The Path to Precision Medicine: From Discovery to Patient Care

Alan R. Shuldiner, MD
Vice President
Regeneron Genetics Center &
Professor (part-time), University of Maryland School of Medicine
The Reality of Therapeutic Development in 2017

- Despite increased investment in R+D in the pharmaceutical industry, the number of new molecular entities is not increasing.
- >90% of molecules that enter Phase I clinical trials fail to demonstrate sufficient safety and efficacy to gain regulatory approval.
- Most failures occur in Phase II clinical trials:
  - 50% due to lack of efficacy
  - 25% due to toxicity
- Pre-clinical models may be poor predictors of clinical benefit.
- Compounds supported by human genetics evidence are substantially more likely to succeed.
The Power of Human Genetics to Accelerate Target Identification, Validation and Drug Development

2003
- Family studies identify PCSK9 GOF as a causal FH gene

2006
- Population studies identify PCSK9 LOF variants conferring ~88% reduction in CHD

2012
- Clinical proof of concept

2008
- Null APOC3 mutation enriched in Amish points to cardio-protective effects

2014
- Two population studies identify variants conferring ~40% reduction in CHD

2015
- Clinical proof of concept
Congenital Insensitivity to Pain (CIP) and the *SCN9A* Gene: Human Genetics Provides Insights Into New Pain Drug Targets

- CIP → pain free burns, fractures, childbirth, etc
- Extremely rare: <1/1,000,000 prevalence
- Mutations in *SCN9A* cause insensitivity to pain
- Efforts to mimic the effects of pain insensitivity through therapeutics blocking the corresponding protein are being pursued
Application of Human Genetics to Accelerate Novel Target Identification and Clinical Development

The RGC applies large-scale, fully-integrated human genetics approaches to advance science, guide the development of therapeutics, and improve patient outcomes.

“Do Well by Doing Good”

**Indication Discovery**
Identify new indications for drug targets and programs

**Target Discovery**
Identify new drug targets and pathways

**Biomarker**
Develop pharmacogenetic markers to predict drug response

**Derisking**
Confirm lack of “on-target adverse side effects” in drug target LoF carriers

**Human Genetics**

**Mouse Genetics**

**Genetic Classifier**
Responders
Non-Responders
Engine Driving Identification of Rare Loss of Function Mutations that Inform Human Biology and Drug Development: Ultra High-Throughput Sequencing and Analysis at the Regeneron Genetics Center

**Technologies and Capabilities**

- Automated biobank with 1.4M+ sample capacity
- Custom fully-automated exome and targeted sequencing sample preparation workflows
- Currently exome sequencing >2,500 exomes per week
  - >250,000 exomes completed
- Among the first “genome center in the cloud” with fully automated analysis pipelines
Maximizing Discovery Opportunities by Leveraging Human Genetics Resources Across Genetic Trait Architecture and Phenotypes

50+ Academic collaborators – Over 250,000 exomes sequenced

Integrated approaches across genetic trait architectures . . .

- General Population
- Phenotype Specific Cohorts
- Family Studies
- Founder & Special Populations

... will power genomic discovery
Geisinger-Regeneron DiscovEHR Collaboration

Two organizations focused on making genomic data medically actionable

Goal: Build comprehensive genotype-phenotype resource combining de-identified genomic and clinical data from >250,000 people to aid drug development and implementation of genomic medicine into patient care

• Geisinger: Integrated health care system
  – 1.6M participants
  – Amongst earliest adopters of EHRs (1996) and leaders in clinical informatics
    • Longitudinal EHR data: Median of ~18 outpatient visits per patient over 13.4 years

• Recruitment ongoing
  – 119,000 patients consented into MyCode-DiscovEHR cohort
  – >90,000 sequenced at the Regeneron Genetics Center
  – Large unselected populations as well as targeted efforts in diseases of interest and deeply phenotyped patients
    • Cardiac catheterization lab (~8,000)
    • Bariatric surgery (~4,000) - one of the largest in the world
HUMAN GENETICS

Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study

Frederick E. Dewey,1* Michael F. Murray,2 John D. Overton,1 Lukas Habegger,1 Joseph B. Leader,4 Samantha N. Fettorli,2 Colm O'Dushlaine,1
Christoph V. Van Hout,1 Jeffrey Staples,1 Claudia Gonzaga-Jauregui,1 Raghu Metpally,2 Sarah A. Pendergrass,2 Monica A. Giovanni,2 H. Lester Kirchner,2
Suganthi Balasubramanian,1 Noura S. Abul-Husn,1 Dustin N. Hartzel,2
Daniel R. Lavage,2 Korey A. Kost,1 Jonathan S. Packer,1 Alexander E. Lopez,1
John Penn,2 Semanti Mukherjee,1 Nehal Gosalia,1 Manoj Kanagaraj,1 Alexander H. Li,1
Lyndon J. Mitnail,1 Lance J. Adams,2 Thomas N. Person,1 Kavita Praveen,1
Anthony Marcketta,1 Matthew S. Lebo,7 Christina A. Austin-Tse,3
Heather M. Mason-Saure,3 Shannon Bruse,1 Scott Mellis,4 Robert Phillips,4
Neil Stahl,5 Andrew Murphy,5 Aris Economides,1 Kimberly A. Skelding,2
Christopher D. Still,2 James R. Elmore,2 Ingrid B. Borecki,1 George D. Yancopoulos,4
F. Daniel Davis,4 William A. Faucci,2 Omri Gottesman,1 Marylyn D. Ritchie,2
Alan R. Shuldiner,3 Jeffrey G. Reid,1 David H. Ledbetter,2 Aris Baras,1 David J. Carey2*

