Next Generation Sequencing (NGS) What and Why?

Gil Stelzer and Noa Wigoda

5.11.19

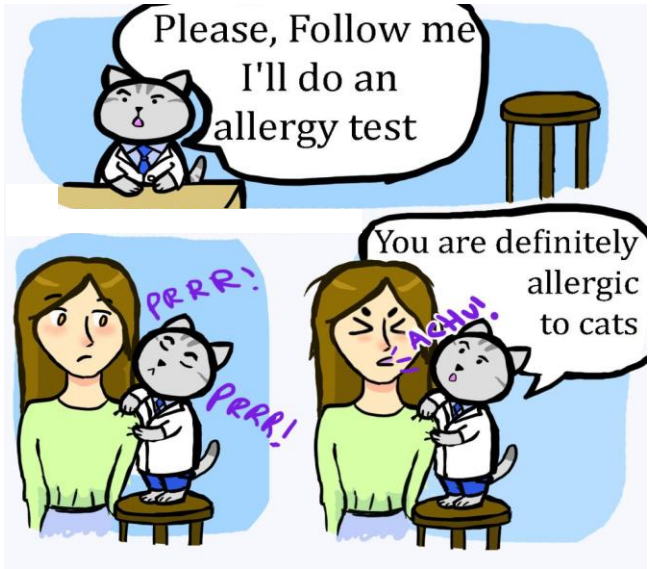
An Introduction to deep-sequencing analysis for biologists

Outline

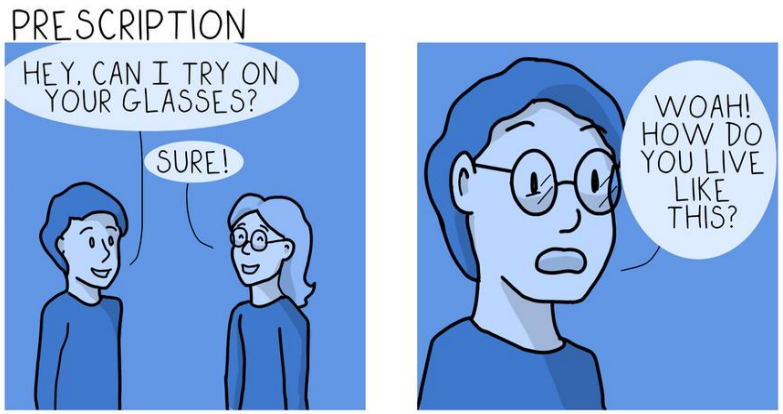
- Vocabulary
- History
- Biology

Vocabulary

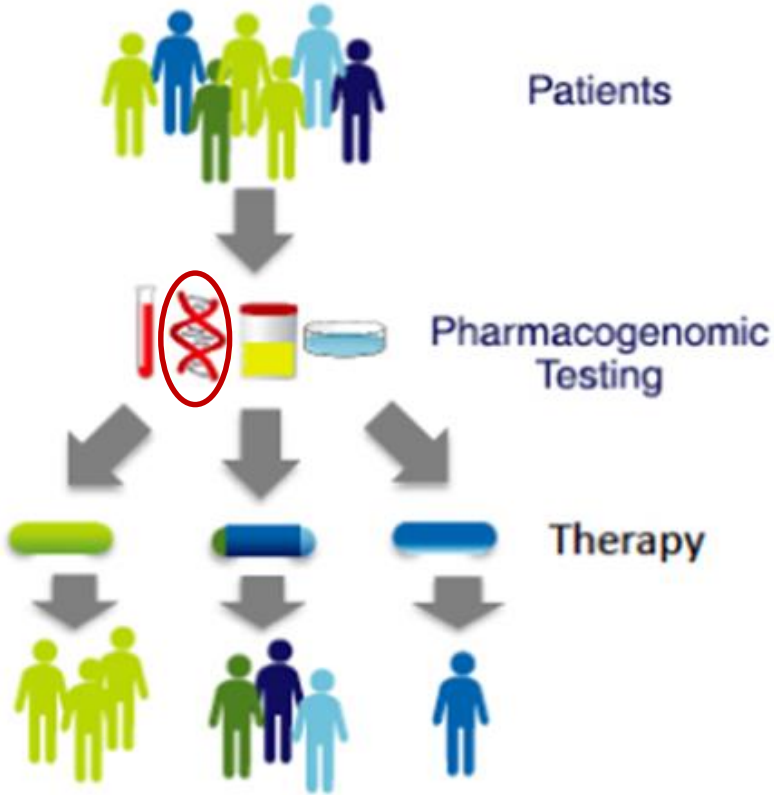
https://pharma.elsevier.com/pharma-rd/personalized-vs-precision-medicine/



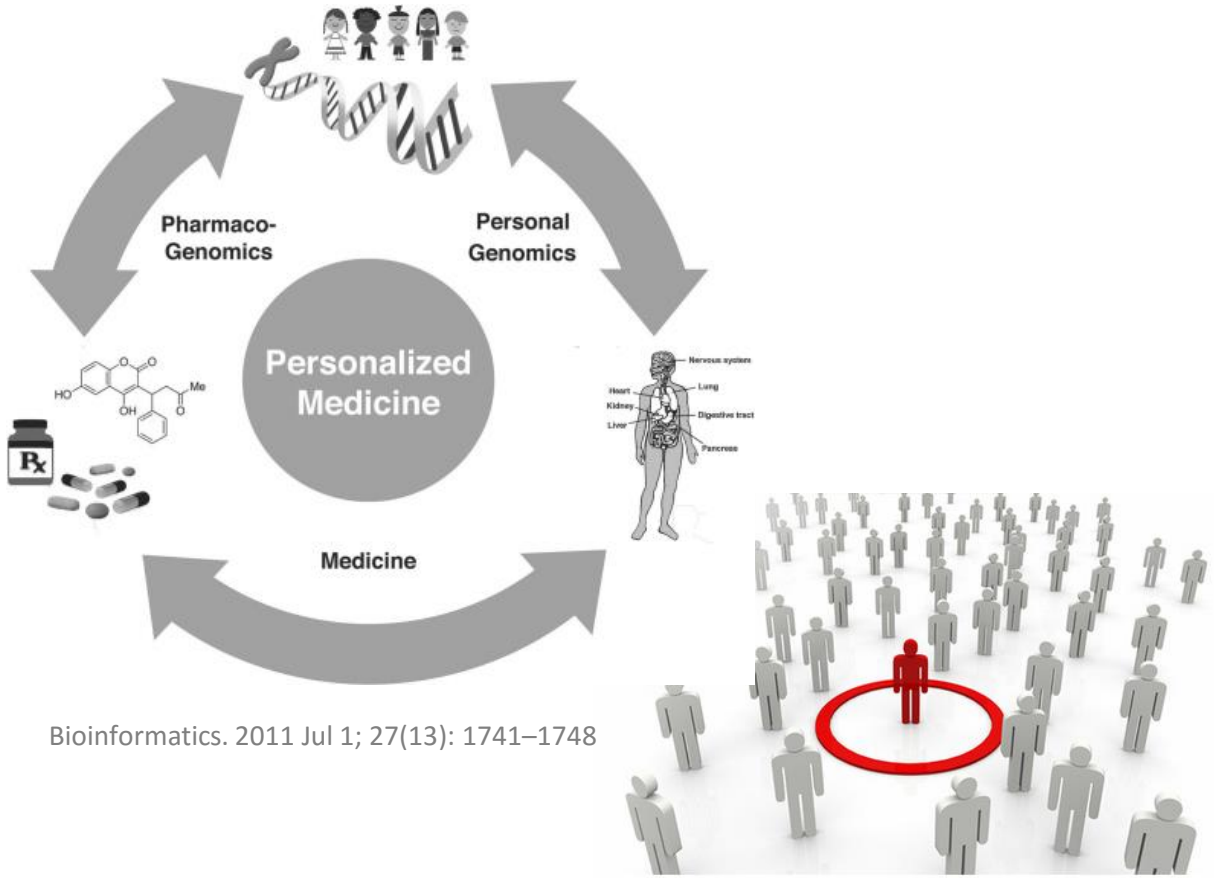
https://www.deviantart.com/guly09/art/Dr-Cat-allergy-test-756983251



