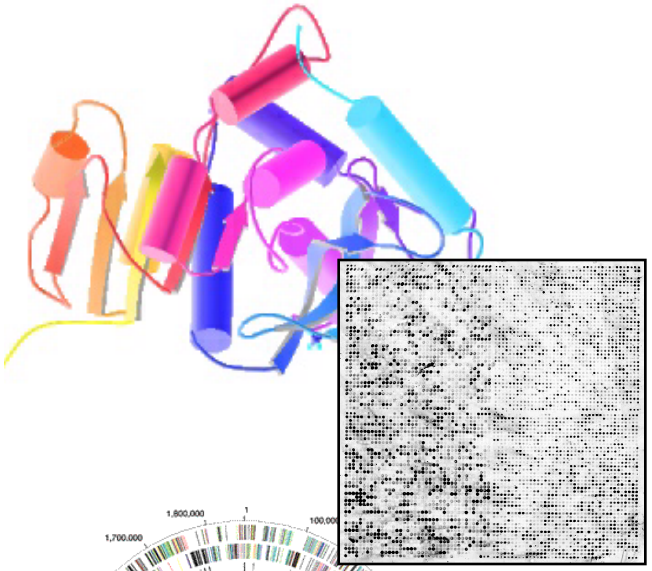
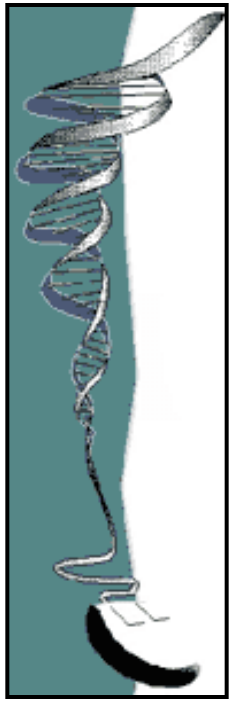
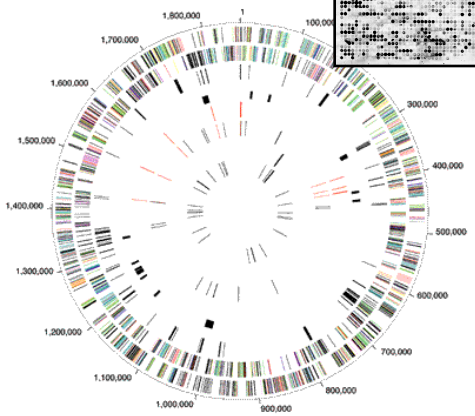
Biomed. Data Sci.
Personal Genomes Intro.



