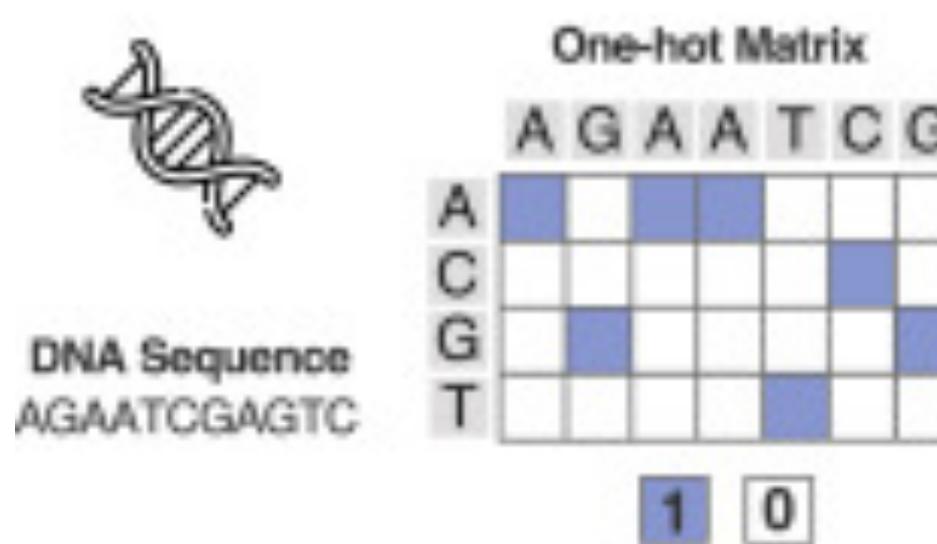


Bioenformatik

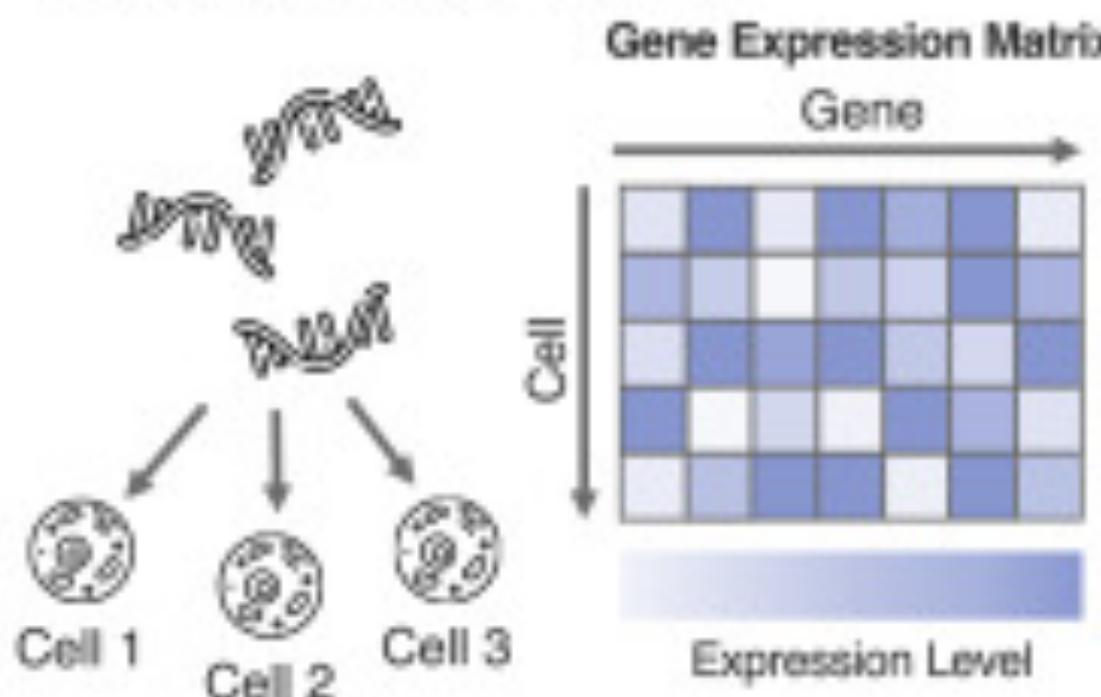
Tugce Bilgin Sonay, Temmuz 2023, Boğaziçi Üniversitesi

Bioenformatik

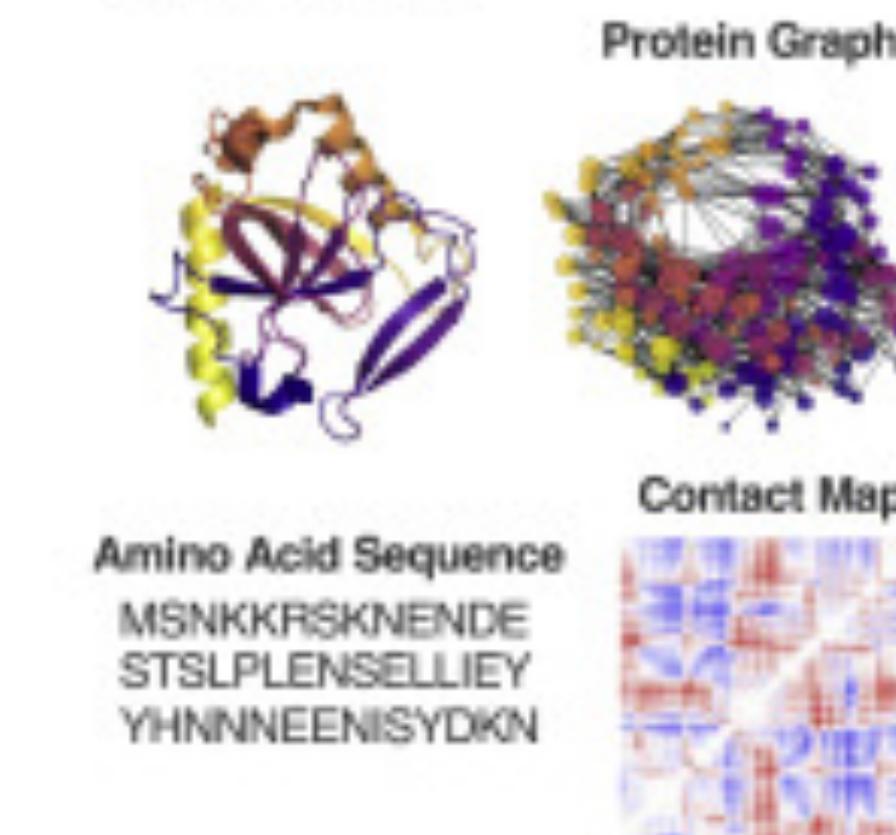
A DNA



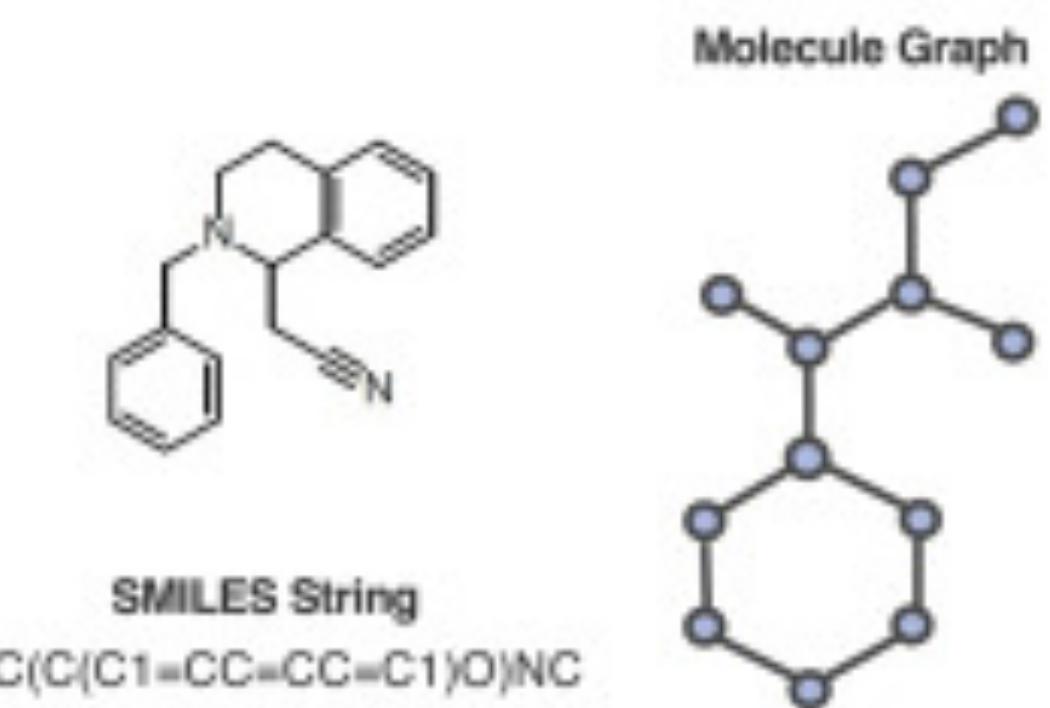
B Gene Expression



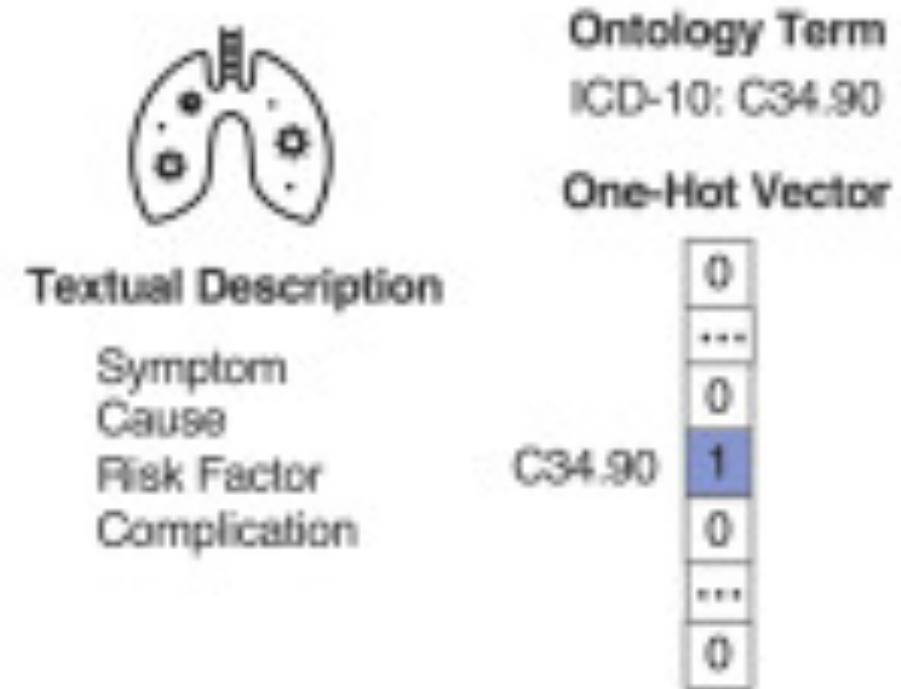
C Protein



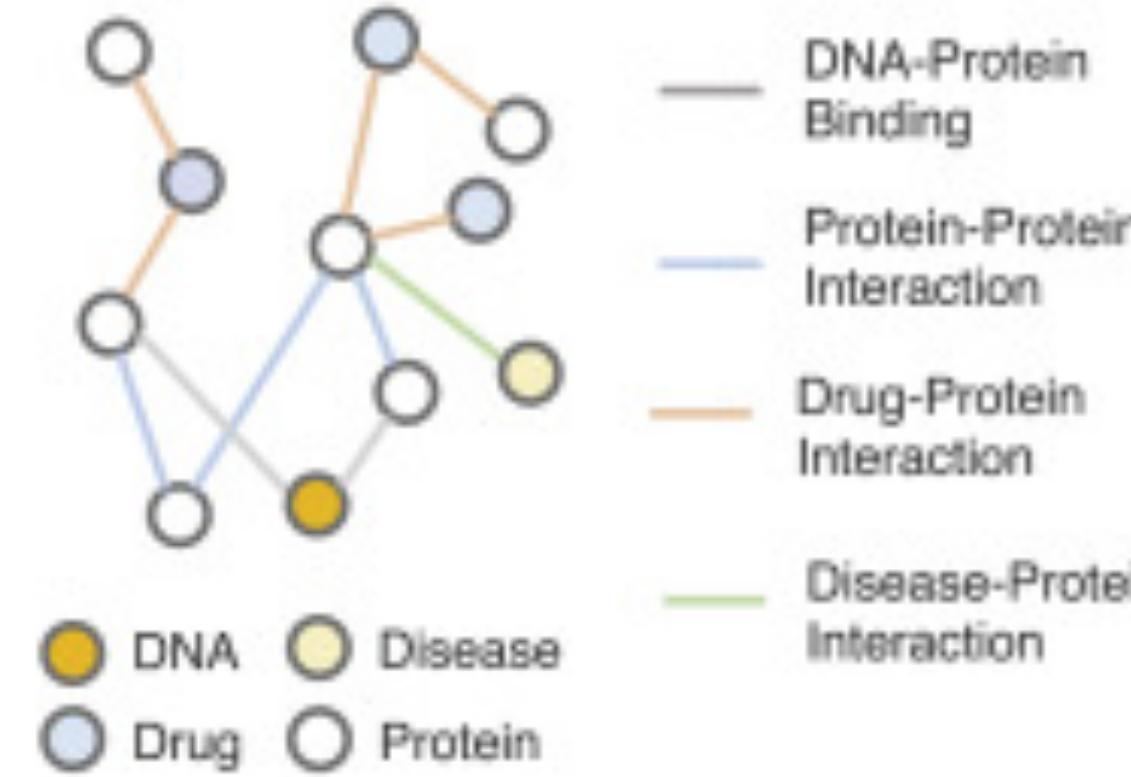
D Compound



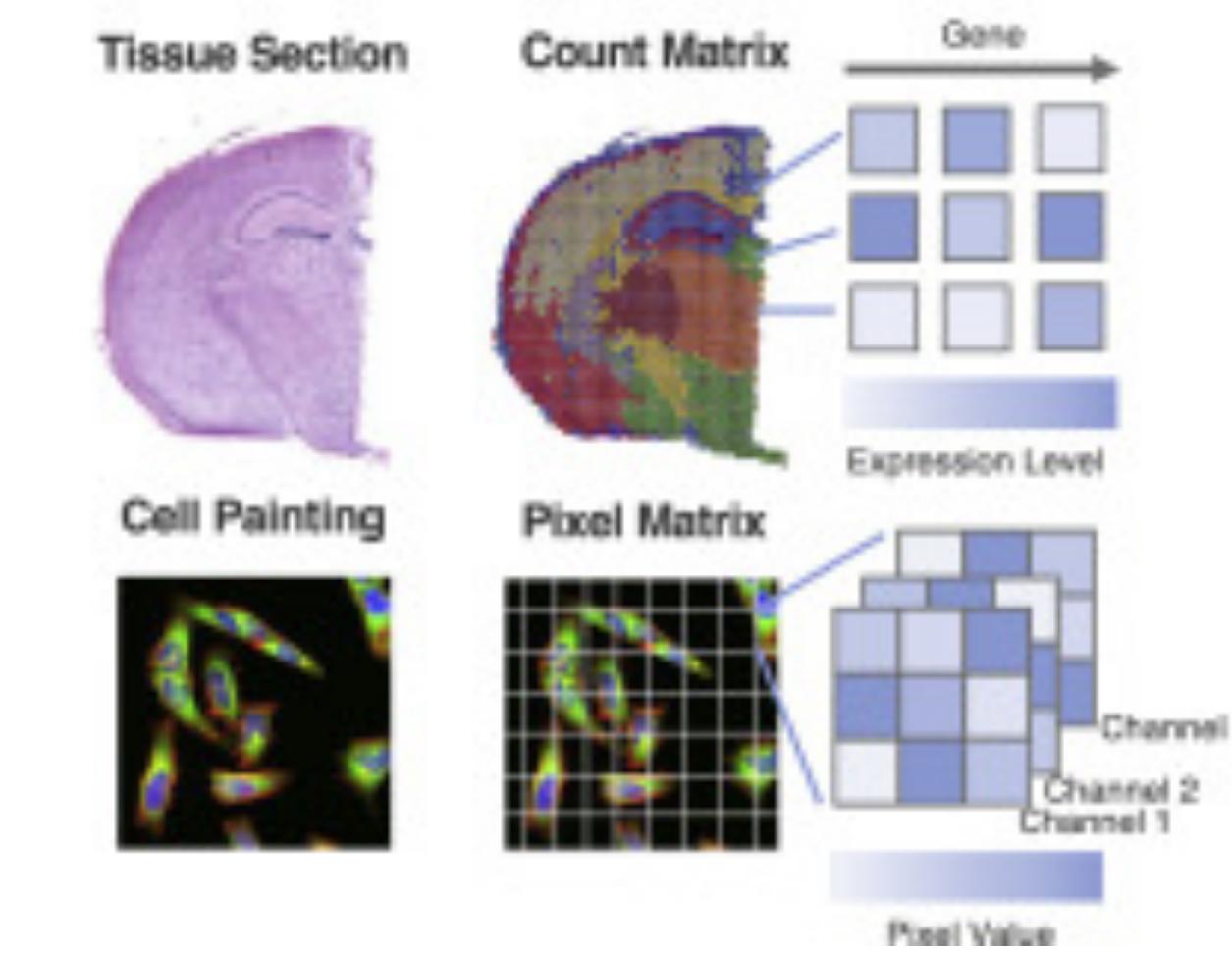
E Disease



F Biomedical Networks



G Spatial Data



H Text



Dinleri, devletleri, toprakları, hatta yaşları bile yok

Bazılarının hayatı suya yazılmış. Mesela Pasifik Okyanusu'nun batı kıyısında, suyun dibinde yaşayan [Bajau Çingeneleri](#)... Onları bulmaya karar verdığimde yol arkadaşım olması için formanımın yanına gittim. "Gelir misin?" dedim, "Fark etmez" dedi. Bu arada meseleye ilişkin tek teknik bilgi, Hüseyin'e Filipinli deniz korsanları tarafından kaçırılma riskinden bahsettirmi ve o da doğru olmaz." İşte 'suyun kaldırma kuvveti' ekseninde biçimlenen gayet ıslak maceramızda...

