar ataxia	AD	4	linkage + X cases / 1 kindred	*TGM6*	21106500
familial ALS	AD	2	linkage + 2 cases / 1 kindred	*VCP*	21145000
dilated cardiomyopathy	AD		linkage + 4 cases / 1 kindred	*BAG3*	21353195
hidradenitis suppurativa	AD	3	linkage + 2 cases[3] / 1 kindred	*NCSTN*	21430701
spinocerebellar ataxia	AD	4	linkage + 4 cases / 1 kindred	*TGM6*	21106500
primary failure tooth eruption	AD	4	linkage + 4 cases[4] / 1 kindred	*PTH1R*	21404329
TARP[1]	XLR	2	linkage + 2 cases / 2 kindreds	*RBM10*	20451169
X-linked luecoencephalopathy	XLR	2	linkage + 1 case[3] / 1 kindred	*MCT8*	21415082
			Homozygosity mapping		
DFNB82 (hearing loss)	AR	1	1 case / 1 kindred	*GPSM2*	20602914
CNS malformations	AR	2	2 cases / 1 kindred	*WDR62*	20729831
Seckel	AR	1	1 case / 1 kindred	*CEP152*	21131973
NPHP-related ciliopathy[2]	AR	1	1 case / 1 kindred	*SDCCAG8*	20835237
autoimmune LPS	AR	1	1 case / 1 kindred	*FADD*	21109225
3MC	AR	1	1 case / 1 kindred	*MASP1*	21035106
complex I deficiency	AR	1	1 case / 1 kindred	*ACAD9*	21057504
non-syndromic MR	AR	2	2 obligate carrier parents	*TECR*	21212097
Ochoa	AR	1	1 case / 1 kindred	*HPSE2*	21450525
			Identification of de novo mutations		
sporadic MR	complex	30	10 parent-child trios	multiple	21076407
autism	complex	60	20 parent-child trios	multiple	

Why Exomes?

Advantages:

- More interpretable
- Easier to follow up
- Larger effect size
- Cheaper and sample size counts

Disadvantages:

Miss non-coding variants and some coding
We do genomes when we need to!

Some of the Challenges in Exome Analysis

- Undercalling of coding variants
  (SNVs, indels, and CNVs)

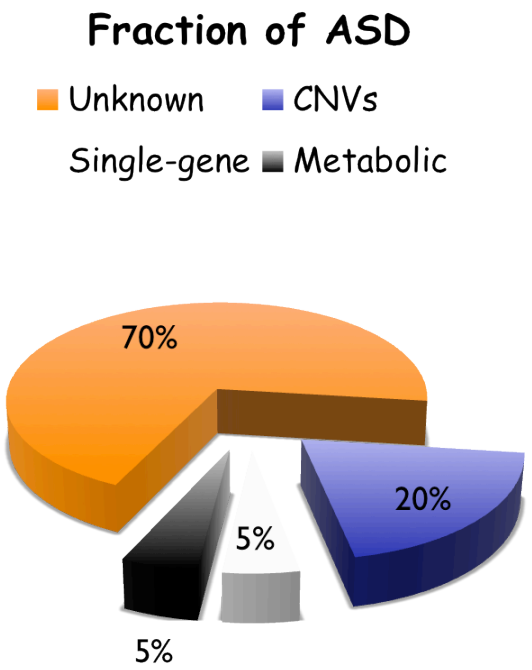
- Causal non-coding

- Soft phenotyping and/or modifiers

- Genetic heterogeneity at all levels

Genetics of Autism

- Strong genetic component ~70-90%
- Linkage and GWAS have uncovered few consistent genes or regions
- Likely widespread heterogeneity
- How do we get at the 70% of unknown causes?