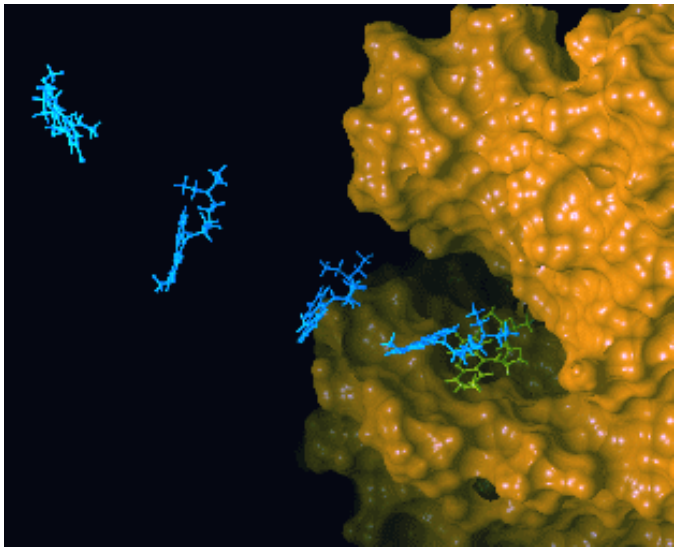


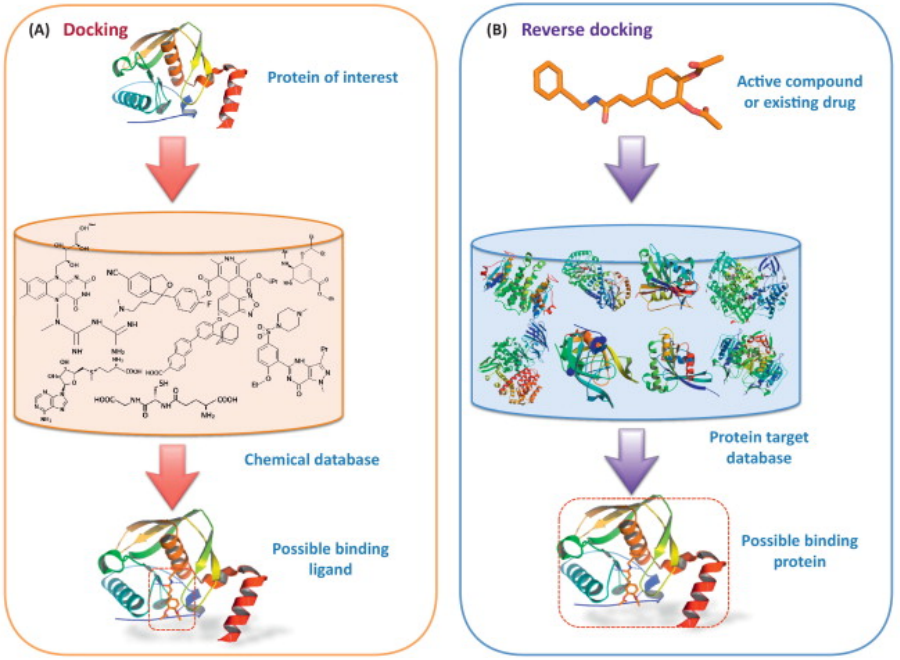


Mark Gerstein, Yale University
GersteinLab.org/courses/452
(last edit in spring ‘21)

