

# The 100 000 Genomes Project Relevant for Transfusion Experts?

Willem H Ouwehand FMedSci

Professor of Experimental Haematology

Director NIHR BioResource – Rare Diseases

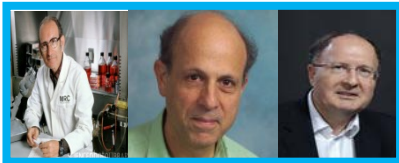
# Did Cambridge discoveries

*change blood transfusion practice and the safety of the blood supply?*



Francis Crick, Ronald Fisher, Robin Coombs

Double helix, Rh blood groups, Crossmatch for safer transfusion



Cesar Milstein, Michael Neuberger, Greg Winter

Safer blood by better ABO & RhD matching/ ~50% of new drugs



Fred Sanger, John Sulston, Shankar Balasubramanian

From slow to fast reading of the DNA code

Better matching for blood & platelets and organs & stem cells

# What is genomics doing for patients

1. 100 000 Genomes Project

2. ThromboGenomics Test

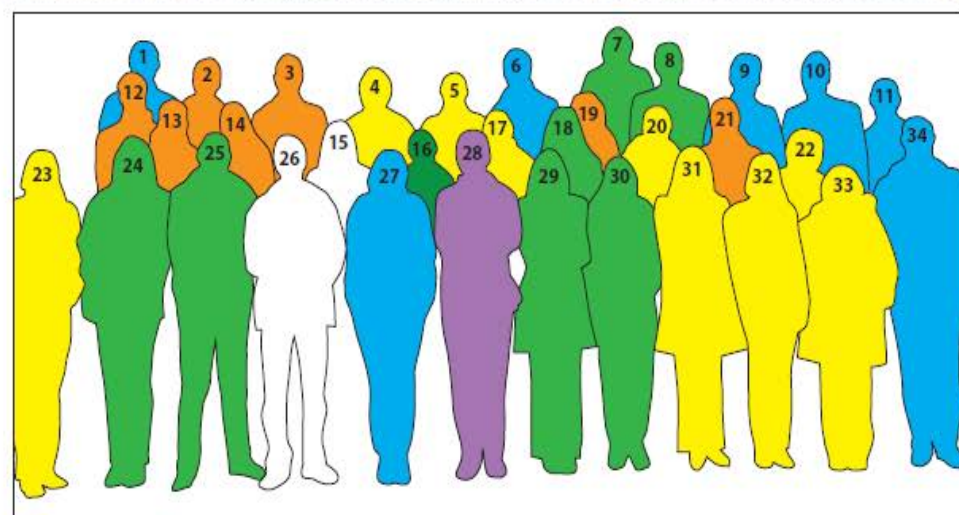
3. Cheap Genotyping Test for Donors





- |                   |                          |                      |
|-------------------|--------------------------|----------------------|
| 1 AUGUSTO RENDON  | 13 JOSE GUERRERO         | 25 DANIEL HAMPSHIRE  |
| 2 PETER SMETHURST | 14 ANNE KELLY            | 26 CEDRIC GHEVAERT   |
| 3 MARCEL REICHEN  | 15 REBECCA CARDIGAN      | 27 TIPHAINE MARTIN   |
| 4 TONY ATTWOOD    | 16 FRANCES BURDEN        | 28 WILLEM OUWEHAND   |
| 5 YAGNESH UMRANIA | 17 JENNIFER SAMBROOK     | 29 FIZZAH CHOUDRY    |
| 6 JOHN ORD        | 18 MYRTO KOSTADIMA       | 30 SAMANTHA FARROW   |
| 7 KATE DOWNES     | 19 AMANDA EVANS          | 31 DEBORAH WHITEHORN |
| 8 MATTIA FRONTINI | 20 ABI CRISP-HIHN        | 32 SOPHIA COE        |
| 9 ERNEST TURRO    | 21 STEVE GARNER          | 33 JENNY JOLLEY      |
| 10 STEFAN GRAF    | 22 RUTENDO MAPETA        | 34 STUART MEACHAM    |
| 11 GRAHAM KIDDLE  | 23 NICOLA FOAD           |                      |
| 12 SJOERT JANSEN  | 24 EWA BIECZYK-MACZYNSKA |                      |

BLUE GROUP: Clinical Bioinformatics & Statistical Genetics  
 ORANGE GROUP: Platelet Biology & Genomics  
 GREEN: Epigenomics of Blood Cells (BluePrint)  
 YELLOW: BioResources (Cambridge BioResource, NIHR BioResource, INTERVAL)  
 WHITE: Visitors



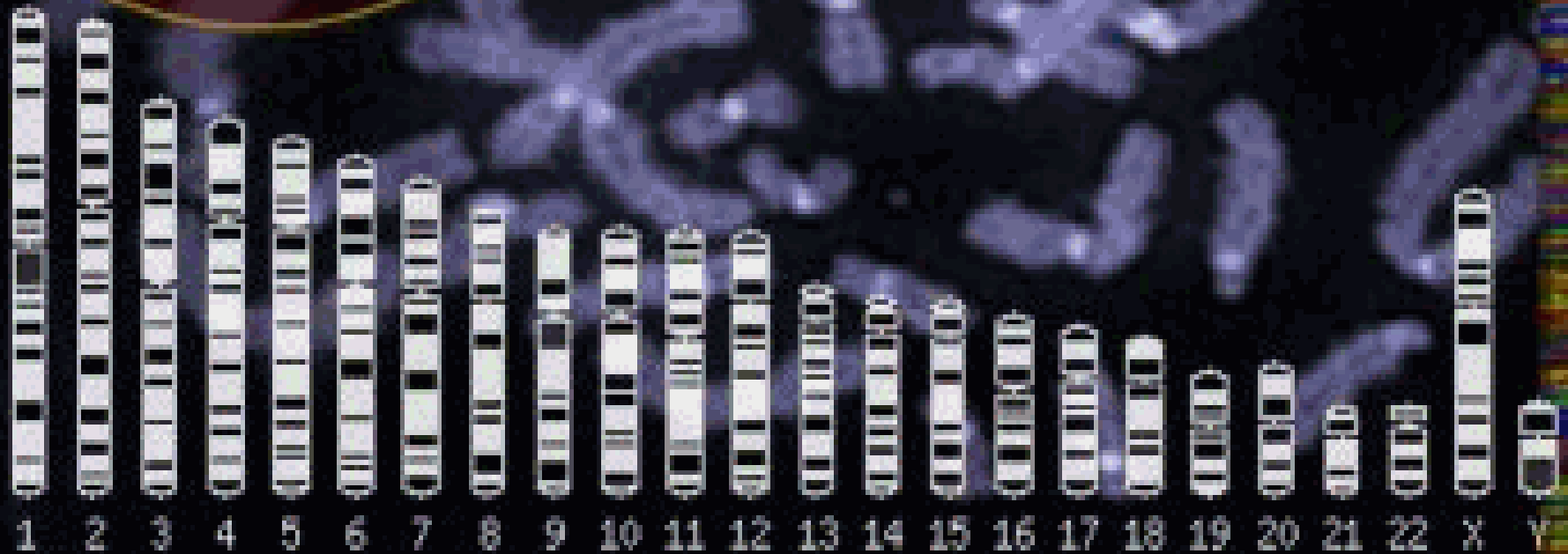


# Human Genome Project

## Ensembl

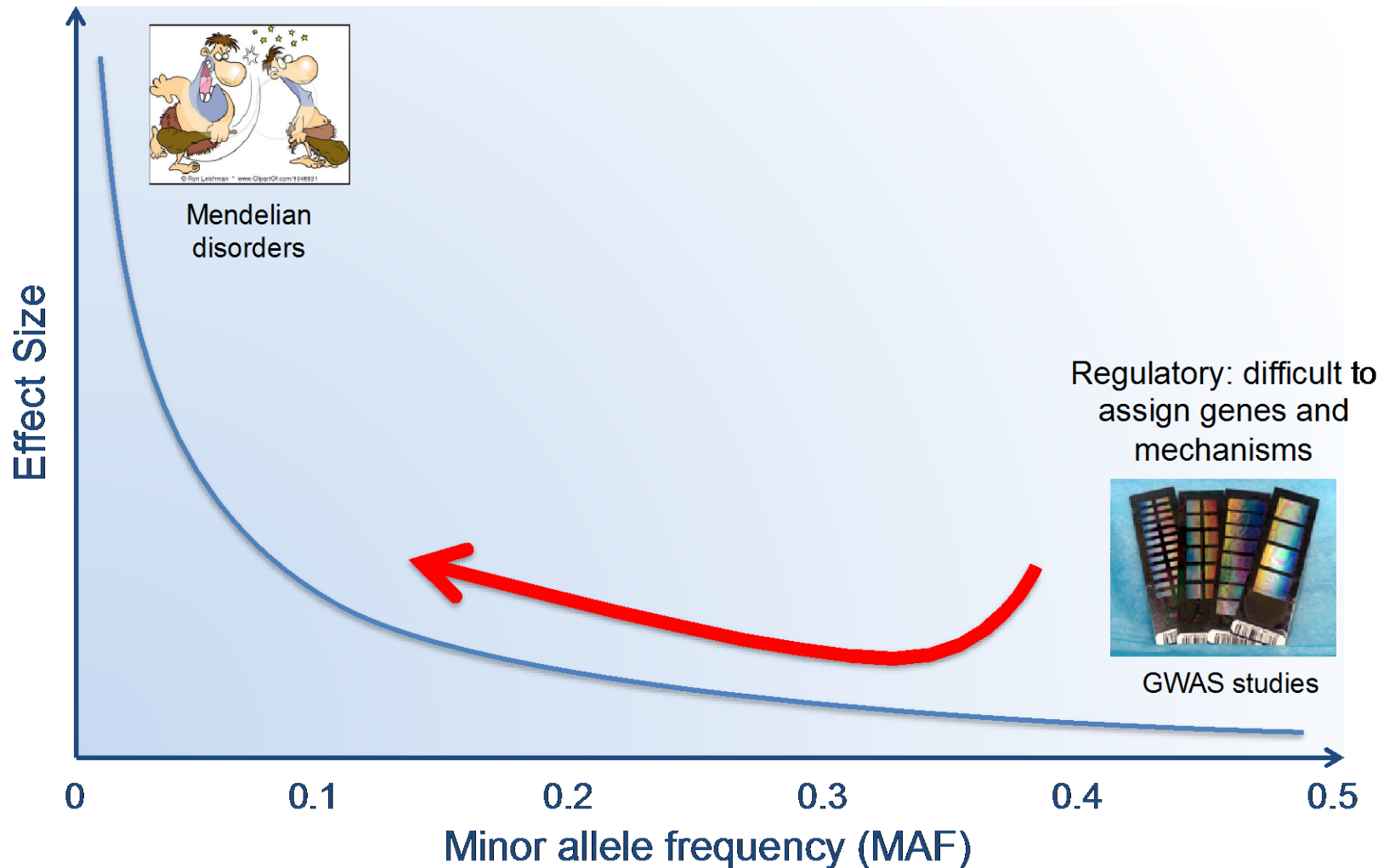
- 20,805 Genes
- 150M Single Nucleotide Variants
- 4.3M Structural Variants