*Modified from Schaaf and Zoghbi 2011

Apply a *de novo* variant approach

Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome

Sarah B Ng[1,7], Abigail W Bigham[2,7], Kati J Buckingham[2], Mark C Hannibal[2,3], Margaret J McMillin[2], Heidi I Gildersleeve[2], Anita E Beck[2,3], Holly K Tabor[2,3], Gregory M Cooper[1], Heather C Mefford[2], Choli Lee[1], Emily H Turner[1], Joshua D Smith[1], Mark J Rieder[1], Koh-ichiro Yoshiura[4], Naomichi Matsumoto[5], Tohru Ohta[6], Norio Niikawa[6], Deborah A Nickerson[1], Michael J Bamshad[1–3] & Jay Shendure[1]

Ng et al.
Nat Genet, Aug 2010

A *de novo* paradigm for mental retardation

Lisenka E L M Vissers[1,2], Joep de Ligt[1,2], Christian Gilissen[1], Irene Janssen[1], Marloes Steehouwer[1], Petra de Vries[1], Bart van Lier[1], Peer Arts[1], Nienke Wieskamp[1], Marisol del Rosario[1], Bregje W M van Bon[1], Alexander Hoischen[1], Bert B A de Vries[1], Han G Brunner[1,3] & Joris A Veltman[1,3]

