Summarizing and Anticipating the Next Decade with NRY, mtDNA and Autosomal DNA

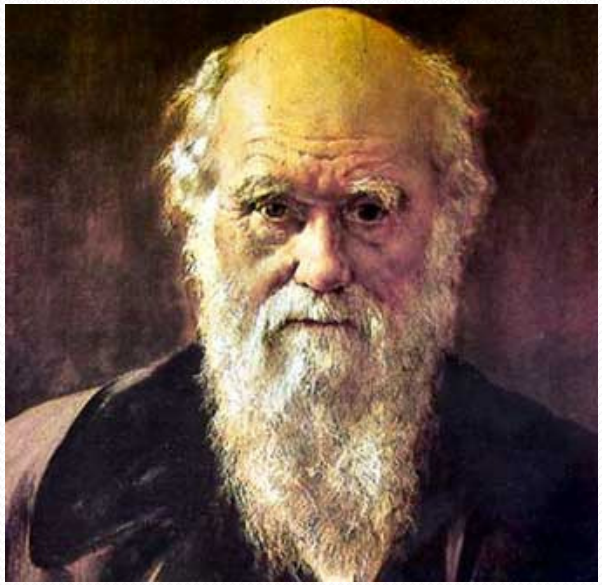
Doron M Behar
Family Tree DNA
6th International Conference on Genetic Genealogy

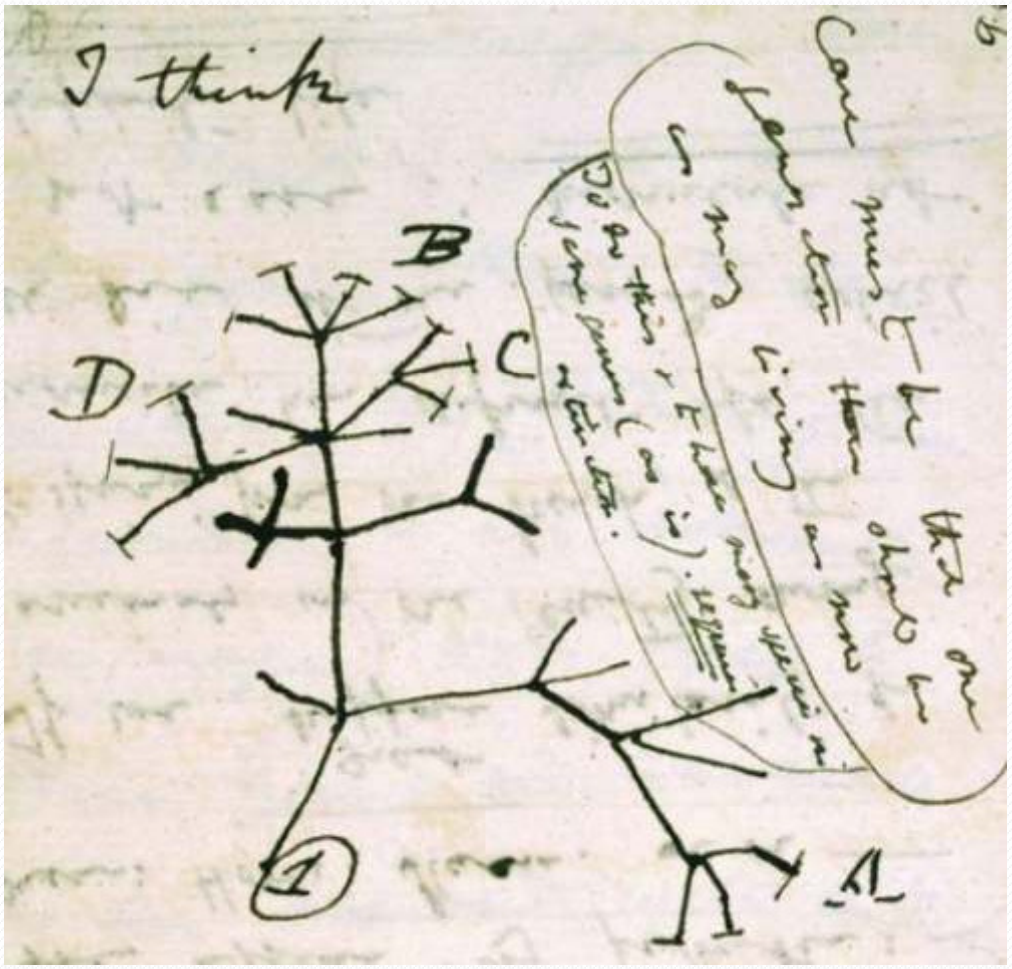
Historical landmarks(1/2)

- 1859 - Charles Robert Darwin: On the origin of species
- 1866 - Gregor Johann Mendel: Experiments on Plant Hybridization
- 1905 – William Bateson: Coin the term “genetics”
- 1910 - Thomas Hunt Morgan: Genes are on chromosomes
- 1944 - Oswald Theodore Avery, Colin McLeod and Maclyn McCarty - identified the DNA molecule

“I think”

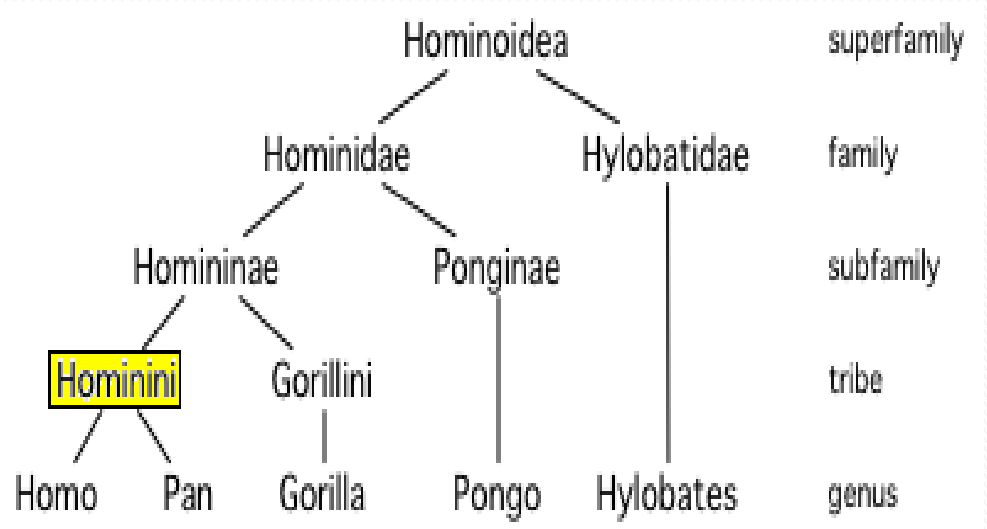
Darwin's first diagram of an evolutionary tree from his 'First Notebook on Transmutation of Species' (1837)."



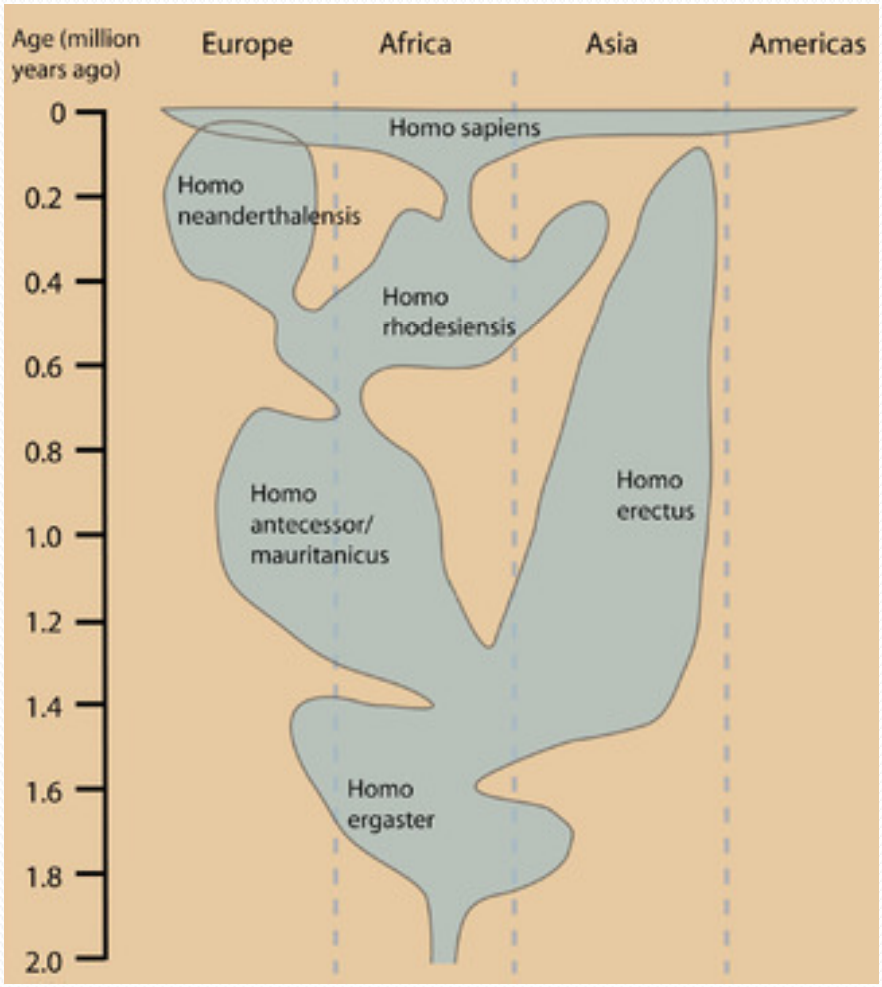


Hominini Tribe

Kingdom:	Animalia
Phylum:	Chordata
Class:	Mammalia
Order:	Primates
Family:	Hominidae
Subfamily:	Homininae
Tribe:	Hominini
Genus:	Homo
Species:	H. sapiens



Homo Genus


Age (million years ago)
Europe
Africa
Asia
Americas
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
Homo sapiens
Homo neanderthalensis
Homo rhodesiensis
Homo antecessor/ mauritanicus
Homo erectus
Homo ergaster

Historical landmarks(2/2)

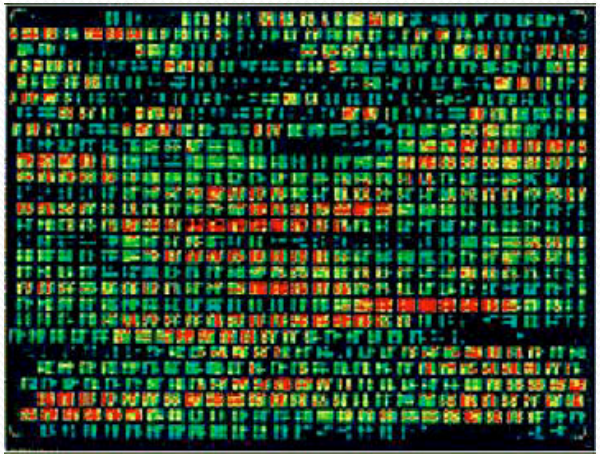
- 1953 –Rosalind Franklin, James D. Watson, Fancis Crick: the double helix
- 1977 – Frederick Sanger: DNA sequencing
- 1983 – Kary Banks Mullis: Polymerase Chain Reaction
- 2003 – Human Genome Project and Celera Genomics: the human genome
- 2005 – HapMap project
- 2008 - 1000 Genomes Project

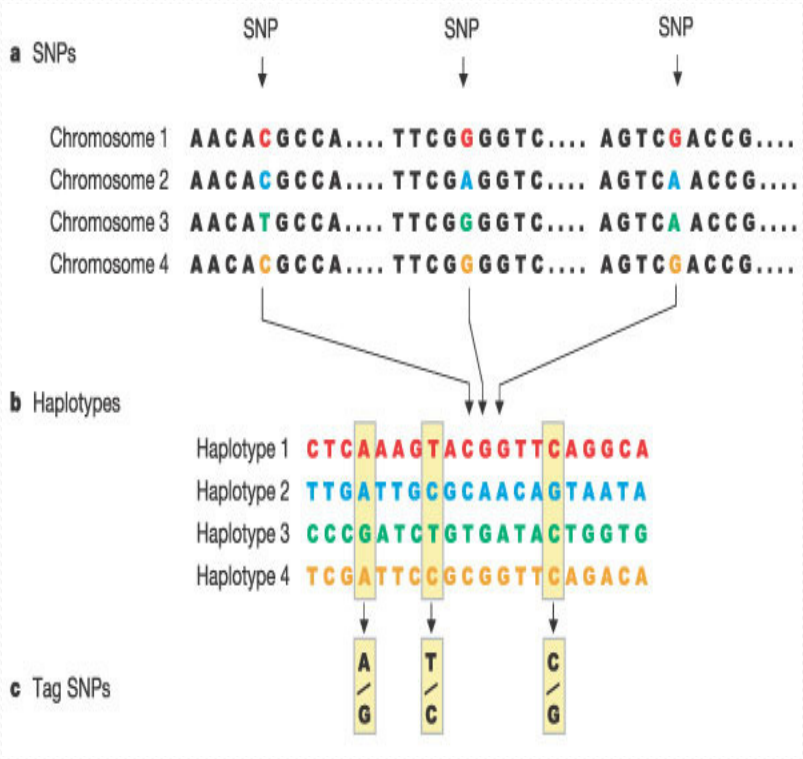
Human Genome Project

- A 13 years old project completed in 2003:
  - *identify* all the 20,000-25,000 human genes
  - *determine* the sequences of human DNA
  - *store* this information in databases,
  - *improve* tools for data analysis
  - *transfer* related technologies to the private sector
  - *address* the ethical, legal, and social issues
- 200 separate principal investigators for 18 countries
- Reported to have cost $3 billion

HapMap project

- DNA sequences differ in every 1,200 base pairs
- Approximately 10 million SNPs estimated to occur commonly in the human genome
- The number of tag SNPs are only 600,000 bp
- Genotype only these SNPs to locate genes involved in medically important traits.





