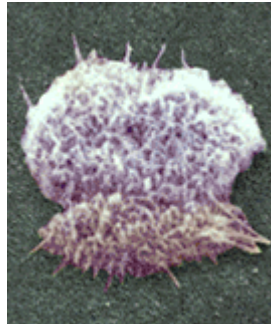


Sukupuolijutut 3

Ihmisen (nisäkkään) seksuaalikromosomi on Y



Y-kromosomi?

Y: male development
Y: genetics
Nature Genetics
Nature Reviews Genetics
 Genome Gateway
 Encyclopaedia of the human genome
 The mouse genome



Links

American Society of Human Genetics
 Baylor University Human Genome Sequencing Center
 European Society of Human Genetics
 Stanford Human Genome Center
 The RIKEN Institute
 The Sanger Institute
 The Whitehead Centre for Genome Research
 Washington University Genome Sequencing Center

NPG Subject areas

Access material from all our publications in your subject area:

Biotechnology



The Y chromosome — with the genes to make a man — has been sequenced. Often regarded as a genetic wasteland, the sequence reveals that we may have underestimated its powers. Here, *Nature* presents the research, as well as news, reviews and analysis. As with all of *Nature's* genome content, these articles are available free online.

News and Views

Tales of the Y chromosome

H. F. Willard

Determining the sequence of the human Y chromosome presented a daunting challenge to genome researchers. But the task is now done, and the secrets revealed justify the effort. **FREE**

Article

The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes **FREE**

H. Skaletsky *et al.*

Letter

Abundant gene conversion between arms of palindromes in human and ape Y chromosomes **FREE**

S. Rozen *et al.*

Review

The human Y chromosome: an evolutionary marker comes of age **FREE**

Nature Reviews Genetics 4, 598 - 612 (2003)

Nature Science Update

Y chromosome sequence completed **FREE**

DNA readout reveals genetic palindromes safeguard male-defining chromosome.

Y chromosomes rewrite British history **FREE**

Anglo-Saxons' genetic stamp weaker than historians suspected.

archive

All articles here are available free to registered users



NPG is committed to offering scientists the finest outlets in print & online

Providing authors with:

a fair deal

authors retain copyright
 no submission charges

advanced online publication

international exposure

in print & online
 HINARI
 SciDevNet

Find out more today

<http://www.nature.com/nature/focus/ychromosome/index.html>

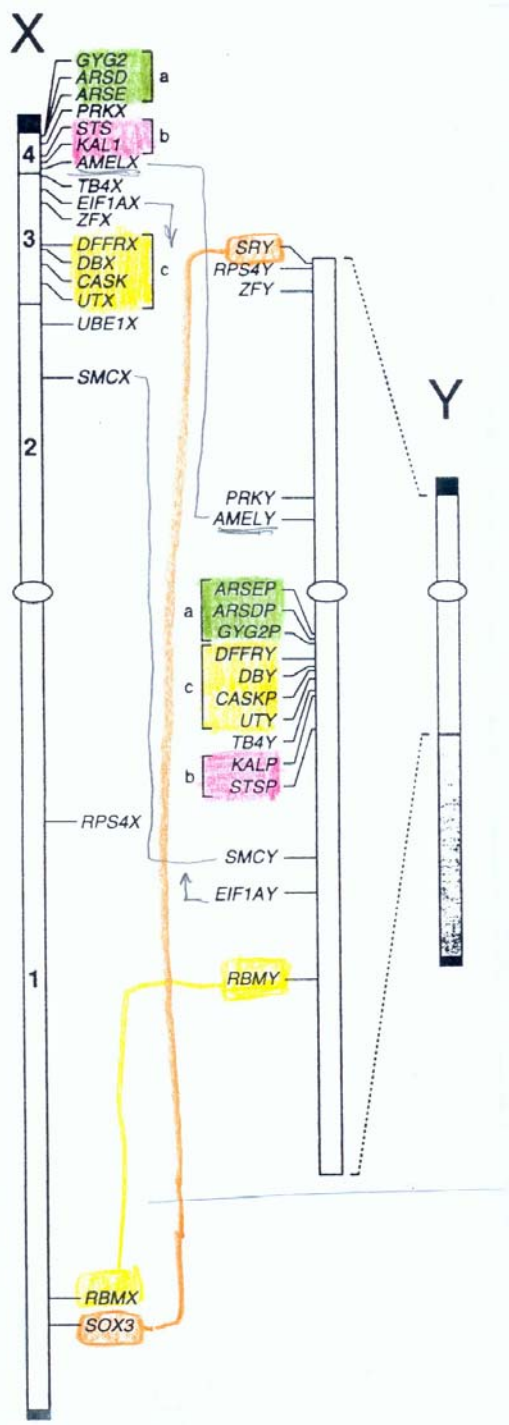
Ihmisen X-kromosomi on normaali kromosomi

Y on erikoinen

- kooltaan vain kolmannes
- geenejä vain sadannes (eli 1%)
- silti konjugoituu vielä X:n kanssa meioosissa

Kehityksessä erottuu neljä hyppäysvaihetta, joissa X/Y rekombinaatio on estynyt tietyllä alueella todennäköisesti inversion takia ja johtanut Y:n rapautumiseen

Miten ihmeessä se voidaan tietää??



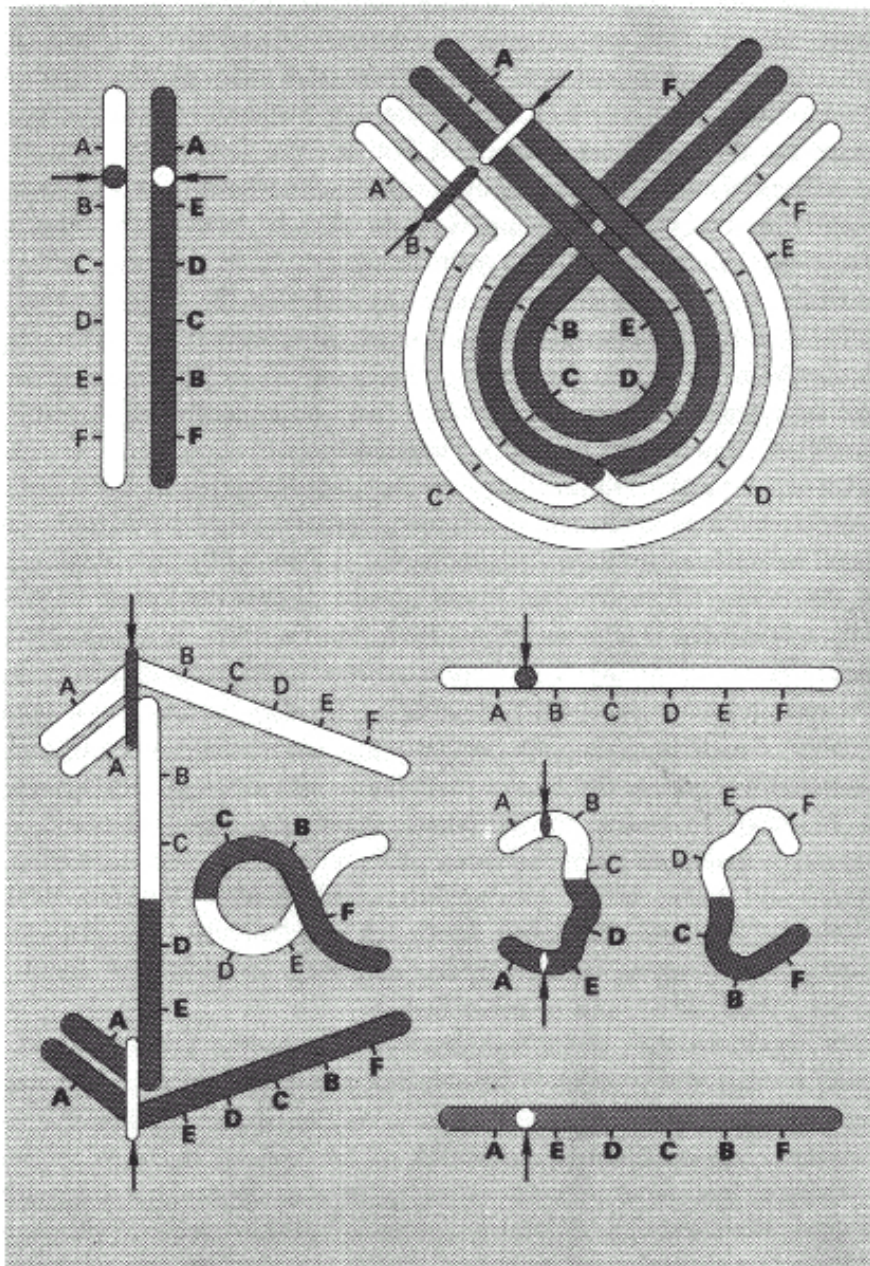
Lahn ja Page etsivät ja löysivät Y-kromosomista yhdeksäntoista X:n kanssa homologista geeniä, joista osa on ns. pseudogeenejä: oikean näköisiä mutta selvästi rappeutuneita (STOP-kodoneita ja frameshiftejä). Jotkut ovat vielä vanhassa ryhmityksessä.

Verrattaessa homologien erilaistumista ne jaettiin 4 ikäkerrostumaan.

Vanhimmassa kerroksessa ovat vanhat alleelit SOX3 ja **SRY** (proteiinissa 29% ero).

Ero on tapahtunut noin 240-320 miljoonaa vuotta sitten, pian lintujen ja nisäkkäiden erottua (silloin tosin molemmat olivat aika lailla *dinosauruksia*)

Y:ssä syntyneet inversiot ovat estäneet rekombinaation X:n kanssa 4 vaiheessa, joista viimeisinkin tapahtui jo 30-50 miljoonaa vuotta sitten.



Muistutus:

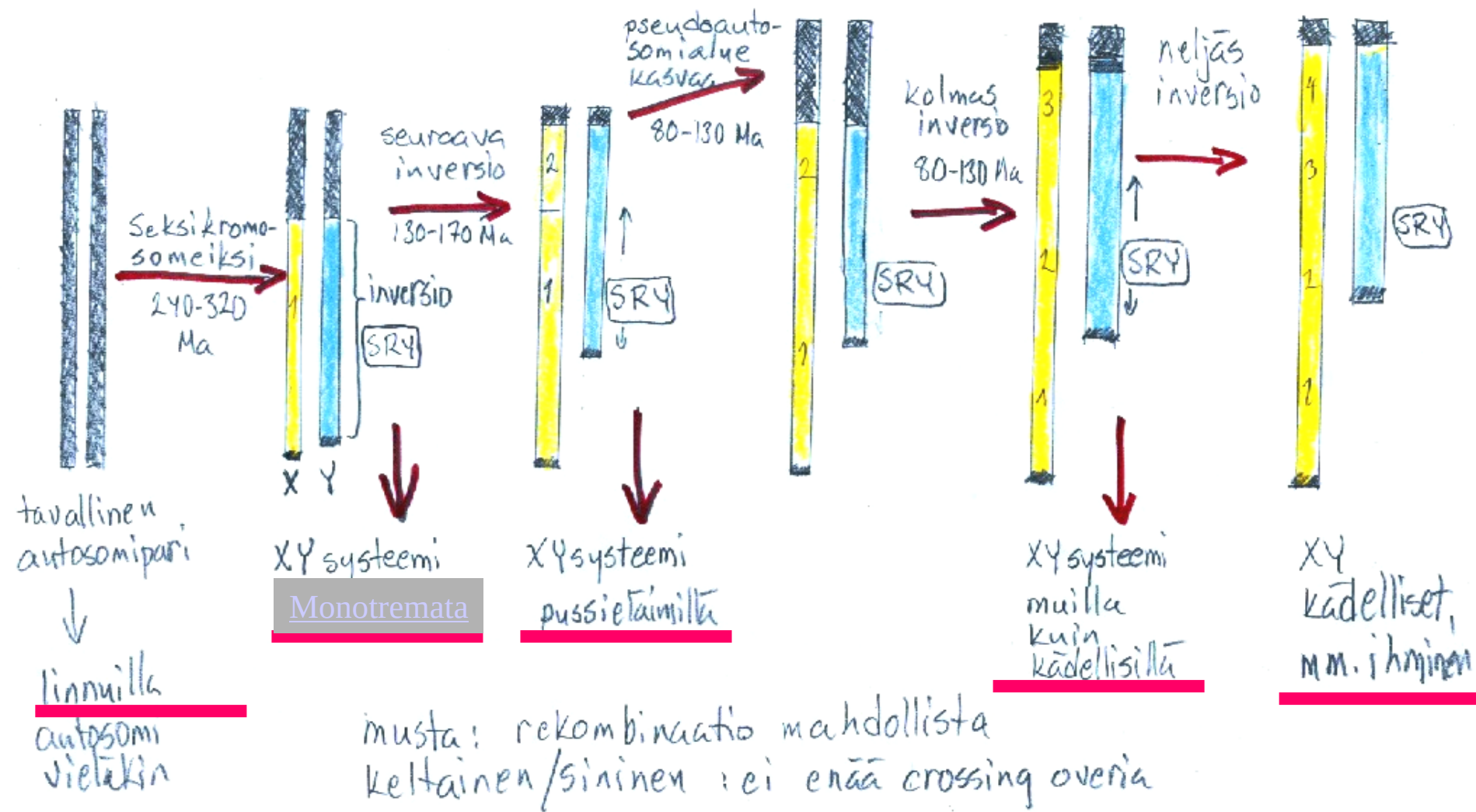
Parasentrisen inversio

Crossing over -juosteissa on 2 tai 0 sentromeeria, joten ne eivät joudu gameetteihin ikinä

FIGURE 5.5

Crossing over in a heterozygote for a paracentric inversion. The centromeres are marked by arrows.