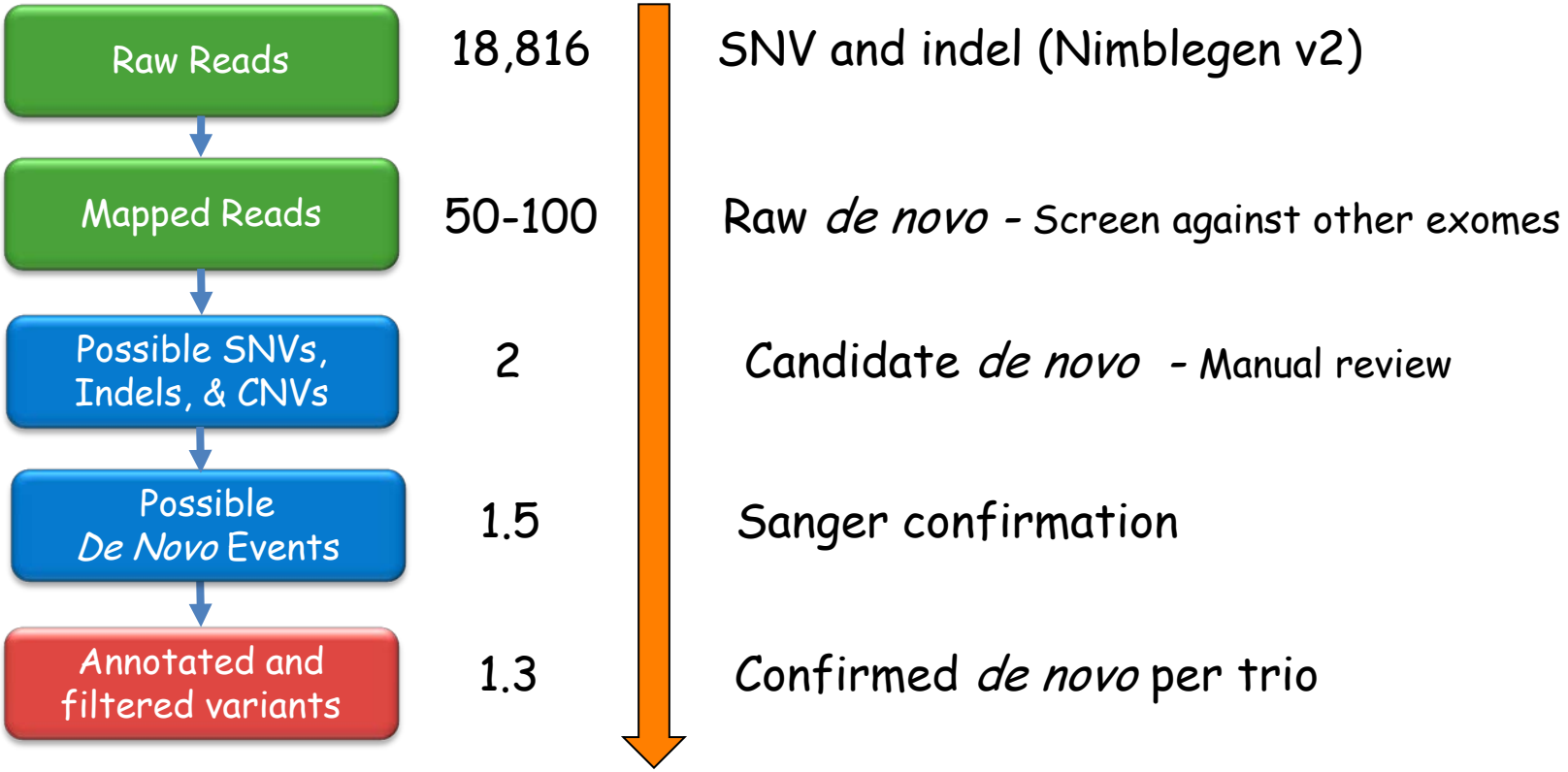
Vissers et al.
Nat Genet, Nov 2010

Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations

Brian J O'Roak[1], Pelagia Deriziotis[2], Choli Lee[1], Laura Vives[1], Jerrod J Schwartz[1], Santhosh Girirajan[1], Emre Karakoc[1], Alexandra P MacKenzie[1], Sarah B Ng[1], Carl Baker[1], Mark J Rieder[1], Deborah A Nickerson[1], Raphael Bernier[3], Simon E Fisher[2,4], Jay Shendure[1] & Evan E Eichler[1,5]

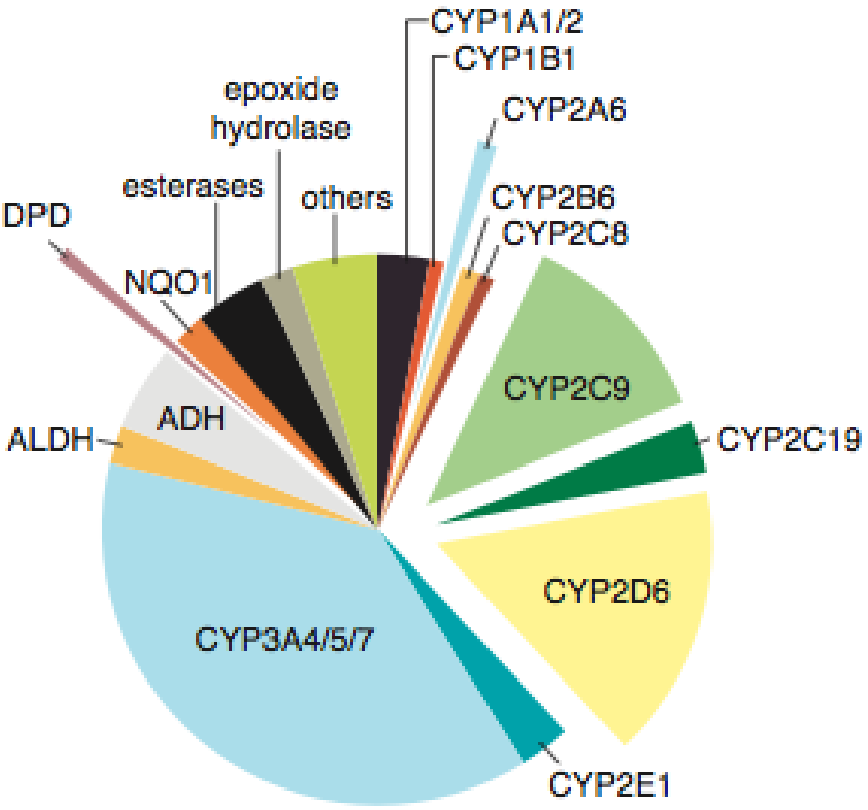
O' Roak et al
Nat. Genet. May, 2011

Trio Based Exome Sequencing



Drug Metabolism: Cytochrome P450s

- Oxidize many biological substances using heme cofactor
- Small handful of *CYP*s responsible for 75% of drug responsiveness in humans
- Genetic variation in drug response responsible for up to 30% of all ADRs

