Am. J. Hum. Genet. 67:182–196, 2000

A Short Tandem Repeat–Based Phylogeny for the Human Y Chromosome

Peter Forster,[1,2] Arne Röhl,[1] Petra Lünnemann,[1] Catrin Brinkmann,[1] Tatiana Zerjal,[3] Chris Tyler-Smith,[3] and Bernd Brinkmann[1]

[1]Institut für Rechtsmedizin, University of Münster, Münster; [2]McDonald Institute for Archaeological Research, University of Cambridge, Cambridge; and [3]Department of Biochemistry, University of Oxford, Oxford

Human Y-chromosomal short tandem repeat (STR) data provide a potential model system for the understanding of autosomal STR mutations in humans and other species. Yet, the reconstruction of STR evolution is rarely attempted, because of the absence of an appropriate methodology. We here develop and validate a phylogenetic-network approach. We have typed 256 Y chromosomes of indigenous descent from Africa, Asia, Europe, Australia, and highland Papua New Guinea, for the STR loci DYS19, DXYS156Y, DYS389, DYS390, DYS392, and DYS393, as well as for five ancient biallelic mutation events: two poly (A) length variants associated with the YAP insertion, two independent SRY-1532 mutations, and the 92R7 mutation. We have used our previously published pedigree data from 11,000 paternity-tested autosomal STR-allele transfers to produce a two-class weighting system for the Y-STR loci that is based on locus lengths and motif lengths. Reduced-median-network analysis yields a phylogeny that is independently supported by the five biallelic mutations, with an error of 6%. We find the earliest branch in our African San (Bushmen) sample. Assuming an age of 20,000 years for the Native American DYS199 T mutation, we estimate a mutation rate of 2.6×10^{-4} mutations/20 years for slowly mutating Y STRs, ~10-fold slower than the published average pedigree rate.

Introduction

The history of the human Y chromosome is currently being unraveled by the study of three types of loci: (*a*) biallelic markers—that is, point mutations and the YAP insertion (Hammer 1995; Underhill et al. 1997), (*b*) the MSY1 VNTR, or "minisatellite," locus (Jobling et al. 1998), and (*c*) a battery of short tandem repeat (STR), or "microsatellite," loci (Kayser et al. 1997). The biallelic markers remain stable for millions of years and are thus the markers of choice for outgroup rooting of the human Y-chromosomal tree by use of other primate Y chromosomes (Jobling et al. 1997). Furthermore, they are amenable to DNA hybridization–detection techniques (Underhill et al. 1997; Anagnostopoulos et al. 1999). In contrast, the MSY1 minisatellite (Jobling et al. 1998) has a very high mutation rate and therefore is especially useful for short time ranges (Foster et al. 1998). STRs are arguably the most interesting type of locus for some applications. The factors affecting STR mutation on the Y chromosome can readily be studied over long time periods, because of the absence of recombination, and they may become relevant for an understanding of the development of autosomal STR-expansion diseases. Furthermore, the mutation rates of some Y STRs are fast enough for them to be determined directly in pedigree studies. In contrast, dates for mutations on the biallelic Y tree will have unknown confidence limits, because of inevitable assumptions about prehistoric demography. An independent STR system is all the more desirable after the controversy surrounding the alleged 200-fold disparity between the "evolutionary" and the "pedigree" mutation rate in mtDNA (Howell and Mackey vs. Macaulay et al. [1997]). Finally, for ancient-DNA and forensic applications, multiallelic loci are desirable for the exclusion of sample contamination. On an interdisciplinary level, a successful STR analysis in humans will be much easier to extend to work on other species (chimpanzees, cattle, sheep, etc.), since STRs are easier to find and cheaper to score than are, for example, biallelic markers.

In practice, researchers are reluctant to use Y STRs for reconstructing the Y-chromosomal tree, and they either restrict themselves to distance analyses with Y STRs (Seielstad et al. 1999) or, instead, use a biallelic skeleton tree on which STR information is added to give a measure of diversity—and, hence, time depth—for the nodes of the biallelic tree (e.g., see Hurles et al. 1999). However, by entering all biallelic information at the outset, one loses the option of verifying the accuracy of the phylogenetic parameter settings generating the STR subtree. In the case of

Received January 25, 2000; accepted for publication May 1, 2000; electronically published May 25, 2000.

Address for correspondence and reprints: Prof. Bernd Brinkmann, Institut für Rechtsmedizin, University of Münster, Von-Esmarch-Strasse 62, 48149 Münster, Germany. E-mail: brinkma@uni-muenster.de

0002-9297/2000/6701-0021$02.00

our phylogenetic-network algorithms (Bandelt et al. 1999), the parameters will determine both whether the most parsimonious (MP) trees are contained in the network and whether, in fact, any MP tree represents the evolutionary tree. There are several STR-specific problems that can lead phylogenetic methods astray, such as real parallel mutations (due to fast mutation rates) and "virtual" parallelisms (due to either inadequate DNA-sequence resolution of substructured loci or violation of the single-repeat mutation model; see Weber and Wong 1993). Confusion is furthermore caused either by the application of minimum-spanning trees and networks (Cooper et al. 1996), which, by definition, do not allow reconstruction of ancestral nodes (Bandelt et al. 1999), or by one's disregarding the prevalence of single-repeat mutations when coding the input data (Pérez-Lezaun et al. 1997).

The aim of our study therefore is to identify and empirically support phylogenetic-network settings that will yield biologically meaningful phylogenies, independent of whether STRs for either the global Y tree or a subtree (defined, e.g., by biallelic markers) are entered. In contrast to previous studies, we employ a higher-than-usual DNA-sequence resolution for the Y-STR loci DYS389 and DYS390; we draw on our autosomal pedigree data on 11,000 allele transfers, to justify a weighting scheme for the set of Y-STR loci; and we employ our reduced median (RM) network method (Bandelt et al. 1995), rather than our median joining (MJ) network method (Bandelt et al. 1999), for deep phylogenetic reconstruction. Finally, we reverse the traditional approach by using ancient Y-chromosomal biallelic markers (YAP, SRY-1532, and 92R7) to empirically test the accuracy of our global Y-chromosomal tree based on Y STRs.

Subjects and Methods

DNA samples (from saliva or blood) were processed from the following 256 male individuals: 23 Australian Aborigines and 15 Highland Papuans (described in Forster et al. 1998); and 35 Han Chinese from Shenyang, 44 Japanese from Shiga, 20 Khalkh Mongolians, 28 Ovambos (Namibian Bantu), 6 San (Bushmen) from Namibia, 5 Western Pygmies (described in Siffert et al. 1999), 1 Eastern Pygmy (Efe tribe, Central Africa), 40 Turks from Adana, and 39 Germans from Münster (described in Nata et al. 1999). In the case of the Germans, individuals were selected on the basis of German surnames, which originated ~700 years ago.