Sukupuolikromosomien kehitys



Semmonen on laki, että jos ei ole rekombinaatiota, geenit rappeutuvat

Tästä eteenpäin voit selailla kuvia *omalla vastuulla*.

Y-kromosomin erikoislaatu (haploidia, sukupuolirajoittuneisuus) tekevät siitä hyvän, mieshistoriaa syvältä kaivelevan markkerin.

Havaintoaineistot eivät vielä ole kovin kattavia maantieteellisesti, mutta tilanne kehittyy koko ajan

The Genetic Legacy of Paleolithic *Homo sapiens* *sapiens* in Extant Europeans: A Y Chromosome Perspective

**Ornella Semino,^{1,2*†} Giuseppe Passarino,^{2,3†} Peter J. Oefner,⁴
Alice A. Lin,² Svetlana Arbuzova,⁵ Lars E. Beckman,⁶
Giovanna De Benedictis,³ Paolo Francalacci,⁷
Anastasia Kouvatsi,⁸ Svetlana Limborska,⁹ Mladen Marcikiae,¹⁰
Anna Mika,¹¹ Barbara Mika,¹² Dragan Primorac,¹³
A. Silvana Santachiara-Benerecetti,¹ L. Luca Cavalli-Sforza,²
Peter A. Underhill²**

A genetic perspective of human history in Europe was derived from 22 binary markers of the nonrecombining Y chromosome (NRY). Ten lineages account for >95% of the 1007 European Y chromosomes studied. Geographic distribution and age estimates of alleles are compatible with two Paleolithic and one Neolithic migratory episode that have contributed to the modern European gene pool. A significant correlation between the NRY haplotype data and principal components based on 95 protein markers was observed, indicating the effectiveness of NRY binary polymorphisms in the characterization of human population composition and history.

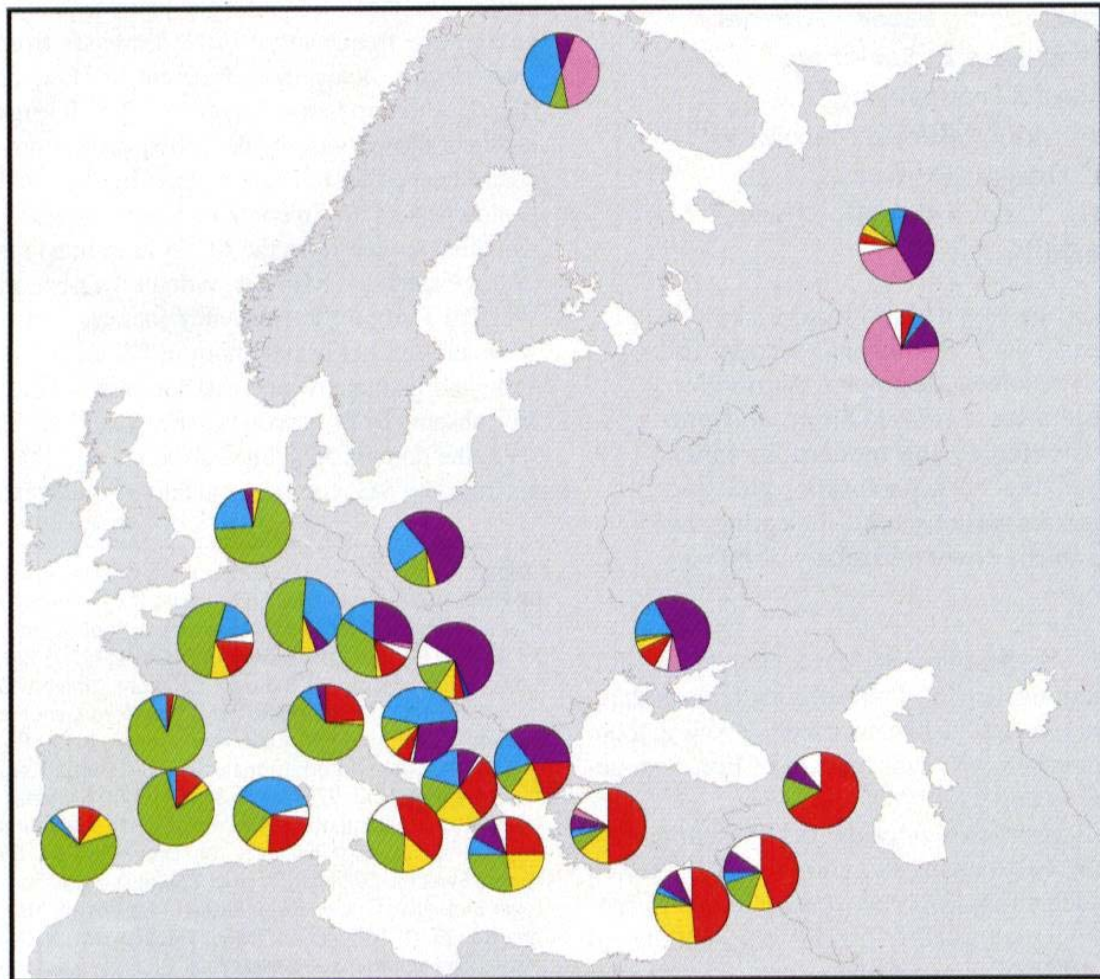
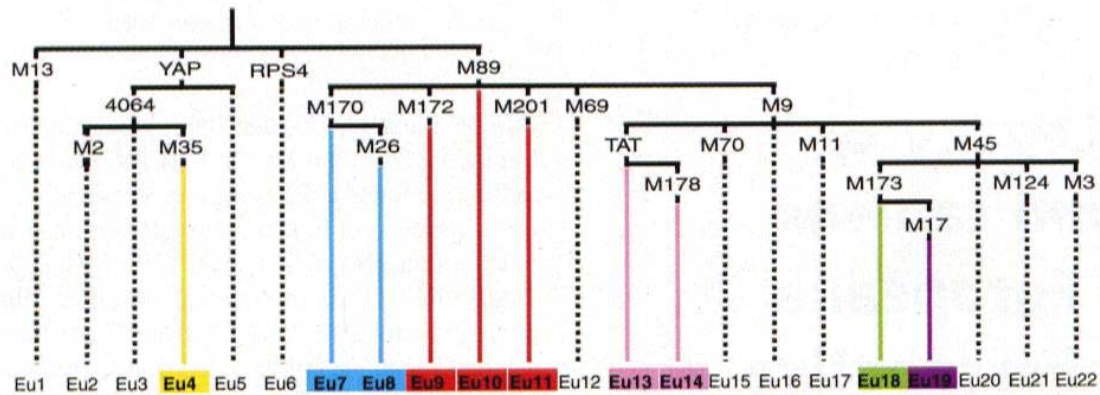
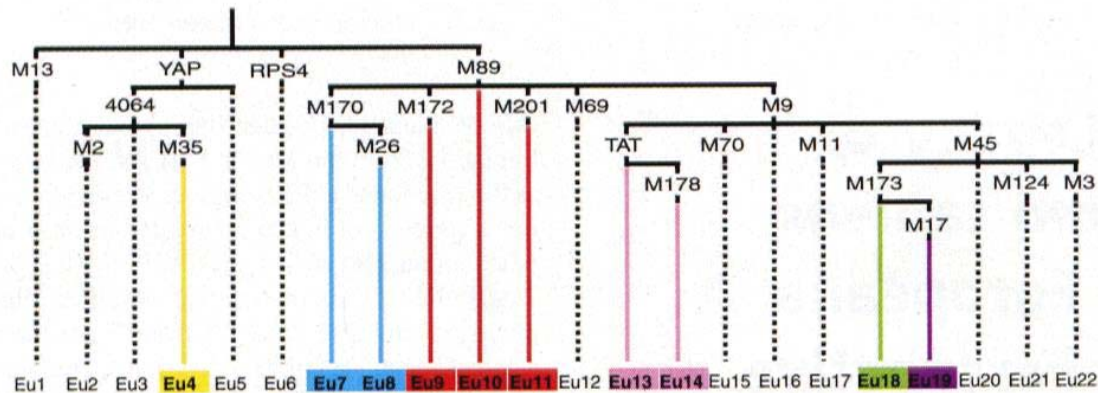


Fig. 1. (Top) Maximum parsimony phylogeny of the NRY markers found in Europe and elsewhere. YAP (32), TAT (14), RPS4 [= RPS4YC711T (33)], and 4064 [= SRY4064 (34)] were previously described. The remaining polymorphisms were identified with DHPLC (11, 12, 27) and are deposited in the National Center for BioTechnology Information (NCBI) dbSNP database (www.ncbi.nlm.nih.gov/SNP). The phylogeny is rooted with the use of great ape sequences. **(Bottom)** The 19 haplotypes observed (Table 1) were pooled into six classes represented by different colors: Yellow indicates haplotype Eu4; blue includes Eu7 and Eu8, which both involve the M170 mutation; red groups three separate haplotypes for reasons explained in the text; pink includes haplotypes Eu13 and Eu14, which both involve the TAT mutation; and green indicates Eu18 and purple indicates Eu19, which despite sharing the M173 mutation are distinguished because they represent a distinct dichotomy in European phylogeography. The other nine observed haplotypes, which catalog the remaining <5% of the total samples, are shown as black dashed lines and are represented in the white sector of relevant pie charts. Three haplotypes, Eu2, Eu5, and Eu21, were not detected. The pie sectors are proportional to the relative frequencies of haplotypes or clades in each population. The two Basque samples have been pooled.