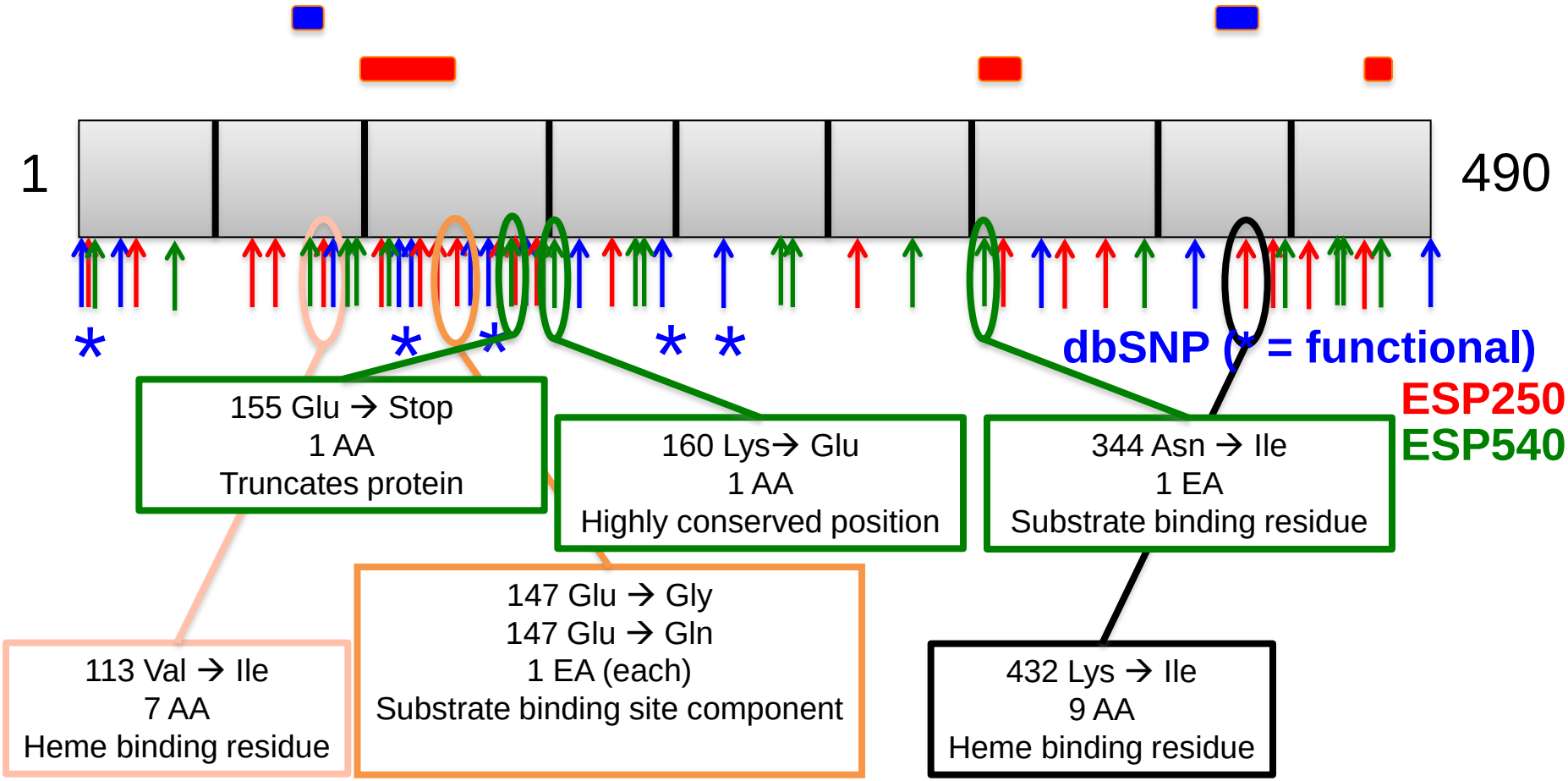


Evans & Relling *Science* (199

Coding Variation in *CYP2C19* (Plavix, Warfarin, Valium)

Heme binding site

Substrate binding site



NHLBI Grand Opportunity Exome Sequencing Project (ESP)

SeattleGO
Debbie Nickerson
Mark Rieder
Jay Shendure
Phil Green
Josh Akey
Mike Bamshad
Carlos Bustamante
Evan Eichler
Suzanne Leal
Bryan Paeper
Peggy Robertson
Josh Smith
Emily Turner

BroadGO
David Altshuler
Stacey Gabriel
Goncalo Abecasis
Mark Depristo
Deborah Farlow
Kiran Giramella
Youna Hu
Goo Jun
Hyun Min Kang
Sekar Kathiresan
Shamil Sunyaev
Cristen Willer
Chenyi Xue

