

Laboratory of Systems Biology and Genetics

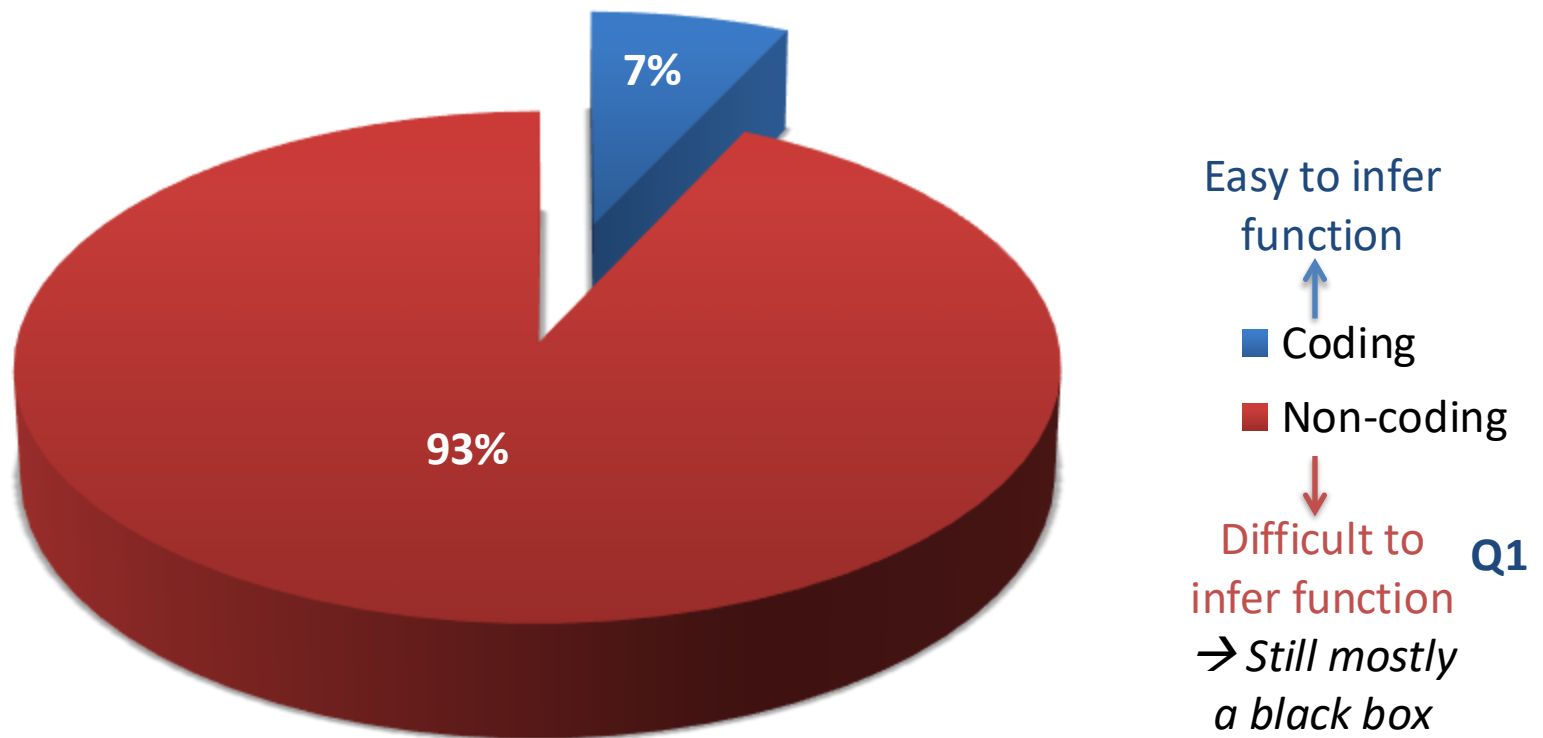
Lecture 4: Regulatory variation & the dawn of precision medicine



**All SNPs linked to quantitative traits
/ diseases so far**

The majority of trait-associated variants (SNPs) map to non-coding regions

>1,200 GWAS → 6,500 disease- or trait-predisposing SNPs

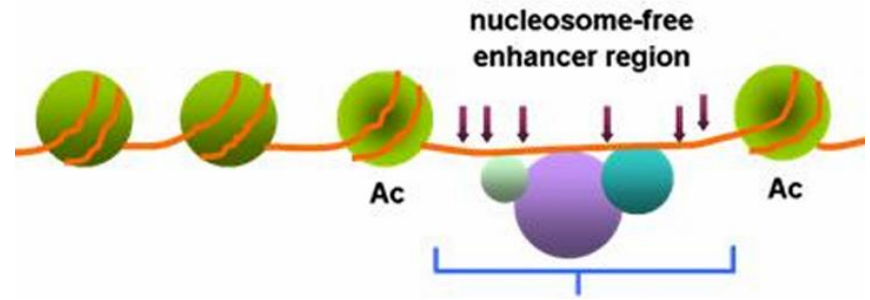
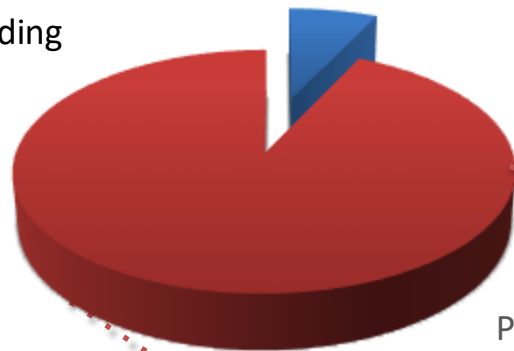


The majority of trait-associated variants (SNPs) map to

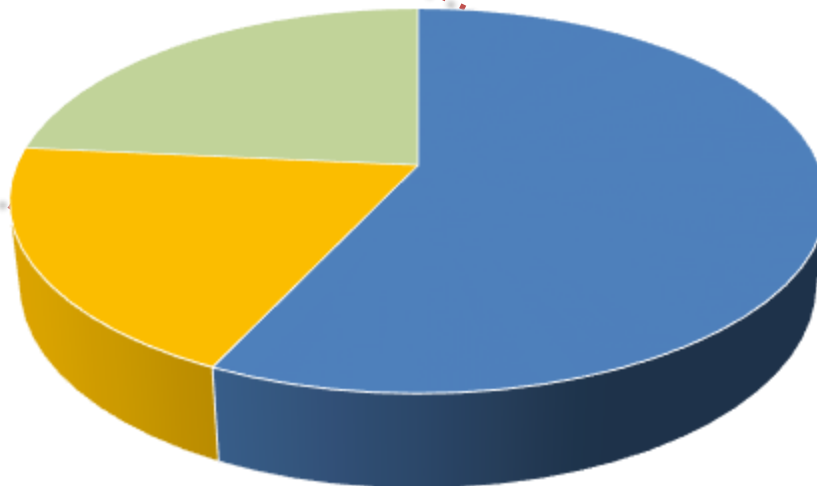
Q1 *functional* non-coding regions

■ Coding

■ Non-coding



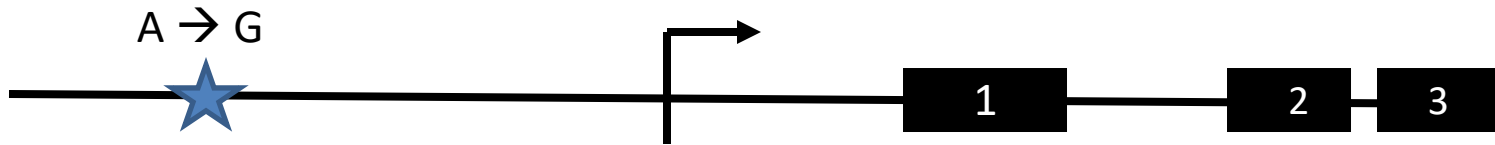
Partition of non-coding GWAS SNPs according to their location



■ GWAS SNPs in nucleosome free regions (NFR)

■ GWAS SNPs in perfect LD with those NFR SNPs

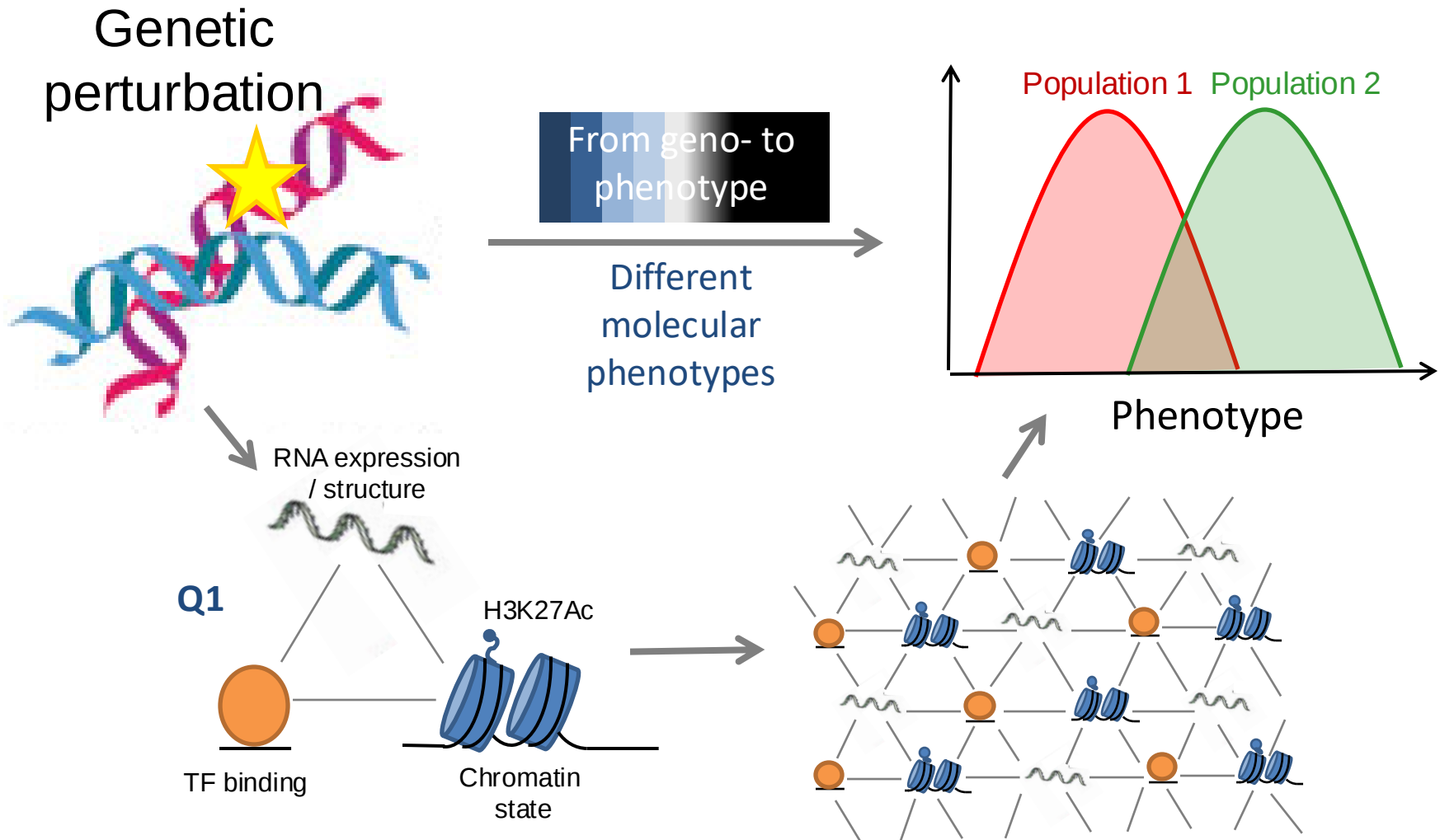
Regulatory polymorphisms



- humans are heterozygous at more functional *cis*-regulatory sites than at amino acid positions ([Rockman and Wray. Mol. Biol. Evol., 2002: 19, 1991](#)).
 - Case study with the CC chemokine receptor 5, a major chemokine co-receptor of HIV-1 necessary for viral entry into cells
 - G to A SNP of *CCR5* at –2459 nt
 - *CCR5* density – low (homozygous GG), intermediate (GA), and highest (homozygous –AA) (correlates with disease progression, i.e. fastest in AA individuals)
- ([Salkowitz et al., Clin. Immunol., 2003: 108, 234](#))
- Potential site for NF-κB/Rel binding?

Whole genome association studies

From association to molecular mechanism



How genomic variation and especially regulatory variation results into phenotypic variation is still a big black box. That is why we need to bridge this gap by looking at how genomic variation causes molecular variation which then leads to organismal phenotypic variation. 5

Whole genome association studies

Gene expression variation: mapping eQTLs

Transcript abundance = a quantitative trait that can be mapped with considerable power

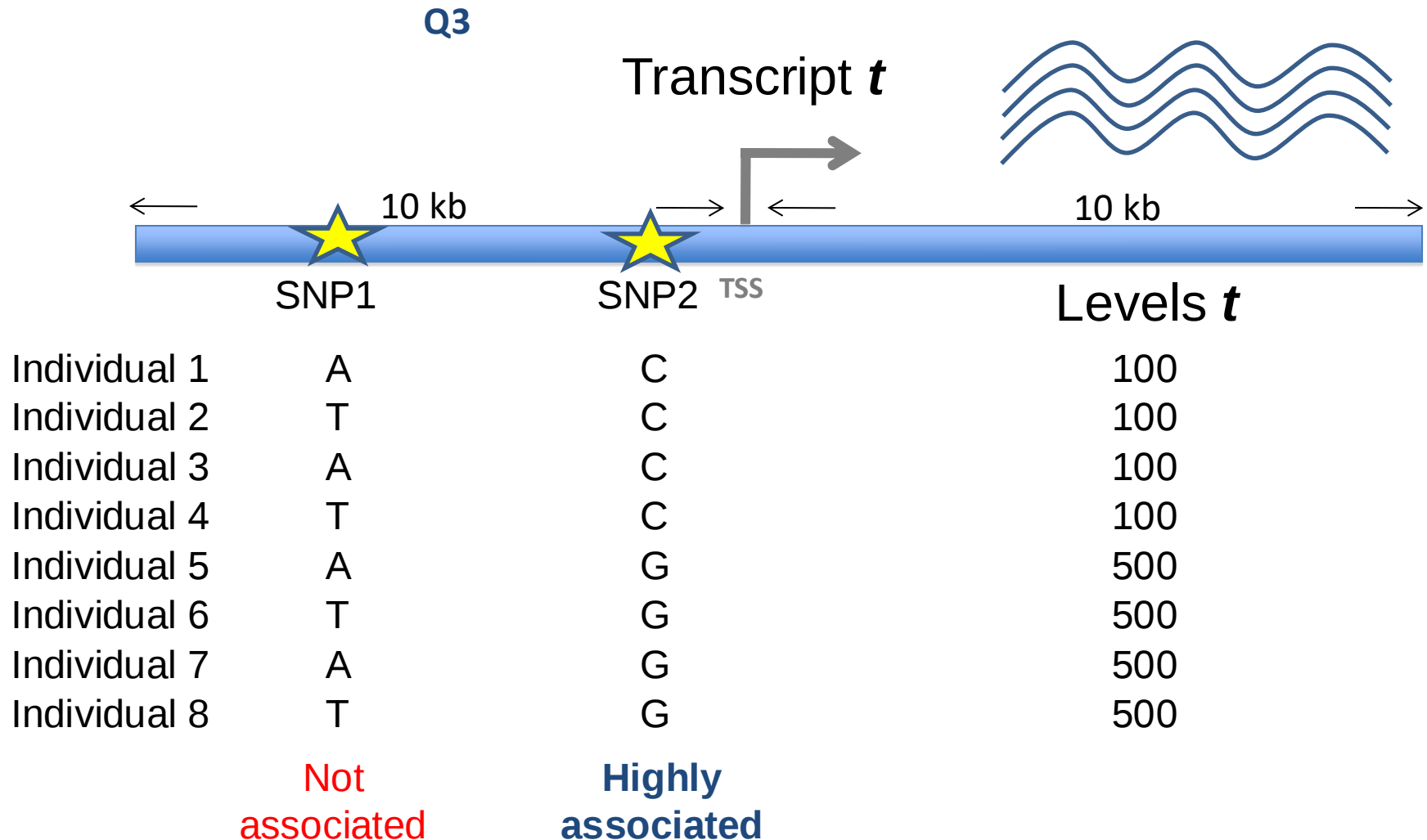


Heritability (H^2) = genetic variance over total trait variance with 0 = no genetic effects and 1 = all variance is under genetic control

