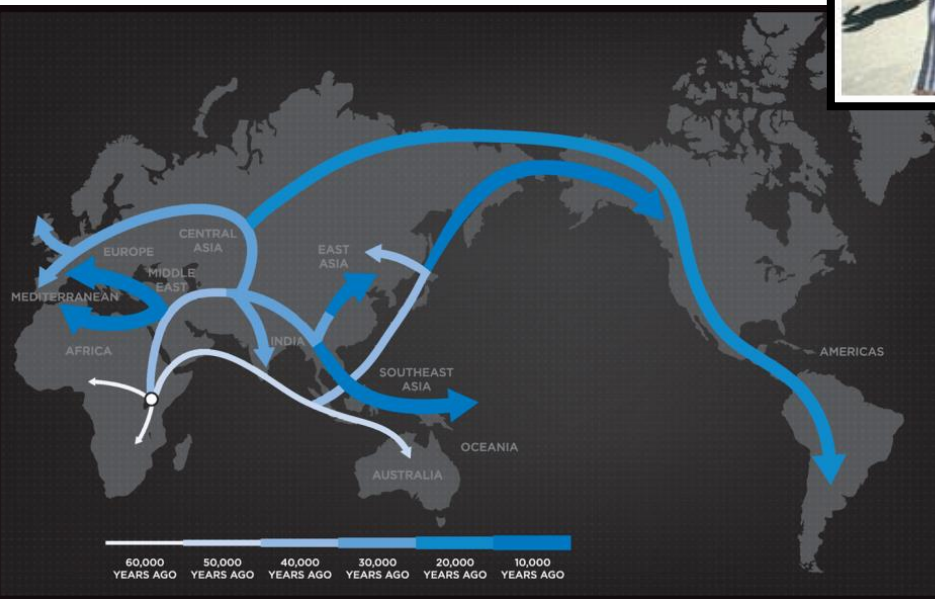


# Tracing Human Origins, Migration, and Settlement With Modern and Ancient DNA

THE  
GENOGRAPHIC  
PROJECT

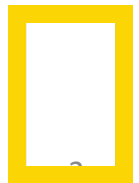
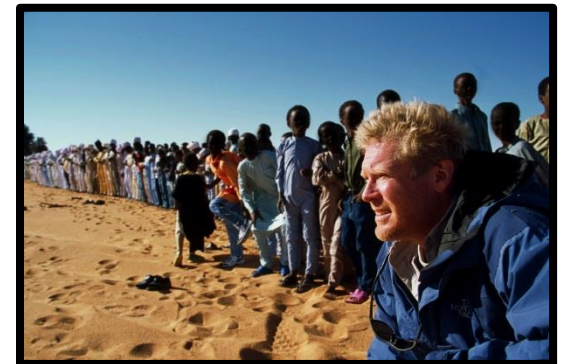
**Dr. Miguel G. Vilar**  
**Lead Scientist, Genographic**  
**Feb 22, 2018**



  
**NATIONAL  
GEOGRAPHIC**

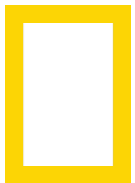
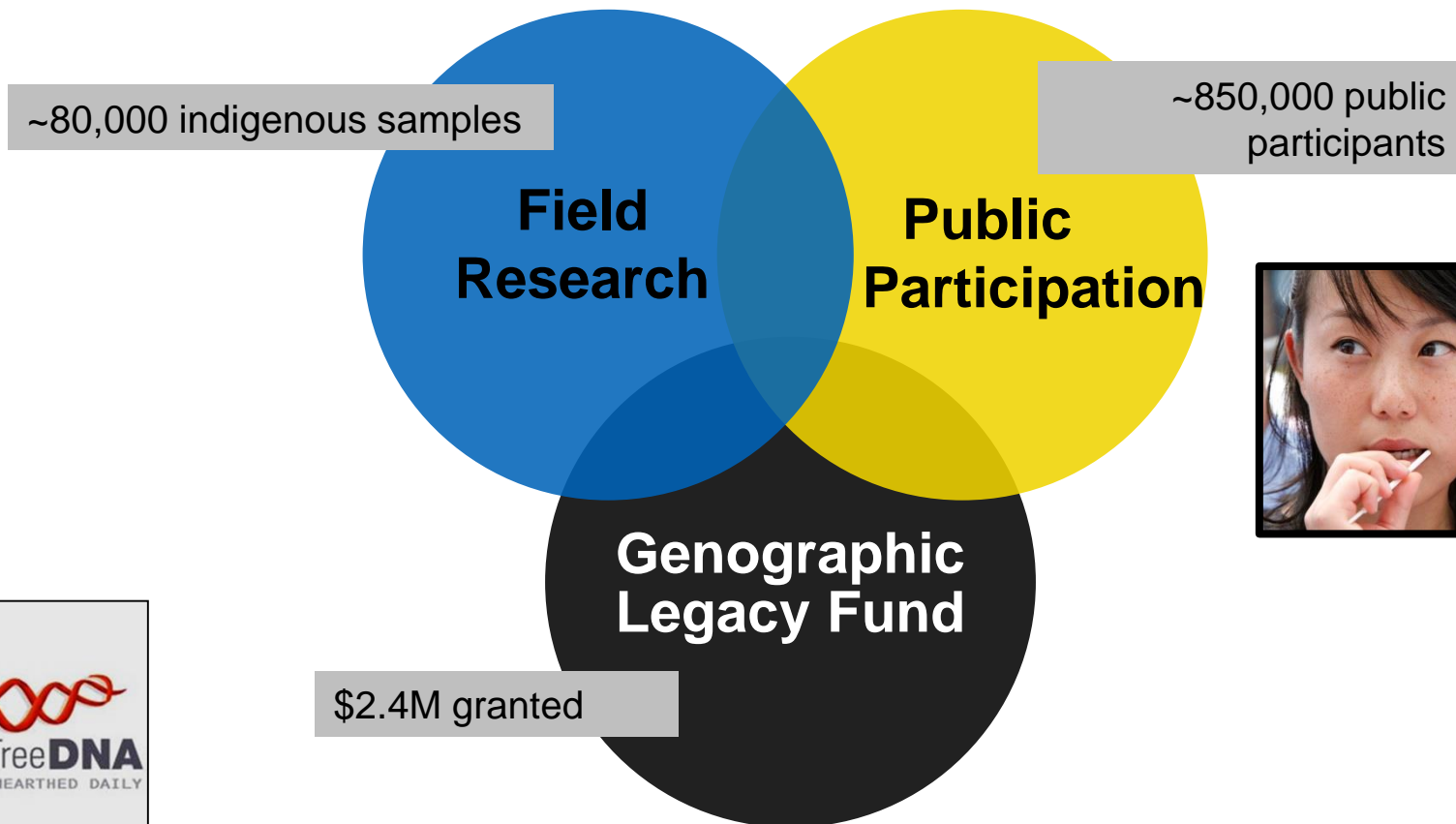
# What is the Genographic Project?

- An on going global study that aims to map human origins, settlement and migration by analyzing DNA samples
- One of the first and largest Citizen Science projects in the country
- Launched in 2005
  - Short DNA sequences
- Redone in 2012 as Geno 2.0
  - DNA Chips
    - Biogeographic Regions
    - Hominin Percentages
- Redone 8/15 & 12/16 as Geno 2.0 Next Generation
  - Illumina Omni Express
  - Helix (Next Generation Sequencing)



# What is The Genographic Project?

**Twelve years of scientific success**



# Genographic Funds: Cultural and linguistic preservation

- Awarded over 100 grants totaling over \$2.4 million, across 5 continents



# Education & Outreach Programs



- More than 5,000 students, across hundreds of schools
- Topics: History, biology, geography, languages



Spain  
May 2013



New Zealand Feb 2014



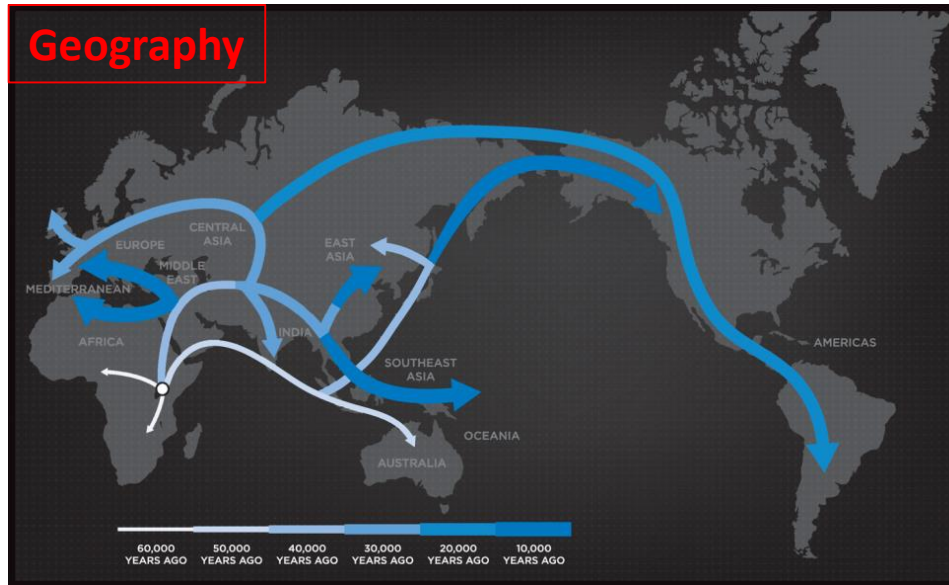
Ireland Jul 2013

# Genographic: The Science

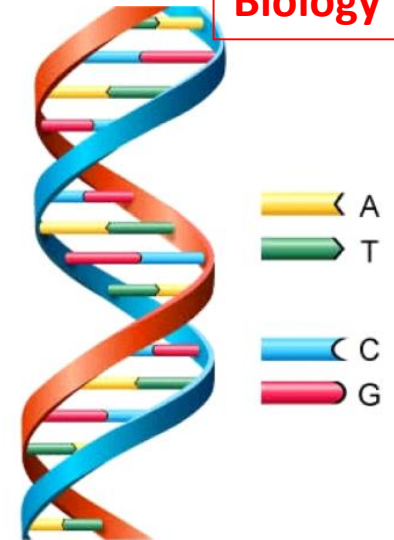
## Anthropology



## Geography



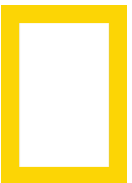
## Biology



What is your own history, how can you learn about it?

“The greatest history book ever written is the one hidden in our DNA.”

## History



# Genographic Scientific Research Centers (2005- 2015)



82,000 indigenous samples  
12 research centers, 6 continents  
Published ~76 academic papers

# Genographic Scientific Grants (2012-2016)

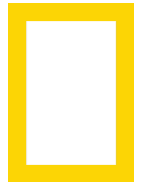
Peopling of the  
Caribbean

Origin and spread  
of Indo-European

Tracing migration  
routes along the  
Mediterranean

Recovering Australia's  
lost history

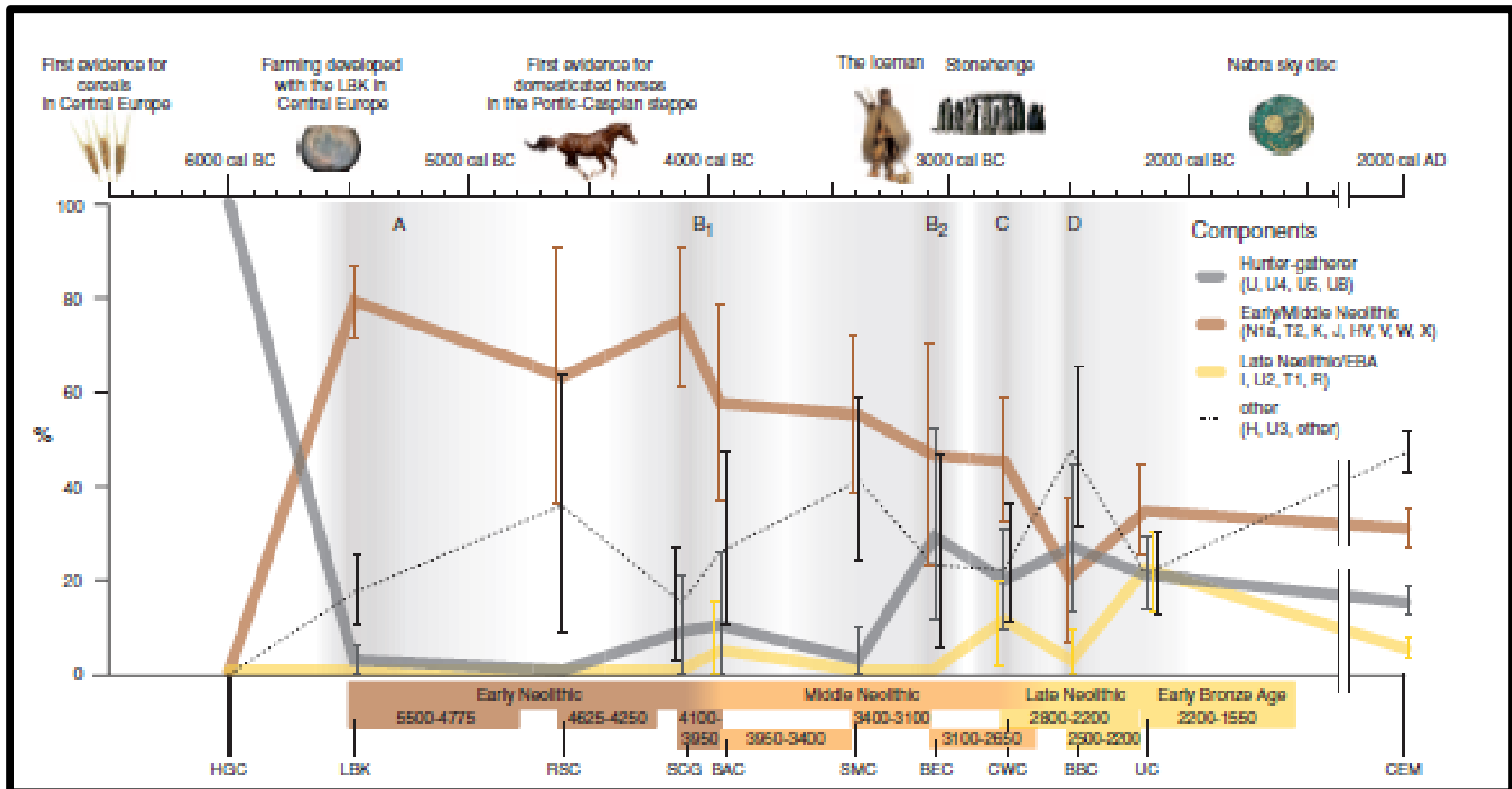
**14 new grantees: >\$725,000**





# Genographic in Academia

- +70 Academic Manuscripts (Science, Nature, PNAS, PLoS Genetics, ...)
- >100 professional conference presentations



# Genographic Project Research: Sample Publications

PLOS ONE

RESEARCH ARTICLE

## Genetic Heritage of the Balto-Slavic Speaking Populations: A Synthesis of Autosomal, Mitochondrial and Y-Chromosomal Data

