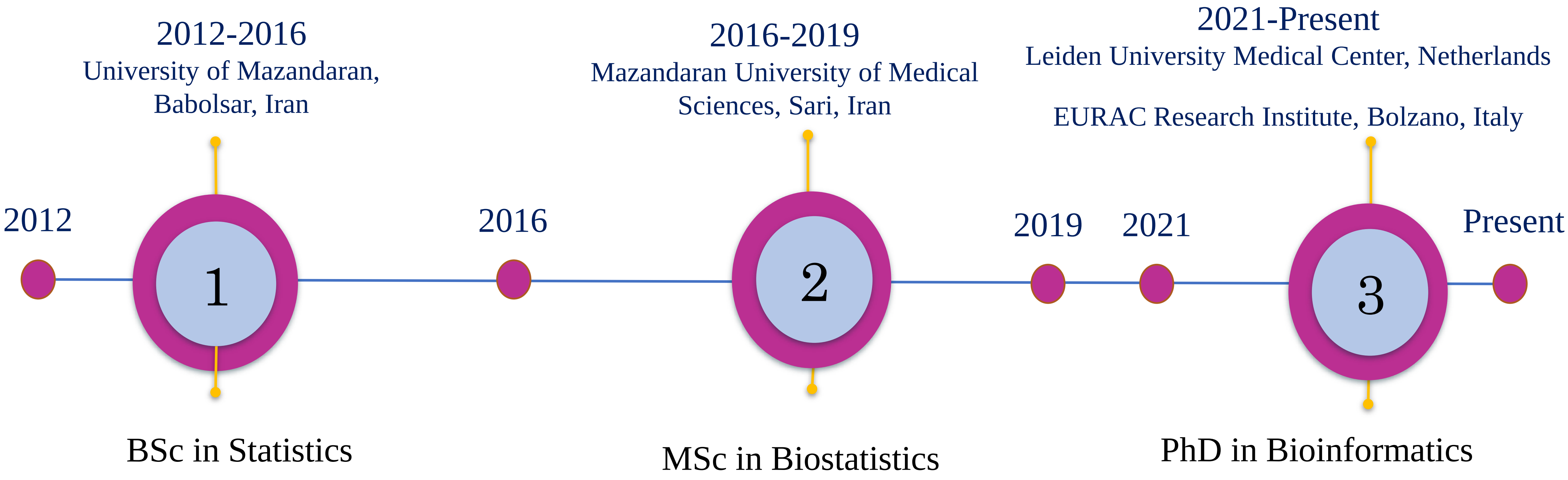
GWAS and Population Genetics

Dariush Ghasemi
PhD Fellow at
Leiden University Medical Center
EURAC Research Institute

July 2022

Education



2012-2016
University of Mazandaran, Babolsar, Iran

BSc in Statistics

2016-2019
Mazandaran University of Medical Sciences, Sari, Iran

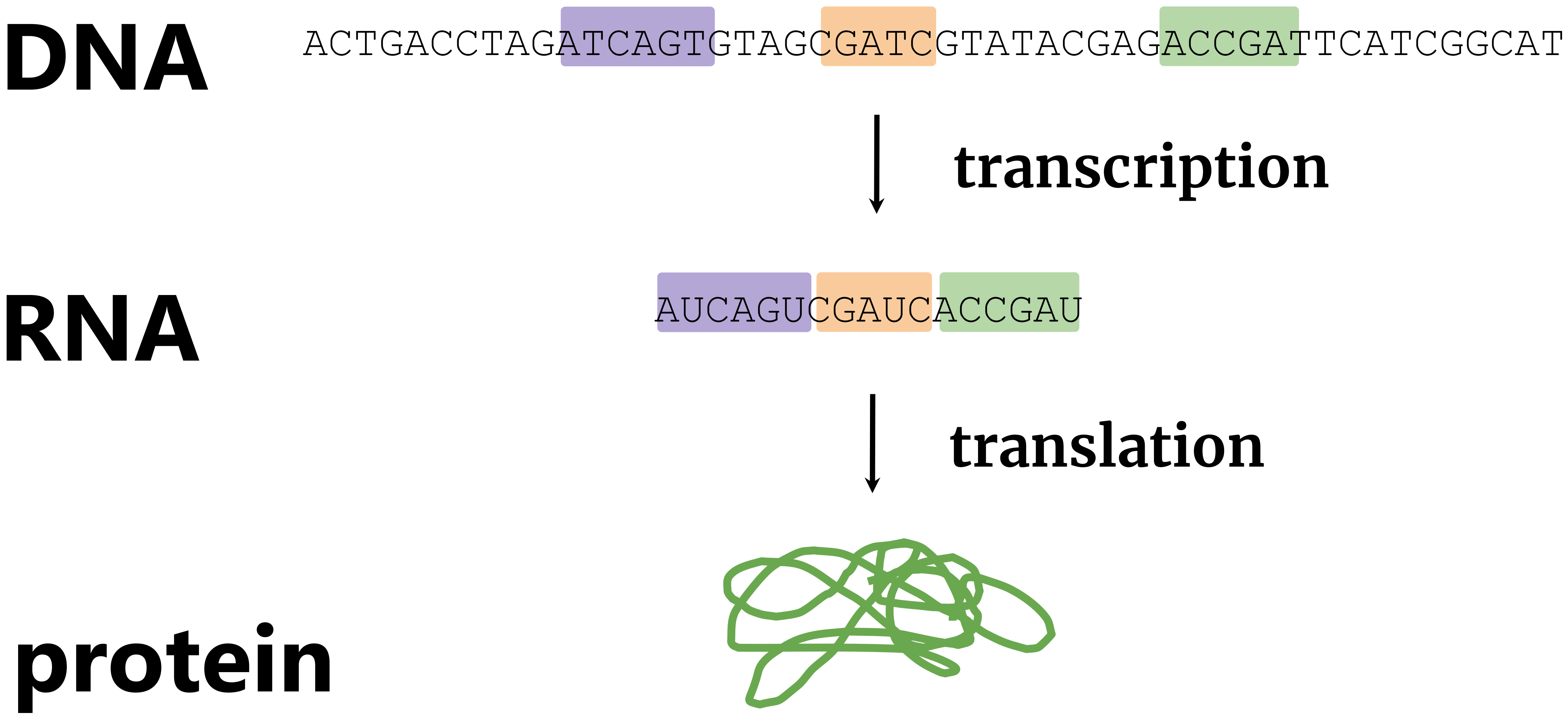
MSc in Biostatistics

2021-Present
Leiden University Medical Center, Netherlands

EURAC Research Institute, Bolzano, Italy

PhD in Bioinformatics

Genome



DNA

ACTGACCTAGATCAGTGTAGCGATCGTATACGAGACCGATTCATCGGCAT

transcription

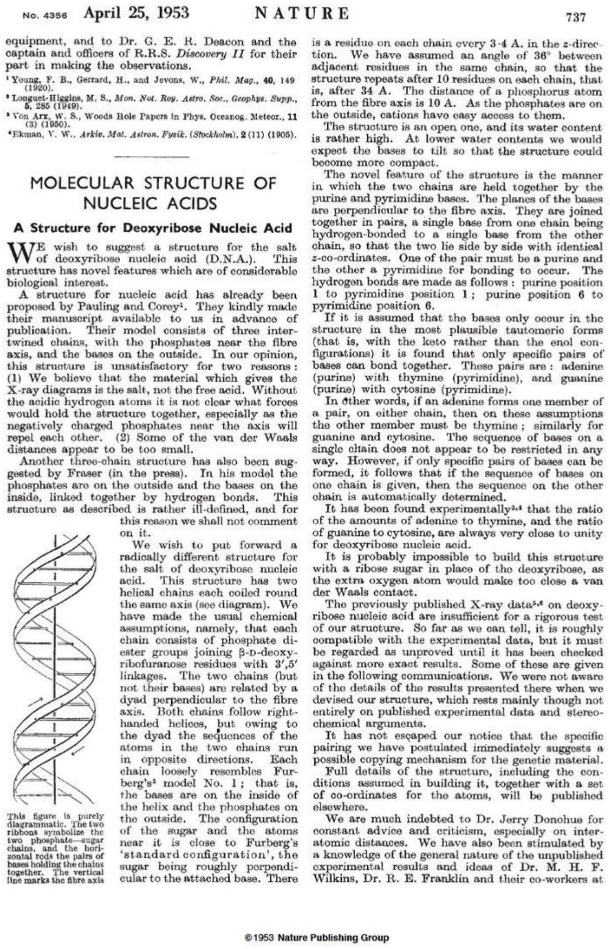
RNA

AUCAGUCGAUCACCGAU

translation

protein

Structure of a DNA



No. 4356 April 25, 1953 NATURE 737

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).

[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).

[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).

[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik.* (*Stockholm*), **2** (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

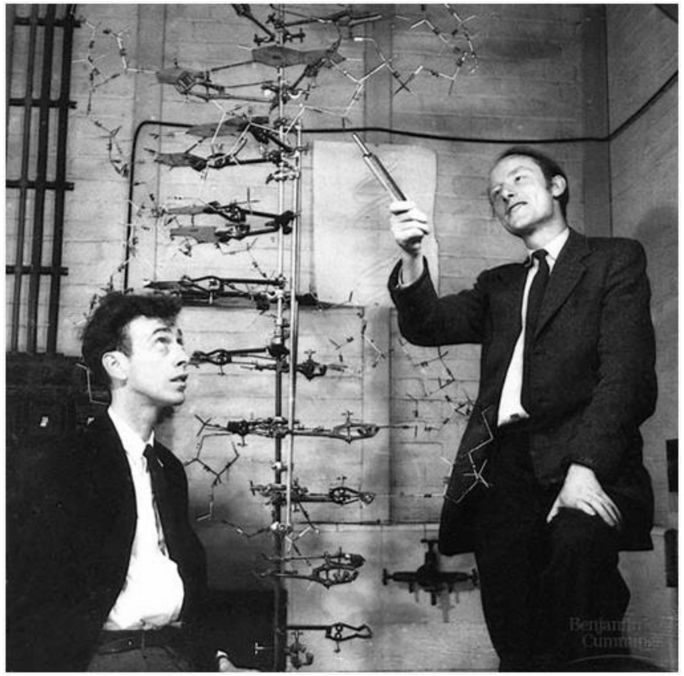
It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

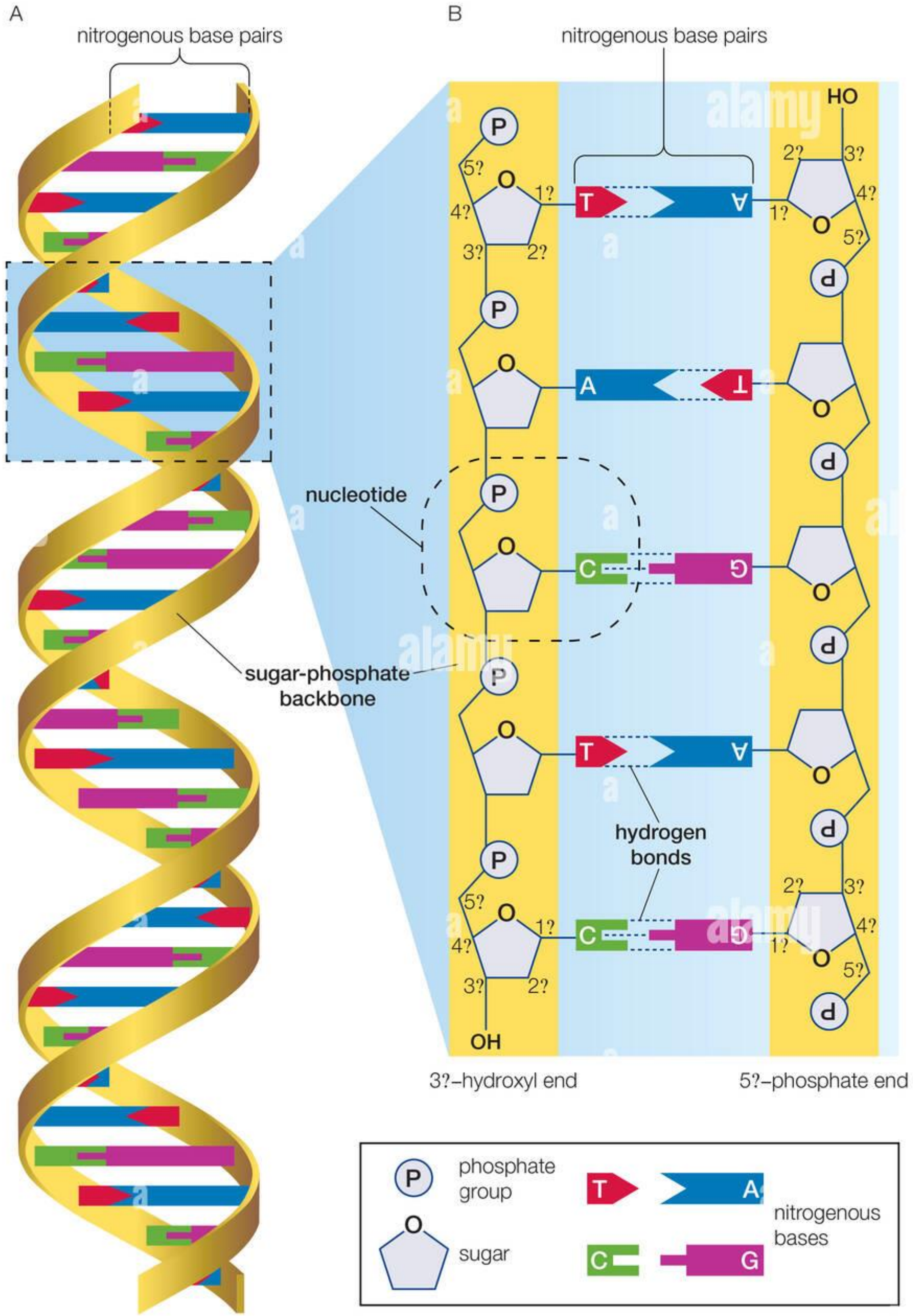
The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at