DNA was extracted by a modified Chelex protocol (Wiegand et al. 1993). The YAP locus was typed by PCR with the primers 5′-CAG GGG AAG ATA AAG AAA TA-3′ and 5′-ACT GCT AAA AGG GGA TGG AT-3′, according to the method of Hammer and Horai (1995), and the long and short forms (Hammer 1995) were distinguished on polyacrylamide gels (see below). The SRY-1532 locus (originally "SRY10831"; Whitfield et al. 1995) was PCR amplified with the primers 5′-TCC TTA GCA ACC ATT AAT CTG G-3′ and 5′-AAA TAG CAA AAA ATG ACA CAA GGC-3′, and the resulting 167-bp product was screened, by digestion with *Dra*III (A was uncut; G was cut), for the SRY-1532 polymorphism. The 92R7 locus (Kwok et al. 1996) was PCR amplified with the primers 5′-TGC ATG AAC ACA AAA GAC GTA-3′ and 5′-GCA TTG TTA AAT ATG ACC AGC-3′, and, to screen for the polymorphism, the resulting 55-bp product was digested with *Hin*dIII (C was partially cut; T was uncut). The Y-chromosomal STRs were amplified as described by Kayser et al. (1997) and Rolf et al. (1999). Both the biallelic and the STR amplicons were sized on silver-stained (Budowle et al. 1991) 8% polyacrylamide gels, with the exceptions of the 92R7 digestion products, which were sized on 14% gels, and the DYS389 and DYS390 loci, which were directly sequenced by the Big Dye deoxy terminator cycle-sequencing method (PE Biosystems) and ABI Prism 310 capillary electrophoresis (PE Biosystems).

The phylogenetic analysis was performed by means of the RM-network method (Bandelt et al. 1995) as implemented in Network 2.0, which includes a new character-compatibility check, and also by the MJ-network method (Bandelt et al. 1999), available at the Life Sciences and Engineering Technology Solutions Web site. The "frequency>1" option was used, which selects those Y types that occur more than once in the data set, for a specified battery of STRs (STR loci in the data set can be deactivated or activated by setting their weight to 0 or >0, respectively). To prevent the small

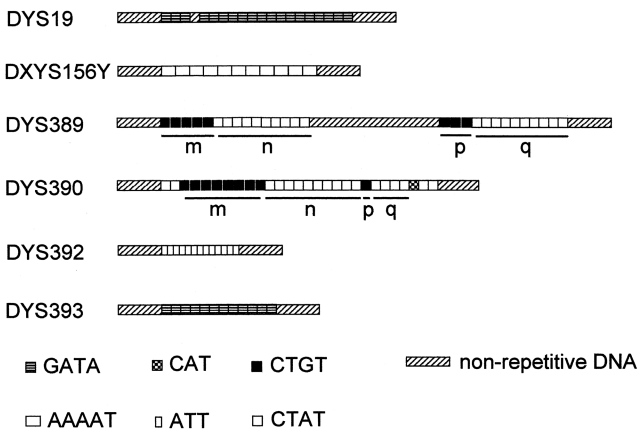

Figure 1 Structure of Y-STR loci typed in the study. The non-repetitive block of DNA in DYS389 is not to scale.

Table 1

Y-STR and Biallelic Polymorphism Types

	No. of														
Individual[a]	DYS19	DXYS156Y	DYS389m	DYS389n	DYS389p	DYS389q	DYS390m	DYS390n	DYS390p	DYS390q	DYS392	DYS393	YAP[b,c]	SRY1532[b]	92R7[b]
H1	14	12	4	12	3	10	8	10	1	4	15	13	0	1	1
H3	15	13	4	13	3	9	8	10	1	4	13	12	0	1	1
H4	15	11	5	11	3	10	8	10	1	4	11	14	0	1	1
H5	17	13	4	11	3	10	7	10	1	4	14	12	0	1	1
H7	13	12	5	12	3	11	8	11	1	4	14	14	0	1	1
H8	16	11	5	12	3	10	8	10	1	4	11	15	0	1	1
H9	16	11	5	11	3	10	8	10	1	4	11	14	0	1	1
H15	15	12	4	12	3	10	8	10	1	4	14	13	0	1	1
H16	16	13	4	12	3	10	8	11	1	4	12	12	0	1	1
H17	15	12	4	12	3	10	8	9	1	4	14	13	0	1	1
H18	16	11	5	11	3	10	8	11	1	4	11	15	0	1	1
H19	17	13	4	12	3	9	8	12	1	4	13	12	0	1	1
H20	16	13	5	10	3	10	8	11	1	4	13	13	0	1	1
H22	17	11	5	12	3	10	8	10	1	4	11	14	0	1	1
H24	15	13	4	12	3	9	8	10	1	4	12	12	0	1	1
H25	16	13	4	14	3	9	8	11	1	4	13	12	0	1	1
H26	15	13	4	13	3	9	8	11	1	4	13	12	0	1	1
H28	16	11	5	12	3	10	8	10	1	4	11	14	0	1	1
H29	16	11	4	14	3	10	8	11	1	4	11	15	0	1	1
H32	14	12	5	11	3	11	7	11	1	4	11	12	0	1	1
H33	15	13	4	12	3	11	8	11	1	4	12	12	0	1	1
H35	15	12	4	12	3	10	9	11	1	4	13	13	0	1	1
H37	15	11	4	12	3	10	8	10	1	4	11	14	0	1	1
H38	16	13	4	14	3	9	8	11	1	4	13	12	0	1	1
H40	16	11	5	11	3	10	8	11	1	4	11	13	0	1	1
H45	16	13	4	13	3	9	7	10	1	4	13	12	0	1	1
H47	15	11	5	11	3	10	8	12	1	4	11	13	0	1	1
H49	14	13	4	12	3	9	8	10	1	4	12	12	0	1	1
H52	14	12	4	14	3	9	8	11	1	4	14	13	0	1	1
H53	16	13	4	13	3	9	8	11	1	4	13	12	0	1	1
H58	15	13	4	12	3	9	8	10	1	4	13	13	0	1	1
H59	13	11	4	13	3	9	8	10	1	4	12	12	0	1	1
H60	15	13	4	13	3	9	8	10	1	4	13	12	0	1	1
H64	15	12	4	12	3	11	9	10	1	4	13	14	0	1	1
H68	13	13	4	14	3	9	8	12	1	4	13	13	0	1	1
J1	14	12	4	12	3	10	8	10	1	4	14	13	0	ND	ND
J2	15	12	4	12	3	10	8	9	1	4	13	13	0	1	1
J5	17	11	5	12	3	12	9	11	1	4	11	13	1L	1	1
J8	15	12	4	12	3	11	9	10	1	4	13	13	0	1	ND
J12	17	11	5	13	3	11	9	10	1	4	11	13	1L	1	1
J13	16	11	5	11	3	11	9	10	1	4	11	13	1L	1	1