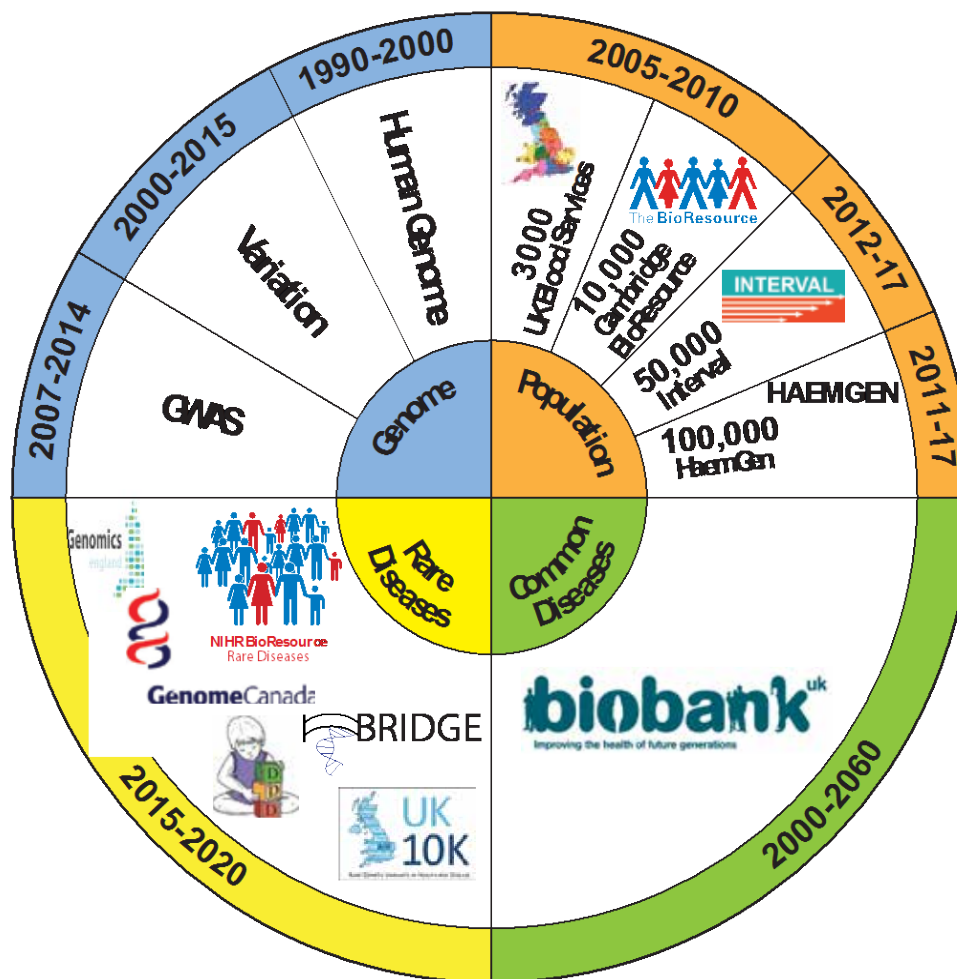
3 billion basepairs



# To discover clinically-informative variants

*for common and rare diseases*









# Embracing Genome Medicine

2001



**Reference  
genome**

2007



**First  
personal  
genomes**

2008



**First  
African  
genome**

2010



**1K**

2015



**10K**

2018



**100K**

2025

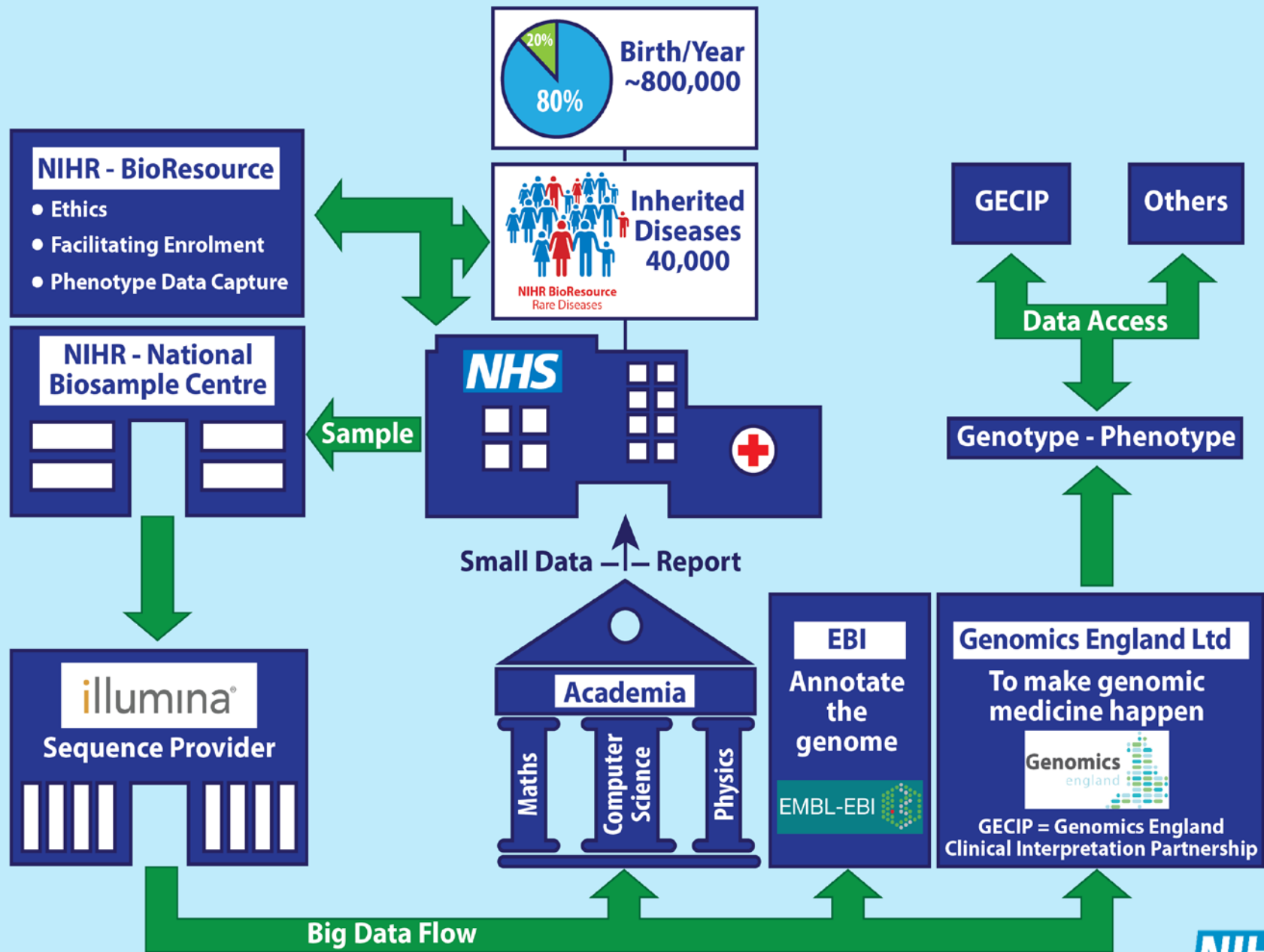


**2M**

**Sanger sequencing**

**Next-generation sequencing**

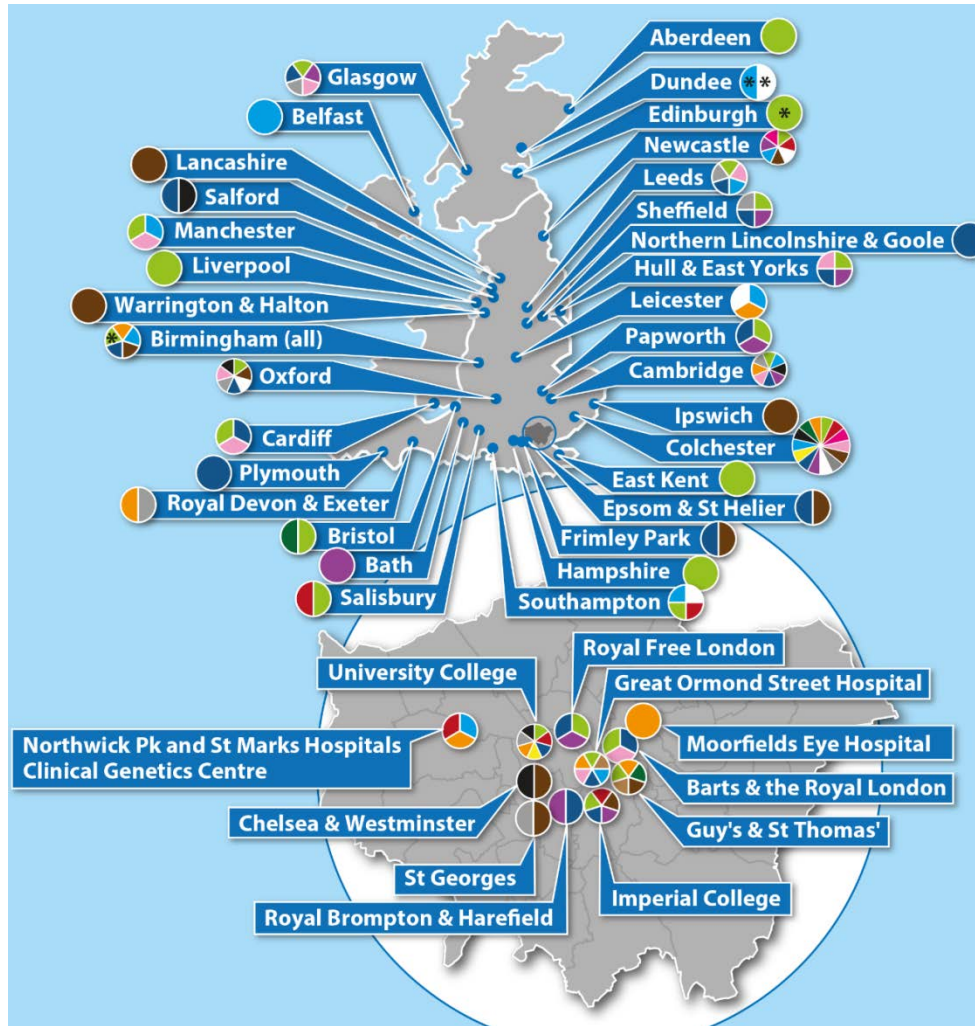
# The NHS 100,000 Genome Project





# The Rare Diseases Pilot study

*introducing whole genome sequencing into the NHS*



- Started in 2013
- **10,000** patients
- Many rare disorder categories
- **50** NHS Trusts
- Whole Genome Sequencing in a clinically accredited lab
- Clinical feedback + research



