

# Reconstructing the evolutionary history of species from genomes: learning about the past to predict the future

Vitor Sousa

**cE3c** – center for ecology, evolution and environmental changes  
Investigador cE3c, Professor Auxiliar Convidado  
Faculdade de Ciências da Universidade de Lisboa



**Encontro Ciência**

10 July 2019

[vmsousa@fc.ul.pt](mailto:vmsousa@fc.ul.pt)



# Reconstructing the evolutionary history of **humans** from genomes: learning about the past **to predict the future**

Vitor Sousa

**cE3c** – center for ecology, evolution and environmental changes  
Investigador cE3c, Professor Auxiliar Convidado  
Faculdade de Ciências da Universidade de Lisboa



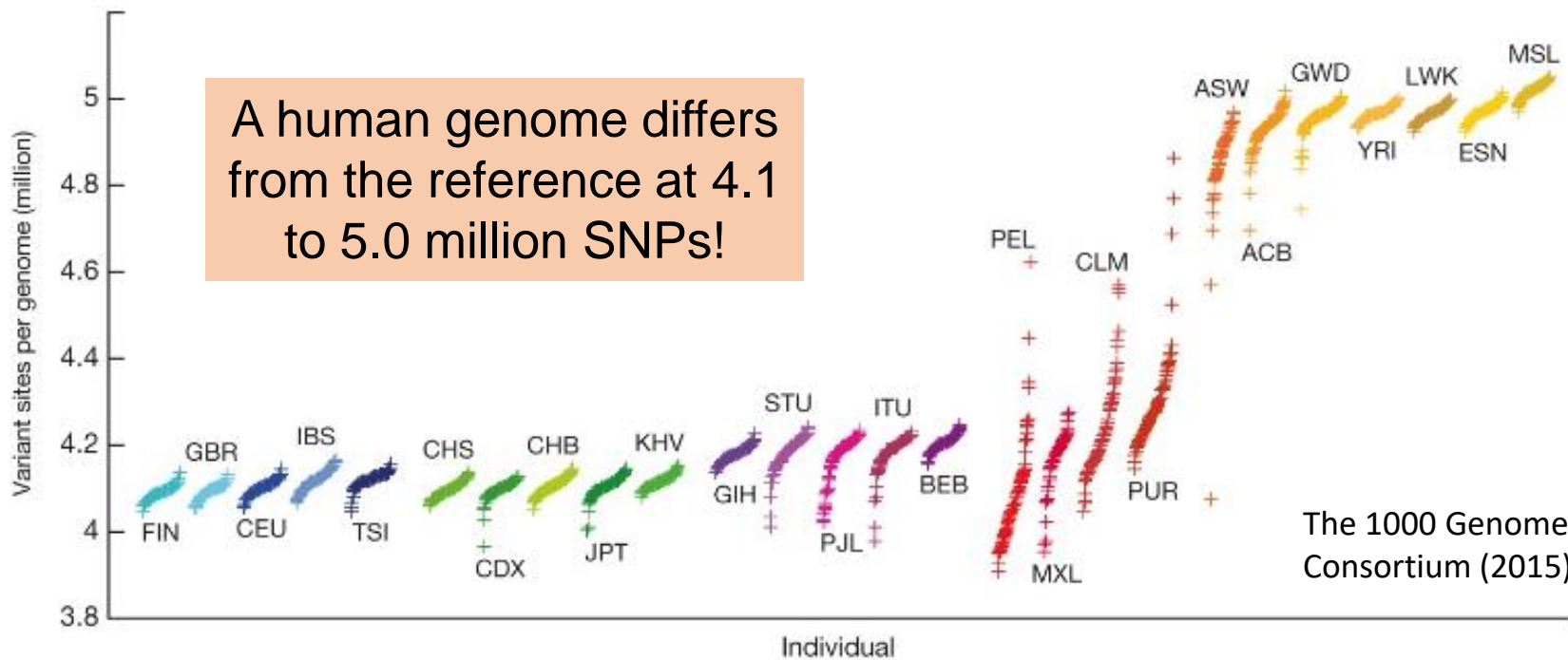
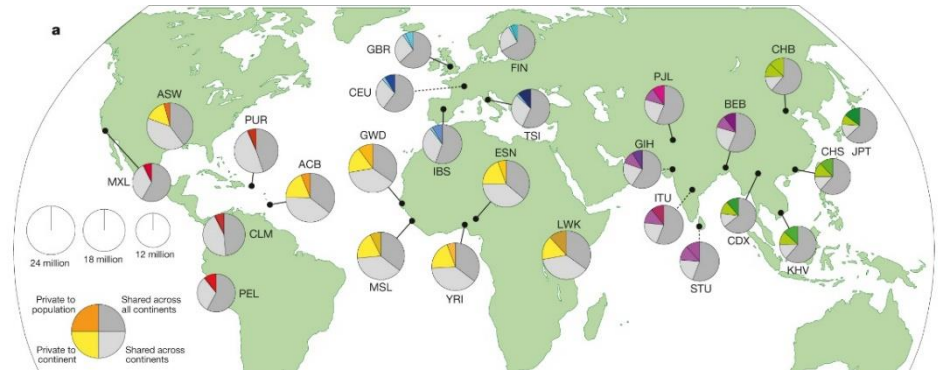
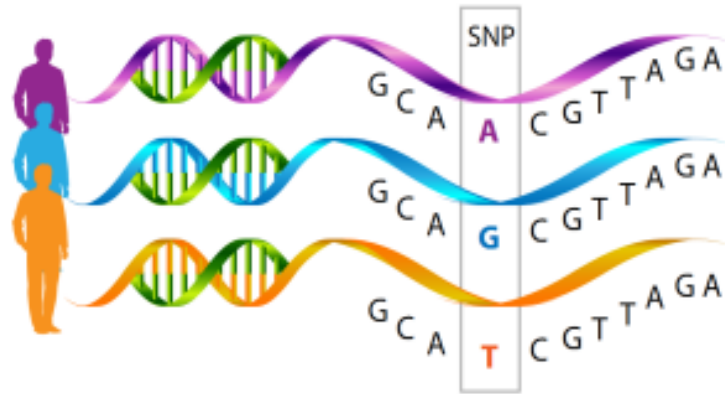
**Ciências**  
**ULisboa**

**Encontro Ciência**  
**Saúde de Qualidade**  
10 July 2019

[vmsousa@fc.ul.pt](mailto:vmsousa@fc.ul.pt)