1,000 Genomes Project

Samples and ELSI Group

Leena Peltonen (co-chair) Sanger Institute
Bartha Knoppers (co-chair) University of Montreal
Aravinda Chakravarti (co-chair) Johns Hopkins
Gonçalo Abecasis University of Michigan
Richard Gibbs Baylor College of Medicine
Lynn Jorde University of Utah
Eric Juengst Case Western Reserve University
Jane Kaye Oxford University
Alastair Kent Genetic Interest Group
Rick Kittles University of Chicago
Jim Mullikin National Human Genome Research Institute
Mike Province Washington University in St. Louis
Charles Rotimi Howard University
Yeyang Su Beijing Genomics Institute
Chris Tyler-Smith Sanger Institute
Ling Yang Beijing Genomics Institute

Steering Committee

Richard Durbin (co-chair) Sanger Institute
David Altshuler (co-chair) Broad / MGH / Harvard
Gonçalo Abecasis University of Michigan
Aravinda Chakravarti Johns Hopkins
Andrew Clark Cornell University
Francis Collins National Human Genome Research Institute
Peter Donnelly Oxford University
Paul Flicek European Bioinformatics Institute
Stacey Gabriel Broad Institute
Richard Gibbs Baylor College of Medicine
Bartha Knoppers University of Montreal
Eric Lander Broad Institute
Elaine Mardis Washington University in St. Louis
Gil McVean Oxford University
Debbie Nickerson University of Washington
Leena Peltonen Sanger Institute
Stephen Sherry National Center for Biotechnology Information
Rick Wilson Washington University in St. Louis
Huanming (Henry) Yang Beijing Genomics Institute

Production Group

Elaine Mardis (co-chair) Washington University in St. Louis
Stacey Gabriel (co-chair) Broad Institute
Richard Durbin Sanger Institute
Richard Gibbs Baylor College of Medicine
David Jaffe Broad Institute
Ruiqiang Li Beijing Genomics Institute
Donna Muzny Baylor College of Medicine
Chad Nusbaum Broad Institute
Aarno Palotie Sanger Institute
Dan Turner Sanger Institute
Jun Wang Beijing Genomics Institute
Richard Gibbs Baylor College of Medicine
Sol Katzman UCSC
Hoda Khouri National Center for Biotechnology Information
Martin Shumway National Center for Biotechnology Information
Jun Wang Beijing Genomics Institute
George Weinstock Baylor College of Medicine
(Broad representative)

Data Flow Group (being formed)

Paul Flicek (co-chair) European Bioinforma
Stephen Sherry (co-chair) National Center
Ewan Birney European Bioinformatics Institu
Clive Brown Sanger Institute
David Dooling Washington University in St.
Richard Gibb
Sol Katzman
Hoda Khouri

Funders

Alan Schafer Wellcome Trust
Francis Collins National Human Genome Research Institute
Lisa Brooks National Human Genome Research Institute
Audrey Duncanson Wellcome Trust
Adam Felsenfeld National Human Genome Research Institute
Mark Guyer National Human Genome Research Institute
Ruth Jamieson Wellcome Trust
Jane Peterson National Human Genome Research Institute
Zhiwu Ren National Planning and Development Committee
Jian Wang Beijing Genomics Institute

Analysis Group

Gil McVean (co-chair) Oxford University
Gonçalo Abecasis (co-chair) University of Michigan
David Altshuler Broad / MGH / Harvard
Paul de Bakker Broad / BWH / Harvard
Brian Browning University of Auckland
Sharon Browning University of Auckland
Carlos Bustamante Cornell University
David Carter Sanger Institute
Aravinda Chakravarti Johns Hopkins
Andrew Clark Cornell University
Don Conrad Sanger Institute
Mark Daly Broad / MGH / Harvard
Manolis Dermitzakis Sanger Institute
Peter Donnelly Oxford University
Richard Durbin Sanger Institute
Evan Eichler University of Washington
Paul Flicek European Bioinformatics Institute
Bryan Howie Oxford University
Matt Hurles Sanger Institute
David Jaffe Broad Institute
Lynn Jorde University of Utah
Hoda Khouri National Center for Biotechnology Information
Eric Lander Broad Institute
Charles Lee Brigham and Women's Hospital
Guoqing Li Beijing Genomics Institute
Heng Li Sanger Institute
Ruiqiang Li Beijing Genomics Institute
Yingrui Li Beijing Genomics Institute
Yun Li University of Michigan
Jonathan Marchini Oxford University
Gabor Marth Boston College
Steve McCarroll Broad Institute
Jim Mullikin National Human Genome Research Institute
Simon Myers Oxford University
Rasmus Nielsen University of California, Berkeley
Aarno Palotie Broad / Harvard
Jonathan Pritchard University of Chicago
Mike Province Washington University in St Louis
Molly Przeworski University of Chicago
Shaun Purcell Broad / MGH / Harvard
Noah Rosenberg University of Michigan
Pardis Sabeti Broad / Harvard
Paul Sch
Steven S
Jonathan
Matthew
Simon Tavare University of Southern California
Chris Tyler-Smith Sanger Institute
Jun Wang Beijing Genomics Institute
David Wheeler Baylor College of Medicine
Hongkun Zheng Beijing Genomics Institute

www.1000genomes.org

Sciencexpress Report

Association of Trypanolytic ApoL1 Variants with Kidney Disease in African-Americans

Giulio Genovese,[1,2]* David J. Friedman,[1,3]* Michael D. Ross,[4] Laurence Lecordier,[5] Pierrick Uzureau,[5] Barry I. Freedman,[6] Donald W. Bowden,[7,8,9,10,11,12] Carl D. Langefeld,[9,10,11,12] Taras K. Oleksyk,[13] Andrea Uscinski Knob,[4] Andrea J. Bernhardy,[1] Pamela J. Hicks,[7,8,9,10,11,12] George W. Nelson,[15] Benoit Vanhollebeke,[5] Cheryl A. Winkler,[14] Jeffrey B. Kopp,[15] Etienne Pays,[5]† Martin R. Pollak[1,16]†

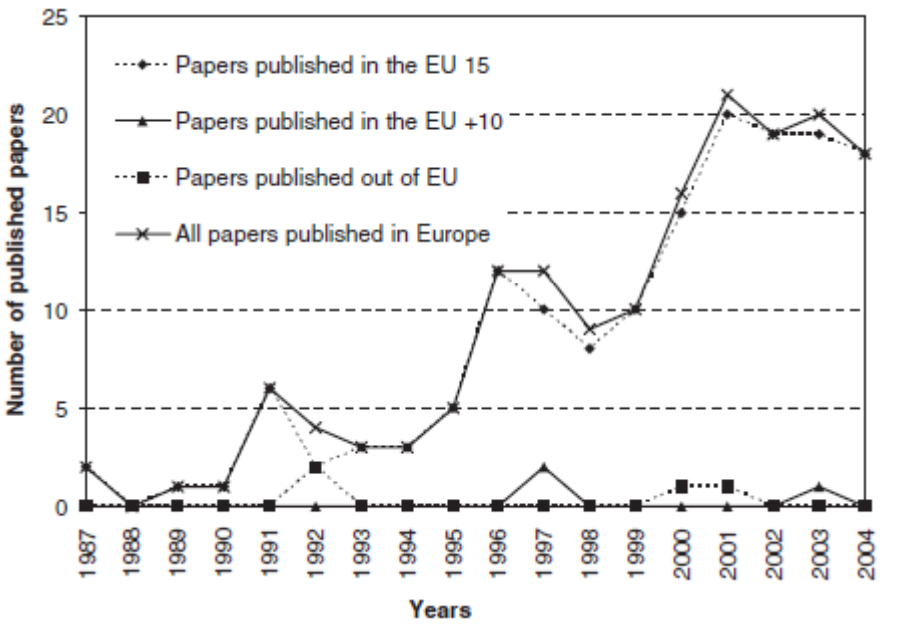
Hum Genet
DOI 10.1007/s00439-010-0861-0

SHORT REPORT

Missense mutations in the *APOL1* gene are highly associated with end stage kidney disease risk previously attributed to the *MYH9* gene

Shay Tzur · Saharon Rosset · Revital Shemer · Guennady Yudkovsky · Sara Selig · Ayele Tarekegn · Endashaw Bekele · Neil Bradman · Walter G. Wasser · Doron M. Behar · Karl Skorecki

Publication rate



European Journal of Public Health, Vol. 17, Supplement 1, 2007

doi:10.1093/eurpub/ckm069

Genetic epidemiology literature in Europe—an overview

R. Ádány, Z. Pocsai

mtDNA – The first steps

RFLP based analysis

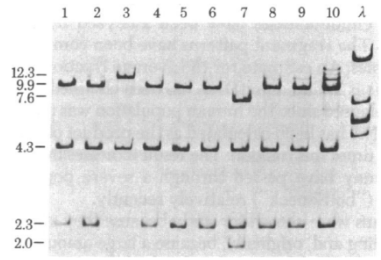


FIG. 1. A 1.2% agarose gel electrophoresis of *Hpa* I digests of human mtDNA samples 1–10. The band patterns of samples 3 and 7 differ from those of the other samples. The DNA fragment sizes, shown on the left in bp × 10^{-3}, were estimated from a calibration curve using *Eco*RI-digested λ DNA (above) and *Hin*dIII-digested PM2 DNA (not shown) as size standards. Faint bands are due to incomplete digestion.

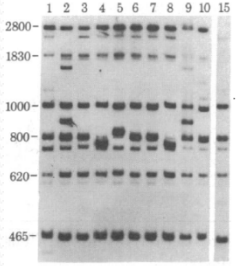


FIG. 2. A 3.5% polyacrylamide gel of *Mbo* I digests of human mtDNA samples 1–10 and 15. All atypical fragments observed were ≥450 bp, hence the portion of the gel containing smaller fragments is not shown. The band pattern shown by samples 3, 6, and 7 was typical for the majority of the 21 samples analyzed. Samples 2 and 9 show identical atypical patterns, as do samples 4 and 8. The DNA fragment sizes, in bp, are shown on the left. The faint bands are completely digested fragments that, because of the labeling method, do not label well.

Table 3. Distribution of restriction endonuclease polymorphisms in mtDNAs from 21 humans

Endonuclease	Morph in sample																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
A. *Bam*HI		2							2												
B. *Hpa* I			2				3								3						
C. *Kpn* I										2											
D. *Hin*cII			2																		
E. *Alu* I	2			3, 4	5, 6	2	7	3		6	2		2		7, 8			9	2		
F. *Hae* III				2	3			2		3	3	3, 4		5		4	6, 7	3		8	9
G. *Hha* I	2		3						3		4										5
H. *Hin*fI				2	3		3, 4, 5	2		3				3, 6, 7	3, 4		3	8		9	3, 10
I. *Hpa* II		2, 3					4						5								
J. *Mbo* I	2	3		4	5			4	3	6					7						
K. *Taq* I							2, 3							4	5				6		7

The samples were monomorphic for the following enzymes: *Eco*RI, *Hin*dIII, *Xba* I, *Pst* I, *Pvu* II, *Sac* I, and *Xho* I (see ref. 10 and text). The characteristics of the numbered polymorphisms are as explained in Table 2 and the text. A blank indicates morph no. 1.

