Taller Biodiversidad: Herramientas del estudio de la diversidad genética

Análisis genético de la diversidad humana: el Genográfico

Pós-Graduação
Genética
UFMG

Fabrício R. Santos

Universidade Federal de Minas Gerais
7 de Setembro de 1927

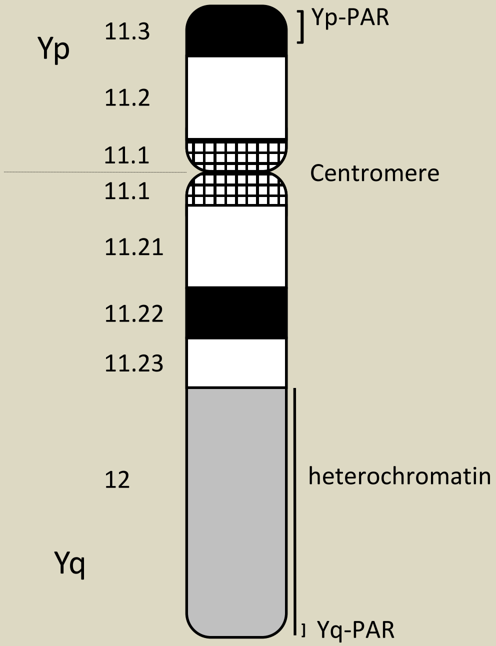
NATIONAL GEOGRAPHIC™ IBM
A research partnership of National Geographic and IBM.

THE WAITT FAMILY FOUNDATION
Global field science supported by the Waitt Family Foundation.

Biólogo, Genetista (fsantos@icb.ufmg.br), Profesor de Evolución
Departamento de Biología General, ICB, UFMG
Belo Horizonte, MG, Brasil

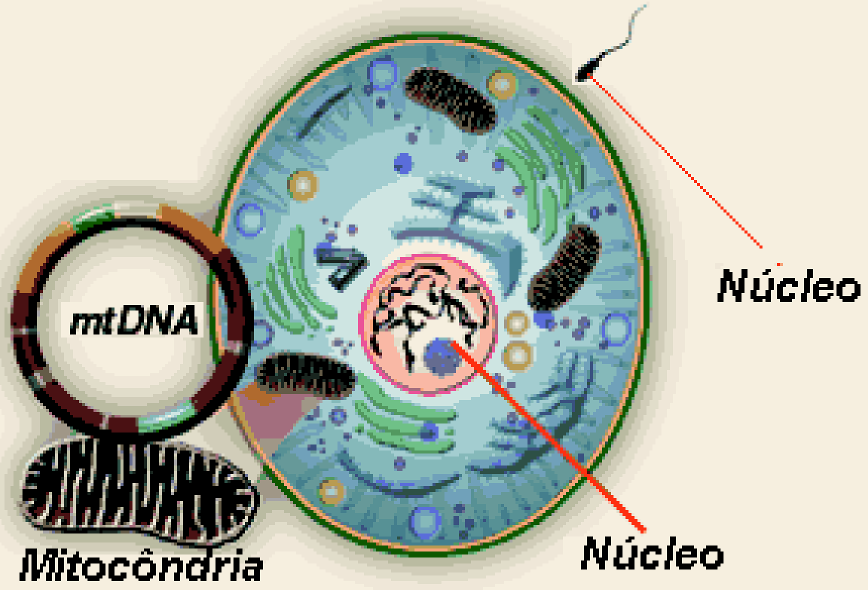
Uniparental markers and genealogical history

Paternal origin
Y chromosome

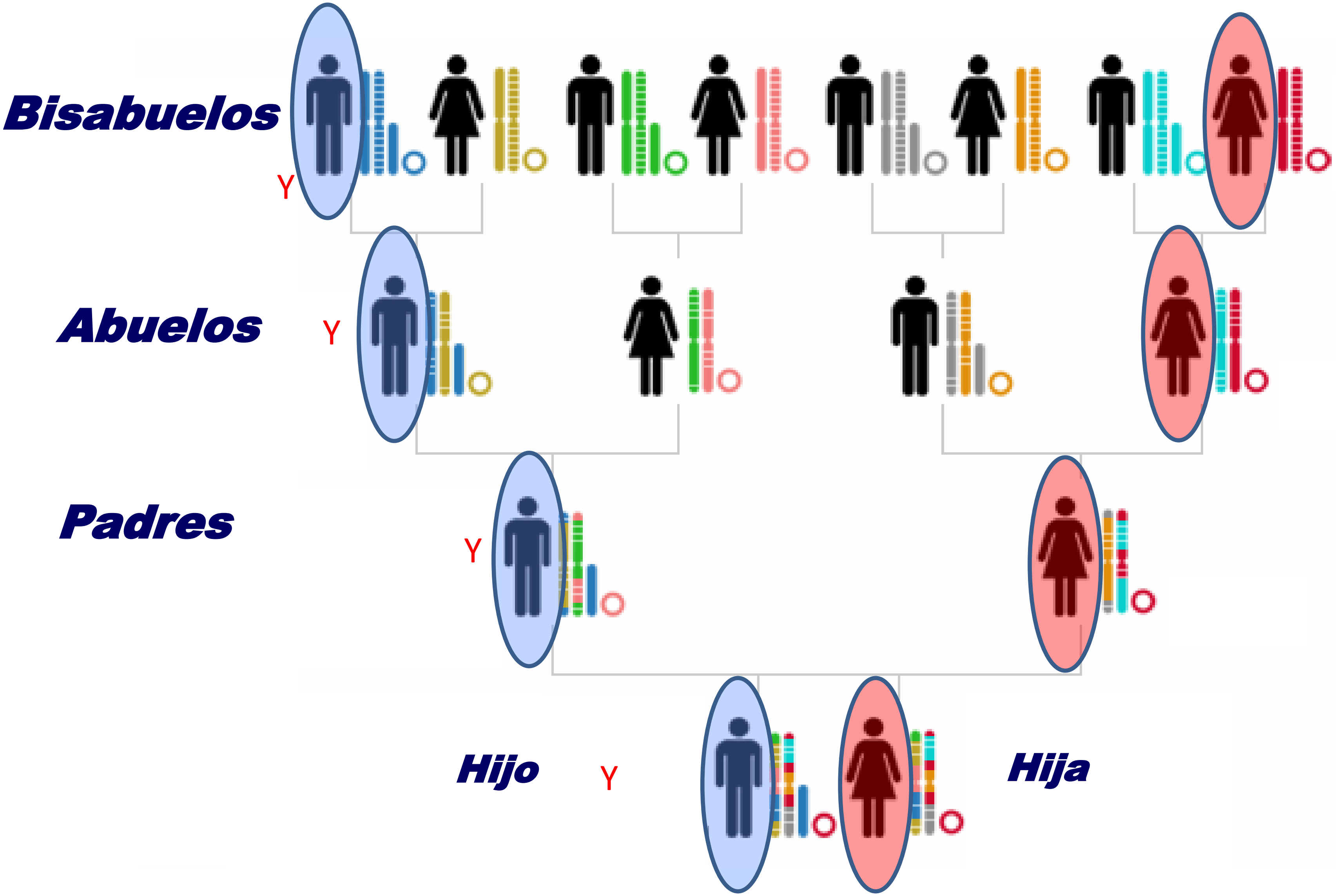


Males
Father to son inheritance

Maternal origin
mtDNA



Females and males
Mother to children inheritance



Bisabuelos

Abuelos

Padres

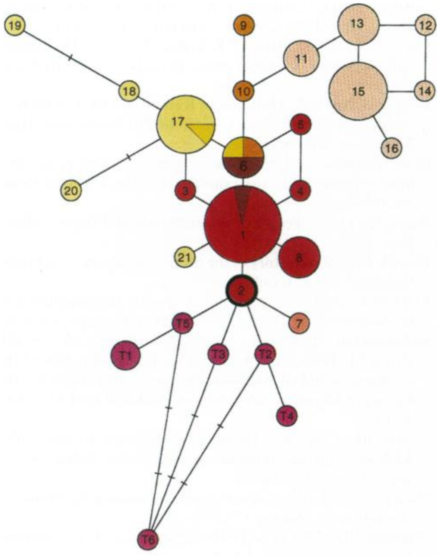
Hijo

Hija

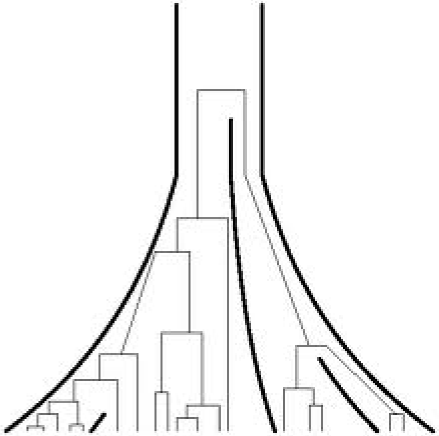
Advantages of studying Y and mtDNA

- No recombination, thus lineages can be directly linked through common ancestors, i.e., we can use an unbiased genealogical approach (descent with modification);
- Admixture is no problem because non-indigenous lineages can be readily recognized;
- Divergence based dating is straightforward, although presenting large confidence limits;
- The effective size is about ¼ of any autosomal locus, thus it is much more sensitive to drift and demographic shifts;
- There is strong geographic and cultural correlation worldwide in Y and mtDNA than any other known single autosomal locus in the genome, although Y is larger than mtDNA (usually).

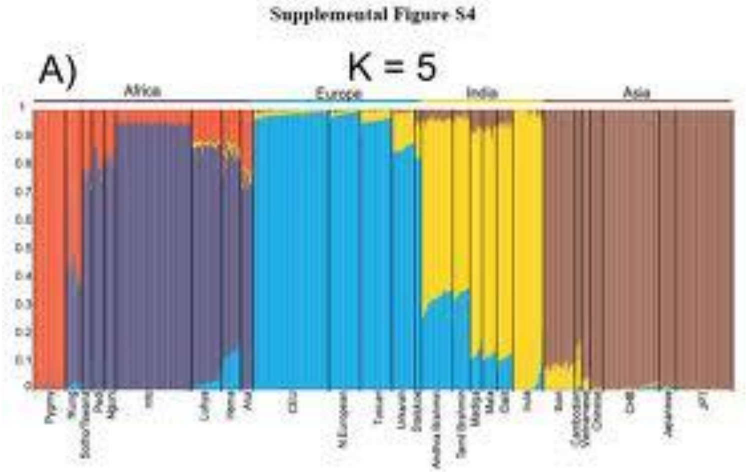
Redes filogeográficas



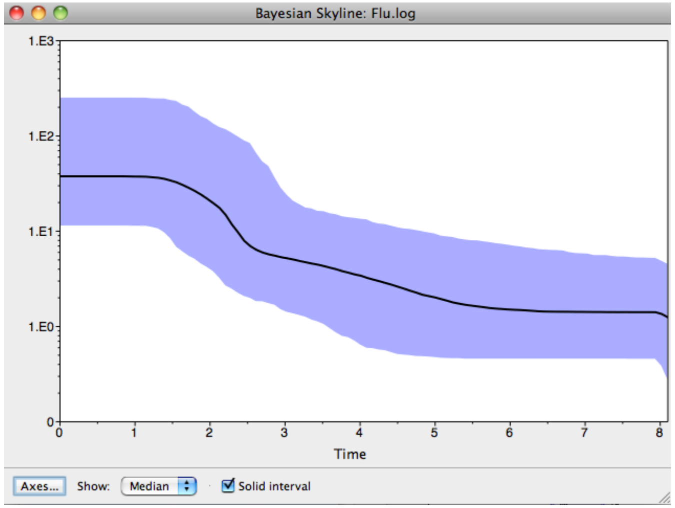
Fechamientos de divisiones poblacionales