Major Application I:
Designing Drugs from Structural Targets

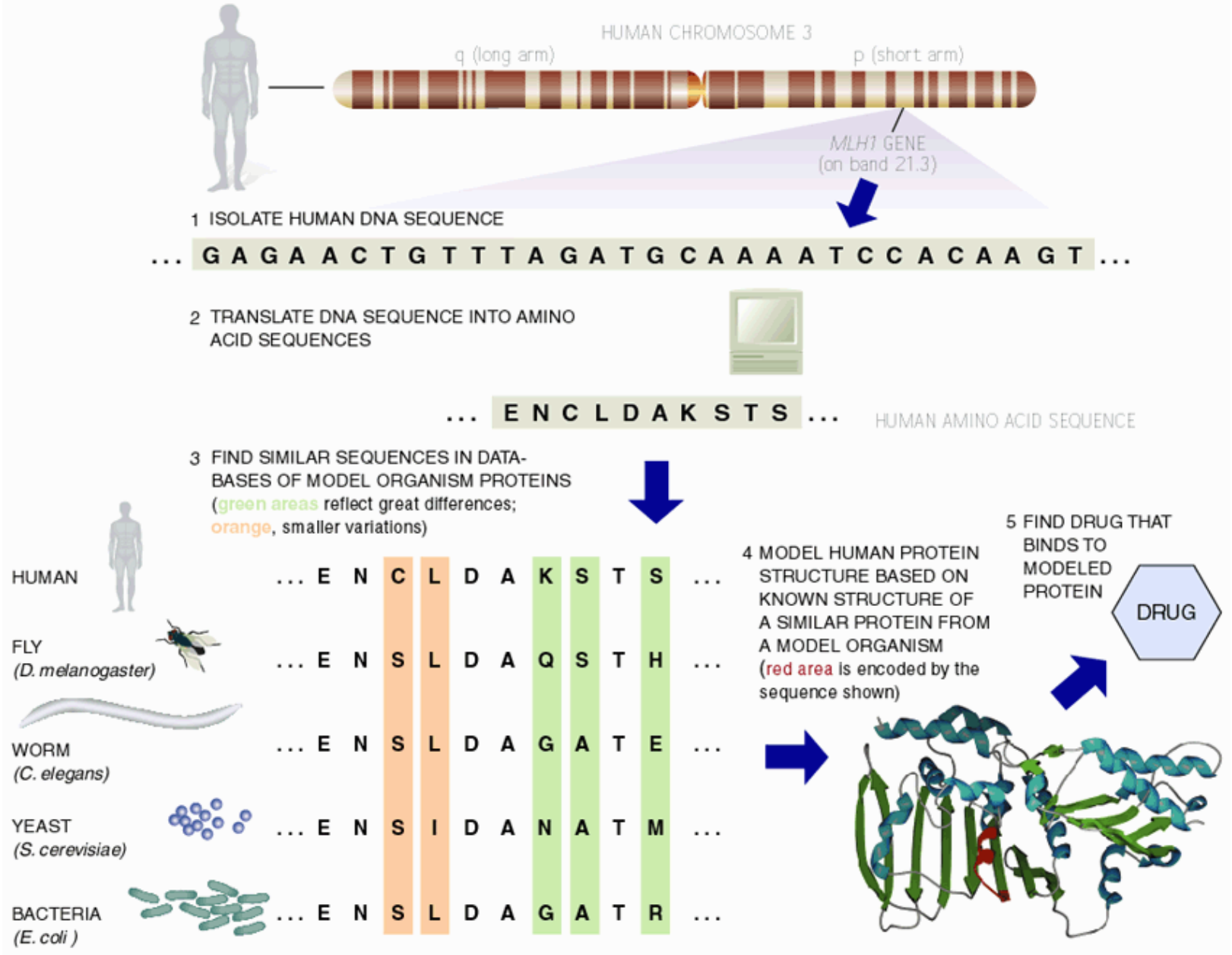
- Understanding how structures bind other molecules
- Designing inhibitors using docking, structure modeling
- *In silico* screens of chemical and protein databases





(From left to right, figures adapted from Olsen Group Docking Page at Scripps, Zheng et al. Trends in Pharmacological Sciences 2013)

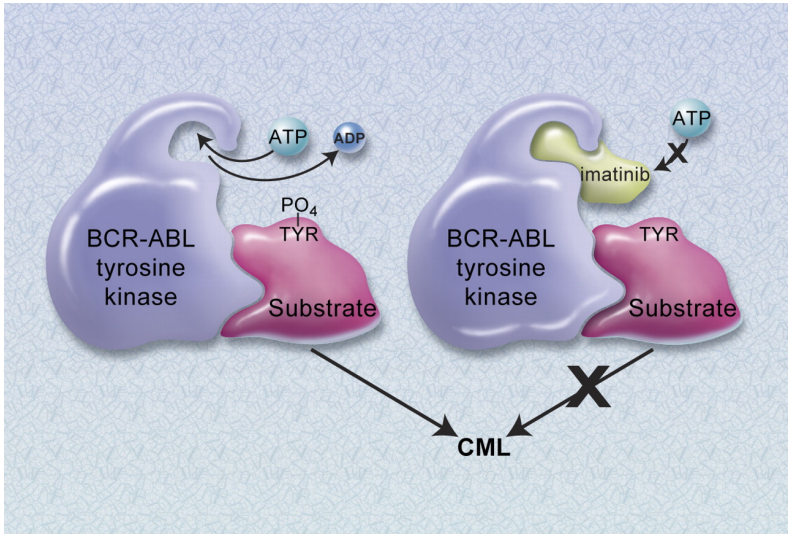
Major Application II: Finding Homologs, to Find Experimentally Tractable Gene Targets

