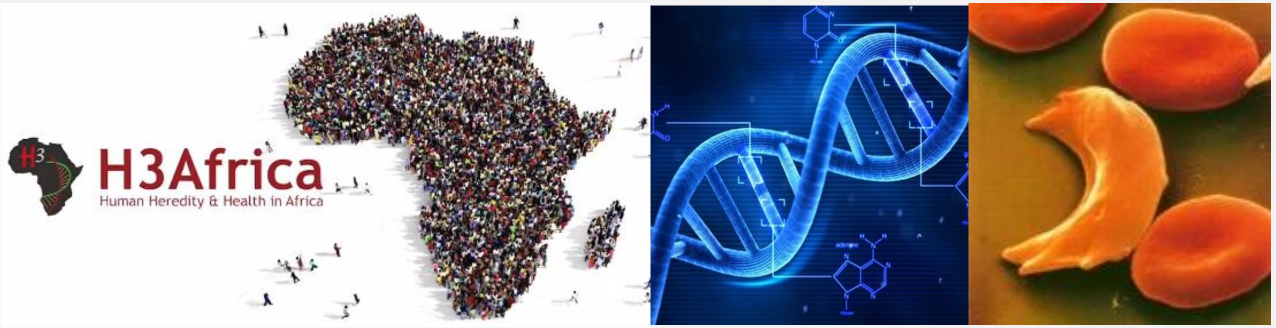
The Importance of Studying Genetic Variation In Different Population Groups

Neil Hanchard, MD, DPhil
neil.Hanchard@nih.gov



Senior Investigator, Center for Precision Health Research
Head, Childhood Complex Disease Genomics Section
National Human Genome Research Institute
National Institutes of Health

Adjunct Associate Professor
Molecular and Human Genetics
Baylor College of Medicine
Hanchard@bcm.edu

NIH National Human Genome Research Institute

The Forefront of Genomics®

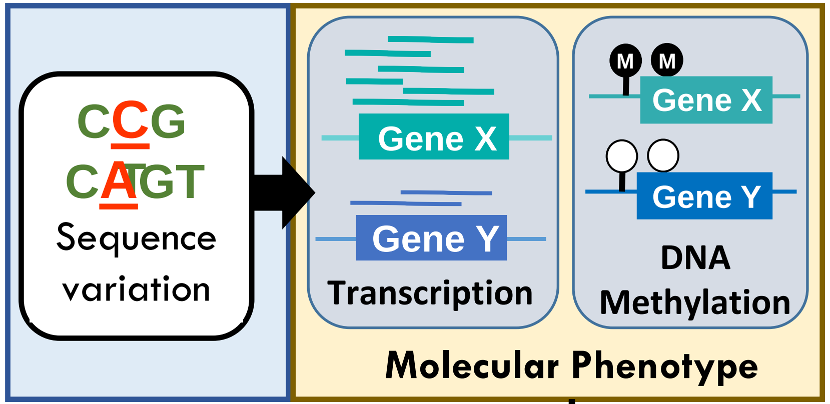
Outline

Importance of Genomics in Different Populations

H3Africa & The Collaborative African Genomics Network

Transfusion Antibodies in Sickle Cell Disease

Childhood Complex Disease Genomics Section

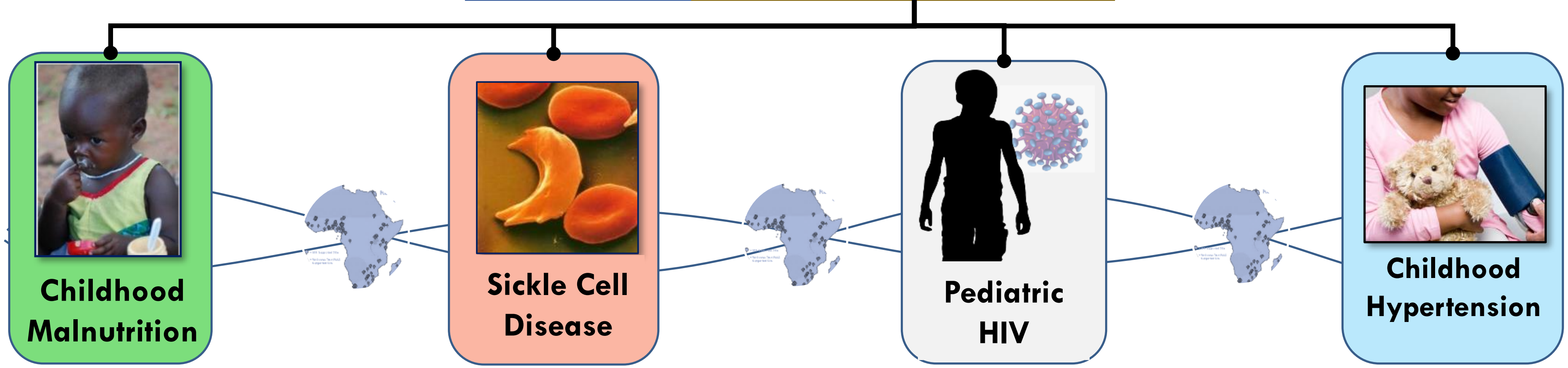


Sequence variation

Transcription

DNA Methylation

Molecular Phenotype



Childhood Malnutrition

Sickle Cell Disease

Pediatric HIV

Childhood Hypertension

Why do some children get Kwashiorkor when severely malnourished?

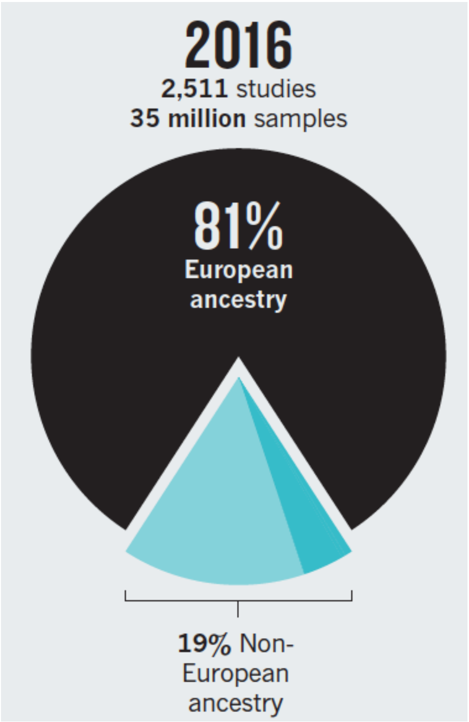
Why do some individuals make alloantibodies after red cell transfusions?

Why do some HIV positive children develop AIDS quickly and others slowly?

Are there single gene defects underlying early-onset essential hypertension?

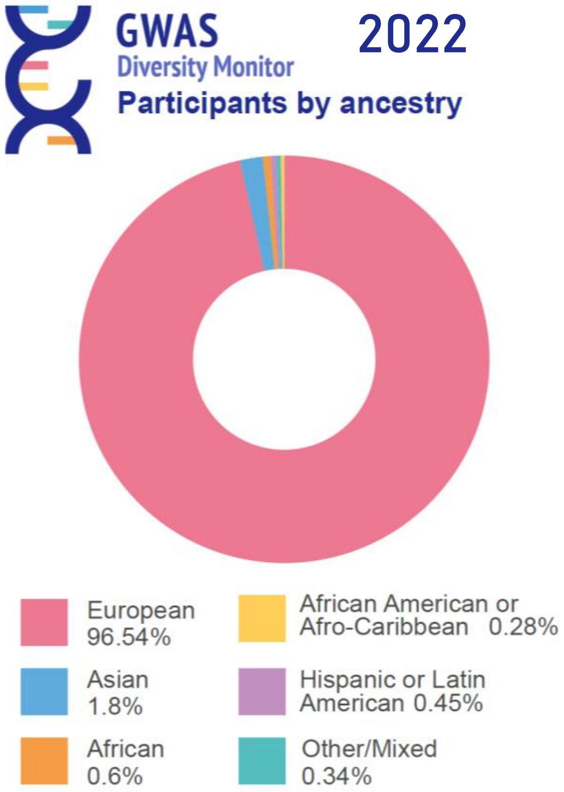
Diversity in human genomics

Genomic studies of diseases and traits have predominantly been conducted in European ancestry populations



Genomics is failing on diversity

An analysis by **Alice B. Popejoy** and **Stephanie M. Fullerton** indicates that some populations are still being left behind on the road to precision medicine.



Mills MC & Rahal C, 2020, Nature Genetics

ARTICLE

OPEN

https://doi.org/10.1038/s41586-018-0579-z

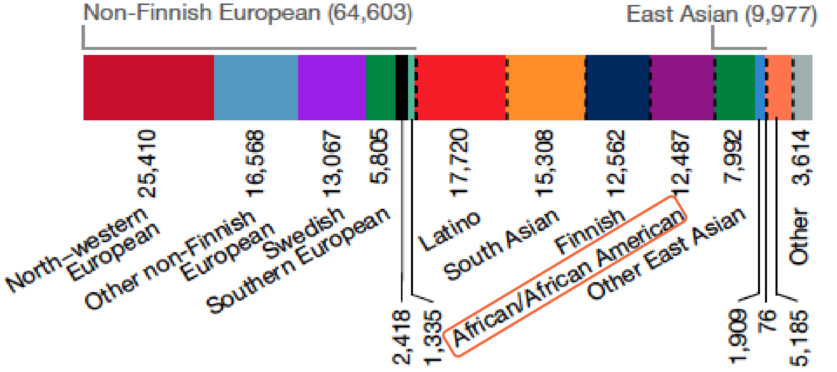
The UK Biobank resource with deep phenotyping and genomic data

Self-reported ethnicity		Representation (%)	
White		94.06	
	British		88.07
	Irish		2.63
	Any other white background		3.36

Bycroft et. al, 2018, Nature

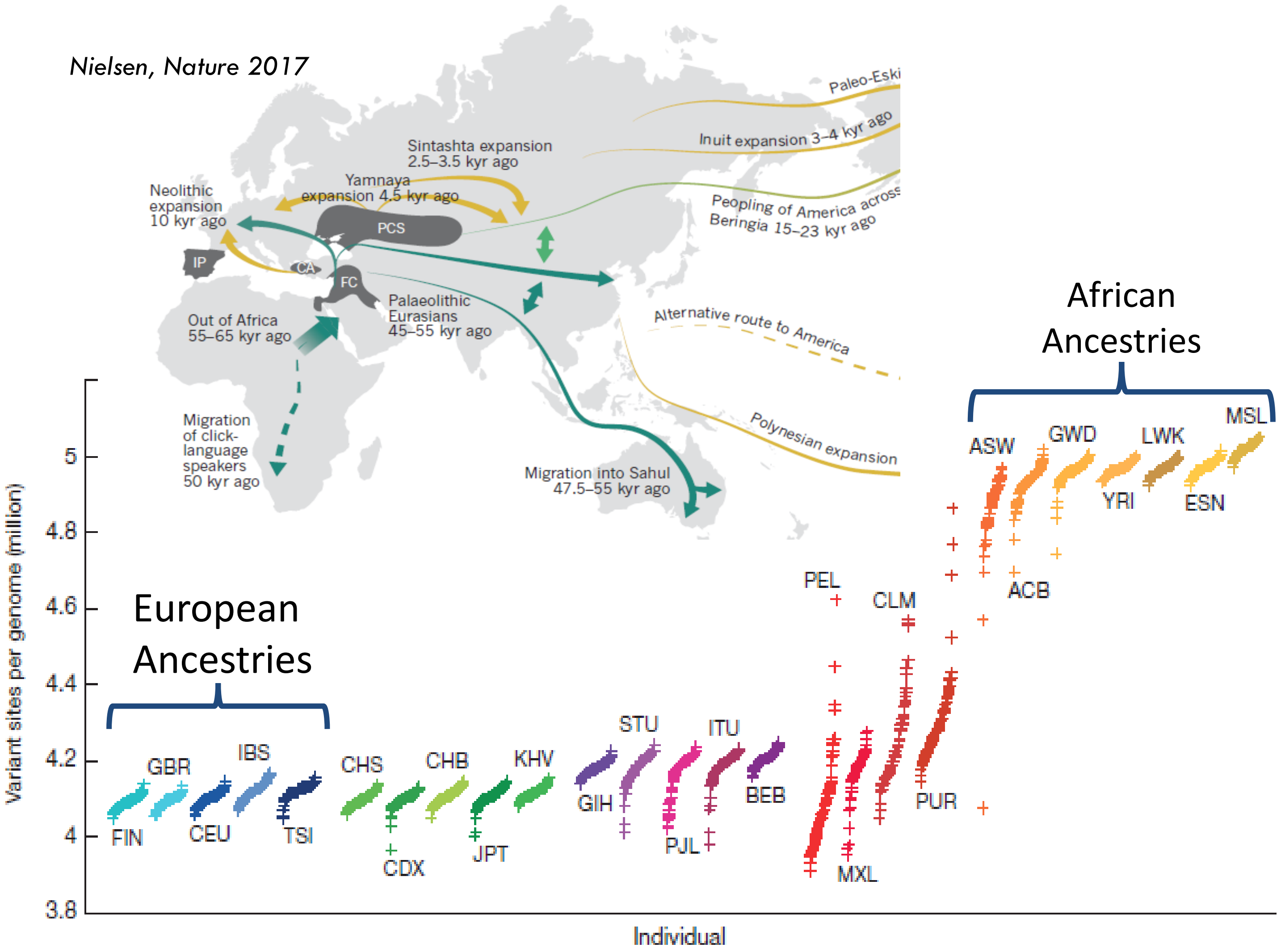
Article

The mutational constraint spectrum quantified from variation in 141,456 humans



Karczewski et. al, 2020, Nature

Genetic Variation in Africa

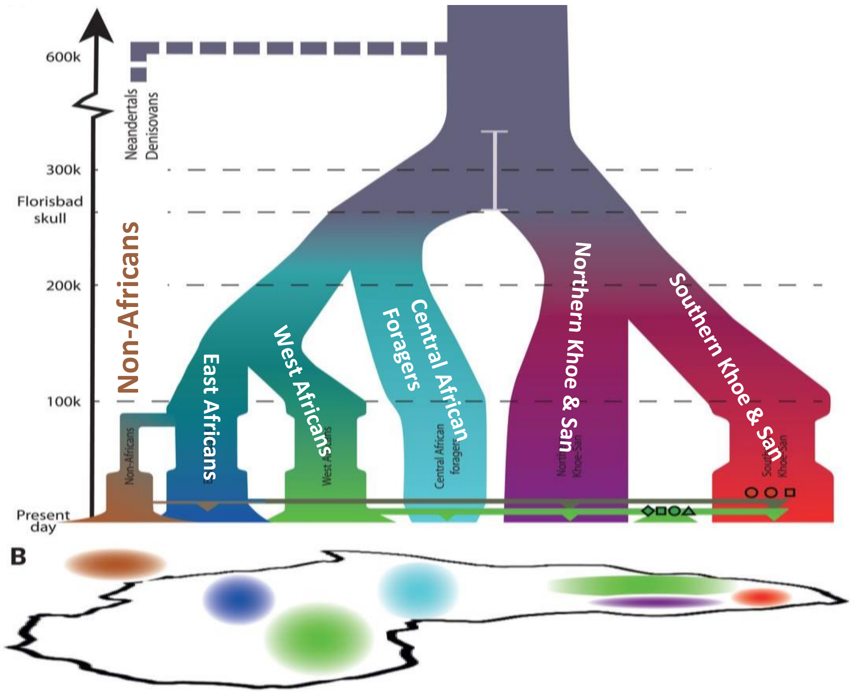


Nielsen, Nature 2017

European Ancestries

African Ancestries

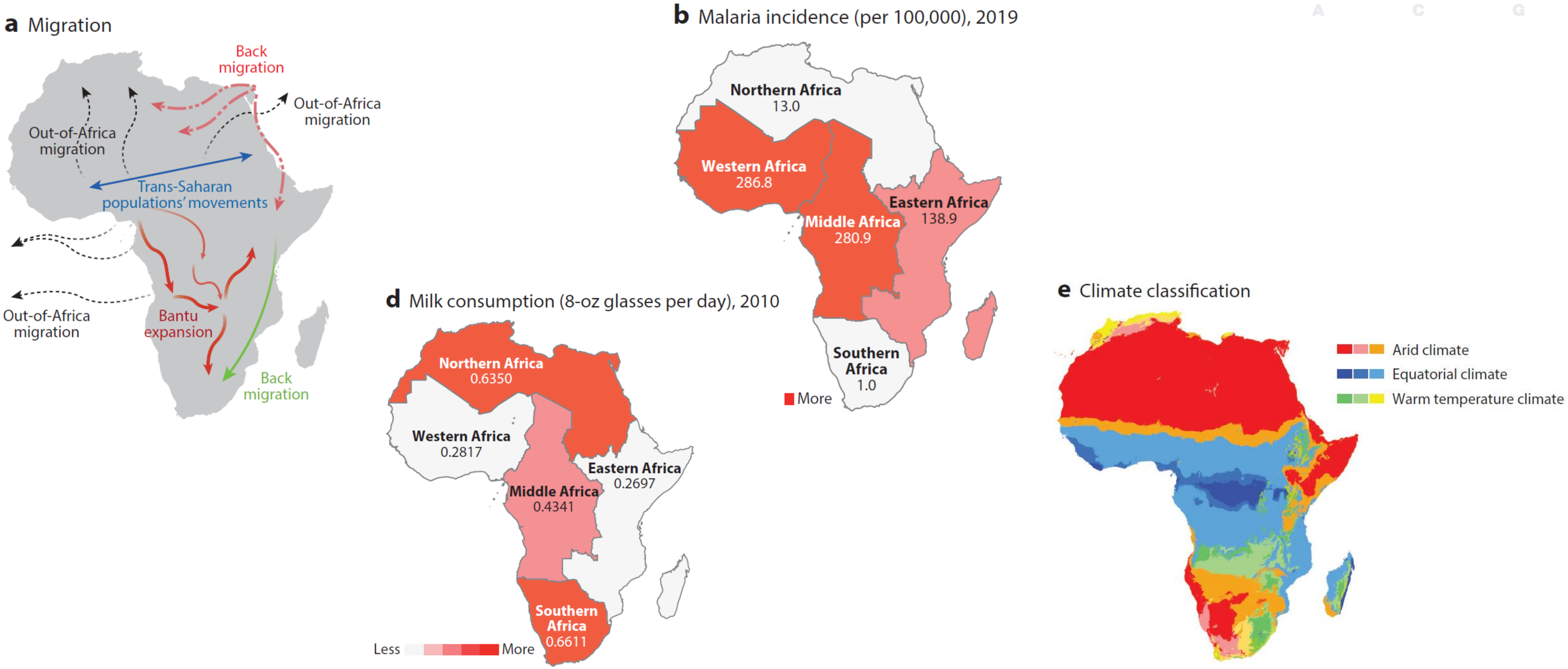
1000 Genomes Project Consortium, Nature. 2015



Schlebusch, Science, 2017 (adapted)

NIH
NHGRI

Drivers of Genetic Variation in Africa



a Migration

b Malaria incidence (per 100,000), 2019

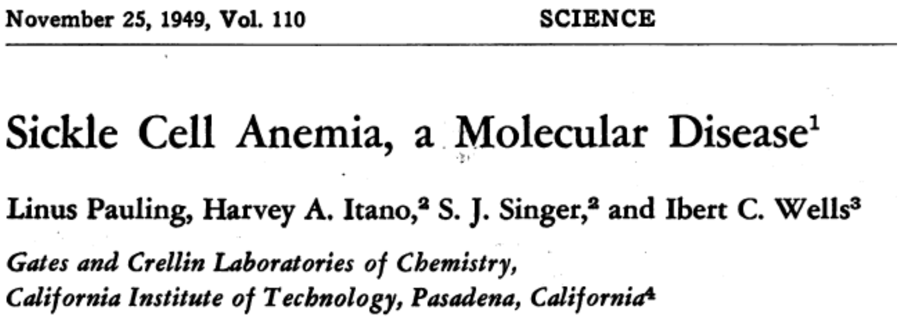
d Milk consumption (8-oz glasses per day), 2010

e Climate classification

Wonkam, et al, Ann Rev Gen Genom, 2022

African Diversity in Genetic Studies

- Individuals of African ancestry contribute significantly to gene mapping…

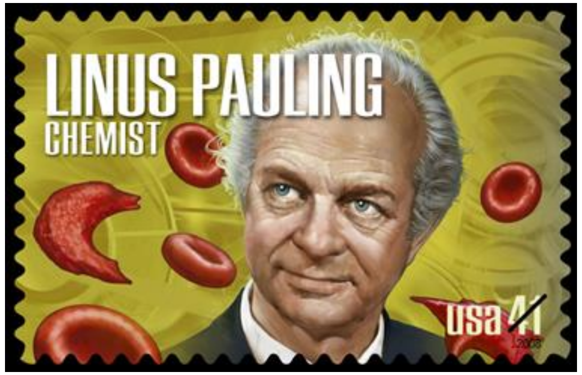


November 25, 1949, Vol. 110 SCIENCE

Sickle Cell Anemia, a Molecular Disease[1]

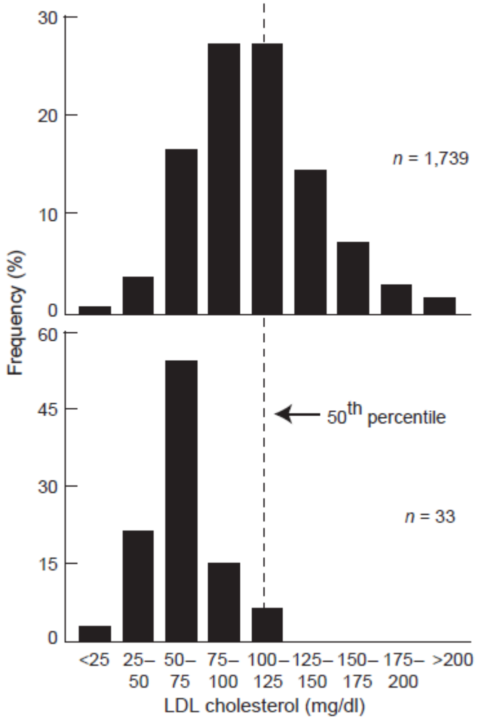
Linus Pauling, Harvey A. Itano,[2] S. J. Singer,[2] and Ibert C. Wells[3]

Gates and Crellin Laboratories of Chemistry,
California Institute of Technology, Pasadena, California[4]



Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in *PCSK9*

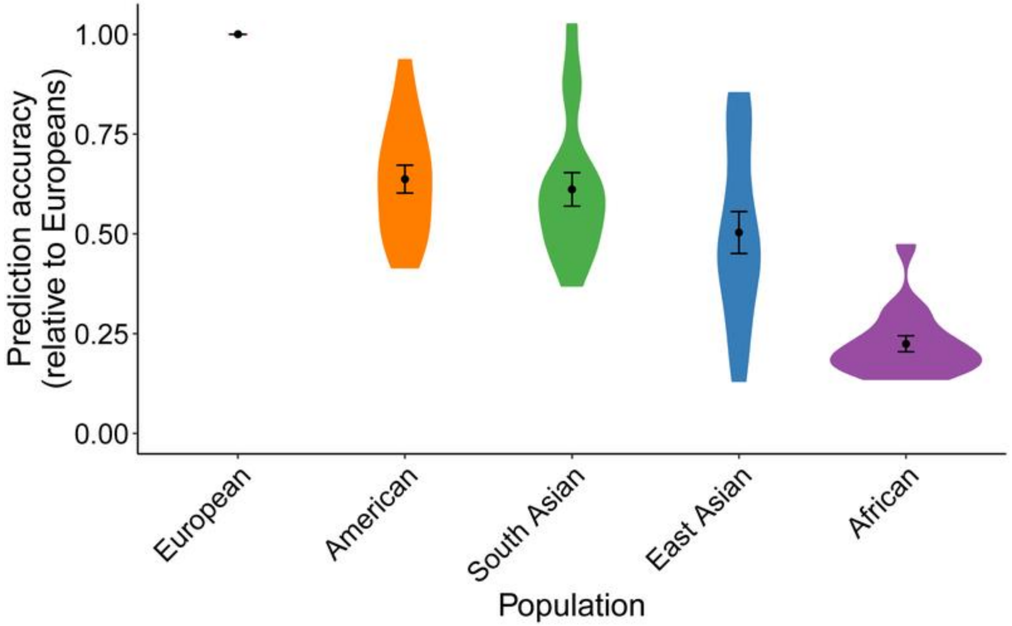
Jonathan Cohen[1–3], Alexander Pertsemlidis[2,3], Ingrid K Kotowski[4], Randall Graham[1], Christine Kim Garcia[1–3] & Helen H Hobbs[1–4]



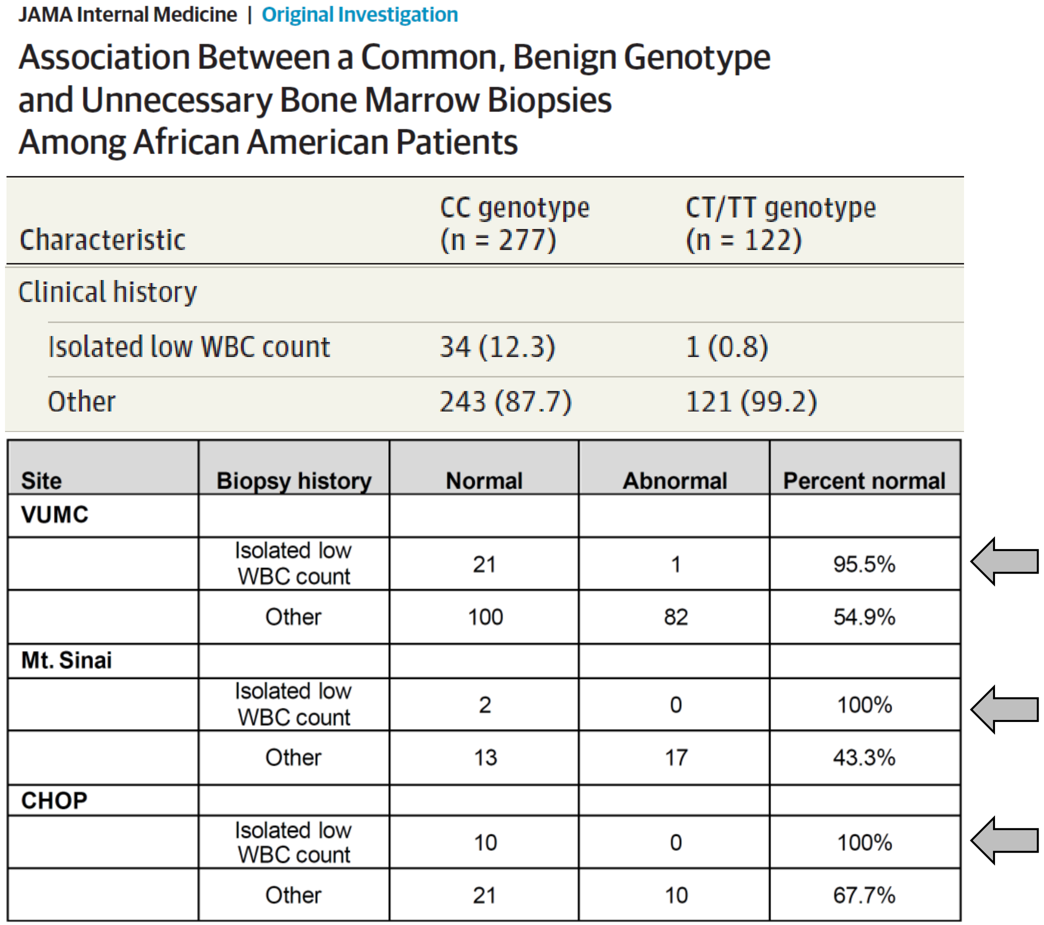
Drug
Evolocumab (Repatha)
Alirocumab (Praluent)
Bococizumab

African Diversity in Genetic Studies

- … but may not be able to benefit equitably from it….



