BIMM 194

Clinical Applications of Genomic Technologies

Lecture 3

Barry Grant

UC San Diego

http://thegrantlab.org/bimm194

Major impact areas for genomic medicine

- Cancer: Identification of driver mutations and drugable variants, Molecular stratification to guide and monitor treatment, Identification of tumor specific variants for personalized immunotherapy approaches (precision medicine).

- Genetic disease diagnose: Rare, inherited and so-called ‘mystery’ disease diagnose.

- Health management: Predisposition testing for complex diseases (e.g. cardiac disease, diabetes and others), optimization and avoidance of adverse drug reactions.

- Health data analytics: Incorporating genomic data with additional health data for improved healthcare delivery.

Major impact areas for genomic medicine

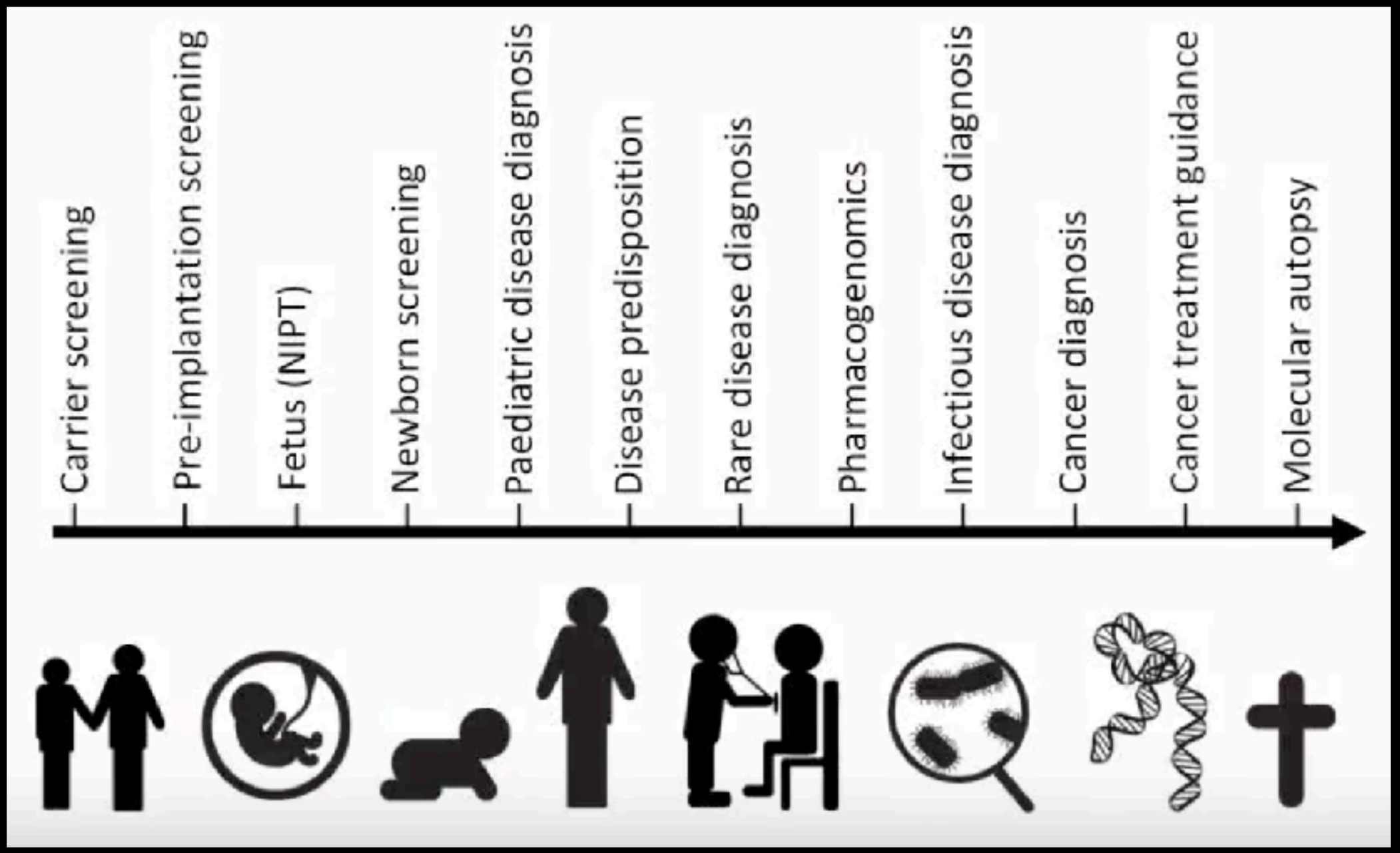
- Cancer: Identification of driver mutations and drugable variants, Molecular stratification to guide and monitor treatment, Identification of tumor specific variants for personalized immunotherapy approaches (precision medicine).

- Genetic disease diagnose: Rare, inherited and so-called ‘mystery’ disease diagnose.

- Health management: Predisposition testing for complex diseases (e.g. cardiac disease, diabetes and others), optimization and avoidance of adverse drug reactions.

- Health data analytics: Incorporating genomic data with additional health data for improved healthcare delivery.

Genomics in the whole of life healthcare



Rapid progress of genome sequencing

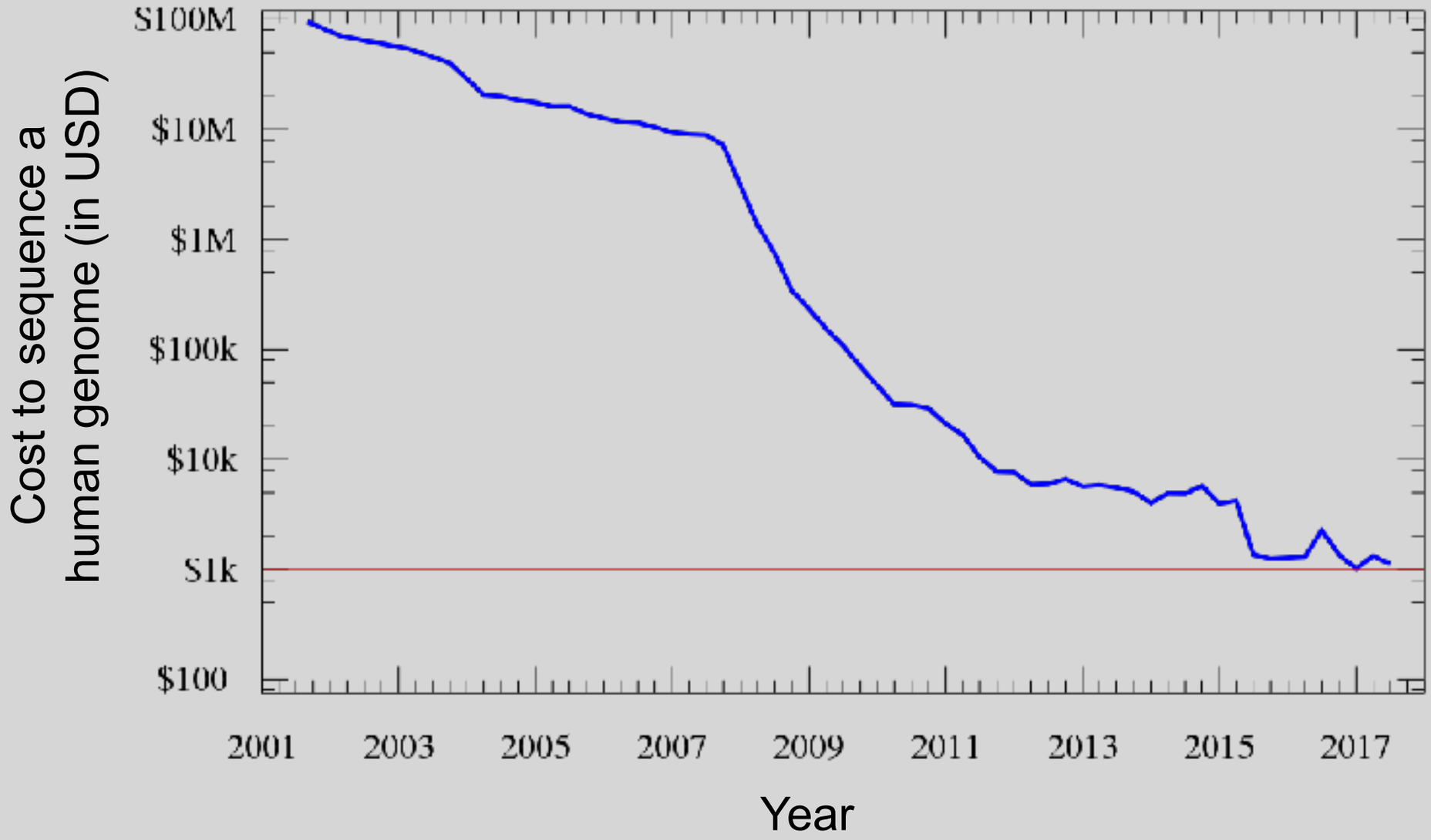
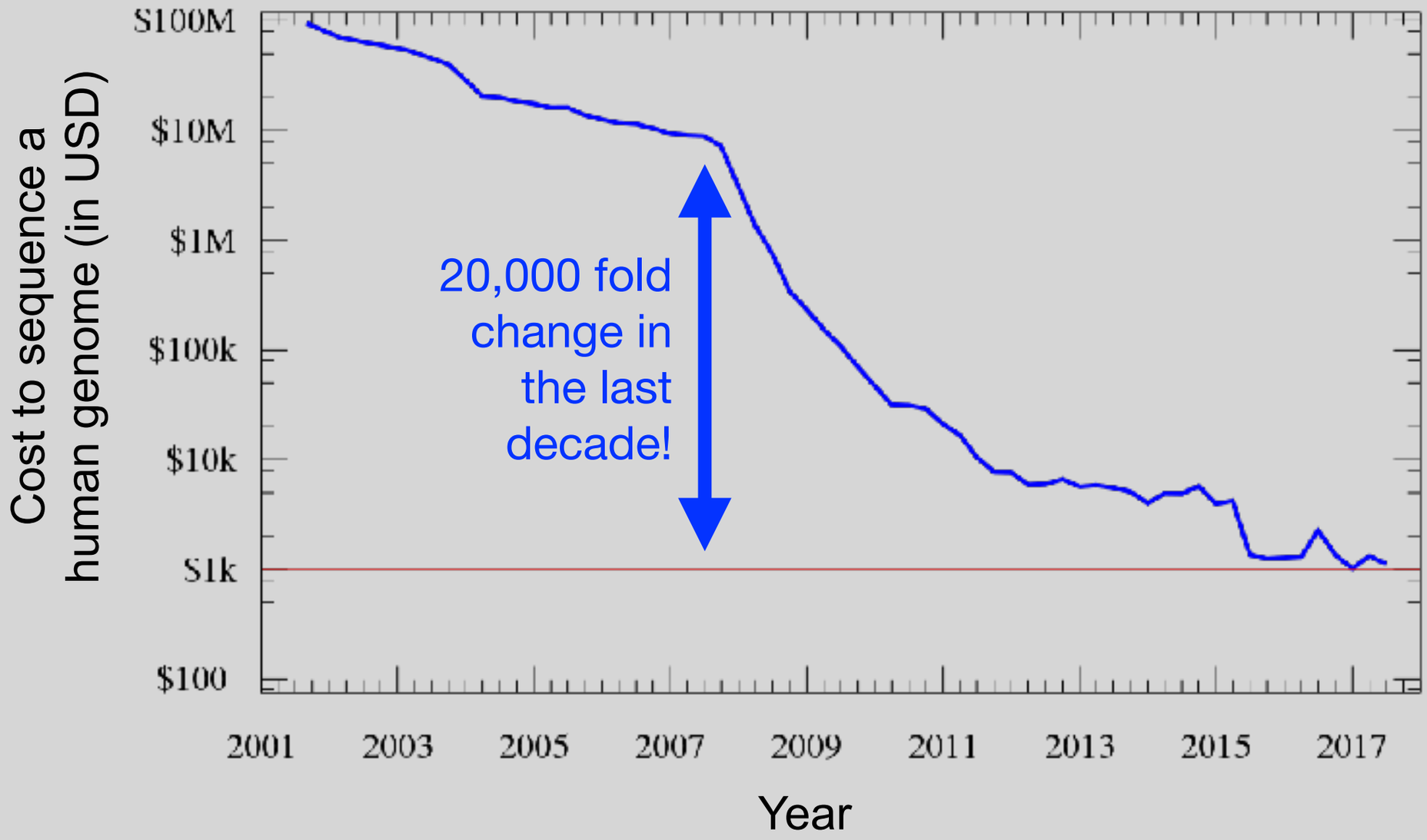


