

# Local Ancestry Inference (LAI)

Katia Bougiouri

**Popgen Summer Course 2025**

Copenhagen

August 7, 2025

# Outline of today

- What is local ancestry inference? How does it differ from global ancestry?
- What can we use LAI for?
- Methods for inferring local ancestry
- Break
- Exercises

# What is genetic ancestry?

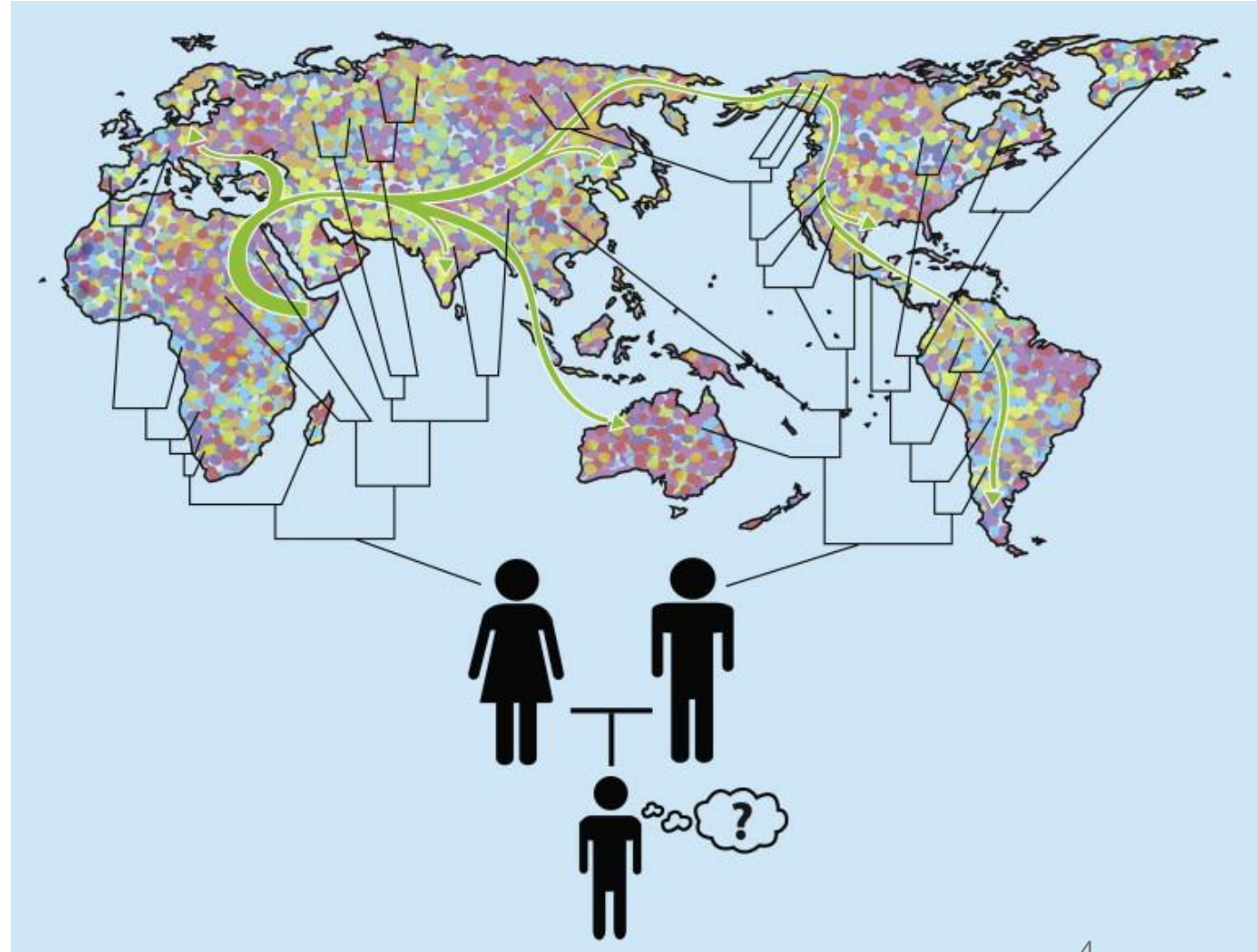
- In general: genetic ancestry is defined by the stretches of the genome that we inherit from our ancestors
- Many definitions, interpretations and misinterpretations
- Continental ancestry the most common for humans
- Compare to “reference” databases of individual samples from diverse populations and genetic regions
- Depends on what specific “time slice” we’re focusing on

European	92.3%
British & Irish United Kingdom	26.7%
French & German Netherlands	8.9%
Scandinavian Norway	7.1%
Ashkenazi Jewish	0.7%
Iberian	0.6%
Eastern European	0.6%
Broadly Northwestern European	38.6%
Broadly Southern European	2.2%
Broadly European	6.8%
East Asian & Native American	3.7%
Native American	2.9%
Broadly East Asian	0.1%
Broadly East Asian & Native American	0.7%
Western Asian & North African	2.8%
Western Asian	2.7%
Broadly Western Asian & North African	0.1%
South Asian	0.5%
Broadly South Asian	0.5%
Unassigned	0.7%



# What is genetic ancestry?

- In general: genetic ancestry is defined by the stretches of the genome that we inherit from our ancestors
- Many definitions, interpretations and misinterpretations
- Continental ancestry the most common for humans
- Compare to “reference” databases of individual samples from diverse populations and genetic regions
- Depends on what specific “time slice” we’re focusing on





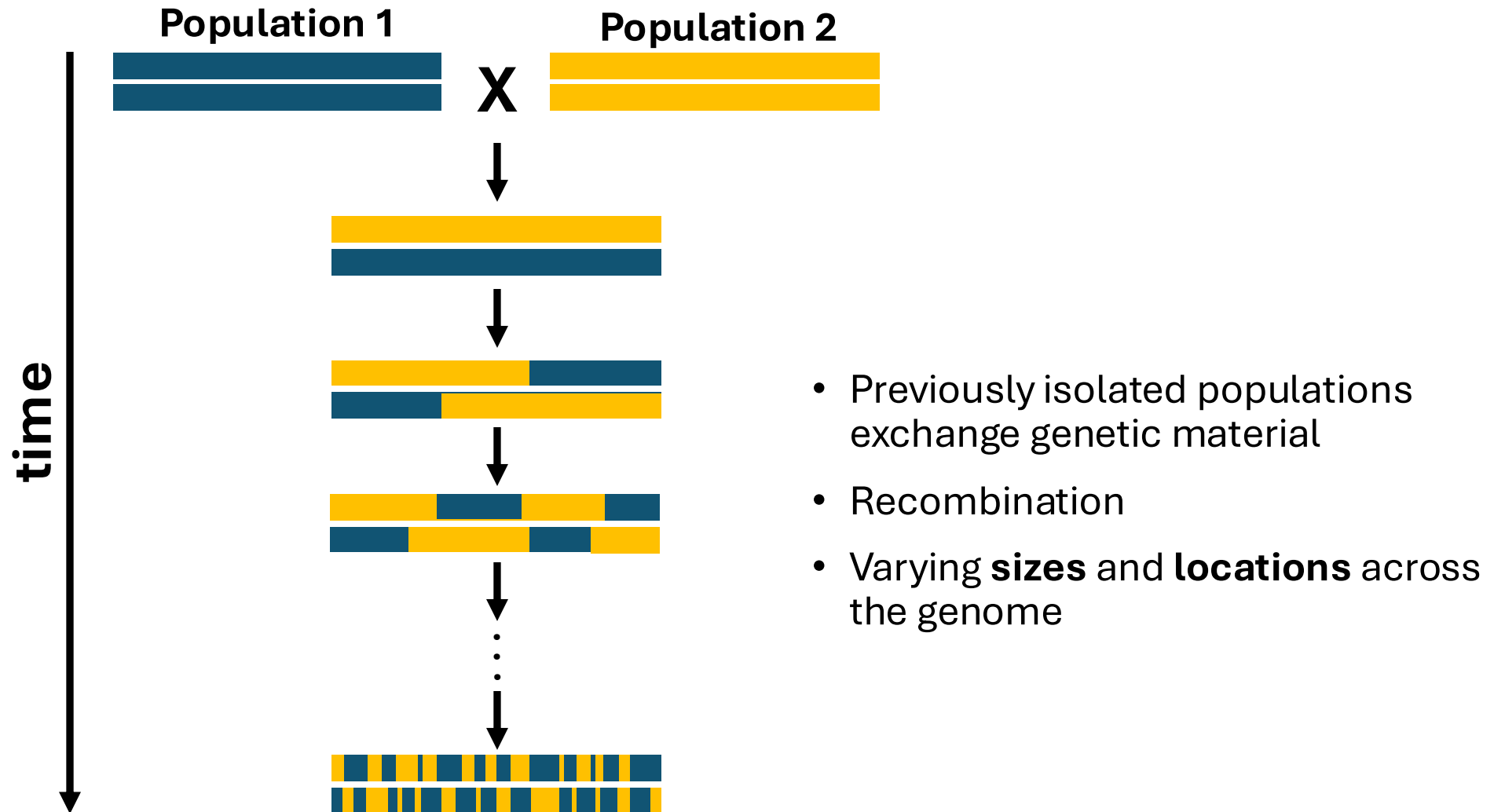


# Genomes are mosaics...

...of segments inherited from different  
ancestral populations at different time points



# Admixture





# Global Ancestry vs Local Ancestry



- **Proportion** of ancestry contributed by different populations
- Averaged across the genome (overall admixture proportions)
- E.g.: Proportion of ANC1=30%