Hasan Cem ARAP TARLI





**11.
AYKUT KENCE
EVRİM
KONFERANSI**
18-19ŞUBAT 2017
www.aykutkence.com



Dr. Tuğçe Bilgin Sonay

Sunum Konusu: DNA'nın en oynak elemanı: ardışık kısa tekrar dizileri ve fenotipik evrim

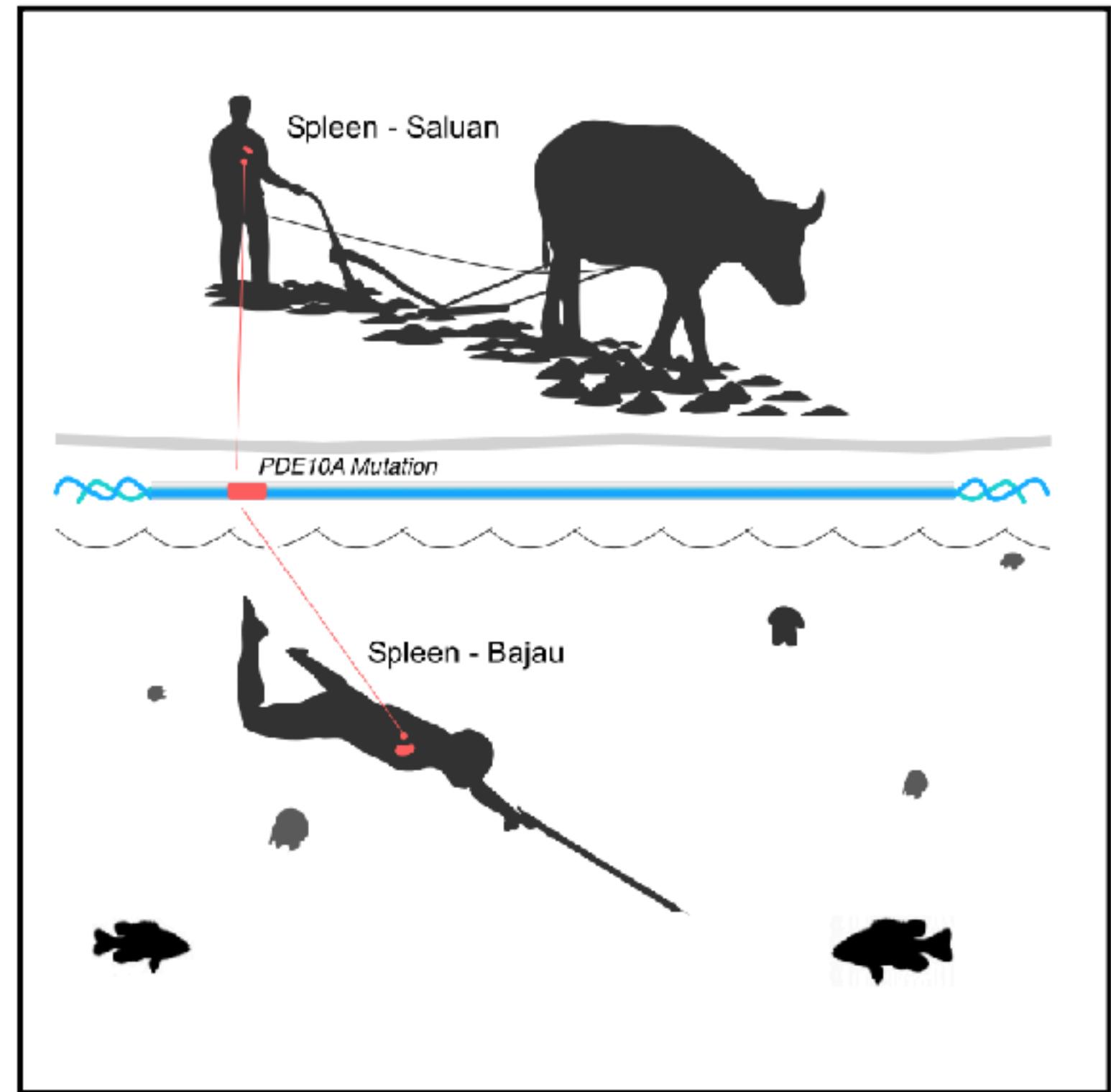
Yeditepe Üniversitesi Moleküler Biyoloji ve Genetik bölümünden mezun olan Dr. Tuğçe Bilgin Sonay, Zürich Üniversitesi'nde Evrimsel Biyoloji üzerine doktorasını bitirdi, yine aynı yerde araştırma görevlisi olarak çalışıyor. Evrim dısında psikoloji ve ikisiin kesişim noktaları olan bilimci ve empatinin ortaya çıkış gibi konulara ilgi duyuyor.

ODTÜ KKM - Kemal Kurdaş Konferans Salónu



Physiological and Genetic Adaptations to Diving in Sea Nomads

Graphical Abstract



Authors

Melissa A. Ilardo, Ida Moltke,
Thorfinn S. Korneliussen, ...,
Suhartini Salingkat, Rasmus Nielsen,
Eske Willerslev

Correspondence

rasmus_nielsen@berkeley.edu (R.N.),
ewillerslev@snm.ku.dk (E.W.)

In Brief

Genetic and physiological adaptations enable the remarkable breath-holding ability of marine nomads.



Highlights

- The Bajau, or “Sea Nomads,” have engaged in breath-hold diving for thousands of years
- Selection has increased Bajau spleen size, providing an oxygen reservoir for diving