Image source: https://en.wikipedia.org/wiki/Carlson_curve

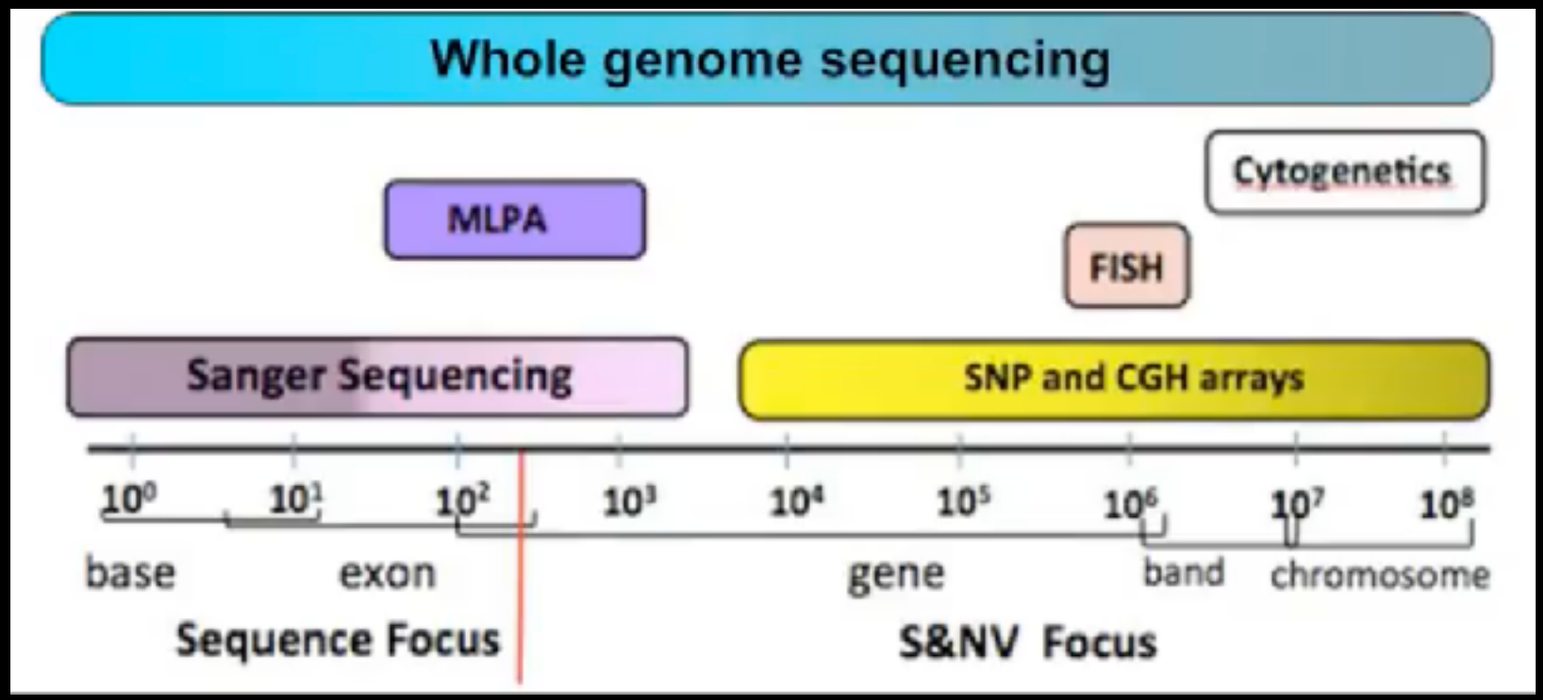
Rapid progress of genome sequencing



MRI: $4k

Image source: https://en.wikipedia.org/wiki/Carlson_curve

Whole genome sequencing transforms genetic testing



- 1000s of single gene tests
- Structural and copy number variation tests
- Permits hypothesis free diagnosis

Today's Menu

Rare genetic disease diagnose	Solving mystery diseases with exome and whole genome sequencing strategies
Complex genetic diseases	Genomic approaches for complex disease predisposition testing & diagnose
Pharmacogenomics	Guiding treatments by informing on drug metabolism, side-effects & interactions.

Today's Menu

Rare genetic disease diagnose	Solving mystery diseases with exome and whole genome sequencing strategies
Complex genetic diseases	Genomic approaches for complex disease predisposition testing & diagnose
Pharmacogenomics	Guiding treatments by informing on drug metabolism, side-effects & interactions.

Solving mystery diseases

- Diseases with a genetic origin effect 16 million people in the US and 23% of all pediatric admissions to hospital are for ‘rare' genetic disorders.

- Most are “mystery diseases” in terms of their genetic origin

- Before the recent adoption of exom and genome sequencing these patients faced extensive periods of testing and inappropriate treatment (with cost estimates of $5 million per person) before the basis of their disease was understood.

- Sequencing can thus help realize enormous savings in healthcare costs and spare patients and their families unnecessary, stressful, and time-consuming testing.

How many Mendelian diseases are there?

- As of 01/10/18 ~7,800 Mendelian diseases have been described.

- For 3,963 of these, the likely disease gene is known.

- For many genes, different genetic variants can have distinct effects on the encoded protein, leading to distinct disease characteristics.

- Indeed, the 3,963 unique diseases that have been solved affect only 2,776 genes because different mutations in the same gene can cause different disease characteristics.

How many Mendelian diseases are there?

- It is probable that many more Mendelian diseases will be “solved” as genomic analysis becomes more integrated into clinical practice.

- There are ~20,000 protein coding genes and and variants in many of these genes would be expected to cause human disease.

- **Q:** How are genes responsible for genetic diseases currently identified?
  - **Exome** or **whole genome sequencing**

Exome sequencing strategies to identify new disease genes

- Exome sequencing has been the most widely used sequencing approach for the identification of new disease-causing genes

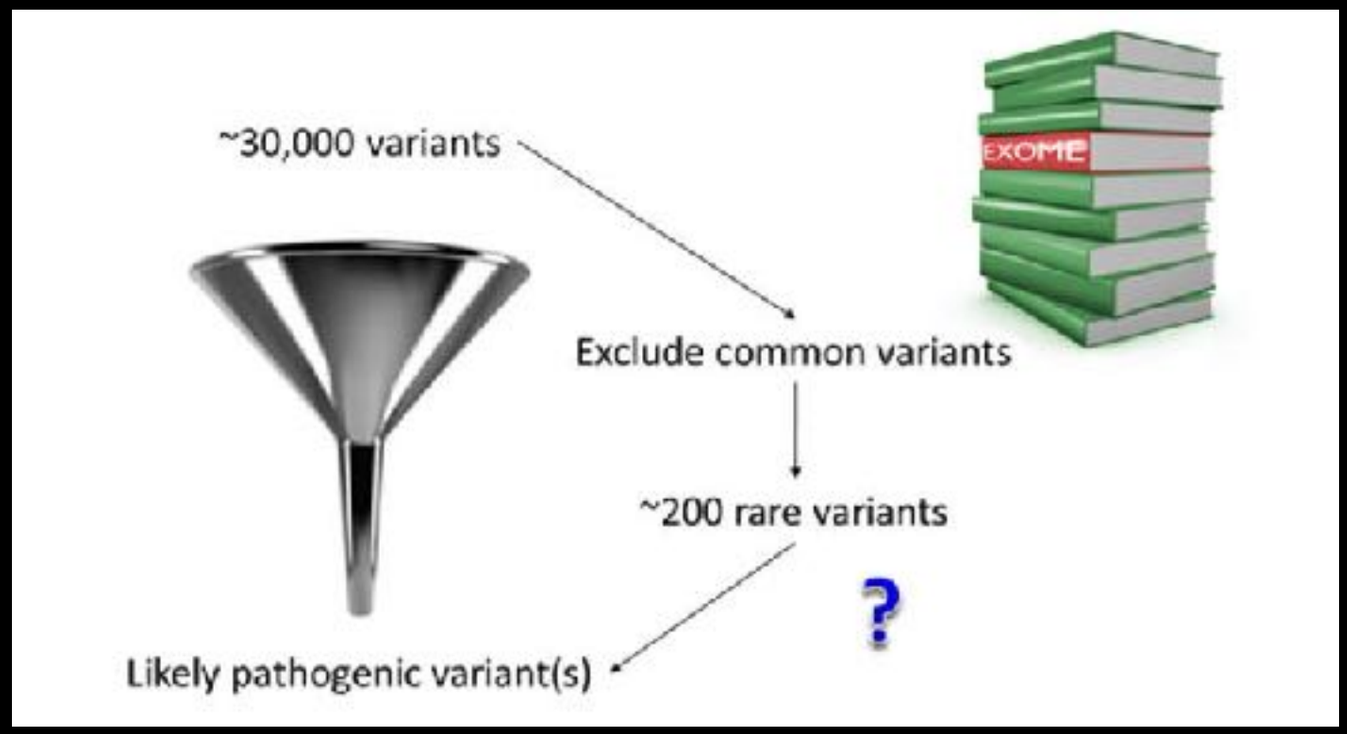
- The coding regions (i.e. the exons) of the genes account for just about 2% of the human genome

- However it is estimated that approximately 85% of the disease-causing mutations fall within a coding region.

Strategies for exam analysis

- Exome sequencing typically identifies ~200 novel/rare variants in each individual.

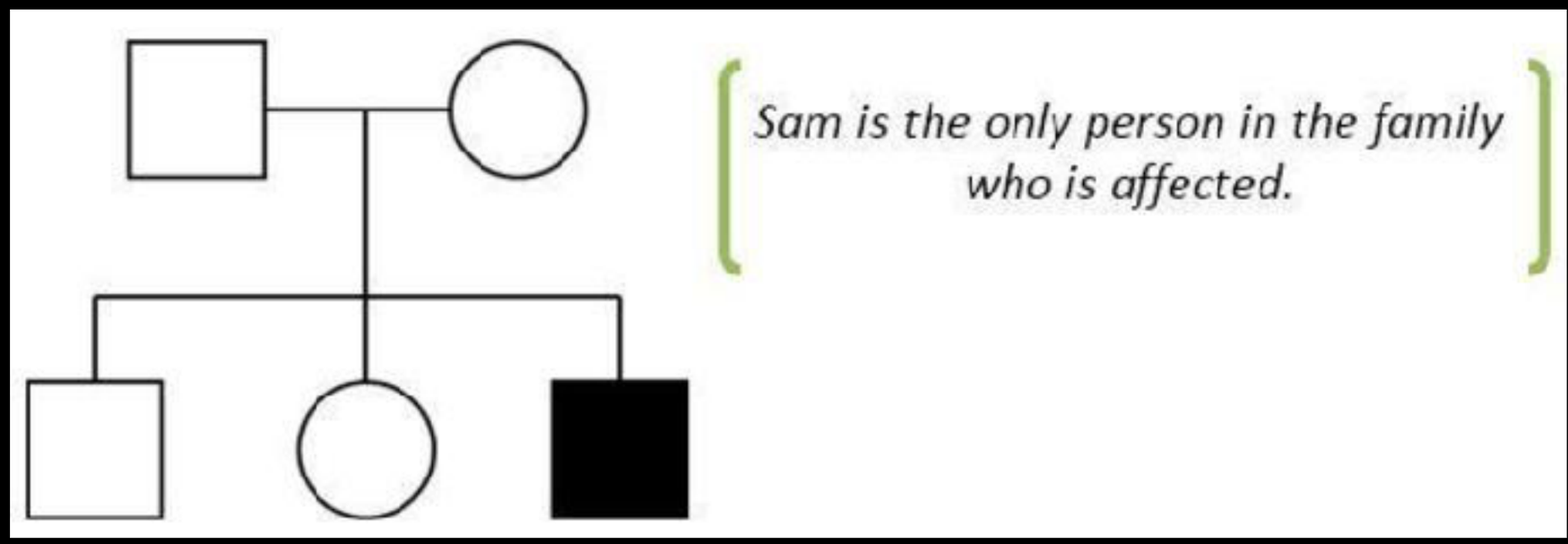
- Therefore to recognize the new disease gene variants among the hundreds of variants of no clinical significance it is necessary to define a clear strategy.