https://www.reddit.com/r/comics/comments/a58y8g/prescription_oc/

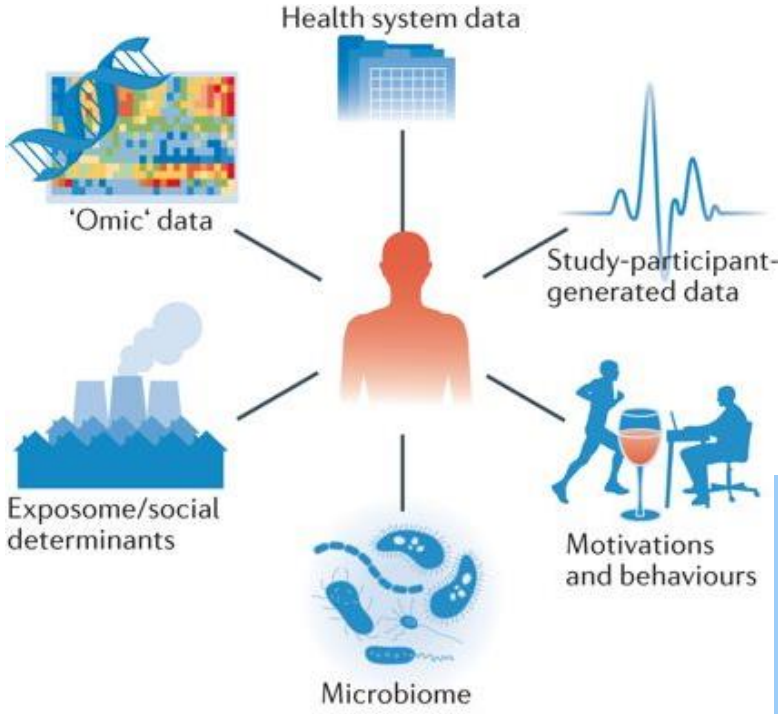


- Personalized

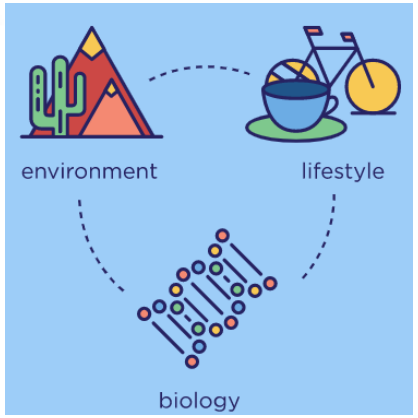


Bioinformatics. 2011 Jul 1; 27(13): 1741–1748

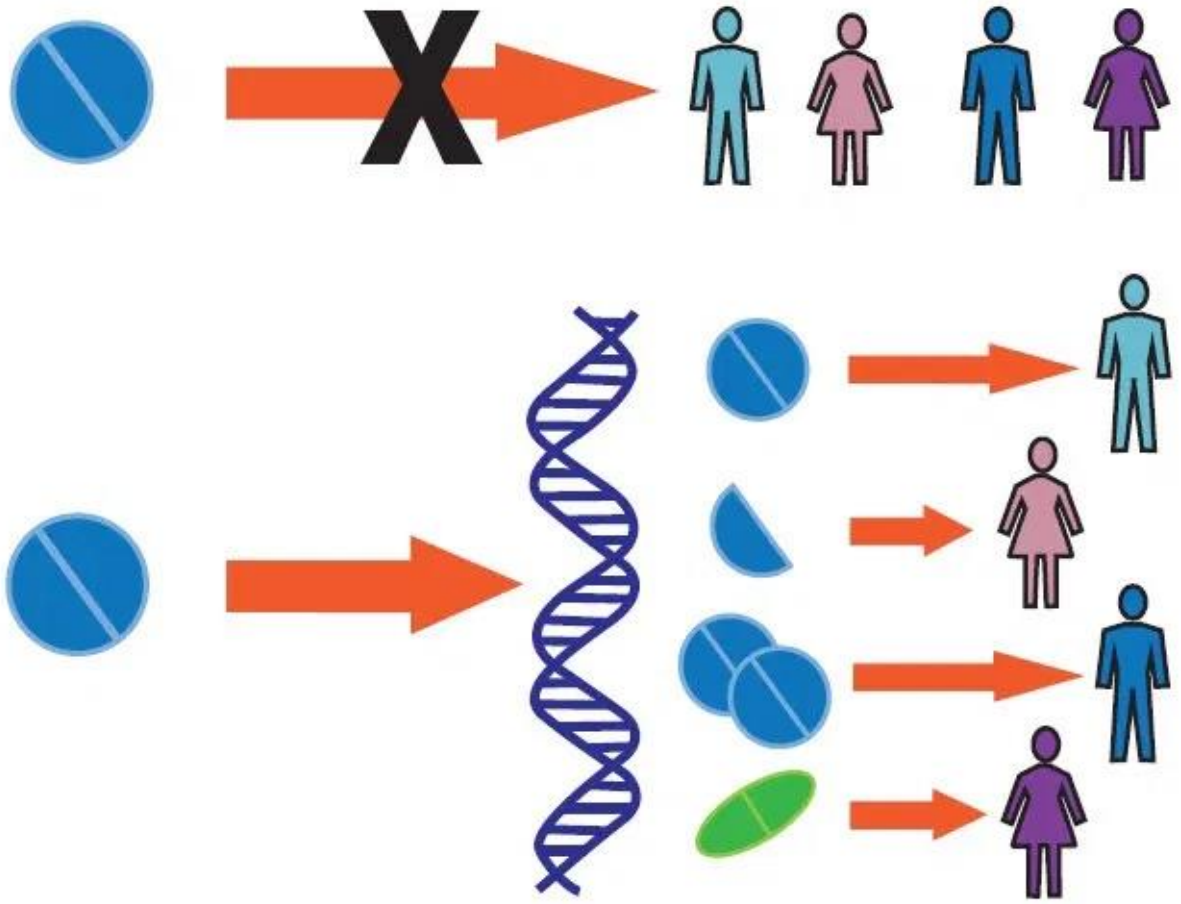
- Precision



Nature Reviews Cardiology volume13, pages591–602 (2016)



The disease mechanism in an individual patient must be diagnosed first, before selecting a therapy, regardless of whether we call this approach personalized or precision medicine



https://rnsights.com/the-push-for-personalized-medicine/

Medical histories

Primitive medicine

Humans in the Stone Age identified edible plants and plants that seem to cure or soothe ailments.

Ancient medicine:

Indian medicine

Chinese medicine

Egyptian medicine

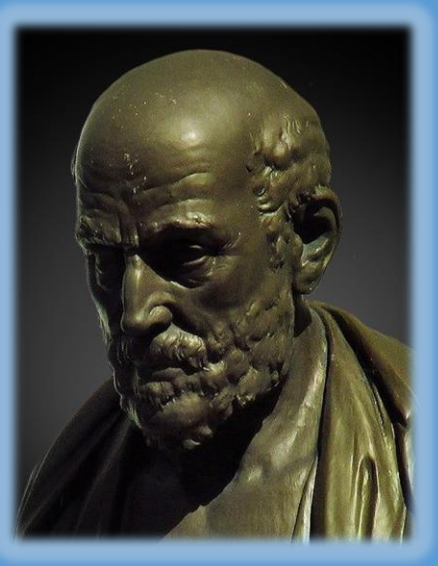
Greek medicine

The first medical interventions were often individualized but ineffective, because doctors lacked an understanding of disease biology.

Age old favorite home remedy ingredients; garlic, ginger, onion, peppers, and herbs.

Amber Dance Nature volume 537, pages S52–S53 (2016)

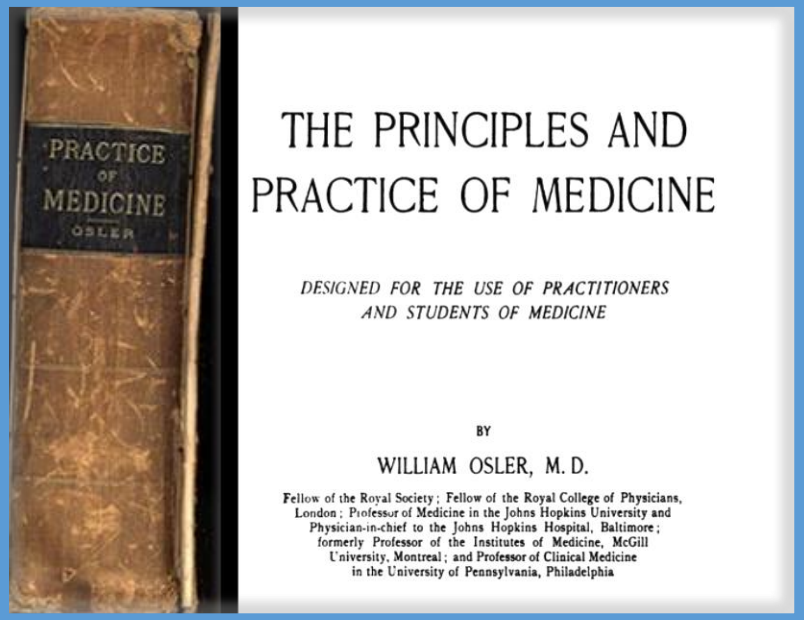
“It’s far more important to know what person the disease has than what disease the person has.”



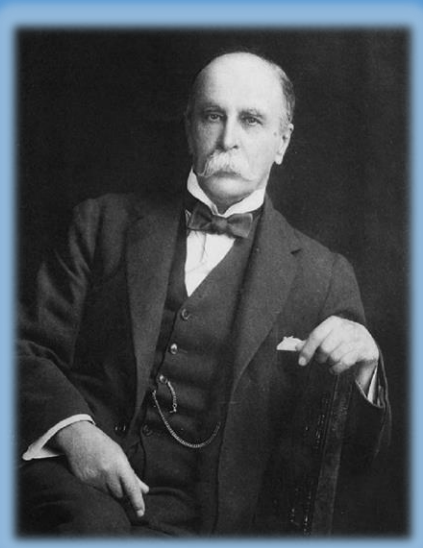
Hippocrates (c. 370 BC)

The notion that individuals presenting with the same signs and symptoms may require different treatment is not new….

"If it were not for the great variability among individuals, medicine might as well be a science and not an art."



Reflecting the lack of objective data available to make decisions that are tailored to individual patients



Sir William Osler, 1892

Johns Hopkins School of Medicine

Human genome

- 3 billion nucleotides (~3×10^9 base pairs) within the 23 chromosomes
- about 1.5% (~30 Mb) of protein-coding sequences, codes for ~20,000 genes
- 6770 genes are associated to disease-causing mutations (DM) or probable/possible pathological mutation (DM?) variants, according to the Human Gene Mutation Database (HGMD®).

It is estimated that 85% of the disease-causing mutations are located in coding regions of the genome.

Stenson, P.D., Mort, M., Ball, E.V. et al. Hum Genet (2017) 136: 665.
https://doi.org/10.1007/s00439-017-1779-6

NIH National Human Genome Research Institute

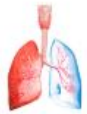
Begin your search here

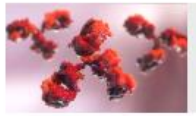
About Genomics | Research Funding | Research at NHGRI | Health | Careers & Training | News & Events | About NHGRI

List of Genetic Disorders

This list of genetic, orphan and rare diseases is provided for informational purposes only and is by no means comprehensive.