Proc. Natl. Acad. Sci. USA
Vol. 77, No. 6, pp. 3605–3609, June 1980
Genetics

Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis

(human evolution/intraspecific variation/population genetics)

WESLEY M. BROWN

Department of Biochemistry, University of California, Berkeley, California 94720

Communicated by Ruth Sager, March 27, 1980

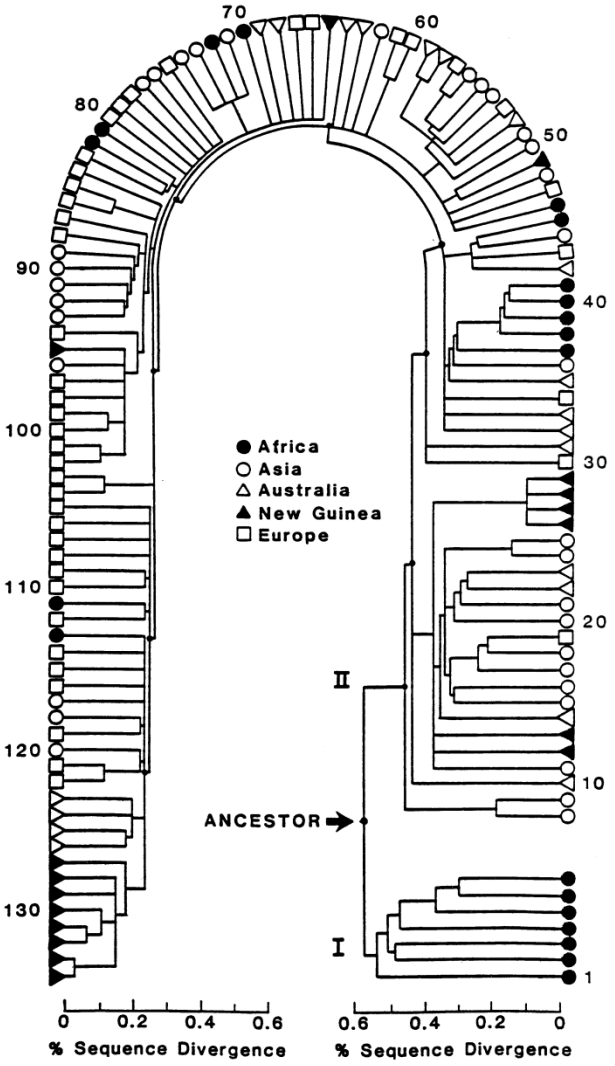
African Eve

Mitochondrial DNA and human evolution

Rebecca L. Cann*, Mark Stoneking & Allan C. Wilson
Department of Biochemistry, University of California, Berkeley, California 94720, USA
* Present address: Department of Genetics, University of Hawaii, Honolulu, Hawaii 96822.

Nature **325**, 31 - 36 (01 January 1987)

Mitochondrial DNAs from 147 people, drawn from five geographic populations have been analysed by restriction mapping. All these mitochondrial DNAs stem from one woman who is postulated to have lived about 200,000 years ago, probably in Africa. All the populations examined except the African population have multiple origins, implying that each area was colonised repeatedly.



Using the HVSI

Table 3

Variants in 82 mtDNA Sequence Haplotypes Occurring More Than Once in Europe and Their Geographical Distribution

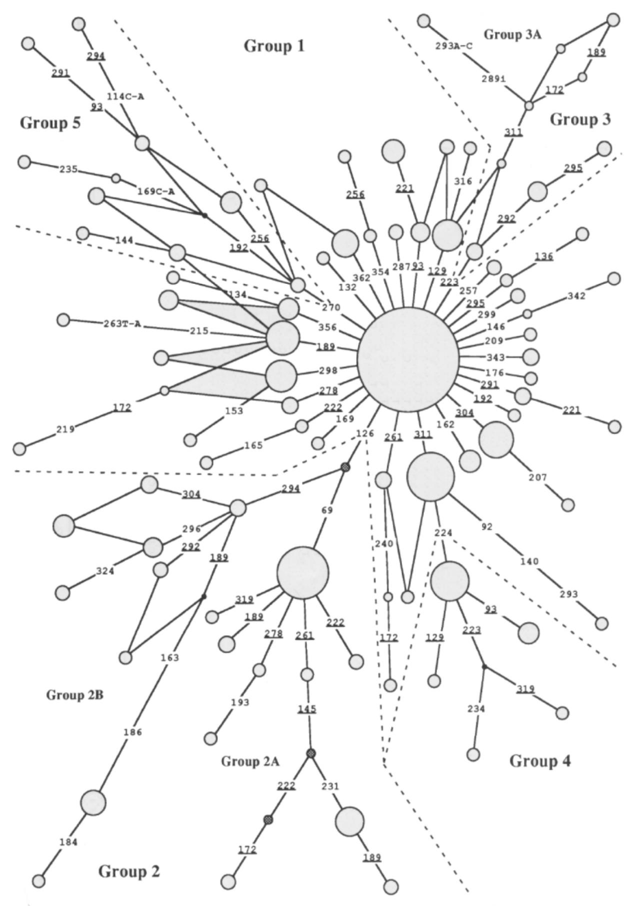
Group and Haplotype	Variants	Geographical Distribution																
	00011111111111111111111122222222222222222222 2 222222223333333333																	
	69912233344445666677888990011222233345566778 8 999999990111244556																	
	92346924604563235926469237959123414506713087 9 123456894169423462	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	All	Europe
	i																	
1:																		
1	CTTCTGACTTTGAGAAACTCCCTCCATAACCCTTCAACCCTCCC ... CCACCCTATTAGTTACTT	.	5	6	11	5	14	8	.	23	13	17	16	4	8	14	144	139
2	?.C...	.	.	.	.	.	.	.	.	.	2	1	2	.	.	.	5	5
3	?....A......................................	.	.	.	.	.	1	2	.	.	1	1	4	.	.	3	12	12
4	?.....G.....................................	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	2	2
5	?.............G.............................	.	.	2	.	.	1	.	1	.	.	.	.	.	.	1	5	5
6	?................T..........................	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	2	2
7	?..................T........................	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	2	2
8	?.....................C.....................	.	.	1	1	1	8	.	.	.	2	1	.	.	.	2	16	16
9	?......................T....................	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	2	2
10	?.........................C.................	.	.	.	.	.	.	.	.	1	.	.	.	.	.	1	2	2
11	?.............................T.............	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	2	2
12	?.....................................T.....	.	.	.	.	.	.	.	.	.	.	3	.	.	.	.	3	3
13	?......................................T....	.	.	.	2	1	1	.	.	.	.	.	.	.	.	.	4	4
14	?...T.	.	.	.	.	.	.	.	.	3	.	.	1	.	.	.	4	4
15	?..T	.	.	.	3	.	.	.	.	.	.	.	.	.	.	.	3	3
16	?... ... T.................	.	.	.	.	.	.	.	.	.	1	.	2	.	.	1	4	4
17	?... ... .T.............	.	.	.	.	1	.	1	.	.	.	.	.	.	.	.	2	2
18	?... ... .C...........	.	.	.	3	1	1	1	.	2	1	.	1	1	1	2	14	14
19	?... ... .G..........	.	.	.	.	.	.	.	1	.	.	.	.	.	.	1	2	2
20	?... ... .C.........	.	1	.	2	1	2	.	1	7	1	1	1	.	.	2	19	18
21	?... ... .C........	2	.	1	9	3	3	2	.	2	2	3	.	2	.	.	29	27

Am. J. Hum. Genet. 59:185–203, 1996

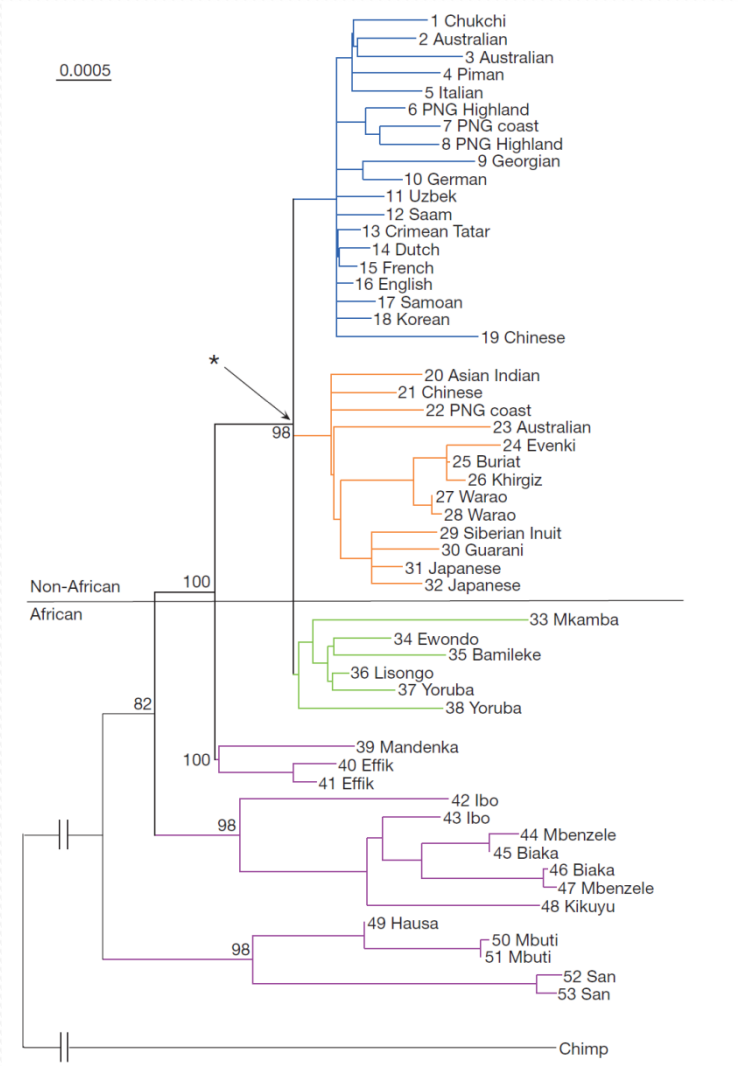
Paleolithic and Neolithic Lineages in the European Mitochondrial Gene Pool

Martin Richards,[1] Helena Côrte-Real,[1] Peter Forster,[3] Vincent Macaulay,[1]
Hilde Wilkinson-Herbots,[5] Andrew Demaine,[6] Surinda Papiha,[7] Robert Hedges,[2]
Hans-Jürgen Bandelt,[4] and Bryan Sykes[1]

[1]Department of Cellular Science, Institute of Molecular Medicine, and [2]Research Laboratory for Archaeology, University of Oxford, Oxford; [3]Heinrich-Pette-Institut für Experimentelle Virologie und Immunologie and [4]Mathematisches Seminar, University of Hamburg, Hamburg; [5]Department of Statistical Science, University College London, London; [6]Department of Medicine, University of Plymouth, Plymouth; and [7]Division of Human Genetics, University of Newcastle, Newcastle-upon-Tyne



Complete mtDNA sequences



Mitochondrial genome variation and the origin of modern humans

Max Ingman*, Henrik Kaessmann†, Svante Pääbo† & Ulf Gyllensten*

** Department of Genetics and Pathology, Section of Medical Genetics, Rudbeck Laboratory, University of Uppsala, S-751 85 Uppsala, Sweden*
† Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22, D-04103 Leipzig, Germany

