Future of Genomics
2015: Precision Medicine Initiative

“I want the country that eliminated polio and mapped the human genome to lead a new era of medicine — one that delivers the right treatment at the right time. In some patients with cystic fibrosis, this approach has reversed a disease once thought unstoppable. Tonight, I'm launching a new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes — and to give all of us access to the personalized information we need to keep ourselves and our families healthier.” President Barak Obama, State of the Union, 2015
Precision Medicine Initiative

1 million participant research

The Precision Medicine Initiative®

WHAT IS IT?

**Precision medicine** is an emerging approach for disease prevention and treatment that takes into account people's individual variations in genes, environment, and lifestyle.

The Precision Medicine Initiative® will generate the scientific evidence needed to move the concept of precision medicine into clinical practice.

Precision Medicine Initiative

• Develop ways to measure risk for a range of diseases based on environmental exposures, genetic factors and interactions between the two;
• Identify the causes of individual differences in response to commonly used drugs (commonly referred to as pharmacogenomics);
• Discover biological markers that signal increased or decreased risk of developing common diseases;
• Use mobile health (mHealth) technologies to correlate activity, physiological measures and environmental exposures with health outcomes;
• Develop new disease classifications and relationships;
• Empower study participants with data and information to improve their own health; and
• Create a platform to enable trials of targeted therapies.
Cost of Sequencing

Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcosts. Accessed 3/25/16.

Complete Genome Information

Song & Snyder, Molecular Systems Biology 9:640 (2013)
GWAS

SNP-trait associations with p-value \( \leq 5.0 \times 10^{-8} \), in GWAS Catalog

[http://www.ebi.ac.uk/gwas/diagram (4/2016)]
GWAS

Digestive system disease 422
Cardiovascular disease 170
Metabolic disease 142
Immune system disease 454
Nervous system disease 423
Liver enzyme measurement 65
Lipid or lipoprotein measurement 303
Inflammatory marker measurement 35

Hematological measurement 302
Body weights and measures 388
Cardiovascular measurement 217
Other measurement 1486
Response to drug 113
Biological process 111
Cancer 402
Other disease 315
Other trait 208

http://www.ebi.ac.uk/gwas/diagram (4/2016)
Genetic Screening

- **When?**
  - Pre-conception
  - Pre-natal testing
  - Newborns
  - Predictive Diagnostic screening (before symptoms?)
  - Predisposition Screening

- **What for?**
  - Disease risk
  - Carriers
  - Diagnosis
  - Treatment
  - Pharmacogenomics

- **Patient rights to know**
  - Genetic Counseling
  - Family Considerations
  - Confusion
  - Changing Information
  - What is actionable?
  - Who owns the data?

- **Protection of Subjects**
  - Discrimination
  - Genetic Information Non-discrimination Act (GINA, 2009)
  - Personal Information
  - Who owns the data?
Pharmacogenomics

- Pharmacogene = a gene involved in the response to a drug
- Pharmacogenetics = the study of genetic influence on drug response, typically one or only a few genes involved
- Pharmacogenomics = the study of how genomic variation influences drug response, looking at variation across the genome

https://www.pharmgkb.org/page/overview
PharmGKB offers different information

- Variant Annotations (Research-level annotations of individual publications describing the relationship between genetic variants and drugs; these are created on a paper-by-paper basis)
- Drug-Centered Pathway
- Very Important Pharmacogene Summaries
- Clinical Annotations (Genotype-based pharmacogenomic relationships summarizing all variant annotations regarding the same genetic variant-drug association)
- Pharmacogenomics-Based Drug-Dosing Guidelines
- Drug Labels with Pharmacogenomic Information

https://www.pharmgkb.org/page/overview
Pharmacogenomics

- pharmacokinetic (PK) pathways
  - what the body does to the drug
    - absorption
    - distribution
    - metabolism
    - elimination
- pharmacodynamic (PD) pathways
  - what the drug does to the body

https://www.pharmgkb.org/page/overview
Pharmacogenomics

Warfarin Pharmacodynamics

- Widely used anticoagulant drug
- Highly effective at antagonising the vitamin K dependent clotting pathway
- Used for a wide range of diseases and conditions
  - Atrial fibrillation
  - Heart valve replacement.

https://www.pharmgkb.org/page/overview
Pharmacogenomics

Warfarin Pharmacodynamics

- Narrow therapeutic window and
- Wide inter-individual variability
- Under-anticoagulation can result in thrombosis
- Over-anticoagulation can result in dangerous bleeding episodes.
- Dosing
  - determined empirically
  - often based on age
  - underlying condition
  - genetics

https://www.pharmgkb.org/page/overview
Pharmacogenomics

Warfarin Pathway, Pharmacodynamics

Entities in the Pathway

Genes (15)