Alena Kushniarevich<sup>1,2\*</sup>, Olga Utevska<sup>3,4\*</sup>, Marina Chuhryaeva<sup>4,5</sup>, Anastasia Agdzhoyan<sup>4,5</sup>, Khadzizat Dibirova<sup>4,5</sup>, Ingrida Uktveryte<sup>6</sup>, Märt Möls<sup>7</sup>, Lejla Mulahasanovic<sup>8,9</sup>, Andrey Pshenichnov<sup>5</sup>, Svetlana Frolova<sup>4</sup>, Andrey Shanko<sup>5</sup>, Ene Metspalu<sup>1,10</sup>, Maere Reidla<sup>1,10</sup>, Kristiina Tambets<sup>1</sup>, Erika Tamm<sup>1,10</sup>, Sergey Koshelev<sup>11</sup>, Valery Zaporozhchenko<sup>1,5</sup>, Lubov Atramentova<sup>9</sup>, Vaidutis Kučinskas<sup>5</sup>, Oleg Davydenko<sup>2</sup>, Olga Goncharova<sup>15</sup>, Irina Evseeva<sup>6,14</sup>, Michail Churnosov<sup>12</sup>, Elvira Pocheshchova<sup>13</sup>, Bayazit Yunusbayev<sup>1,16</sup>, Elza Khusnutdinova<sup>16,17</sup>, Damir Marjanović<sup>18,19</sup>, Pavao Rudan<sup>1</sup>, Siiri Roots<sup>1</sup>, Nick Yankovsky<sup>1</sup>, Phillip Endicott<sup>20</sup>, Alexei Kassian<sup>1,21,22</sup>, Anna Dybo<sup>21</sup>, The Genographic Consortium<sup>1</sup>, Chris Tyler-Smith<sup>23</sup>, Elena Balanovska<sup>4</sup>, Mait Metspalu<sup>1</sup>, Toomas Kivisild<sup>1,10,24</sup>, Richard Villems<sup>1,10,25\*</sup>, Oleg Balanovsky<sup>4,5\*</sup>

AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 00:00-00 (2015)

## Antiquity and Diversity of Aboriginal Australian Y-Chromosomes

Nano Nagle,<sup>1</sup> Kaye N. Ballantyne,<sup>2,3</sup> Mannis van Oven,<sup>3</sup> Chris Tyler-Smith,<sup>4</sup> Yali Xue,<sup>4</sup> Duncan Taylor,<sup>5,6</sup> Stephen Wilcox,<sup>7</sup> Leah Wilcox,<sup>1</sup> Rust Turkalov,<sup>7</sup> Roland A.H. van Oorschot,<sup>2</sup> Peter McAllister,<sup>8</sup> Lesley Williams,<sup>9</sup> Manfred Kayser,<sup>3</sup> Robert J. Mitchell,<sup>1\*</sup> and The Genographic Consortium

European Journal of Human Genetics (2015), 1-7  
© 2015 Macmillan Publishers Limited. All rights reserved 1018-4813/15  
www.nature.com/ejhg

ARTICLE

## Origins, admixture and founder lineages in European Roma

Begoña Martínez-Cruz<sup>1,15,16</sup>, Isabel Mendizabal<sup>1,15,17</sup>, Christine Harmant<sup>2,3</sup>, Rosario de Pablo<sup>4</sup>, Mihai Ioana<sup>5,6</sup>, Dora Angelicheva<sup>7</sup>, Anastasia Kouvatzi<sup>8</sup>, Halyna Makukh<sup>9</sup>, Mihai G Netea<sup>10</sup>, Horolma Pamjaj<sup>11</sup>, Andrea Zalán<sup>11</sup>, Ivailo Tournev<sup>12,13</sup>, Elena Marushiakova<sup>14</sup>, Vesselin Popov<sup>14</sup>, Jaume Bertranpetit<sup>1</sup>, Luba Kalaydjieva<sup>7</sup>, Lluís Quintana-Murci<sup>2,3</sup>, David Comas<sup>4,1</sup> and the Genographic Consortium<sup>18</sup>

The Roma, also known as 'Gypsies', represent the largest and the most widespread ethnic minority of Europe. There is increasing evidence, based on linguistic, anthropological and genetic data, to suggest that they originated from the Indian subcontinent, with subsequent bottlenecks and undetermined gene flow from/to hosting populations during their diaspora. Further support comes from the presence of Indian uniparentally inherited lineages, such as mitochondrial DNA M and Y-chromosome H haplogroups, in a significant number of Roma individuals. However, the limited resolution of most genetic studies so far, together with the restriction of the samples used, have prevented the detection of other non-Indian founder lineages that might have been present in the proto-Roma population. We performed a high-resolution study of the uniparental genomes of 753 Roma and 984 non-Roma hosting European populations, with non-Roma samples as a reference, during their diaspora. We found a novel uniparental maternal (H7, J1b3, J1c1, M1) haplogroup, whose classification allows us to identify a founder lineage not negligible, is substantially different from the other H haplogroups consistently points to a North-Western European population.

PLOS ONE

RESEARCH ARTICLE

## Genetic Diversity in the Lesser Antilles and Its Implications for the Settlement of the Caribbean Basin

Jada Benn Torres<sup>1\*</sup>, Miguel G. Vilar<sup>2,3</sup>, Gabriel A. Torres<sup>1</sup>, Jill B. Galeski<sup>2</sup>, Ricardo Bharath Hernandez<sup>4</sup>, Zolia E. Browne<sup>5</sup>, Marlon Stevenson<sup>6</sup>, Wendell Walters<sup>5,6</sup>, Theodore G. Schurr<sup>2</sup>, The Genographic Consortium<sup>1</sup>

1 Department of Anthropology, University of Notre Dame, Notre Dame, Indiana, United States of America, 2 Department of Anthropology, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, 3 Missions Programs, National Geographic Society, Washington, D.C., United States of America, 4 Santa Rosa First Peoples Community, Arima, Trinidad and Tobago, 5 The Grenada Heritage Foundation Inc., Kingstown, St. Vincent and the Grenadines, 6 Sandy Bay Village, St. Vincent and the Grenadines

\* Membership of the Genographic Consortium is provided in the Acknowledgments.  
\* [jbentor@nd.edu](mailto:jbentor@nd.edu)

OPEN ACCESS

Citation: Benn Torres J, Vilar MG, Torres GA, Galeski JB, Bharath Hernandez R, Browne ZE, et al. (2015) Genetic Diversity in the Lesser Antilles and Its Implications for the Settlement of the Caribbean Basin. PLOS ONE 10(10): e0139192. doi:10.1371/journal.pone.0139192

Editor: Francesc Calafell, Universitat Pompeu Fabra, SPAIN

Received: July 13, 2014

Accepted: September 10, 2015

Published: October 8, 2015

Copyright: © 2015 Benn Torres et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

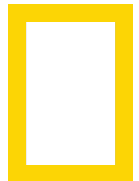
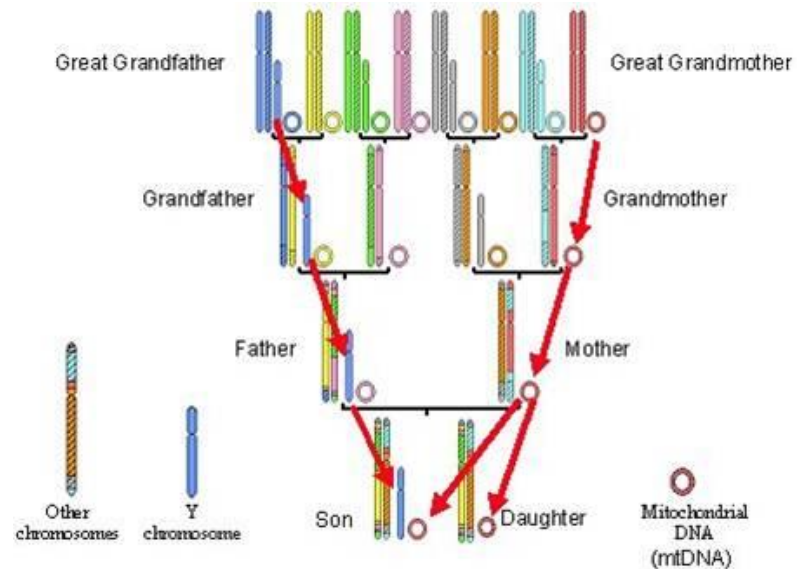
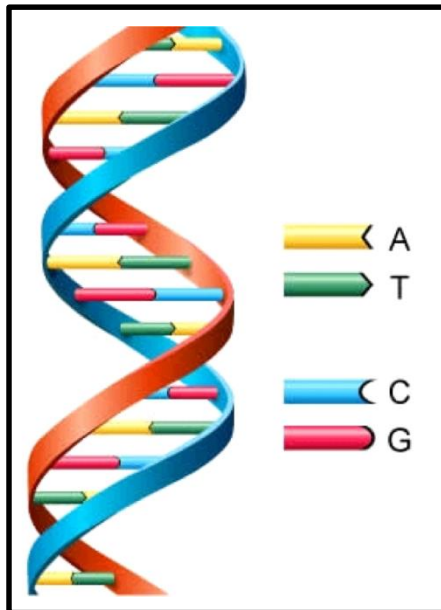
Data Availability Statement: Data are available from the NCBI GenBank with accession numbers K777741-56. A complete list of accession numbers can be found in the Supporting Information files.

Abstract

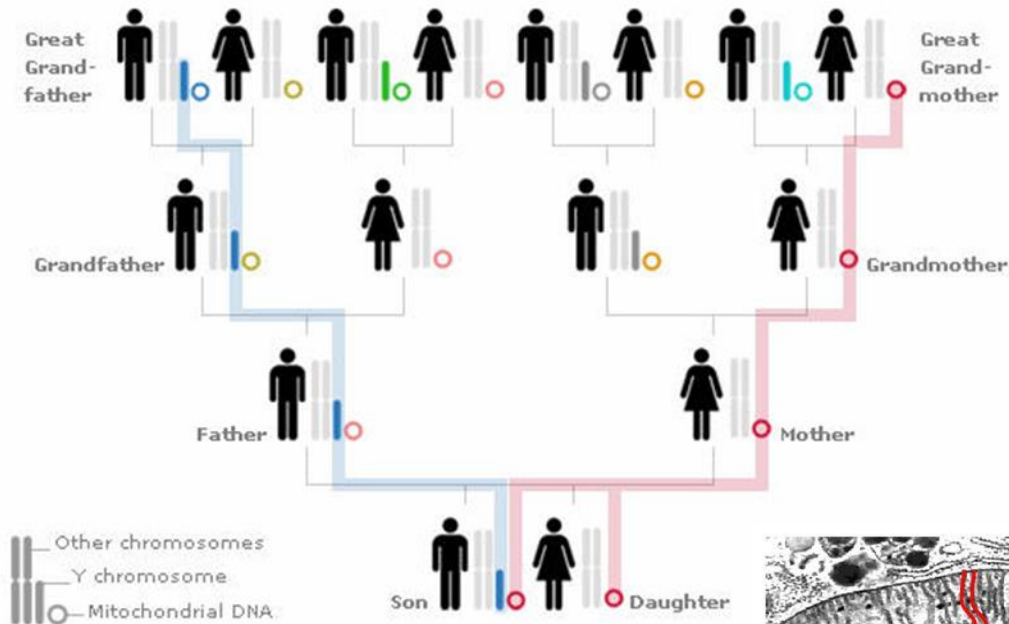
Historical discourses about the Caribbean often chronicle West African and European influence to the general neglect of indigenous people's contributions to the contemporary region. Consequently, demographic histories of Caribbean people prior to and after European contact are not well understood. Although archeological evidence suggests that the Lesser Antilles were populated in a series of northward and eastern migratory waves, many questions remain regarding the relationship of the Caribbean migrants to other indigenous people of South and Central America and changes to the demography of indigenous communities post-European contact. To explore these issues, we analyzed mitochondrial DNA and Y-chromosome diversity in 12 unrelated individuals from the First Peoples Community in Arima, Trinidad, and 43 unrelated Garifuna individuals residing in St. Vincent. In this community-sanctioned research, we detected maternal indigenous ancestry in 42% of the participants, with the remainder having haplotypes indicative of African and South Asian maternal ancestry. Analysis of Y-chromosome variation revealed paternal indigenous American ancestry indicated by the presence of haplogroup Q-M3 in 28% of the male participants from both communities, with the remainder possessing either African or European haplogroups. This finding is the first report of indigenous American paternal ancestry among indigenous populations in this region of the Caribbean. Overall, this study illustrates the role of the region's first peoples in shaping the genetic diversity seen in contemporary Caribbean populations.

# What is the Science?

- Genome: Totality of genetic information of a species
  - Three billion base pairs or DNA markers (ie. instruction book)
    - A book with 23 chapters we call chromosomes
    - 22 of them are autosomes, two are sex chromosomes (X and Y)
  - Each base pair (bp) of DNA is represented by a letter (A, C, G, T)
  - Autosomal DNA is Inherited bi-parentally (>99%), re-shuffled each generation
- What makes each person unique are the mutations, or variations in the DNA

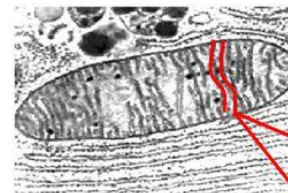
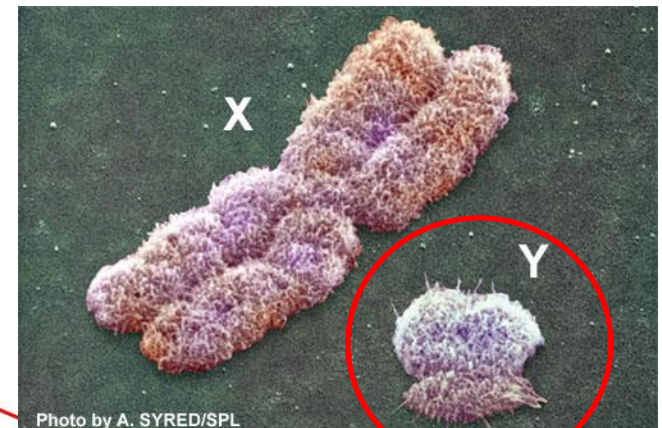


# Uni-parental genome (>1%)



## Y Chromosome DNA

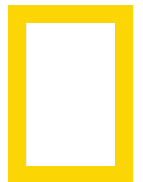
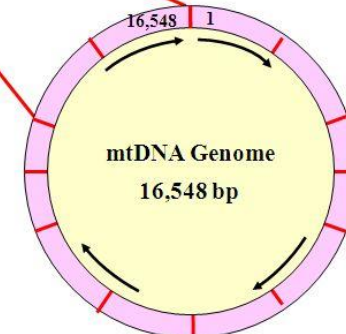
- Passed on from father to son
- Does not mix (mostly)