The DiscovEHR collaboration between the Regeneron Genetics Center and Geisinger Health System couples high-throughput sequencing to an integrated health care system using longitudinal electronic health records (EHRs). We sequenced the exomes of 50,726 adult participants in the DiscovEHR study to identify ~4.2 million rare single-nucleotide variants and insertion/deletion events, of which ~176,000 are predicted to result in a loss of gene function. Linking these data to EHR-derived clinical phenotypes, we find clinical associations supporting therapeutic targets, including genes encoding drug targets for lipid lowering, and identify previously unidentified rare alleles associated with lipid levels and other blood level traits. About 3.5% of individuals harbor deleterious variants in 76 clinically actionable genes. The DiscovEHR data set provides a blueprint for large-scale precision medicine initiatives and genomics-guided therapeutic discovery.
In-Depth, Longitudinal Health Records Enriched for Age-Related Diseases and Phenotypes

Patients by Years of Clinical Data

Most Prevalent Labs in GHS EHR

Most Prevalent Office Visit Dx in GHS EHR
The RGC and GHS Have Developed A Large Number of High-quality, EHR-derived Phenotypes For Genetic Analyses

A constantly growing library of more than 8,000 quantitative and binary traits are available for high-throughput and in-depth genotype-first and phenotype-first analyses:

**Binary and Quantitative Trait Matrices:**
Include PheWAS, Immune, Lab Traits, DEXA, Echo, EKG, Ocular Measures, PFT’s, Vitals and Anthropometrics

**Deep Dive Datasets:**
Examples include Coronary Artery Disease and Lipids, COPD and Asthma, Bariatric Traits and Liver Histology, Gout
Sequence Variants Identified Using Whole Exome Sequencing of 50,726 DiscovEHR Participants (Dewey et al, Science 2016)

In 50K Exomes:
- 92% (n=17,409) of genes with at least 1 heterozygous pLOF
- 7% (n=1,313) of genes with at least 1 homozygous pLOF

Each individual:
- Heterozygous pLOF for ~21 genes
- Homozygous pLOF for ~1 gene

<table>
<thead>
<tr>
<th>Variant type</th>
<th>All variants</th>
<th>Allele frequency ≤ 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single nucleotide variants</td>
<td>4,028,206</td>
<td>3,947,488</td>
</tr>
<tr>
<td>Insertion/deletion variants</td>
<td>224,100</td>
<td>218,785</td>
</tr>
<tr>
<td>Predicted loss of function variants</td>
<td>176,365</td>
<td>175,393</td>
</tr>
<tr>
<td>Nonsynonymous variants</td>
<td>2,025,800</td>
<td>2,002,912</td>
</tr>
<tr>
<td>Total</td>
<td>4,252,306</td>
<td>4,166,273</td>
</tr>
</tbody>
</table>
### Lipid therapy targets harbor LOFs with nominally significant or directionally consistent clinical associations that recapitulate drug effects

**Table: Lipid therapy targets**

<table>
<thead>
<tr>
<th>Target</th>
<th>Agent</th>
<th>Action</th>
<th>Phase</th>
<th>Clinical effect</th>
<th>LOF carriers</th>
<th>LDL-c</th>
<th>HDL-c</th>
<th>Triglycerides</th>
<th>Total cholesterol</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPARA</td>
<td>Fenofibrate</td>
<td>Agonist</td>
<td>Approved</td>
<td>Decreased triglycerides, increased HDL</td>
<td>2</td>
<td>0.8</td>
<td>9 mg/dl</td>
<td>0.2</td>
<td>-28%</td>
</tr>
<tr>
<td>HMGCRC</td>
<td>Atorvastatin, rosuvastatin, pravastatin, simvastatin</td>
<td>Antagonist</td>
<td>Approved</td>
<td>Decreased LDL, total cholesterol, increased HDL</td>
<td>12</td>
<td>0.7</td>
<td>-4 mg/dl</td>
<td>0.3</td>
<td>9%</td>
</tr>
<tr>
<td>NPC1L1</td>
<td>Ezetimibe</td>
<td>Antagonist</td>
<td>Approved</td>
<td>Decreased LDL</td>
<td>121</td>
<td>0.03</td>
<td>-7 mg/dl</td>
<td>0.07</td>
<td>-4%</td>
</tr>
<tr>
<td>APOB</td>
<td>Mipomersen</td>
<td>Antagonist</td>
<td>Approved</td>
<td>Decreased LDL</td>
<td>80</td>
<td>0.0003</td>
<td>-15 mg/dl</td>
<td>0.06</td>
<td>6%</td>
</tr>
<tr>
<td>MTPP</td>
<td>Lomitapide</td>
<td>Antagonist</td>
<td>Approved</td>
<td>Decreased LDL</td>
<td>24</td>
<td>0.9</td>
<td>1 mg/dl</td>
<td>0.4</td>
<td>4%</td>
</tr>
<tr>
<td>HCAR3</td>
<td>Niacin</td>
<td>Agonist</td>
<td>Approved</td>
<td>Increased HDL, decreased triglycerides, LDL</td>
<td>107</td>
<td>0.4</td>
<td>-3 mg/dl</td>
<td>0.4</td>
<td>-2%</td>
</tr>
<tr>
<td>CETP</td>
<td>Anacetrapib, evacetrapib</td>
<td>Antagonist</td>
<td>Phase 3</td>
<td>Increased HDL</td>
<td>37</td>
<td>0.3</td>
<td>-6 mg/dl</td>
<td>2.010^-6</td>
<td>23%</td>
</tr>
<tr>
<td>PCSK9</td>
<td>Alirocumab, evolocumab, bococizumab</td>
<td>Antagonist</td>
<td>Phase 3</td>
<td>Decreased LDL</td>
<td>52</td>
<td>8.10^-9</td>
<td>-25 mg/dl</td>
<td>0.3</td>
<td>3%</td>
</tr>
<tr>
<td>APOC3</td>
<td>APOC3 inhibitors</td>
<td>Antagonist</td>
<td>Phase 2</td>
<td>Decreased triglycerides, increase HDL</td>
<td>226</td>
<td>0.3</td>
<td>-3 mg/dl</td>
<td>1.510^-43</td>
<td>28%</td>
</tr>
<tr>
<td>ACLY</td>
<td>ATP citrate lyase inhibitors</td>
<td>Antagonist</td>
<td>Phase 2</td>
<td>Decreased LDL</td>
<td>13</td>
<td>0.2</td>
<td>-14 mg/dl</td>
<td>1.0</td>
<td>0%</td>
</tr>
<tr>
<td>ANGPTL3</td>
<td>ANGPTL3 inhibitors</td>
<td>Antagonist</td>
<td>Phase 2</td>
<td>Decreased triglycerides, LDL, HDL</td>
<td>150</td>
<td>0.0004</td>
<td>-10 mg/dl</td>
<td>0.0002</td>
<td>-8%</td>
</tr>
</tbody>
</table>