Bases:

Adenine

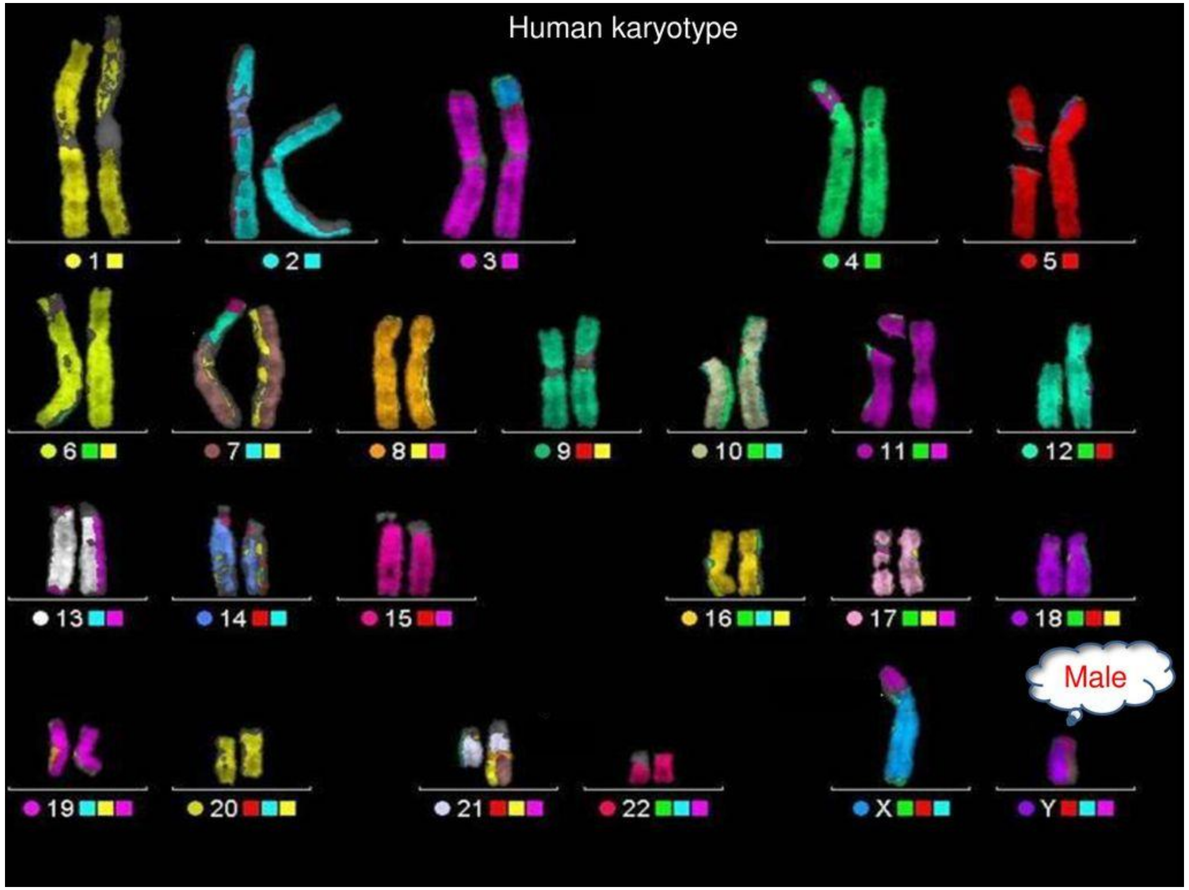
Thymine

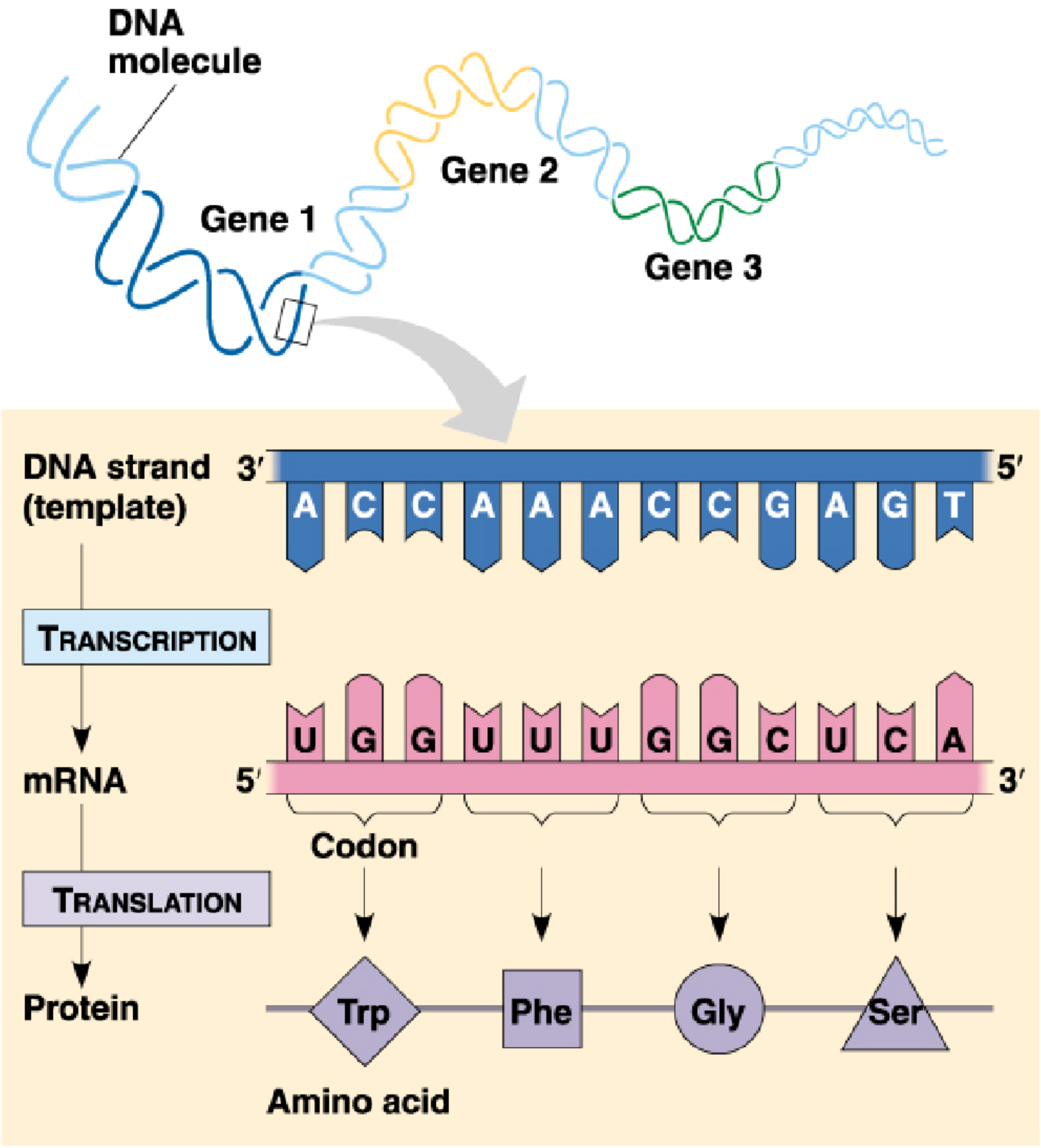
Cytosine

Guanine

1953 Watson & Crick

Genome





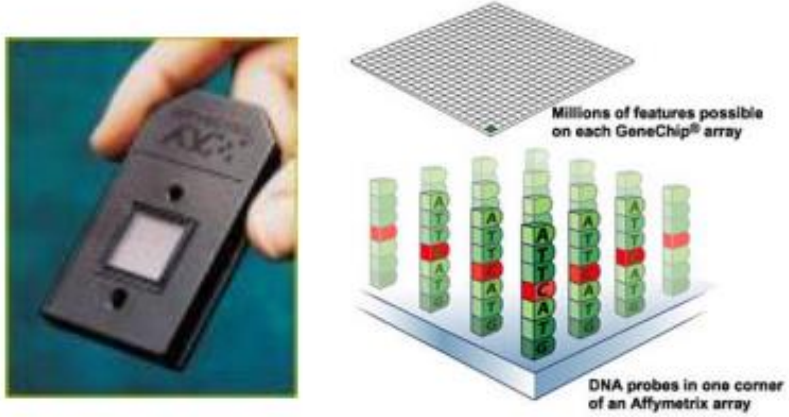
©1999 Addison Wesley Longman, Inc.

Twenty years of technological progress

nature
the human genome
Nuclear fission
Five-dimensional energy landscapes
Seafloor spreading
The view from under the Arctic ice
Career prospects
Sequence creates new opportunities
naturejobs
genomics special

Mapping the human genome
2001

Characterizing common variation
2003 - 2007


Millions of features possible on each GeneChip® array
DNA probes in one corner of an Affymetrix array

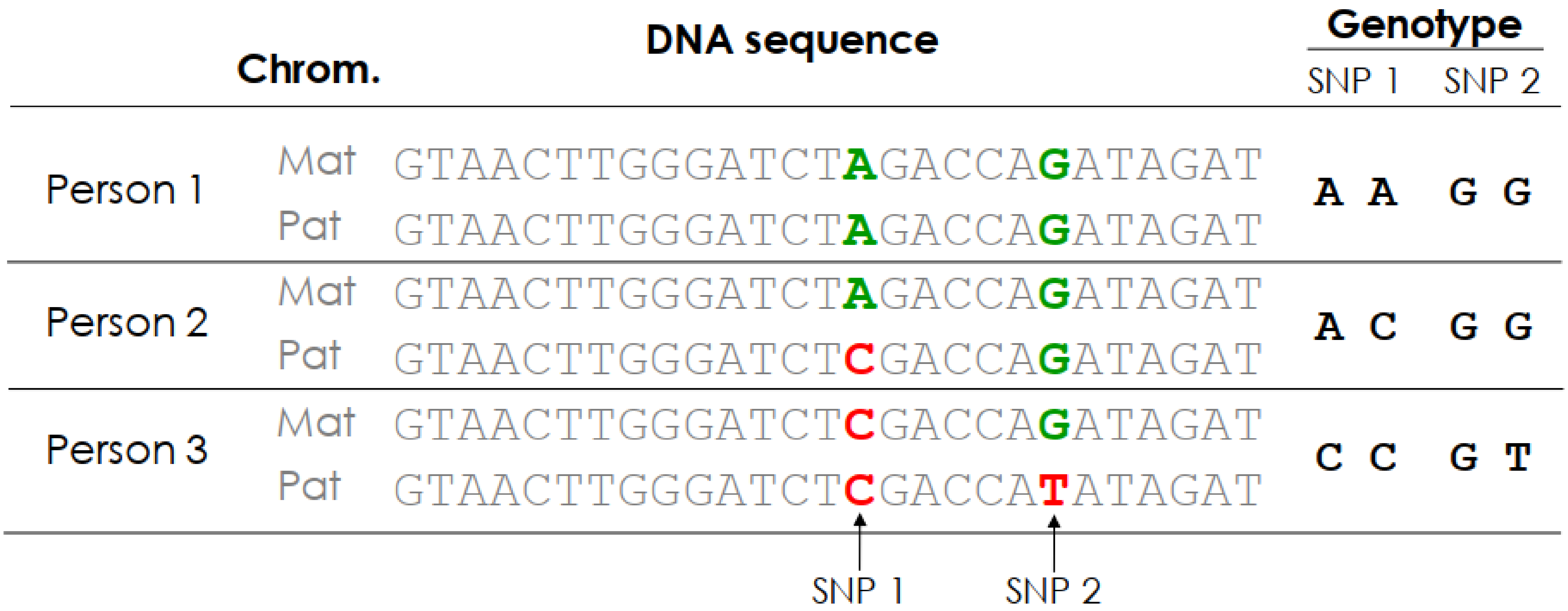
Genome-wide association (GWAS)
2005 - present