1. Exome trio strategy

- Consider Sam who was diagnosed with diabetes at birth.
  - born small for gestational age
  - multiple heart defects that needed surgical correction
  - imaging of his abdomen showed that he had no pancreas
  - no one in his family had diabetes or heart defects

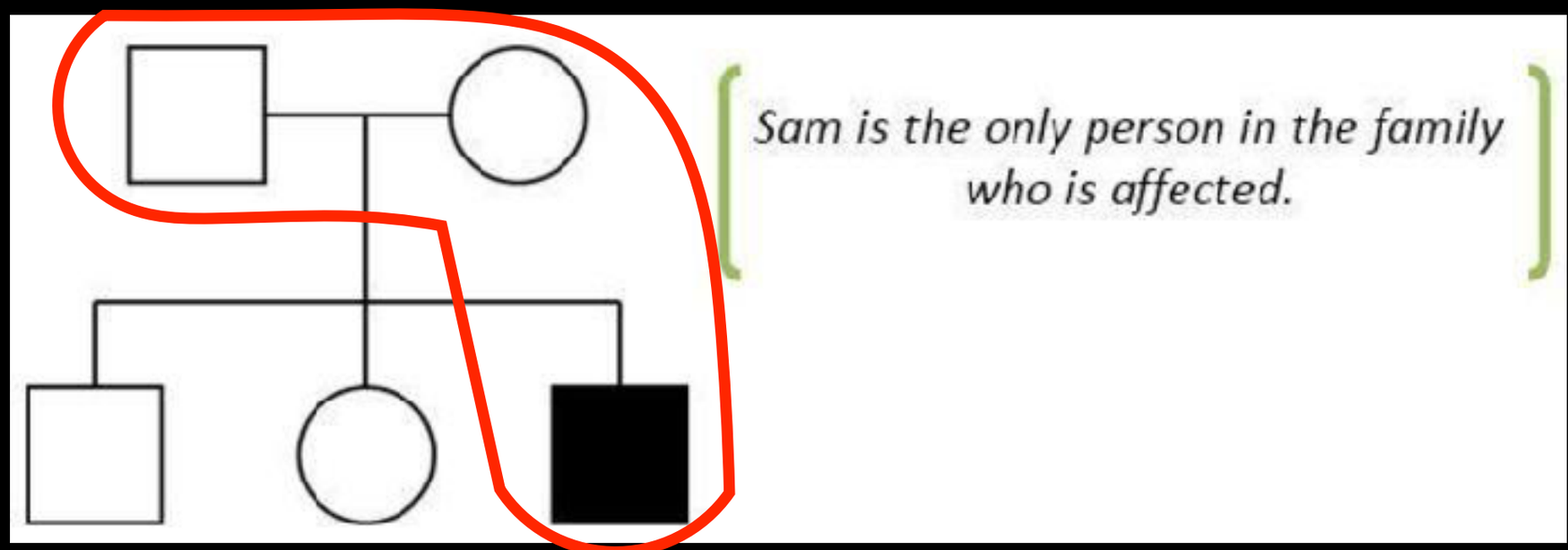
- This would be Sam's pedigree:



1. Exome trio strategy

- In this example it is likely that the mutation has arisen spontaneously in the patient and has not been inherited by either unaffected parent (geneticists call this a *de novo* mutation).

- We can use exome sequencing of Sam (the patient) and his parents and exclude all the variants that Sam has inherited from either parent and look at the mutations that are found only in Sam.



1. Exome trio strategy

- Using this approach, we typically find between 0 and 4 *de novo* coding mutations in each patient.

- In Sam's case only one *de novo* mutation in the **GATA6** gene was found.

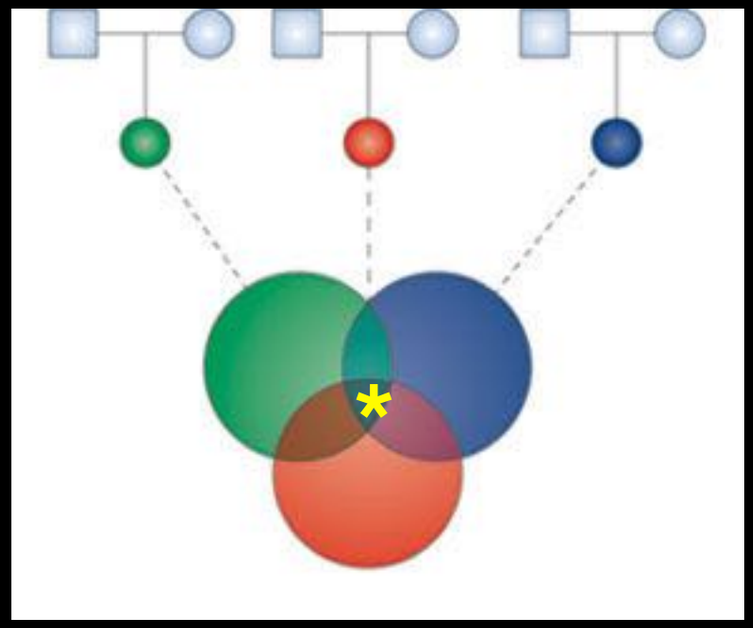
- After finding this mutation in Sam, GATA6 was recognized as being the gene most frequently mutated in patients with **neonatal diabetes** and who are born without a pancreas (a condition called **pancreatic agenesis**).

GATA6 haploinsufficiency causes pancreatic agenesis in humans, Lango Allen H et al 2011, Nature Genetics

2. Shared phenotype strategy

- Sometimes there are multiple patients who clearly have the same rare syndrome that is likely to have a monogenic basis.

- In these cases the disease-causing gene can be identified by performing exome sequencing on multiple affected individuals and looking for either the same genetic variant or different variants within the same gene.



Exome sequencing as a tool for Mendelian disease gene discovery
Nature Reviews Genetics 12, 745–755 (2011)

The need for whole genome sequencing…

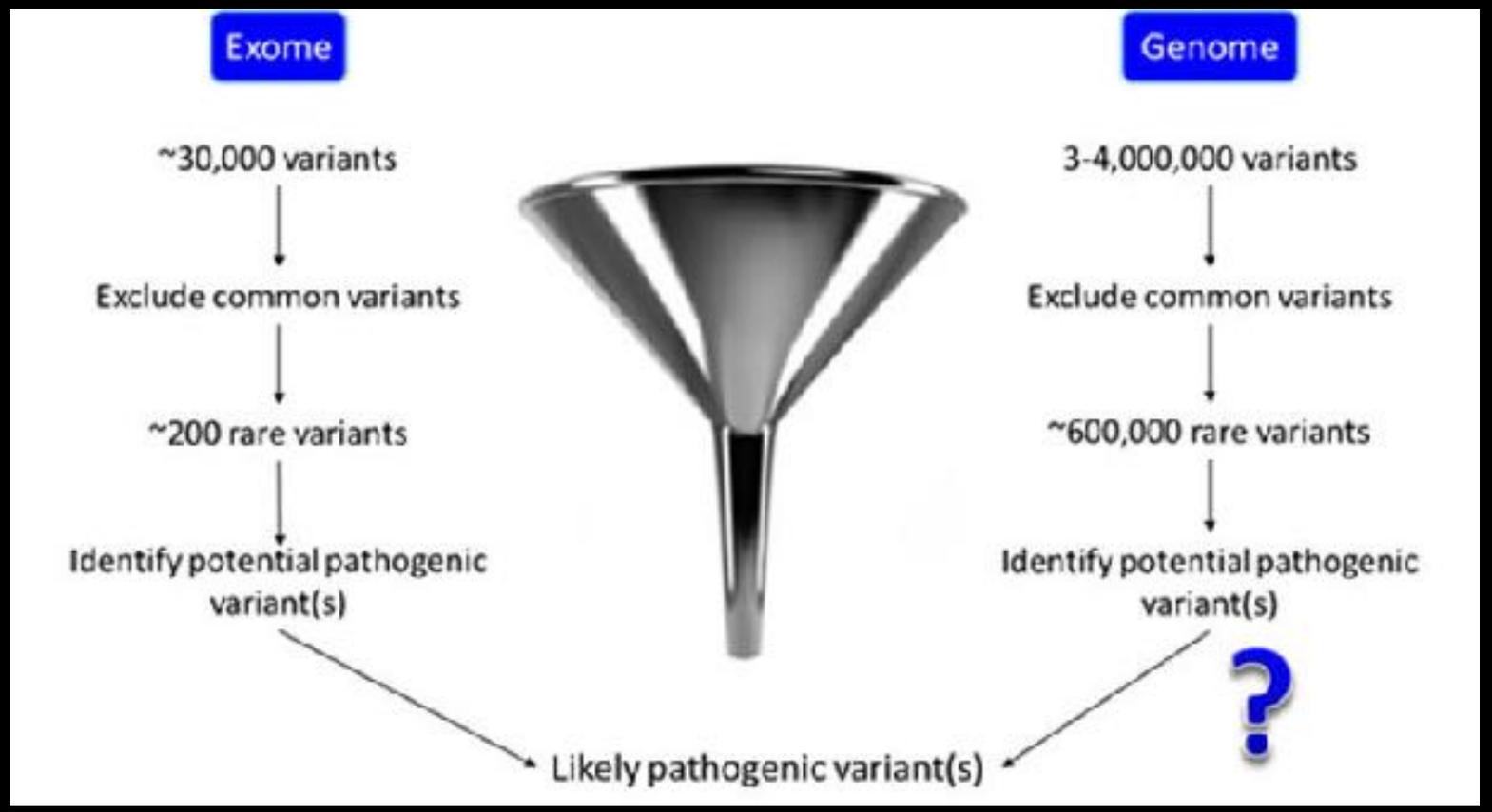
- But, even if 85% of the disease-causing variants fall within the exome, sometimes the causal mutation is located outside the coding regions.

- In these cases, exome sequencing simply isn’t enough… We need whole genome sequencing!

- As the cost of sequencing the whole genome is falling rapidly, this approach will likely becoming the method of choice for identifying mutations causing disease.

Identifying mutations causing disease outside the exome is challenging

- Each person's genome contains on average ~3,500,000 variants! You can imagine now how difficult it is to pinpoint the single genetic change which is the cause of a patient's disease among all these variants.



Identifying mutations causing disease outside the exome is challenging

- Predicting the effect of a variant outside of a coding region is extremely challenging because:
  - 99% of all the variants found in any individual are located in a non-coding region
  - the knowledge of what the non-coding regions do is very limited

- For these reasons there are only a very small number of non-coding disease mutations that have been identified to date.