Martin, Nature Genetics, 2019


JAMA Internal Medicine | Original Investigation

Association Between a Common, Benign Genotype and Unnecessary Bone Marrow Biopsies Among African American Patients

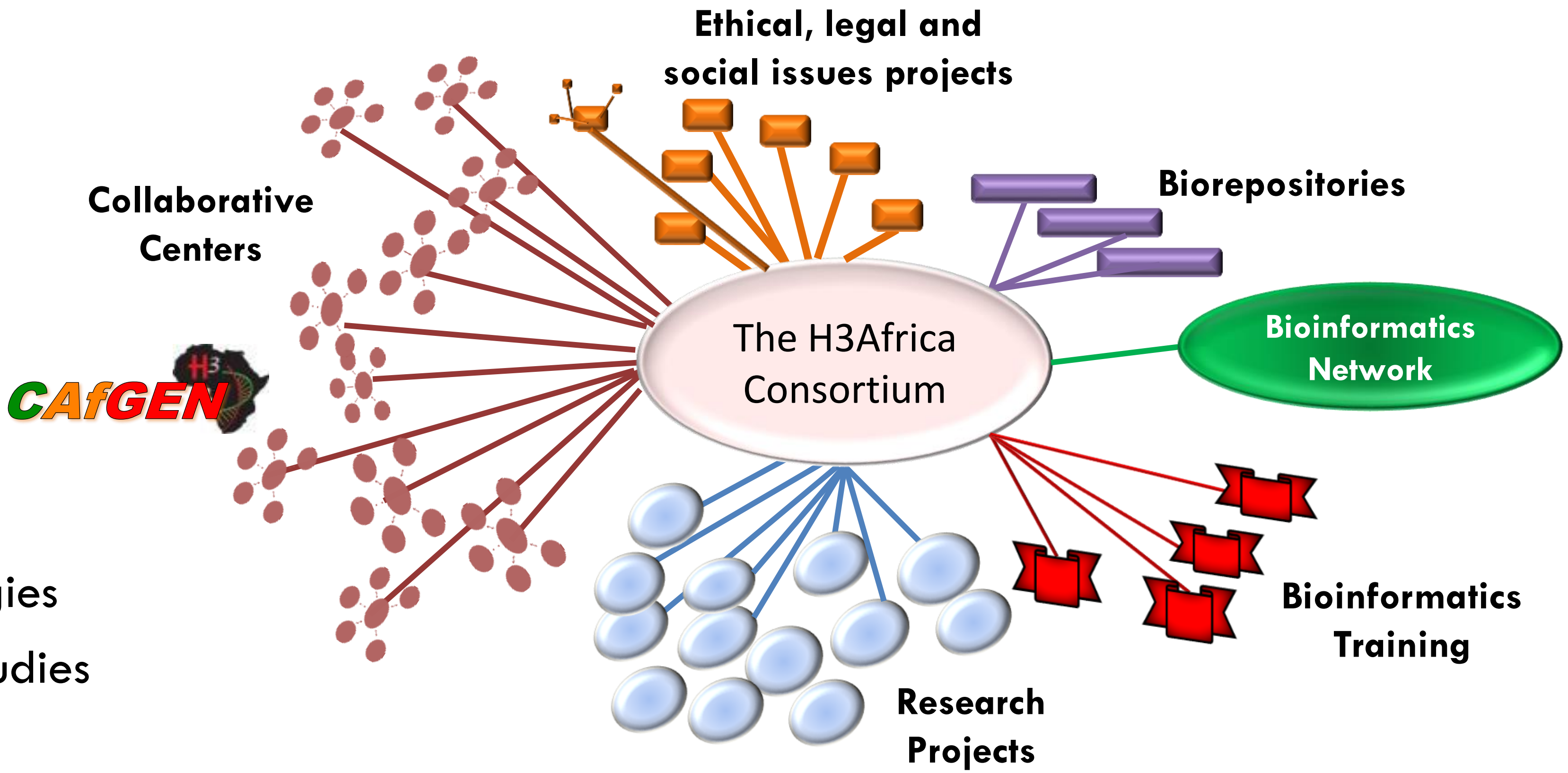
Characteristic	CC genotype (n = 277)	CT/TT genotype (n = 122)
Clinical history		
Isolated low WBC count	34 (12.3)	1 (0.8)
Other	243 (87.7)	121 (99.2)

Site	Biopsy history	Normal	Abnormal	Percent normal
VUMC				
	Isolated low WBC count	21	1	95.5%
	Other	100	82	54.9%
Mt. Sinai				
	Isolated low WBC count	2	0	100%
	Other	13	17	43.3%
CHOP				
	Isolated low WBC count	10	0	100%
	Other	21	10	67.7%

Van Dreist, JAMA Int Med, 2021

Human Heredity and Health in Africa (H3Africa) Consortium

- Access to new genomic technologies
- Integrate genomic and clinical studies
- Facilitate training at all levels
- Establish research infrastructure



Ethical, legal and social issues projects

Collaborative Centers

Biorepositories

CAfGEN

The H3Africa Consortium

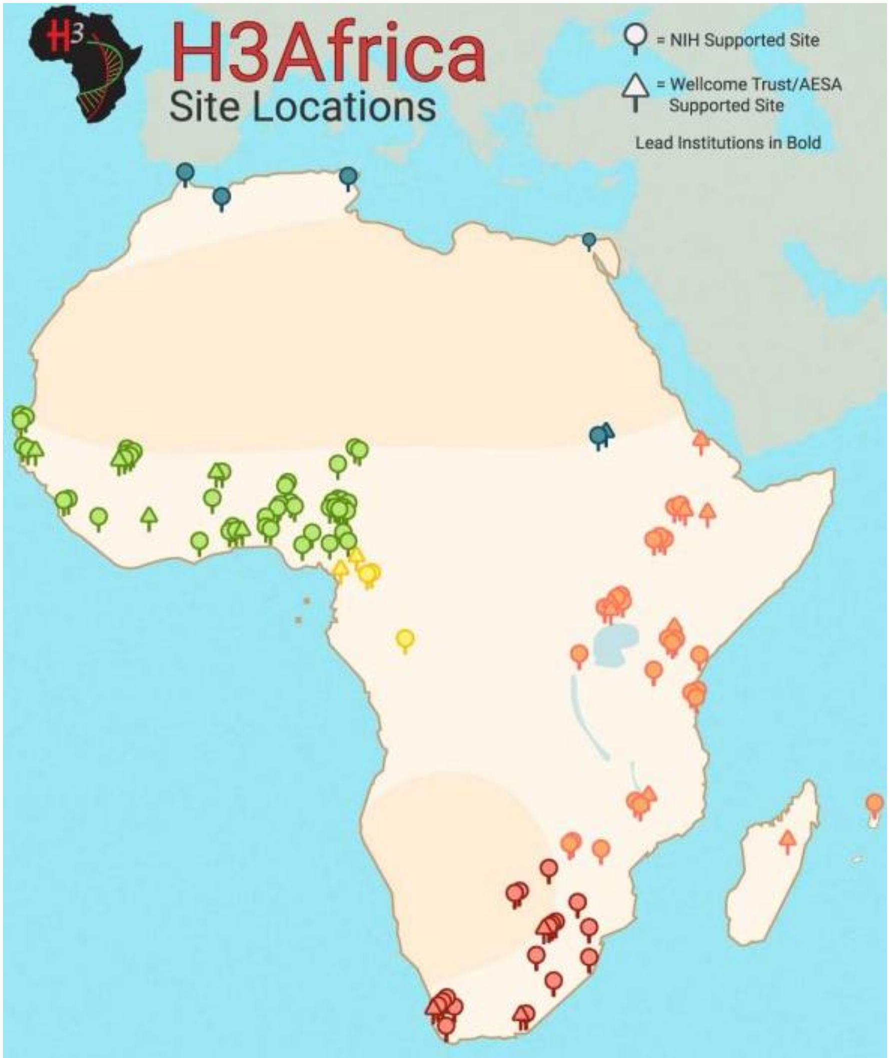
Bioinformatics Network

Bioinformatics Training

Research Projects

<u>Pricipal Investigators must be from an African country</u>

Human Heredity and Health in Africa (H3Africa) Consortium




Courtesy of the AWIGEN project

34 African Countries involved

51 H3Africa Projects supported

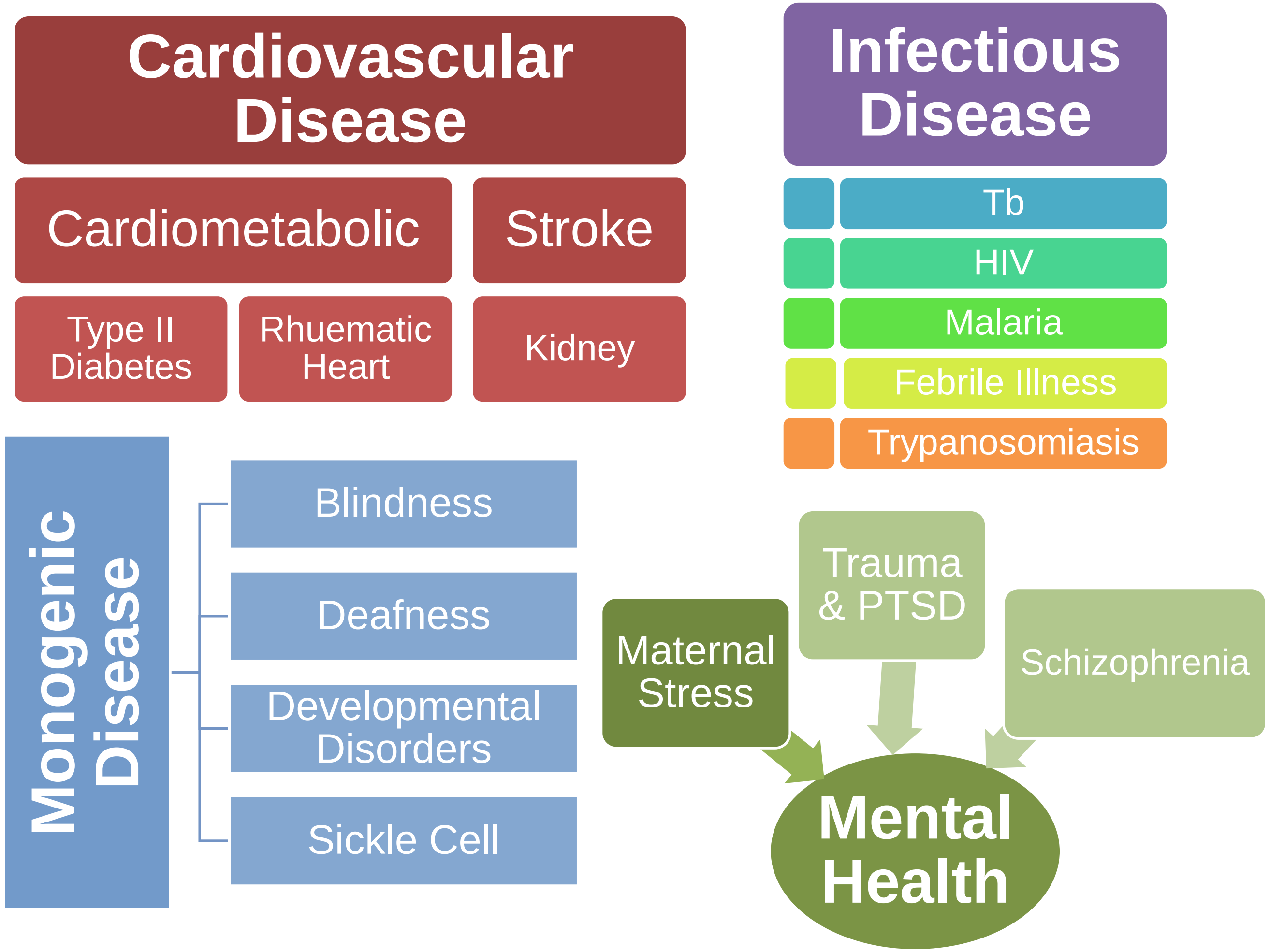
27 Policies, Guidelines & SOPs

1528 Number of members

H3Africa human genetics projects span a wide spectrum of diseases

118065

Research Participants Recruited



Cardiovascular Disease

- Cardiometabolic
  - Type II Diabetes
  - Rhuematic Heart
- Stroke
  - Kidney

Infectious Disease

- Tb
- HIV
- Malaria
- Febrile Illness
- Trypanosomiasis

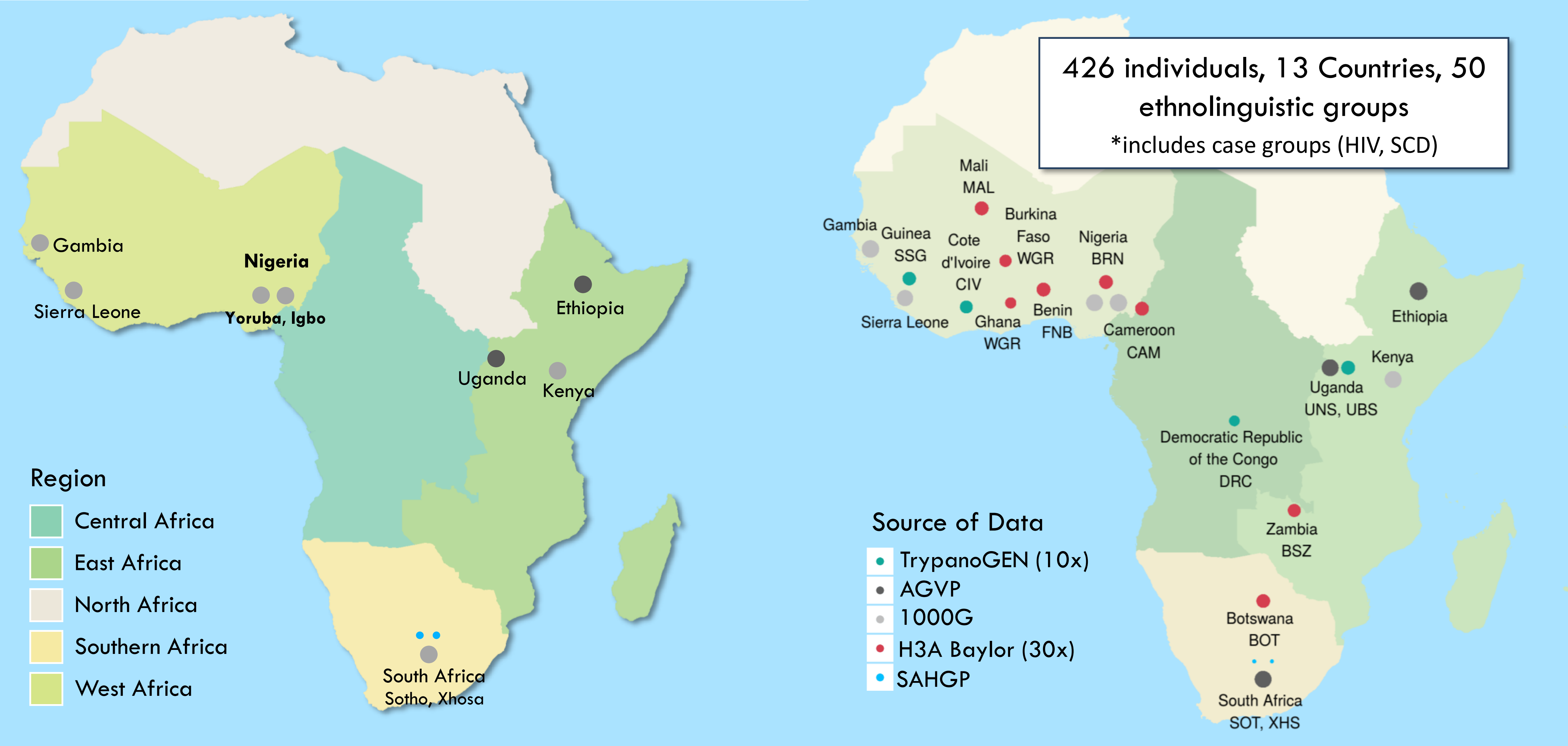
Monogenic Disease

- Blindness
- Deafness
- Developmental Disorders
- Sickle Cell

Mental Health

- Maternal Stress
- Trauma & PTSD
- Schizophrenia

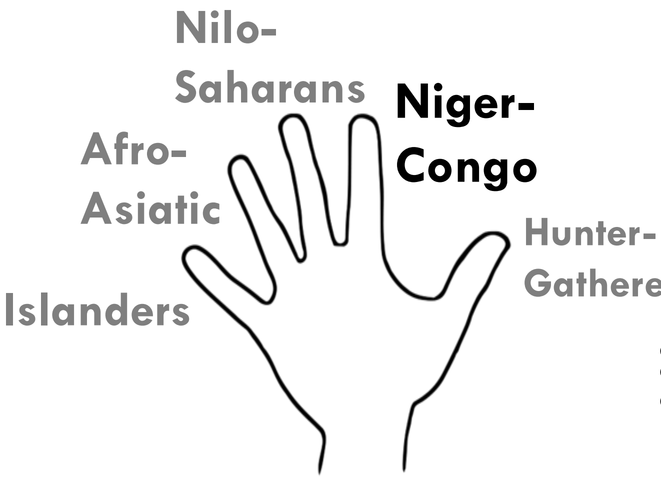
Whole genome sequencing efforts in Africa

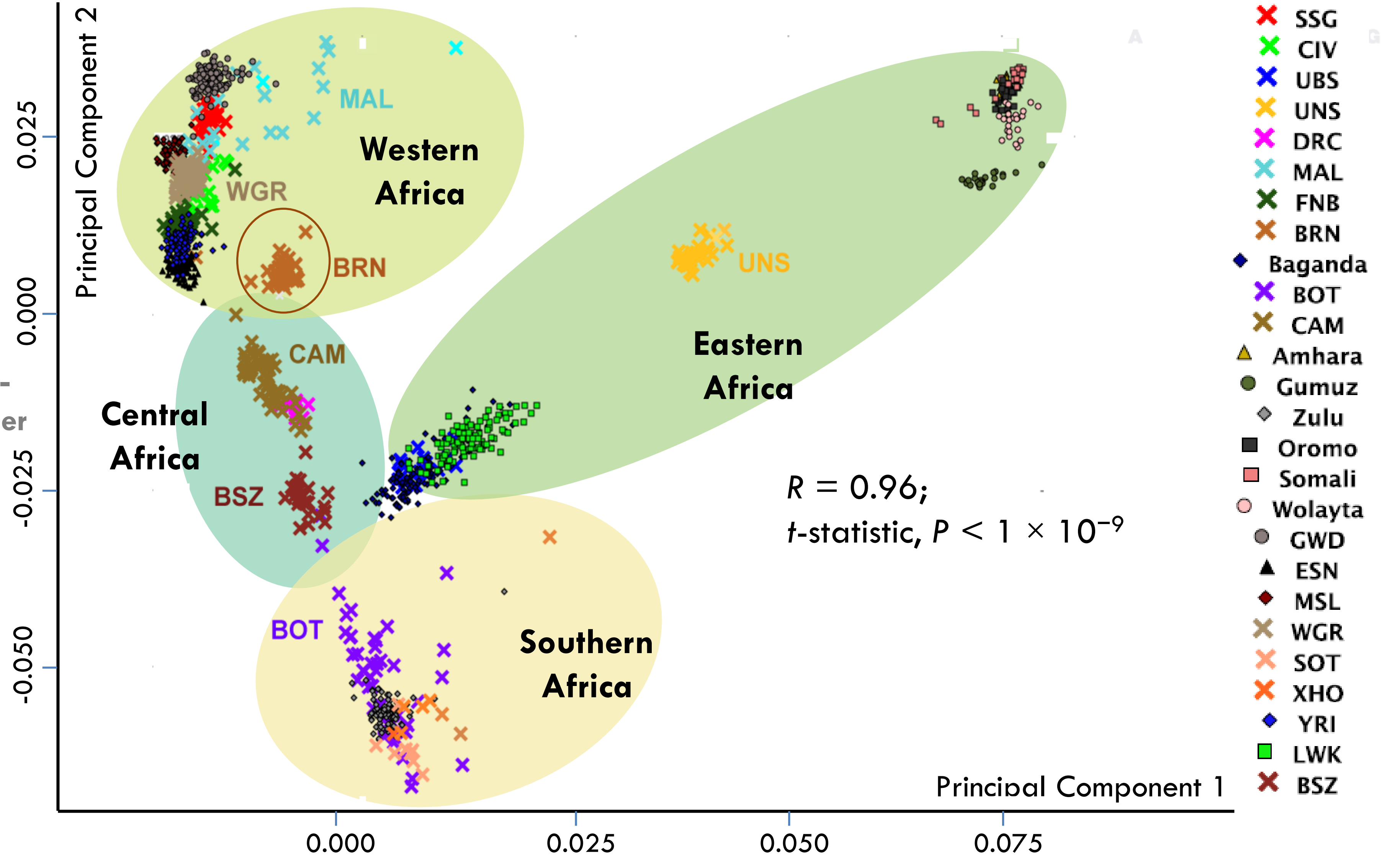


426 individuals, 13 Countries, 50 ethnolinguistic groups

*includes case groups (HIV, SCD)