Mitochondrion

## Mitochondrial DNA

- Inherited strictly from mother
- Very common, doesn't mix
- Highly mutable

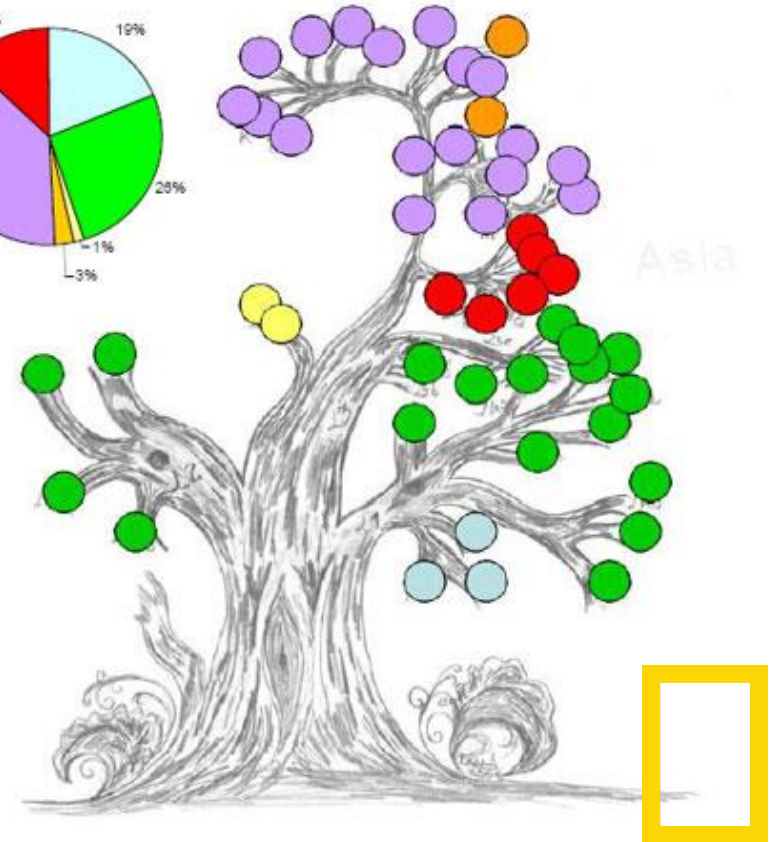
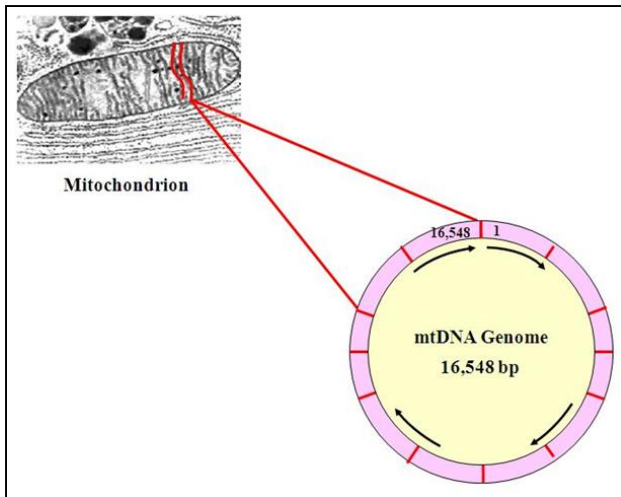
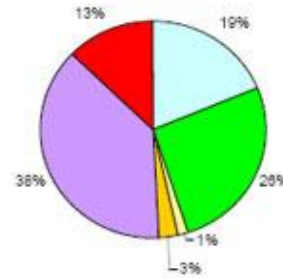


# Mitochondrial DNA (mtDNA)—maternally-inherited

- Many mutations (changes or variations in the DNA)
- mtDNA is a small, numerous molecule

-Hypervariable region (~1,000 bp) has no genes, but include many mutations that define specific branches of humanity, or haplogroups

-Lineages that share mutations are part of the same haplogroup, or branch of the tree.

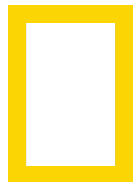
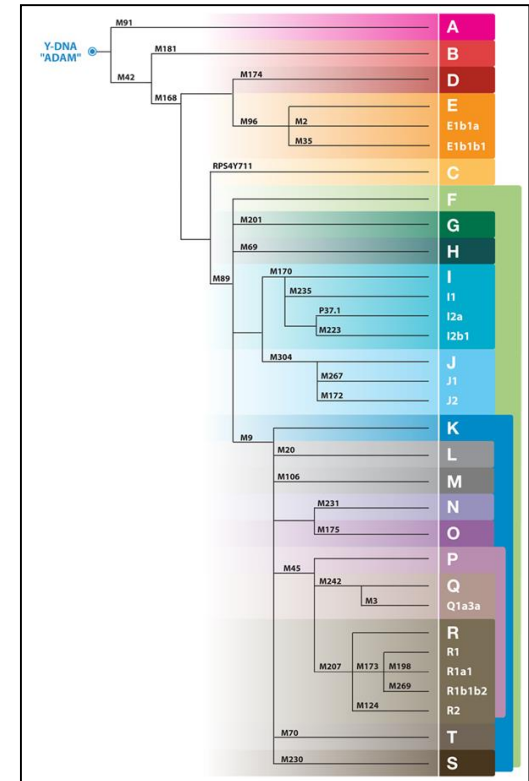
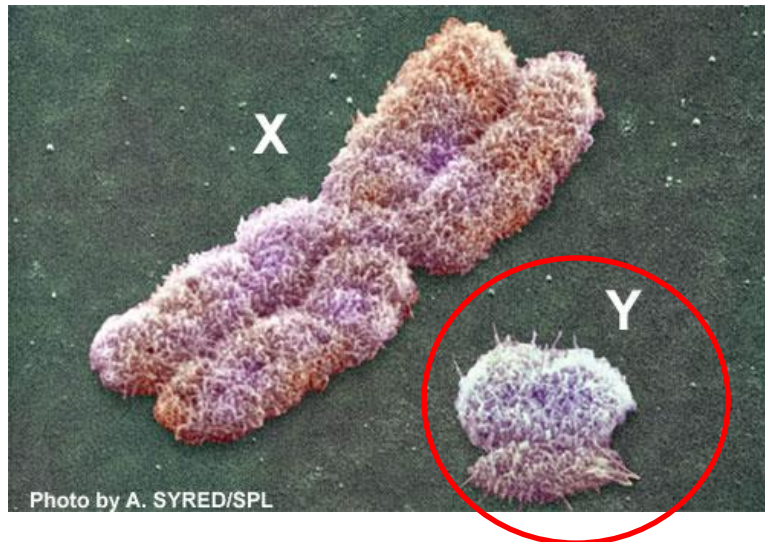


# Y Chromosome DNA (Y-DNA) – Paternally inherited

- Many and various types of mutations
- Smallest nuclear chromosome (59 Million basepairs)

-It holds very few genes, most of the genes that make men and boys male

-As in mtDNA, lineages that share mutations are part of the same haplogroup, or branch of the tree.



# mtDNA Haplogroups

Haplogroup	Subgroup	Mutations
A		16223, 16290, 16319
	A2	16111, 16223, 16290, 16319, 16362
B		16189, 16217
C		16223, 16298, 16327
	C1	16223, 16298, 16325, 16327
D		16223, 16362
	D1	16223, 16325, 16362

-A, B, C, D, X2a

-H-K, N, T-X2b

-L0, L1, L2, and L3

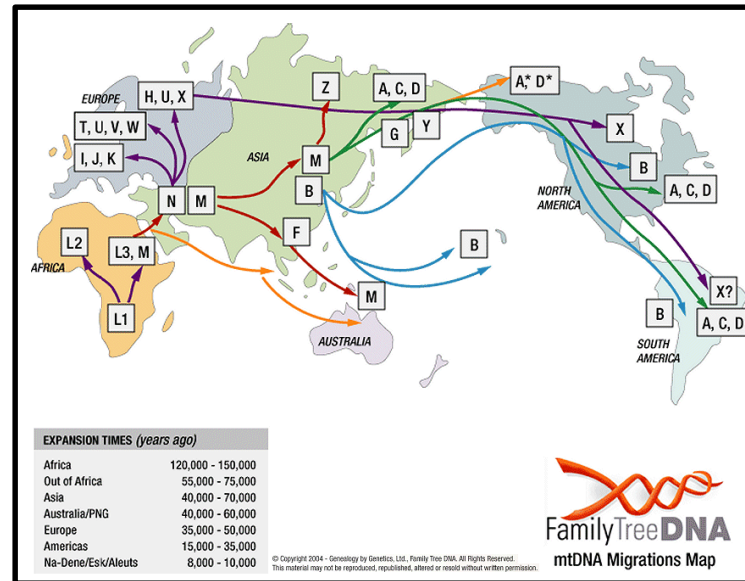
-A-G, M-S, Y-Z

Native American

European

African

Asian and Australian



# Y Chromosome Haplogroups

-Q y C3

-E, G, I, J, R, T

-A, B, E

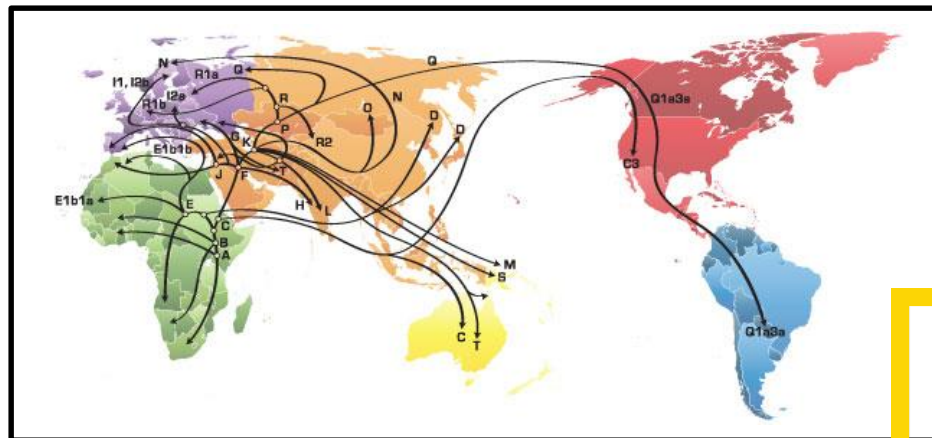
-C, D, H, J, L, Q, S

Native American

European

African

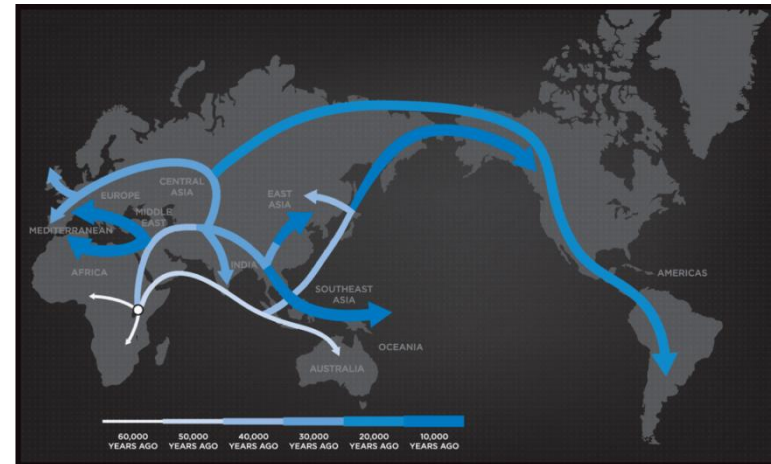
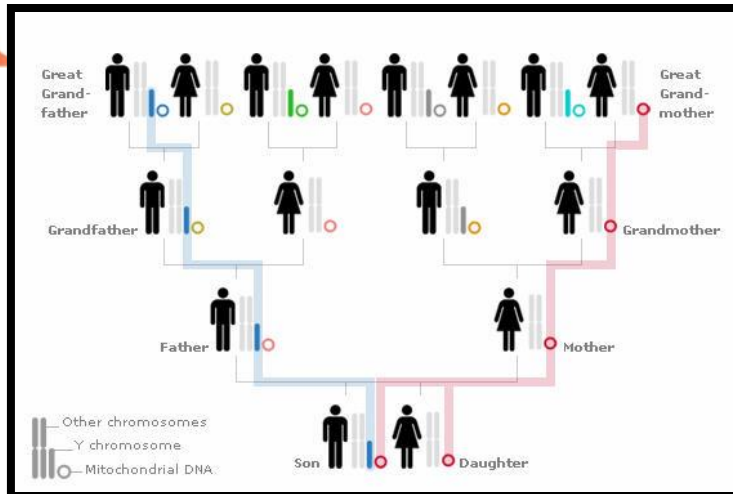
Asian and Australian



# How does DNA become a map?

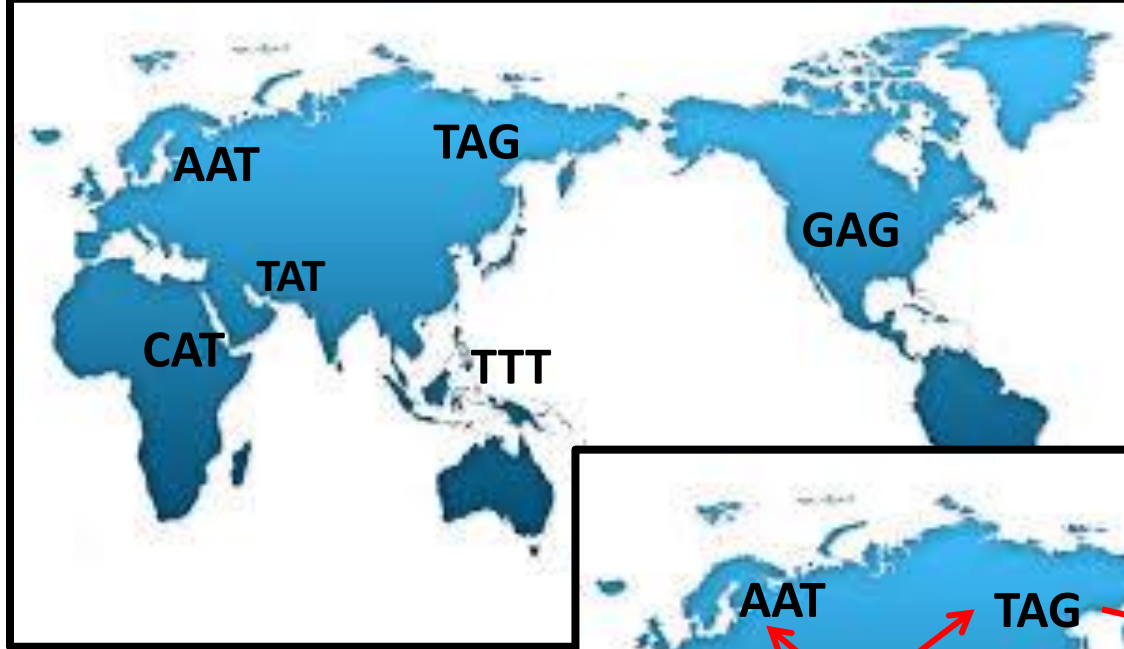


Yellow arrow pointing left: A  
Green arrow pointing right: T  
Blue arrow pointing left: C  
Red arrow pointing right: G