# There is lots of variation at the genome level



The 1000 Genomes Project Consortium (2015) Nature

# Evolutionary history of populations is complex

## Genomic processes

- Mutation
- Recombination

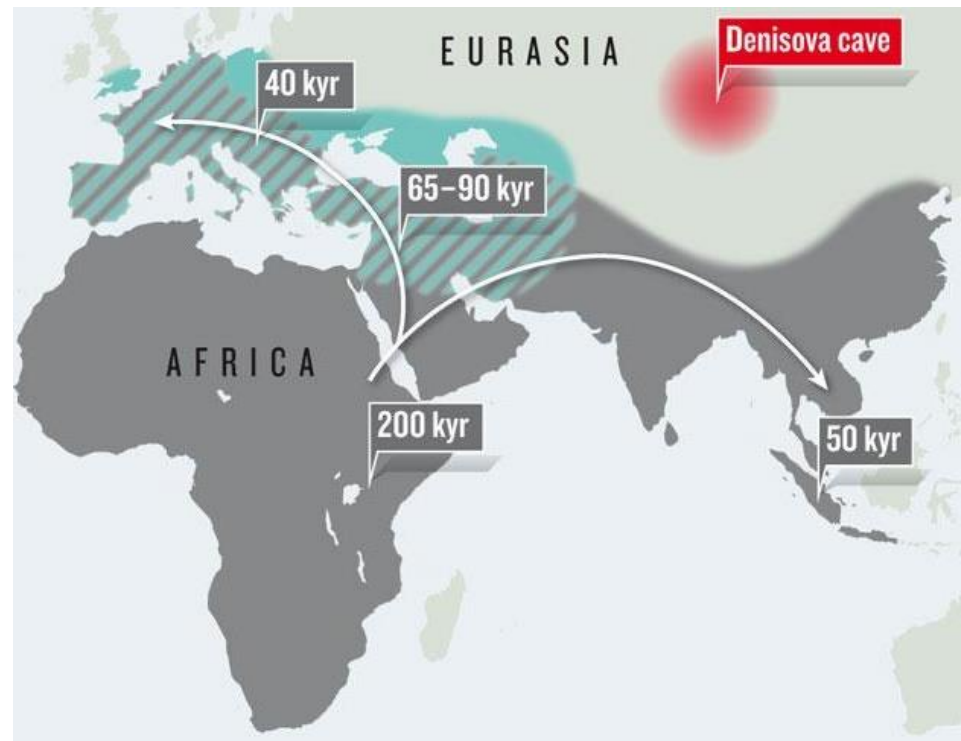
## Demography

- Effective population sizes
- Population split times
- Migration rates

## Selection

Natural selection:

- Beneficial mutations involved in adaptation
- Deleterious mutations with negative effect



Callaway (2009) *Nature*

# Bioinformatics and Population genomics

- Example I: New method to infer demographic history
- Example II: Using simulations to obtain predictions

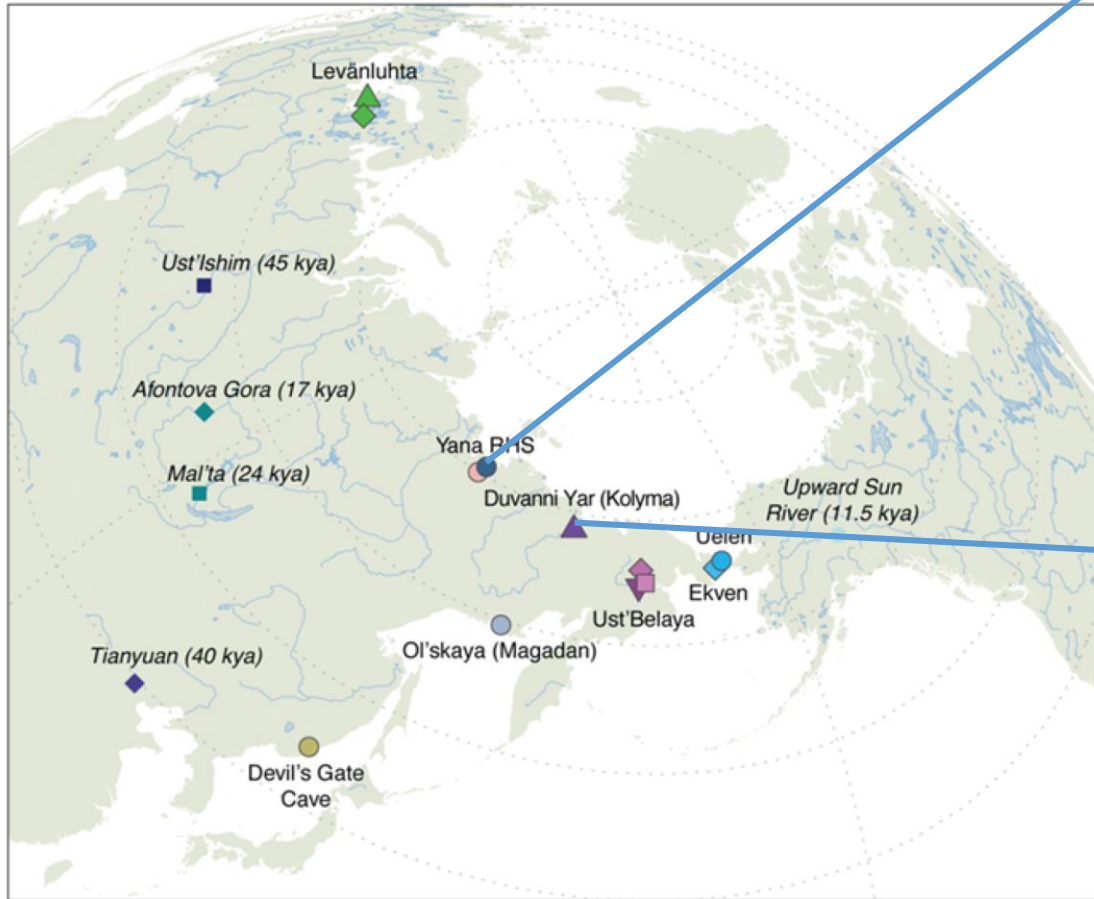
# The population history of northeastern Siberia since the Pleistocene

Martin Sikora<sup>1,43\*</sup>, Vladimir V. Pitulko<sup>2,43\*</sup>, Vitor C. Sousa<sup>3,4,5,43</sup>, Morten E. Allentoft<sup>1,43</sup>, Lasse Vinner<sup>1</sup>, Simon Rasmussen<sup>6,41</sup>, Ashot Margaryan<sup>1</sup>, Peter de Barros Damgaard<sup>1</sup>, Constanza de la Fuente<sup>1,42</sup>, Gabriel Renaud<sup>1</sup>, Melinda A. Yang<sup>7</sup>, Qiaomei Fu<sup>7</sup>, Isabelle Dupanloup<sup>8</sup>, Konstantinos Giampoudakis<sup>9</sup>, David Nogués-Bravo<sup>9</sup>, Carsten Rahbek<sup>9</sup>, Guus Kroonen<sup>10,11</sup>, Michaël Peyrot<sup>11</sup>, Hugh McColl<sup>1</sup>, Sergey V. Vasilyev<sup>12</sup>, Elizaveta Veselovskaya<sup>12,13</sup>, Margarita Gerasimova<sup>12</sup>, Elena Y. Pavlova<sup>2,14</sup>, Vyacheslav G. Chasnyk<sup>15</sup>, Pavel A. Nikolskiy<sup>2,16</sup>, Andrei V. Gromov<sup>17</sup>, Valeriy I. Khartanovich<sup>17</sup>, Vyacheslav Moiseyev<sup>17</sup>, Pavel S. Grebenyuk<sup>18,19</sup>, Alexander Yu. Fedorchenko<sup>20</sup>, Alexander I. Lebedintsev<sup>18</sup>, Sergey B. Slobodin<sup>18</sup>, Boris A. Malyarchuk<sup>21</sup>, Rui Martiniano<sup>22</sup>, Morten Meldgaard<sup>1,23</sup>, Laura Arppe<sup>24</sup>, Jukka U. Palo<sup>25,26</sup>, Tarja Sundell<sup>27,28</sup>, Kristiina Mannermaa<sup>27</sup>, Mikko Putkonen<sup>25</sup>, Verner Alexandersen<sup>29</sup>, Charlotte Primeau<sup>29</sup>, Nurbol Baimukhanov<sup>30</sup>, Ripan S. Malhi<sup>31,32</sup>, Karl-Göran Sjögren<sup>33</sup>, Kristian Kristiansen<sup>33</sup>, Anna Wessman<sup>27,34</sup>, Antti Sajantila<sup>25</sup>, Marta Mirazon Lahr<sup>1,35</sup>, Richard Durbin<sup>22,36</sup>, Rasmus Nielsen<sup>1,37</sup>, David J. Meltzer<sup>1,38</sup>, Laurent Excoffier<sup>4,5\*</sup> & Eske Willerslev<sup>1,36,39,40\*</sup>