Q3 eQTL is a locus that induces a heritable change in gene expression

Whole genome association studies

Gene expression variation: mapping cis-eQTLs



Whole genome association studies

Gene expression variation: mapping cis-eQTLs

Q4



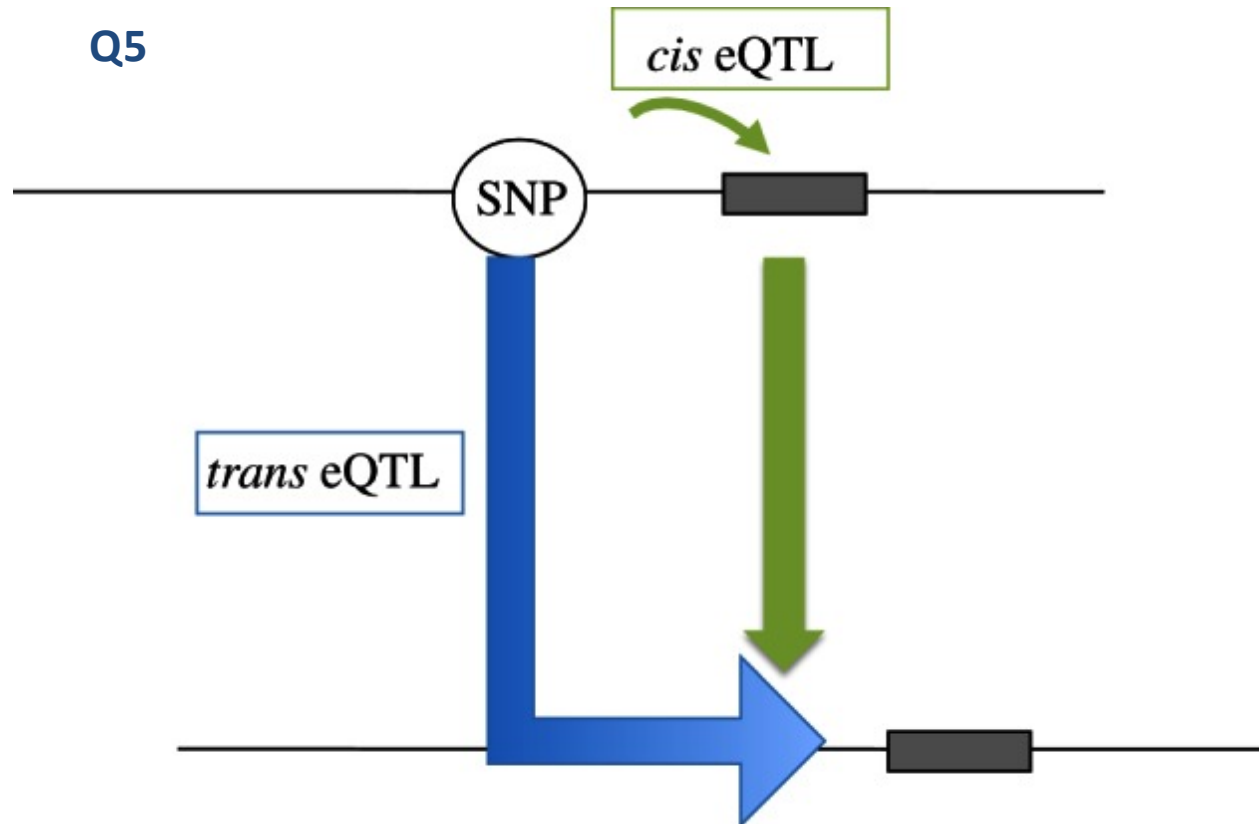
cis-eQTLs are highly abundant in the human genome!

- GTEx (Genotype Tissue Expression Consortium): **cis-eQTLs for about 18k genes with a total of 4,278,636 genetic variants** (~50 tissues from up to ~1000 postmortem donors)
- cis-eQTLs are enriched about 1.5-fold among GWAS QTLs (compared to all variants tested in GWAS)

GTEx, Science, 2020

Many eQTLs are not *cis* but *trans*

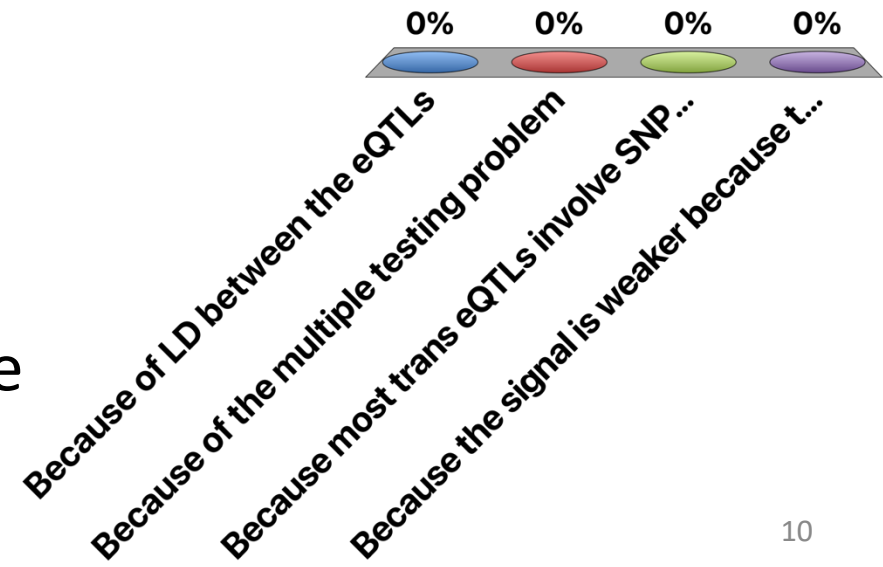
(*trans*: they operate at a distance)



Need many more samples / individuals

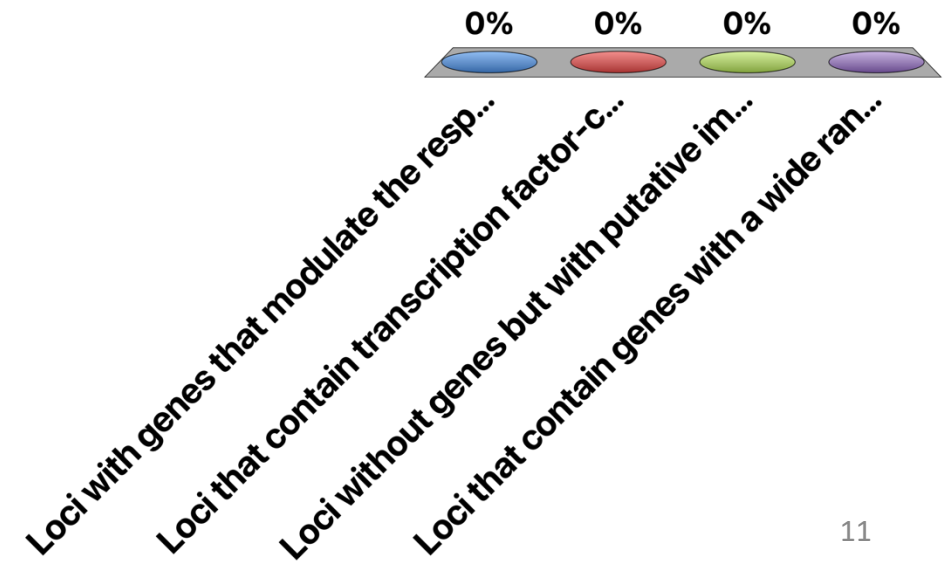
Q5 to map *trans* eQTLs, why?

- A. Because of LD between the eQTLs
- B. Because of the multiple testing problem
- C. Because most *trans* eQTLs involve SNPs with a low MAF
- D. Because the signal is weaker because these variants modulate gene expression at a far distance



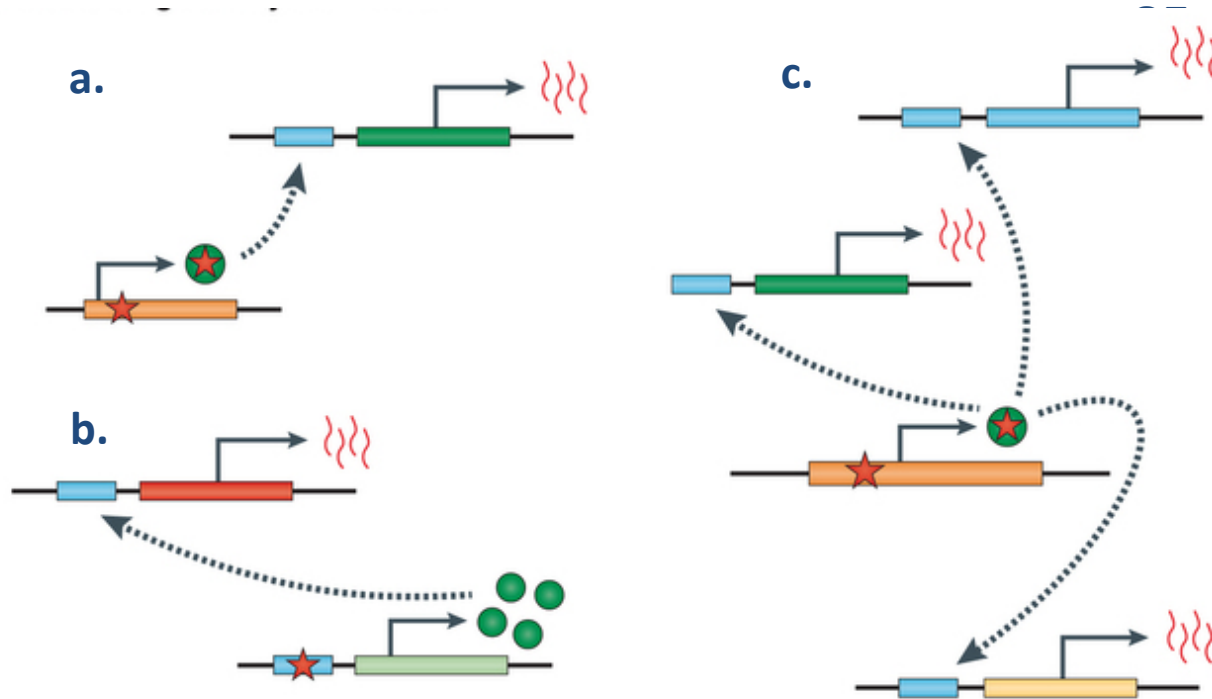
Trans-eQTLs map to which loci?

- A. Loci with genes that modulate the response to the environment
- B. Loci that contain transcription factor-coding genes
- C. Loci without genes but with putative important regulatory elements
- D. Loci that contain genes with a wide range of molecular functions



Whole genome association studies

Q6 Gene expression: mapping *trans* eQTLs



Albert and Kruglyak, Nature Reviews genetics, 2015

Trans-eQTLs are due to polymorphisms that alter the function **(a)** or expression **(b)** of a diffusible factor. **(c)** shows that the effect can act on many genes at once in *trans* (e.g. changes in actin levels)

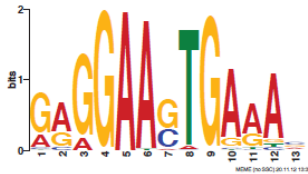
Whole genome association studies

Mapping tfQTLs: bridging genome variation with gene regulation

ChIP-seq of PU.1 in lymphoblastoid cell lines (LCLs) of several unrelated individuals

PU.1 *de novo* motif

$P=4.7e-2240$
(1000/1000)