- New techniques, and projects such ENCODE and the Epigenome Roadmapas, aim at finding the elements of the genome which regulate (switch on and off) genes, such as enhancers and promoters.

Nicholas Volker: an early poster child for whole-genome sequencing

- Nicholas Volker was healthy until he was 2 years old. Then he developed a cut that would not heal. His condition dramatically worsened. He developed sepsis and his intestine became dangerously inflamed, necessitating a hundred surgeries including the removal of his colon. He could not eat or drink and required parenteral nutrition. No one knew the cause, but it seemed certain that he was dying.

- By sequencing his genome he was found to have a mutation in the XIAP gene associated with immune function.

- He then received a cord blood transplant and made a remarkable recovery and is now doing well at 6 years old.

https://www.medscape.com/viewarticle/822094
https://tinyurl.com/NicholasVolker

Beery twins

- Non-identical twins born as “floppy” babies, had seizures and delayed motor skills. They were diagnosed with cerebral palsy and several otters conditions.

- Years later genome sequencing revealed a mutation in both copes of the SPR gene in both twins.

- SRP is involved in producing dopamine and serotonin.

- Administrating of dopamine and serotonin supplements improved their health so they are now symptom-free.

http://www.nature.com/news/2011/110615/full/news.2011.368.html

Currently disease causing mutations are found in only ~30% of cases

- For the majority of these cases finding disease causing mutations often does not lead to effective treatments.

- However, the information can still be helpful for guiding patient management, reproductive choices and future certainty. For example:

  - Can bring relief for patients and their families

  - Can be helpful for planning future pregnancies (e.g. IVF and genetic testing for embryo selection)

  - Predicting the possible disease course and long-term prognosis

The personal impact

" We had to experiment with a lot of medicines that as a parent you wouldn't normally want to give to a child"

"… then we did one WGS test, got a diagnosis, and now we have hope …"

Genetic testing and IVF

Example of Gerstmann-Straussler-Scheinker (GSS): Transmissible Spongiform Encephalopathy

- 26 yo finds she has GSS, likely develop dementia and die 30-50yo
- Underwent IVF and genetic testing for embryo selection and now has healthy twins
- "P*eople who carry a gene like GSS have a moral duty to use preimplantation diagnosis to spare the next generation*" Janet Malek, bioethicist, Brody School of Medicine (as reported by NYT article)
- Are we willing to argue that such people should not be allowed to exist?

What would our world be like without these people

Side-Note:

- Woody Guthrie: Huntington's disease
- Frederic Chopin: Cystic Fibrosis
- Miles Davis: Sickle Cell Anemia
- John F. Kennedy: Addison's disease
- Maurice Ravel: Frontotemporal dementia
- Lou Gehrig: ALS
- Ronald Reagan: Alzheimer's disease
- Charles K. Kao (Nobel prize in physics, father of fiber optics and broad band): Alzheimer's
- Stephen Hawking: ALS

etc…

When Genome Sequencing just isn't enough

- NGS does not always identify the causative mutation. So where are these missing mutations hiding?

- The most commonly used short read technologies can only call mutations in 88-95% of the genome.

- The remaining 5-12% of the genome is either not sequenced to a high enough quality to allow for mutation detection or is impossible to map as a result of repetitive DNA sequences.

- But technology is moving fast and single molecule real-time sequencing platforms that generate reads of around 10,000 bases are filling in some of these gaps and identifying new variants.

The first direct RNA sequencing by nanopore

Side-Note:

- For example this new nanopore sequencing method was just published **last week**: https://www.nature.com/articles/nmeth.4577

- "Sequencing the RNA in a biological sample can unlock a wealth of information, including the identity of bacteria and viruses, the nuances of alternative splicing or the transcriptional state of organisms. However, current methods have limitations due to short read lengths and reverse transcription or amplification biases. Here we demonstrate nanopore direct RNA-seq, a highly parallel, real-time, single-molecule method that circumvents reverse transcription or amplification steps."

Can we detect all mutations?

- The sensitivity of mutation detection depends upon the mutation type as well as the read depth.

- Base substitutions (SNVs) are most easily detected but insertions and deletions (InDels) are more difficult because of capture bias (in targeted methods but not genome sequencing) and mapping issues.

- Bioinformatic tools to detect copy number variants (CNVs), chromosome rearrangements (structural variants: SVs) and insertions are still in their infancy and their sensitivity has not been established.

Are we looking in the right place?

- Even if it were possible to accurately call all variants in the entire genome it is still likely that some mutations would be missed as we might just not be looking in the right place.

- For example some mutations may arise spontaneously after conception (so called “post-zygotic”, ‘somatic’ or ‘acquired’ mutations) which will result in varying levels of the mutation between tissues.

- For these individuals sequencing DNA extracted from the blood may not detect the causative mutation. It is therefore important to consider the most appropriate source of DNA for sequencing studies.

Searching for something that is not there to be found

- It is important to remember that genetic disease does not always result from a change in the DNA sequence.

- A number of diseases are known to result from defects in the methylation status of DNA; an epigenetic mechanism used to control gene expression.

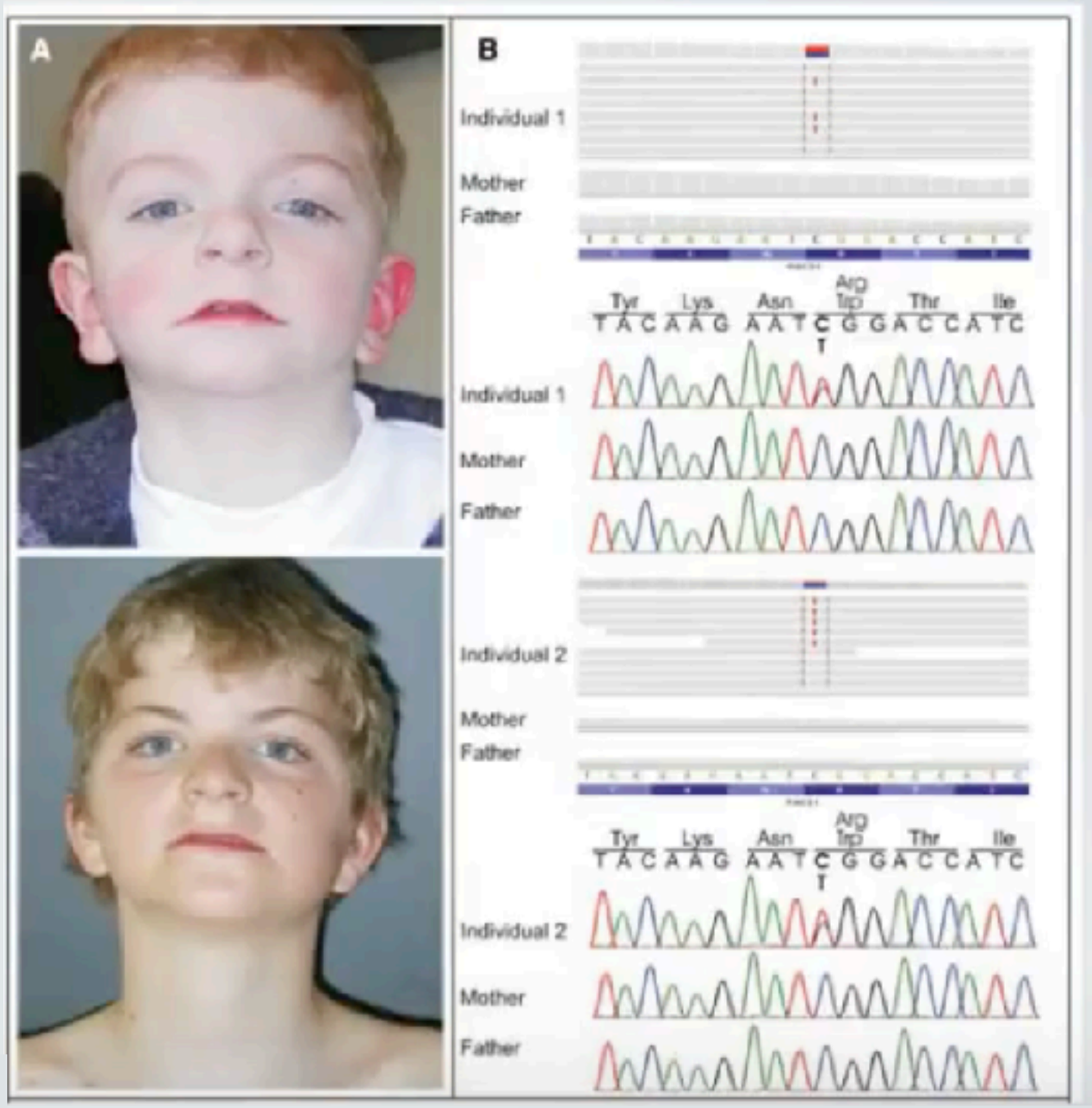
- Abnormalities in methylation cannot be detected by conventional sequencing and require a different NGS analysis approach.

N.B: Phenotype usually drives variant interpretation

Two unrelated boys with unexplained Intellectual Disability and:

- Low anterior hairline,
- Highly arched eyebrows,
- …
- Downturned mouth corners,
- Diastema of the teeth
- and Low-set ears