# NIHR BioResource - Rare Diseases

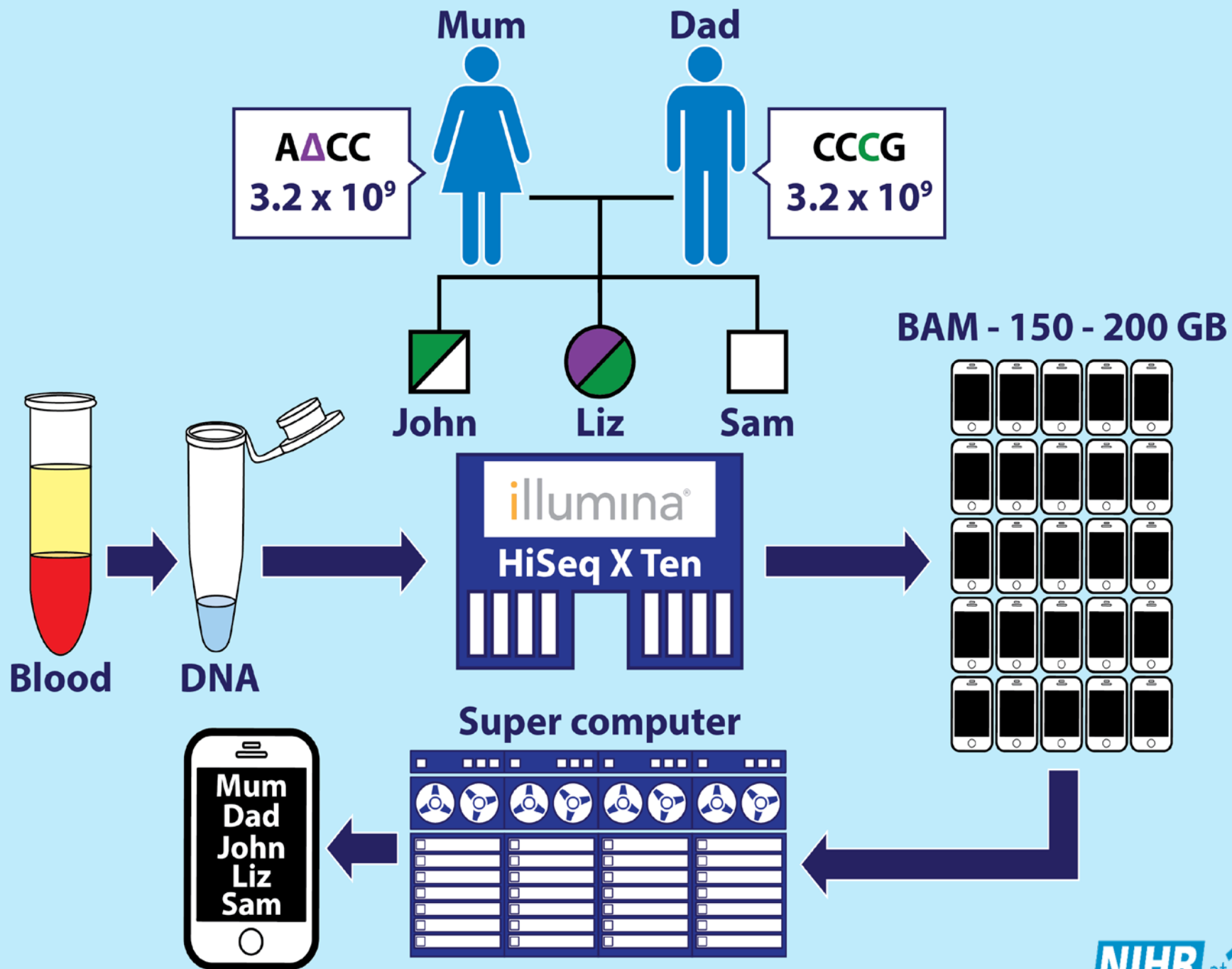


**NHS**  
National Institute for  
Health Research



The NIHR National BioResource is a collaboration between seven NIHR Biomedical Research Centres and Units, working together to create a unique resource of patients and healthy volunteers willing to participate in biomedical research.







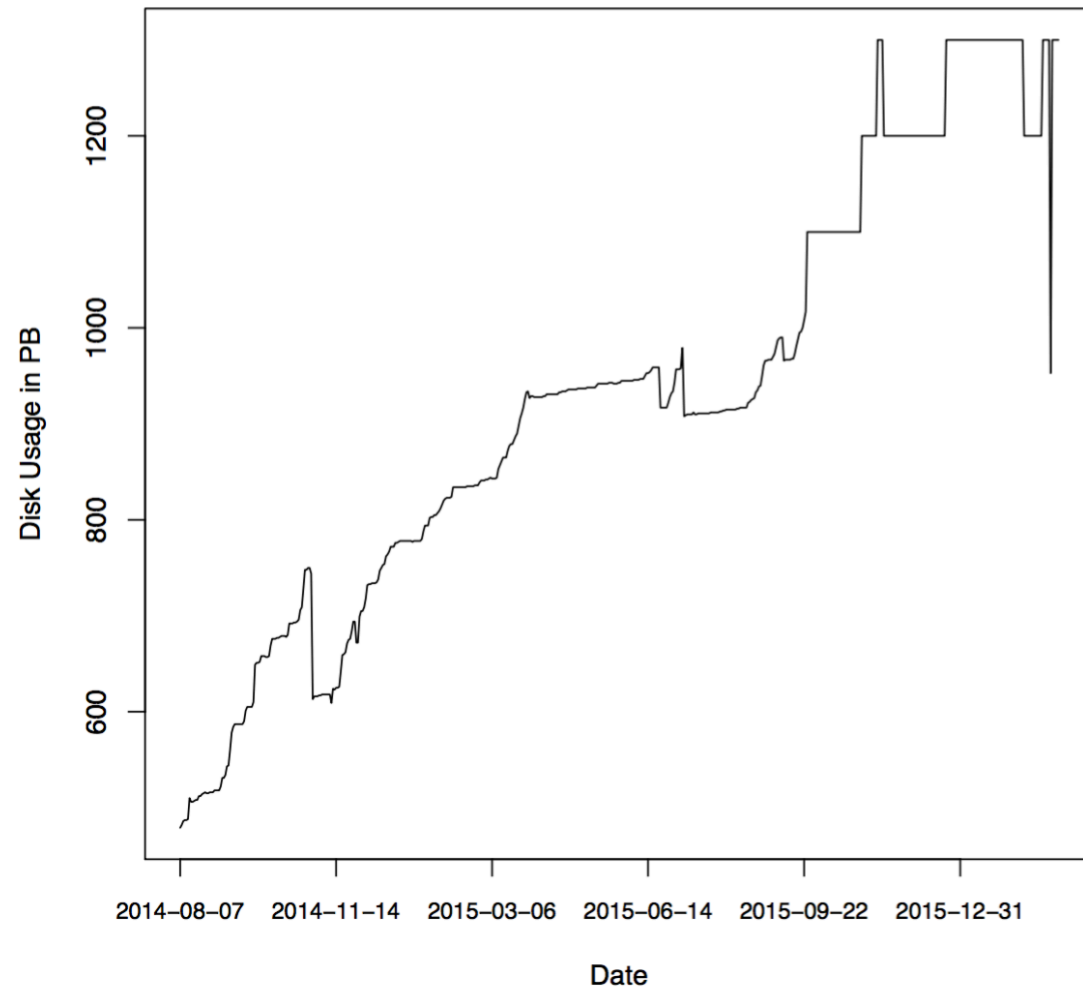
# Example of a DNA variant

*with annotation from the Human Gene Mutation database*

```
AN=11684;WGS10K_AN=11684;AC=2;WGS10K_AC=2;AF=0.000171174;WGS10K_AF=0.000171174;PASSRATE=1;
CALLRATE=0.999661;ANN=G|missense_variant|MODERATE|KCNJ10|ENSG00000177807|Transcript|ENST00000368089
|protein_coding|2/2||ENST00000368089.3:c.194G>C|ENSP00000357068.3:p.Arg65Pro|421/5293|194/1140|65/379|
R/P|cGc/cCc|rs137853066|1||-
1|SNV|HGNC|6256|YES||CCDS1193.1|ENSP00000357068|IRK10_HUMAN|Q9BXC5_HUMAN|UPI000012D8A6||delet
erious(0)|probably_damaging(1)|hmmpanther:PTHR11767&hmmpanther:PTHR11767:SF21&Pfam_domain:PF01007&
Gene3D:1.10.287.70&PIRSF_domain:PIRSF005465&Superfamily_domains:SSF81324&Prints_domain:PR01320|||||
||pathogenic||1|||||30|6.475330|1|||||C/C|C|C/C|||||ENSG00000177807|ENSP00000357068|ENST0
0000368089|||||0.95798|D|-
3.83|5.17|5.17|0.70671|0|0.51073|0.610034|0|0.34075|0.602189|0|0.14214|0.542086||0.000000|0.84324|D|0.0
00000|D|0.97488|0.9240|D|0.98055|1.0470|M|0.85792|2.78|R65P|0.81033|simple_aae|A|1|0.90998|D|-
6.36|D|0.89865|1.0|D|0.97092|1.0|9|0.91219|D|0.0|16.2022|0.81789|0.0:1.0:0.0:0.0||65|P78508|IRK10_HUMA
N|IRK10_HUMAN|R65P|P|65|R|G||5|rs137853066|SeSAME_syndrome|0|2|AEFDBI|D|0.87002|0.98806|KCNJ10|1
|158278753|1|160012129|0|0.20241|0.516011|1.000000|0.88722|1.000000|0.90892|0.807000|0.32744|0.871000
|0.44667|160042339|C|CGC|rs137853066|GO:0051935:L-
glutamate_uptake_involved_in_synaptic_transmission&GO:0021554:optic_nerve_development&GO:0009637:respons
e_to_blue_light|,G|upstream_gene_variant|MODIFIER|KCNJ10|ENSG00000177807|Transcript|ENST00000509700|pr
ocessed_transcript||||-/593|||||rs137853066|1|44|-
1|SNV|HGNC|6256|||||pathogenic||1|||||30|6.475330|||||;HGMD_
CLASS=DM;HGMD_MUT=ALT;HGMD_DB=rs137853066;HGMD_DNA=NM_002241.4:c.194G>C;HGMD_GENE=KCNJ10;H
GMD_PROT=NP_002232.2:p.R65P;HGMD_PHEN="Epilepsy_ataxia_sensorineural_deafness_and_tubulopathy";PhyloP_
UCSC=7.653;phastCons_UCSC=1;GenePhenotypes="Variation:dbSNP_ClinVar:rs137853066:SESAME_SYNDROME","Vari
ation:OMIM:rs137853066:SEIZURES,_SENSORINEURAL_DEAFNESS,_ATAXIA,_MENTAL_RETARDATION,_AND_ELECTROLY
TE_IMBALANCE","Variation:HGMD-
PUBLIC:CM092649:Annotated_by_HGMD_but_no_phenotype_description_is_publicly_available";GERP_UCSC=5.17;HG
MD_ID=CM092649
```