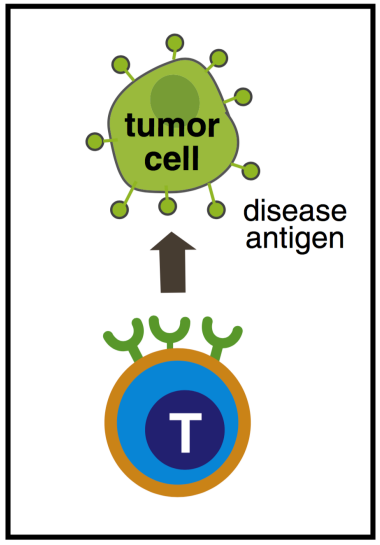


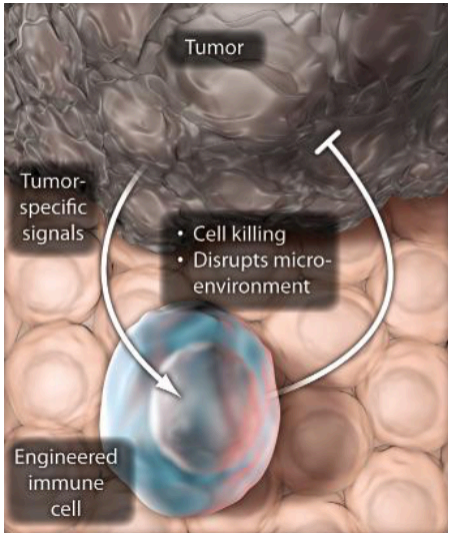
[Adapted from Sci. Am.]

Major Application III: Customizing treatment in oncology

- Identifying disease causing mutations in individual patients
- Designing targeted therapeutics
  - e.g. BCR-abl and Gleevec
  - Cancer immunotherapies targeting neoantigens



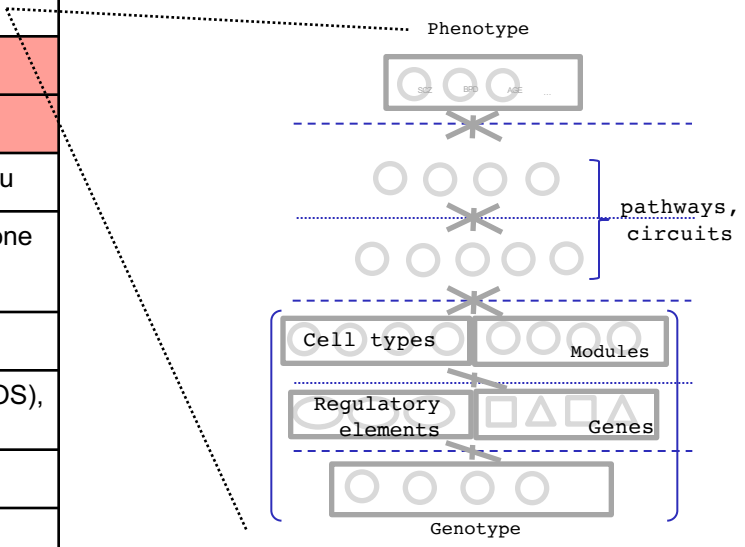




(From left to right, figures adapted from Druker BJ. Blood 2008 and the Lim Lab at UCSF)

Major Application IV:
Finding molecular mechanisms & drug targets for diseases we know little about (Neuro-psychiatic Diseases)

Disease	Heritability*	Molecular **Mechanisms**
Schizophrenia	**81%**	-
Bipolar disorder	70%	-
Alzheimer's disease	58 - 79%	Apolipoprotein E (APOE), Tau
Hypertension	30%	Renin–angiotensin–aldosterone
Heart disease	34-53%	Atherosclerosis, VCAM-1
Stroke	32%	Reactive oxygen species (ROS), Ischemia
Type-2 diabetes	26%	Insulin resistance
Breast Cancer	25-56%	BRCA, PTEN