J14	17	11	5	12	3	11	9	11	1	4	11	13	1L	1	1
J17	15	12	4	12	3	11	8	9	1	4	13	13	0	1	1
J19	17	11	5	12	3	11	9	11	1	4	11	13	1L	1	1
J20	15	12	4	12	3	11	8	9	1	4	13	13	0	1	1
J21	15	12	4	11	3	11	8	10	1	4	13	13	0	1	1
J27	14	14	4	12	3	8	8	10	1	4	13	12	0	1	1
J29	15	12	4	12	3	11	8	9	1	4	14	13	0	1	1
J30	16	11	4	13	3	10	9	11	1	4	13	13	1	1	1
J31	15	11	5	13	3	11	9	12	1	4	11	13	1L	1	1
J39	15	11	5	11	3	11	9	12	1	4	11	12	1L	1	1
J42	16	13	4	13	3	9	8	12	1	4	13	12	0	1	1
J43	16	11	5	13	3	10	9	10	1	4	11	13	1L	1	1
J45	15	12	4	12	3	10	8	9	1	4	13	13	0	1	1
J53	16	12	5	11	3	11	8	10	1	4	13	13	0	1	1
J54	14	12	4	11	3	11	8	9	1	4	14	14	0	1	1
J56	17	13	4	12	3	9	8	12	1	4	13	12	0	1	1
J60	17	11	5	12	3	10	9	10	1	4	11	13	1L	1	1
J63	16	13	4	13	3	9	8	11	1	4	13	12	0	1	1
J64	17	11	5	11	3	10	10	11	1	4	11	13	1L	1	1
J65	15	11	5	10	3	11	8	10	1	4	11	15	0	1	1
J66	16	11	5	11	3	9	8	10	1	4	11	14	0	1	1
J70	15	12	4	12	3	11	8	9	1	4	13	13	0	1	1
J75	17	11	5	11	3	10	9	9	1	4	11	14	1L	1	1
J76	15	11	5	12	3	11	9	11	1	4	11	13	1L	1	1
J79	17	11	5	12	3	11	9	11	1	4	11	13	1L	1	1
J80	16	13	4	14	3	9	8	13	1	4	14	12	0	1	1
J83	13	12	4	12	3	12	8	9	1	4	14	13	0	1	1
J97	16	13	4	13	3	9	8	11	1	4	13	12	0	1	1
J98	17	11	5	12	3	10	9	11	1	4	11	13	1L	1	1
J102	16	13	4	13	3	9	8	11	1	4	13	12	0	1	1
J107	17	13	4	14	3	9	8	10	1	4	13	12	0	1	1
J115	16	13	4	13	3	9	8	10	1	4	12	12	0	1	1
J118	15	12	4	13	3	11	8	10	1	4	13	13	0	1	1
J120	16	12	4	11	3	11	7	10	1	4	13	13	0	1	1
J125	17	11	5	12	3	10	10	11	1	4	11	13	1L	1	1
J127	17	11	5	12	3	11	9	11	1	4	11	13	1L	1	1
J129	15	12	4	11	3	11	8	9	1	4	13	13	0	1	1
J130	17	13	4	13	3	10	8	11	1	4	13	13	0	1	1
Ov1	15	11	6	12	3	10	8	8	1	4	11	13	1	1	1
Ov4	15	11	6	12	3	10	8	8	1	4	11	13	1	0	1
Ov5	15	11	6	12	3	9	8	8	1	4	12	13	1	1	1
Ov13	16	11	6	10	3	10	8	8	1	4	11	13	1	1	1
Ov17	16	11	6	10	3	10	8	8	1	4	11	15	1	1	1
Ov21	14	11	6	10	3	9	8	12	1	4	11	13	1L	1	1

(*continued*)

Table 1 Continued

Individual[a]	No. of DYS19	DXYS156Y	DYS389m	DYS389n	DYS389p	DYS389q	DYS390m	DYS390n	DYS390p	DYS390q	DYS392	DYS393	YAP[b,c]	SRY1532[b]	92R7[b]
Ov25	16	11	7	11	3	11	8	11	1	4	11	13	0	1	1
Ov27	14	11	6	9	3	9	8	17	0	0	11	12	ND	1	1
Ov30	16	11	6	11	3	10	8	9	1	4	11	14	1	1	1
Ov40	15	11	6	10	3	10	8	8	1	4	11	14	1	1	1
Ov41	15	11	6	11	3	10	8	8	1	4	11	14	1	1	1
Ov52	15	11	6	12	3	10	8	8	1	4	11	13	1	1	1
Ov54	13	11	6	12	3	7	8	11	1	4	11	14	1	1	1
Ov56	16	11	6	10	3	10	8	8	1	4	11	15	1	1	1
Ov60	15	11	6	11	3	10	8	8	1	4	11	14	1	1	1
Ov65	15	11	6	12	3	9	8	8	1	4	11	13	1	1	1
Ov68	15	11	6	11	3	9	8	8	1	4	11	14	1	1	1
Ov69	16	11	6	11	3	10	8	8	1	4	11	14	1	1	1
Ov71	15	11	6	11	3	10	8	8	1	4	11	13	1	1	1
Ov72	15	11	6	11	3	10	8	8	1	4	11	13	1	1	1
Ov73	15	11	6	12	3	9	8	8	1	4	11	13	1	1	1
Ov74	17	11	6	11	3	10	8	8	1	4	11	15	1	1	1
Ov78	15	11	7	12	3	10	8	11	1	4	11	14	0	1	1
Ov84	15	11	6	12	3	11	8	8	1	4	11	13	1	1	1
Ov87	16	12	5	12	3	9	7	10	1	4	11	12	0	1	1
Ov107	15	11	6	11	3	10	8	8	1	4	11	13	1	1	1
Ov108	15	11	6	12	3	10	8	8	1	4	11	13	1	1	1
Ov181	14	11	6	10	3	9	8	11	1	4	11	13	1L	1	1
P4	14	12	5	12	3	10	8	11	1	4	14	13	0	1	1
P10	14	12	5	12	3	10	8	11	1	4	14	13	ND	ND	ND
P12	14	12	5	12	3	10	8	11	1	4	14	13	ND	ND	ND
P13	14	12	5	12	3	10	8	11	1	4	14	13	ND	ND	ND
P14	14	12	5	12	3	10	8	11	1	4	14	13	ND	ND	ND
P16	14	12	5	12	3	10	8	11	1	4	14	14	ND	ND	ND
P19	14	12	5	12	3	12	8	11	1	4	14	13	ND	ND	ND
P23	15	11	5	11	3	10	8	12	0	0	11	14	ND	ND	ND
P24	15	11	5	13	3	11	8	12	0	0	11	14	0	1	1
P35	15	12	5	12	3	9	8	10	1	4	13	13	ND	ND	ND
P49	15	12	5	12	3	9	8	10	1	4	13	13	ND	ND	ND
P51	14	12	5	12	3	10	8	12	1	4	14	13	ND	ND	ND
P56	14	11	5	11	3	10	8	12	1	4	11	14	ND	ND	ND
P82	15	11	5	11	3	10	8	12	0	0	11	14	ND	ND	ND
P105	15	11	5	11	3	11	8	14	0	0	11	13	ND	ND	ND
T4	14	12	5	12	3	9	8	10	1	4	11	13	0	1	1
T5	15	12	5	12	3	9	8	8	1	4	11	14	0	1	1
T9	14	13	4	13	3	10	8	11	1	4	13	14	0	1	1
T12	15	12	5	15	3	11	8	11	1	4	11	13	0	1	0
T15	16	12	5	13	3	10	8	12	1	4	11	13	0	0	1
T22	13	11	6	11	3	9	8	10	1	4	12	13	1	1	1