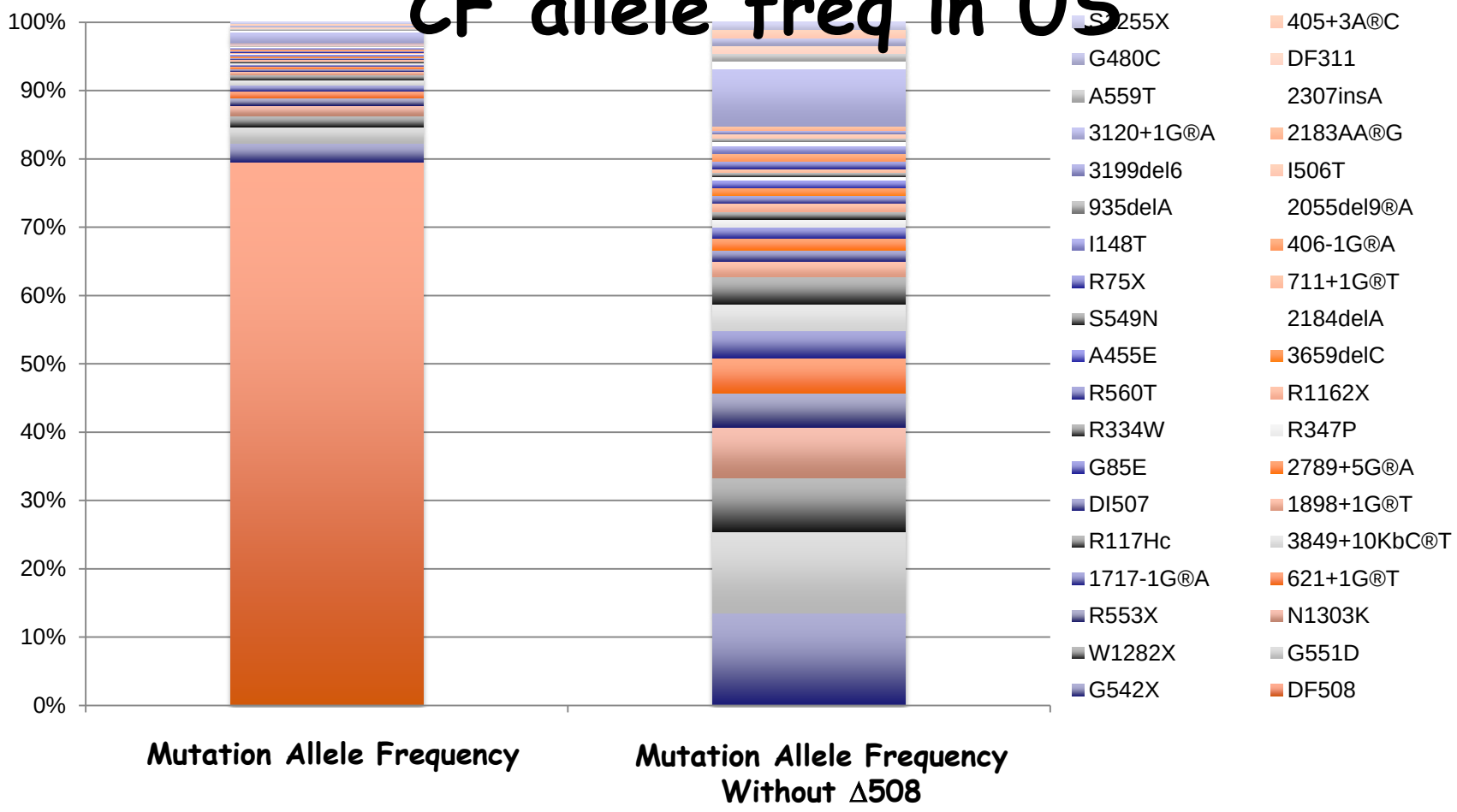
LungGO
Mike Bamshad
Kathleen Barnes
Mary Emond
Ron Gibson
Mike Knowles
Rasika Mathias
Ed Silverman
Holly Tabor
Fred Wright
Mark Wurfel