*https://www.snpedia.com/index.php/Heritability

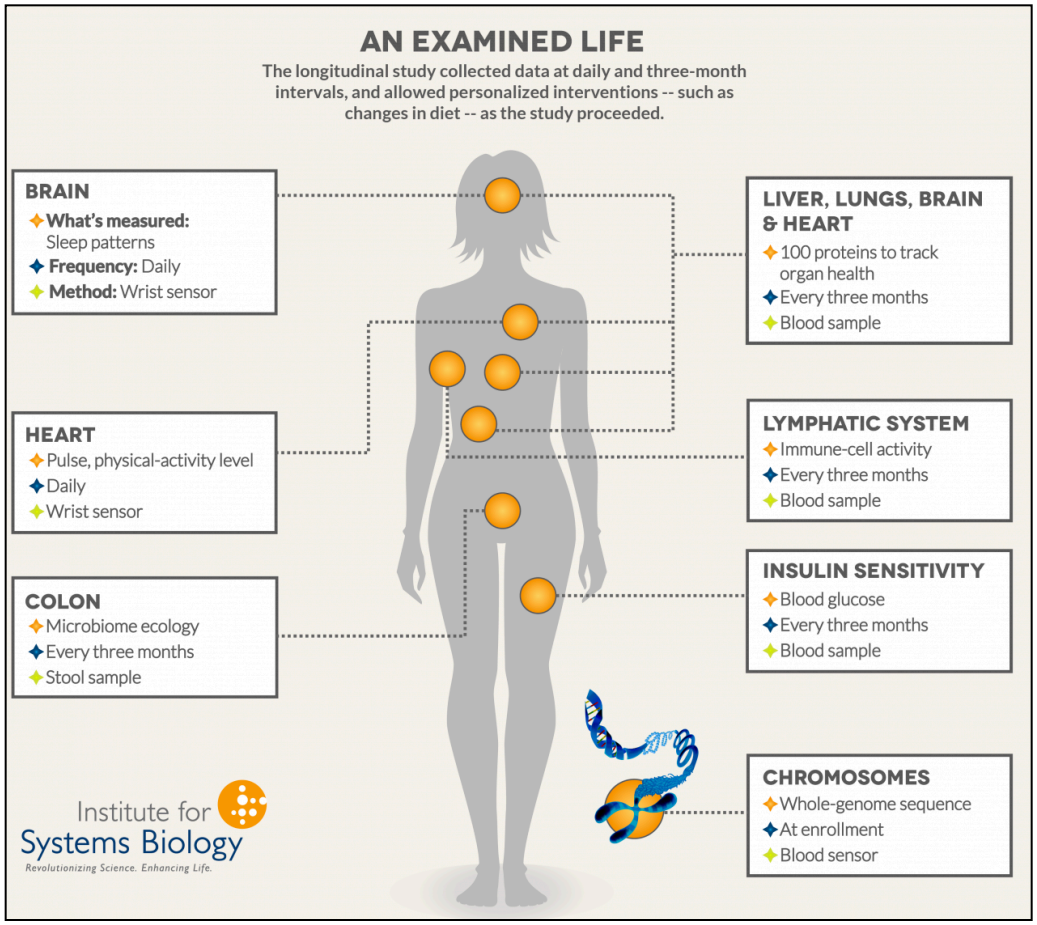
Many psychiatric conditions are highly heritable
- Schizophrenia: up to 80%

But we don't understand basic molecular mechanisms underpinning this association
- (in contrast to many other diseases such as cancer & heart disease)

Moreover, current models substantially underestimate heritability using genetic data
- Schizophrenia : ~25%

Thus, interested in developing predictive models of psychiatric traits which:
- Use observations at intermediate (molecular levels) levels to inform latent structure.
- Use the predictive features of these "molecular endo phenotypes" to begin to suggest actors involved in mechanism

