



# Human population genomics:

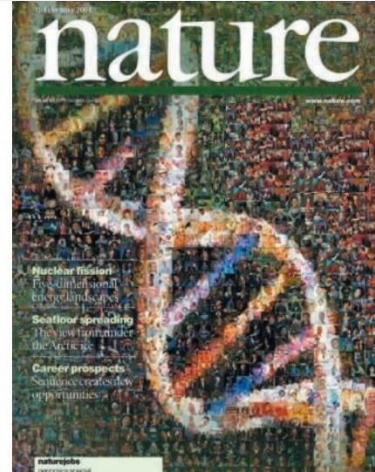
Whole genome sequencing of 175 Mongolians uncovers population-specific genetic architecture and gene flow throughout North and East Asia

兰天明

数字化地球研究所

2021.11.04

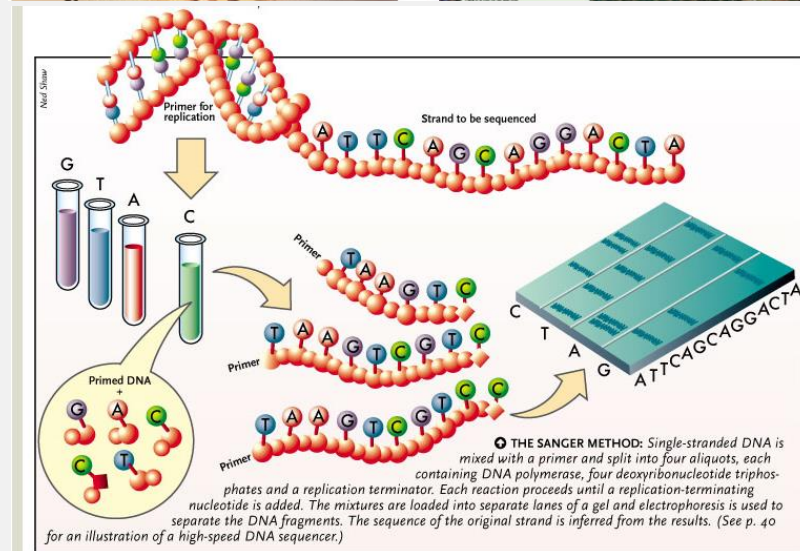
# “HGP” and BGI



- Provide human genome reference to the world

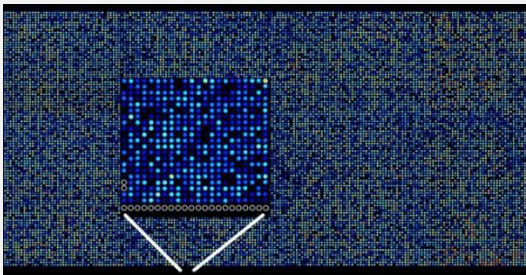


1%



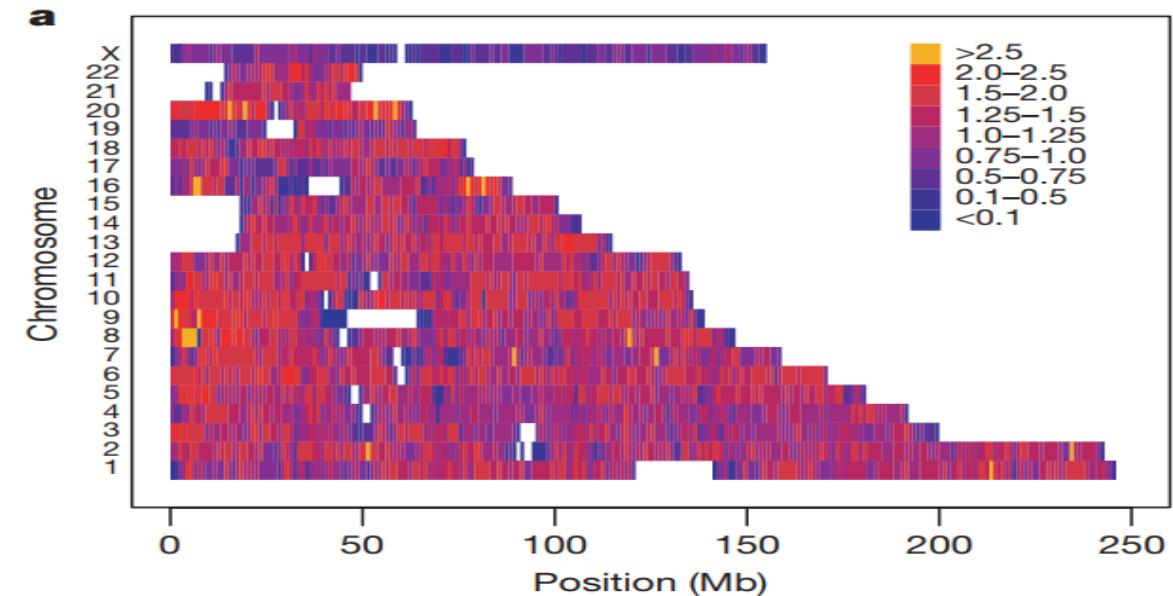
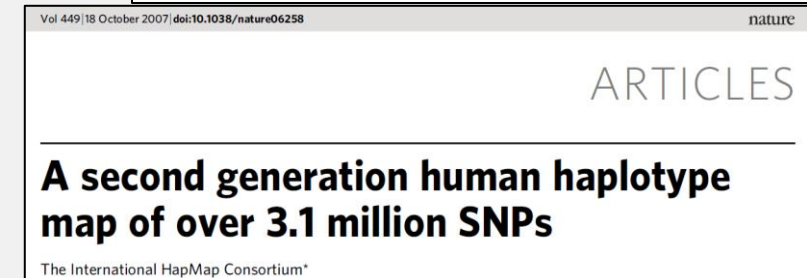
**THE SANGER METHOD:** Single-stranded DNA is mixed with a primer and split into four aliquots, each containing DNA polymerase, four deoxyribonucleotide triphosphates and a replication terminator. Each reaction proceeds until a replication-terminating nucleotide is added. The mixtures are loaded into separate lanes of a gel and electrophoresis is used to for an illustration of a high-speed DNA sequencer.)

# The HapMap Project



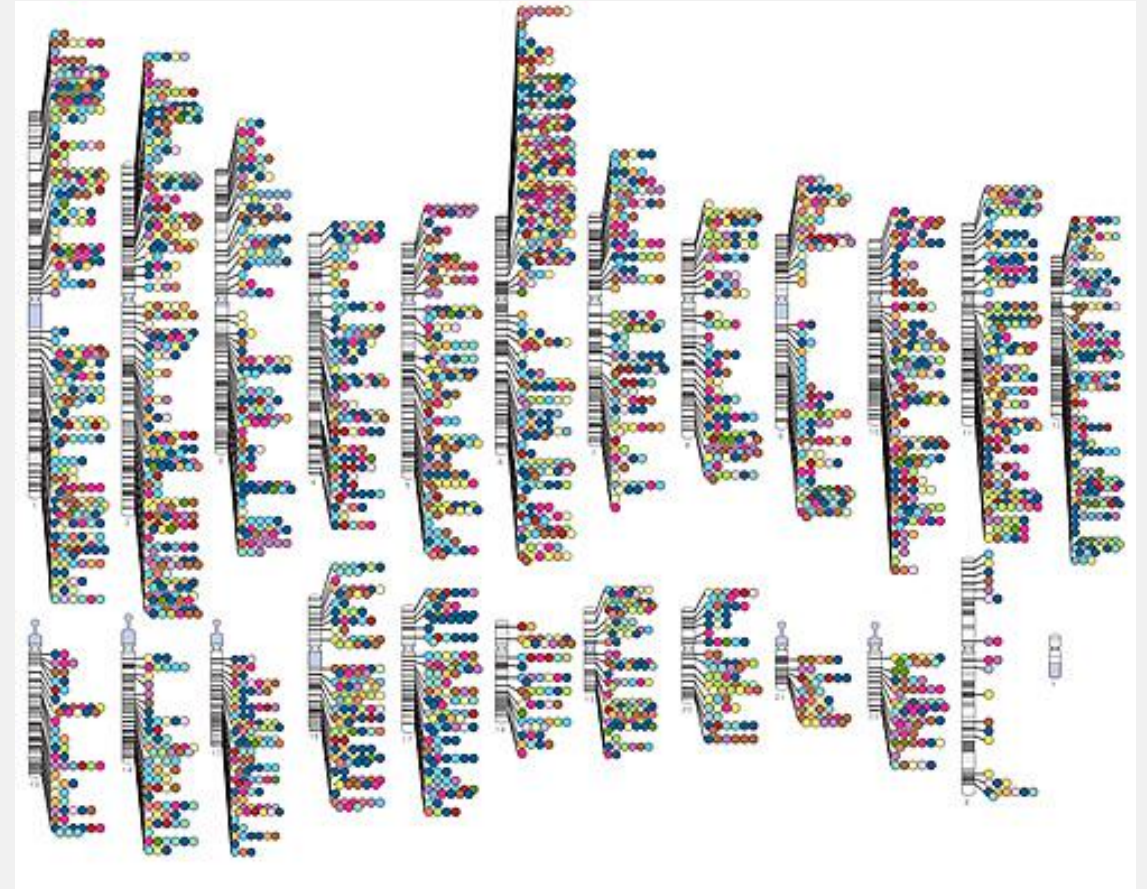
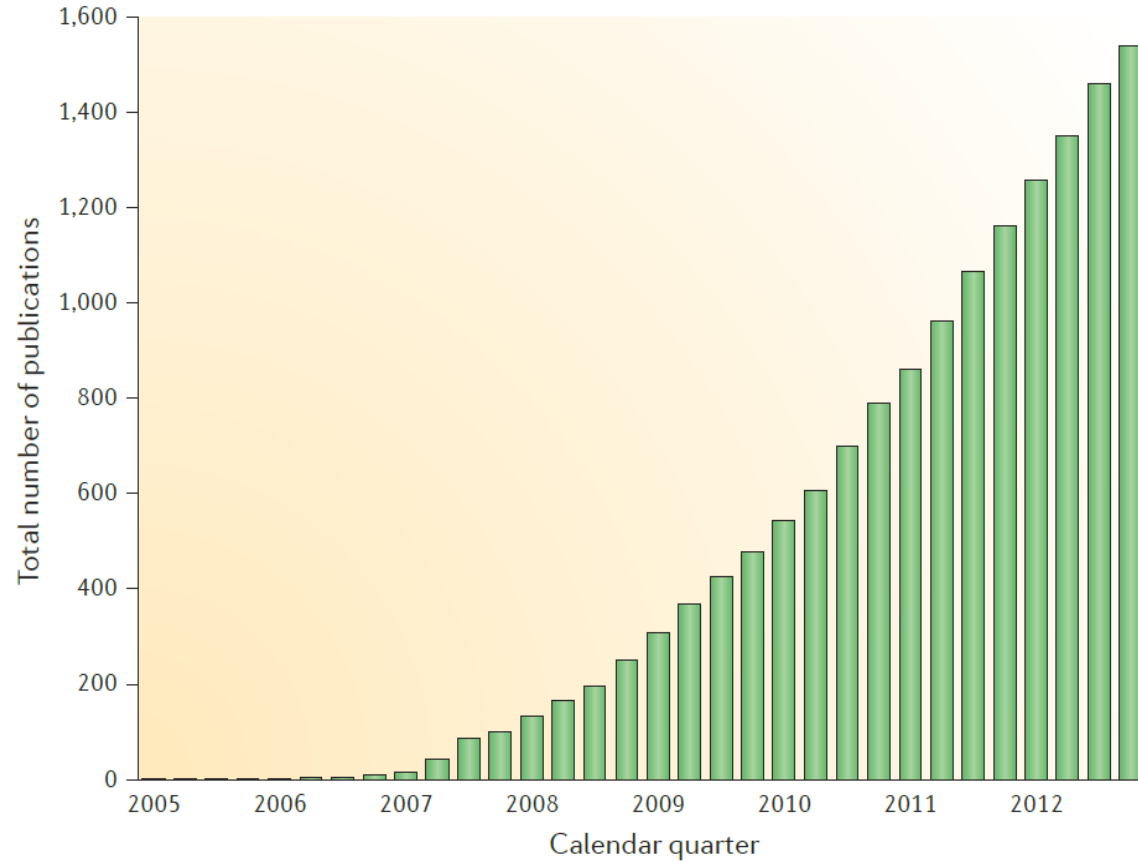
- Provide the first global genetic variation map for humans