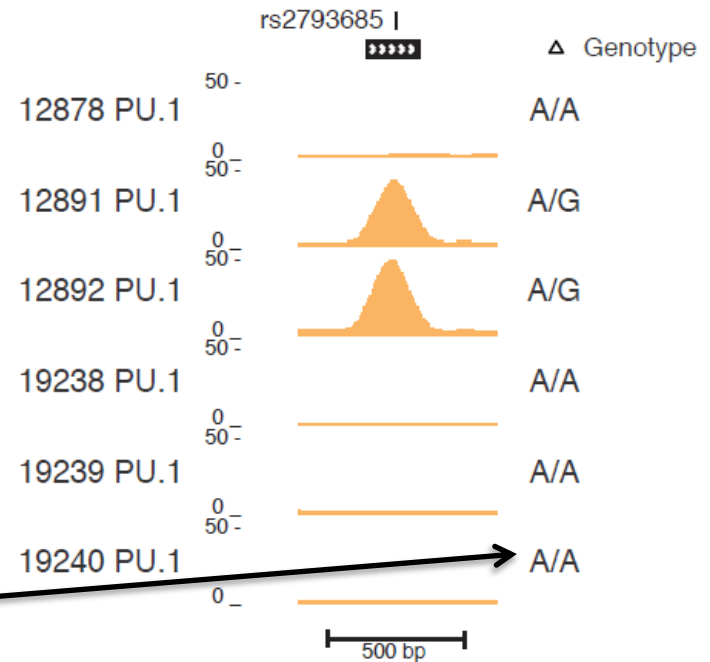
Q7

tfQTLs: Polymorphisms that induce
heritable variation in TF DNA binding

PU.1 (EICE)
motif



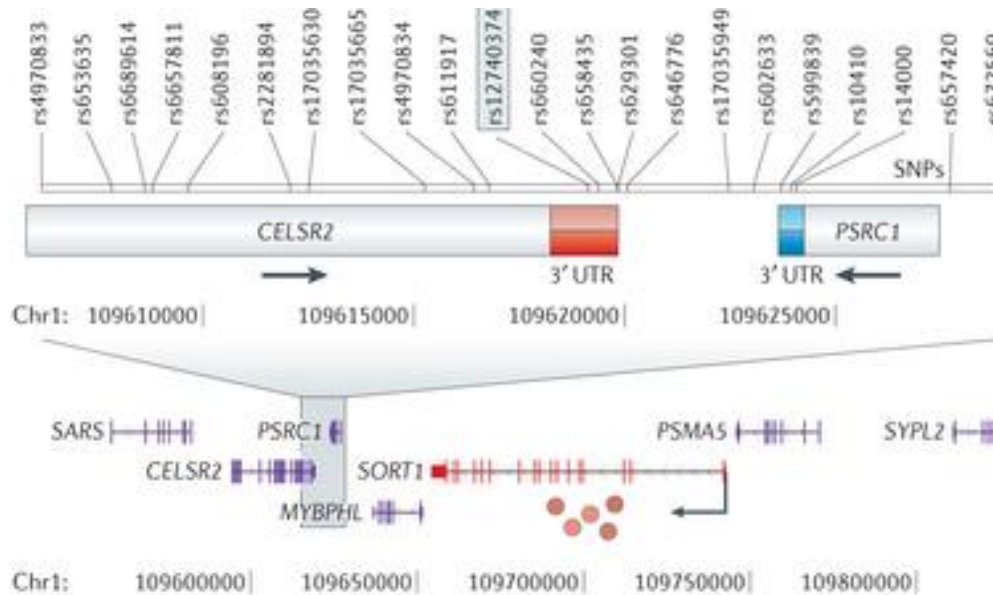
Example of a PU.1
motif-disrupting SNP



Whole genome association studies

Solving the molecular mechanism underlying GWAS variants using “xQTLs”

The variant rs12740374 is a GWAS QTL linked to variation in **myocardial infarction** susceptibility

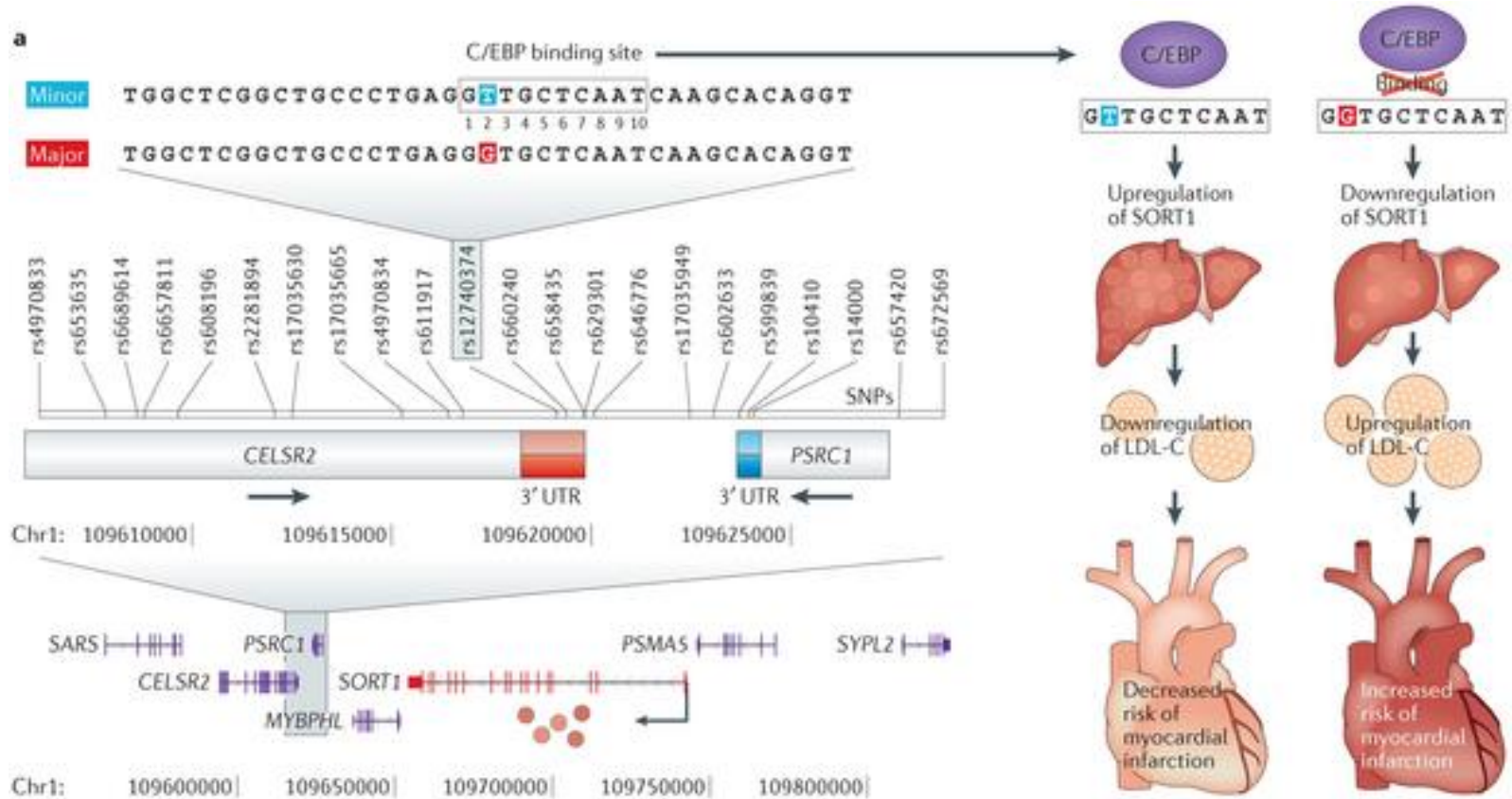


The variant is located in the 3'UTR of the *CELSR2* gene, so should we investigate *CELSR2* for its impact on myocardial infarction, or how do we dissect the underlying molecular mechanism?

Albert and Kruglyak, Nature Reviews Genetics, 2015

Whole genome association studies

Solving the molecular mechanism underlying GWAS variants using “xQTLs”



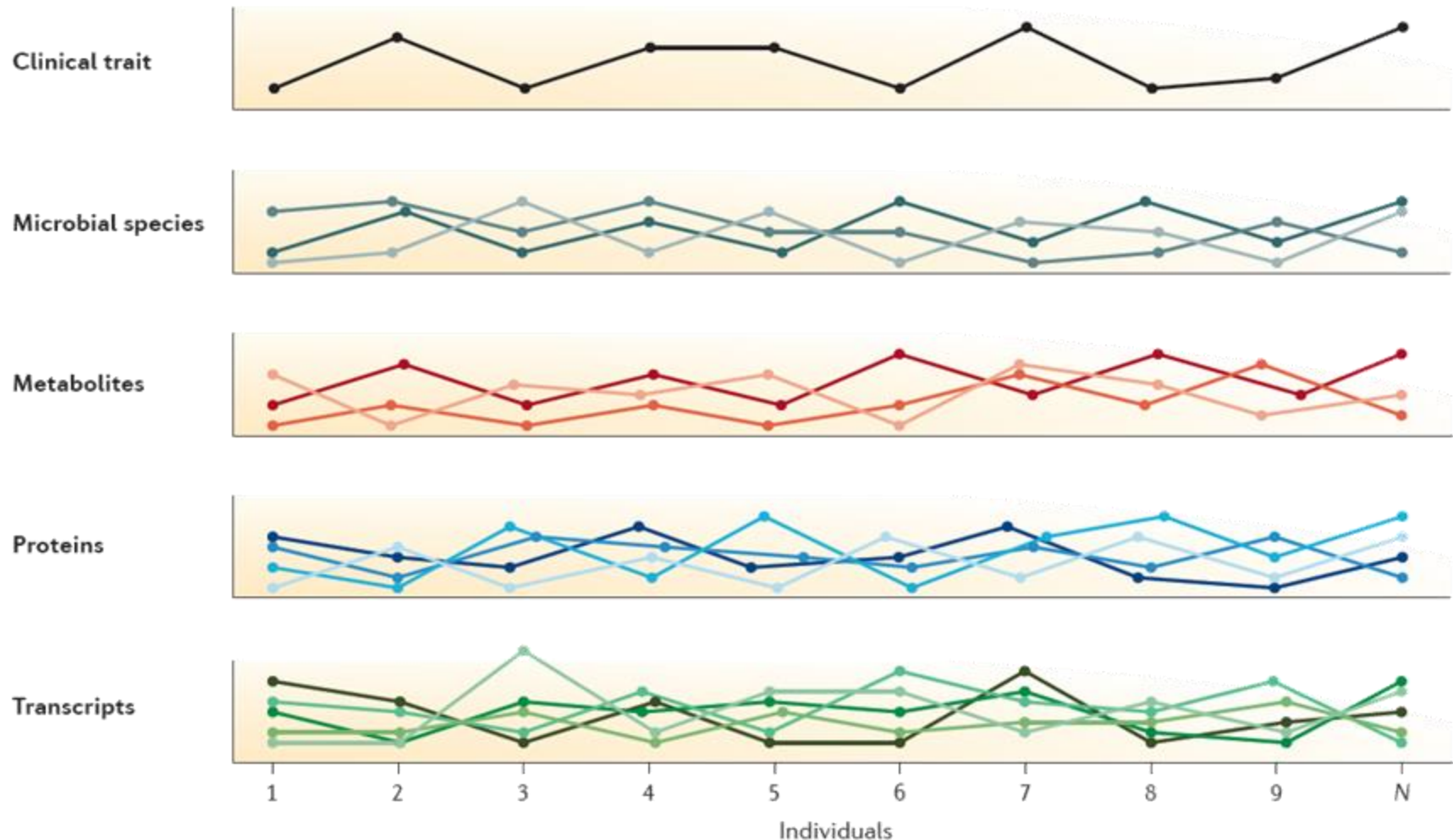
Albert and Kruglyak, Nature Reviews Genetics, 2015

- Minor allele of rs12740374 creates a TF binding site for CCAAT/enhancer-binding protein (C/EBP) → tfQTL (C/EBP)
- Binding of C/EBP at this site leads to increased expression of the sortilin 1 (*SORT1*) gene (40 kb downstream) in liver cells → eQTL (*SORT1*) (chromatin conformation!)

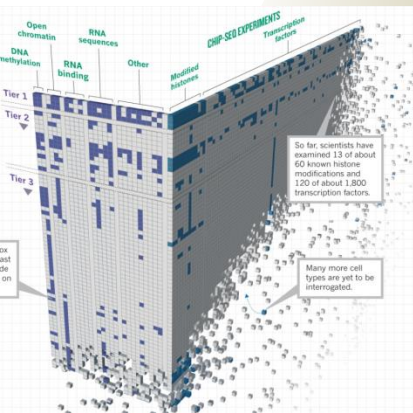
Q8 → overlapping these “x”QTLs uncovers the molecular mechanism

Whole genome association studies

The future: Mapping xQTLs: bridging genome variation with phenotypic variation at all levels



From 1 to...



Q9 How can genomic analyses improve the practice of personalized medicine?

What is most obvious to you?

Entering the age of personalized medicine

Toward the elucidation of each person's genetic make-up

Necessary for:

Q9 1) DNA-based risk assessment for common complex disease

Breast Cancer → One of the better established disease prediction models

→ Women with **BRCA1 mutation** ~65% chance of developing breast or ovarian cancer before the age of 70. Why not all?

→ Genotype-dependent (e.g. certain alleles of the *TNRC9* gene increase the effect of having the mutant *BRCA1* allele)

The New York Times

The Opinion Pages

WORLD U.S. N.Y. / REGION BUSINESS TEC.

OP-ED CONTRIBUTOR

My Medical Choice

By ANGELINA JOLIE

Published: May 14, 2013 | 1712 Comments

LOS ANGELES



- Has Mother who died at 57 of breast cancer
- Had double Mastectomy
- Carries BRCA1 allele

Entering the age of personalized medicine

Toward the elucidation of each person's genetic make-up

Necessary for:

1) DNA-based risk assessment for common complex disease

- Breast Cancer → One of the better established disease prediction models
- Women with **BRCA1 mutation** ~65% chance of developing breast or ovarian cancer before the age of 70. Why not all?
 - Genotype-dependent (e.g. certain alleles of the *TNRC9* gene increase the effect of having the mutant *BRCA1* allele)
 - *How to calculate the risk?*
 - Several different types of prediction models exist:
 - Incorporate breast and ovarian cancer in 1st- and 2nd degree relatives
 - Age of cancer onset
 - Incorporate racial / ethnic backgrounds (e.g. Ashkenazi Jews)
 - Other variables defining the person's personal and family history
 - Tools: Logistic regression, Bayesian etc.

Entering the age of personalized medicine

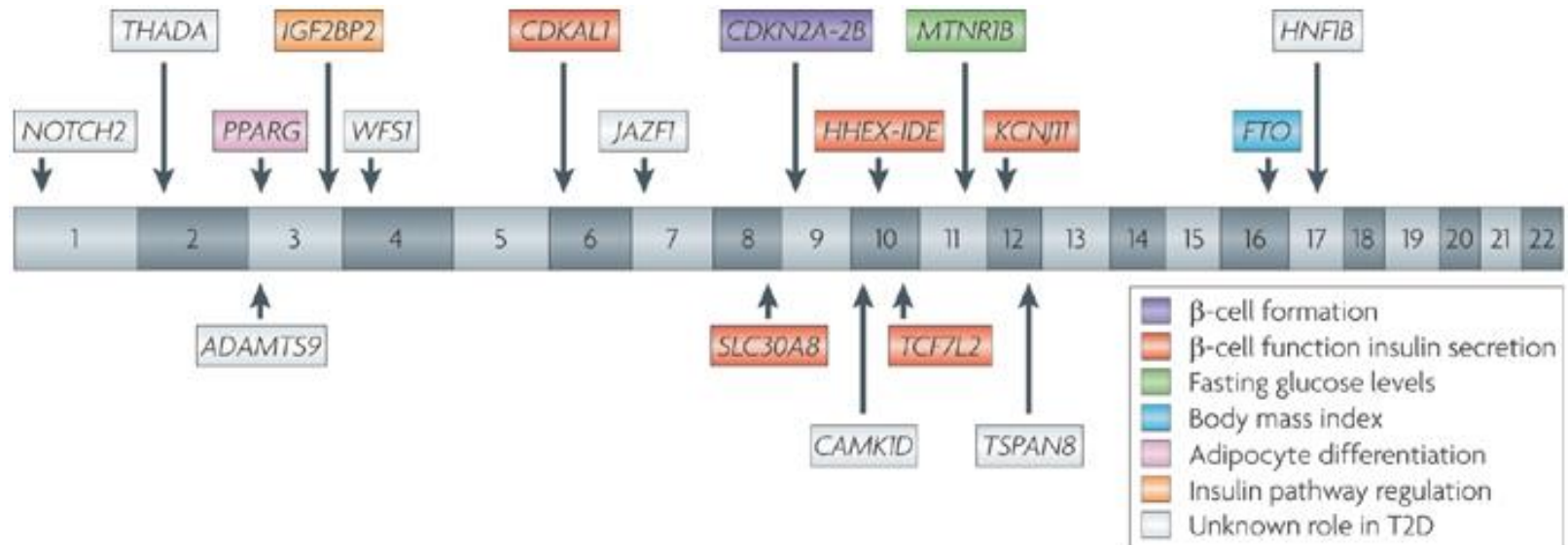
Toward the elucidation of each person's genetic make-up

Q9

Necessary for:

2) Identification of novel molecular signatures for disease diagnosis, prognosis, or drug design

e.g. Type II Diabetes → many possible disease predisposition markers
→ how to calculate probability?