WGS in ‘new’ African populations expands the genomic landscape



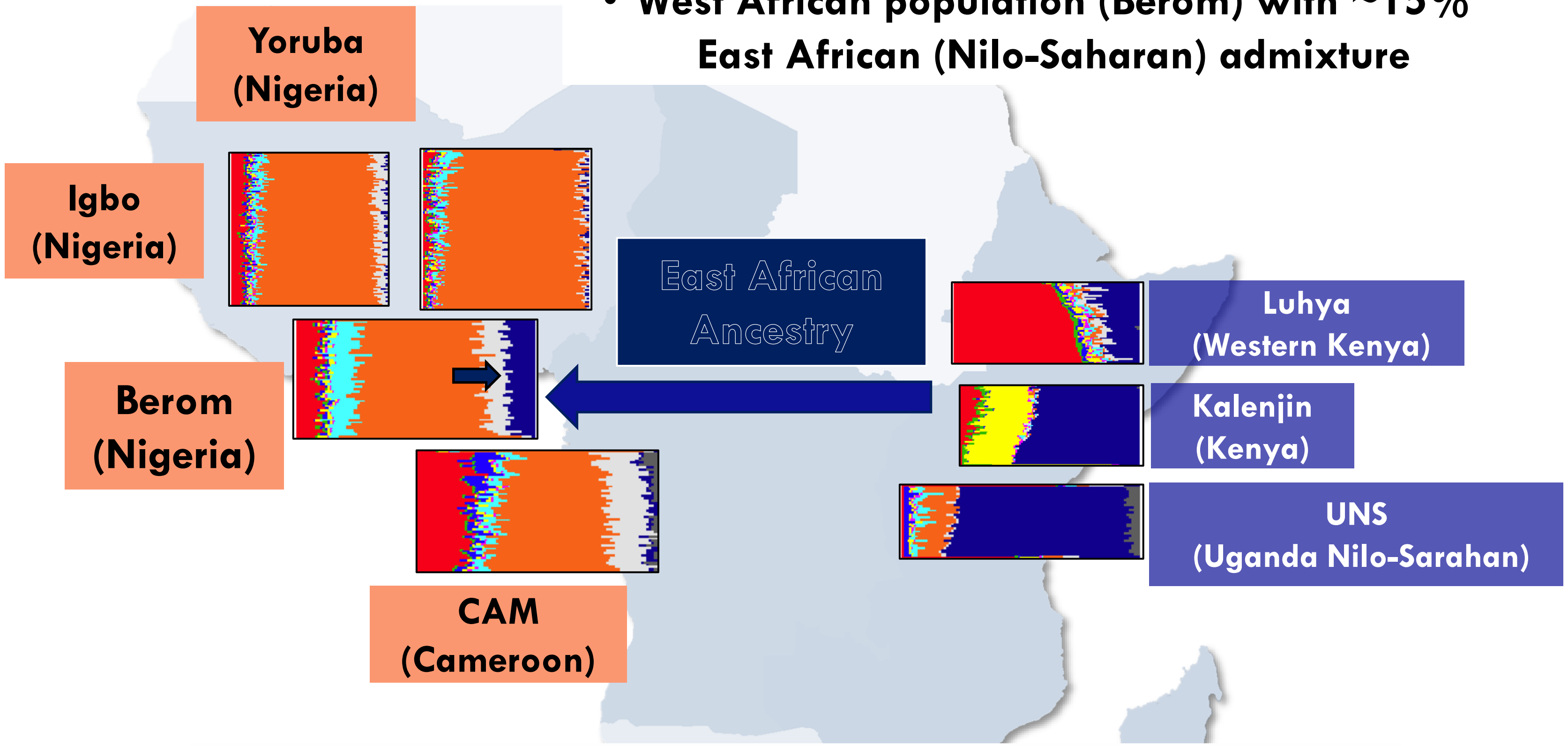


$R = 0.96$;
t-statistic, $P < 1 \times 10^{-9}$

Choudhury, Nature, Oct 2020

African populations show complex patterns of admixture

- **West African population (Berom) with ~15% East African (Nilo-Saharan) admixture**



Yoruba (Nigeria)

Igbo (Nigeria)

Berom (Nigeria)

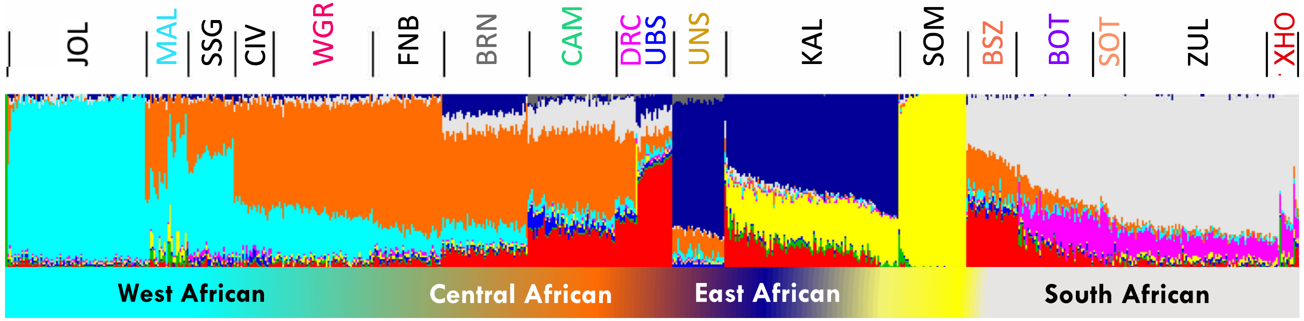
CAM (Cameroon)

East African Ancestry

Luhya (Western Kenya)

Kalenjin (Kenya)

UNS (Uganda Nilo-Sarahan)

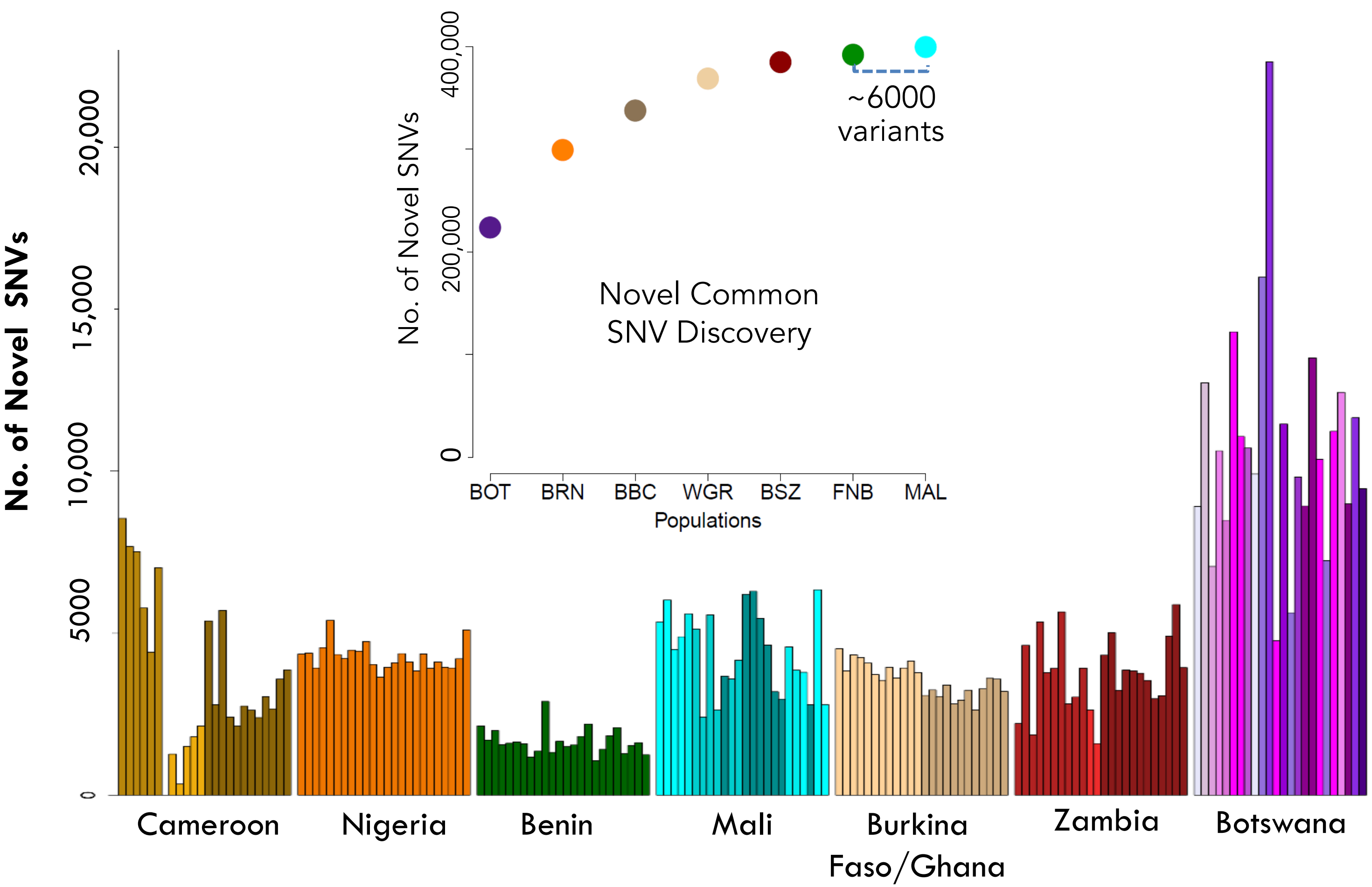


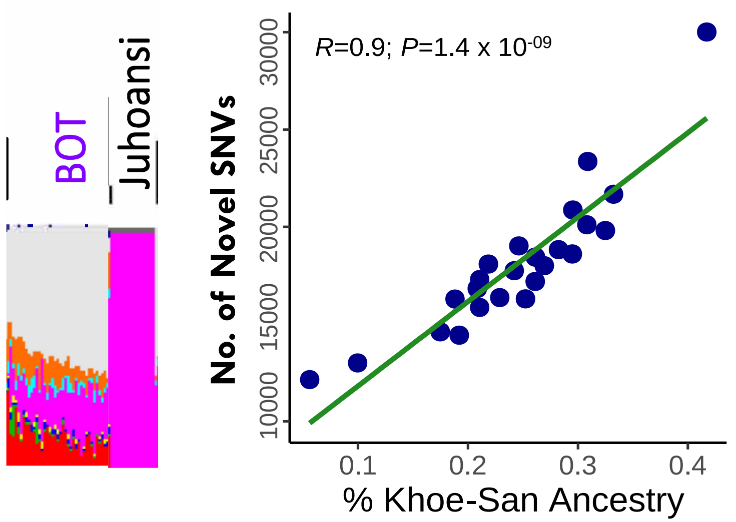
Choudhury, Nature, Oct 2020

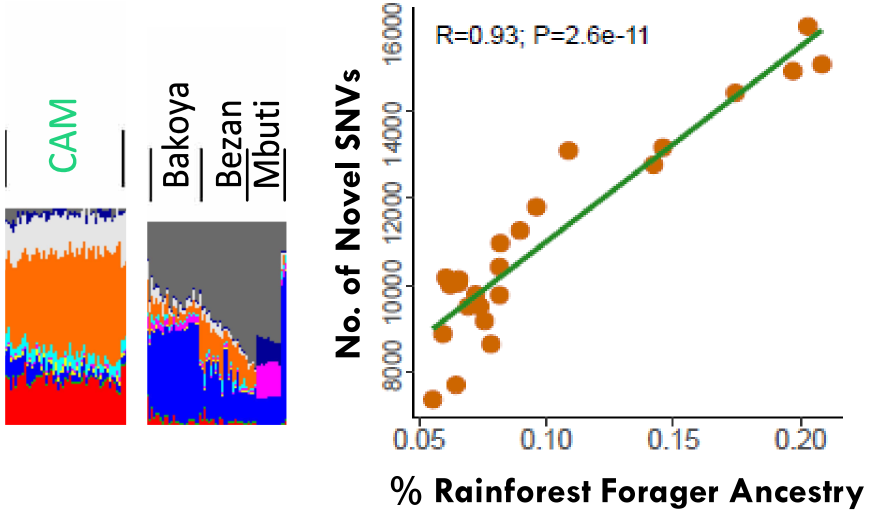
Novel variant discovery in African genomes

~3.4 million ‘novel’ single nucleotide variants (SNVs)

- Strong correlation with new previously unsampled groups

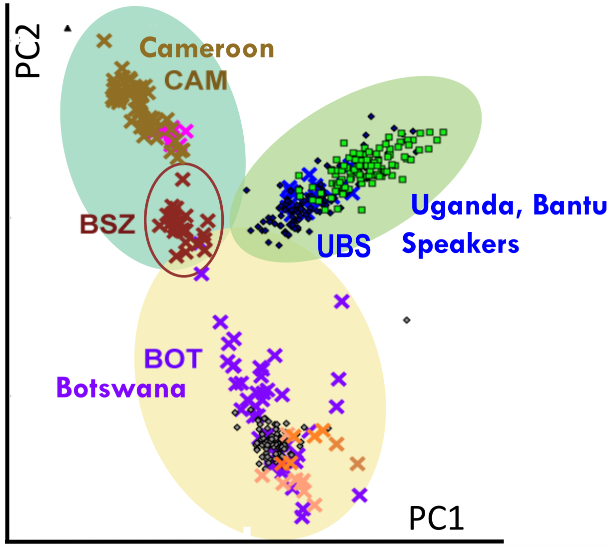


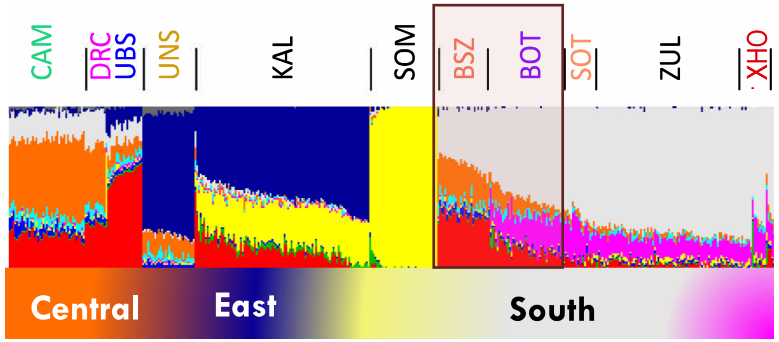


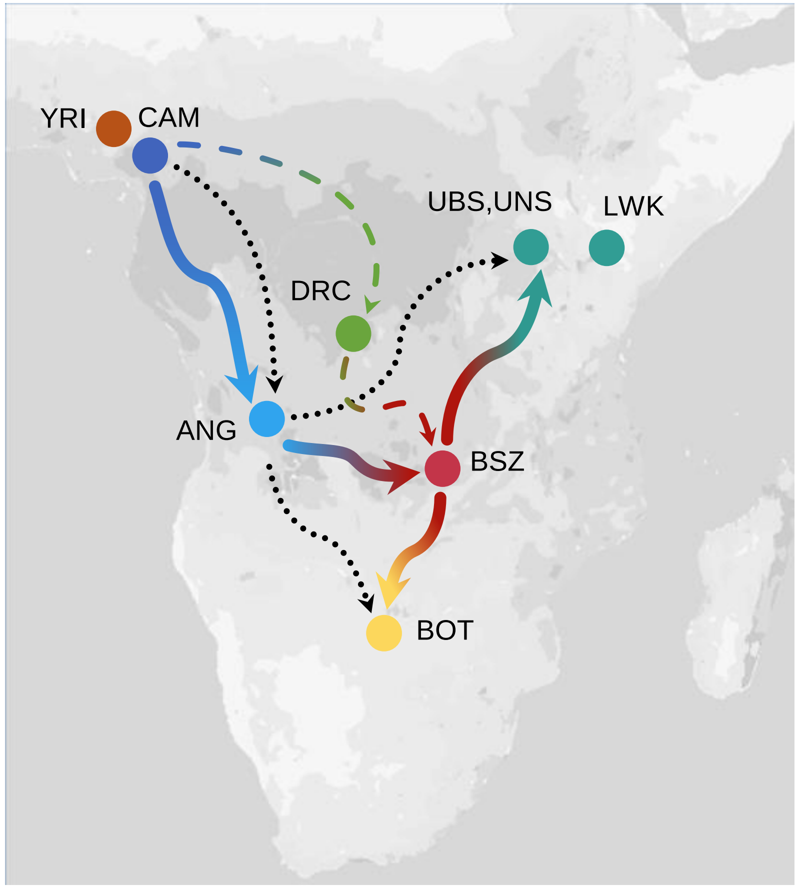


Choudhury, Nature, Oct 2020

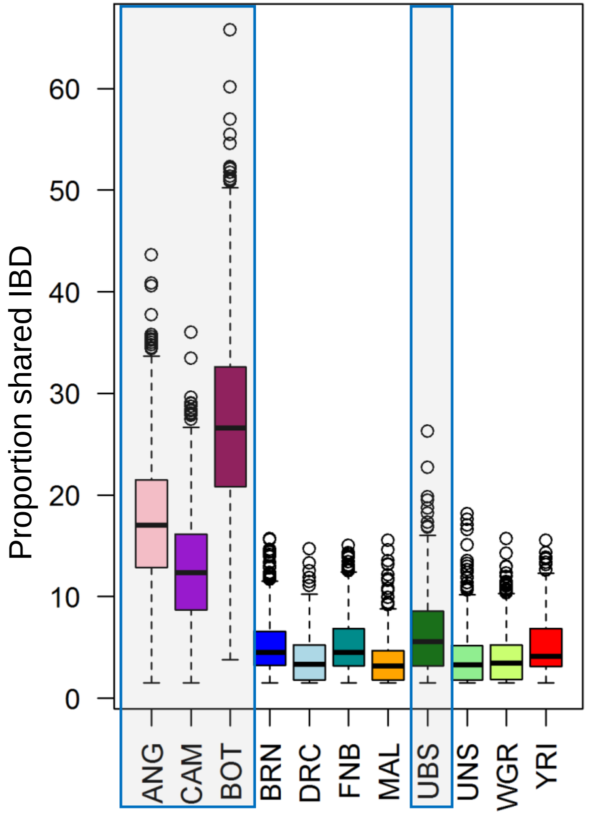
Unsampled populations inform historical migrations







Bantu-speaker Migration Routes

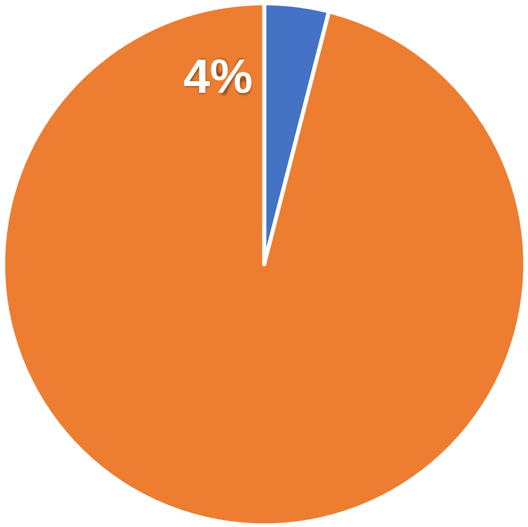


Bantu Speakers from Zambia (**BSZ**) share most with Botswana (**BOT**), Angola (**ANG**), and Uganda (**UBS**)

Choudhury, Nature, Oct 2020

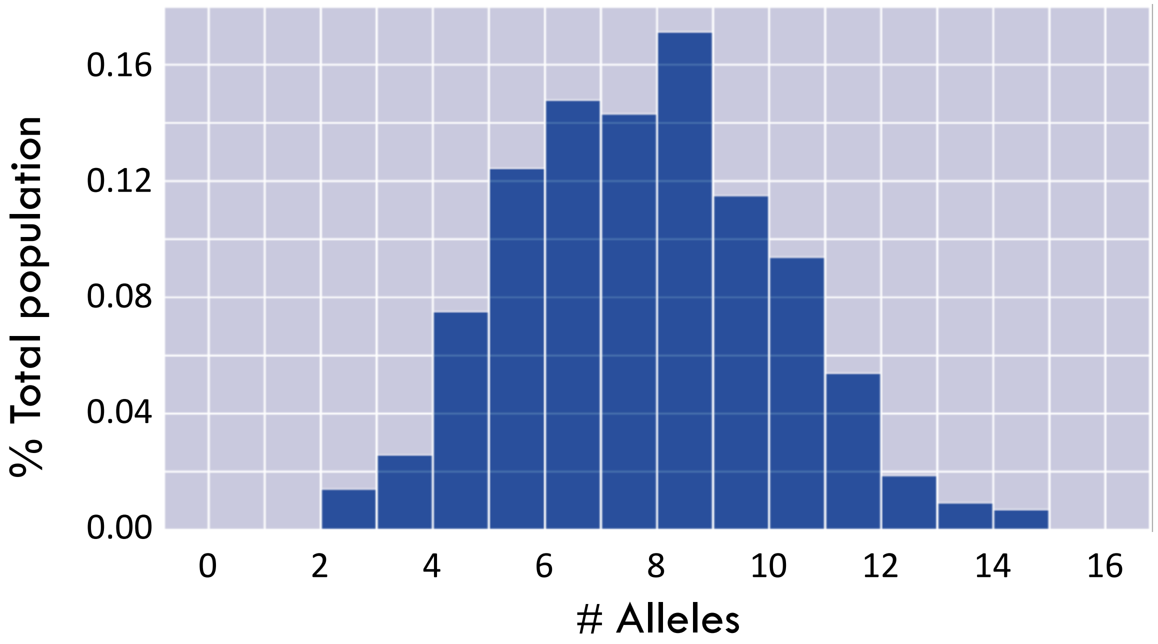
African genomes inform medically relevant variation

- American College of Medical Genetics and Genomics (**ACMG**) list of medically-actionable genes
- **ClinVar** - Curated archive of likely variant significance from submitted findings