8/11 Lipid therapy targets harbor LOFs with nominally significant or directionally consistent clinical associations that recapitulate drug effects.
Inactivating Variants in ANGPTL4 and Risk of Coronary Artery Disease

Frederick E. Dewey, M.D., Viktoria Gusarova, Ph.D., Colm O'Dushlaine, Ph.D., Omri Gottesman, M.D., Jesus Trejos, M.S., Charleen Hunt, Ph.D., Cristopher V. Van Hout, Ph.D., Lukas Habegger, Ph.D., David Buckler, Ph.D., Ka-Man V. Lai, Ph.D., Joseph B. Leader, Ph.D., Michael F. Murray, M.D., Marylyn D. Ritchie, Ph.D., H. Lester Kirchner, Ph.D., David H. Ledbetter, Ph.D., John Penn, M.S., Alexander Lopez, M.S., Ingrid B. Borecki, Ph.D., John D. Overton, Ph.D., Jeffrey G. Reid, Ph.D., David J. Carey, Ph.D., Andrew J. Murphy, Ph.D., George D. Yancopoulos, M.D., Ph.D., Aris Baras, M.D., Jesper Gromada, Ph.D., D.M.Sc., and Alan R. Shuldiner, M.D.
Loss-of-Function Carriers in a ANGPTL4 Have Favorable Lipid Phenotypes and Are Protected From CAD (Dewey et al, NEJM 2016)
Hypolipidemic Effects of Anti-ANGPTL4 Antibody in Mice and Monkeys (Dewey et al, NEJM 2016)
Hypolipidemic Effects of Anti-ANGPTL4 Antibody in Mice and Monkeys (Dewey et al, NEJM 2016)

AE: Some mice and one monkey developed abdominal lymphadenopathy and chylos ascites
**ANGPTL4 p.E40K Human Homozygotes do not Exhibit Increased Rates of Lymphatic Abdominal Pathology in DiscovEHR**

In chart review of 17 p.E40K homozygotes, 5 had CT abdominal imaging, and 4/5 had explicit mention of normal abdominal lymphatics, 1/5 had no mention of lymphatic abnormalities.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Non-carriers (n=41,777)</th>
<th>E40/K40 heterozygotes (n=1,661)</th>
<th>K40 homozygotes (n=17)</th>
<th>pLOF carriers (n=75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disordrs of lymphoid system</td>
<td>3,831 (9.2)</td>
<td>154 (9.3)</td>
<td>N (%)</td>
<td>P*</td>
</tr>
<tr>
<td>Disorder of lymph node</td>
<td>1,661 (4.0)</td>
<td>70 (4.2)</td>
<td>0 (0.0)</td>
<td>0.9</td>
</tr>
<tr>
<td>Lymphadenitis</td>
<td>295 (7.1)</td>
<td>12 (7.2)</td>
<td>0 (0.0)</td>
<td>0.7</td>
</tr>
<tr>
<td>Mesenteric lymphadenitis</td>
<td>12 (0.03)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0.5</td>
</tr>
<tr>
<td>Granulomatous lymphadenitis</td>
<td>5 (0.01)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0.7</td>
</tr>
<tr>
<td>Ascites</td>
<td>308 (0.7)</td>
<td>11 (0.7)</td>
<td>1 (5.9)</td>
<td>0.8</td>
</tr>
<tr>
<td>Peritonitis</td>
<td>282 (0.7)</td>
<td>17 (1.0)</td>
<td>0 (0.0)</td>
<td>0.1</td>
</tr>
<tr>
<td>Malabsorption</td>
<td>3,291 (7.9)</td>
<td>142 (8.6)</td>
<td>0 (0.0)</td>
<td>0.3</td>
</tr>
<tr>
<td>Abdominal discomfort</td>
<td>15,183 (36.3)</td>
<td>612 (37.0)</td>
<td>4 (35.2)</td>
<td>0.2</td>
</tr>
<tr>
<td>Diarrhea symptom</td>
<td>6,099 (14.6)</td>
<td>222 (13.4)</td>
<td>2 (10.2)</td>
<td>0.2</td>
</tr>
</tbody>
</table>