T26	14	12	5	11	3	9	8	10	1	4	11	12	0	1	1
T29	13	11	7	12	3	9	8	11	1	4	11	12	1	1	1
T32	17	12	5	13	3	10	8	11	1	4	11	12	0	0	0
T33	14	12	4	11	3	10	8	10	1	4	14	13	0	1	1
T38	14	12	5	11	3	10	8	11	1	4	13	12	0	1	1
T39	15	12	5	11	3	9	8	10	1	4	11	12	0	1	1
T40	14	12	4	13	3	10	8	10	1	4	14	13	0	1	1
T43	13	12	6	11	3	10	8	12	1	4	11	13	1	1	1
T44	14	12	5	11	3	10	8	10	1	4	11	12	0	1	1
T45	14	12	5	13	3	10	8	12	1	4	11	13	0	0	0
T46	14	12	5	12	3	10	8	12	1	4	11	13	0	0	0
T47	15	12	5	11	3	9	8	12	1	4	11	12	0	1	1
T48	15	12	5	12	3	9	8	12	1	4	11	13	0	0	0
T50	13	12	5	11	3	10	8	11	1	4	13	12	0	1	0
T51	15	12	5	12	3	10	8	10	1	4	13	13	0	1	0
T52	14	12	5	12	3	10	8	10	1	4	11	12	0	1	1
T53	16	12	5	13	3	10	8	11	1	4	11	12	0	0	0
T54	14	12	5	11	3	10	8	10	1	4	11	12	0	1	1
T56	16	12	5	11	3	9	8	8	1	4	11	13	0	1	1
T57	14	12	5	13	3	10	8	10	1	4	11	12	0	1	1
T59	13	12	6	11	3	10	8	11	1	4	11	13	1	1	1
T60	15	12	5	12	3	9	8	11	1	4	11	13	0	0	0
T61	15	12	6	11	3	10	8	12	1	4	11	13	1	1	1
T64	13	11	7	12	3	10	8	11	1	4	11	13	1	1	1
T65	14	12	5	11	3	10	8	10	1	4	11	12	0	1	0
T66	13	11	6	11	3	9	8	10	1	4	12	13	1	1	0
T68	14	12	6	11	3	9	8	9	1	4	12	12	1	1	1
T72	13	11	5	14	3	10	8	12	1	4	11	13	0	0	0
T83	17	12	5	13	3	10	8	11	1	4	11	13	0	0	0
T89	14	12	5	11	3	10	8	11	1	4	11	12	0	ND	1
T91	15	12	5	12	3	9	8	11	1	4	11	12	0	1	1
T94	14	13	5	11	3	10	8	15	0	0	13	12	0	1	0
T95	13	12	5	11	3	11	8	9	1	4	11	13	0	1	1
W1	15	12	5	11	3	8	8	11	1	4	11	12	0	1	1
W2	15	12	5	11	3	8	8	11	1	4	11	12	0	1	1
W6	14	11	6	10	3	9	7	13	1	4	11	13	1L	1	1
W8	15	11	5	11	3	11	8	10	1	4	11	15	0	1	1
W9	15	11	6	11	3	10	8	8	1	4	11	14	1	1	1
E2	15	11	6	11	3	11	8	8	1	4	11	16	ND	ND	1
M235	15	11	5	10	3	10	8	9	1	4	11	13	0	1	ND
M236	17	12	4	11	3	9	8	9	1	4	12	12	0	1	1
M237	16	11	5	11	3	10	8	11	1	4	11	13	0	1	1
M238	14	12	4	12	3	11	9	10	1	4	13	14	0	ND	1
M239	14	12	5	11	3	11	8	11	1	4	11	12	0	1	1

(*continued*)

Table 1 Continued

Individual[a]	No. of DYS19	DXYS156Y	DYS389m	DYS389n	DYS389p	DYS389q	DYS390m	DYS390n	DYS390p	DYS390q	DYS392	DYS393	YAP[b,c]	SRY1532[b]	92R7[b]
M240	16	11	5	11	3	10	8	12	1	4	11	13	0	1	1
M241	14	14	4	12	3	9	8	12	1	4	14	12	0	1	1
M242	15	11	5	11	3	11	8	12	1	4	11	13	0	1	1
M243	14	12	4	11	3	10	8	10	1	4	14	13	0	1	1
M244	16	13	4	13	3	11	8	11	1	4	13	12	0	1	1
M245	14	12	5	12	3	10	8	10	1	4	16	13	0	1	0
M246	18	12	5	13	3	10	8	12	1	4	11	13	0	0	0
M247	16	11	5	11	3	10	8	12	1	4	11	13	0	1	1
M248	15	13	4	12	3	9	8	11	1	4	13	12	0	1	1
M249	15	12	5	10	3	10	8	10	1	4	11	13	0	1	1
M250	15	11	5	10	3	10	8	10	1	4	11	14	0	1	1
M251	15	11	5	12	3	8	8	9	1	4	11	14	0	1	1
M252	15	11	5	11	3	10	8	11	1	4	11	13	0	1	1
M253	14	12	4	12	3	11	8	10	1	4	14	14	0	1	1
M254	16	11	5	11	3	10	8	12	1	4	11	13	0	1	1
A32	15	11	5	12	3	11	4	10	1	4	11	13	ND	ND	ND
A34	15	11	5	12	3	11	4	10	1	4	11	13	ND	ND	ND
A38	15	11	5	13	3	11	4	10	1	4	11	13	ND	ND	ND
A39	15	11	5	11	3	10	4	9	1	4	11	13	ND	ND	ND
A40	15	11	5	11	3	11	4	9	1	4	11	13	ND	ND	ND
A42	15	11	5	12	3	10	4	10	1	4	11	13	0	1	1
A43	15	12	5	11	3	10	8	12	1	4	13	13	0	1	ND
A44	15	12	5	13	3	11	8	12	1	4	13	12	ND	ND	ND
A45	15	12	5	11	3	10	8	12	1	4	13	13	ND	ND	ND
A47	15	12	5	12	3	11	8	11	1	4	13	12	ND	ND	ND
A48	17	11	5	9	3	10	8	9	1	4	11	13	ND	ND	ND
A50	15	11	5	12	3	11	4	10	1	4	11	13	ND	ND	ND
A51	15	12	5	11	3	10	8	12	1	4	13	13	ND	ND	ND
A52	14	12	5	11	3	10	8	13	1	4	13	13	ND	ND	ND
A56	15	11	5	12	3	10	8	9	1	4	11	14	ND	ND	ND
A60	15	11	5	12	3	11	4	10	1	4	11	13	ND	ND	ND
A68	15	11	5	11	3	10	8	9	1	4	11	13	ND	ND	ND
A71	16	12	5	12	3	10	8	12	1	4	11	13	ND	ND	ND
A74	14	12	5	11	3	10	8	10	1	4	12	13	ND	ND	ND
A80	16	12	5	13	3	10	8	14	1	4	13	13	ND	ND	ND
A84	15	11	5	12	3	11	4	10	1	4	11	13	ND	ND	ND
A88	16	11	5	12	3	10	8	9	1	4	11	13	ND	ND	ND
A89	15	11	5	11	3	10	8	9	1	4	11	13	ND	ND	ND
B5	13	12	5	10	3	10	7	9	1	4	13	11	ND	ND	ND
B7	12	12	5	10	3	11	7	9	1	4	13	11	0	0	1
B11	12	12	5	10	3	11	7	9	1	4	13	11	ND	ND	ND
B14	16	11	5	11	3	11	8	8	0	4	11	13	0	1	1
B15	13	12	5	10	3	10	7	9	1	4	13	11	ND	ND	ND