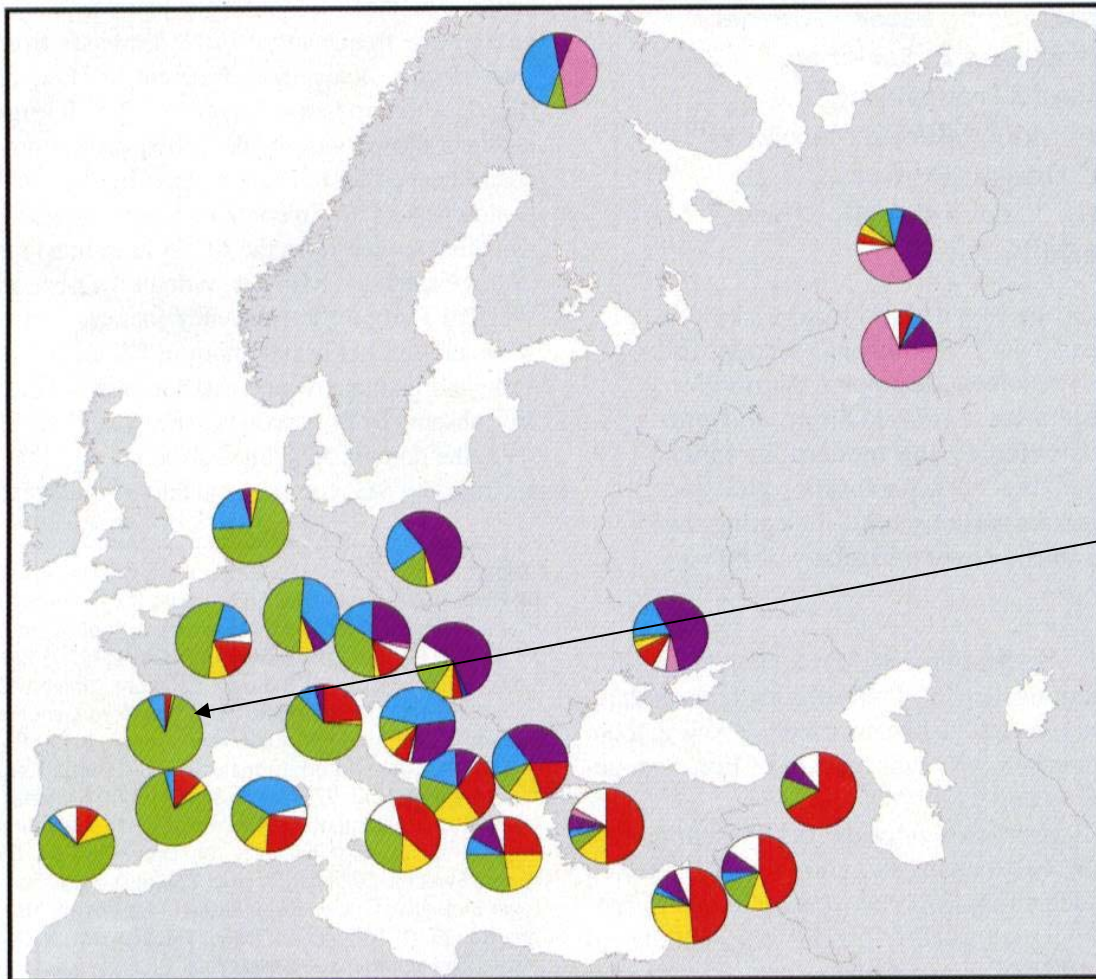


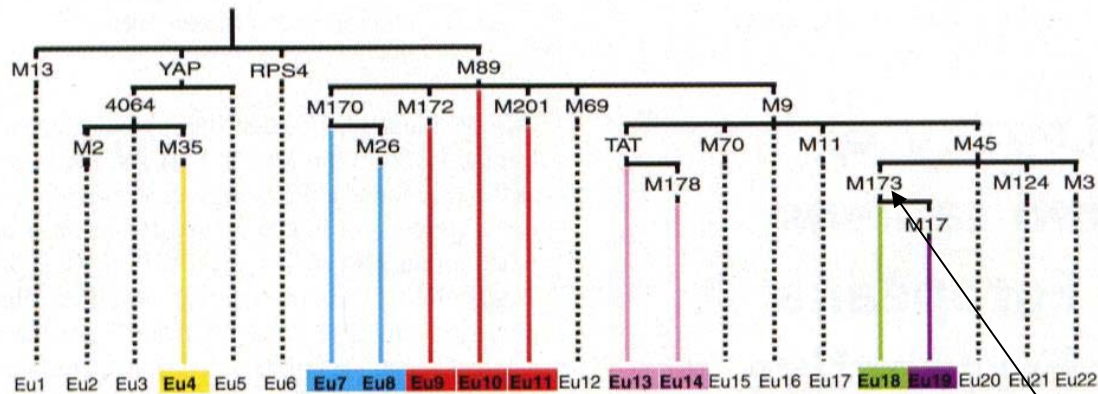
Vihreä Haplotyyppi Eu18

Keskus Iberian
jääkausirefugiossa,
niin kuin monilla
eläimillä ja kasveilla

[mtDNA haploryhmät
V ja H]

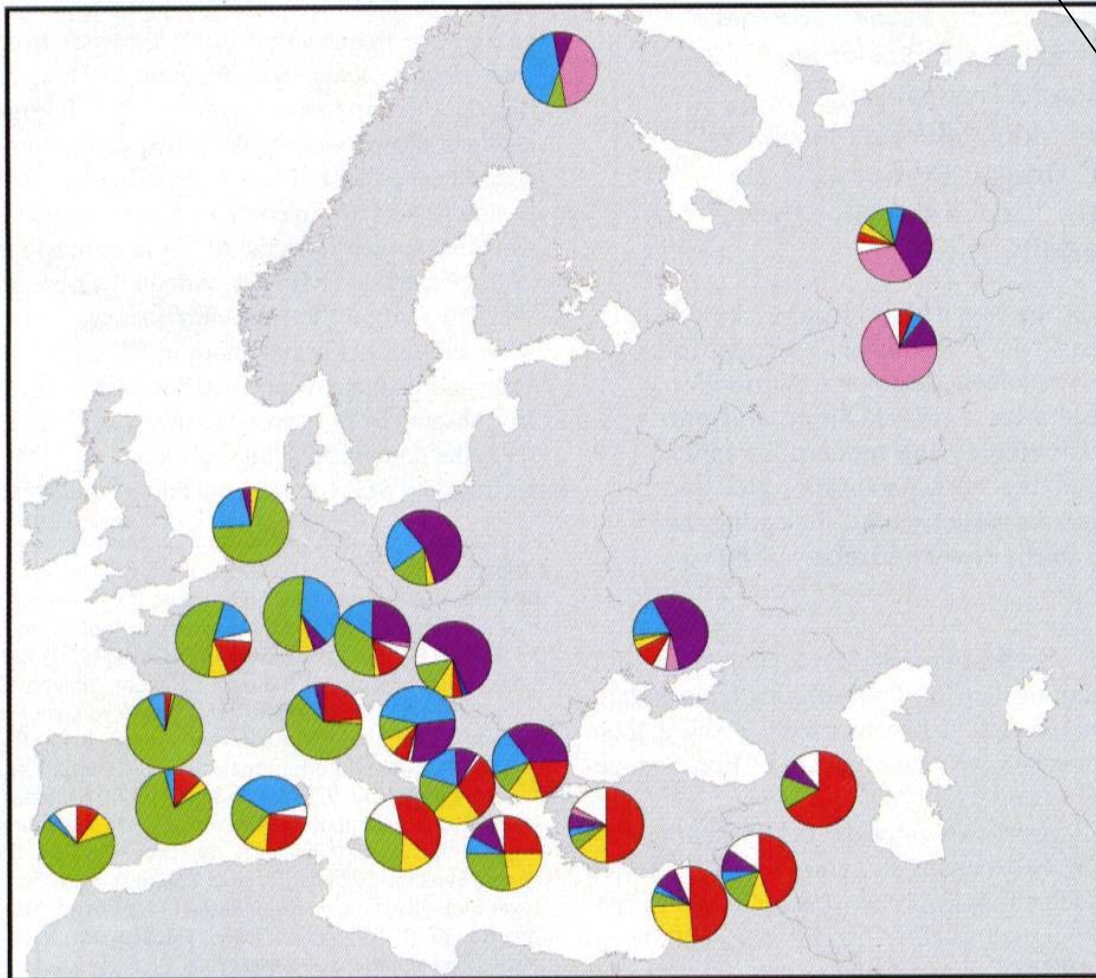
Korrelaatio baskien
kieleen on selvä



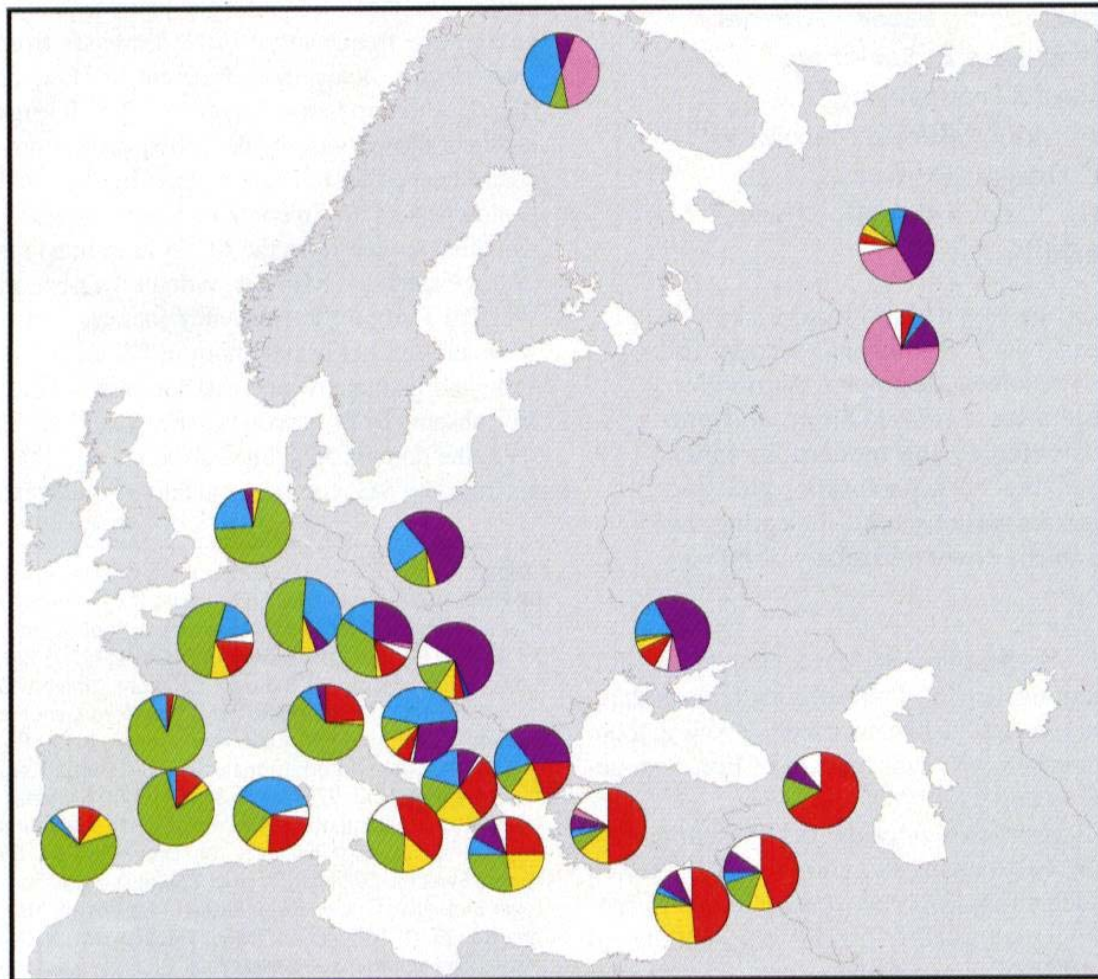
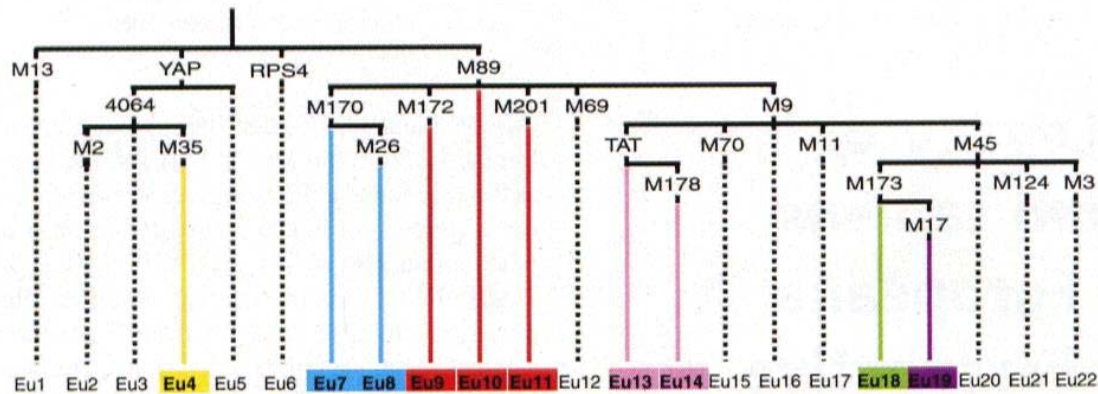


Vihreä Haplotyyppi Eu18

Keskus Iberian
jääkausirefugiossa,
niin kuin monilla
eläimillä ja kasveilla



M173, vihreän ja
violetin yhteinen juuri
on 30 000 vuotta
vanha, ehkä Aurignac-
setlementin ajoilta