Global ancestry



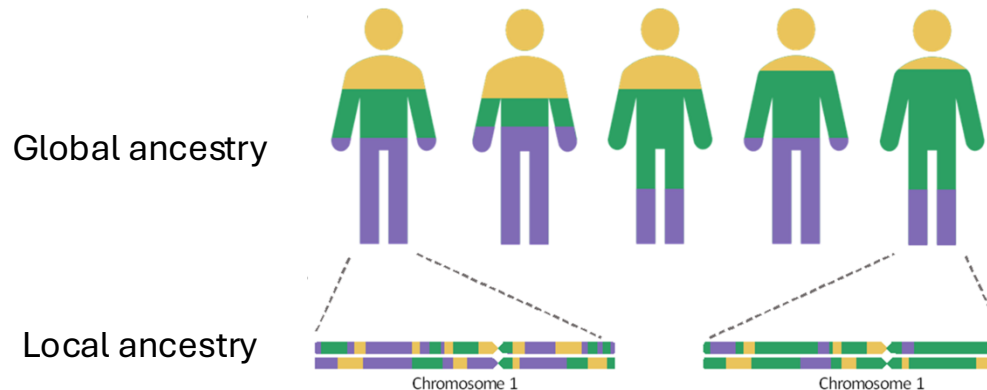


# Global Ancestry vs Local Ancestry



- **Proportion** of ancestry contributed by different populations
- Averaged across the genome (overall admixture proportions)
- E.g.: Proportion of ANC1=30%

- **Regional** ancestry of a genomic segment
- Infer ancestral population of origin at each position
- E.g.:  
chr1:1-4,868,294 ANC1  
chr1:4,868,295-8,275,948 ANC2





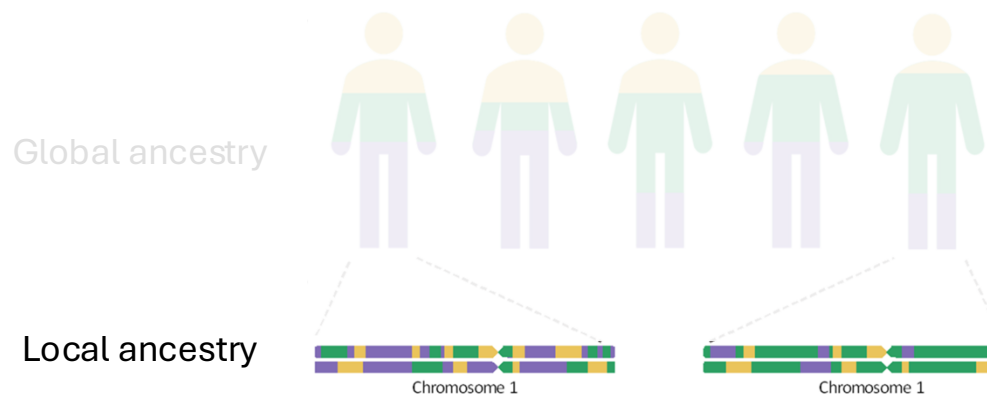


# Global Ancestry vs Local Ancestry



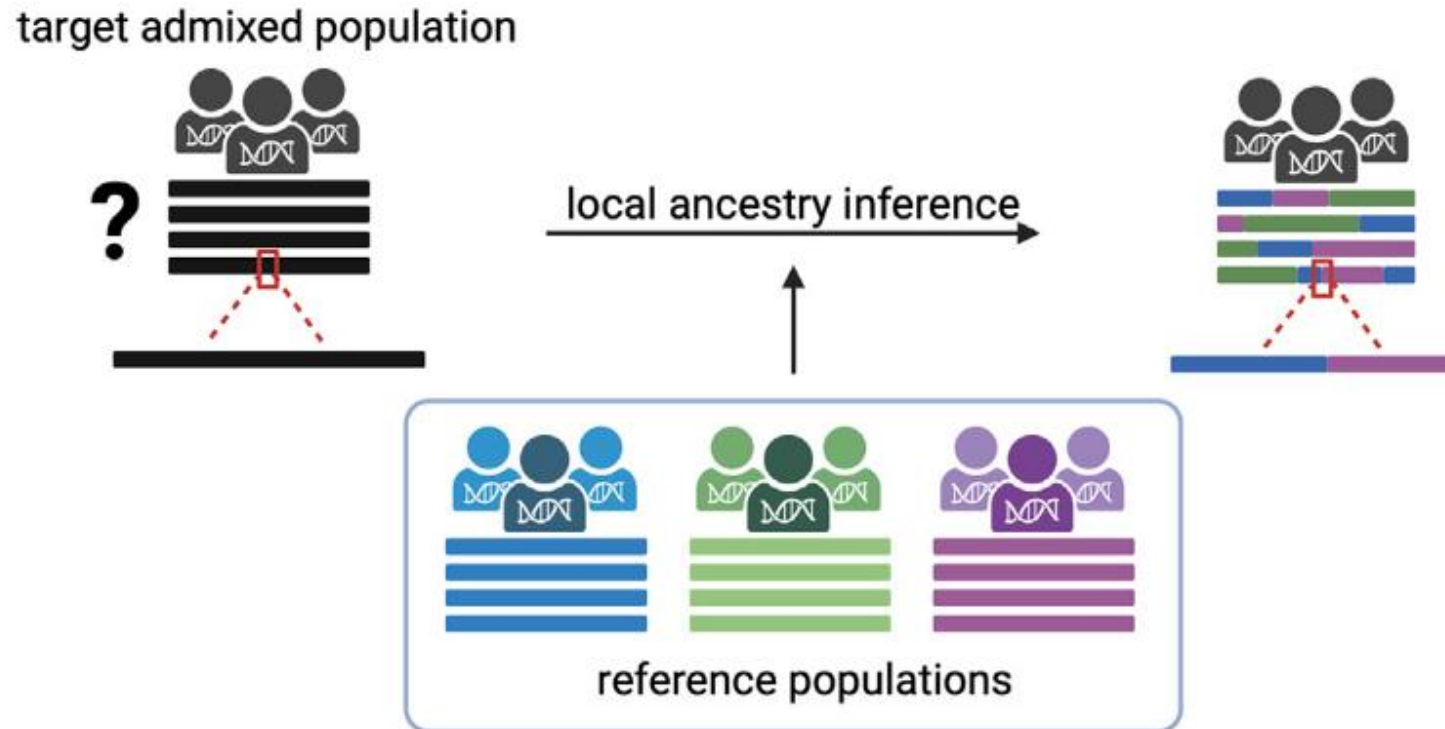
- Proportion of ancestry contributed by different populations
- Averaged across the genome (overall admixture proportions)
- E.g.: Proportion of ANC1=30%

- Regional ancestry of a genomic segment
- Infer ancestral population of origin at each position
- E.g.:  
chr1:1-4,868,294 ANC1  
chr1:4,868,295-8,275,948 ANC2



# Local Ancestry Inference (LAI)

“Chromosome painting”



# Applications of LAI

**Evolutionary Biology**

*Association between  
phenotypes and local ancestry  
("Admixture mapping")*

*History of migration and  
admixture*

*Natural selection and  
adaptation*

**Population genetics**

*More precise estimation  
of polygenic risk scores*

*Pharmacogenomics*

*Assist GWAS and QTL  
studies to identify  
ancestry-differentiated  
risk variants*

**Medical genetics**

# Applications of LAI

**Evolutionary Biology**

*Association between  
phenotypes and local ancestry  
("Admixture mapping")*

*History of migration and  
admixture*

**Population genetics**

*More precise estimation  
of polygenic risk scores*

*Natural selection and  
adaptation*

*Pharmacogenomics*

*Assist GWAS and QTL  
studies to identify  
ancestry-differentiated  
risk variants*

**Medical genetics**



# Applications of LAI

OPEN  ACCESS Freely available online

 **PLOS** | GENETICS

## Reconstructing the Population Genetic History of the Caribbean

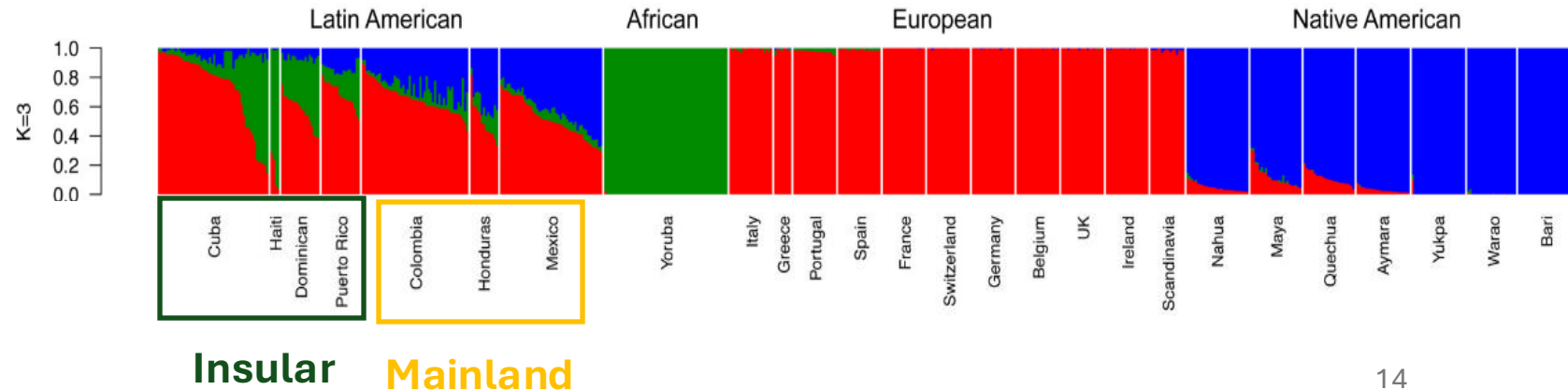
**Andrés Moreno-Estrada<sup>1</sup>, Simon Gravel<sup>1,2</sup>, Fouad Zakharia<sup>1</sup>, Jacob L. McCauley<sup>3</sup>, Jake K. Byrnes<sup>1,4</sup>, Christopher R. Gignoux<sup>5</sup>, Patricia A. Ortiz-Tello<sup>1</sup>, Ricardo J. Martínez<sup>3</sup>, Dale J. Hedges<sup>3</sup>, Richard W. Morris<sup>3</sup>, Celeste Eng<sup>5</sup>, Karla Sandoval<sup>1</sup>, Suehelay Acevedo-Acevedo<sup>6</sup>, Paul J. Norman<sup>7</sup>, Zulay Layrisse<sup>8</sup>, Peter Parham<sup>7</sup>, Juan Carlos Martínez-Cruzado<sup>6</sup>, Esteban González Burchard<sup>5</sup>, Michael L. Cuccaro<sup>3</sup>, Eden R. Martin<sup>1,3\*</sup>, Carlos D. Bustamante<sup>1,1\*</sup>**