Pygmies and Bantus

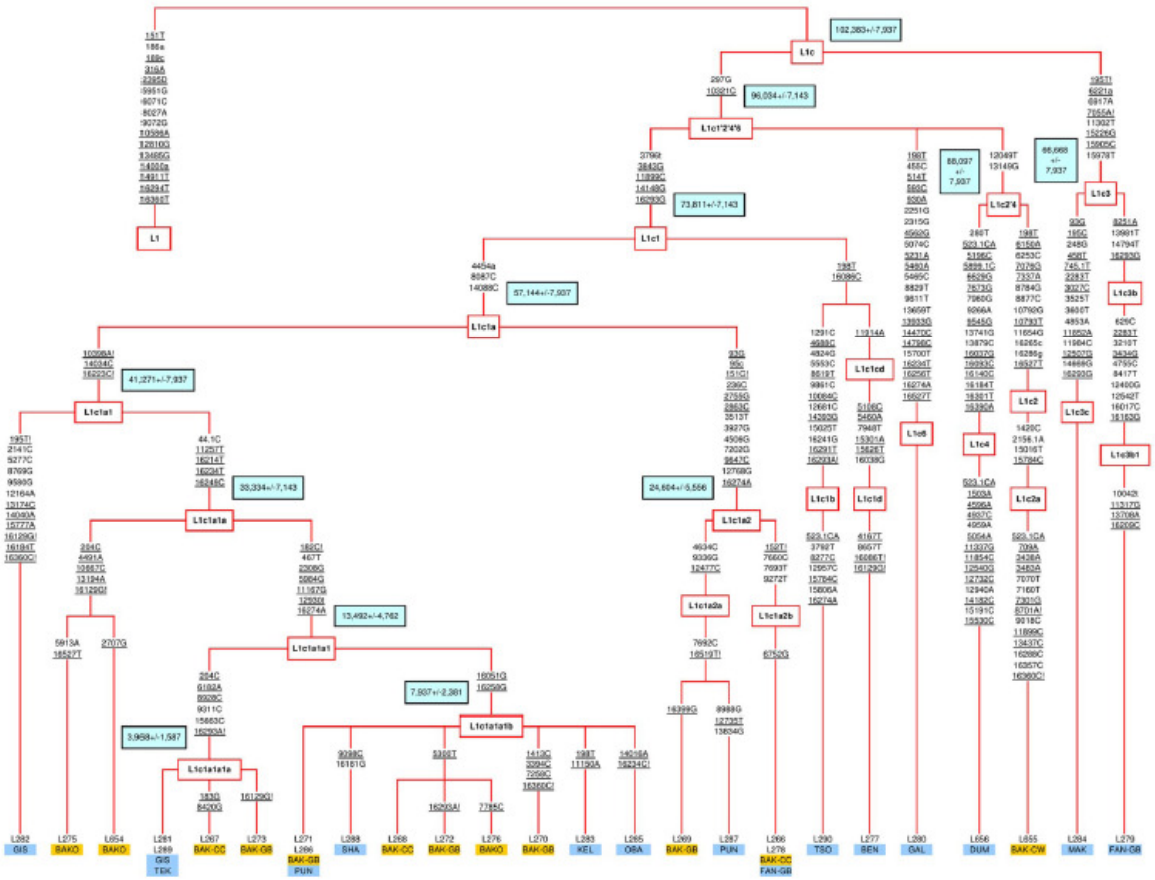


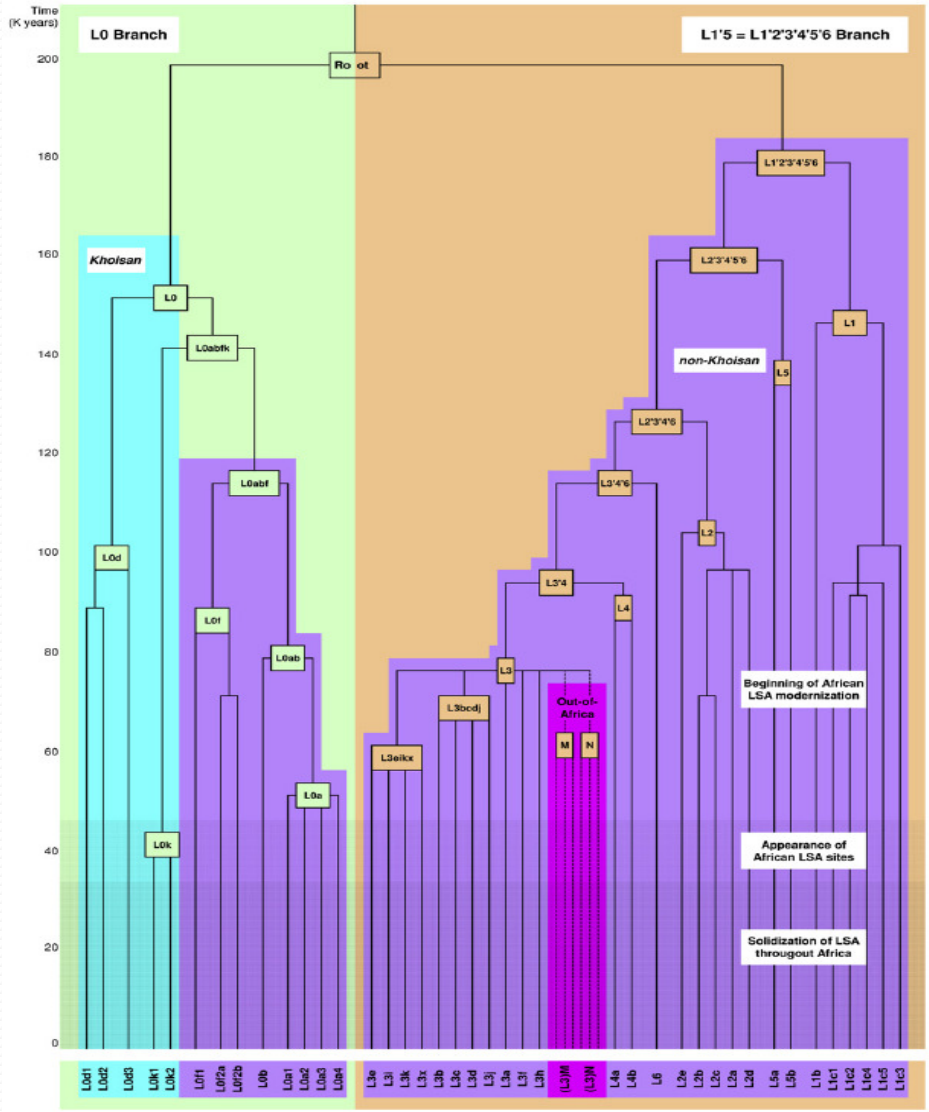
Fig. 1. Phylogenetic tree of complete mtDNA sequences belonging to haplogroup L1c. The tree is rooted on Hg L1 and shows subhaplogroup affiliations. Mutations are shown on the branches. Transitions are labeled in uppercase letters, transversions are indicated in lowercase letters, deletions are indicated by a "d" after the deleted nucleotide position, and insertions are indicated by a dot, followed by the number and type of inserted nucleotides. Underlined nucleotide positions occur at least twice in the tree. The exclamation mark (!) at the end of a nucleotide position denotes a reversion to the ancestral state in the relative pathway from the rCRS (36). Individuals highlighted in orange correspond to PHG and those in blue to Bantu-speaking AGR. Population codes for each individual are as in Table 1. Coalescence age estimates for the main subhaplogroups are also reported.

Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter–gatherers and Bantu-speaking farmers

Lluis Quintana-Murci[a,b], Hélène Quach[a], Christine Harmant[a], Francesca Luca[a], Blandine Massonnet[a], Etienne Patin[a], Lucas Sica[c], Patrick Mouguiama-Daouda[d], David Comas[e], Shay Tzur[f], Oleg Balanovsky[g], Kenneth K. Kidd[h], Judith R. Kidd[h], Lolke van der Veen[d], Jean-Marie Hombert[d], Antoine Gessain[i], Paul Verdu[j], Alain Froment[j], Serge Bahuchet[j], Evelyne Heyer[j], Jean Dausset[a,k], Antonio Salas[l], and Doron M. Behar[f]