BGLAP, CALU, CYP4F2, EPHX1, F10, F2, F7, F9, GAS6, GGCX, MGP, PROC, PROS1, PROZ, VKORC1

Relationships in the Pathway

<table>
<thead>
<tr>
<th>Arrow From</th>
<th>Arrow To</th>
<th>Controllers</th>
<th>PMID</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGLAP, MGP</td>
<td>BGLAP, MGP</td>
<td>GGCX</td>
<td>16270630, 16493479</td>
</tr>
<tr>
<td>GAS6</td>
<td>GAS6</td>
<td>GGCX</td>
<td>16270630, 16493479</td>
</tr>
<tr>
<td>GGCX</td>
<td>GGCX</td>
<td>CALU</td>
<td>15075329, 16493479</td>
</tr>
<tr>
<td>NADH</td>
<td>NAD+</td>
<td>EPHX1, VKORC1</td>
<td></td>
</tr>
<tr>
<td>Vitamin K (epoxidized)</td>
<td>Vitamin K (reduced)</td>
<td>EPHX1, VKORC1</td>
<td>14765194, 15358623, 15900282, 16270630, 16493479</td>
</tr>
<tr>
<td>Vitamin K (reduced)</td>
<td>Hydroxy-Vitamin K1</td>
<td>CYP4F2</td>
<td>17341693, 18250228, 19297519</td>
</tr>
<tr>
<td>Vitamin K (reduced)</td>
<td>Vitamin K (epoxidized)</td>
<td>GGCX</td>
<td>15900282, 16270630, 16493479</td>
</tr>
<tr>
<td>VKORC1</td>
<td>VKORC1</td>
<td>warfarin</td>
<td>15358623, 16270630, 16493479</td>
</tr>
<tr>
<td>F10, F2, F7, F9, PROC, PROS1, PROZ</td>
<td>F10, F2, F7, F9, PROC, PROS1, PROZ</td>
<td>GGCX</td>
<td>16270630, 16493479</td>
</tr>
</tbody>
</table>

Download data in TSV format. Other formats are available on the Downloads/LinkOuts tab.

https://www.pharmgkb.org/page/overview
Pharmacogenomics


https://www.pharmgkb.org/page/overview
Pharmacogenomics

Pathway
Warfarin Pathway, Pharmacodynamics

Related Publications

Reference


https://www.pharmgkb.org/page/overview
Direct to Consumer
Personalized Genomics

We bring the world of genetics to you.

- Receive 60+ personalized genetic reports
- Understand what your DNA says about your health, traits and ancestry
- Access interactive tools to share, compare and discover more with friends and family

order now $199

23andme.com
Direct to Consumer Personalized Genomics

Genetic reports. Backed by science.

Our rigorous quality standards:

- Our Carrier Status Tests meet FDA criteria for being scientifically and clinically valid
- All saliva samples are processed in CLIA-certified and CAP-accredited labs
- Genotyping is a well-established and reliable platform for analyzing DNA
- Our scientists and medical experts use a rigorous process to develop the reports
- Your personalized reports are based on well-established scientific and medical research

23andme.com
Direct to Consumer
Personalized Genomics

Environment
pollution, sun exposure

Genetic
DNA, ethnicity

Lifestyle
smoking, exercise

Why genetics is only part of the story.

When it comes to your health and traits, DNA is only part of the story. Other variables come into play, including non-genetic factors, such as your environment and lifestyle.
Direct to Consumer
Personalized Genomics

What is in the kit?

- saliva collection kit
- specimen bag
- step by step instructions
- funnel lid
- saliva collection tube
- tube cap
- tube container

23andme.com
Direct to Consumer
Personalized Genomics

SHOW RESULTS FOR George Mias

**Health Risks (122)**

<table>
<thead>
<tr>
<th>Risk Condition</th>
<th>Your Risk</th>
<th>Average Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psoriasis</td>
<td>16.8%</td>
<td>11.4%</td>
</tr>
<tr>
<td>Parkinson's Disease</td>
<td>2.2%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Scleroderma (Limited Cutaneous Type)</td>
<td>0.08%</td>
<td>0.07%</td>
</tr>
</tbody>
</table>

**DECREASED RISKS**

<table>
<thead>
<tr>
<th>Risk Condition</th>
<th>Your Risk</th>
<th>Average Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gout</td>
<td>17.1%</td>
<td>22.8%</td>
</tr>
<tr>
<td>Alzheimer's Disease</td>
<td>4.3%</td>
<td>7.2%</td>
</tr>
</tbody>
</table>

**Inherited Conditions (53)**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cystic Fibrosis</td>
<td>Variant Present</td>
</tr>
<tr>
<td>Hemochromatosis (HFE-related)</td>
<td>Variant Present</td>
</tr>
<tr>
<td>ARSACS</td>
<td>Variant Absent</td>
</tr>
<tr>
<td>Agenesis of the Corpus Callosum with Peripheral Neuropathy (ACC PN)</td>
<td>Variant Absent</td>
</tr>
<tr>
<td>Alpha-1 Antitrypsin Deficiency</td>
<td>Variant Absent</td>
</tr>
</tbody>
</table>

See all 53 carrier status...