Estructuración poblacional


Supplemental Figure S4

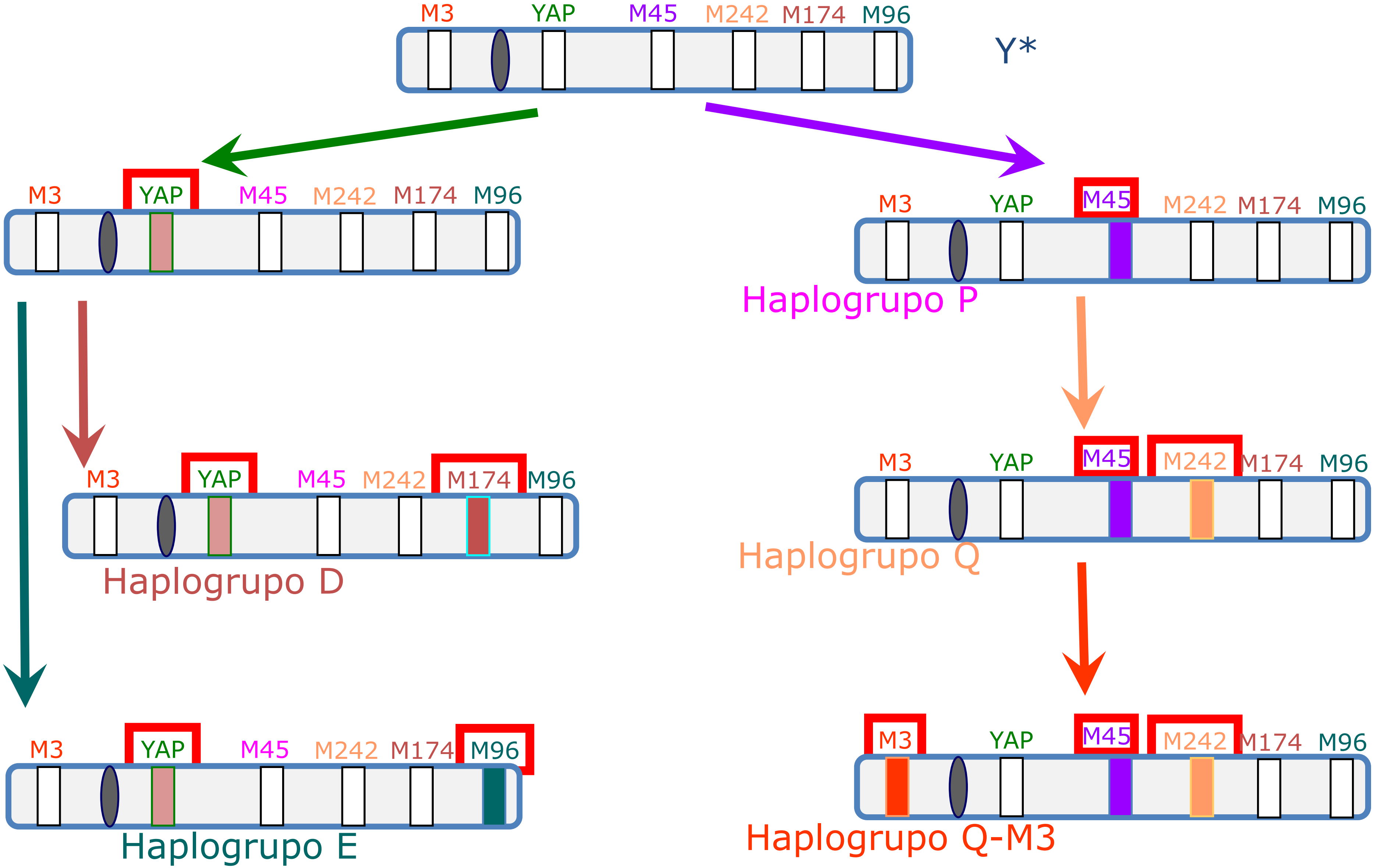
Reconstrucciones paleodemográficas

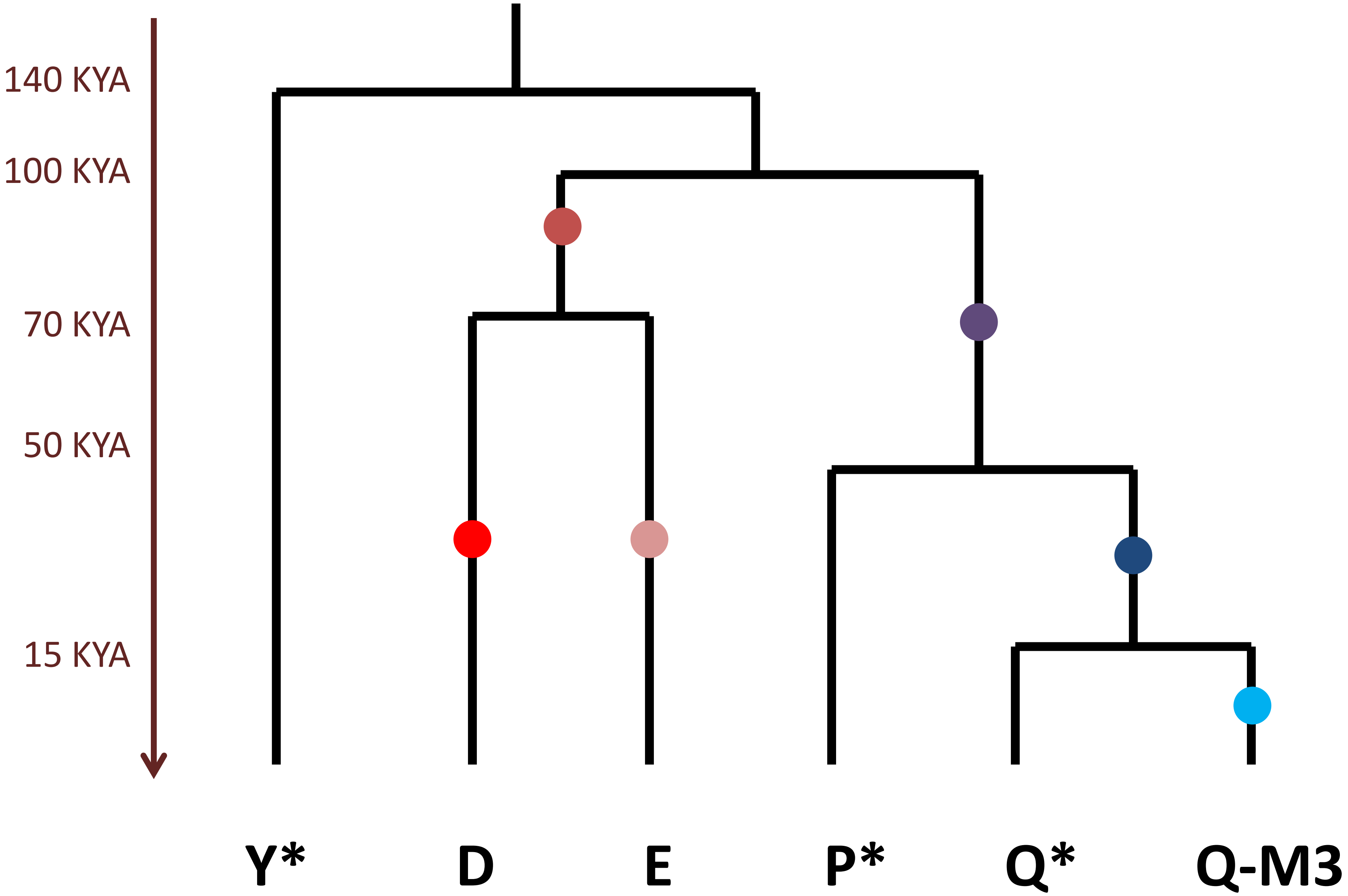

Bayesian Skyline: Flu.log

Show: Median

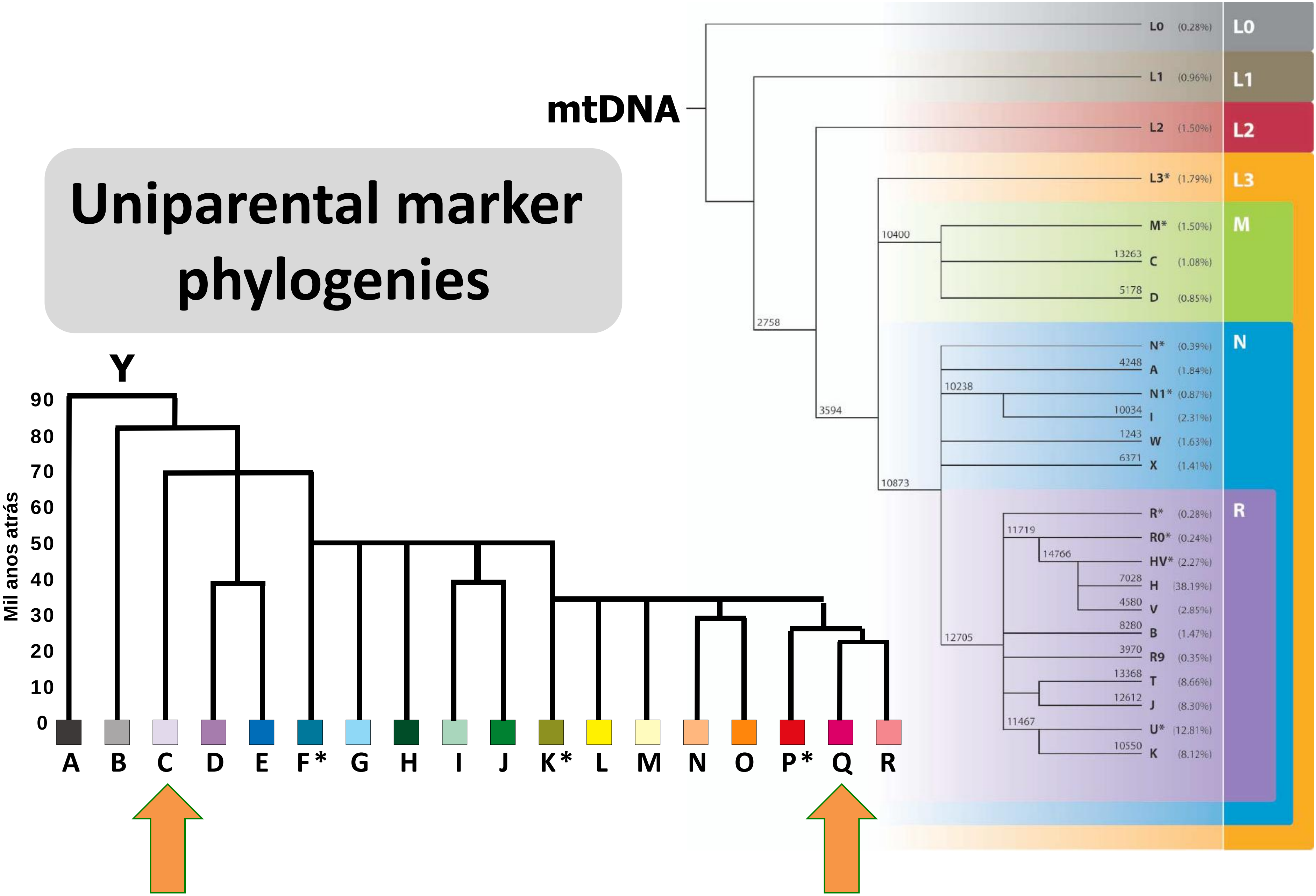
☑ Solid interval

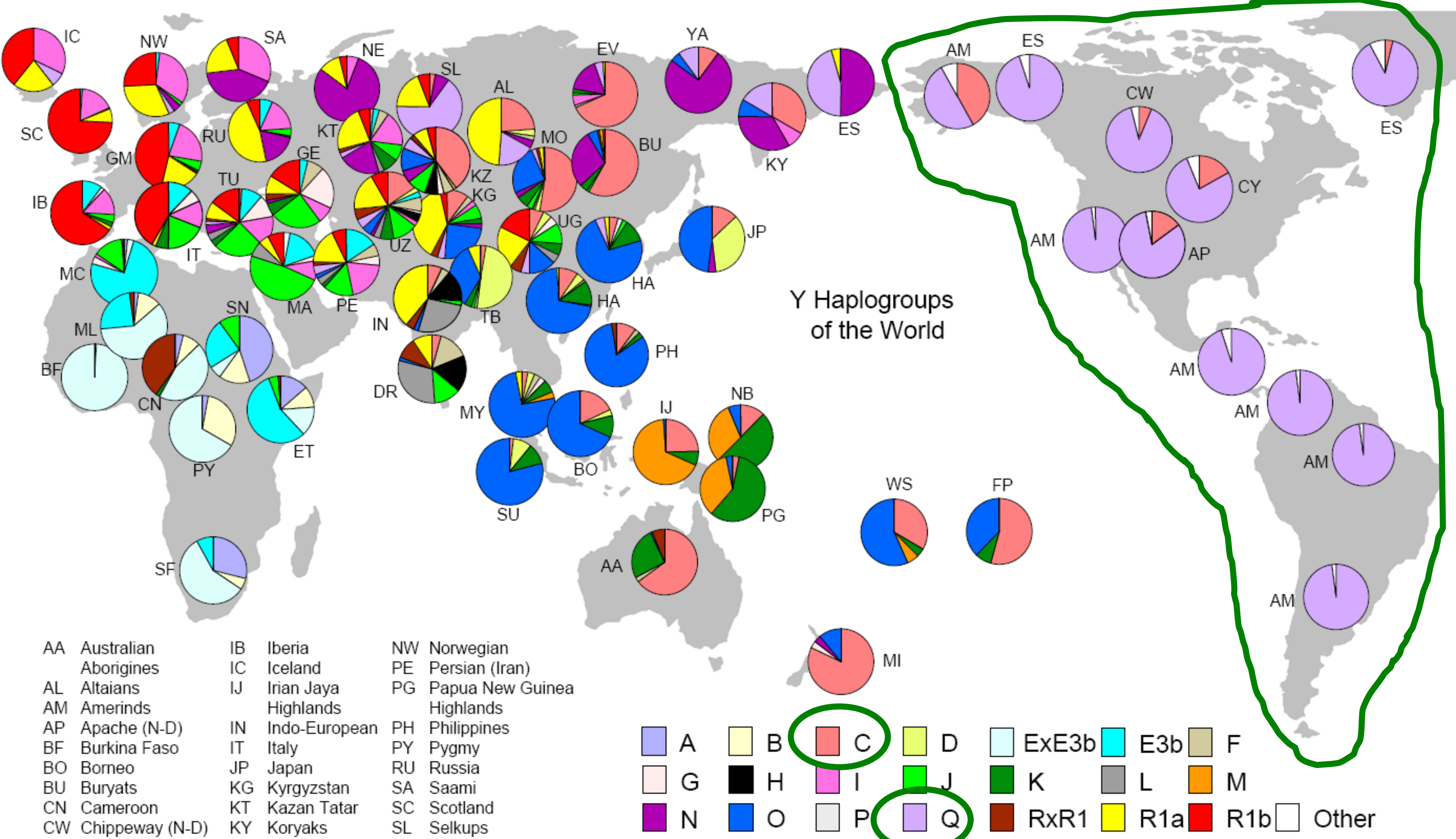
Generation of Y haplogroups through mutation accumulation




140 KYA
100 KYA
70 KYA
50 KYA
15 KYA
Y*
D
E
P*
Q*
Q-M3

Uniparental marker phylogenies

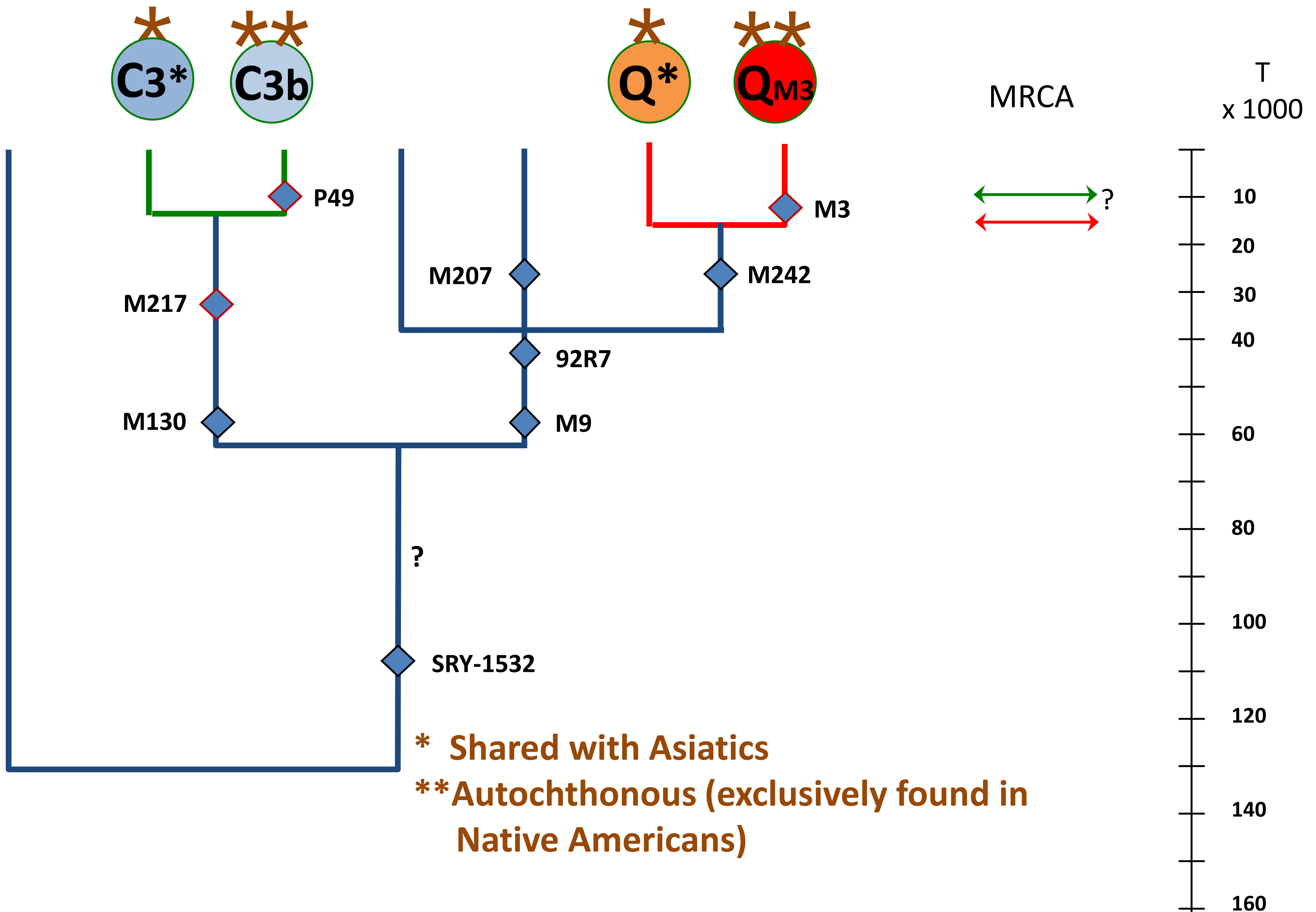




AA	Australian Aborigines
AL	Altaians
AM	Amerinds
AP	Apache (N-D)
BF	Burkina Faso
BO	Borneo
BU	Buryats
CN	Cameroon
CW	Chippeway (N-D)
CY	Cheyenne
DR	Dravidian
ES	Eskimos
ET	Ethiopia
EV	Evenks
FP	French Polynesia
GE	Georgia-Armenia
GM	Germany
HA	Han Chinese
IB	Iberia
IC	Iceland
IJ	Irian Jaya Highlands
IN	Indo-European
IT	Italy
JP	Japan
KG	Kyrgyzstan
KT	Kazan Tatar
KY	Koryaks
KZ	Kazakhstan
MA	Mideast Arabs
MC	Morocco
MI	Maori
ML	Mali
MO	Mongols
MY	Malaysia
NB	New Britain
NE	Nenets
NW	Norwegian
PE	Persian (Iran)
PG	Papua New Guinea Highlands
PH	Philippines
PY	Pygmy
RU	Russia
SA	Saami
SC	Scotland
SL	Selkups
SF	South Africa
SN	Sudan
SU	Sumatra
TB	Tibet
TU	Turkish
UG	Uygurs
UZ	Uzbek
WS	Western Samoa
YA	Yakuts

The data in this map is supposed to represent the situation before the recent European expansion beginning about 1500 AD. In some cases such as some Native American tribes and the Maori this can be done reliably because STR typing was done. In other cases, especially in America, it is guesswork. The “Other” sectors in America indicate this. Native American groups are labeled by language group as Amerind, Na-Dene (N-D), and Eskimo. F, K, L, and P are in some cases “catchall” groups because some researchers did not use enough markers for a full haplotype determination.