They shared the exact same mutation in PACS1



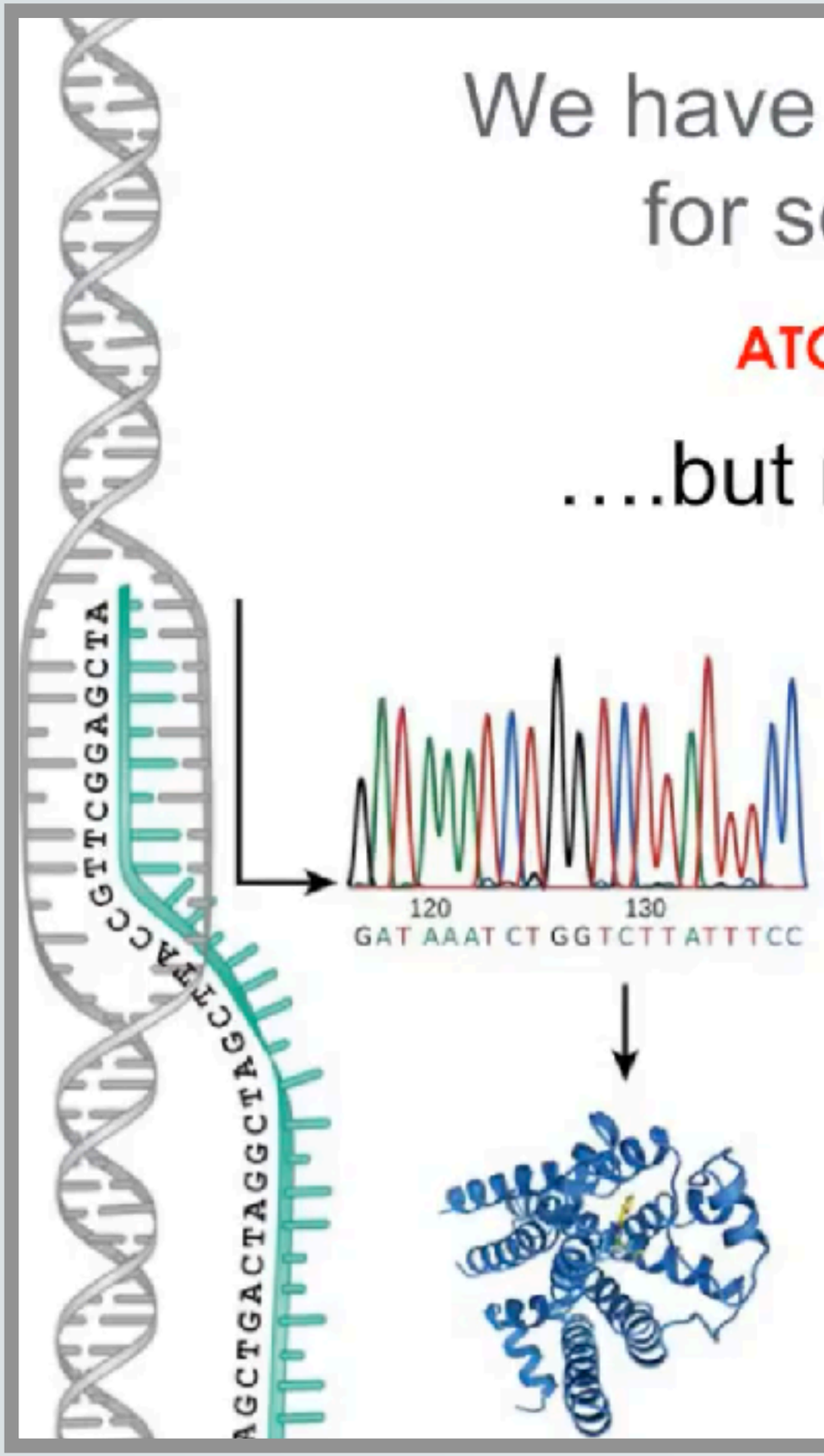
pheno + geno = new syndrome

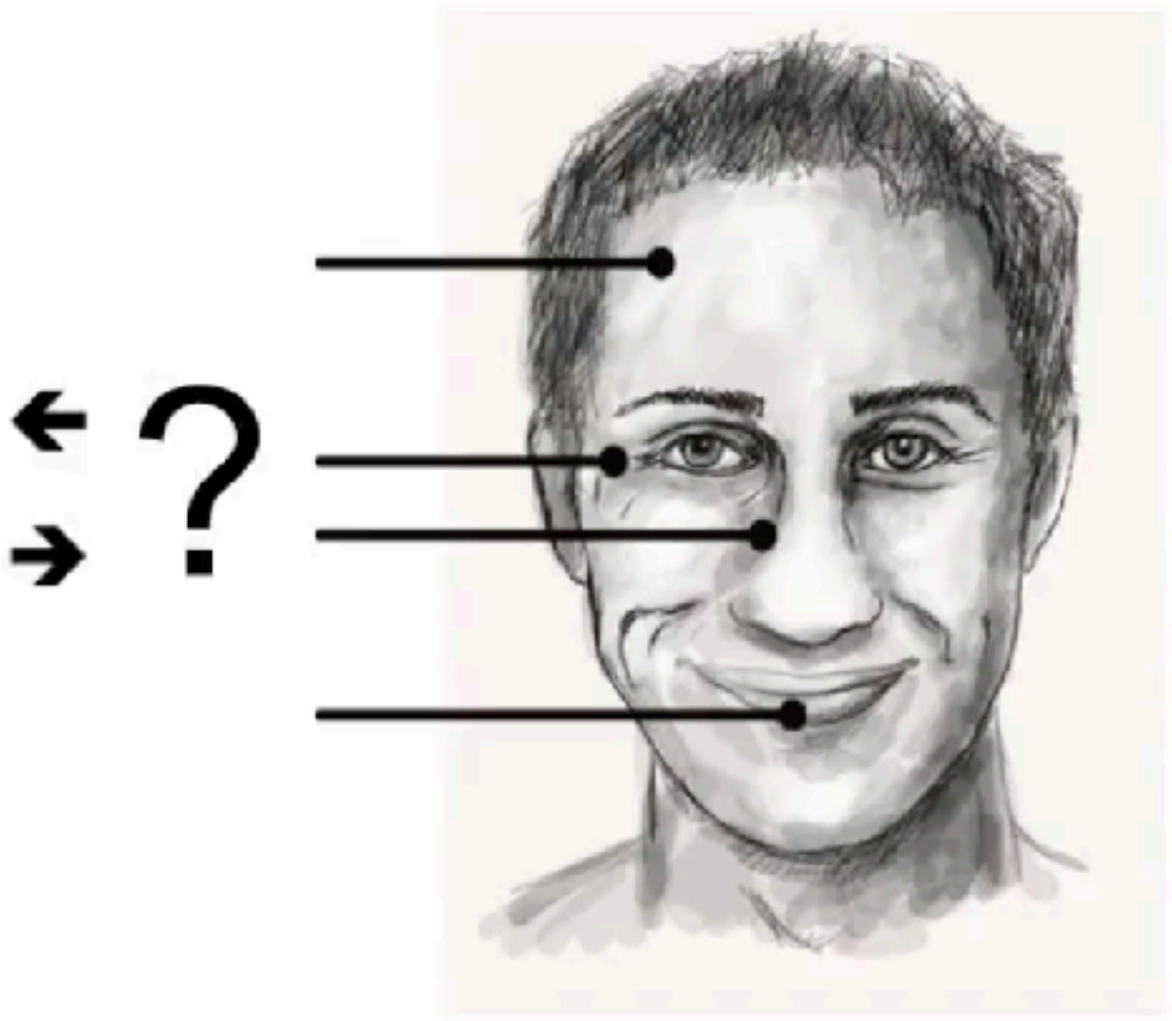
https://en.wikipedia.org/wiki/PACS1

We have a common language for sequence data….