# How do Mutations Become Migration?



CAT	East Africa
TAT	Southwest Asia
TAG	Siberia
GAG	Americas



# Field Work in Puerto Rico in 2010

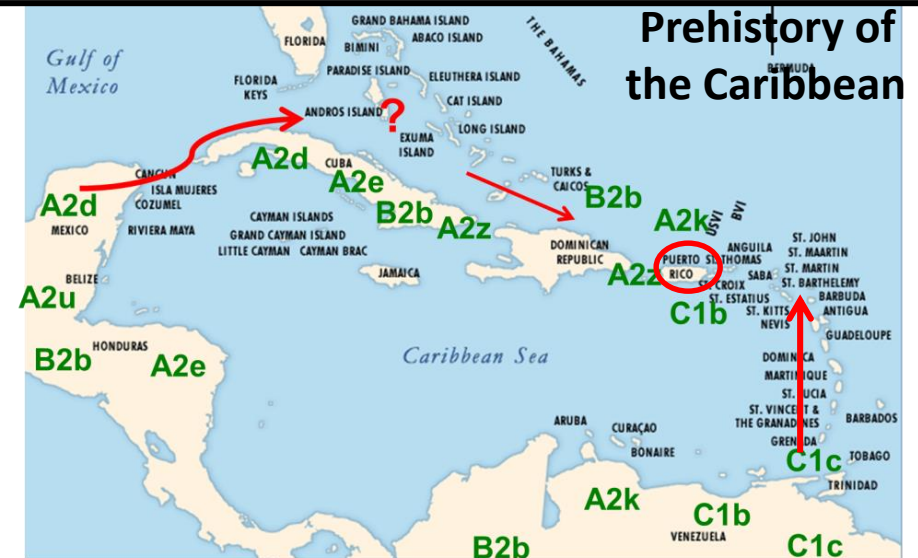
## Genetic Diversity in Puerto Rico and Its Implications for the Peopling of the Island and the West Indies

Miguel G. Vilar,<sup>1</sup> Carlalynne Melendez,<sup>2</sup> Akiva B. Sanders,<sup>1</sup> Akshay Walia,<sup>1</sup> Jill B. Gaiieski,<sup>1</sup> Amanda C. Owings,<sup>1</sup> Theodore G. Schurr,<sup>1\*</sup> and The Genographic Consortium

<sup>1</sup>Department of Anthropology, University of Pennsylvania, Philadelphia  
<sup>2</sup>National Geographic Society, IBM and the Waitt Family Foundation, University of Pennsylvania. *Liga Guakia Taina-Ke, Humacao, Puerto Rico*

## Origins and genetic legacies of the Caribbean Taino

Hannes Schroeder<sup>a,b,1</sup>, Martin Sikora<sup>a</sup>, Shyam Gopalakrishnan<sup>a</sup>, Lara M. Cassidy<sup>c</sup>, Pierpaolo Maisano Delser<sup>c,d</sup>, Marcela Sandoval Velasco<sup>a</sup>, Joshua G. Schraiber<sup>a</sup>, Simon Rasmussen<sup>f</sup>, Julian R. Homburger<sup>g</sup>, Maria C. Avila-Arcos<sup>h</sup>, Morten E. Allentoft<sup>a</sup>, J. Víctor Moreno-Mayar<sup>a</sup>, Gabriel Renaud<sup>a</sup>, Alberto Gómez-Carballea<sup>i,j</sup>, Jason E. Laffoon<sup>b,k</sup>, Rachel J. A. Hopkins<sup>l</sup>, Thomas F. G. Higham<sup>l</sup>, Robert S. Carr<sup>m</sup>, William C. Schaffer<sup>n,o</sup>, Jane S. Day<sup>p</sup>, Menno Hoogland<sup>b</sup>, Antonio Salas<sup>i,j</sup>, Carlos D. Bustamante<sup>g</sup>, Rasmus Nielsen<sup>a,q</sup>, Daniel G. Bradley<sup>c</sup>, Corinne L. Hofman<sup>b</sup>, and Eske Willerslev<sup>a,d,r,1</sup>

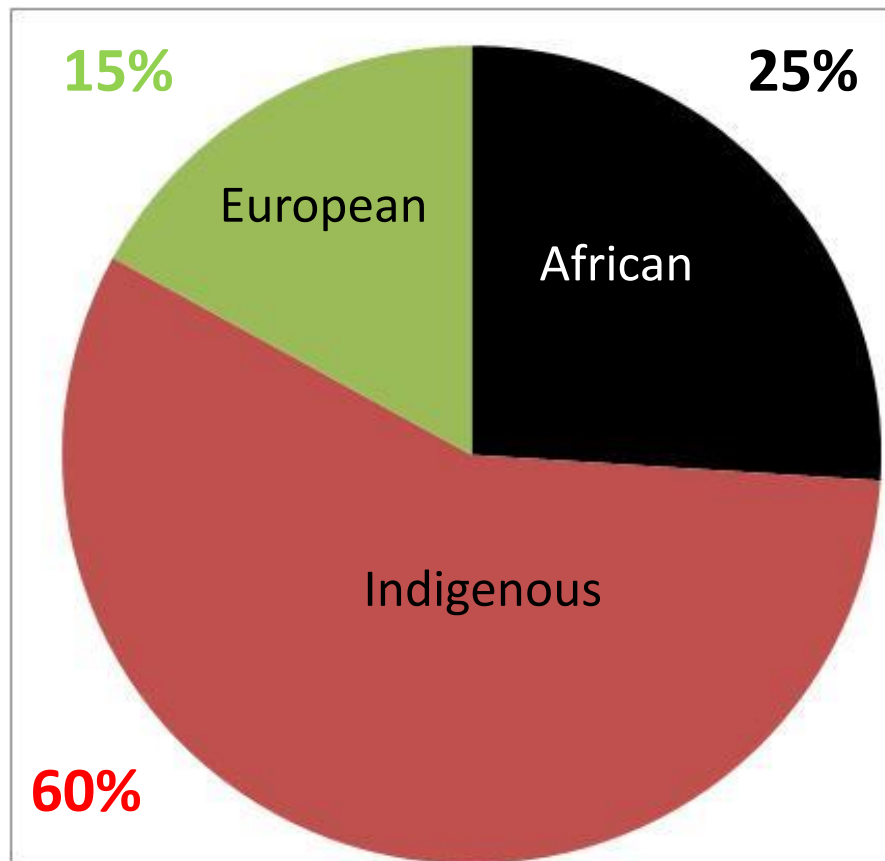


## Prehistory of the Caribbean

The Caribbean was settled in various waves originating from both Mesoamerica and South America. Further sampling, more complete mtDNA genome sequencing, and analysis of other DNA markers are still needed to clarify these patterns

# Genetic Diversity in Puerto Rico: mtDNA lineages

- Sampled from 326 participants on two expeditions to the island (2010-2011)
- We categorized diversity by looking at the origin of ancestral haplogroups



**European groups:**  
H, J, K, T, U, V

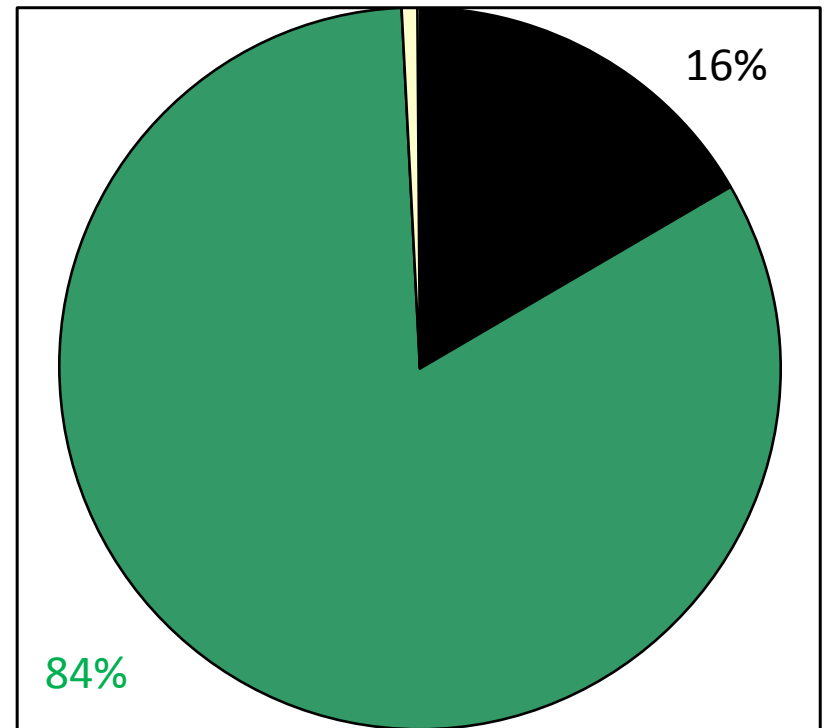
**Indigenous groups:**  
A, B, C, D, X

**African groups:**  
L0, L1, L2, L3

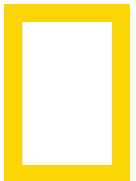


# Genetic Diversity in Puerto Rico: Y chromosome lineages

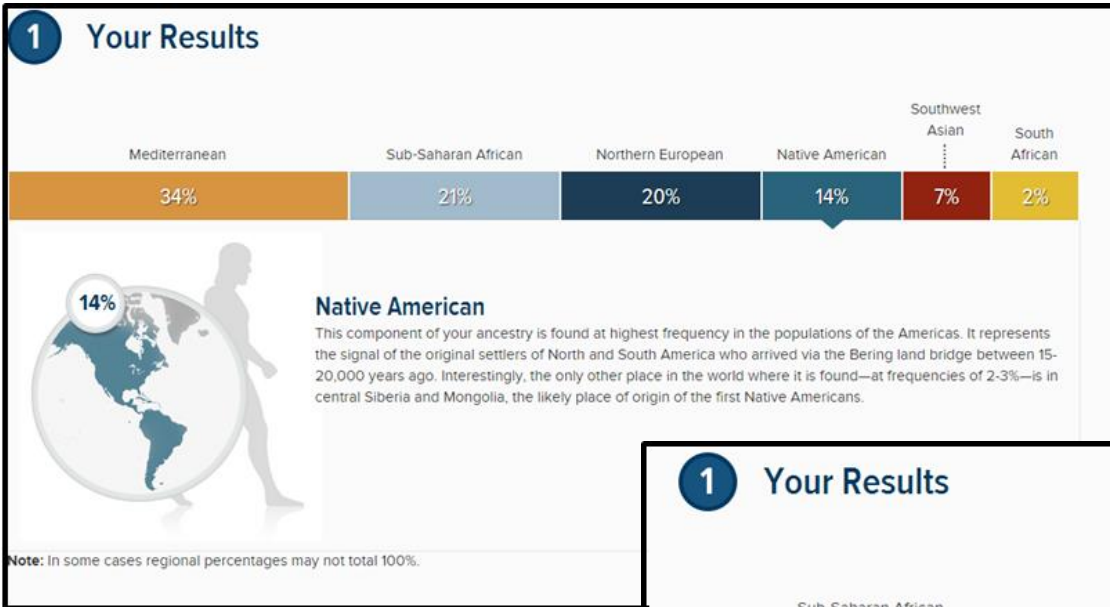
- 121 male participants
  - 41 are Haplogroup E
    - 20 E1b1a (Sub-Saharan African)
    - 21 E1b1b (Med/ N African)
  - 7 are Haplogroup G (G2a) (Med)
  - 10 are Haplogroup I
    - 5 are I2 (Med)
    - 5 are I1
  - 6 are Haplogroup J (J2a1) (Med)
  - 54 are Haplogroup R1
    - 53 R1b (Western Europe)
    - 1 R1a
  - 2 are Haplogroup T (Med/ N African)



101 are European, Middle Eastern, North African (I, J, R, T)	84%
20 were Sub-Saharan African (A, B, E)	16%
Indigenous lineages (Q)	0%