ATP7B
BRCA1
BRCA2
RYR1
KCNQ1
PKP2

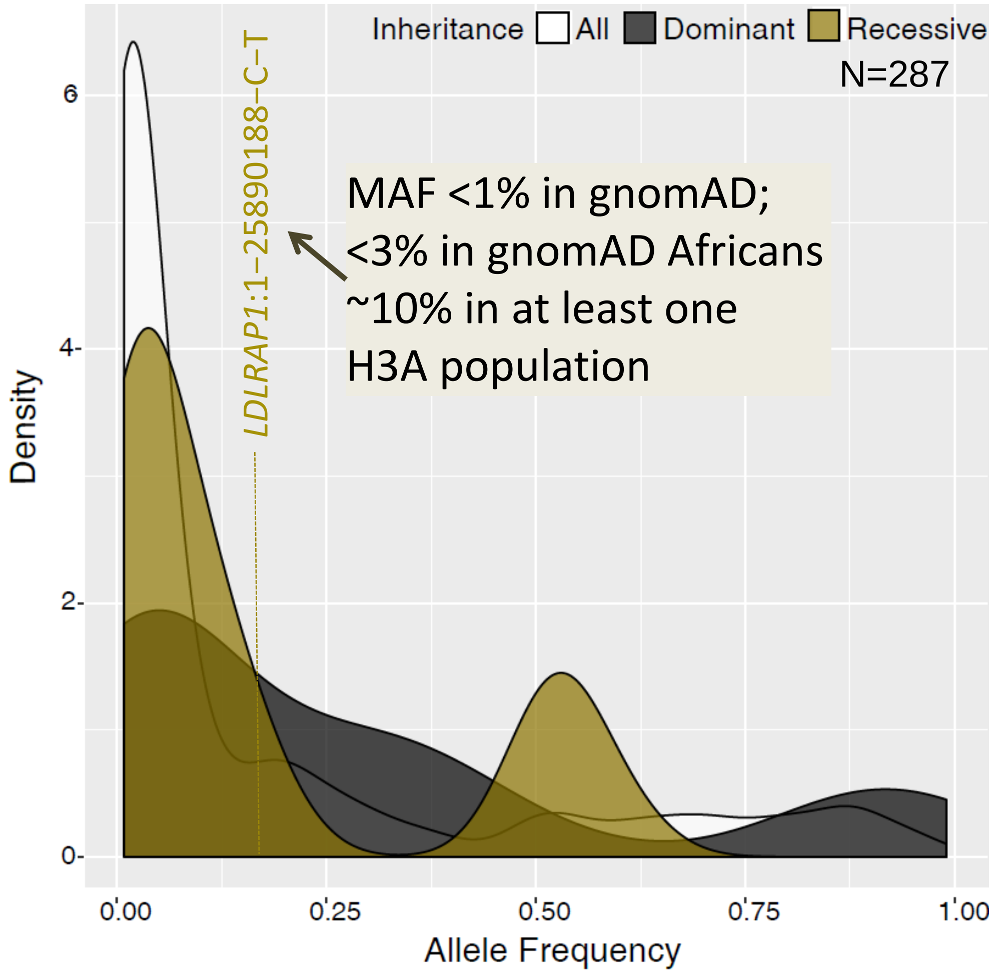
Only 8 ACMG variants in 8 Individuals



High individual burden of ClinVar pathogenic variants

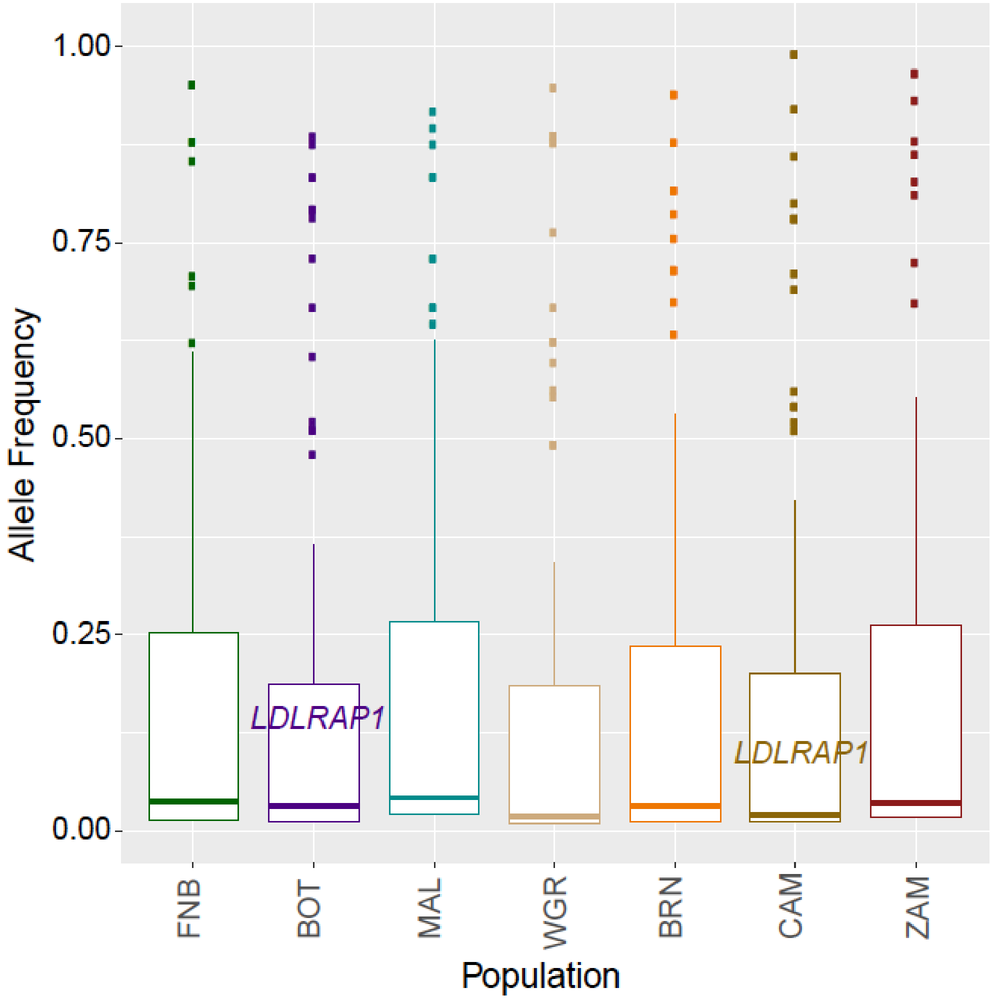
Choudhury, Nature, Oct 2020

Pathogenic SNVs in ClinVar have different frequencies across Africa



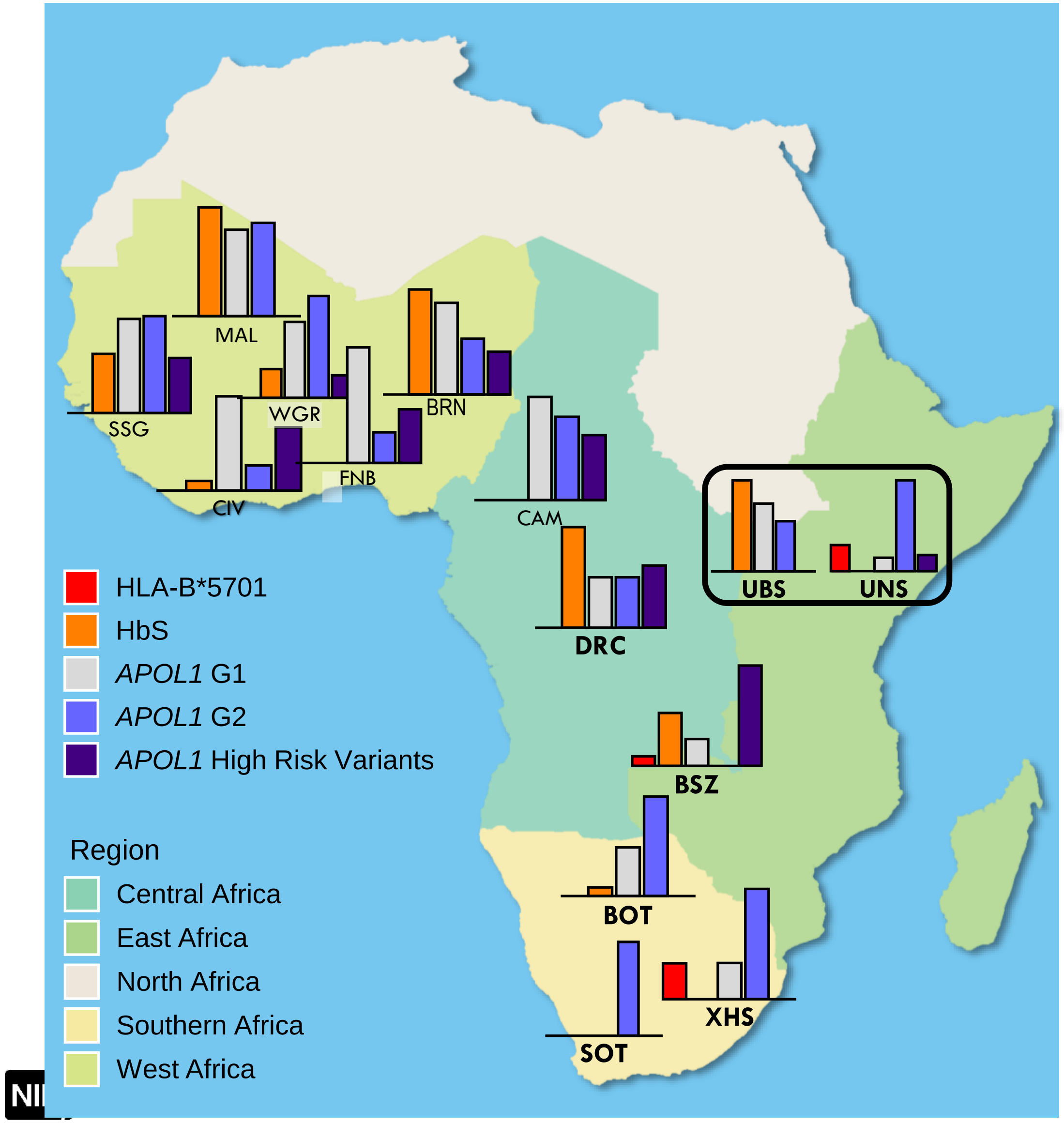
MAF <1% in gnomAD;
<3% in gnomAD Africans
~10% in at least one H3A population

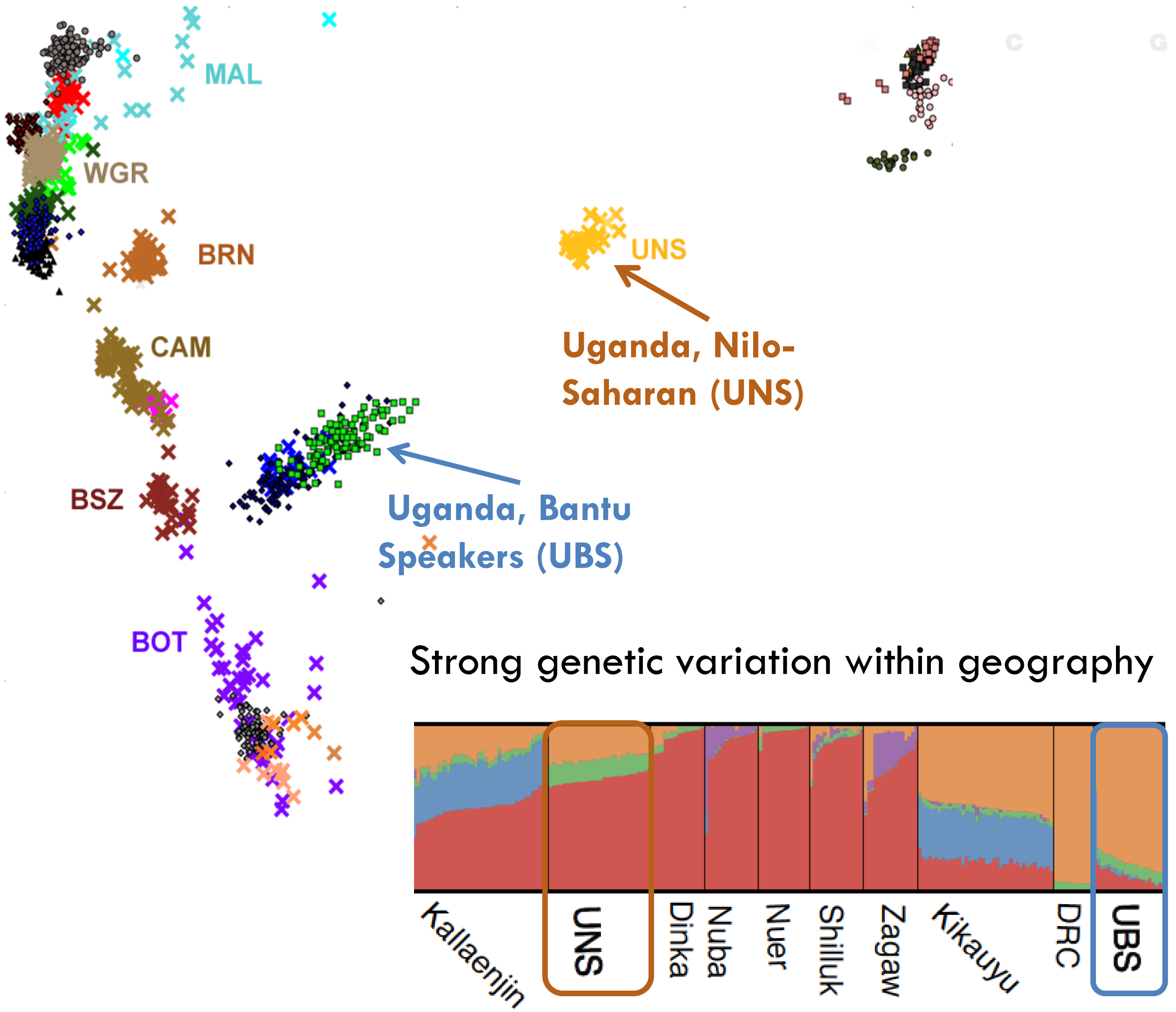
LDLRAP1 - Autosomal recessive hypercholesteremia



Choudhury, Nature, Oct 2020

African medically-relevant variants





Uganda, Nilo-Saharan (UNS)

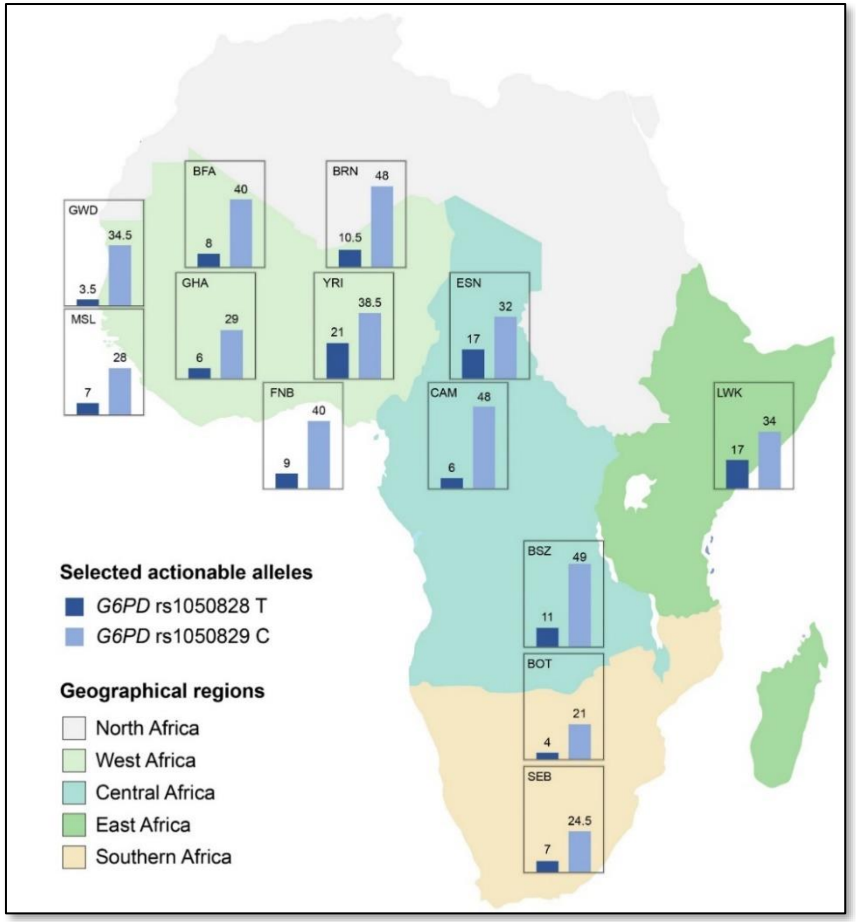
Uganda, Bantu Speakers (UBS)

Strong genetic variation within geography

Choudhury, Nature, Oct 2020

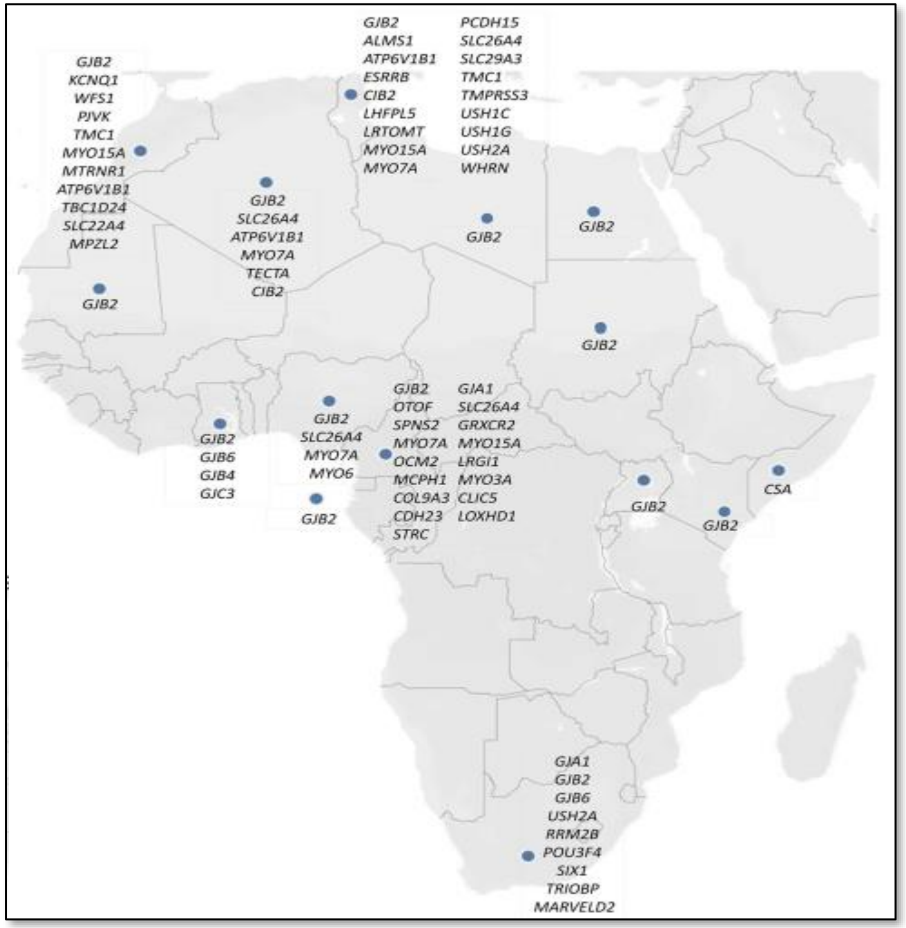
African medically-relevant variants

- The frequency of G6PD deficiency varies widely across the continent



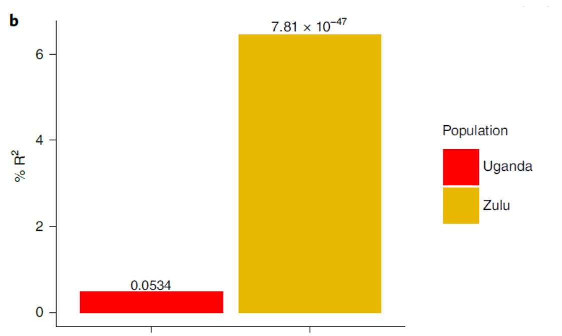
Sitabule, Clin Pharm & Ther, 2022

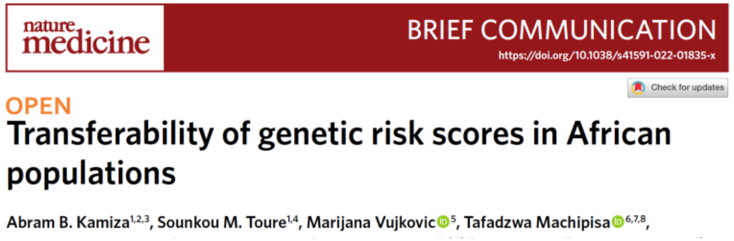
- The most common cause of non-syndromic hearing loss outside of Africa (*GJB2*) is not always the most common across Africa



Adadey, Hum Gen, 2022

Implications for common variant mapping across Africa




nature medicine

BRIEF COMMUNICATION

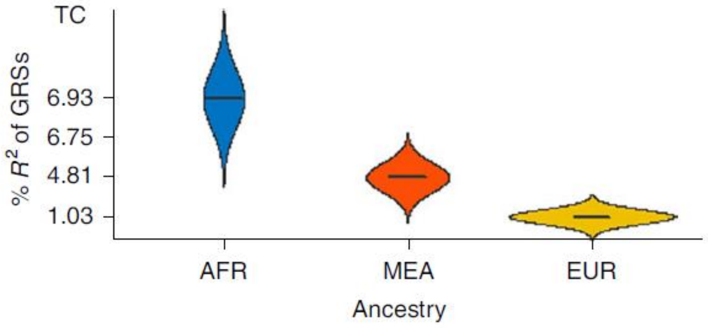
https://doi.org/10.1038/s41591-022-01835-x

Check for updates

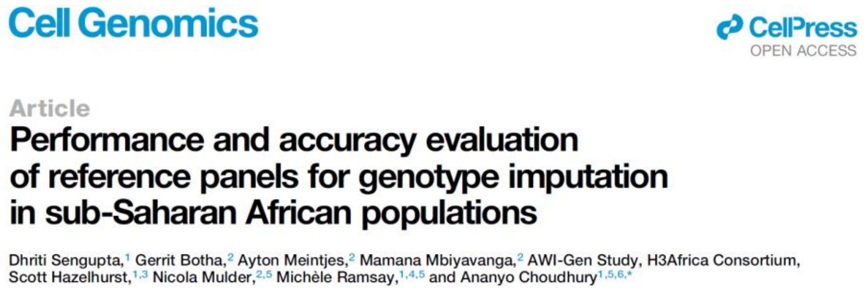
OPEN

Transferability of genetic risk scores in African populations

Abram B. Kamiza[1,2,3], Sounkou M. Toure[1,4], Marijana Vujkovic[5], Tafadzwa Machipisa[6,7,8],



- In Africans, genetic risk scores created from African GWAS work better than European, but not everywhere in Africa

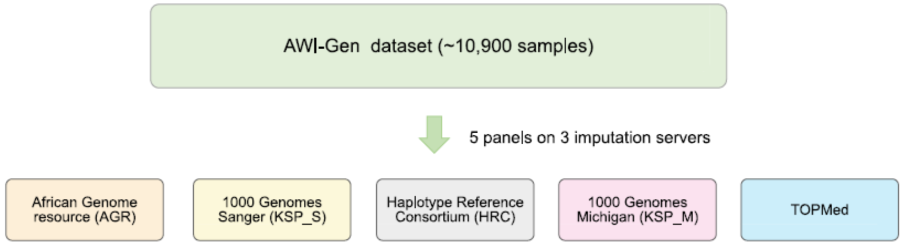

Cell Genomics

CellPress OPEN ACCESS

Article

Performance and accuracy evaluation of reference panels for genotype imputation in sub-Saharan African populations

Dhriti Sengupta,[1] Gerrit Botha,[2] Ayton Meintjes,[2] Mamana Mbiyavanga,[2] AWI-Gen Study, H3Africa Consortium, Scott Hazelhurst,[1,3] Nicola Mulder,[2,5] Michèle Ramsay,[1,4,5] and Ananyo Choudhury[1,5,6,*]


AWI-Gen dataset (~10,900 samples)

5 panels on 3 imputation servers

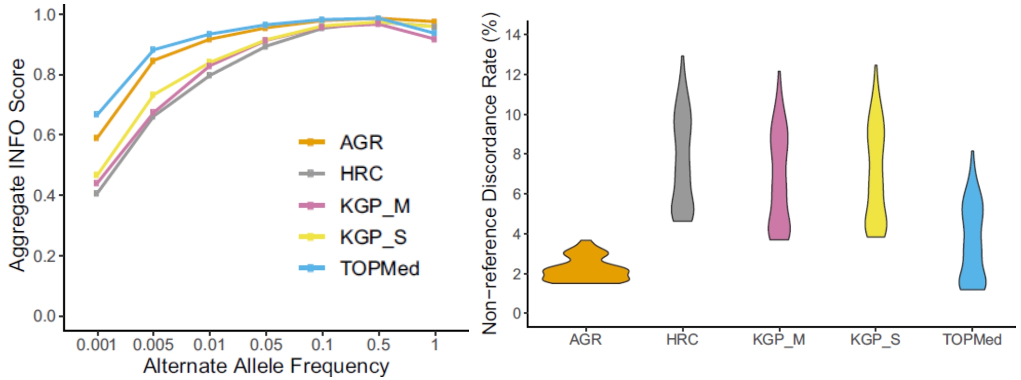
African Genome resource (AGR)

1000 Genomes Sanger (KSP_S)

Haplotype Reference Consortium (HRC)

1000 Genomes Michigan (KSP_M)

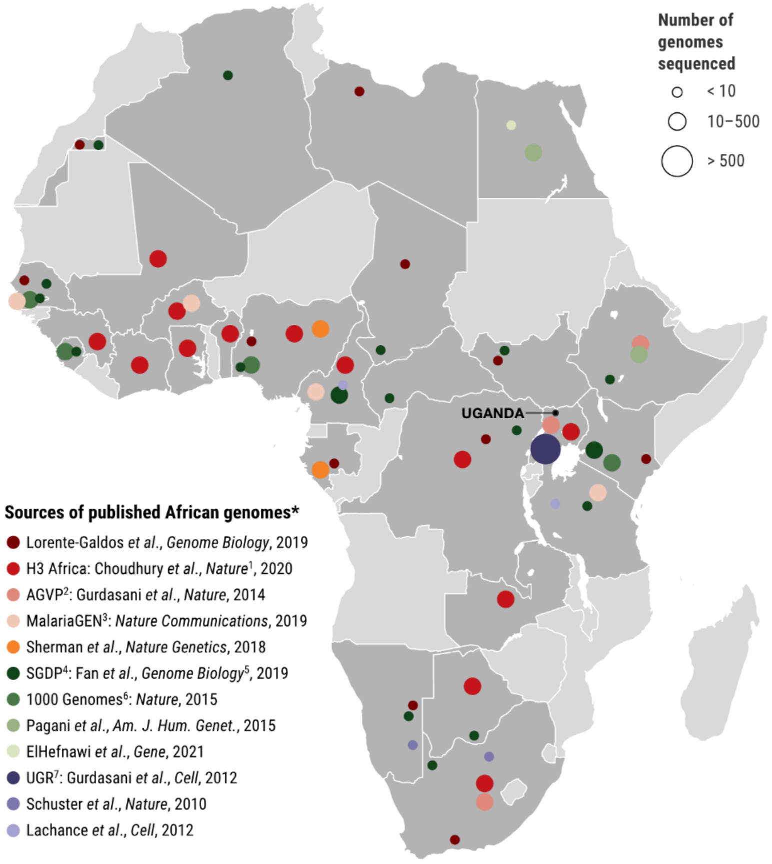
TOPMed



- African imputation panel is smaller than best African diaspora panel, but is more accurate

More work to be done…

Kai Kruse – The true size of Africa



Science News, Elizabeth Pennisi, 2021

Setting the agenda in research

Comment

Christian Happi at Redeemer's University in Ede, Nigeria, plans to sequence human genomes.