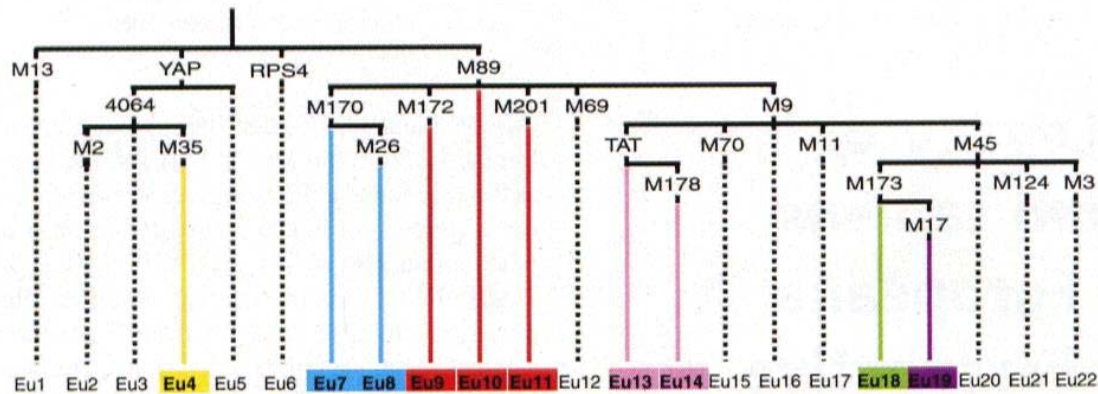


Violetti

Haplotyyppi Eu19

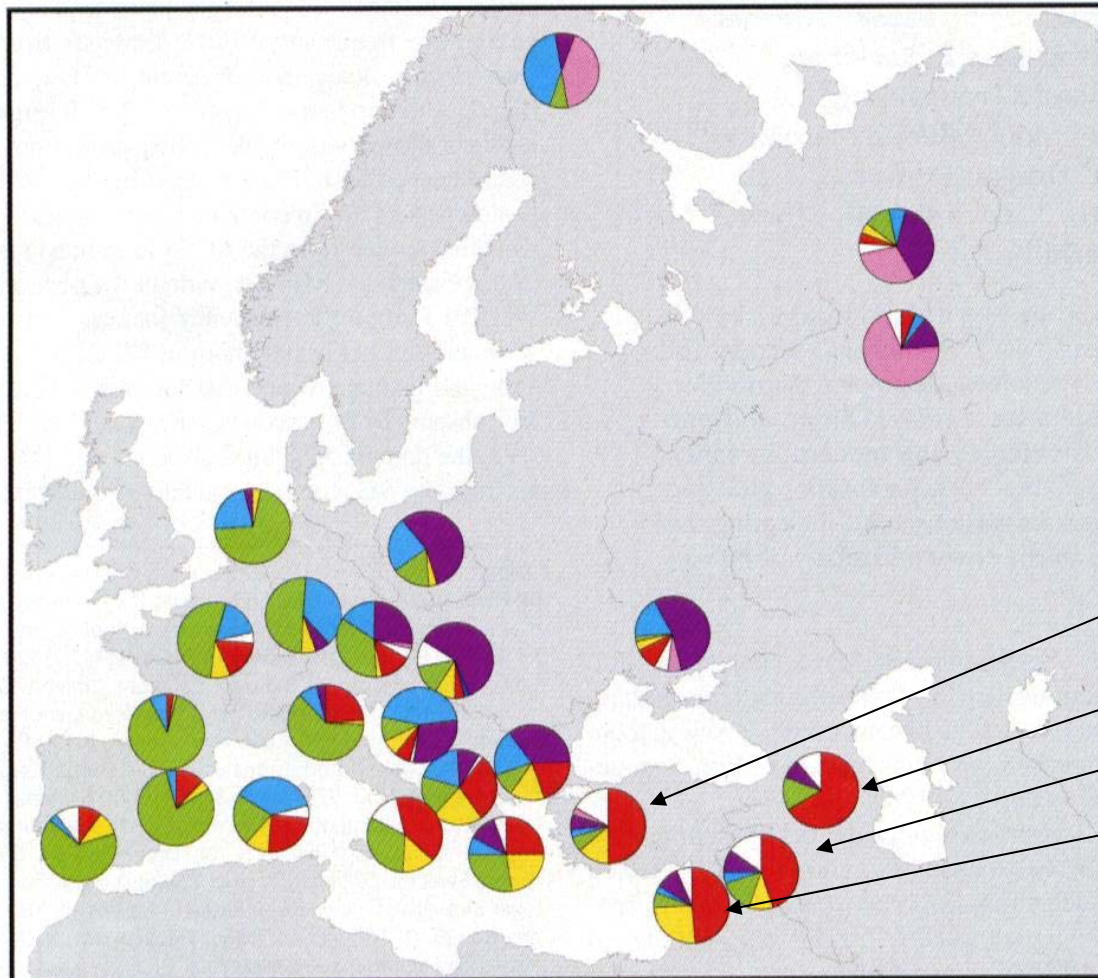
Keskus nykyisen Ukrainan seudun glasiaalirefugioissa, josta se on levinnyt Eurooppaan, mutta myös itään (Pohjois-Intia ja Pakistan)

Korrelaatiit:
Yamnaia-kulttuuri ja indoeurooppalainen kieliperhe

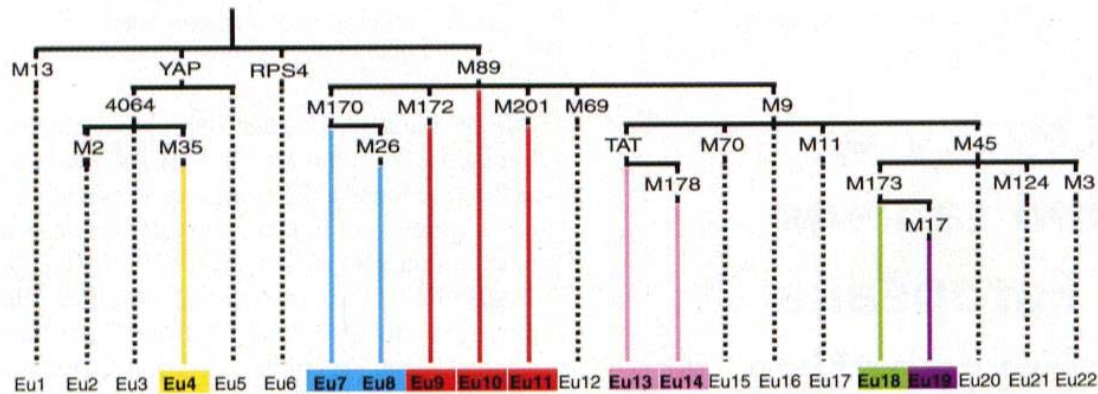


Punainen Haplot Eu9-11

Punaiseksi on väritetty laajan M89-merkityn haaran basaalityyppejä, joiden levinneisyyden painopiste on Lähi-Idässä



Turkki
Georgia
Syyria
Libanon

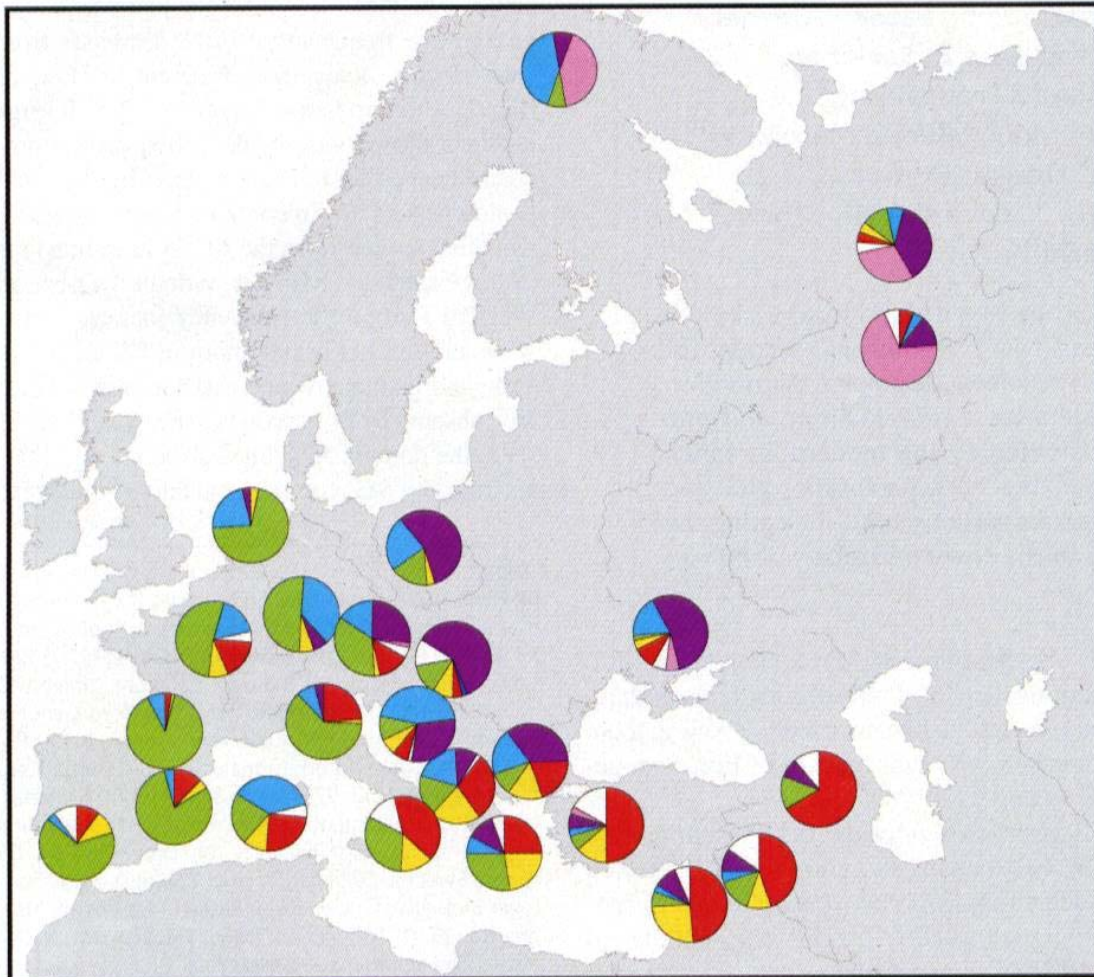


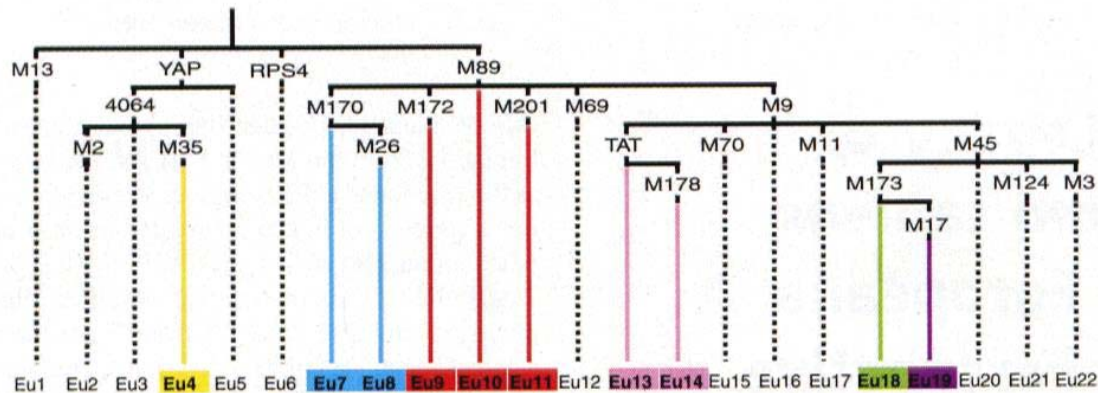
Sininen Haplot Eu7 ja Eu8

Estimoitu ikä (M170)
on 22 000 vuotta

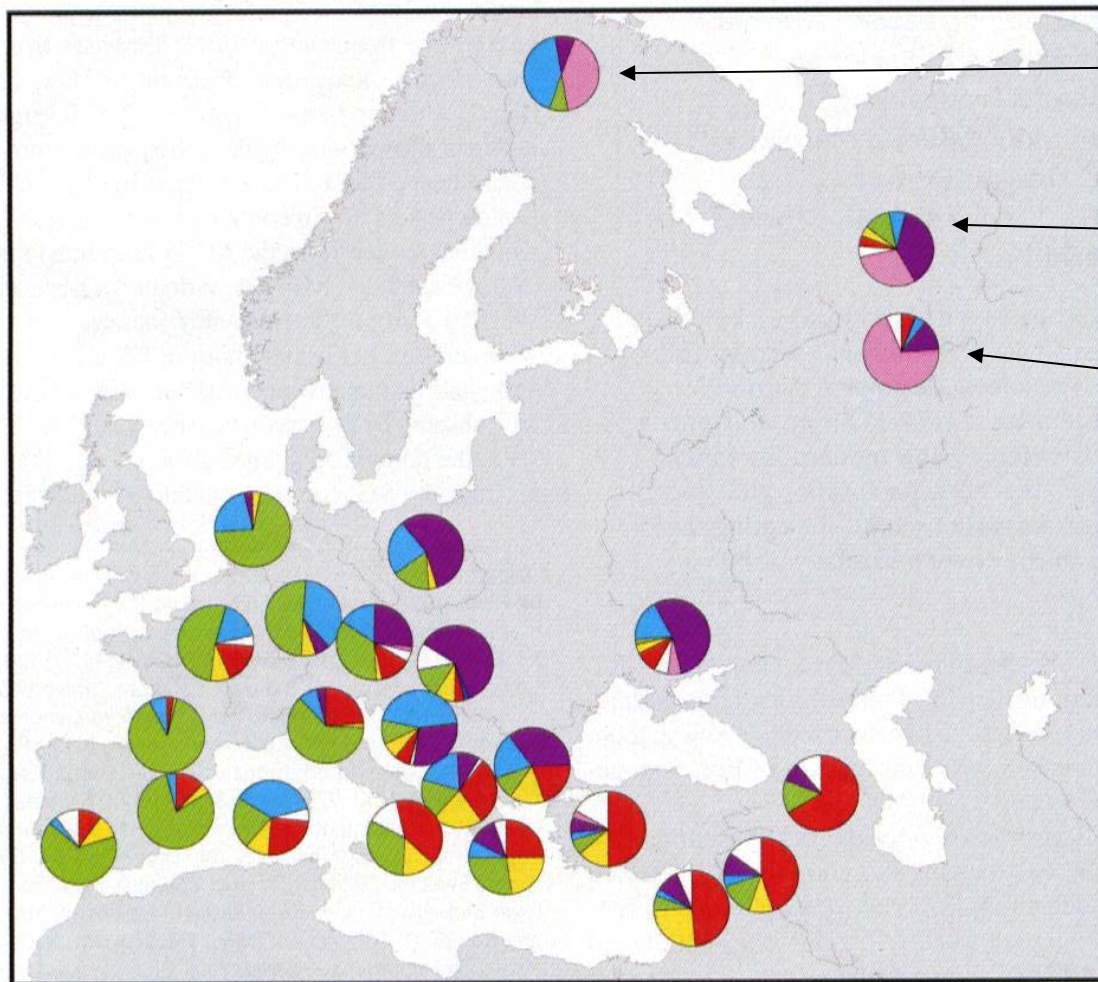
Levinneisyys rajoittunut
Eurooppaan (Eu7)