10%





# GWAS studies after HapMap project



# Initiation of the 1000 Genomes Project



- ✓ ~8 years
- ✓ More than 2500 human individuals
- ✓ 26 global human populations
- ✓ High resolution genetic variation catalogue for global humans

## ARTICLE

doi:10.1038/nature09534

### A map of human genome variation from population-scale sequencing

The 1000 Genomes Project Consortium\*

## ARTICLE

doi:10.1038/nature11632

### An integrated map of genetic variation from 1,092 human genomes

The 1000 Genomes Project Consortium\*

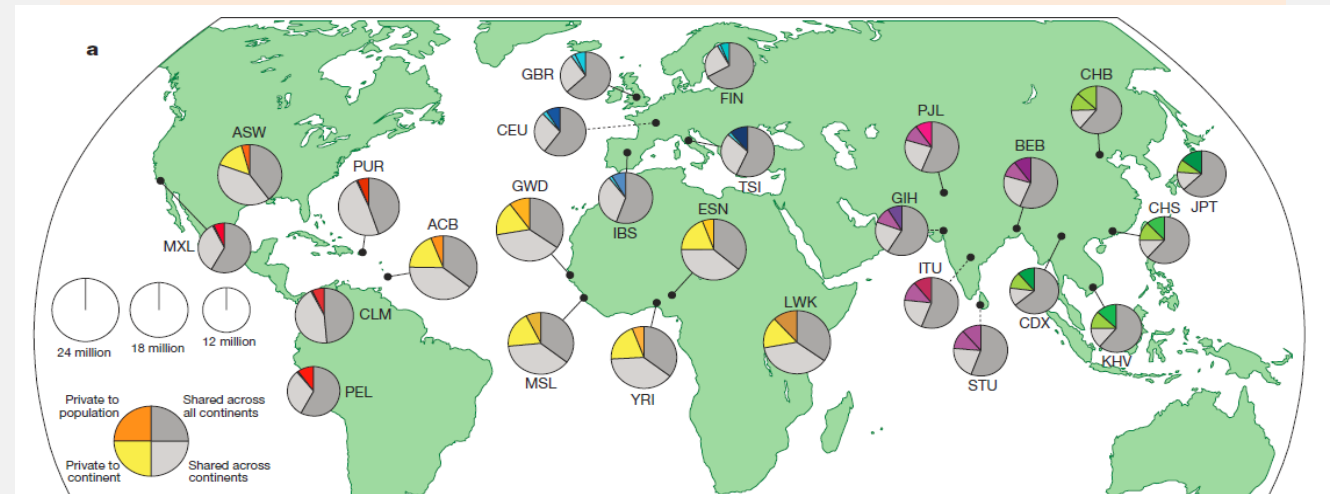
## ARTICLE

OPEN

doi:10.1038/nature15393

### A global reference for human genetic variation

The 1000 Genomes Project Consortium\*





# UK10K

## Rare Genetic Variants in Health and Disease

# ARTICLE

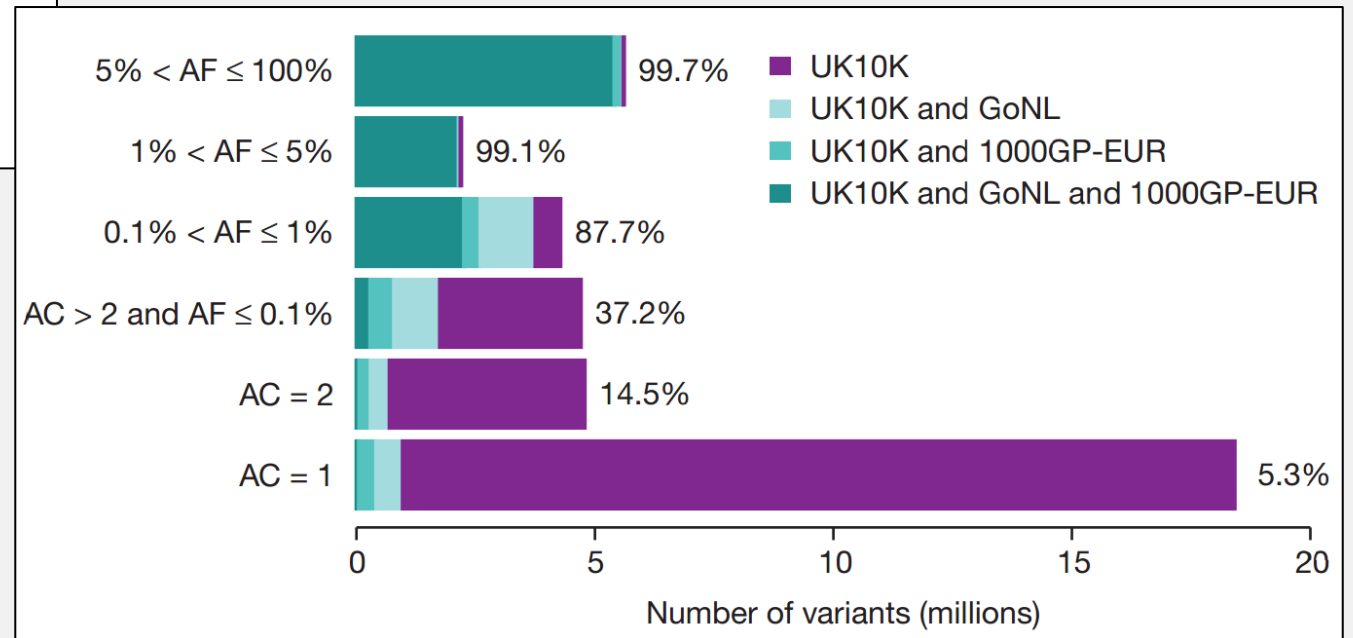
OPEN

doi:10.1038/nature14962

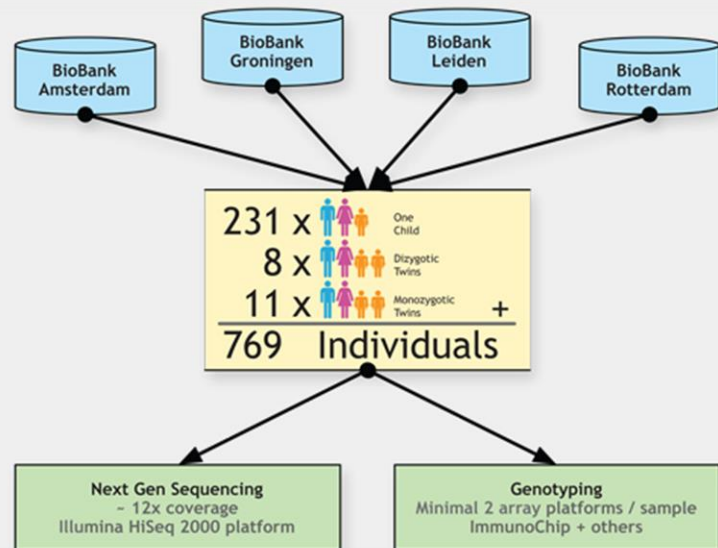
# The UK10K project identifies rare variants in health and disease

The UK10K Consortium\*

- ✓ Nearly **10,000 individuals** from population-based and disease collections
- ✓ The contribution of rare and low-frequency variants to human traits







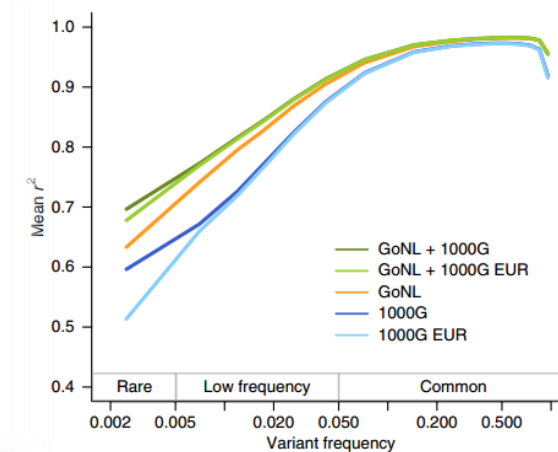
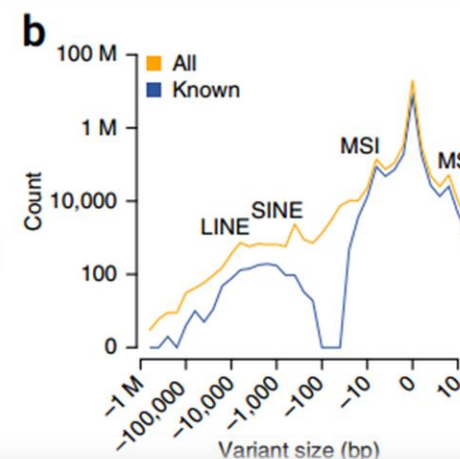
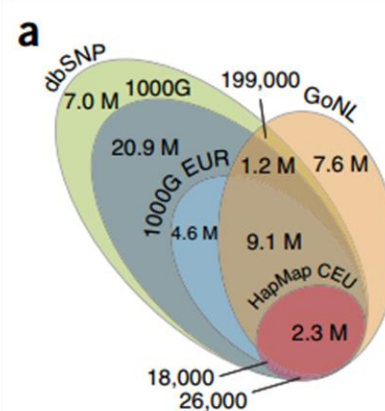
- ✓ 250 Dutch parent-offspring families
- ✓ A high quality reference panel for Netherlands population.
- ✓ The GoNL Project illustrates single population whole-genome sequencing can provide detailed characterization of genetic variation and may guide the design of future population studies.