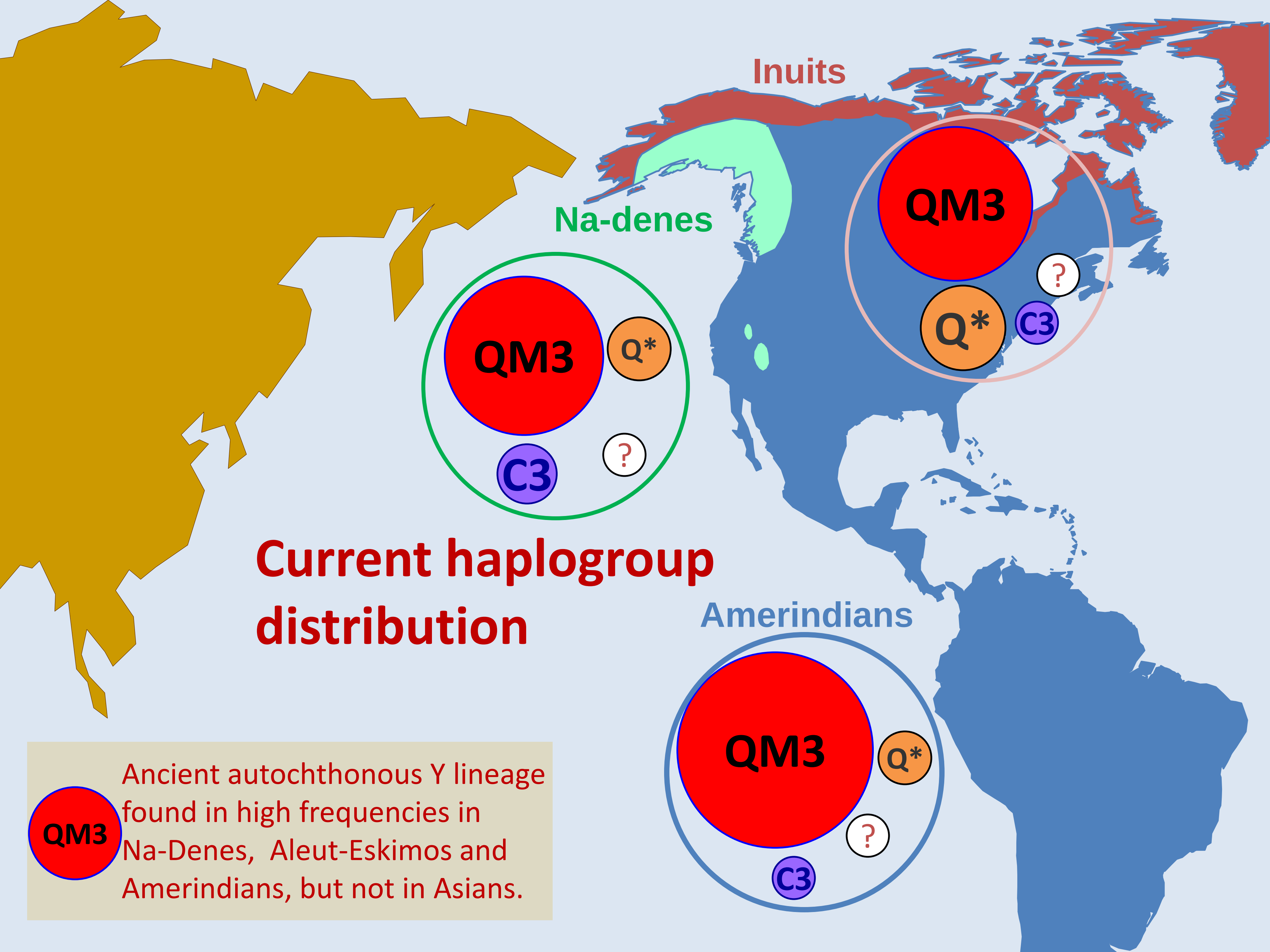
Estimates of T_{MRCA} for Native American Y Lineages



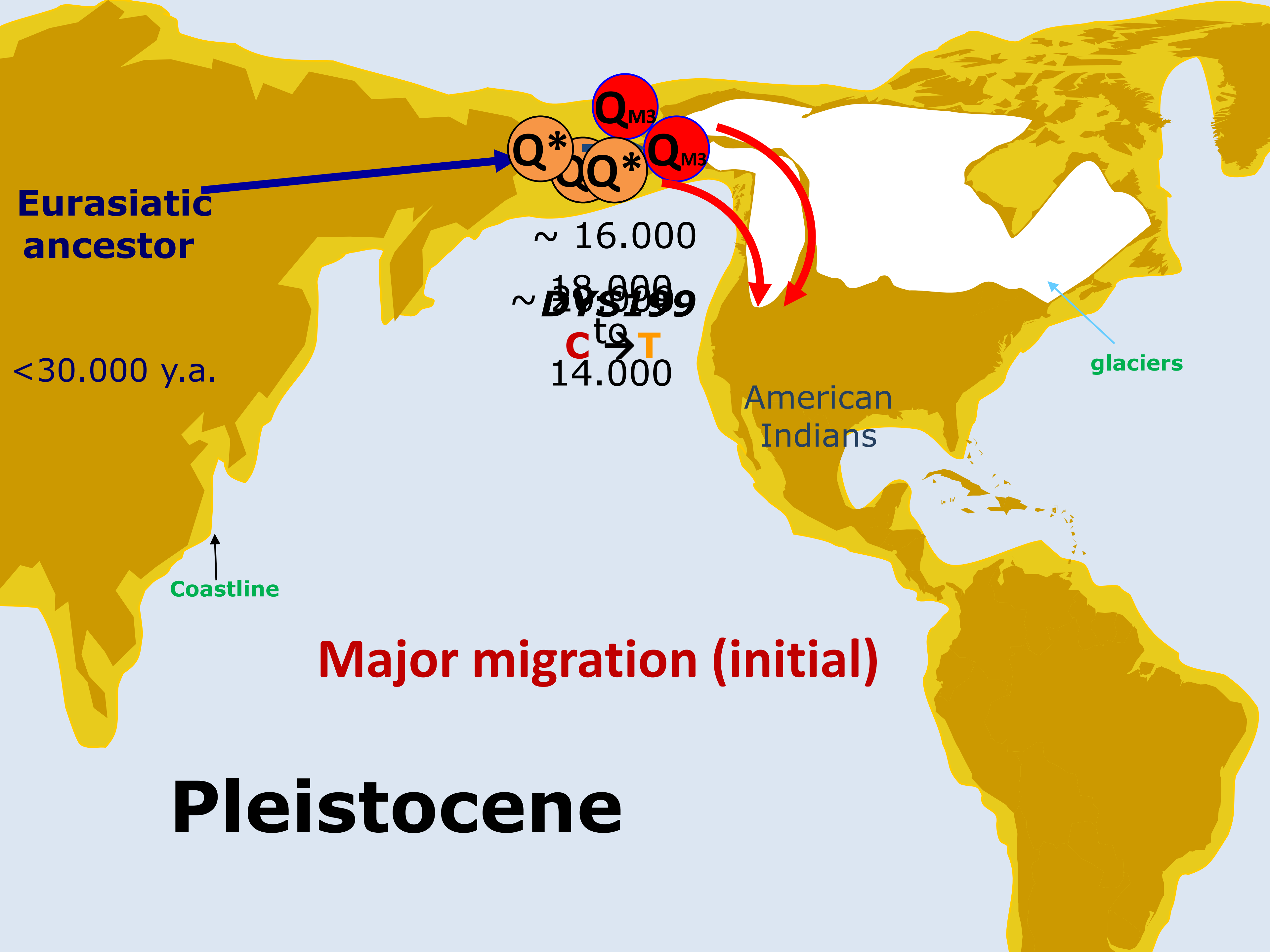
* Shared with Asiatics

**Autochthonous (exclusively found in Native Americans)

Current haplogroup distribution

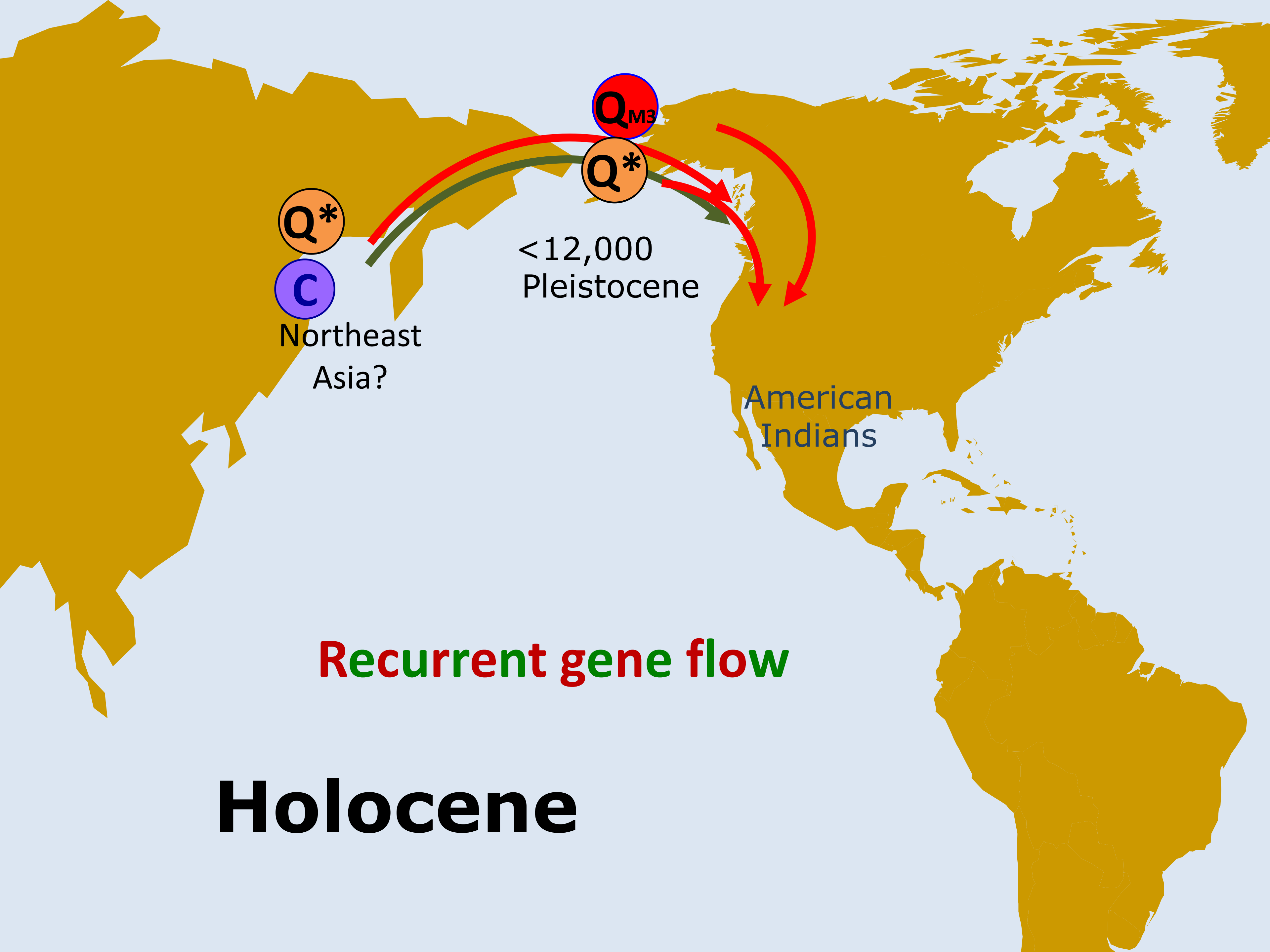


QM3 Ancient autochthonous Y lineage found in high frequencies in Na-Denes, Aleut-Eskimos and Amerindians, but not in Asians.



Major migration (initial)

Pleistocene



Recurrent gene flow

Holocene

Sudamérica

- **Bolivia**: Susana Revollo
- **Perú**: Ricardo Fujita
- **Ecuador**: César Paz-y-Miño
- **Brasil**: Fabrício Santos
- **Guiana Francesa**: Benoit de Thoisy
- **Colombia**: Guillermo Barreto y William Usaquén
- **Chile, Costa Rica ???**

Samples analyzed (2007 - 2011)

Country	Communities /ethnic groups	# Participants
PERU	88 / >25	1037 (33 women)
BOLIVIA	85 / 29	973 (56 women)
ECUADOR	22 / 5	123 (only men)
BRAZIL	39 / 23	193 (9 women)
TOTAL	**234 / >80**	**2326*** **(98 women)**

2326 participants
234 sampling sites

Peru (1037/88)
Bolivia (973/85)
Ecuador (123/22)
Brazil (193/39)

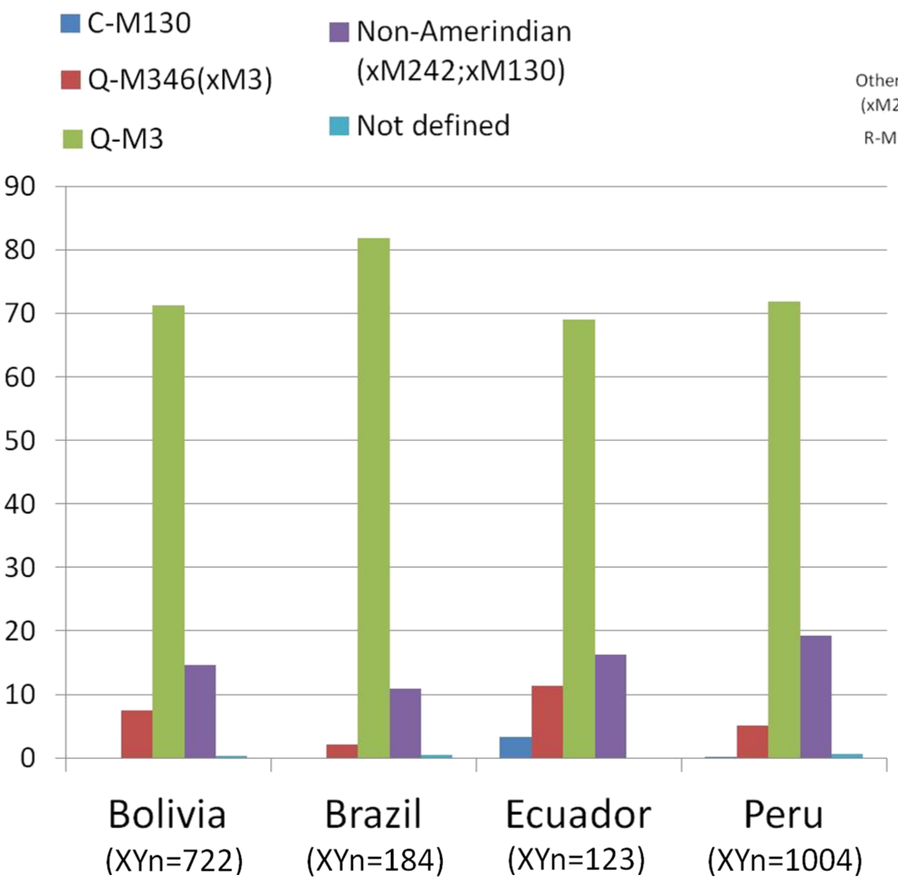
+ **1000** native individuals from Brazil and Peru

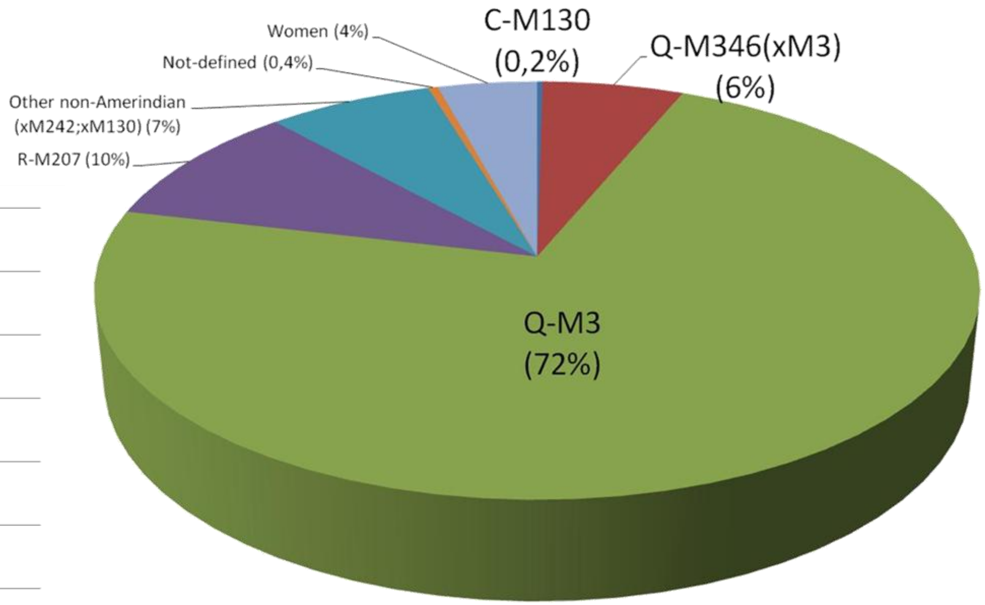
+ Tad Schurr - North American data

+ Salzano´s group data

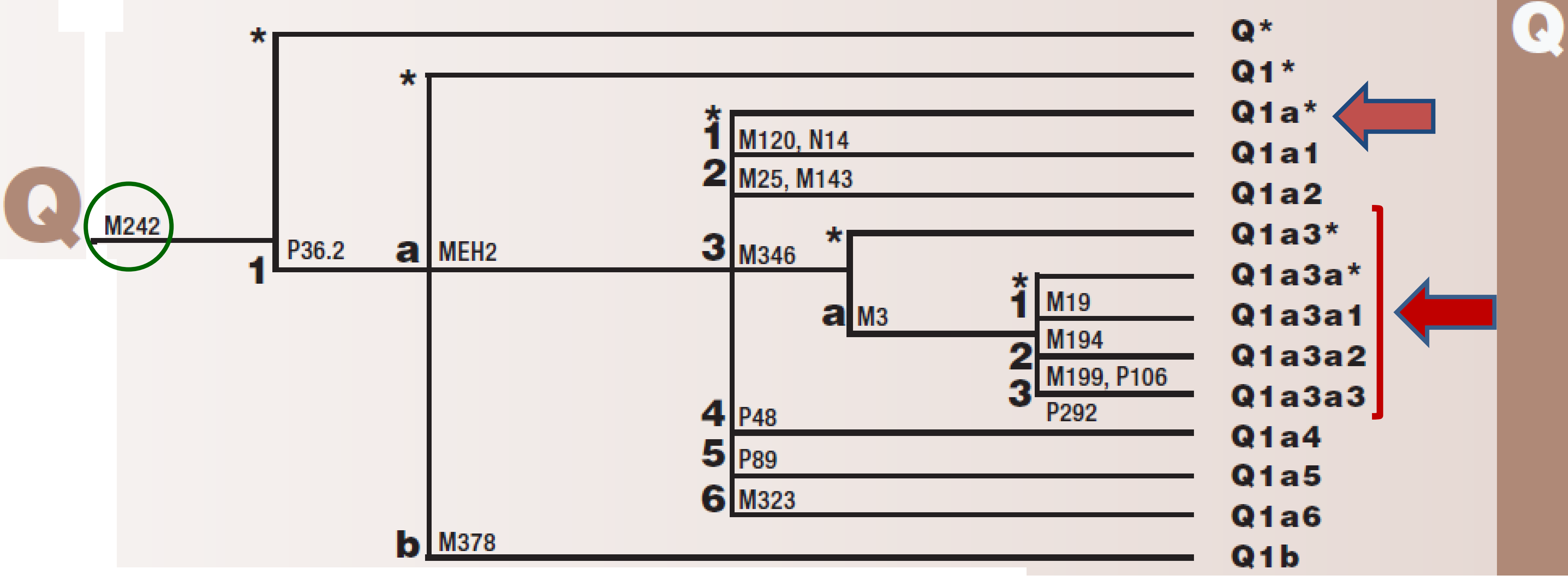
+ Published data

Y chromosome Hg Frequencies





- 92% of male samples were genotyped
- 82% of male samples are Amerindian (1665 Q and 5 C)
- 57% of non-Q/C samples are R-M207