Each person may have a different genetic pre-disposition,
resulting in a different prognosis or drug treatment

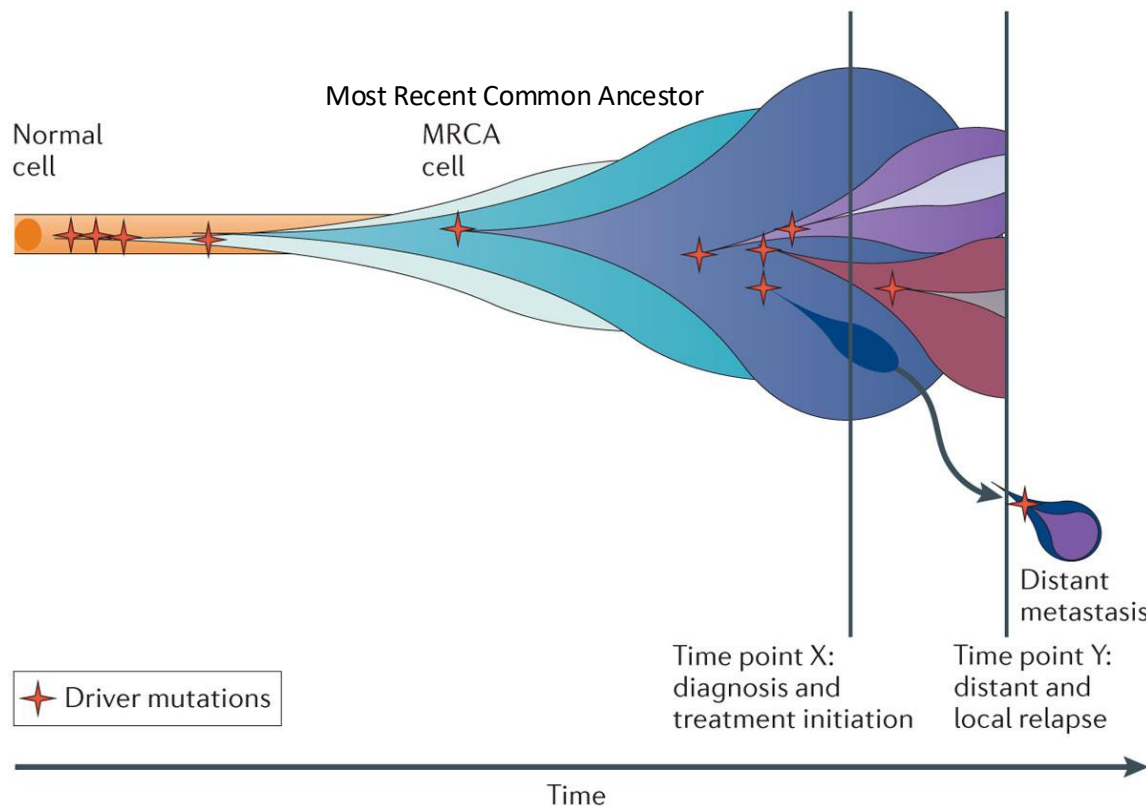
Entering the age of personalized medicine

Toward the elucidation of each person's genetic make-up

Necessary for:

2) Identification of novel molecular signatures for disease diagnosis, prognosis, or drug design

Another striking example: cancer



Q10 Cancers are genomically diverse and dynamic entities:

- Unique clones (colored bubbles) arise because of accumulating driver mutations in MRCA cell progeny
- Ongoing linear and branching evolution results in multiple subclones which drive disease relapse and metastasis.
- The dynamic clonal architecture is shaped by mutation and competition between subclones given specific environmental selection pressures, including cancer treatments.

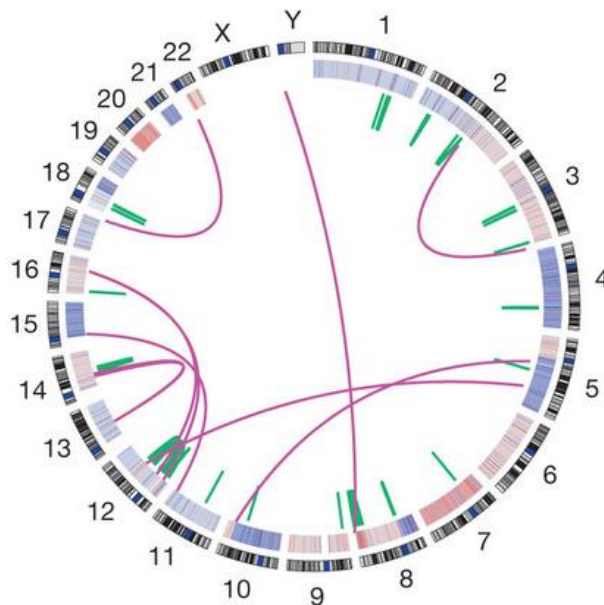
Entering the age of personalized medicine

Toward the elucidation of each person's genetic make-up

Necessary for:

2) Identification of novel molecular signatures for disease diagnosis, prognosis, or drug design

Another striking example: colorectal cancer genome of one patient (wildtype versus cancer cells)



Interchromosomal translocations

Intrachromosomal translocations

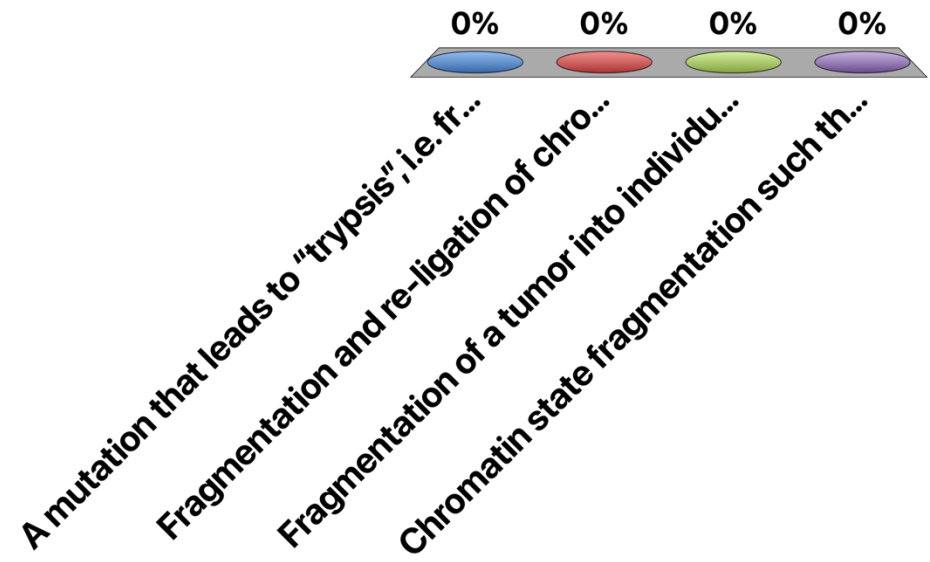
Amplifications and deletions

Individual nucleotide mutations not shown

The molecular diversity of human cancer is staggering (scary!)
(Patients with = disease) $\neq$ (Patients with = underlying biological disorder)

What is chromotrypsis (cancer context)?

- A. A mutation that leads to “trypsis”, i.e. fragmentation of a cell?
- B. Fragmentation and re-ligation of chromosomes
- C. Fragmentation of a tumor into individual cells resulting in cell spread and metastasis
- D. Chromatin state fragmentation such that all genes on a particular chromosome become activated



Entering the age of personalized medicine

Toward the elucidation of each person's genetic make-up

Necessary for:

3) A DNA-guided therapy and dose selection

A person's genetic make-up significantly affects the efficacy of a drug

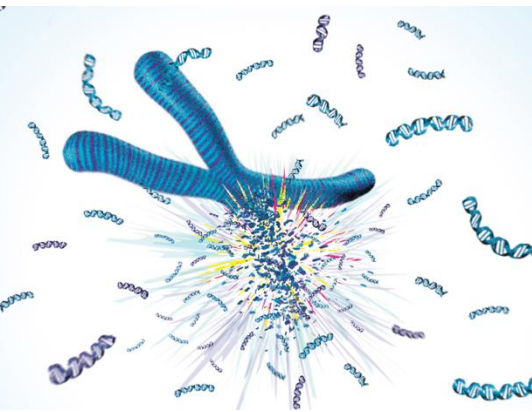
- Polymorphisms in the *VKORC1* and *CYP2C9* genes dictate the effective dose levels of the anti-coagulant **Warfarin**
- Polymorphisms in the *UGT1A1* gene correlate with increased toxicity of the anti-colon cancer drug **Irinotecan**
- Polymorphisms in the *MTHFR* gene are associated with increased toxicity of **Methotrexate** used to treat Crohn's disease
- Polymorphisms in the *CYP2D6* gene dictates the probability of relapse in women with breast cancer treated with **Tamoxifen**

Entering the age of personalized medicine

Toward the elucidation of each person's genetic make-up

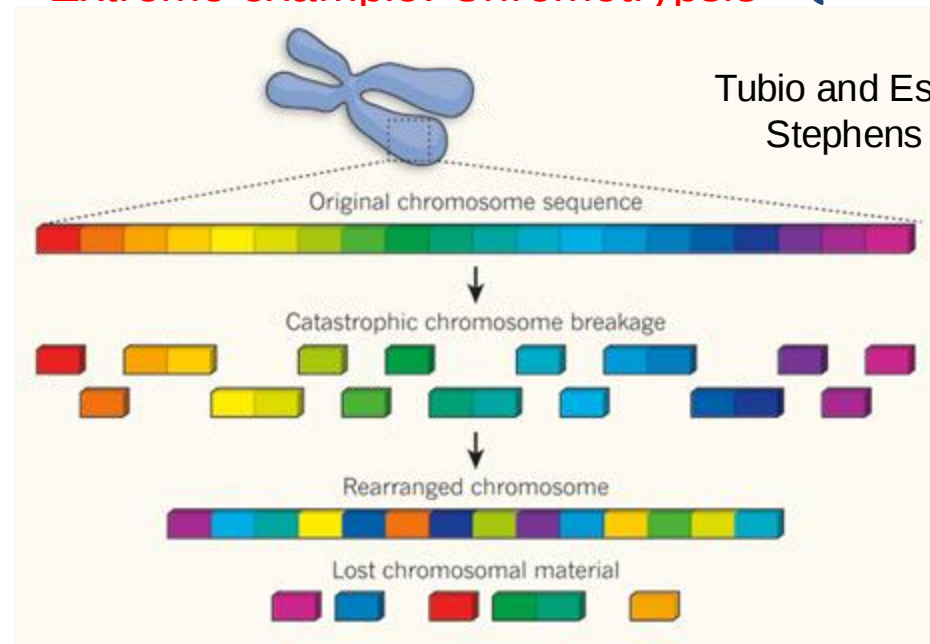
Necessary for:

2) Identification of novel molecular signatures for disease diagnosis, prognosis, or drug design



EMBL/P. Riedinger

Extreme example: Chromotrypsis Q10



Tubio and Estevill, Nature, 2011;
Stephens et al., Cell, 2011

- 1000's of clustered chromosomal rearrangements in a single catastrophic event in confined genomic regions (both cancer (2-3%) and congenital diseases)
- This phenomenon opposes the conventional theory that cancer is the gradual acquisition of genomic rearrangements and somatic mutations over time

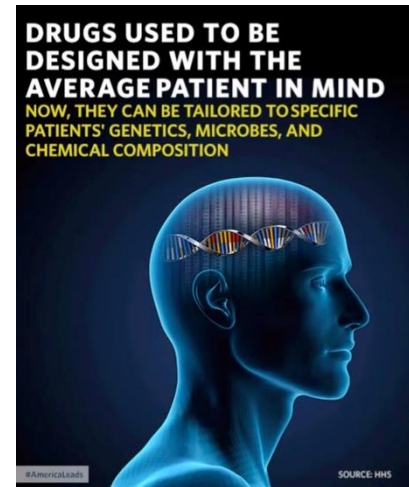
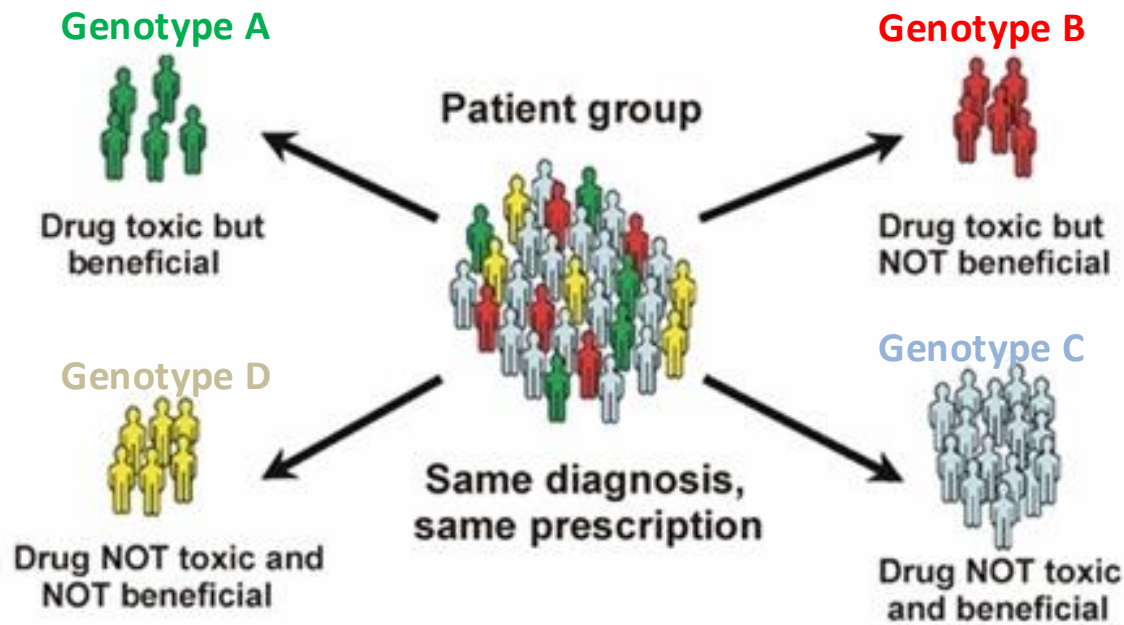
Entering the age of personalized medicine