***Nature*** (2019)



# Siberia and colonization of the Americas



Yana RHS (31,600 years ago)  
Whole-genome depth of coverage 25x



Kolyma (9,800 years ago)  
Whole-genome depth of coverage 14x



# Hypothesis: Continuity vs Replacement of populations

**Data:** Ancient and present-day samples; 625 blocks of 1Mb (~1.5 Million SNP)

**Method:** Composite likelihood - *fastsimcoal2*  
(Excoffier et al, 2013 Plos Genetics)

Europe (Sardinia)	Ancient North Siberians (Yana)	Ancient Paleo- siberian (Kolyma)	Neo- siberian (Even)	East Asia (Han)
----------------------	---	---	----------------------------	-----------------------

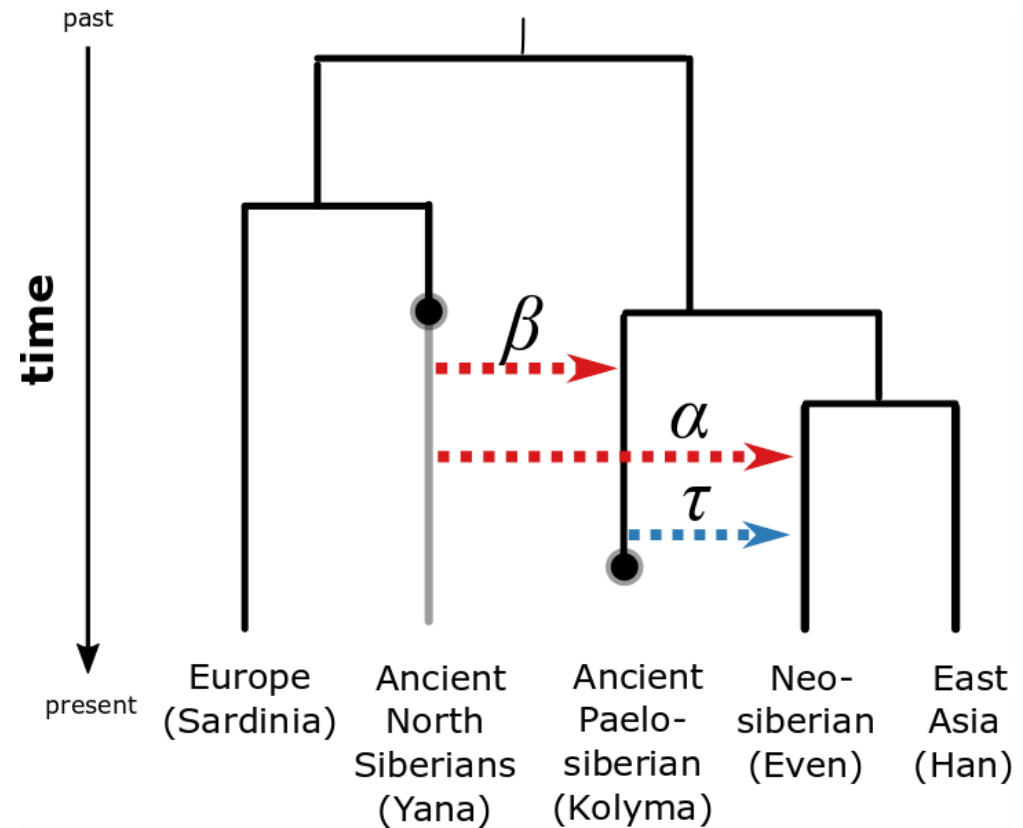


# Hypothesis: Continuity vs Replacement of populations

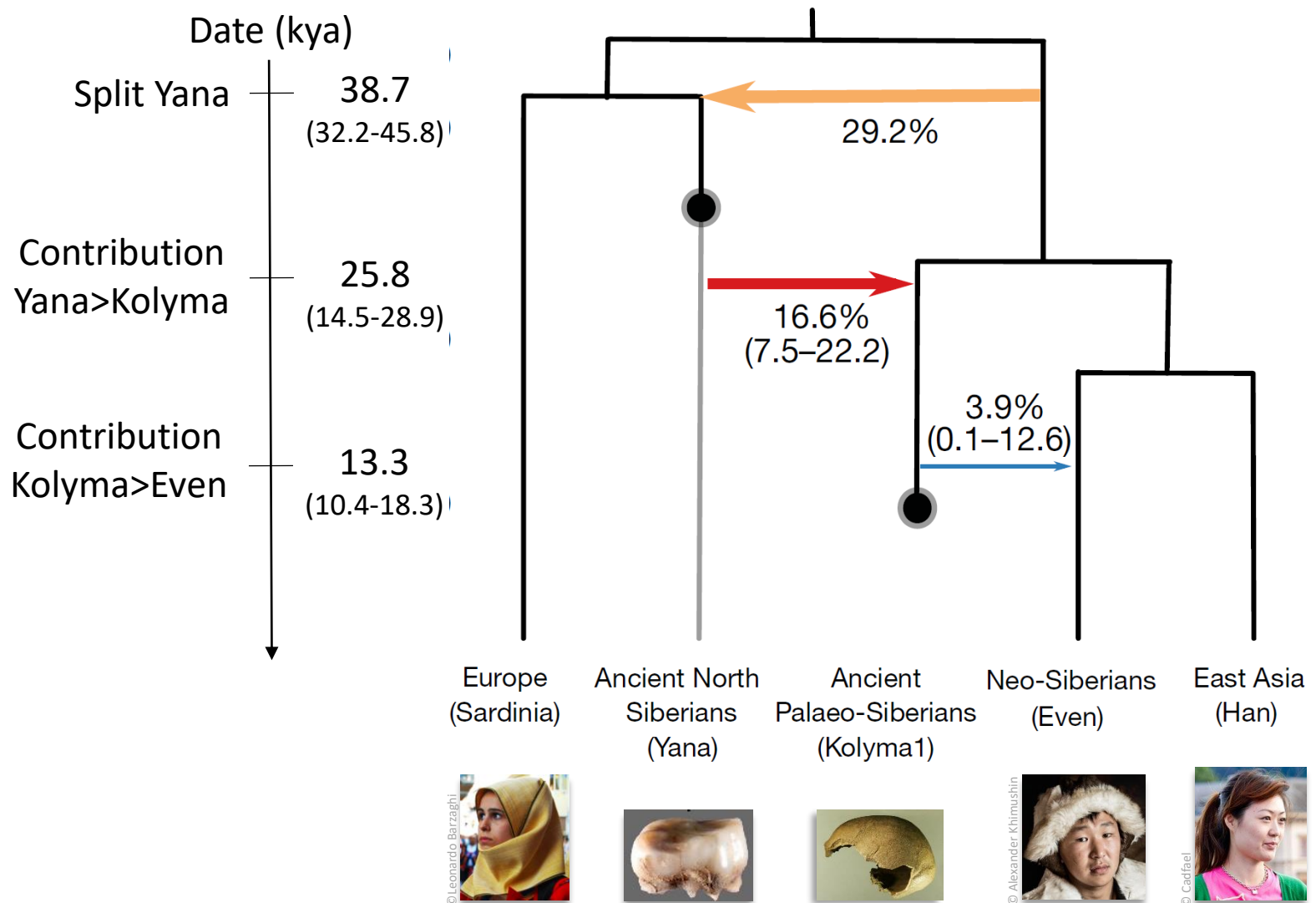
For instance:

$\beta = 1$  indicates continuity:  
Kolyma descends from Yana

$\beta = 0$  indicates replacement  
of Yana by Kolyma



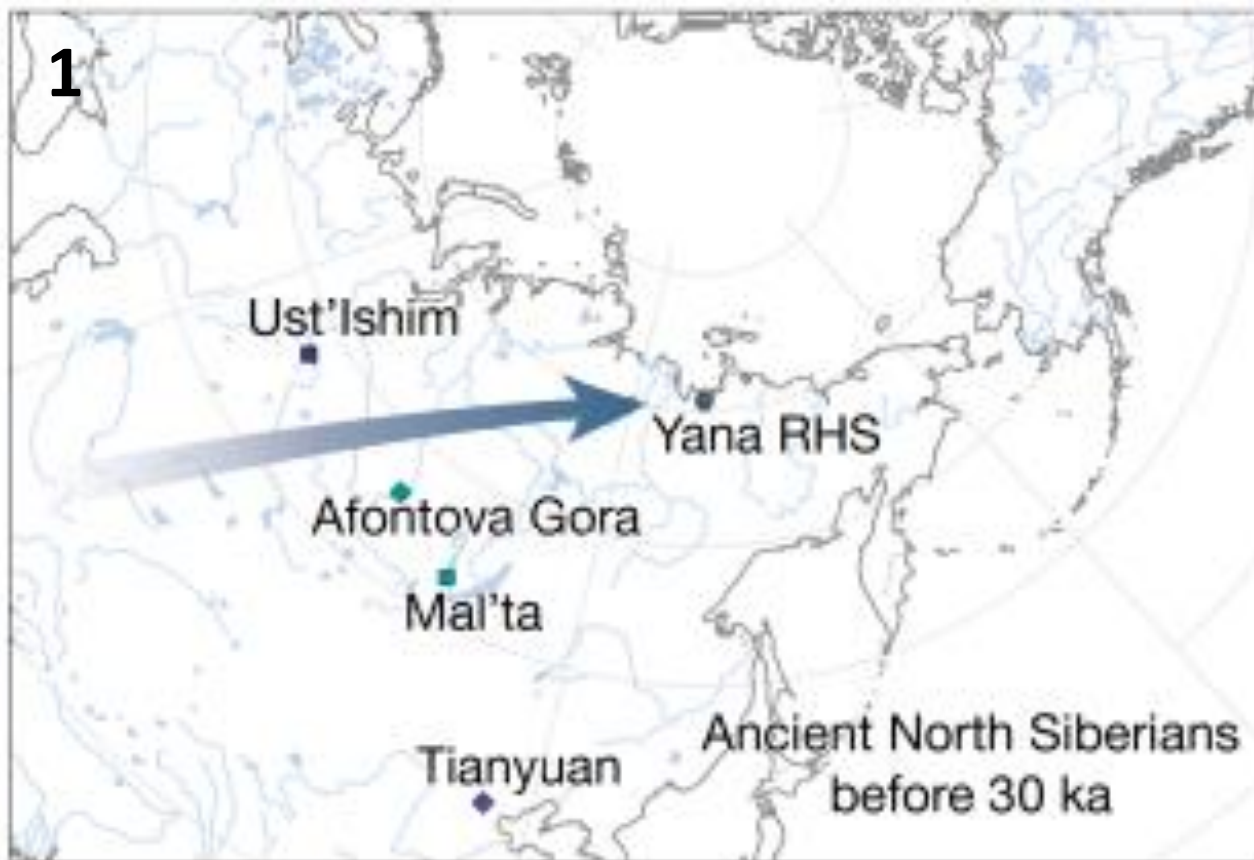
# Estimates indicate replacement with gene flow



# Summary: 3 migration waves



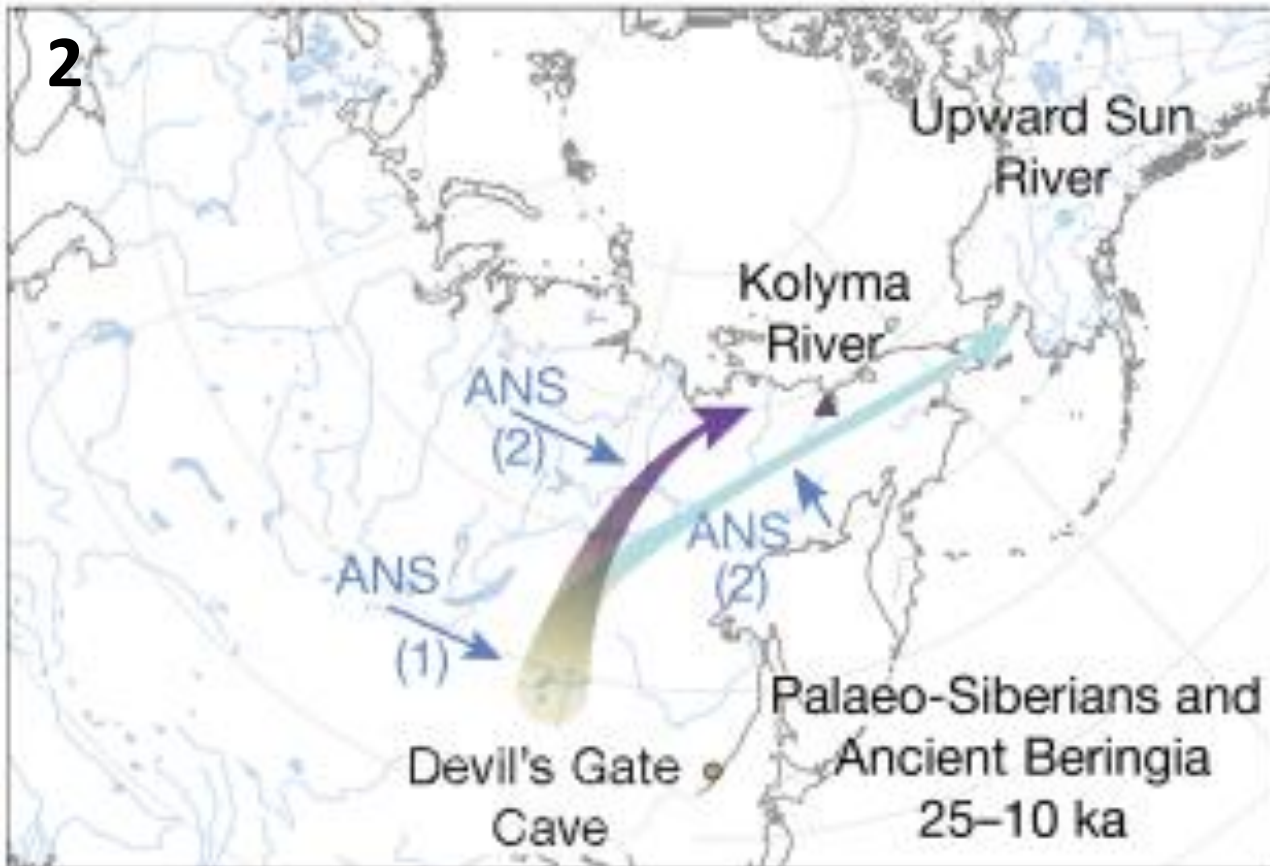
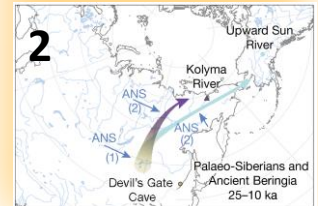
- Ancient North Siberians (Yana) reached Siberia before 30 ka (thousand-years ago)