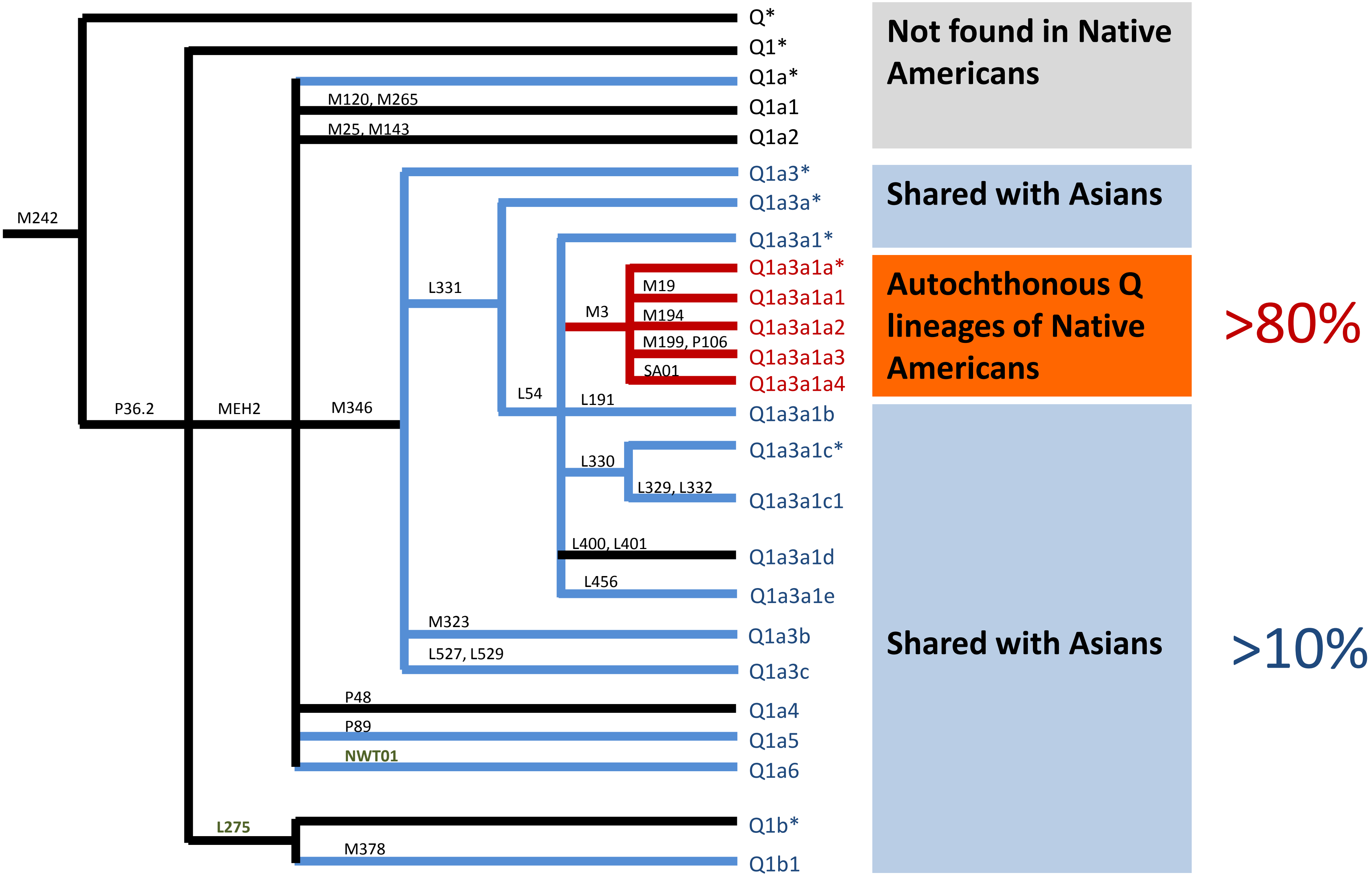
Q1a*(Q-MEH2*)

PALEOESKIMO
(Rasmussen *et al*. 2010)

Q1a3 (Q-M346 and Q-M3)

CONTEMPORANEOUS NATIVE AMERICANS

Q haplogroup in Native Americans



Not found in Native Americans

Shared with Asians

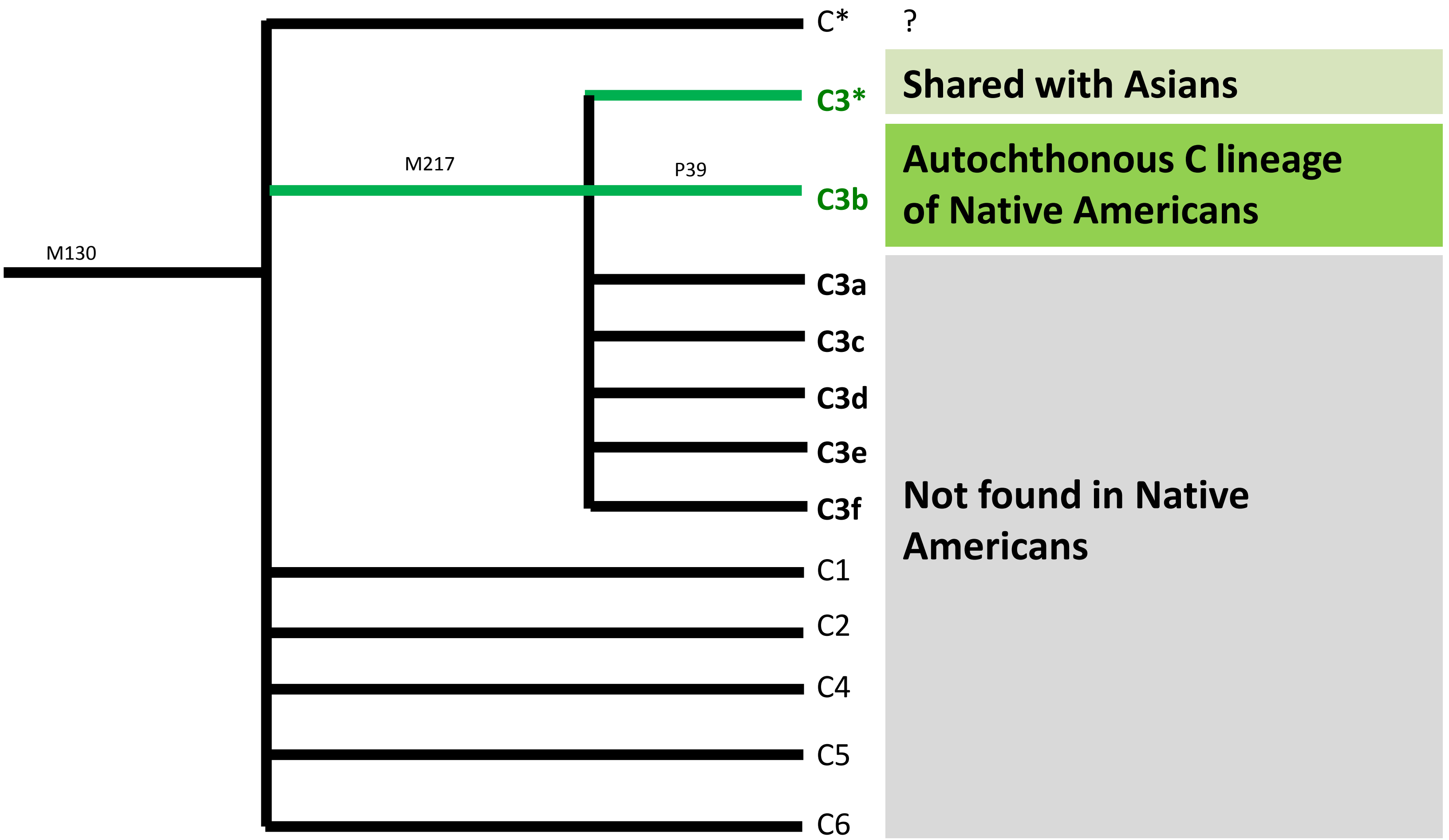
Autochthonous Q lineages of Native Americans

>80%

Shared with Asians

>10%

C haplogroup in Native Americans



?

Shared with Asians

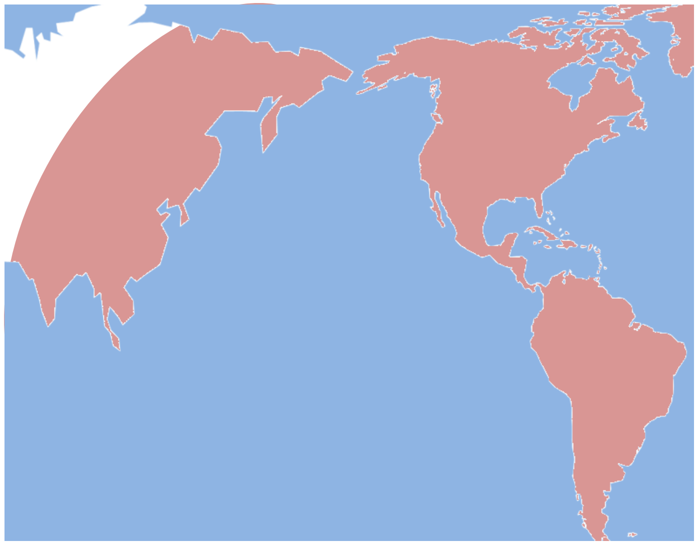
Autochthonous C lineage of Native Americans

<3%

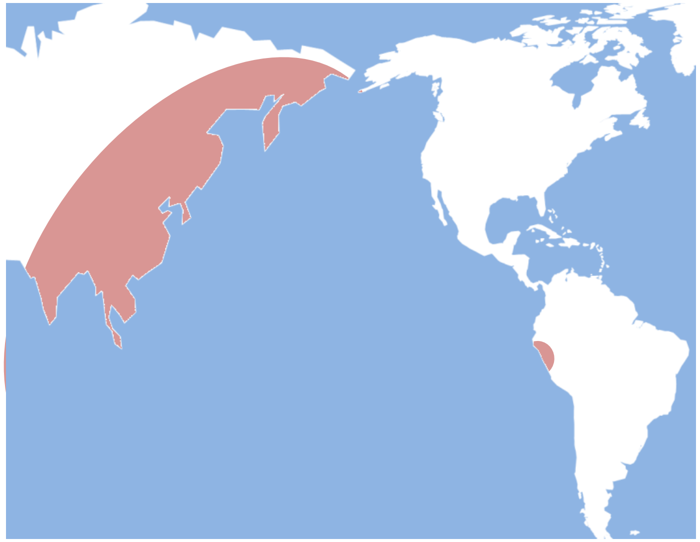
Not found in Native Americans

Minor Y lineages: small gene flow during Holocene?

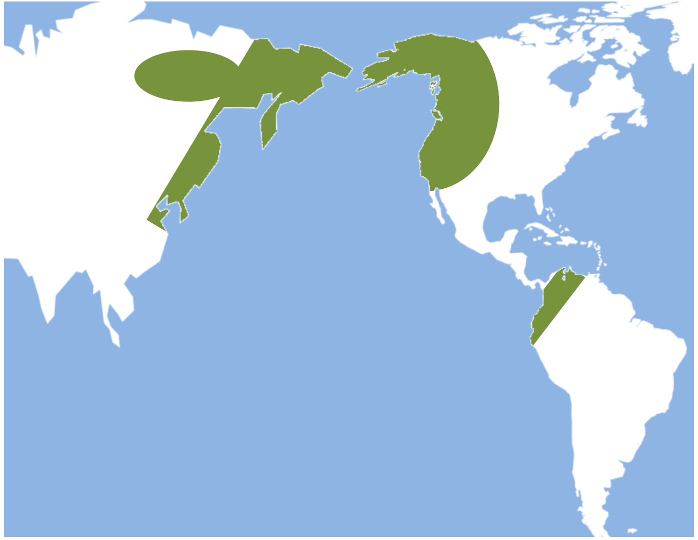
Q1a3*(M346)



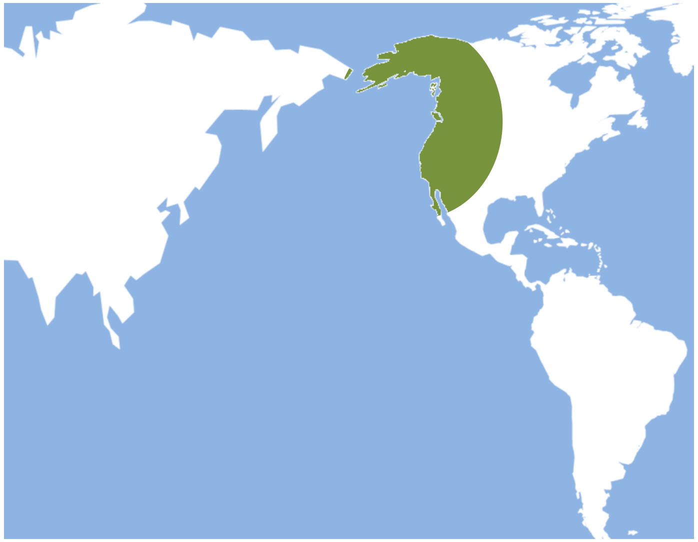
Q1b1(M378)