Toward the elucidation of each person's genetic make-up

Necessary for:

3) A DNA-guided therapy and dose selection

New strategy in clinical trials to test the efficacy of a new drug:

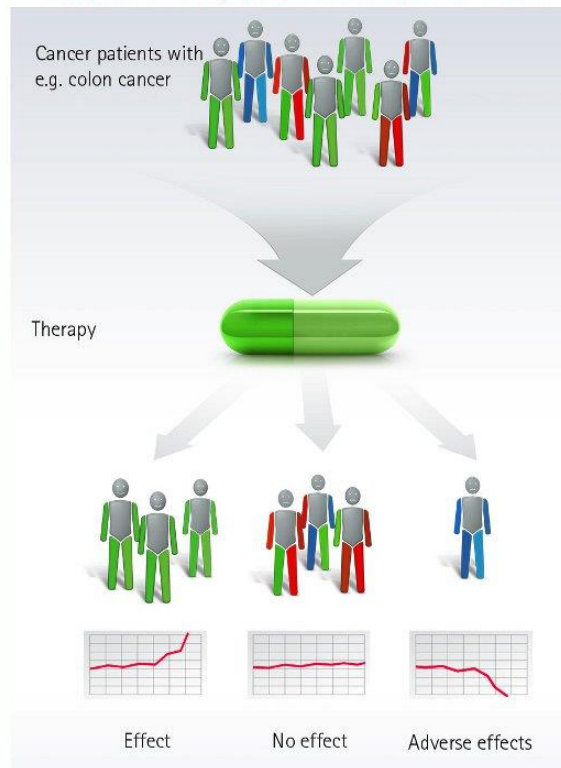


Q9

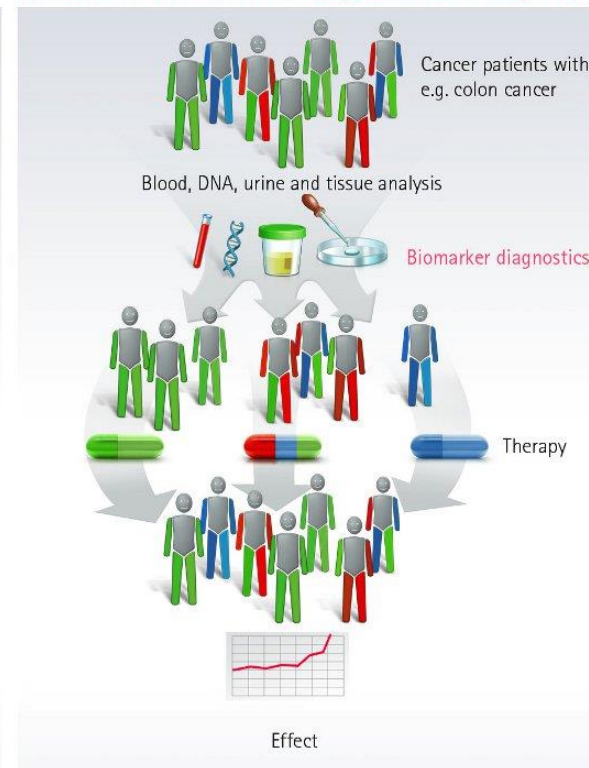
Drug may be as efficient as existing drug on a whole, but optimal for a specific subpopulation:
Genotyping is now essential in trials

Personalized medicine: tailored treatments

Medicine of the present: one treatment fits all



Medicine of the future: more personalized diagnostics



*Different people respond differently to the same therapy: while one treatment brings about the desired success in one group of patients with e.g. colon cancer, it does not change the condition of other groups at all, or even leads to adverse effects (left). The reason: the genetic makeup and metabolic profile of each individual patient influences the effect of a drug. Personalized medicine takes these individual patterns of cellular and metabolic products into account in the diagnostic phase: **biomarker diagnostics** separates patients into groups with similar characteristics, and provides information on the best individual treatment. This should enable all patients to benefit from their own, "personal" therapy.*

<http://pharma.bayer.com/en/innovation-partnering/research-and-development-areas/oncology/personalized-medicine/#&gid=1&pid=1>

Q9 How can genomic analyses improve the practice of personalized medicine?

Summary

- A. DNA-based risk assessment of disease
- B. Uncovering the patient-specific mechanism underlying disease
- C. Pharmacology / drug selection and dosage

From precision medicine to genetic engineering



Entering the age of personalized medicine

From a person's genetic make-up to gene therapy

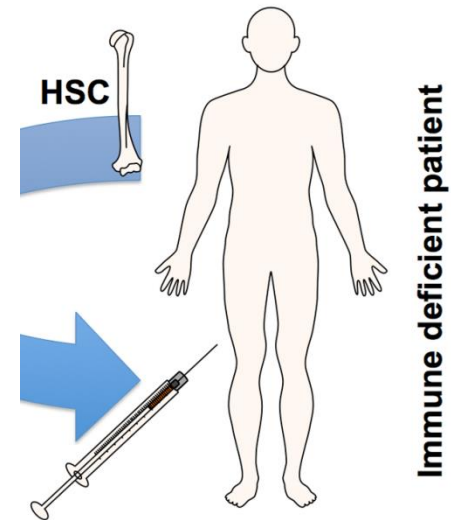
RIP: «Bubble boy» David



Entering the age of personalized medicine

From a person's genetic make-up to gene therapy

RIP: «Bubble boy» David



- Received unmatched bone marrow transplant
- Died of lymphoma (12 yrs old)

SCID: severe combined immunodeficiency

Entering the age of personalized medicine

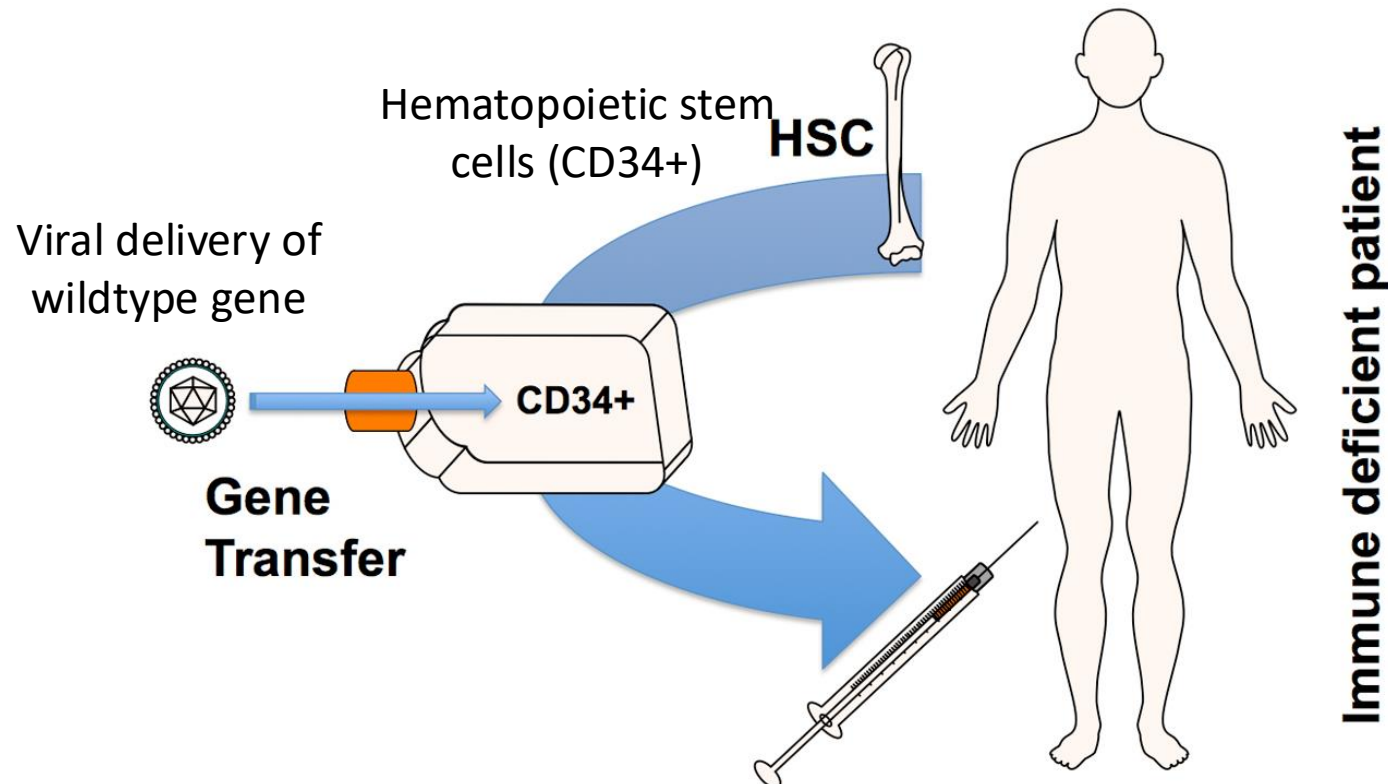
From a person's genetic make-up to gene therapy

RESEARCH ARTICLE

IMMUNODEFICIENCY Gaspar et al., Science Translational Medicine, 2011

Q11

Long-Term Persistence of a Polyclonal T Cell Repertoire After Gene Therapy for X-Linked Severe Combined Immunodeficiency

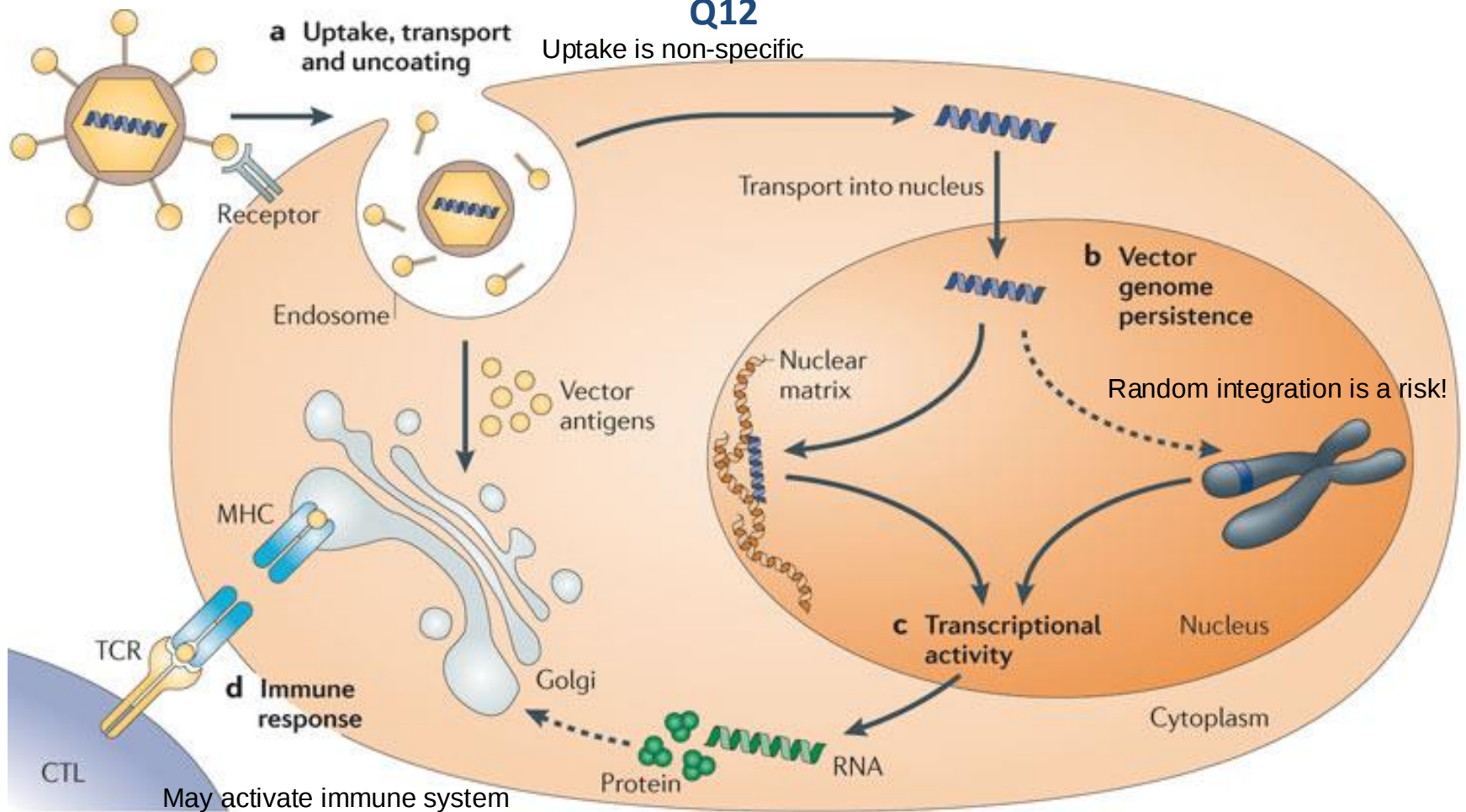


Entering the age of personalized medicine

From a person's genetic make-up to gene therapy

But a long road ahead....: challenges

Q12



Entering the age of personalized medicine

From a person's genetic make-up to gene therapy

Genome engineering: **CRISPR**, a new kid on the block



The image is a blue banner for the Nobel Prize in Chemistry 2020. At the top left is the Nobel Prize logo and the text "NOBELPRISET I KEMI 2020" and "THE NOBEL PRIZE IN CHEMISTRY 2020". At the top right is the logo of the Kungliga Vetenskapsakademien (The Royal Swedish Academy of Sciences). Below the header are two portraits of the laureates. On the left is Emmanuelle Charpentier, a woman with dark curly hair, wearing a blue jacket. On the right is Jennifer A. Doudna, a woman with blonde hair, wearing a black top. Below each portrait is their name, birth information, and affiliation.