Fenotipik Elastikiyet



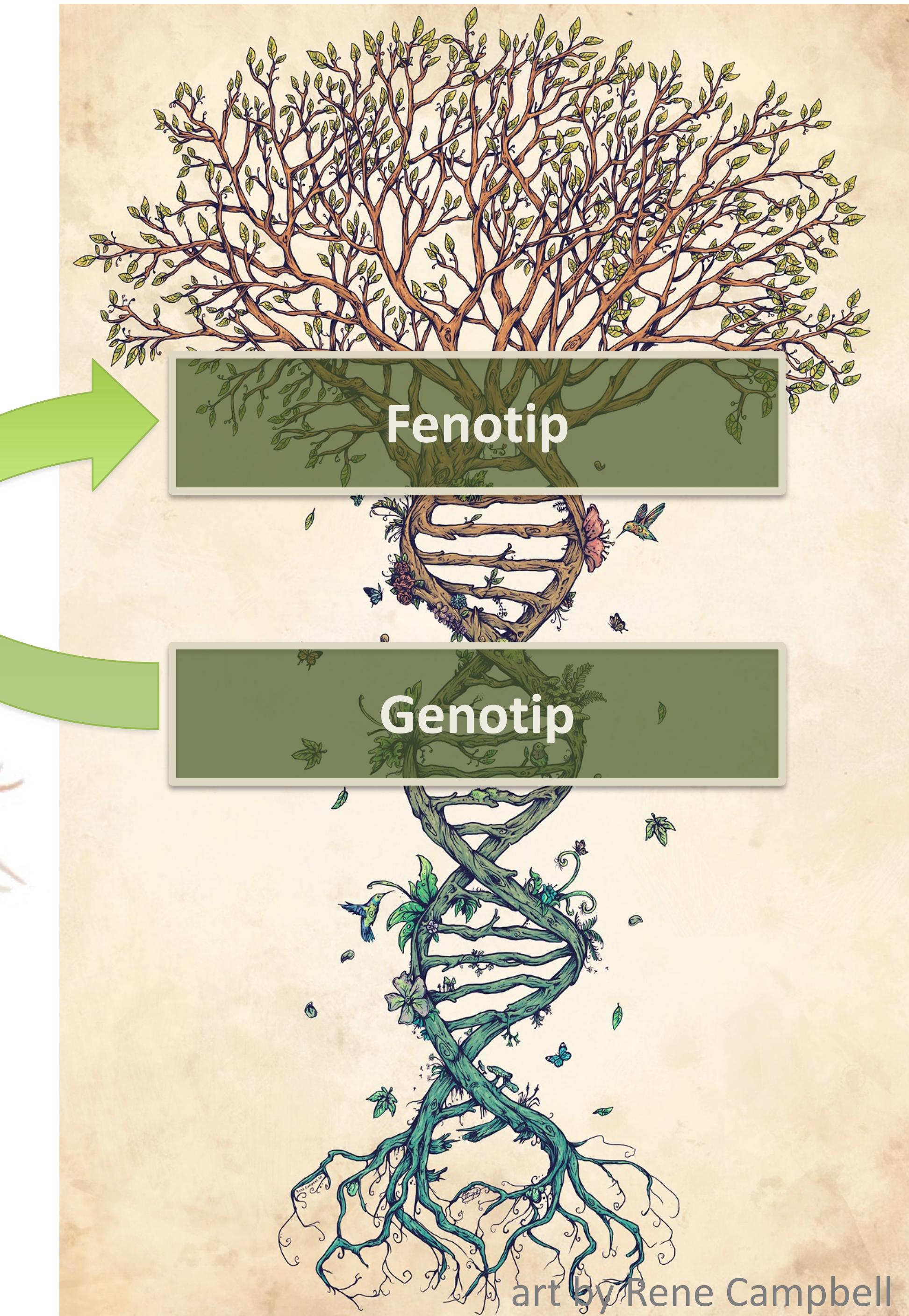
17° C



25° C

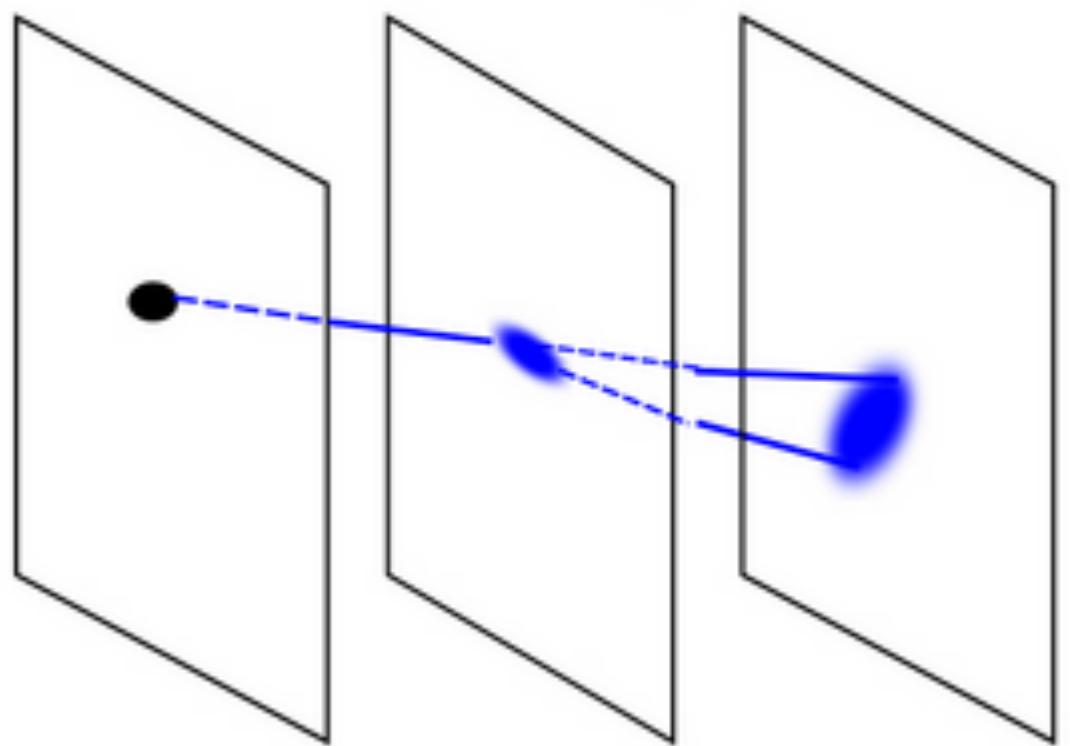


28° C

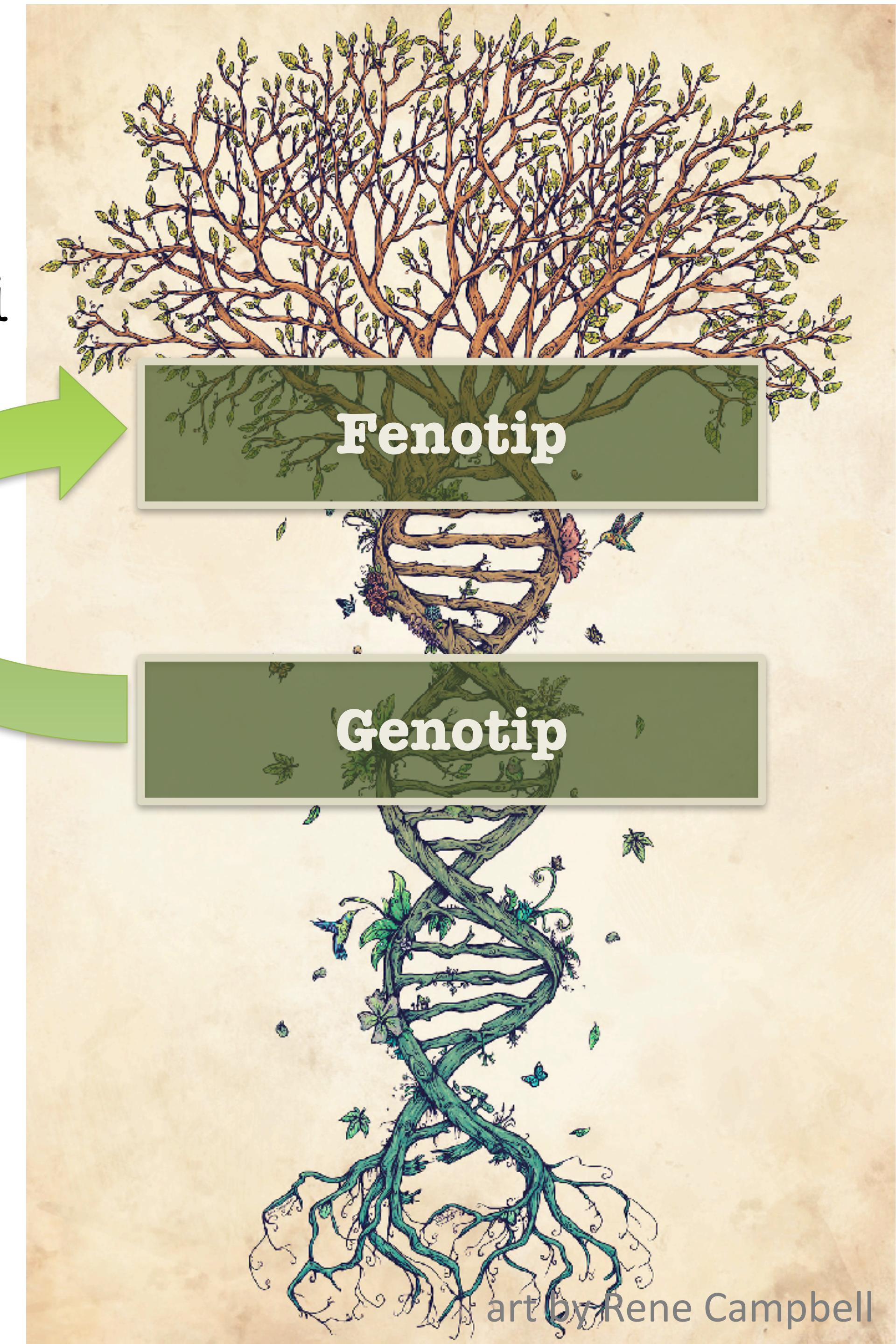


art by Rene Campbell

Genotip-Fenotip Haritalari



Alberch, 1991



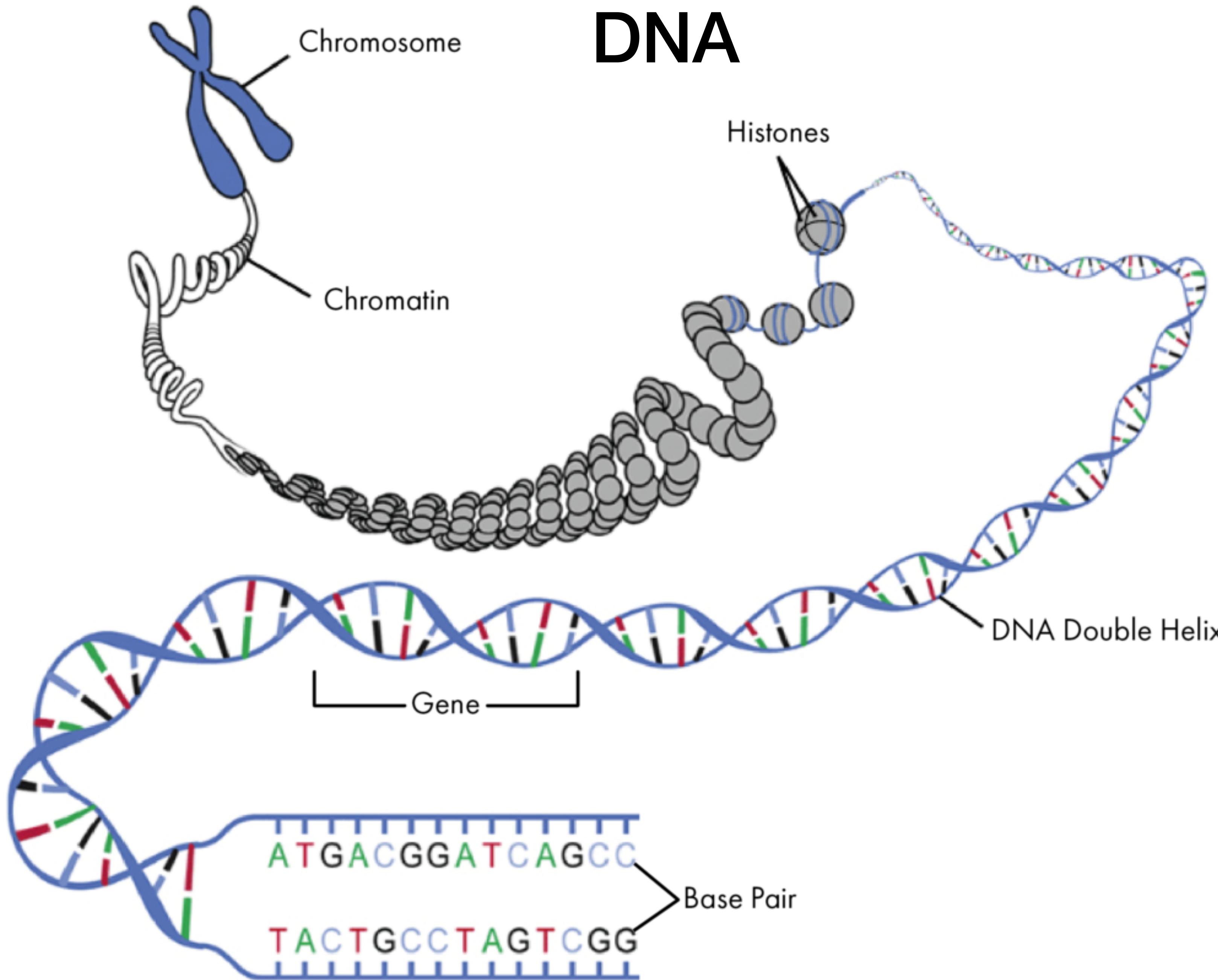
art by Rene Campbell



Yuzde 99 aynı

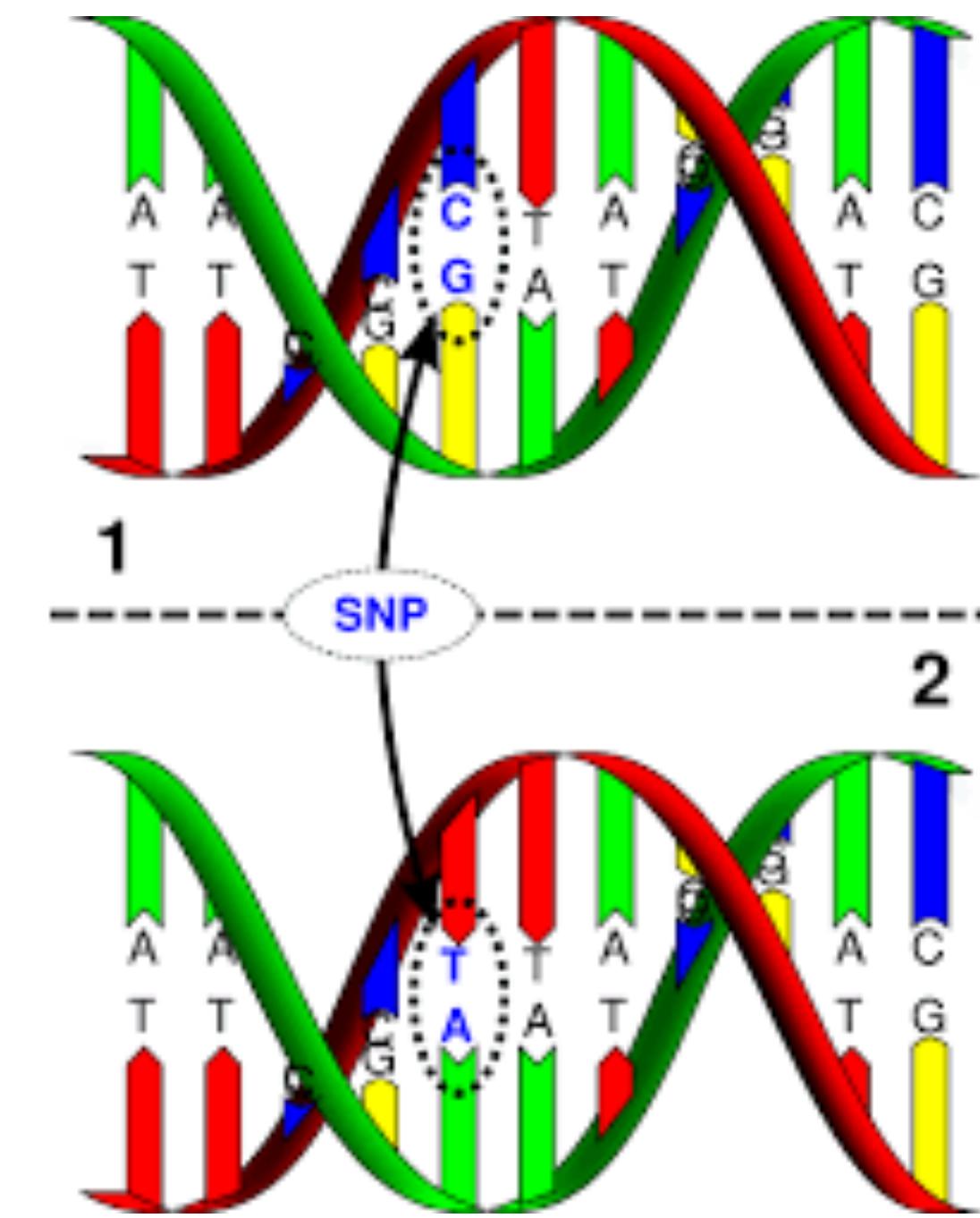


DNA



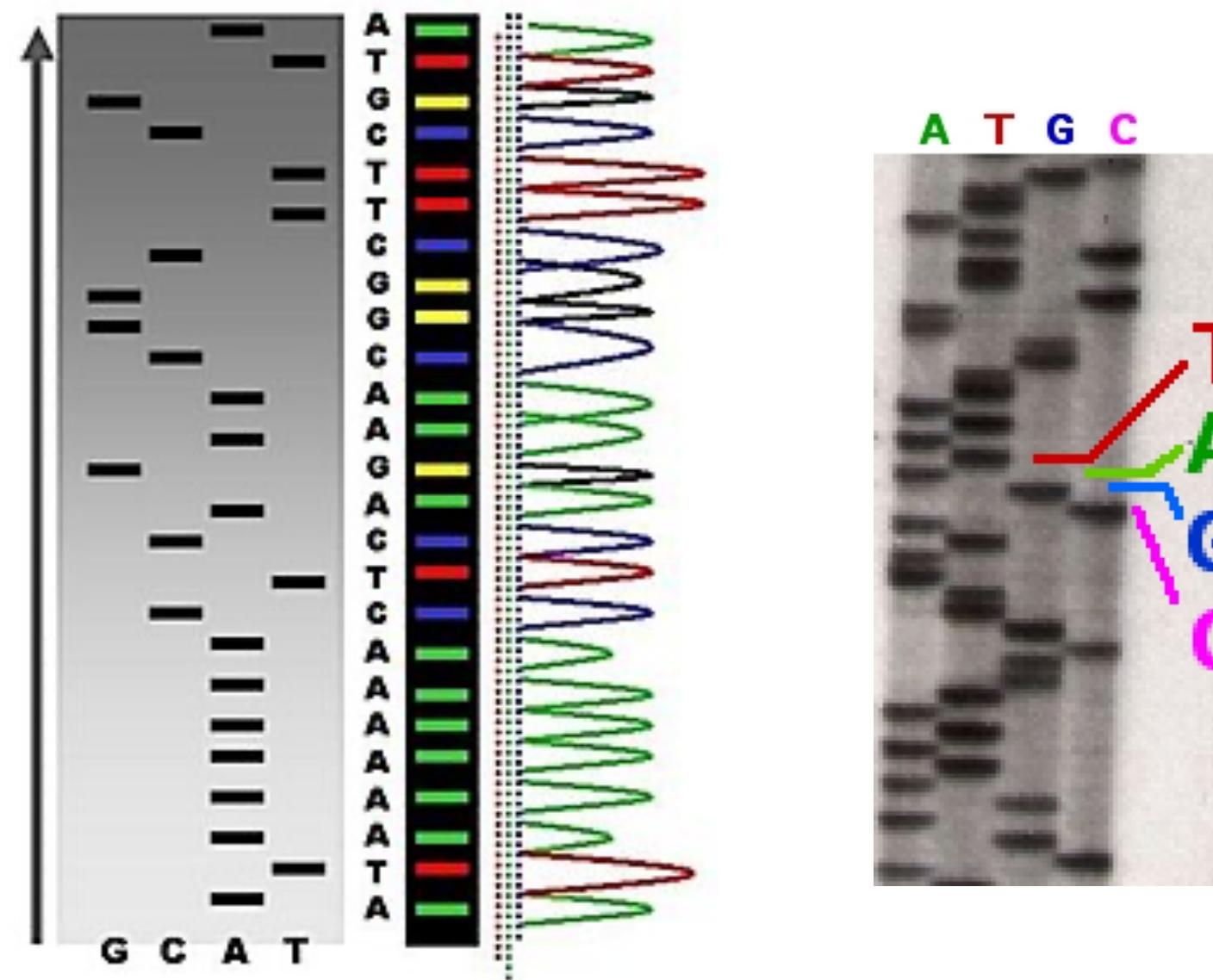
Mutasyonlar:

- **Tek Nukleotit Polimorfizmi - SNP**
10-100 milyon nukleotitte bir mutasyon



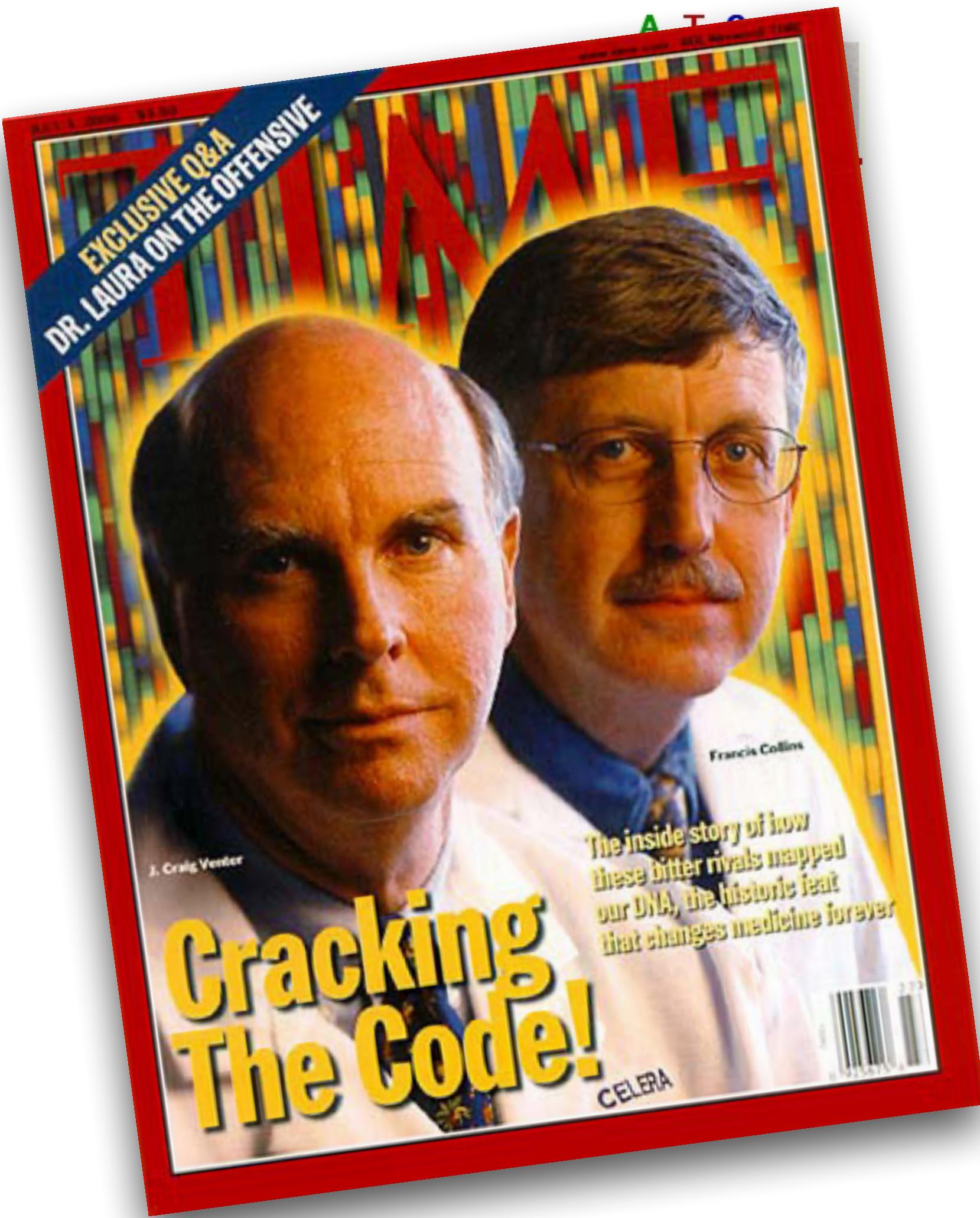
DNA Dizisini nasıl biliyoruz?

Sanger Dizileme



- 1977: İlk bakteri genomu
- 1986: Renk zorluğu genleri
- 1987: CRISPR dizisi
- 1994: İlk kanserojenler
- 1995: H.influenza genomu
- 1996: İlk ökaryotik genom
- 1998: İlk cokhuceli genomu

DNA Dizisini nasıl biliyoruz?



- 1977: İlk bakteri genomu
- 1986: Renk zorluğu genleri
- 1987: CRISPR dizisi
- 1994: İlk kanserojenler
- 1995: H.influenza genomu
- 1996: İlk ökaryotik genom
- 1998: İlk cokhucreli genomu

Insan genomu?

If you printed all your DNA
(6.4 billion letters), it would fill
4,200 books*

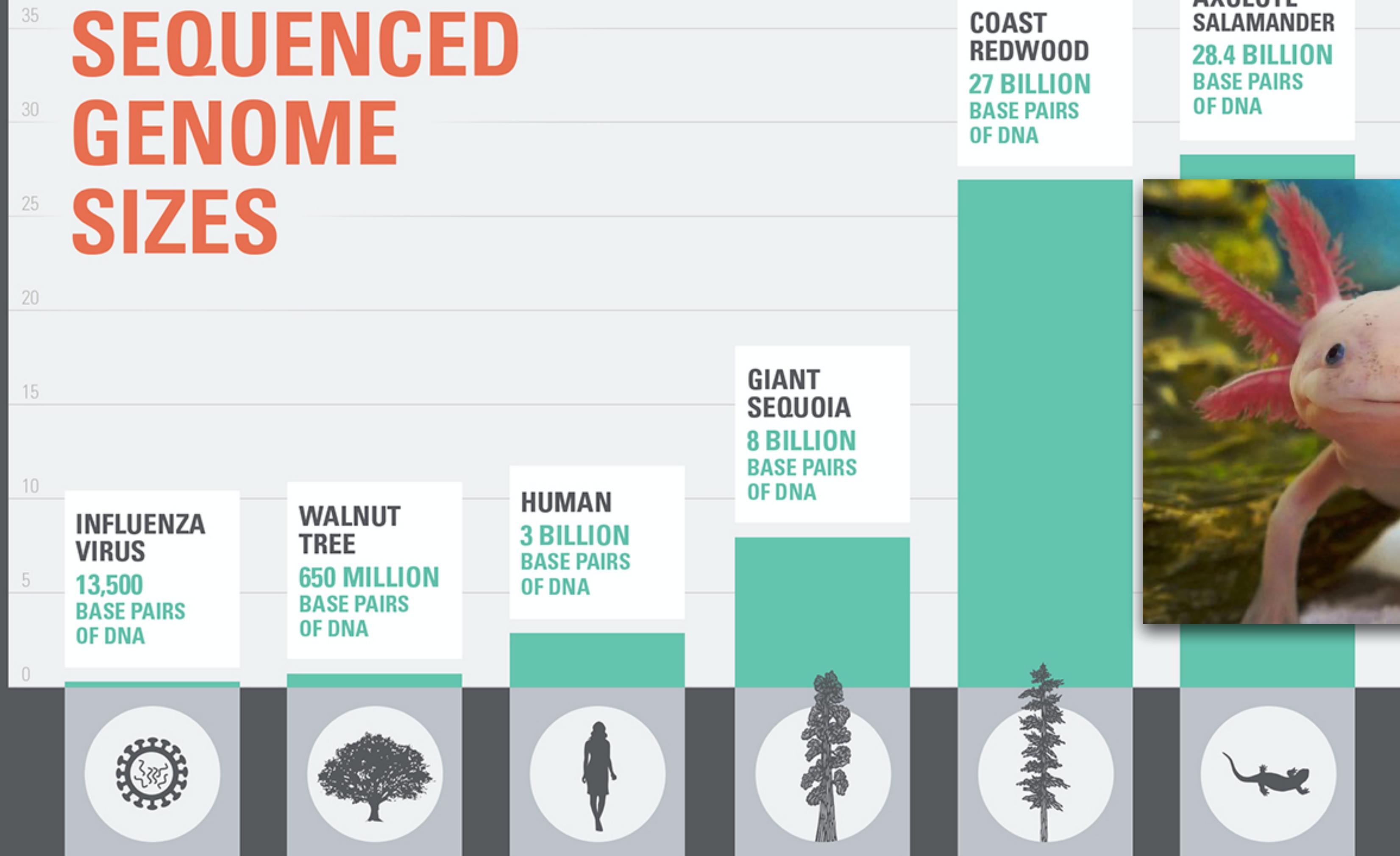


Veritas
sequences your
whole genome

(equivalent to
4,200 books)

* Assuming that a book, like Darwin's *Origin of Species*, has 500 pages.

SEQUENCED GENOME SIZES



The coast redwood and giant sequoia genomes were sequenced by Save the Redwoods League, University of California, Davis, and Johns Hopkins University. Learn more at SaveTheRedwoods.org/RedwoodGenome.