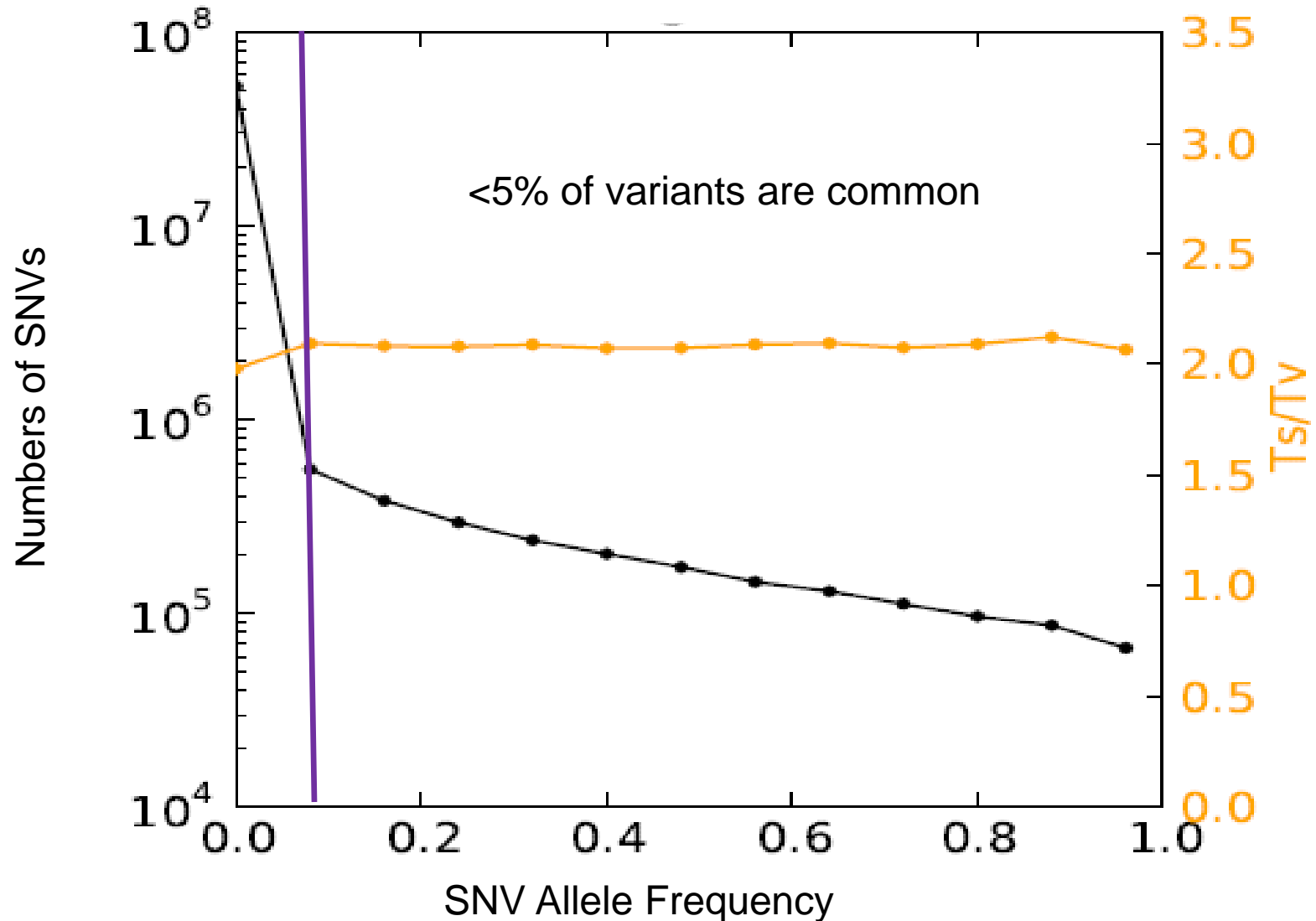
# Whole Genome Data growth to 1.4 Petabyte

*Cambridge University High Performance Computer – 90 largest after the Chinese army*



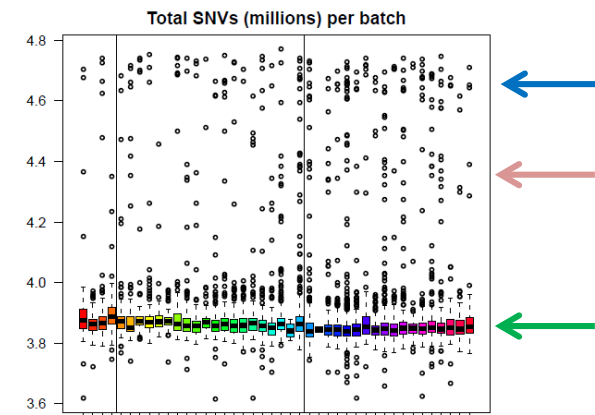
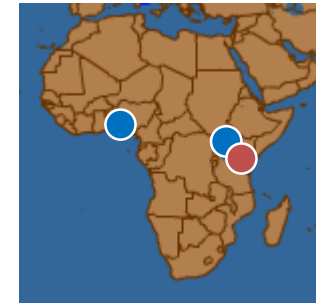
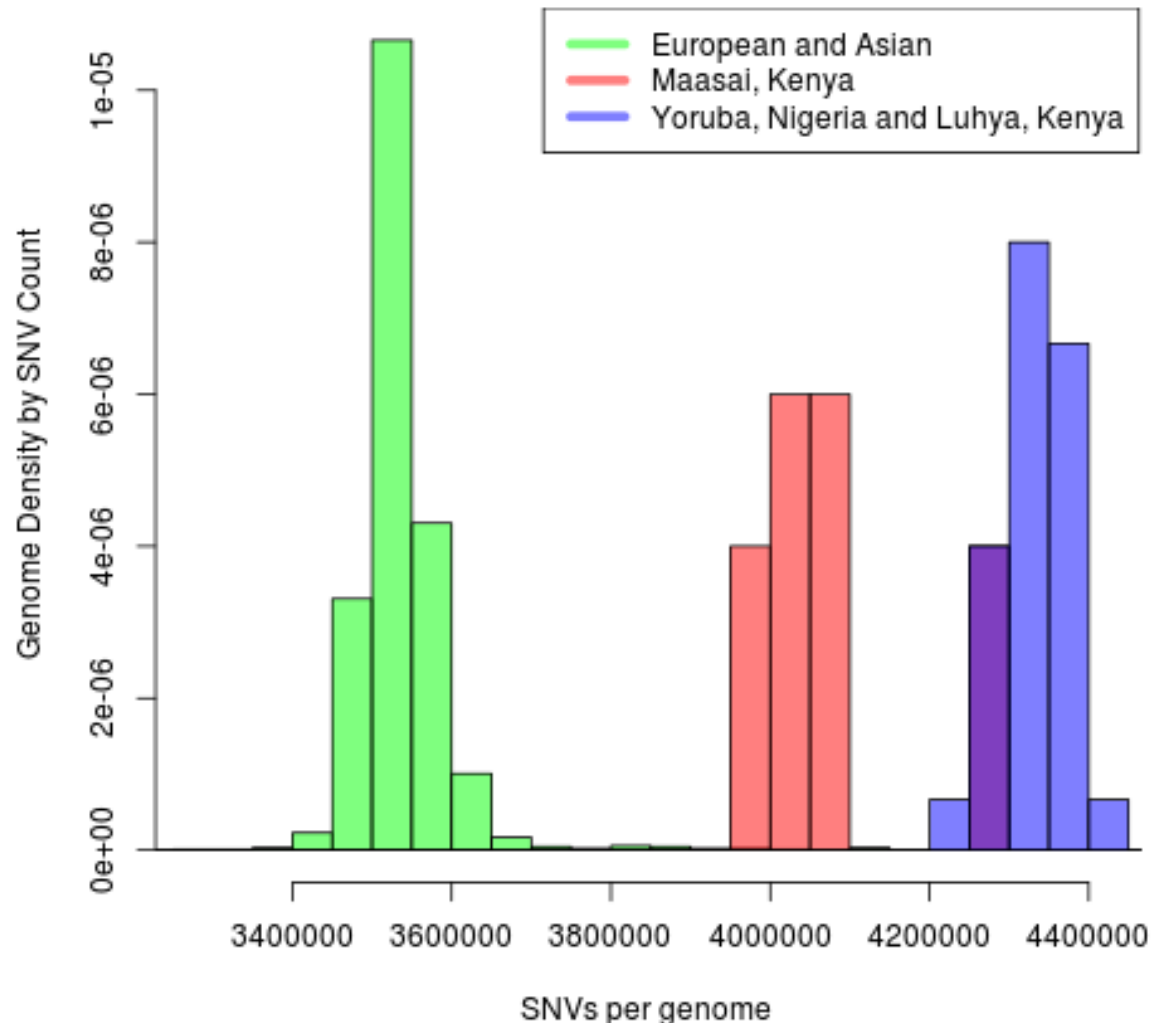
# Whole Genome Sequence of 7,200 DNA samples

*Reveals 55M unique SNVs with most being extremely rare*





# The vast majority of DNA samples sequenced *are from individuals of European ancestry – important for reporting*



Total number of SNVs per genome depends on the population the genome originated from :  
Populations are based on closest HapMap population using PCA analysis

# How good are we in delivering a conclusive diagnosis?

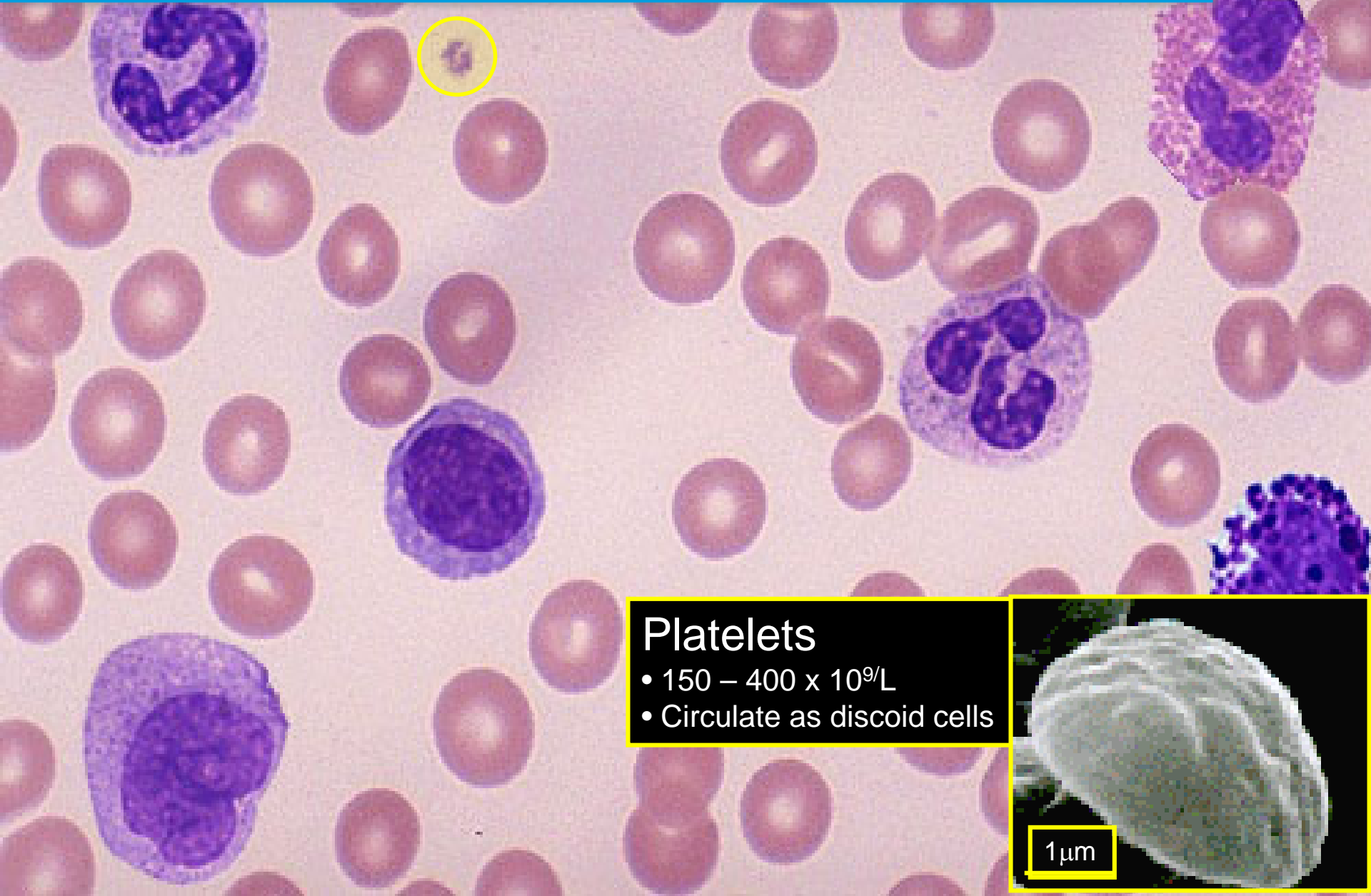
*analysis of the results for the first 4,000 patient samples*