Emmanuelle Charpentier
Born in France, 1968
Max Planck Unit for the Science of Pathogens, Germany

Jennifer A. Doudna
Born in the USA, 1964
University of California, Berkeley, USA
Howard Hughes Medical Institute

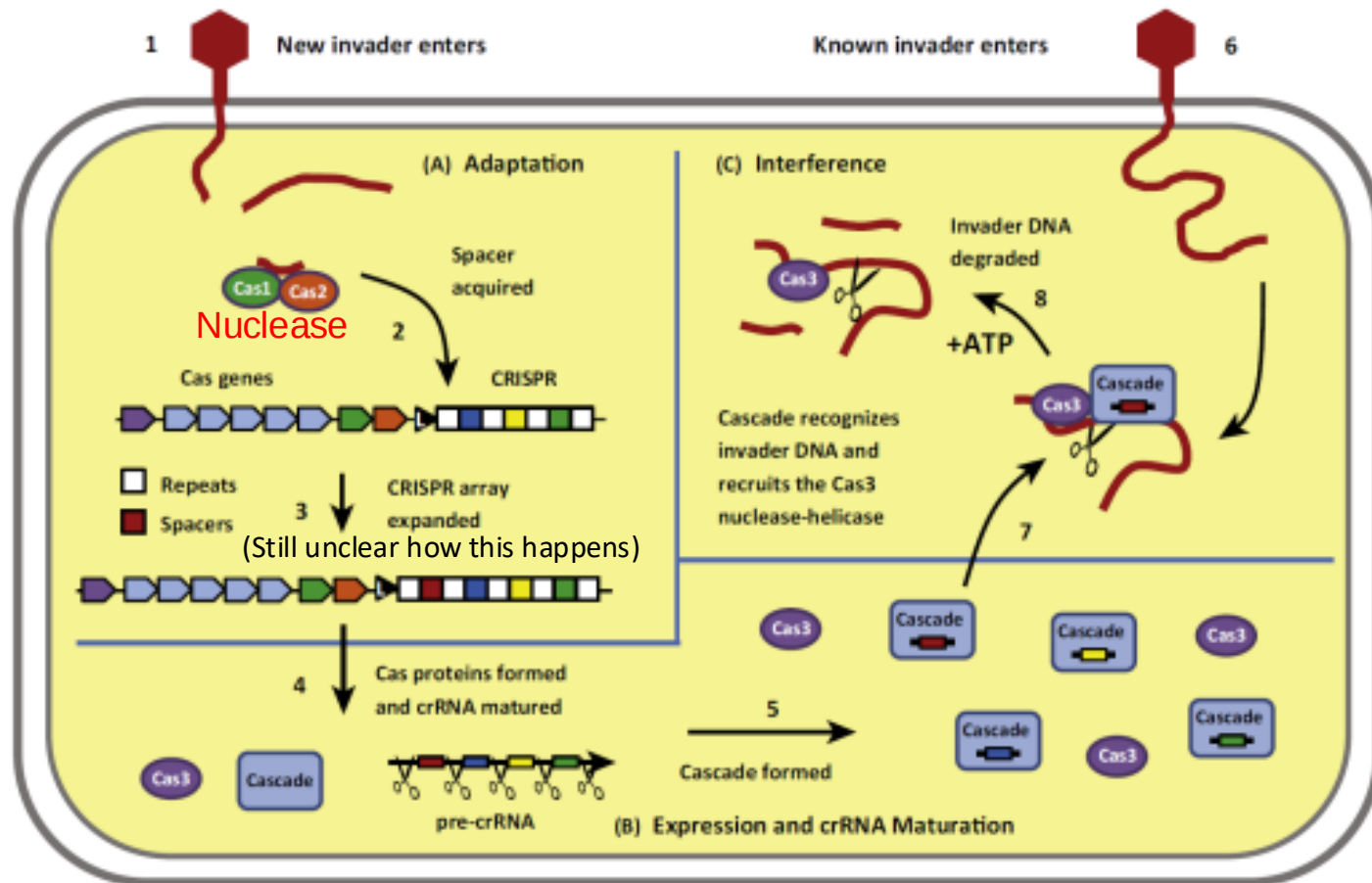
Entering the age of personalized medicine

From a person's genetic make-up to gene therapy

Genome engineering: **CRISPR**, a new kid on the block
Q13

CRISPRs (Clustered Regularly Interspaced Short Palindromic Repeats) → DNA loci that contain multiple, short, direct repetitions of base sequences (based on *Streptococcus thermophilus*, but in many bacteria)
CAS= CRISPR-associated proteins

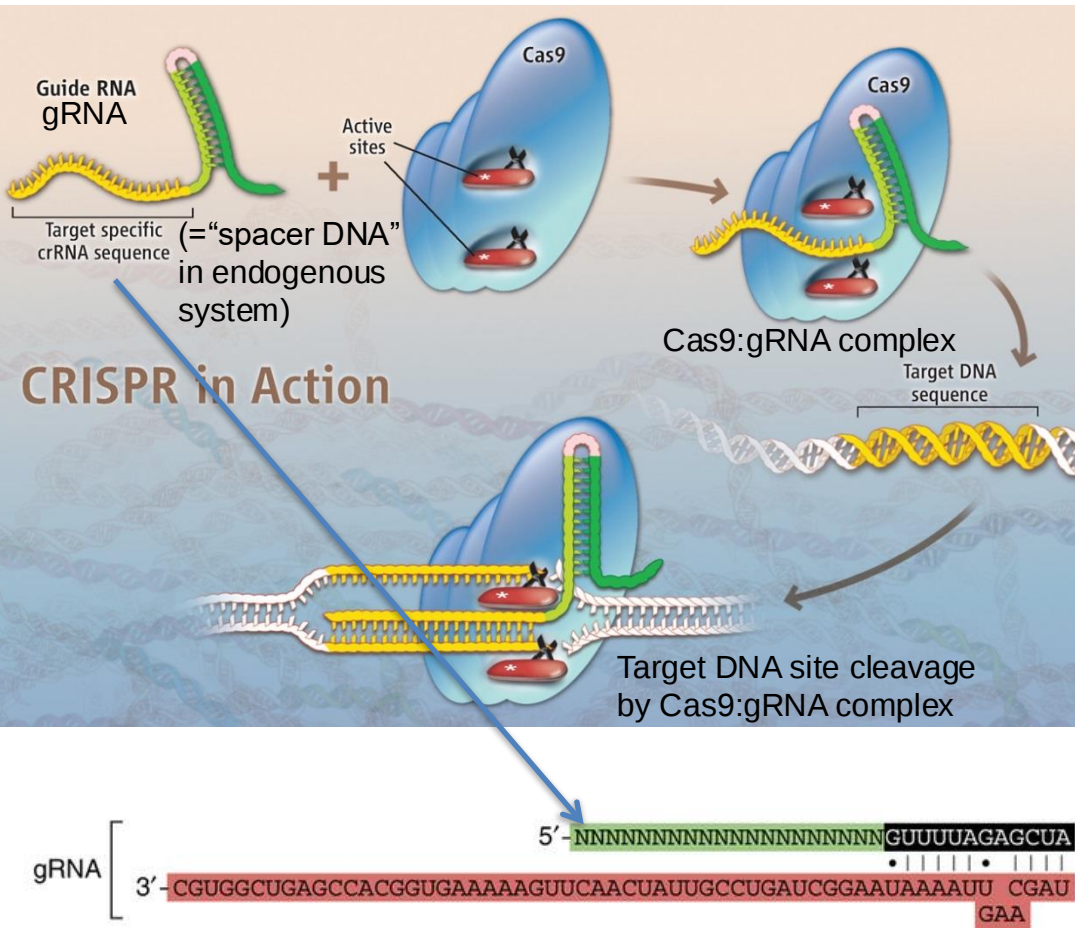
Each **repeat** = series of bases followed by the same series in reverse and then by 30 or so base pairs known as "spacer DNA" → short segments of DNA from a virus and serve as a 'memory' of past exposures.



Entering the age of personalized medicine

From a person's genetic make-up to gene therapy

Genome engineering: **CRISPR**, biotech application



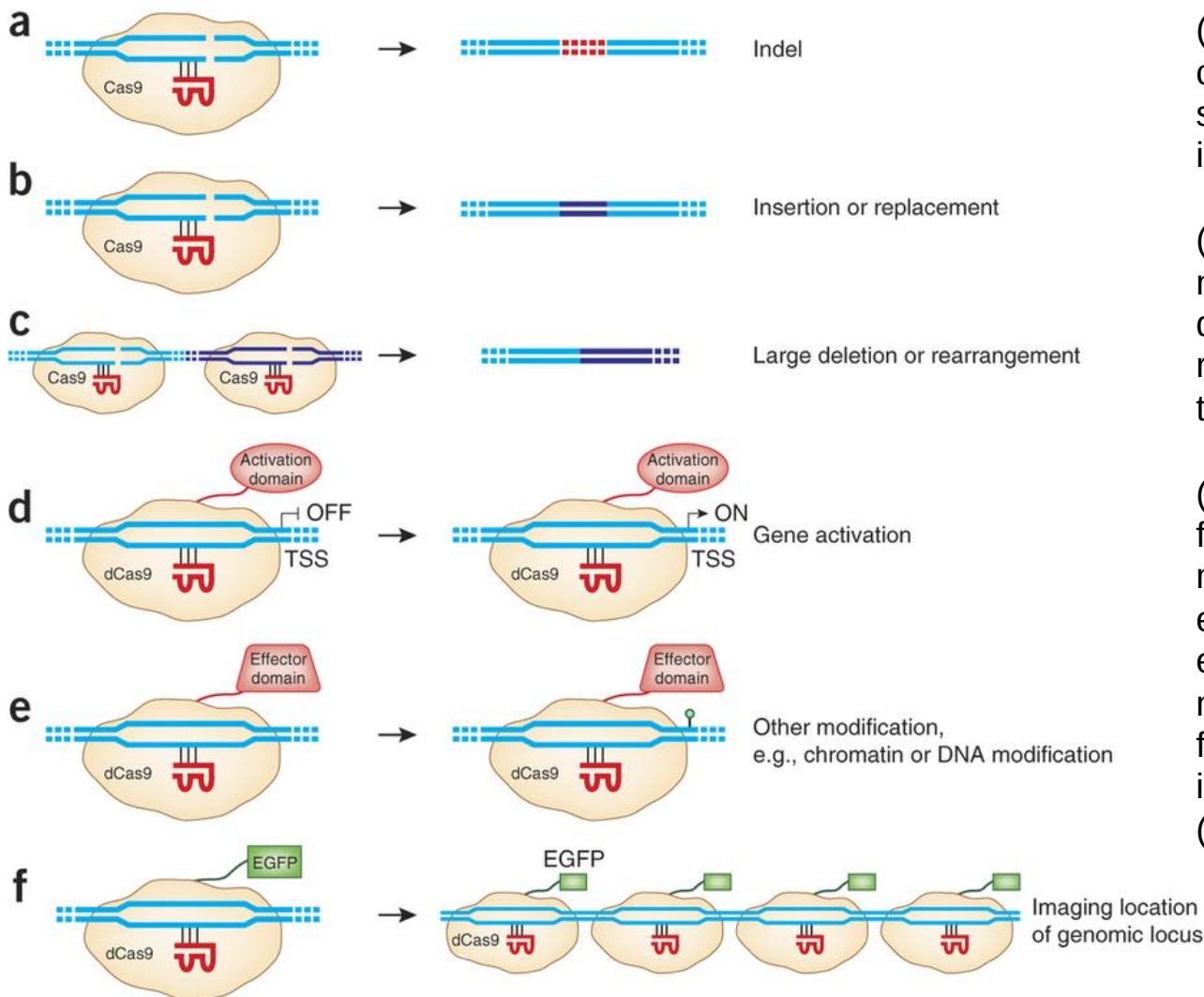
With just a guide RNA and a protein called Cas9, researchers first showed that the CRISPR system can home in on and cut specific DNA, knocking out a gene or enabling part of it to be replaced by substitute DNA.

Entering the age of personalized medicine

From a person's genetic make-up to gene therapy

Q13

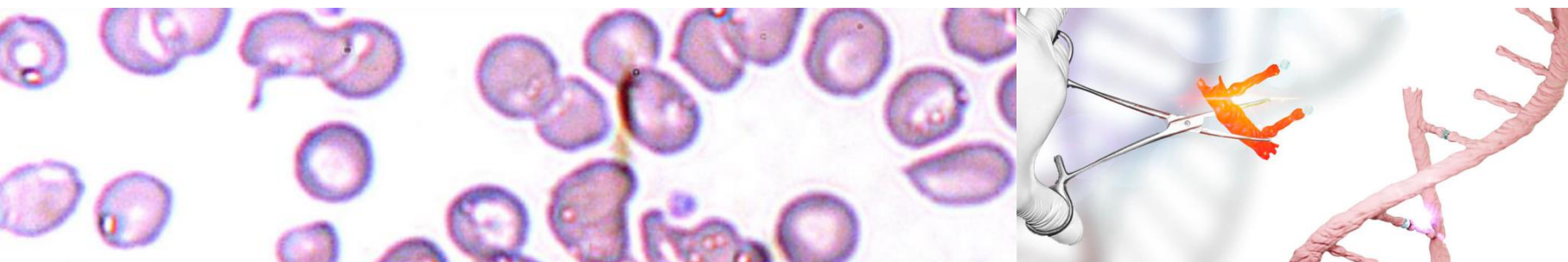
Genome engineering: **CRISPR**, biotech application



(a,b) gRNA-directed Cas9 nuclease can induce indel mutations (a) or specific sequence replacement or insertion (b).

(c) Pairs of gRNA-directed Cas9 nucleases can stimulate large deletions or genomic rearrangements (e.g., inversions or translocations).

(d-f) gRNA-directed dCas9 can be fused to activation domains (d) to mediate upregulation of specific endogenous genes, heterologous effector domains (e) to alter histone modifications or DNA methylation, or fluorescent proteins (f) to enable imaging of specific genomic loci. (TSS, transcription start site)



[Home](#) / [News & Opinion](#)

US Companies Launch CRISPR Clinical Trial

The Germany-based study will test an ex vivo genome-editing therapy for the inherited blood disorder β -thalassemia.

Sep 3, 2018

CATHERINE OFFORD



1.1K



10



84

The therapy (CTX001) will treat a deficiency in the production of hemoglobin in adults.

- Will NOT target genetic mutations responsible for this deficiency
- Will cleave a gene called *BCL11A* that represses the production of fetal hemoglobin
- CTX001 will be tested **ex vivo**—blood cells will be removed from the patient, edited, and then replaced.

CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia

Haydar Frangoul, M.D., David Altshuler, M.D., Ph.D., M. Domenica Cappellini, M.D., Yi-Shan Chen, Ph.D., Jennifer Domm, M.D., Brenda K. Eustace, Ph.D., Juergen Foell, M.D., Josu de la Fuente, M.D., Ph.D., Stephan Grupp, M.D., Ph.D., Rupert Handgretinger, M.D., Tony W. Ho, M.D., Antonis Kattamis, M.D., et al.