*Versus sequenced non-carriers
**ANGPTL4 p.E40K and Loss of Function Variants are Associated with Reduced Odds of Type 2 Diabetes: A new indication for ANGPTL4 inhibition?**

- **Take home points:**
  - The p.E40K variant was associated with ~15% reduced odds of diabetes per allele
  - Loss of function variant carriers had 58% reduced odds of diabetes

### Disease Allele Frequency: Cases | Allele Frequency: Controls | Odds Ratio* (95% CI) | P* | Heterozygous loss-of-function variants (n = 75)
| Allele Frequency: Cases | Allele Frequency: Controls | Odds Ratio* (95% CI) | P* |
|------------------------|--------------------------|----------------------|----|--------------------------|--------------------------|----------------------|----|
| Type 2 diabetes        | 1.84 (355 hets, 6 homs)  | 2.06 (1,053 hets, 14 homs) | 0.86 (0.76-0.99) | 0.03 | 0.05 (10 hets) | 0.11 (58 hets) | 0.42 (0.19-0.83) | 0.01 |

Abbreviations: AF, allele frequency; hets, heterozygotes; homs, homozygotes; CAF, cumulative allele frequency; OR, odds ratio

*Adjusted for age, age^2, sex, principal components of ancestry, and BMI.

O'Dushlaine, et al. ,in preparation
Insights From Whole Exome Sequencing in Mendelian Diseases Collaborations (CUMC, CSC & TSK)

Total variants in proband

Filter for exome variants

Filter for rarity

Filter for transmission

Prioritize remaining genes

<table>
<thead>
<tr>
<th>Families/samples</th>
<th>Families/samples analyzed</th>
<th>Families with known causative variants</th>
<th>Families with novel variants in known disease genes</th>
<th>Families with novel disease genes</th>
<th>Families with multiple candidate genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>sequenced</td>
<td>756/5747</td>
<td>23 (15)</td>
<td>92 (32)</td>
<td>126</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Claudia Gonzaga-Jauregui
Gene Discovery in Familial Pediatric Onset Pulmonary Arterial Hypertension: *TBX4* Implicated in Multiple PAH Families

<table>
<thead>
<tr>
<th>Family/Proband ID</th>
<th>VARIANT</th>
<th>INHERITANCE</th>
<th>DELETERIOUS PREDICTION</th>
<th>GHS ALLELE FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1</td>
<td>c.C1054T; p.R352X</td>
<td>de novo</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Family 2</td>
<td>c.664delA; p.T222fs</td>
<td>maternal</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Family 3</td>
<td>c.C577T; p.Q193X</td>
<td>maternal</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Family 4</td>
<td>c.537_546del; p.P180Ifs*45</td>
<td>de novo</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Family 5</td>
<td>c.1070_1070delC; p.S359Lfs*20</td>
<td>de novo</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Family 6</td>
<td>c.C293G; p.P98R</td>
<td>de novo</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Family 7</td>
<td>c.1115dupC; p.Pro372fs</td>
<td>paternal</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Family 8</td>
<td>c.498_500delCTC; p.166_167delIS</td>
<td>paternal</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Singleton 1</td>
<td>c.702+1G&gt;A (splicing)</td>
<td>unknown</td>
<td>damaging</td>
<td>0</td>
</tr>
<tr>
<td>Singleton 2</td>
<td>25.9kb deletion</td>
<td>unknown</td>
<td>damaging</td>
<td>0</td>
</tr>
</tbody>
</table>

- ~71 families and 192 singletons recruited through CUMC; enriched for pediatric onset PAH
- Rare, deleterious variants in *TBX4* identified in 16 different cases (10 families & 6 singletons)
Founder Populations: Stacking the Deck for Discovery of Novel Genes for Aging and Age-Related Phenotypes

- **Principle 1: Genetic Homogeneity:**
  - Gene pool of entire population derives from a small number of founders

- **Principle 2: Drift:**
  - Rare (single copy) founder LOF alleles can increase in frequency
    - Opportunity for novel large-effect gene discovery
    - Opportunities to identify modifier genes

- **Principle 3: Consanguinity and large families:**
  - Further opportunity to identify homozygotes for enriched LOF alleles

- **Principle 4: Homogeneous lifestyle**
  - Fewer confounding influences
  - Geographically localized → Genotype-first call-back studies
Why Study Complex Diseases in the Amish?
Why Study Complex Diseases in the Amish?

- A cultural isolate – traditional dress, no electricity, phones, cars
- Genetically homogeneous closed founder population
Why Study Complex Diseases in the Amish?

- A cultural isolate – traditional dress, no electricity, phones, cars
- Genetically homogeneous closed founder population
  - Complex genetics less complex
  - Enrichment of rare large-effect mutations (founder effect)
- Western/Central European in origin
Why Study Complex Diseases in the Amish?