Project	Diagnostic rate	Prescreening	Tier 1 genes	Genes reported	Unique variants	Novel variants
RETINOPATHY						
NEUROLOGICAL						
HYPERTENSION						
HAEMOSTASIS						
IMMUNE						

# How are we doing in discovering new genes for rare diseases

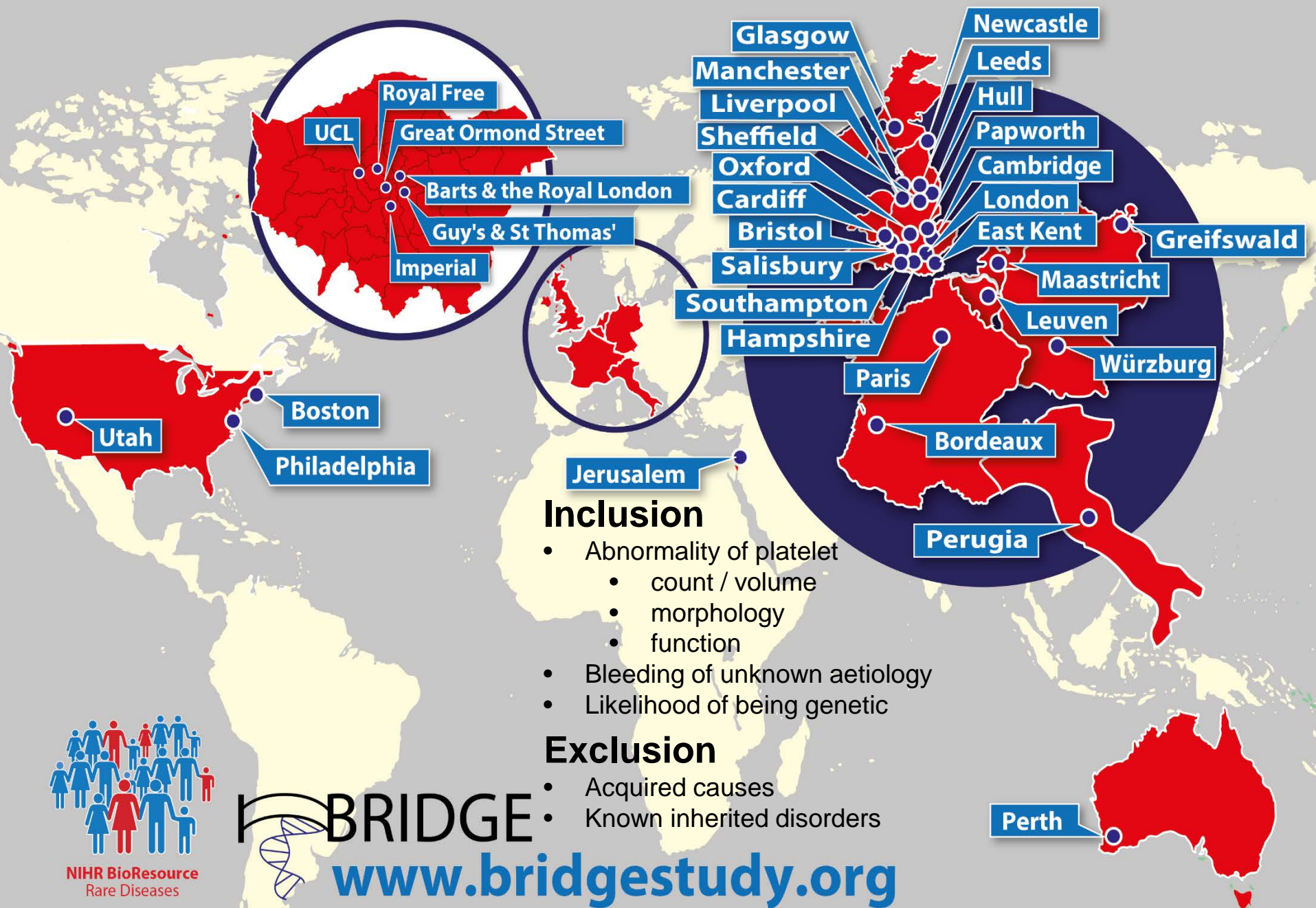
*85% of patients with 'funny looking platelets' have no diagnosis*



## Platelets

- $150 - 400 \times 10^9/L$
- Circulate as discoid cells





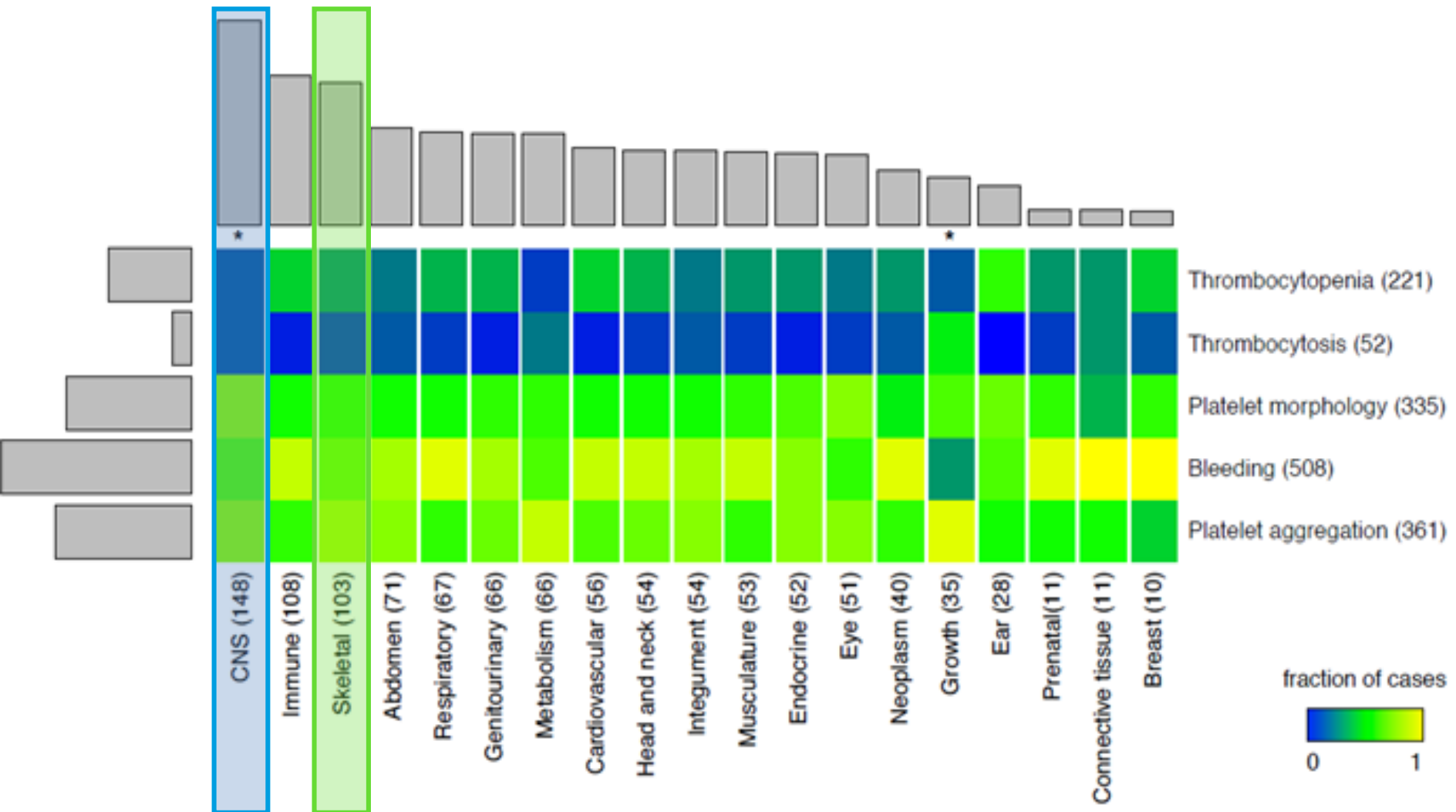






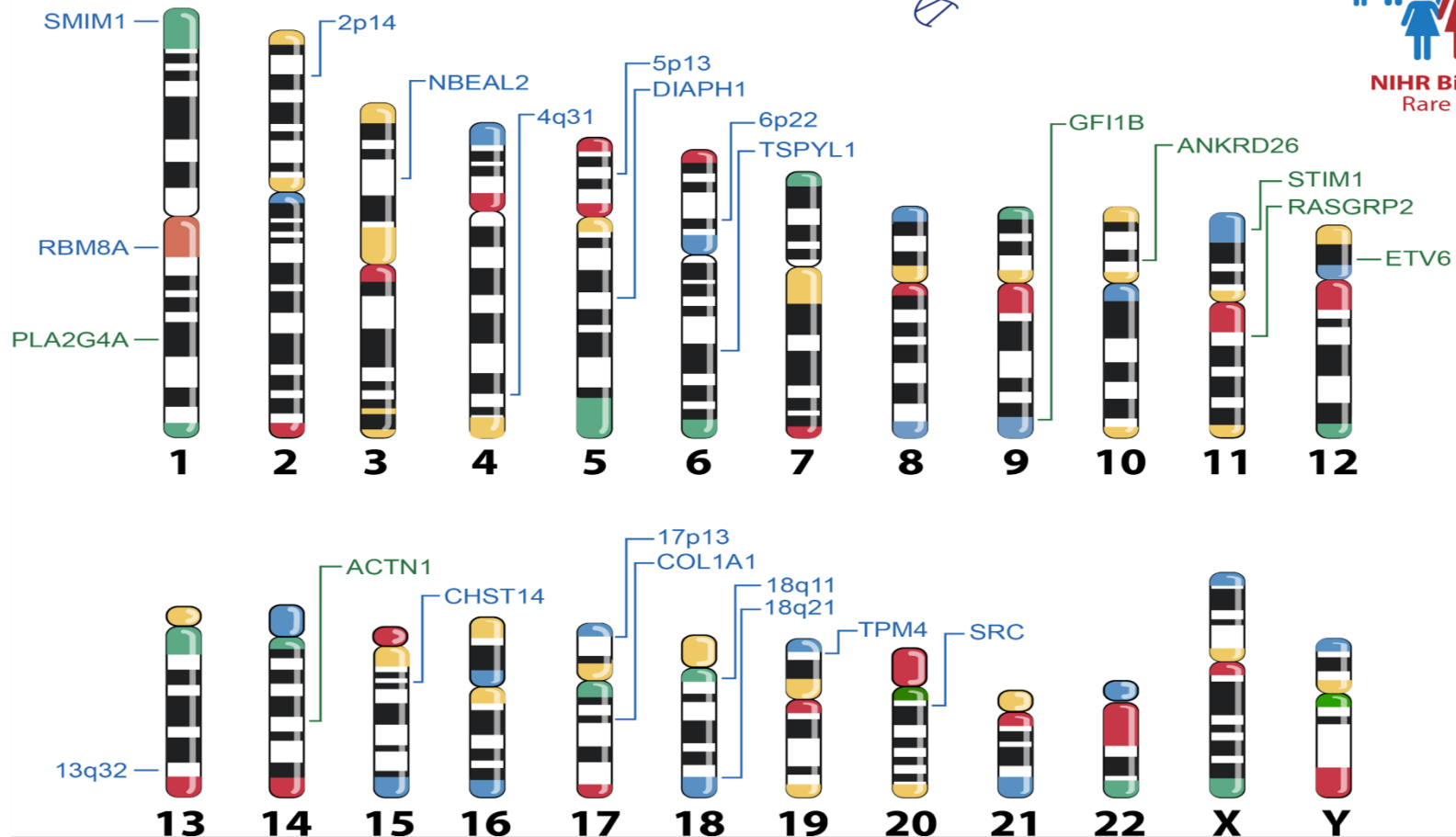
# Capturing the complexity of human disease by HPO terms