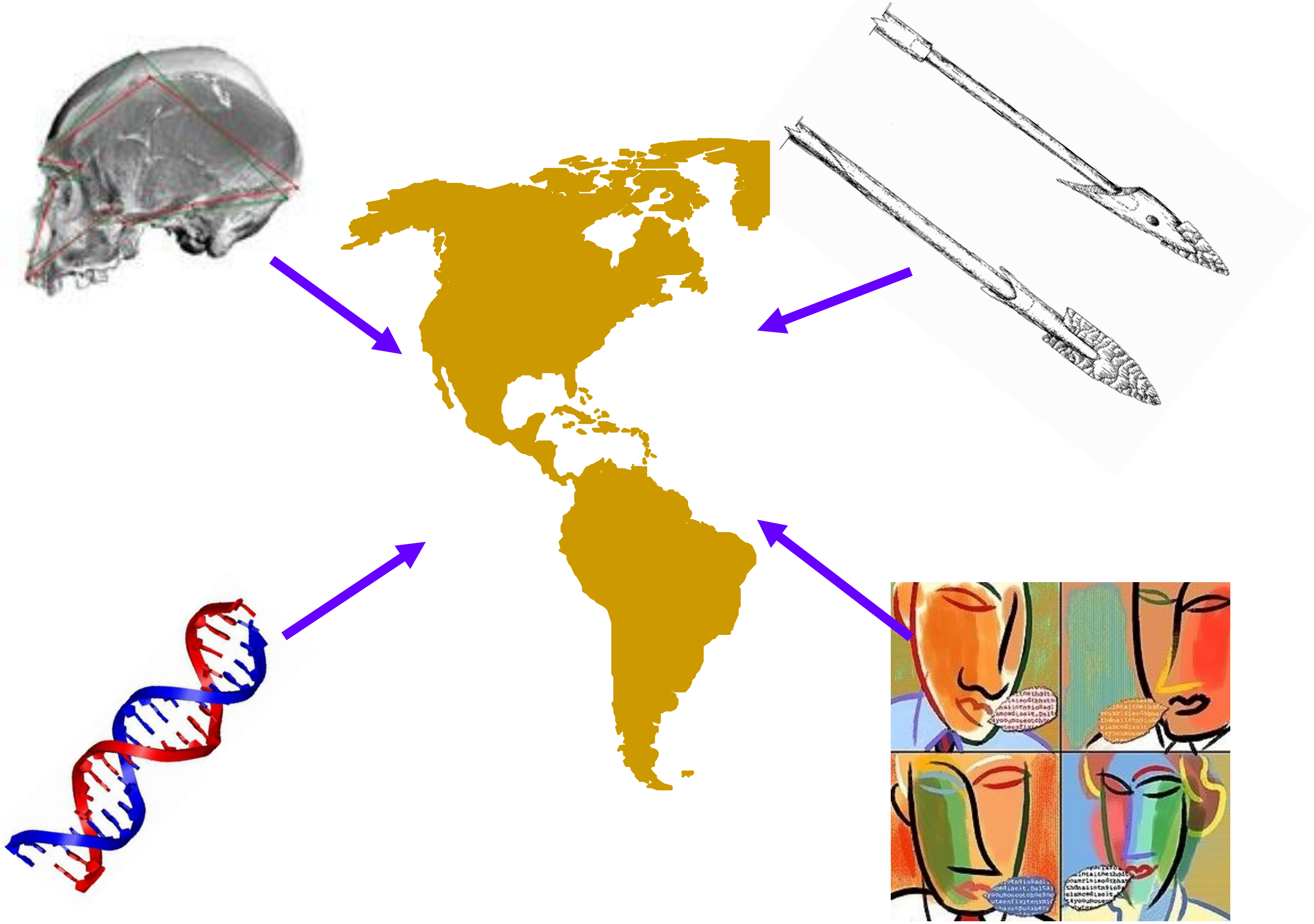
C3*(M217)



C3b(P39)



Integrative history

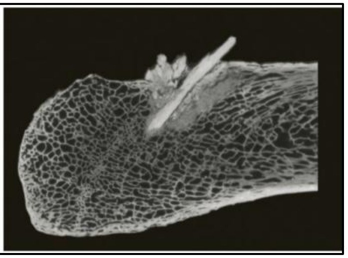


Pleistocenic evidences in America

~13,500 y.o. (11,500 rc. y.) human skull (Luzia) found in Lagoa Santa (Brazil)

14,800 y.o. human settlement found in Monte Verde (Chile)

13,800 y.o. hunted mammoth found in Washington State (US)

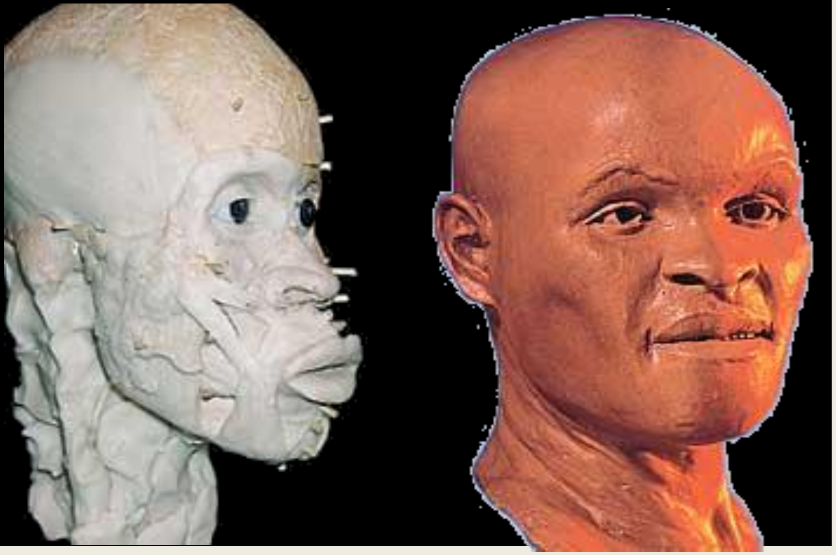


14,300 y.o. feces and tools found in Oregon State (US)

Facial reconstructions of the first natives

Luzia

Lapa Vermelha, MG, Brazil
13.500 YA



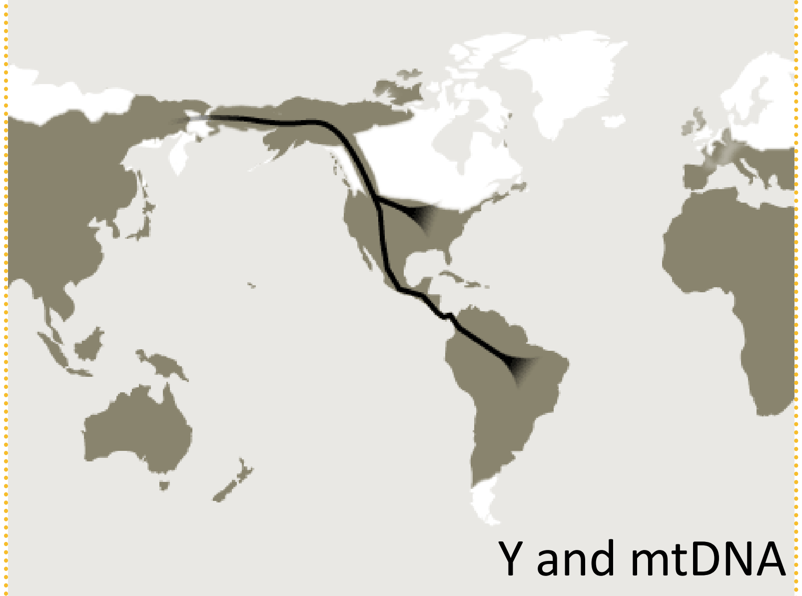
Spirit Cave man

Nevada, USA
10.600 YA

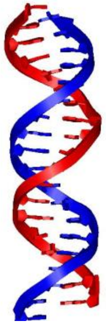
Paleoindians (>8.000 YA) – plesiomorphic traces (pre-mongoloids)

Hypotheses: bones *vs* molecules

DNA: Single major wave (Out of Beringia)

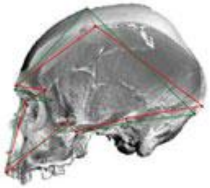

Y and mtDNA

- Most are descendants from a source population from Northeast Asia in the Pleistocene.
- Confinement in Beringia during late Pleistocene



Bones: Two biological components


W. Neves *et al*.

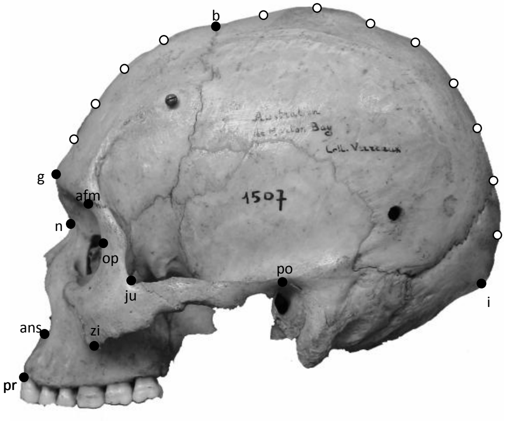


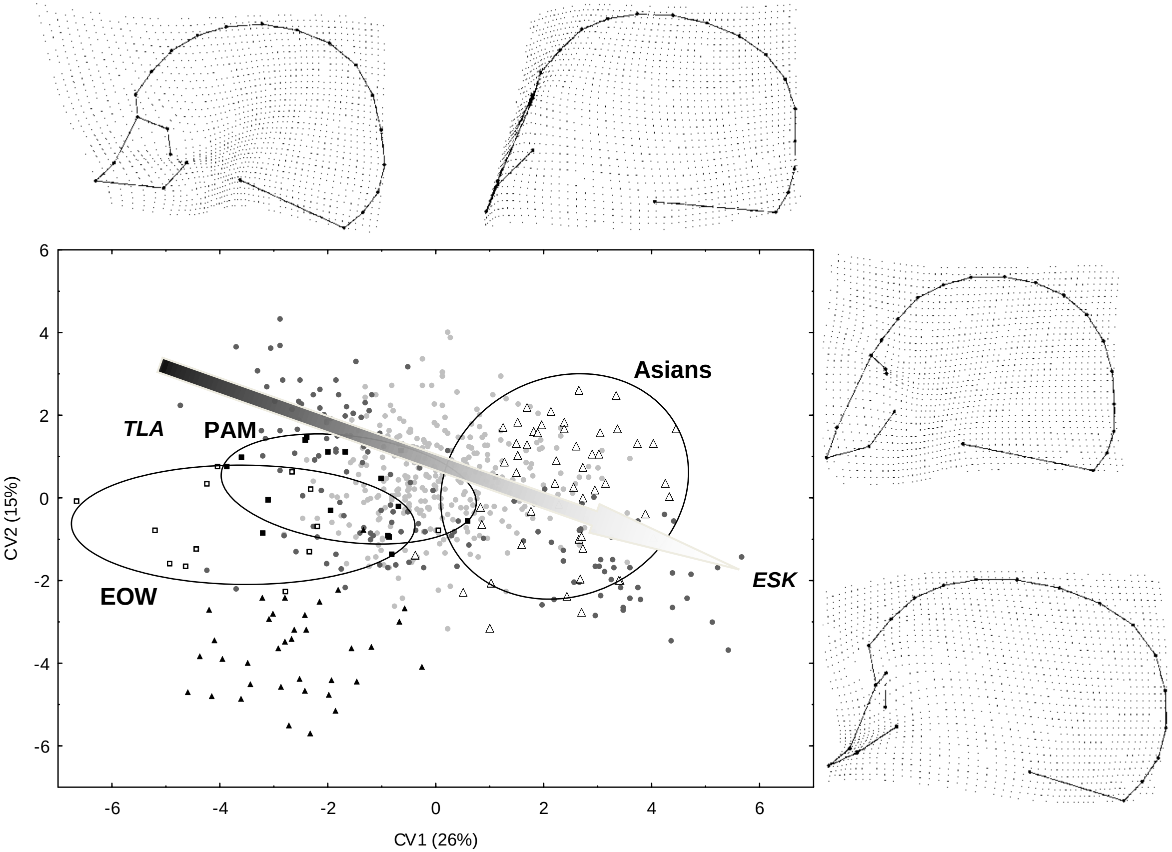
- Double origin:
  - Southeast Asians: Paleoamericans (pre-mongoloids)
  - Northeast Asians: most of modern native Americans (mongoloids)
- Admixture between both components
- Paleoamerican relicts (Tierra del Fuego, Baja California etc)

The Peopling of America: Craniofacial Shape Variation on a Continental Scale and its Interpretation From an Interdisciplinary View

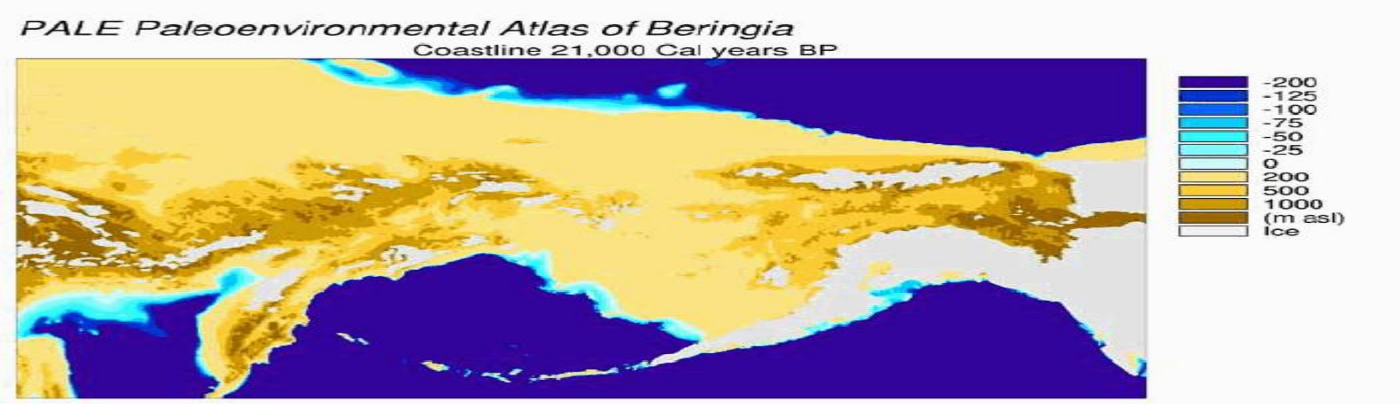
Rolando González-José,[1]* Maria Cátira Bortolini,[2] Fabrício R. Santos,[3] and Sandro L. Bonatto[4]

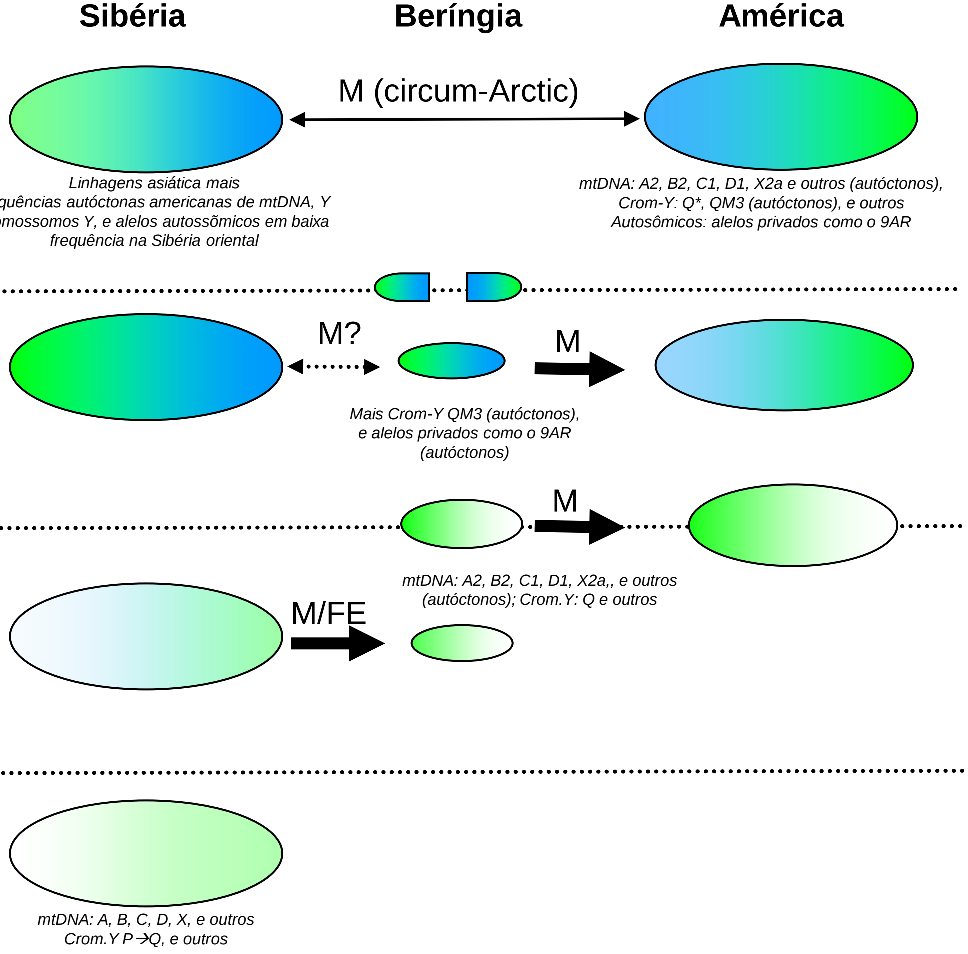
AJPA 2008 Oct;137(2):175-187





Light gray circles: modern South Americans; dark gray circles: modern North Americans; solid triangles: Australians; open triangles: modern Asians; open squares: EOW; solid squares: LS+PAM





Sibéria **Beríngia** **América**

Q* + Q-M3
C3* + C3b

T_3
12-0 KYBP
Formação do Estreito de Bering

M (circum-Arctic)

Linhagens asiática mais Sequências autóctonas americanas de mtDNA, Y cromossomos Y, e alelos autossômicos em baixa frequência na Sibéria oriental

mtDNA: A2, B2, C1, D1, X2a e outros (autóctonos), Crom-Y: Q, QM3 (autóctonos), e outros Autosômicos: alelos privados como o 9AR*

Q* + Q-M3
C3

T_2
18-12 KYBP
Ao final da Glaciação

M?