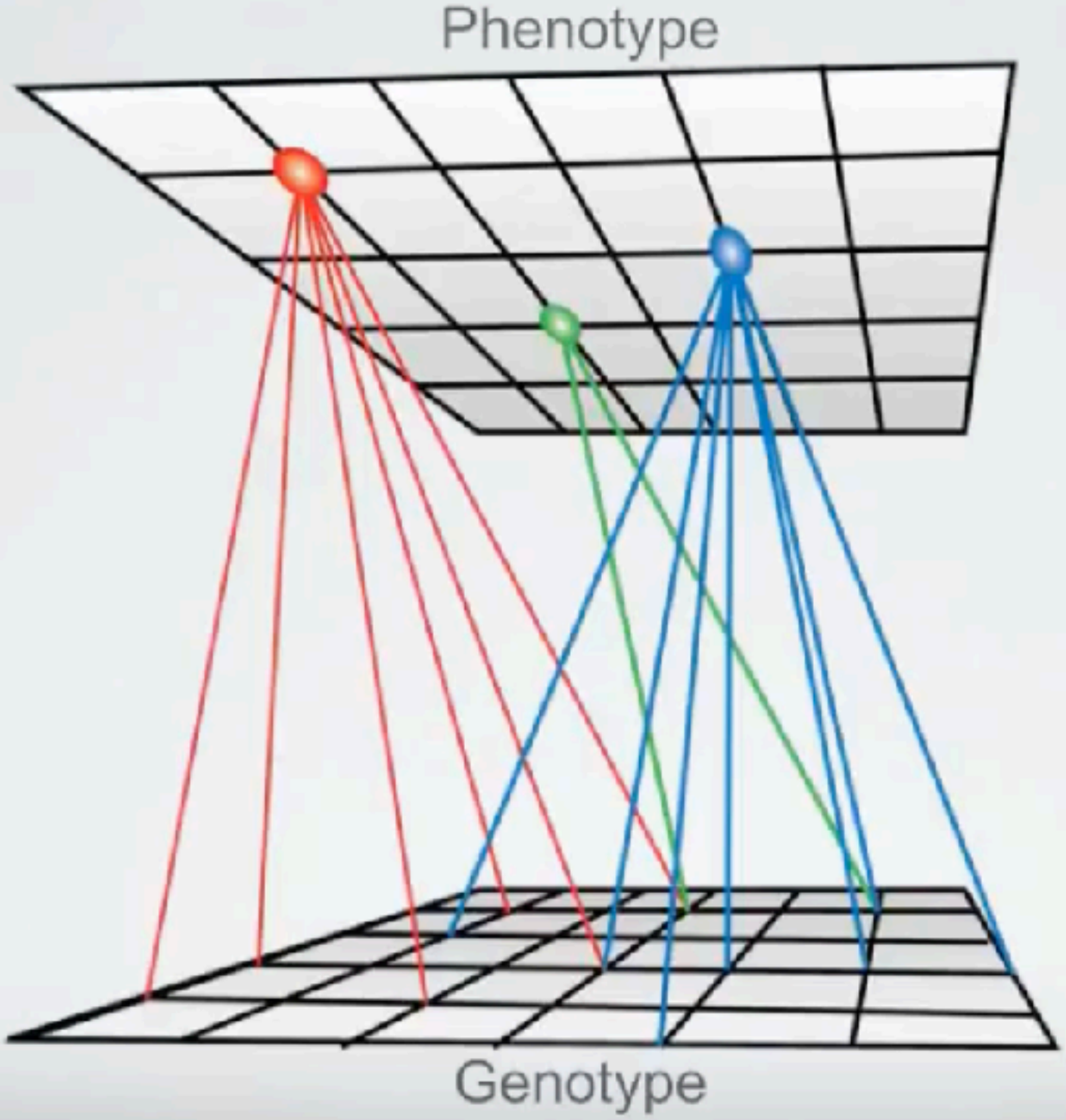
ATCTTAGCACGTTAC…

….but not for phenotypes

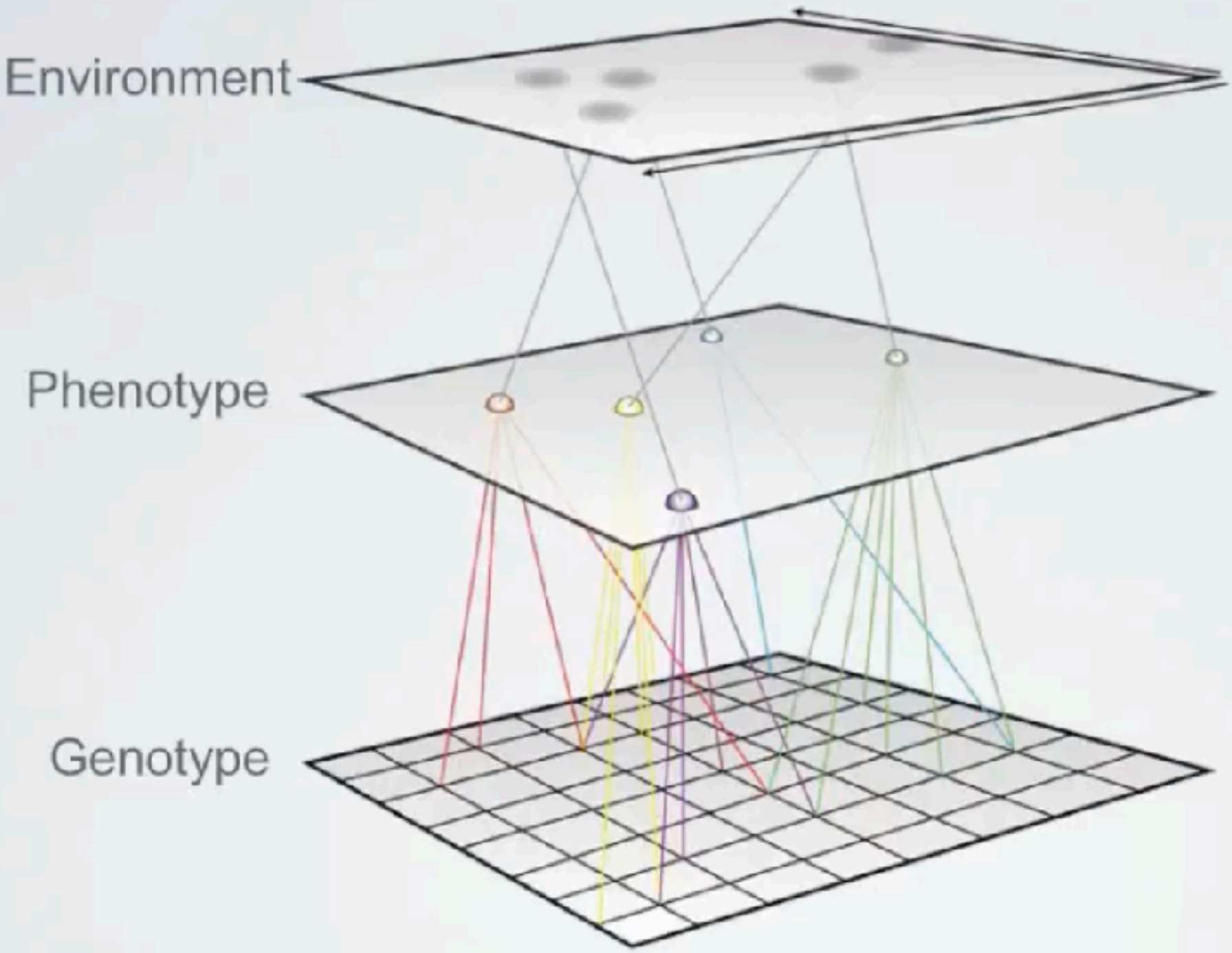




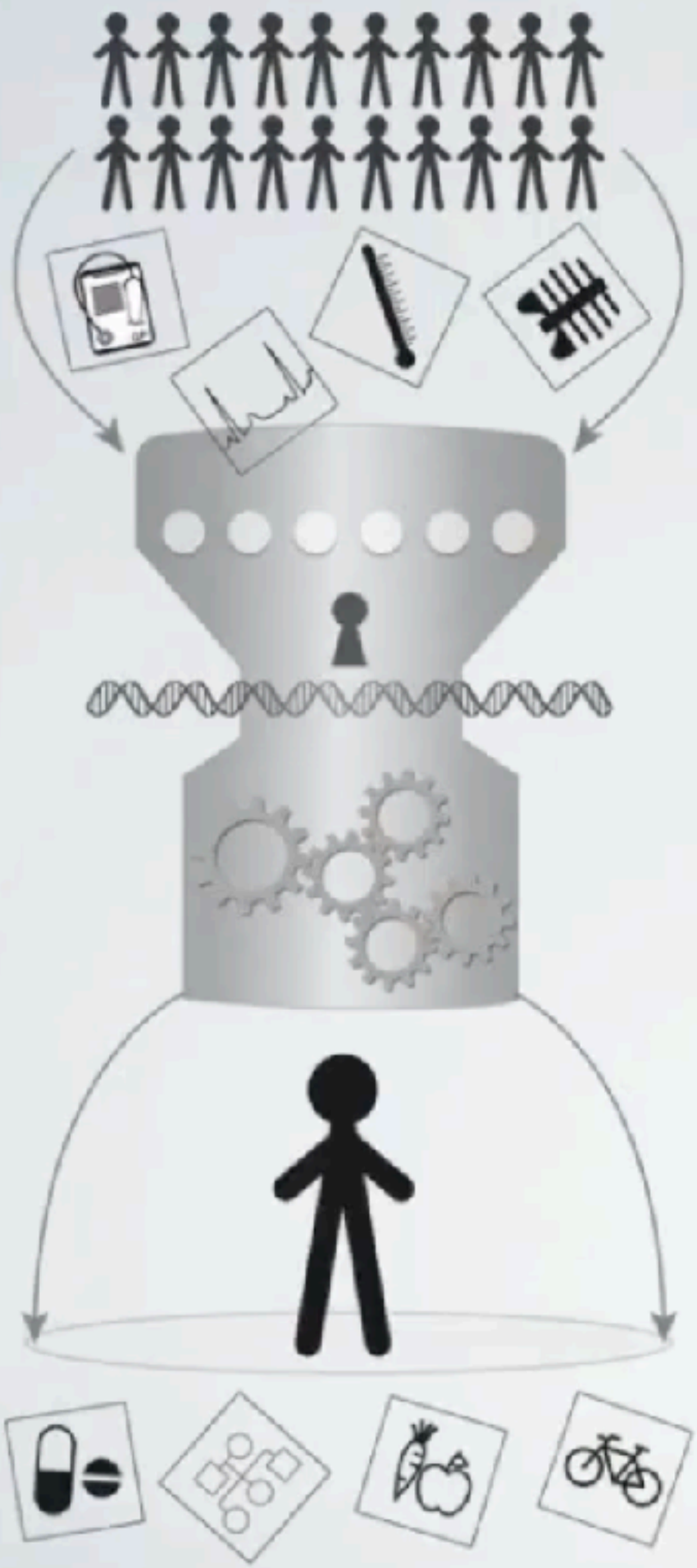
Unlocking the value of the genome requires "sequencing" of the phenome



… and ideally "sequence" the environment too



Explosive growth in personal health monitoring technology



Using population health analytics to inform precision healthcare

1. Standardised consent

2. Standardised clinical data

3. Analytics to infer new genetic associations with clinical characteristics

4. Machine learning algorithms to optimise clinical decision-making based on population-health data

5. Consumer awareness to maximise effectiveness

Today's Menu

Rare genetic disease diagnose	Solving mystery diseases with exome and whole genome sequencing strategies
Complex genetic diseases	Genomic approaches for complex disease predisposition testing & diagnose
Pharmacogenomics	Guiding treatments by informing on drug metabolism, side-effects & interactions.

Predisposition testing for complex disease

- Complex genetic diseases are likely caused by a combination of genetic, environmental, and lifestyle factors.

- Examples include Alzheimer's disease, asthma, Parkinson's disease, multiple sclerosis, osteoporosis, kidney diseases, autoimmune diseases, and many more

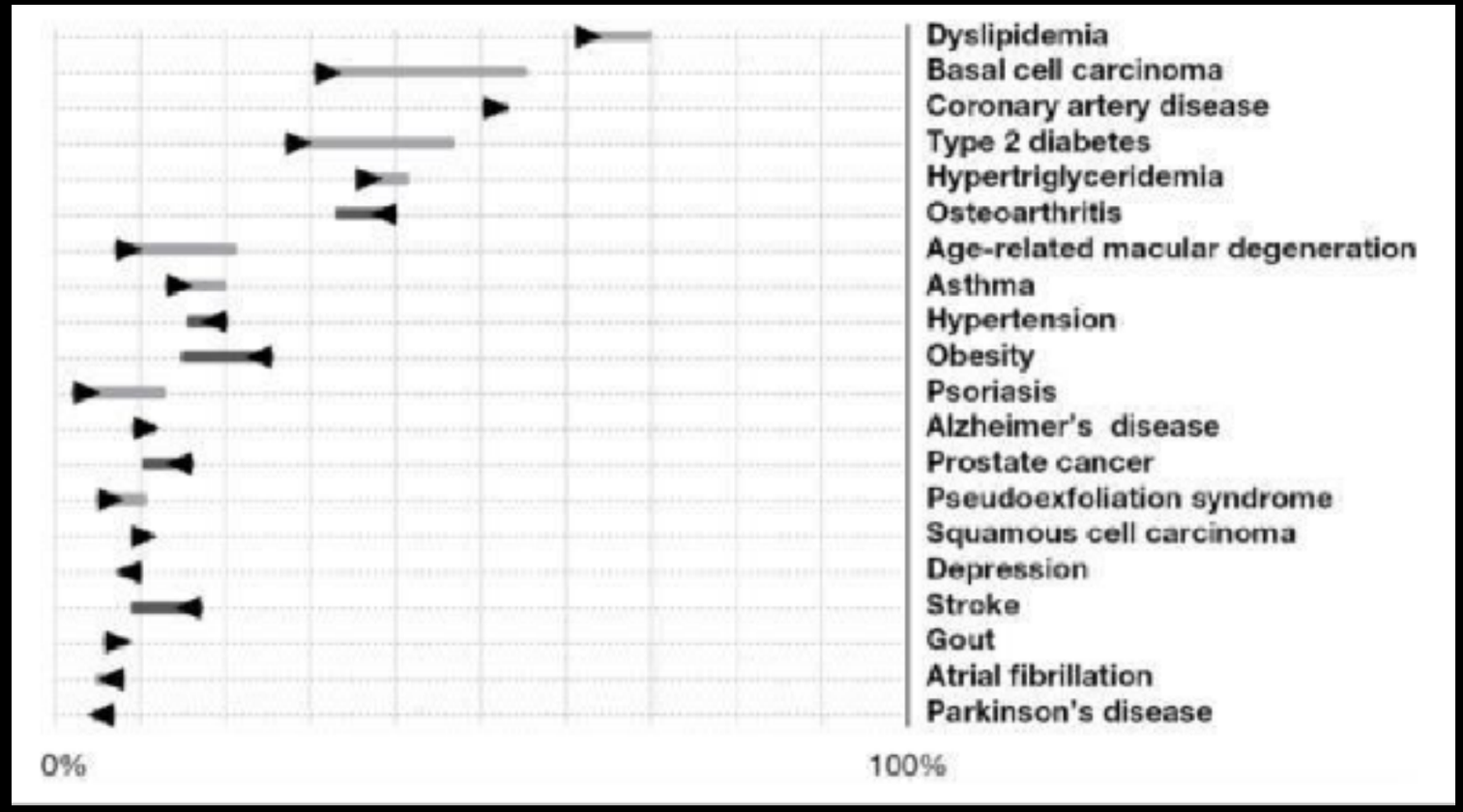
- Our understanding of the contributing ‘risk’ factors, their relative importance and their interactions with each other is generally limited.

- Dissecting these complex genetic diseases is the next frontier in modern genomic medicine

Current approaches

- Whole-genome and exome-sequencing can be combined with transcriptome sequencing (RNA-Seq) to assess expression levels and the expression of mutated transcripts and splice variants.

- A simple additive model, based on the assumption that risk factors have independent effects on disease risk, is often used.



Portion of Michael Snyder's 'risk-o-gram'

Direct to Consumer (DTC) Good, bad or ugly?