1596–1601 | PNAS | February 5, 2008 | vol. 105 | no. 5

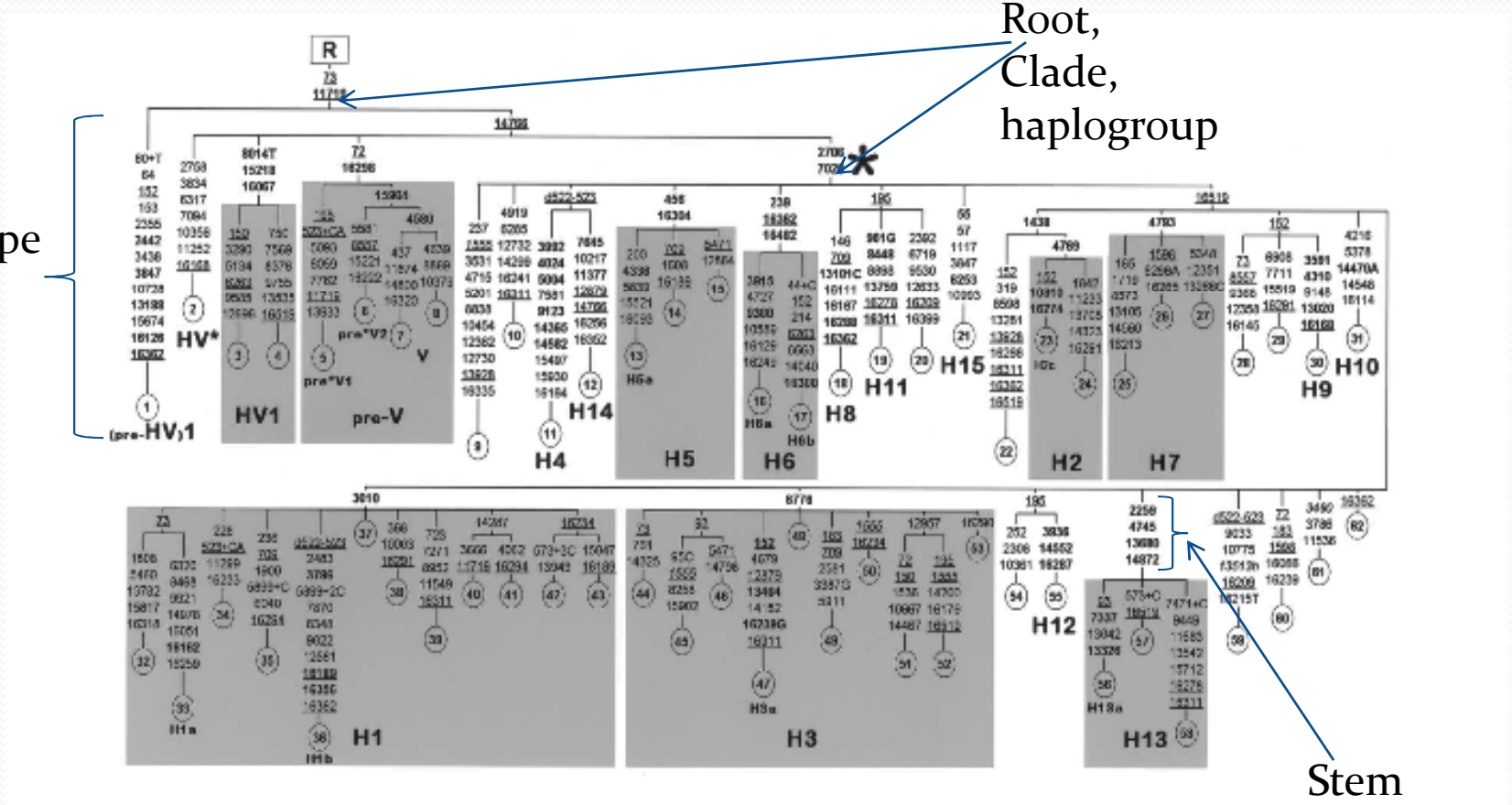
The Root Position



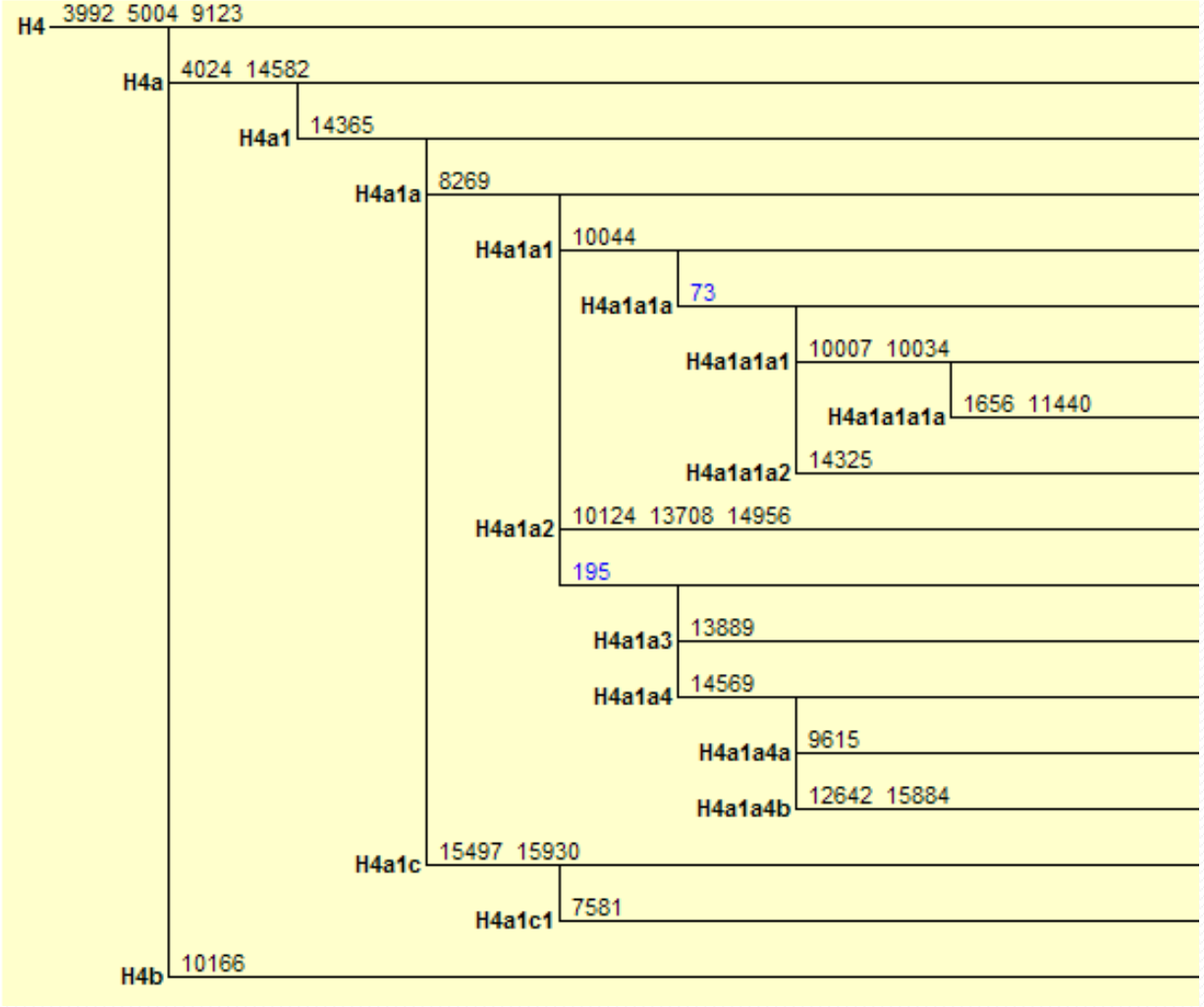
The Dawn of Human Matrilineal Diversity

Doron M. Behar,[1,13,*] Richard Villems,[2,13] Himla Soodyall,[3] Jason Blue-Smith,[4] Luisa Pereira,[5,6] Ene Metspalu,[2] Rosaria Scozzari,[7] Heeran Makkan,[3] Shay Tzur,[1] David Comas,[8] Jaume Bertranpetit,[8] Lluis Quintana-Murci,[9] Chris Tyler-Smith,[10] R. Spencer Wells,[4] Saharon Rosset,[11,12] and The Genographic Consortium[14]

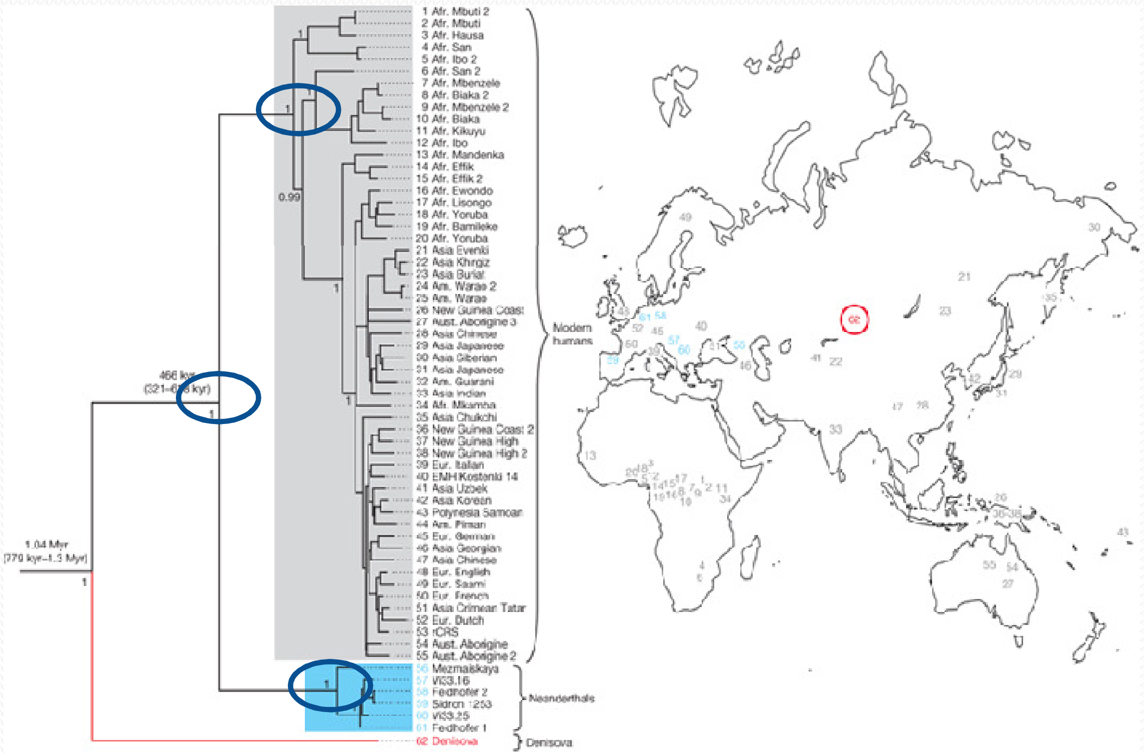
Revising the terms



Refining H4


H4
3992 5004 9123
H4a
4024 14582
H4a1
14365
H4a1a
8269
H4a1a1
10044
H4a1a1a
73
H4a1a1a1
10007 10034
H4a1a1a1a
1656 11440
H4a1a1a2
14325
H4a1a2
10124 13708 14956
195
H4a1a3
13889
H4a1a4
14569
H4a1a4a
9615
H4a1a4b
12642 15884
H4a1c
15497 15930
H4a1c1
7581
H4b
10166

Phylogenetic tree of Homo complete mtDNAs



J Krause *et al. Nature* **464**, 894-897 (2010)

NRY – The first steps

Ann. Hum. Genet. (1993), **57**, 55–64
Printed in Great Britain

The common, Near-Eastern origin of Ashkenazi and Sephardi Jews supported by Y-chromosome similarity

A. S. SANTACHIARA BENERECETTI[1,2], O. SEMINO[2], G. PASSARINO[3], A. TORRONI[2*], R. BRDICKA[4], M. FELLOUS[5] AND G. MODIANO[6]

[1] *Dipartimento di Biologia Cellulare, Università della Calabria, Cosenza, Italy*
[2] *Dipartimento di Genetica e Microbiologia 'A. Buzzati Traverso', Università di Pavia, Via Abbiategrasso 207, 27100 Pavia, Italy*
[3] *ISMEC CNR, Cosenza, Italy*
[4] *University of Prague, Czechoslovakia*
[5] *Institut Pasteur, Paris, France*
[6] *Dipartimento di Biologia, Università 'Tor Vergata', Roma, Italy*

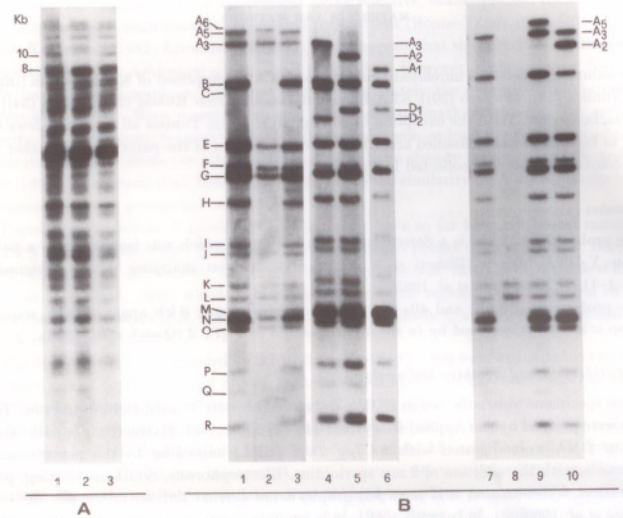


Fig. 1. Electrophoretic patterns of human male DNAs digested with *Taq*I and hybridized: (*A*) by p12f2 (allelic fragments of 10 and 8 kb). The 10 kb band in lane 1 and 3 is not present in lane 2 where the 8 kb fragment, independent of this polymorphism, appears to be reinforced. (*B*) by p49a,f. With the exception of K and L, which are also present in female patterns (lane 8), all of the numerous bands identified by these probes are Y-specific (Ngo *et al.* 1986). Lanes 1, 3 and 9, haplotype 49 ($A_{3,5}C_6D_6F_1I_1$); lane 2, haplotype 62 ($A_{3,6}C_6D_6F_1I_6$, with the absence of the BHPR bands); lane 4, haplotype 15 ($A_3C_1D_2F_1I_1$); lane 5, haplotype 8 ($A_2C_6D_1F_1I_1$); lane 6, haplotype 42 ($A_1C_6D_1F_1I_1$); lane 7, haplotype 11 ($A_2C_6D_6F_1I_1$); lane 10, haplotype 52 ($A_{2,3}CO_6D_6F_1I_1$).

In conclusion, the analysis of these Y-specific polymorphisms has given further impressive support to the common origin of Sephardim and Ashkenazim, has indicated a low contribution of foreign Y-chromosomes to the Ashkenazi gene pool but has also shown a striking similarity between Jews and people from the same land of origin. Is this surprising? Didn't Ishmael and Isaac each receive a copy of Abraham's Y-chromosome?

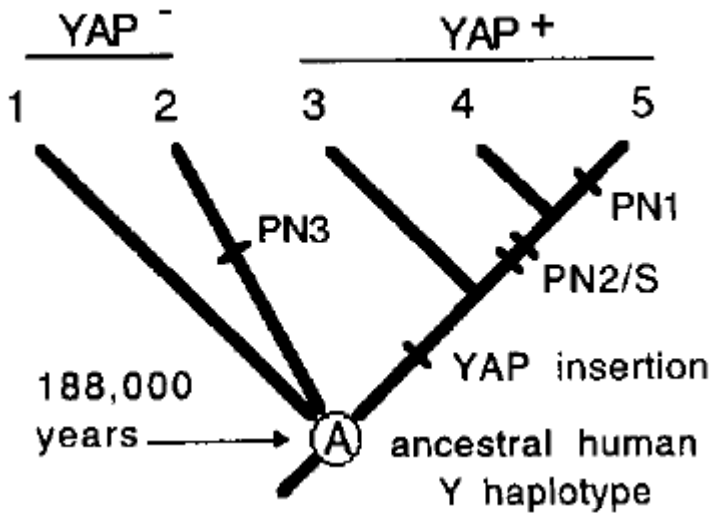
Adam

A recent common ancestry for human Y chromosomes

Michael F. Hammer

Laboratory of Molecular Systematics and Evolution, Biosciences West, University of Arizona, Tucson, Arizona 85721, USA