**1<sup>st</sup> migration wave**

# Summary: 3 migration waves

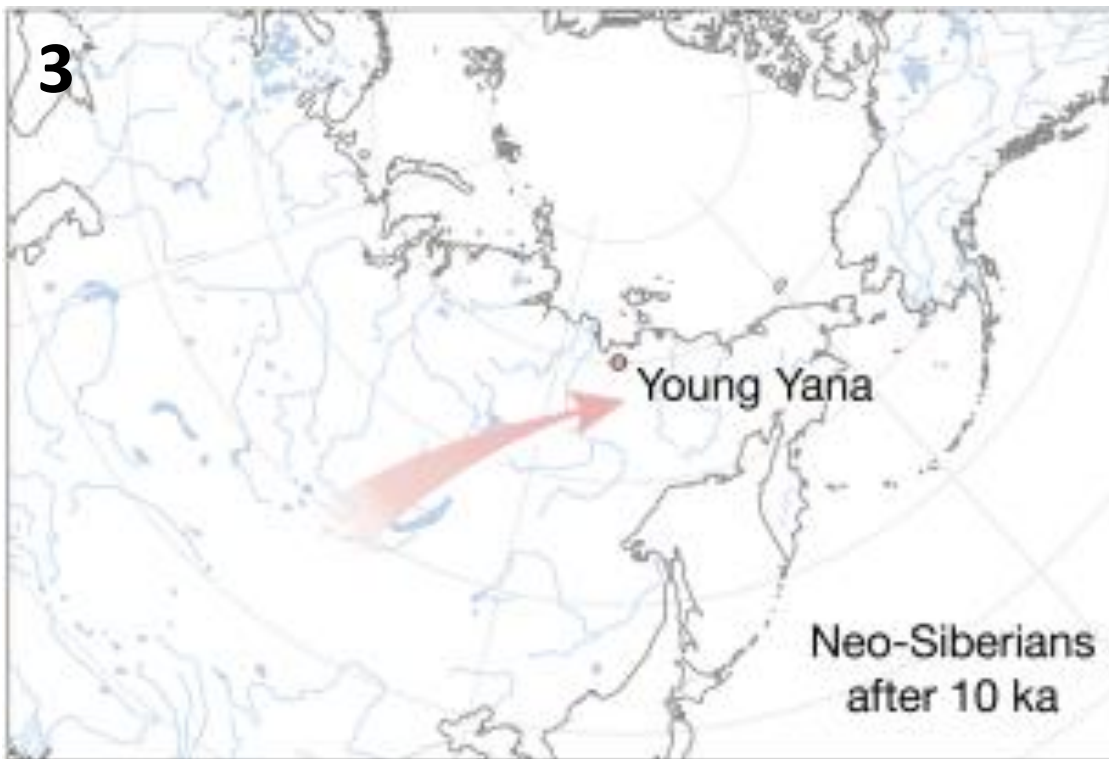
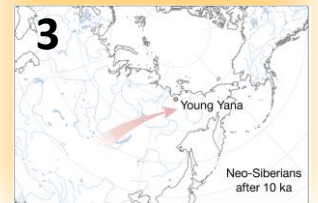
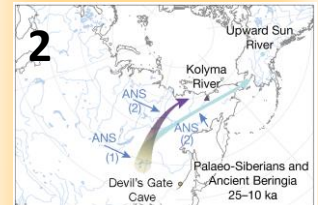
- Ancient North Siberians (Yana) reached Siberia before 30 kya
- Paleo-Siberians (Kolyma) migrated after Last Glacial Maximum (26.5 ka)
- Native-Americans are closer to Kolyma, with 20% of Yana contribution



**2<sup>nd</sup> migration wave**

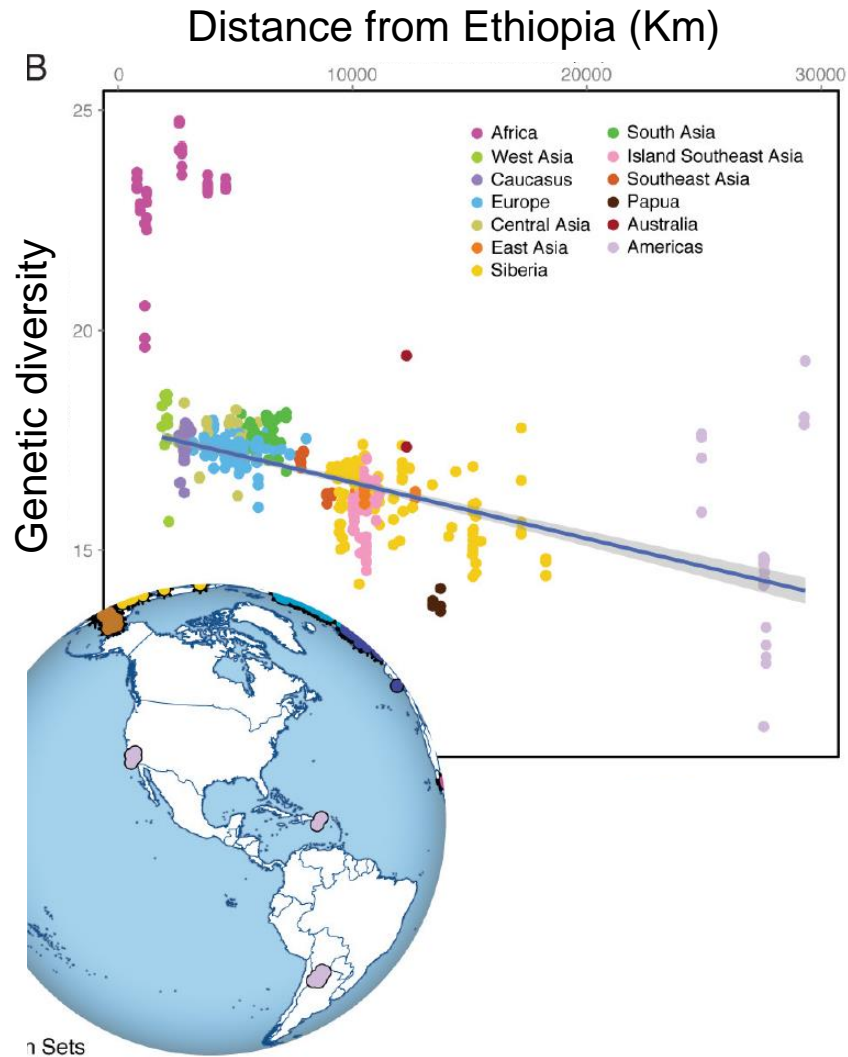
# Summary: 3 migration waves

- Ancient North Siberians (Yana) reached Siberia before 30 ka
- Paleo-Siberians (Kolyma) likely migrated after Last Glacial Maxima
- Native-Americans are closer to Kolyma, with 20% of Yana contribution
- Paleo-Siberians (Kolyma) were replaced by Neo-Siberians, likely associated with the cooler period “Younger Dryas” (12.8-11.5 ka)