Eu8 sisältää uuden
mutaation M26 ja
löytyy baskeilta ja
sardinialaisilta





TAT -markkerin
painopiste on Aasiassa



saamelaiset

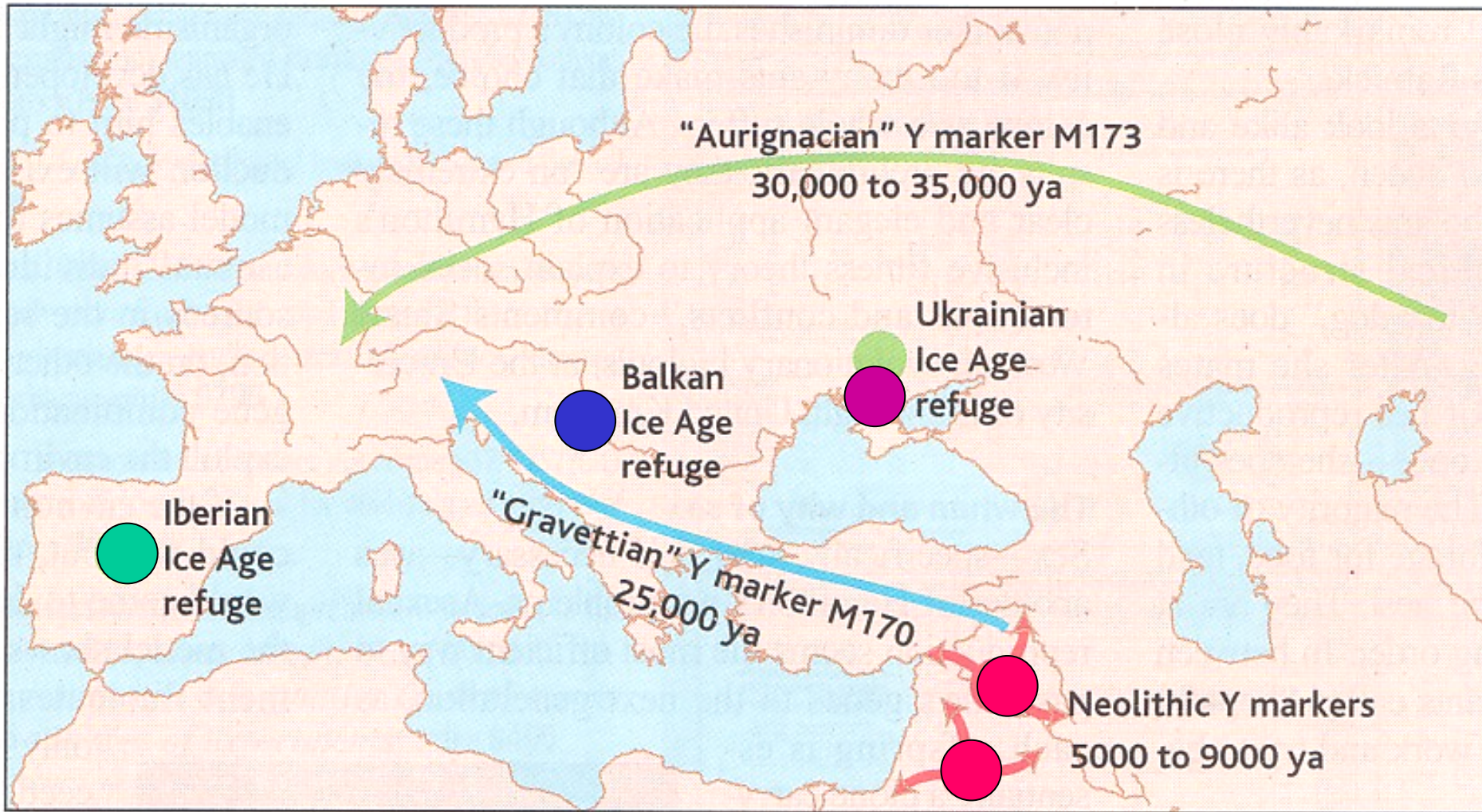
udmurtit

marit

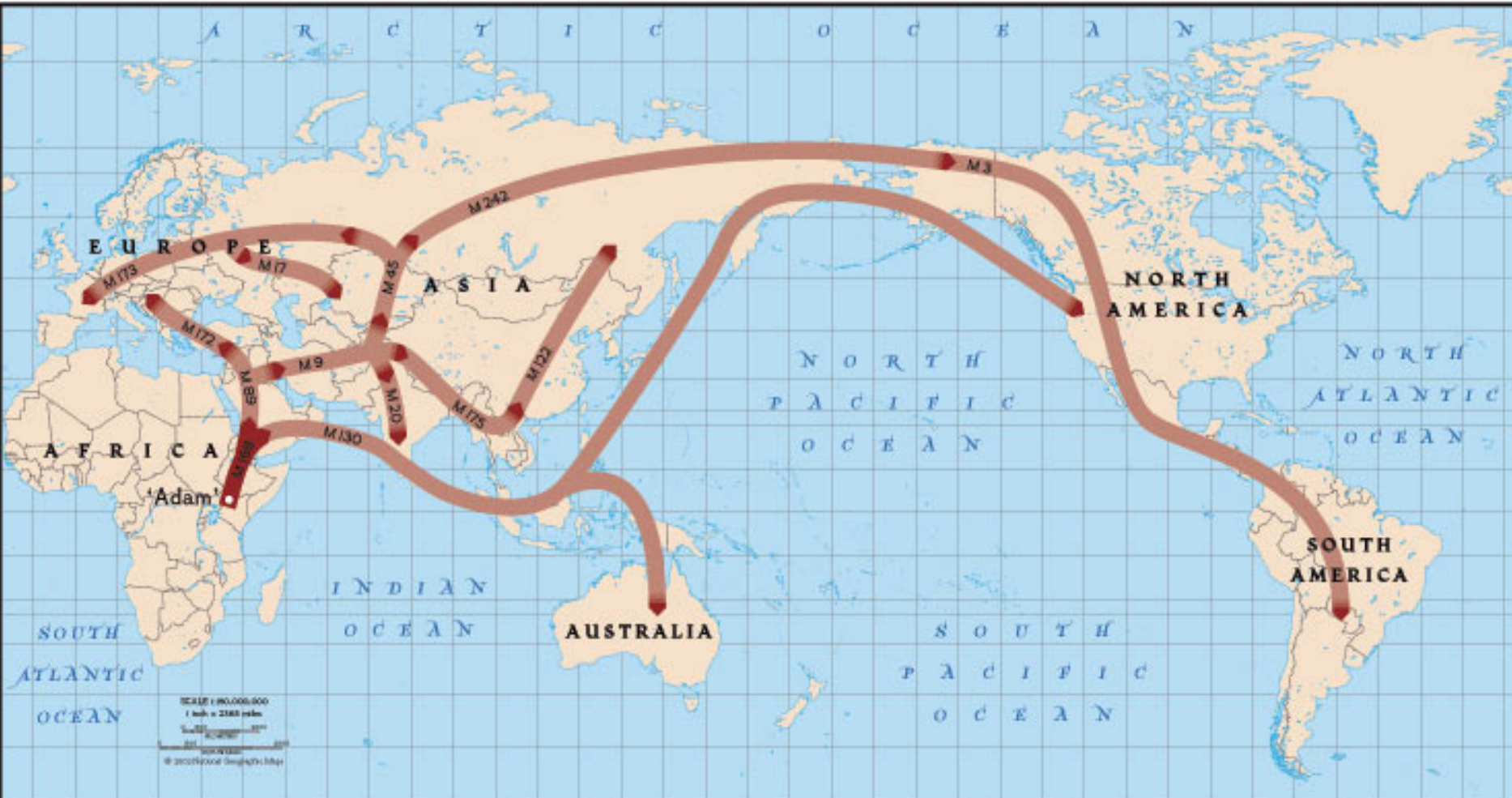
Table 1. Frequencies (in percent) of the haplotypes found in the examined European populations.

Population†	n	Haplotypes*																			
		Eu1	Eu3	Eu4	Eu6	Eu7	Eu8	Eu9	Eu10	Eu11	Eu12	Eu13	Eu14	Eu15	Eu16	Eu17	Eu18	Eu19	Eu20	Eu21	
Andalusian Spanish	29			10.3		3.4		6.9	3.4					6.9		3.4	65.5				
Basque	45			2.2		2.2	4.4		2.2								88.9				
French Basque	22						9.1	4.5									86.4				
Catalan	24			4.2		4.2		4.2		8.3							79.2				
French	23			8.7		17.4		13.0	4.3								52.2		4.3		
Dutch	27			3.7		22.2											70.4	3.7			
German	16			6.2		37.5											50.0	6.2			
Czech and Slovakian	45			2.2		15.6		8.9		4.4			2.2	2.2	2.2		35.6	26.7			
Central-northern Italian	50			2.0		8.0		14.0		10.0							62.0	4.0			
Calabrian	37		2.7	13.5				21.6	10.8	8.0					2.7	5.4	32.4		2.7		
Sardinian	77	1.3		10.4	1.3	2.6	35.1	5.2	5.2	14.2				1.3			22.1			1.3	
Croatian	58			6.9		44.8		5.2		1.7					1.7		10.3	29.3			
Albanian	51		2.0	21.6		19.6		23.5	4.0	2.0							17.6	9.8			
Greek	76		1.3	22.4	1.3	7.9		21.0	1.3	2.6				1.3		1.3	27.6	11.8			
Macedonian	20			15.0		20.0		15.0	5.0								10.0	35.0			
Polish	55			3.6		23.6											16.4	56.4			
Hungarian	45			8.9		11.1		2.2		2.2						2.2	13.3	60.0			
Ukrainian	50			4.0		18.0		6.0		4.0	2.0		6.0	2.0			2.0	54.0	2.0		
Georgian	63							33.3	3.2	30.1				1.6	1.6	1.6	14.3	7.9		6.3	
Turkish	30		3.3	13.3		3.3		40.0	3.3	6.6			3.3	3.3	3.3	3.3	6.6	6.6	3.3		
Lebanese	31			25.8	3.2	3.2		29.0	16.1	3.2						3.2	6.4	9.7			
Syrian	20		10.0	10.0		5.0		15.0	30.0		5.0						15.0	10.0			
Saami	24					41.7							41.7				8.3	8.3			
Udmurt	43			4.7		7.0			4.7			2.3	27.9		4.6		11.6	37.2			
Mari	46					4.3			6.5			4.3	65.2		6.5			13.0			
Total	1007																				

*The haplotypes are defined by the following markers and the respective derived alleles: Eu1, M13-C; Eu3, YAP⁺, 4064-A; Eu4, YAP⁺, 4064-A, M35-C; Eu6, RPS4-T; Eu7, M89-T, M170-C; Eu8, M89-T, M170-C, M26-A; Eu9, M89-T, M172-G; Eu10, M89-T; Eu11, M89-T, M201-T; Eu12, M89-T, M69-C; Eu13, M89-T, M9-G, TAT-C; Eu14, M89-T, M9-G, TAT-C, M178-T; Eu15, M89-T, M9-G, M70-C; Eu16, M89-T, M9-G; Eu17, M89-T, M9-G, M11-G; Eu18, M89-T, M9-G, M45-A, M173-C; Eu19, M89-T, M9-G, M45-A, M173-C, M17(delG); Eu20, M89-T, M9-G, M45-A; Eu21, M89-T, M9-G, M45-A, M124-T. Haplotypes Eu2, Eu5, and Eu22 were not observed. †Several samples were previously described (10, 11, 28). Samples not previously examined included 23 French, 16 Germans, 39 northern Italians, 45 Sardinians, 58 Croats, 20 Macedonians from northern Greece, 55 Poles, 50 Ukrainians, 20 Syrians, 24 Saami, 43 Udmurts, and 46 Mari.



Men on the move. Y chromosome data reveal three major migrations into Europe, which researchers tie to known archaeological cultures. At 40,000 years ago (ya), the Aurignacian people moved in (green), followed by the Gravettians 25,000 years ago (blue), and finally the Neolithic farmers (red) 9000 years ago.