Next generation sequencing
2009 - present

Technologies

	Arrays	Exomes	Genomes
Upside	Hits + epidemiology	Gene identification	Comprehensive capture
Downside	Hit interpretation	Limited Scope	Cost (small N)

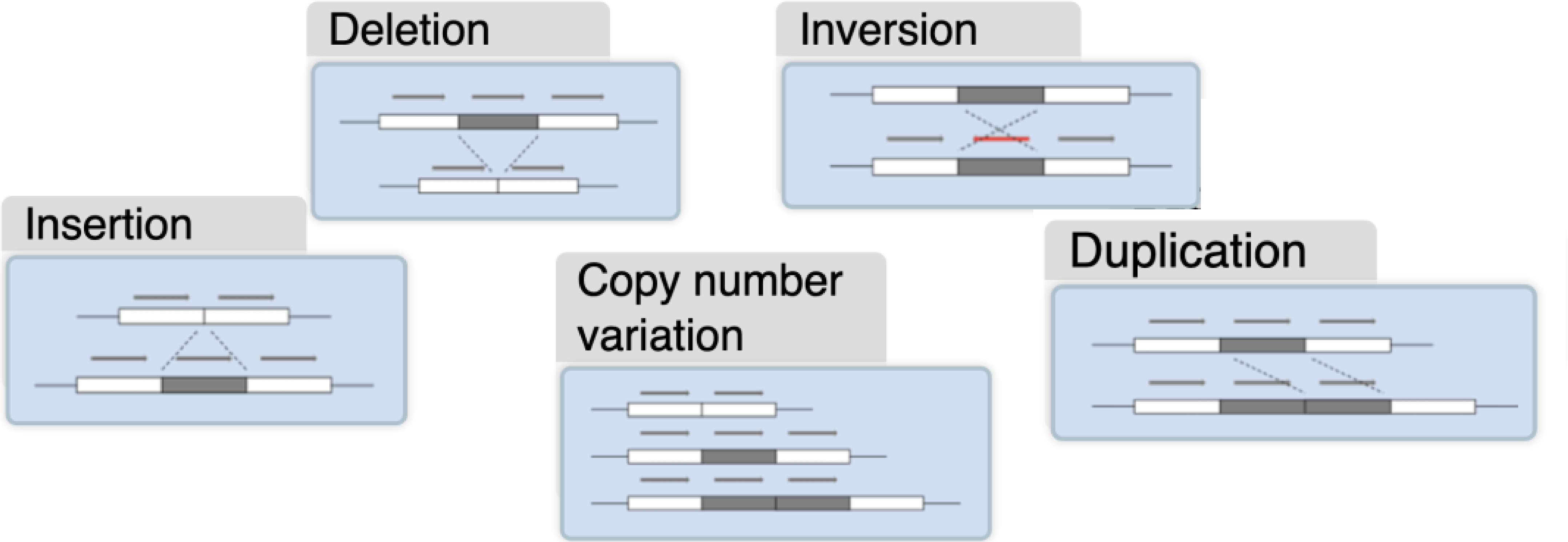
Single Nucleotide Polymorphism (SNP)



	Chrom.	DNA sequence	Genotype	
			SNP 1	SNP 2
Person 1	Mat	GTAACTTGGGATCTAGACCAGATAGAT	A A	G G
	Pat	GTAACTTGGGATCTAGACCAGATAGAT		
Person 2	Mat	GTAACTTGGGATCTAGACCAGATAGAT	A C	G G
	Pat	GTAACTTGGGATCTCGACCAGATAGAT		
Person 3	Mat	GTAACTTGGGATCTCGACCAGATAGAT	C C	G T
	Pat	GTAACTTGGGATCTCGACCATATAGAT		

- Mutation that arose at some point in demographic history
- Typically, each SNP has two alleles (bases)
- Each SNP is eventually given an “rs” number rs214621

Structural Mutations



Deletion

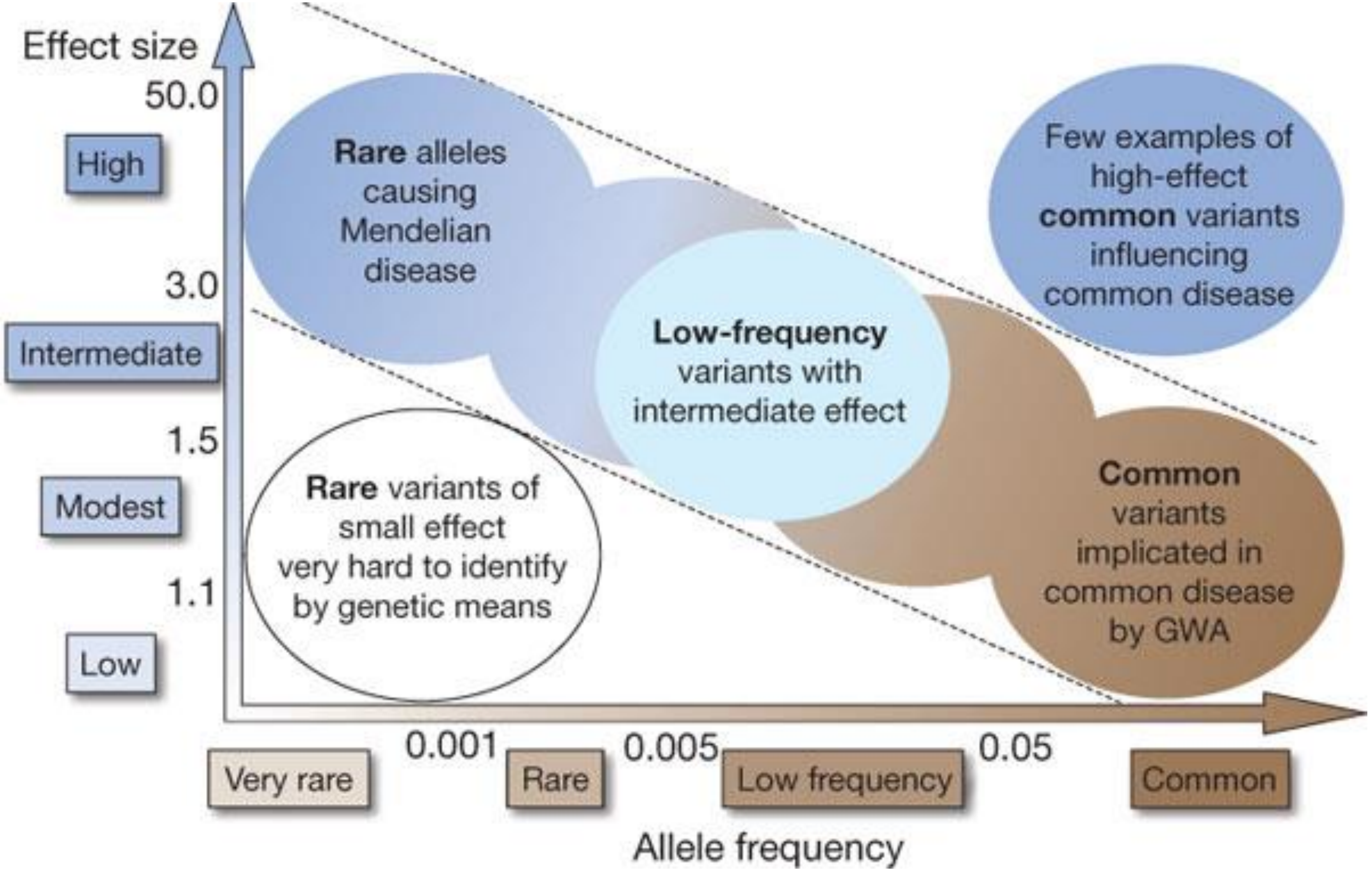
Inversion

Insertion

Copy number variation

Duplication

Variants' Distribution in Population



Effect size
50.0
High
3.0
Intermediate
1.5
Modest
1.1
Low
Rare alleles causing Mendelian disease
Few examples of high-effect **common** variants influencing common disease
Low-frequency variants with intermediate effect
Rare variants of small effect very hard to identify by genetic means
Common variants implicated in common disease by GWA
0.001
0.005
0.05
Very rare
Rare
Low frequency
Common
Allele frequency

Simplest Regression Model of Association

$$Y_i = \alpha + \beta X_i + e_i$$

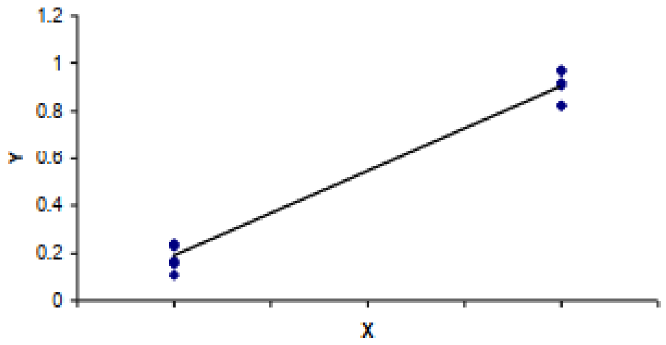
where

$Y_i =$ trait value for individual i

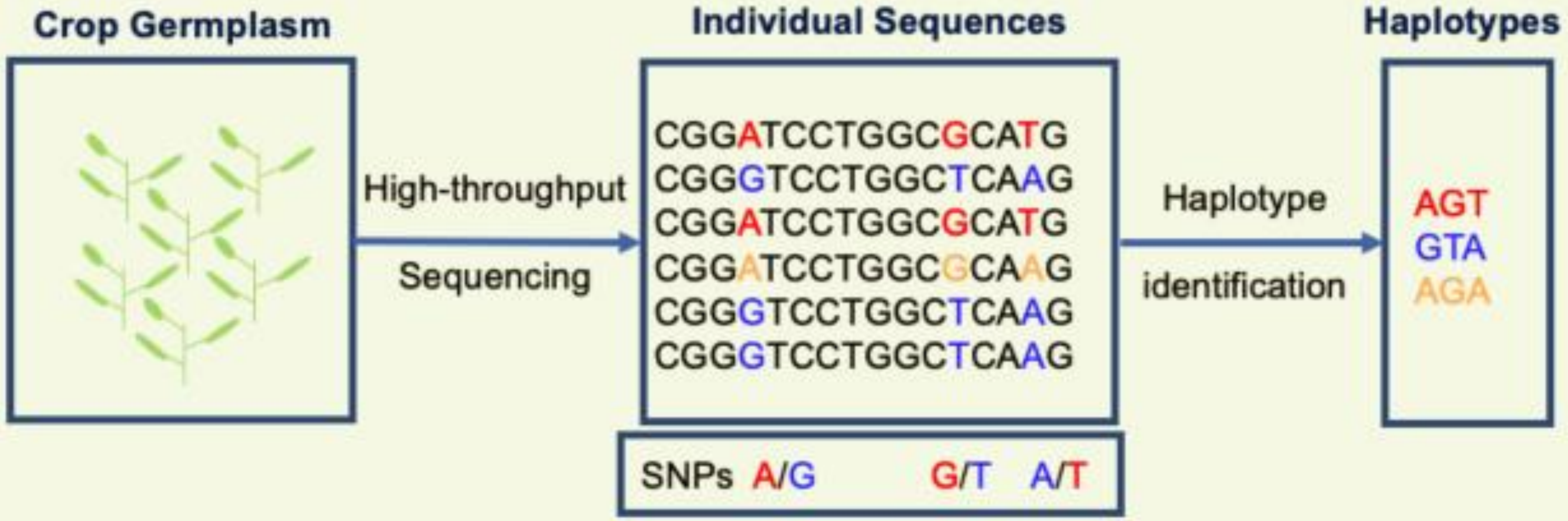
$X_i =$ 1 if allele individual i has allele ‘A’