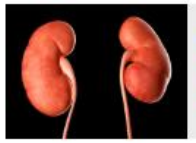

About Achondroplasia ›


About Alpha-1 Antitrypsin Deficiency ›


About Antiphospholipid Syndrome ›

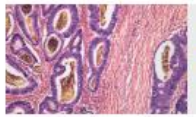

About Attention Deficit Hyperactivity Disorder ›


About Autism ›


About Autosomal Dominant Polycystic Kidney Disease ›


About Breast Cancer ›

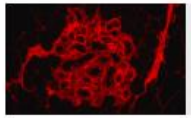

About Charcot-Marie-Tooth Disease ›

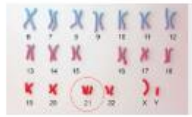

About Colon Cancer ›


About Cri du Chat Syndrome ›

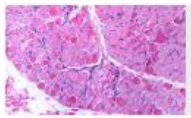

About Crohn's Disease ›

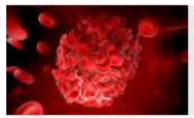

About Cystic Fibrosis ›


About Dercum Disease ›


About Down Syndrome ›


About Duane Syndrome ›

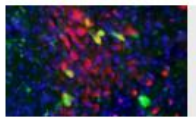

About Duchenne Muscular Dystrophy ›


About Factor V Leiden Thrombophilia ›


About Familial Hypercholesterolemia ›


About Familial Mediterranean Fever ›


About Fragile X Syndrome ›


About Gaucher Disease ›

https://www.genome.gov/For-Patients-and-Families/Genetic-Disorders

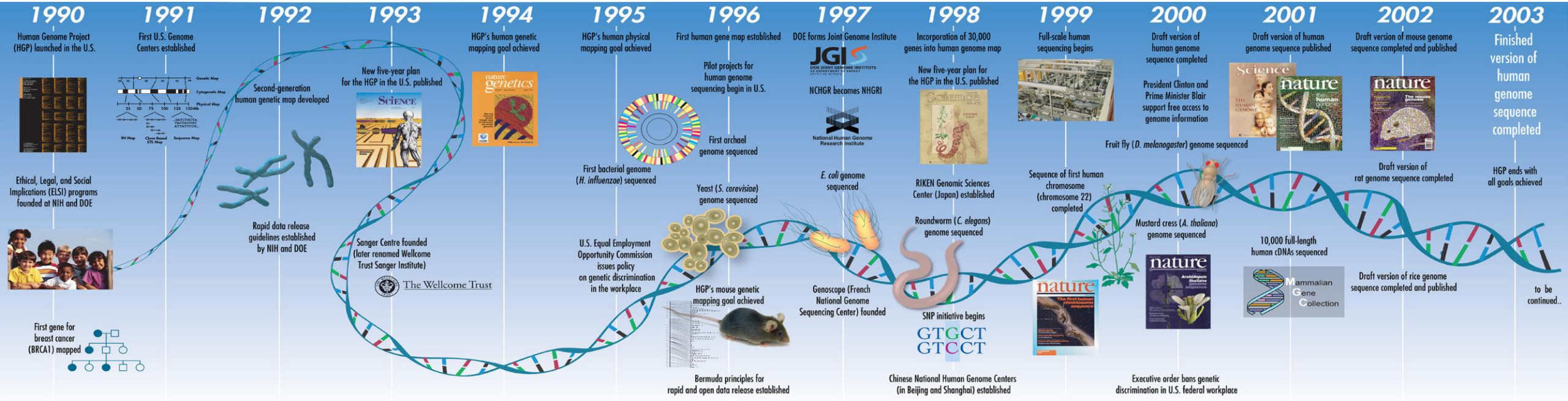
Human Genome Project

The most famous sequencing project of the Human genome,

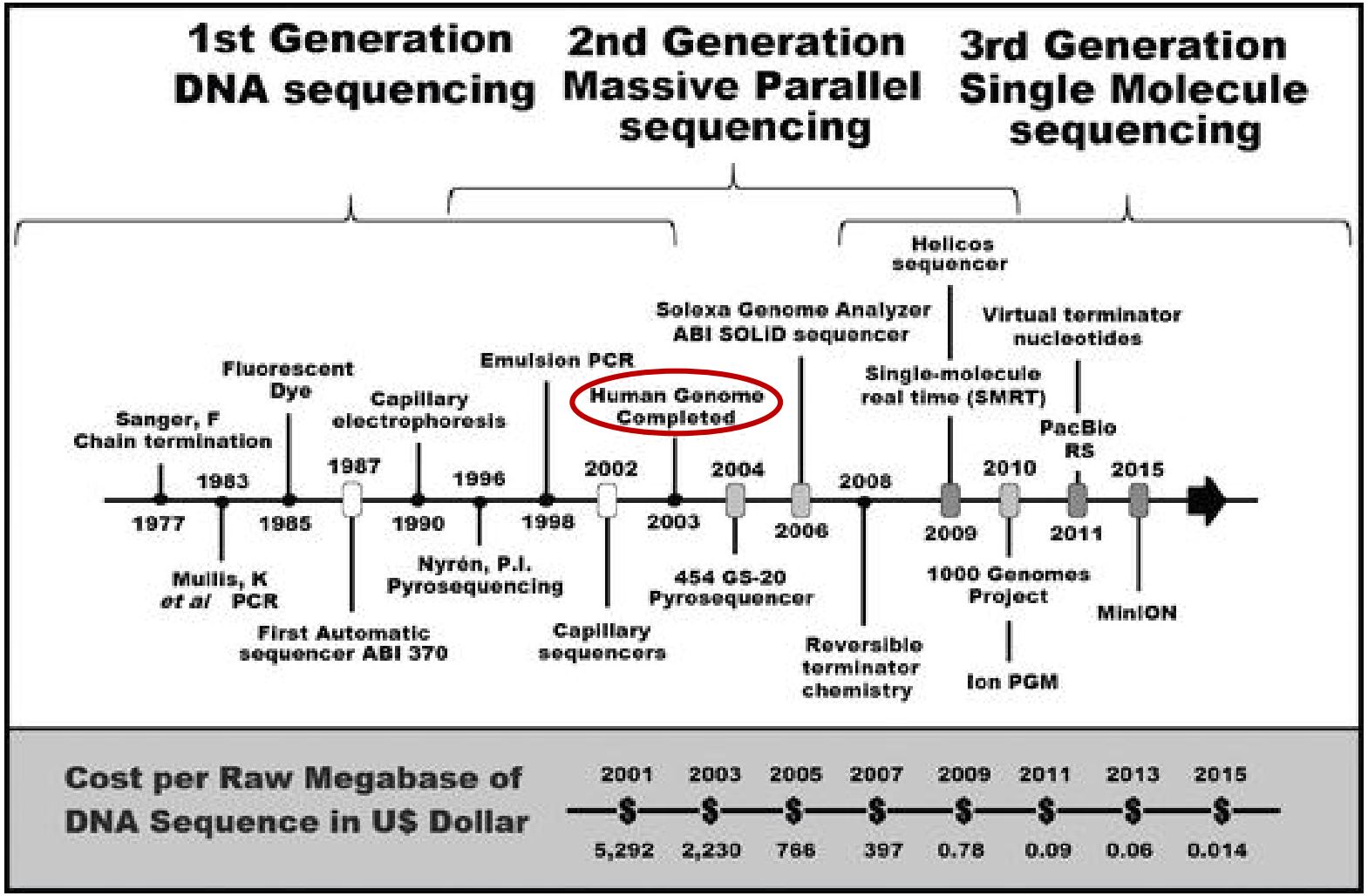
an international collaborative research effort

produced in 13 years 3 billion sequenced bases

with an estimated cost of ~ $2.7 billion



Timeline of DNA sequencing Methods

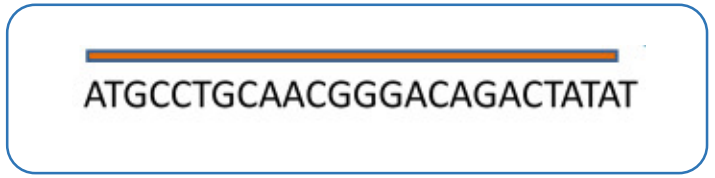


https://www.intechopen.com/media/chapter/55706/media/F1.png

Generations of DNA sequencing Methods

First generation

1977 Sanger developed the method called ‘Chain-termination’.
To date, Sanger is still the gold-standard method in diagnostic tests.



Second generation

The era of the parallel massive sequencing on a micro scale.
called: NGS – Next Generation Sequencing.
Dramatically increased the amount of DNA sequenced.
High-throughput short-read sequencing.

cost-effective and efficient

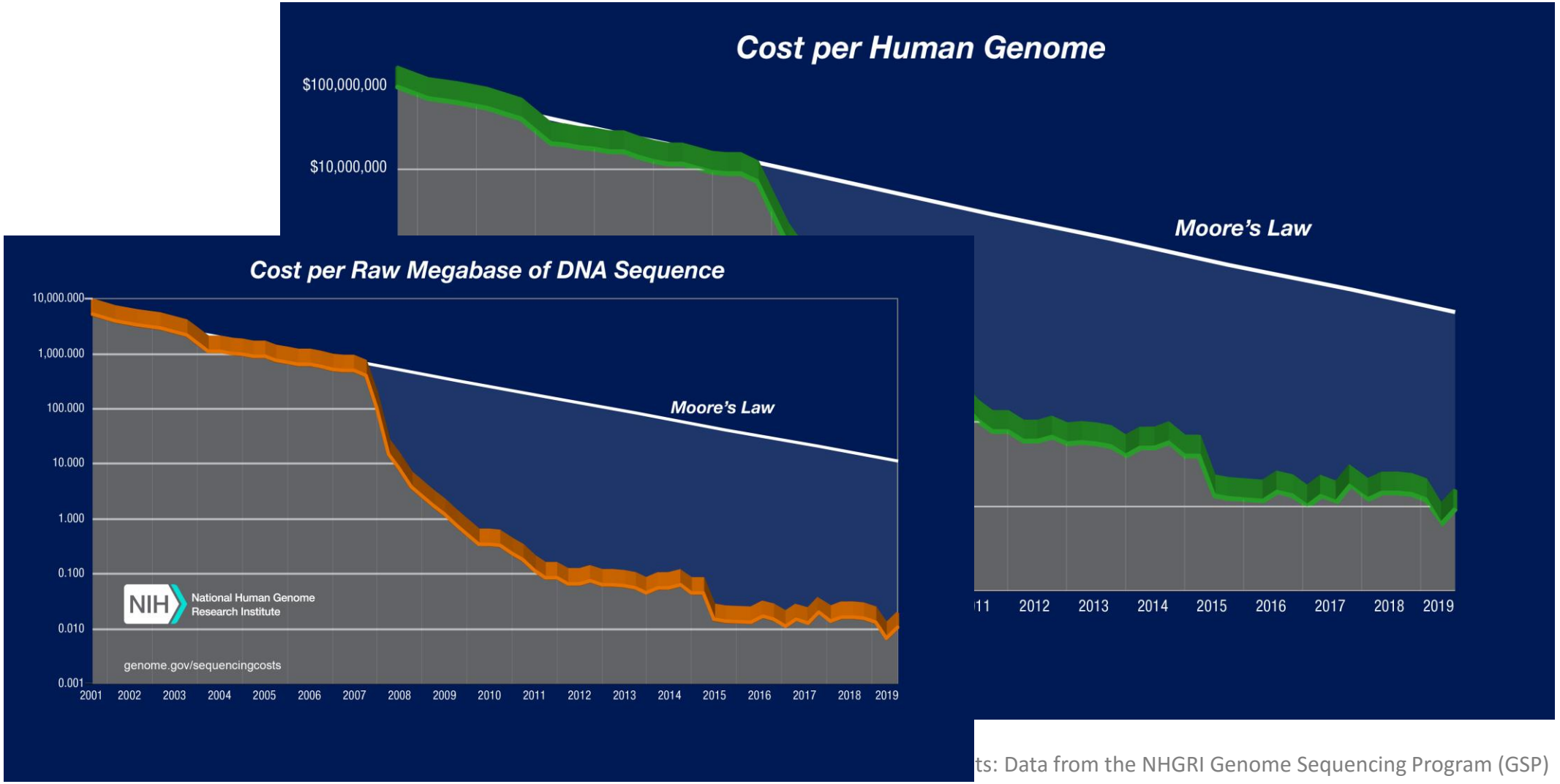
Third generation

Technology of single-molecule sequencing (SMS), which has no need to amplify the DNA.

- Pacific Bioscience - ‘single molecule real time’ (SMRT), very long reads.
- Oxford Nanopore Technologies -reads are incredibly long (500 kb), process is extremely fast, no need for special nucleotides.

https://www.intechopen.com/books/applications-of-rna-seq-and-omics-strategies-from-microorganisms-to-human-health/application-of-next-generation-sequencing-in-the-era-of-precision-medicine#B41

Timeline of DNA sequencing costs



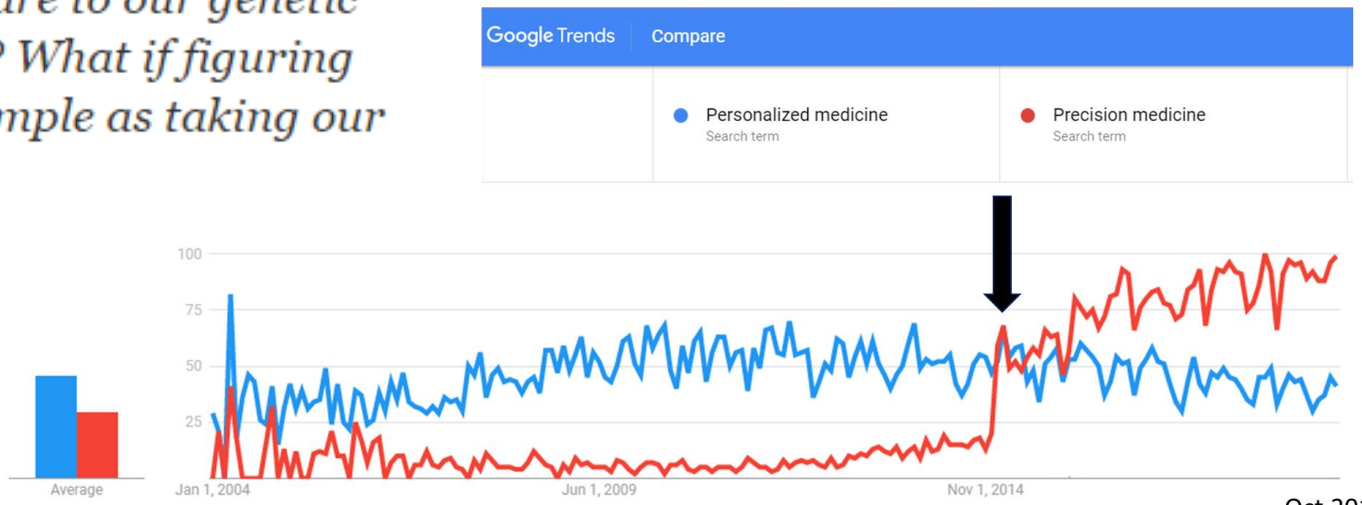
ts: Data from the NHGRI Genome Sequencing Program (GSP)

https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data

FACT SHEET: Obama Administration Announces Key Actions to Accelerate Precision Medicine Initiative

> *Doctors have always recognized that every patient is unique, and doctors have always tried to tailor their treatments as best they can to individuals. You can match a blood transfusion to a blood type — that was an important discovery. What if matching a cancer cure to our genetic code was just as easy, just as standard? What if figuring out the right dose of medicine was as simple as taking our temperature?*
>
> - President Obama, January 30, 2015

A $215-million project to study
the genomes and health status of
1 million volunteers
and develop the required
databases and privacy standards

