Using Omics and Big Data to Manage Health and Disease

Michael Snyder
Stanford University
May 19, 2017

Conflicts: Personalis, Genapsys, SensOomics, Qbio

Drivers of Big Data

1) DNA Sequencing
Human Genome Cost <$2K

2) Mass Spectrometry

3) Wearables

General Goals

1) Understand how individuals change over time and during periods of health and disease at high resolution

2) Understand how different “omes” (microbiome, metabolome, proteome, genome) relate to one another dynamically

3) Understand how individual responses are similar and differ from one another when faced with specific perturbations

4) Identify factors that can affect and help manage the health of an individual

Personal Omics Profiling

Biosensors

Omics Measurements

Billions of Measurements!

Health Is a Product of Genome & Exposome

Genome

Exercise

Pathogens

Food

Stress

Medical Tests, Questionaires

Personal Omics Profile

84 months; >250 Timepoints; 10 Viral Infections

Chen et al., Cell 2012, unpublished
**Genome Sequence** (Illumina, Complete Genomics)

**Glucose levels**

![Graph showing glucose levels over time](image)

**Extended Time Line**

![Graph showing HgA1c levels over time](image)

**Molecules and Biochemical Pathways that Change During Acquisition of Diabetes**

- Insulin Biosynthetic Pathway
- Glucose Regulation of Insulin Secretion
- Platelet Plug Formation

**Epigenetics: DNA Methylation**

- Affected by nutrition, lifestyle factors, aging, and environment
- Causes gene silencing

![DNA methylation](image)

**Gene Inactivation by Mutation and Methylation: PDE4 involved in eosinophilia**

- PDE4 DIP Gene
- Father: Inactivated by mutation
- Mother: Inactivated by DNA Methylation

![Gene expression](image)

**Longitudinal Profiling of 100 individuals (Prediabetics & Healthy) over periods of health, stress and disease**

- Year 1: Viral infection
- Year 2: Stress
- Year 2: Diet change

![Longitudinal profiling](image)
# Genome Sequencing – First 60 People

- Eleven have important pathogenic mutations:
  - SHBD (2X): high freq. of neuroendocrine tumors
  - PROC: Affects coagulation
  - HNF1A: MODY mutation
  - ABCG8: Hyperinsulinemic hypoglycemia
  - MUTYH: Colon cancer
  - SLC7A9: Cystinuria
  - RBM20: Dilated cardiomyopathy
  - CHEK2: Breast cancer
  - APC: Colon cancer
  - BRCA1: Breast & ovarian cancer

- All have reportable carrier mutations and/or pharmacogenetic variants

Shannon Rego et al.  Personalis, Inc

---

## A subset of individuals undergo a dietary perturbation.

23 participants:

- 13 Insulin resistant
- 10 healthy controls (BMI matched)

Brian Piening, Wenyu Zhou, Gucci Gu, Kevin Contrepois

---

## Metabolic differences between IR and IS

Univariate analysis: Wilcoxon test P value < 0.05 and fold change > 1.5

---

## Example data: Short-term weight gain

Integrative c-means clustering: pattern recognition across RNA-seq, proteome, metabolome, microbiome, cytokines
Blood cytokine profiles: 20 subjects at baseline

Normalized log2 plasma concentration

Microbial abundance pattern group by individual, not by dietary supplement

Distance matrix by Manhattan methods and Hierarchical clustering by Ward method

Sensors: Measure Many Things

Overview of the project

Circadian and diurnal patterns

43 People
Activity Phenotypes: 4 Patterns

SpO2 Levels
Drops During Airline Flights
Finger device
91-96%: 65.4%
90 or less: 5.1%

SpO2 levels drops on airplane flights:
18 Subjects

SpO2 Measurements
Associated With Fatigue

SpO2 Measurements
Can Adapt on Flights >7hrs

SpO2 Measurements: Unusually Low on One Flight to Norway

Li, Dunn et al.
PloS Biol 2017

Li, Dunn et al.
PloS Biol 2017

Li, Dunn et al.
PloS Biol 2017

Li, Dunn et al.
PloS Biol 2017
Elevated Heart Rate and Skin Temperature During Lyme Disease

Lyme Disease

Early detection of Lyme disease

Detects All Days of Illness

Change-of-Heart Algorithm

Dects Periods of Illness at High Resolution using Wearable Device

HR differences in IR and Insulin Sensitivity

Smart Phone = Control Center
Overall Summary

1) Personal genome sequencing is here. It can be used to predict disease risk and manage health.
2) Multi-omics analyses are valuable for determining pathways and biochemical activities involved in human disease.
3) Longitudinal profiles are very valuable for understanding personal disease states.
4) Everyone’s profile is different.
5) Wearables will be useful for managing health.
6) Individuals will be responsible for their own health.

Acknowledgements

Snyder Lab
Wenyu Zhou
Brian Planing
Kevin Contraposi
Tejaswini Mithra
Kim Kukurba
Shannon Rapp
Jessica Siwal
Hannah Rost
Varsha Rao
Liang Liang
Tejas Mithra
Christina Yeh
Hassan Chab
Eric Wei

Wearables
Xiao Li
Jessie Dunn
Sophia Miyake...
Dennis Sleith
Heather Hall

Wainstock Lab
George Wainstock
Erica Sodergran
Shana Leopold
Daniel Spakowitz
Blake Hanson
Eddy Baulista
Lauren Peterson
Lei Chen
Benjamin Leopold
Sai Lak
Purna Vats

Jon Bernstein
McLaughlin Lab
Tracy McLaughlin
Colleen Craig
Candice Allstor
Dale Penisman
Elizabeth Colbert

NIH
Lita Proctor
Salvatore Sechi
Jon LoTempio
And other

Genomics and Personalized Medicine
What Everyone Needs to Know®

Michael Snyder
Available from Amazon