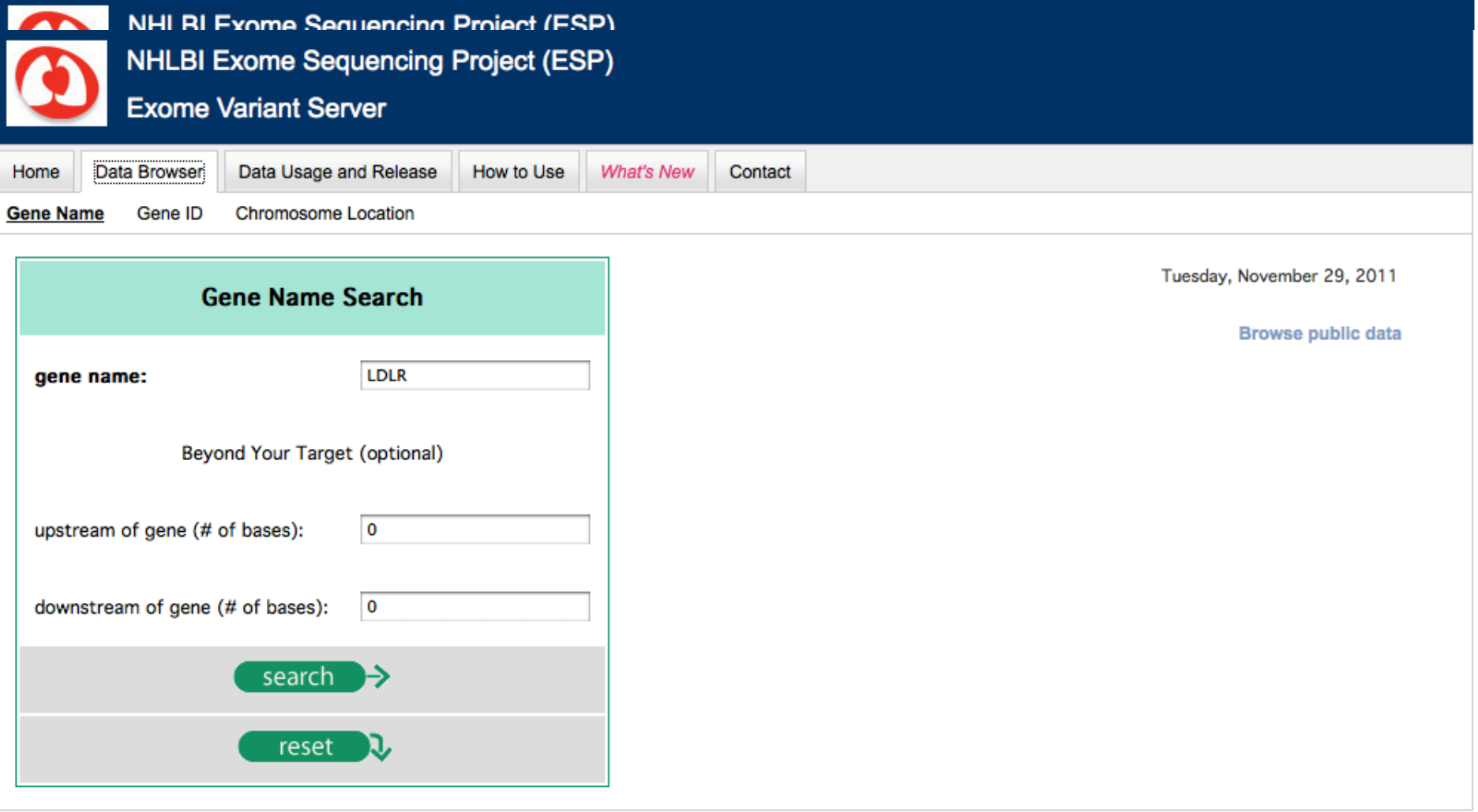
HeartGO
Stephen Rich
Larry Atwood
Eric Boerwinkle
Myron Gross
Leslie Lange
Alanna Morrison
Christopher O'Donnell
Bruce Psaty
Wendy Post
Alex Reiner
Jerome Rotter
Herman Taylor
Russell Tracy
James Wilson