**1** Department of Genetics, Stanford University School of Medicine, Stanford, California, United States of America, **2** Department of Human Genetics and Genome Quebec Innovation Centre, McGill University, Montreal, Québec, Canada, **3** Center for Genetic Epidemiology and Statistical Genetics, John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, Florida, United States of America, **4** Ancestry.com DNA, LLC, San Francisco, California, United States of America, **5** Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, California, United States of America, **6** Department of Biology, University of Puerto Rico at Mayaguez, Mayaguez, Puerto Rico, **7** Department of Structural Biology, Stanford University School of Medicine, Stanford, California, United States of America, **8** Center of Experimental Medicine “Miguel Layrisse”, IVIC, Caracas, Venezuela

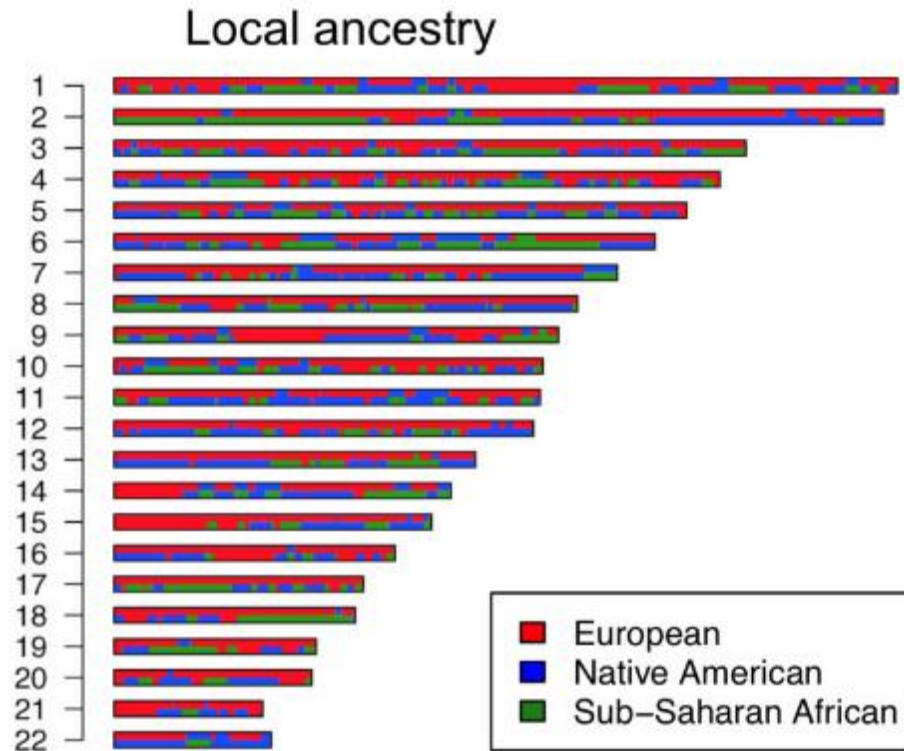
# Reconstructing the Population Genetic History of the Caribbean



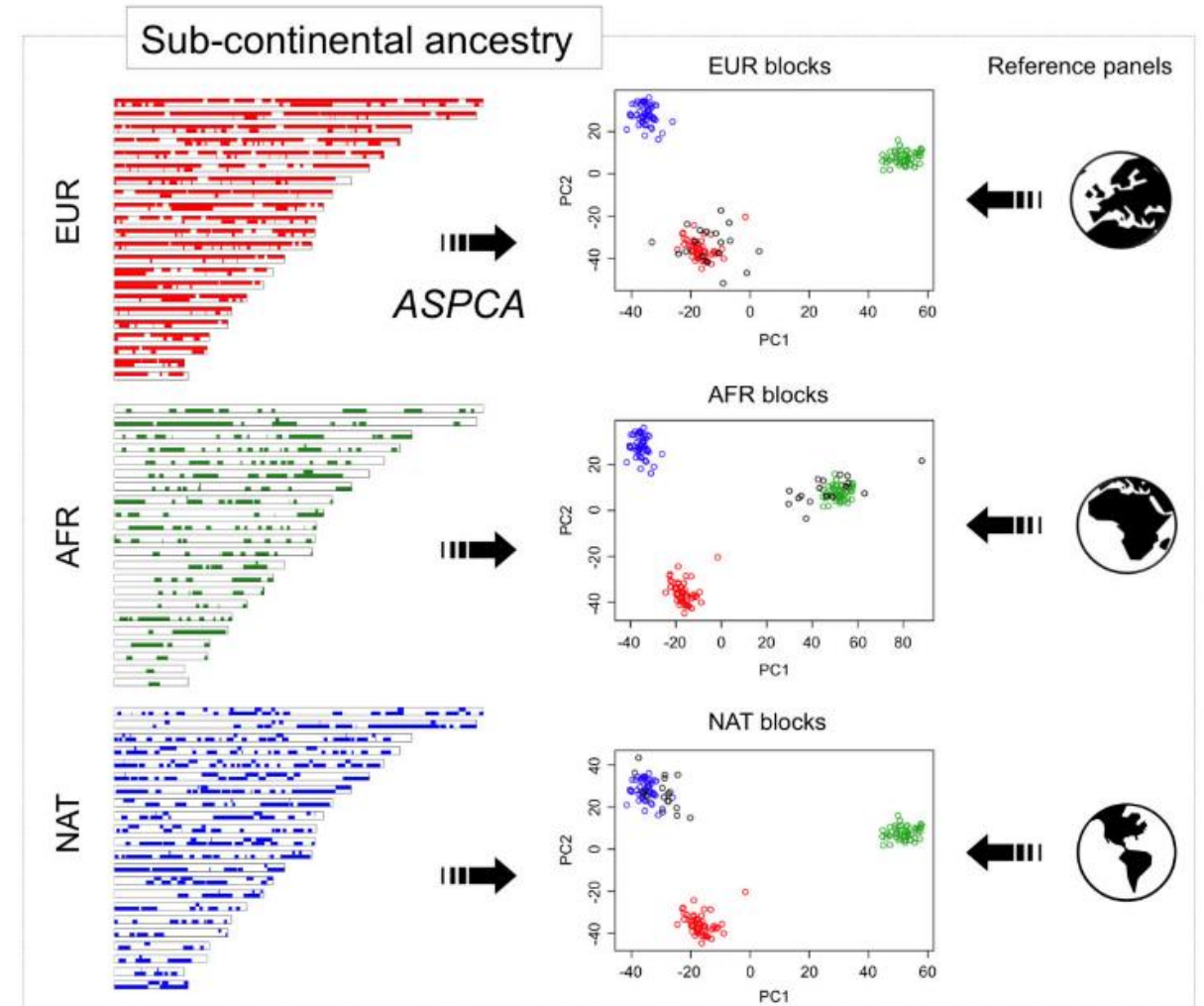
- Complex recent population history
- Severe bottlenecks of Native pre-Columbian populations
- Encounter between Native Americans and Europeans in the Caribbean before the mainland
- Initial destination for the trans-Atlantic slave trade



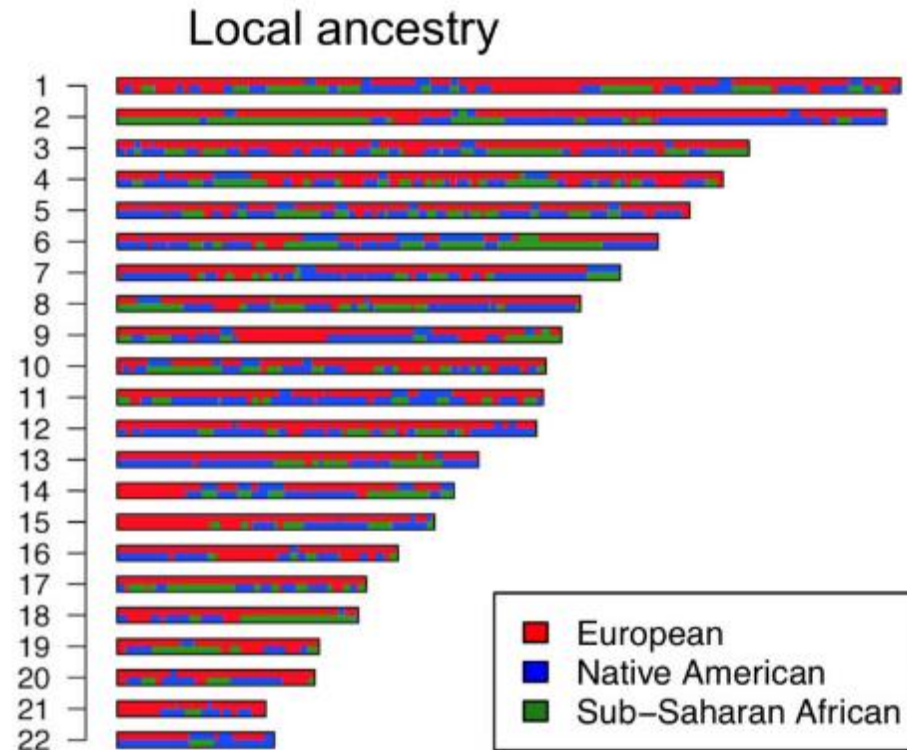
# Reconstructing the Population Genetic History of the Caribbean