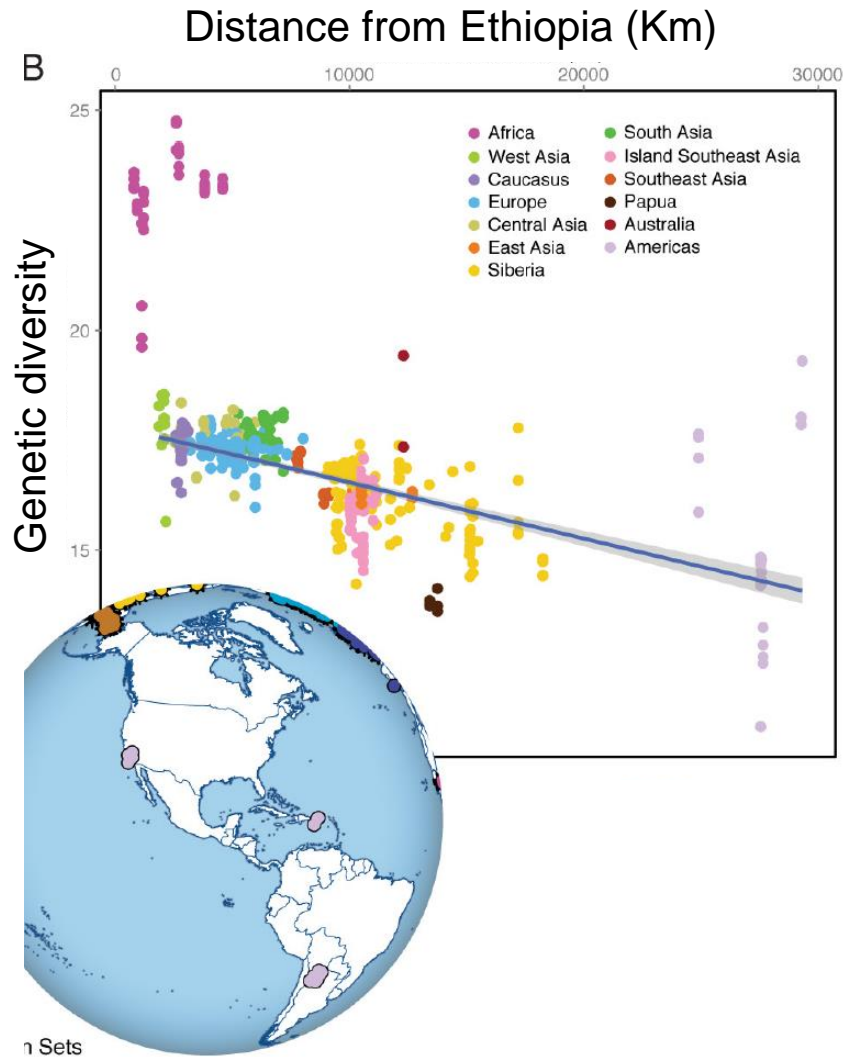


**3<sup>rd</sup> migration wave**

# Effect of population range expansions

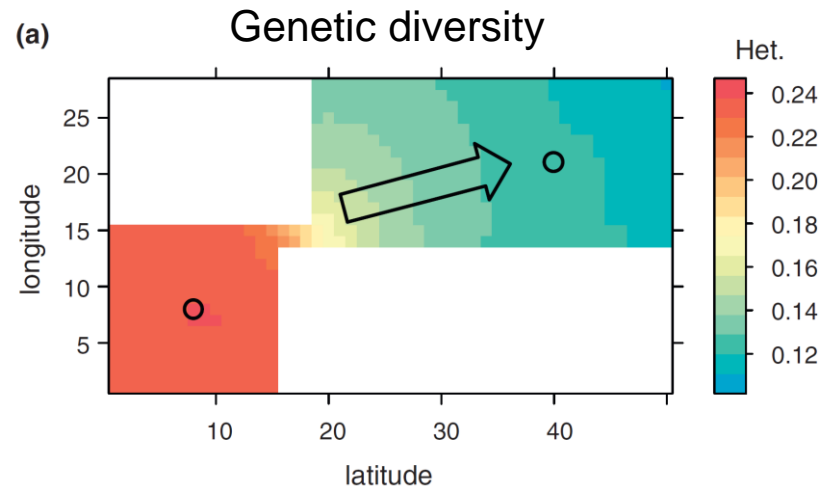


# Effect of population range expansions



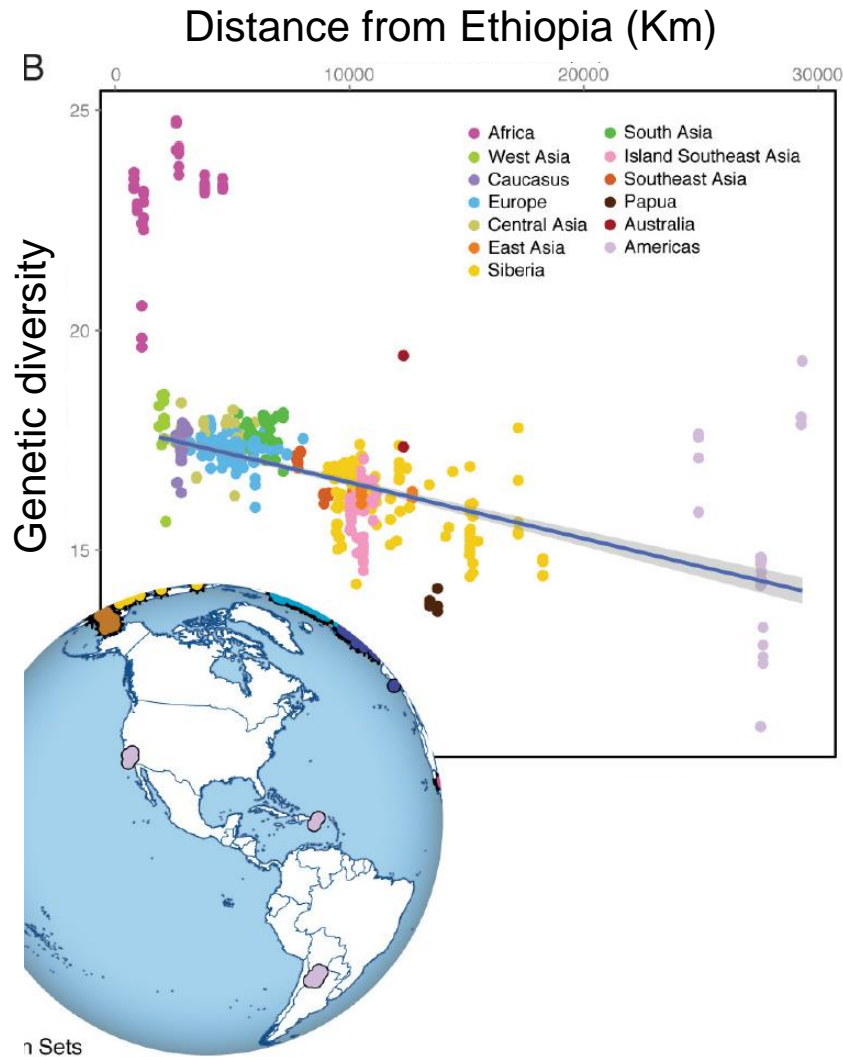
Pagani et al (2016) Nature

## Results from models (simulations)



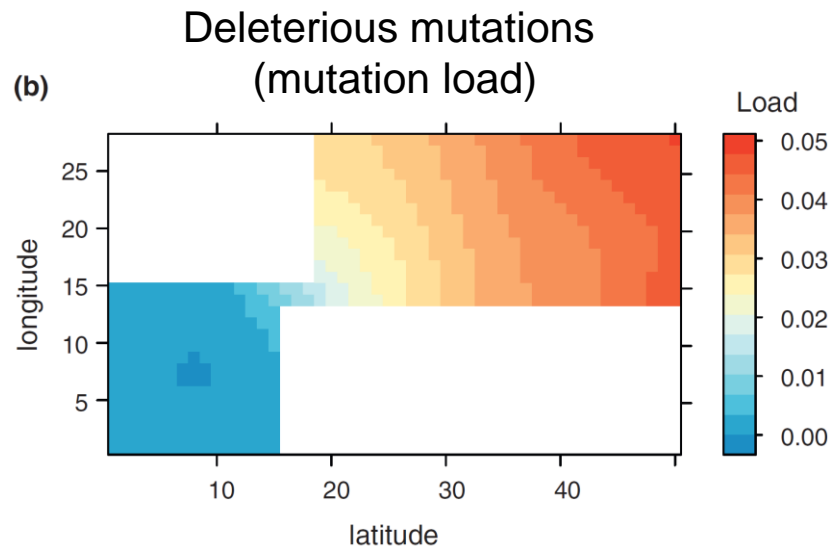
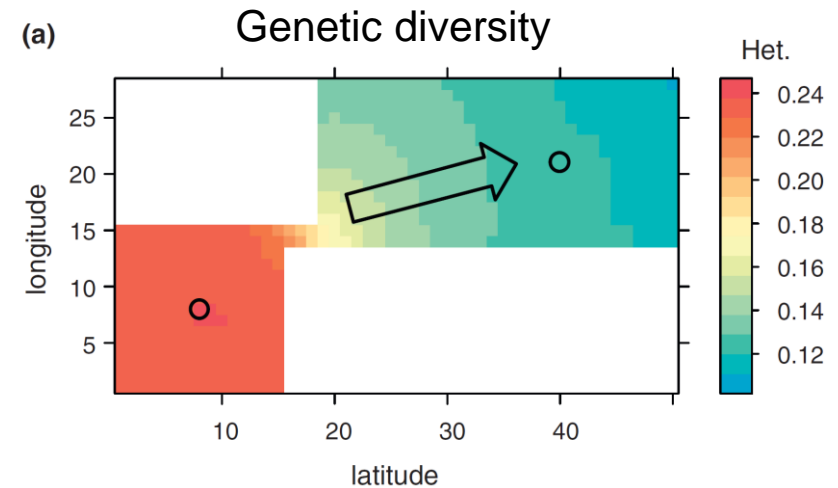
Sousa et al (2014) Curr Opin Genet Dev

# Effect of population range expansions



Pagani et al (2016) Nature

## Results from models (simulations)



Sousa et al (2014) Curr Opin Genet Dev

Bioinformatics and genomic data allow us to reconstruct the evolutionary history of populations