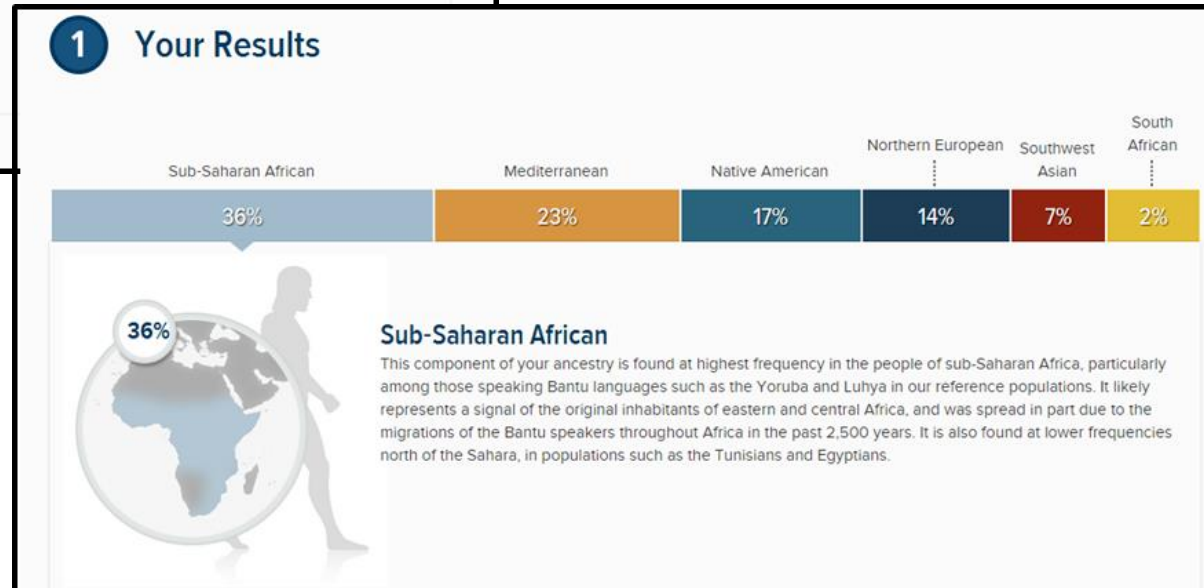


# Peopling of the Caribbean: Genetic Diversity in Puerto Rico



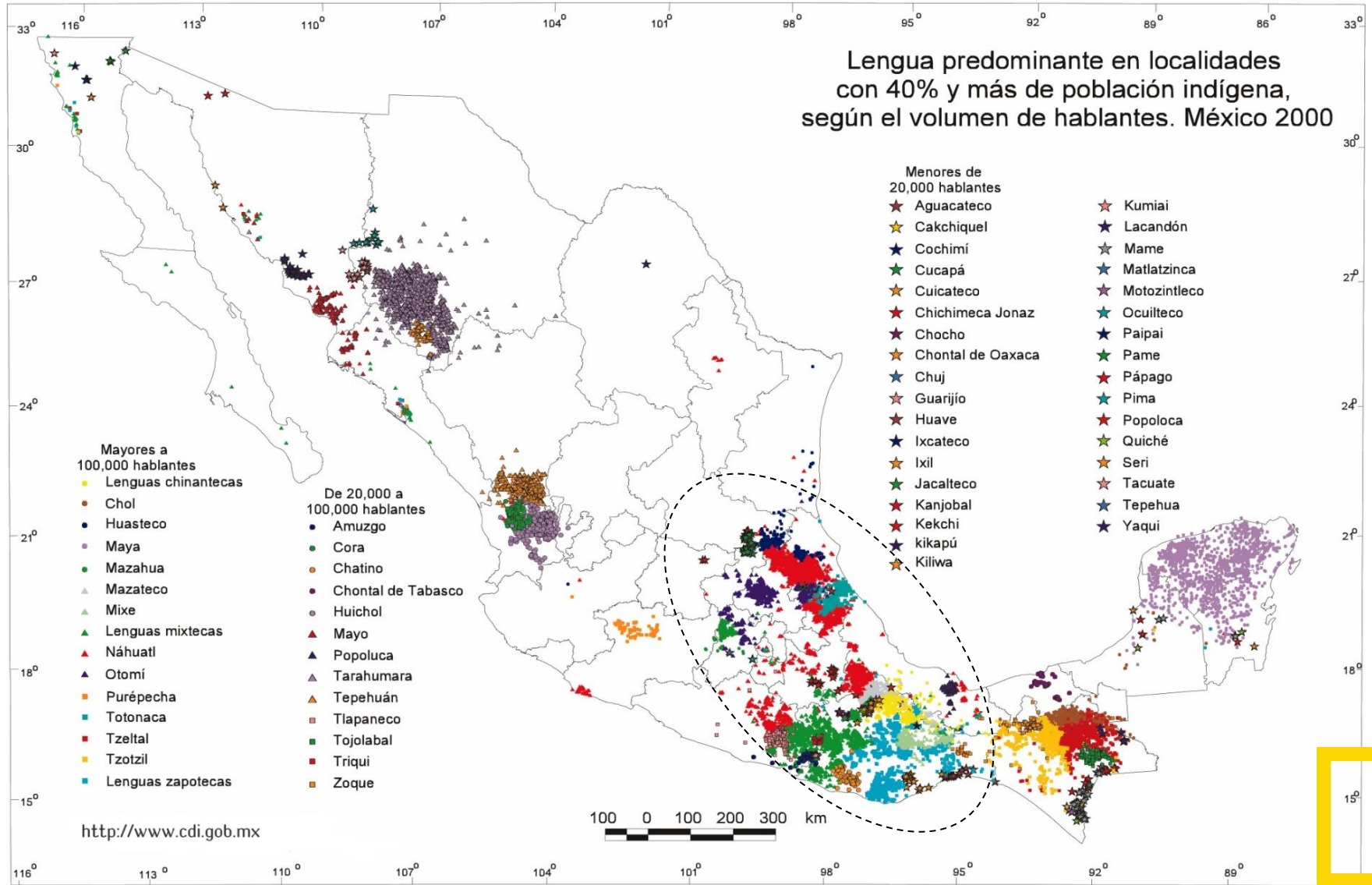
**Ancestral breakdown (Avg)**

European	40 - 60%
African	15 - 25%
Indigenous (Taino)	12 - 15%

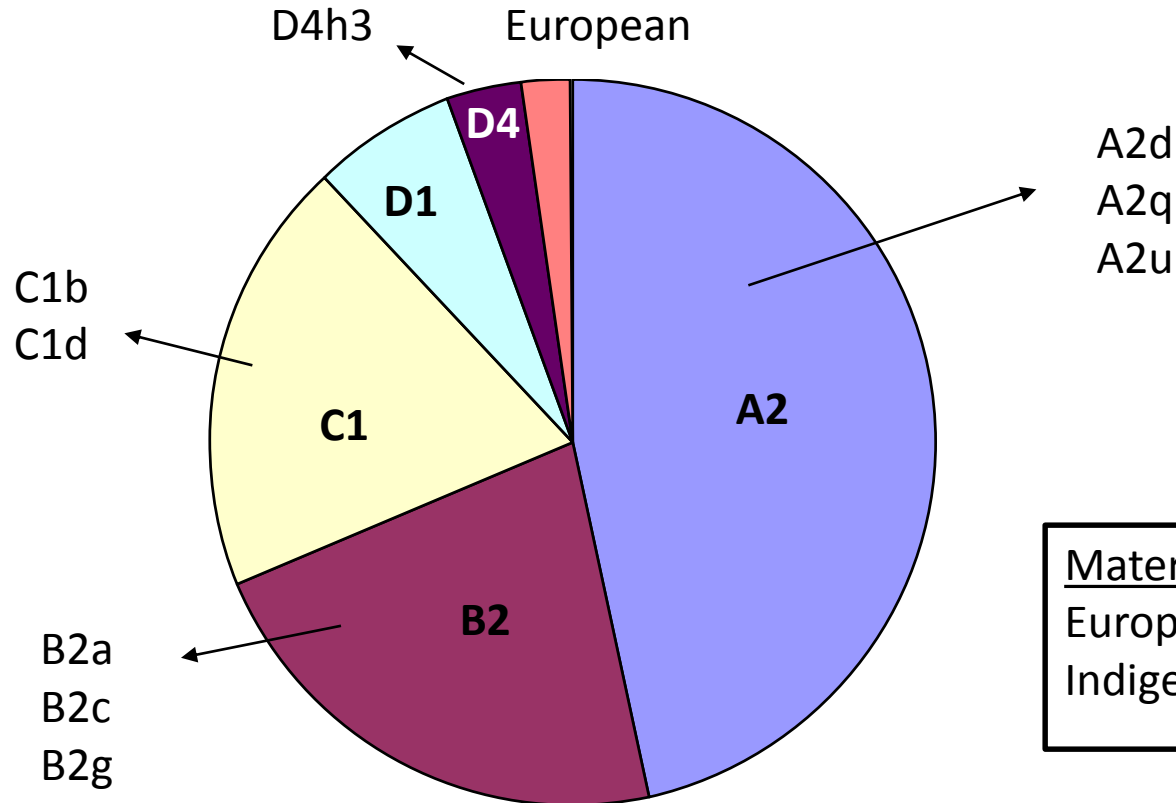


# Larger scale work in Mexico (2011-2014)

## 12 Million Indigenous Mexicans, 62 languages



# First Group, Otomi: 97% Indigenous



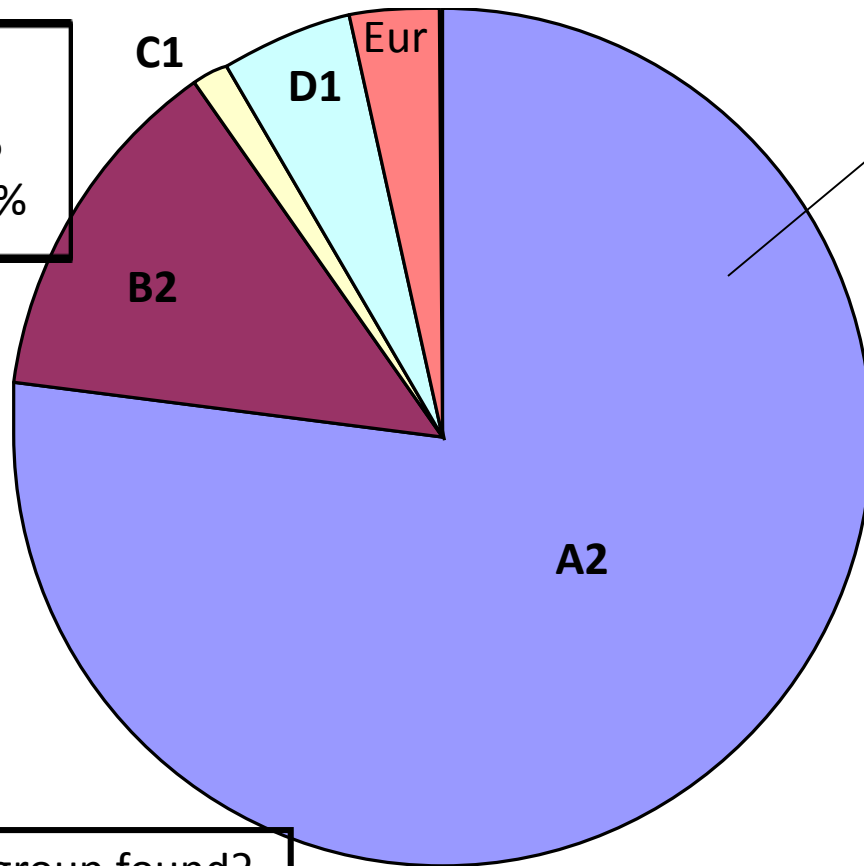
<u>Maternal Lineages:</u>	
European (H, J, K, T, U, V)	3%
Indigenous (A, B, C, D)	97%

<u>Where else in the Americas is the haplogroup found?</u>	
A2d:	Commonly found in northern Mexico
A2q:	Mesoamerica
A2u:	Mesoamerica
B2a:	Southwestern USA
B2c:	Western USA
C1b:	Throughout Americas
C1d:	Throughout Americas
D1:	Throughout Americas
D4h3:	Pacific Coast

# Second group, Nahua: 96% Indigenous

## Maternal Lineages:

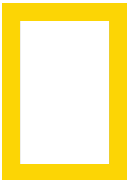
European	4%
Indigenous	96%



A2g  
A2h  
A2q  
A2u

## Where else is the haplogroup found?

- A2g: South America
- A2h: Central and South America
- A2q: Mesoamerica
- A2u: Mesoamerica





# Genographic work on Mexicans and Mexican-Americans

What does it mean to be Mexican?

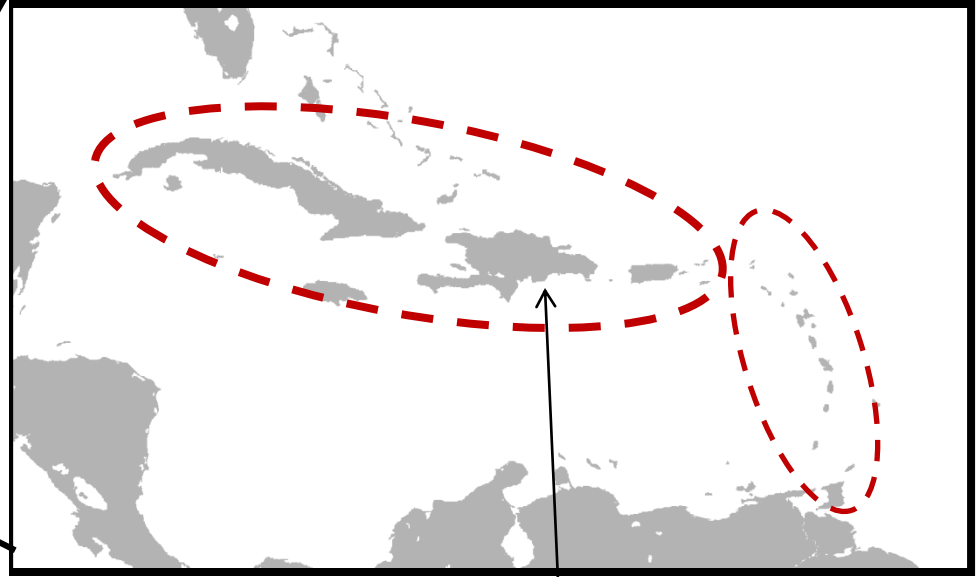
-Based on 15 Indigenous Mexicans and 20 Mexican-American participants

Indigenous Mexicans: 70% - 100% Native-American DNA

Mexican-Americans: 30 - 60% Native American DNA

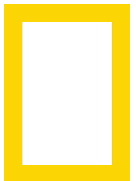


# Dominican Republic (2015-now)



What does it mean to be Dominican?

Dominican Republic:  
1,000 Samples : 40 people from 25 villages  
What is the genetic make up of a Dominican?



# What does it mean to be Hispanic?

## Mexican-Americans:

30-60% Native American DNA  
20-40% European DNA  
5-10% African DNA

## Indigenous Mexicans:

75-100% Native American DNA  
0-30% European

## Dominicans:

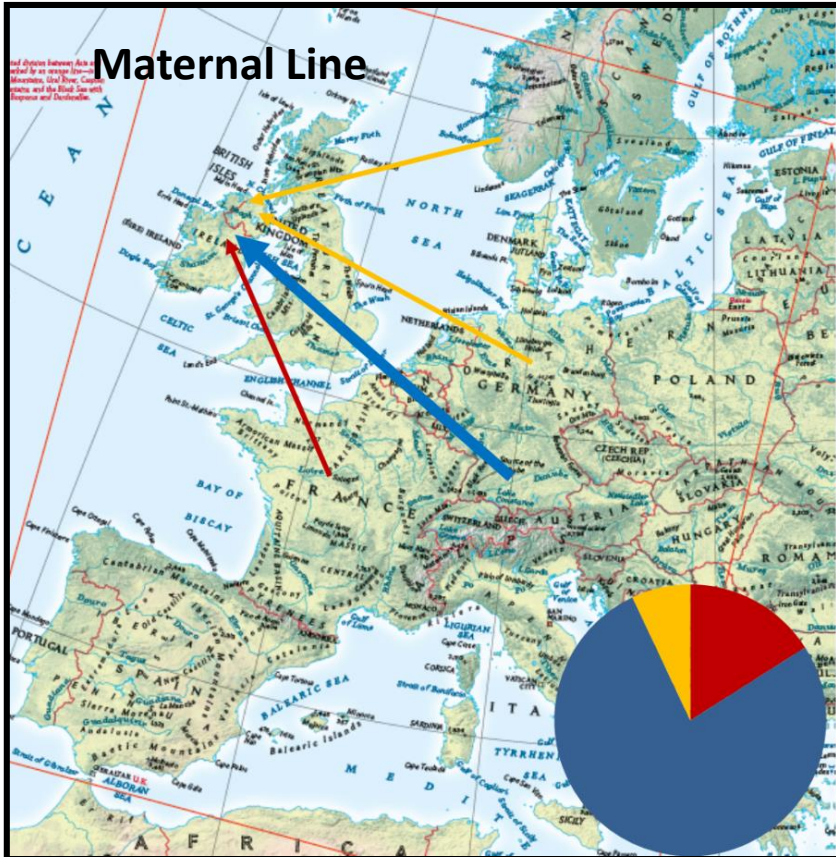
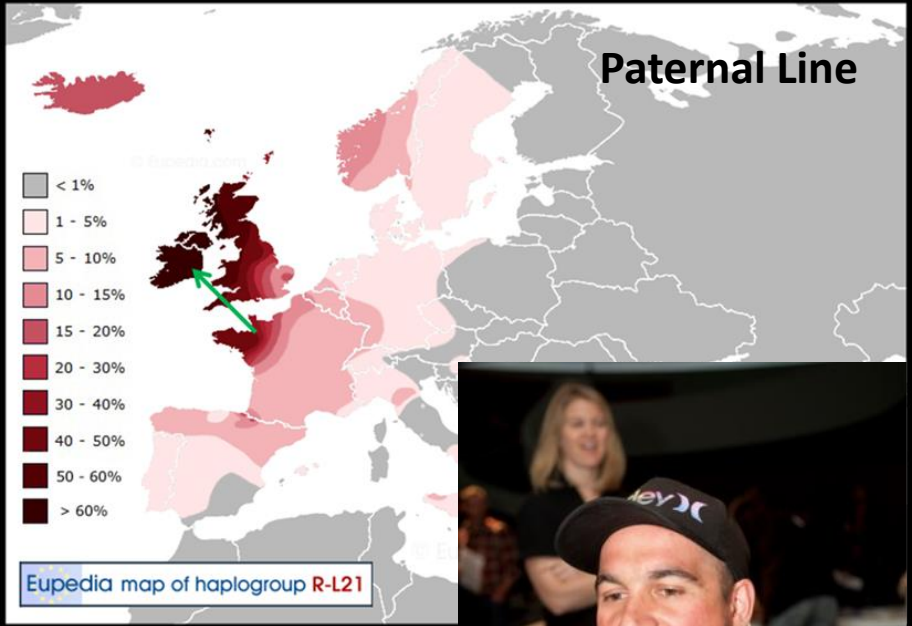
?

