The Regeneron-Geisinger Collaboration as a Model

George D. Yancopoulos, MD, PhD
President, Regeneron Laboratories
Chief Scientific Officer & Founding Scientist, Regeneron
Precision Medicine Initiative
The dream is to go from genes to a new way of promoting health

**Prognosis:**
Using your genome to better predict health risks, earlier and more precisely

**Better diagnosis & treatment selection:**
Using your genome to better discriminate between related disease states and appropriately deliver more precise treatments – the right treatments for the right patients!

**Discover New Targets for Drug Development:**
Using your genome find genes that cause or protect against disease, and then develop therapeutics based on these new “genetically-defined” drug targets
Human Genetics to Drugs: CAPS treatment with IL1 blockers was one of first successful examples with biologics*, PCSK9 perhaps the most highlighted

*Excludes replacement therapies for deficiency diseases (e.g. Genzyme, Biomarin, etc)

<table>
<thead>
<tr>
<th>Disease/Indication</th>
<th>Gene/Target</th>
<th>Drug/Candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryopyrin Associated Periodic Syndromes (CAPS)</td>
<td>CIAS1/ NLRP3</td>
<td>Arcalyst: IL1 Trap (Regeneron) Canakinumab: IL1 Ab (Novartis)</td>
</tr>
<tr>
<td>Hypercholesterolemia &amp; Cardiovascular Disease</td>
<td>PCSK9</td>
<td>Alirocumab: PCSK9 Ab (Regeneron) Evolocumab: PCSK9 Ab (Amgen)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>PPARG</td>
<td>Thiazolidinediones, PPARγ agonist</td>
</tr>
<tr>
<td>Pain</td>
<td>SCN9A (NaV1.7)</td>
<td>In development (Pfizer, Regeneron, Xenon, etc)</td>
</tr>
<tr>
<td>Cystic fibrosis</td>
<td>CFTR</td>
<td>Kalydeco (Vertex)</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>CCR5</td>
<td>Selzentry/maraviro (Pfizer)</td>
</tr>
<tr>
<td>Benign prostatic hyperplasia, male pattern baldness</td>
<td>SRD5A2 (5α reductase)</td>
<td>Finasteride (Merck)</td>
</tr>
</tbody>
</table>
Anti-PCKS9 Antibodies for lowering LDL-C
From Genetics to Drug: Following up the genetics, Regeneron leads therapeutics race


2005/6  PCSK9 LOF mutations in human populations associated with low cholesterol & decreased CV risk (J. Cohen & H. Hobbs, UTSW)

2006/7  Mouse modeling confirms causation, provides test system for Abs

2010  Regeneron first to treat human patients with PCSK9 blocking Ab

2012  Regeneron first to publish human data w PCSK9 Ab (Alirocumab) in NEJM

2013  Regeneron reports first Phase 3 data with PCSK9 Ab (ODYSSEY MONO)

2015  Regeneron has earliest potential FDA Approval, July 2015
Regeneron-Geisinger Foundational Collaboration to Sequence >250,000 People in the Geisinger Health System

“Scientifically and medically, it’s pretty exciting,” said Dr. Leslie G. Biesecker, chief of the genetic disease research branch at the government’s National Human Genome Research Institute, who is familiar with the project. “As far as I’m aware, it’s the largest clinical sequencing undertaking in this country so far by a long shot.” He added that the move of sequencing into general health care “is going to change medicine.”

- First major collaboration of its kind between a leading biotechnology company and a leading health care provider/payer
- Geisinger: >2.5 million patient health system with “cradle to grave” records; amongst earliest adopters of EHRs, and leaders in clinical informatics
- Goals:
  - First, build the world’s most comprehensive genotype-phenotype resource combining de-identified genomic and clinical data from >250,000 people to aid drug development and genomic medicine
  - Then, turn the genomic data into medically actionable information by translating genetic discoveries into improved patient outcomes and therapeutics
Regeneron’s Innovative Technologies & Automation Enable Ultra High-Throughput Sequencing & Analysis

- **Automated Biobank (1.4M Samples)**
- **Library Prep Automation (>200,000 Samples/Yr)**
- **Illumina Fleet (>80,000 Exomes/Yr)**
- **Cloud Based Informatics & Analysis**

### Key Technologies and Capabilities

- Automated biobank with 1.4M+ sample capacities
- Custom fully-automated exome and targeted sequencing sample preparation workflows
- Exome sequencing capabilities of >80,000 individuals per year and targeted sequencing projects (several hundred gene panel) of >100,000 individuals completed in a few months
- First genome center “in the cloud” with fully automated analysis pipelines
Large Population-Based Study with Geisinger: Integrating Genetics and Longitudinal Health Records

- Greater than 2.5 million people in Geisinger’s catchment area
- ~18 outpatient visits and 7 years of longitudinal records per patient on average and >1.6M outpatient visits at Geisinger per year

**Engaging Geisinger community as Volunteers & Participants in this effort:**

- >80,000 consented to date (>90% consent rate, broad consent for research & data sharing) → already >50,000 samples collected for genomic analysis → **40,000 already sequenced**
- Breakdown of consented population includes 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
- **Key differentiator:** Geisinger is the world’s largest longitudinal “live” population in which iterative call back phenotyping and sample collection are operationalized
In-Depth, Longitudinal Health Records Provide a Rich Resource for Genetic Discoveries