## ARTICLES

nature  
genetics

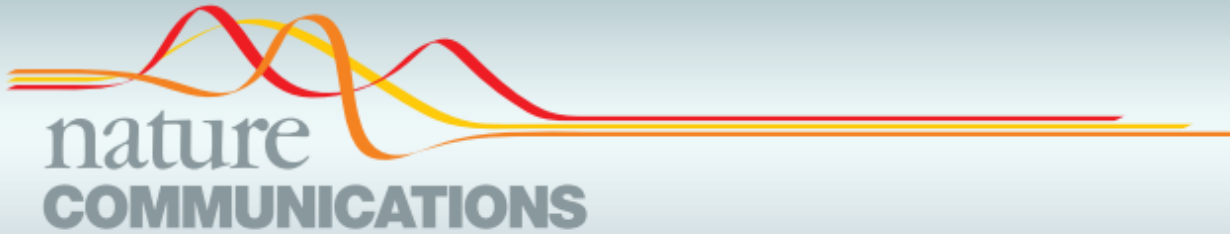
Whole-genome sequence variation, population structure and demographic history of the Dutch population

The Genome of the Netherlands Consortium\*



*The Genome of the Netherlands Consortium.  
Nature Genetics. 2014*

- Through a high-coverage sequencing (32.4 X on average), 21.2 million, including 12 million novel, single-nucleotide variants (SNVs) at an estimated false discovery rate of <1.0% were identified.
- the population-specific variant panels, which covers 99.0% SNVs of minor allele frequency ≥0.1%



## ARTICLE

Received 22 Nov 2014 | Accepted 7 Jul 2015 | Published 21 Aug 2015

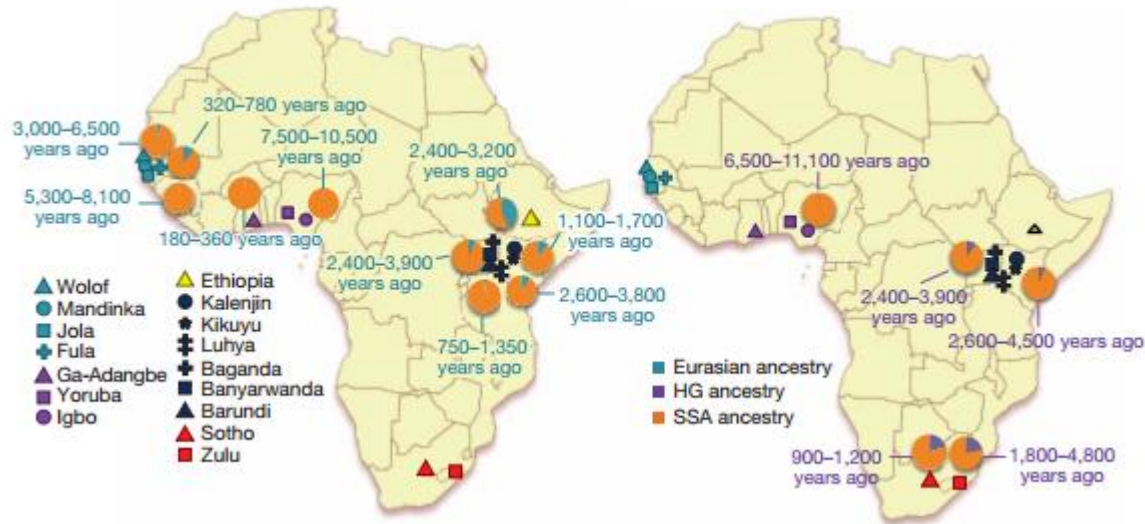
DOI: 10.1038/ncomms9018

**OPEN**

# Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals



# African Genome Variation Project



- The African Genome Variation Project represents dense genotypes from 1,481 individuals and whole-genome sequences from 320 individuals across sub-Saharan Africa.
- Using whole-genome sequencing, we demonstrate further improvements in imputation accuracy

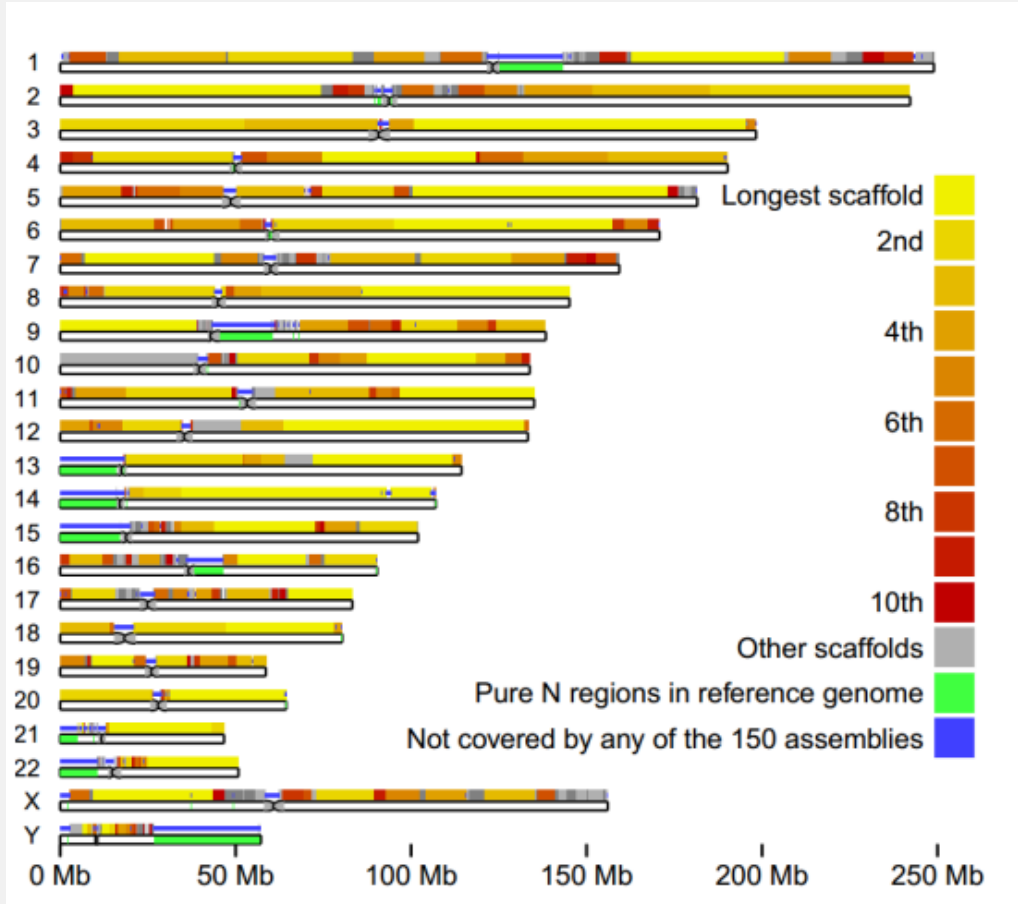
## ARTICLE

**OPEN**

doi:10.1038/nature13997

# The African Genome Variation Project shapes medical genetics in Africa

# Danish Genomes Project



## LETTER

OPEN

doi:10.1038/nature23264

### Sequencing and *de novo* assembly of 150 genomes from Denmark as a population reference

- ✓ *de novo* assemblies for 150 individuals (50 trios) .
- ✓ Provides a regional reference genome that we expect will improve the power of future association mapping studies and hence pave the way for precision medicine initiatives, which now are being launched in many countries.

## THE PRECISION MEDICINE INITIATIVE



Rewriting Life

### U.S. to Develop DNA Study of One Million People

An Obama initiative seeks to channel a torrent of gene information into treatments for cancer, other diseases.

### Significance

Large-scale initiatives toward personalized medicine are driving a massive expansion in the number of human genomes being sequenced. Therefore, there is an urgent need to define quality standards for clinical use. This includes deep coverage and sequencing accuracy of an individual's genome. Our work represents the largest effort to date in sequencing human genomes at deep coverage with these new standards. This study identifies over 150 million human variants, a majority of them rare and unknown. Moreover, these data identify sites in the genome that are highly intolerant to variation—possibly essential for life or health. We conclude that high-coverage genome sequencing provides accurate detail on human variation for discovery and clinical applications.

## Deep sequencing of 10,000 human genomes

Amalio Telenti<sup>a,b,1</sup>, Levi C. T. Pierce<sup>a,c,1</sup>, William H. Biggs<sup>a,1</sup>, Julia di Iulio<sup>a,b</sup>, Emily H. M. Wong<sup>a</sup>, Martin M. Fabani<sup>a</sup>, Ewen F. Kirkness<sup>a</sup>, Ahmed Moustafa<sup>a</sup>, Naisha Shah<sup>a</sup>, Chao Xie<sup>d</sup>, Suzanne C. Brewerton<sup>d</sup>, Nadeem Bulsara<sup>a</sup>, Chad Garner<sup>a</sup>, Gary Metzker<sup>a</sup>, Efren Sandoval<sup>a</sup>, Brad A. Perkins<sup>a</sup>, Franz J. Och<sup>a,c</sup>, Yaron Turpaz<sup>a,d</sup>, and J. Craig Venter<sup>a,b,2</sup>



- Han Chinese
- Tibetans
- Yao population
- Dai population
- **Mongolian populations**
- ...

## LETTER

doi:10.1038/nature13408

### Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA

Emilia Huerta-Sánchez<sup>1,2,3\*</sup>, Xin Jin<sup>1,4\*</sup>, Asan<sup>1,5,6\*</sup>, Zhuoma Bianba<sup>7\*</sup>, Benjamin M. Peter<sup>2</sup>, Nicolas Vinckenbosch<sup>2</sup>, Yu Liang<sup>1,5,6</sup>, Xin Yi<sup>1,5,6</sup>, Mingze He<sup>1,8</sup>, Mehmet Somel<sup>9</sup>, Peixiang Ni<sup>1</sup>, Bo Wang<sup>1</sup>, Xiaohua Ou<sup>1</sup>, Huasang<sup>1</sup>, Jiangbai Luosang<sup>1</sup>, Zha Xi Ping Cuo<sup>10</sup>, Kui Li<sup>11</sup>, Guoyi Gao<sup>12</sup>, Ye Yin<sup>1</sup>, Wei Wang<sup>1</sup>, Xiuqing Zhang<sup>1,13,14</sup>, Xun Xu<sup>1</sup>, Huanming Yang<sup>1,15,16</sup>, Yingrui Li<sup>1</sup>, Jian Wang<sup>1,16</sup>, Jun Wang<sup>1,15,17,18,19</sup> & Rasmus Nielsen<sup>1,2,20,21</sup>



**Sequencing of  
Adaptation to h**  
Xin Yi, et al.  
*Science* **329**, 75  
DOI: 10.1126/scienc

Low pass genomes of 141,431 Chinese reveal patterns of viral infection, novel phenotypic associations, and the genetic history of China



(GIGA)<sup>n</sup>  
SCIENCE

GigaScience, 6, 2017, 1–7

doi: 10.1093/gigascience/glx067

Advance Access Publication Date: 31 July 2017

Data Note

# Mongolian populations

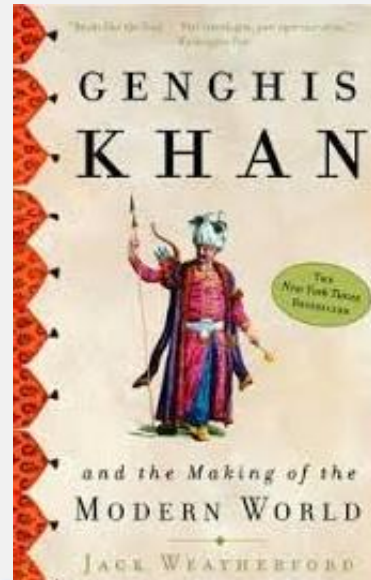
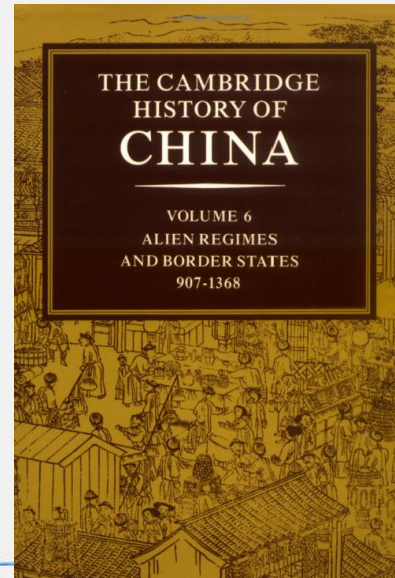
# Background and motivations

- Mongolians is a big and widely distributed ethnic group.
- Over 10 million Mongolians currently reside in Mongolia, Northern China, and other neighboring countries.