**Traits (63)**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol Flush Reaction</td>
<td>Does Not Flush</td>
</tr>
<tr>
<td>Bitter Taste Perception</td>
<td>Unlikely to Taste</td>
</tr>
<tr>
<td>Blond Hair</td>
<td>&lt;1% Chance</td>
</tr>
<tr>
<td>Earwax Type</td>
<td>Wet</td>
</tr>
<tr>
<td>Eye Color</td>
<td>Likely Brown</td>
</tr>
</tbody>
</table>

See all 63 traits...

**Drug Response (25)**

<table>
<thead>
<tr>
<th>Drug Response</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proton Pump Inhibitor (PPI) Metabolism (CYP2C19-related)</td>
<td>Rapid</td>
</tr>
<tr>
<td>Warfarin (Coumadin*) Sensitivity</td>
<td>Increased</td>
</tr>
<tr>
<td>Phenytoin Sensitivity (Epilepsy Drug)</td>
<td>Increased</td>
</tr>
<tr>
<td>Sulfonylurea Metabolism</td>
<td>Greatly reduced</td>
</tr>
<tr>
<td>Abacavir Hypersensitivity</td>
<td>Typical</td>
</tr>
</tbody>
</table>

See all 25 drug response...
Direct to Consumer
Personalized Genomics

Parkinson's Disease

Parkinson's disease is a disorder of the brain's motor system caused by a loss of dopamine-producing brain cells. Approximately one and a half million Americans have the disease, and about 50,000 new patients are diagnosed each year. The main symptoms are trembling in the hands, arms, legs, jaw, and face; stiffness of the limbs and trunk; slowed movement; and impaired balance and coordination. Symptoms of Parkinson's disease usually come on gradually and affect people over the age of 50, although there are rare forms that progress more quickly and strike at a younger age. Though very little is known about the genetics of Parkinson's, mutations in a gene known as LRRK2 have been found to greatly increase a person's likelihood of developing the condition.

The following results are based on established research for 10 reported markers, updated April 26th, 2012.

Your Results

Show information for George Mias assuming European ethnicity and an age range of 30-79

George Mias
2.2 out of 100 men of European ethnicity who share George Mias's genotype will develop Parkinson's Disease between the ages of 30 and 79.

Average
1.6 out of 100 men of European ethnicity will develop Parkinson's Disease between the ages of 30 and 79.

What does the Odds Calculator show me?

Use the ethnicity and age range selectors above to see the estimated incidence of Parkinson's Disease due to genetics for men with George Mias's genotype. The 23andMe Odds Calculator assumes that a person is free of the condition at the lower age in the range. You can use the name selector above to see the estimated incidence of Parkinson's Disease for the genotypes of other people in your account.

The 23andMe Odds Calculator only takes into account effects of markers with known associations that are also on our genotyping chip. Keep in mind that aside from genetics, environment and lifestyle may also contribute to one's risk for Parkinson's Disease.

Understanding Your Results

The heritability of Parkinson's is relatively low but a recent study estimated it to be about 27% in European populations. This means that environment generally plays a larger role than genetics in determining a person's risk for the disease. However, a small fraction of Parkinson's cases are attributed to rare mutations in a small number of genes, including the G2019S mutation in LRRK2, which is included in this report. People with the LRRK2 G2019S mutation have a much higher than average risk of developing Parkinson's disease during their lifetimes. This report also discusses other genetic factors that are associated with higher risk for PD in European and Asian populations. (sources)
Direct to Consumer
Personalized Genomics

Marker Effects

What does this chart show?
The chart shows the approximate effects of the selected person’s genotype at the 10 reported markers. Higher, red bars indicate increased risk from the average, while lower, green bars indicate decreased risk from the average. The light gray bars show the maximum possible effects for the possible genotypes at the marker.

Mouse over individual bars to view additional information about each marker. Click on a bar to view detailed information about that marker below. You can read more about all markers in the technical report.

LRRK2 Marker: rs34637584

Mutations in the LRRK2 gene are one of the most common known genetic causes of Parkinson’s disease (PD).

More than 50 variants are known in the LRRK2 gene. Several of these have been associated with PD. This variant reported by 23andMe, rs34637584, also known as the G2019S mutation, is the best-studied LRRK2 SNP related to Parkinson’s in individuals with European ancestry.