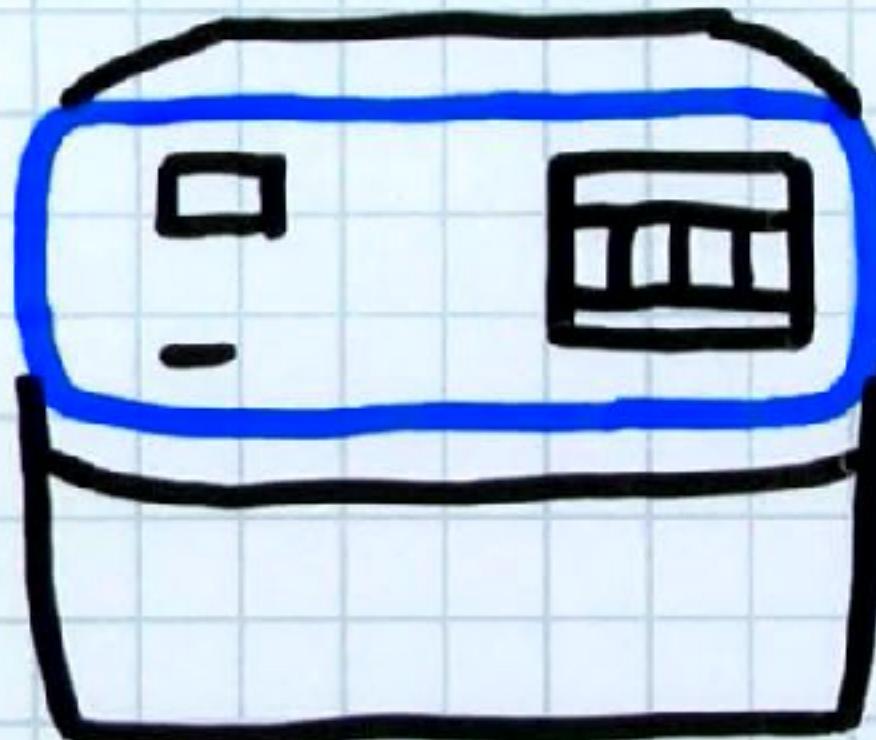


Mari Asano Schwankie

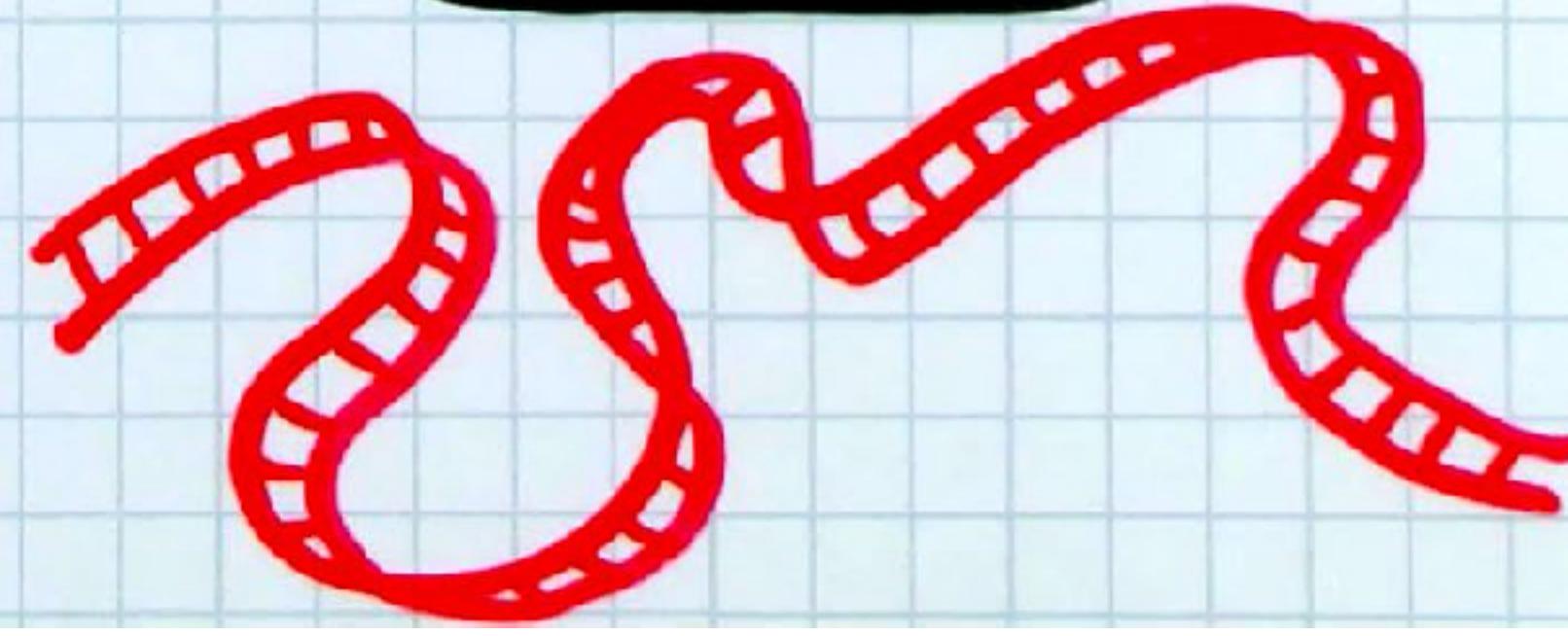
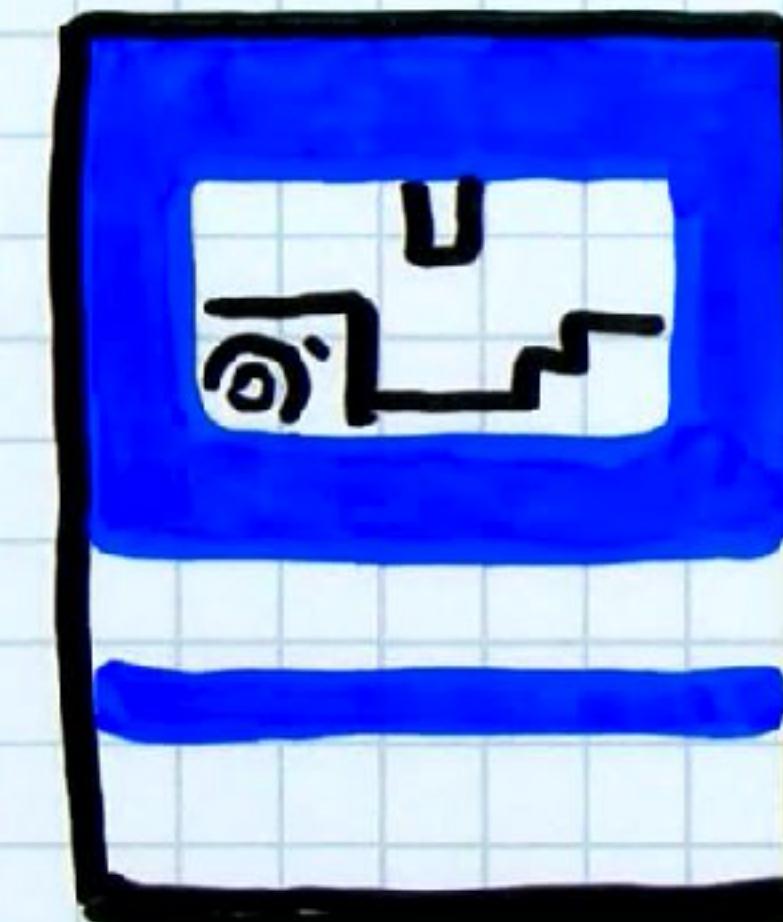
DNA Dizileme Teknolojileri

Yeni Nesil Dizileme

NGS
MASSIVELY
PARALLEL

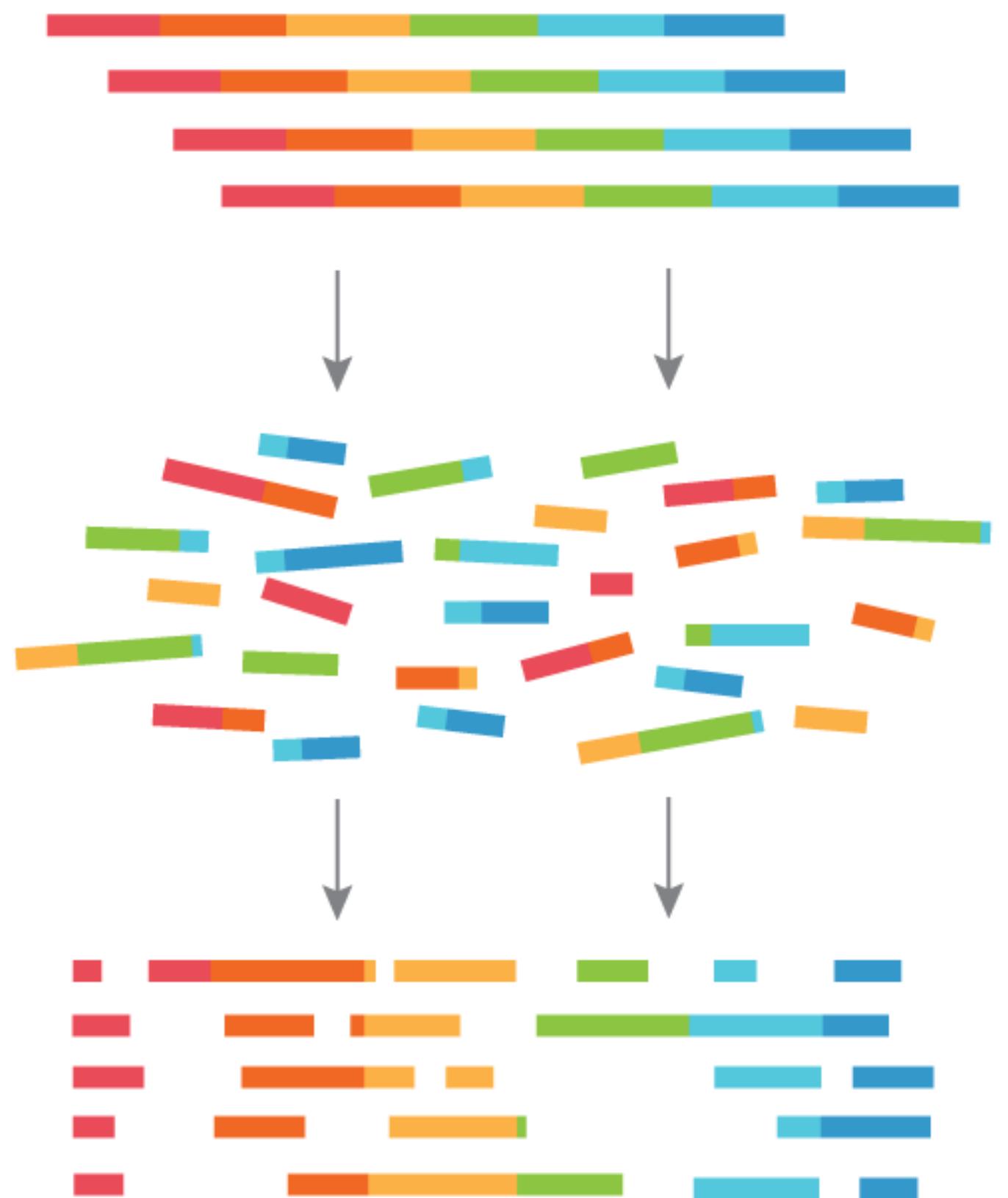


SANGER



DNA Dizileme Teknolojileri

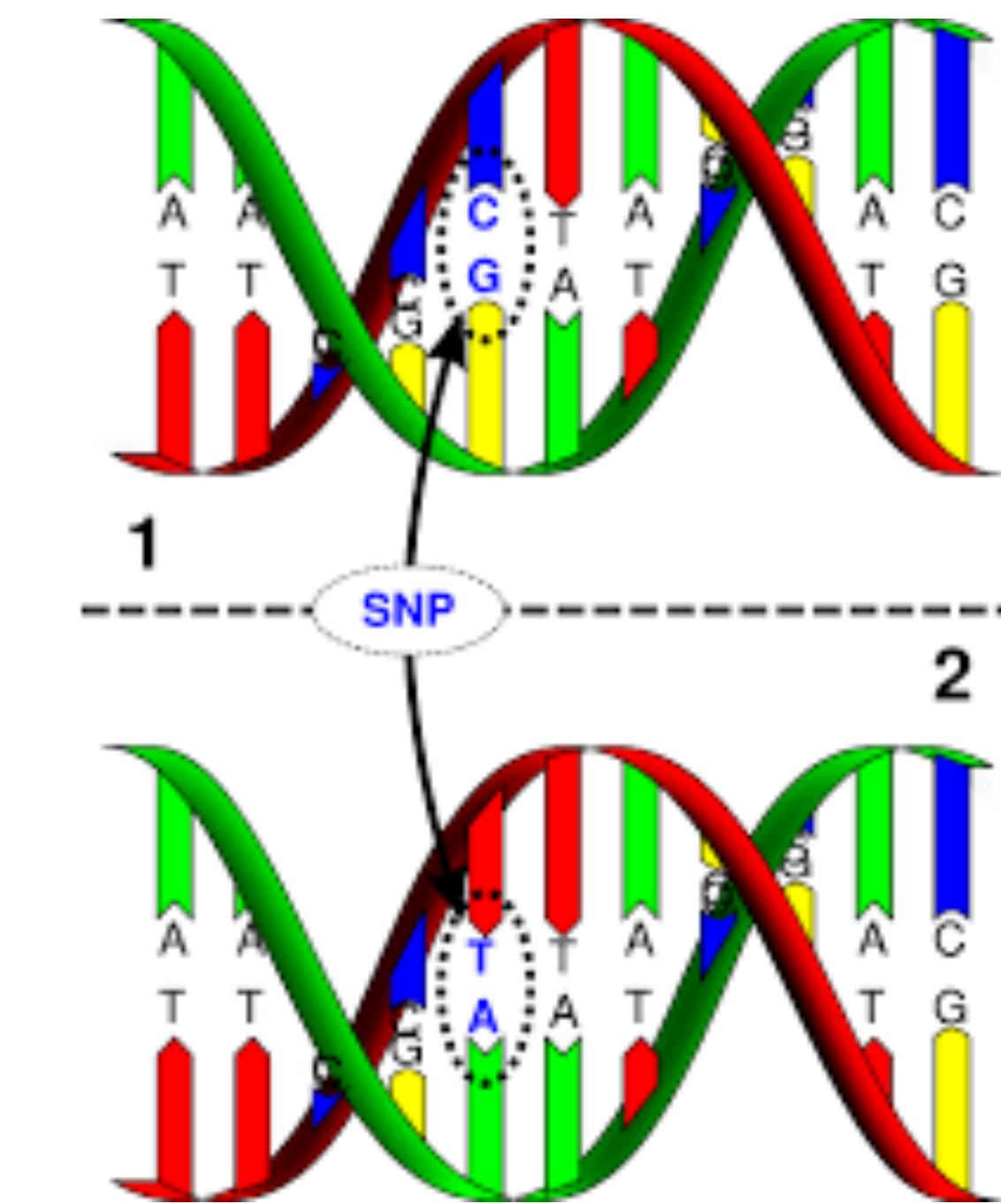
Yeni Nesil Dizileme



ATGTTCCGATTAGGAAACCTATCTGTAAGTGTTCATTCACTAAAAGGGAGGAAA

Mutasyonlar:

- **Tek Nukleotit Polimorfizmi - SNP**
10-100 milyon nukleotitte bir mutasyon
- **Yapisal Cesitlilik** - silinmeler ve eklenmeler (indel) insan genomunun yuzde besi



referans:
birey:

C A C A G T G C G C - T		
C A C C G T G - G C A T		
	substitusyon (TNP/ SNP)	sillinme (indel)
		eklenme

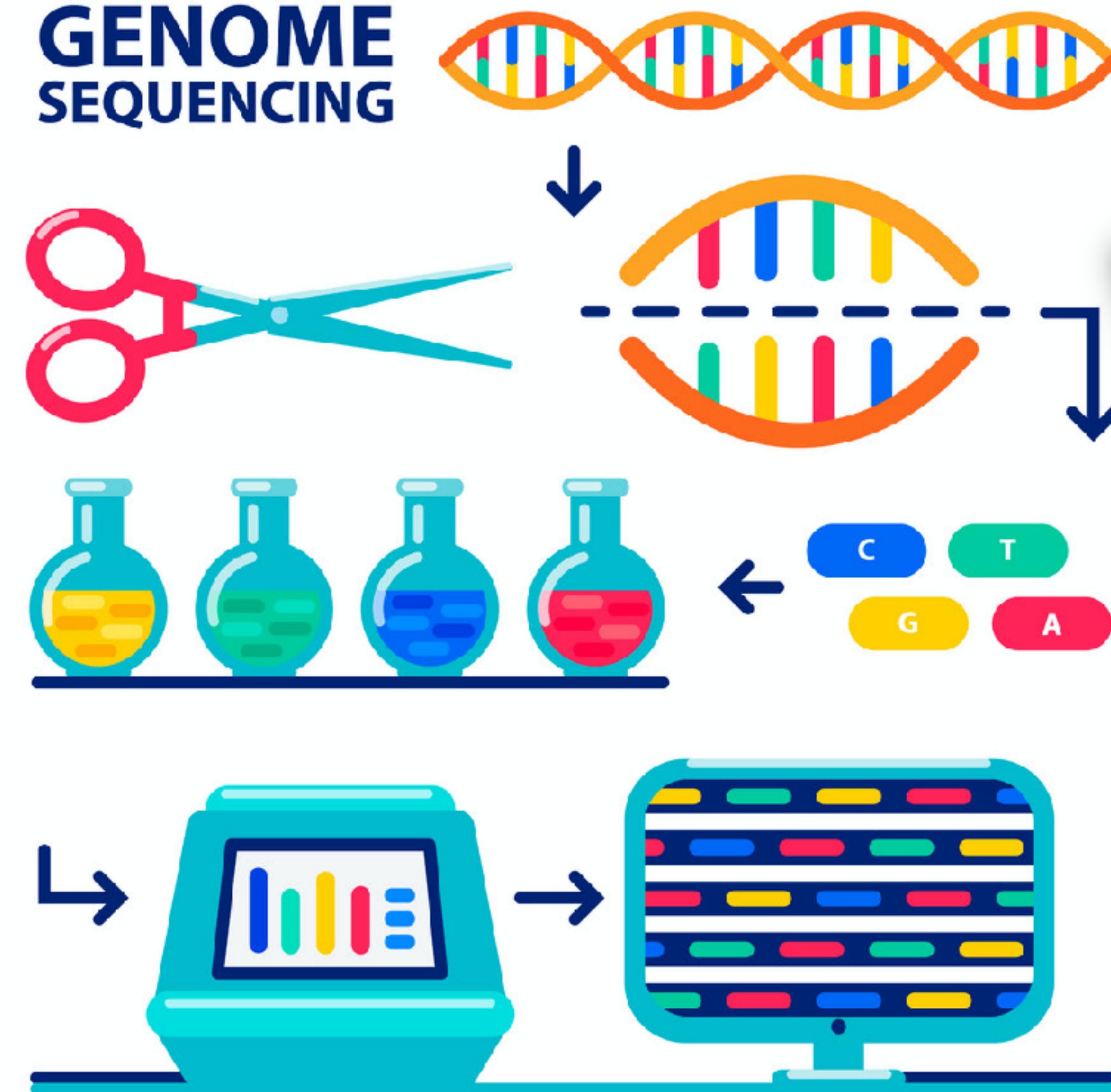


Human Genome Project

1990–2003

13 yıl
\$3 milyar
20 ülke

GENOME SEQUENCING

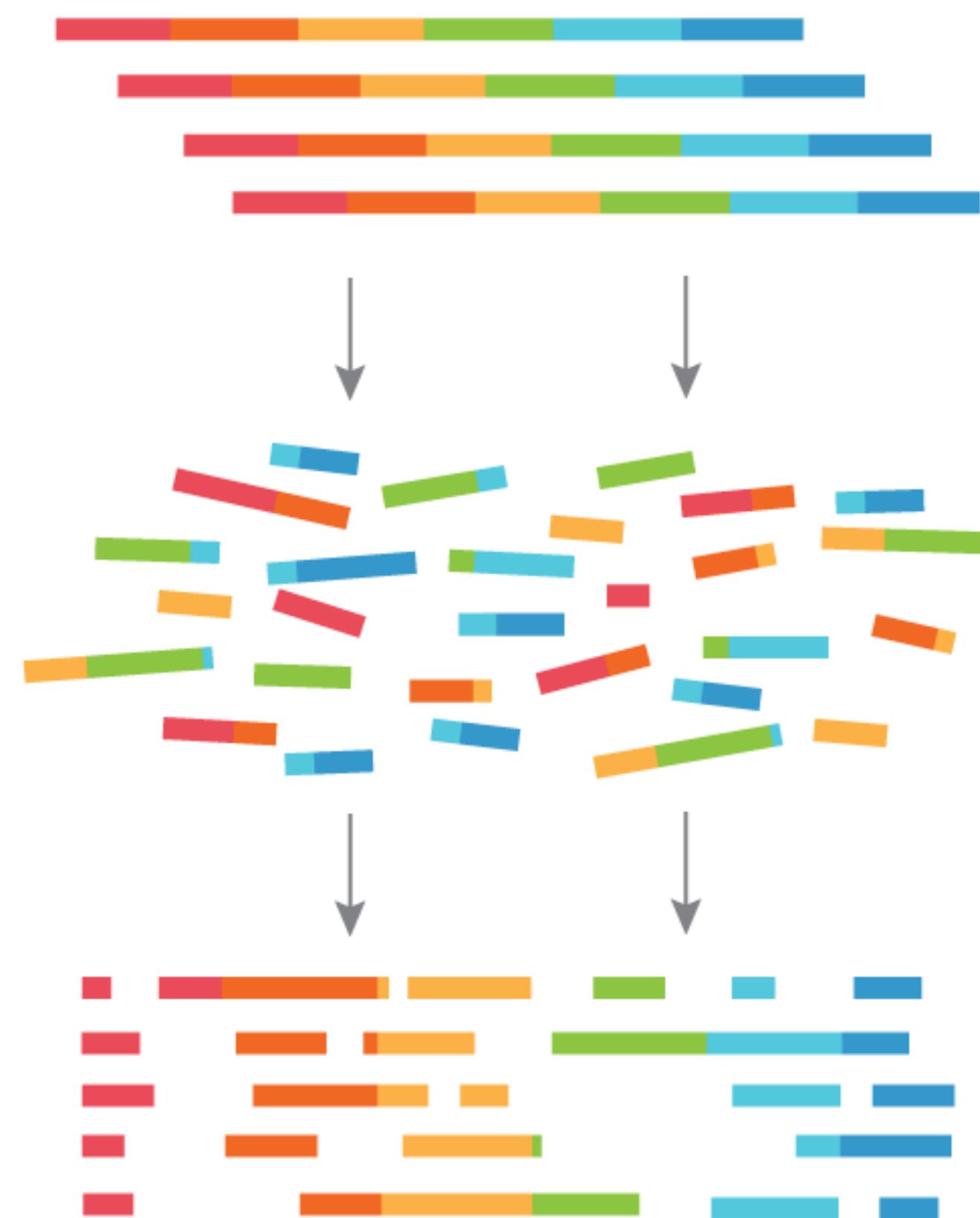


İnsan Genom Projesinden neler öğrendik?