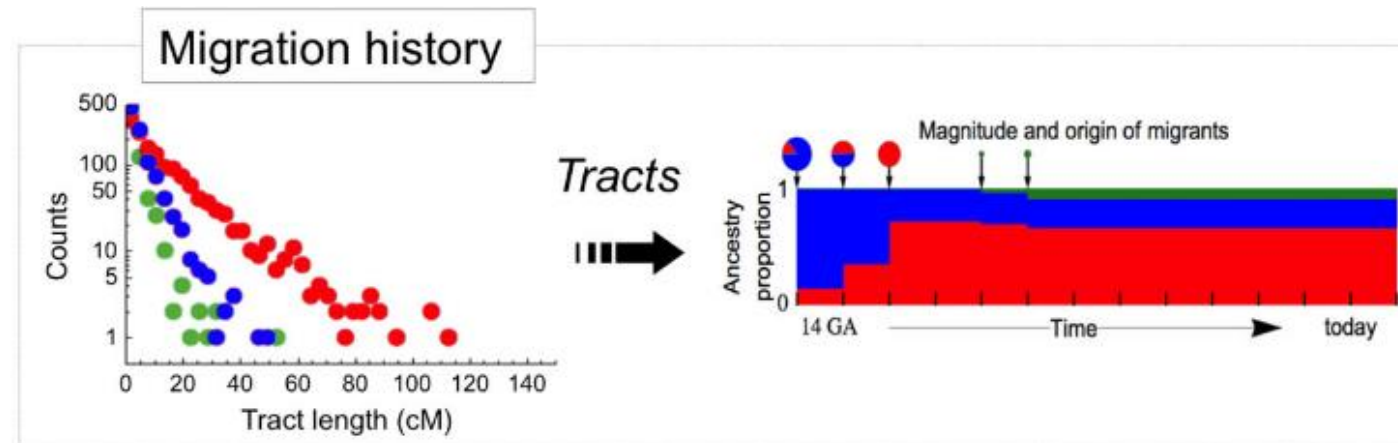
PCAdmix (Brisbin et al, 2012)



# Reconstructing the Population Genetic History of the Caribbean



PCAdmix (Brisbin et al, 2012)



Tracts (Gravel, 2012)

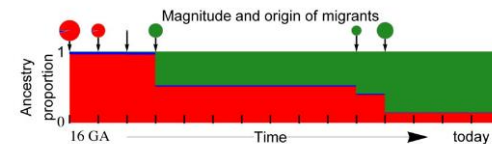
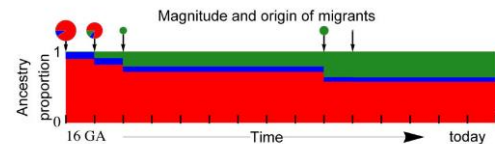
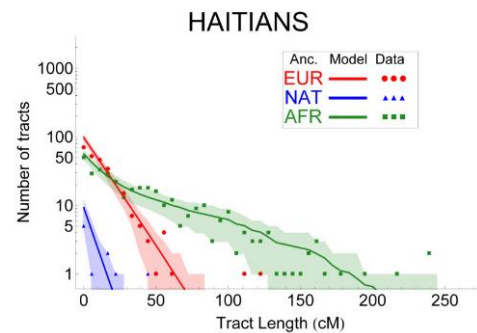
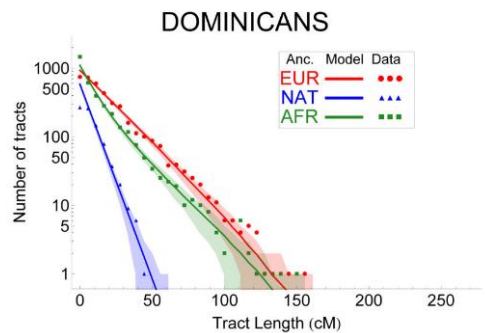
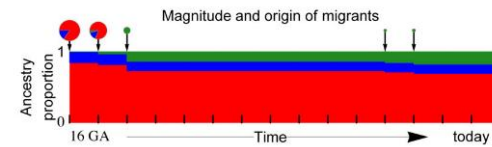
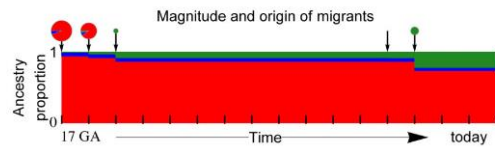
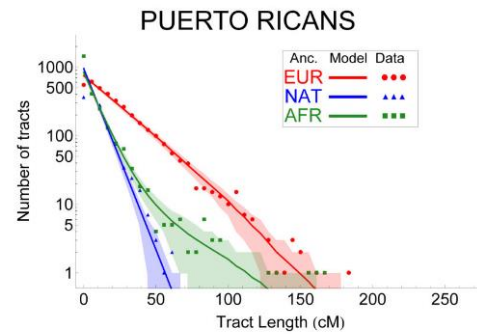
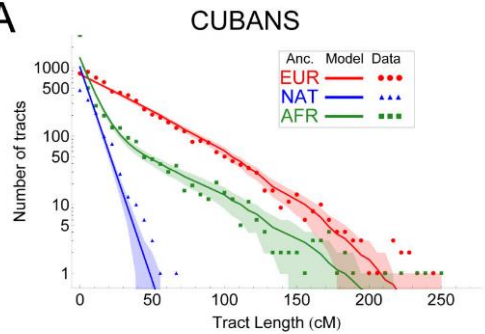