B17	15	11	6	11	3	10	8	8	1	4	11	15	ND	ND	ND
G1.24	14	12	5	11	3	10	9	11	1	4	13	13	ND	ND	ND
G1.36	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G1.39	14	12	5	12	3	10	8	11	1	4	15	13	0	1	0
G1.45	15	12	5	12	3	11	8	11	1	4	13	13	0	1	1
G1.51	14	12	5	11	3	10	8	10	1	4	13	13	0	1	0
G1.54	14	12	5	11	3	10	8	10	1	4	13	13	0	1	0
G1.57	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G1.78	14	12	5	13	3	10	8	11	1	4	13	13	0	1	0
G2.03	14	12	5	11	3	9	8	10	1	4	13	13	0	1	0
G2.06	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G2.15	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G2.18	14	12	5	11	3	10	8	10	1	4	13	13	0	1	0
G2.36	15	12	5	11	3	9	8	9	1	4	11	13	0	1	0
G2.39	14	12	5	11	3	10	8	12	1	4	14	13	0	ND	0
G2.63	15	12	5	12	3	10	8	13	1	4	11	13	0	0	0
G2.69	16	12	5	13	3	10	8	11	1	4	11	13	0	1	1
G2.81	15	12	5	12	3	10	8	10	1	4	11	12	0	1	1
G3.03	16	12	5	11	3	9	8	8	1	4	11	14	0	1	1
G3.06	15	12	5	11	3	10	8	12	1	4	14	13	0	ND	ND
G3.18	14	12	5	11	3	11	8	11	1	4	13	13	0	1	0
G3.39	15	12	5	13	3	10	8	12	1	4	11	13	0	ND	0
G4.03	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G4.48	14	12	5	11	3	10	8	11	1	4	13	13	0	ND	0
G4.54	14	12	5	13	3	10	8	10	1	4	13	12	0	1	0
G4.57	16	12	5	13	3	10	8	10	1	4	11	13	0	1	1
G4.78	15	12	5	12	3	9	7	10	1	4	11	12	0	1	1
G5.15	14	12	5	13	3	9	8	9	1	4	11	13	0	1	1
G5.48	14	12	5	11	3	10	8	10	1	4	13	13	0	1	0
G5.51	14	11	5	11	3	10	8	10	1	4	13	13	0	1	0
G5.57	14	12	5	11	3	9	8	9	1	4	11	13	0	1	1
G5.66	15	10	5	11	3	10	8	10	1	4	12	14	0	1	1
G6.45	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G6.54	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G6.57	14	12	5	11	3	10	8	9	1	4	13	13	0	1	1
G6.66	16	12	5	11	3	10	8	12	1	4	11	13	0	0	0
G6.69	14	12	5	11	3	10	8	11	1	4	13	13	0	1	0
G6.75	15	11	6	11	3	9	8	8	1	4	11	14	ND	ND	ND
G6.78	16	12	5	13	3	10	8	12	1	4	11	13	0	1	1
G7.09	14	12	5	11	3	10	8	11	1	4	13	12	0	1	0

[a] H = Han, G = German, T = Turk, J = Japanese, Ov = Ovambo, W = Western Pygmy, E = Eastern Pygmy, A = Australian Aboriginal, P = Papuan, B = San (Bushman), and M= Mongolian.

[b] ND = not determined.

[c] L = long YAP variant.

sample of San and Pygmies from being excluded, they were entered as having frequency 2 throughout the analyses. In all RM analyses the RM reduction threshold was set to the default value of 2; that is, character conflicts (displayed as network reticulations) are resolved into subtrees only if the (equivalent) weight of at least two characters opposes one conflicting character weight (Bandelt et al. 1995, 1999). In the absence of theoretical techniques to meaningfully evaluate the accuracy of intraspecies phylogenies (equivalent to bootstrapping for interspecies phylogenies; see Bandelt et al. 1995), the accuracy of the proposed STR phylogeny was empirically determined by superimposing biallelic markers on it, and the discrepancy was quantified by the ratio of the number of links that have to be utilized twice (i.e., the links that have to be added) to the number of links in the proposed STR tree. Relative time estimates were calculated by means of ρ, the average distance to the node of interest (Morral et al. 1994), as measured in single-repeat differences, and were transformed to absolute time estimates by multiplication with the pedigree and evolutionary mutation rates proposed below. As a phylogeny-based statistic, ρ offers the advantage of being unbiased by demographic processes (Forster et al. 1996). The sampling error of ρ was approximated as $\sqrt{(\rho/n)}$, where n denotes the sample size (Torroni et al. 1998).

Results and Discussion

Locus Sequencing

The structures of the Y-chromosomal STR loci analyzed here are shown in figure 1. Note that we have renamed the subsegments of DYS389 as "DYS389m," "-n," "-p," and "-q" (contra Rolf et al. 1998) and that we have aligned the nomenclature for the subsegments of the structurally similar locus DYS390 as DYS390m, -n,- p, and -q (contra Forster et al. 1998). The loci DYS19, DXYS156Y, DYS392, and DYS393 were characterized by length sizing, whereas the complex loci DYS389 and DYS390 were directly sequenced. The resulting Y types for the African, European, Asian, Australian, and Papuan samples are given intable 1. The sequences for DYS389 confirm the findings of Rolf et al. (1998)—that is, that the DYS389m segment is highly informative for identification of indigenous descent, presumably because the short size of DYS389m (4–6 repeats) slows its mutation rate and thus retains ancient alleles. In contrast, the DYS389n and -q alleles are not continent specific and obscure the phylogenetic information of DYS389m when DYS389 is typed by the popular forensic DYS389I/II system (Kayser et al. 1997). Furthermore, the sequences for the four DYS390 segments confirm and extend our previously published findings (Forster et al. 1998) that DYS390m and DYS390p alleles discriminate Asians, Papuans, and Australians from each other and from Whites and Africans. The value of the four DYS390 segments has been neglected, probably because of ascertainment bias: in Whites, all segments but DYS390n are nearly invariant (table 1).