- 20,000 genimiz varmış
- Kromozomlarımız farklı uzunluklardaymış
- Genomun çoğu protein kodlamamış
- Dizi = fonksiyon değilmiş
- Anlık veri paylaşımı
- Etiksel tartışmalar

DNA Dizileme Teknolojileri

Yeni Nesil Dizileme



ATGTTCCGATTAGGAAACCTATCTGTAAGTGTTCATTCACTAAAGGGAGGAAA

2003 Human Genome Project



2006 1st individual genome



2007 1st NGS Genome



2008 1st 30x genome



2010 1st sub-10K genome



2014 1st \$1,000 genome



DNA Dizileme Teknolojileri

Yeni Nesil Dizileme

\$3,000,000,000 | 2003 Human Genome Project



\$20,000,000 | 2006 1st individual genome



\$2,000,000 | 2007 1st NGS Genome



\$200,000 | 2008 1st 30x genome



\$10,000 | 2010 1st sub-10K genome



\$1,000 | 2014 1st \$1,000 genome





deniz
diyeti



sut
sindirim



malaryaya
dayanıklılık



koleraya
dayanıklılık



arsenikli
ortam



soguk
hava



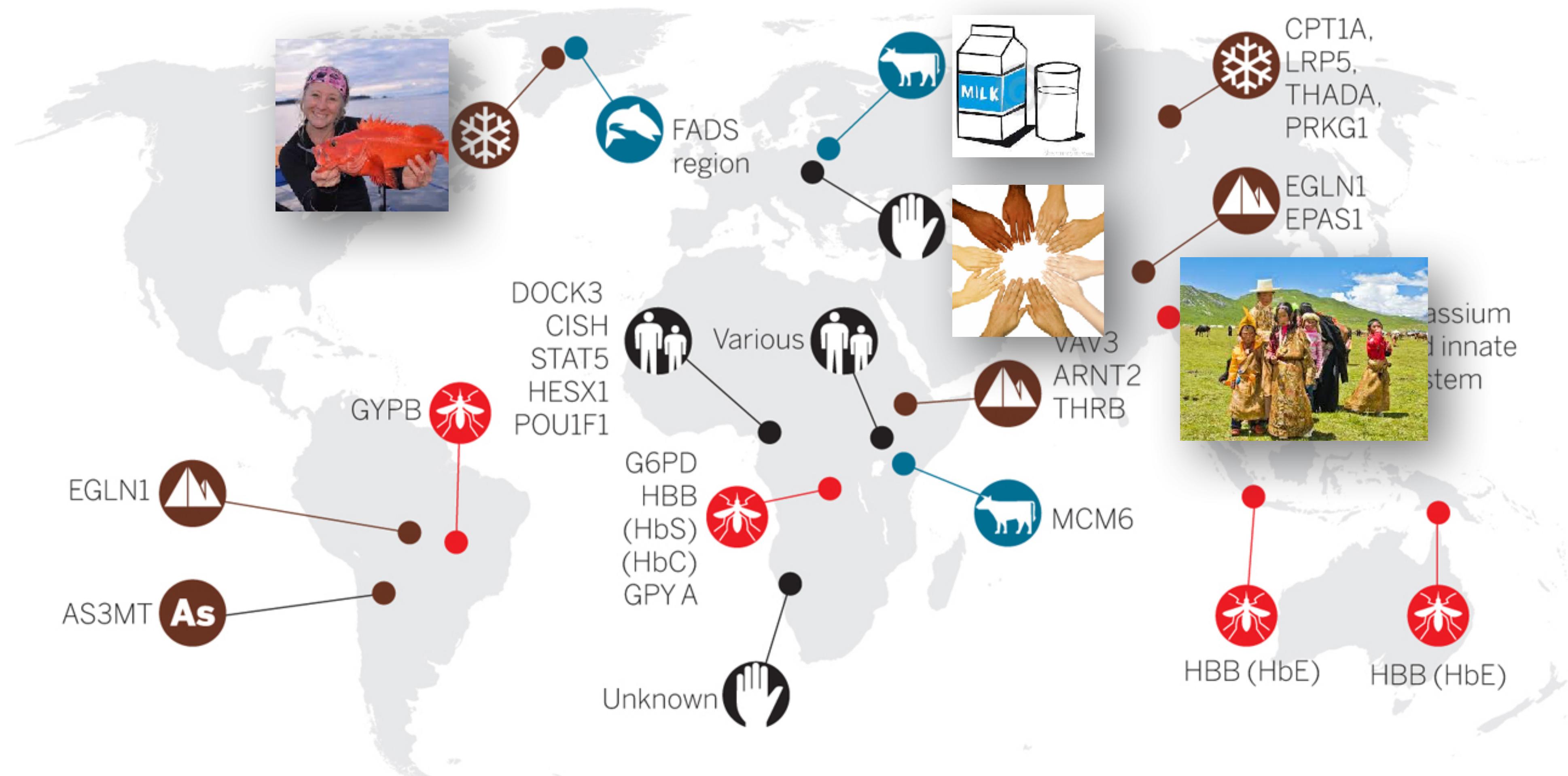
irtifa



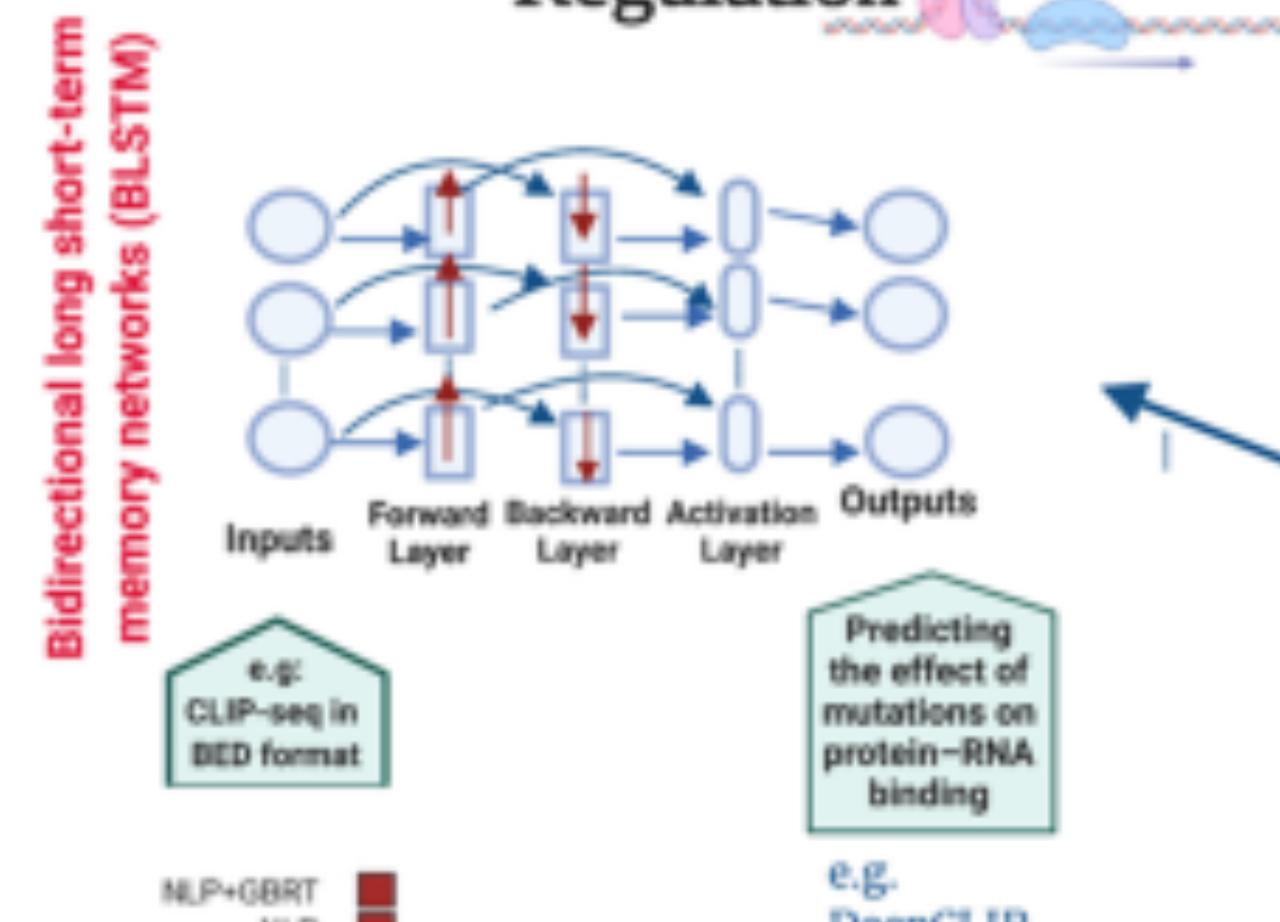
acik renk
ten



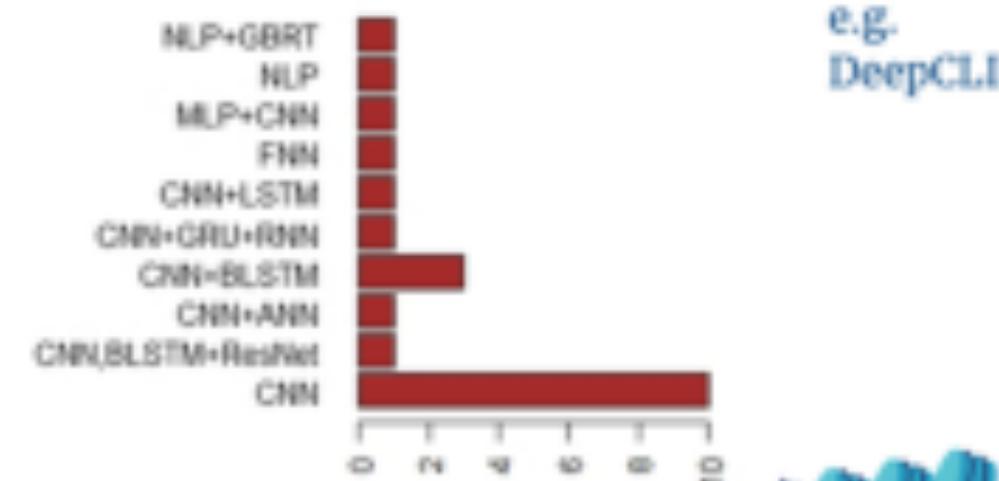
kisa
boy



Gene Expression & Regulation

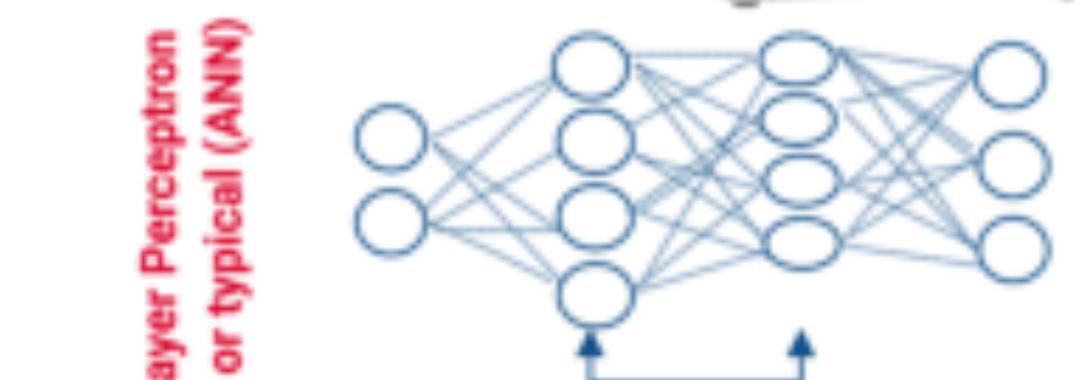


e.g:
CLIP-seq in BED format

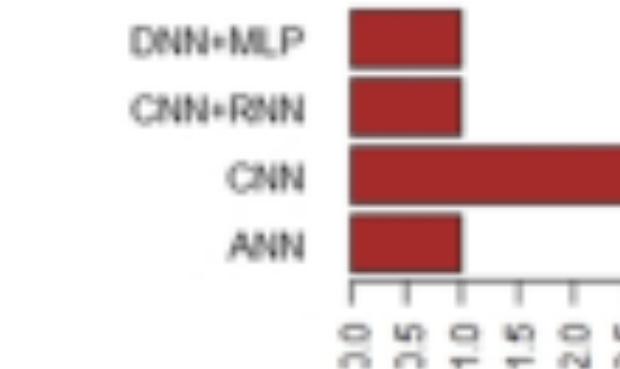


Predicting the effect of mutations on protein-RNA binding
e.g. DeepCLIP

Variants calling & Annotation

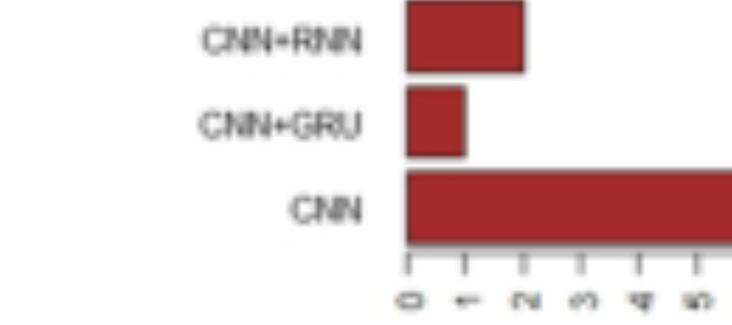
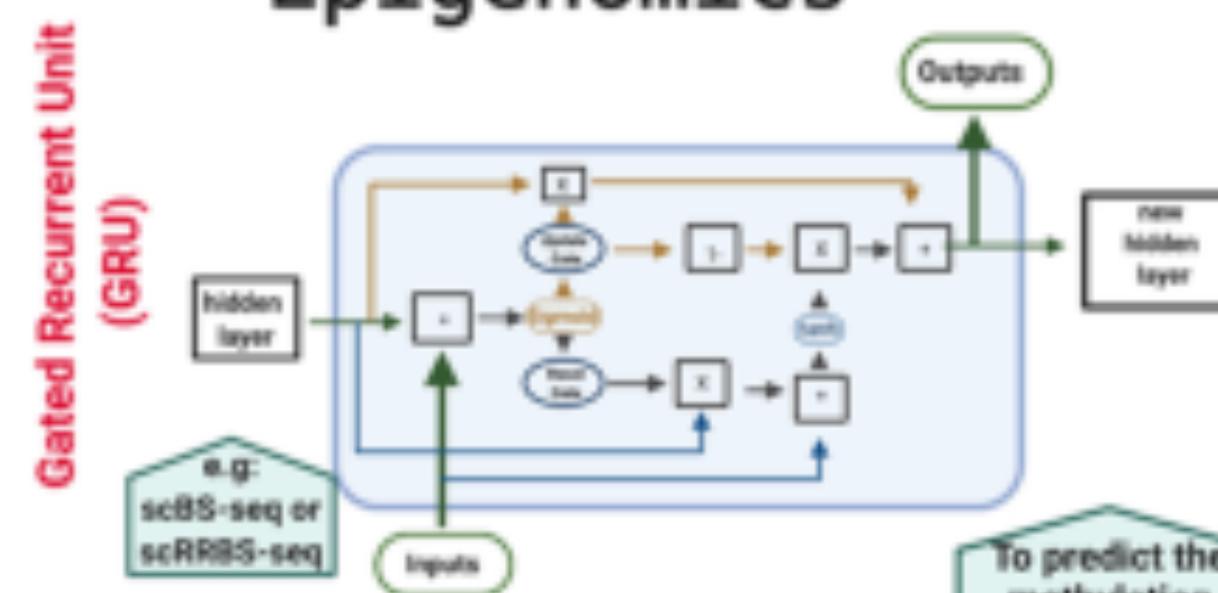


e.g:
WES data in VCF



Classify true and false variants from WES data
e.g. GARFIELD - NGS

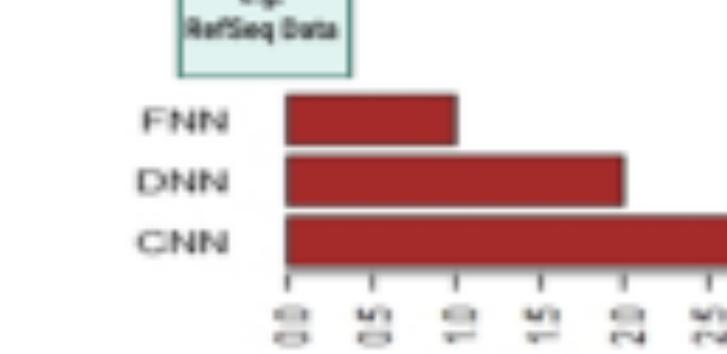
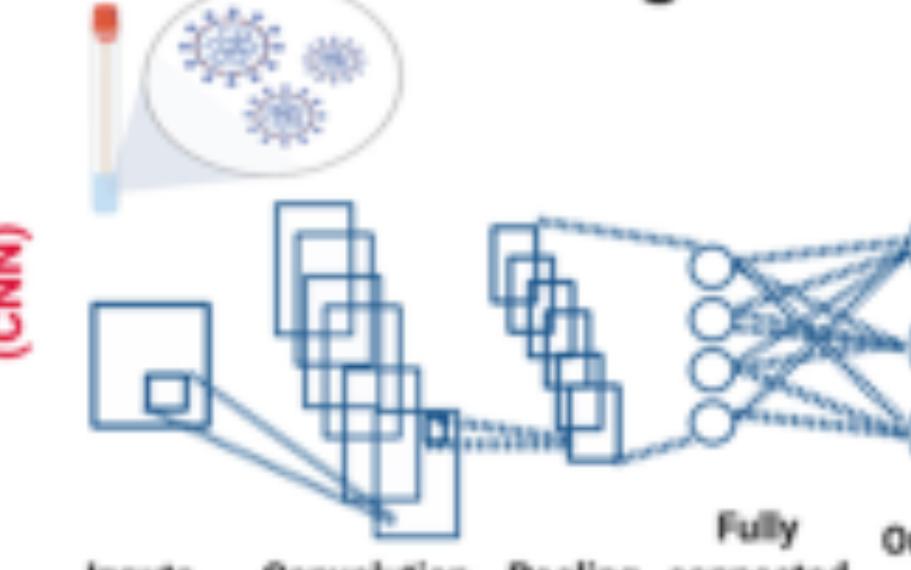
Epigenomics