M

Mais Crom-Y QM3 (autóctonos), e alelos privados como o 9AR (autóctonos)

M

mtDNA: A2, B2, C1, D1, X2a,, e outros (autóctonos); Crom.Y: Q e outros

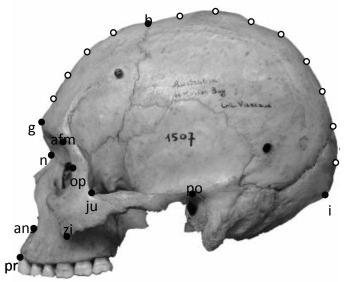
Q + C

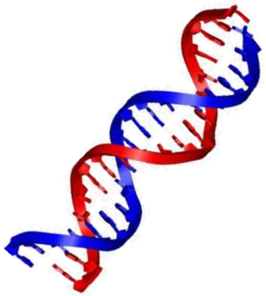
T_1
26-18 KYBP
Formação da Beríngia (LGM)

M/FE

T_0
?-26 KYBP
Começo da Glaciação

mtDNA: A, B, C, D, X, e outros
Crom.Y P→Q, e outros





Peopling of South America

Lathrap 1970
and others

> 12 K YA

Paleoindians

hunter-gatherers:
Macro-Ge, Maku etc

< 6 K YA (*farmers´ expansions*)

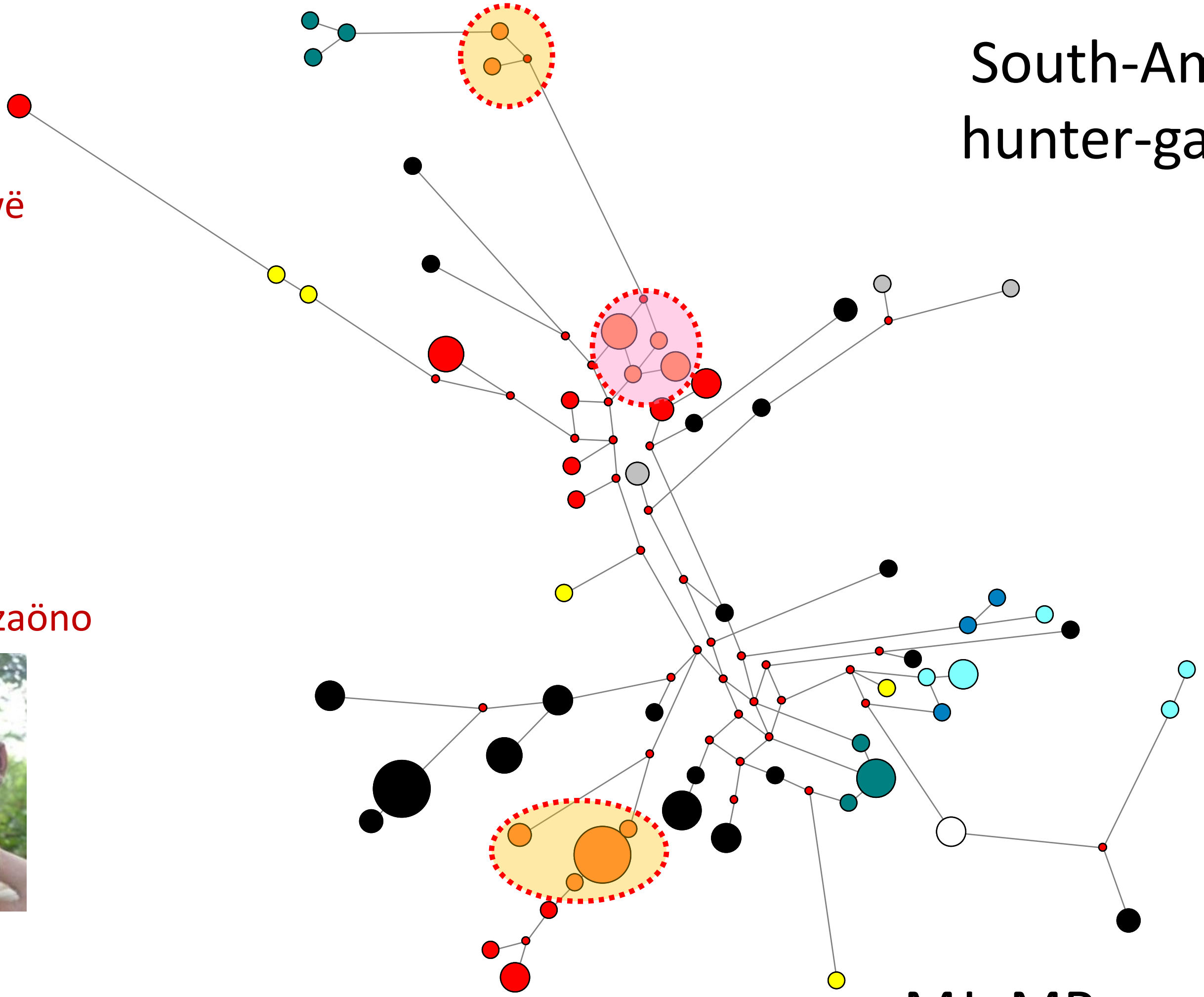
Andeans

Arawaks (Equatorial)

Tupians (Equatorial)


Ice sheets boundary
Land bridge route
Beringia
Ice-Free Corridor
Laurentide Ice Sheet
Cordilleran Ice Sheet
Coastal route
Meadowcroft Rockshelter
Folsom
Clovis
Modern coastline
Atlantic Ocean
Ancient coastline
Pacific Ocean
Monte Verde
Andes Ice Sheet
Migration routes
Ice sheet
Ancient coastline
Modern coastline

South-American hunter-gatherers



MJ+MP

Populations

- ⬜ Nambikwara
- 🟧 Xavante
- 🟨 Kaingang
- 🟥 Kayapo
- ⬜ Hupda
- 🟦 Daw
- ⬛ Chiquitano
- ⬜ Nahua
- 🟩 Maxacali

fst=0.26

🟨 Xavante Owawë

Xavante Puretzaöno

Origin of the Uros

Uros

“peoples of the lake”

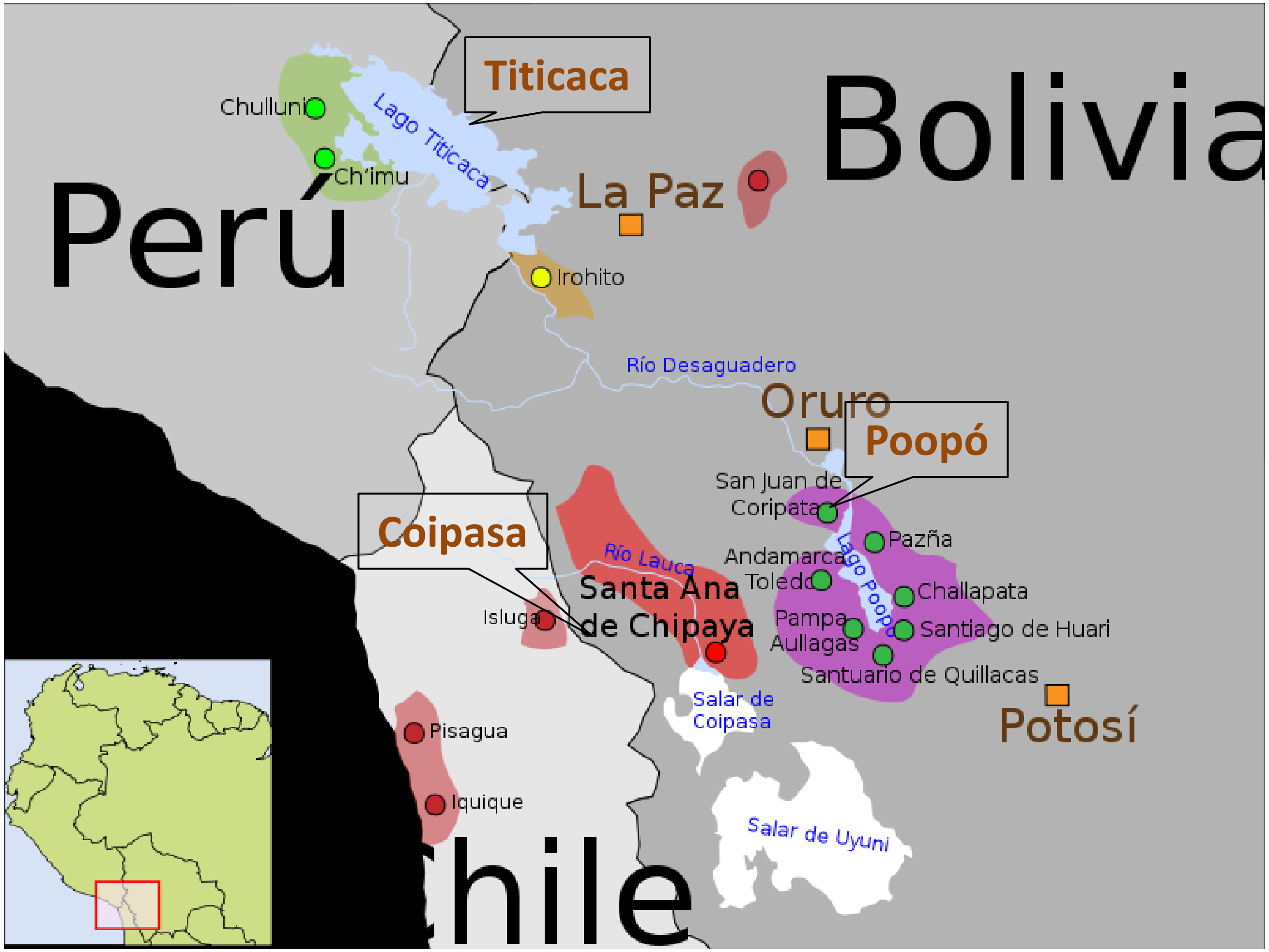
Bolivia

- 🔴 Uru-Chipayas (Coipasa)
- 🟣 Uru-Moratos (Poopó)
- 🟡 Uru-Iruhitos (Desaguadero)

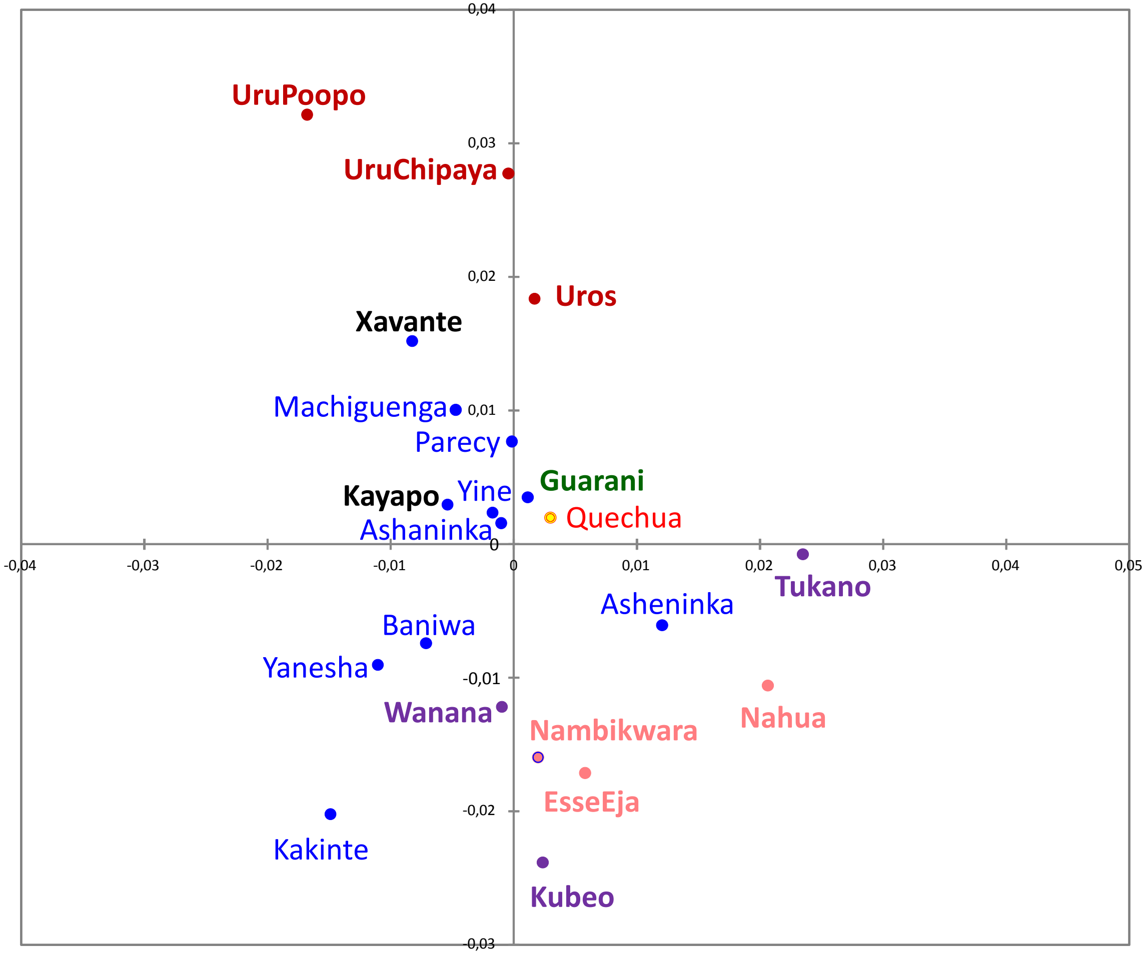
Peru

- 🟢 Uros (Puno)

🟥 Recent migrants

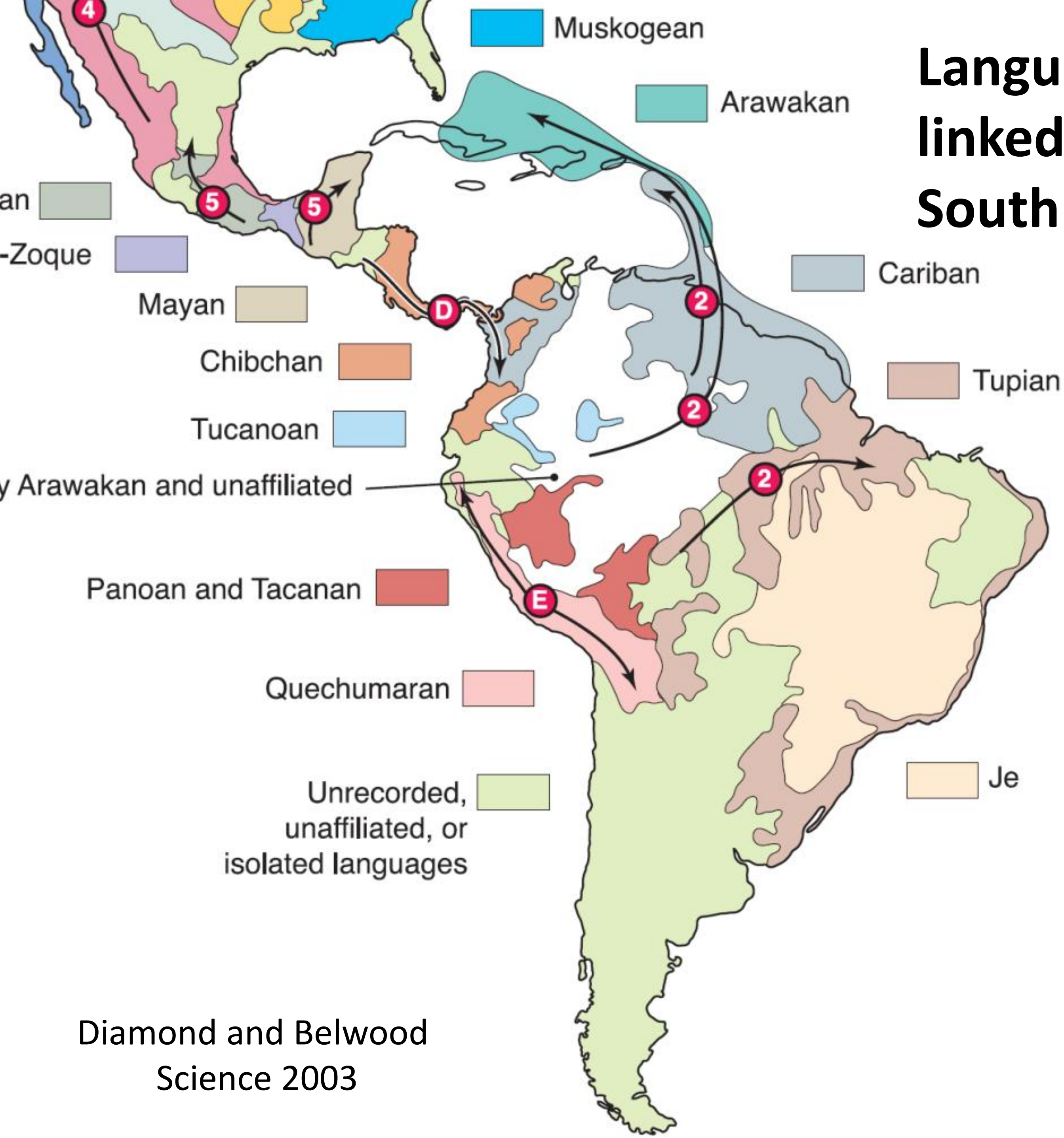


MDS of HVS data ABCD mtDNA haplogroups



Nei genetic distances (GenAlEx + XLSTAT)