Parkinson’s is a fairly rare disease. The average person has a 1-2% chance of developing the disease during his or her lifetime. The chance that a person with the G2019S mutation will develop Parkinson’s is much higher than average and increases with age. One recent study found that people with the G2019S mutation have a 28% chance of developing Parkinson’s by the age of 59, 51% by the age of 69 and 74% by the age of 79. However, estimates of PD risk due to the G2019S mutation vary greatly. While it is well established that the mutation’s effect is very strong, there is no consensus about its exact magnitude.

Of all people with Parkinson’s, few have the G2019S mutation, but it is present at high levels in patients from some ethnic groups. Up to 40% of people with PD who are of Arab-Berber ancestry and 20% of Ashkenazi Jewish people with PD have this mutation.

Scientists do not know why only some people with the G2019S mutation get PD. There may be unknown effects due to other genes or environmental factors.

Citations

Direct to Consumer  
Personalized Genomics

<table>
<thead>
<tr>
<th>Gene or Region</th>
<th>SNP</th>
<th>Genotype</th>
<th>Adjusted Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRRK2</td>
<td>rs34637584</td>
<td>GG</td>
<td>0.98</td>
</tr>
<tr>
<td>LRRK2</td>
<td>rs34778348</td>
<td>GG</td>
<td>NA (not applicable)</td>
</tr>
<tr>
<td>GBA</td>
<td>i4000415</td>
<td>TT</td>
<td>0.99</td>
</tr>
<tr>
<td>SNCA</td>
<td>rs356220</td>
<td>CT</td>
<td>1.02</td>
</tr>
<tr>
<td>MAPT</td>
<td>rs393152</td>
<td>AA</td>
<td>1.09</td>
</tr>
<tr>
<td>PARK16</td>
<td>rs947211</td>
<td>GG</td>
<td>1.04</td>
</tr>
<tr>
<td>PARK16</td>
<td>rs823156</td>
<td>AA</td>
<td>NA (not applicable)</td>
</tr>
<tr>
<td>BST1</td>
<td>rs4698412</td>
<td>AA</td>
<td>1.13</td>
</tr>
<tr>
<td>DGKQ</td>
<td>rs11248060</td>
<td>CC</td>
<td>0.94</td>
</tr>
<tr>
<td>STK39</td>
<td>rs2390669</td>
<td>AC</td>
<td>1.15</td>
</tr>
</tbody>
</table>

23andme.com (Older style reports pre FDA)
Detected the following variants: 2789+5G>A

<table>
<thead>
<tr>
<th>23andMe Name</th>
<th>Other Name(s)</th>
<th>DNA Change</th>
<th>Genotype</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>i3000001</td>
<td>DeltaF508</td>
<td>CTT to –</td>
<td>CTT, CTT</td>
<td>Has one mutation in the <strong>CFTR</strong> gene linked to cystic fibrosis. A person with one of these mutations typically does not have cystic fibrosis, but may pass the mutation to offspring. May still have other mutations in the CFTR gene (not reported here). Variants detected: 2789+5G&gt;A</td>
</tr>
<tr>
<td>i4000292</td>
<td>DeltaI507</td>
<td>ATC to –</td>
<td>ATC, ATC</td>
<td></td>
</tr>
<tr>
<td>i4000294</td>
<td>G85E</td>
<td>G to A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000296</td>
<td>R334W</td>
<td>C to T</td>
<td>CC</td>
<td></td>
</tr>
<tr>
<td>i4000297</td>
<td>R347P/H</td>
<td>G to C, A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000291</td>
<td>A455E</td>
<td>C to A</td>
<td>CC</td>
<td></td>
</tr>
<tr>
<td>i4000299</td>
<td>V520F</td>
<td>G to T</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000300</td>
<td>G542X</td>
<td>G to T</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000301</td>
<td>S549N</td>
<td>G to A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000305</td>
<td>G551D</td>
<td>G to A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000306</td>
<td>R553X</td>
<td>C to T</td>
<td>CC</td>
<td></td>
</tr>
<tr>
<td>i4000307</td>
<td>R560T</td>
<td>G to C</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000308</td>
<td>R1162X</td>
<td>C to T</td>
<td>CC</td>
<td></td>
</tr>
<tr>
<td>i4000309</td>
<td>W1282X</td>
<td>G to A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000311</td>
<td>N1303K</td>
<td>C to G</td>
<td>CC</td>
<td></td>
</tr>
<tr>
<td>i4000313</td>
<td>394delTT</td>
<td>TT to –</td>
<td>TT, TT</td>
<td></td>
</tr>
<tr>
<td>i4000314</td>
<td>621+1G&gt;T</td>
<td>G to T</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000315</td>
<td>711+1G&gt;T</td>
<td>G to T</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000316</td>
<td>1078delT</td>
<td>T to –</td>
<td>T, T</td>
<td></td>
</tr>
<tr>
<td>i4000317</td>
<td>1717-1G&gt;A</td>
<td>G to A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000318</td>
<td>1898+1G&gt;A</td>
<td>G to A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000320</td>
<td>2789+5G&gt;A</td>
<td>G to A</td>
<td>AG</td>
<td></td>
</tr>
<tr>
<td>i4000321</td>
<td>3120+1G&gt;A</td>
<td>G to A</td>
<td>GG</td>
<td></td>
</tr>
<tr>
<td>i4000322</td>
<td>3659delC</td>
<td>C to –</td>
<td>C, C</td>
<td></td>
</tr>
<tr>
<td>i4000324</td>
<td>3905insT</td>
<td>– to T</td>
<td>–, –</td>
<td></td>
</tr>
<tr>
<td>i4000325</td>
<td>3849+10kbC&gt;T</td>
<td>C to T</td>
<td>CC</td>
<td></td>
</tr>
</tbody>
</table>