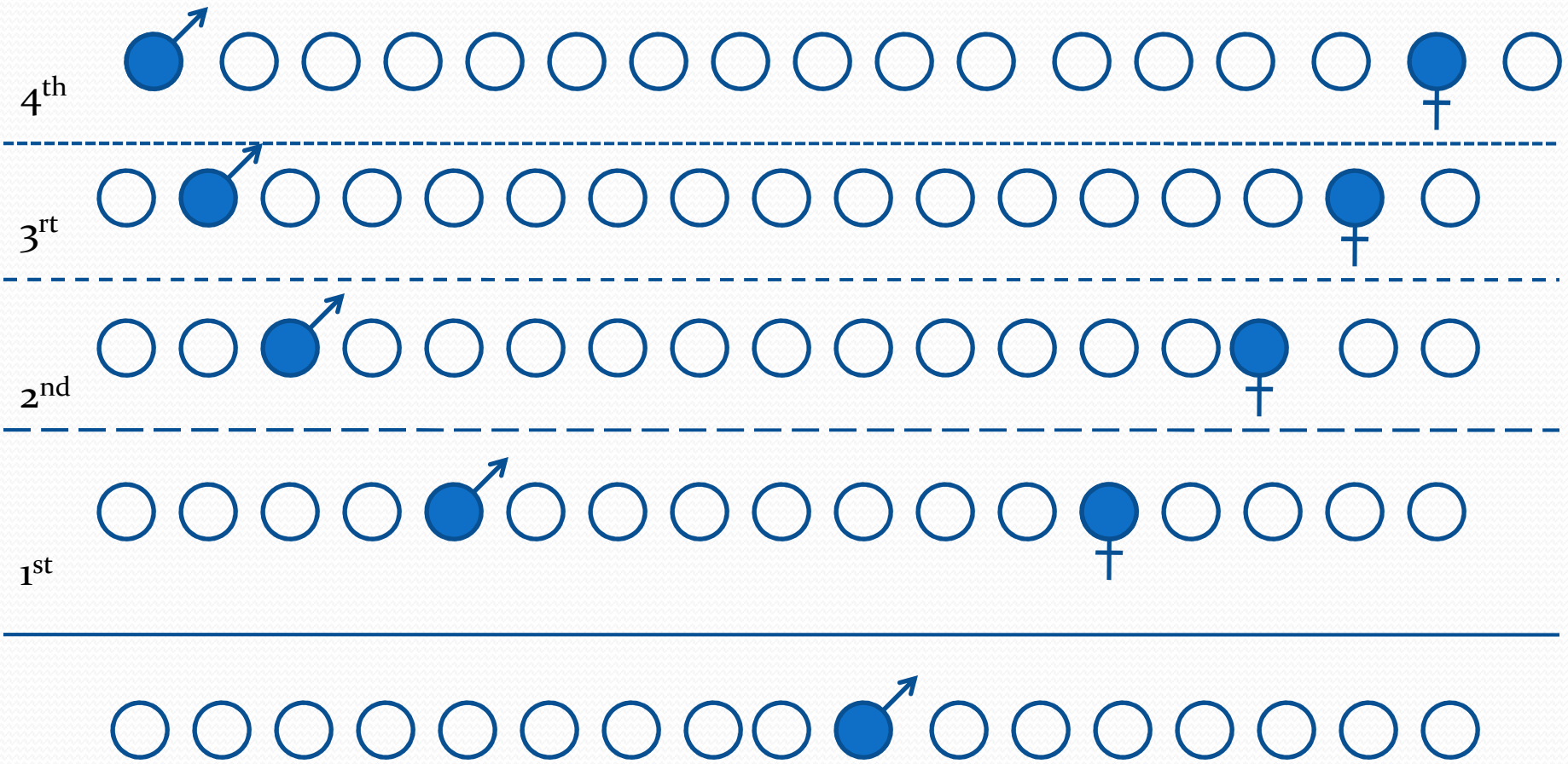
NATURE · VOL 378 · 23 NOVEMBER 1995



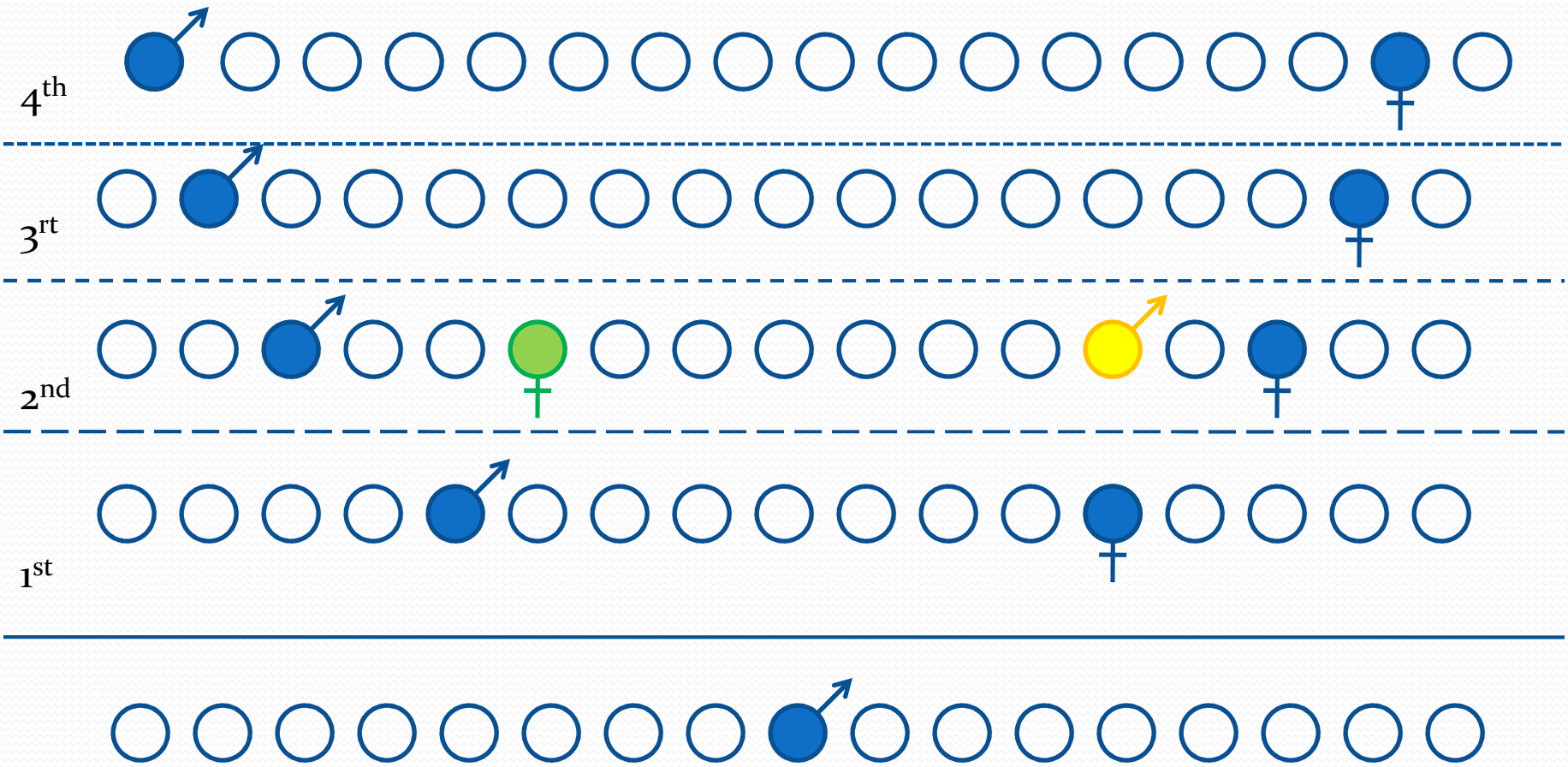
distribution[11]). This age agrees with estimates of the ancestral human mtDNA molecule[12,13]. An implication of these results (Fig. 2) is that the current distribution of Y chromosomes carrying the YAP element[14,15] reflects migrations of male *H. sapiens* from Africa to Europe and Japan.

A coherent picture now emerges that places both our common ancestral Y chromosome and mitochondrial DNA molecule in the late middle Pleistocene only slightly before the hypothesized