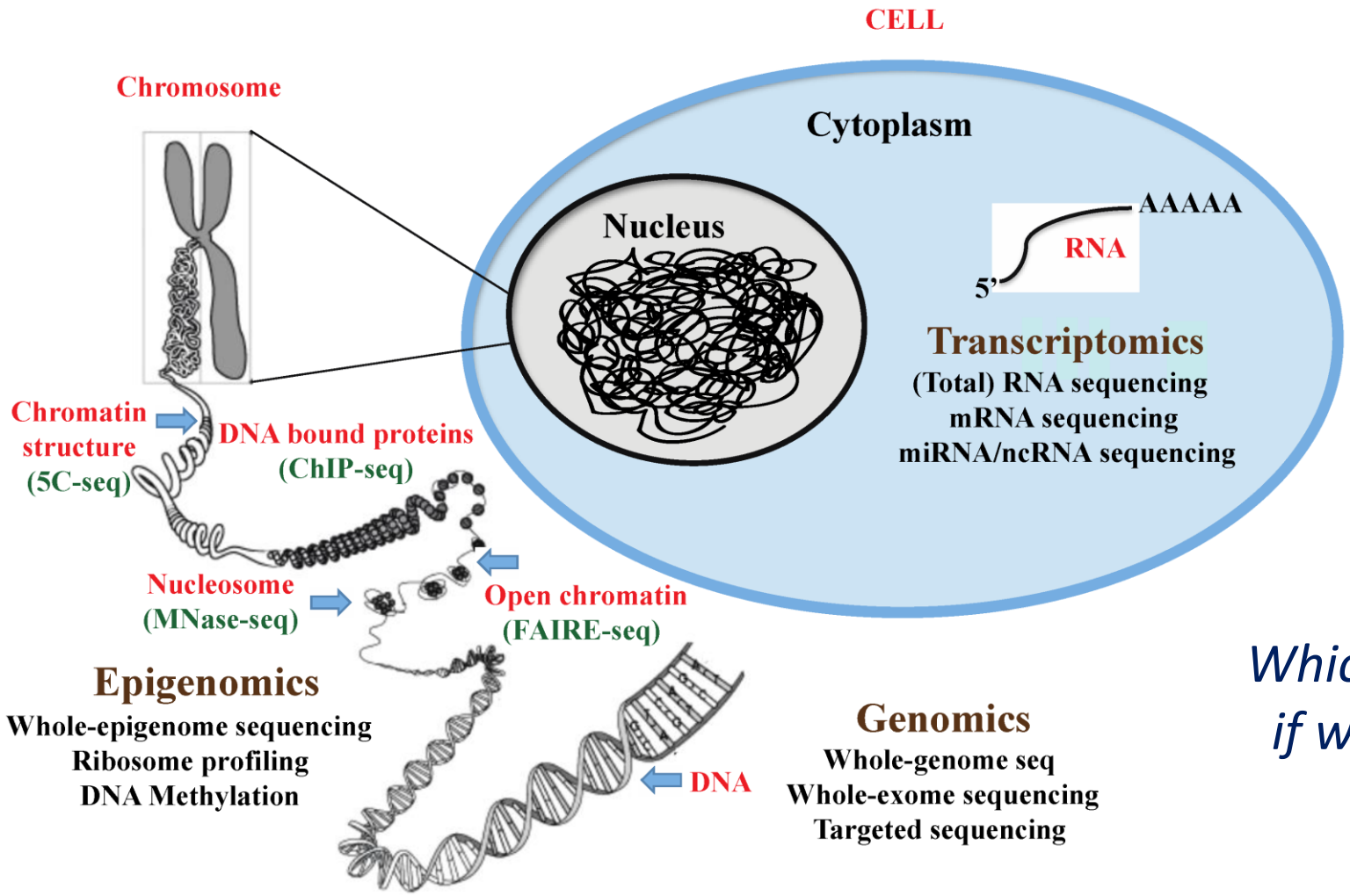


Oct,2019

Precision Medicine
One size does not fit all

https://www.youtube.com/watch?v=ljmPtJhyobk

Biology….
Applications of next-generation sequencing



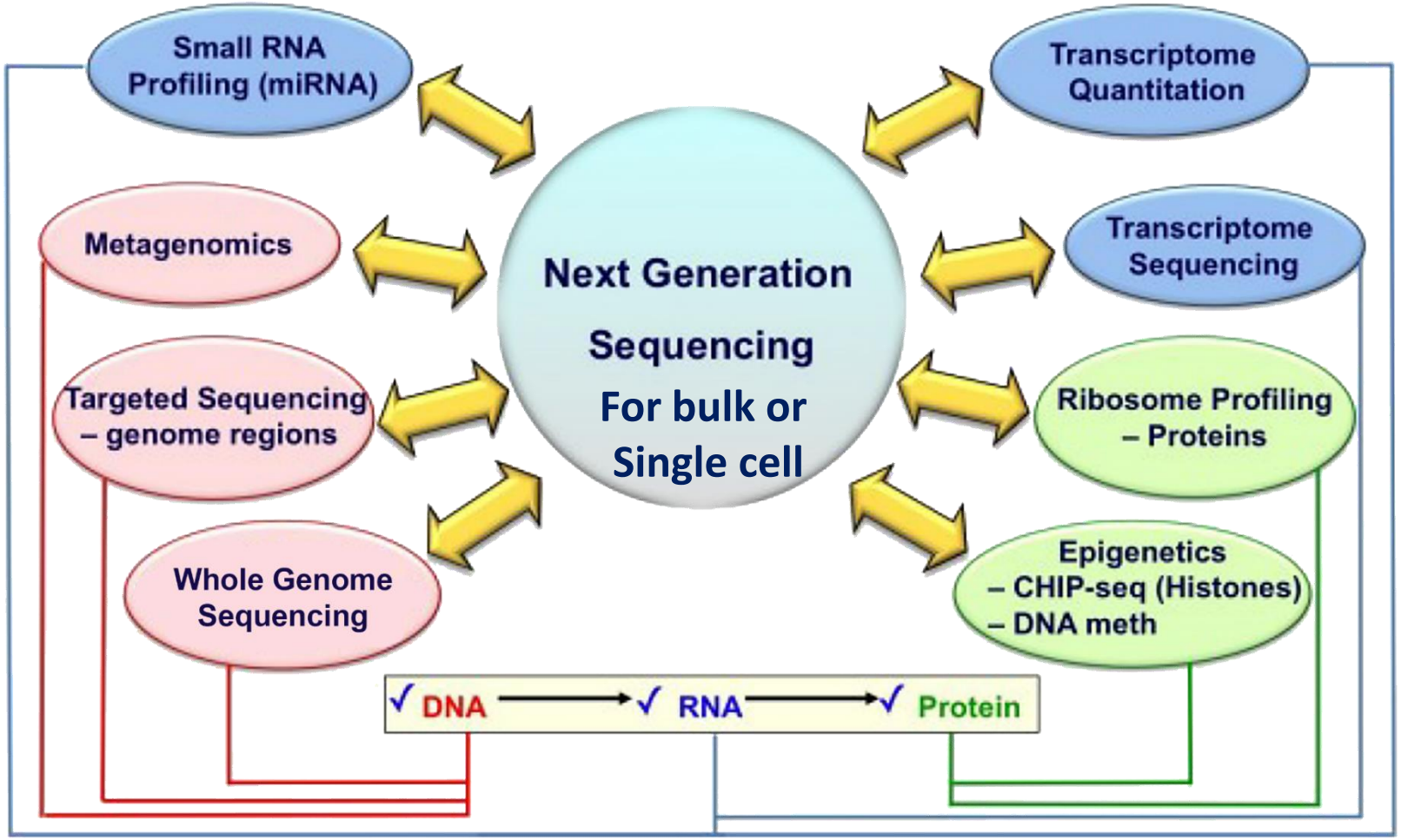
Know the biological origin of your sample - it matters!

Which application will we choose if we want to check differential gene expression?

Strahan R, Uppal T, Verma SC. Next-Generation Sequencing in the Understanding of Kaposi's Sarcoma-Associated Herpesvirus (KSHV) Biology. Viruses. 2016; 8(4):92.

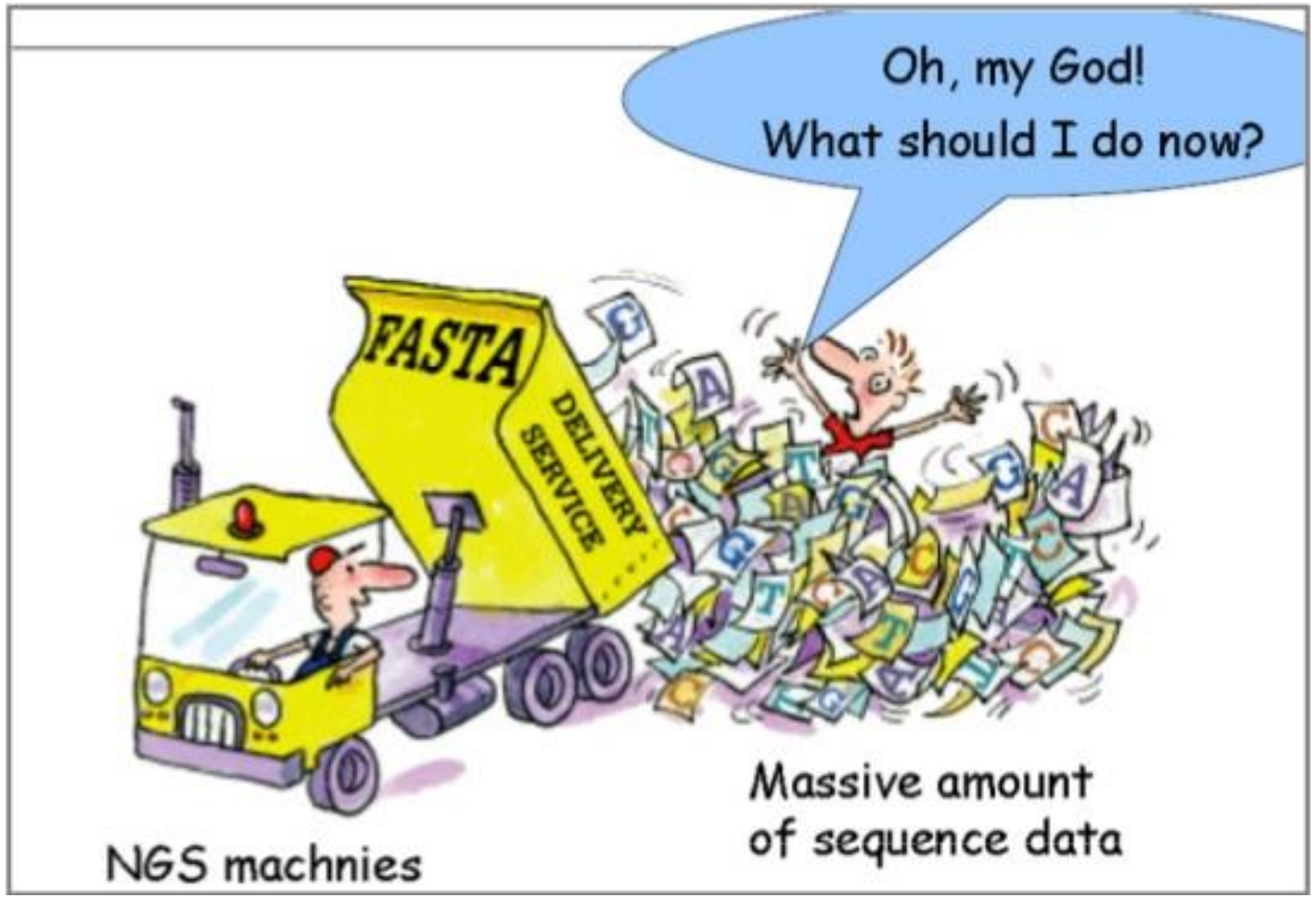
Applications of next-generation sequencing

Which application will we choose if we want to check whether we are carriers of a genetic disease?