# Reconstructing the Population Genetic History of the Caribbean

## Insular Caribbean

## Mainland Caribbean

A



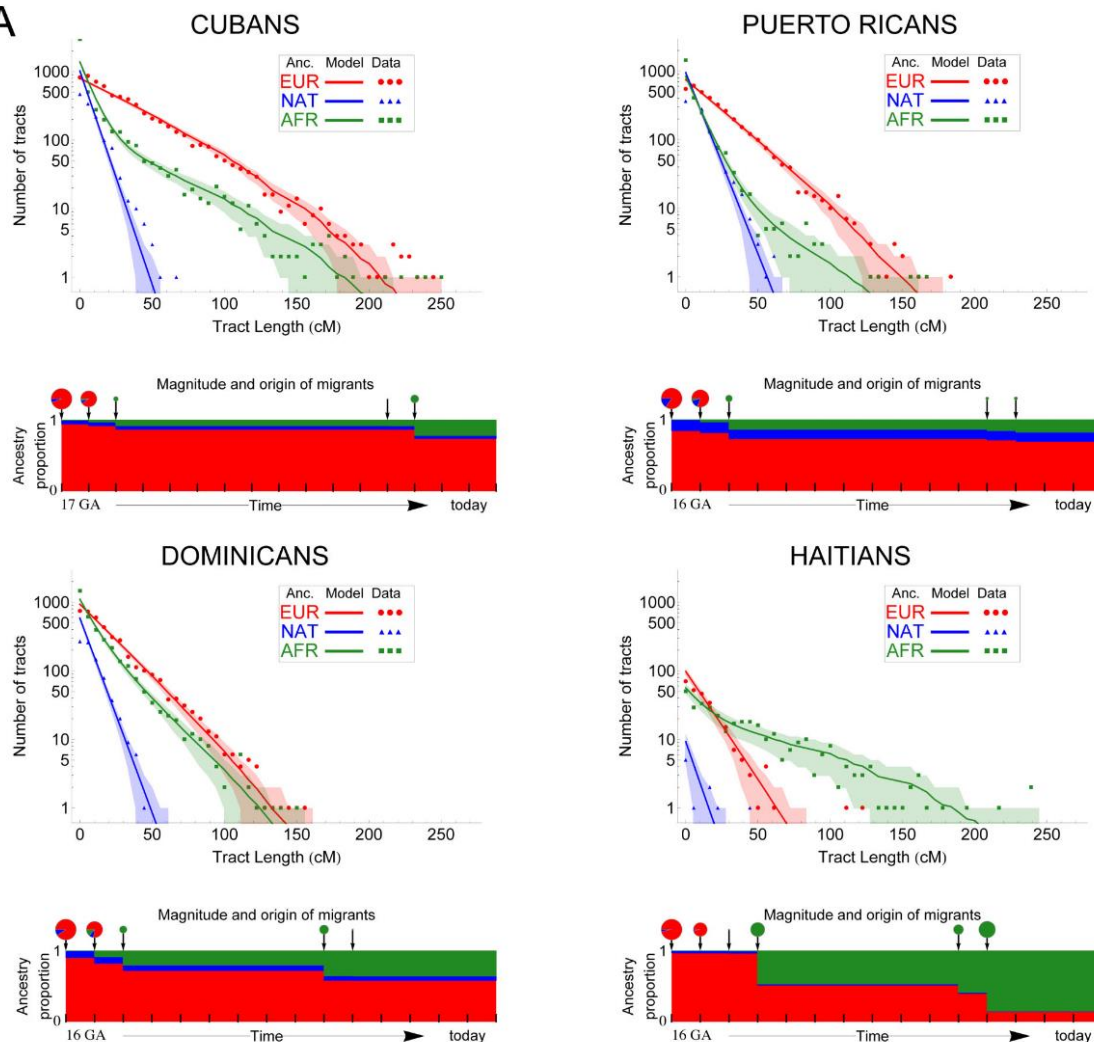
Insular: Two pulses of African ancestry

# Reconstructing the Population Genetic History of the Caribbean

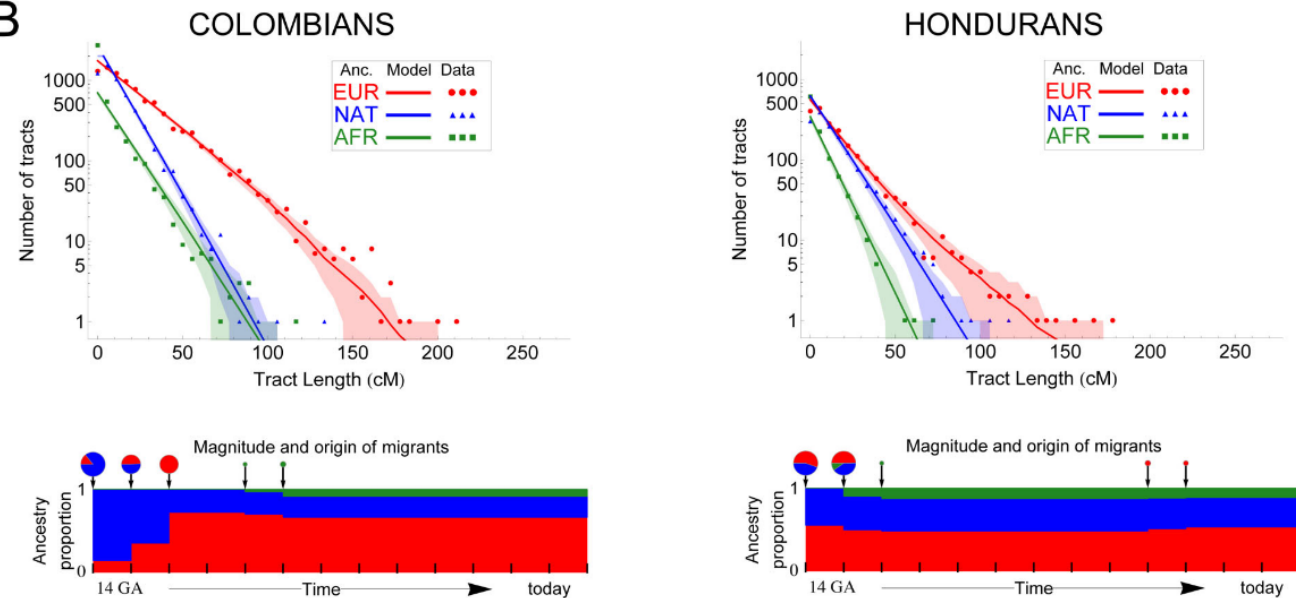
## Insular Caribbean

## Mainland Caribbean

A



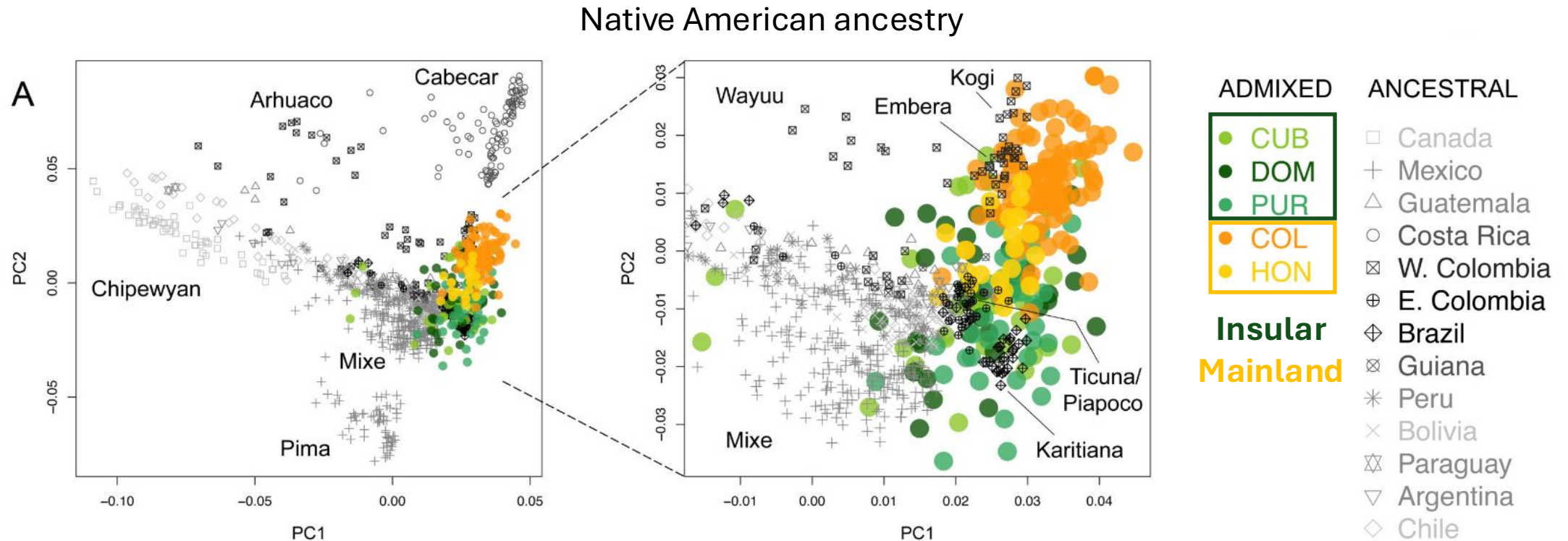
B



Insular: Two pulses of African ancestry  
Mainland: Two pulses of European ancestry

Onset of admixture is older in insular populations (16-17 generations ago) vs mainland populations (14 generations ago)

# Reconstructing the Population Genetic History of the Caribbean

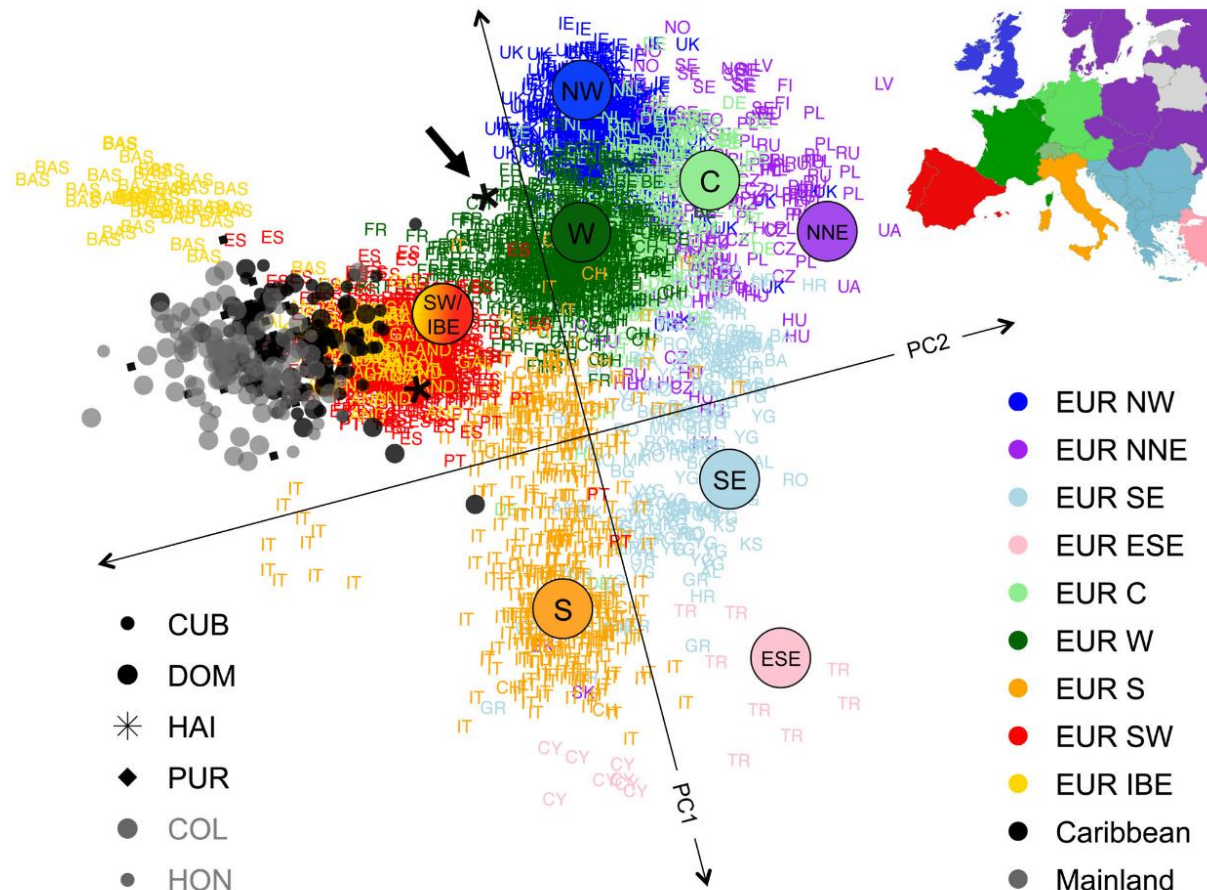


Mainland (Columbia & Honduras) cluster with groups from Western Columbia and Central America

Most Caribbean islanders cluster with Amazonian groups from Eastern Columbia, Brazil and Guiana