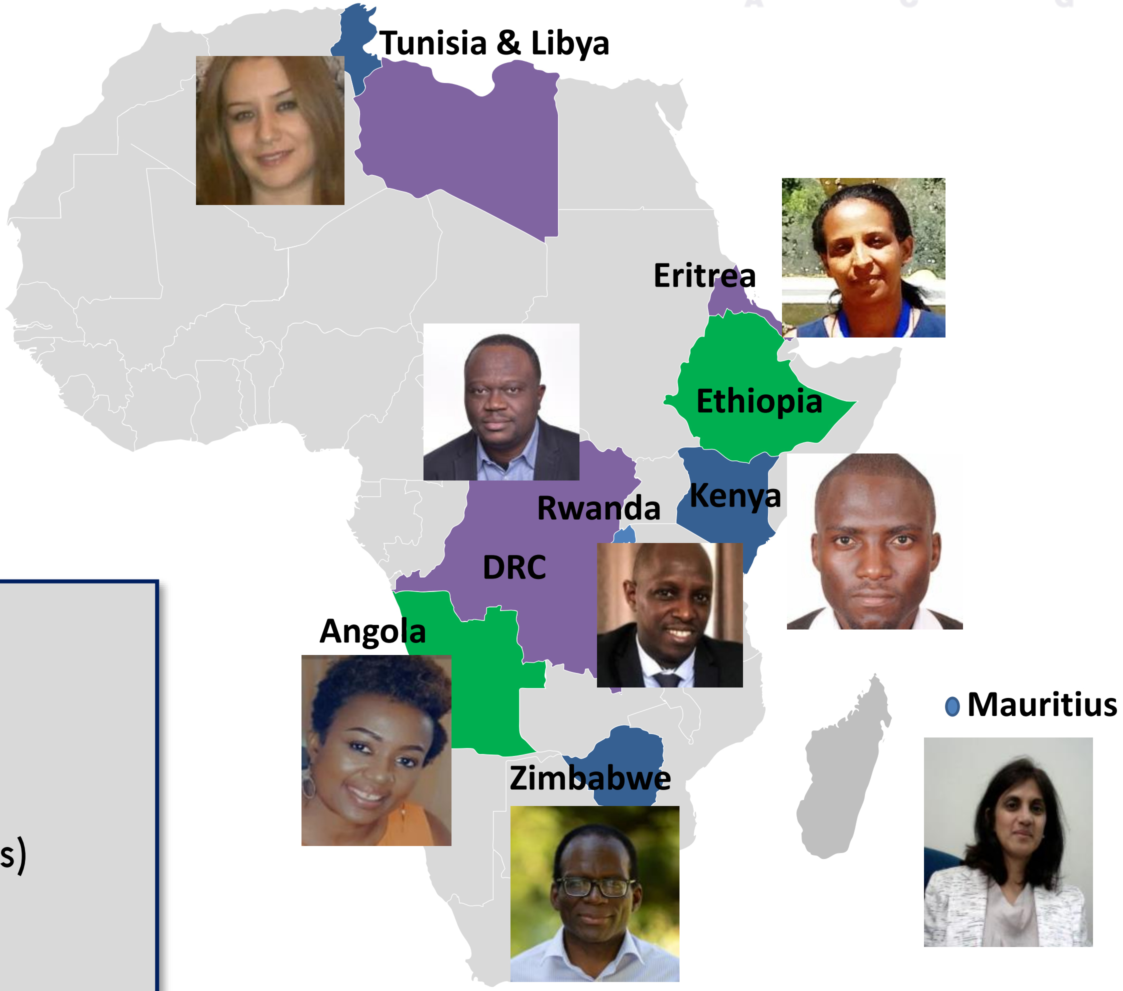
Sequence three million genomes across Africa

Ambroise Wonkam

Assessing Genomic Diversity in Africa - AGenDA

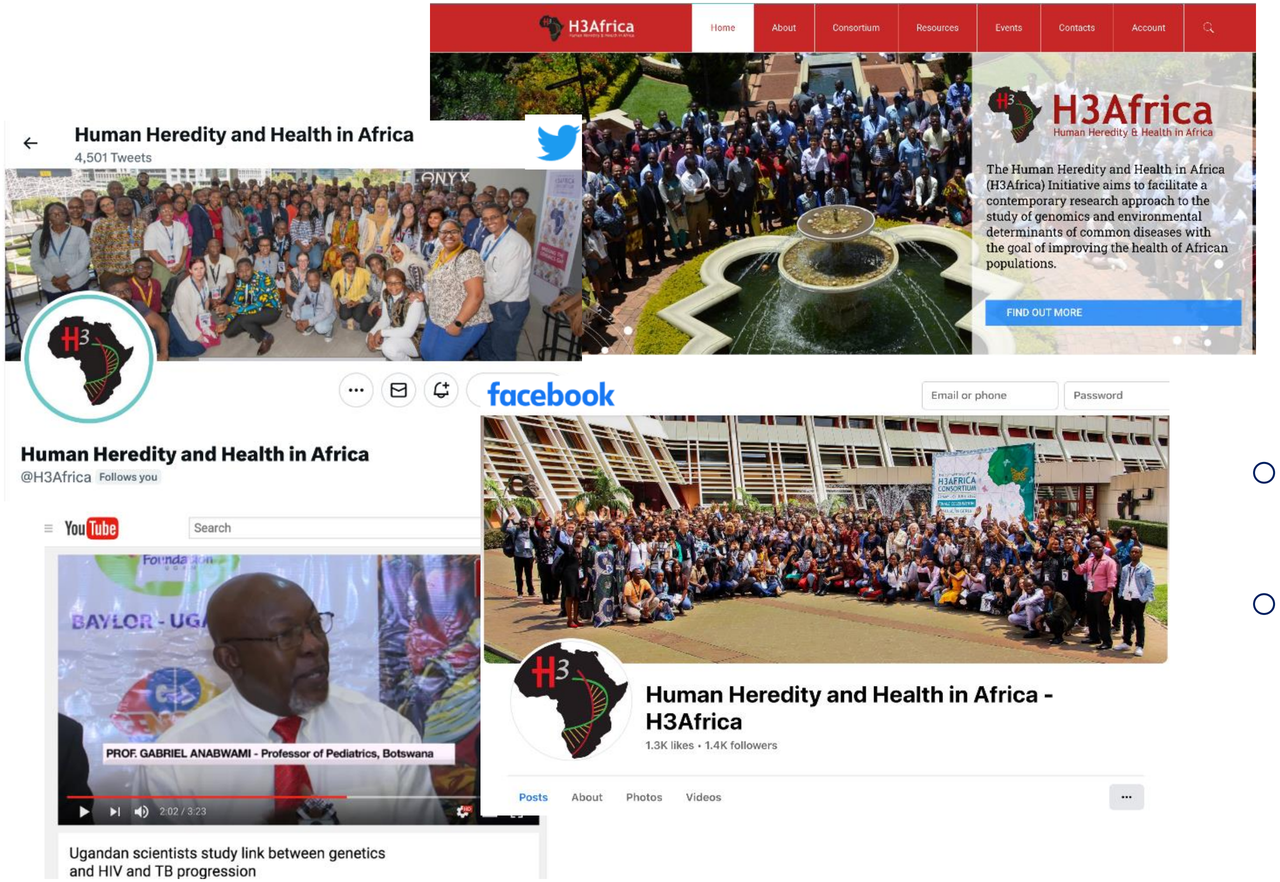
- Sequence ~1200+ from under-represented African groups
- Make deidentified data available
- Integrate new data to reference panels

- **10** Ethics Committee approvals
- **1** Partnership agreement (Illumina)
- **1** MTA with Illumina (Export permit)
- **7** MTA with study sites (Import permits)
- **1** Data Governance and Partnership agreement with all 10 groups



Laying the Groundwork

Media & Education



H3Africa
Human Heredity & Health in Africa

The Human Heredity and Health in Africa (H3Africa) Initiative aims to facilitate a contemporary research approach to the study of genomics and environmental determinants of common diseases with the goal of improving the health of African populations.

Human Heredity and Health in Africa
4,501 Tweets

Human Heredity and Health in Africa
@H3Africa Follows you

PROF. GABRIEL ANABWAMI - Professor of Pediatrics, Botswana

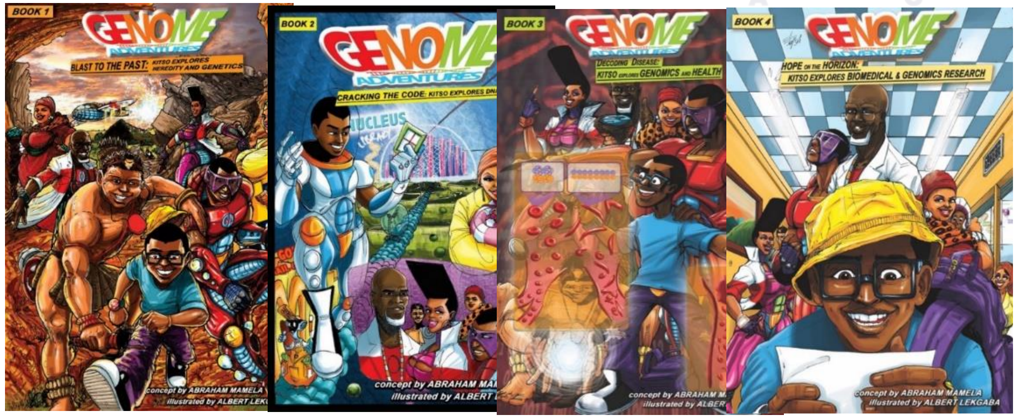
Ugandan scientists study link between genetics and HIV and TB progression

Human Heredity and Health in Africa - H3Africa
1.3K likes • 1.4K followers

Harvard Law & International Development Society
a community dedicated to challenges at the intersection of law, policy and development

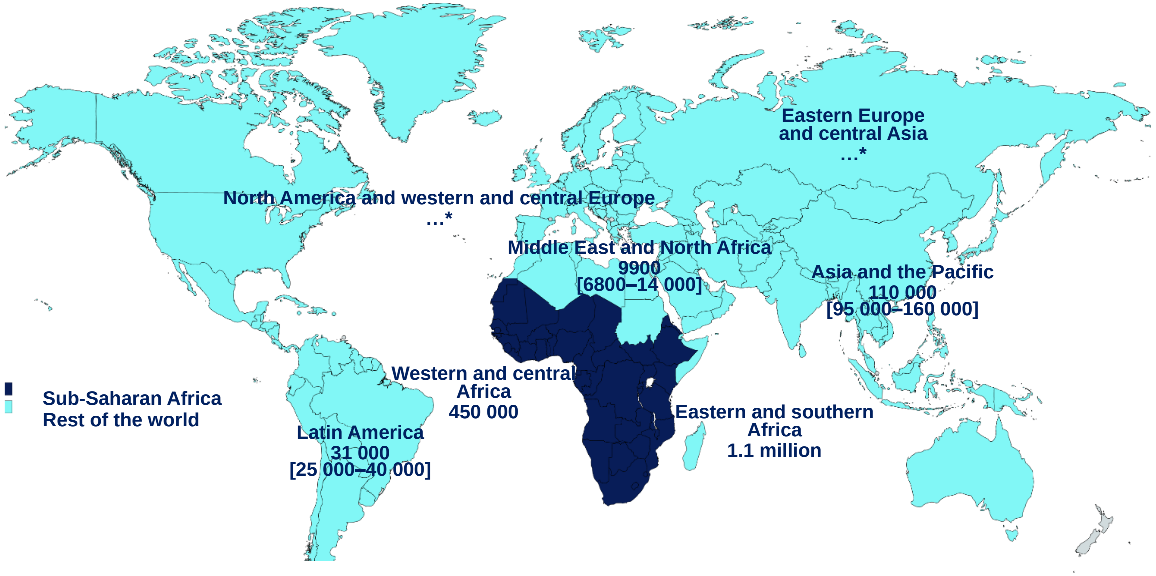
Developing best practices for addressing the legal, ethical, and societal implications of genomics research and biobank governance with pediatric populations in Africa

Community Engagement & Ethics



- Tells the story of DNA and genetics through the eyes of young man who is recruited to a genetic study
- Translated into 5 commonest languages in Africa
  - Arabic, Swahili, Hausa, French, Portuguese
  - Setswana and Luganda

Genetics of Pediatric HIV disease progression in Africa



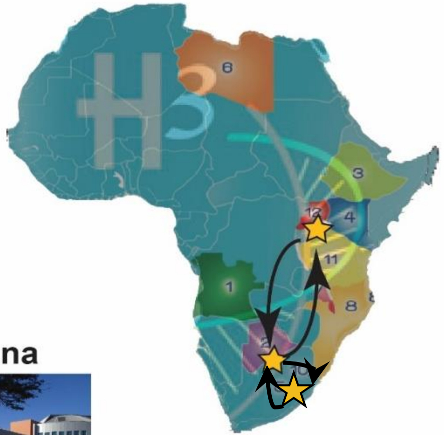
Children <15yrs living with HIV

Baylor College of Medicine

Houston, Texas, USA

Makerere University

Baylor Uganda Children's Foundation



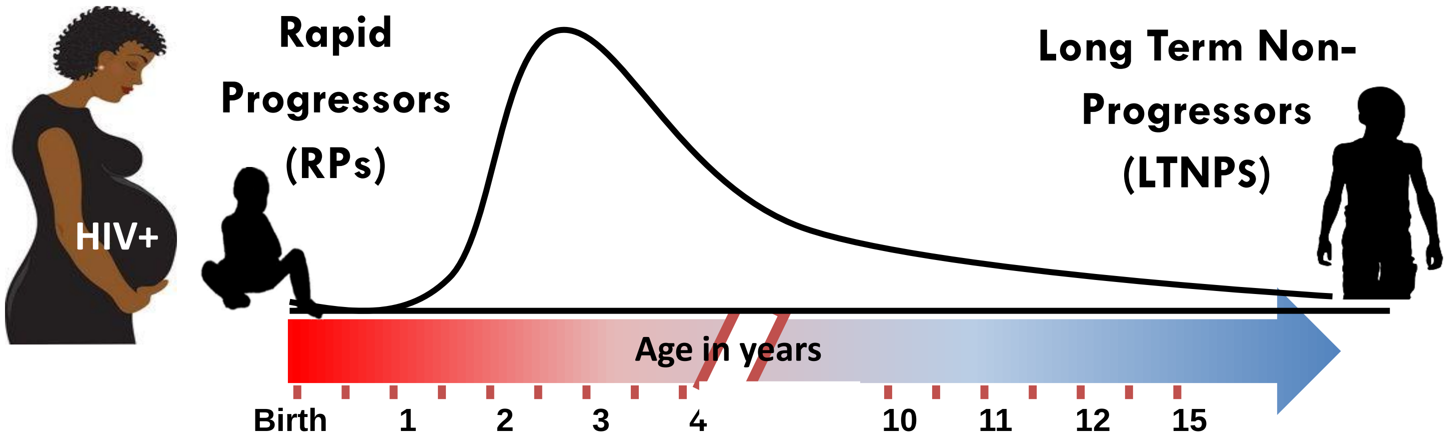
University of Botswana

Botswana-Baylor Children's Clinical COE

Baylor-Swaziland Centre of Excellence

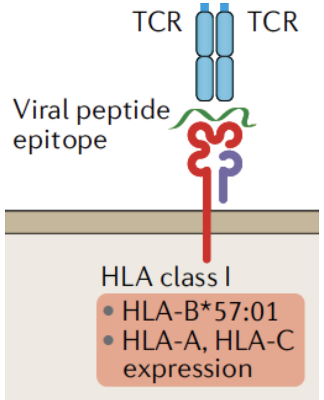
Collaborative African Genomics Network

Disease progression in HIV +ve African children

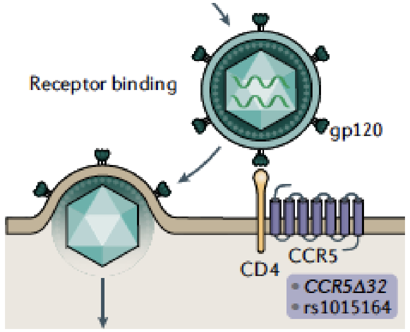


European Adults

- **HLA** class I variants
  - HLA-B*57, B*27



- **Chemokines** (e.g. *CCR5* Δ32)

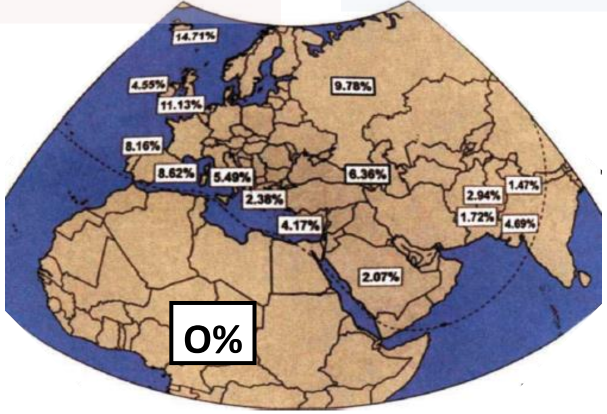


Children ≠ Adults

- Low immune activation
- Low CCR5 expression on memory T-cells

Population Genetics

- Different HLA-types in African adult progression
- *CCR5* Δ32 rare in Africa



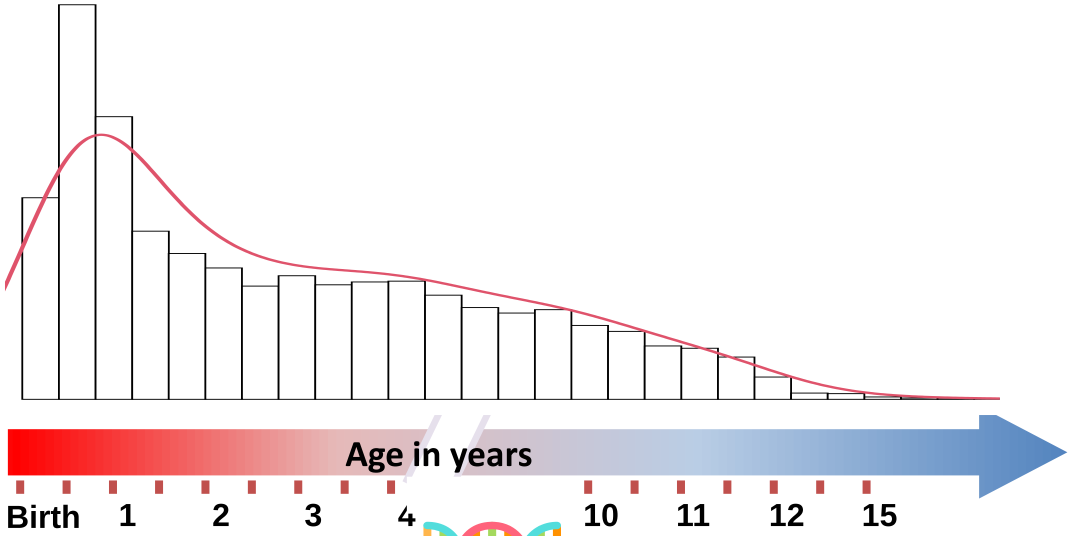
Martinson et al, Nat Gen, 1997

Disease progression in HIV +ve African children

- 14,246 ART naïve children
- Retrospective EHR

Rapid Progressors (RPs)

Phase I: n= 500
Phase II: n=500



Long Term Non-Progressors (LTNPs)

Phase I: n=500
Phase II: n=500

DNA Sequencing

840 (Exome) 400 (Genome)

Coding Variant Association

- Rare variants
- Copy number variants
- KIR
- HLA class I


Marion Amajul


John Mukisa


Thabo Diphoko


Samuel Kyobe

Unmapped Reads

Virome Analysis


Savannah Mwesigwa

Missing Variation


Gaseene Sebetso

SNP genotyping

(n=1550)

Genome wide Association Analysis


Lesedi Williams

NIH NHGRI

CAfGEN

HLA-C is putatively associated with pediatric disease progression

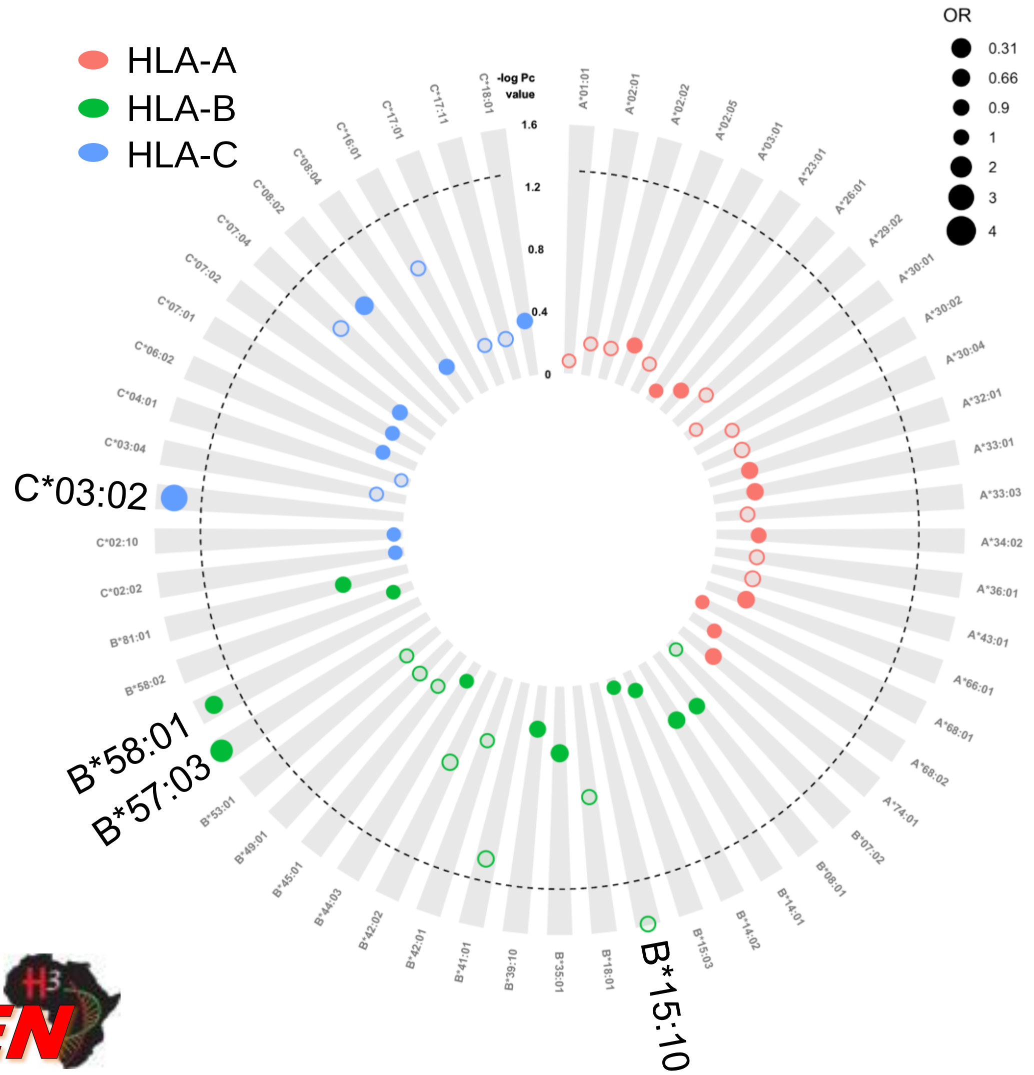
Samuel Kyobe

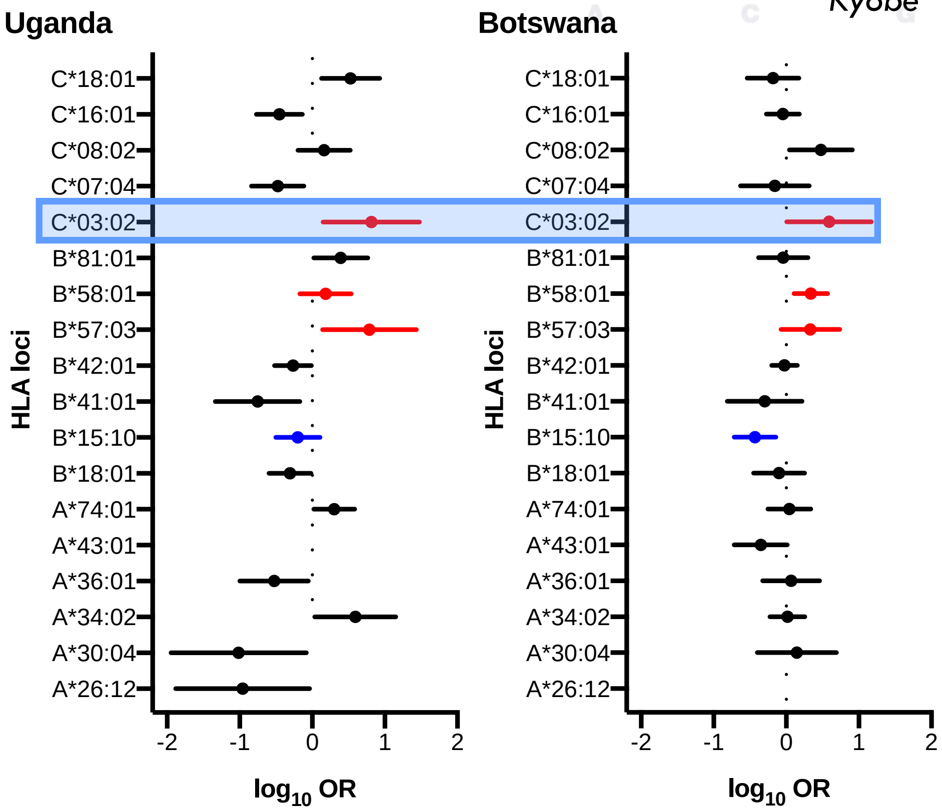
HLA Allelotypes

WES DATA
N = *748*

HLA Typing

- HLAreporter



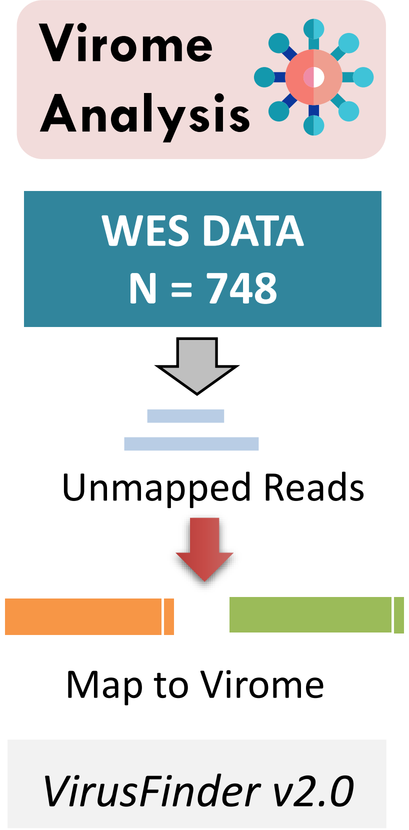


- **HLA-C*03:02 not previously associated with LTNP**
- ~5 -12% of variance in progression