Total population		
(10 million (2010))		
Regions with significant populations		
	China (Inner Mongolia)	5,981,840 (2010) <sup>[1]</sup>
	Mongolia	2,921,287 <sup>[2]</sup>
	Russia	647,417 <sup>[3]</sup>
	South Korea	34,000 <sup>[4]</sup>
	United States	15,000 - 18,000 <sup>[5]</sup>
	Kyrgyzstan	12,000 <sup>[6]</sup>
	Czech Republic	7,515 <sup>[7]</sup>
	Japan	5,401 <sup>[8]</sup>
	Canada	5,350 <sup>[9]</sup>
	Germany	3,852 <sup>[8]</sup>
	United Kingdom	3,701 <sup>[8]</sup>
	France	2,859 <sup>[8]</sup>
	Turkey	2,645 <sup>[8]</sup>
	Kazakhstan	2,523 <sup>[8]</sup>
	Austria	1,955 <sup>[10]</sup>
	Malaysia	1,500 <sup>[8]</sup>

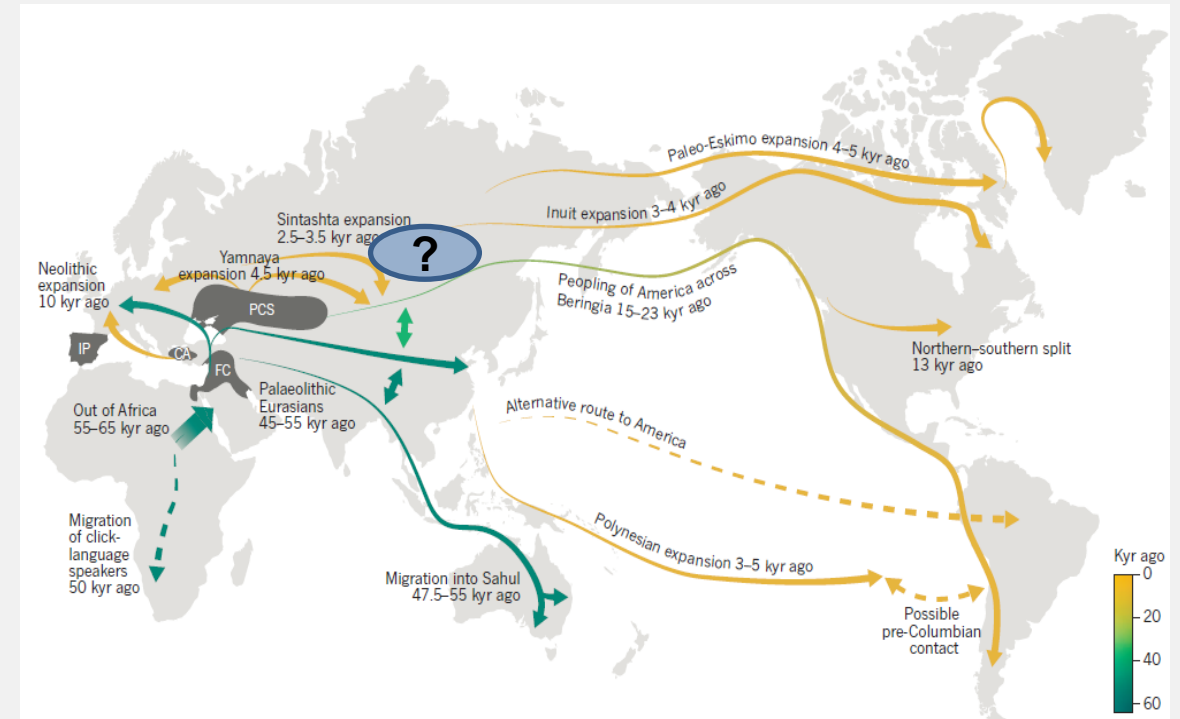


- Great impacts from Genghis Khan and his Empire to the world, especially Eurasians.
  - Society
  - Culture
  - Language
  - Religions
  - ...



# Background and motivations

- Mongolians is an important and interesting ethnic group
- Intricate population history and structure
  - ✓ Broadly defined ethnic group, including Khalkha Mongols, Buryats, Oirats, Abaga Mongols, Horchin Mongols, and so on.
- Important roles in human evolution
  - ✓ Early human evolution (Peopling of Native Americans)
  - ✓ Modern human evolution (Eurasians)

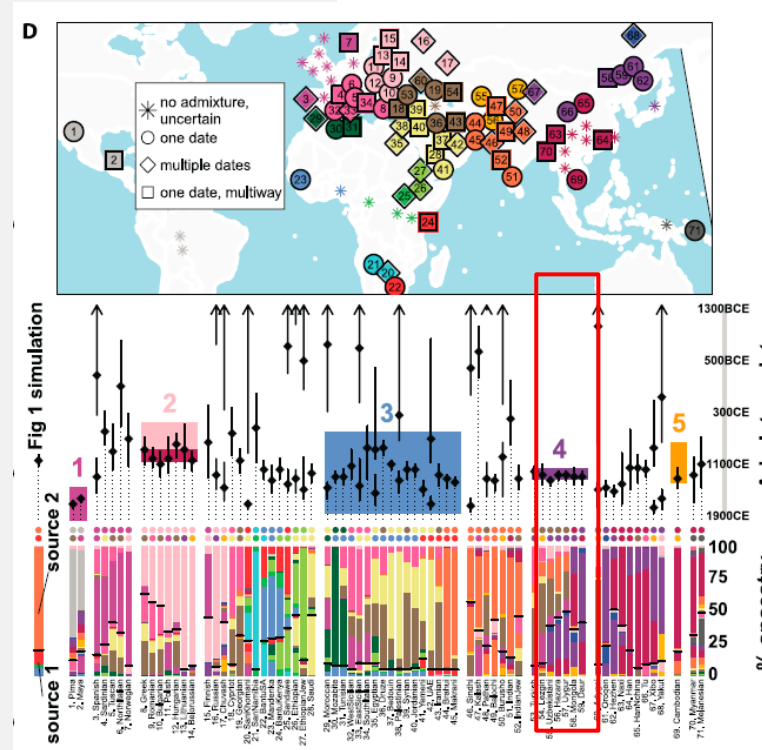


R. Nielsen, et al. (2017), *Nature*, 541:302-310.



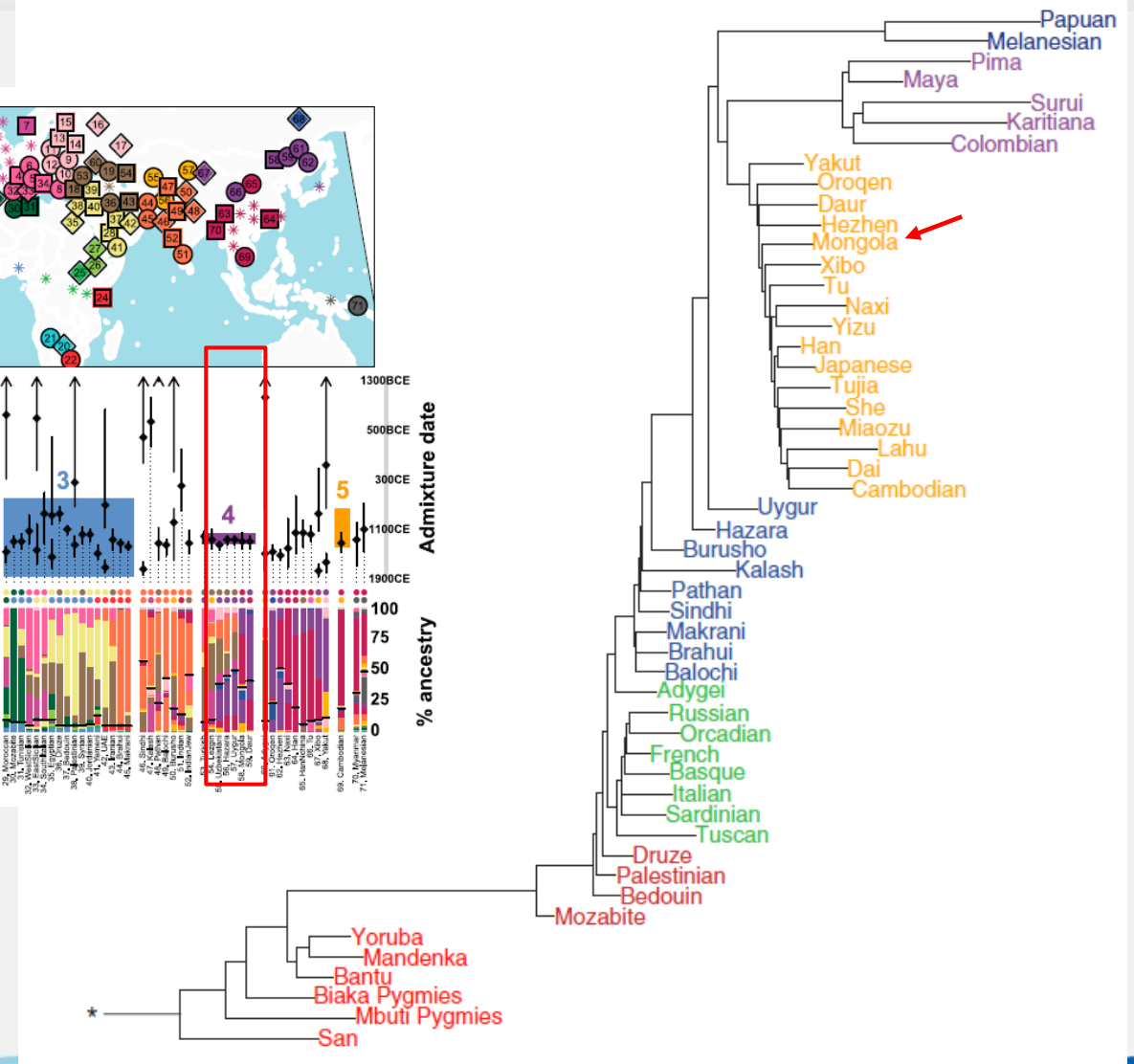
# Background and motivations

- Mongolians are under explored and learned limitedly
- Studies focus on local genomic regions
  - Y chromosome or Mitochondrial genome
  - Limited to single or several genes/mutations
- Only as supporting part



G. Hellenthal, et al. (2014). *Science*, 343: 747-751.

J. Z. Li, et al. (2008). *Science*, 319: 1100-1104.





# Project Design

Samples 175 Mongolian DNA samples, including 6 tribes.