Language expansions linked to agriculture in South America



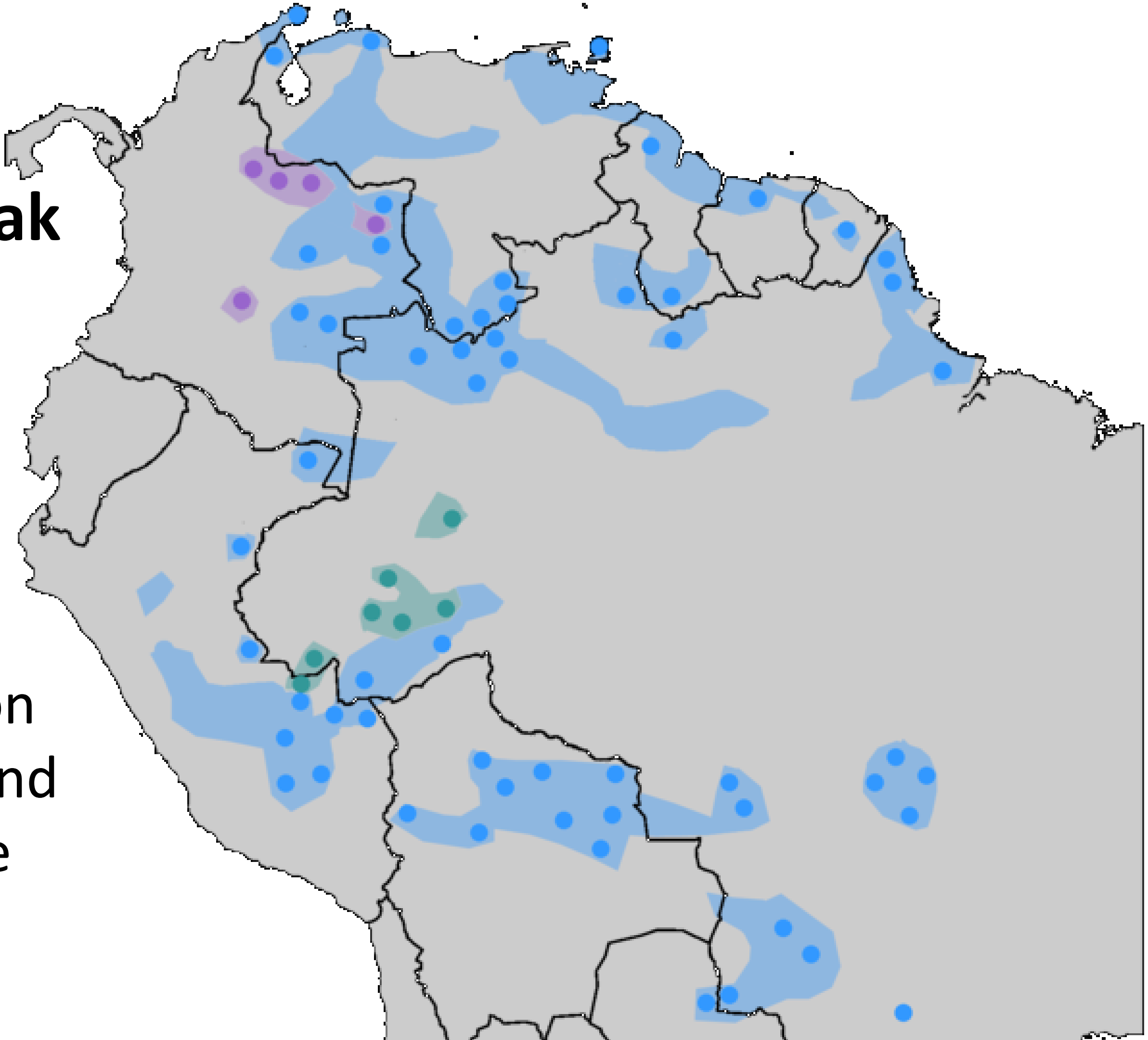
2 – *Arawak, Tupi and Caribe*

E – *Quechua and Aymara*

Diamond and Belwood
Science 2003

Expansion of the Arawak indigenous groups throughout South America

A pre-Columbian connection between Amazon natives and indigenous groups from the Caribbean and the Andes



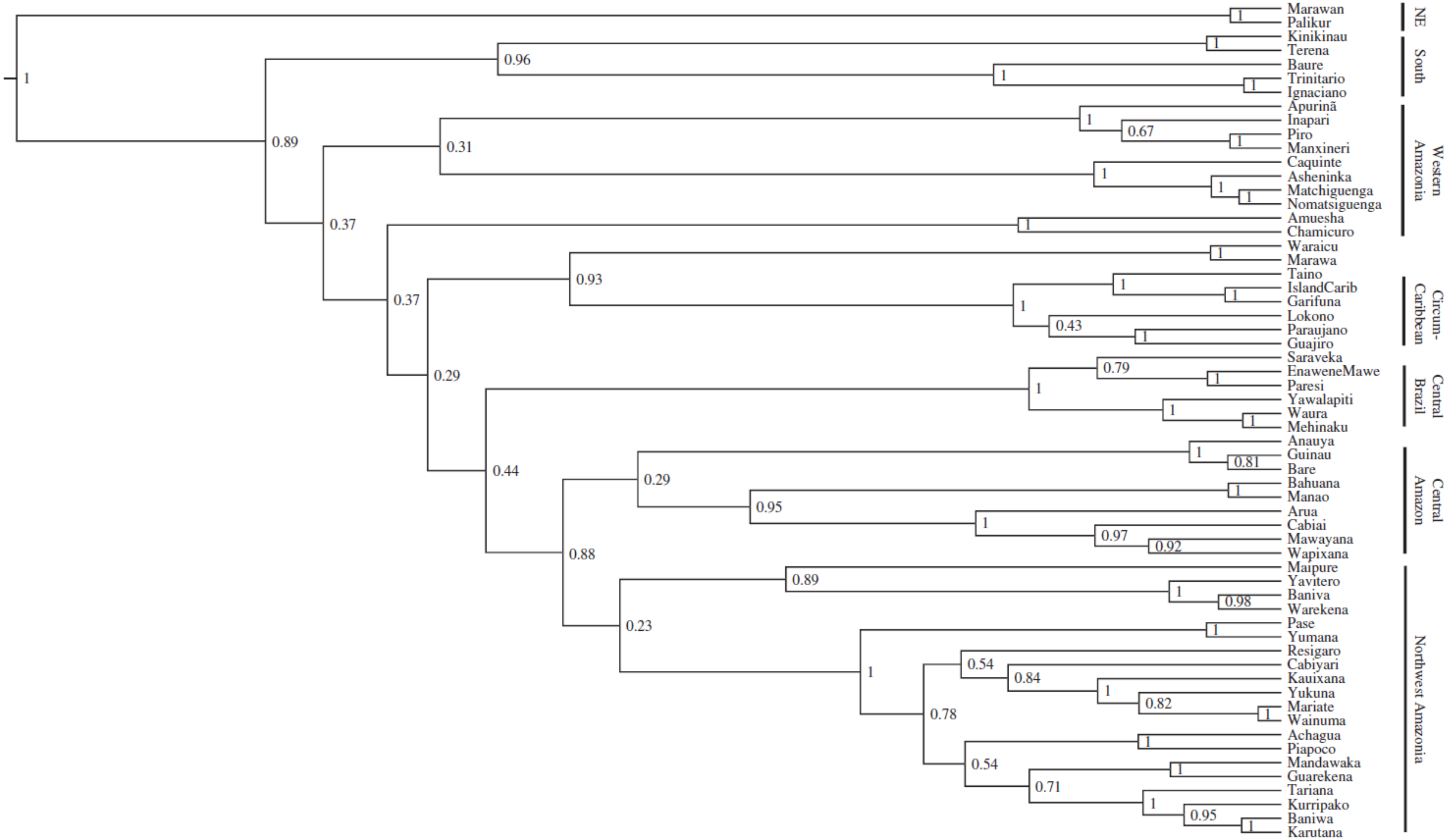
PROCEEDINGS OF THE ROYAL SOCIETY B

FirstCite® e-publishing

Proc. R. Soc. B
doi:10.1098/rspb.2010.2579
Published online

Bayesian phylogeography of the Arawak expansion in lowland South America

Robert S. Walker[1,*] and Lincoln A. Ribeiro[2]



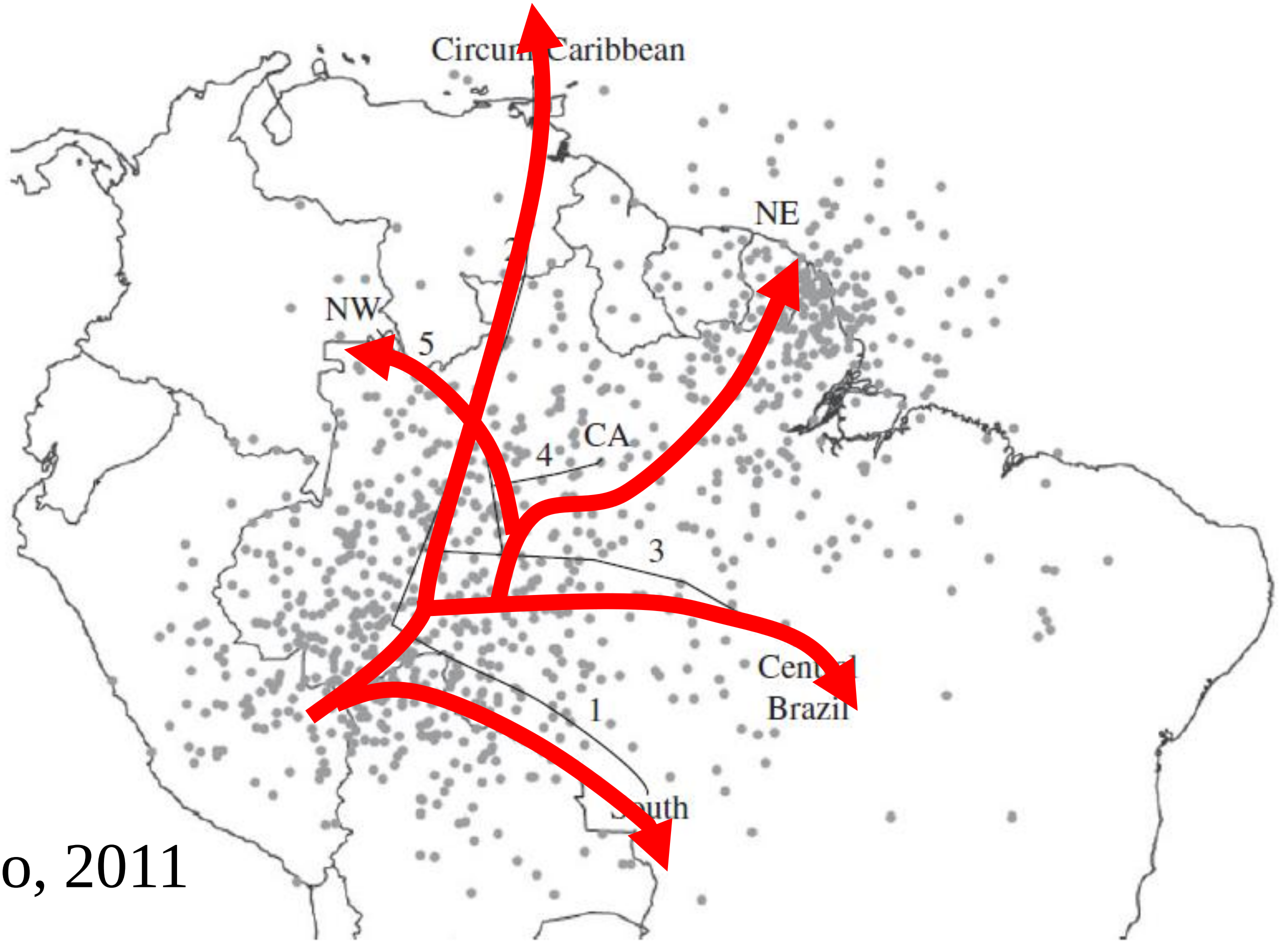
PROCEEDINGS
OF
THE ROYAL
SOCIETY
B

FirstCite®
e-publishing

Proc. R. Soc. B
doi:10.1098/rspb.2010.2579

Bayesian phylogeography of the Arawak expansion in lowland South America

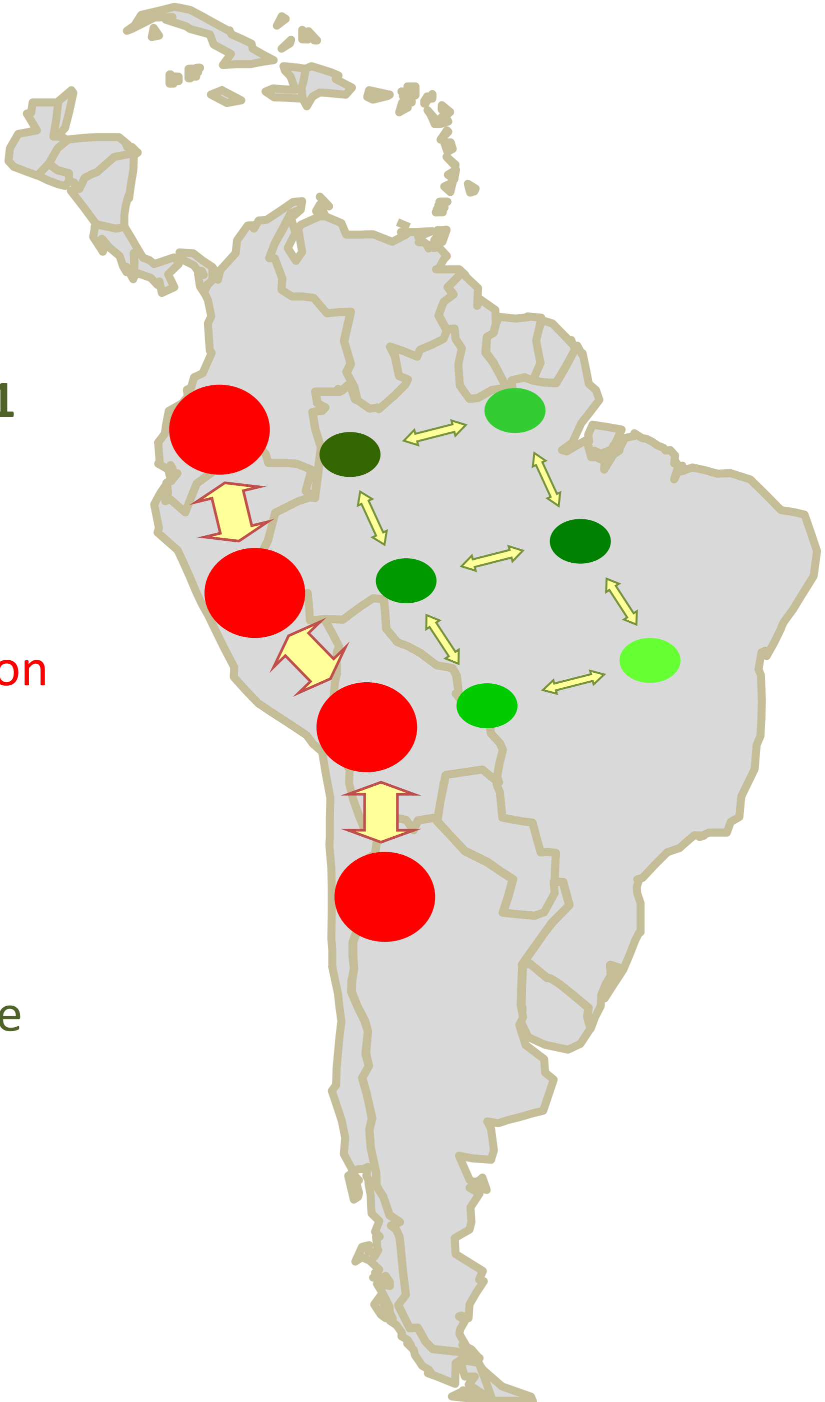
Robert S. Walker[1,*] and Lincoln A. Ribeiro[2]