- Evidence of at least three migration waves in Siberia, likely associated with climatic events

Sikora et al (2019) Nature

- Range expansions lead to accumulation of deleterious mutations at the edge of expansion

Sousa et al (2014) Curr Opin Genet Dev

# Acknowledgements

Martin Sikora  
Laurent Excoffier  
Isabelle Dupanloup  
Stephan Peischl  
Eske Willerslev



## Evolutionary Genetics group at cE3c



UID/BIA/00329/2015-2018  
UID/BIA/00329/2019



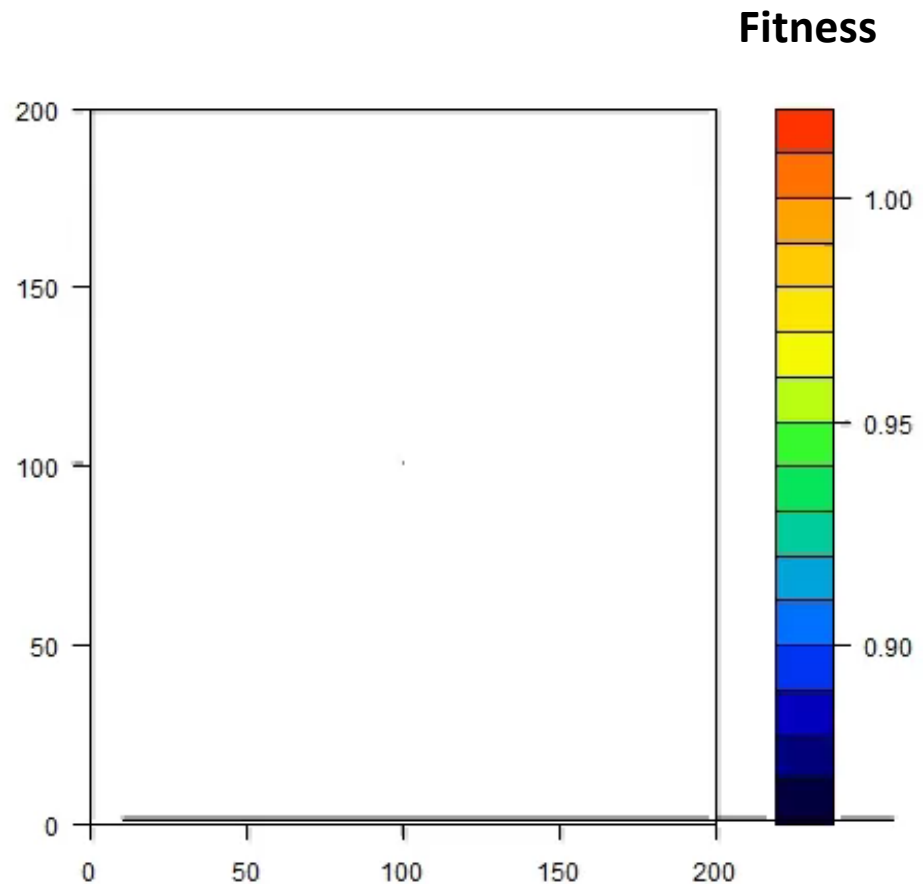
MCSA 2018-2020:  
MAPgenome  
(N.799729)

# Modeling range expansions with simulations

Simulations of spatial expansions show accumulation of deleterious mutations at the edge of expansion