0 otherwise

i.e., test of mean differences between ‘A’ and ‘not-A’ individuals



Haplotypes



Crop Germplasm

High-throughput

Sequencing

Individual Sequences

CGGATCCTGGCGCATG
CGGGTCCTGGCTCAAG
CGGATCCTGGCGCATG
CGGATCCTGGCGCAAG
CGGGTCCTGGCTCAAG
CGGGTCCTGGCTCAAG

SNPs A/G G/T A/T

Haplotype

identification

Haplotypes

AGT
GTA
AGA

Reference Panels for Genomic Imputation

27 October 2005 | www.nature.com/nature
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
nature
THE HAPMAP PROJECT
Chapter and verse on human genetic variation

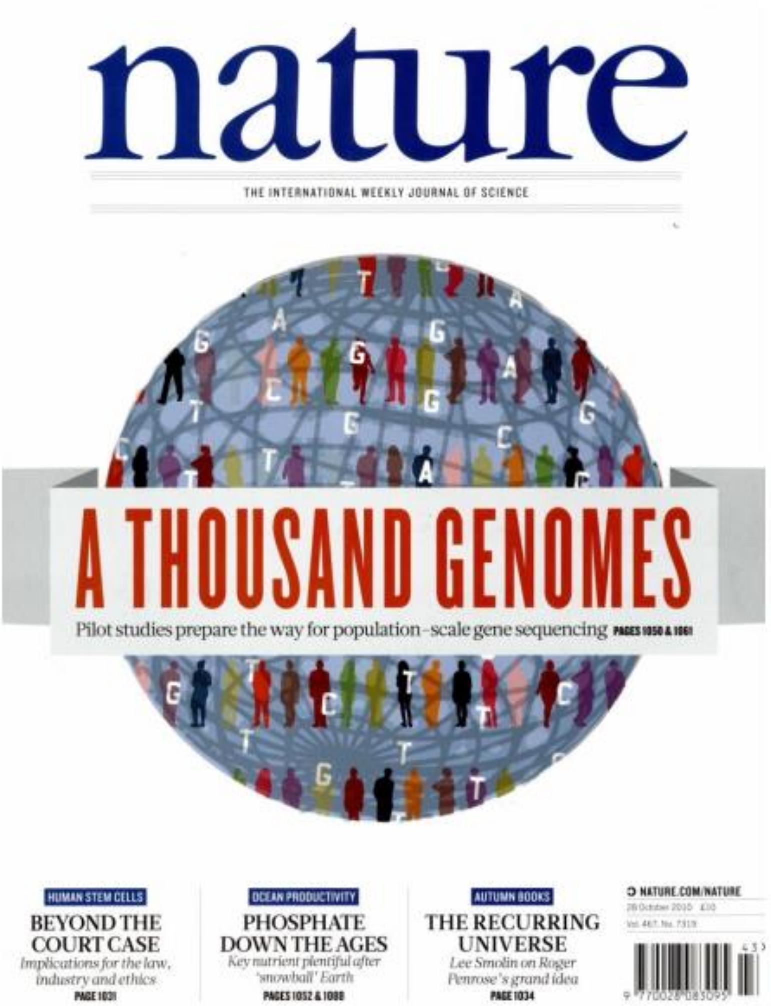
HapMap (haplotype map) Project

270 individuals:

30 parent-offspring trios of the Yoruba from Ibadan, Nigeria (YRI)
30 trios of Utah residents with European ancestry (CEU)
45 individuals from Beijing, China (CHB)
45 individuals from Tokyo, Japan (JPT)

The International HapMap Consortium (2005). A haplotype map of the human genome. *Nature*.

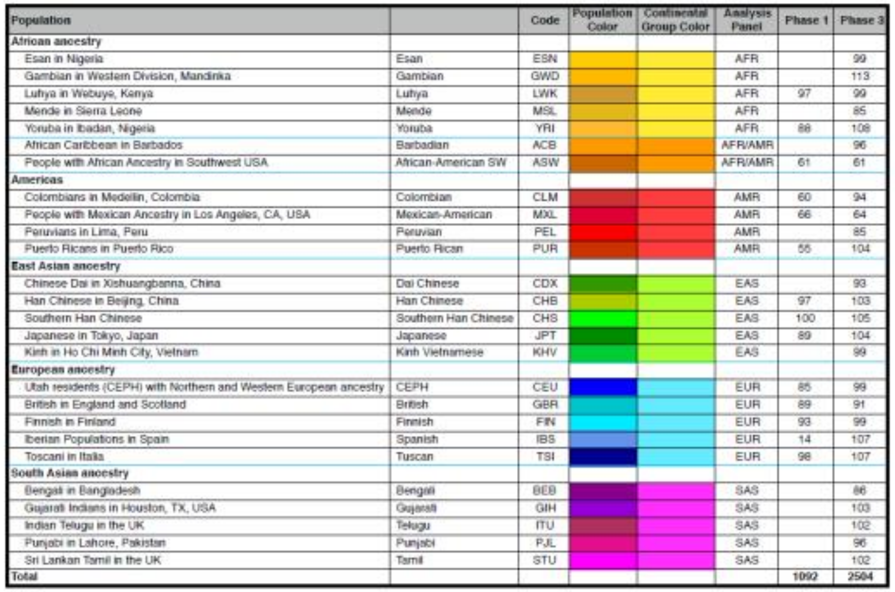
Reference Panels for Genomic Imputation


nature
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
A THOUSAND GENOMES
Pilot studies prepare the way for population-scale gene sequencing PAGES 1050 & 1061
HUMAN STEM CELLS
BEYOND THE COURT CASE
Implications for the law, industry and ethics
PAGE 1031
OCEAN PRODUCTIVITY
PHOSPHATE DOWN THE AGES
Key nutrient plentiful after 'snowball' Earth
PAGES 1052 & 1088
AUTUMN BOOKS
THE RECURRING UNIVERSE
Lee Smolin on Roger Penrose's grand idea
PAGE 1034
NATURE.COM/NATURE
28 October 2010 £10
Vol. 467, No. 7319

1000 Genomes Project

Phase 1: 1,092 individuals from 14 populations..

Phase 3: 2,504 individuals from 26 populations (~500 samples form each 5 continental ancestry groups, with ~5 populations for each group)

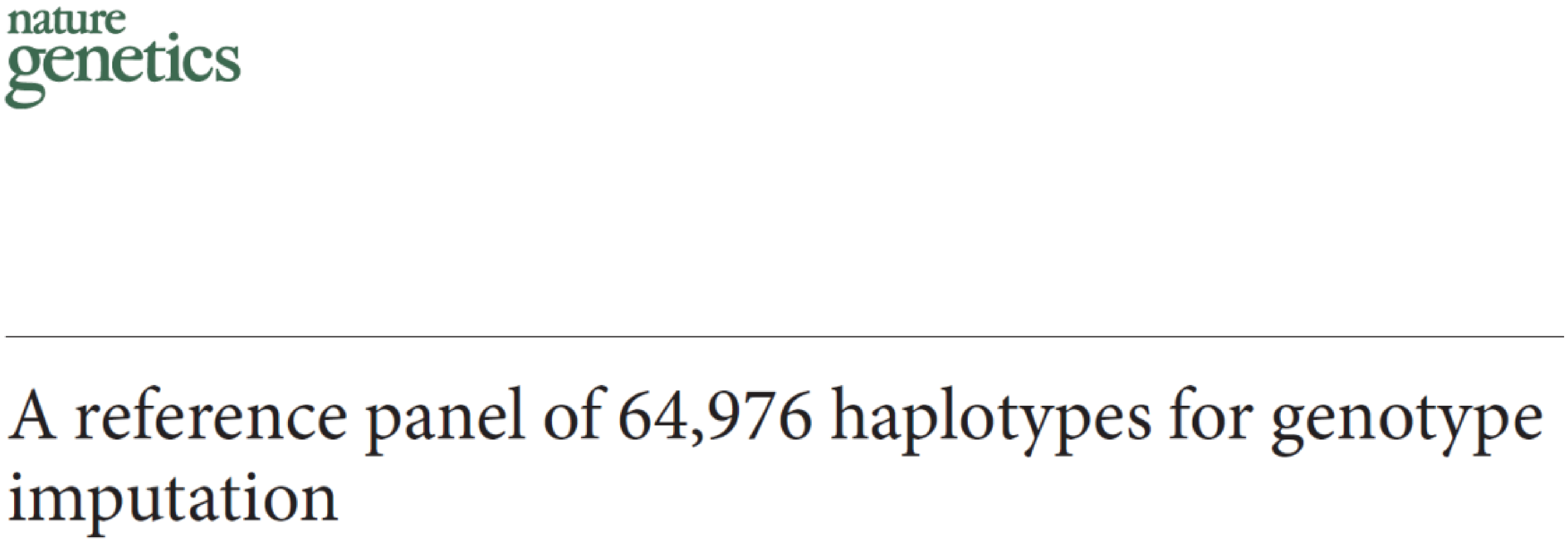


Population		Code	Population Color	Continental Group Color	Analysis Panel	Phase 1	Phase 3
African ancestry							
Esan in Nigeria	Esan	ESN	🟨	🟨	AFR		99
Gambian in Western Division, Mandinka	Gambian	GWD	🟧	🟨	AFR		113
Luhya in Webuye, Kenya	Luhya	LWK	🟫	🟨	AFR	97	99
Mende in Sierra Leone	Mende	MSL	🟨	🟨	AFR		85
Yoruba in Ibadan, Nigeria	Yoruba	YRI	🟧	🟨	AFR	88	108
African Caribbean in Barbados	Barbadian	ACB	🟧	🟧	AFR/AMR		96
People with African Ancestry in Southwest USA	African-American SW	ASW	🟫	🟧	AFR/AMR	61	61
Americas							
Colombians in Medellin, Colombia	Colombian	CLM	🟥	🟥	AMR	60	94
People with Mexican Ancestry in Los Angeles, CA, USA	Mexican-American	MXL	🟥	🟥	AMR	66	64
Peruvians in Lima, Peru	Peruvian	PEL	🟥	🟥	AMR		85
Puerto Ricans in Puerto Rico	Puerto Rican	PUR	🟥	🟥	AMR	55	104
East Asian ancestry							
Chinese Dai in Xishuangbanna, China	Dai Chinese	CDX	🟩	🟩	EAS		93
Han Chinese in Beijing, China	Han Chinese	CHB	🟩	🟩	EAS	97	103
Southern Han Chinese	Southern Han Chinese	CHS	🟩	🟩	EAS	100	105
Japanese in Tokyo, Japan	Japanese	JPT	🟩	🟩	EAS	89	104
Kinh in Ho Chi Minh City, Vietnam	Kinh Vietnamese	KHV	🟩	🟩	EAS		99
European ancestry							
Utah residents (CEPH) with Northern and Western European ancestry	CEPH	CEU	🟦	🟦	EUR	85	99
British in England and Scotland	British	GBR	🟦	🟦	EUR	89	91
Finnish in Finland	Finnish	FIN	🟦	🟦	EUR	93	99
Iberian Populations in Spain	Spanish	IBS	🟦	🟦	EUR	14	107
Toscani in Italia	Tuscan	TSI	🟦	🟦	EUR	98	107
South Asian ancestry							
Bengali in Bangladesh	Bengali	BEB	🟪	🟪	SAS		86
Gujarati Indians in Houston, TX, USA	Gujarati	GIH	🟪	🟪	SAS		103
Indian Telugu in the UK	Telugu	ITU	🟪	🟪	SAS		102
Punjabi in Lahore, Pakistan	Punjabi	PJL	🟪	🟪	SAS		96
Sri Lankan Tamil in the UK	Tamil	STU	🟪	🟪	SAS		102
Total						1092	2504

The 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*
The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature*

Reference Panels for Genomic Imputation

The Haplotype Reference Consortium (HRC)



nature genetics

A reference panel of 64,976 haplotypes for genotype imputation

The Haplotype Reference Consortium (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*.