Walker & Ribeiro, 2011

South American Evolutionary dinamics of populations

Tarazona-Santos et al. Am.J.H.G. 2001



Andes

Large Ne

High gene flow

homogeneization

Eastern

Small Ne

Low gene flow

drift/divergence

2326 participants
234 sampling sites

Peru (1037/88)
Bolivia (973/85)
Ecuador (123/22)
Brazil (193/39)

+ 1000 native individuals from Brazil and Peru

+ T.Schurr - North American Genographic data

+ Salzano´s group data

+ Published data

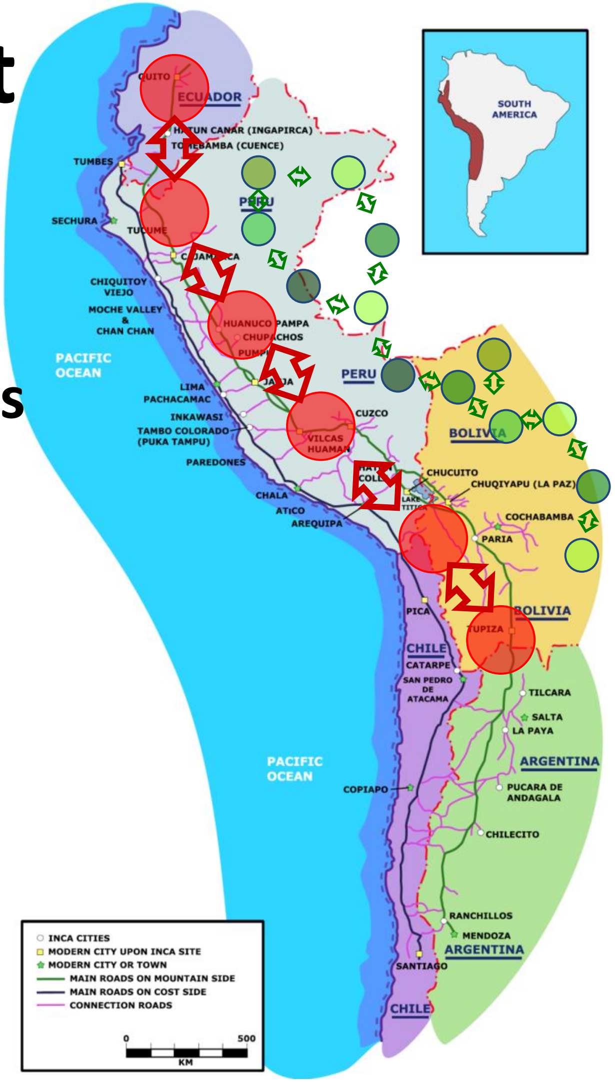
Andes settlement

Agricultural expansion

Andean societies

Qhapaq Ñan

Postcolumbian movements



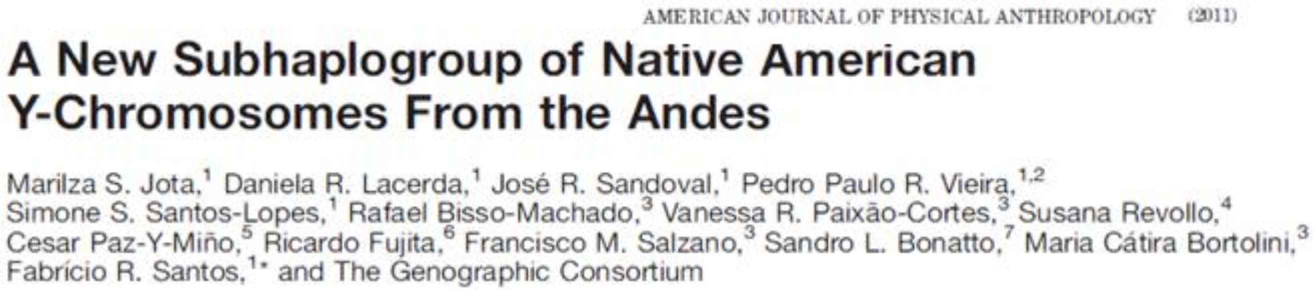
Andes

- Low Fst’s
- Large N_e
- Small drift

Amazon

- High Fst’s
- Small N_e
- Large drift

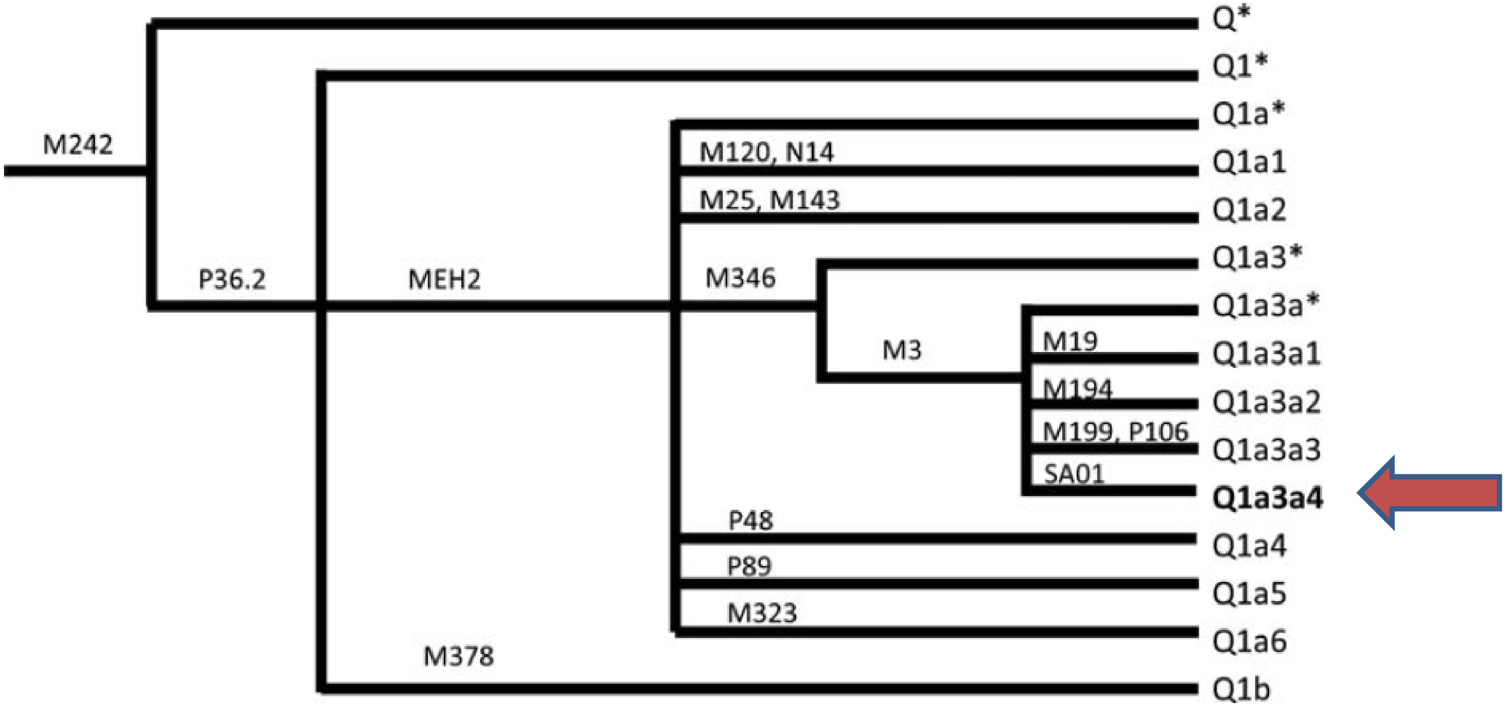
Genographic South-American Center (in preparation)



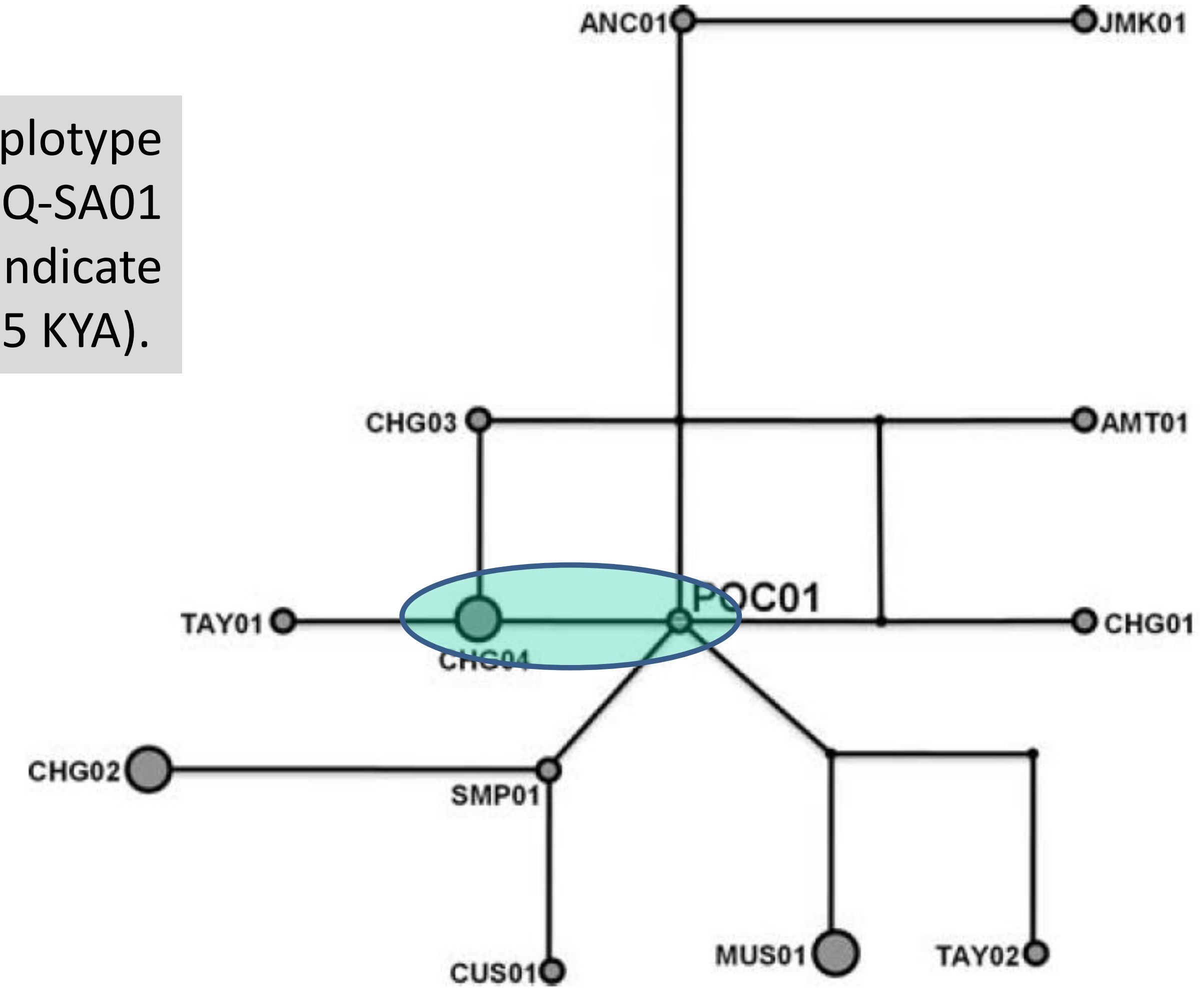
AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY (2011)

A New Subhaplogroup of Native American Y-Chromosomes From the Andes

Marilza S. Jota,[1] Daniela R. Lacerda,[1] José R. Sandoval,[1] Pedro Paulo R. Vieira,[1,2] Simone S. Santos-Lopes,[1] Rafael Bisso-Machado,[3] Vanessa R. Paixão-Cortes,[3] Susana Revollo,[4] Cesar Paz-Y-Miño,[5] Ricardo Fujita,[6] Francisco M. Salzano,[3] Sandro L. Bonatto,[7] Maria Cátira Bortolini,[3] Fabrício R. Santos,[1*] and The Genographic Consortium



15 Y STR haplotype network of the Q-SA01 lineage (N=16) indicate a recent origin (~5 KYA).



Origin of the SA01 SNP ~5 KYA

The SA01 Y-STR network indicates a southwards dispersion, which may have been spread together farming (or Wary/Quechuan culture) in the Andes.

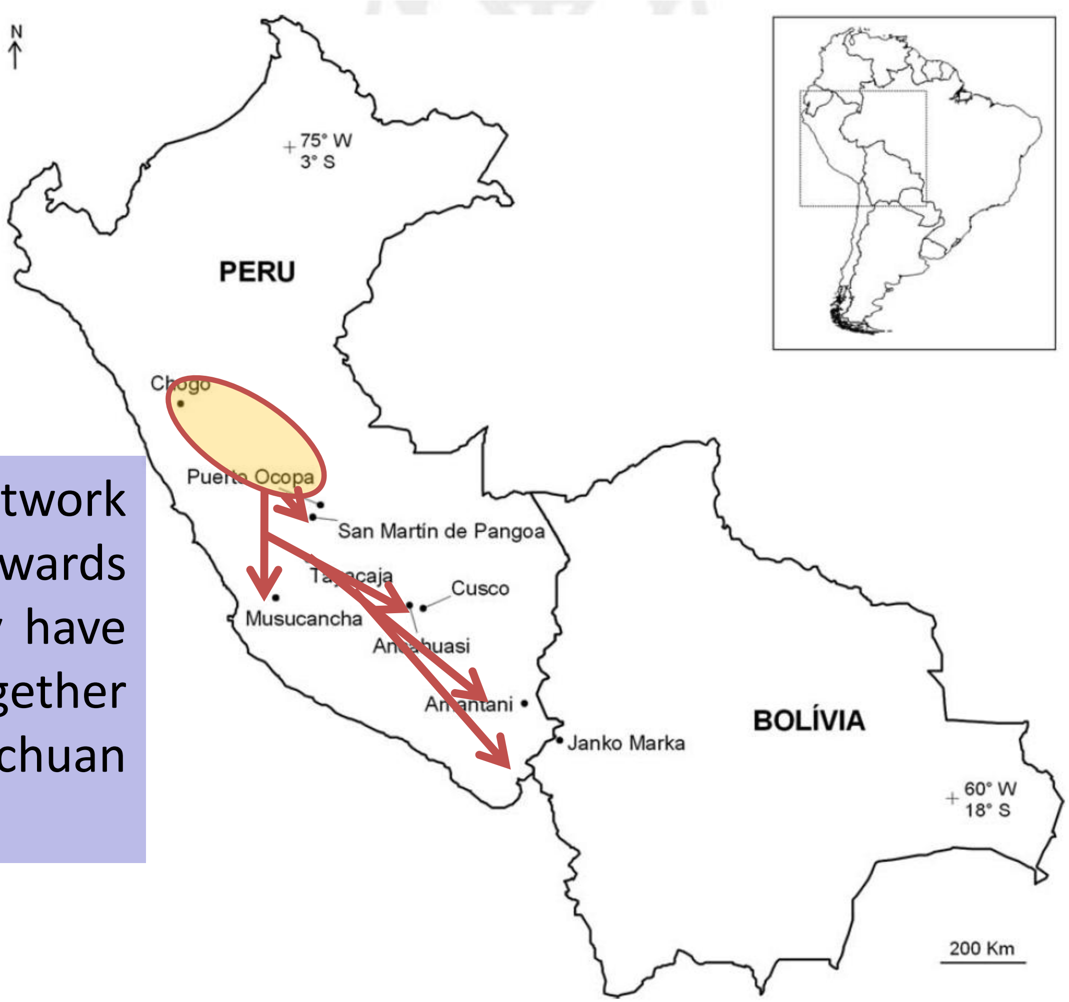


Fig. 2. Sample locations for individuals bearing the derived allele at marker SA01.

A LANDMARK STUDY OF THE HUMAN JOURNEY

THE GENOGRAPHIC PROJECT

I Reunión Genográfico-Sudamérica - Cusco – Abril de 2009

Tariano
Baniwa
Aymara
Descendiente Arawak?

Kayapó

Apoyo

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INCIPIT VITA NOVA
7 DE SETEMBRO DE 1927

VERITAS LIBERABIT VOS

UNIVERSITAS MAJOR PACENSIS DIVI ANDREAE

UFRJ
UNIVERSIDADE DO BRASIL

udla

25 ANOS
FAPEMIG

CNPq
Conselho Nacional de Desenvolvimento
Científico e Tecnológico

CAPES

NATIONAL GEOGRAPHIC™

IBM.

A research partnership of National Geographic and IBM.

THE WAITT FAMILY FOUNDATION

Global field science supported by the Waitt Family Foundation.

thanks

fsantos@icb.ufmg.br