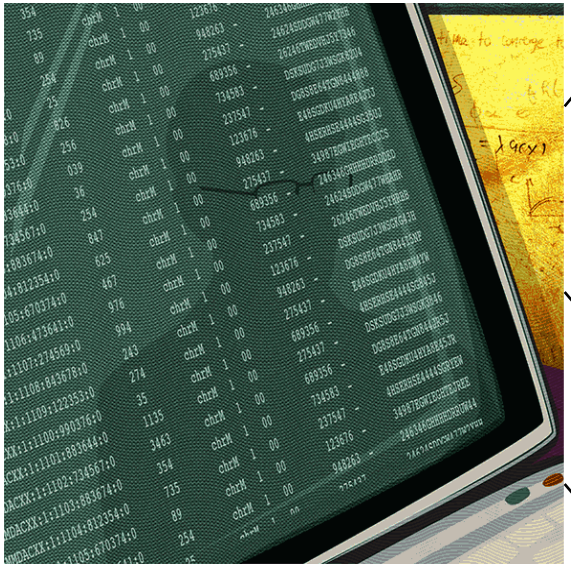
Major Application V: Holistic Personal Genome Characterization, in Normal Individuals



(Figure from Institute for Systems Biology)

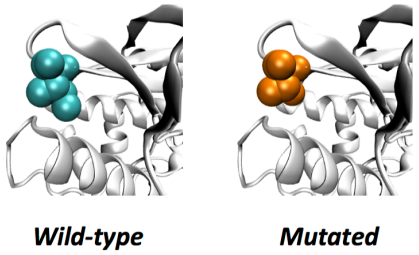
- Mental disease & cancer are two extremes with respect to genomics (CEN, 92: 26)
  - Many other conditions in between, often involving interaction with the environment
- Pers. Genome Characterization
  - Identify mutations in personal genomes (SNPs, SVs, &c)
  - Estimate phenotypic (deleterious or protective) impact of variants.
  - Compare one person to wider population.
- Track changes over time & consider interaction w/ environment
  - Transcriptome studies
  - Longitudinal health studies (e.g. 100K wellness project, Framingham Heart Study)