Kyobe, Front in Gen, 2021

Anelloviridae in Pediatric HIV disease progression

Savannah Mwesigwa



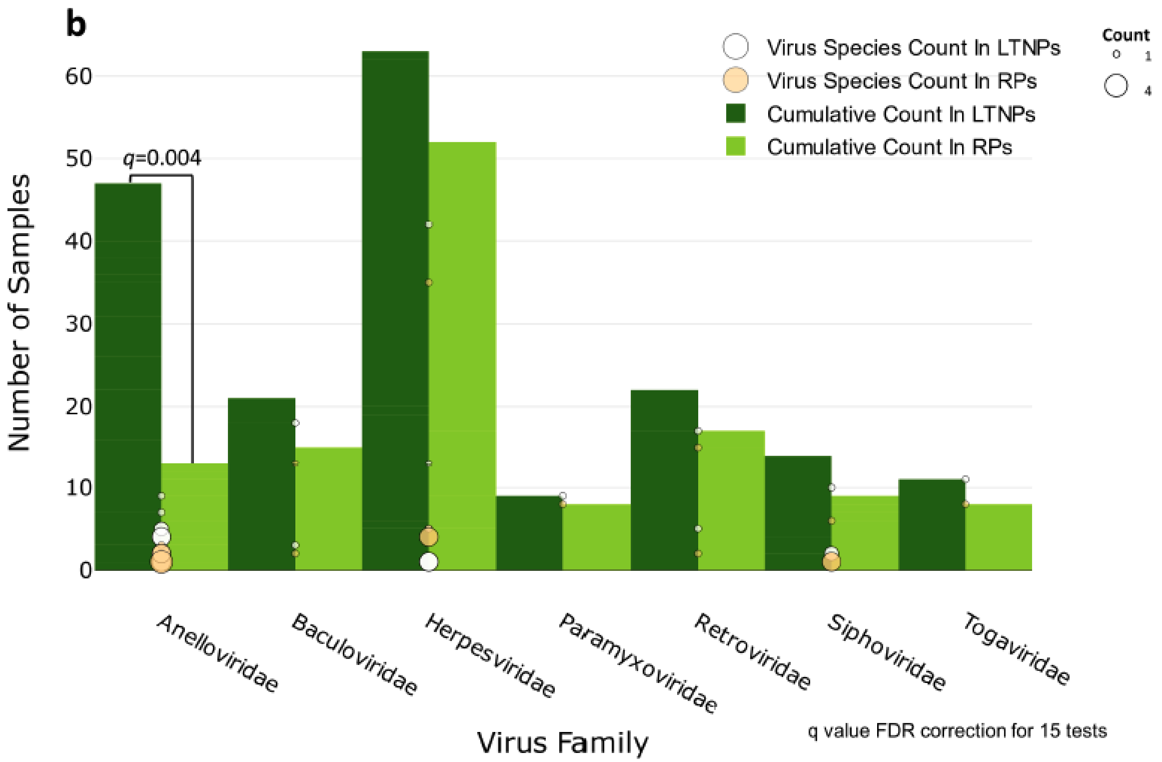
Virome Analysis

WES DATA
N = 748

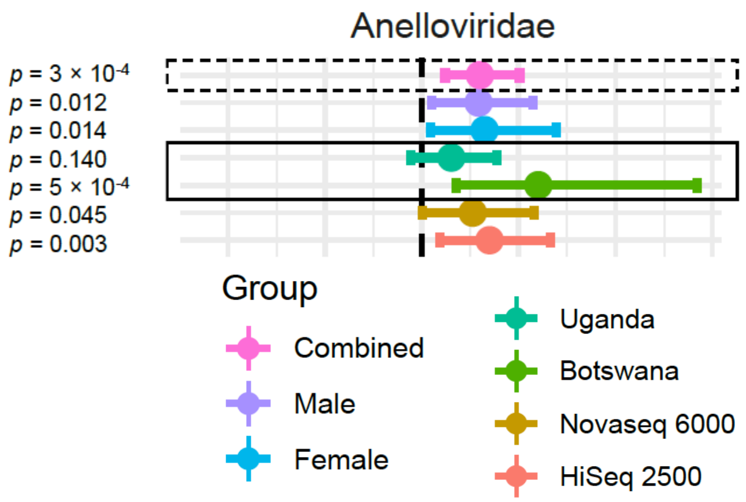
Unmapped Reads

Map to Virome

VirusFinder v2.0



q value FDR correction for 15 tests



- Anelloviridae - T-lymphotropic virus strongly associated with immune quiescence
- Up to 90% of adults already infected, but higher in LTNP children than adults
- Unclear whether cause or effect

NIH NHGRI

CAfGEN

Mwesigwa, npj Genome Med, 2021

Rare variants contribute to disease progression

Marion Amajul

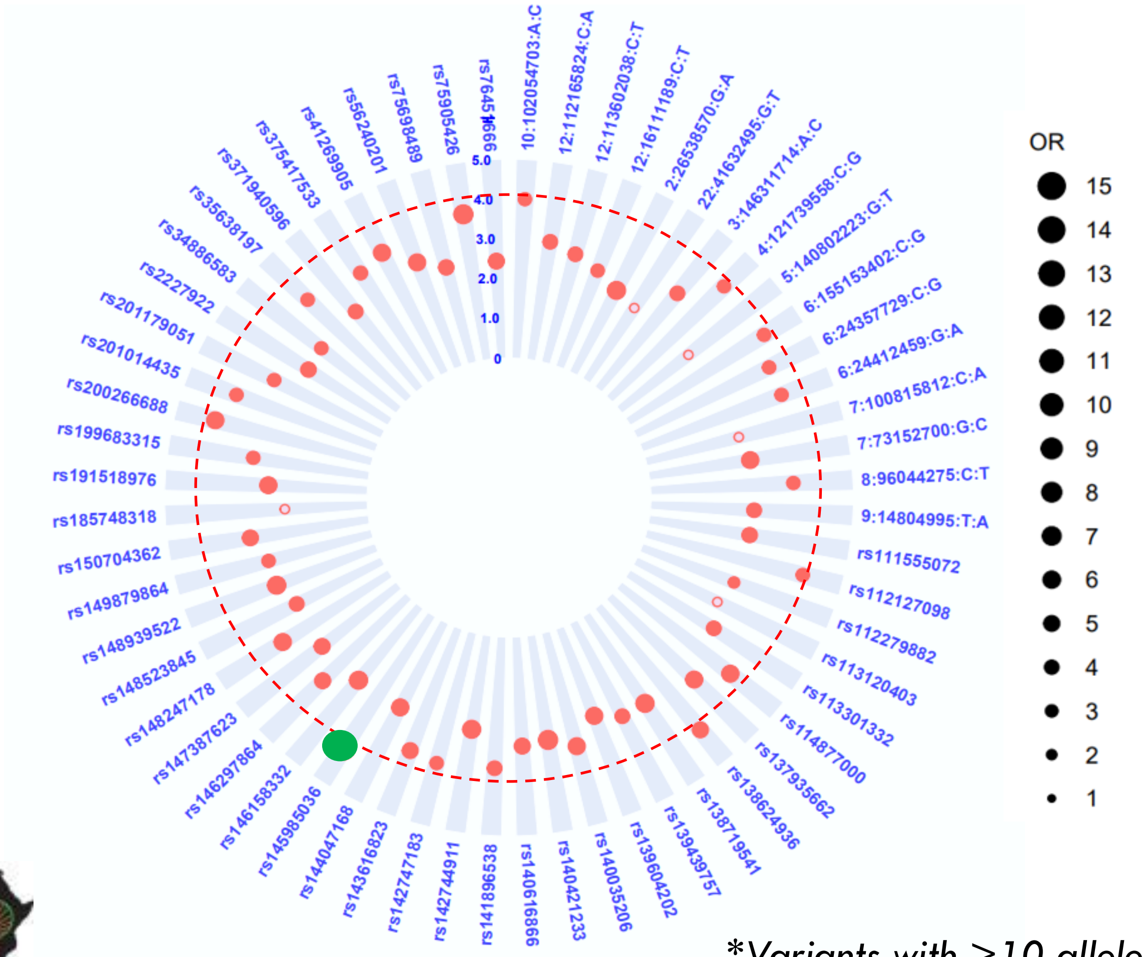
Coding variant enrichment

WES DATA
N = 807

Annotation & Classification

Rare (MAF <1%), damaging, coding variation enriched in cases

- ***CCR5-Δ32*** base pair deletion allele <u>not</u> observed
- <u>No</u> association in chemokine or chemokine receptor genes

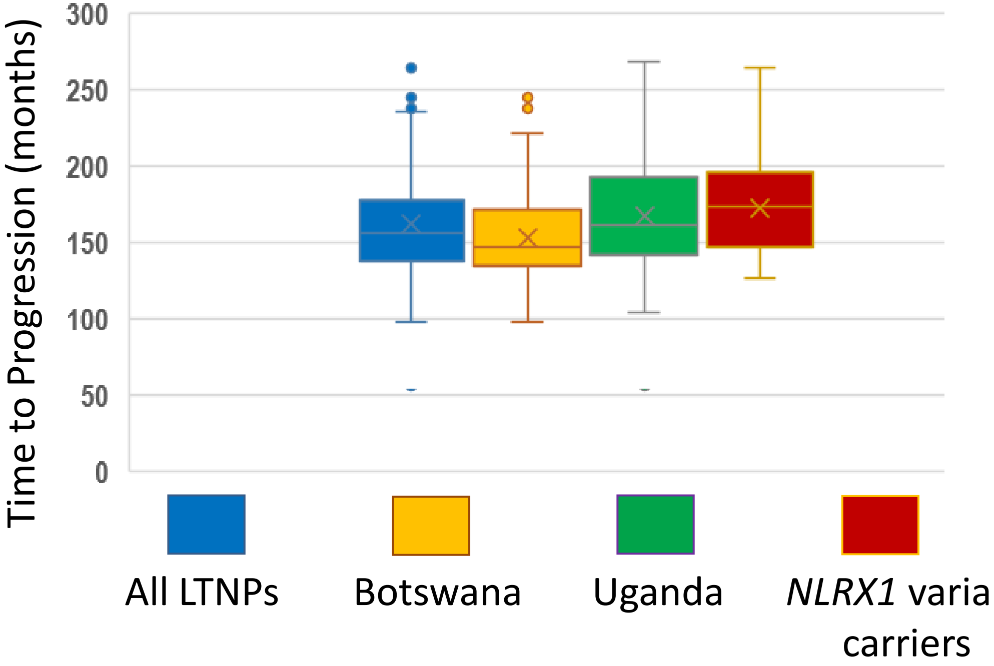


**Variants with >10 alleles*

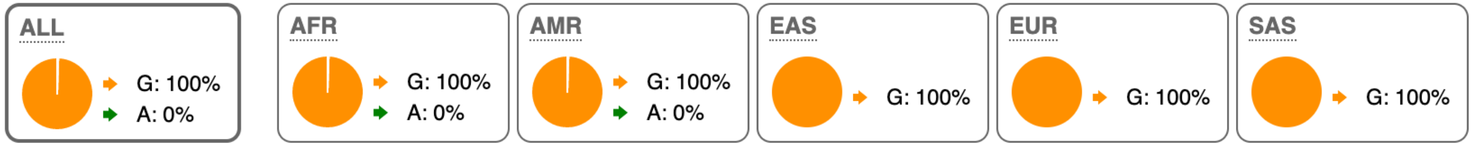
CHR	Variant ID	Raw P	Adj. P
11	rs145985036	0.000099	0.025
17	rs140616866	0.0018	0.803

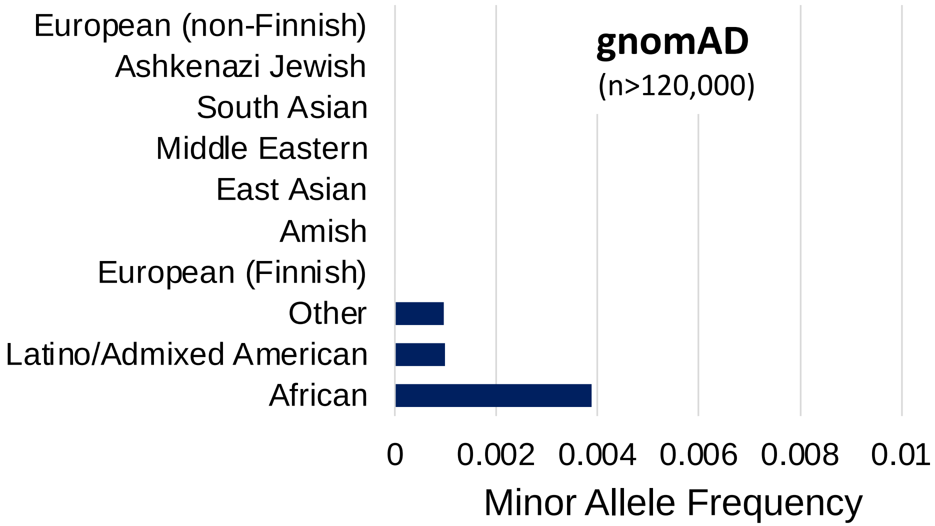
A rare variant in *NLRX1* is enriched among LTNPs

						No. of alt. alleles			Uganda (n=463)		Botswana (n=350)	
Chr	**Ref**	**Alt**	**Gene**	**rsID**	**Nominal P**	**Total**	**LTNP**	**RP**	LTNP	RP	LTNP	RP
									11	1	**3**	0
11	**G**	**A**	***NLRX1***	**rs145985036**	**0.00032**	**15**	**14**	**1**	**0.029**	0.008	**0.003**	0



1000 Genomes Project Phase 3 allele frequencies



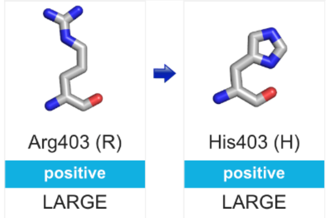


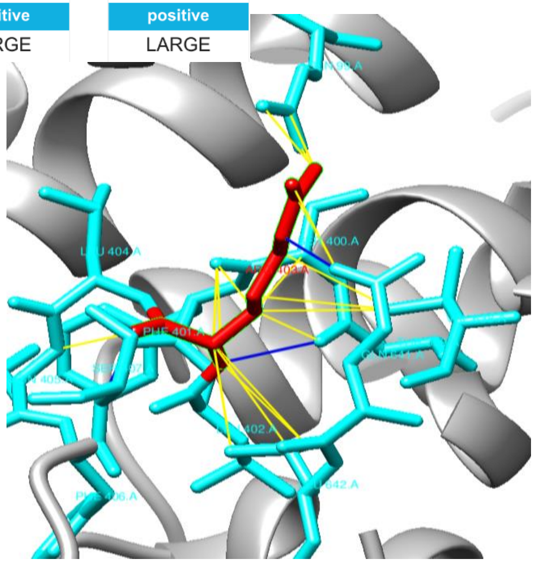
NIH NHGRI

CAfGEN

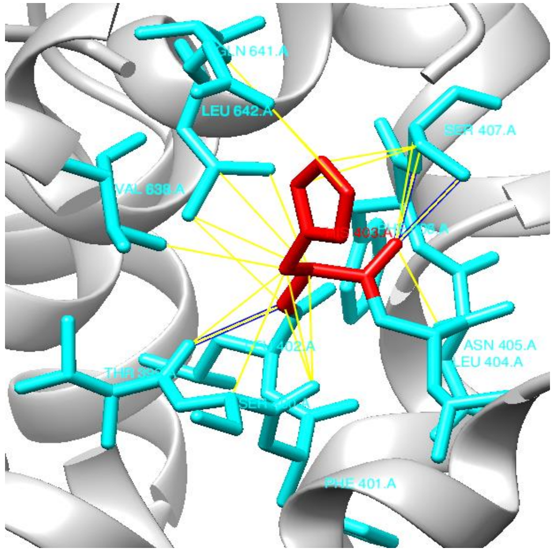
Enriched variant is predicted to damage the resulting protein

		Insilico pathogenicity predictors				
Gene	**Protein Change**	**SIFT**	**Polyphen**	**MutTaster**	**PROVEAN**	**CADD**
NLRX1	**p.R403H**	**Damaging**	**Damaging**	**Damaging**	**Damaging**	**24.6**



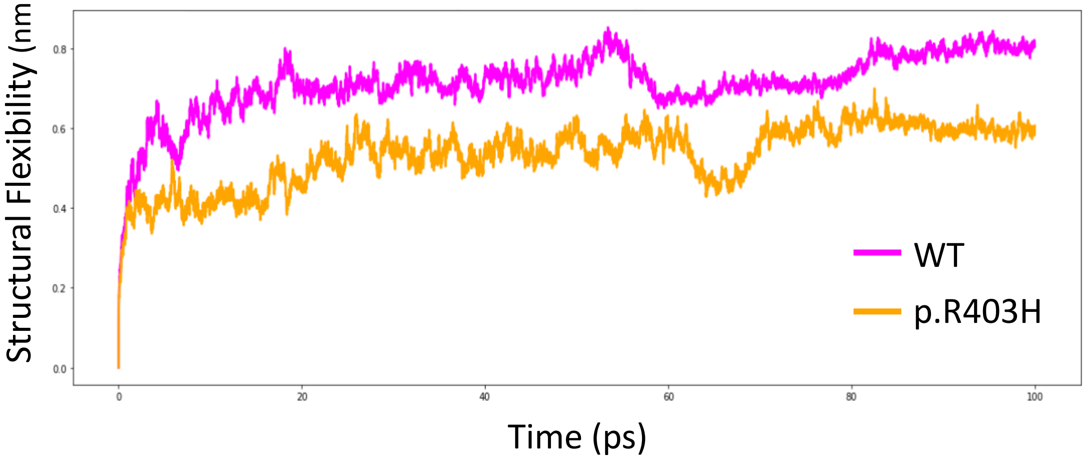


Wild type R403



Mutant R403H

Increased rigidity of mutant protein vs wild type



NIH
NHGRI
CAfGEN
H3

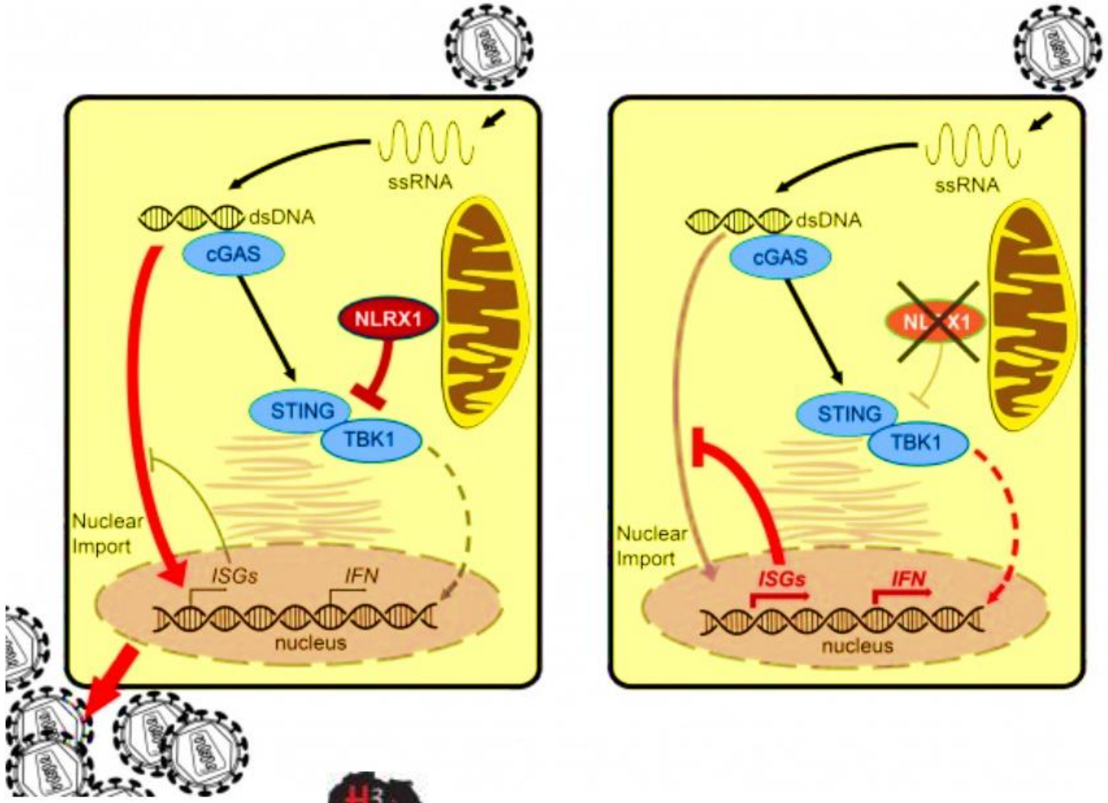
NLRX1 and HIV1 infection

Cell Host & Microbe

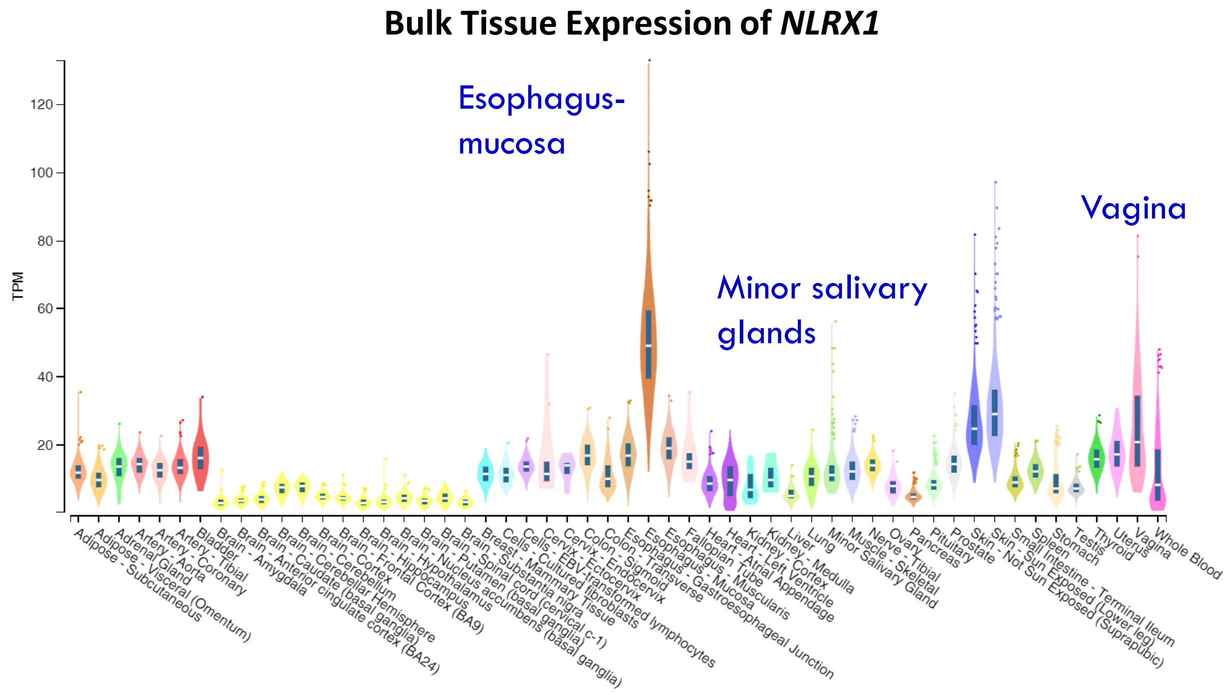
Article

NLRX1 Sequesters STING to Negatively Regulate the Interferon Response, Thereby Facilitating the Replication of HIV-1 and DNA Viruses

Haitao Guo,[1,13] Renate König,[2,3,13] Meng Deng,[1,4] Maximilian Riess,[3] Jinyao Mo,[5] Lu Zhang,[1] Alex Petrucelli,[1] Sunnie M. Yoh,[2] Brice Barefoot,[7] Melissa Samo,[7] Gregory D. Sempowski,[7] Aiping Zhang,[8] Anamaris M. Colberg-Poley,[8,9] Hui Feng,[1] Stanley M. Lemon,[1,5,10] Yong Liu,[1,11] Yanping Zhang,[1,11] Haitao Wen,[1] Zhigang Zhang,[1] Blossom Damania,[1,10] Li-Chung Tsao,[1] Qi Wang,[1] Lishan Su,[1,10] Joseph A. Duncan,[1,5,6] Sumit K. Chanda,[2,*] and Jenny P.-Y. Ting[1,10,12,*]



- A major regulator of IFN-Y in innate immunity
- NLRX1 attenuates immune reactions, thus enabling HIV replication
- NLRX1 reduction potentiates innate immunity and blocks HIV replication
- High expression in mucus membranes of GI and reproductive tract