- A cultural isolate – traditional dress, no electricity, phones, cars
- Genetically homogeneous closed founder population
  - Complex genetics less complex
  - Enrichment of rare large-effect mutations (founder effect)
- Western/Central European in origin
- Very large extended pedigrees (mean sibship size = 7)
  - Extensive genealogical records (Fisher Book, AGD)
  - Geographically localized
- Homogeneous lifestyle (e.g., diet, minimal use of medications)
- Generalizability of findings
Old Order Amish Demography has Increased the Number of Some Minor Allele Homozygotes

• Alleles have drifted to higher frequency in the Amish, R3527Q \textit{APOB}; R19X \textit{APOC3}, L28P \textit{DRK1B}, etc.
• Finite population size results in consanguinity, \~3.5\% of each Amish genome is autozygous
• Example: allele drifts from 1\% to 5\% in the Amish, \~40 fold increase in minor allele homozygotes
The Path to Personalized Medicine

Disease/Trait With Genetic Component

- Diagnostics/Newborn screening
- Early Prevention
- Pharmacogenomics/Nutrigenomics
- Gene Therapy

Identify Gene

- Understand Basic Biological Defect
  - New Drug Targets (Target validation)

rx
The Path to Personalized Medicine

Disease/Trait With Genetic Component

Identify Gene

Diagnostics/Newborn screening

Early Prevention

Pharmacogenomics/Nutrigenomics

Understand Basic Biological Defect

New Drug Targets

(Target validation)

Gene Therapy

Rx
The Path to Personalized Medicine

Disease/Trait With Genetic Component

Identify Gene

Understand Basic Biological Defect

New Drug Targets
(Target validation)

CYP2C19 and clopidogrel response
Shuldiner, JAMA 2009

R3500Q APOB and LDL
Shen, Arch Int Med 2010

R355W SLC30A8, zinc and diabetes
Marathur, Diabetologia 2015

HSL Null and diabetes
Albert, NEJM, 2015

R19X APOC3 and TG
Pollin, Science 2008

GRB10 and diabetes
Rampersaud, Diabetes 2007

SLC13A1 and sulfate biology
Perry, G3, 2016

Sitosterolemia and CVD
Horenstein, Arterioscler Thromb Vasc Biol 2013

KDR, and cancer
Maitland, Clin Cancer Res 2015

STK39 and BP
Wang, PNAS 2009

PEARI, aspirin response and CAD events
Lewis, Circ Genet 2013

Early Prevention

Diagnostics/Newborn screening

Pharmacogenomics/
Nutrigenomics

Gen

Pearson
Some Cool Findings in the Amish: Many drifted alleles that inform biology and precision medicine

- ~1 in 8 Amish carry R3527Q APOB, a cause of autosomal dominant familial hypercholesterolemia (Shen et al. Arch Int Med 2010)

- ~1 in 25 Amish carry R19X APOC3 and have low triglycerides levels and are protected from CAD (Pollin et al. Science 2008)

- ~1 in 40 Amish carry T224M KCNQ1, which is highly associated with longer QT interval, a risk factor for syncope and sudden death

Novel genes for monogenic diseases that inform biology and therapeutic development (Strauss, Genetics in Medicine 2017)

- ~1 in 20 Amish carry a 19 bp frame-shift mutation in LIPE that increases risk for T2D by 2-fold and causes partial lipodystrophy in homozygotes (Albert et al. NEJM 2014)
The Path to Personalized Medicine

Disease/Trait With Genetic Component

Identify Gene

Diagnostics/Newborn screening

Early Prevention

Pharmacogenomics/
Nutrigenomics

Understand Basic Biological Defect

New Drug Targets
(Target validation)

Gene Therapy

HSL Null and diabetes
Albert, NEJM, 2015

Rare/enriched
Common
Null Mutation in Hormone-Sensitive Lipase Gene and Risk of Type 2 Diabetes

Jessica S. Albert, Ph.D., Laura M. Yerges-Armstrong, Ph.D., Richard B. Horenstein, M.D., Toni I. Pollin, Ph.D., Urmila T. Sreenivasan, M.S., Sumbul Chai, M.S., William S. Blaner, Ph.D., Soren Snitker, M.D., Ph.D., Jeffrey R. O’Connell, Ph.D., Da-Wei Gong, Ph.D., Richard J. Breyer III, M.D., Alice S. Ryan, Ph.D., John C. McLenithan, Ph.D., Alan R. Shuldiner, M.D., Carole Ształryd, Ph.D., and Coleen M. Damcott, Ph.D.

June 2014
19 base pair deletion (p.V1068GfsX19) identified in exon 10 of LIPE

HSL

FABP4

Albert et al, NEJM, 2014
19 base pair deletion (p.V1068GfsX19) identified in exon 10 of LIPE

HSL

FABP4

Albert et al, NEJM, 2014
19 base pair deletion (p.V1068GfsX19) identified in exon 10 of LIPE

Albert et al, NEJM, 2014
19 base pair deletion (p.V1068GfsX19) identified in exon 10 of LIPE

HSL
ACALDPMLDDSVMLARLLRNGLQPVTLVRVEDLPHGFLTLAALCRETRQAAELCVERIR
LVLTNPAGAGPSGETGAGVDGGCGGRH

Albert et al, NEJM, 2014
19 base pair deletion (p.V1068GfsX19) identified in exon 10 of LIPE

Albert et al, NEJM, 2014
19 base pair deletion (p.V1068GfsX19) identified in exon 10 of **LIPE**

Amish carrier frequency = 0.013 (4 homozygotes)
GHS carrier frequency 0.00073 (No homozygotes)

Albert et al, NEJM, 2014
HSL knockout mouse

**General**
- Non-obese
- Male infertility

**Adipose tissue**
- Decreased lipolysis
- Increased DAG
- Decreased cholesterol esterase activity

**Systemic**
- Normoglycemic
- Normoinsulinemic
- Decreased TAG
- Increased HDL
Mating of heterozygous HSL Homo Sapiens

Age (2009):
Number of Children:
Diagnosed diabetes:

78
10
Y

Died at age 56 of leukemia

<table>
<thead>
<tr>
<th>Age (2009)</th>
<th>Number of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>10</td>
</tr>
<tr>
<td>53</td>
<td>2</td>
</tr>
<tr>
<td>52</td>
<td>5</td>
</tr>
<tr>
<td>50</td>
<td>6</td>
</tr>
<tr>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>47</td>
<td>7</td>
</tr>
<tr>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>43</td>
<td>6</td>
</tr>
<tr>
<td>42</td>
<td>10</td>
</tr>
<tr>
<td>39</td>
<td>2</td>
</tr>
</tbody>
</table>

*II*  *ID*  *DD*  *ND*  *not yet determined*

Albert et al, NEJM, 2014
Mating of heterozygous HSL Homo Sapiens

Age (2009): 78
Number of Children: 10
Diagnosed diabetes: Y
Died at age 56 of leukemia

Diagnosed diabetes:
II II
ID ID
DD DD
ND ND*
*not yet determined

Albert et al, NEJM, 2014
p.V1068GfsX19 HSL is a Null Mutation

Triglycerides

- **p = 0.001**

HDL-cholesterol

- **p = 1.6x10^-7**

OR for DM = 2.0; p=0.02

Albert et al, NEJM, 2014
<table>
<thead>
<tr>
<th>Subject number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current age</strong></td>
<td>46</td>
<td>49</td>
<td>56</td>
<td>44</td>
<td>53</td>
<td>55</td>
<td>42</td>
<td>51</td>
<td>47</td>
<td>52</td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td>27</td>
<td>28</td>
<td>27</td>
<td>31</td>
<td>23</td>
<td>23</td>
<td>29</td>
<td>22</td>
<td>24</td>
<td>29</td>
</tr>
<tr>
<td><strong>Body Fat (%)</strong></td>
<td>31.3</td>
<td>33.2</td>
<td>33.6</td>
<td>40.8</td>
<td>36.8</td>
<td>36.3</td>
<td>21.1</td>
<td>13.4</td>
<td>16.2</td>
<td>30.6</td>
</tr>
<tr>
<td><strong>VAT (cm²)</strong></td>
<td>82.7</td>
<td>-</td>
<td>89</td>
<td>74.6</td>
<td>76.6</td>
<td>77.3</td>
<td>64.8</td>
<td>-</td>
<td>43.5</td>
<td>53.4</td>
</tr>
<tr>
<td><strong>Abdominal SAT (cm²)</strong></td>
<td>276</td>
<td>-</td>
<td>318</td>
<td>316</td>
<td>205</td>
<td>212</td>
<td>193</td>
<td>-</td>
<td>79.9</td>
<td>222</td>
</tr>
<tr>
<td><strong>Calf SAT (cm³)</strong></td>
<td>14.7</td>
<td>7.7</td>
<td>33.6</td>
<td>22.6</td>
<td>21.5</td>
<td>8.6</td>
<td>-</td>
<td>13.8</td>
<td>23.4</td>
<td></td>
</tr>
<tr>
<td><strong>Abdominal SAT:Calf SAT</strong></td>
<td>18.8</td>
<td>-</td>
<td>41.4</td>
<td>9.4</td>
<td>9.1</td>
<td>9.9</td>
<td>22.3</td>
<td>-</td>
<td>5.8</td>
<td>9.5</td>
</tr>
<tr>
<td><strong>Liver Fat (%)</strong></td>
<td>9.4</td>
<td>25.3</td>
<td>3.7</td>
<td>2.8</td>
<td>2.8</td>
<td>3.5</td>
<td>-</td>
<td>2.4</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td><strong>Adiponectin (µg/ml)</strong></td>
<td>4.3</td>
<td>3.6</td>
<td>10.9</td>
<td>28</td>
<td>17.5</td>
<td>4.7</td>
<td>22.4</td>
<td>10.5</td>
<td>17.9</td>
<td></td>
</tr>
<tr>
<td><strong>Leptin (ng/ml)</strong></td>
<td>4.6</td>
<td>4.8</td>
<td>9.2</td>
<td>18.9</td>
<td>6.3</td>
<td>9</td>
<td>1.8</td>
<td>0.6</td>
<td>0.8</td>
<td>4.6</td>
</tr>
<tr>
<td><strong>Fasting triglycerides (mg/dl)</strong></td>
<td>148</td>
<td>152</td>
<td>256</td>
<td>81</td>
<td>140</td>
<td>109</td>
<td>193</td>
<td>53</td>
<td>73</td>
<td>146</td>
</tr>
<tr>
<td><strong>HDL Cholesterol (mg/dl)</strong></td>
<td>52</td>
<td>53</td>
<td>40</td>
<td>53</td>
<td>69</td>
<td>70</td>
<td>46</td>
<td>65</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td><strong>Fasting insulin (µU/ml)</strong></td>
<td>7.9</td>
<td>20.3</td>
<td>15</td>
<td>7.6</td>
<td>2.6</td>
<td>9.1</td>
<td>5.6</td>
<td>3.3</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td><strong>Fasting glucose (mg/dl)</strong></td>
<td>98</td>
<td>125</td>
<td>256</td>
<td>102</td>
<td>84</td>
<td>86</td>
<td>202</td>
<td>90</td>
<td>93</td>
<td>85</td>
</tr>
<tr>
<td><strong>HOMA-IR</strong></td>
<td>1.9</td>
<td>6.3</td>
<td>9.5</td>
<td>1.9</td>
<td>0.5</td>
<td>1.9</td>
<td>2.8</td>
<td>0.7</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td><strong>2-hr OGTT glucose (mg/dl)</strong></td>
<td>204</td>
<td>-</td>
<td>-</td>
<td>140</td>
<td>119</td>
<td>101</td>
<td>470</td>
<td>112</td>
<td>150</td>
<td>107</td>
</tr>
<tr>
<td><strong>Diagnosed diabetes (Age)</strong></td>
<td>Y (46)</td>
<td>Y (46)</td>
<td>Y (48)</td>
<td>IGT</td>
<td>N</td>
<td>N</td>
<td>Y (40)</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td><strong>Number of offspring:</strong></td>
<td>0</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