Reference Panels for Genomic Imputation

Trans-Omics for Precision Medicine (TOPMed)

nature

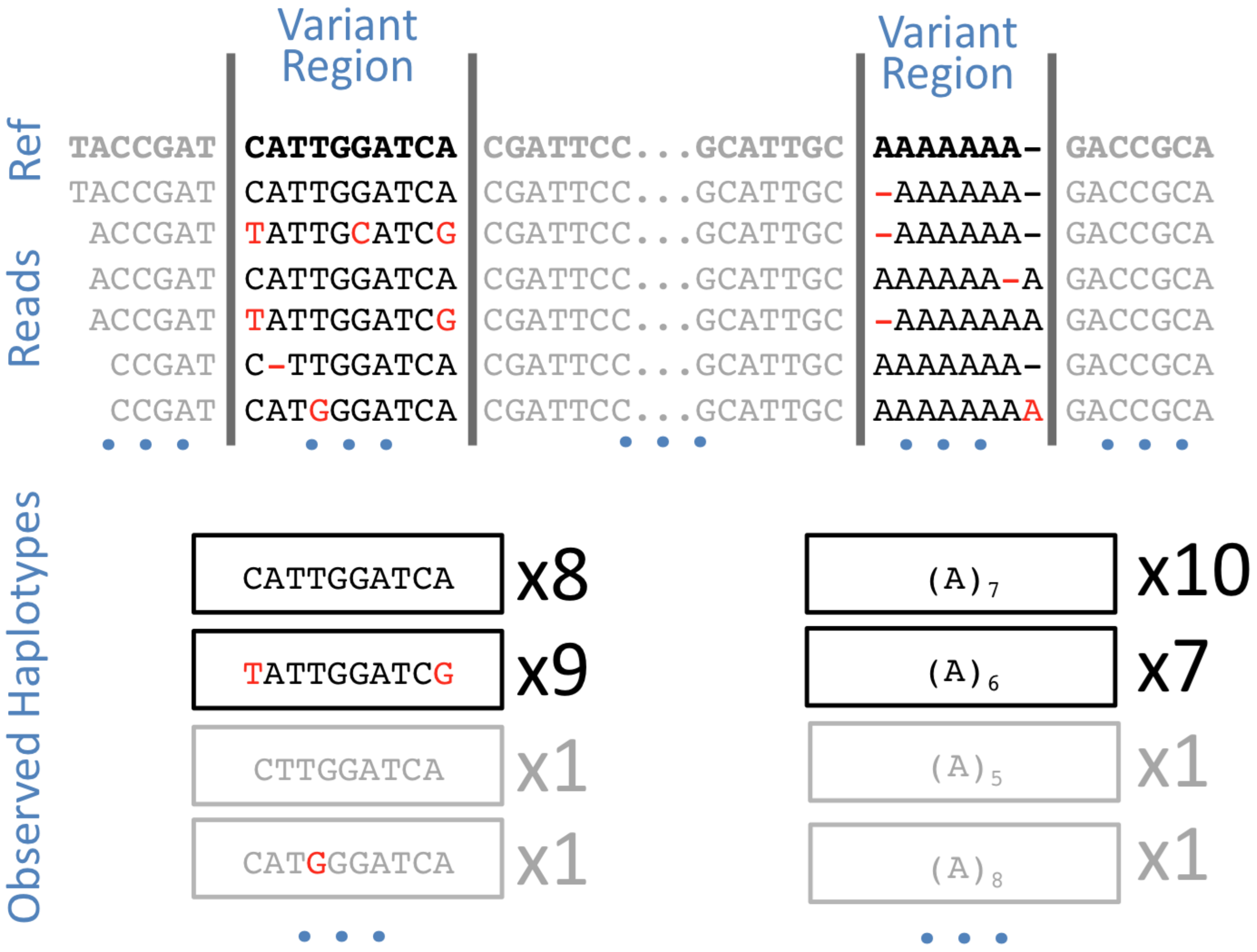
Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program

Taliun, D., Harris, D.N., Kessler, M.D. *et al.* (2021). Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature*

Step1: Variant Identification (sequencing)

Software:

- freeBayes
- GATK



		Variant Region		Variant Region	
Ref	TACCGAT	CATTGGATCA	CGATTCC...GCATTGC	AAAAAAA-	GACCGCA
Reads	TACCGAT	CATTGGATCA	CGATTCC...GCATTGC	-AAAAAA-	GACCGCA
	ACCGAT	TATTGCATCG	CGATTCC...GCATTGC	-AAAAAA-	GACCGCA
	ACCGAT	CATTGGATCA	CGATTCC...GCATTGC	AAAAAA-A	GACCGCA
	ACCGAT	TATTGGATCG	CGATTCC...GCATTGC	-AAAAAAA	GACCGCA
	CCGAT	C-TTGGATCA	CGATTCC...GCATTGC	AAAAAAA-	GACCGCA
	CCGAT	CATGGGATCA	CGATTCC...GCATTGC	AAAAAAAA	GACCGCA
	...	...	...	...	...

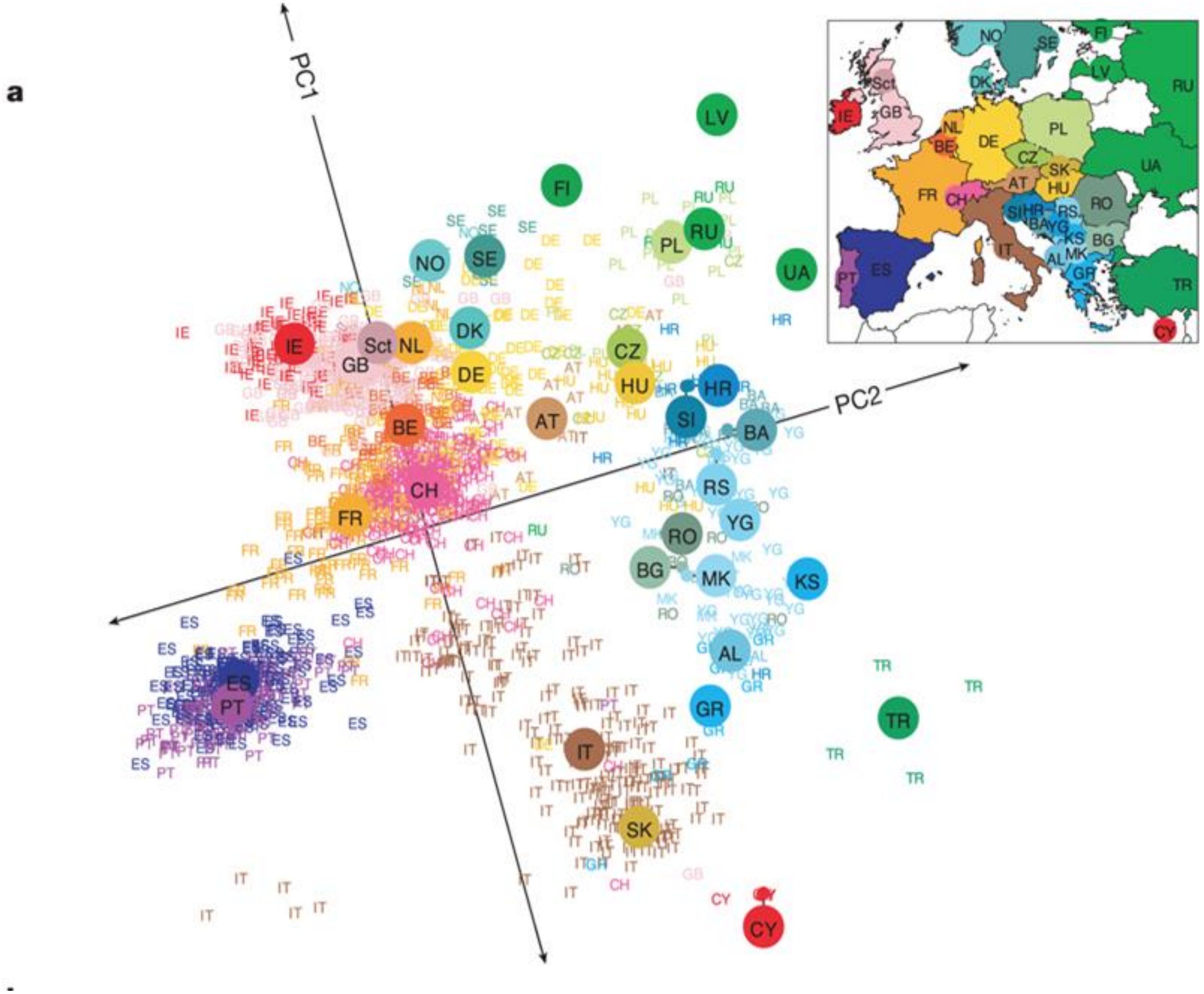
Observed Haplotypes			
CATTGGATCA	x8	$(A)_7$	x10
TATTGGATCG	x9	$(A)_6$	x7
CTTGGATCA	x1	$(A)_5$	x1
CATGGGATCA	x1	$(A)_8$	x1
...		...	

https://github.com/ekg/freebayes

Step2: Population Stratification

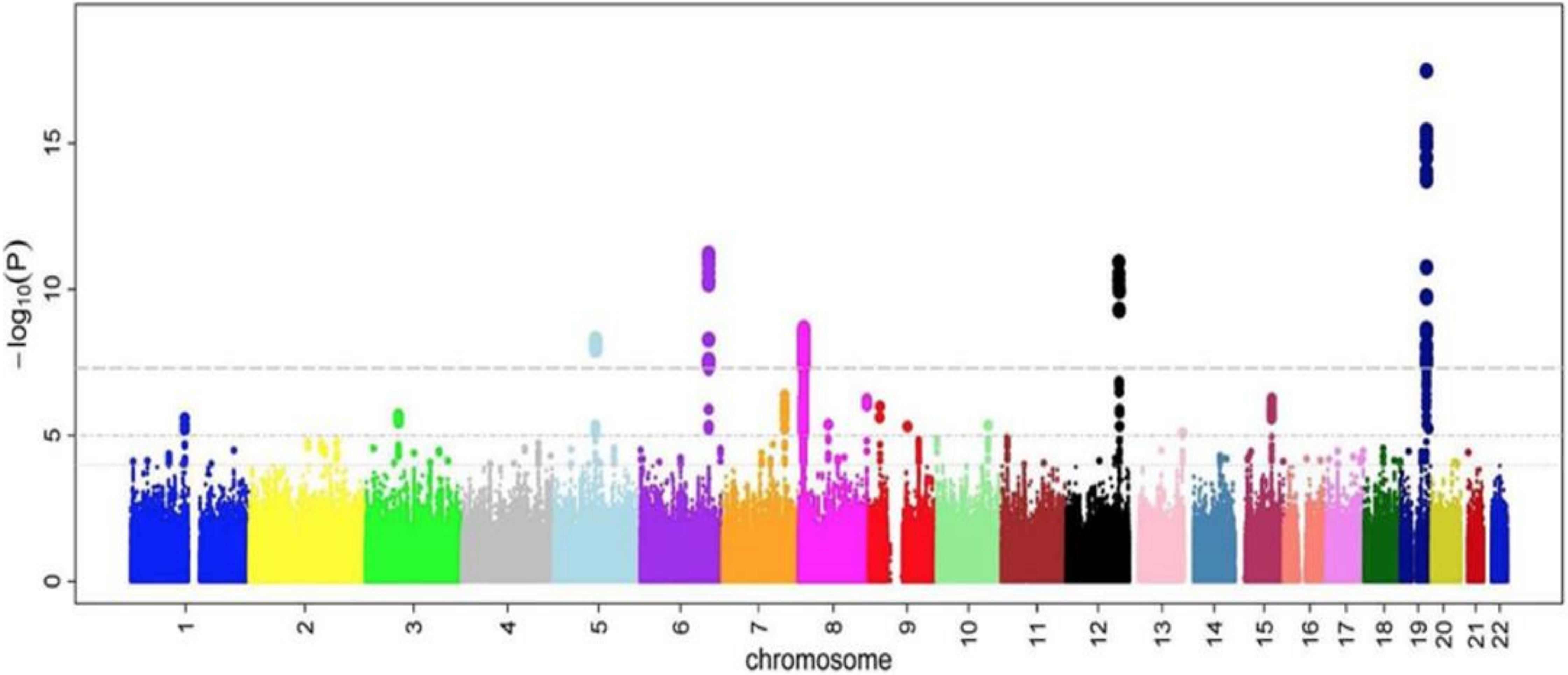
Software:

- EIGENSOFT
- snpStats



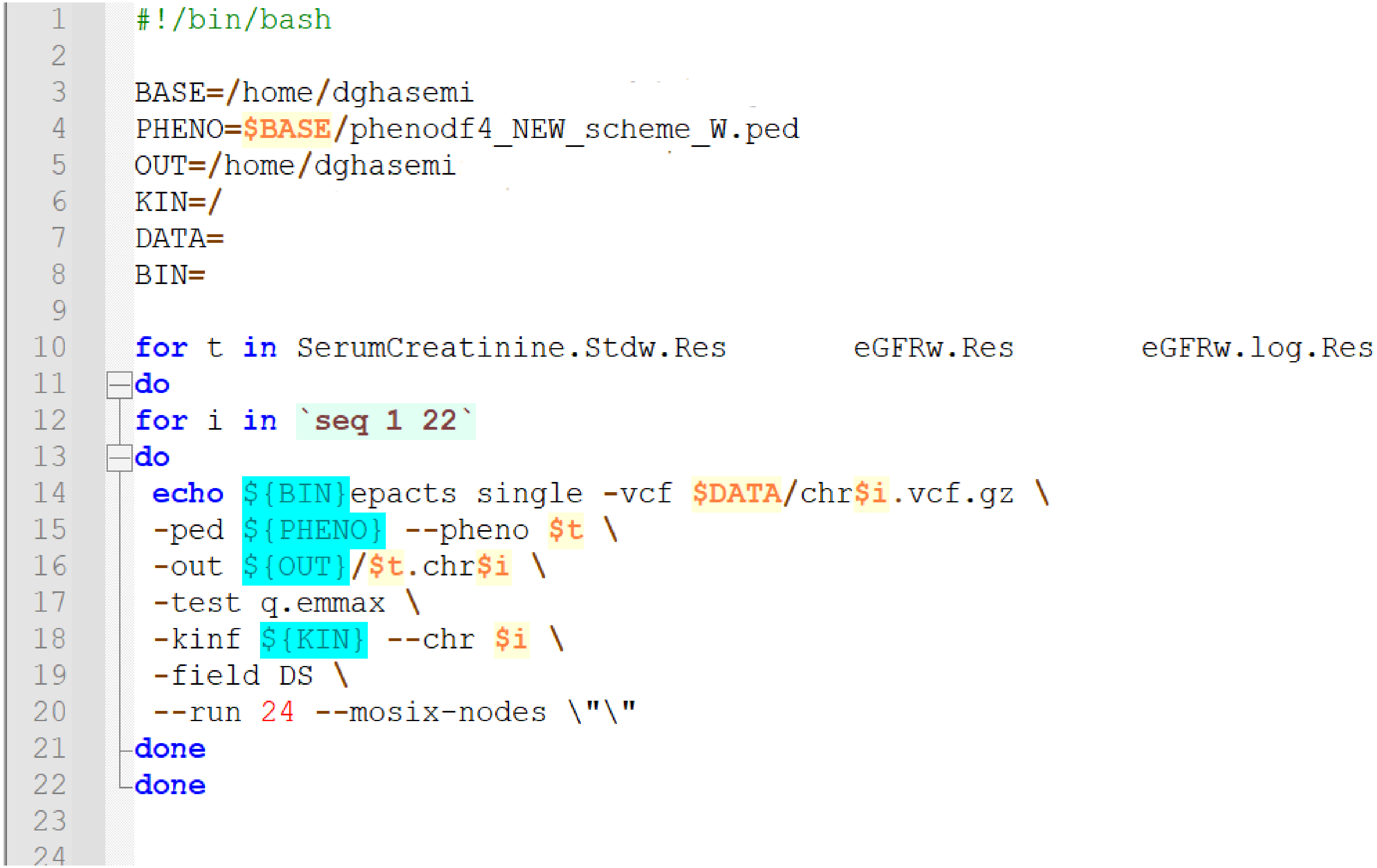
http://www.nature.com/nature/journal/v456/n7218/full/nature07331.html

Step3: Statistical Tests



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

Step3: Statistical Tests | Performing GWAS via EPACTS

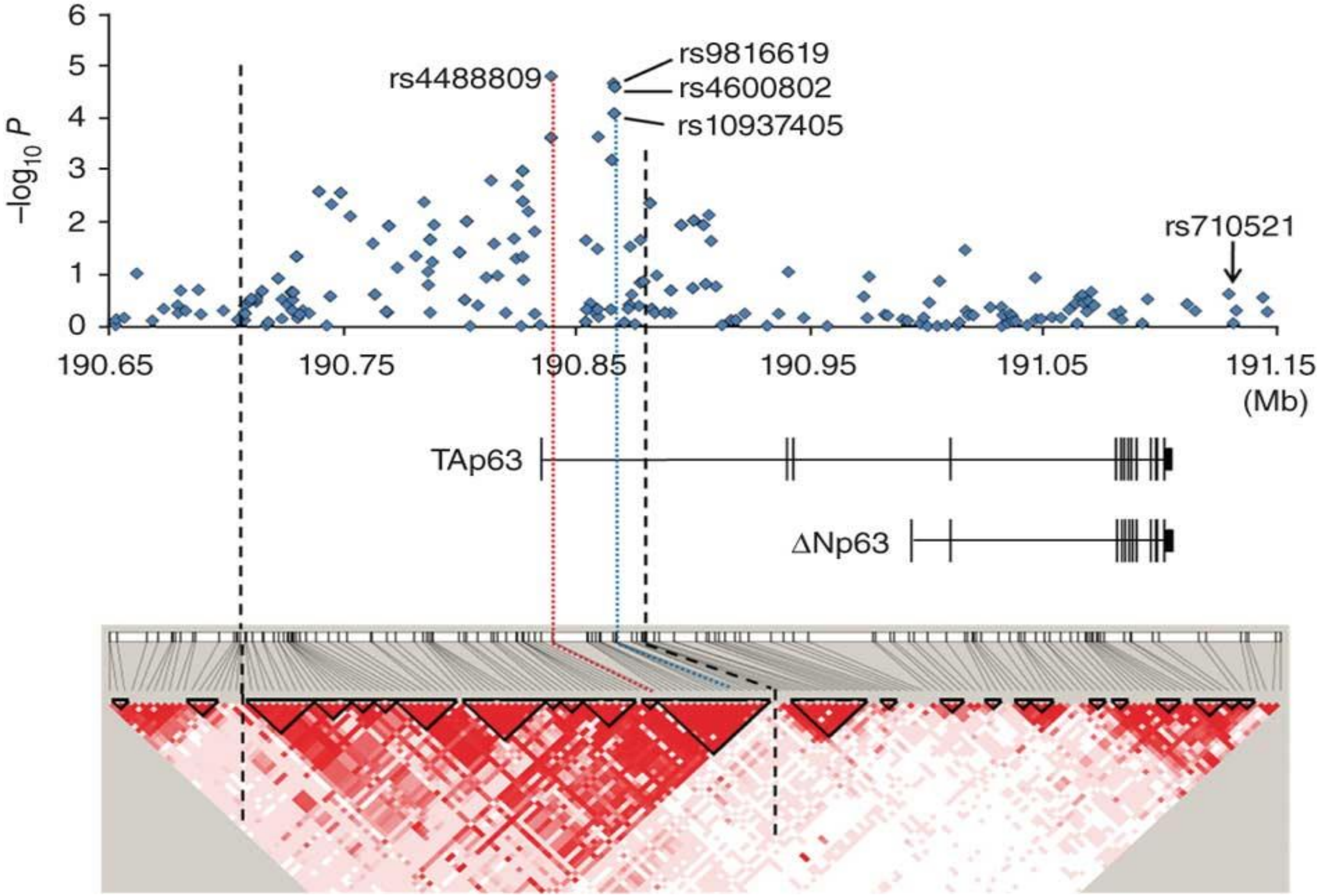


```
#!/bin/bash

BASE=/home/dghasemi
PHENO=$BASE/phenodf4_NEW_scheme_W.ped
OUT=/home/dghasemi
KIN=/
DATA=
BIN=

for t in SerumCreatinine.Stdw.Res       eGFRw.Res       eGFRw.log.Res
do
for i in `seq 1 22`
do
 echo ${BIN}epacts single -vcf $DATA/chr$i.vcf.gz \
 -ped ${PHENO} --pheno $t \
 -out ${OUT}/$t.chr$i \
 -test q.emmax \
 -kinf ${KIN} --chr $i \
 -field DS \
 --run 24 --mosix-nodes \"\"
done
done


```

Step4: Examine local region

Software:

- PLINK
- Annotating Genomic Variants Workflow



http://www.nature.com/ng/journal/v42/n10/full/ng.667.html

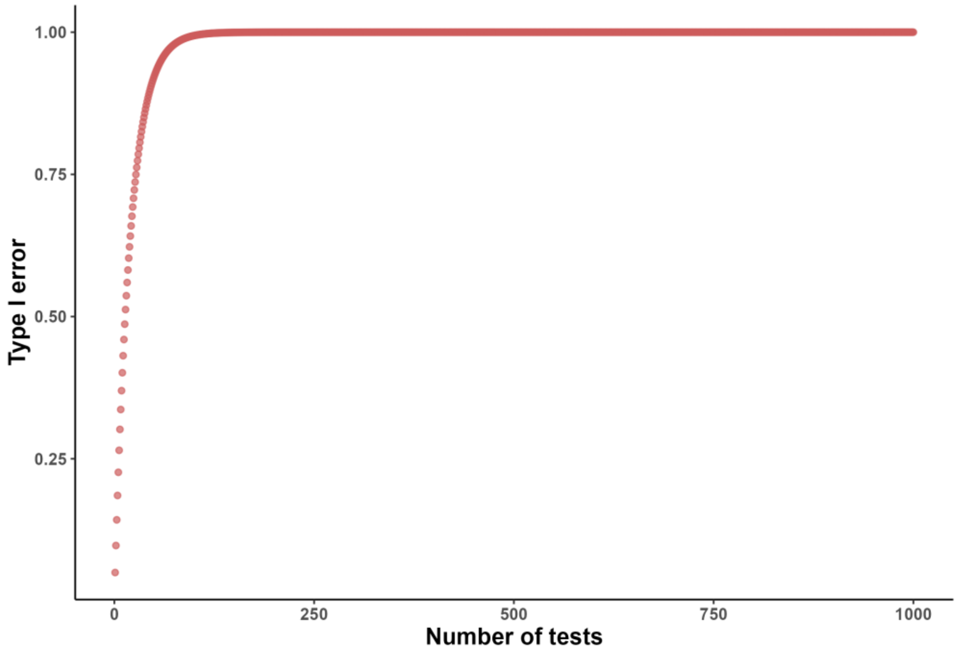
Multiple Testing Adjustments

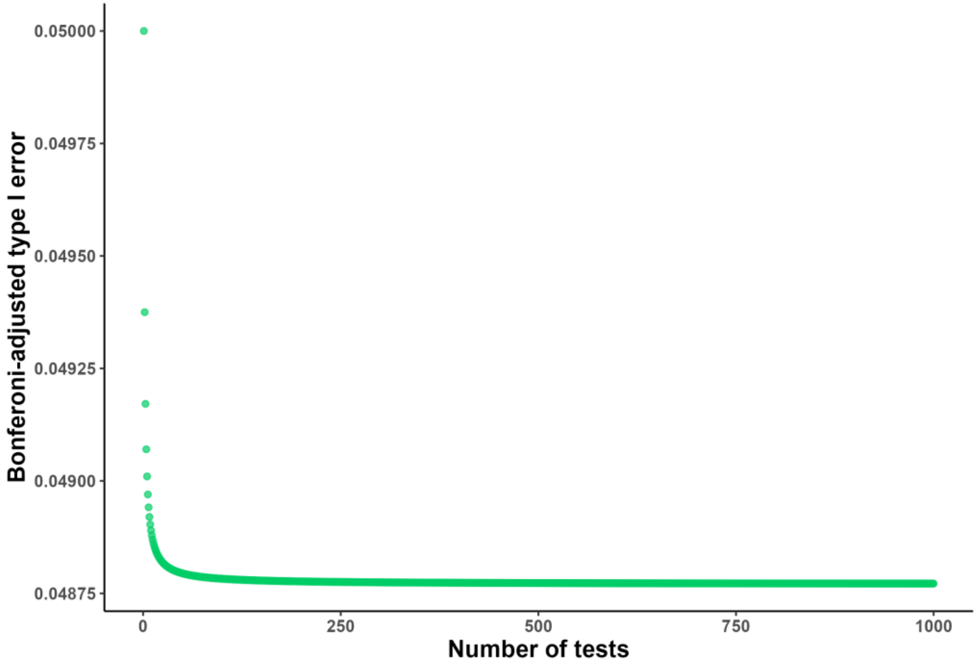
- Measure 10,000 genes
- Calculate 10,000 p-values
- Call genes “significant” if p-value < 0.05
- Expected Number of False Positives:

  10,000 × 0.05 = 500 False Positives

- Bonferroni Correction:

  P-values less than α/m are significant





PhD Project: Kidney function Biomarkers in CHRIS Study Participants

- Database: CHRIS study

Cooperative Health Research In South Tyrol study

. 13 municipalities
. 28,000 inhabitants (adults)
. 13,393 participants (2011-2018)
. Mean age 46 (18-93)
. 55% females

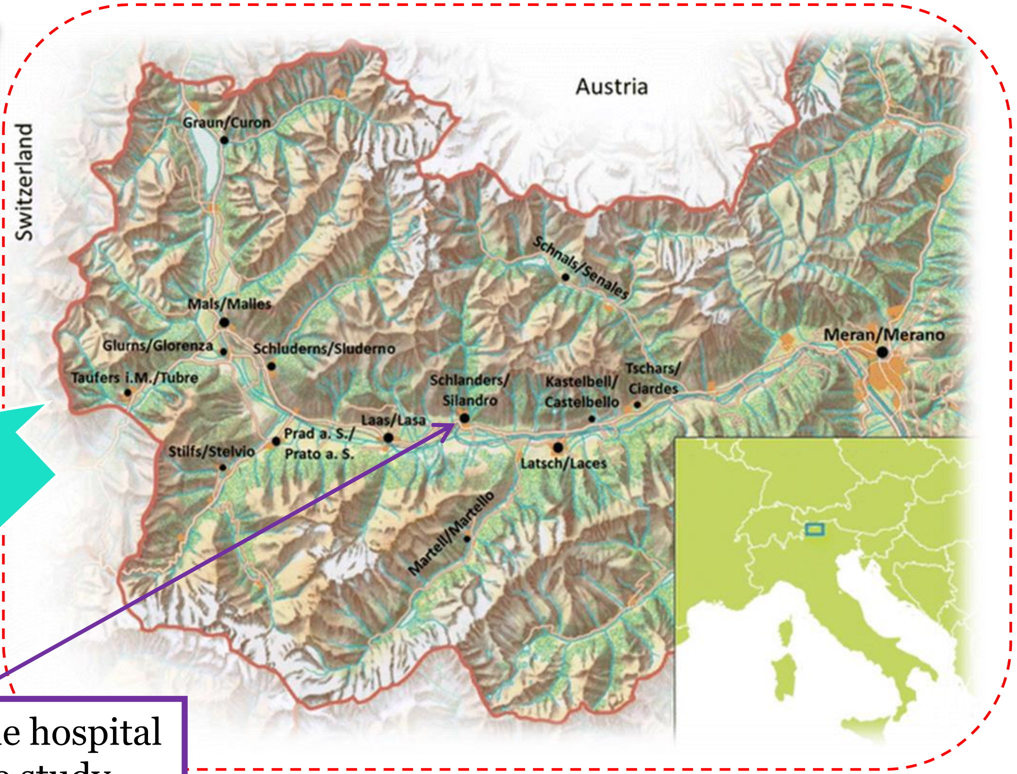
N= 10,500

10,500 GWAS

3,600 WES

Integrated via genotype imputation

CHRIS
eurac research
Südtiroler Gesundheitsstudie
Studio sulla salute in Alto Adige

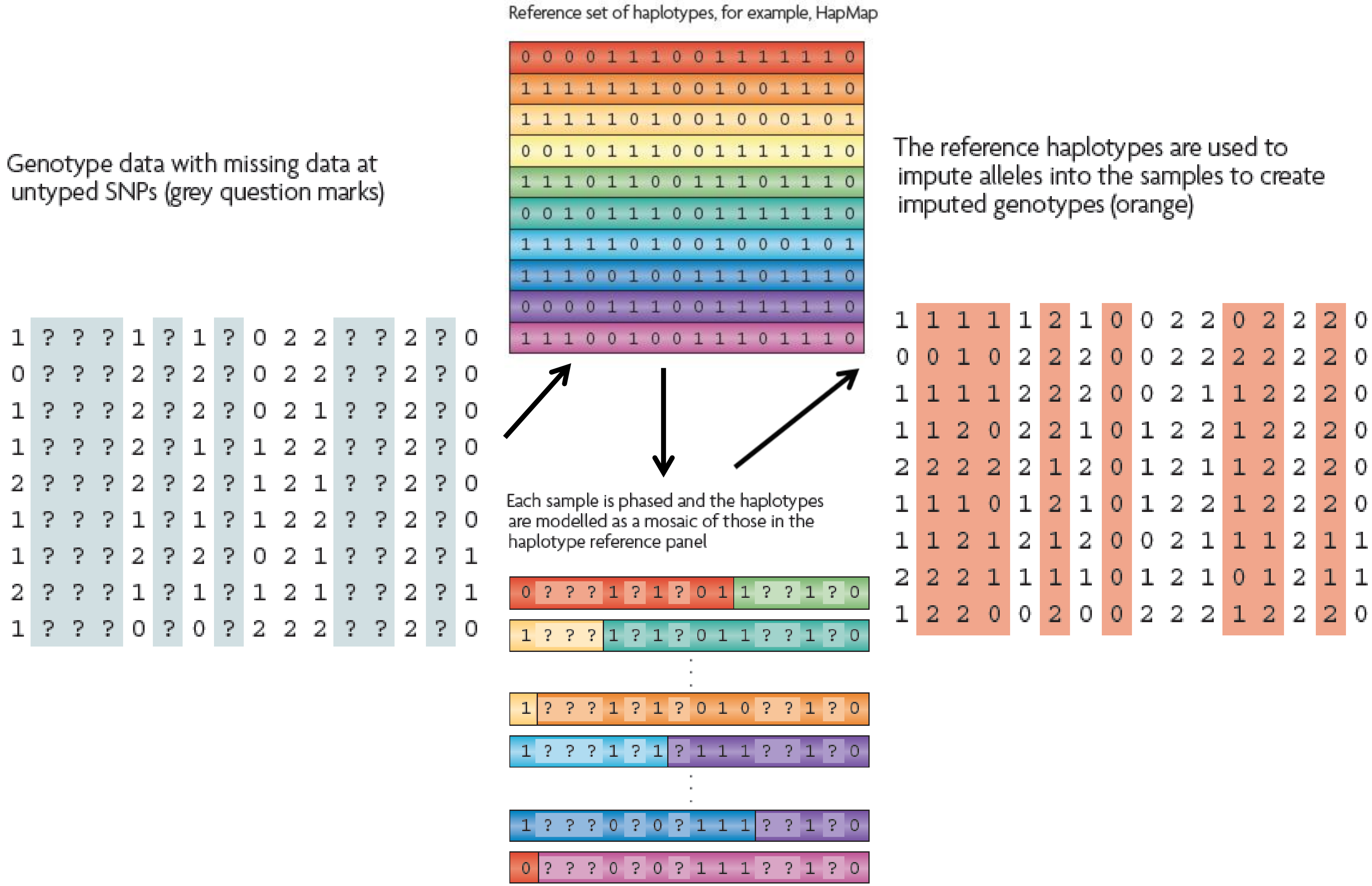


Schlanders: one hospital
A single site study

South Tyrol, Italy

Phenotypes: Serum Creatinine, Urinary Creatinine, UACR, Serum Albumin, Urinary Albumin, eGFR

What is imputation? (Marchini & Howie 2010)



Genotype data with missing data at untyped SNPs (grey question marks)

```
1 ? ? ? 1 ? 1 ? 0 2 2 ? ? 2 ? 0
0 ? ? ? 2 ? 2 ? 0 2 2 ? ? 2 ? 0
1 ? ? ? 2 ? 2 ? 0 2 1 ? ? 2 ? 0
1 ? ? ? 2 ? 1 ? 1 2 2 ? ? 2 ? 0
2 ? ? ? 2 ? 2 ? 1 2 1 ? ? 2 ? 0
1 ? ? ? 1 ? 1 ? 1 2 2 ? ? 2 ? 0
1 ? ? ? 2 ? 2 ? 0 2 1 ? ? 2 ? 1
2 ? ? ? 1 ? 1 ? 1 2 1 ? ? 2 ? 1
1 ? ? ? 0 ? 0 ? 2 2 2 ? ? 2 ? 0
```

Reference set of haplotypes, for example, HapMap

```
0 0 0 0 1 1 1 0 0 1 1 1 1 1 1 0
1 1 1 1 1 1 1 0 0 1 0 0 1 1 1 0
1 1 1 1 1 0 1 0 0 1 0 0 0 1 0 1
0 0 1 0 1 1 1 0 0 1 1 1 1 1 1 0
1 1 1 0 1 1 0 0 1 1 1 0 1 1 1 0
0 0 1 0 1 1 1 0 0 1 1 1 1 1 1 0
1 1 1 1 1 0 1 0 0 1 0 0 0 1 0 1
1 1 1 0 0 1 0 0 1 1 1 0 1 1 1 0
0 0 0 0 1 1 1 0 0 1 1 1 1 1 1 0
1 1 1 0 0 1 0 0 1 1 1 0 1 1 1 0
```

Each sample is phased and the haplotypes are modelled as a mosaic of those in the haplotype reference panel

```
0 ? ? ? 1 ? 1 ? 0 1 1 ? ? 1 ? 0
1 ? ? ? 1 ? 1 ? 0 1 1 ? ? 1 ? 0
                  ⋮
1 ? ? ? 1 ? 1 ? 0 1 0 ? ? 1 ? 0
1 ? ? ? 1 ? 1 ? 1 1 1 ? ? 1 ? 0
                  ⋮
1 ? ? ? 0 ? 0 ? 1 1 1 ? ? 1 ? 0
0 ? ? ? 0 ? 0 ? 1 1 1 ? ? 1 ? 0
```

The reference haplotypes are used to impute alleles into the samples to create imputed genotypes (orange)

```
1 1 1 1 1 2 1 0 0 2 2 0 2 2 2 0
0 0 1 0 2 2 2 0 0 2 2 2 2 2 2 0
1 1 1 1 2 2 2 0 0 2 1 1 2 2 2 0
1 1 2 0 2 2 1 0 1 2 2 1 2 2 2 0
2 2 2 2 2 1 2 0 1 2 1 1 2 2 2 0
1 1 1 0 1 2 1 0 1 2 2 1 2 2 2 0
1 1 2 1 2 1 2 0 0 2 1 1 1 2 1 1
2 2 2 1 1 1 1 0 1 2 1 0 1 2 1 1
1 2 2 0 0 2 0 0 2 2 2 1 2 2 2 0
```

Genomic Imputation:
1- Starting Data

Genotyped sample

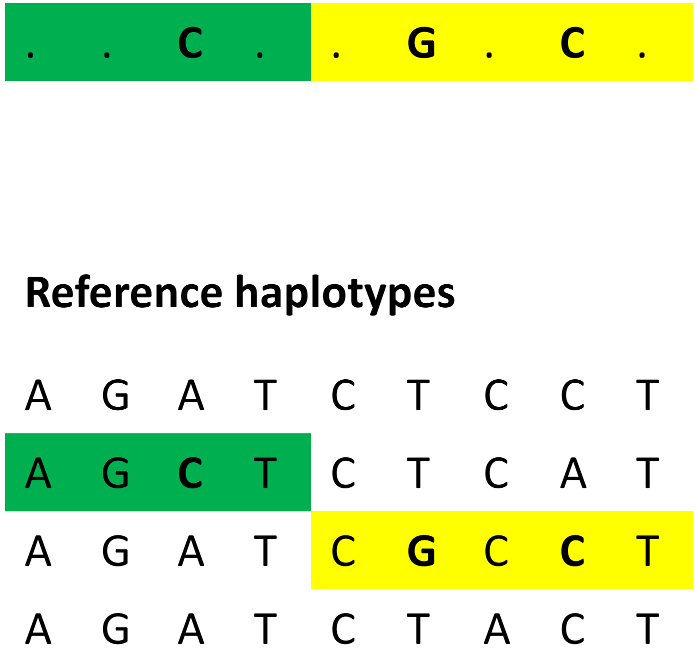
.	.	**C**	.	.	**G**	.	**C**	.