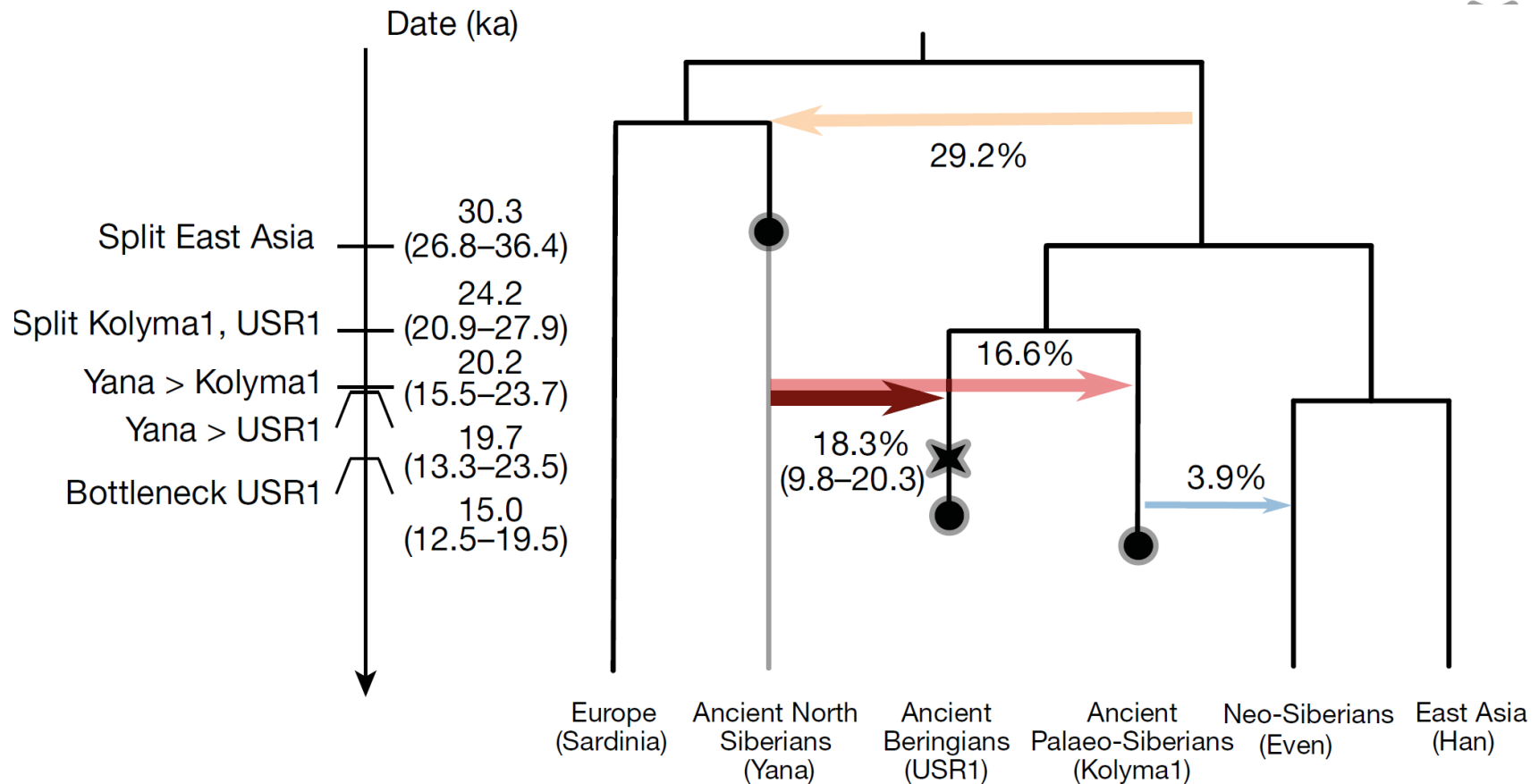
Experimentally verified in different organisms, e.g. in bacteria and humans

Bosshard et al 2017 Genetics;  
Peischl et al 2018 Genetics



Video from Stephan Peischl (Univ. Bern)

# Estimates consistent with replacement with gene flow



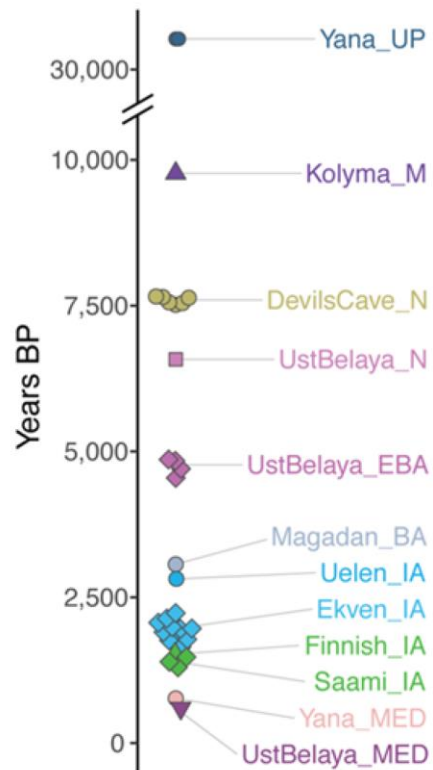
- Kolyma is the closest population to Native Americans (USR1 and Karitiana)
- Native Americans with a contribution of up to 20% from Yana

# Yana genetic legacy is widespread across Eurasia

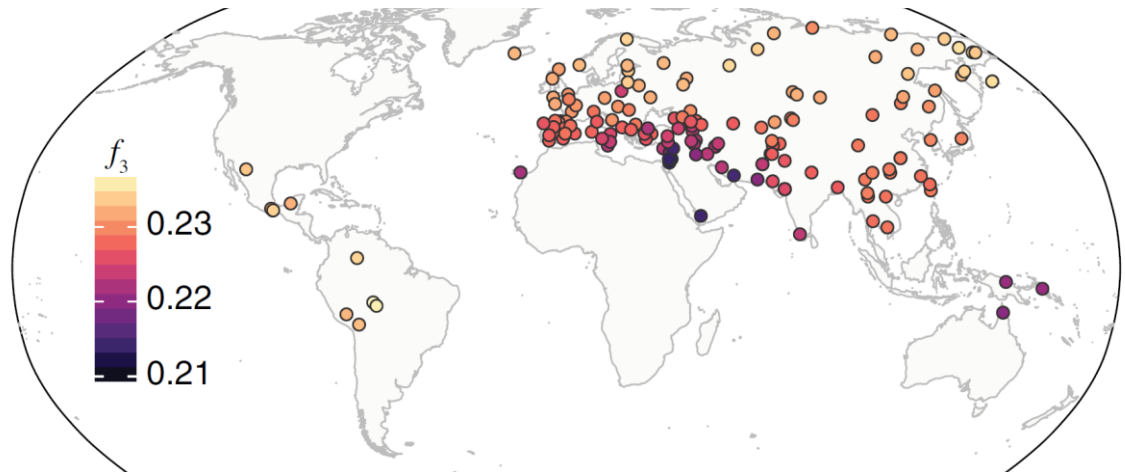
## Genetic similarities

$$f_3(\text{Mbuti}; X, Y)$$

**b**



Yana (31,600 years)



Kolyma (9,800 years)