# Reconstructing the Population Genetic History of the Caribbean

European ancestry



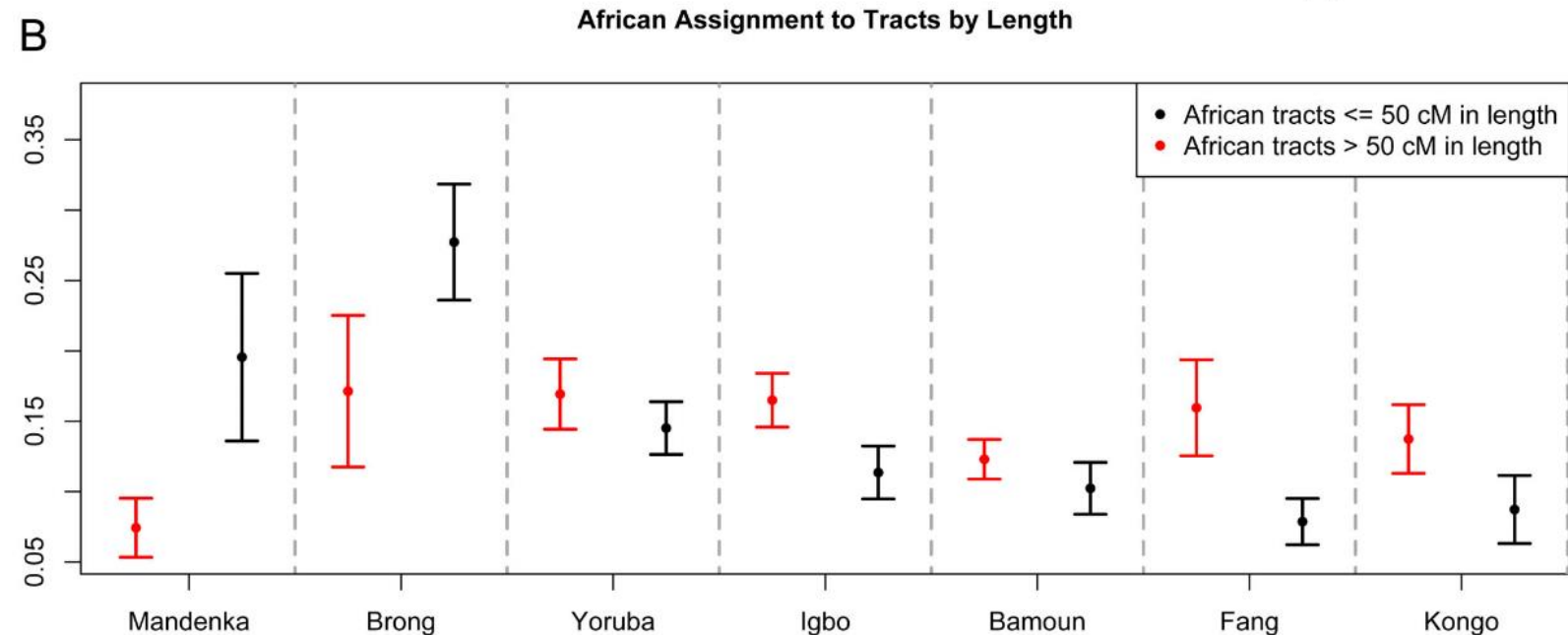
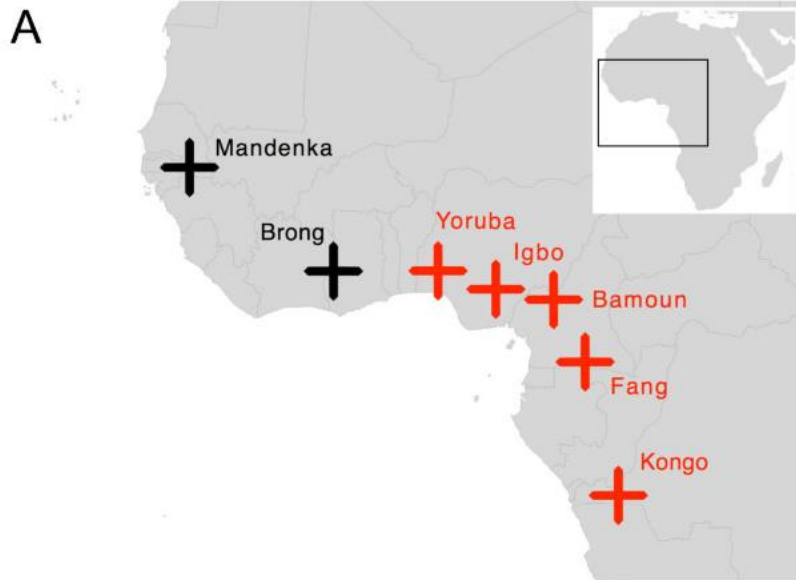
Clusters with samples from the Iberian Peninsula

Deviation of Latino samples from the centroid of the Iberian cluster → bottleneck & drift of Europeans



# Reconstructing the Population Genetic History of the Caribbean

## African ancestry



Higher proportion of African ancestry in Caribbean than mainland

Two pulses of African ancestry in Caribbean islands

Shorter ancestry tracts more likely from northern West Africa (older event)

Longer ancestry tracts more likely from central West Africa coast (more recent event)



Great Lakes wolf



Red wolf

# Applications of LAI (non-human)

“A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids”



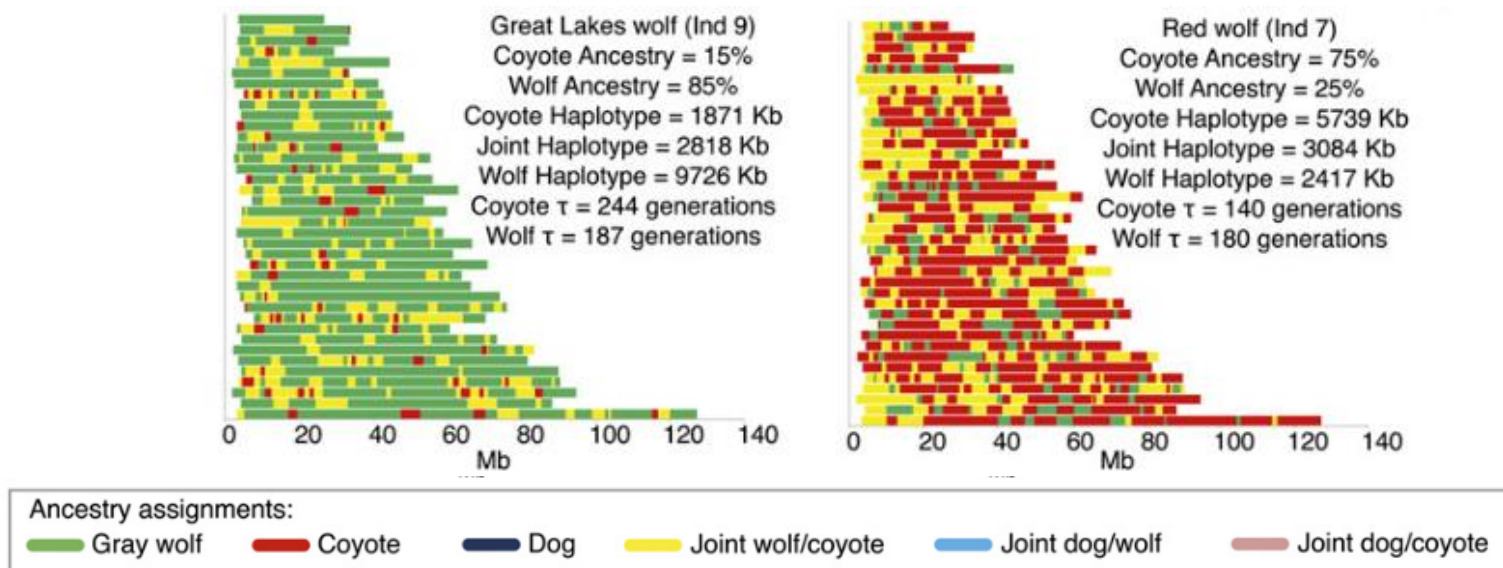
Great Lakes wolf



Red wolf

# Applications of LAI (non-human)

“A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids”



- LAI approach to look at past admixture between North American canids
- Great Lake wolf and red wolf are admixed varieties derived from gray wolves and coyotes respectively

# Applications of LAI (non-human)

“Genome-wide local ancestry and the functional consequences of admixture in African and European cattle populations”



Taurine cattle



Indicine cattle

Diverged 150-500kya

- Domestication and spread of cattle has led to gradients of taurine and indicine ancestry
- LAI to look at hybrid African and admixed Southern European cattle populations





Taurine cattle

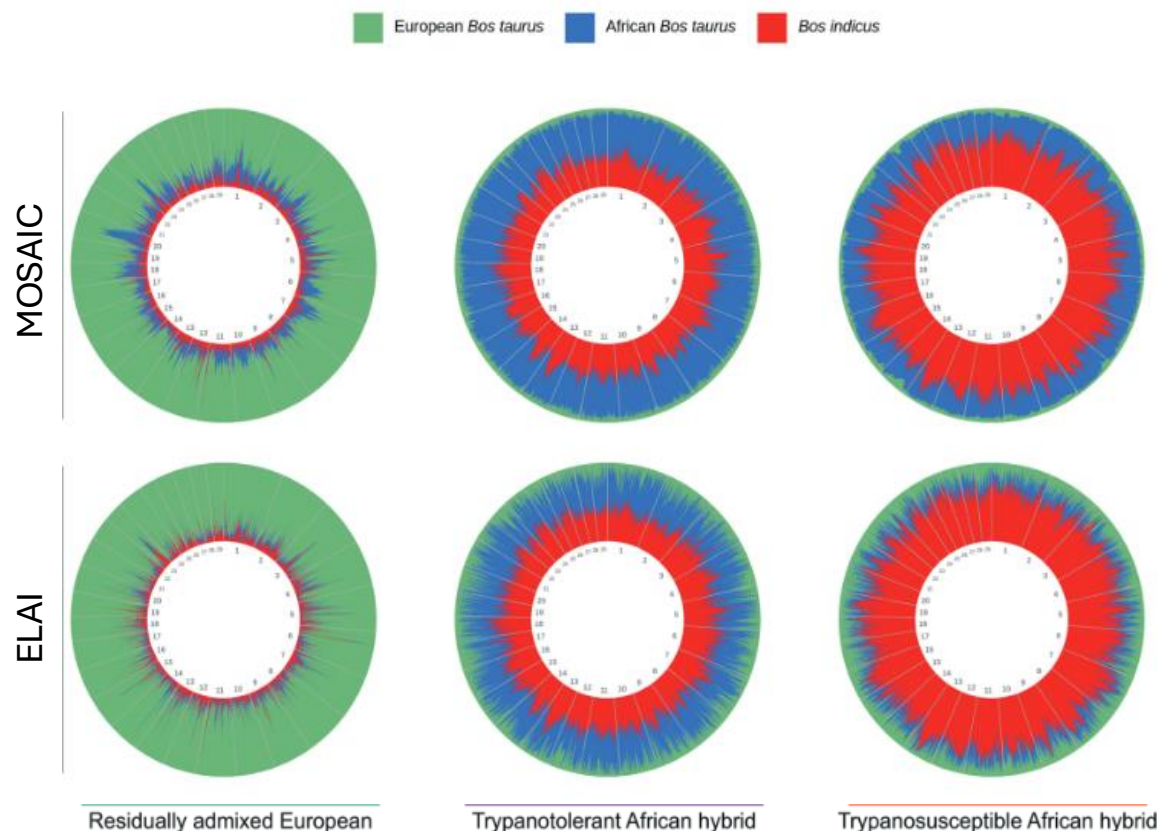


Indicine cattle

Diverged 150-500kya

# Applications of LAI (non-human)

“Genome-wide local ancestry and the functional consequences of admixture in African and European cattle populations”