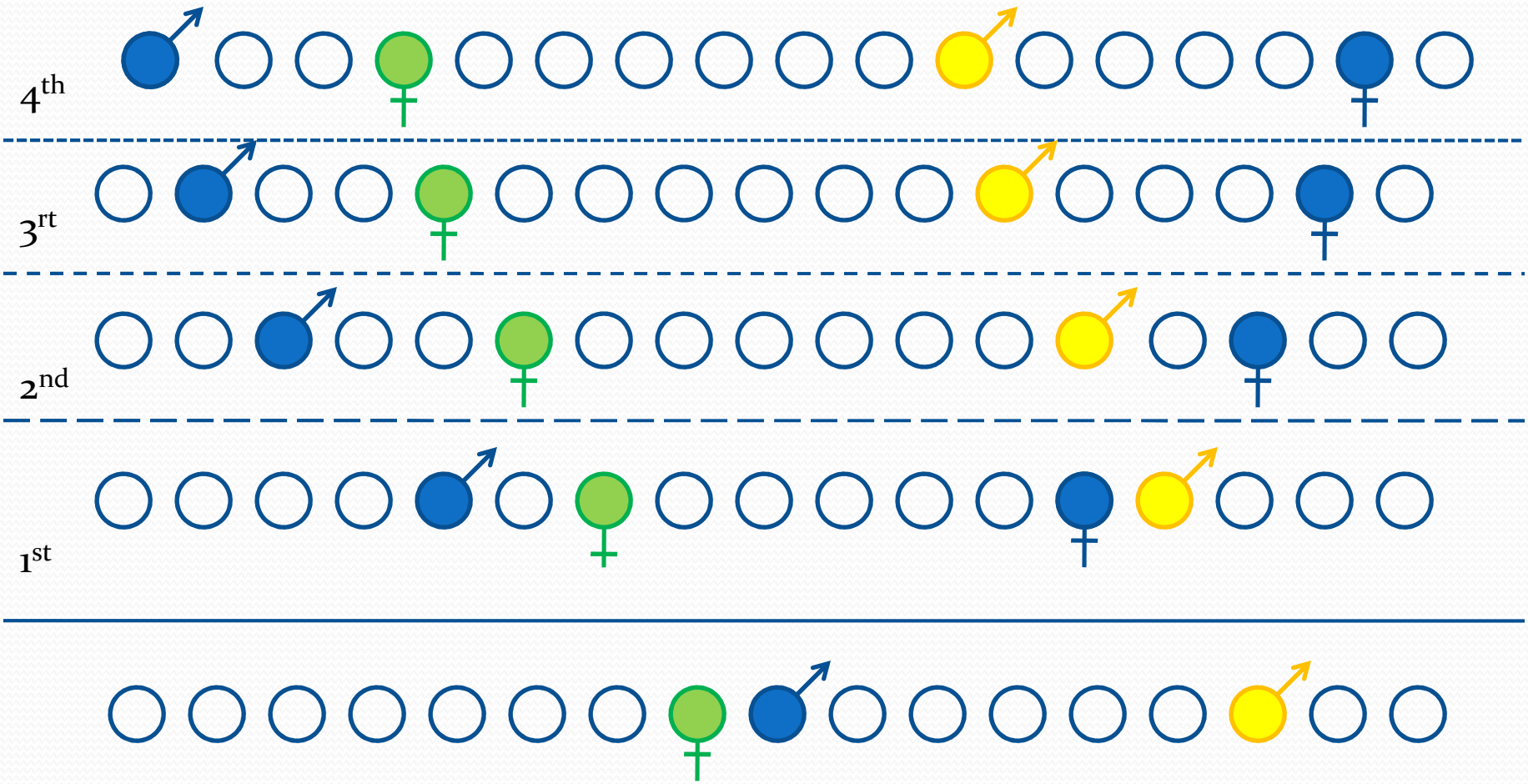
Ancestry at the population level



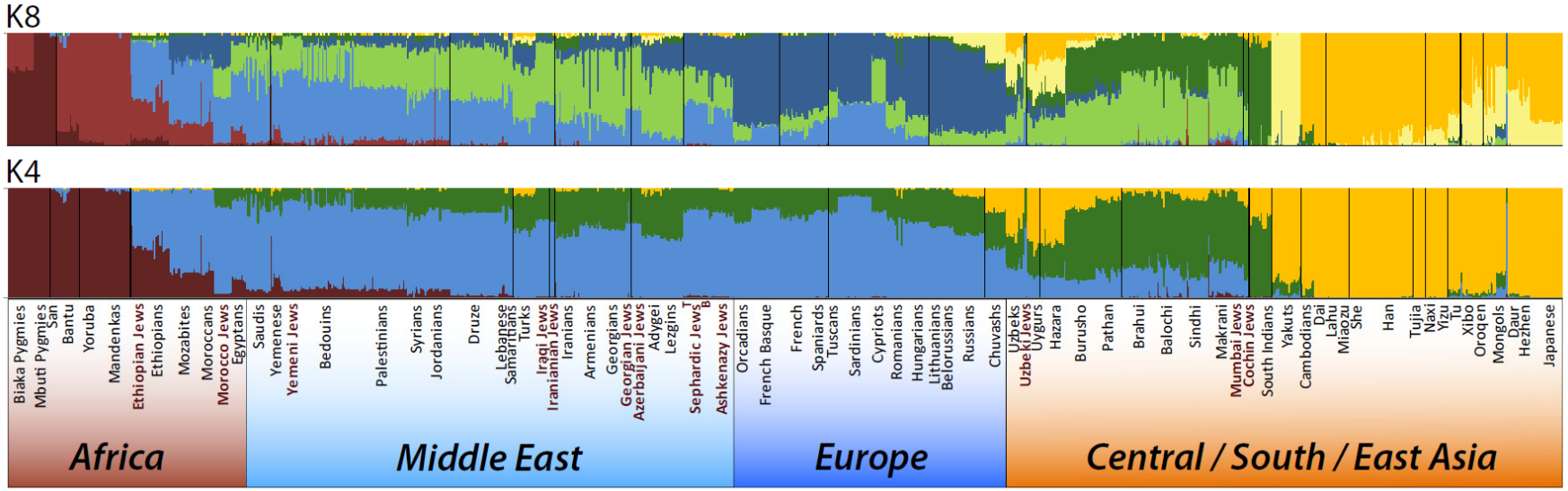
Ancestry at the population level



Ancestry at the population level



Admixture analysis



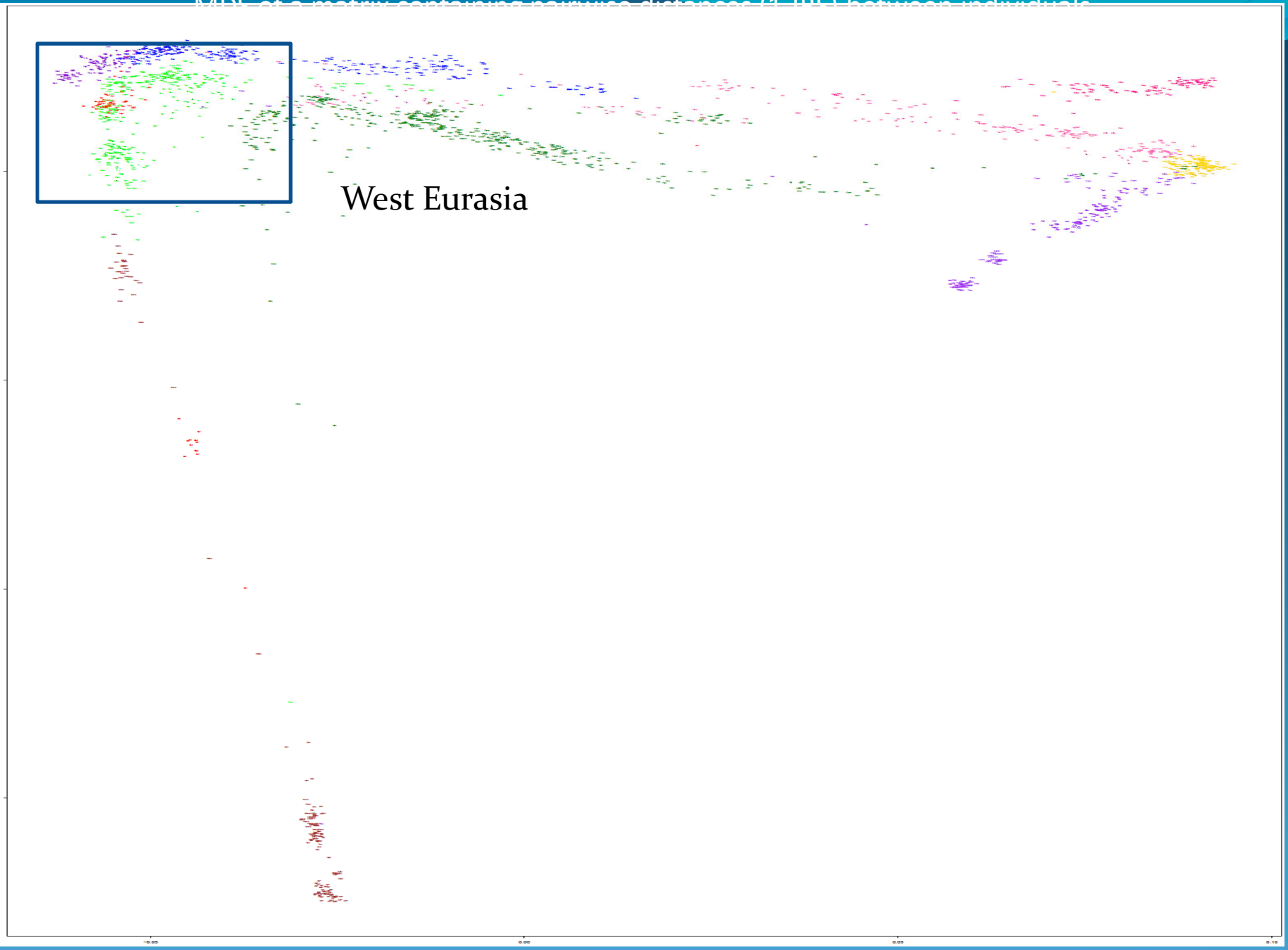
doi:10.1038/nature09103 nature

LETTERS

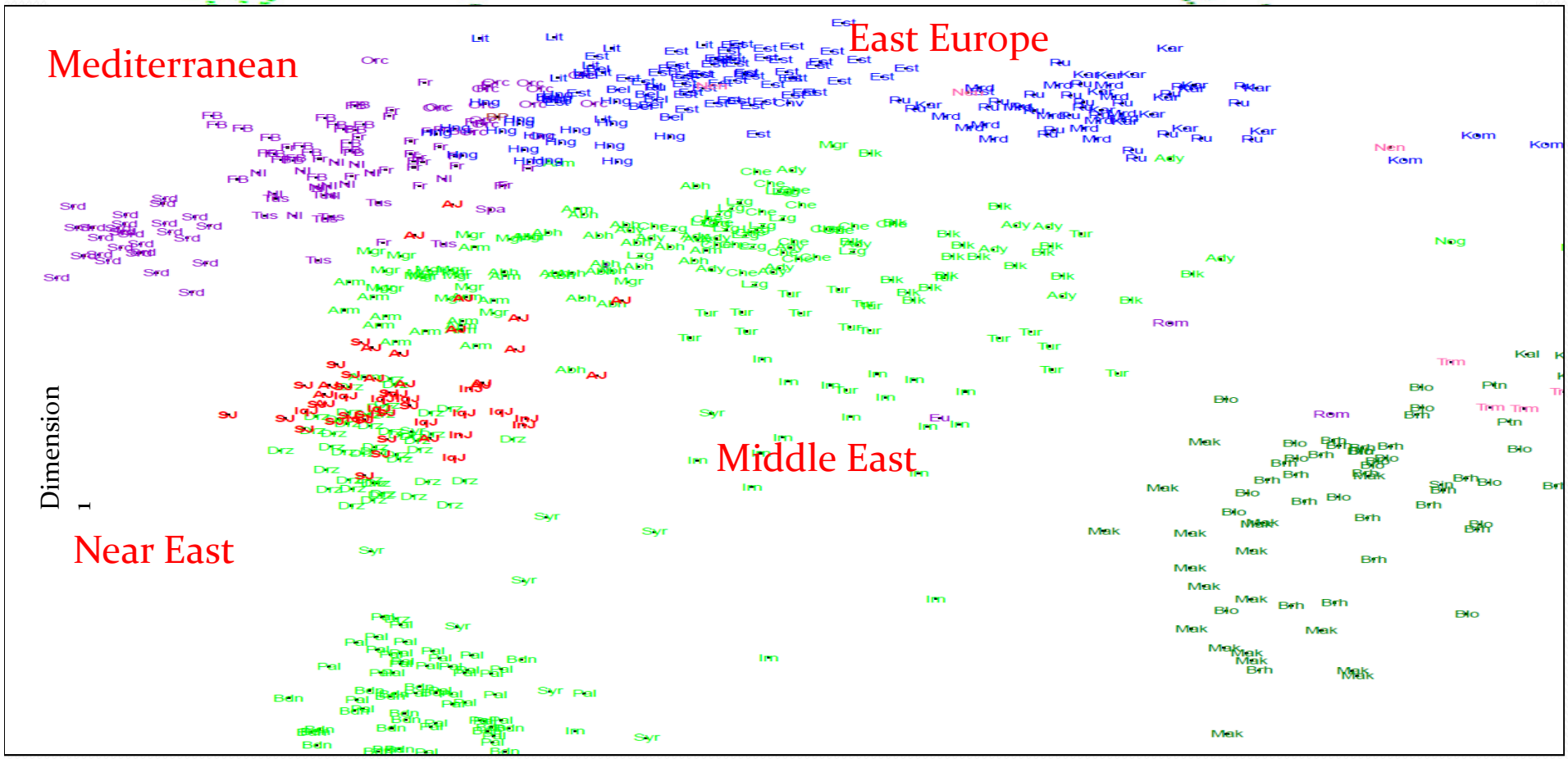
The genome-wide structure of the Jewish people

Doron M. Behar[1,2*], Bayazit Yunusbayev[2,3*], Mait Metspalu[2*], Ene Metspalu[2], Saharon Rosset[4], Jüri Parik[2], Siiri Rootsi[2], Gyaneshwer Chaubey[2], Ildus Kutuev[2,3], Guennady Yudkovsky[1,3], Elza K. Khusnutdinova[3], Oleg Balanovsky[6], Ornella Semino[7], Luisa Pereira[8,9], David Comas[10], David Gurwitz[11], Batsheva Bonne-Tamir[11], Tudor Parfitt[12], Michael F. Hammer[13], Karl Skorecki[1,5] & Richard Villems[2]

Principal Components Analysis


West Eurasia
Dimension 2

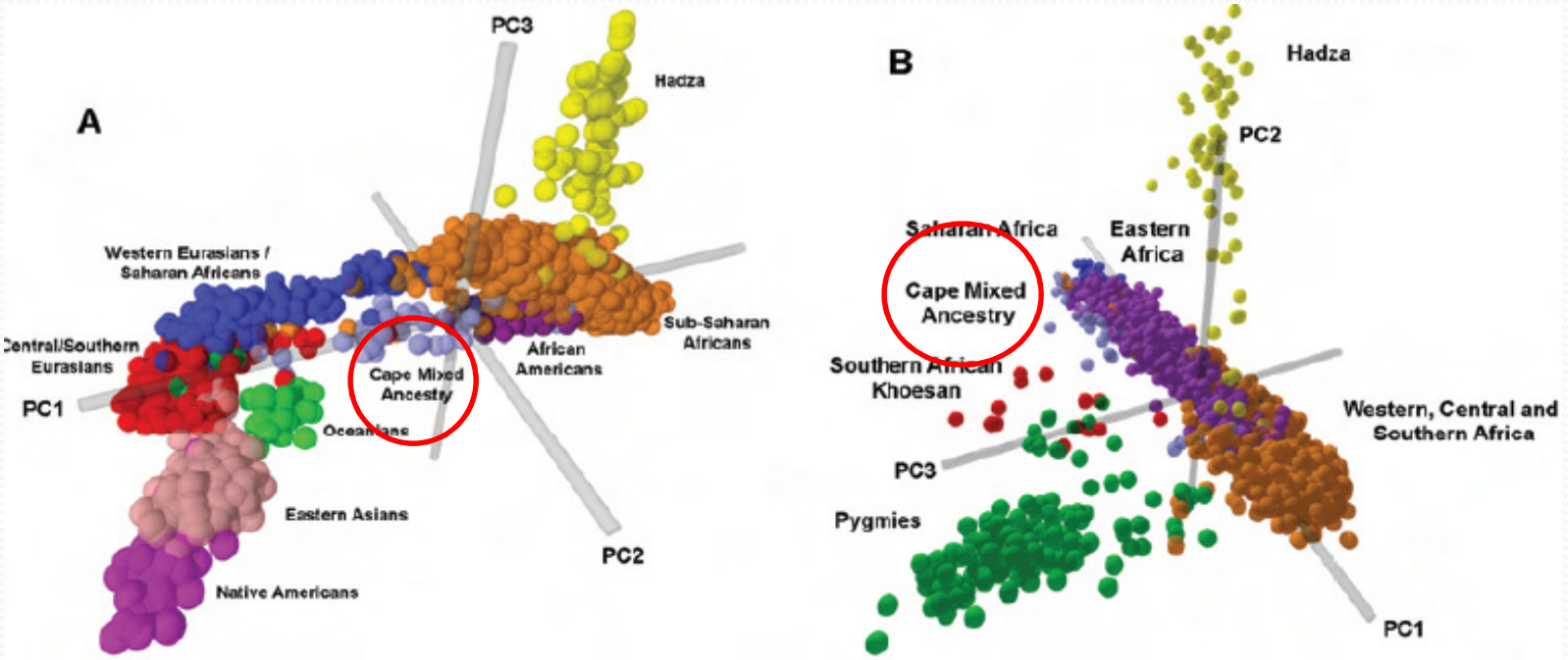
MDS of a matrix containing pairwise distances (1-IBS) between individuals