*autism and immunological & skeletal disorders in patients who consulted their haematologist?*



# 24 New Genes since 2011

and many more to follow

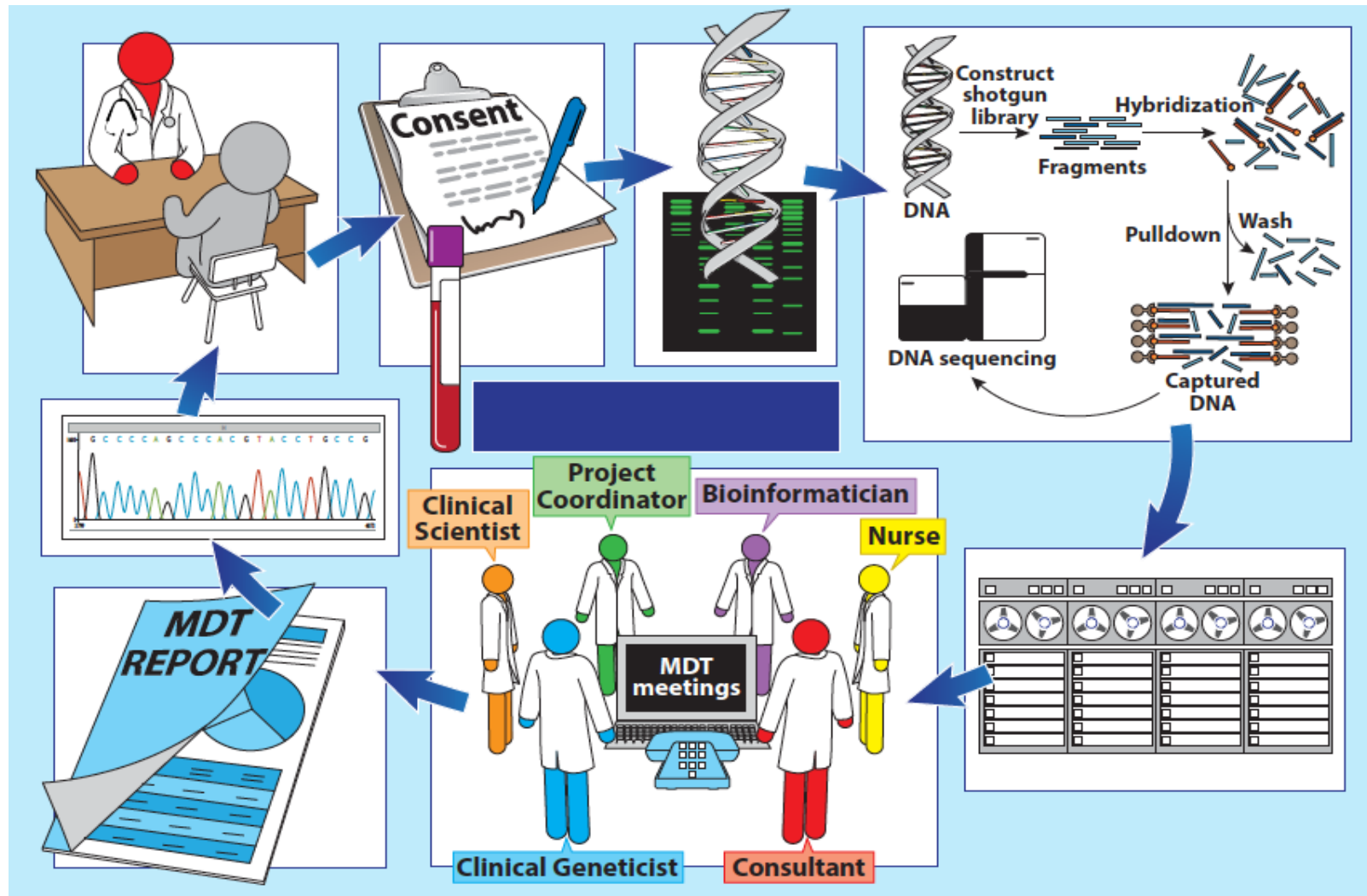


Blue = BRIDGE-BPD; Green = Other groups

Albers *et al*, Nat Genetics 2011; Albers *et al*, Nat Genetics 2012; Cvejic *et al*, Nat Genetics 2013; Chen *et al*, Science 2014; Westbury *et al*, Genome Medicine 2015; Green *et al*, AJHG 2016; Stritt *et al*, Blood 2016; Stritt *et al*, Nature Communications 2016; Turro *et al*, Science Translational Medicine 2016; Simeoni *et al*, Blood 2016; Lentaigne *et al*, Blood 2016; Sivapalaratnam *et al*, Blood (revision)

# Translating the discovery of new genes into the clinic

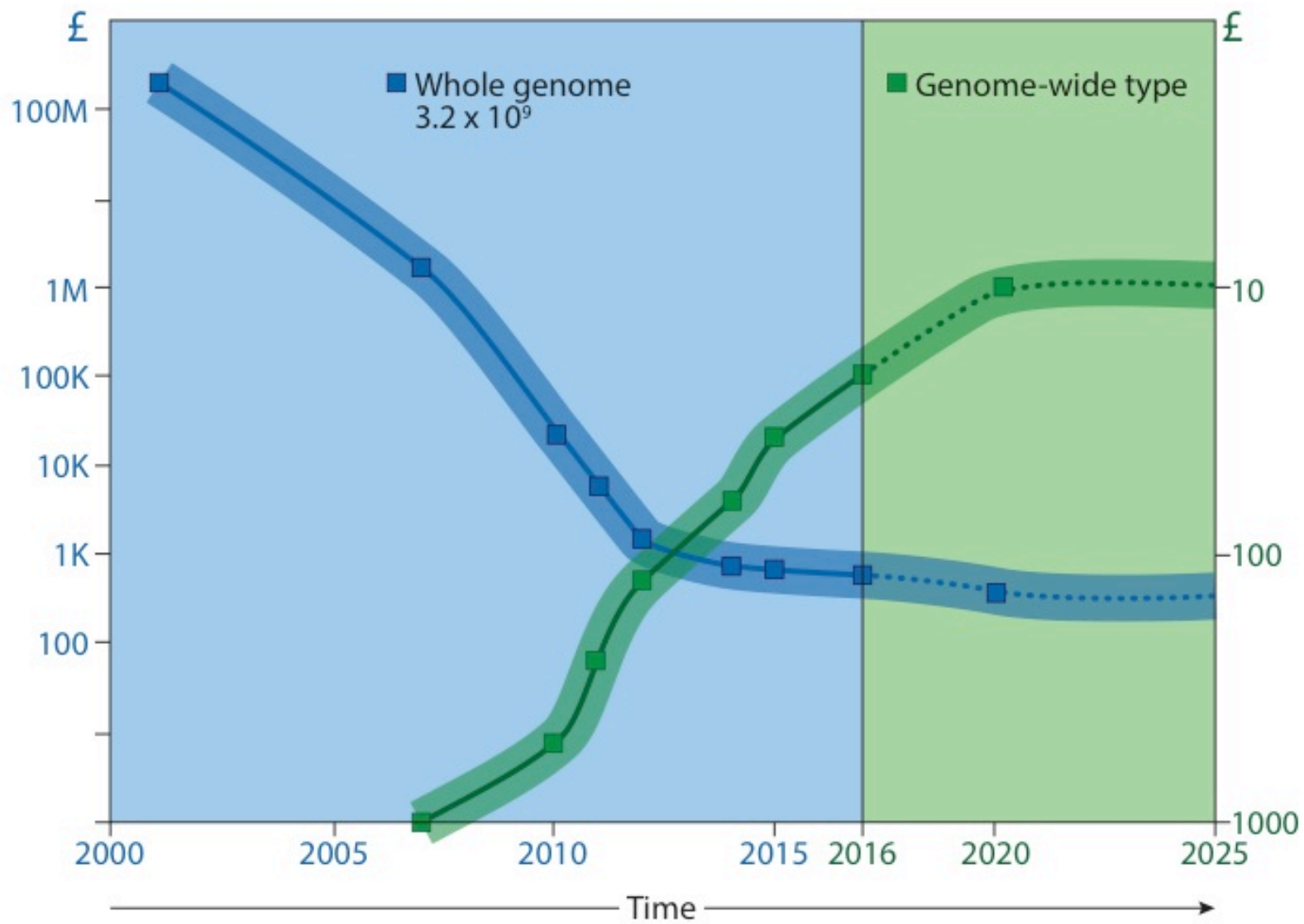
*ThromboGenomics test for patients with platelet, thrombotic and bleeding disorders*