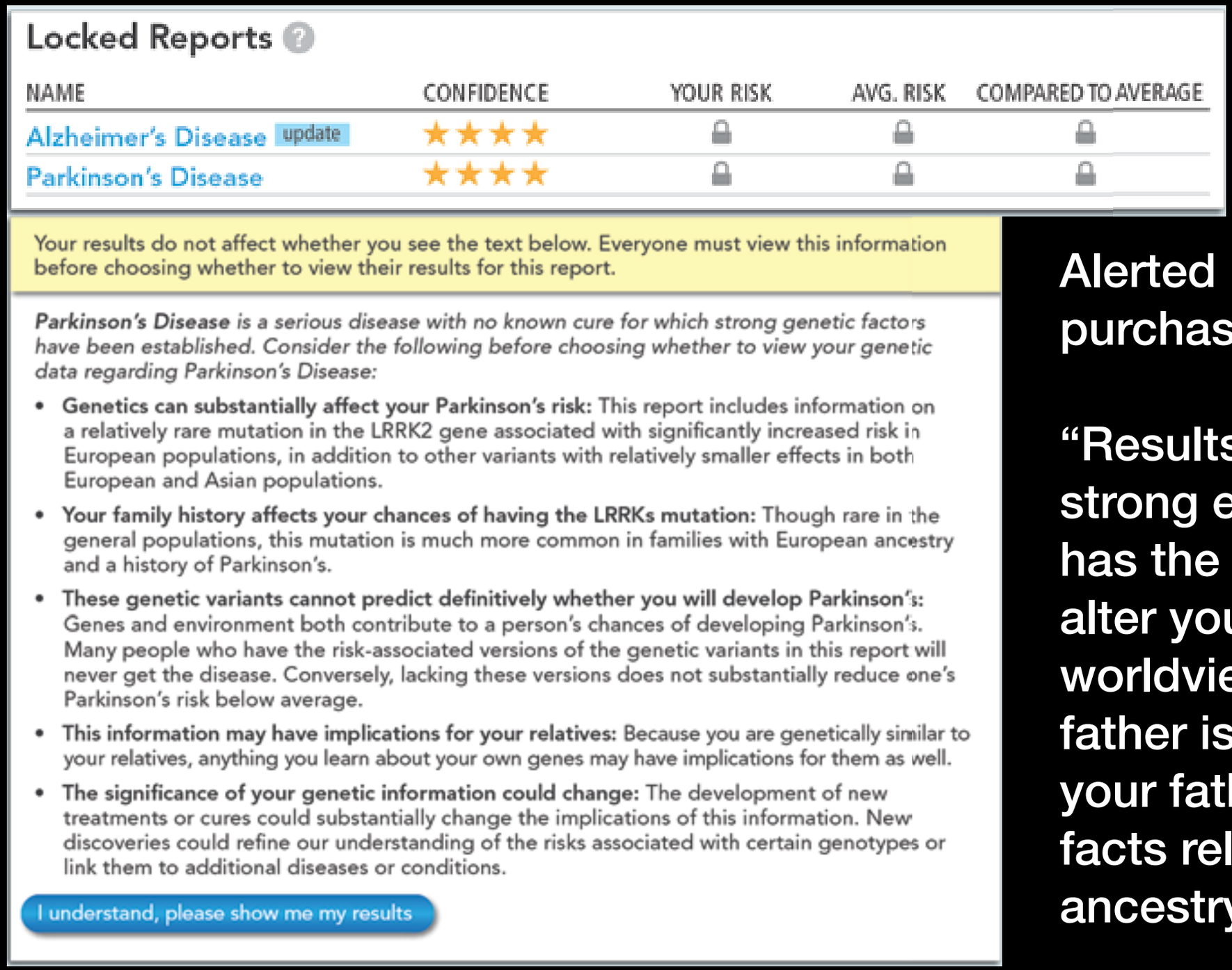
- 23andMe: Personal Genome Service (PSG): Single-nucleotide polymorphism chip capable of identifying mutations in genes associated with 254 specific diseases and conditions.

- Consumer's right to know
  - Medical (governmental) paternalism
  - Right to information about ourselves (medical records)

- Raw genetic data accessible to the consumer.

- Biobank of genetic information: used and sold for medical research and patentable discoveries.

23andMe Consent



Locked Reports

NAME	CONFIDENCE	YOUR RISK	AVG. RISK	COMPARED TO AVERAGE
Alzheimer's Disease update	★★★★	🔒	🔒	🔒
Parkinson's Disease	★★★★	🔒	🔒	🔒

Your results do not affect whether you see the text below. Everyone must view this information before choosing whether to view their results for this report.

***Parkinson's Disease** is a serious disease with no known cure for which strong genetic factors have been established. Consider the following before choosing whether to view your genetic data regarding Parkinson's Disease:*

- **Genetics can substantially affect your Parkinson's risk:** This report includes information on a relatively rare mutation in the LRRK2 gene associated with significantly increased risk in European populations, in addition to other variants with relatively smaller effects in both European and Asian populations.
- **Your family history affects your chances of having the LRRKs mutation:** Though rare in the general populations, this mutation is much more common in families with European ancestry and a history of Parkinson's.
- **These genetic variants cannot predict definitively whether you will develop Parkinson's:** Genes and environment both contribute to a person's chances of developing Parkinson's. Many people who have the risk-associated versions of the genetic variants in this report will never get the disease. Conversely, lacking these versions does not substantially reduce one's Parkinson's risk below average.
- **This information may have implications for your relatives:** Because you are genetically similar to your relatives, anything you learn about your own genes may have implications for them as well.
- **The significance of your genetic information could change:** The development of new treatments or cures could substantially change the implications of this information. New discoveries could refine our understanding of the risks associated with certain genotypes or link them to additional diseases or conditions.

I understand, please show me my results

Alerted before purchase that:

"Results may evoke strong emotions and has the potential to alter your life and worldview (e.g. your father is not genetically your father, surprising facts related to your ancestry…"

23andMe and the FDA

- “Immediately discontinue marketing Personalized Genome Service (PGS)”
- “Some of the uses for which PGS is intended are particularly concerning, such as assessments for BRCA-related genetic risk and drug responses because of the potential health consequences that could result from false positives or false negative assessments for high risk indications such as these.”
- Fair criticism?

Incidental or additional findings in genomic testing

- NGS test have the potential to uncover genetic information that may not be so welcome

- Analysis of a gene that is known to cause a disease other than the one for which the patient is being tested could reveal additional unexpected information

- Current policy is often not to report an incidental finding of carrier status, but to focus solely on the clinical reason for referral.

- Whilst there is growing evidence that many patients are keen to receive information about additional, clinically actionable findings, the results of studies such as the 100,000 Genomes Project are eagerly awaited in order to aid the understanding of the overall risks and benefits of receiving such information.

Side-Note:

What do you think?

- Should patients be offered the option of learning about “additional findings”?

- If you were invited to participate in a research project by having your genome sequenced what would you want to know?

- What would you not want to know?

- Would you discuss this with your immediate relatives before making a decision?

Attitudes of nearly 7000 health professionals, genomic researchers and publics toward the return of incidental results from sequencing research, Middleton A et al 2015, European Journal of Human Genetics

Variants of uncertain significance: innocent until proven guilty

- Sometimes, it is impossible to say with certainty whether we think a variant identified in a disease gene is causing the condition in question.

- In this situation we often use the term “variant of uncertain significance” or VUS.

- In some cases, a decision will be made to manage the VUS as a pathogenic mutation.

- However, in the majority of cases, a VUS will be managed as benign unless more evidence can be gathered to reclassify it as pathogenic.

Today's Menu

Rare genetic disease diagnose	Solving mystery diseases with exome and whole genome sequencing strategies
Complex genetic diseases	Genomic approaches for complex disease predisposition testing & diagnose
Pharmacogenomics	Guiding treatments by informing on drug metabolism, side-effects & interactions.

Pharmacogenomics

- Genomics can have a direct impact on drug treatments for mystery diseases (e.g. the case of the Beery twins from earlier) and complex diseases such as cancer (e.g. genetically tailored Breast cancer treatments).

- Our DNA also affects drug metabolism, risk for adverse event and side-effects, as well as drug-drug interactions.

- Pharmacogenomics aims use knowledge of genetic variants to predict clinical response variability, risk for adverse events, genotype-specific dosing and treatment strategies.

Pharmacogenomics

- There are several hundred identified genes that affect drug response. Many encode drug-metabolizing enzymes (e.g. cytochrome P450s), transporters (e.g. ABC transporters) or immune responses (e.g. HLA variants)

Gene / Polymorphism	Drug	Indication	Adverse Reaction
HLA-B57:01	Flucloxacillin	Bacterial Infection	Drug induced liver injury (DILI)
HLA-B57:01	Abacavir	HIV/AIDS	Hypersensitivity
SLC01B1	Simvastatin	High LDL cholesterol	Statin Induced Myopathy
HLA-B15:02	Phenytoin	Epilepsy	Stevens Johnson Syndrome
IL28B	Peg-interferon-α and ribavirin	Hepatitis C	Lack of response
CYP2C9, VKORC1	Warfarin	Anticoagulant	Internal Bleeding
*CYP2C19*2*	Clopidogrel	Antithrombotic	Lack of response

Warfarin dosage

- Warfarin is an anti-clotting agent administered to patients at risk of developing blood clots in their hearts (e.g. from cardiac arrhythmias or mechanical value replacements)

- Two genes VKORC1 and CYP2C9 influence warfarin action.

- G1639A VKORC1 leads to less protein product and less activation of clotting proteins. Hence less warfarin is required in these patients. D36Y VKORC1 has decreased ability to bind to warfarin and hence more warfarin is required in these patients.

- I359L CYP2C9 leads to slower warfarin metabolism and these patients should be administered lower doses.

ORIGINAL ARTICLE

A Randomized Trial of Genotype-Guided Dosing of Warfarin

Munir Pirmohamed, Ph.D., F.R.C.P., Girvan Burnside, Ph.D., Niclas Eriksson, Ph.D., Andrea L. Jorgensen, Ph.D., Cheng Hock Toh, M.D., Toby Nicholson, F.R.C.Path., Patrick Kesteven, M.D., Christina Christersson, M.D., Ph.D., Bengt Wahlström, M.D., Christina Stafberg, M.D., J. Eunice Zhang, Ph.D., Julian B. Leathart, M.Phil., Hugo Kohnke, M.Sc., Anke H. Maitland-van der Zee, Pharm.D., Ph.D., Paula R. Williamson, Ph.D., Ann K. Daly, Ph.D., Peter Avery, Ph.D., Farhad Kamali, Ph.D., and Mia Wadelius, M.D., Ph.D., for the EU-PACT Group*

ABSTRACT

BACKGROUND

The level of anticoagulation in response to a fixed-dose regimen of warfarin is difficult to predict during the initiation of therapy. We prospectively compared the effect of genotype-guided dosing with that of standard dosing on anticoagulation control in patients starting warfarin therapy.

METHODS

We conducted a multicenter, randomized, controlled trial involving patients with atrial fibrillation or venous thromboembolism. Genotyping for *CYP2C9*2*, *CYP2C9*3*, and *VKORC1* (−1639G→A) was performed with the use of a point-of-care test. For patients assigned to the genotype-guided group, warfarin doses were prescribed according to pharmacogenetic-based algorithms for the first 5 days. Patients in the control (standard dosing) group received a 3-day loading-dose regimen. After the initiation period, the treatment of all patients was managed according to routine clinical practice. The primary outcome measure was the percentage of time in the therapeutic range of 2.0 to 3.0 for the international normalized ratio (INR) during the first 12 weeks after warfarin initiation.

RESULTS

A total of 455 patients were recruited, with 227 randomly assigned to the genotype-guided group and 228 assigned to the control group. The mean percentage of time in the therapeutic range was 67.4% in the genotype-guided group as compared with 60.3% in the control group (adjusted difference, 7.0 percentage points; 95% confidence interval, 3.3 to 10.6; P<0.001). There were significantly fewer incidences of excessive anticoagulation (INR ≥4.0) in the genotype-guided group. The median time to reach a therapeutic INR was 21 days in the genotype-guided group as compared with 29 days in the control group (P<0.001).

CONCLUSIONS

Pharmacogenetic-based dosing was associated with a higher percentage of time in the therapeutic INR range than was standard dosing during the initiation of warfarin therapy. (Funded by the European Commission Seventh Framework Programme and others; ClinicalTrials.gov number, NCT01119300.)

Benefits of using a Pharmacogenetic based dosing strategy

Pirmohamed et al. N Engl J Med (2013)

Statin induced myopathy

- Statins are considered to be safe, well tolerated and the most efficient drugs for the treatment of hypercholesterolemia, one of the main risk factor for atherosclerosis (used for lowering cholesterol).

- A rare side effect is a sensation of muscle burning and or muscle weakness (myalgia)

- Multiple mutations in the liver drug transporter SLCO1B1 have now been associated elevated levels of statins in the body and an increased risk of statin-induced myopathy.

- Genotyping these variants may help to achieve the benefits of statin therapy more safely and effectively.