To predict the methylation states from single-cell data
e.g. DeepCpG

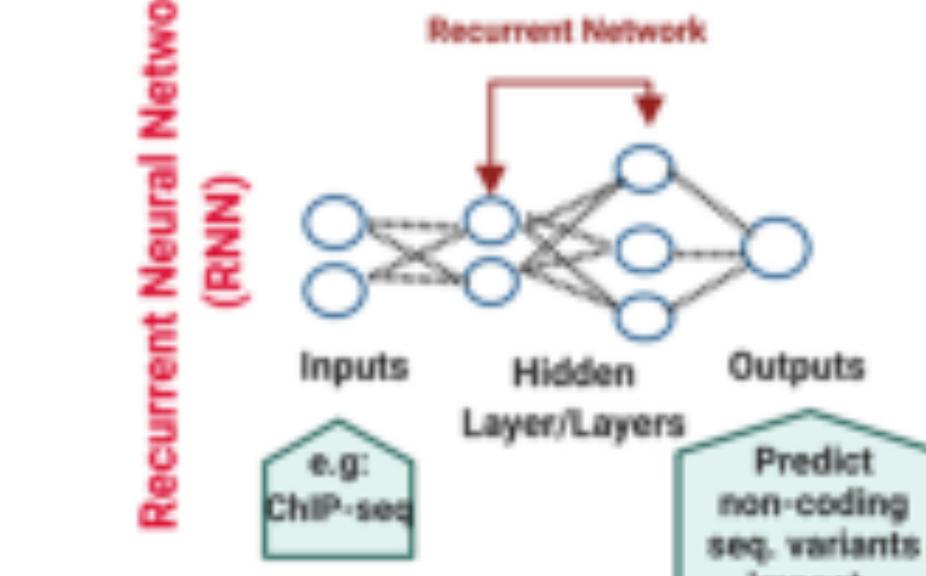
Pharmacogenomic

Convolutional Neural Network (CNN)

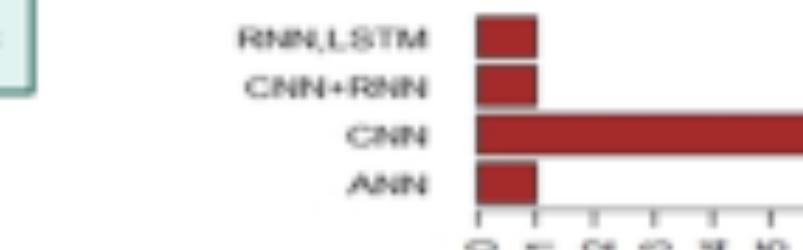


Predict the beta-lactamase (BLs) using protein or genome sequence datasets
e.g. DeepBL

Disease Variants



e.g:
ChIP-seq



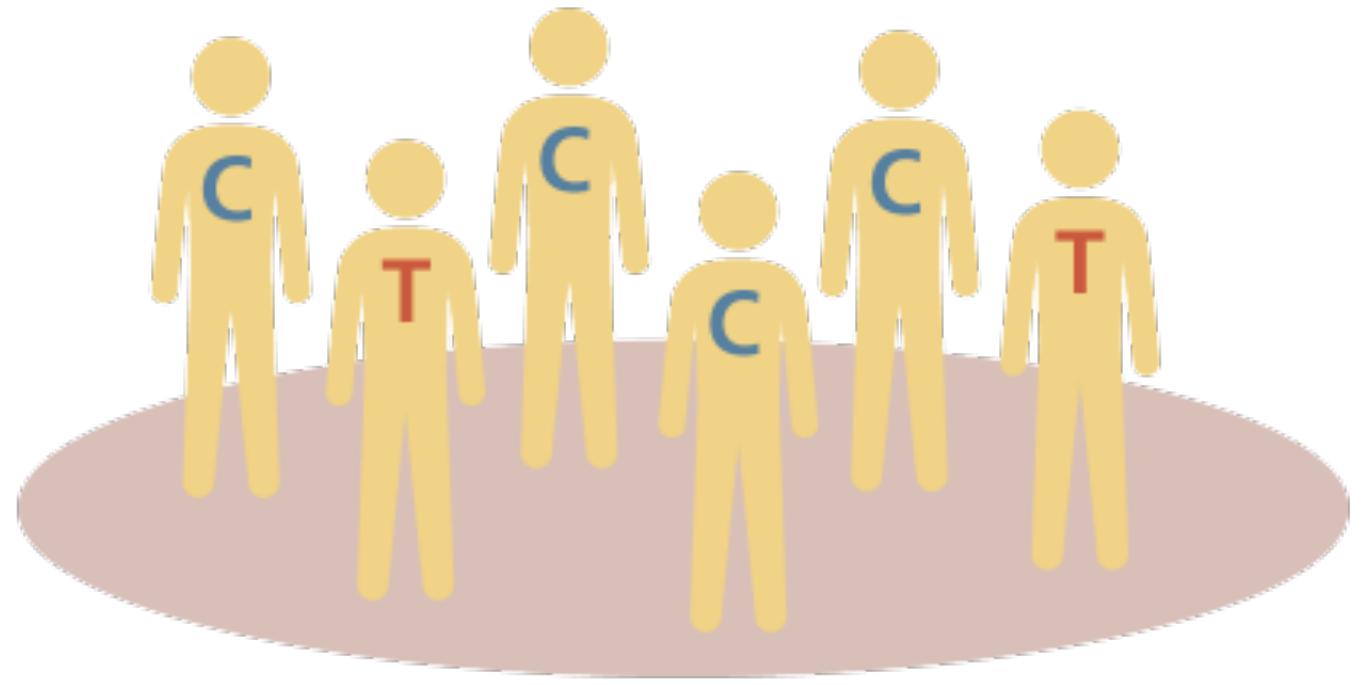
Predict non-coding seq. variants impact
e.g. DeepMLO

Genomlar Arası Bağlantılar (GWAS)

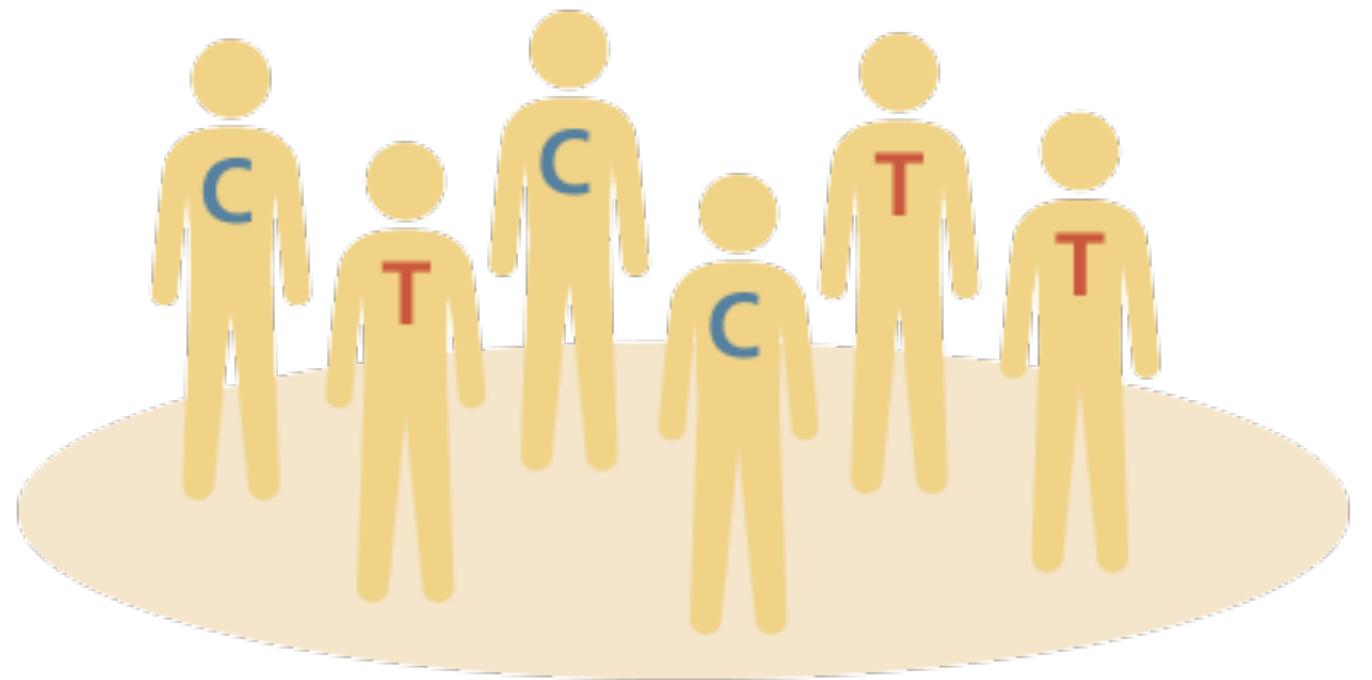
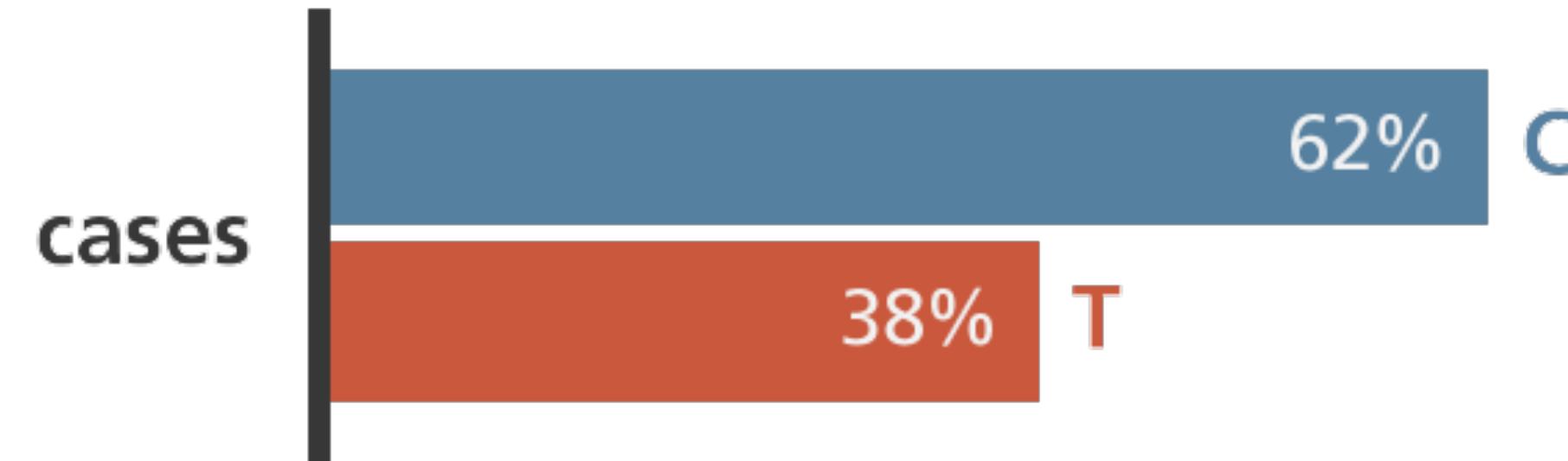


6734 kişi:
DNA dizisi
Hastalıklar, boy kilo endeksi, alışkanlıklar, etnik köken

Genomlar Arası Bağlantılar (GWAS)



cases (n=1,000)
people with heart disease



controls (n=1,000)
people without heart disease



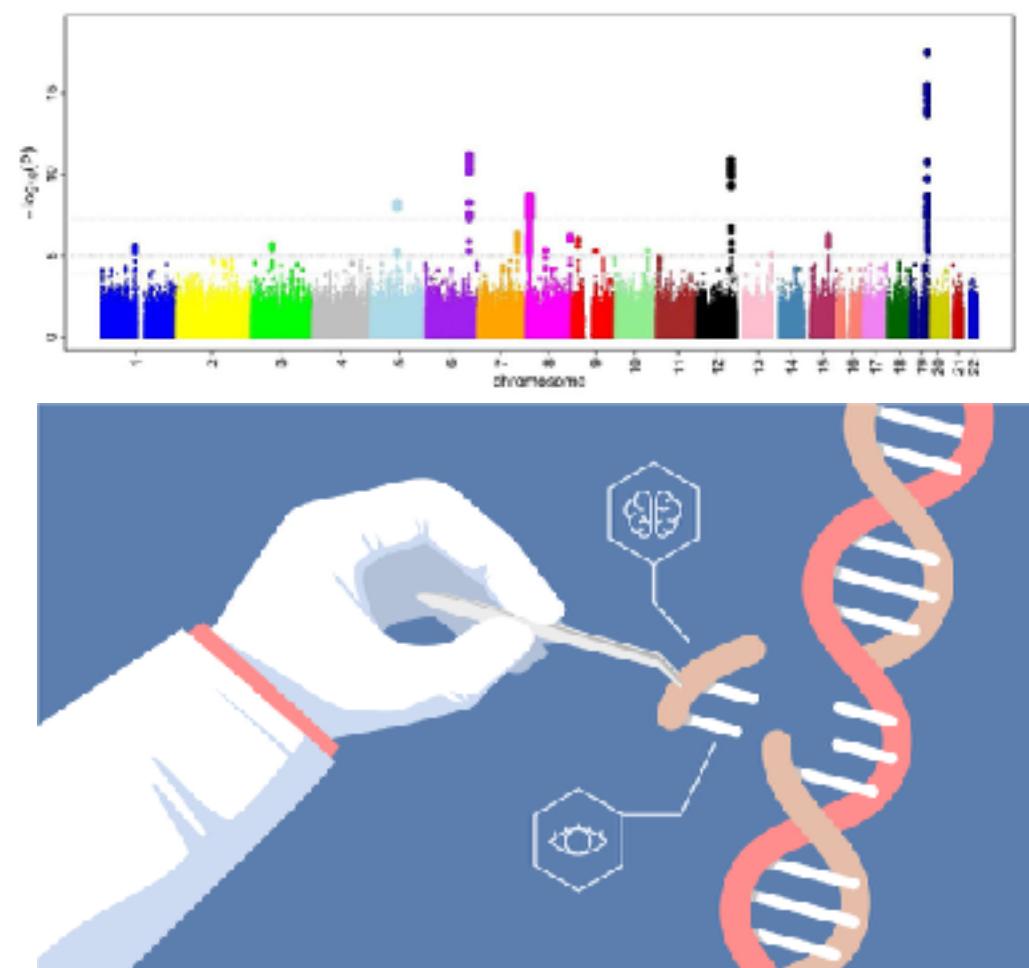
Kanser Genomigi



**Cogu mutasyonun aslında
hiç bir seye etkisi yoktur**

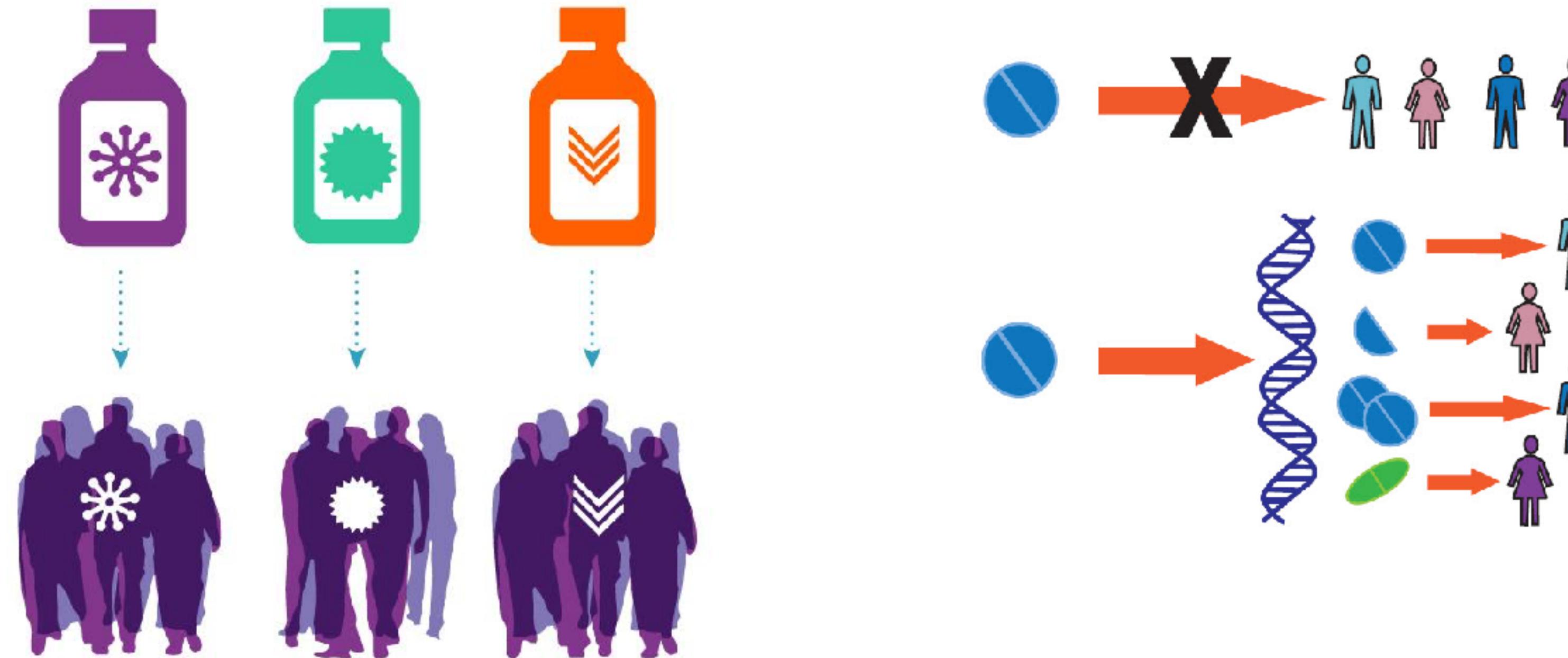
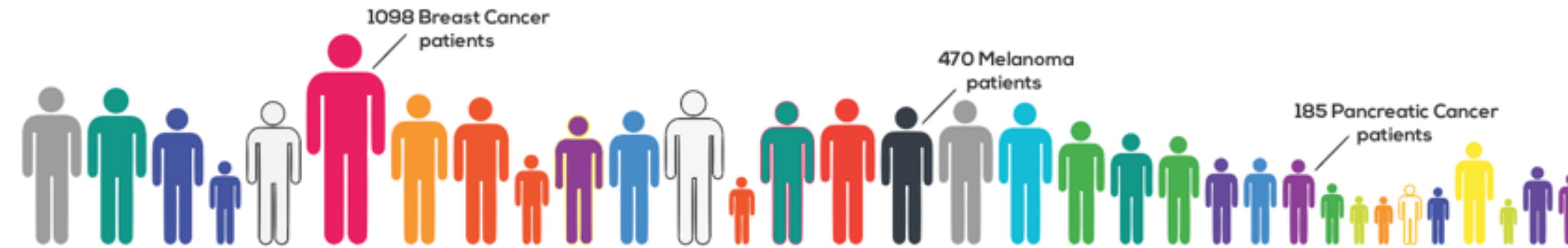
E hangi mutasyonlar?