11:10

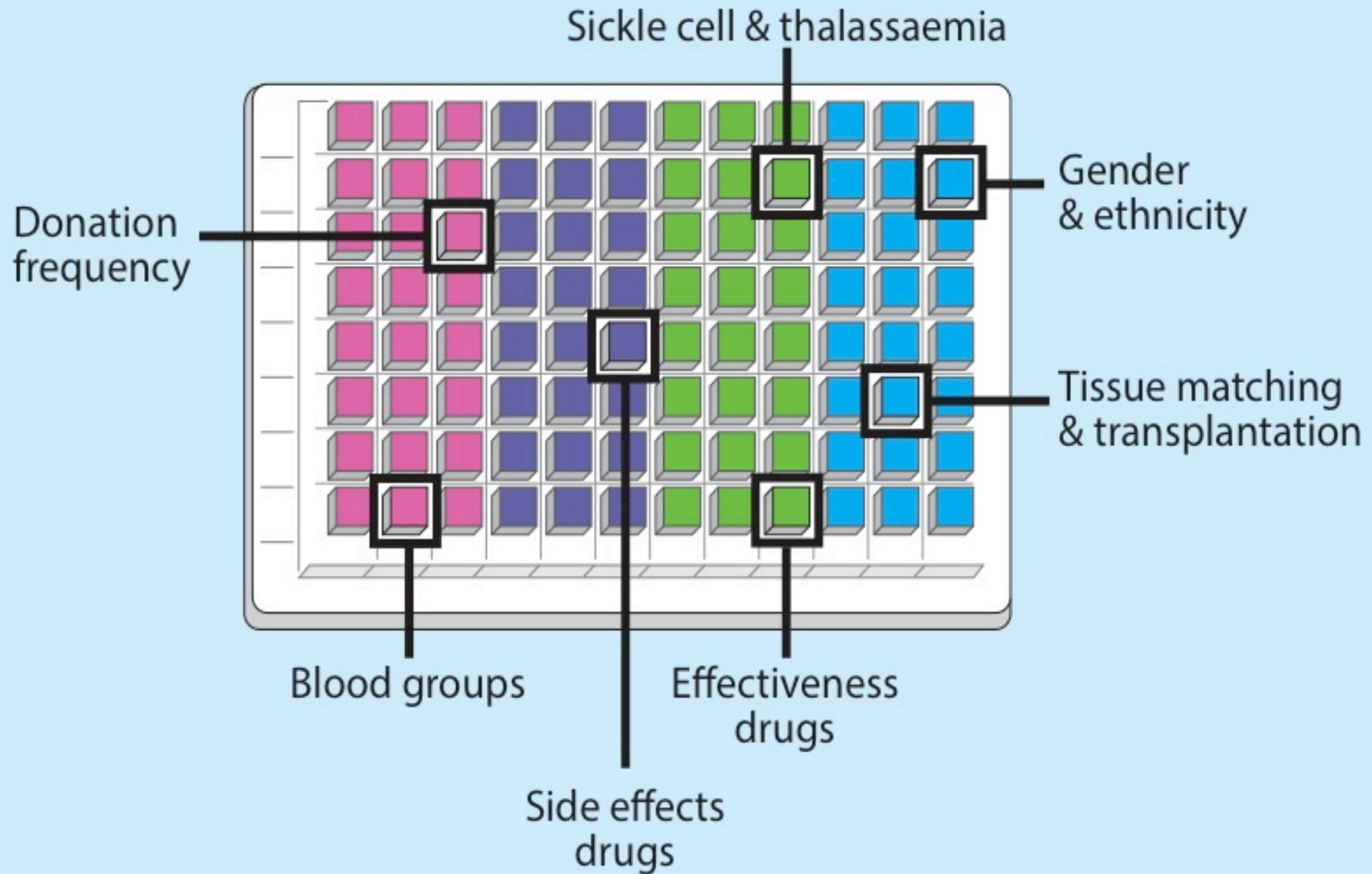
## BLOOD TEST BREAKTHROUGH





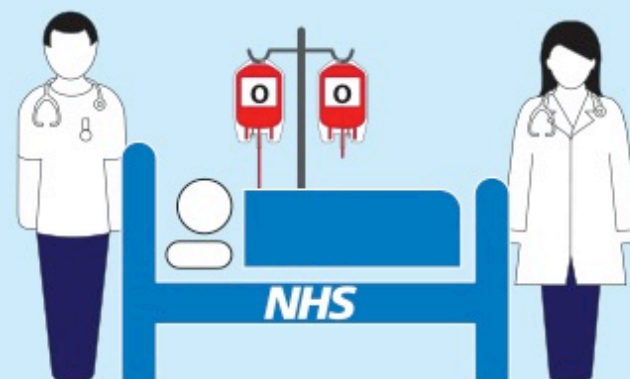
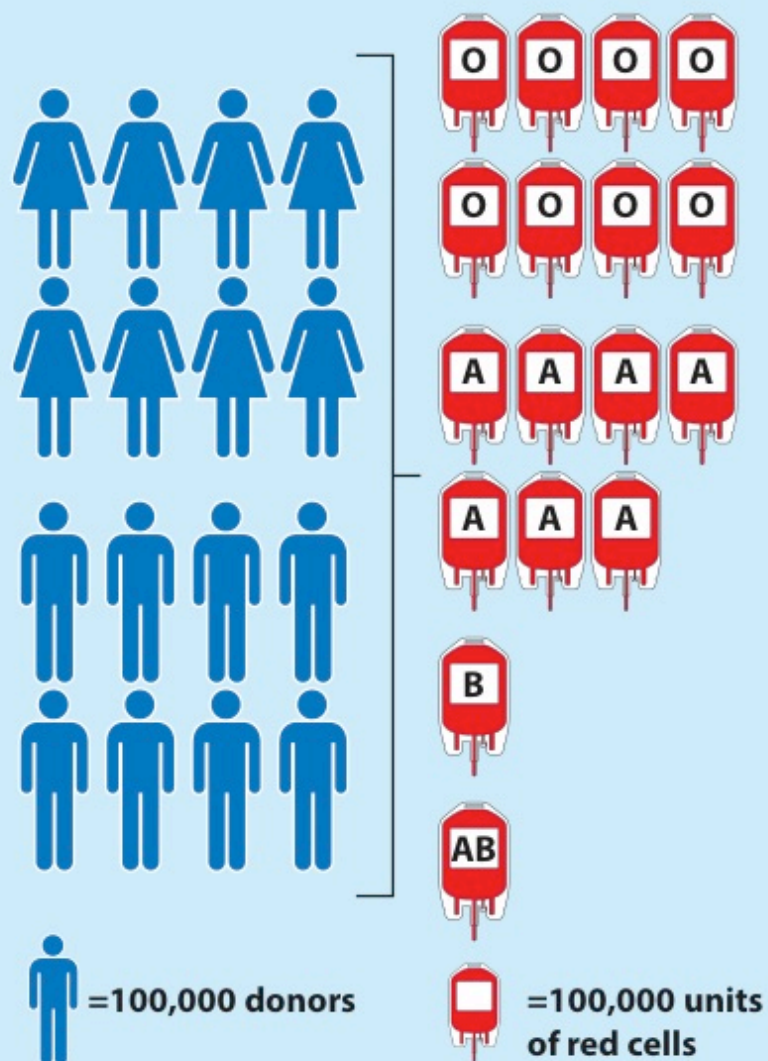


# 1,000,000 DNA variants

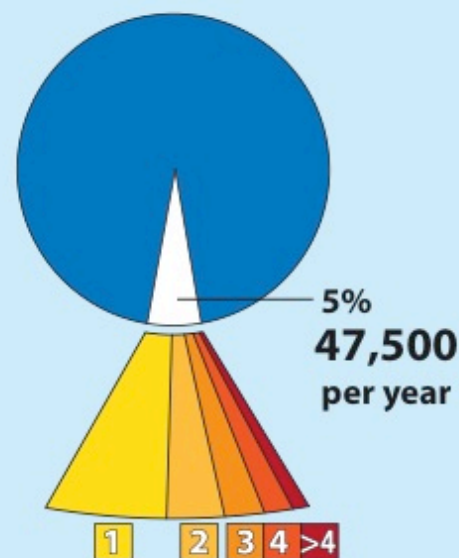


# Blood services are the largest match-makers

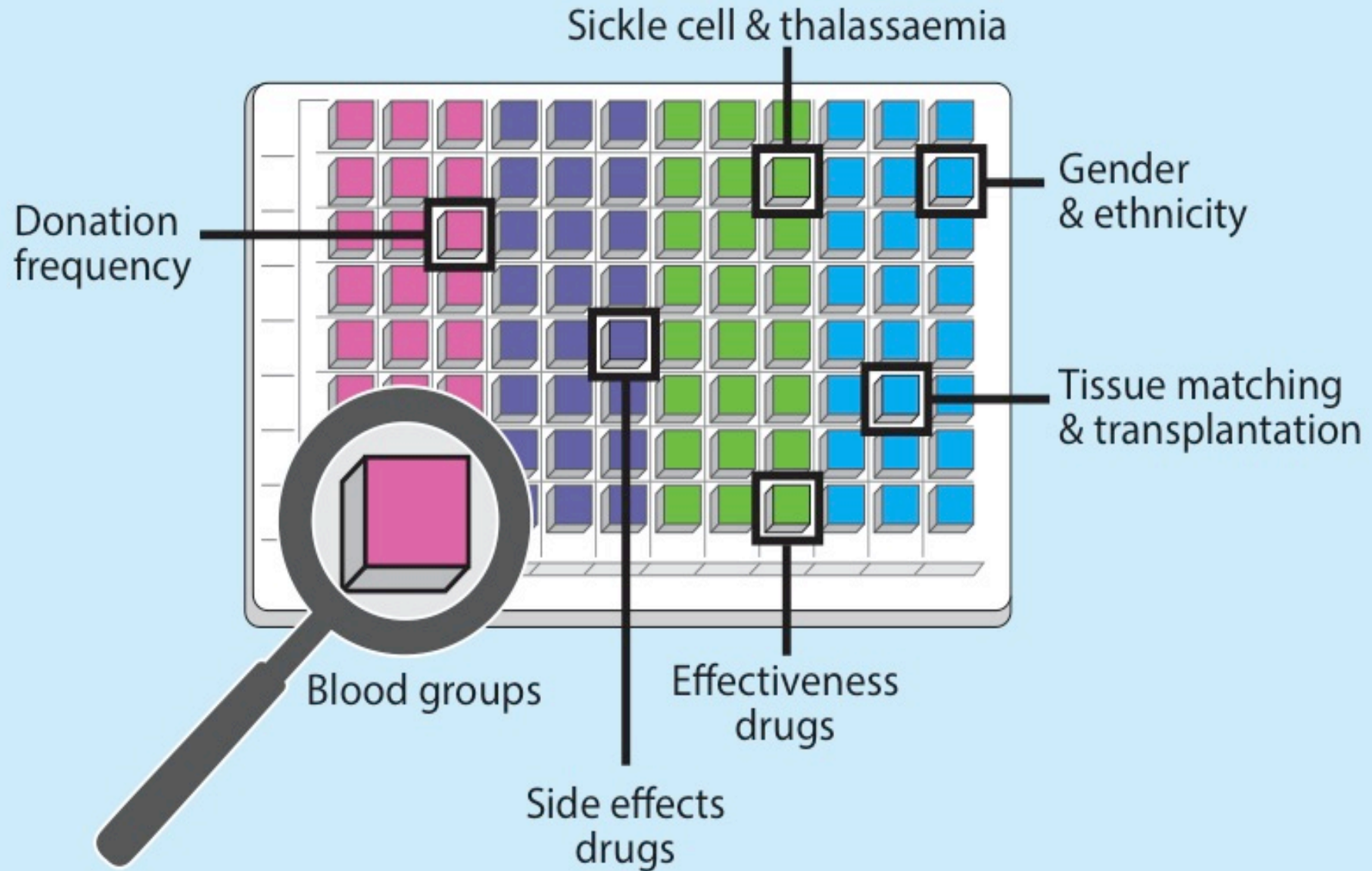
*should we type all donors for all the clinically relevant blood cell alloantigens?*