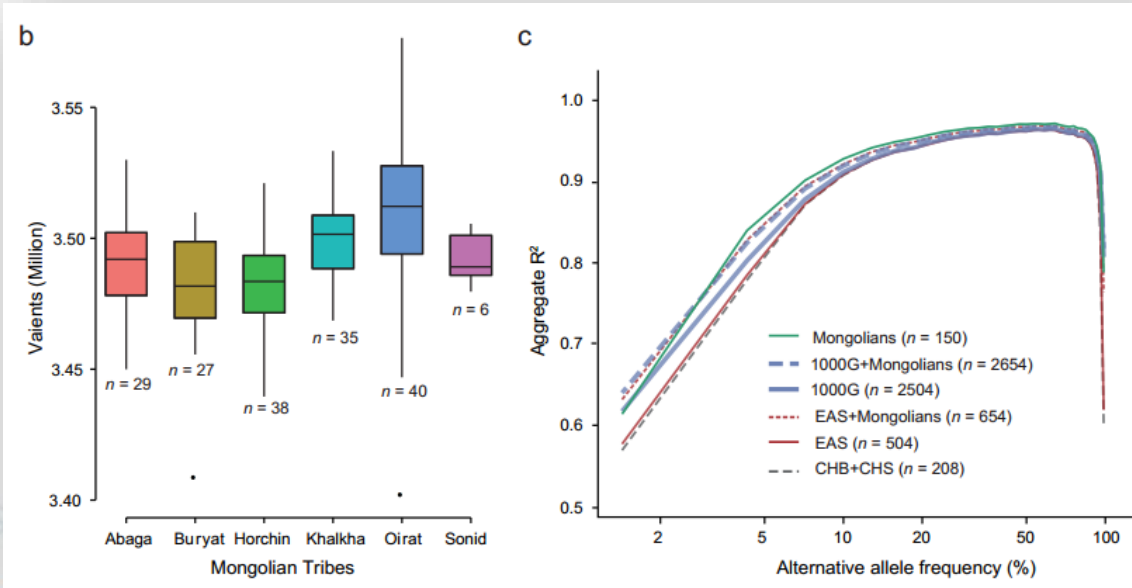
Research Strategies Whole genome sequencing based population analysis (21X);

Scientific questions

- How is the evolutionary relationship between Mongolian population and other human populations, especially for Asia populations?
- Is there an extremely high genetic diversity when compared with other human population?
- Direction and extent of genes flows between Mongolia and Europe Population?
- How did the rise of Genghis Khan influence the genetic structure of Mongolia population and Eurasians?
- Are all the Mongolian tribes shared the same population history?
- How does the Mongolian population give suggestions about the migration of East Asian?

# 1. Reference panel with the best imputation accuracy

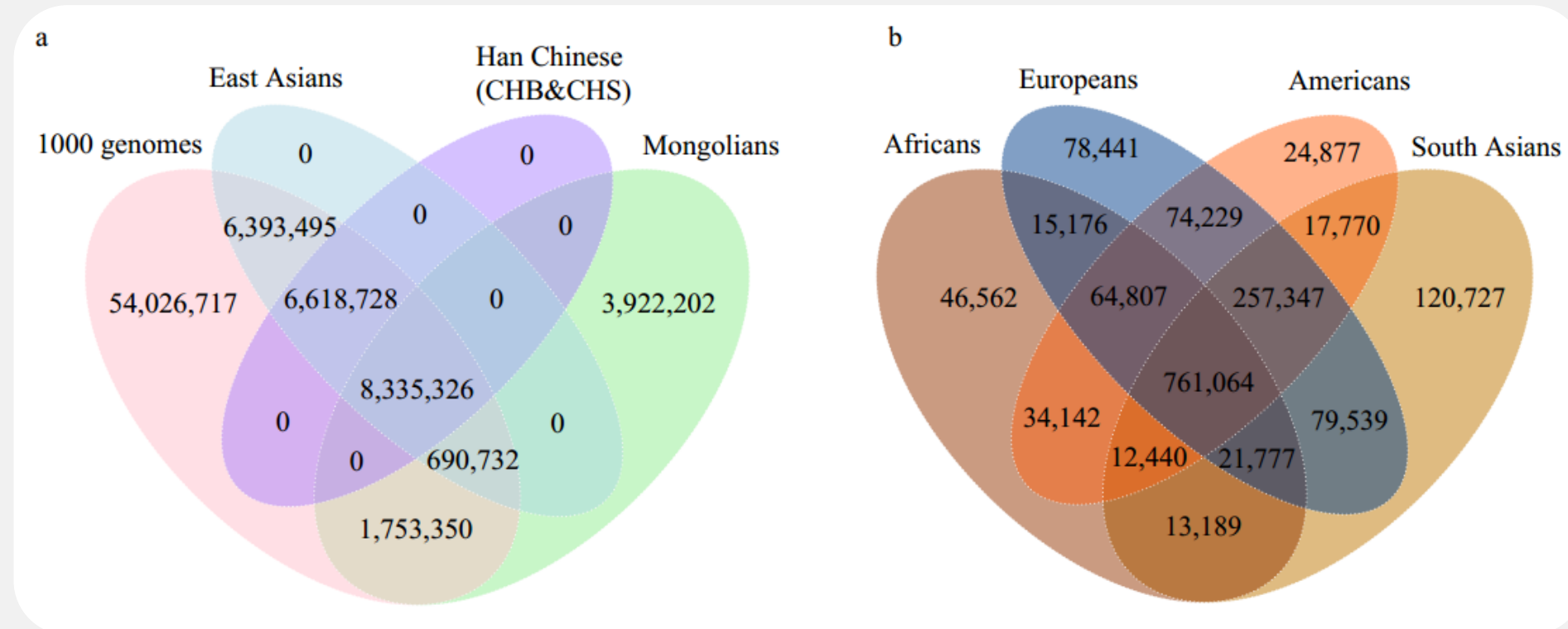
- We masked the genotypes of 25 randomly selected Mongolians to create a “missing” sample, then masked the SNP positions that not existed on the 2.5M Omni chip.
- We imputed the complete “missing” genotypes using different reference haplotype panels (impute2).



- Using a panel of Mongolians (remaining 150) or Mongolians+1000G as the reference was superior to other combinations tested (Fig. c).

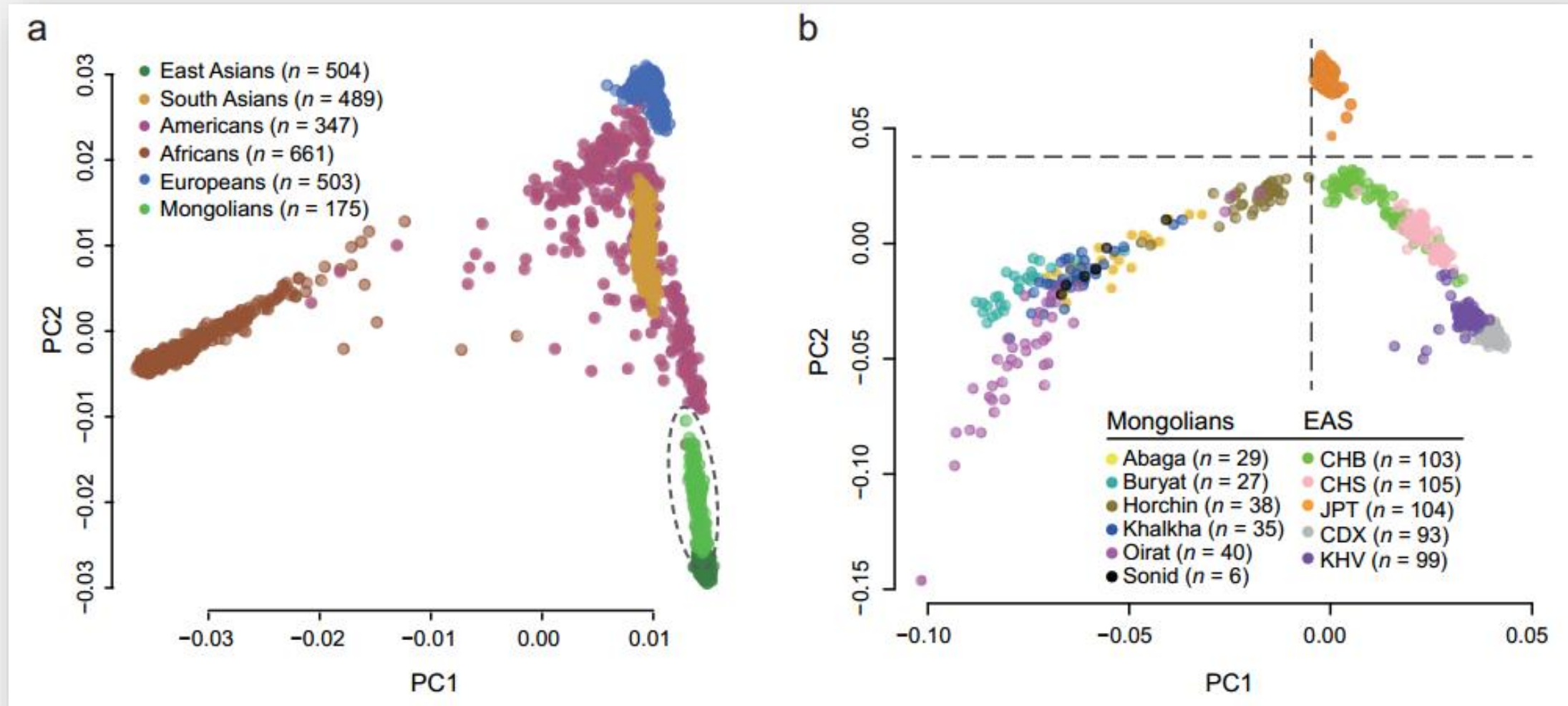
## 2. Europeans share a large proportion of variants with Mongolians

- We identified 15.2 million SNPs. Compared to the variants in 1000G, 3.9 million SNPs are novel; another 11.5% are not reported in any East Asian populations (EAS) but are represented in one or more non-EAS populations (Fig. a). Of the non-EAS groups, Europeans share the greatest proportion of variants with Mongolians (Fig. b).