- Introgressed ancestry from African taurine, European taurine and Asian indicine.
- Increased African taurine ancestry in admixed trypanotolerant cattle compared to trypanosusceptible cattle
- Enrichment of genes associated with haemoglobin and oxygen transport in retained African taurine ancestry in admixed trypanotolerant cattle

# Applications of LAI

**Evolutionary Biology**

*Association between  
phenotypes and local ancestry  
("Admixture mapping")*

*History of migration and  
admixture*

**Population genetics**

*More precise estimation  
of polygenic risk scores*

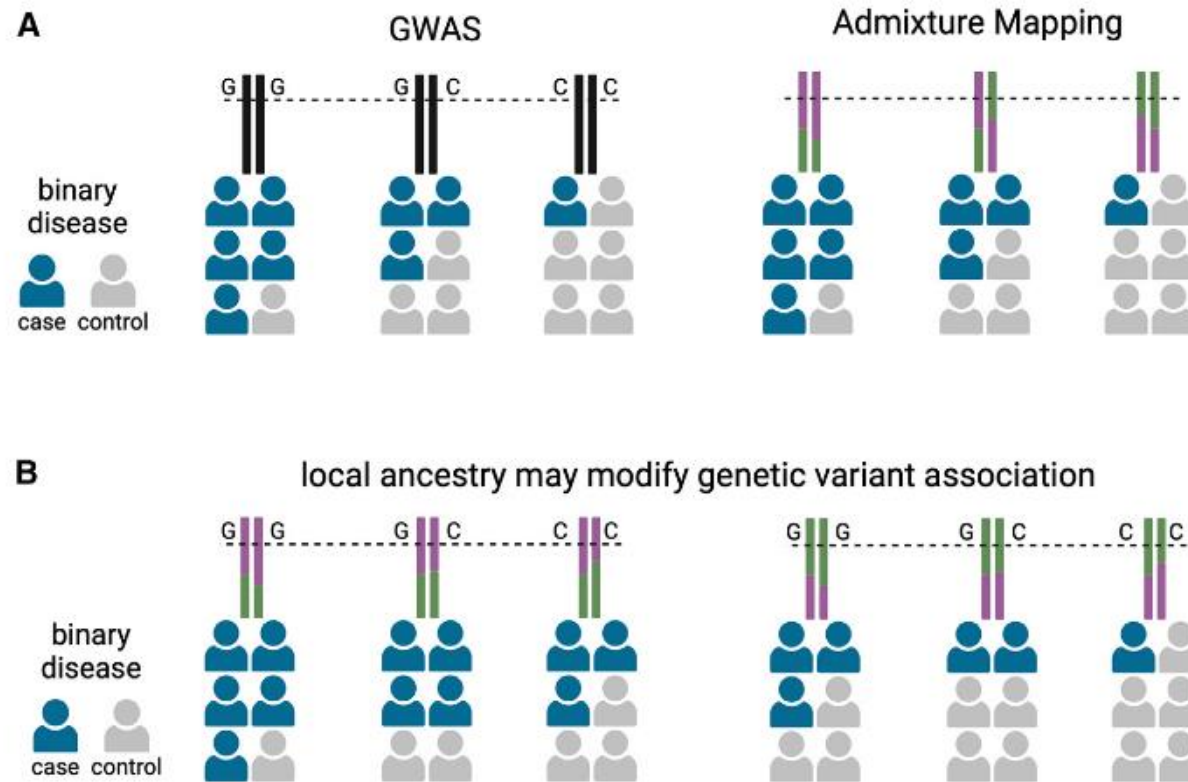
*Natural selection and  
adaptation*

*Assist GWAS and QTL  
studies to identify  
ancestry-differentiated  
risk variants*