**Cystic Fibrosis**

Cystic fibrosis (CF) is caused by mutations in a **gene** called **CFTR**. This gene codes for a **protein** that helps move salt and water through cells. Mutations in CFTR result in the build-up of thickened mucus and other secretions that can damage the lungs, pancreas, and other organs. The disease is inherited in a **recessive** manner, meaning that a person must inherit a mutated copy of the CFTR gene from both parents in order to develop the disease. Cystic fibrosis is most common in populations with European or Jewish ancestry, where about one out of every 25 to 30 people carries a CFTR mutation.

The following results are based on ★★★★ Established Research for 26 reported markers, updated May 29th, 2014.

23andme.com (Older style reports pre FDA)
Direct to Consumer
Personalized Genomics

Warfarin (Coumadin®) Sensitivity

Warfarin is an anticoagulant (also known as a blood thinner). It is used to treat and prevent blood clots. Blood clots can block blood flow and cause a heart attack or stroke.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>Combination</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1799853</td>
<td>CT</td>
<td>CYP2C9*2/*3, VKORC1 -1673/3673 AG</td>
<td>Likely to be more sensitive to warfarin based on genetics. Genetic information may only be useful when determining an initial dose of warfarin. Many other factors also influence warfarin sensitivity. If you are taking warfarin, keep taking it as directed by your doctor.</td>
</tr>
<tr>
<td>rs1057910</td>
<td>AC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9923231</td>
<td>CT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sulfonylurea Metabolism

Sulfonylurea drugs are commonly used to treat type 2 diabetes, a disease that affects about 350 million people worldwide. Genetic as well as non-genetic factors can influence how a person responds to these drugs. This report covers two genetic variants associated with the ability to clear to sulfonylurea drugs from the body. Decreased drug clearance can result in better chances for successful treatment but may also increase the risk of side effects.

The following results are based on 4★ Established Research for 2 reported markers.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>Combination</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1799853</td>
<td>CT</td>
<td>CYP2C9 *2/*3</td>
<td>Greatly reduced ability to clear sulfonylurea drugs from the body. Clearance may affect treatment effectiveness, likelihood of side effects, and optimal dose.</td>
</tr>
<tr>
<td>rs1057910</td>
<td>AC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Direct to Consumer
Personalized Genomics

MATERNAL LINE: H

Locations of haplogroup H before the widespread migrations of the past few hundred years.

H originated in the Near East and then expanded after the peak of the Ice Age into Europe, where it is the most prevalent haplogroup today. It is present in about half of the Scandinavian population and is also common along the continent’s Atlantic coast.

PATERNAL LINE: E1b1b1a2*

E1b1b1a2* is a subgroup of E1b1b1a

Locations of haplogroup E1b1b1a before the widespread migrations of the past few hundred years.

E1b1b1a is most common in northern Africa and southern Europe. It arose about 23,000 years ago in eastern Africa and spread into the Mediterranean region after the Ice Age. E1b1b1a, a subgroup of E1b1b, expanded out of the Near East 8,000 years ago into northern Africa and southern Europe. Today it is one of the most common haplogroups in those regions.
Direct to Consumer
Personalized Genomics
Global Similarity

DNA Relatives

23andme.com (Older style reports pre FDA)
Direct to Consumer
Personalized Genomics

Got Neanderthal DNA?

An estimated 2.6% of your DNA is from Neanderthals.