- In silico adaylar - istatistikî tahmin
- In vitro mutajenez - deneysel evrim



Kisisel Tip

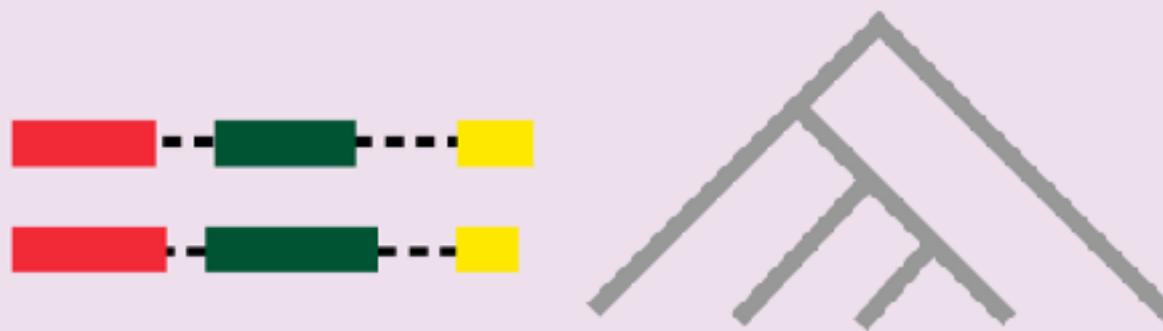
Matched tumor & normal tissues from more than **11,000** patients, representing **33** cancer types.



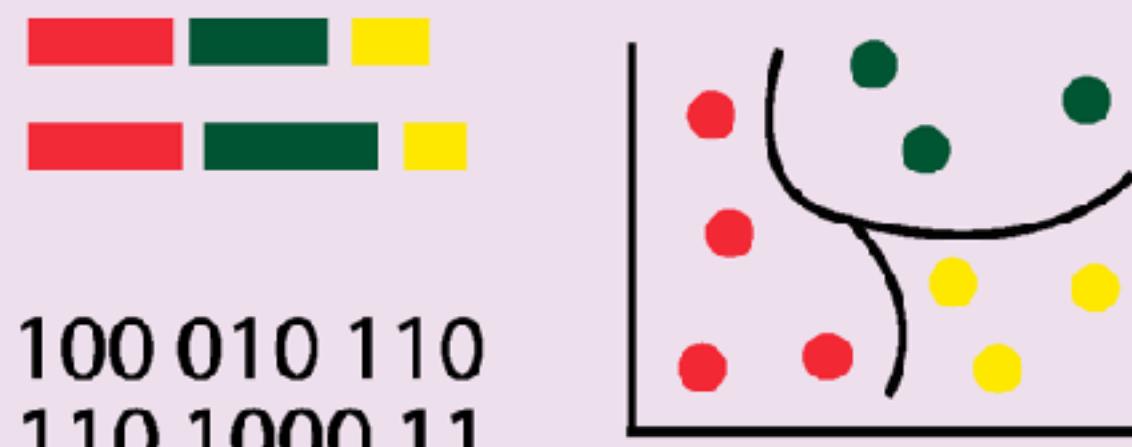
Machine and deep learning integration with bioinformatics

Molecular evolution

Phylogenetic inference

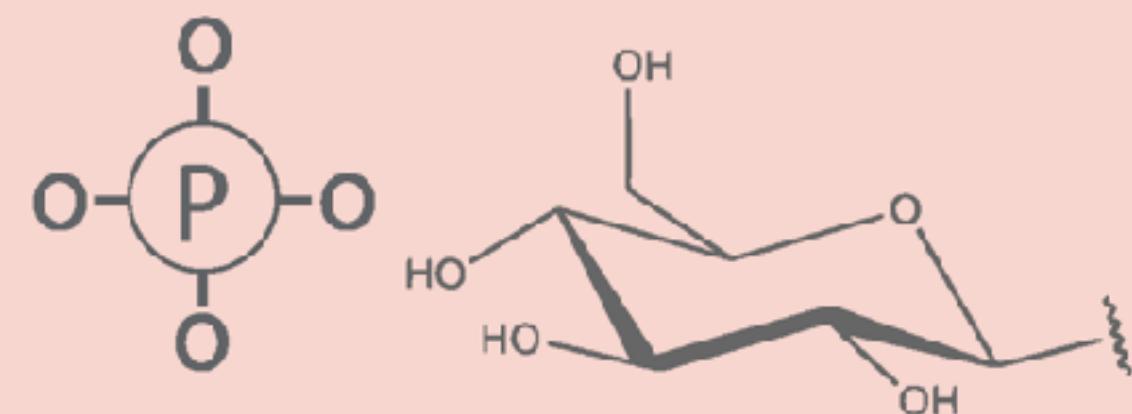


Alignment-free sequence classification

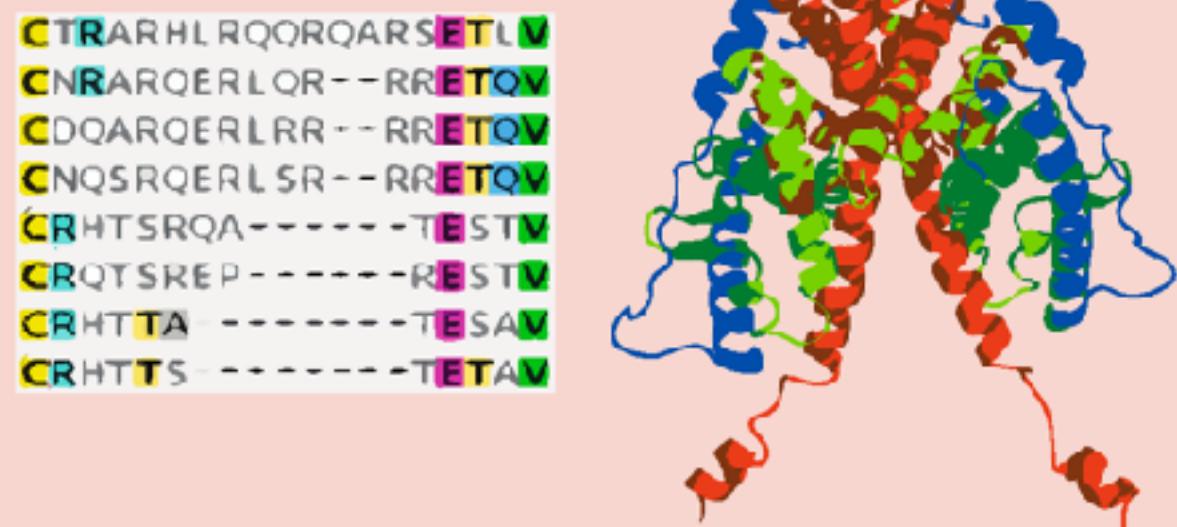


Protein structure Analysis

Post translational modification

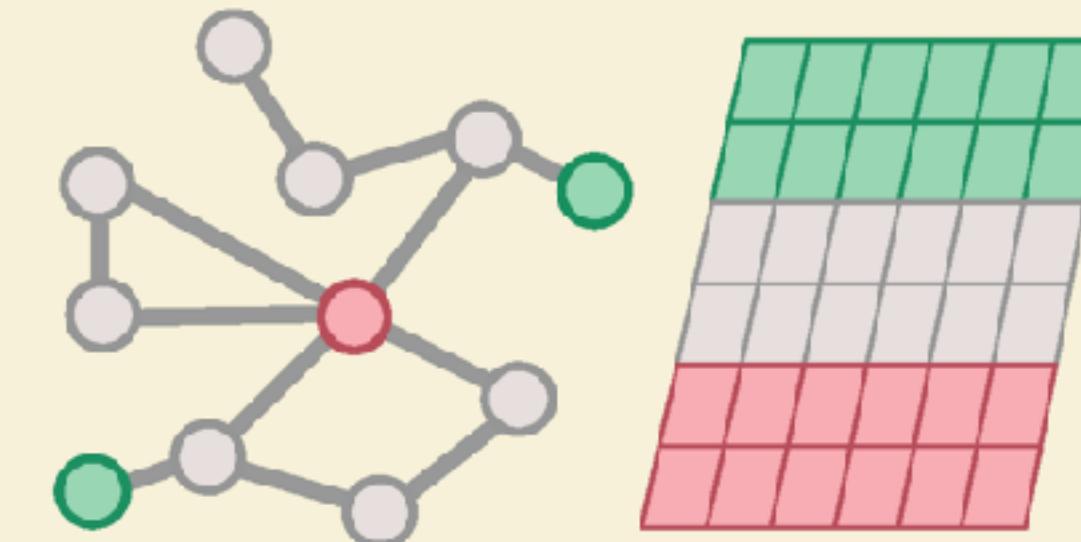


Folding and structure

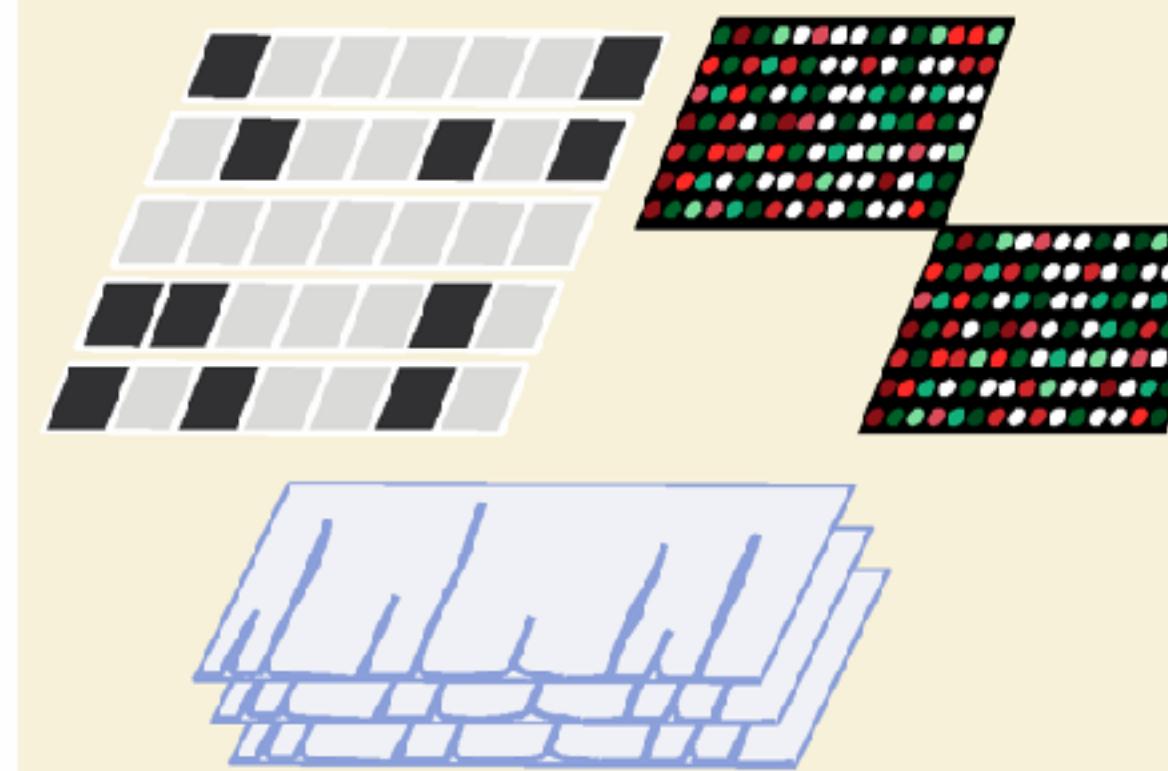


Systems biology

Biological Networks

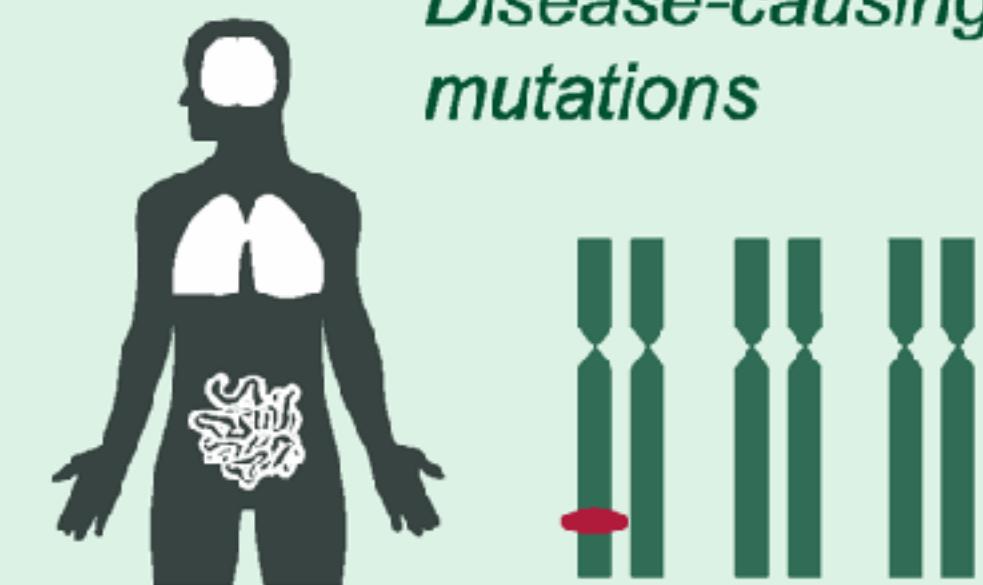


Multi-Omics integration

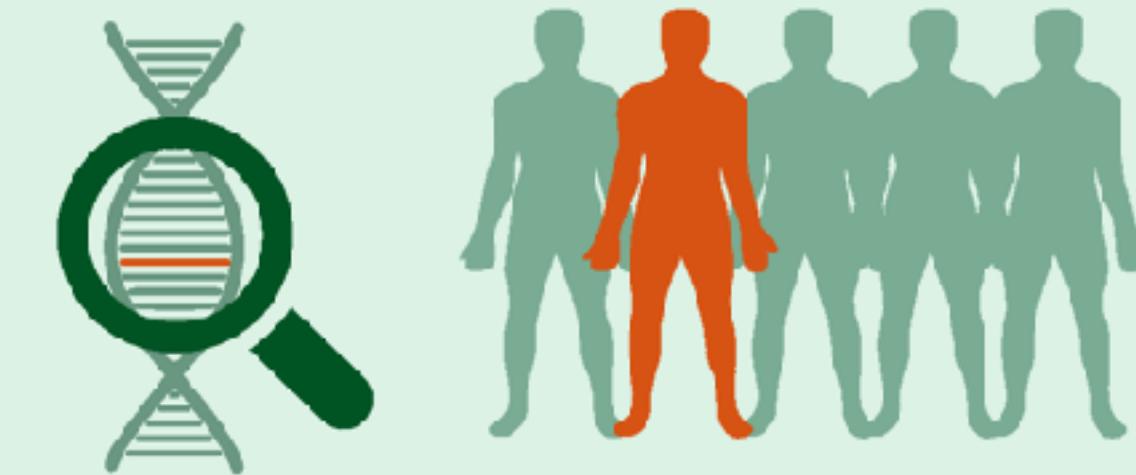


Genomics for Disease Research

Disease-causing mutations



Biomarkers discovery



- Inference of tree topology
- Sequence classification
- Viral sequence identification
- Functional annotation

- Phosphorylation site prediction
- Protein glycosylation prediction
- Protein contact maps prediction
- Structural homology prediction

- Biological networks construction
- Biological interactions prediction
- Pathway dynamics prediction
- Platform integration frameworks

- Disease associated genes and mutations
- Biomarkers
- Precision medicine applications