EARLY MAN MIGRATION

M 168: 50,000 years ago
M 130: 50,000 years ago
M 89: 45,000 years ago
M 9: 40,000 years ago

M 175: 35,000 years ago
M 45: 35,000 years ago
M 173: 30,000 years ago

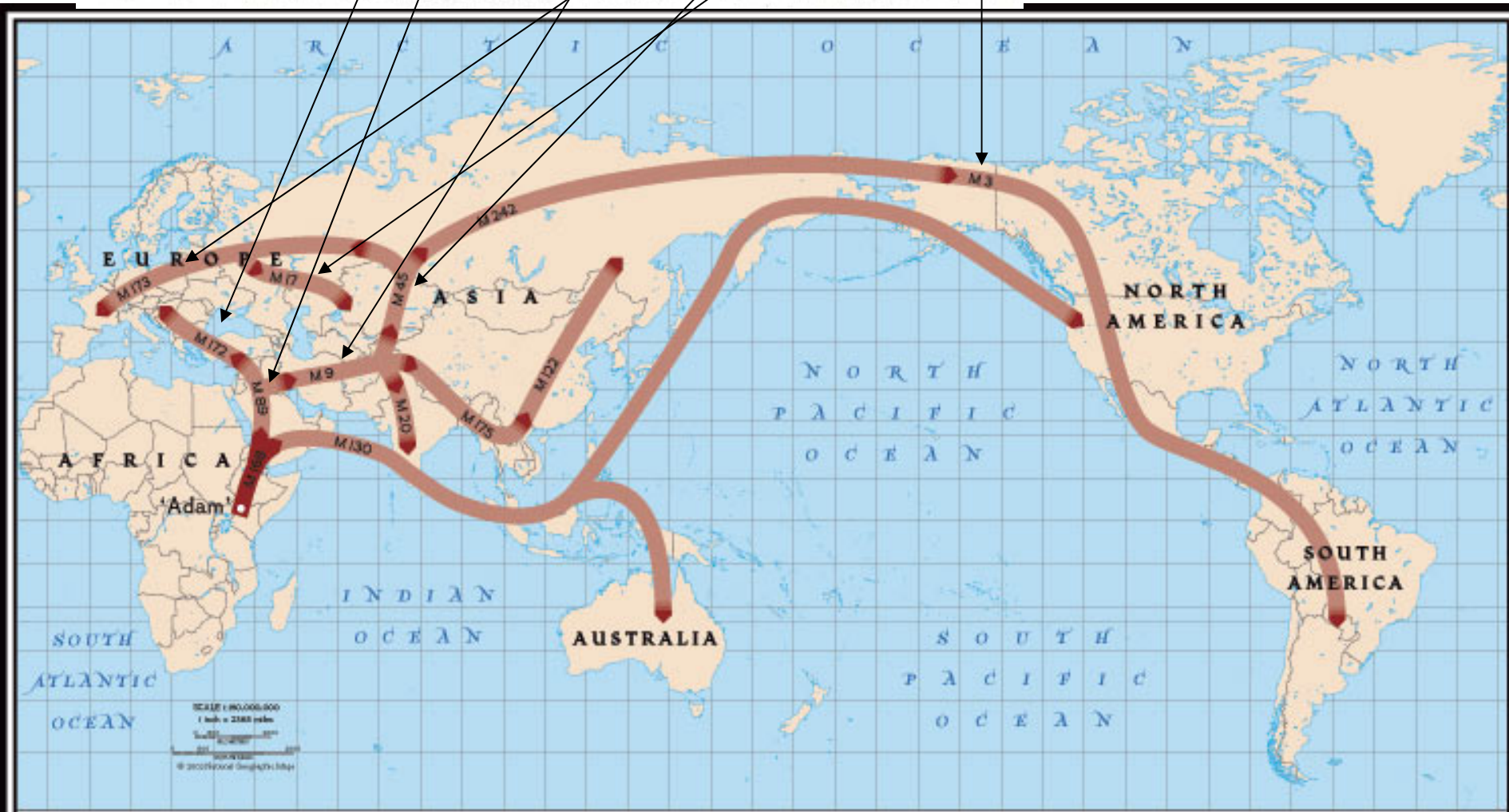
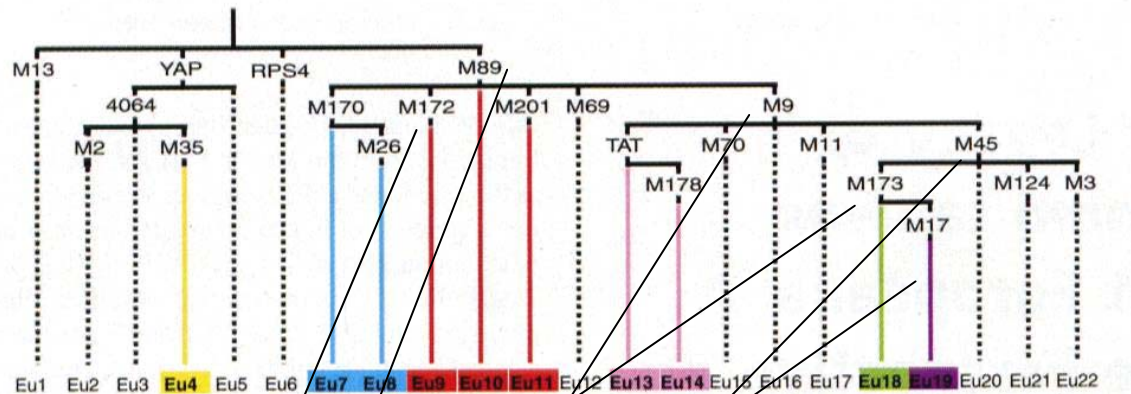
M 20: 30,000 years ago
M 242: 20,000 years ago
M 122: 10,000 years ago

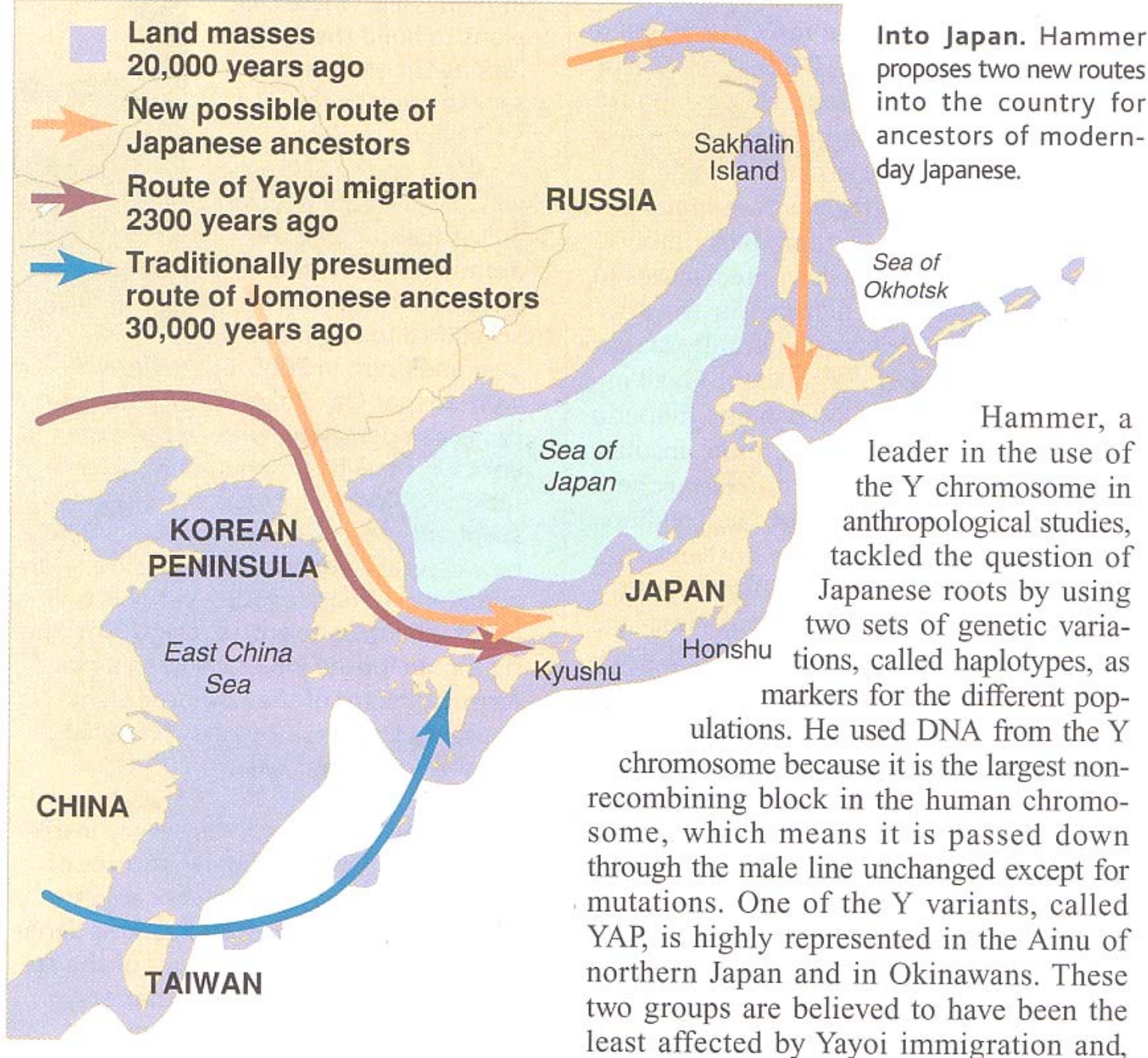
M 3: 10,000 years ago
M 172: 10,000 years ago
M 17: 10,000 years ago

Journey of Man

The markers above represent genetic markers found in DNA as identified by Dr. Spencer Wells in the film *Journey of Man*. Tracing our ancestry back to the people who left Africa some 50,000+ years ago, Dr. Wells tells the story of a remarkable human journey that began in Africa (M168), and divided and expanded into the rest of the world, connecting everyone to a global family tree. Watch *Journey of Man* in December on National Geographic Channel, and visit www.nationalgeographic.com/channel/intl to find information specific to your region.

**NATIONAL
GEOGRAPHIC
CHANNEL**





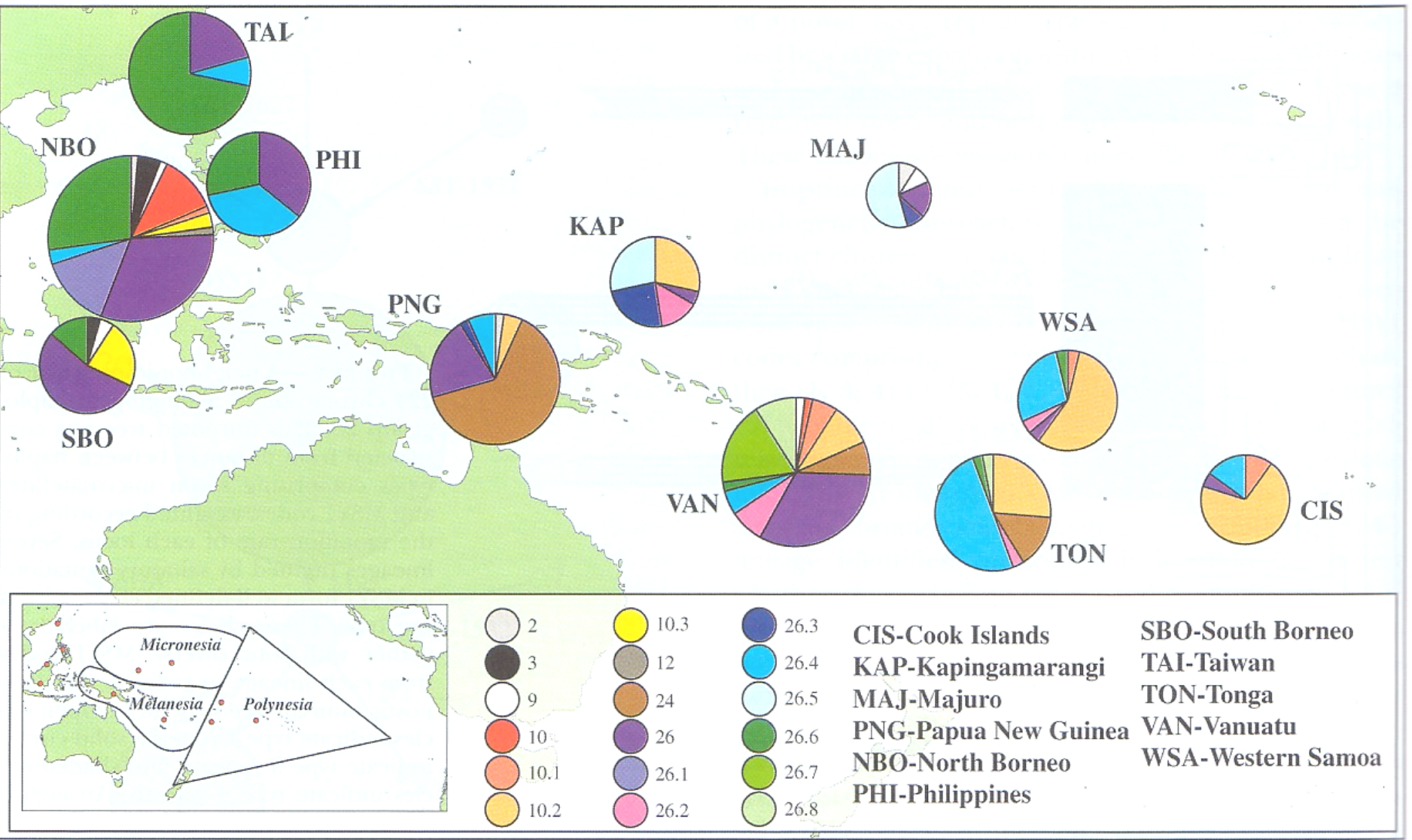


FIGURE 3.—Map of Oceania and SE Asia indicating Y chromosomal lineage frequencies in each of the 11 populations. Circle area is proportional to sample size. The inset map indicates the three geographical regions of the Pacific into which each population falls.