George Mias (you)

Average European user

MODERN HUMANS
- Higher brow
- Narrower shoulders
- Slightly taller

NEANDERTHALS
- Heavy eyebrow ridge
- Long, low, bigger skull
- Prominent nose with developed nasal chambers for cold-air protection
INTEGRATIVE PERSONAL OMICS PROFILING

I. Whole Genome Sequencing
Disease Risk Evaluation
Medical History & Environment
Pharmacogenomic Evaluation

RISK EVALUATION

II. Transcriptomics
Proteomics
Metabolomics
Clinical Tests
Autoantibodyomics
Microbiomics
New omics

LONGITUDINAL OMICS PROFILING OF MULTIPLE PHYSIOLOGICAL STATES

Healthy → Infected → Recovery → Healthy

III. iPOP Database

INTEGRATION OF MULTIPLE OMICS AND TEMPORAL RESPONSES MATCHED AGAINST iPOP DATABASES
Personalized Medicine

- Personalized
- Determine risks
- Monitor
- Integrate

Mias and Snyder, Quantitative Biology 1(1) p. 71 (2013).
Integrative Personal Omics Profiling (iPOP)

I. Genome Sequencing

• Whole Genome Sequencing
  ‣ Illumina (120-fold coverage)
  ‣ Complete Genomics (150-fold coverage)
  ‣ Exome Sequencing (Nimblegen, Agilent and Illumina) (80-100-fold coverage)

• Variants identified
  ‣ $\sim 3.3 \times 10^6$ Single Nucleotide Variants (SNVs)
  ‣ $\sim 2 \times 10^5$ Small insertions and deletions (InDels)
  ‣ Structural variants (SVs > 1Kb changes)

I. Genome Sequencing

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parkinson's disease</td>
<td>18</td>
</tr>
<tr>
<td>Gout</td>
<td>1</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Alzheimer's disease</td>
<td>13</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>36</td>
</tr>
<tr>
<td>Depression</td>
<td>1</td>
</tr>
<tr>
<td>Pseudoexfoliation syndrome</td>
<td>1</td>
</tr>
<tr>
<td>Psoriasis</td>
<td>14</td>
</tr>
<tr>
<td>Stroke</td>
<td>4</td>
</tr>
<tr>
<td>Asthma</td>
<td>13</td>
</tr>
<tr>
<td>Obesity</td>
<td>11</td>
</tr>
<tr>
<td>Hypertension</td>
<td>5</td>
</tr>
<tr>
<td>Age related macular degeneration</td>
<td>12</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>1</td>
</tr>
<tr>
<td>Hypertriglyceridemia</td>
<td>4</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>28</td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>6</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>30</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>1</td>
</tr>
<tr>
<td>Open angle glaucoma</td>
<td>5</td>
</tr>
</tbody>
</table>

Rong Chen, Atul Butte

Ashley et al. Lancet 375, 1525-1535 (2010)
I. Genome Sequencing

Rong Chen, Atul Butte

Ashley et al. Lancet 375, 1525-1535 (2010)
Data Timeline:

- Day 0
  - HRV onset
- Days 289-292
  - RSV onset
- High Glucose

Data Timeline:

- Timeline Events
  - Human Rhinovirus (HRV)
  - Respiratory Syncytial Virus (RSV)
  - High Glucose (Type II Diabetes - as per physician)

Data Timeline: Infections

Day 0
HRV onset

Days 289-292
RSV onset

High Glucose

0 4 21 116 185 186 255 289 290 292 294 297 301 307 311 322 329 369 380 400

Data Timeline: Cytokines

Data Timeline: Glucose Levels

Glycated HbA1c (%)

<table>
<thead>
<tr>
<th>Day</th>
<th>HRV</th>
<th>RSV</th>
<th>Lifestyle Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4</td>
<td>329</td>
<td>369</td>
<td>6.7 (329)</td>
</tr>
<tr>
<td>6.7</td>
<td>369</td>
<td>496</td>
<td>4.9 (496)</td>
</tr>
</tbody>
</table>

Data Timeline: Autoantibody-ome

- ProtoArrays (Invitrogen)
Data Timeline: Autoantibody-ome

Higher in Test Subject Vs. Healthy Group (40pts)

- **ARRDC3** arrestin domain containing 3
- **EIF3E** eukaryotic translation initiation factor 3
- **PAQR4** progestin and adipoQ receptor family member 1
- **DOK6** docking protein 6 (insulin receptor docking)
- **GOSR1** golgi SNAP receptor complex member 1
- **BTK** Bruton agammaglobulinemia tyrosine kinase
- **ASPA** aspartoacylase

**Blue** intersect with RNA expression

**Yellow** insulin related


Mias & Snyder, Quantitative Biology 1(1) p. 71 (2013).
II. Dynamics: Data Analysis Framework

- Integration of Dynamic Omics
  - Transcriptome
  - Proteome
  - Metabolome

1. Preprocessing
2. Common Classification Scheme
   - i. Overall trends (autocorrelation)
   - ii. Spikes at specific timepoints
3. Clustering and Enrichment Analysis

II. Dynamics: Data Analysis Framework

1. Preprocessing
2. Common Classification Scheme
   i. Overall trends (autocorrelation)
   ii. Spikes (maxima) at specific timepoints
   iii. Spikes (minima)
3. Clustering and Enrichment Analysis