Analyzing Carl Zimmer's genome

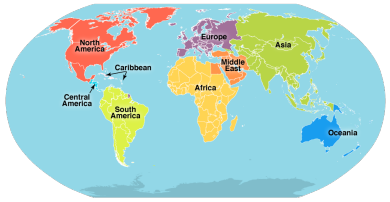


SNV AAGCT → ACGCT

Protein Structure

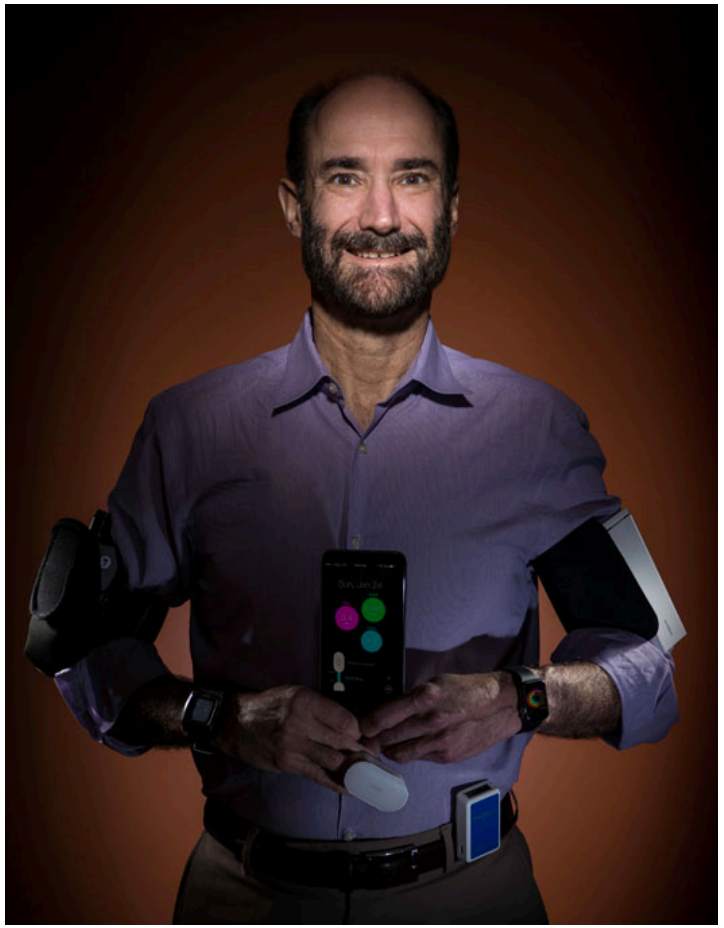


Ancestry



Expanding personalized medicine beyond the genome.

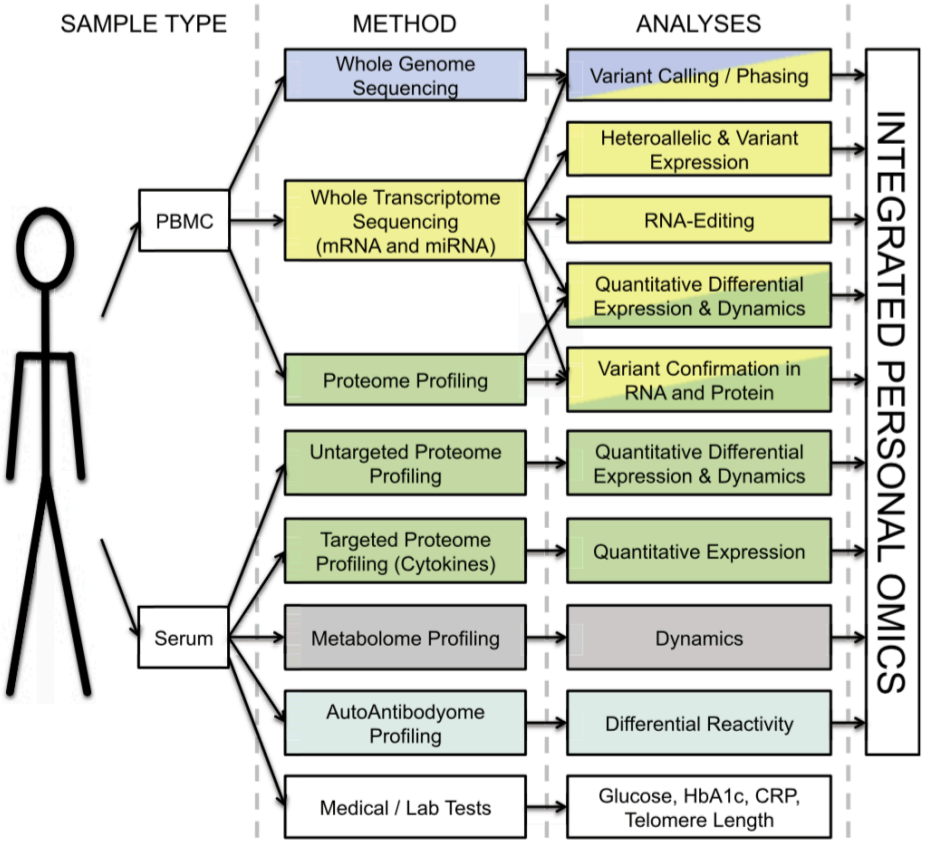
- An integrated personal omics profile (iPOP) is an example of a more comprehensive version of personalized medicine.
- Michael Snyder had his genome sequenced and collected many other large scale datasets over an extended period of time.



https://med.stanford.edu/news/all-news/2017/01/wearable-sensors-can-tell-when-you-are-getting-sick.html

Integrated personal omics profile (iPOP)

- Numerous types of data were collected, primarily from blood samples. The datasets include:
  - Transcriptomic
  - Proteomic
  - Metabolomic
  - Cytokine profiling
  - Autoantibody profiling
  - Medical exams



Chen et al. Cell 2012