Y-kromosomin rakenteesta ja sisällöstä (todella ylimääräistä)

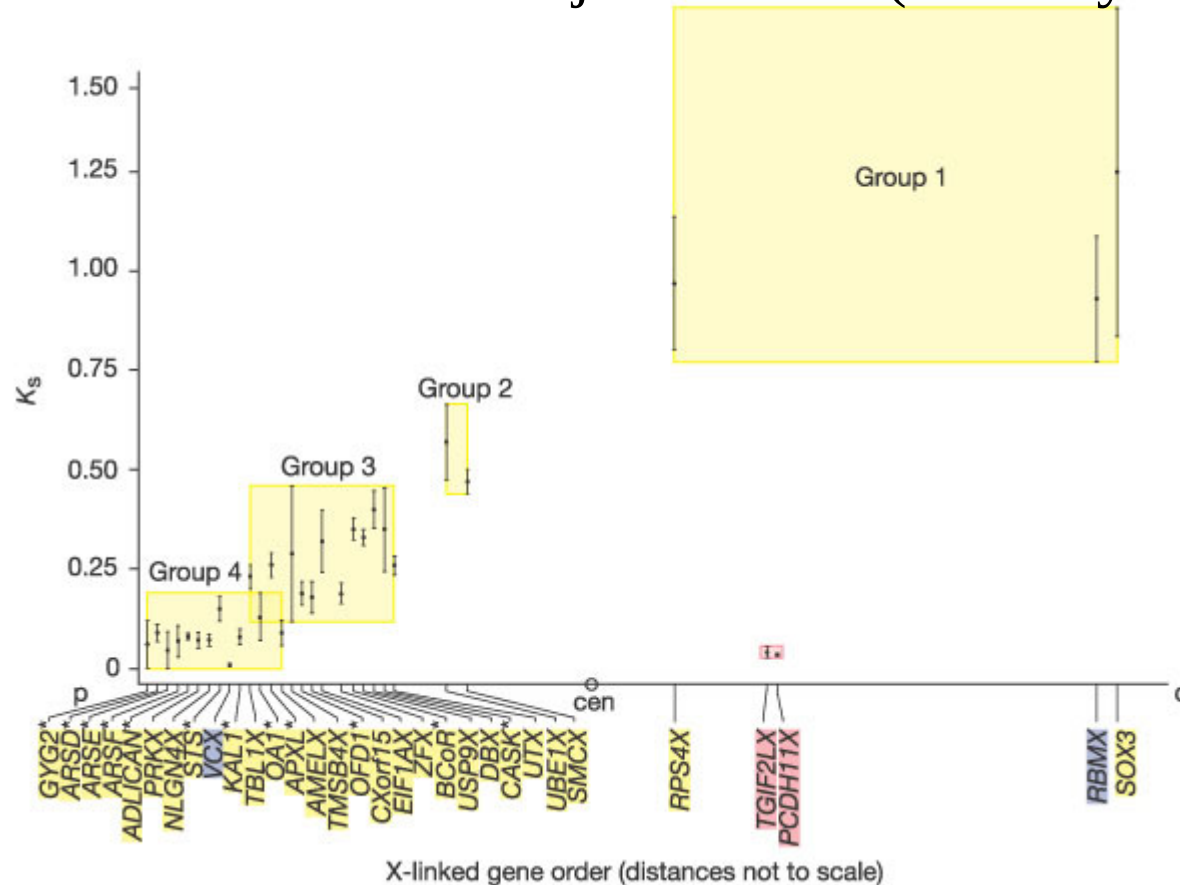


Figure 7 Plot of K_s (Supplementary Table 5) versus X-linked gene order for 31 X-Y gene (or gene/pseudogene) pairs. Colour highlighting of X-linked gene names indicates whether Y homologues are X-degenerate (yellow), ampliconic (blue) or X-transposed (pink). Within the plot, four yellow rectangles denote four previously defined 'evolutionary strata', or groups of genes²⁶; a small pink rectangle highlights two X-transposed genes. Genes in the X chromosome are ordered according to the NCBI sequence assembly of November 2002; distances between genes are not drawn to scale. Standard errors for K_s values are shown.

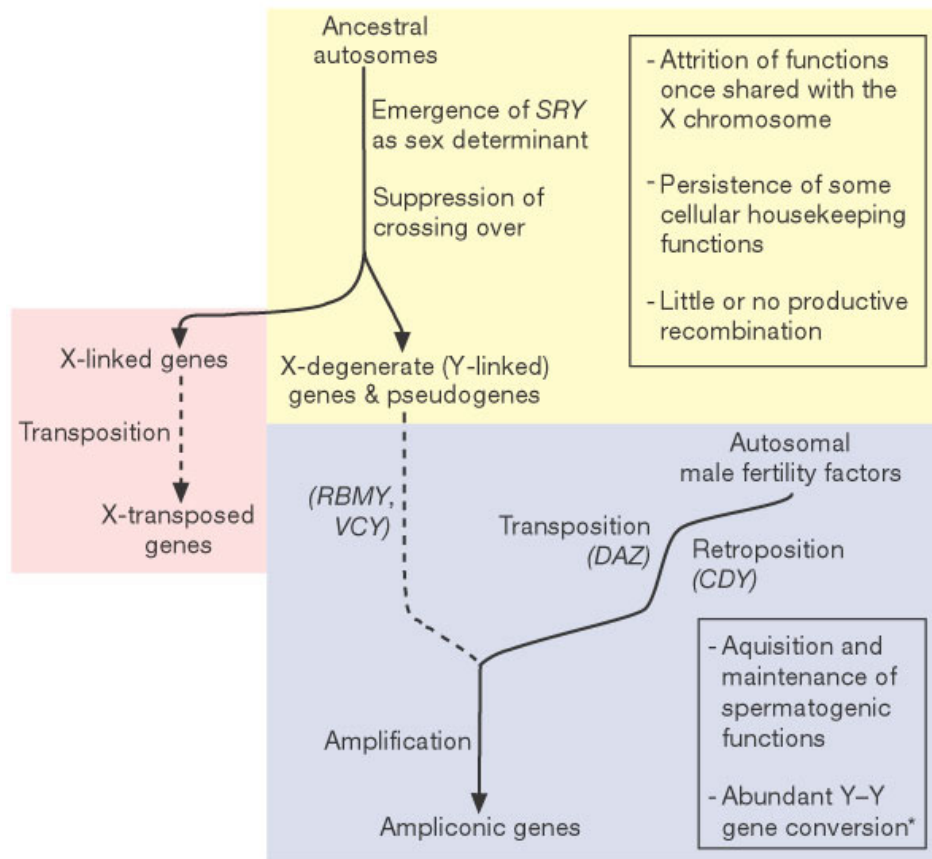


Figure 6 Molecular evolutionary pathways and processes that gave rise to genes in three MSY euchromatic sequence classes. X-degenerate genes and pseudogenes (yellow background) derived from an autosomal pair that was ancestral to both the X and Y chromosomes (and that was enlarged by subsequent fusion with other autosomes or autosomal segments⁵⁰). X-transposed genes (pink background) derived from X-linked genes, which in turn derived from the ancestral autosomal pair. Ampliconic genes (blue background) were derived through three converging processes: amplification of X-degenerate genes (for example, *RBMV*, *VCY*); transposition and amplification of autosomal genes (*DAZ*); and retroposition and amplification of autosomal genes (*CDY*). Boxes enumerate dominant themes in X-degenerate (yellow) and ampliconic (blue) gene evolution. The asterisk indicates that Y–Y gene conversion is apparently common in the 61% of ampliconic sequences that exhibit intrachromosomal identities of 99.9%.

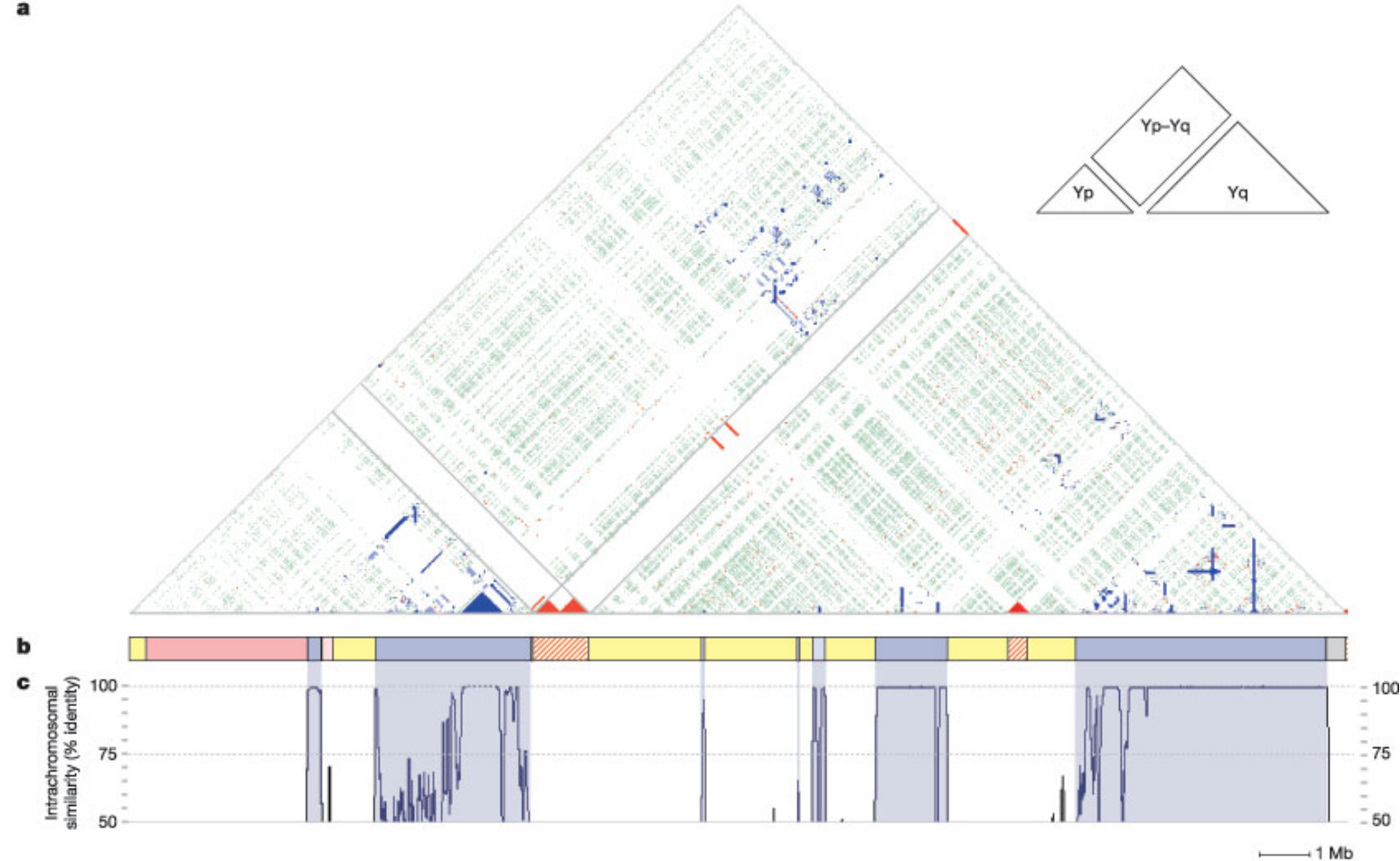


Figure 5 Sequence similarities within the MSY. **a**, Triangular dot plot in which the MSY's sequence is compared to itself. Within the plot, each dot represents a match of >65% within a window of 2,000 nucleotides. Green dots represent matches of this quality between LINE1 elements; red dots represent matches between heterochromatic sequences; blue dots represent matches between all other sequences. Direct repeats appear as horizontal lines, inverted repeats as vertical lines, and palindromes as vertical lines that nearly intersect the baseline. Long arrays of tandem repeats appear as pyramids. The inset indicates that the large triangular plot contains two smaller triangles (one revealing sequence similarities within Yp, and one revealing similarities within Yq) and a rectangle (revealing similarities between Yp and Yq). **b**, MSY schematic, as in Fig. 1b. **c**, Plot of intrachromosomal sequence similarity, which serves to identify ampliconic sequences (blue). Using a 50-kb sliding window and 1-kb steps, each MSY euchromatic sequence was compared to all other available MSY euchromatic sequences. (Long interspersed repeats were excluded before analysis.) At each point along the length of the MSY, the highest sequence similarity (expressed as per cent nucleotide identity) was identified. All such values >50% are shown. An expanded version of this plot is shown in Fig. 2f.