II. Dynamics: Data Analysis Framework

- (1) Data Preprocessing: Transcriptome

II. Dynamics: Data Analysis Framework

- (1) Data Preprocessing: Proteome

Tracking Different Timepoints with Chemical Labeling (TMT 6plex)

http://www.thermo.com
II. Dynamics: Data Analysis Framework

- (1) Data Preprocessing: Proteome

Tracking Different Timepoints with Chemical Labeling (TMT 6plex)

- Common reference ratio
  - Reproducible
  - Sets can be combined: 3,731 proteins followed over 14 timepoints

II. Dynamics: Data Analysis Framework

- (1) Data Preprocessing: Proteome

II. Dynamics: Data Analysis Framework

- (2) Common Framework Data Classification

<table>
<thead>
<tr>
<th></th>
<th>Transcriptome (HRV+RSV)</th>
<th>Proteome (RSV)</th>
<th>Metabolome (RSV)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>19714</td>
<td>3731</td>
<td>4228</td>
</tr>
<tr>
<td><strong>Autocorr</strong></td>
<td>4922</td>
<td>257</td>
<td>475</td>
</tr>
<tr>
<td><strong>Spike Max</strong></td>
<td>3718</td>
<td>1240</td>
<td>577</td>
</tr>
<tr>
<td><strong>Spike Min</strong></td>
<td>7891</td>
<td>1194</td>
<td>884</td>
</tr>
</tbody>
</table>

II. Dynamics: Data Analysis Framework

Raw Datasets:
- Transcriptome
- Proteome
- Metabolome

Data Preprocessing:
- QC
- Normalization
- Statistical Simulation

Fourier Space
- Common Framework
- Data Classification: Autocorrelated (I) and Spike Sets (II) & (III)

Clustering and Enrichment Analysis

Classified Data Clustering

Cluster Selection
- based on fusion coefficient analysis
- Mathematica

Gene-based Annotation (RefSeq + Uniprot)
- Metabolite KEGG annotation

BiNGO GO enrichment analysis: MF, BP, CC

Known Reactome FI pathway analysis

per Cluster Data Visualization and Cytoscape Data Integration

III. Dynamics: Transcriptome

• (3) Clustering and Enrichment Analysis

III. Dynamics: Transcriptome

- (3) Clustering and Enrichment Analysis

(II) Spike Maxima Data Clusters

HRV
RSV

Days after HRV infection

Gene Ontology analysis,
$7 \times 10^{-3} > p > 5 \times 10^{-12}$
- antigen processing and presentation
- immune system process
- immune response
- antigen processing & presentation of peptide or polysaccharide antigen via MHC class II defense response
- inflammatory response

E.g. Pathway analysis,
FDR $< 1 \times 10^{-3}$
- Allograft rejection(K)
- Intestinal immune network for IgA production(K)
- Graft-versus-host disease(K)
- Type I diabetes mellitus(K)
- Phagosome(K)
- Asthma(K)
- Autoimmune thyroid disease(K)
- Viral myocarditis(K)
- Endocytosis(K)
- Fc gamma R-mediated phagocytosis(K)

III. Dynamics: Transcriptome

- (3) Clustering and Enrichment Analysis

E.g. Pathway analysis, FDR < 1x10^{-3}
- TCR signaling in naive CD4+ T cells (N)
- Signalling by NGF (R)
- TRAIL signaling pathway (N)
- p75(NTR)-mediated signaling (N)
- TNF receptor signaling pathway (N)
- Glypican pathway (N)
- Neurotrophin signaling pathway (K)
- IL1-mediated signaling events (N)
- T cell receptor signaling pathway (K)
- TCR signaling in naive CD8+ T cells (N)
- Plasma membrane estrogen receptor signaling (N)
- Proteoglycan syndecan-mediated signaling events (N)
- EphrinB-EPHB pathway (N)
- Trk receptor signaling mediated by PI3K and PLC-gamma (N)
- BMP receptor signaling (N)
- Regulation of cytoplasmic and nuclear SMAD2/3 (N)
- TGF-beta receptor signaling (N)
- Neurotrophic factor-mediated Trk receptor signaling (N)
- Insulin signaling pathway (K)
- IFN-gamma pathway (N)
- p38 signaling mediated by MAPKAP kinases (N)
- BCR signaling pathway (N)