Bulk Tissue Expression of *NLRX1*

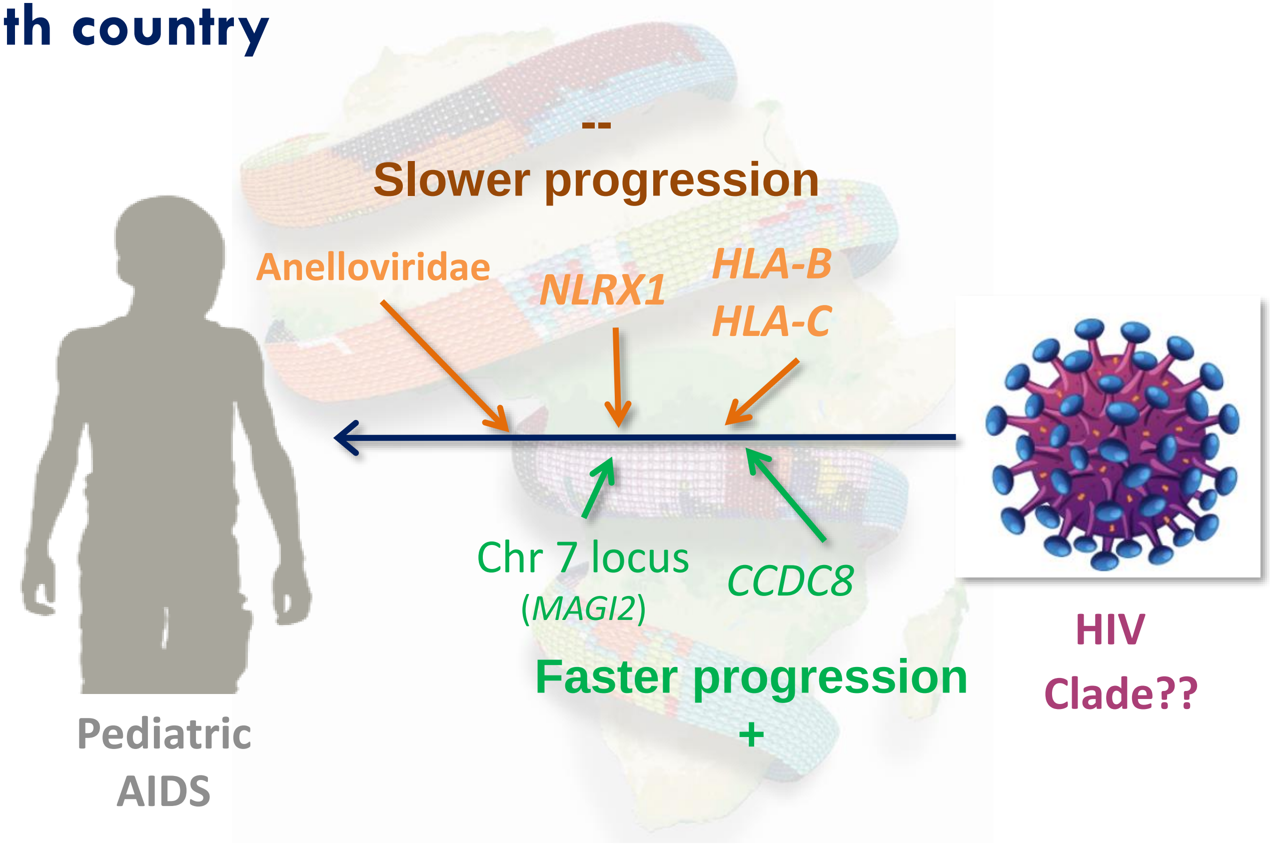
Esophagus-mucosa

Minor salivary glands

Vagina

Genetic Variation in Pediatric HIV in Africa

- **Complex host genetic contribution to disease progression**
- **Variation with country**



--

Slower progression

Anelloviridae

NLRX1

HLA-B

HLA-C

Chr 7 locus (*MAGI2*)

CCDC8

Faster progression

\+

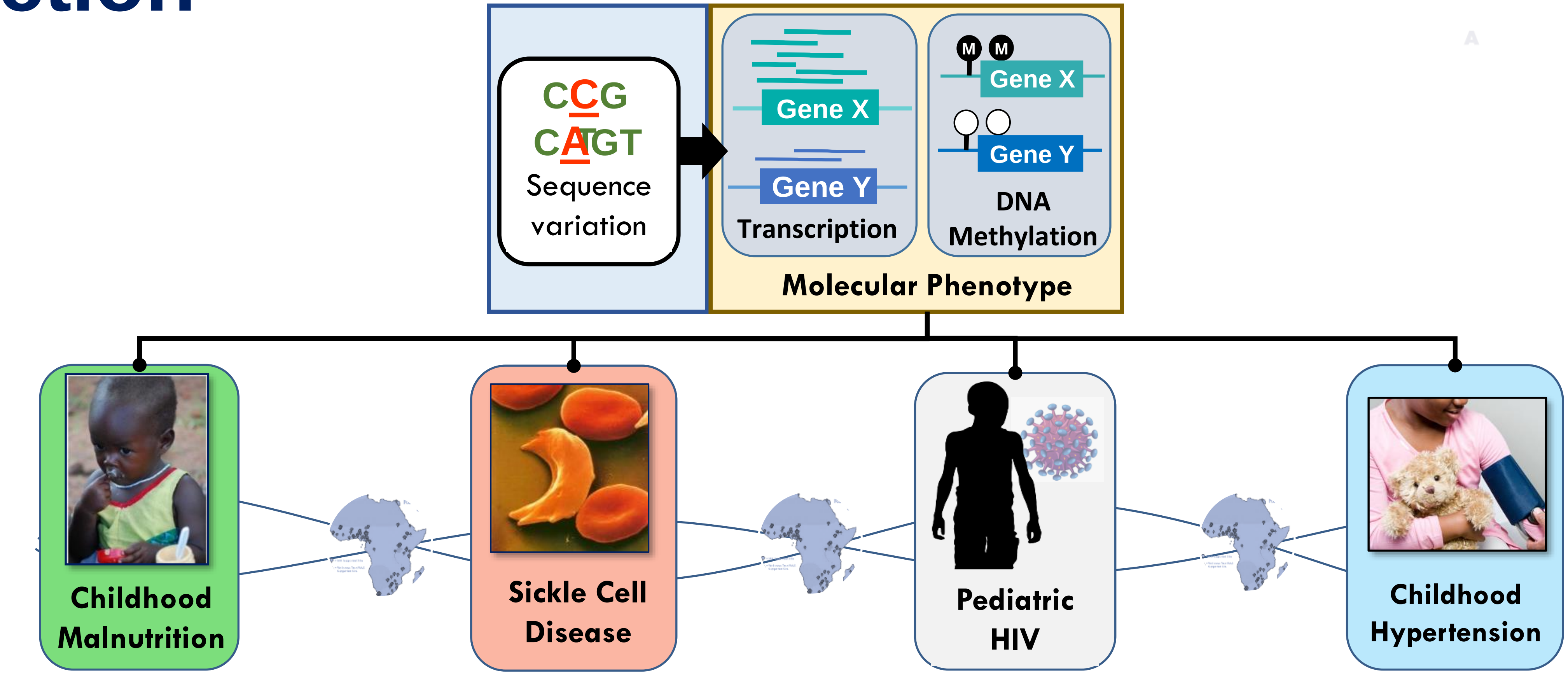
Pediatric AIDS

HIV Clade??

NIH NHGRI

CAfGEN

Childhood Complex Disease Genomics Section



CCG
CATGT
Sequence variation

Gene X
Gene Y
Transcription

M M
Gene X
Gene Y
DNA Methylation

Molecular Phenotype

Childhood Malnutrition

Sickle Cell Disease

Pediatric HIV

Childhood Hypertension

Why do some children get Kwashiorkor when severely malnourished?

Why do some individuals make alloantibodies after red cell transfusions?

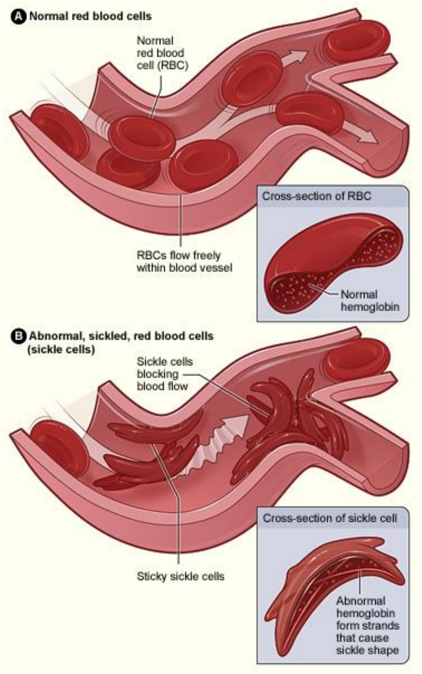
Why do some HIV positive children develop AIDS quickly and others slowly?

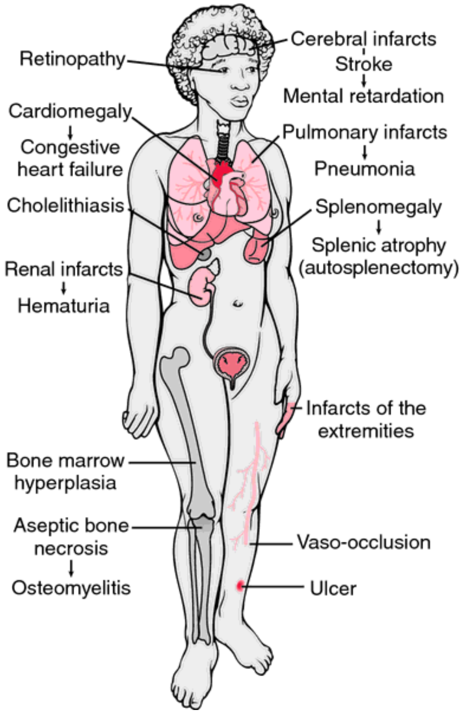
Are there single gene defects underlying early-onset essential hypertension?

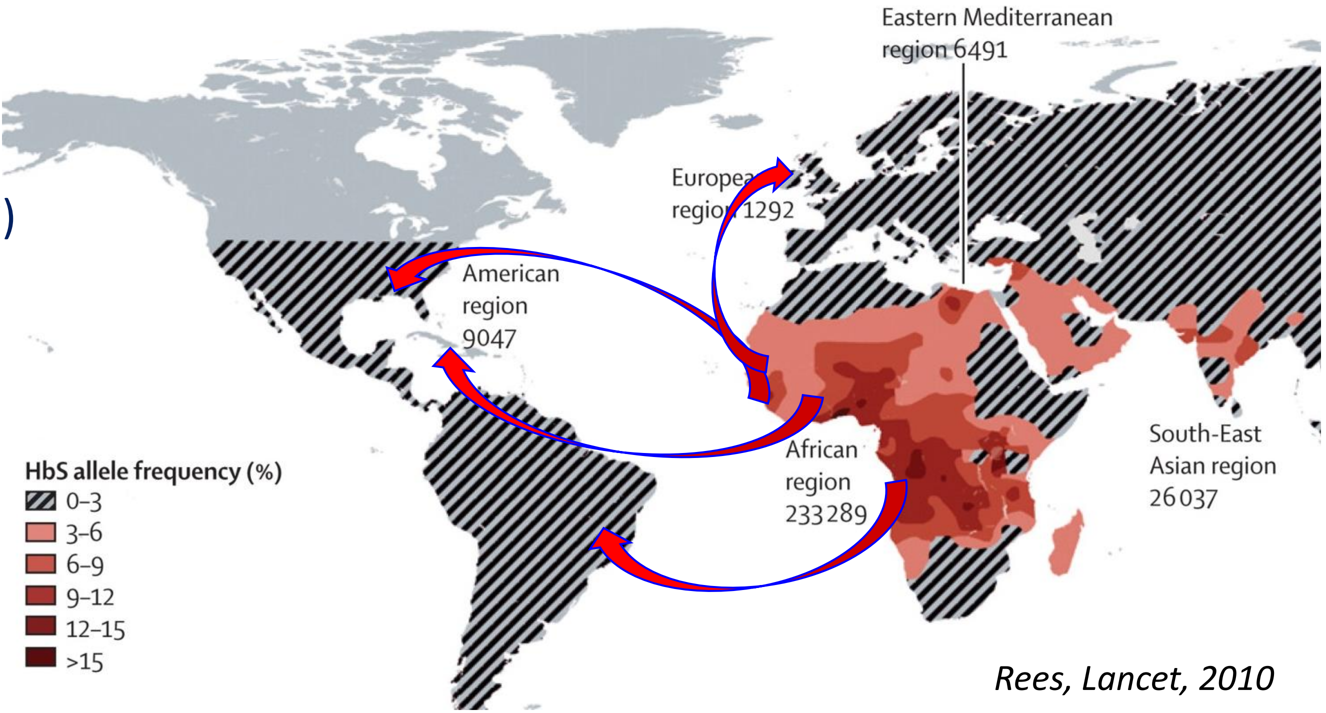
NIH
NHGRI

Sickle cell Disease is the most common single-gene disorder globally

- Caused by single point mutation in the beta-globin (*HBB*)
- Recessive disorder (both gene copies affected)
- Leads to abnormal hemoglobin (Hemoglobin S)







Rees, Lancet, 2010

Highest allele frequency in Africa

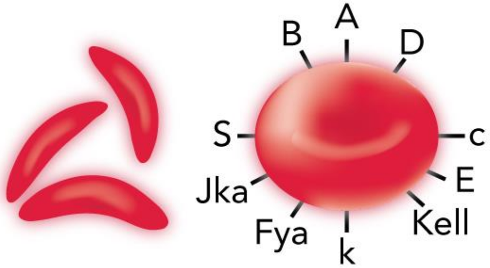
- Carrier rate of up to 30%
- Protective effect from severe Malaria

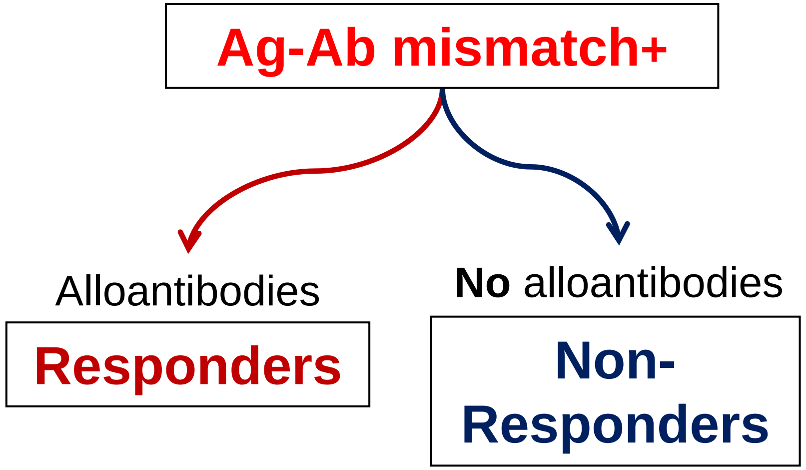
Spread with movement of peoples:

- Mediterranean, North & South America, Europe, Caribbean

RBC alloimmunization in sickle cell disease (SCD)

- Transfusion of Red Blood Cells (RBCs) remains important in managing SCD
- >10% of SCD patients will develop alloantibodies after transfusions → **Hemolysis**
- Alloantibodies create a challenge to find compatible blood





Ag-Ab mismatch+

Alloantibodies

Responders

No alloantibodies

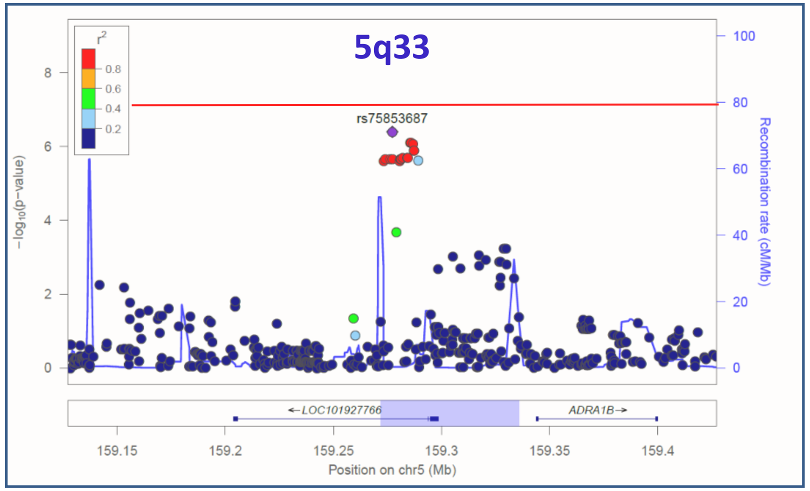
Non-Responders

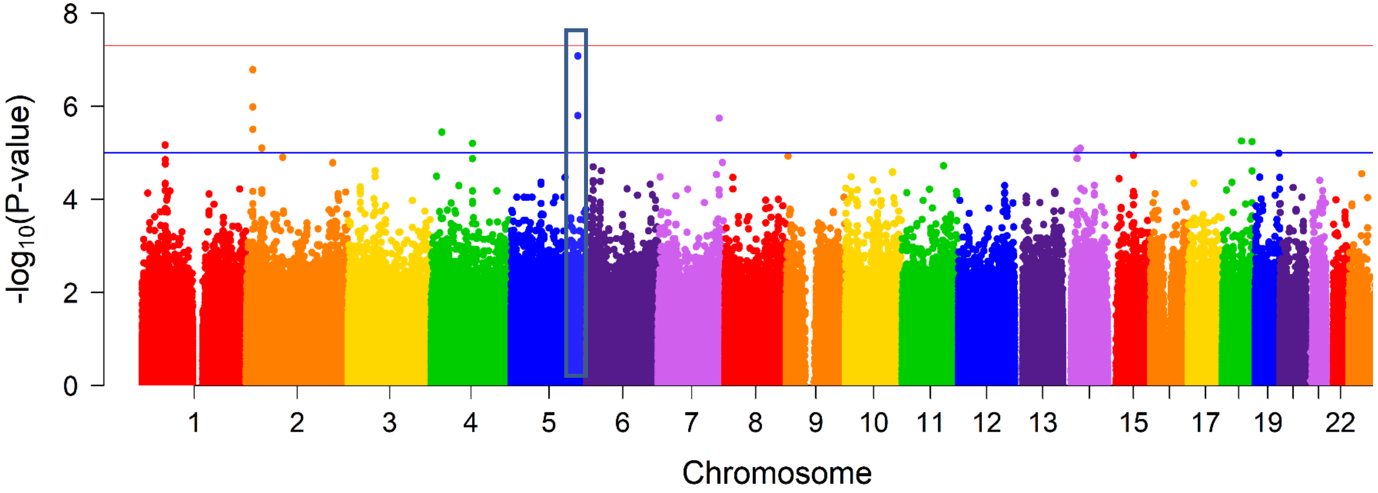
Genome-wide association study (GWAS) of alloimmune 'responder' status

- 288 SCD chronic transfusion recipients
  - Responders (n=154)
  - Non-responders (n=134)
- 2.5 million SNPs genotyped

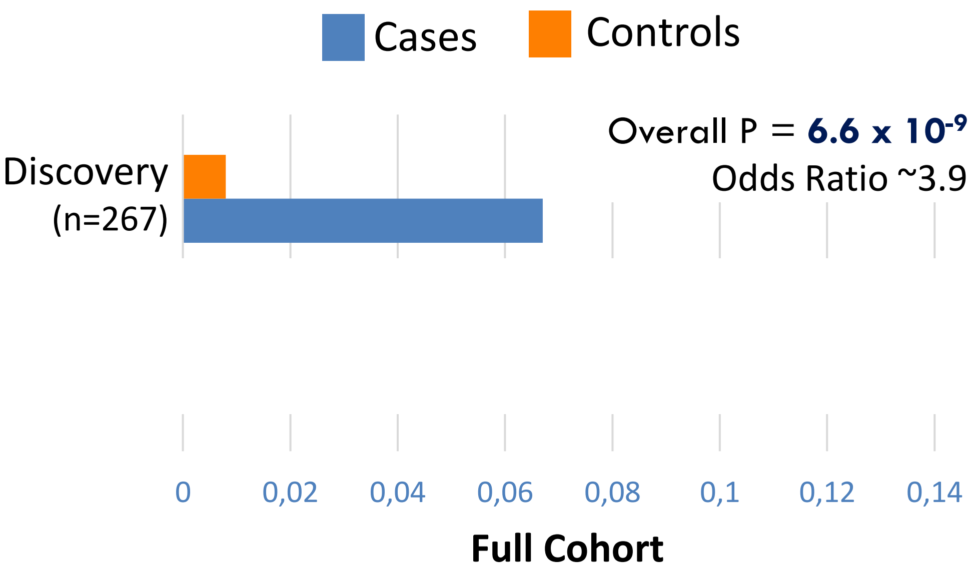
Lesedi Williams

GWAS of SCD alloimmunization





Frequency of 'A' allele



Overall P = 6.6×10^{-9}

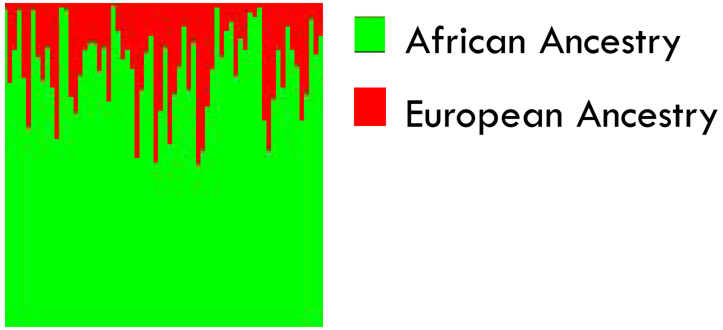
Odds Ratio ~3.9

Williams, Blood Advances, 2018

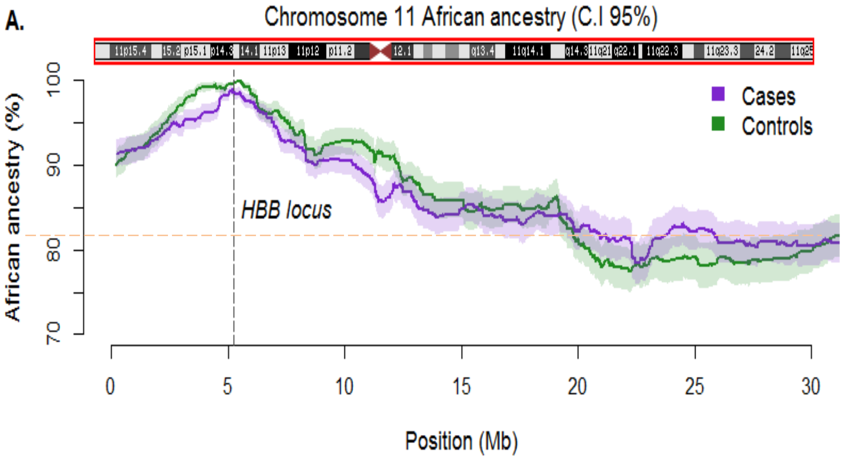
GWAS of SCD alloimmunization

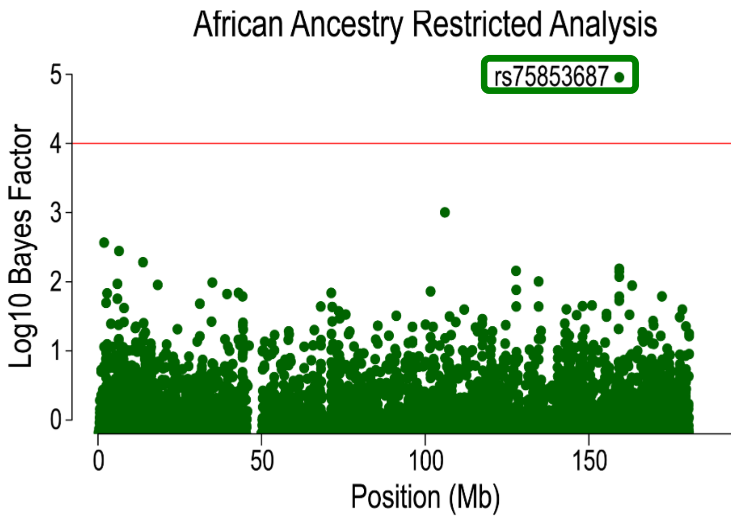
Admixture among African Americans

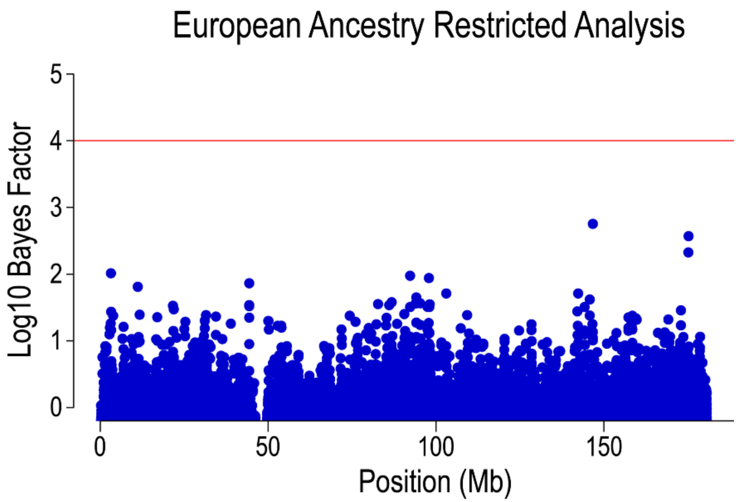
- Mean European ancestry:
  - In African Americans ~20-25%