Article Figures/Media

Metrics

January 21, 2021

N Engl J Med 2021; 384:252-260

DOI: 10.1056/NEJMoa2031054

Chinese Translation 中文翻译

36 References 433 Citing Articles Letters

- Does NOT target genetic mutations responsible for this deficiency
- Mutates a binding site for GATA1 in a key *BCL11A* enhancer that results in *BCL11A* downregulation and derepression of fetal hemoglobin

*For all CRISPR-based current clinical trials, see
<https://crisprmedicineneeds.com/clinical-trials/>*

CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia

Haydar Frangoul, M.D., David Altshuler, M.D., Ph.D., M. Domenica Cappellini, M.D., Yi-Shan Chen, Ph.D., Jennifer Domm, M.D., Brenda K. Eustace, Ph.D., Juergen Foell, M.D., Josu de la Fuente, M.D., Ph.D., Stephan Grupp, M.D., Ph.D., Rupert Handgretinger, M.D., Tony W. Ho, M.D., Antonis Kattamis, M.D., et al.

Article Figures/Media

Metrics

January 21, 2021

N Engl J Med 2021; 384:252-260

DOI: 10.1056/NEJMoa2031054

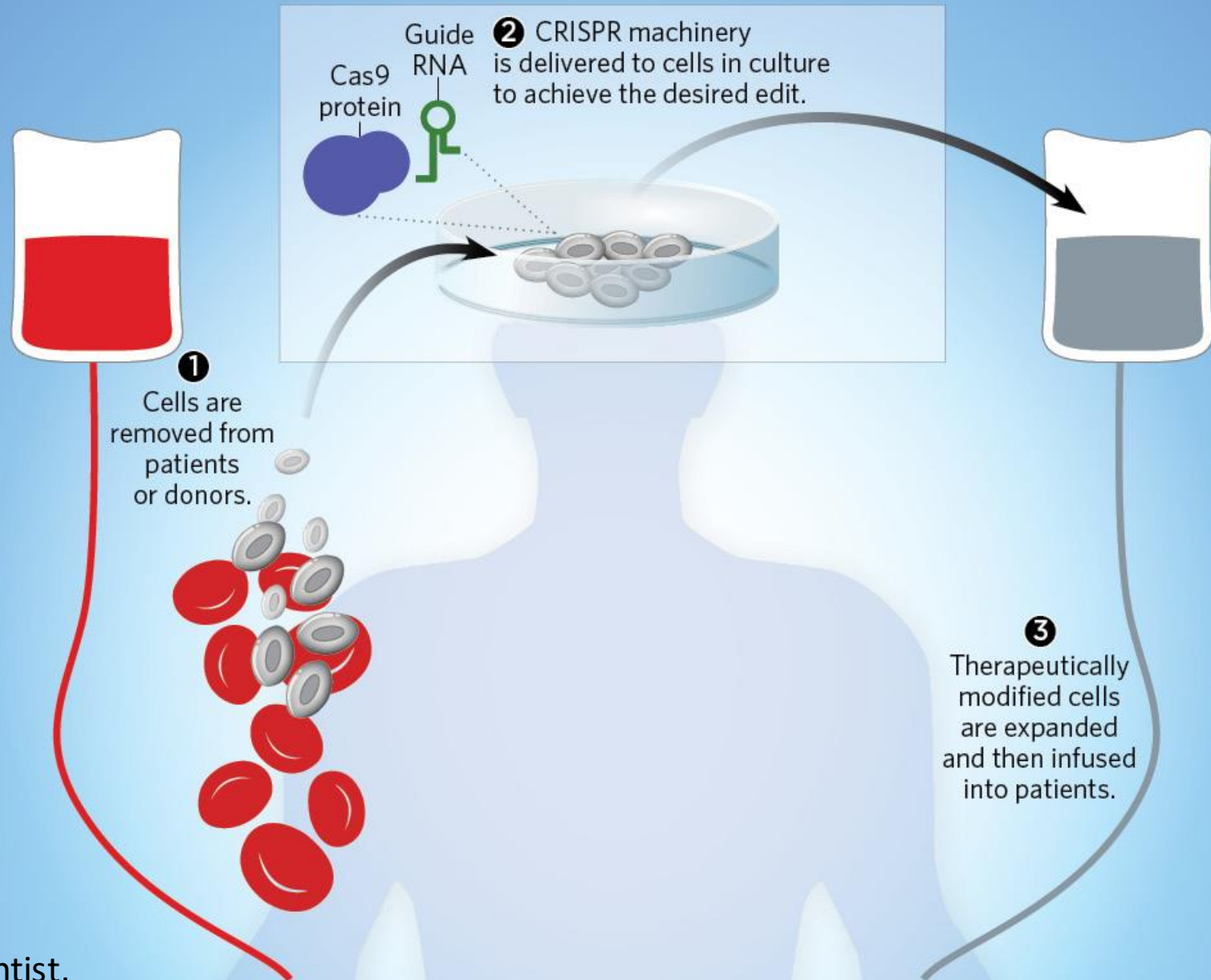
Chinese Translation 中文翻译

36 References 433 Citing Articles Letters

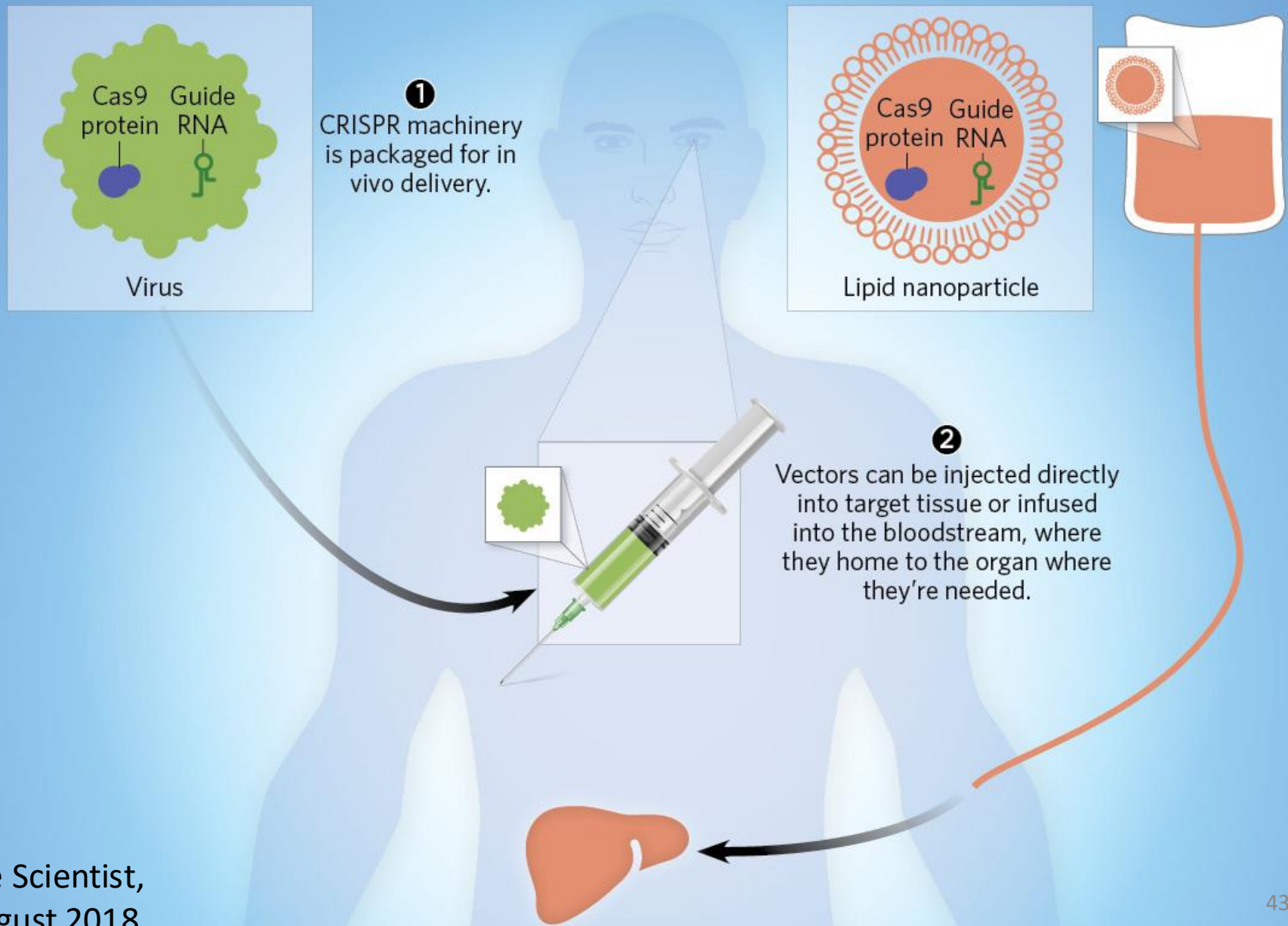
The UK's regulator has approved the world's first CRISPR–Cas9 gene editing therapy, which aims to cure sickle cell disease and transfusion-dependent β -thalassemia. Casgevy (exagamglogene autotemcel) is a first-of-its-kind treatment made by Vertex Pharmaceuticals and CRISPR Therapeutics in Zug, Switzerland. It comes just 11 years after Jennifer Doudna and Emmanuelle Charpentier [invented the technology](#). The green light from the Medicines and Healthcare Products Agency represents a major scientific achievement for Vertex and CRISPR and a landmark for the biotech industry.

Approved: Jan 2024!

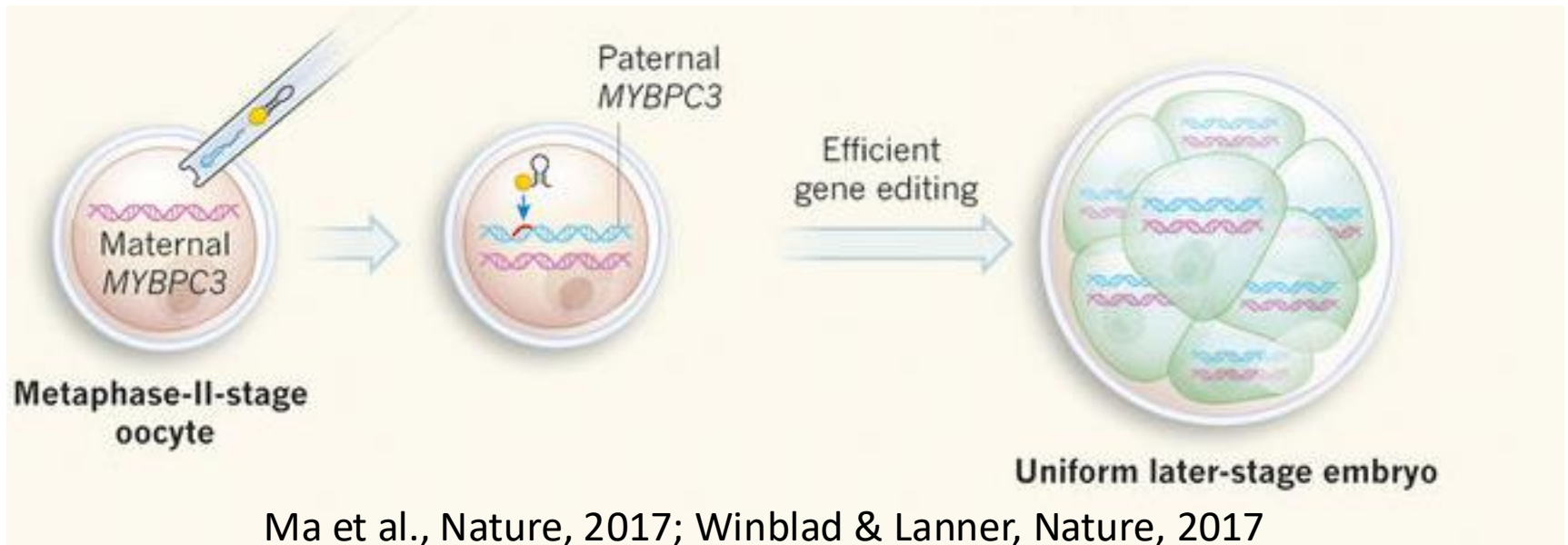
Ex vivo CRISPR therapy strategy



In vivo CRISPR therapy strategy



Correction of a pathogenic gene mutation in human embryos

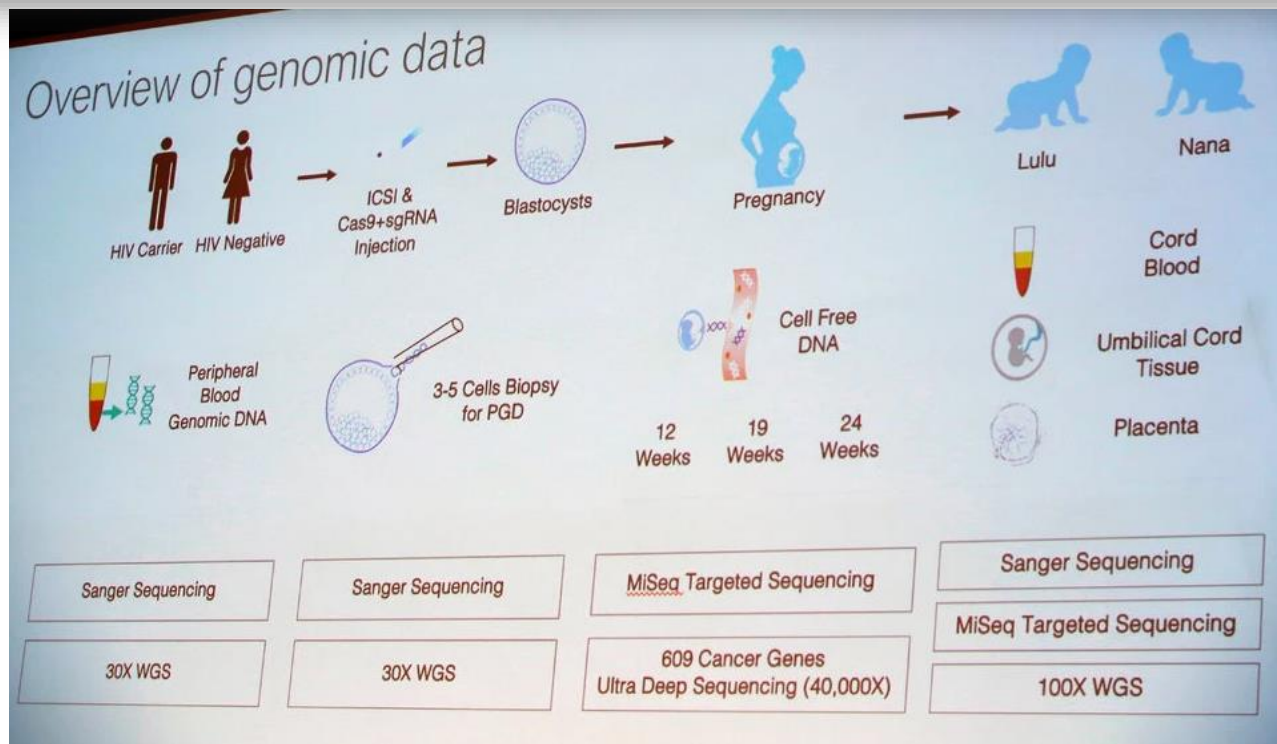


Aim: correct a mutation in the *MYBPC3* gene (which is associated with heart disease) in human embryos:

- Gene-editing components and sperm were injected into oocytes that contained non-mutated versions of *MYBPC3*
- Half the sperm used had a *MYBPC3* mutation
- 42 of 58 embryos tested (72.4%) did not have the *MYBPC3* mutation, indicating that at least “half” (around 25%) of the embryos were rescued

Why Are Scientists So Upset About the First Crispr Babies?