Reference haplotypes

A	G	A	T	C	T	C	C	T
A	G	C	T	C	T	C	A	T
A	G	A	T	C	G	C	C	T
A	G	A	T	C	T	A	C	T

Genomic Imputation:

2- Identify shared regions of chromosome



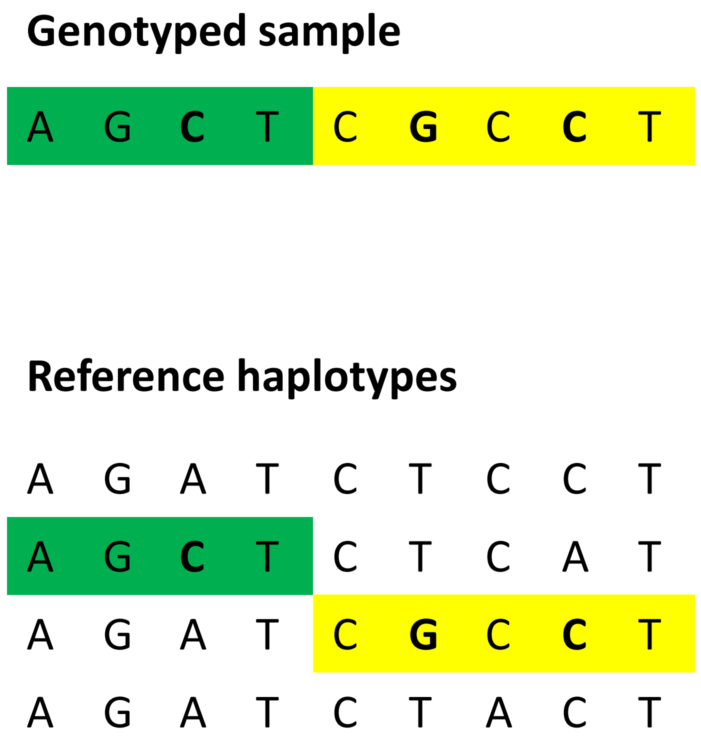
Genotyped sample

.	.	**C**	.	.	**G**	.	**C**	.

Reference haplotypes

A	G	A	T	C	T	C	C	T
A	G	**C**	T	C	T	C	A	T
A	G	A	T	C	**G**	C	**C**	T
A	G	A	T	C	T	A	C	T

Genomic Imputation: 3.
Fill in missing genotypes



Genotyped sample

A G **C** T C **G** C **C** T

Reference haplotypes

A G A T C T C C T

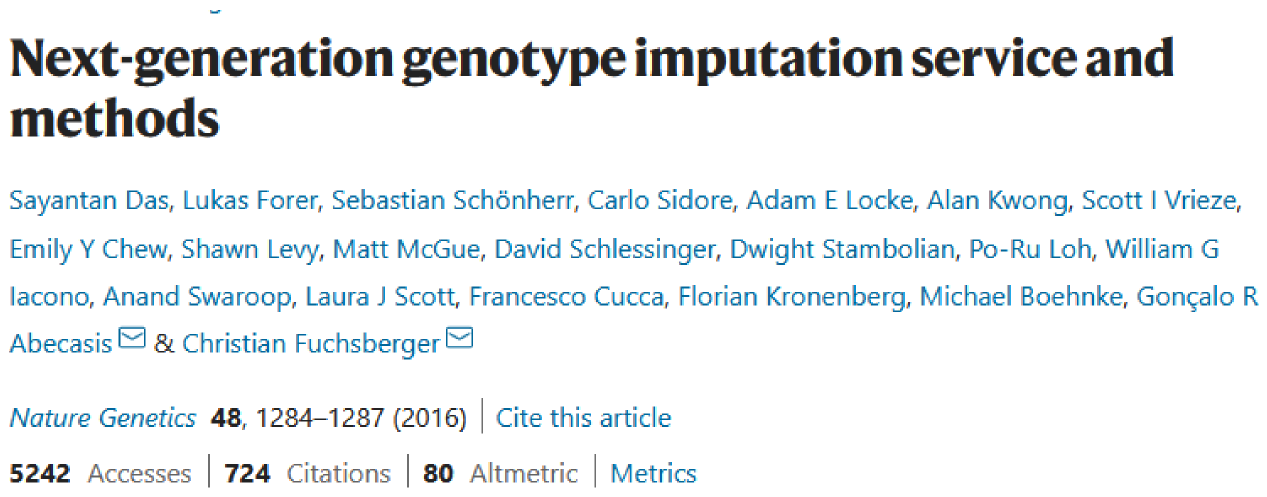
A G **C** T C T C A T

A G A T C **G** C **C** T

A G A T C T A C T

Genomic Imputation:
Minimac4

- https://github.com/statgen/Minimac4
- Building on the work from Gonçalo Abecasis, Christian Fuchsberger and colleagues
- Analysis options
  - SAIGE
  - BoltLMM
  - plink2



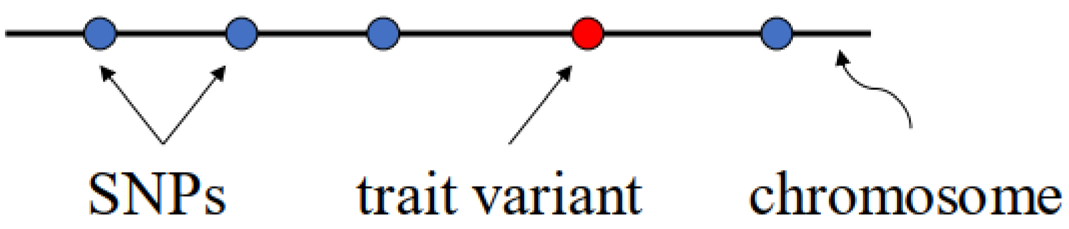
Next-generation genotype imputation service and methods

Sayantan Das, Lukas Forer, Sebastian Schönherr, Carlo Sidore, Adam E Locke, Alan Kwong, Scott I Vrieze, Emily Y Chew, Shawn Levy, Matt McGue, David Schlessinger, Dwight Stambolian, Po-Ru Loh, William G Iacono, Anand Swaroop, Laura J Scott, Francesco Cucca, Florian Kronenberg, Michael Boehnke, Gonçalo R Abecasis ✉ & Christian Fuchsberger ✉

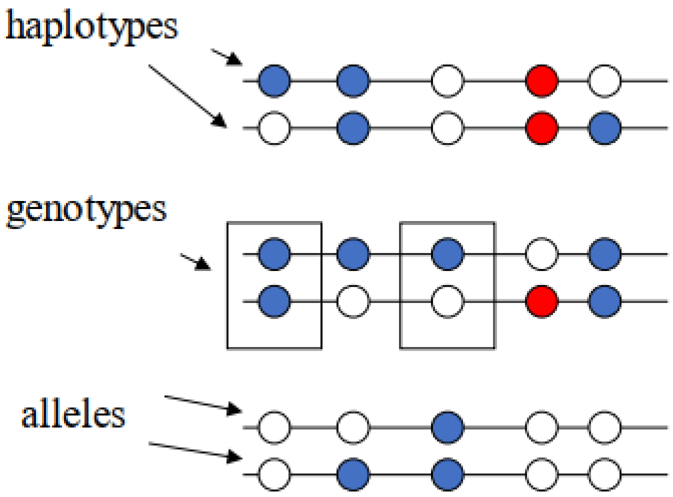
Nature Genetics **48**, 1284–1287 (2016) | Cite this article

5242 Accesses | **724** Citations | **80** Altmetric | Metrics

Definitions

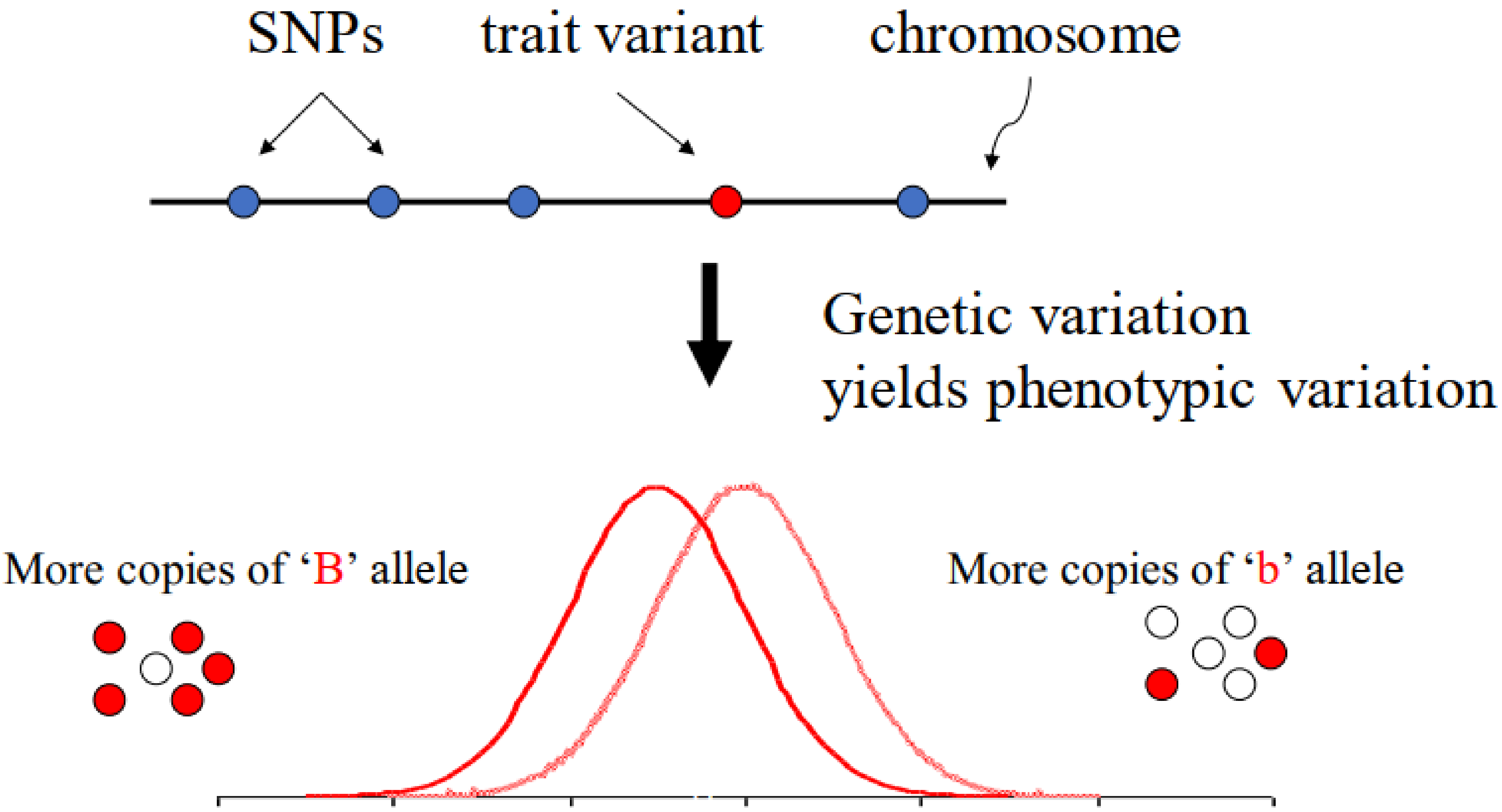


Population Data



Affection	$\text{Trait}_1 \ldots \text{Trait}_n$	
A	10.3	75.66
A	9.9	-99
U	15.8	101.22

Allelic Associations


SNPs
trait variant
chromosome
Genetic variation
yields phenotypic variation
More copies of 'B' allele
More copies of 'b' allele

Biostatistics & Epidemiology

Cristian Pattaro
Group Leader

-Biostatistics
-Epidemiology
-Genetics

Luisa Foco
Senior Researcher

-Statistical genetics
-Genetic epidemiology
-Biomedical statistics

Roberto Melotti
Senior Researcher

-Epidemiology
-Biomedical statistics

Fabiola Del Greco M.
Senior Researcher

-Causal inference
-Statistical genetics
-Epidemiology

Ryosuke Fujii
Post Doc Researcher

-Epidemiology
-Genetics

Martin Gögele
Researcher

-Biodemography
-Epidemiology
-Study design

Giulia Barbieri
PhD Student

-Epidemiology
-Biomedical statistics

Maeregu W. Arisido
Post Doc Researcher

-Biostatistics
-Statistical methods

Rebecca Lundin
Senior Researcher

-Public Health
-Epidemiology

Daniele Bottigliengo
Post Doc Researcher

-Causal inference
-Biostatistics

Dariush Ghasemi S.
PhD Student

-Biostatistics
-Machine Learning
-Statistical genetics

Laura Barin
Post Doc Researcher

-Epidemiology
-Public Health
-Biostatistics

Daniele Giardiello
PhD Student

-Biostatistics
-Cancer epidemiology

Thank You for the Attention