Albert et al, NEJM, 2014
Increased liver and visceral fat; decreased LE fat in DD homozygotes

**MRI**

- **Liver**
  - (25.3% vs 2.8% fat)

- **L4-L5/visceral fat**
  - (89.0 cm² vs. 77.3 cm²)

- **Calf**
  - 7.7 cm² vs 21.5 cm²

**CT**

- **DD**
  - Age 56
  - BMI 27 kg/m²

- **II**
  - Age 55
  - BMI 23 kg/m²

- **Spleen:liver ratio**
  - $p=2.4 \times 10^{-4}$

Albert et al, NEJM, 2014
Unexpected effect of HSL deficiency on adipocyte function

WAT cytokine release

Infiltration by macrophages

Expression of PPARγ-regulated genes

↓ Lipolysis

↓ Adipocyte insulin sensitivity

↓ Fat cell size

Albert et al, NEJM, 2014
Unexpected effect of HSL deficiency on adipocyte function

**Implications:**

*In addition to its metabolic function, HSL may have an endocrine function to generate endogenous ligands for PPAR\(\gamma\)*

Opportunity to identify endogenous ligand(s) for PPAR\(\gamma\) as next diabetes therapeutic

**HSL KO (or deficient) humans with diabetes should respond well to PPAR\(\gamma\) agonists (thiazolidinediones)**

**HSL activators as therapeutic to decrease TG and increase insulin sensitivity and glucose tolerance in patients with diabetes**
Building the World’s Largest Founder Population Collection for Discovery of Novel Disease-associated Genetic Variants:

**Discovery Research Investigating Founder Population Traits (DRIFT) Goals**

- Catalog population-specific allelic architecture
- Understand the biological and functional consequences of specific mutations identified
  - Genotype – Phenotype associations (especially of rare LoF/GoF mutations enriched in a given population)
  - Replicate/extension in larger general population
  - “Genotype-first” call back studies
- Share and establish best practice approaches to relieve disease burden in these populations
The Future of Precision Medicine is Here Today

The GHS Genome-First Return of Results Program
Informed consent specifies intent to return results that are medically actionable after CLIA confirmation

Geisinger will NOT return results that are NOT medically actionable

Geisinger experts will decide what to return
  - Geisinger “76” (56 ACMG + 20) genes causative of 27 diseases that are medically actionable
    - e.g., hereditary breast and ovarian cancer; Lynch syndrome; familial hypercholesterolemia; hypertrophic cardiomyopathy

Initial analyses indicate that ~3.5% of study participants will test positive for an actionable variant
Informed consent specifies intent to return results that are medically actionable after CLIA confirmation

Geisinger will NOT return results that are NOT medically actionable

Geisinger experts will decide what to return
- Geisinger “76” (56 ACMG + 20) genes causative of 27 diseases that are medically actionable
  - e.g., hereditary breast and ovarian cancer; Lynch syndrome; familial hypercholesterolemia; hypertrophic cardiomyopathy
## Geisinger 76: 1 in 25 Patients Have an Actionable Result

<table>
<thead>
<tr>
<th>GENOMIC CONDITION</th>
<th>Number or patients diagnosed</th>
<th>CLINICAL RISK</th>
<th>DISEASE-ALTERING INTERVENTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familial Hypercholesterolemia (FH)</td>
<td>1 in 250</td>
<td>Early-onset Coronary Artery Disease and Stroke</td>
<td>Targeted screening and aggressive medical management</td>
</tr>
<tr>
<td>Hereditary Breast and Ovarian Cancer Syndrome</td>
<td>1 in 400</td>
<td>Early-onset Breast, Ovarian, and Prostate Cancers</td>
<td>Targeted screening with prophylactic medical and surgical intervention</td>
</tr>
<tr>
<td>Lynch Syndrome</td>
<td>1 in 440</td>
<td>Early-onset Colon and Uterine Cancers</td>
<td>Targeted screening and management of pre-cancerous changes</td>
</tr>
<tr>
<td>TOTAL</td>
<td>&gt; 1 in 100</td>
<td>Multiple Cancers and Cardiovascular Diseases</td>
<td>Life-saving screening and intervention before development of disease</td>
</tr>
</tbody>
</table>

Other conditions: cardiomyopathy, long QT syndrome, malignant hyperthermia, arrhythmogenic right ventricular cardiomyopathy, MEN2, tuberous sclerosis, hereditary pheochromocytomas and paragangliomas
Barbara Barnes’ MyCode Story

- 57 Year old grandmother bringing up three grandchildren ages 3, 5, and 14
- Found to have a pathogenic BRCA1 mutation
  - “Okay, so what do we do next? I have 15 more years to go until they’re raised.”
- Genetic counseling and workup
  - Negative mammogram
  - Elected to have preventive bilateral salpingo-oophorectomy
  - Stage 1 cancer found in one fallopian tube
  - Completing chemotherapy with expected excellent outcome
  - Daughter tested for BRCA1
Kim Mummert’s MyCode Story