### 3. Possibly high genetic diversity and population differentiation



### 3. Possibly high genetic diversity and population differentiation

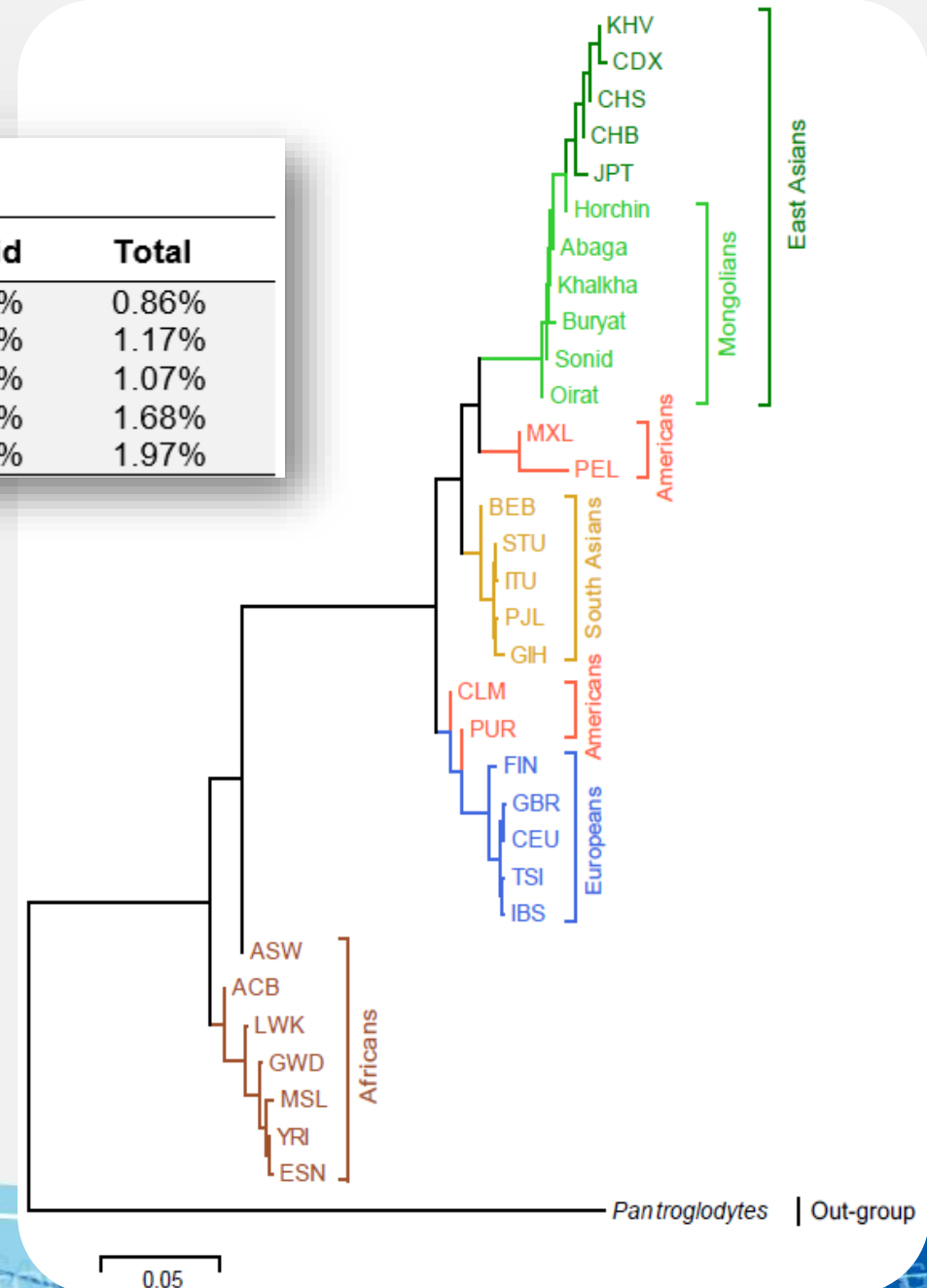
Table S3 Differentiation between Mongolians and global populations ( $F_{ST}$ )

		Abaga	Buryat	Horchin	Khalkha	Oirat	Sonid	Total
East Asians	CHB	0.98%	1.82%	0.25%	1.11%	1.22%	1.15%	0.86%
	CHS	1.29%	2.19%	0.49%	1.45%	1.56%	1.47%	1.17%
	JPT	1.21%	1.88%	0.66%	1.28%	1.41%	1.33%	1.07%
	KHV	1.80%	2.68%	1.08%	1.94%	2.00%	1.93%	1.68%
	CDX	2.12%	2.97%	1.35%	2.26%	2.32%	2.28%	1.97%

In general,  $F_{ST}$  showd that the differentiation within Mongolian is less than that with EAS. However, some cases showed deeper differentiations when compared with EAS, such as Horchin.

Table S4 Differentiation among Mongolian tribes ( $F_{ST}$ ).

	Abaga	Buryat	Horchin	Khalkha	Oirat	Sonid
Abaga	--	0.41%	0.35%	0.09%	0.14%	0.13%
Buryat	0.41%	--	0.92%	0.26%	0.34%	0.39%
Horchin	0.35%	0.92%	--	0.38%	0.49%	0.41%
Khalkha	0.09%	0.26%	0.38%	--	0.07%	0.01%
Oirat	0.14%	0.34%	0.49%	0.07%	--	0.05%
Sonid	0.13%	0.39%	0.41%	0.01%	0.05%	--



## 4. Population Demography

- We randomly selected four individuals using multiple sequentially Markovian coalescent method (MSMC) to estimate population demography.
- It appears the Mongolian tribes started to differentiate ~10 kya.
- Among Mongolian tribes, Horchin, Buryat, and the rest each had a different growth pattern since 10 kya.
- Pairwise estimation of separation time indicates Mongolians and other East Asians separated 4.5-13 kya, whereas the tribes of Mongolians started to differentiate much late.

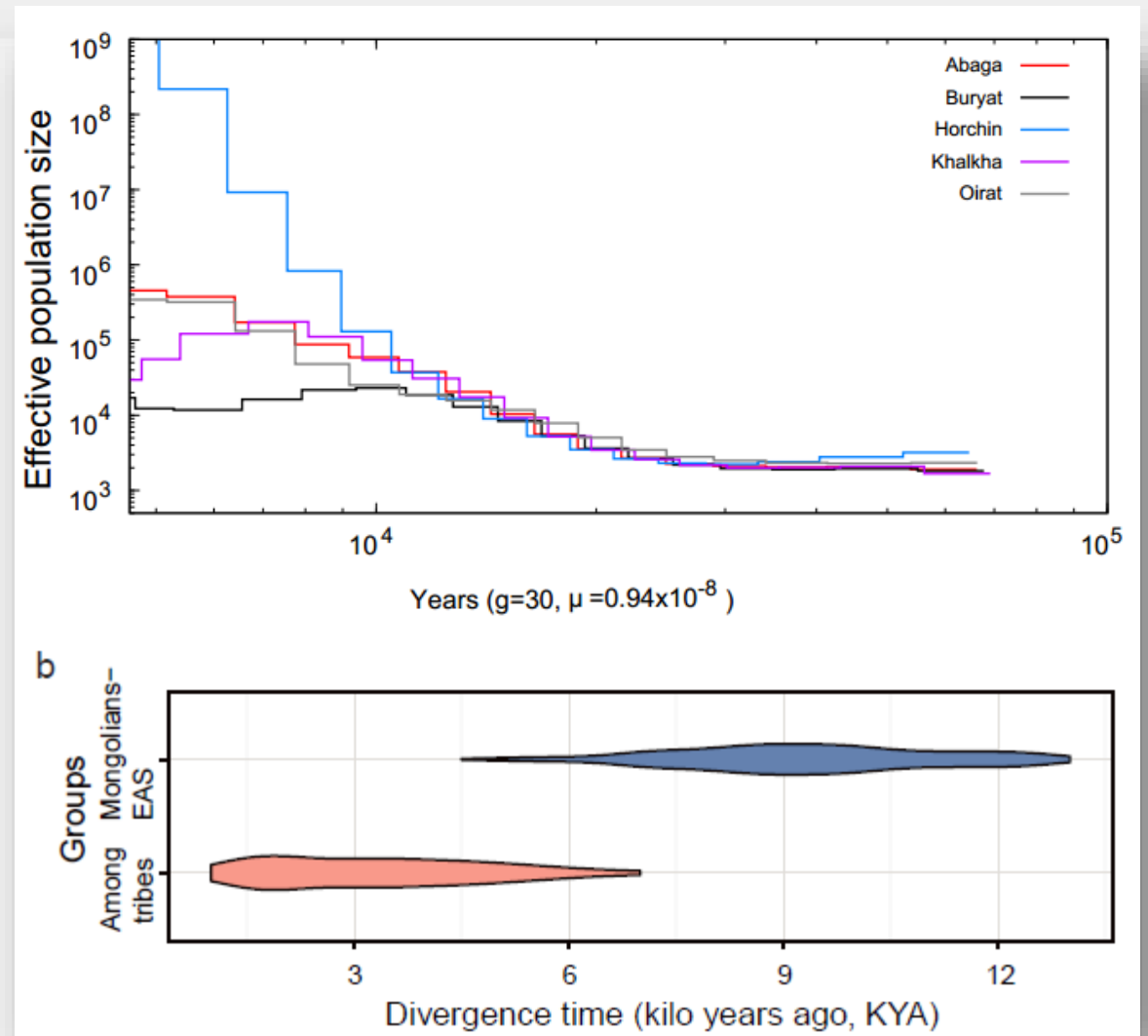
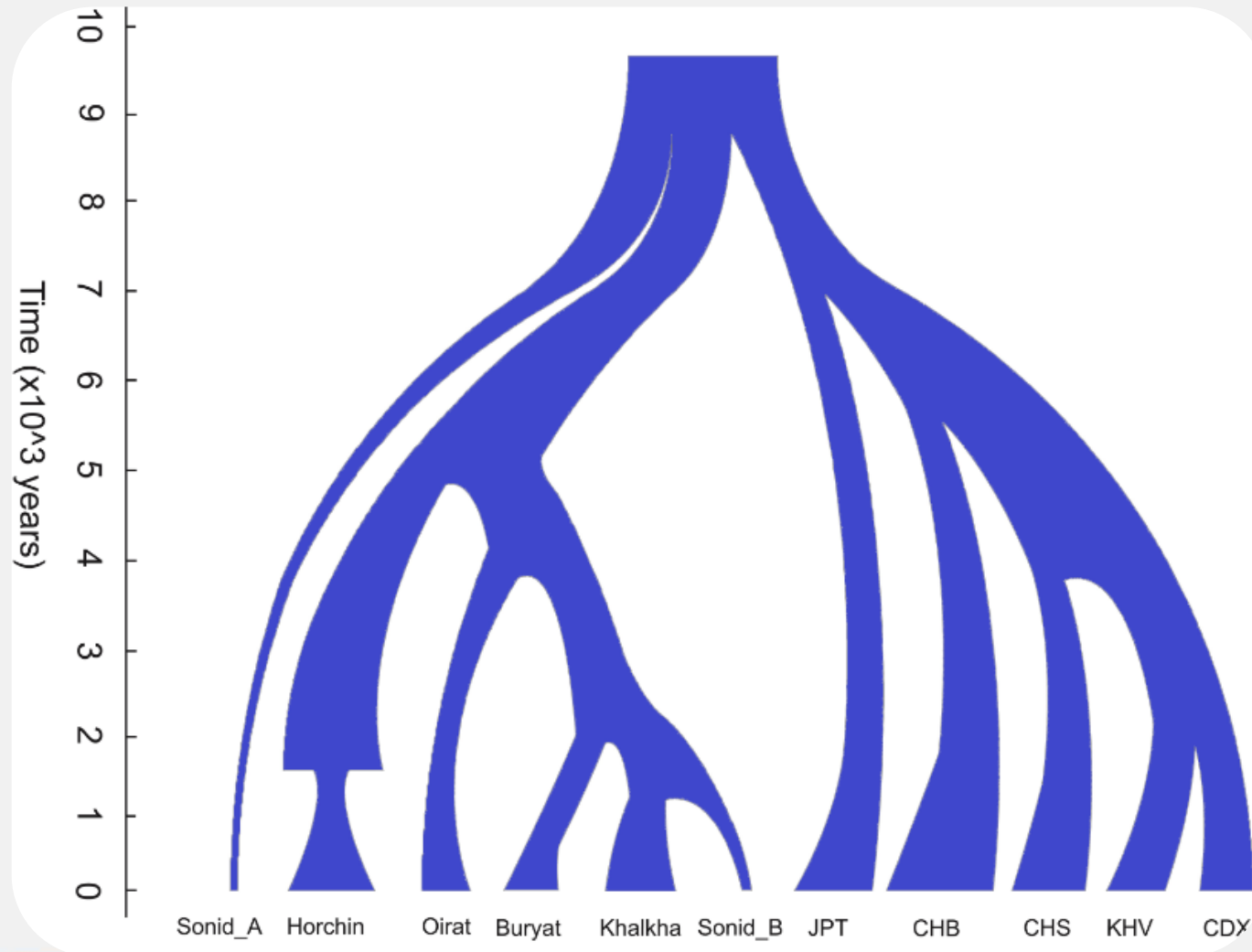