E.g. Pathway: IFN-gamma

III. Dynamics: Integrated Omics

• (3) Clustering and Enrichment Analysis

Days after RSV infection

Autocorrelated Data Clusters

E.g. Pathway: Protein processing in endoplasmic reticulum

E.g. Pathway: Phagosome

E.g. Pathway: Lysosome

E.g. Pathway: Insulin

III. Dynamics: Integrated Omics

- (3) Clustering and Enrichment Analysis

**E.g. Pathway Analyses, FDR <5x10⁻²**
- Spliceosome (K)
- Glucose Regulation of Insulin Secretion (R)
- Formation and Maturation of mRNA Transcript (R)
- Oxidative phosphorylation (K)
- Electron Transport Chain (R)
- Parkinson's disease (K)
- Huntington's disease (K)
- Influenza Life Cycle (R)
- Metabolism of non-coding RNA (R)
- Transport of Mature Transcript to Cytoplasm (R)
- Protein export (K)
- Pyruvate metabolism and TCA cycle (R)

**E.g. Metabolites in Cluster**
- 3R-hydroxy-5Z-dodecenoic acid
- 5,6-DiHETE-EA
- 7-Ethoxycoumarin
- Lauric acid
- 1-O-(1Z-tetradecenyl)-2-(9Z-octadecenoyl)-sn-glycerol
- 1alpha-hydroxy-26,27-dinorvitamin D3
- 25-carboxylic acid / 1alpha-hydroxy-26,27-dinorcholecalciferol
- 12-oxo-9-octadecynoic acid
- GPCho(O-16:0/O-4:0[U])
- 19-hydroxy-17-o xoandrost-5-en-3-beta-yl sulfate - 11.538899

**E.g. Pathway:**
Glucose Regulation of Insulin Secretion

**Chen*, Mias*, Li-Pook-Than*, Jiang* et al. Cell 148, 1293 (2012).**
III. Dynamics: Integrated Omics

• (3) Clustering and Enrichment Analysis

E.g. Pathway Analyses, FDR <5x10^-2

<table>
<thead>
<tr>
<th>Pathway/Signaling Event</th>
<th>Metabolites in Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAIL signaling pathway</td>
<td>4,8,12,15,19,21-tetracosahexaenoic acid</td>
</tr>
<tr>
<td>Class I PI3K signaling events</td>
<td>Idebenone Metabolite</td>
</tr>
<tr>
<td>Glypican pathway</td>
<td>p-Cymene</td>
</tr>
<tr>
<td>TNF receptor signaling pathway</td>
<td>Cer(d18:0/24:0)</td>
</tr>
<tr>
<td>Signaling by NGF</td>
<td>Normetanephrine sulfate</td>
</tr>
<tr>
<td>p75(NTR)-mediated signaling</td>
<td>gamma-Hydroxyphenylbutazone</td>
</tr>
<tr>
<td>TCR signaling in naive CD4+ T cells</td>
<td>glucuronide</td>
</tr>
<tr>
<td>Apoptosis</td>
<td>Methoxyhydroxymethylhydrocinnamic acid</td>
</tr>
<tr>
<td>IGF1 pathway</td>
<td>5,8-Dihydroxy-3,4-dihydroxybutoxytril</td>
</tr>
<tr>
<td>Arf6 signaling events</td>
<td>3-Methylsuberic acid - 7.6812496</td>
</tr>
<tr>
<td>Insulin Pathway</td>
<td>5-Methytrihydroxydopamyltril-Lglutamate</td>
</tr>
<tr>
<td>IFN-gamma pathway</td>
<td>3-keto palmitic acid - 12.719786</td>
</tr>
<tr>
<td>Signaling by Wnt</td>
<td>Cholylglycine</td>
</tr>
<tr>
<td>Neurotrophin signaling pathway</td>
<td>Protoporphyrin + 12.516361</td>
</tr>
<tr>
<td>IL2- &amp; IL1 mediated signaling events</td>
<td>Desmethyldehydroamiodipine</td>
</tr>
<tr>
<td>Fc gamma R-mediated phagocytosis</td>
<td>Chenodeoxycholyltaurine</td>
</tr>
</tbody>
</table>

E.g. Metabolites in Cluster

- 4,8,12,15,19,21-tetracosahexaenoic acid
- Idebenone Metabolite
- p-Cymene
- Cer(d18:0/24:0)
- Normetanephrine sulfate
- gamma-Hydroxyphenylbutazone
- glucuronide
- Methoxyhydroxymethylhydrocinnamic acid
- 5,8-Dihydroxy-3,4-dihydroxybutoxytril
- 3-Methylsuberic acid - 7.6812496
- 5-Methytrihydroxydopamyltril-Lglutamate
- 3-keto palmitic acid - 12.719786
- Cholylglycine
- Protoporphyrin + 12.516361
- Desmethyldehydroamiodipine
- Chenodeoxycholyltaurine
Precision Medicine

- Personalized
- Determine risks
- Monitor
- Integrate

Mias and Snyder, Quantitative Biology 1(1) p. 71 (2013).
A Framework for Personalized & Precision Medicine

Mias & Snyder, Clinical Pharmacology and Therapeutics 93, p29 (2013)