The Cape Mixed Ancestry population

- Commonly self identified as South African “Cape Coloured”
- Established later than 1652
- Make up 9% of the South African population
- Historical records suggest the fusion of indigenous Khoi and San, tribal Bantu-speaking populations, European settlers and slaves’ descendants from Java, India, Mozambique and Madagascar

Cape, Biparental level



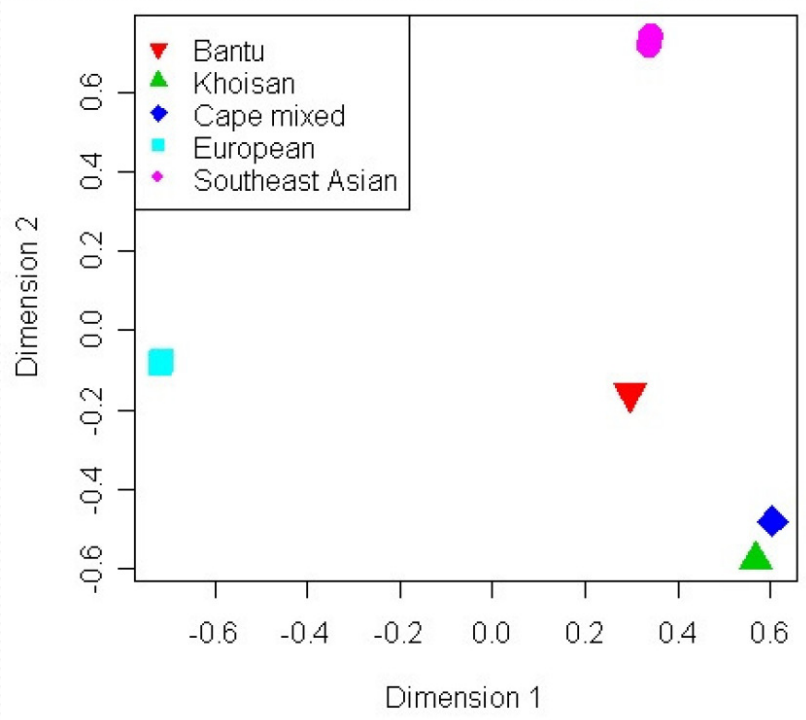
The Genetic Structure and History of Africans and African Americans

Sarah A. Tishkoff,[1,2]* Floyd A. Reed,[1]†‡ Françoise R. Friedlaender,[3]‡ Christopher Ehret,[4] Alessia Ranciaro,[1,2,5]§ Alain Froment,[6]§ Jibril B. Hirbo,[1,2] Agnes A. Awomoyi,[1]|| Jean-Marie Bodo,[7] Ogobara Doumbo,[8] Muntaser Ibrahim,[9] Abdalla T. Juma,[9] Maritha J. Kotze,[10] Godfrey Lema,[11] Jason H. Moore,[12] Holly Mortensen,[1]¶ Thomas B. Nyambo,[11] Sabah A. Omar,[13] Kweli Powell,[1]# Gideon S. Pretorius,[14] Michael W. Smith,[15] Mahamadou A. Thera,[8] Charles Wambebe,[16] James L. Weber,[17] Scott M. Williams[18]

www.sciencexpress.org / 30 April 2009 / Page 1 / 10.1126/science.1172257

Cape, Uniparental level

Maternal Ancestry

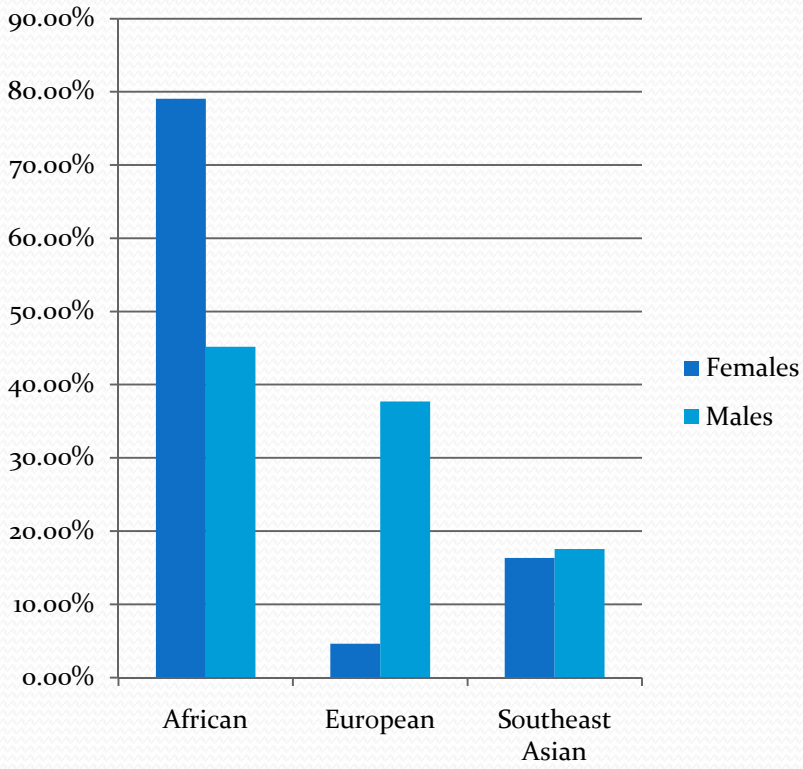


Strong Maternal Khoisan Contribution to the South African Coloured Population: A Case of Gender-Biased Admixture

Lluis Quintana-Murci,[1,2,*] Christine Harmant,[1,2] Hélène Quach,[1,2] Oleg Balanovsky,[3] Valery Zaporozhchenko,[3] Connie Bormans,[4] Paul D. van Helden,[5] Eileen G. Hoal,[5] and Doron M. Behar[1,6]

The American Journal of Human Genetics *86*, 611–620, April 9, 2010

Sexual asymmetry



Biparental vs Uniparental

- Biparental markers contain far larger amount of variation as well as larger population effective size
- Uniparental markers contain far less amount of variation but are superior with respect to the ability to:
  - draw solid hierarchical phylogenies.
  - trace migrations patterns
  - estimate coalescence and expansion ages
  - infer sexually asymmetric parental demographic histories

What have we achieved?

1	2	3	4	5
Critical mass of uniparental and biparental markers	Critical mass of samples	A combined NRY, mtDNA and autosomal approach	The “big picture” is resolved	**Now we need the details!**

The Human Genome Promise

- Molecular medicine
- Energy sources and environmental applications
- Risk assessment
- Bioarchaeology, anthropology, evolution, and human migration
- DNA forensics (identification)
- Agriculture, livestock breeding, and bioprocessing

Molecular Medicine

- *Improved diagnosis of disease*
- *Earlier detection of genetic predispositions to disease*
- *Rational drug design*
- *Gene therapy and control systems for drugs*
- *Pharmacogenomics "custom drugs"*

The Race

- The **Archon X Prize** for Genomics:

  The $10 million prize (US), donated by diamond prospector Stewart Blusson, is to be awarded to "the first Team that can build a device and use it to sequence 100 human genomes within 10 days or less, with an accuracy of no more than one error in every 100,000 bases sequenced, with sequences accurately covering at least 98% of the genome, and at a recurring cost of no more than $10,000 (US) per genome."

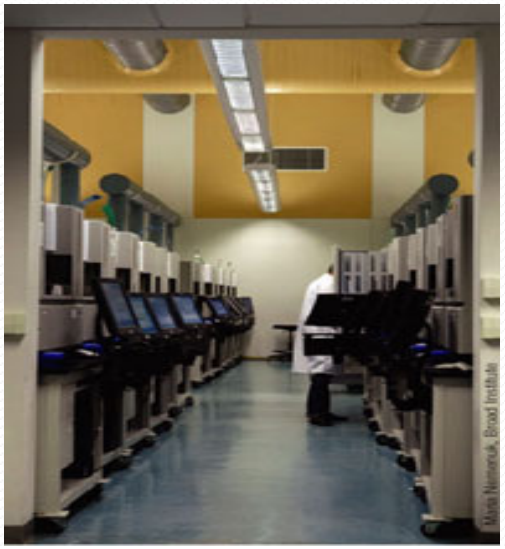
Answering the challenge

- Illumina
- Sequenom
- 454 Life Sciences
- Pacific Biosciences
- Intelligent Bio-Systems
- Genome Corp
- ION Torrent Systems
- Helicos Biosciences
- Compete genomics
- Halcyon Molecular
- IBM
- GE Global

Sequencing

Conventional

- Sanger based technology
- A 30 years old monopoly
- 600-1000 bp per read
- Reaction time is a few hours
- 10,000,000 reactions to sequence the genome



Next generation

- Developed in 2005 by 454 Life Sciences
- A throughput equivalent to 50 Applied Biosystem's 3730XL capillary sequencers at about one-sixth of the cost

Full Genome Sequencing prices

- Conventional sequencing - $10M
- January 2009, Knome complete.Com - $68,500
- June 2009,Illumina - $48,000
- August 2009, Helicos - <$50,000
- November 2009, Complete Genomics - $1,700

Claims

- April 2009, Complete Genomics: “plan to be able to sequence one million full genomes *per year* by 2013”
- June 2009, Illumina: “during the next five years, perhaps markedly sooner, the price point for full genome sequencing will fall from $48,000 to under $1,000”.
- August 2009, Pacific Biosciences: “will sequence 10,000 full genomes by the end of 2010”.
- August 2009, GE Global Research: “is also now in the race to commercialize full genome sequencing as they are currently working on creating a service that will deliver a full genome for $1,000 or less”.
- September 2009, Halcyon Molecular: “will be able to provide full genome sequencing in under 10 minutes for less than $100 per genome”.
- October 2009, IBM: “they were also in the heated race to provide full genome sequencing for under $1,000, with their ultimate goal being able to provide their service for $100 per genome”.

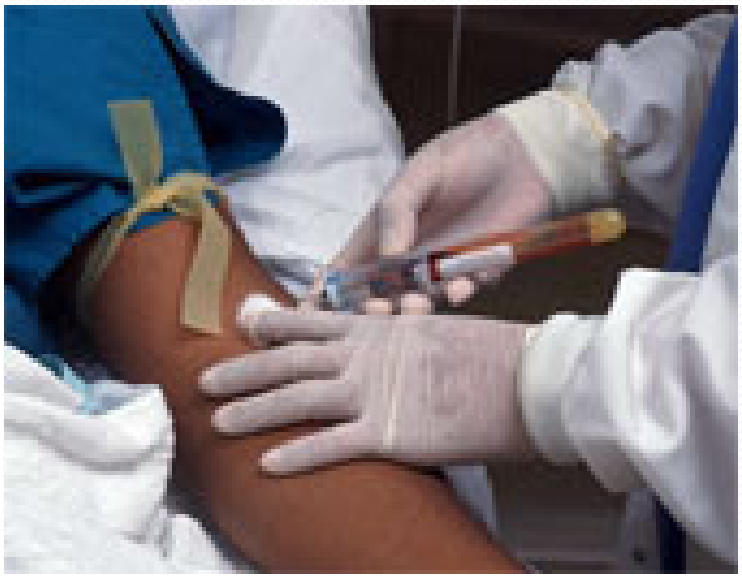
Common Uses of DNA Products

- Ancestry:
  - Parental tracing: Anthropological, Genealogical
  - Admixture analysis
- Forensic:
  - Paternity testing
  - Direct identification of a subject
  - Predicting traits
- Personalized Medicine:
  - Mendelian disorders
  - Strong traits
  - Weak associated traits

The Changed Paradigm

Healthcare providers

- Mendelian disorders
- Paternity testing
- Direct identification of a subject



Direct-to-consumer

- Paternal tracing
- Admixture analysis
- Weak associated traits
- Strong traits
- Mendelian disorders
- Paternity testing

The major players

Industry

- Elimination of technological barriers
- Towards a $1,000 personal full genome sequence
- A thriving direct-to-consumer market

Institutional

- Ethical, Legal and Social issues are under debate
- Minimal penetration of the knowledge to clinical standards of care
- No appropriate education of healthcare providers
- Bioinformatics tools are lacking

Conclusions

- It is possible that in this decade it will be easier, cheaper, more accurate and more informative to genotype genes over specific mutations.
- It is likely that the goal of $1,000 per full genome sequence will be met in this decade.
- The ethical, legal and social issues allowing the *transfer of the information to the public* will be clarified, legislated and approved.

Thank you!