**850,000**  
transfusion episodes

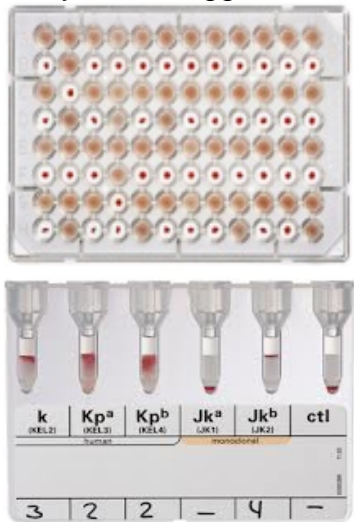


# 1,000,000 DNA variants

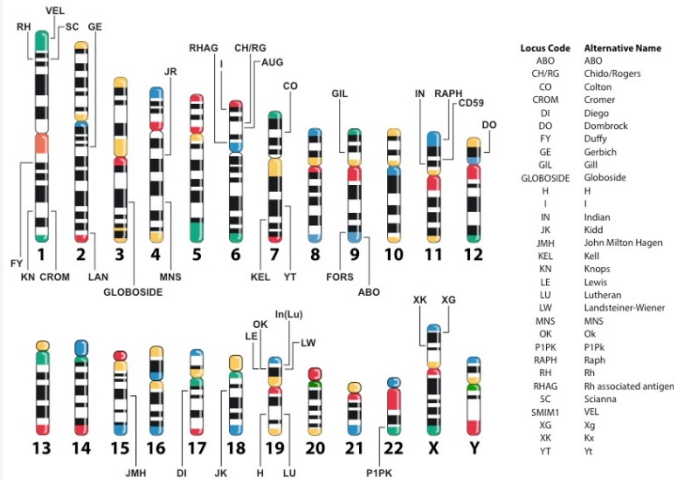




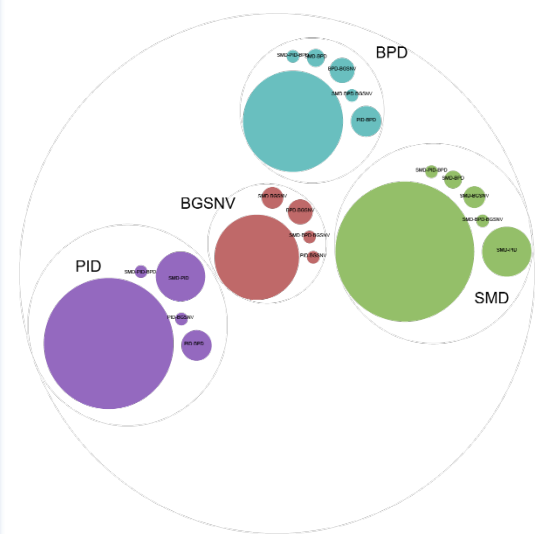
100+ years of agglutination



30 years of blood group genetics



4 years of knowledge curation

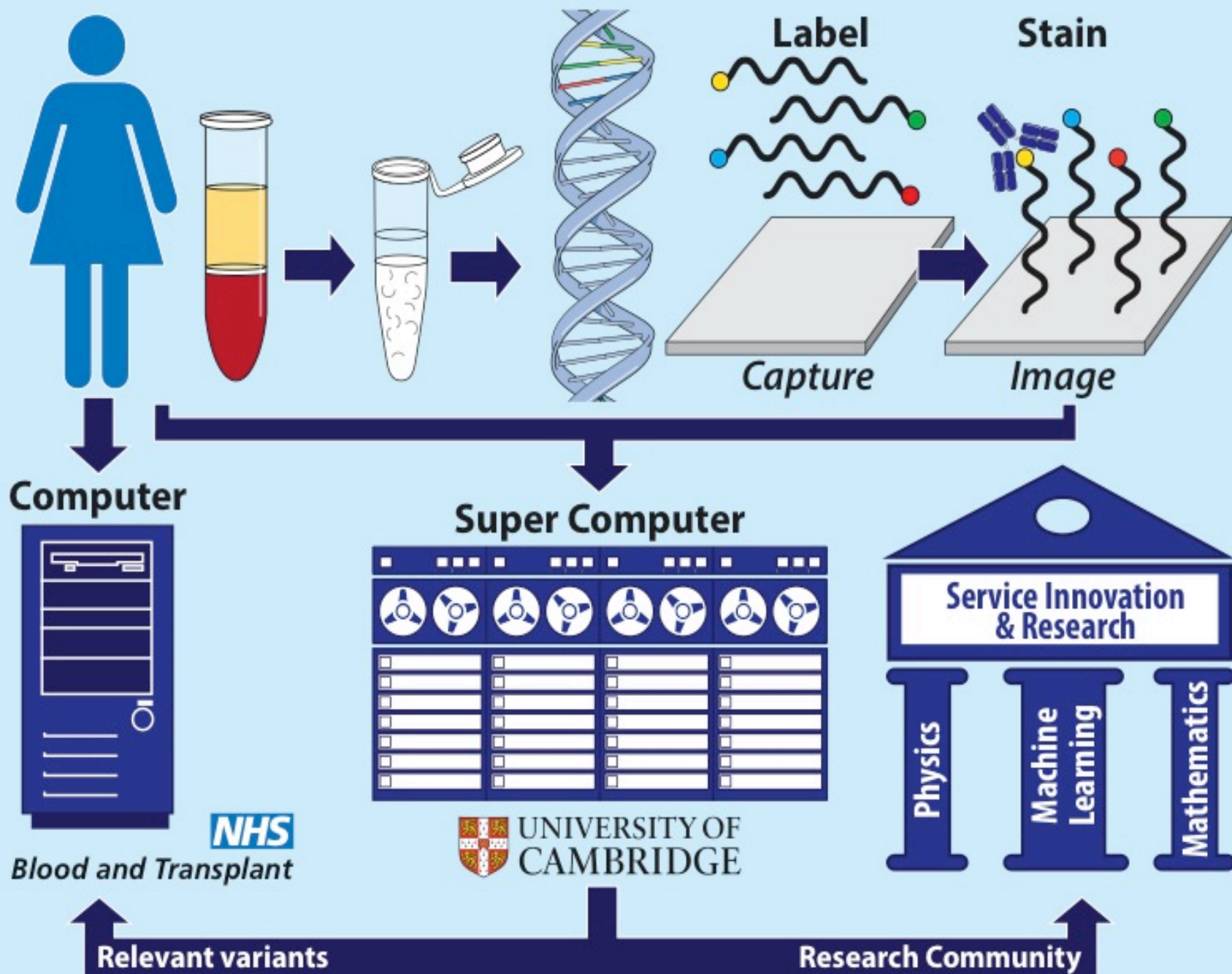


**Moving knowledge from expensive text books to public databases**  
*this is being delivered through a multi-disciplinary partnership approach*

A collage of logos for the institutions and organizations involved in the project. The logos include:

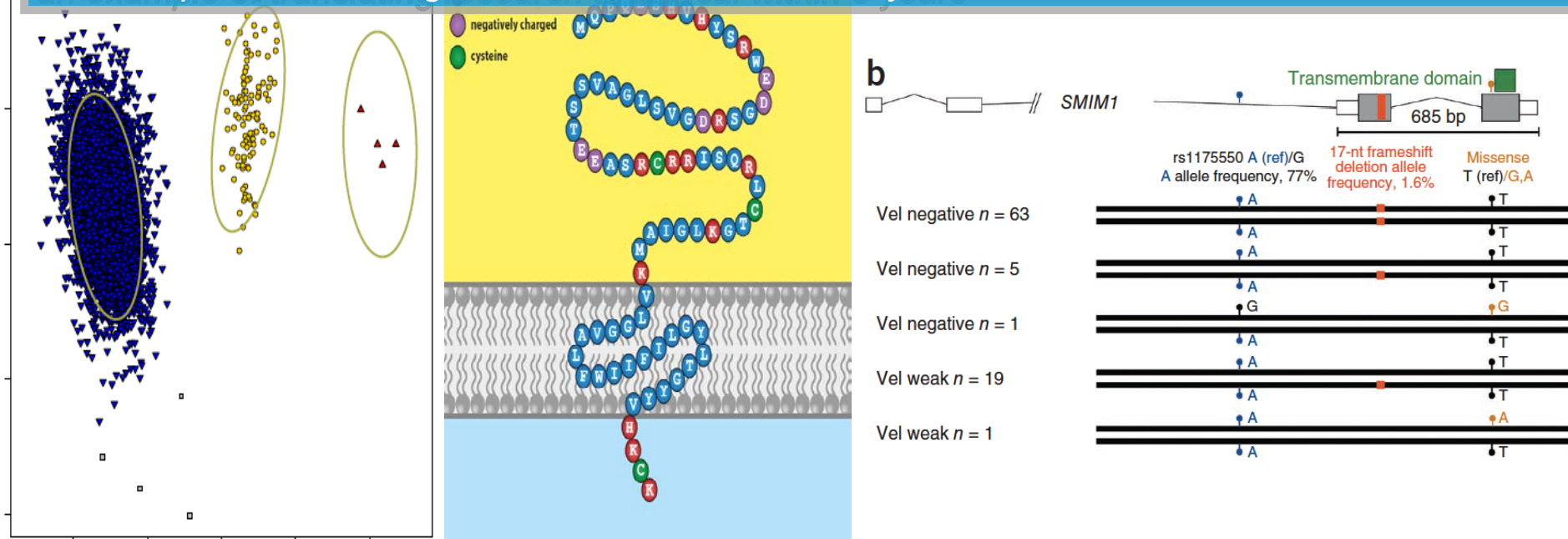
- UNIVERSITY OF OXFORD
- UNIVERSITY OF CAMBRIDGE
- IBGRL Reference Services
- EMBL-EBI
- LRG
- MRC Biostatistics Unit
- ANTHONY NOLAN
- AFFYMETRIX
- Genomics england
- NIHR BioResource
- Sanquin Blood Supply
- NHS Blood and Transplant





# Can blood groups be ascertained by genotyping

*an example of translating research to clinical within 3 years*



LETTERS

## *SMIM1* underlies the Vel blood group and influences red blood cell traits

Ana Cvejic<sup>1,2,21</sup>, Lonneke Haer-Wigman<sup>3,4,21</sup>, Jonathan C Stephens<sup>1,5,6,21</sup>, Myrto Kostadima<sup>7</sup>,

# Conclusions & Opportunities

1. 100 000 Genomes Project delivers improvements for patients with rare diseases

2. ThromboGenomics DNA test for patients with rare inherited bleeding, thrombotic and platelet disorders

3. Cheap genotyping test may allow NHSBT to further improve the matching between donors and patients



## The BRIDGE Bleeding, Thrombotic and Platelet Disorders Consortium

## Funding

Imperial College  
London

KU LEUVEN

UNIVERSITY OF UTAH  
SCHOOL OF MEDICINE

CH

Royal Free London  
NHS Foundation Trust

NHS



Universitätsmedizin  
GRIFFSWALD



UMC Utrecht

The Children's Hospital  
of Philadelphia

University of  
BRISTOL

THE UNIVERSITY OF  
WESTERN  
AUSTRALIA



UNIVERSITY OF  
LIVERPOOL

CRPP

Beth Israel Deaconess  
Medical Center

Salisbury  
NHS Foundation Trust

Hampshire Hospitals  
NHS Foundation Trust

CHARITÉ

A teaching hospital of  
Harvard Medical School

EMBL-EBI

University Hospital Southampton  
NHS Foundation Trust

southern  
haemophilia  
network

Maastricht UMC+  
azM | Maastricht University



Rudolf Virchow  
Zentrum  
DFG Research Center  
for Experimental Biomedicine  
University of Würzburg

Oxford University Hospitals  
NHS Trust

Cambridge University Hospitals  
NHS Foundation Trust



East Kent Hospitals University  
NHS Trust

CATGO  
Cambridge Translational  
Genomics Laboratory

ASSISTANCE  
PUBLIQUE

HÔPITAUX  
DE PARIS

Great Ormond Street  
Hospital for Children  
NHS Trust

Central Manchester  
University Hospitals  
NHS Foundation Trust



Bwrdd Iechyd Prifysgol  
Abertawe Bro Morgannwg  
University Health Board

Hôpitaux  
Universitaires  
Est Parisien

TROUSSEAU  
LA ROCHE-GUYON

The Newcastle upon Tyne Hospitals  
NHS Foundation Trust

Barts Health  
NHS Trust

NHS  
National Institute for  
Health Research

wellcome trust



NHS  
Blood and Transplant

MRC | Medical  
Research  
Council