WHISP
Rebecca Jackson
Chris Carlson
Kari North
Ulrike Peters
Chris Bizon
Nora Franceschini
Malia Fullerton
Li Hsu
Charles Kooperberg
Ethan Lange
Leslie Lange
Yun Li
Danyu Lin
Keri Monda
Alex Reiner
Kira Taylor

CFTR

CF allele freq in US




NHLBI Exome Sequencing Project (ESP)
Exome Variant Server
Home
Data Browser
Data Usage and Release
How to Use
What's New
Contact
Gene Name
Gene ID
Chromosome Location
Gene Name Search
gene name:
LDLR
Beyond Your Target (optional)
upstream of gene (# of bases):
0
downstream of gene (# of bases):
0
search
reset
Tuesday, November 29, 2011
Browse public data

NHLBI Exome Sequencing Project (ESP)

Exome Variant Server

SNP Results | Coverage Results

Gene Name: LDLR

Gene ID: 3949

Chromosome 19: 11200038 - 11244506 (+)

Select Data Set(s)

Check at least one data set below.

Select	Number Variations	Population
☑	144	EuropeanAmerican
☑	121	AfricanAmerican

Display Results

display snp summary

reset

.0008 FH DENVER 2

LDLR, ASP283ASN [dbSNP:rs121908030]

In an African American patient with FH (143890), Leitersdorf and Hobbs (1990) found a change of aspartic acid-283 (GAC) to asparagine (AAC).

.0009 FH AFRIKANER 2

LDLR, VAL408MET [dbSNP:rs28942078]

This and the asp206-to-glu mutation (see 606945.0006) are frequent among Afrikaners with FH (143890). A GTG-to-ATG mutation is responsible (Leitersdorf et al., 1989). In a study of 138 chromosomes of Afrikaner FH patients, Kotze et al. (1991) found that 31 (23.3%) had this mutation. Schuster et al. (1993) found the same mutation in a German family and showed that it existed on the same 7-RFLP haplotype as did the mutation described in South Africa and in the Netherlands, suggesting a common European origin. Similarly, Defesche et al. (1993) found the val408-to-met mutation in 19 (1.5%) of 1,268 FH patients of Dutch descent. In 9 of the patients carrying this mutation on one allele, the LDLR DNA haplotype was that observed in a South African FH patient homozygous for the same mutation. The remaining 10 Dutch FH patients all shared a common haplotype, partly identical to the Afrikaner haplotype, which could have arisen from a single recombinational event. With the exception of the family reported by Schuster et al. (1993), this mutation has been described only in persons of Dutch ancestry.

.0010 FH ALGERIA

LDLR, ALA410THR [dbSNP:rs28942079]

A GCT-to-ACT change is responsible for this variant (Zuliani and Hobbs, 1990).

Lots of people

Gail Jarvik
David Veenstra
Wylie Burke
S. Malia Fullerton
Debbie Nickerson
Mark Rieder
Fuki Hisama
Peter Tarczy-Hornoch
William Grady
Wendy Raskind
Arno Motulsky
Brian Browning
Virginia Sybert
Patrick Heagerty
Sara Goering
Donald Patrick
Robin Bennett
GC to write blurbs
Debbie Olson
Peter Byers
Emily Turner
David Crosslin
Emily Hendricks

Martha Horike Pyne
Jane Ranchalis
Beverly Berg-Rood
Brian Comstock
Chris Nefcy
Susan Trinidad
Josh Smith
Bryan Paeper
Jeff Furlong
Peggy Robertson
Katie Igartua
CLIA Compliance Officier
MITS Clincal Comput Dev
Grad student RA for Outcomes

Exome

- 180,000 exons in human genome
- 1% of the human genome
- 30 megabases (Mb)
  - 30M results?
- Estimated to constitute about 85% of the disease-causing mutations

RFA ->UO1 Proposal

- Project 1
  - Clinical Genomics study (Jarvik (PI), Veenstra, Patrick, Regier, Heagerty)
- Project 2
  - WXS (Nickerson, Reider)
  - Return of results process (Burke, Jarvik, et al)
- Project 3
  - Patient and clinician perspectives (Fullerton, Trinidad)