Locus-Specific Mutation Rates

A character- weighting scheme is required for the phylogenetic analysis of loci that differ widely in their mutation rates, in order to identify parallel mutation events during evolution and, hence, to approach the true tree. To empirically determine a biologically justifiable weighting scheme, we refer to our study of 11,000 parent-offspring meioses for autosomal STR loci (Brinkmann et al. 1998). All these autosomal STRs have tetranucleotide-repeat motifs, except for CD4 which has a pentanucleotide-repeat motif. Our results showed that the STR mutation rate is a function of both the mean number of repeats and the motif length: mutations in the pentanucleotide motif CD4 STR were conspicuously absent in our sample, indicating a lower mutation rate. By analogy, we discarded, in the first analysis, the trinucleotide-repeat locus DYS392 (for experimental justification, see below) and the n and q segments of the longest locus, DYS389. The other loci were then subdivided into two weight classes: the pentanucleotide DXYS156Y and the very short segments DYS389m and -p and DYS390m, -p, and -q were assigned to the higher-weight class, and the other tetranucleotide loci were assigned to the lower-weight class. The two classes were weighted by a factor of 2:1, which is the most conservative value when the RM-network algorithm is used, given integer weights and the inherent reduction threshold of 2: network reticulations are resolved into subtrees only if at least two characters (or, more precisely, the equivalent weight) oppose a conflicting character (Bandelt et al. 1995).

Phylogeographic Congruence

The resulting network is shown in figure 2. Because of the elimination of singly occurring Y types, it is based on 187 of the total of 256 individuals. The effectiveness of the weighting is apparent in that DXYS156Y, DYS389m, and DYS390m are represented by 3–9 single-repeat links in the network, whereas the putatively fast loci DYS393, DYS19, and DYS390n are represented by 16–32 links. The length of the estimated MP tree within this network of 89 single-repeat links is 65 single-repeat mutation steps, as estimated by the Network2.0b program. The network assigns the Y types into geographically grouping branches, with the Asian Y types particularly well differentiated from the others, whereas the distinction between Whites and Australian/Papuans is

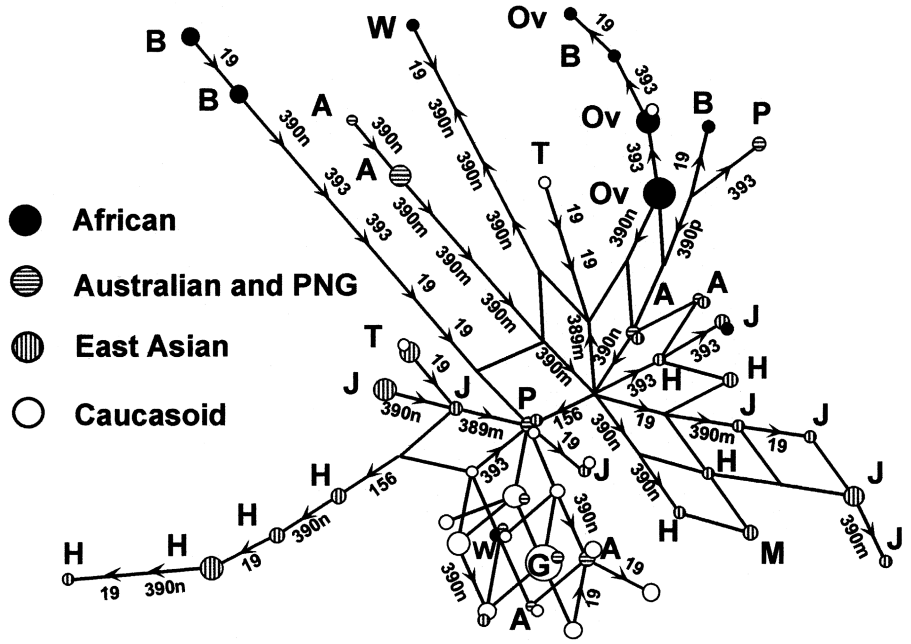

Figure 2 RM network of 187 Y chromosomes from a global sample. The network contains 45 Y-STR types represented more than once in a sample of 256 Y chromosomes (125 Y types). The STR loci were weighted as follows: DYS19, DYS390n and -p, and DYS393, weight 1; DXYS156Y, DYS389m, and DYS390m, weight 2. The length of an MP tree in this network is estimated at 65 single-repeat mutation steps. Note that the placement of the DYS390p branch is ambiguous, since it depends on the coding of DYS390n, -p, and -q. Links are labeled by loci, with the arrows pointing in the direction of locus lengthening. Parallel links in a reticulation represent the same locus. The area of a given circle is proportional to number of individuals. For clarity, one superfluous node in the cluster of White individuals has been omitted. Single-letter abbreviations are as intable 1. For reference, the Papuan in the center of the network has the following repeat lengths: 15 (DYS19), 12 (DXYS156Y), 5 (DYS389m), 8 (DYS390m), 10 (DYS390n), 1 (DYS390p), and 13 (DYS393).

less well resolved, being based only on the fast locus DYS19. Within Y-chromosomal groups, the White Y types cluster closest together, including the majority of Turks, who are nevertheless thought to be partly Asian (Rolf et al. 1999); note the Turkish Y type in the otherwise exclusively Asian branch. The most extreme outliers are represented by Namibian San Y types. If the latter are taken as outgroups for rooting, then the center of the network is indicated as the root, although the central reticulation and the unstructured San branch prevent the pinpointing of the exact node. These geographic plausibilities in the network motivate further testing of the network by means of the set of ancient biallelic markers.

Congruence with Biallelic Markers

Among the oldest and most widespread biallelic markers are the derived alleles of YAP, SRY-1532, and 92R7 (Jobling et al. 1997), and we chose these deep markers as a demanding test for the STR tree, since the STRs would have had ample time to undergo parallelisms if their mutation rate were indeed as high as pedigree studies suggest. These three "biallelic" loci have the resolution of five loci, because SRY-1532 has mutated twice independently in the human Y-chromosome tree (Jobling et al. 1997) and because we have distinguished the long from the short variants of the YAP polymorphism. The superposition of the biallelic markers on the STR network is shown in figure 3: the network does indeed offer a tree replicating the biallelic marker tree. In the case of the YAP marker, the STR tree indicated within the STR network not only separates YAP^- from YAP^+ Y types but also groups together the short YAP variants and suggests that the long YAP variants are ancestral. With SRY-1532 and 92R7, the STR network correctly separates the two independent SRY-1532 mutations in Africans and non-Africans, known, on the basis of the biallelic tree, to be distinct (Jobling et al. 1997). The root of the Y-STR phylogeny is indicated by the SRY-1532 polymorphism (Jobling et al. 1997) and is found in the San, which is in agreement both with the mtDNA trees of Vigilant et al. (1991) and of Watson et al. (1997), which also show that the San harbor phylogenetically deep mtDNA branches, and with the African Y-chromosomal study by Scozzari et al. (1999).