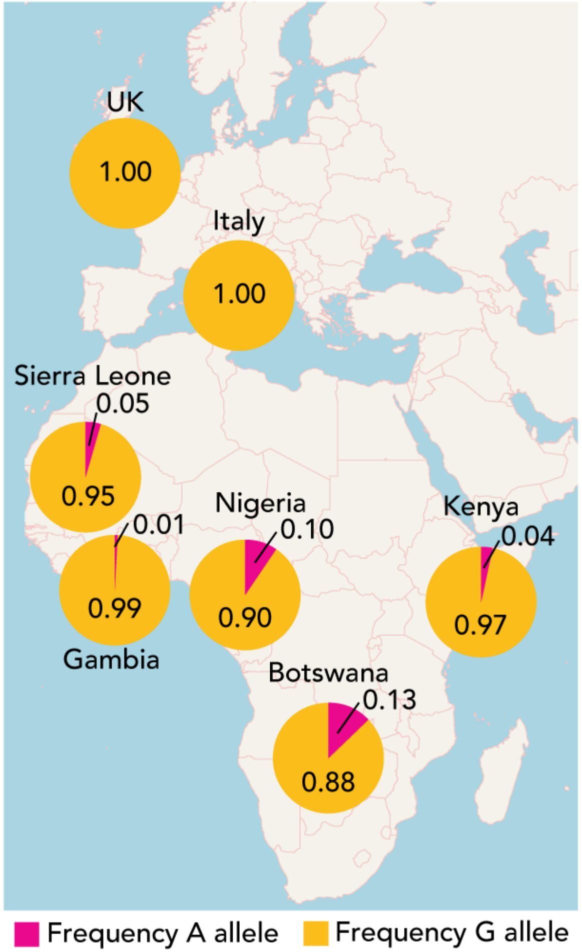
- In our SCD transfusion cohort ~17%





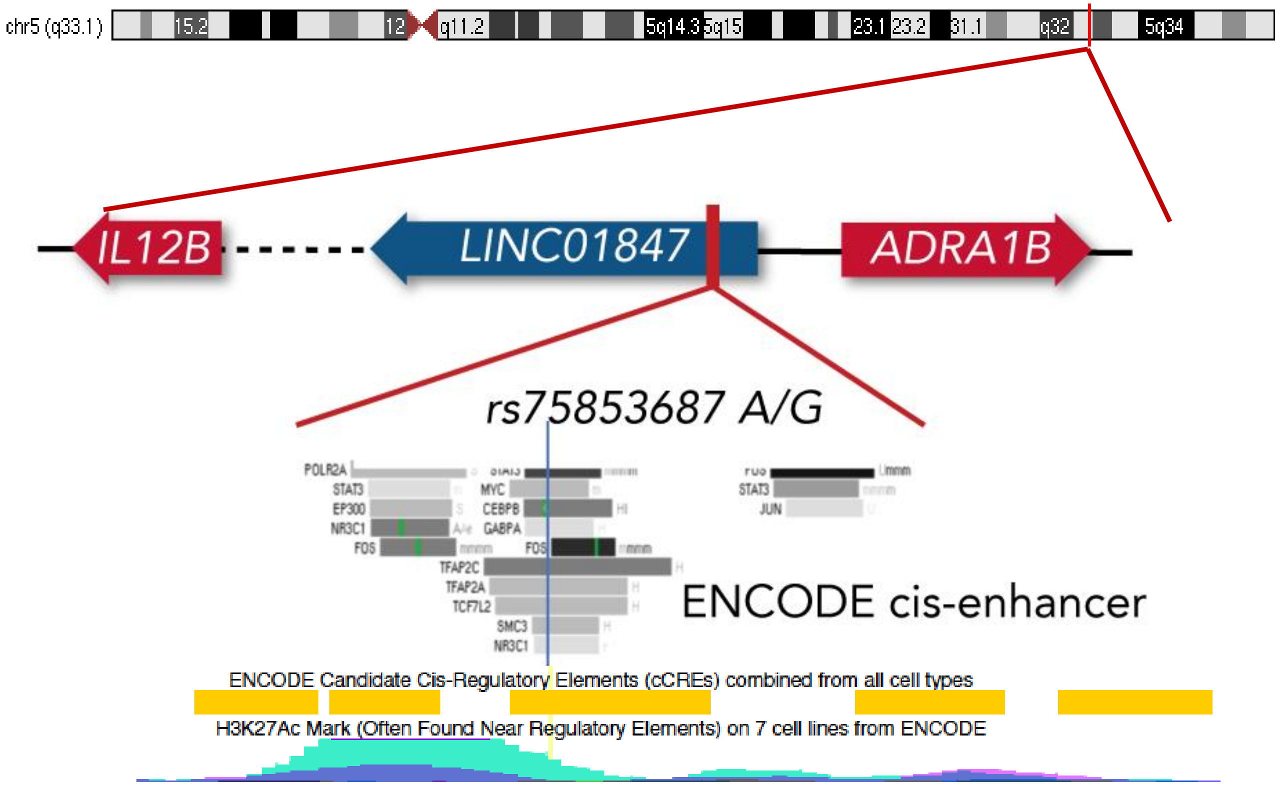


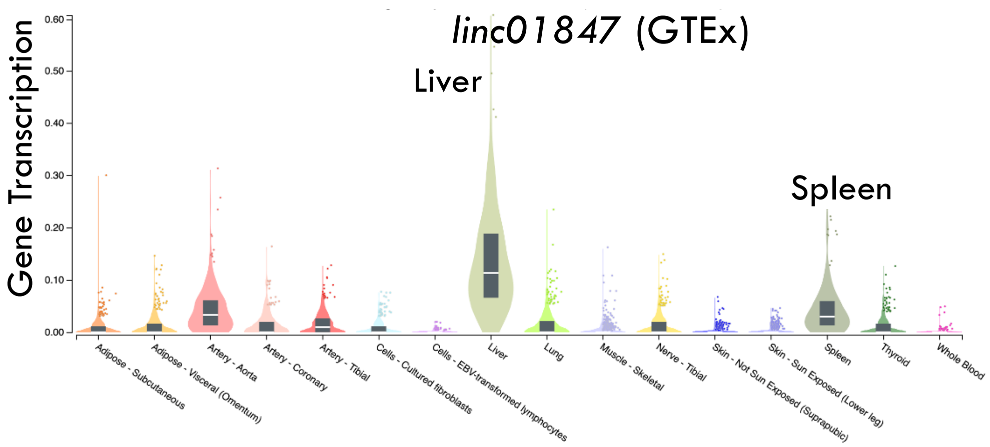
Allele exclusive to haplotype of African origin

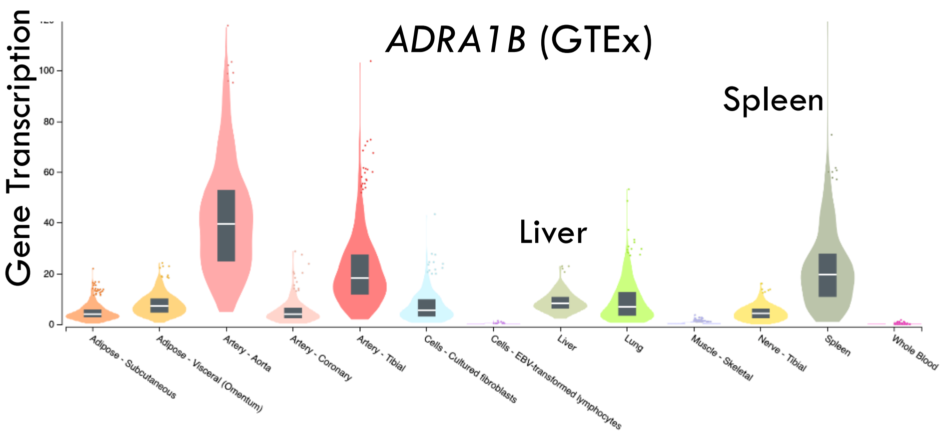


Williams, Blood Advances, 2018

Chromosome 5q33 risk locus

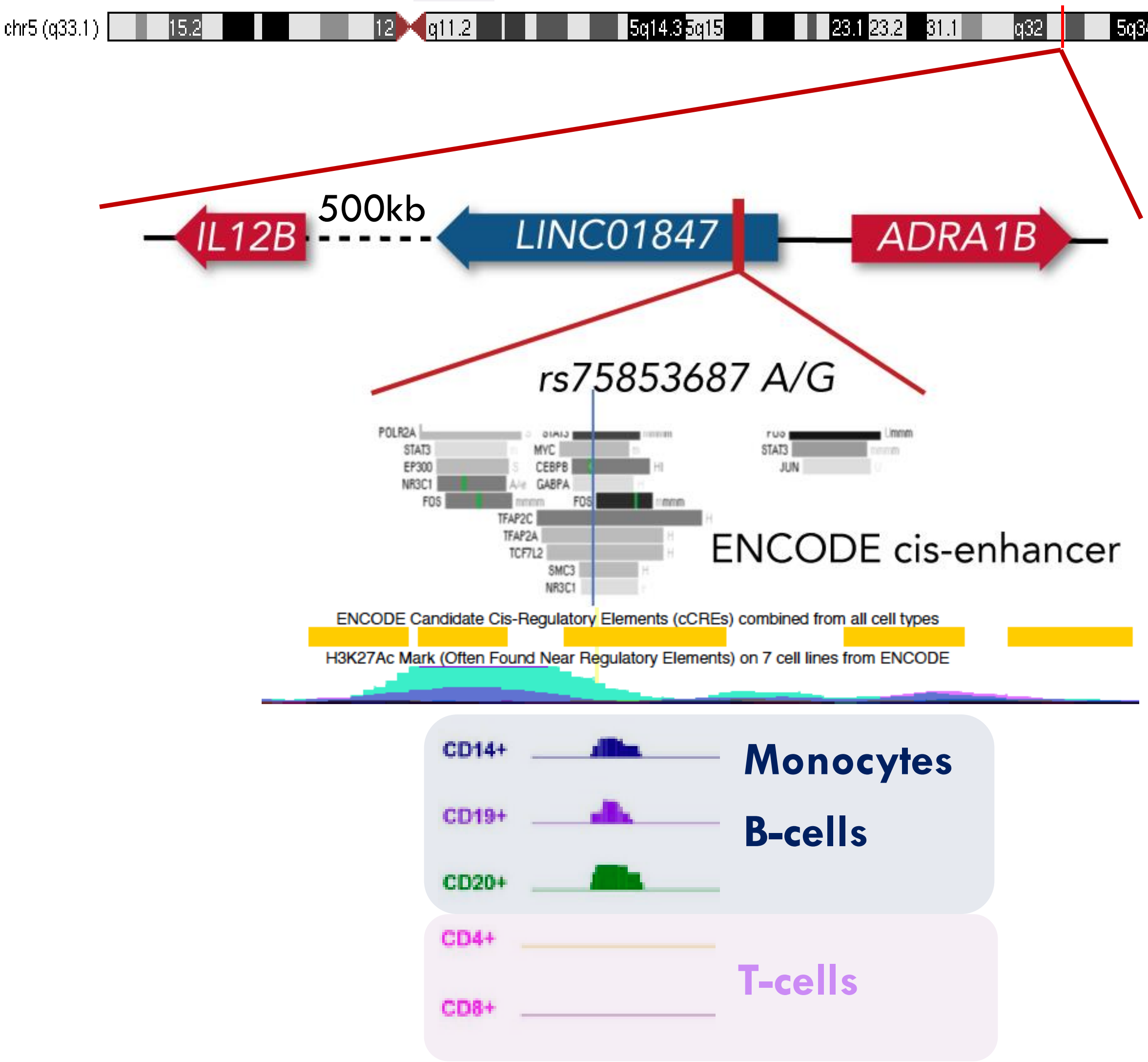


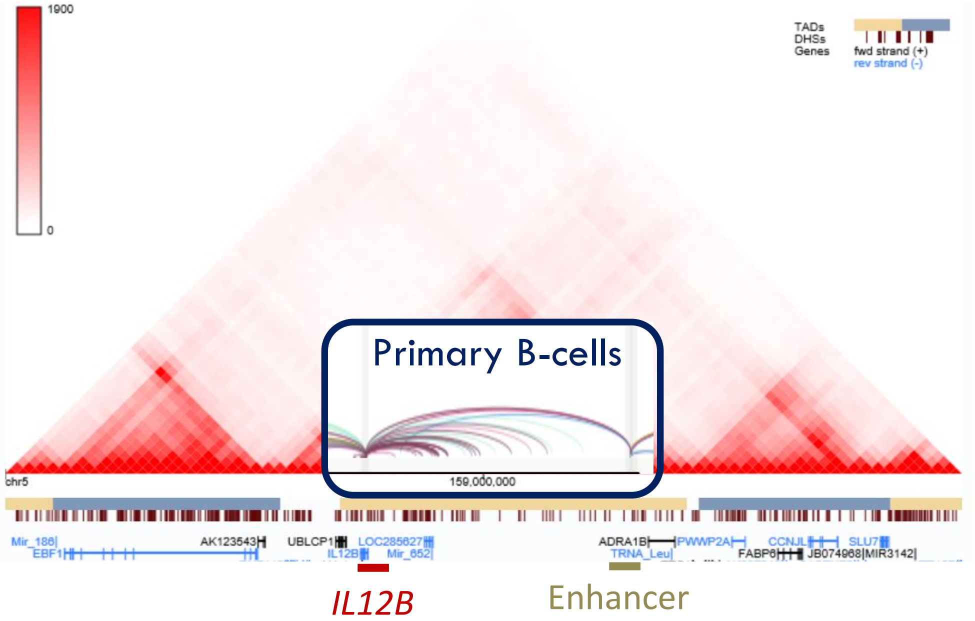




NIH
NHGRI

Chromosome 5q33 risk locus

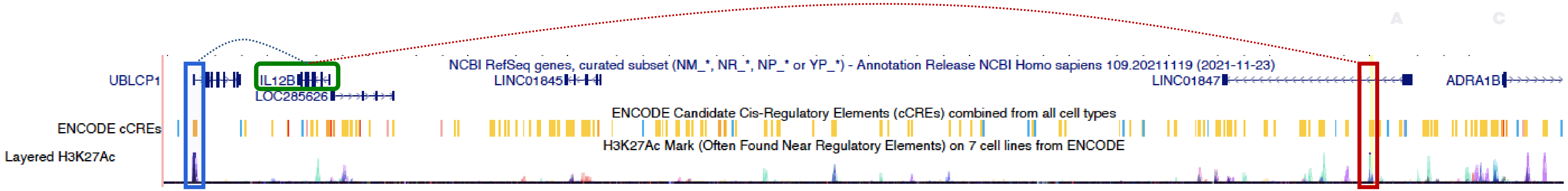




NIH
NHGRI

With Luckey Chance, UVA

Chromosome 5q33 risk locus



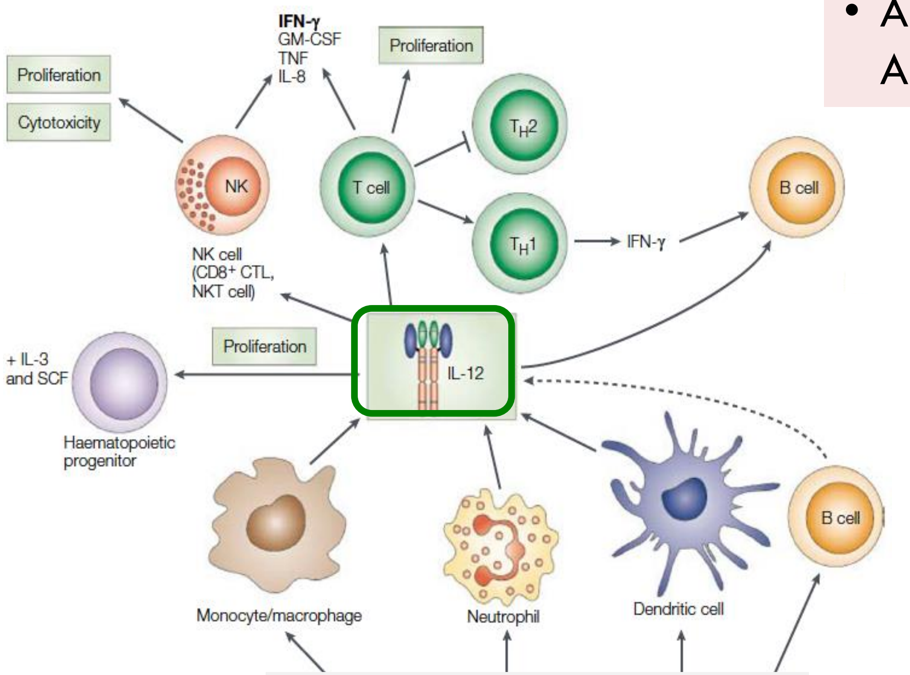
- Tb resistance in East Africa (Uganda & Tanzania)

ARTICLE

A Locus at 5q33.3 Confers Resistance to Tuberculosis in Highly Susceptible Individuals

Rafal S. Sobota,[1,2] Catherine M. Stein,[3,4] Nuri Kodaman,[1,2,3] Laura B. Scheinfeldt,

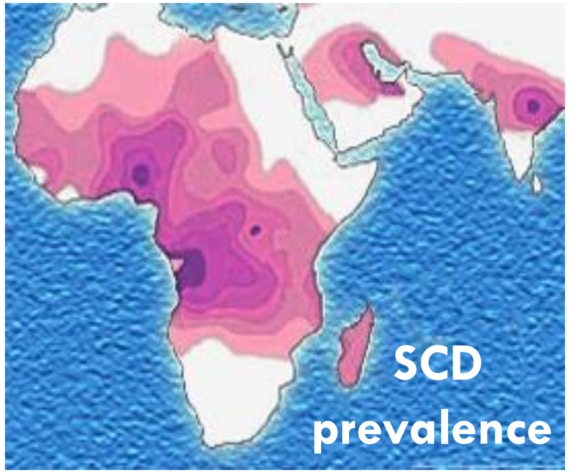
- Alloantibody susceptibility in African Americans with SCD

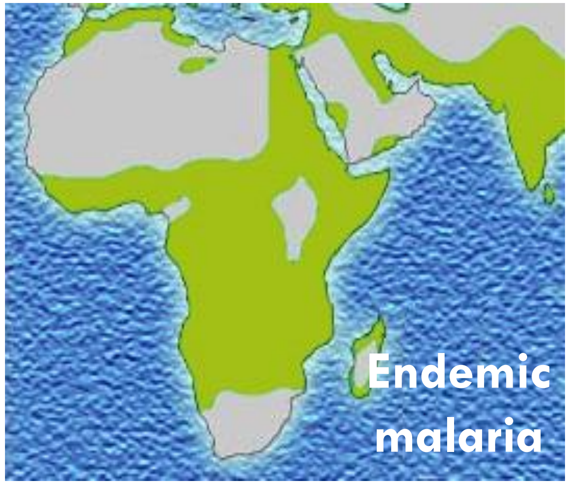


Pathogens

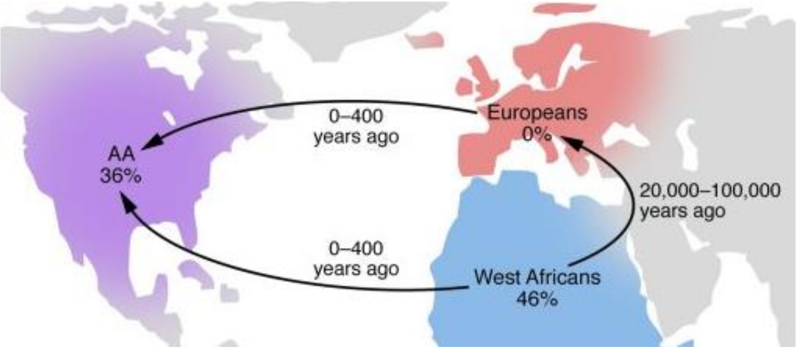
Transatlantic Pleiotropy

HBB

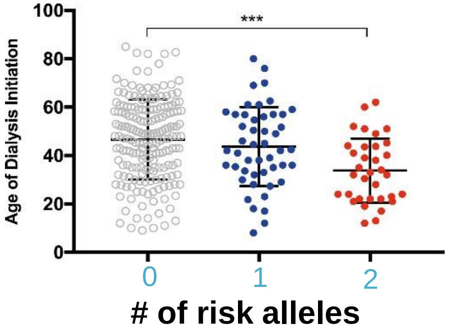




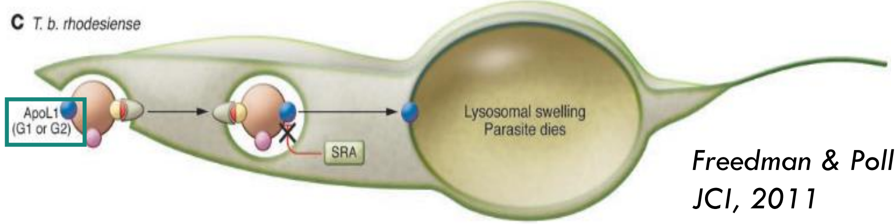
APOL1



Susceptibility to end-stage renal disease



Protection from Trypanosomiasis Infection



Freedman & Pollack,
JCI, 2011

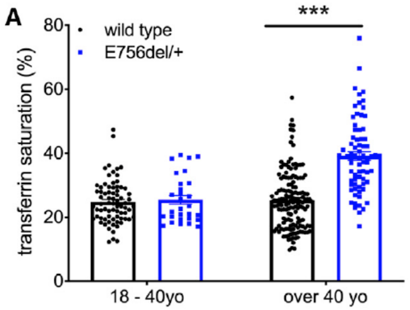
Adult human

Gain-of-function (GOF) mutation

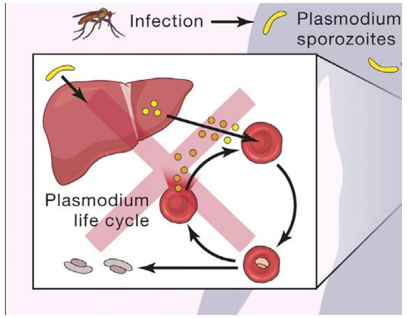
PIEZO1

E756del
rs59446030

Iron accumulation in Afr. Am. Adults



Resistance to severe malaria



Ma, Cell , 2021
Hanchard & Wonkam, Cell, 2021

Summary

- Genomic studies in different populations is important for equity in the development of precision medicine
  - Identify previously unrecognized disorders
  - Better characterize known disorders
  - Inform the practice of medical genetics
  - Guide genetic anthropology
  - Potential to identify pleiotropic genes and variants
- More genome sequencing and genome studies in diverse populations required

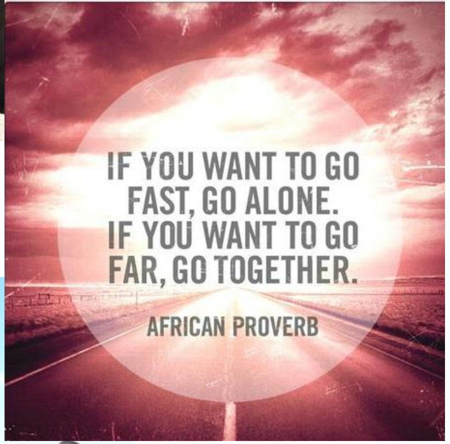
Acknowledgments

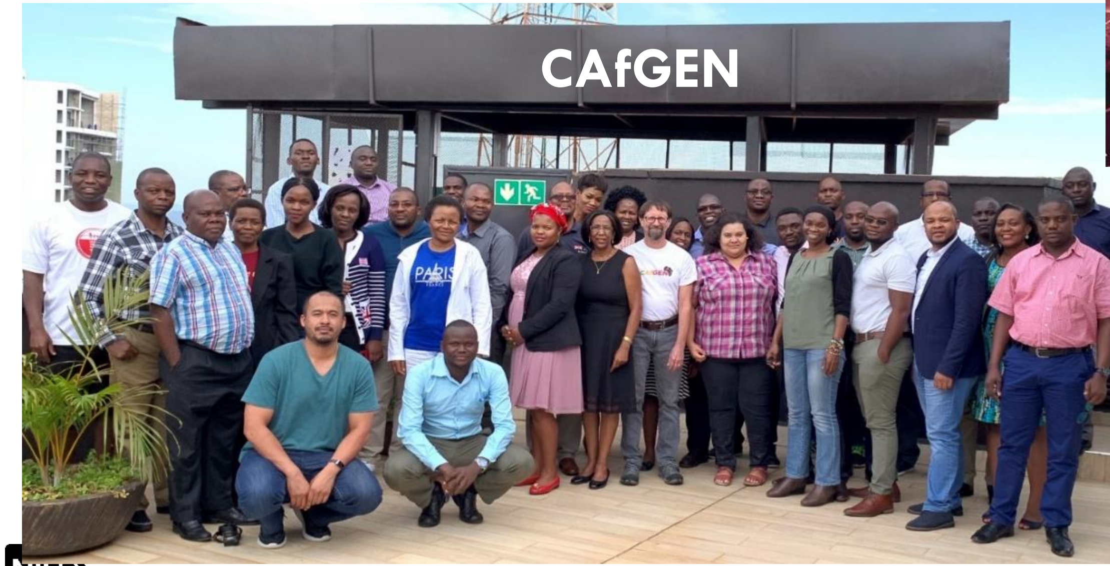
- Aparna Haldipur
- **Pamela Russell**
- Natasha Lie
- Yixing Han
- Thalia Billawali
- Qing Li

NIH - CCDGS

- Edmond Wonkam
- Toluwani Taiwo
- Aarti Jajoo

- Jared Redmond
- Simrah Hamid
- **Allyson Motter**
- Emilyn Banfield
- Sarah Brown





Funding

- BCM Intramural
- NIH/NHGRI/NIAID
- American Association of Blood Banking (AABB)

BCM

- Patricia Hernandez
- Gladys Zapata
- John Belmont
- **Graeme Mardon**
- **Joann Moulds**
- **Alice Chen**