Only because a rogue researcher defied myriad scientific and ethical norms and guidelines. We break it down.



Yuval Noah Harari

New York Times Bestselling
Author of *Sapiens*



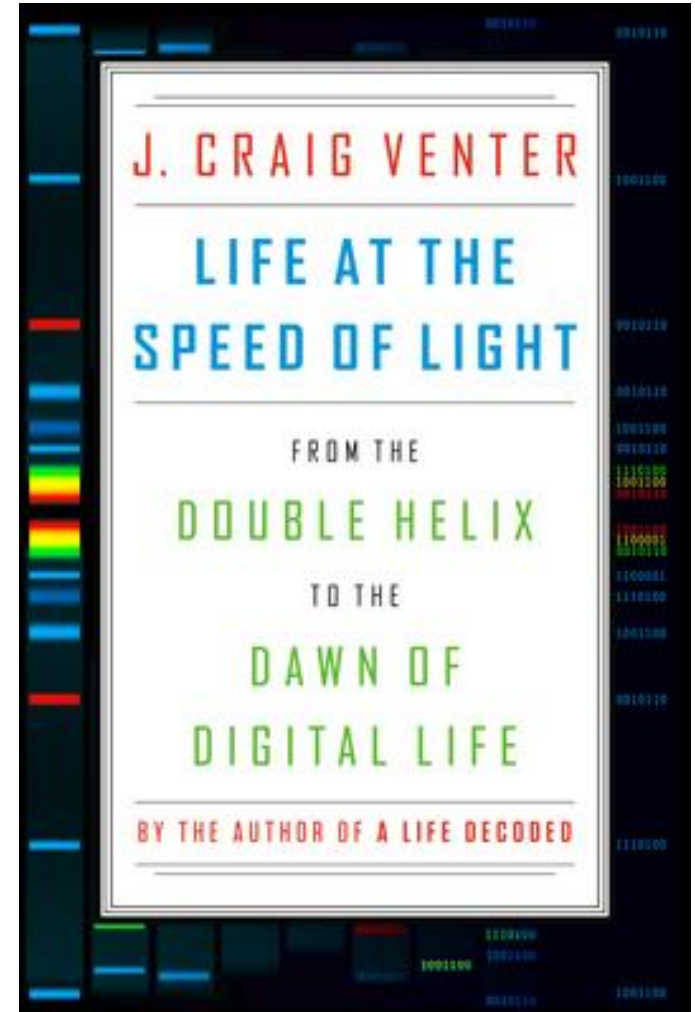
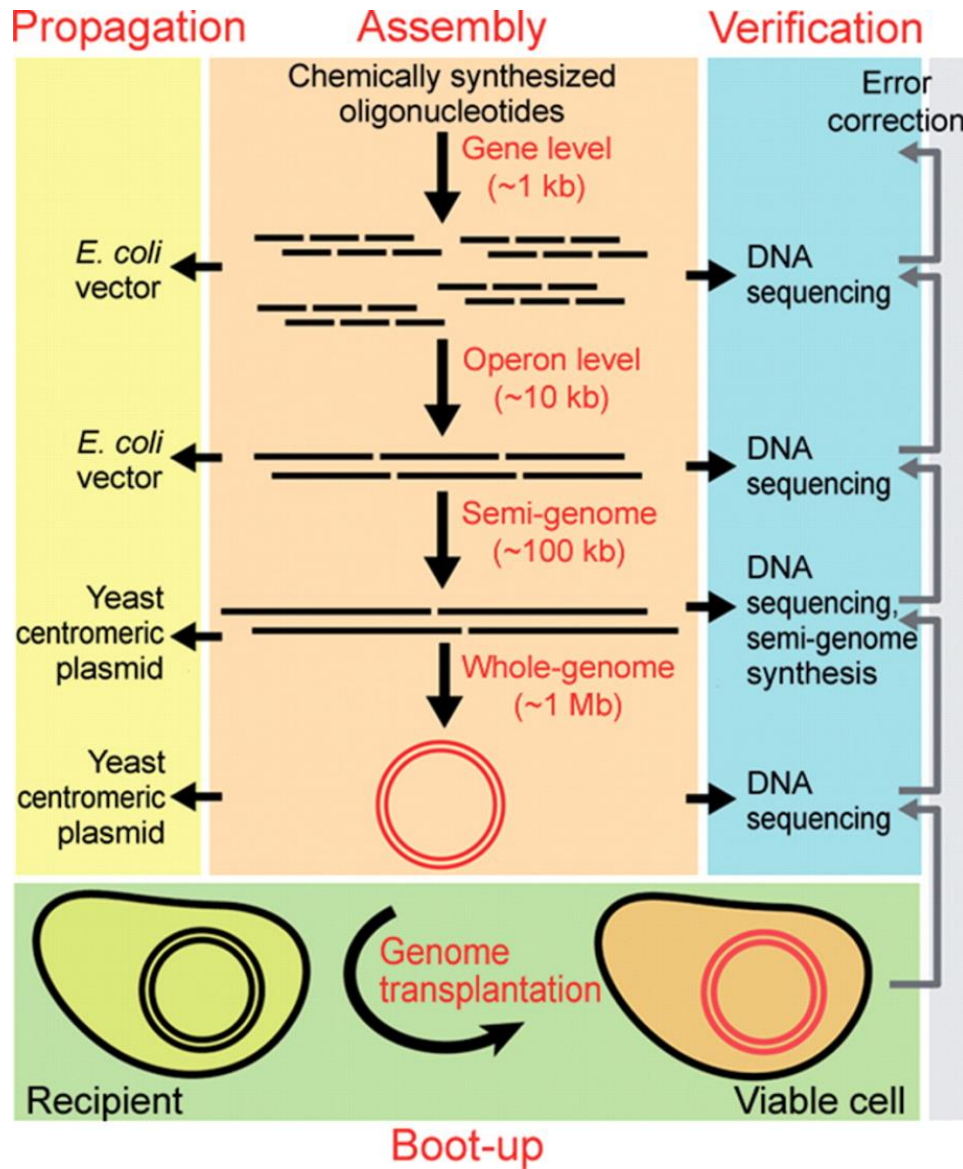
Homo Deus

A Brief History
of Tomorrow

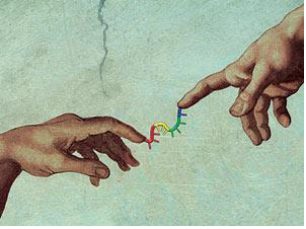
The Synthetic Genome

Goal: Build a bug

17 Oct 2013



Gibson et al., Science, 2010⁴⁷

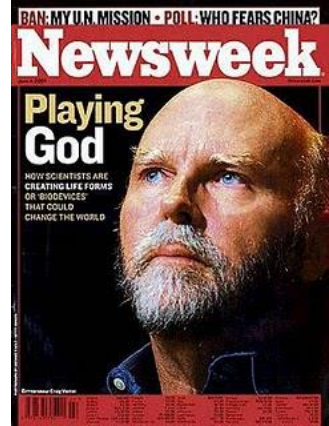


The Synthetic Genome

Synthetic biology and ethical implications

Synthetic “Life,” Ethics, National Security, and Public Discourse

Cho and Relman, Science, 2010



“Synthetic genomics and synthetic biology may necessitate a new model for addressing ethical and policy issues because of the complexity of the biological systems being mimicked and manipulated. The complex interactions of biological parts and their evolution will likely lead to unpredictable, emergent behavior in engineered organisms and ecosystems.”

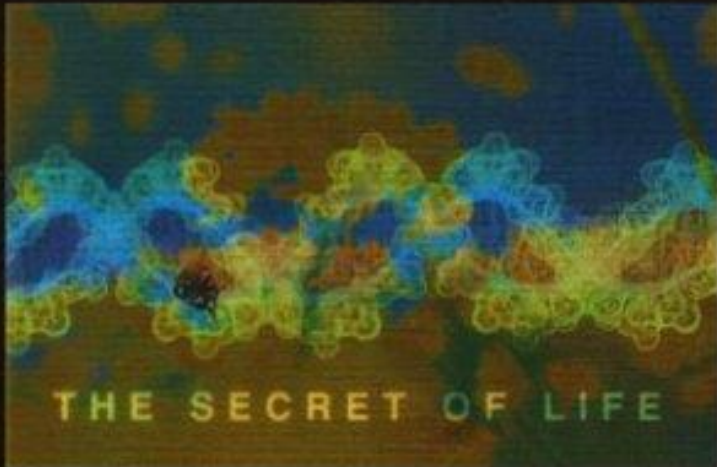
“The greatest challenge in addressing biosecurity and ethical concerns has been, and will be, to design effective oversight mechanisms that avoid undue harm to the overwhelmingly beneficial life sciences enterprise. “

“A realistic assessment of likely benefits is important because it highlights potential issues of distributive justice and fairness, especially with growing skepticism about the practical application of genomics to date, and the tendency toward hype.”

“Further discourse in this area should be informed by perspectives from theology, philosophy, the social sciences, and the general public.”

Other Christmas reading

DNA



THE SECRET OF LIFE

JAMES D. WATSON

WINNER OF THE NOBEL PRIZE AND AUTHOR OF *THE DOUBLE HELIX*

WITH ANDREW BERRY

Genes, Experience, &
What Make Us Human

Nature VIA Nurture

Author of the National
Bestseller *GENOME*

Matt Ridley

Follow-up courses

BIOENG-420 Single Cell Genomics (MA2)

BIO-411: Life Sciences Engineering: genome to function (MA2)

BIO-463 Genomics and Bioinformatics (MA1 or 3)

BIO-468 Scientific literature analysis in Computational molecular biology (MA1)

BIOENG-455 Computational cell biology (MA1)

CS-433 Pattern classification and machine learning (MA1 or 3)

CS-502 Deep learning in Biomedicine (MA1)

BIO-491 New tools and research strategies in personalized health (MA2)

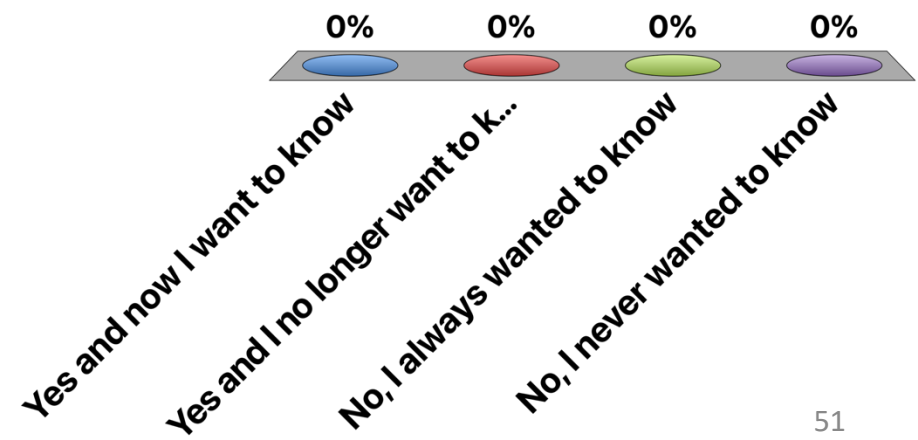
ChE-411 Principles and Applications of Systems Biology (MA1 or 3)

MATH-493 Applied biostatistics (MA2) (further expansion of R)

BIO-455 Introduction au droit et à l'éthique en STV (MA1 or 3)

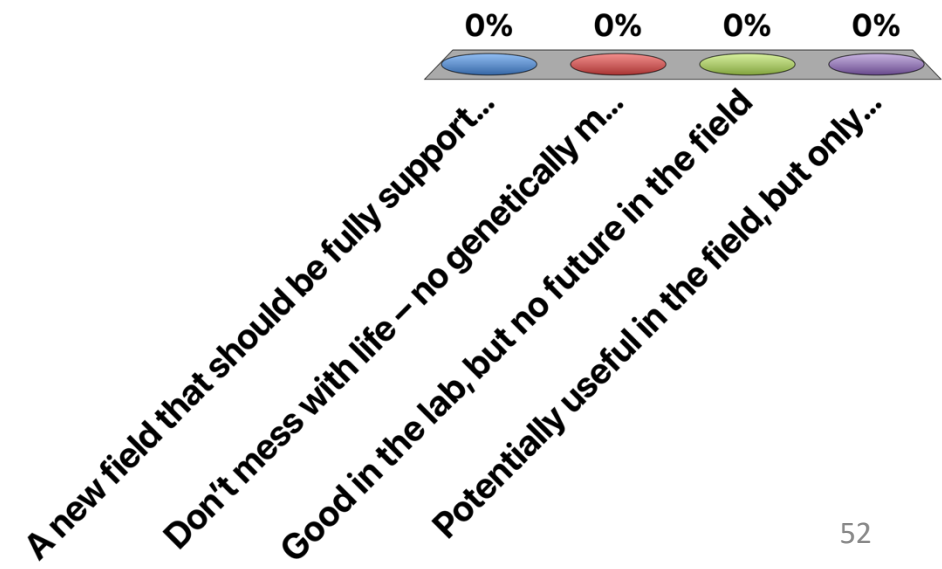
Did this class change your opinion about sequencing your genome?

- A. Yes and now I want to know
- B. Yes and I no longer want to know
- C. No, I always wanted to know
- D. No, I never wanted to know



What about synthetic genomics?

- A. A new field that should be fully supported to move forward
- B. Don't mess with life – no genetically modified organisms for me
- C. Good in the lab, but no future in the field
- D. Potentially useful in the field, but only when «proven» harmless



What do you think are opportunities?
(sequencing your genome)

What are your principal concerns?
(sequencing your genome)

What do you think are opportunities?
(synthetic genomics)

What are your principal concerns?
(synthetic genomics)