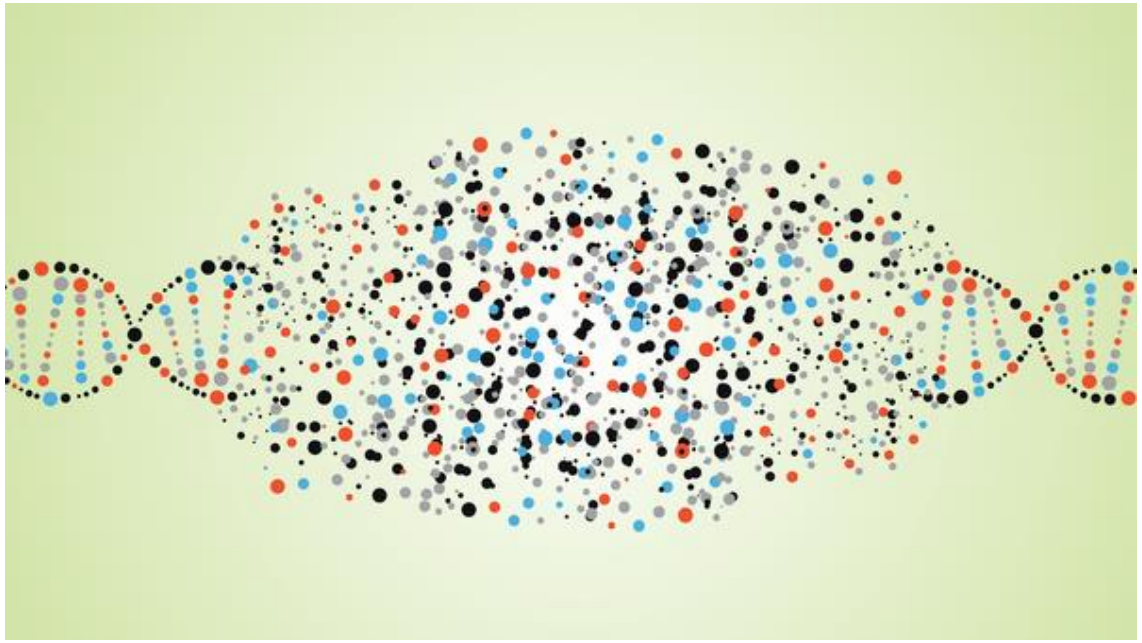
https://www.europeanpharmaceuticalreview.com/article/10409/dna-sequencing-technologies-and-emerging-applications-in-drug-discovery/

Output of next-generation sequencing



As high-throughput analytical tools improve,
allowing researchers to collect more and more data,
the challenge becomes how to interpret it all.

Next-generation sequencing: The genome jigsaw



"To understand why high-throughput gene-sequencing technology often produces frustrating results", says Titus Brown, "imagine that 1,000 copies of Charles Dickens' novel A Tale of Two Cities have been shredded in a wood chipper. Your job is to put them back together into a single book," he says.

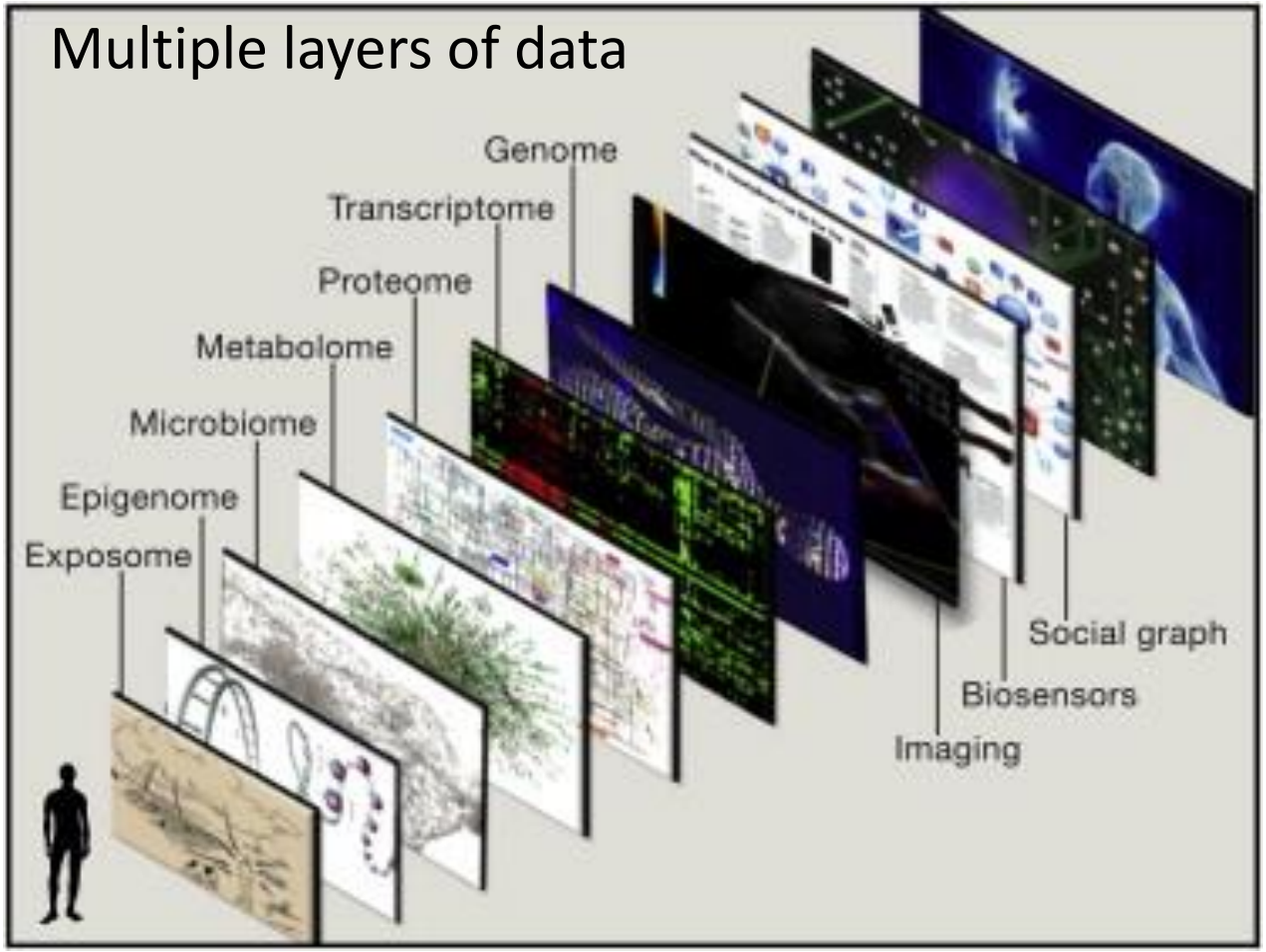
http://www.nature.com/nature/journal/v501/n7466/full/501261a.html

Precision medicine

Precision medicine is based on the individual characteristics of each patient (genetic, environmental, behavioral)

Aims to optimize and customize strategies for prevention, detection and therapy of disease.

- Hereditary rare monogenic diseases
- Metabolic diseases (Obesity, Diabetics, …)
- Infectious diseases (Pathogen tracking)
- Cancer
- Pharmacogenomics (Drug prescription)



http://www.sciencedirect.com/science/article/pii/S0092867414002049#

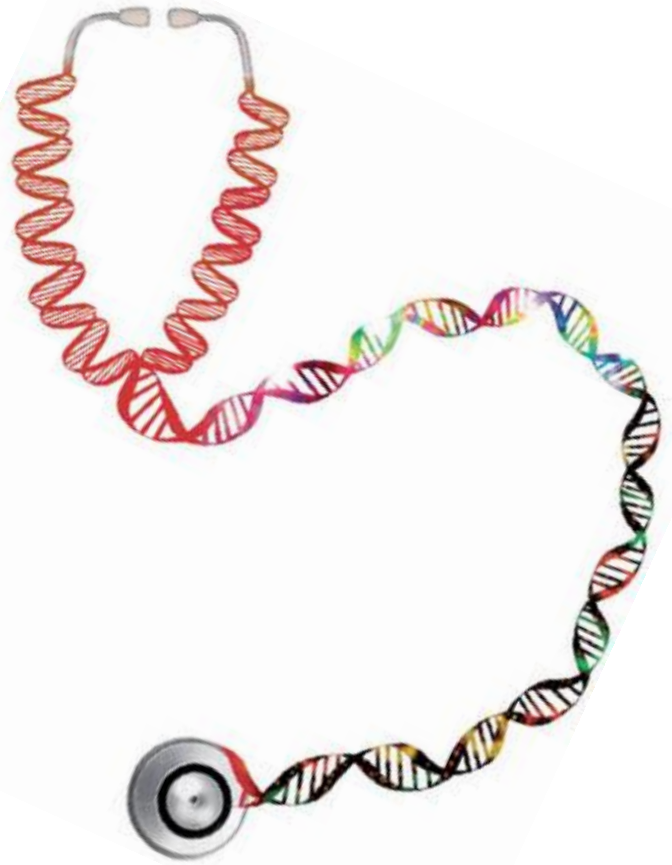
NGS in precision medicine

Changes in DNA sequence could cause genetic disease

It has become easier to sequence patient samples:

- Expressed genes (‘transcriptomes’)
- known exons (‘exomes’)
- complete genomes

NGS allows identifying biomarkers for early diagnosis



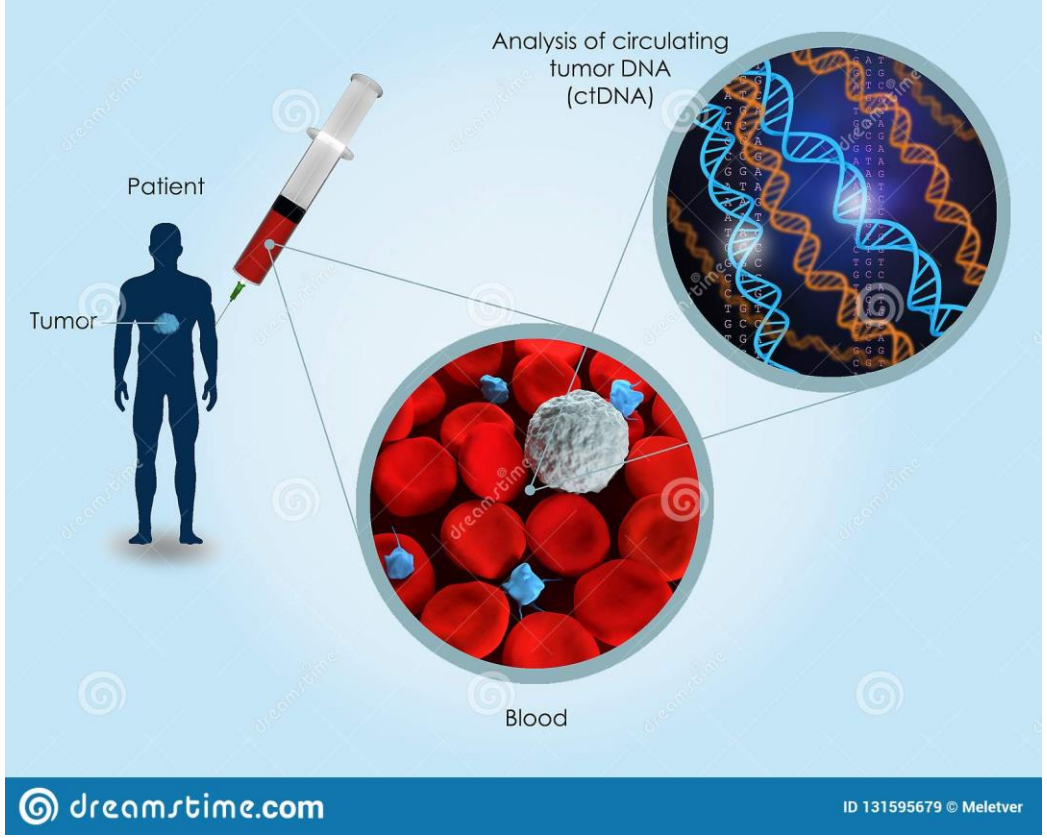
https://www.nature.com/naturejobs/2011/110519/images/nj0342-i1.jpg

Examples of medical applications

- Cancer screening

- Non-Invasive Prenatal Testing

Cancer, example 1-
Circulating tumor DNA



Blood contains two types of cancer-derived materials that are susceptible to detailed molecular analysis:

- Intact circulating tumor cells (CTC)

  extremely rare: ~ <10 cells/mL

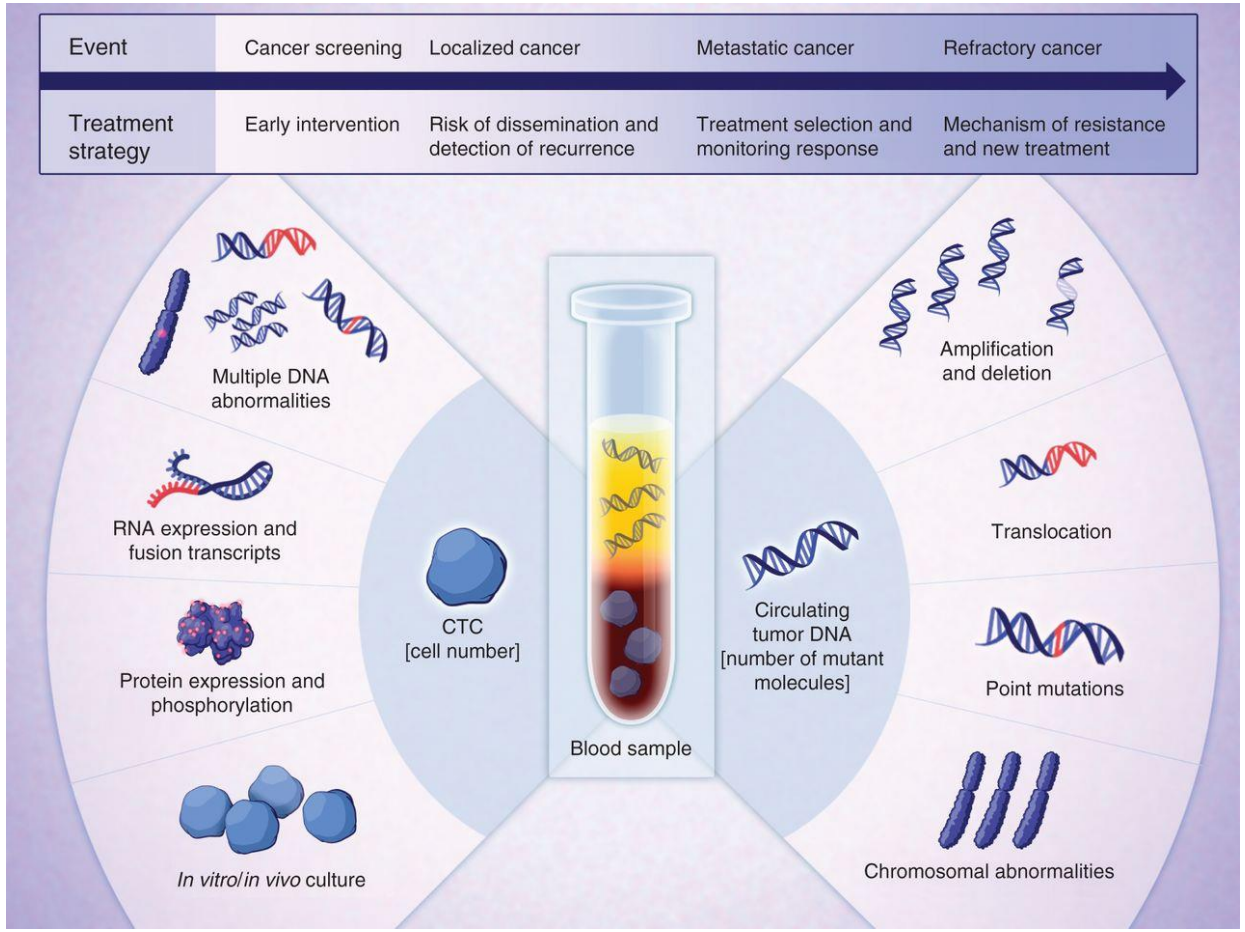
- Cell-free circulating tumor DNA (ctDNA)

  <0.1% to >10% of cfDNA (cell free FNA)

  Healthy individuals have less than 25 ng cfDNA per mL.

Blood-Based Analyses of Cancer: Circulating Tumor Cells and Circulating Tumor DNA
Daniel A. Haber and Victor E. Velculescu

Blood-Based Analyses of Cancer: Circulating Tumor Cells and Circulating Tumor DNA



Relative levels of ctDNA, within an individual,
correlate with tumor burden and response to therapy.

Goal:
Use liquid biopsy to diagnose cancer, at a stage in which it may be curable.

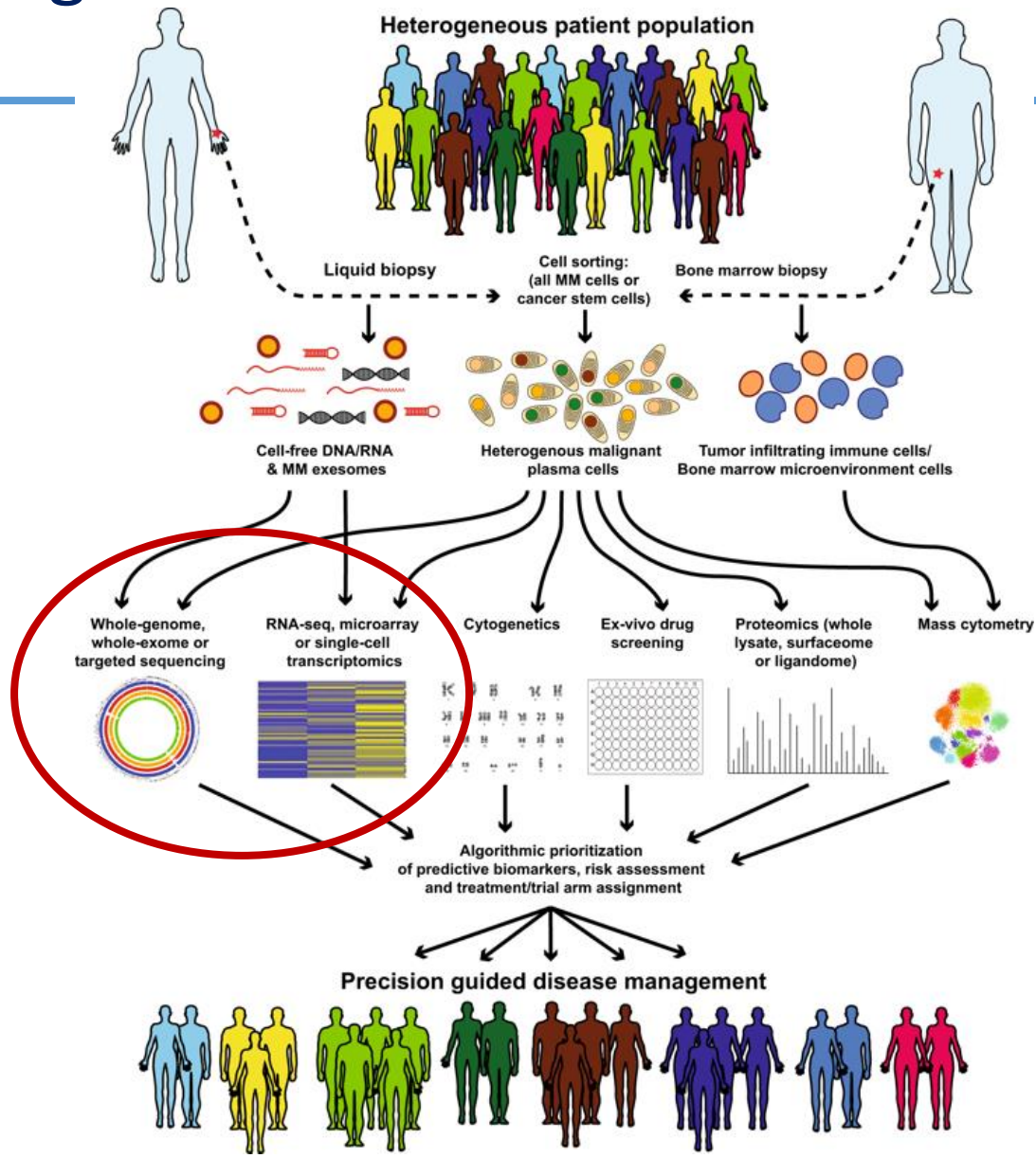
Blood-Based Analyses of Cancer: Circulating Tumor Cells and Circulating Tumor DNA
Daniel A. Haber and Victor E. Velculescu

The future of myeloma precision medicine: integrating the compendium of known drug resistance mechanisms with emerging tumor profiling technologies

Goal

Identify biomarkers that predict efficacy or resistance within an individual's sub-clonally heterogeneous tumor.

NGS applications are part of the effort to identify biomarkers.



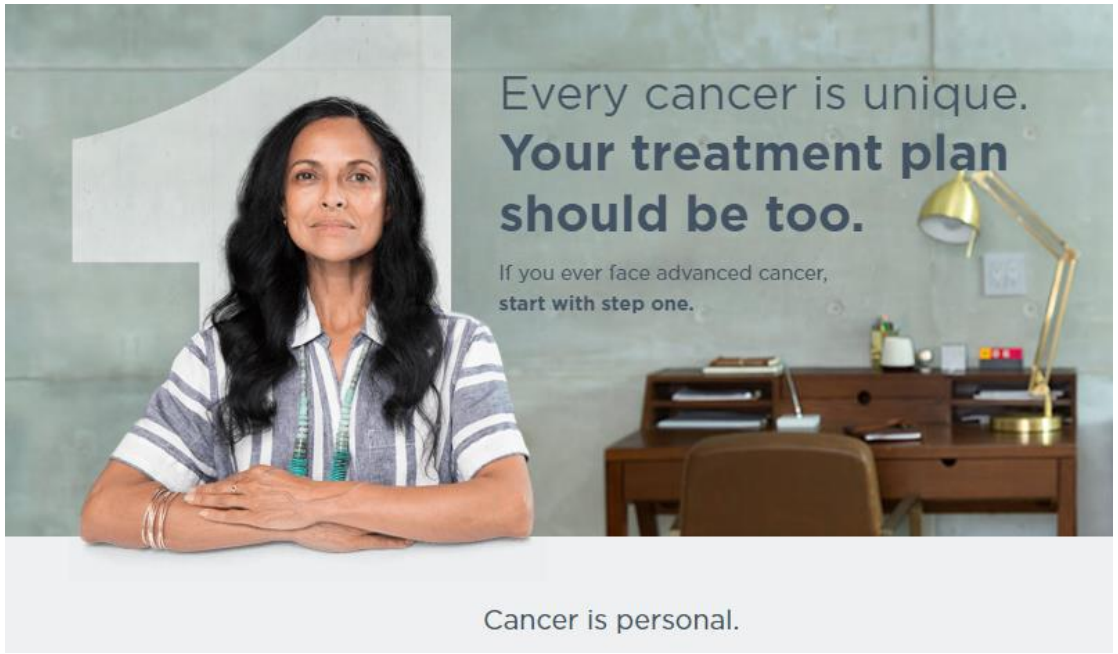
Harding et al., Leukemia volume 33, pages 863–883 (2019)

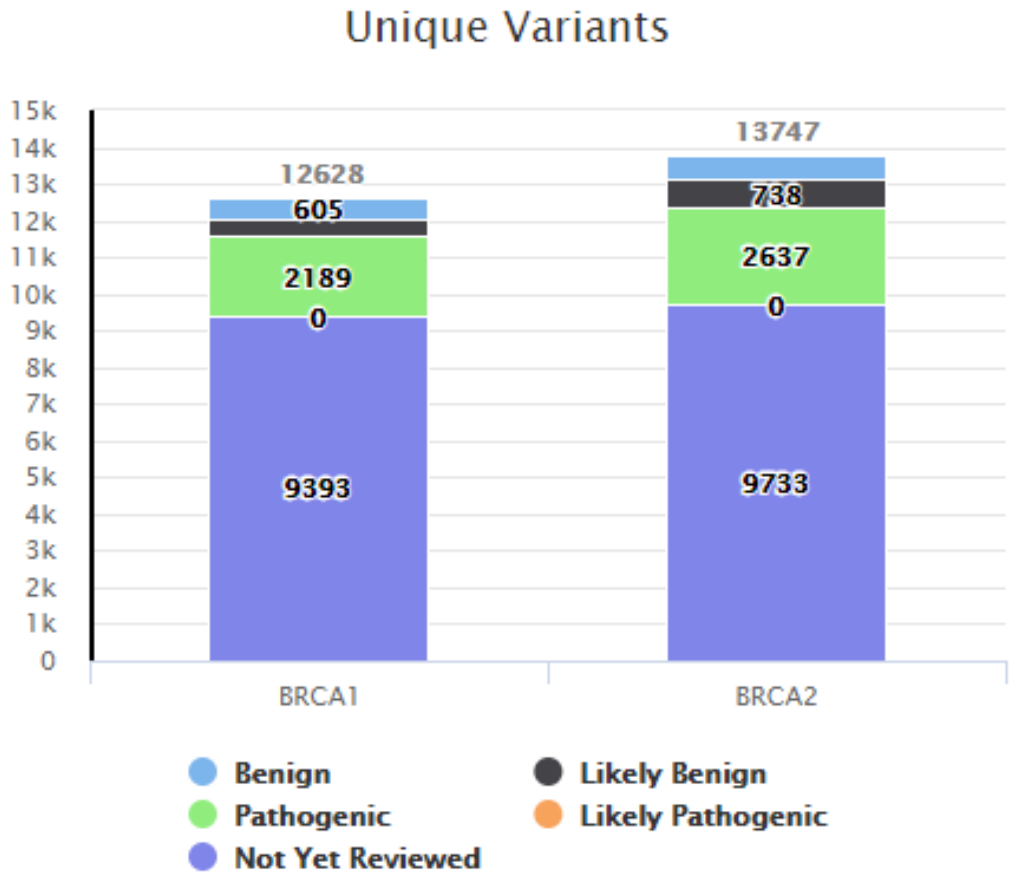
Cancer, example 2-

The BRCA1 and BRCA2 genes typically help to prevent cancer in humans.