## Puerto Ricans:

12-15% Native American DNA  
20-25% African DNA  
50-70% European DNA



# Is it race question, or is it about identity? What does it mean to be Irish?



**Hunter-Gatherers**  
(7,000 BC - 4,000 BC)  
Haplogroups U, U4, U5, U8

**Late Mesolithic/Neolithic**  
(4,000 BC - 2500 BC)  
Haplogroups H, J, K, T2, V, X

**Late Neolithic-Bronze Age**  
(2,500 BC – 1 AD)  
Haplogroup I, T1



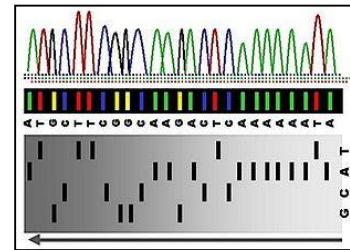
Ireland Jul 2013

# Genographic: Public Participation

- Genographic (Geno 1.0) 2005 - 2012 ~520,000 participants
  - MtDNA: Maternal line
  - Y-Chromosome: Paternal line



- Read DNA Sequences



- Geno 2.0 2012 – 2018 ~450,000 participants
  - Maternal and Paternal ancestry
  - What is your ancestral mixture?
  - 145,000 mutational spots on your genome



# What useful information is in Geno 2.0 data?

## Deep Ancestry

- Thousands of Y-SNPs from nearly every known haplogroup
- Thousands of mtDNA SNPs, most comprehensive look, just short of a complete mtDNA genome

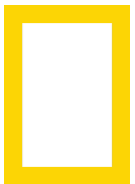


## Regional Ancestry

- Calculated from 18 or 22 Bio-geographical regions
- 60 Comparative Populations

## Hominin Ancestry

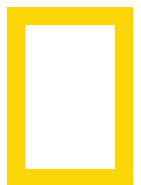
- Neanderthal estimates based on latest analyses



# Geno 2.0 Next Generation w/ Family Finder transfer



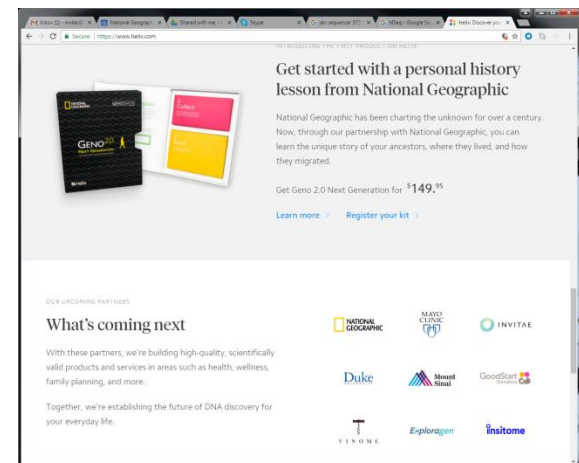
750K SNP custom array  
Based on Illumina Omni Express



# Geno 2.0 Next Generation (w/ Helix)



Next Generation Sequencing: 30-40 Million SNPs  
One Samples and you are done.  
DNA Results Marketplace  
Genealogy  
Health  
Physical Traits  
Personality, Tastes, Likes...







Ancestry:

National Geographic  
Insitome (Neanderthal, Metabolism)

Entertainment:

Wine Explorer  
Slumber Type  
Baby Glimpse  
Personalized Prints

Family:

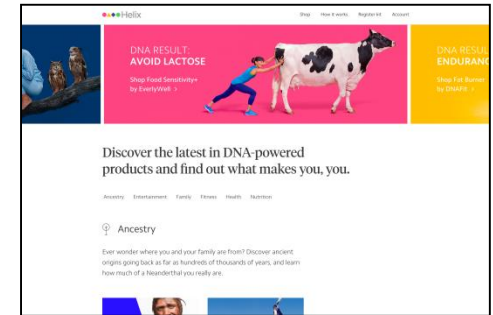
Carrier Check  
Breast Milk

Fitness:

Fat Burner  
Muscle Builder  
Fitness Buddy

Health:

Diabetes Test, Cholesterol Test



Nutrition:

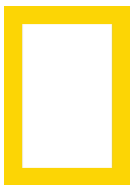
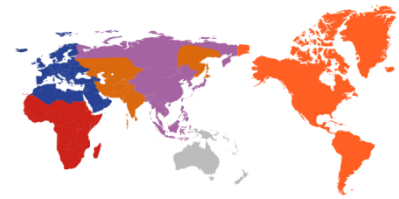
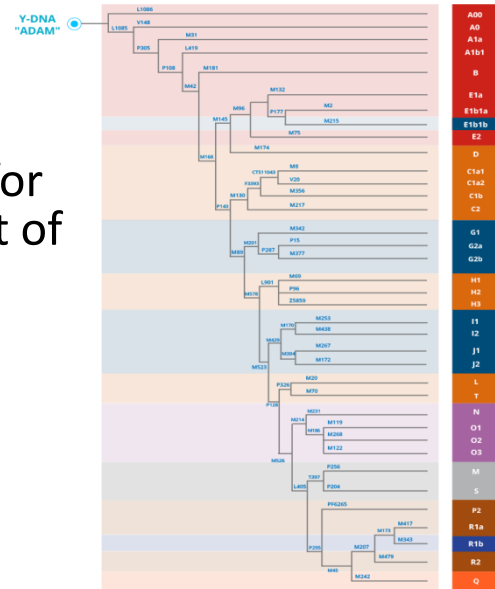
Meal Planner  
Food Sensitivity  
Weight Loss Coaching



# Objectives for 2018: Grants and Genographic Database

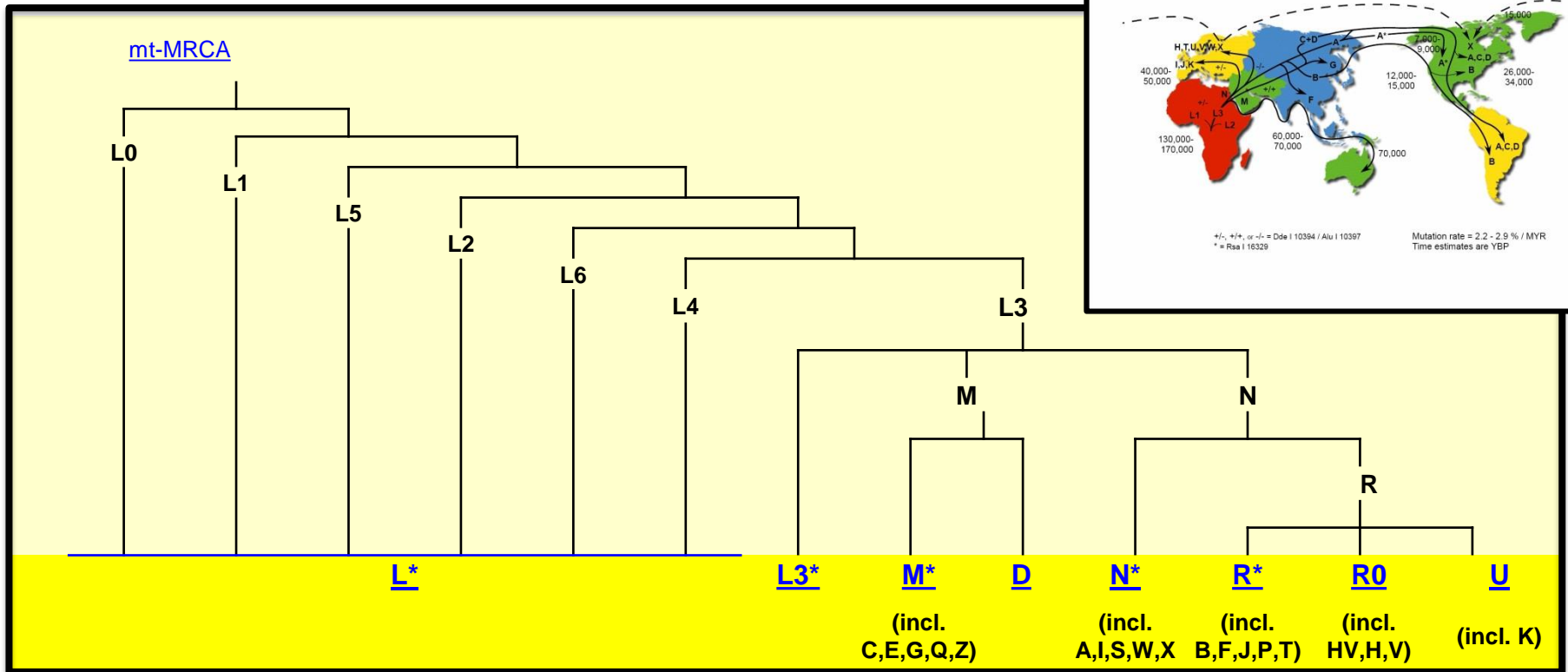
Continue to decode the human story, hidden in our DNA (and now hidden in our Database)

- Employ the best genotyping platform available specifically for Anthropological Genetics, by analyzing a comprehensive set of Y-chromosome, mtDNA, and autosomal DNA AIMs
- Grow the citizen science component, actively engaging the public and scientists in data collection and analysis, and by increasing connectivity and interconnectivity among participants
- Broaden our scientific collaboration by leveraging the Genographic Database, which is one of the lasting scientific legacies of the Project, allowing us to explore global genetic patterns at an unprecedented scale



# Geno 2.0 Next Generation for mtDNA

- Complete mitochondrial DNA genome
  - >95% in the coding region. Haplogroup calls based on Phylotree Build 16 (2014), moving to Build 17 (2016)



# Geno 2.0 Next Generation for Y-DNA

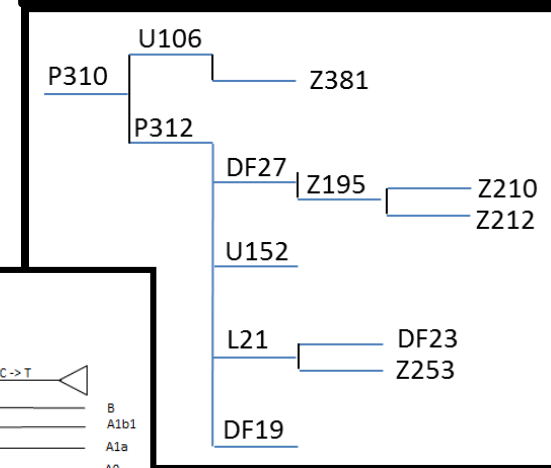
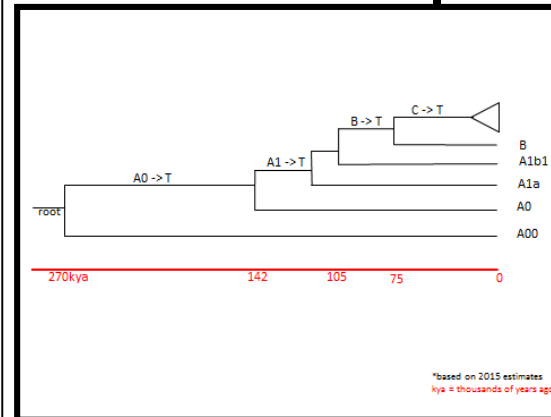
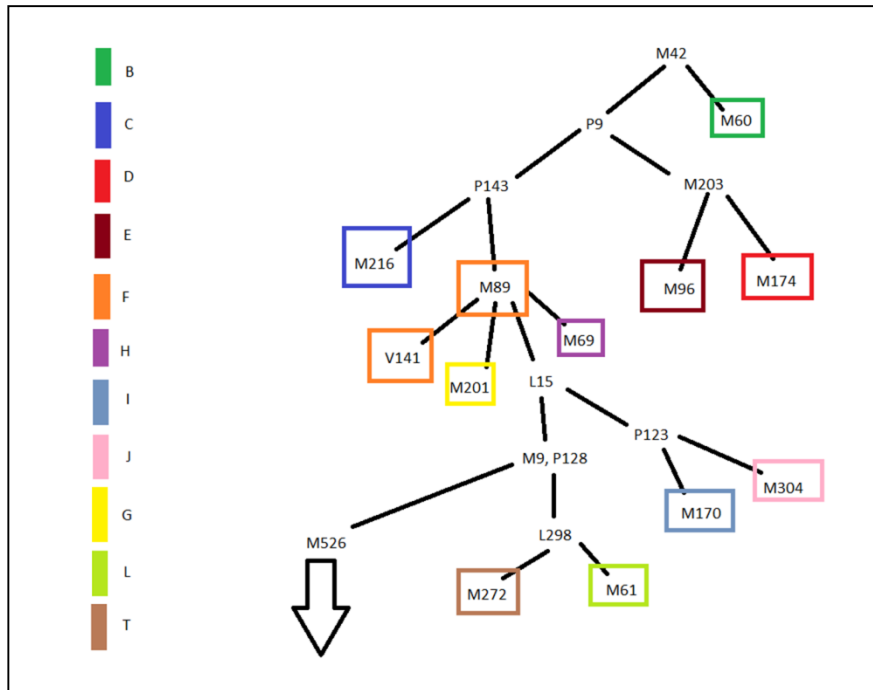
- More Y chromosome DNA SNPs

~17,000 SNPs, many we still placing on the tree.

Hundreds of E, G, I, J R1a and R1b SNPs

New origin (A00)

New bifurcations



# Newer Bio-Geographical Results (Across 30+ Regions)

Genographic results are interactive, grow with the database and updated

## Your Regional Ancestry

(100 Years - 10,000 Years Ago)

We are all more than the sum of our parts, but the results below offer some of the most fascinating and newest information possible with your Geno 2.0 Next Generation test. In this section, we display your affiliations with a set of eighteen world regions. This information is determined from your entire genome, so we're able to see both parents' information, going back six generations, or more. Your percentages reflect both recent influences and ancient genetic patterns in your DNA due to how groups migrated to and from different regions, mixing for hundreds or even thousands of years. Your ancestors may have also mixed with ancient, now extinct hominid cousins, like Neanderthals. If you or your parents have an admixed background, this pattern can get complicated very quickly! Use the reference population matches below to help understand your results.