Abacavir hypersensitivity

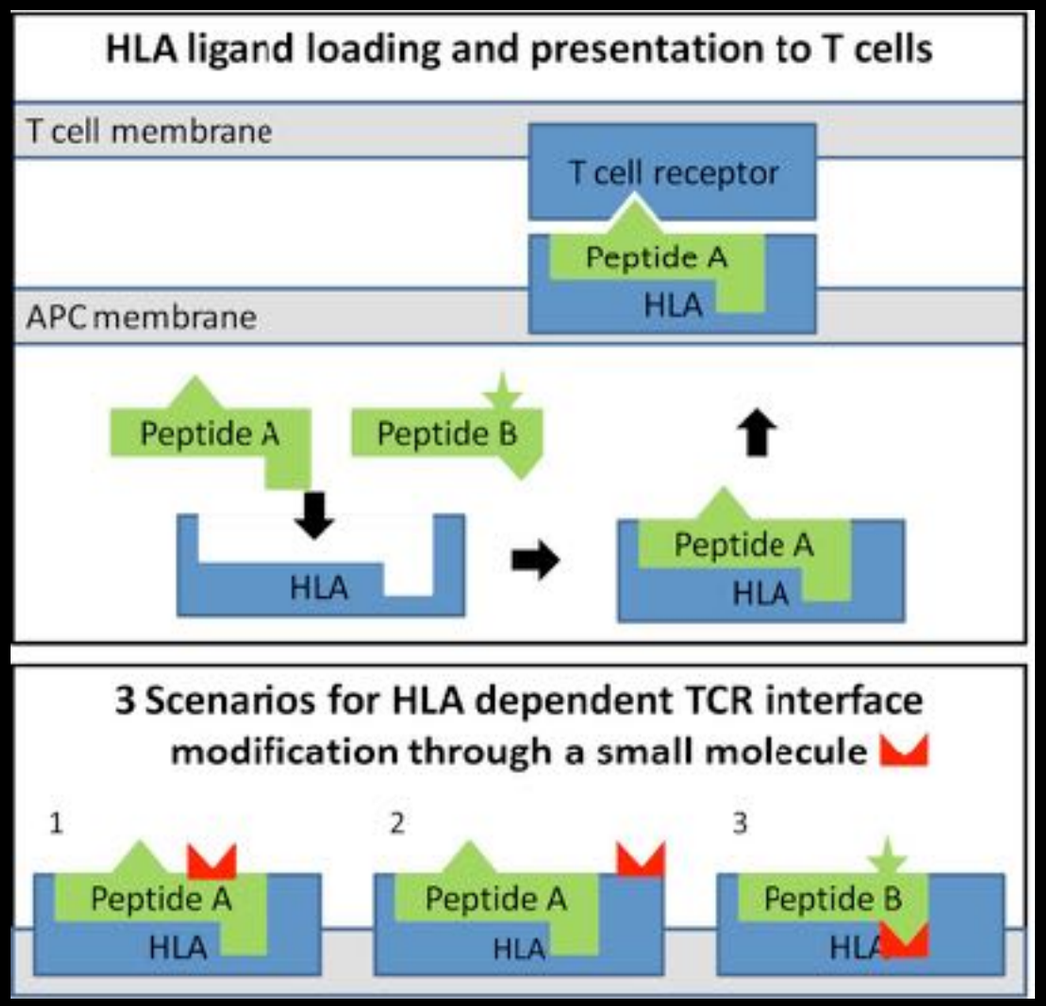
- Abacavir is a nucleoside reverse transcriptase inhibitor used for treating HIV infection.

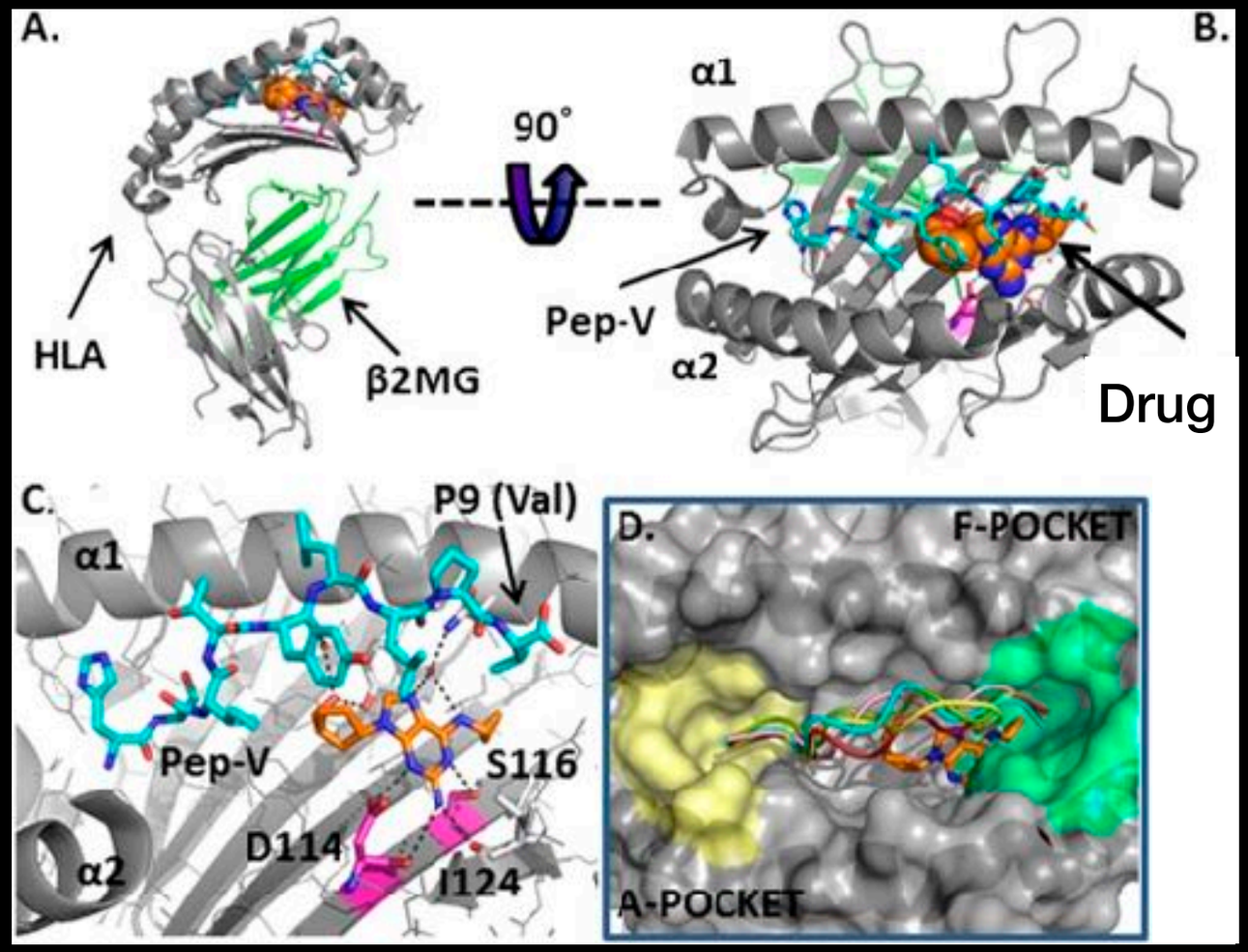
- A rare adverse effect from abacavir is an immune system mediated hypersensitivity reaction, which can be severe and potentially life-threatening.

- Hypersensitivity has been known for some time to be associated with the presence ofHLA-B*5701.

- A screening test for the HLA-B*5701 allele can assist clinicians to identify patients who are at risk of developing a hypersensitivity reaction to abacavir.

Detailed adverse molecular mechanism now known



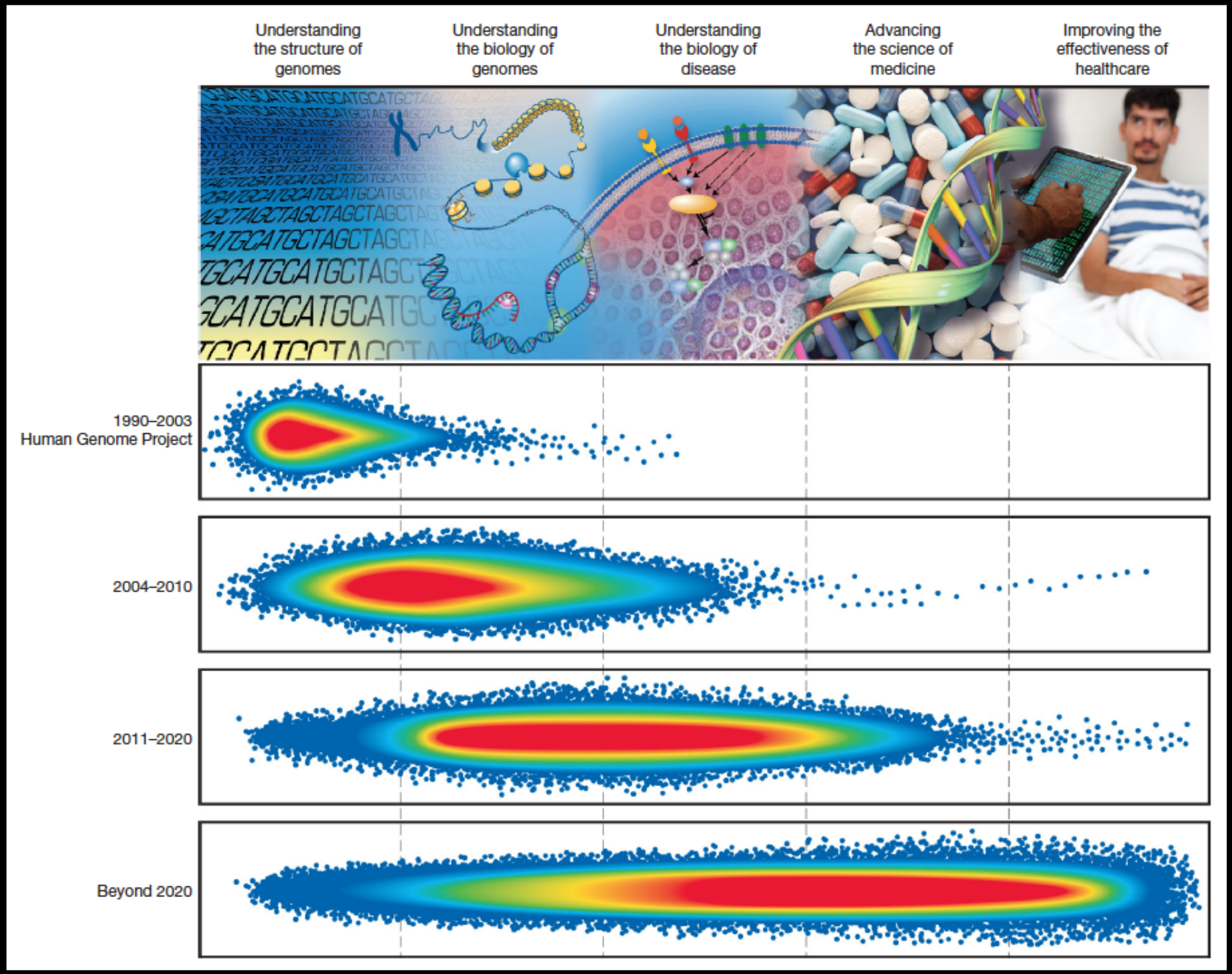


Drug hypersensitivity caused by alteration of the MHC-presented self-peptide repertoire, PNAS (2012)

Summary

- Genomics approaches have been very successful in identifying the genetic basis of rare diseases.
- Genetic testing for complex diseases is difficult but some successes do exist. In many respects this is the next frontier for genomic medicine.
- Pharmacogenomics is an exciting area that is likely to yield many further medically relevant advances in the not too distant future

Charting a course for genomic medicine



Eric Green et al., Nature 2011

Summary

- Genomics approaches have been very successful in identifying the genetic basis of rare diseases.
- Genetic testing for complex diseases is difficult but some successes do exist. In many respects this is the next frontier for genomic medicine.
- Pharmacogenomics is an exciting area that is likely to yield many further medically relevant advances in the not too distant future