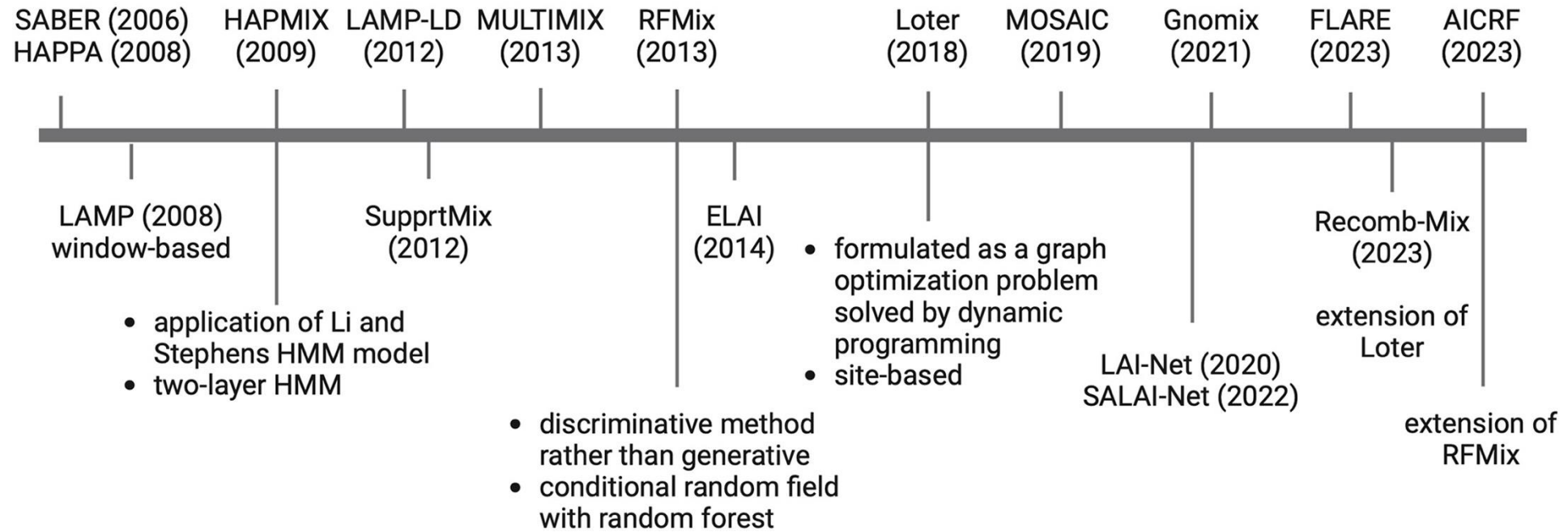
*Pharmacogenomics*

**Medical genetics**

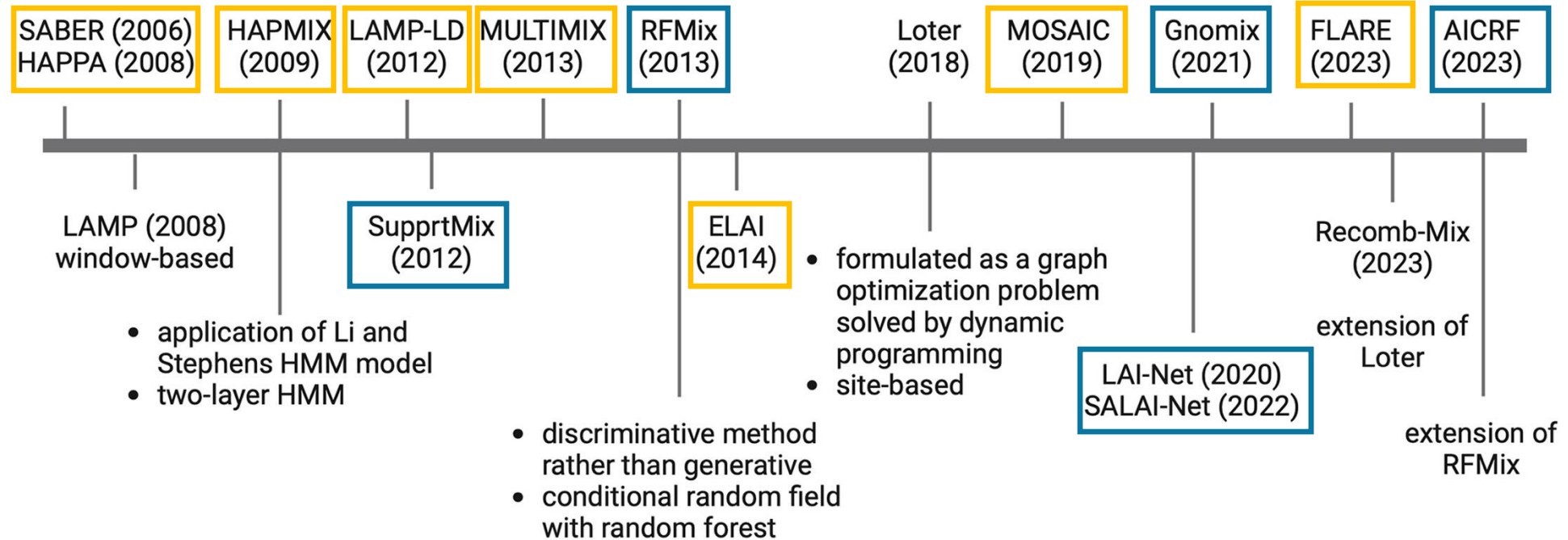
# Applications of LAI



# LAI methods




# LAI methods



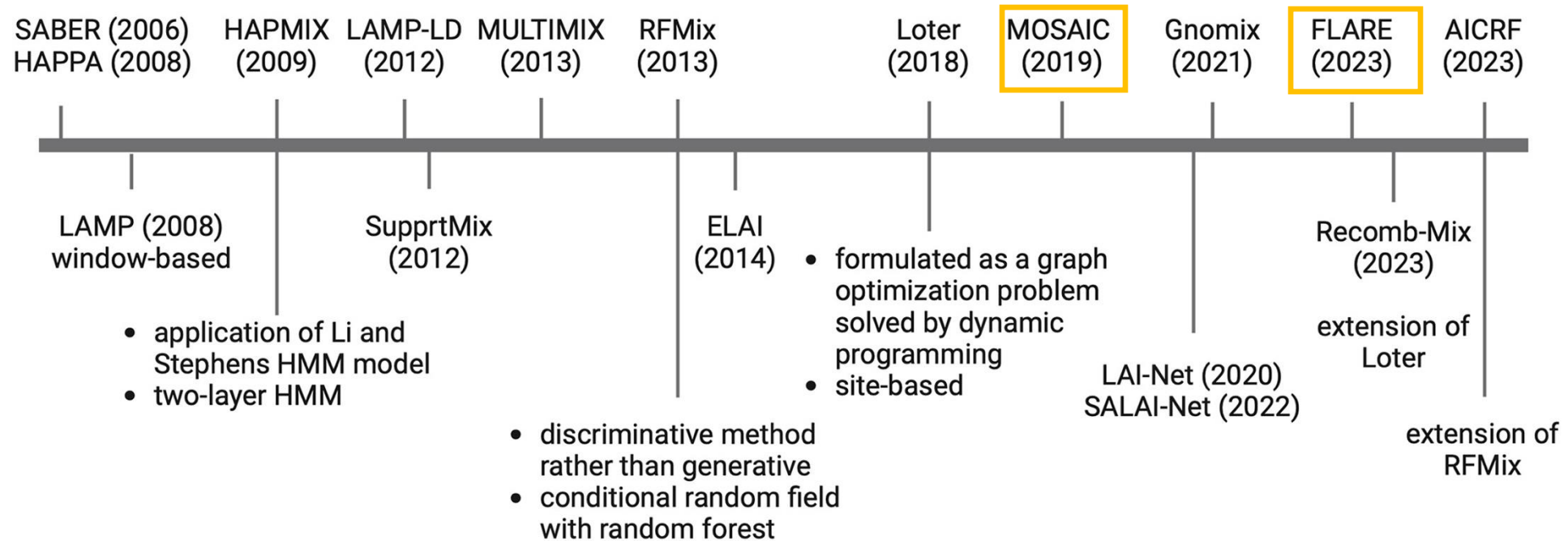
# LAI methods

## Not one to rule them all

- 
- Not obvious how to select the best approach
  - Some faster, some slower
  - Some better for more recent admixture events
  - Some better for older admixture events
  - Some good for non-model species
  - Some good when donor reference populations are not direct surrogates for the original source populations
  - Some better for smaller reference populations sample sizes
  - Some don't require reference populations
- SABER (2006) HAPMIX (2009) LAMP (2008) window-based application of Li and Nei (2001) two-layer HMM (2012) SupprtMix (2012) ELAI (2014) discriminative method (2014) conditional random field with random forest (2014) RFMix (2013) site-based (2013) optimization problem solved by dynamic programming (2013) SALAI-Net (2022) LAFNet (2020) extension of RFMix (2023) Recomb-Mix (2023) extension of (2023) FLARE (2023) AI (2023)



# For today's exercise



# Commonly required input for LAI

- Admixed samples (usually phased)
- Good proxy reference panels (phased)
- Genetic map
- Input admixture time (sometimes)

# Keep in mind when doing LAI

- Reference panel sample sizes
- Ancestral representation present in reference panels
- Evolutionary history of the admixed population
- How far back is your admixture event
- Generation time of your species
- Differentiation of your source populations
- Methods differ, not one for all
- Computational time





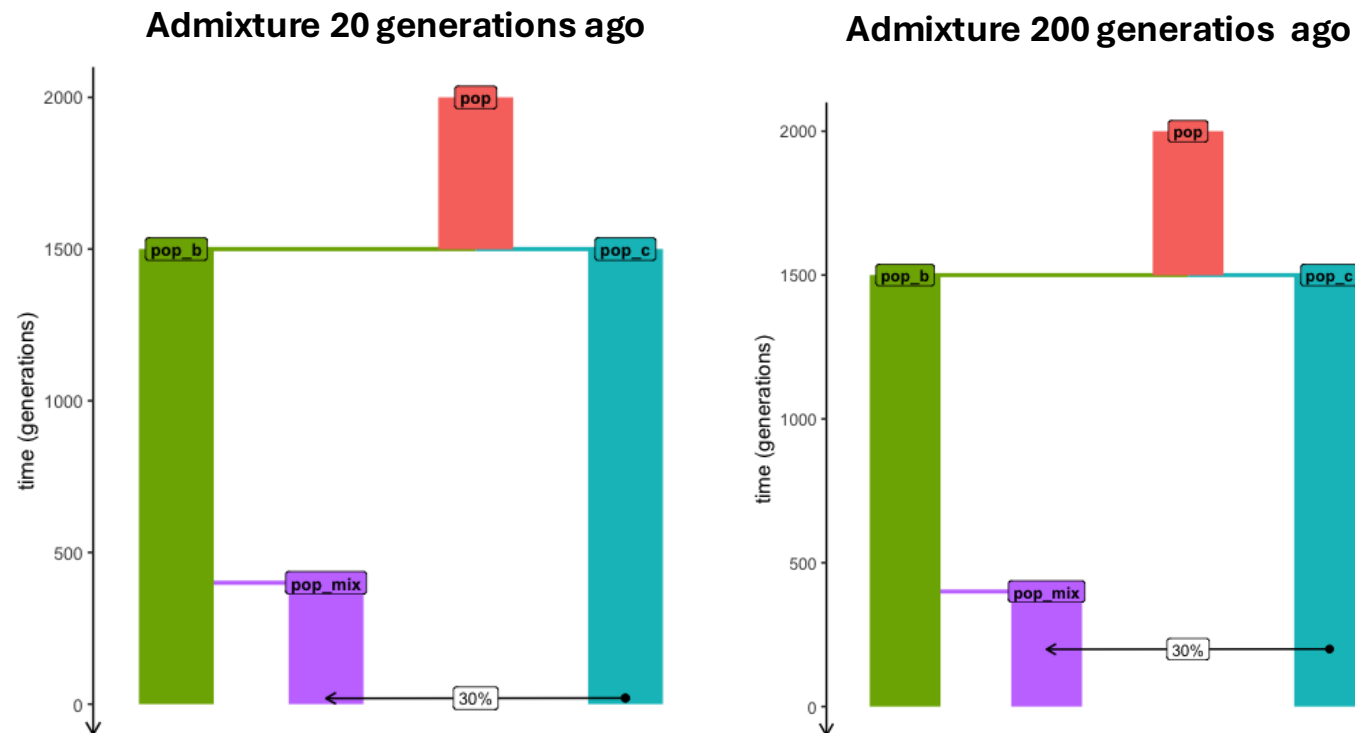
# Time for exercises!





# Time for exercises!

- Go to [https://emily.popgen.dk/user/katia/notebooks/LocalAncestry\\_Exercise.ipynb](https://emily.popgen.dk/user/katia/notebooks/LocalAncestry_Exercise.ipynb)
- We will use simulated data based on two demographic scenarios
- We will use two LAI methods: FLARE and MOSAIC
- We will test how accurate LAI is at different admixture times, sample sizes and reference population divergence



# References

- Royal, Charmaine D., et al. "Inferring genetic ancestry: opportunities, challenges, and implications." *The American Journal of Human Genetics* 86.5 (2010): 661-673
- Suarez-Pajes, Eva, et al. "Genetic ancestry inference and its application for the genetic mapping of human diseases." *International journal of molecular sciences* 22.13 (2021): 6962
- Padhukasahasram, Badri. "Inferring ancestry from population genomic data and its applications." *Frontiers in genetics* 5 (2014): 204
- Sun, Quan, et al. "Opportunities and challenges of local ancestry in genetic association analyses." *The American Journal of Human Genetics* 112.4 (2025): 727-740
- Moreno-Estrada, Andrés, et al. "Reconstructing the population genetic history of the Caribbean." *PLoS genetics* 9.11 (2013): e1003925.
- vonHoldt, Bridgett M., et al. "A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids." *Genome research* 21.8 (2011): 1294-1305
- McHugo, Gillian P., et al. "Genome-wide local ancestry and the functional consequences of admixture in African and European cattle populations." *Heredity* 134.1 (2025): 49-63
- Browning, Sharon R., Ryan K. Waples, and Brian L. Browning. "Fast, accurate local ancestry inference with FLARE." *The American Journal of Human Genetics* 110.2 (2023): 326-335.
- Salter-Townshend, Michael, and Simon Myers. "Fine-scale inference of ancestry segments without prior knowledge of admixing groups." *Genetics* 212.3 (2019): 869-889