GHS Clinical Encounters by Year

Repeat Fasting Lipid Panels in GHS EHR

Greater than 93% of patients with lipid panels have more than one measurement

Most Prevalent Labs in GHS EHR

Most Prevalent Office Visit Dx in GHS EHR
In Depth Review of De-identified Longitudinal Health Records

- **Triglycerides**
- **LDL Cholesterol**
- **Glucose**
- **BMI**

Gastric Bypass - 1/1/2008
deCODE Genetics has pioneered the field in this area

LOF’s of great interest (Human “Knockouts”): comparing deCODE’s LOFs to those in interim analysis of the first 31,000 individuals from Regeneron-Geisinger Collaboration (RGC)

- RGC has identified >25x more predicted LOF variants (in first 31,000 individuals) than deCODE reported (from 2,636 Icelanders)
  - RGC has identified ~51% of unique LOFs observed by deCODE
  - deCODE has identified only 2% of unique LOFs that were observed by RGC

- In first 31,000 individuals, RGC finds ~90% of genes are affected by at least one LOF
  - deCODE finds only ~26% of genes affected

- RGC finds at least one homozygous LOF carrier for 2,610 genes vs 1,171 for deCODE

- Signal to noise
  - RGC: 188,614 unique LOFs out of ~6.1M exonic variants from WES of 31,508 individuals
  - deCODE: 7,399 unique LOFs and ~470K exonic variants out of ~20M WGS variants in 2,636 people
RGC-GHS Have Identified ~25x More LOFs Than Recently Reported by deCODE Genetics

**Number of Unique LOF Variants** (drawn to scale)

<table>
<thead>
<tr>
<th></th>
<th>RGC</th>
<th>deCODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>188,614</td>
<td>7,399</td>
</tr>
<tr>
<td>RGC only</td>
<td>184,829</td>
<td>-</td>
</tr>
<tr>
<td>deCODE only</td>
<td>-</td>
<td>3,614</td>
</tr>
<tr>
<td>RGC &amp; deCODE</td>
<td>3,785</td>
<td></td>
</tr>
</tbody>
</table>

Potential: Growing, Saturated
### RGC Interim analysis of >31,000 Individuals Nominates Multiple New Lipid Genes and Functional Variants

<table>
<thead>
<tr>
<th>Trait</th>
<th>Single marker at P &lt; 5x10^{-8}</th>
<th>LOF only at P &lt; 1x10^{-7}</th>
<th>LOF+rare NS at P &lt; 1x10^{-7}</th>
<th>LOF+strict deleterious NS at P &lt; 1x10^{-7}</th>
<th>LOF+rare deleterious NS at P &lt; 1x10^{-7}</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDL-c</td>
<td>CETP, APOC3, APOC2, LPL, CD300LG, APOA4, LIPG, LIPC, ANGPTL4, BUD13, APOB, HNF4A, FADS1/FADS2, APOC2, ABCA1, LCAT, 8 novel</td>
<td>APOC3, ABCA1</td>
<td>APOC3, 3 novel</td>
<td>APOC3, 2 novel</td>
<td>APOC3, CETP, 4 novel</td>
</tr>
<tr>
<td>LDL-c</td>
<td>BCAM/TOMM40/APOE, PCSK9, LDLR, APOB, SORT1/CELSR2, FADS1/FADS2, 4 novel</td>
<td>2 novel</td>
<td>LDLR</td>
<td>LDLR, APOE</td>
<td>LDLR</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>GCKR, APOA5/APOA4/APOC3, LPL, FADS1, MLXILPL, ANGPTL3/DOCK7, CD300LG, APOB, ANGPTL4, TM6SF2, 5 novel</td>
<td>APOC3</td>
<td>APOC3</td>
<td>APOC3, APOE</td>
<td>APOC3</td>
</tr>
</tbody>
</table>

Colm O'Dushlaine and Cris Van Hout
Families to Populations at the RGC: An Illustrative Example in Severe Obesity

Findings from Family Studies
- Many novel candidate Mendelian disease genes identified through family studies in severe, early-onset obesity

Expanded in Population Studies
- Genes identified through family studies are queried across the Geisinger population to further our understanding and establish associations with BMI/obesity in the general population
- Example: Carriers of loss-of-function variants in a novel candidate gene, first identified through family studies in severe obesity, have higher BMI (~3 BMI units, 15-20lbs) in the GHS population

Colm O'Dushlaine and Claudia Gonzaga-Jauregui
Precision Medicine Initiative
The dream is to go from genes to a new way of promoting health

Prognosis:
Using your genome to better predict health risks, earlier and more precisely

Better diagnosis & treatment selection:
Using your genome to better discriminate between related disease states and appropriately deliver more precise treatments – the right treatments for the right patients!

Discover New Targets for Drug Development:
Using your genome find genes that cause or protect against disease, and then develop therapeutics based on these new “genetically-defined” drug targets