Germline BRCA1/2 status is a critical biomarker to help determine the appropriate therapy for patients with metastatic breast or ovarian cancer.

BRACAnalysis CDx®





Risk of developing cancer over a lifetime for women with a pathogenic BRCA variant:

- BRCA1:
  Breast cancer (by age 70) - up to 55% chance
  Ovarian cancer (in their lifetime)- up to 39% chance

- BRCA2:
  Breast cancer (by age 70) - up to 45% chance
  Ovarian cancer (in their lifetime)- up to 17% chance

https://brcaexchange.org/factsheet

https://www.ncbi.nlm.nih.gov/clinvar/?term=brca2%5Bgene%5D

Fam Cancer. 2017 Jul;16(3):311-317. doi: 10.1007/s10689-016-9955-8.

Detection of false positive mutations in BRCA gene by next generation sequencing.

Suryavanshi M[1], Kumar D[2], Panigrahi MK[2], Chowdhary M[2], Mehta A[3].

Next-generation sequencing shows great promise by allowing rapid mutational analysis of multiple genes in human cancer

but

our results indicate the need for careful sequence analysis to avoid false positive results.

Sanger Sequencing did not confirm these insertions.

The boy who cried wolf……

False positive - A test result that indicates that a person has a specific disease or condition when the person actually does not have the disease or condition.

False negative - A test result that indicates that a person does not have a specific disease or condition when the person actually has the disease or condition.

BRCA-related risk assessment for breast cancer

I made major decisions
— like having a double mastectomy —
based on a false positive.

I was robbed of the chance to
breastfeed my babies,
and it broke my heart.

Maureen Boesen
proud to be BRCA-positive and a previvor.

https://www.huffpost.com/entry/brca-genetic-testing-mastectomy_n_5c6c39fbe4b012225acd80d3

BRCA-related risk assessment for breast cancer

Potential health consequences could result from false positive or false negative assessments.

- False positive
  - undergo prophylactic surgery
  - chemoprevention
  - intensive screening
- False negative
  - failure to recognize an existing risk that may end the patient's life….

Maureen Boesen

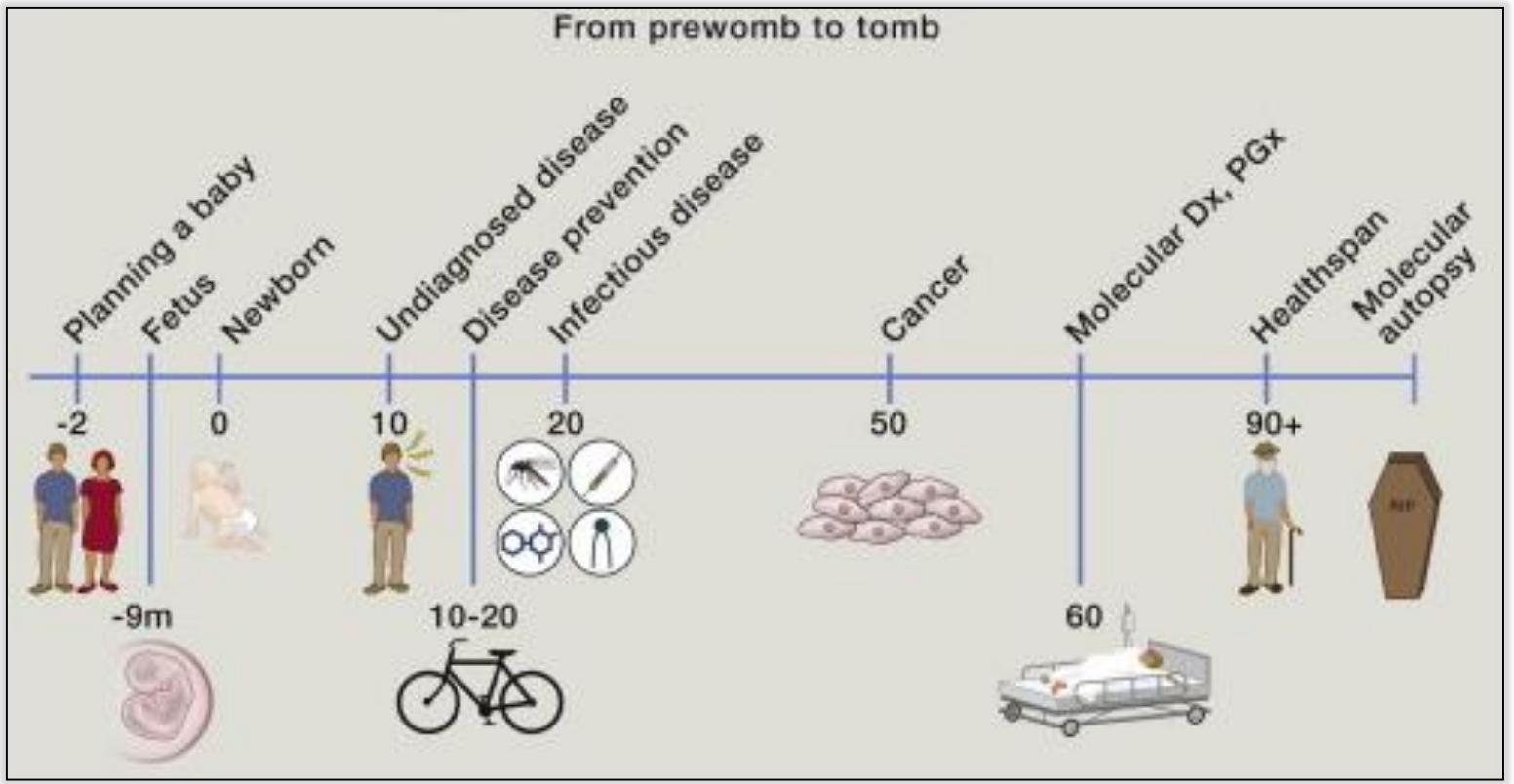
Doctors can be wrong and so can tests.

It is important to validate the results of wet-lab pathogenicity data and of genomic data.

All NGS tools are still prone to sequencing artefacts!

Sanger sequencing is recommended to confirm detected variants.

Timeline of sequencing applications in medicine from prewomb to tomb



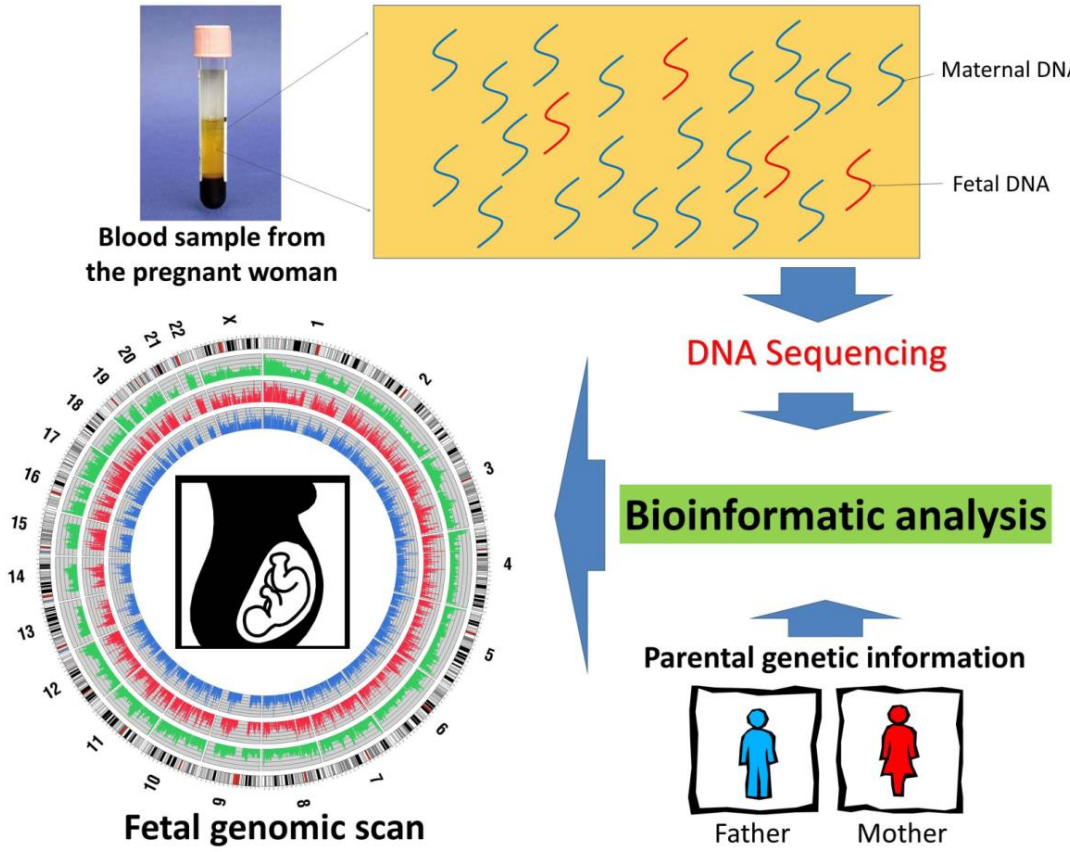
http://www.sciencedirect.com/science/article/pii/S0092867414002049#

In the future, omics information collected at a young age will be used to identify susceptibility to various medical conditions that have actionable prevention strategies.

Defining the genomics of health-span, rather than the traditional focus on diseases, may prove to be especially worthwhile.

Noninvasive Prenatal Diagnosis

During pregnancy, a median of 10% of the DNA in the blood plasma of a pregnant women— called ‘cell-free' DNA — comes from the fetus.