Figure. Mongolian population demography.

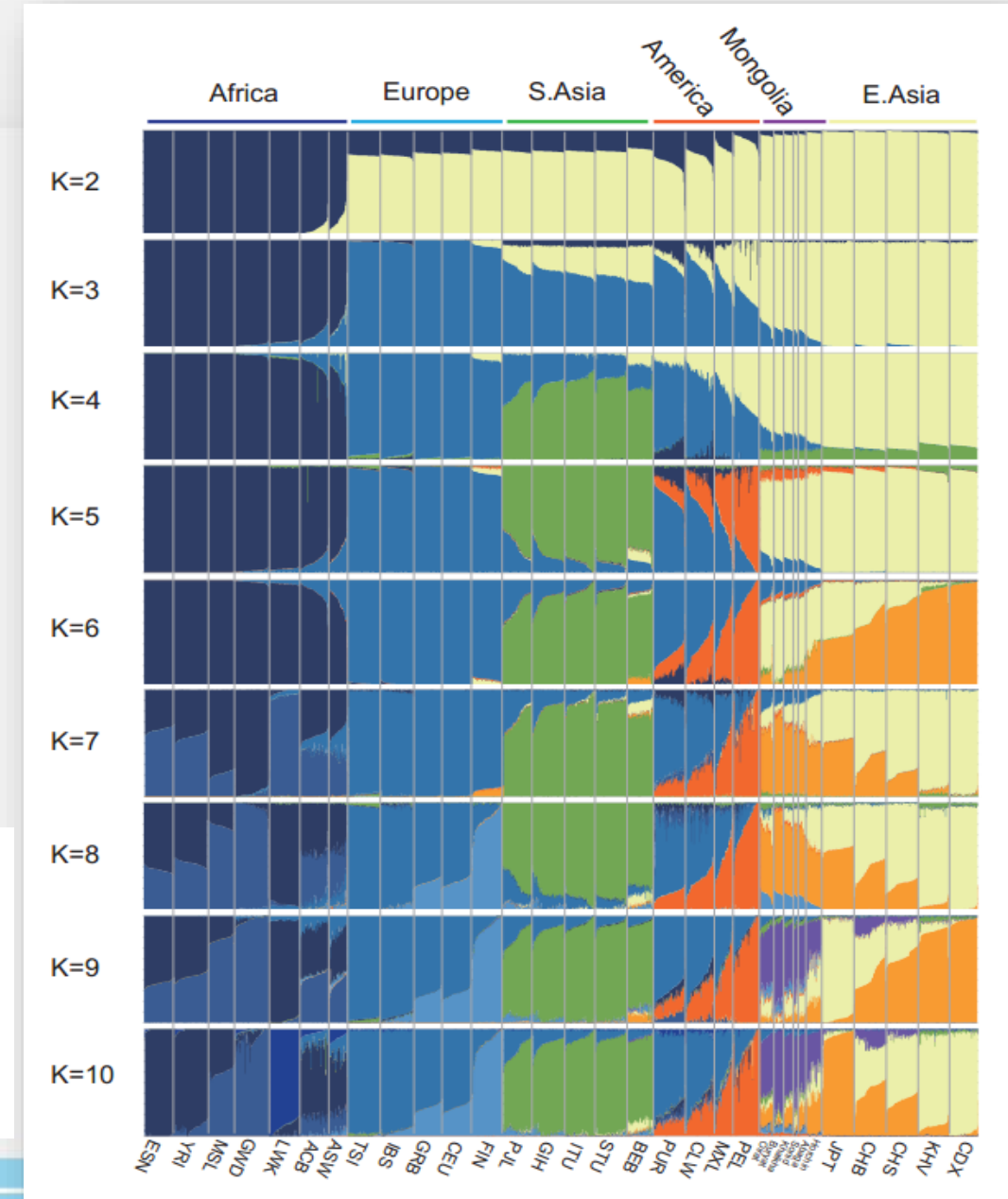
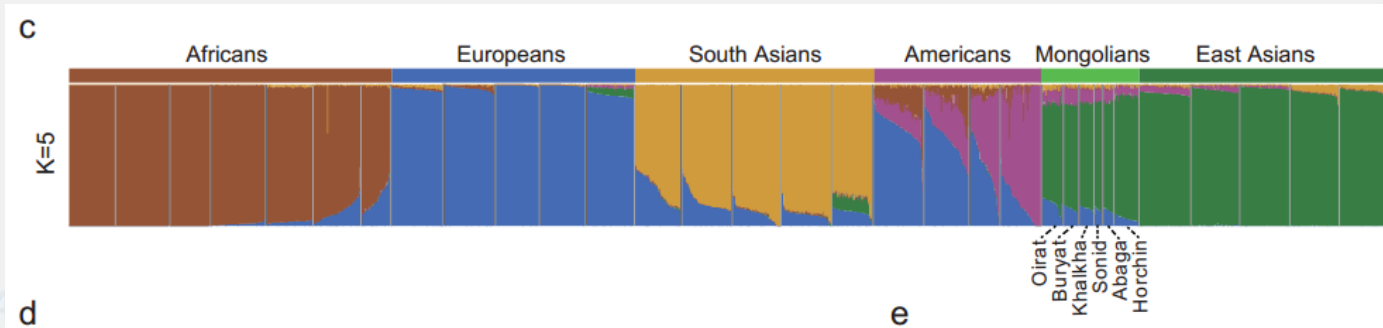


## 4. Population Demography



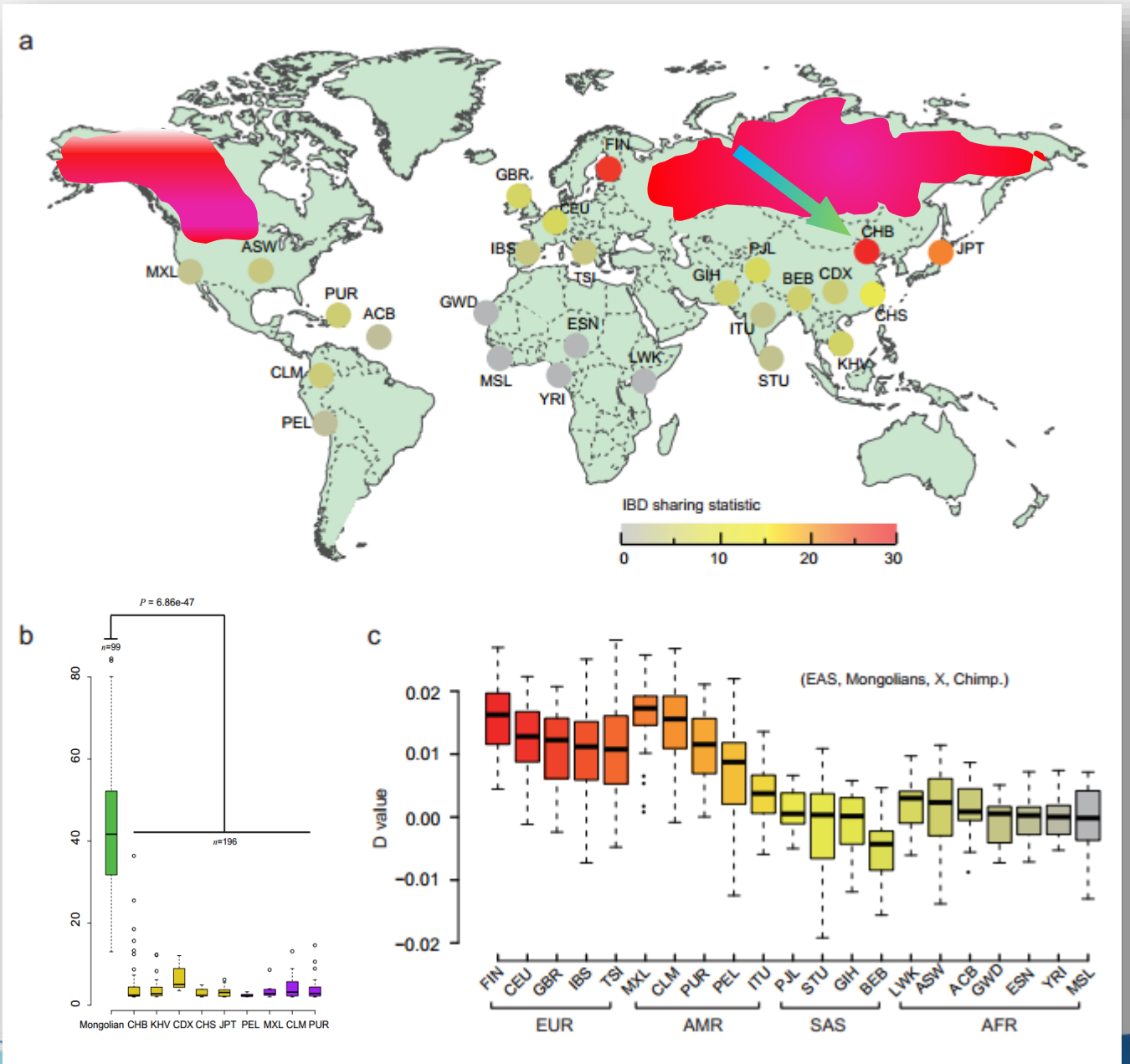
## 5. Genetic Structure

- This figure shows the results of  $K=2$  through 10. At  $K=2$  through 10, the European components can always be found in Mongolia population, but cannot obviously found in other East Asia populations. At  $K=9$ , the new component accounts for a major portion of ancestry for individuals from Mongolia, separating this region from the East Asia.