• 66 year old GHS employee
• Tested positive for a Lynch Syndrome mutation
• Obtained genetic counseling
  – More frequent colonoscopy
  – Two children in their 20’s will be screened
RESEARCH ARTICLE SUMMARY

HUMAN GENETICS

Genetic identification of familial hypercholesterolemia within a single U.S. health care system

Prevalence and Clinical Impact of FH Variants in DiscovEHR

A. Distribution of 229 heterozygous carriers by FH gene

- LDLR (29 variants; 19 pLOF)
- PCSK9 (4 variants)
- APOB (2 variants)

B. Population characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>FH variant positive/total</th>
<th>Estimated prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>All DiscovEHR participants</td>
<td>229/50,726</td>
<td>1:222</td>
</tr>
<tr>
<td>Participants recruited from cardiac catheterization lab</td>
<td>57/6,747</td>
<td>1:118</td>
</tr>
<tr>
<td>Participants recruited from other sites</td>
<td>172/43,979</td>
<td>1:256</td>
</tr>
</tbody>
</table>

C. Participants with severe hypercholesterolemia (LDL-C > 190 mg/dl)
N = 4,435 of 42,696 individuals with LDL-C data available (10.4%)

FH variant positive: 112 of 4,435 (1:40)

D. Currently on statin

- FH variant negative: 38%
- FH variant positive: 58%

Statin-treated with LDL-C < 100 mg/dl

- FH variant negative: 46%
- FH variant positive: 77%
Why GenomeFirst Is Important: Most patients with FH are not diagnosed and treated inadequately

- Only 35 (15.7%) of the 229 FH variant carriers had EHR evidence of a “Pure Hypercholesterolemia” diagnosis or at least one encounter at Lipid Clinic.
- Criteria supporting a clinical diagnosis of FH were found using EHR data in only 55% of variant carriers.
- Active statin use was identified in 58% and high-intensity statin use in 37% of carriers.
- Only 46% of statin-treated carriers had a LDL cholesterol level below 100 mg/dl.
- Genomic screening can prompt the diagnosis of FH patients, the majority of whom are receiving inadequate lipid-lowering therapy.

**Abul-Husn et al, Science 2016**
Summary and Conclusions

- The future of drug discovery and precision medicine will be fueled by human genomic discovery.
- Genetic “experiments of nature” can inform therapeutic target discovery and provide insight into mechanism.
- Return of medically actionable genetic results will require significant health care system resources to realize downstream health and economic benefits.
- Partnerships between industry, academia and health care systems can accelerate genomic discovery and implementation of precision medicine.
Thank you to the thousands of patients, volunteers, collaborators, investigators, and scientists who make this work possible

**RGC Leadership**
- Aris Baras
- Rick Dewey
- Aris Economides
- John Overton
- Jeff Reid
- Alan Shuldiner

**Regeneron Team (all but especially)**
- Noura Abul-Husn
- Xiaodong Bai
- Suganthi Balasubramanian
- Shannon Bruse
- Rostislav Chernomorsky
- Jonathan Chung
- Jan Freudenberg
- Jesper Gromada
- Claudia Gonzaga-Jauregui
- Nehal Gosalia
- Viktoria Gusalova
- Lukas Habegger
- Julie Horowitz
- Marcus Jones
- Alejandra King
- Alexander Lopez
- Evan Maxwell
- Lyndon Mitnaul
- Semanti Mukherjee
- Colm O’Dushlaine
- Chiatogu Onyewu
- Jonathan Packer
- John Penn
- Christopher Sprangel
- Jeffrey Staples
- Tanya Teslovice
- Ricardo Ulloa
- Cristopher Van Hout

**Geisinger Health System**
- David J. Carey
- David H. Ledbetter
- Michael Murray
- Marylyn D. Ritchie
- Marc Williams
- W. Andrew Faucett
- Sarah A. Pendergrass
- Chris Still
- Tooraj Mirshahi
- H. Lester Kirchner
- Joseph B. Leader
- G. Craig Wood
- Lance J. Adams
- Peter Benotti
- Vishal C. Mehra
- Raghu Metpally
- Uyenlinh L. Mirshahi
- Cassandra M Hartle
- F. Daniel Davis
- Daniel R. Lavage
- Dustin Hartzel
- John Snyder
- Neil Manus
- Ryan D. Colonie

**Columbia University**
- Wendy Chung
- Rudy Leibel
Acknowledgements – University of Maryland

Jess Albert
Coleen Damcott
Julie Ducharme
Adam Fisch
Susan Fried
Mao Fu
Amish Gandhi
Da-Wei Gong
Kristen Hairston
Amanda Holmes
Nicole Hoppman
Richard Horenstein
Hong Hu
Yasmin Khan
Jie Liu

Patrick McArdle
Daniel McBride
John McLenithan
Cary McMahan
Braxton Mitchell
Karen Norton
Jeffrey O'Connell
Sandra Ott
Christina Perry
Toni Pollin
Evadnie Rampersaud
Laurie Reinhart
Kathy Ryan
Mona Sabra
Larry Sauder
Jack Shelton

Haiqing Shen
Julia Shi
Kristi Silver
John Sorkin
Soren Snitker
Nanette Steinle
Elizabeth Streeten
Carole Szalryd
Keith Tanner
Magnda Tolea
Robert Vogel
Matthew Weir
Rongze Yang
Diaozhan Yu
Li Zhang
Joe Zhao