### 1 Your Results

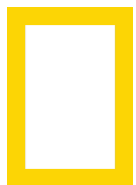


#### Great Britain and Ireland

This component of your ancestry is associated with the western European islands of Great Britain and Ireland, but traces can also be found along the northern and western coasts of continental Europe. As modern humans first entered Europe, this part of the world was uninhabitable and covered in ice sheets. As the ice sheets retreated, settlers moved to the islands. The earliest settlers likely survived on fishing, but farming eventually reached the islands in the past several thousand years. Stone monuments (e.g., Stonehenge) are associated with some of the islands' earliest cultures. Historically, these islands were populated by Celts and later marked the farthest reaches of the Roman Empire, thus genetic connections still exist between these regions. Yet it was Britain's global empire during the 18th and 19th centuries that helped spread this component, as well as the English language, throughout the world.

Today, this ancestral component is seen in people of British and Irish descent, including those throughout the United States, Canada, Australia, and most other former British colonies.

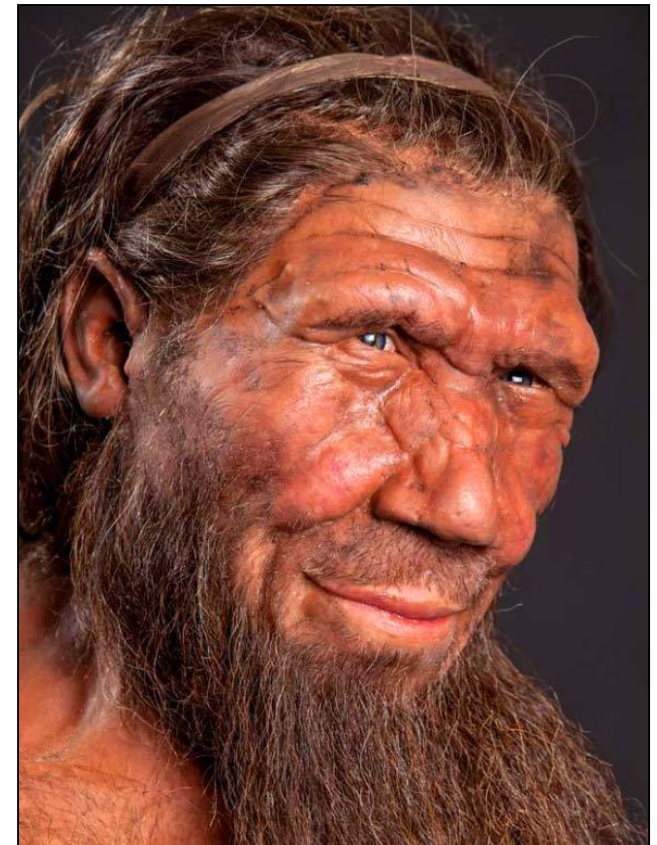
**Note:** In some cases regional percentages may not total 100%.



# Results: Your Hominin Ancestry

And it is also fun

How much Neanderthal do you have?



## YOUR HOMINID ANCESTRY

When our ancestors first migrated out of Africa around 60,000 years ago, they were not alone. At that time, at least two other species of hominid cousins walked the Eurasian landmass: Neanderthals and Denisovans. Most non-Africans are about 2% Neanderthal. Indigenous sub-Saharan Africans have no Neanderthal DNA because their ancestors did not migrate through Eurasia.

2.9%

NEANDERTHAL

LEARN MORE >

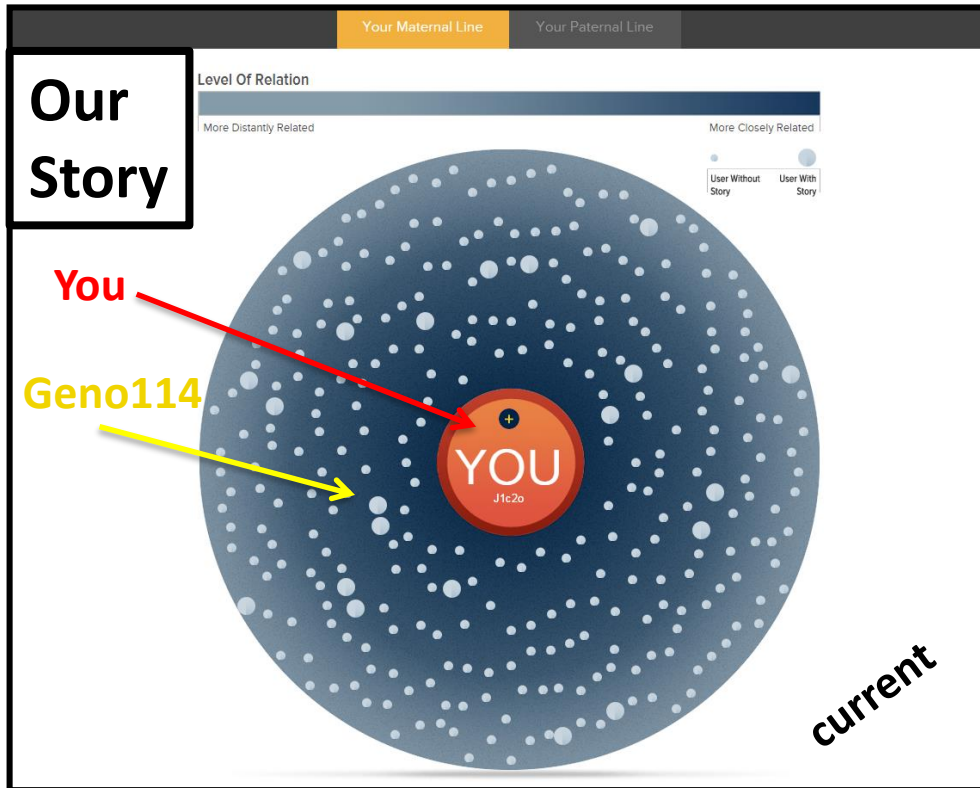
3.9%

DENISOVAN

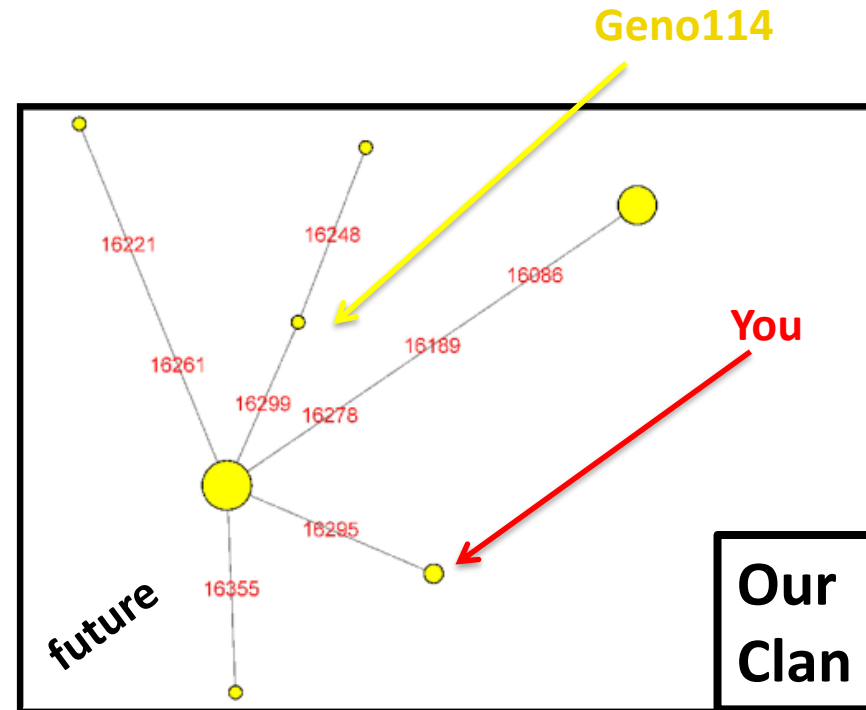
LEARN MORE >



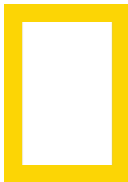
# Geno 2.0 meets Citizen Science



- Share a story, no communication
- Count the number of “relatives”



- Number of SNPs apart
- Mutation rate -> haplogroup age
- How closely related are you?
- Share a story, communicate



# What can be next?

(2018 - 2022 )

**Ancient DNA:** Remove the leaves – Look inside the Tree. We know past cultures through materials, now through DNA  
(Science)

Work with Archaeologists, Curators

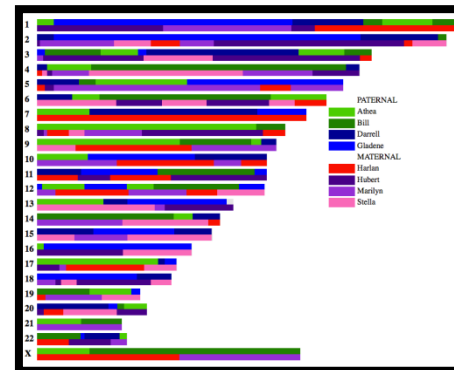
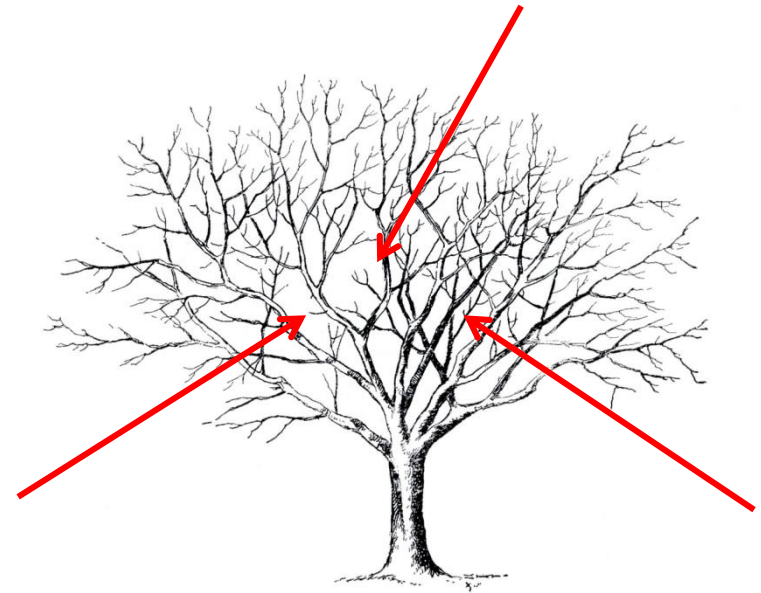
**Database Growth/Universities & Education**

Bring Help to “Read the Leaves” (DAR)

Publish new results from the DAR

(Science) Remake the Maps, Educ. Materials

**Storytelling: What are the New Stories?**





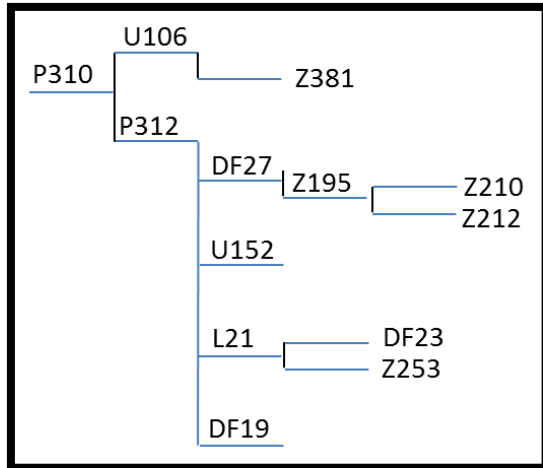
# Growing research in Ancient DNA w/ new Lab Partnerships



48 Ancient DNA samples (500 and 8,000 BCE)  
-Rare and new lineages  
70 Modern samples (Yagan, Kaweshkar, Selknam)



# Bringing Genealogists and Academics Together



Genealogist researching R1b wants to know when/where a SNP arose?

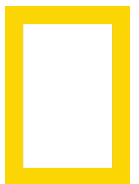
**Genographic DAR**  
Gains access to the Genographic DAR, discovers 500 individuals with same/related SNPs  
  
Grows their Tree  
Identifies possible place of origin of SNP

Contact Genographic ([mvilar@ngs.org](mailto:mvilar@ngs.org))  
To get information about participant and Researcher working in region

Possible research collaboration comes  
  
Jointly apply for NatGeo Grant to sample/analyze new group, population

Grow the Tree  
Grow the Admin Group  
  
**Publish Research**

**Any other Ideas?**

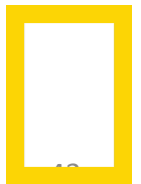
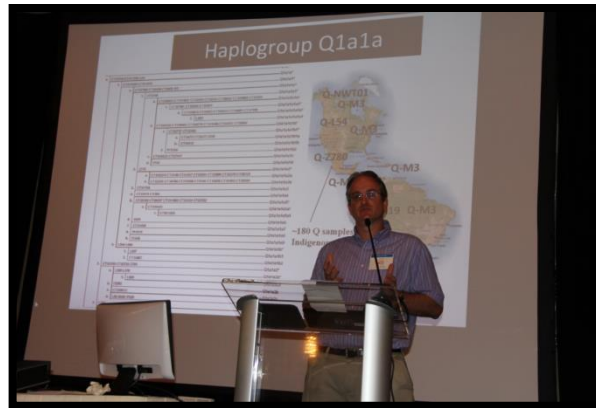


# Bring in Schools and Universities

Howard University, Cornell University, Harvard University, University of Texas  
George Washington University, University of Pennsylvania, University of Maryland

More than 200 Educators buy kits every year

Developing:  
Maps  
Lessons  
Curriculum



# Genographic Project Database or DAR

as of Fall 2017 (estimate)

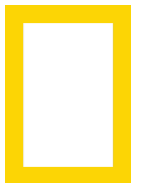
## Current status of DAR (Geno 1.0)

Number of consumer participants	429,777
control Region mtDNA	216,675
Y chromosome (dozen STR)	213,102
Research Center Participants	78,819

## Current status of DAR (Geno 2.0)

Participants with completed Geno 2.0 results	
~215,000	
Participants with Geno 2.0 Next Generation (Array and Sequencing)	
~145,000	
Percentage of people that gave email addresses	87%
Percentage of participants that opted-in to science	65%
Demographic information for most participants	

**Size of the Genographic Project database: > 890,000 individuals**



# Genographic Database (DAR) Search Tool

Username:   
Password:

A LANDMARK STUDY OF THE HUMAN JOURNEY

THE GENOGRAPHIC PROJECT

MIDNA Tree Search

by HG | by mutation

Enter Haplogroup:  
M7c1a1

Markers  
G12372A  
G13759A

Intro Stories | Detailed Stories | Migration Map | Heat Map (for M7)

- root (No Story)
- RSRs (No Story)
- L1-2-3-4-5-6
- L2-3-4-5-6 (No Story)
- L2-3-4-6 (No Story)
- L3-4-6 (No Story)
- L3-4 (No Story)
- L3
- M
- M7
- M7b-c (No Story)
- M7c (No Story)
- M7c1 (No Story)
- M7c1a (No Story)
- M7c1a1 (No Story)

Tree

SNP

Story

To apply to work on the Genographic Project database email me at: [mvilar@ngs.org](mailto:mvilar@ngs.org)

# Genographic Database Search Tool

Username:   
Password:

A LANDMARK STUDY OF THE HUMAN JOURNEY

THE GENOGRAPHIC PROJECT

MIDNA Tree | Y Chromosome Tree

Y Tree Search

by HG | by mutation

Enter SNP:  
P312

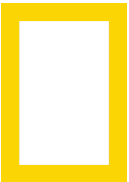
Y Haplogroups  
R1b1a1a1a1a

Intro Stories | Detailed Stories | Migration Map | Heat Map (for P305)

- P305 (from haplogroup root) (No Story)
- M42 (from haplogroup A1)
- M168 (from haplogroup A1b)
- P143 (from haplogroup BT) (No Story)
- M89 (from haplogroup CT)
- M578 (from haplogroup CF) (No Story)
- P128 (from haplogroup F)
- M526 (from haplogroup HLJK) (No Story)
- M45 (from haplogroup LJK)
- M207 (from haplogroup K)
- P231 (from haplogroup MNOPS)
- L278 (from haplogroup MPS)
- M343 (from haplogroup P)
- M269 (from haplogroup P1)
- P310 (from haplogroup R)
- P312 (from haplogroup R1)

The common direct paternal ancestor of all men alive today was born in Africa around 140,000 years ago. He was neither the first human male nor the only male alive in his time. He was the only male whose direct lineage is present in current generations. Most men, including your direct paternal ancestors, trace their ancestry to one of this man's descendants.