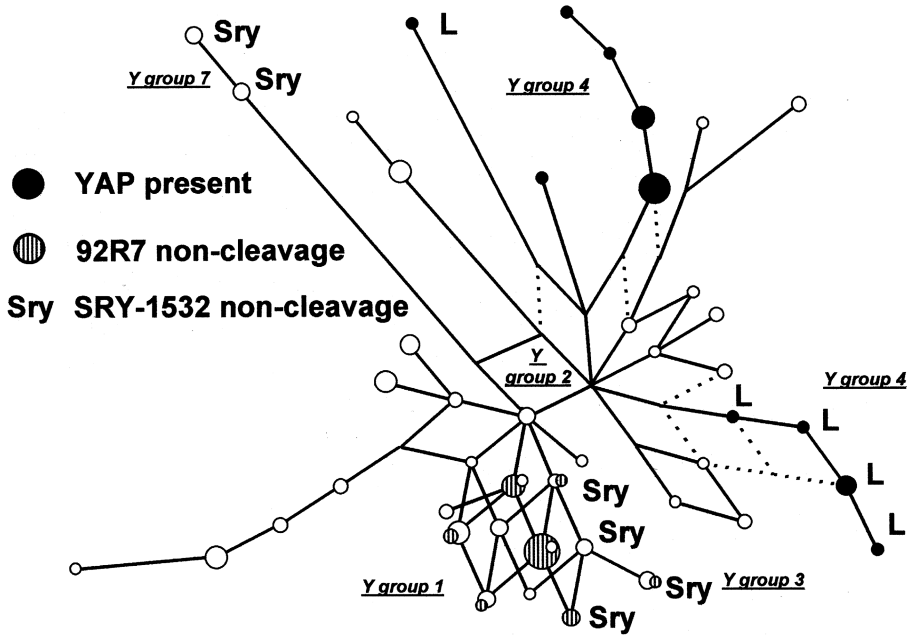

Figure 3 Superposition of biallelic markers on the STR phylogeny. For orientation, the Y-type nomenclature of Jobling et al. (1997) has been provided. Superfluous links not confirmed by the biallelic markers are denoted by the dotted lines. L = long YAP variant.

Furthermore, the SRY-1532 A and 92R7 T types are correctly grouped together in the White sample. However, this cluster is not well resolved internally by the STRs, as is evident from the imperfect separation, by the fast DYS19, of Australians/Papuans from Whites. The phylogenetic error revealed by the biallelic loci can be quantified by counting the number of links that have to be traversed twice in the proposed tree, which amounts to 4 single-mutation links in a total tree length of 65 single-mutation links. The addition of Jobling et al.'s biallelic Y nomenclature to the STR network demonstrates that the Y STRs resolve the deepest node of the biallelic markers (the default group 2 of Jobling et al. 1997) more finely by virtue of DXYS156Y.

To further justify our decision to discard from the analysis the trinucleotide motif STR DYS392, we included it as a low-weight character in the network calculation and otherwise retained the settings noted above. The network (not shown), although superficially tree-like, incorrectly splits the unique mutation events at 92R7 and at SRY-1532 into at least two artifactual events each and produces another obvious artifact in that the San branch off a peripheral German node. It is evident that DYS392 has mutated, at least twice, from 11 to 13 repeats in Y group 1, violating the single-repeat mutation model, in agreement with table 3 of Hurles et al. (1999). Intriguingly, DYS392 appears to be quite stable in Africans (table 1), where it is generally represented by short alleles, whereas in Native Americans (fig. 4), where it is represented by alleles >13 repeats, it is hypervariable but nevertheless has not been observed to transgress the 13–11 repeat jump. The DYS392 allele of 12 repeats is rare anywhere in the world (Kayser et al. 1997). The unusual mutation mechanism of DYS392 is not explicable by its simple sequence structure, and further research will be necessary for an understanding of the phenomenon.

Calibration of the Y-STR Clock

The reconstruction, by means of the Y-STR phylogeny, of deep-branching nodes separating ancient human groups (figs. 2 and 3) appears to contradict the expectation (de Knijff et al. 1997), based on pedigree mutation rates, that Y-STR lines of descent cannot be inferred at >49,000 years (which is actually an upper estimate, since loss of resolution in the reconstruction of ancestral states necessarily precedes loss in genetic-distance resolution as measured by D_{dm} linearity). We can test the plausibility of the pedigree rate ($\sim 3 \times 10^{-3}$/20 years/locus; see Kayser et al. 1997; Bianchi et al. 1998; Jobling et al. 1999) by applying it to the Native American Y tree (fig. 4) for the data of Bianchi et al. (1998), which were preselected for the presence of the pan-American–specific DYS199 variant. The widespread occurrence in America of the DYS199 T variant—and the failure, despite intensive screening, to find it in Asia (Karafet et al. 1997; Santos et al. 1999, but note the mishap in theirt-

able 1)—could be due to either its presence in a founding gene pool for all of America or to later geographic expansion of that Y type across America. The latter possibility is less likely, given the geographically specific derived branches in figure 4 (recent expansion should have affected both the ancestral and the derived types). If we assume that the sharing of DYS199 T by many tribes is due to descent from a common Y founder, then archaeological and mtDNA evidence suggests that this Y type should be ~20,000 years old (Morell 1990; Forster et al. 1996); if, on the other hand, we assume a more recent pan-American expansion of the DYS199 T type, then it is unlikely to postdate the distinct Clovis, Monte Verde, and Monte Alegre cultures of 14,000 calendar years ago (Roosevelt et al. 1996; Saunders 1998). However, application of the pedigree mutation rate yields much younger ages; measuring the ρ diversity value (in a manner analogous to that of Forster al. 1996) from figure 4 yields an age of only 1,775 years, and ρ for each locus separately yields 1,300 years (DYS19), 3,700 years (DYS390), 800 years (DYS391), and 1,300 years (DYS393).

To obtain an alternative "evolutionary" Y-STR mutation rate, we assume an age of 20,000 years for the Native American DYS199 C-T mutation. The phylogenetic mutation rates for each Y STR, determined from the proposed Native American Y tree of figure 4, are given in table 2. The average rate for the tetranucleotide-motif loci (i.e., when DYS392 is excluded) is $2.6 \times 10^{-4} \pm 2.7 \times 10^{-5}$/20 years/locus ($\rho = 92/89$ for the four loci) and, hence, is an order of magnitude slower than the pedigree mutation rate—and is even outside the 95% confidence interval given by Heyer et al. (1997). There are two possible factors that could explain the discrepancy. First, the Y STRs used in pedigree studies include the fast type omitted in our calibration (i.e., DYS389n and -q and DYS392); if we combine the pedigree data of Kayser et al. (1997), Bianchi et al. (1998), and Jobling et al. (1999) and consider only the four putatively slow tetranucleotide loci used in our DYS199 T calibration (i.e., DYS19, DYS390, DYS391, and DYS393), we obtain a mutation rate of 3/2,651 for the four loci—that is, $2.8 \times 10^{-4} \pm 3 \times 10^{-4}$/generation/locus. Although the sample size is unsatisfactory, this rate is clearly compatible with our calibration. Second, the