HUMAN Y CHROMOSOME: AN EVOLUTIONARY MARKER COMES OF AGE

Mark A. Jobling & Chris Tyler-Smith about the authors

Preface

Until recently, the Y chromosome seemed to fulfil the role of juvenile delinquent among human chromosomes — rich in junk, poor in useful attributes, reluctant to socialize with its neighbours and with an inescapable tendency to degenerate. The availability of the near-complete chromosome sequence, plus many new polymorphisms, a highly resolved phylogeny and insights into its mutation processes, now provide new avenues for investigating human evolution. Y-chromosome research is growing up.

Summary

The human Y chromosome is male-sex-determining and haploid, and so escapes recombination for most of its length. Haplotypes, which can be defined by the many binary markers and microsatellites that are available, pass down paternal lineages and change only by mutation.

A small effective population size and the practice of patrilocality accentuate drift, which leads to the marked geographical differentiation of Y haplotypes. This makes the Y chromosome a powerful tool for investigating events in human genetic history. The study of mutation on the Y chromosome clarifies intra-allelic processes in general, and provides specific information about mutation rates that is useful in estimating the coalescent times of lineages. Intrachromosomal paralogous sequences are plentiful and cause pathogenic and non-pathogenic structural rearrangements.

Selection might be important in shaping Y-chromosome diversity in populations, but it has been difficult to identify. Some studies show associations between deleterious phenotypes and particular haplotypes, but these associations are weak; some coalescence times are younger than expected, which indicates recent selection, but these estimates are uncertain, and population phenomena might be an alternative explanation.

The phylogeny of binary Y haplogroups is well established, but the dates of branchpoints are uncertain. Many populations have been poorly sampled, and there is ascertainment bias in the set of available binary markers.

The recent coalescence time, rooting of the Y phylogeny in Africa and evidence for an 'Out-of-Africa' range expansion, all show that modern Y-chromosome diversity arose recently in Africa and replaced Y chromosomes elsewhere. The pattern of Y-chromosome variation broadly fits a model of a southern migration that reached Australia, and a northern migration into Eurasia.

Many features of the patterns of modern Y-chromosome diversity reflect later range expansions and contractions that were driven by changes in climate and lifestyle. Long-term population size, social structures and social selection have also been important.

Future developments in the field are likely to include more markers, and a move towards the unbiased resequencing of samples.

Other parts of the genome might show a 'haplotype-block' structure that is made up of regions of strong linkage disequilibrium. If this is so, then methods pioneered in the analysis of the Y chromosome could be widely applicable.

High mutation rates have driven extensive structural polymorphism among human Y chromosomes

Sjoerd Repping^{1,2}, Saskia K M van Daalen², Laura G Brown¹, Cindy M Korver², Julian Lange¹, Janet D Marszalek¹, Tatyana Pyntikova¹, Fulco van der Veen², Helen Skaletsky¹, David C Page¹ & Steve Rozen¹

Nature Genetics
4 April 2006

Although much structural polymorphism in the human genome has been catalogued¹⁻⁵, the kinetics of underlying change remain largely unexplored. Because human Y chromosomes are clonally inherited, it has been possible to capture their detailed relationships in a robust, worldwide genealogical tree^{6,7}. Examination of structural variation across this tree opens avenues for investigating rates of underlying mutations. We selected one Y chromosome from each of 47 branches of this tree and searched for large-scale variation. Four chromosomal regions showed extensive variation resulting from numerous large-scale mutations. Within the tree encompassed by the studied chromosomes, the distal-Yq heterochromatin changed length ≥ 12 times, the *TSPY* gene array changed length ≥ 23 times, the 3.6-Mb IR3/IR3 region changed orientation ≥ 12 times and the *AZFc* region was rearranged ≥ 20 times. After determining the total time spanned by all branches of this tree (~ 1.3 million years or 52,000 generations), we converted these mutation counts to lower bounds on rates: $\geq 2.3 \times 10^{-4}$, $\geq 4.4 \times 10^{-4}$, $\geq 2.3 \times 10^{-4}$ and $\geq 3.8 \times 10^{-4}$ large-scale mutations per father-to-son Y transmission, respectively. Thus, high mutation rates have driven extensive structural polymorphism among human Y chromosomes. At the same time, we found limited variation in the copy number of Y-linked genes, which raises the possibility of selective constraints.

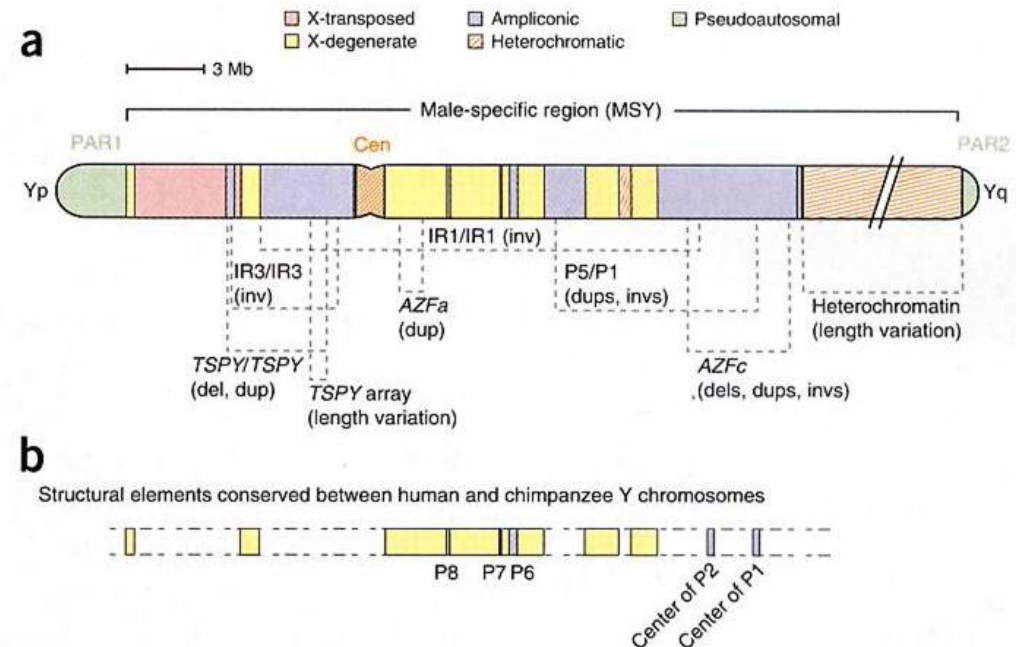


Figure 1 Overview of potential structural variation in the human Y chromosome. At top, the structure of the reference Y chromosome, including short and long arms (Yp and Yq), pseudoautosomal regions 1 and 2 (PAR1 and PAR2) and centromere (Cen). (a) Potential structural polymorphisms for which we assayed (details in **Supplementary Methods**). (b) Structural elements conserved between human and chimpanzee Y chromosomes are shown according to their position in the reference human Y chromosome. These conserved elements consist of the X-degenerate sequence, palindromes P8, P7 and P6 and the centers of palindromes P2 and P1.

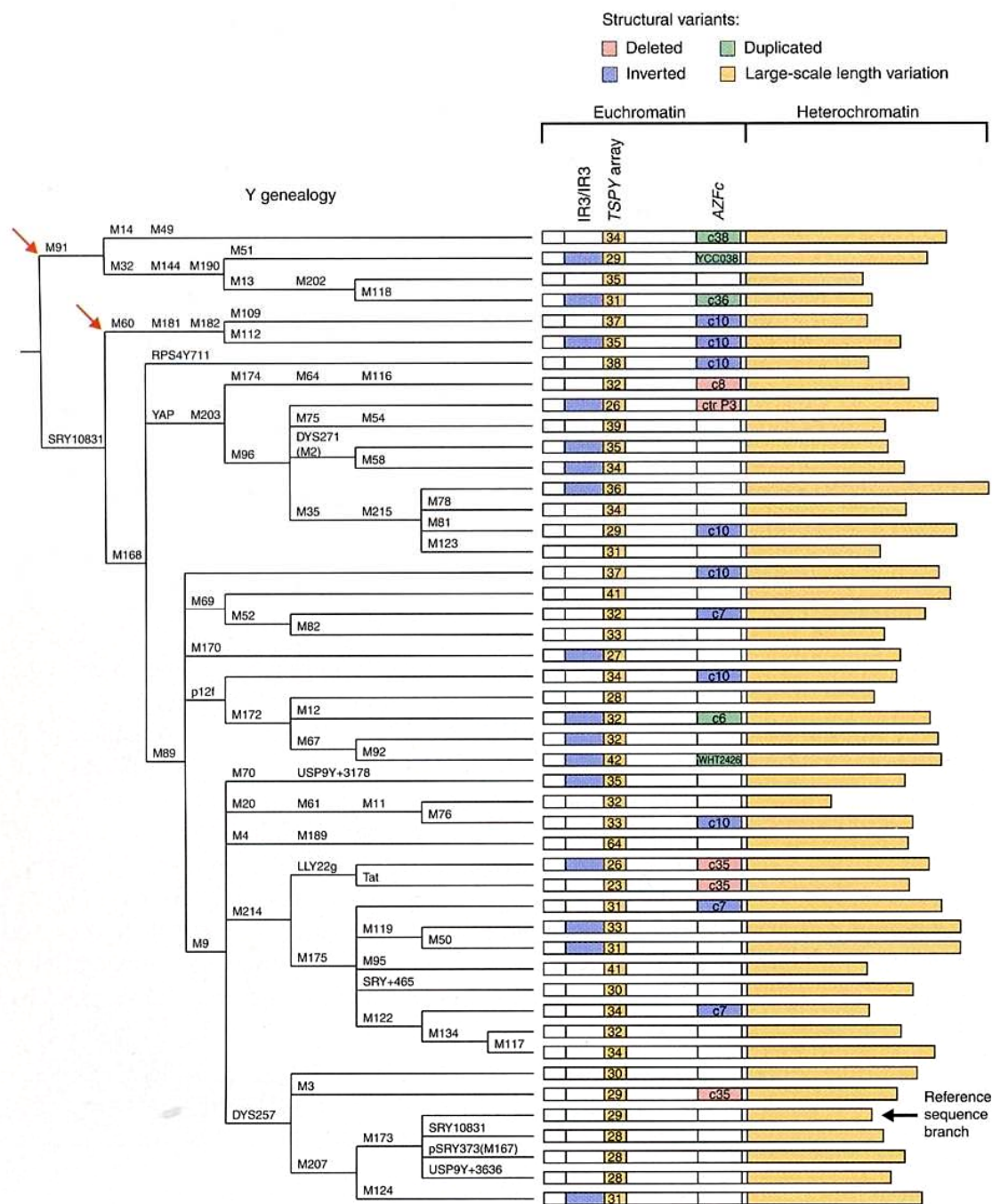


Figure 2 Y chromosome genealogical tree (left) and identified structural polymorphisms (right). Chromosomes were assigned to one of 47 branches by typing for the stable, biallelic polymorphisms indicated (for example, M91 and M60; refs. 6,7). Red arrows indicate major branches confined to Africa⁶. For each branch, the structure of the Y chromosome sampled is schematized, including, at far right, the length of distal-Yq heterochromatin. Within the euchromatin, the presence of a particular structural variant is indicated by a color-coded rectangle. Codes denoting specific AZFc architectures are explained in **Figure 4**, **Supplementary Table 2** and **Supplementary Figures 2–7**. See **Supplementary Figures 3** and **4** for the ‘ctr P3’ deletion and for ‘YCC038’, which contains a small deletion, but in which duplication predominates. The reference Y chromosome belongs to the indicated branch (**Supplementary Methods**), but, as no corresponding cell line exists, its heterochromatin and TSPY array lengths could not be determined. **Supplementary Figure 1** provides sample identifiers and Y-haplotype designations^{6,7}.

composed of low-complexity sequences organized in tandem arrays¹⁶. It ranged in length from 29% to 54% of the metaphase Y chromosome, with a median of 44% (**Figs. 1, 2, 3a** and **5a**). The TSPY array is composed of highly similar 20.4-kb repeat units, each containing a copy of the TSPY gene and of the CYorf16 transcription unit^{11,16,17}. The TSPY array ranged in size from 23 to 64 units (0.47 to 1.3 Mb), with a median of 32 units (0.65 Mb; **Figs. 1, 2, 3b** and **5b**).

The third region, in proximal Yp, was inverted in 16 chromosomes (**Figs. 1, 2, 3c–f, 5d**)^{18,19}. We localized the boundaries