DNA Dizileme Teknolojileri

Yeni Nesil Dizileme

\$3,000,000,000



\$20,000,000



\$2,000,000



\$200,000



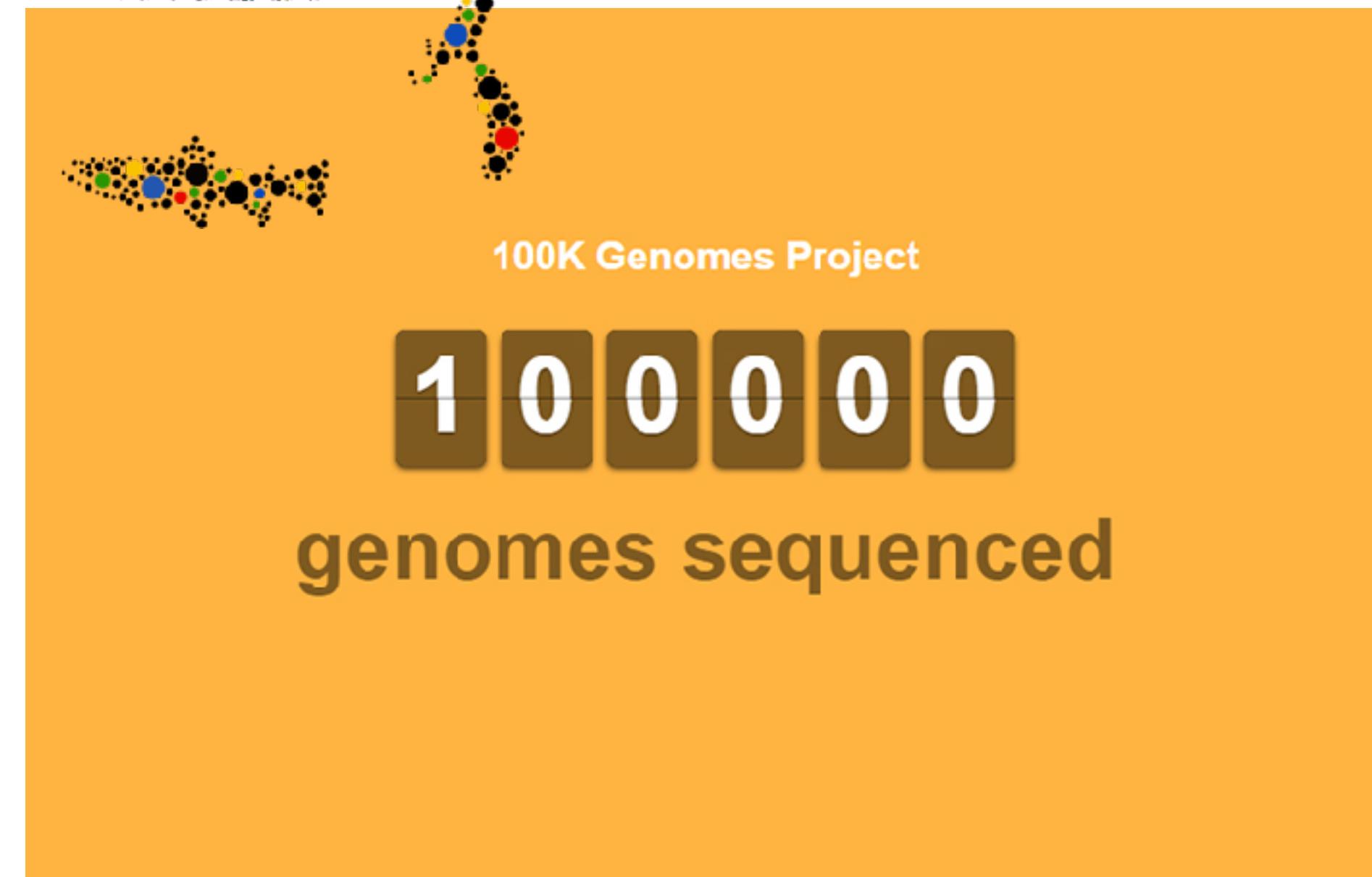
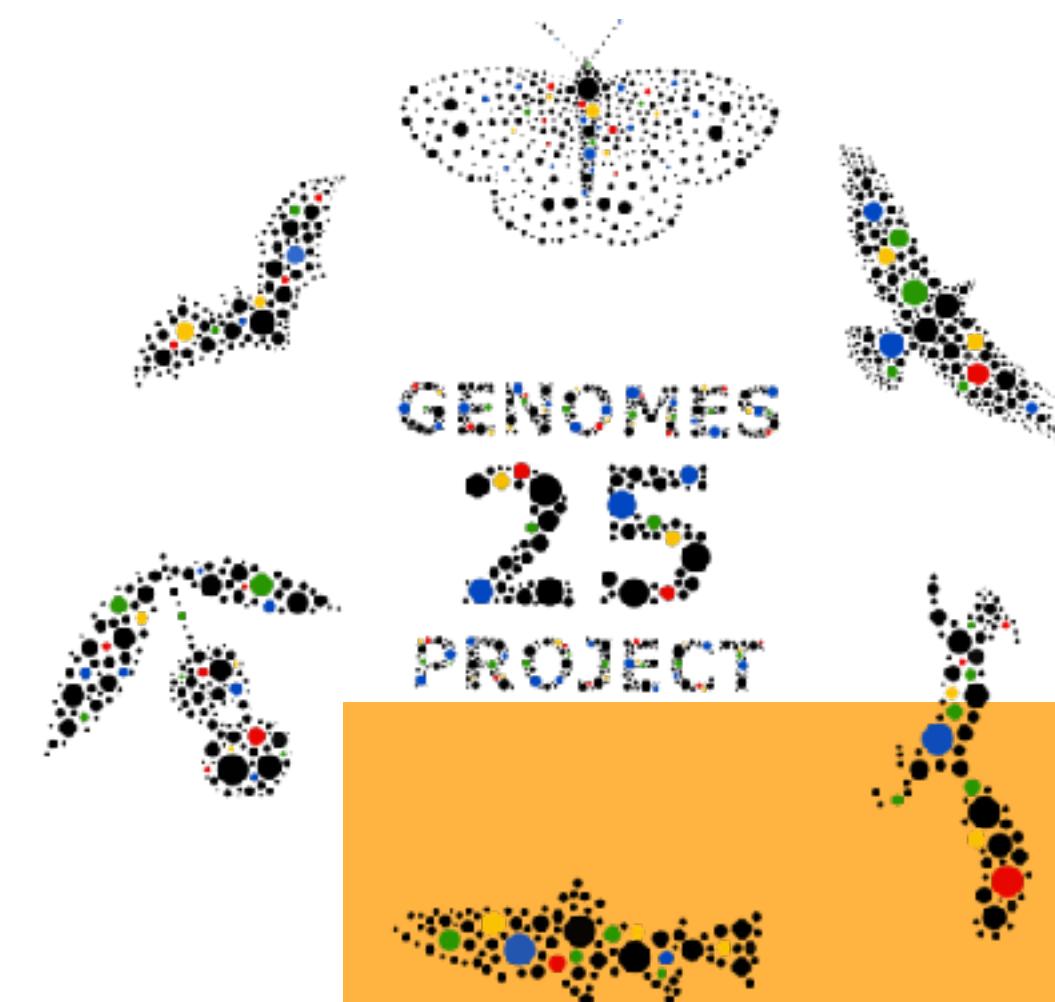
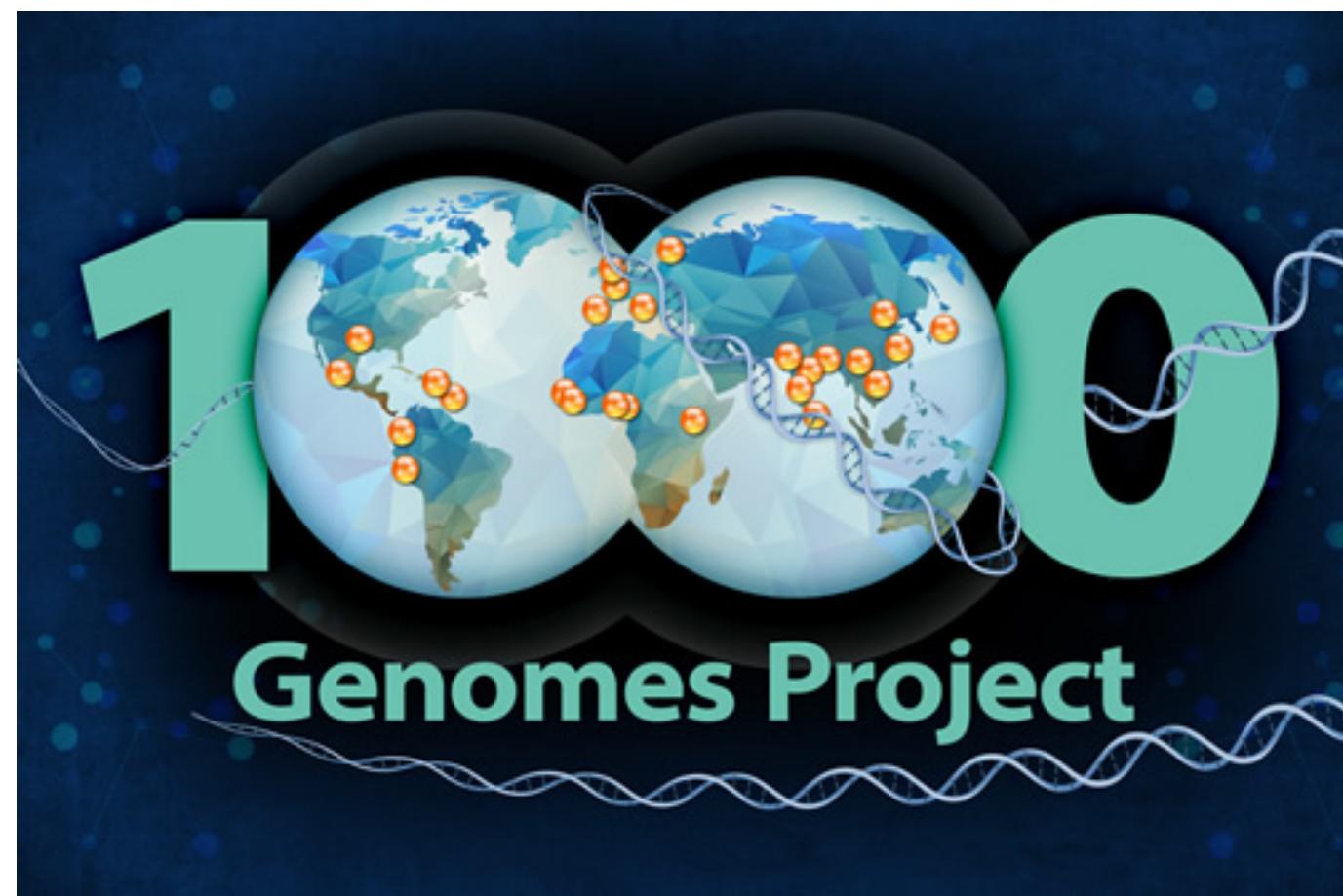
\$10,000



\$1,000



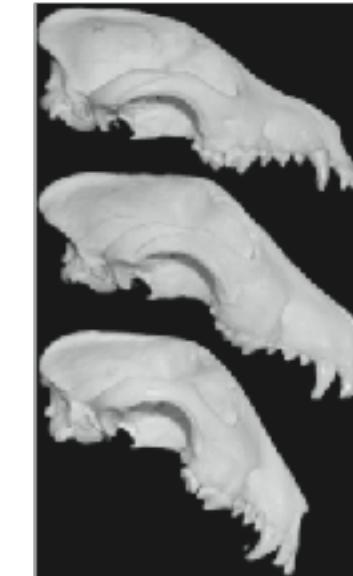
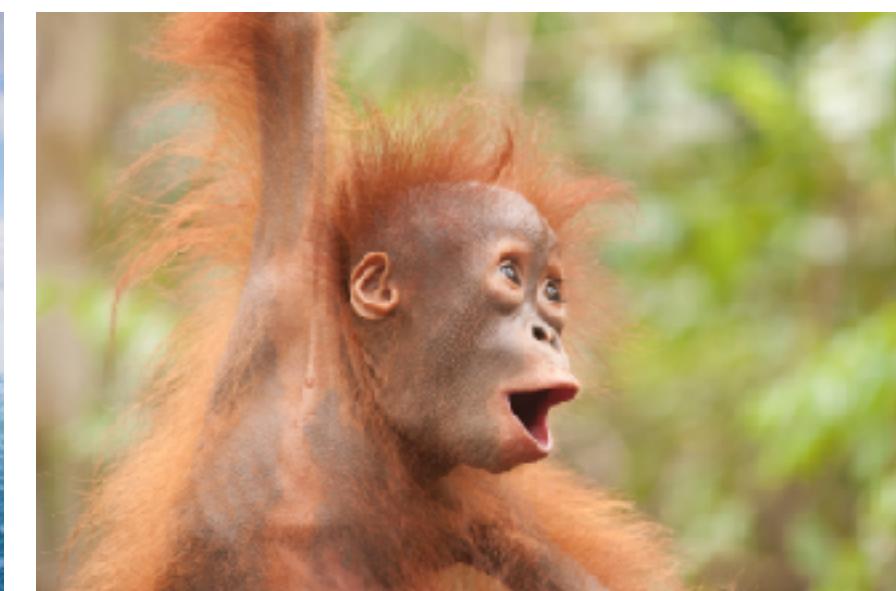
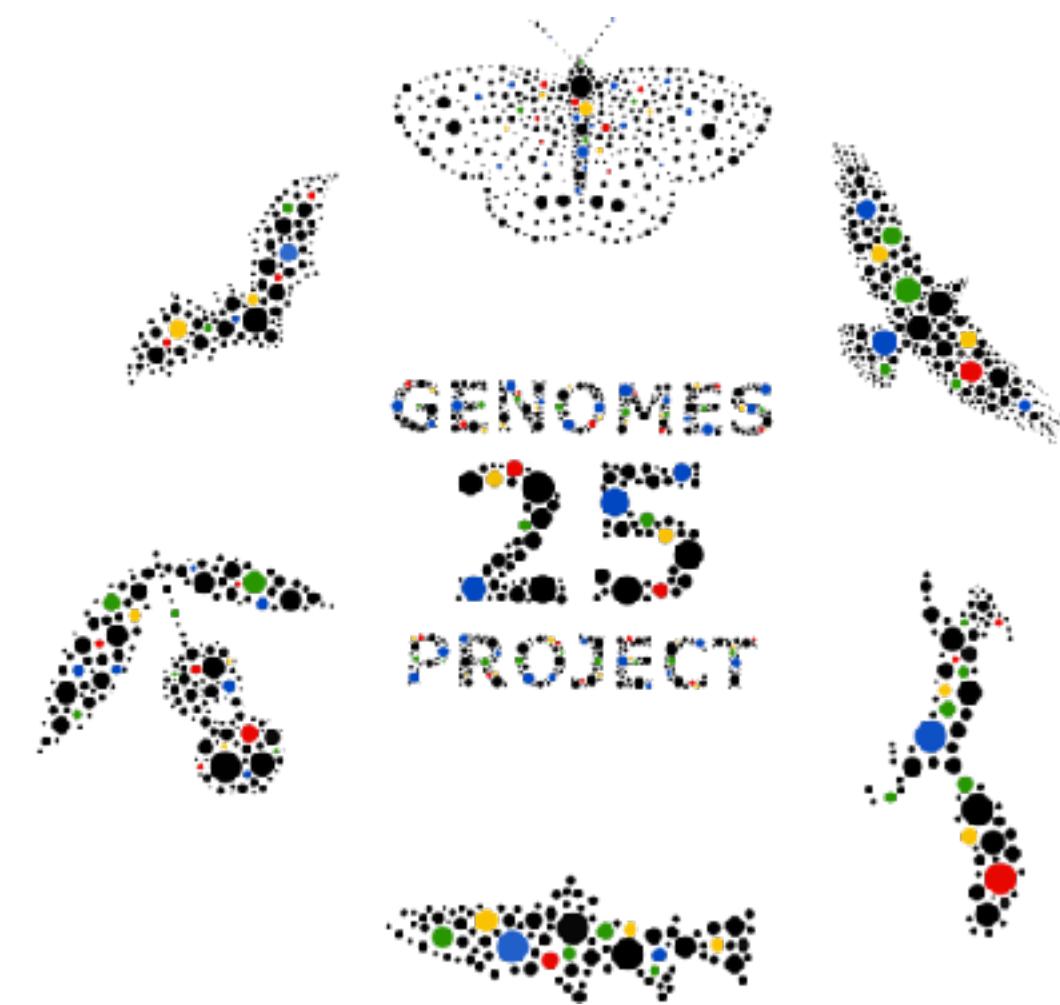
Genom Projeleri



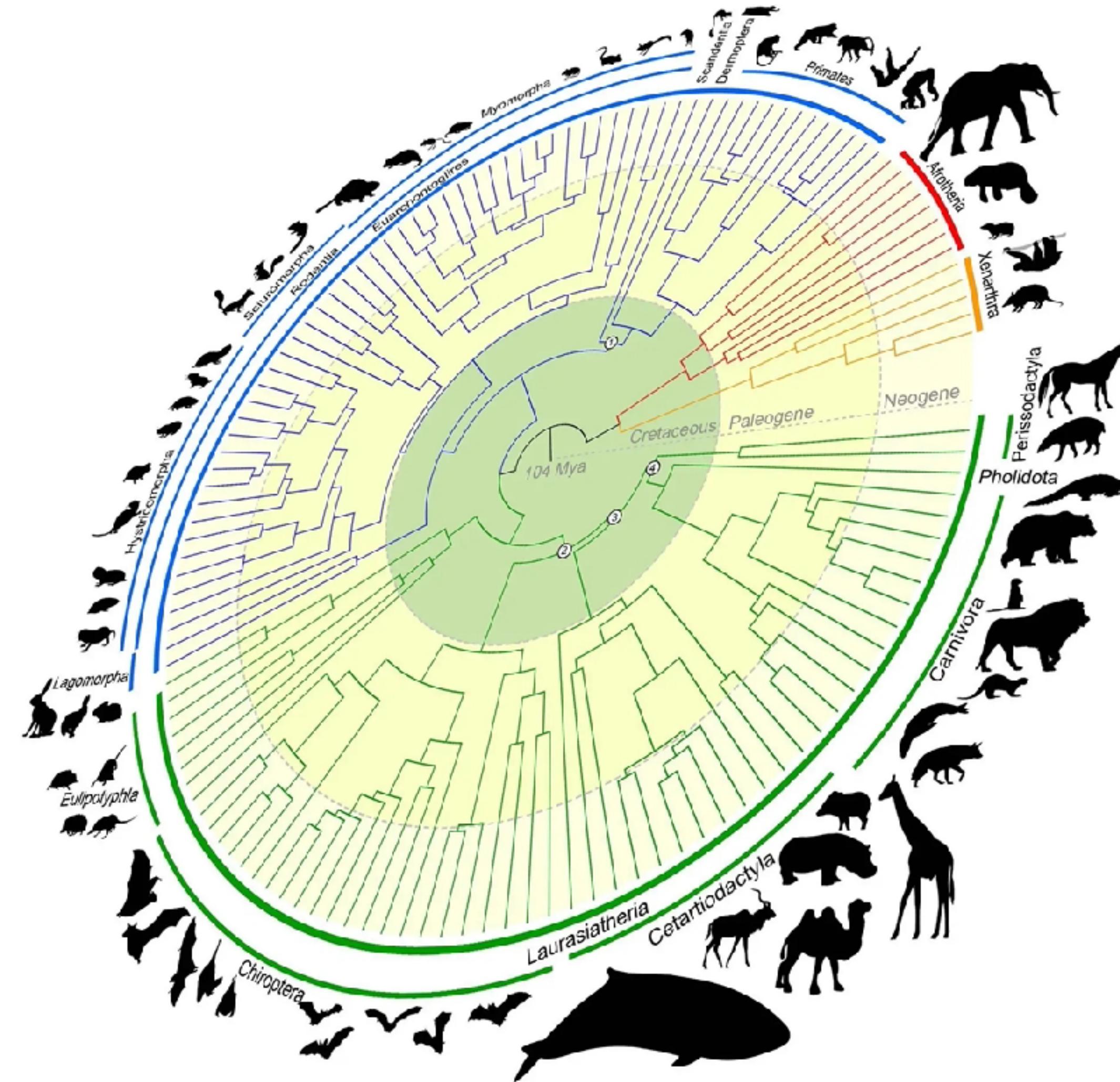
NIH HUMAN
MICROBIOME
PROJECT

Genom Projeleri

Biological Diversity and Conservation



Phylogenetic Trees



Protein Verisetleri

- Protein-protein interaction: <https://string-db.org/>
- Protein databases: <https://www.uniprot.org/> <https://www.rcsb.org/>
- Best friend of a bioinformatician: <https://www.biostars.org>
- Another best friend: <https://stackoverflow.com>

Bioenformatik Forumları

Notebooklar

- Jupyter Notebook, bioinformatician's diary: <https://jupyter.org>

Genomik Analizleri

- Bedtools : <https://bedtools.readthedocs.io/en/latest/>
- USCS Table Browser: <https://genome.ucsc.edu/cgi-bin/hgTables>
- Biomart Ensembl: <https://useast.ensembl.org/index.html>
- ncbi entrez: <https://www.ncbi.nlm.nih.gov/>
- Pan Cancer, TCGA: <https://cancergenome.nih.gov>

Kanser Verisetleri

Gen Verisetleri

- Gene Cards: <https://www.genecards.org>
- DAVID for Gene Ontology terms: <https://david.ncifcrf.gov/summary.jsp>

Genom Verisetleri

- 1000 genomes new name : <https://www.internationalgenome.org>

Kodlama

- R Circos - <https://cran.r-project.org/web/packages/BioCircos/vignettes/BioCircos.html>
- R packages: <https://www.bioconductor.org>
- python packages: Biopython