## 6. Gene Flow

- Different degrees of recent gene flow with global populations, which were characterized by IBDs.
- Significantly more shared IBDs with Finns than that with majority groups, which is also supported by the analyses of *D* test.
- The gene flows widely exist between Finns and Siberians, including Mongolians.





## 6. Gene Flow

- We further explore the direction and time for the recent gene flows between Mongolians and other human populations.
- Mongolians received more recent gene flows (5.6%) from Finns from  $20.2 \pm 1.0$  to  $20.5 \pm 1.0$  generations (**~600 years ago**).
- Finns received gene flows from Mongolians (1.4%) from  $36.0 \pm 7.6$  to  $39.7 \pm 7.4$  generations ago (**~1,200 years ago**).

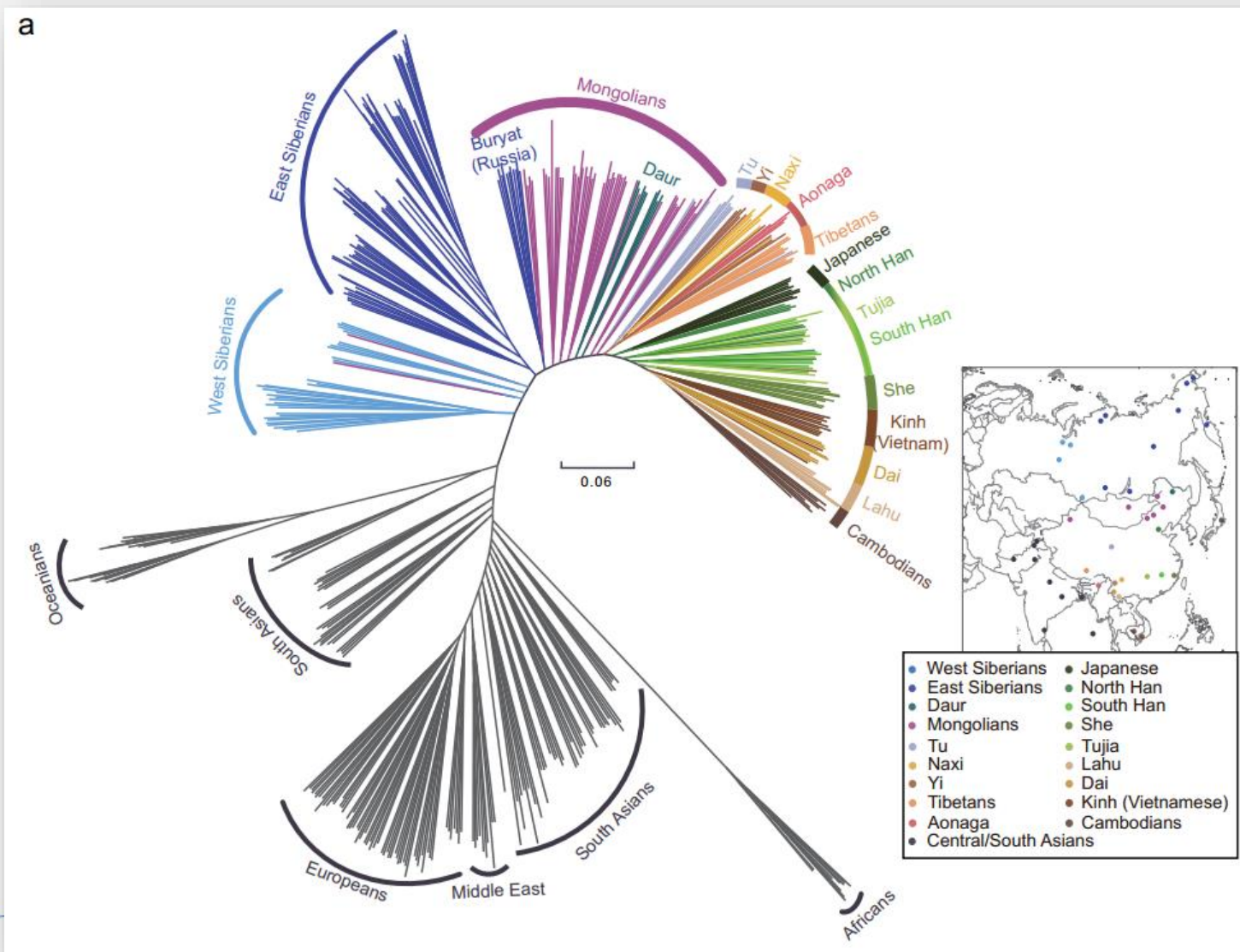
**Table S7**

Test population	Reference population A	Reference population B	Date of admixture (Generation $\pm$ sd)	Date of admixture (Year $\pm$ sd)*	Z score of admixture
FIN	Mongolian	TSI	36.02(7.57)	1080.60(227.10)	3.57
FIN	Mongolian	IBS	37.07(7.18)	1112.10(215.40)	3.76
FIN	Mongolian	CEU	36.70(6.44)	1101.00(193.20)	4.02
FIN	Mongolian	GBR	39.74(7.38)	1192.20(221.40)	3.65
CHB	Mongolian	CHS	21.61(3.49)	648.30(104.70)	6.12
CHB	Mongolian	CDX	23.65(2.91)	709.50(87.30)	7.55
CHB	Mongolian	KHV	20.81(2.35)	624.30(70.50)	8.87
Mongolian	FIN	CHB	20.38(1.08)	611.40(32.40)	18.88
Mongolian	FIN	CHS	20.19(1.04)	605.70(31.20)	19.34
Mongolian	FIN	CDX	20.28(1.08)	608.40(32.40)	18.8
Mongolian	FIN	KHV	20.54(0.96)	616.20(28.80)	21.45

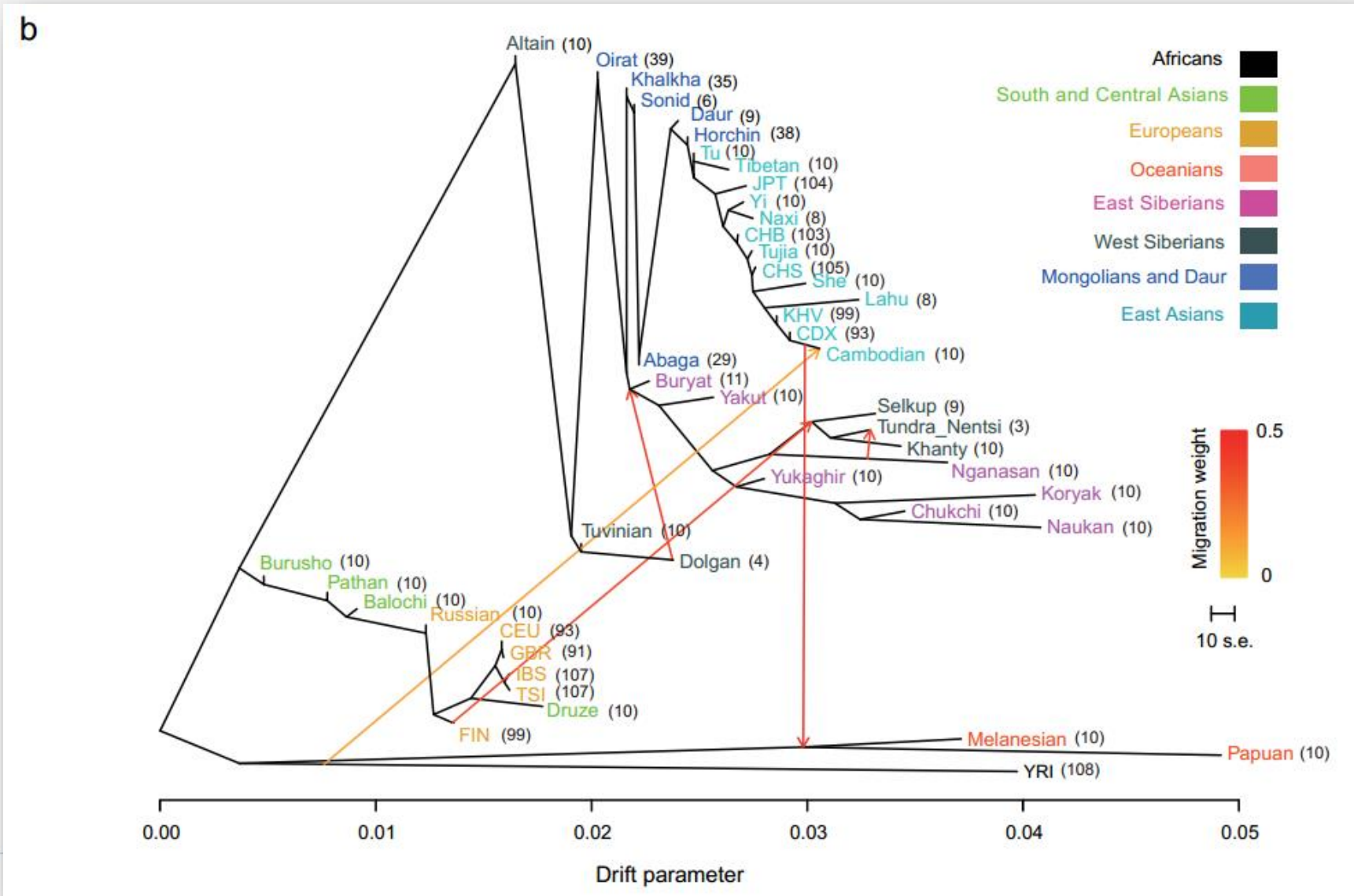
\*30 years per generation.

- This is well in line with the period of Yuan dynasty (1271 – 1368).
- Gene flows continuously occurred between Mongolians and widespread human populations in last over 10,000 years

## 7. Possible North to South Migration



## 7. Possible North to South Migration





# Conclusions

1. We built a high quality reference panel for Mongolian population and even for Northern Asians.
2. Genetic differentiations and diversity within Mongolians were larger than expected, which may result from the different population demography.
3. Gene flows continuously occurred between Mongolians and North Asian human populations over a very long period.
4. We highlight that North, East, and Southeast Asian populations are more aligned with each other than these groups are with South Asian and Oceanian populations.

# Acknowledgement



白海花  
郭小森  
那日苏  
吴奇柱  
周欢敏  
张勇  
罗忻  
张丹丹  
党宁馨

.....



Thanks!