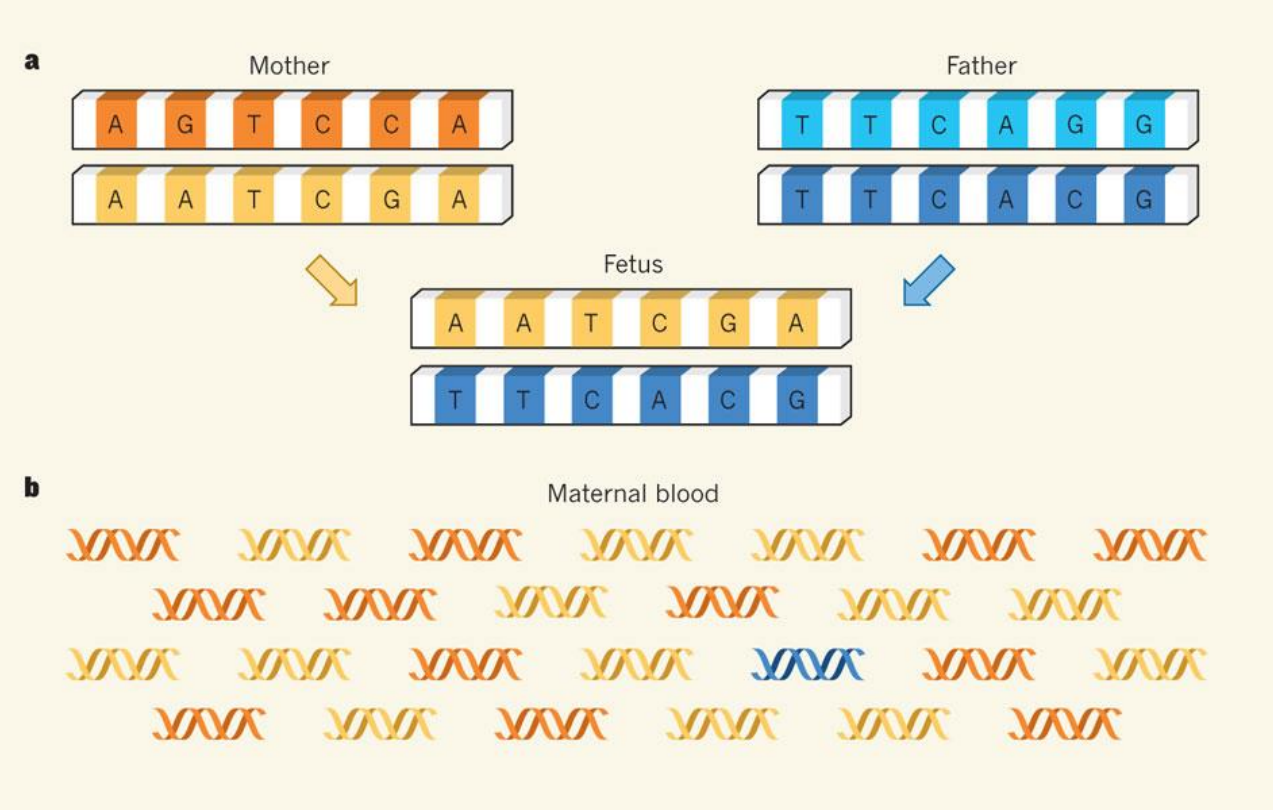


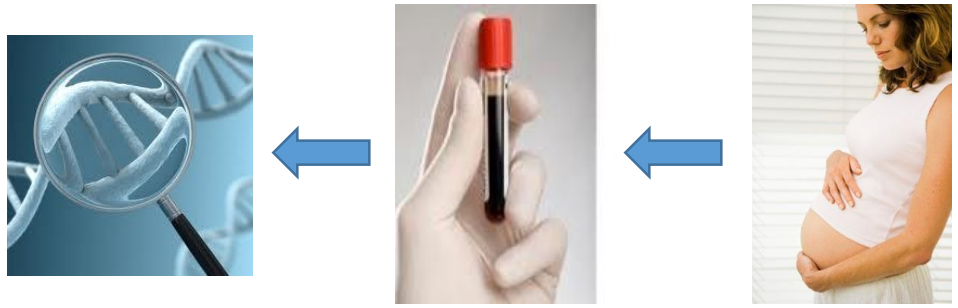
- Custom panels (multi-gene), many genes associated with a specific phenotype are sequenced and analyzed concomitantly, decreasing cost and improving efficiency of genetic diagnostic.
- Whole-exome sequencing (WES) all *coding* regions are sequenced with a relatively deeper depth, compared to WGS, the major advantage of WES is a significant cost reduction.
- Whole-genome sequencing (WGS) complete coverage of the genome, including promoters and regulatory regions.

http://www.smartplanet.com/blog/rethinking-healthcare/babys-genome-found-in-moms-blood/

Noninvasive Prenatal Diagnosis

During pregnancy, a median of 10% of the DNA in the blood plasma of a pregnant women— called ‘cell-free' DNA — comes from the fetus
(more precisely from the placenta).





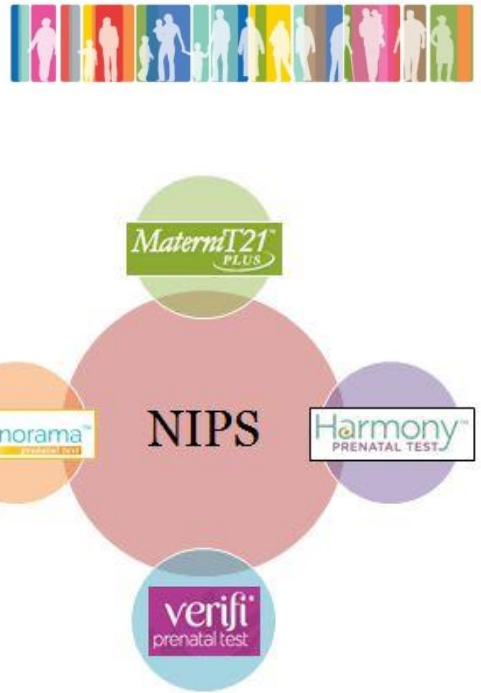
Diagnostic sequencing for fetal indications is best done as a trio analysis, where fetal and both parental samples are sequenced and analyzed together.

Why?

Screening for trisomy

phg foundation making science work for health

Non-invasive prenatal diagnosis using fetal DNA in maternal plasma: From dream to reality

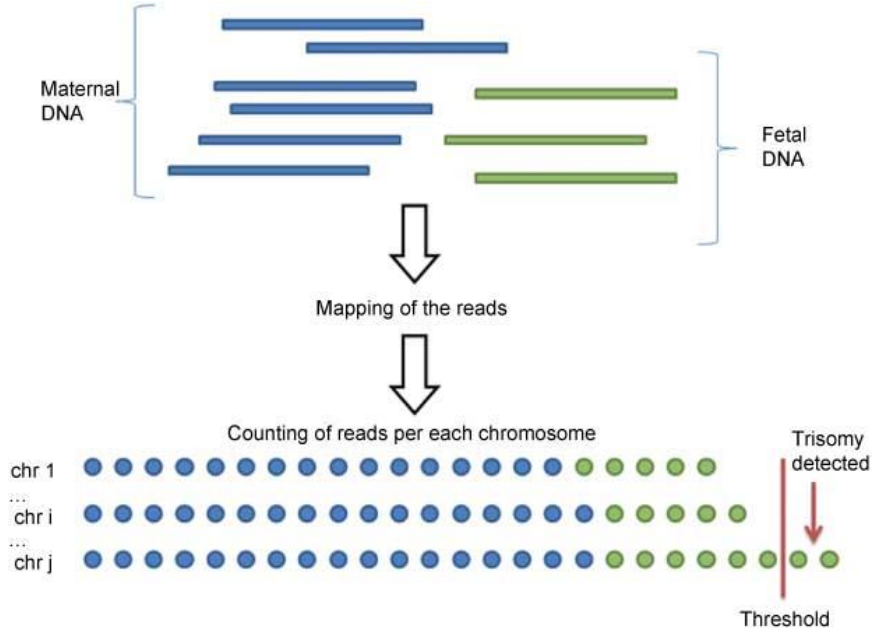


Screening for trisomy 21 (Down syndrome)

with the use of cell-free DNA (cfDNA)

Among the 11,994 women with low-risk pregnancies (maternal age < 35 years) cfDNA testing identified 19 women with trisomy 21, with 6 false positive results.

Limitation: no cfDNA result was returned for ~3% of women

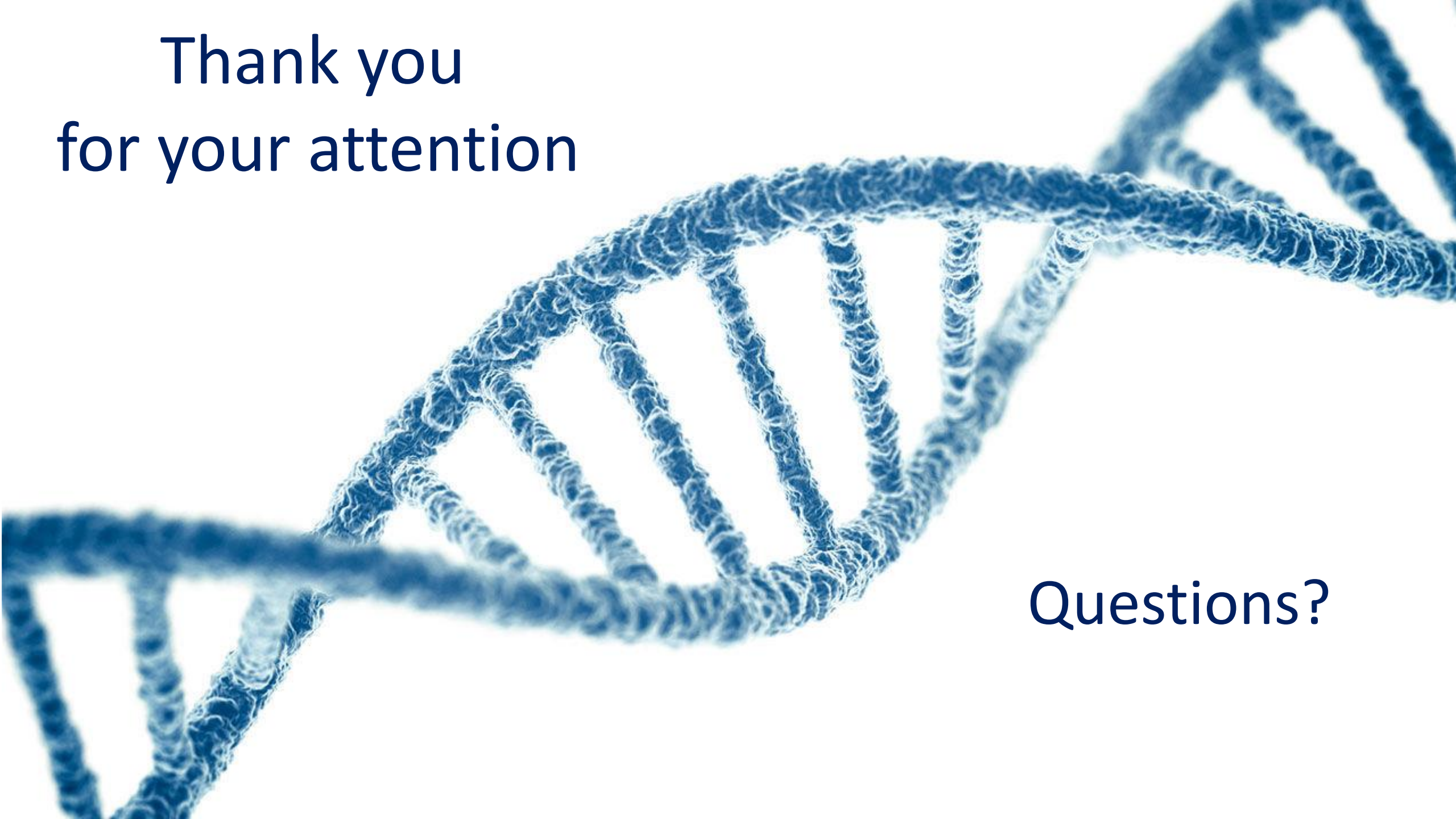


Nat Rev Genet. 2015 Jun;16(6):316-7. doi: 10.1038/nrg3953.

Summary

- NGS technologies sequence millions of small fragments of DNA in parallel, revolutionizing genomic research and impacting clinical practice.
- Multi-omics data is collected from patients, and integrated to create individual molecular profiles.
- Be critical with your experiments! We will cover the following in the course:
  - Experimental design
  - Selection of application
  - Selection of samples
  - Statistical considerations
  - Analysis of results
- NGS-based genomic sequencing will enable “precision medicine”, where patients receive individualized therapy based on their genomic data.

Thank you for your attention



Questions?