Your branch of this lineage took part in out-of-Africa migrations. Your ancestors traveled to West Asia where they lived by hunting wildlife and gathering wild fruits and berries. Over time, groups containing this branch spread west toward Europe and east to Central Asia, then south into the Levant region.



# Genographic Database Search Tool

Welcome, Miguel Vilar!

Logout

Demographic Data | MitDNA Data | Y SNP Data | Y STR Data | Admixture | Autosomal File

Search | My Recordsets | Combine | Records for Chamorro

Genetics

Y Haplogroup:

Y Shorthand:

MitDNA Haplogroup:

Search Type:

Find records where:

is:

what:

## Search the Genographic database by:

mitochondrial haplogroup

Y haplogroup (long form, short form)

Y-SNP

mtDNA position (SNP)

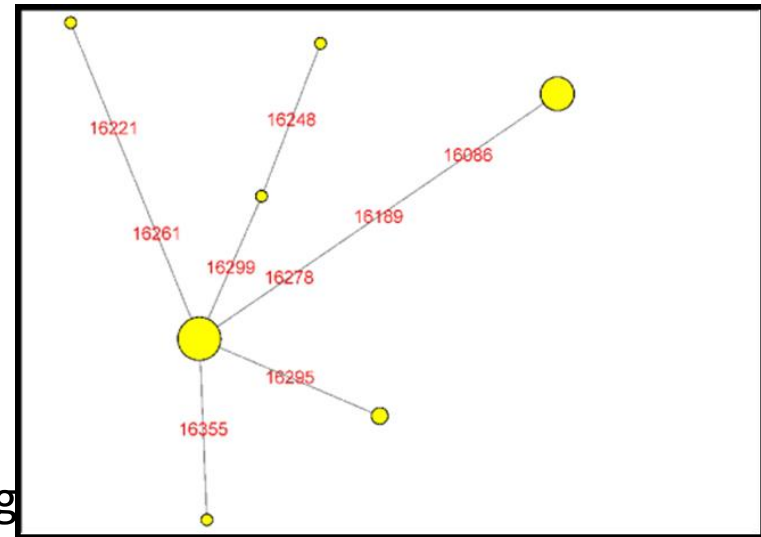
Place of Birth

Current ZipCode

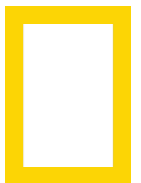
maternal or paternal POB

maternal or paternal ethnicity & language

Same for grandpaternal and grandmaternal



Bringing Academic Tools (Network, Admixture) to Genealogists



# How can Genealogists use the DAR?

- Research your own group: Origin, dispersal, history
  - ie. more than 1,200 R1b-M222
- Research by Ethnicity/Place of Birth, Migration
  - ie. More than 2,000 people with Greek Ancestry
- Research by Zip Code: Where are the haplogroups now?
- Research by SNP: Possibilities are endless
- Help us discover new branches of the Y and mtDNA Tree. On going work...
  - working with Rebekah Canada on a Y haplogroup Q tree
  - working with Roberta Estes on mtDNA haplogroups in Americas
  - working with Marie Rundquist on Y DNA in the Americas
  - I'm also working with 12 field researchers, and >30 data researchers
- Genographic as the path for Genealogists & Academics to work together

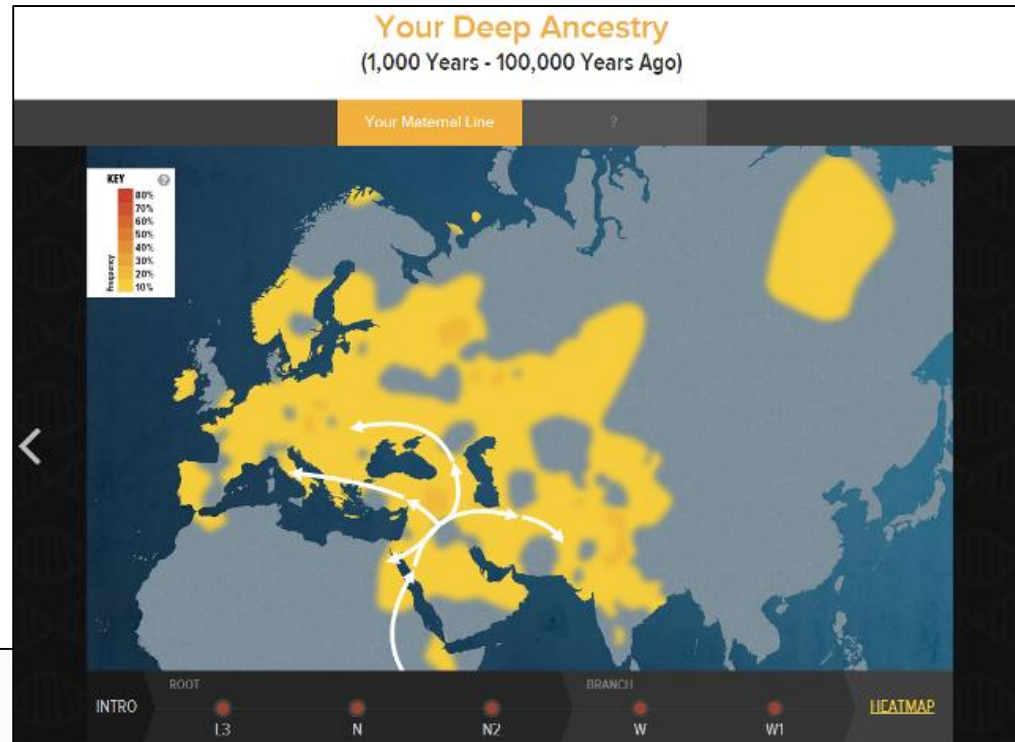




# Participant #1

## Maternal Haplogroup: W1

“A lineage with strongly connected to Balto-Slavic people. Finland (9.6%), Hungary (5.2%), Latvia (4.1%), Macedonia (4%) and Belarus (3.7%). The Finns and the Hungarians are both speakers of an Uralic language, which would imply an Uralic connection with haplogroup W”



## 1 Your Results

Mediterranean

37%

Northern European

33%

Southwest Asian

18%

Sub-Saharan

African

Native American

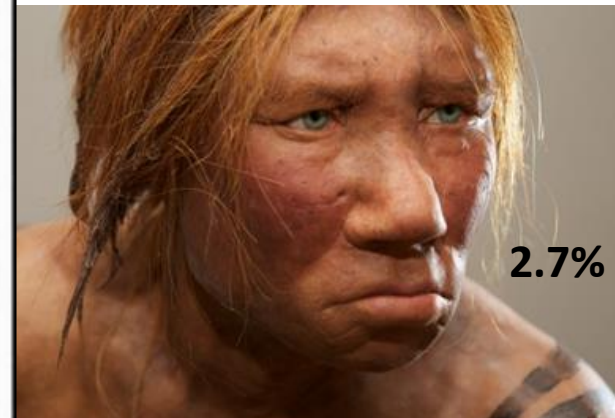
5%

5%



### Mediterranean

This component of your ancestry is found at highest frequencies in southern Europe and the Levant—people from Sardinia, Italy, Greece, Lebanon, Egypt and Tunisia in our reference populations. While not limited to these groups, it is found at lower frequencies throughout the rest of Europe, the Middle East, Central and South Asia. This component is likely the signal of the Neolithic population expansion from the Middle East, beginning around 8,000 years ago, likely from the western part of the Fertile Crescent.



# Participant #2

## Your Deep Ancestry (1,000 Years - 100,000 Years Ago)

Your Maternal Line

Your Paternal Line

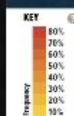
### Maternal Haplogroup H6a

INTRO L3 N R R0 R0 IIIV II H6 H6A H6A1B IIEATMAP

## Your Deep Ancestry (1,000 Years - 100,000 Years Ago)

Your Maternal Line

Your Paternal Line



### Paternal Haplogroup E1b1a

INTRO New! P305 M42 M168 M203 M96 P147 P177 P2 M2 U174 M191 IIEATMAP

## 1 Your Results

Sub-Saharan African

29%

Northern European

28%

Mediterranean

26%

Southwest Asian

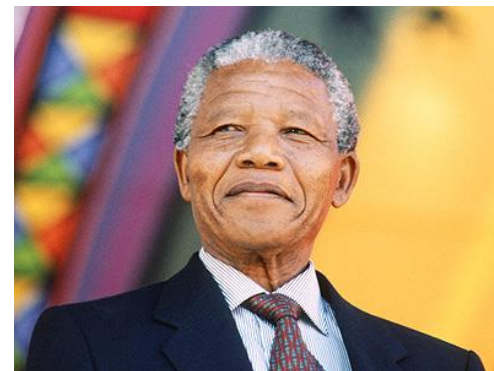
13%



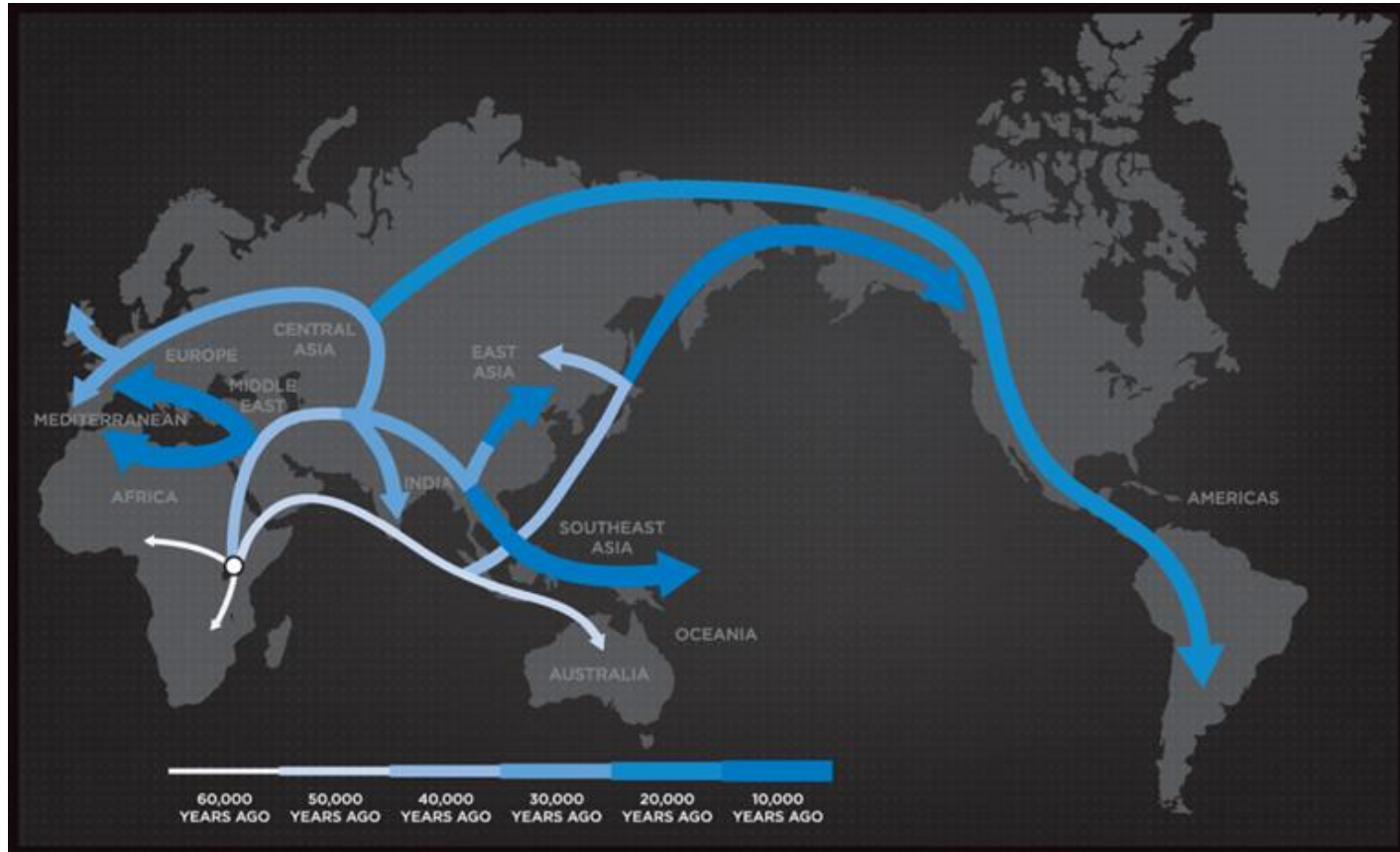
### Sub-Saharan African

This component of your ancestry is found at highest frequency in the people of sub-Saharan Africa, particularly among those speaking Bantu languages such as the Yoruba and Luhya in our reference populations. It likely represents a signal of the original inhabitants of eastern and central Africa, and was spread in part due to the migrations of the Bantu speakers throughout Africa in the past 2,500 years. It is also found at lower frequencies north of the Sahara, in populations such as the Tunisians and Egyptians.

Same lineage as:



# After 13 Years, What is the take away?



# Thank You!



**NATIONAL  
GEOGRAPHIC**

THE  
**GENOGRAPHIC**  
PROJECT