Table 2

Diversity in ρ, for Native American Y-STR Calibration

Locus	Repeats in Founder	ρ	Mutations/20 Years	Sampling Error
DYS19	13	17/89	.00019	.000046
DYS390	24	49/89	.00055	.000079
DYS391	10	11/89	.00012	.000037
DYS392	14	54/89	.00061	.000083
DYS393	13	15/89	.00017	.000044

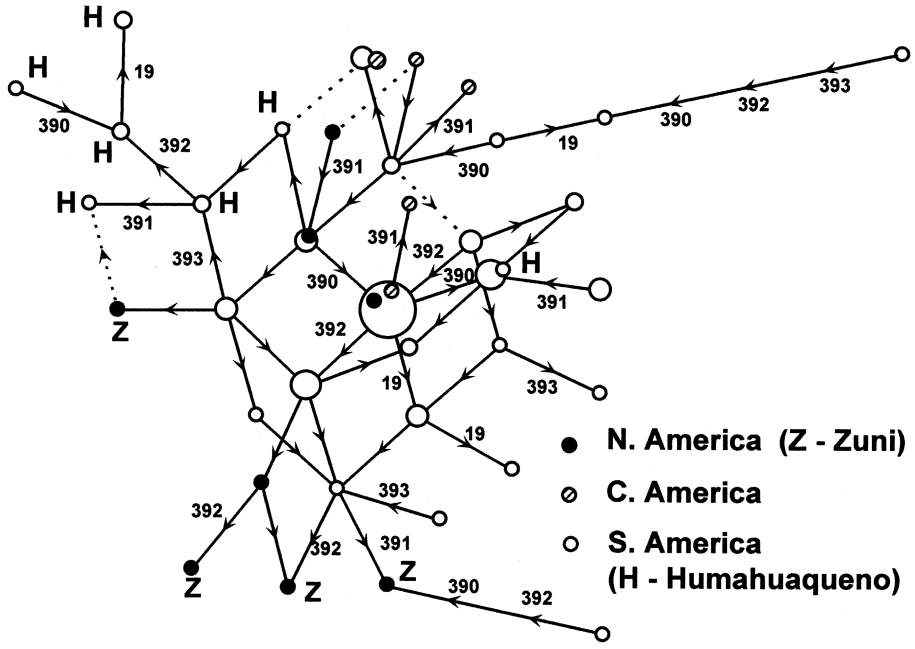

Figure 4 Network of Native American Y types, from Bianchi et al. (1998). The network includes all 89 individuals and was calculated by the MJ method ($\epsilon = .6$), to resolve the high homoplasy caused by DYS392. The weighting scheme was as follows: DYS19, 1; DYS390 (total), .6; DYS391, 1; DYS392, .6; and DYS393, 1. Preprocessing the data with RM before applying MJ produces an identical network.

pedigree studies are performed on fathers in paternity cases who, in Germany, are, on average, >30 years old (Brinkmann et al. 1998). Since mutation rate increases with paternal age, modern paternity cases may not be representative of prehistoric fathers, who may have been younger. A factor that is not taken into account in our calibration is the dependence of mutation rate on allele length, since short alleles are expected to mutate more slowly (Brinkmann et al. 1998). This is seen, for example, not only in the allele-dependent mutational behavior of DYS392, which has been discussed above, but also in the low variability of the short Native American DYS19 allele (Karafet et al. 1997; Carvalho-Silva et al. 1999).

Practical Guidelines

The problems of phylogenetic Y-STR analysis identified and addressed in this study can be summarized as follows. At the laboratory level, it is crucial to type the complex loci DYS389 and DYS390, either by sequencing or by nested PCR (Rolf et al. 1998), rather than by conventional length sizing; otherwise, the information content of the slowly mutating "m" segments is swamped by that of the rapidly mutating "n" and "q" segments, with the effect that the MP tree for the data will not approximate the true tree. Pentanucleotide-repeat loci such as DXYS156Y are similarly informative for ancient descent and must be included, whereas the trinucleotide locus DYS392 mutates rapidly at length >13 and is prone to double-repeat mutations from 11 to 13 repeats. It may be argued that trinucleotide and substructured loci should be discarded altogether, and this may be possible in the future; but the former will remain interesting for the study of triplet-expansion mechanisms, and, because of their high-discrimination capability, the latter are currently relevant in forensics. At the data-analysis level, the RM-network method is superior to the MJ-network method, for longer branches (fig. 3), whereas MJ performs best with predominantly one-step networks (fig. 4). A weighting system such as the two-class system presented here is indispensable for phylogenetic recognition of parallel mutations, as is the application (in the case of large data sets) of the "frequency >1" option, which discards singly occurring Y types. Not only does this option reduce homoplasy (Richards et al. 1996), but it also tends to eliminate incorrectly typed loci, assuming that identical nonsystematic mistakes are unlikely. Errors are even more likely to occur than in mtDNA sequencing, because of the separate PCR amplifications required for Y-STR typing. Finally, in the interpretation of results, the published average pedigree mutation rate for STRs will systematically yield ages that are too young either when applied to a selected stable STR set as presented here or when inadvertently applied to deep phylogenies with unidentified parallelisms. Should there remain a residual discrepancy after mutational saturation is taken into account, it would be interesting to investigate the possibility of dating the age of prehistoric fathers by comparing modern Y-chromosomal with modern mtDNA sequences, since paternal age dependence of STR mutations has been demonstrated in previous pedigree analyses (Brinkmann et al. 1998).

Acknowledgments

We thank Heide Benkmann (Hamburg) and Winfried Siffert (Essen) for samples, and we thank Matthew Hurles (Cambridge) and a reviewer for expert advice. T.Z. is supported by the Wellcome Trust, and C.T.-S. is supported by the Cancer Research Campaign.

Electronic-Database Information

The URL for data in this article is as follows:

Life Sciences and Engineering Technology Solutions, http://www.fluxus-engineering.com (for network methods)

References

Anagnostopoulos T, Green PM, Rowley G, Lewis CM, Giannelli F (1999) DNA variation in a 5-Mb region of the X chromosome and estimates of sex-specific/type-specific mutation rates. Am J Hum Genet 64:508–517

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Bandelt H-J, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753

Bianchi NO, Catanesi CI, Bailliet G, Martinez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera RJ, et al (1998) Characterisation of ancestral and derived Y-chromosome haplotypes of New World native populations. Am J Hum Genet 63:1862–1871

Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62:1408–1415

Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC (1991) Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. Am J Hum Genet 48: 137–144

Carvalho-Silva DR, Santos FR, Hutz MH, Salzano FM, Pena SD (1999) Divergent human Y-chromosome microsatellite evolution rates. J Mol Evol 49:204–214

Cooper G, Amos W, Hoffman D, Rubinsztein DC (1996) Net-

work analysis of human Y microsatellite haplotypes. Hum Mol Genet 5:1759–1766

de Knijff P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110:134–149

Forster P, Harding R, Torroni A, Bandelt H-J (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59:935–945

Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, Brinkmann B (1998) Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. Mol Biol Evol 15:1108–1114

Foster EA, Jobling MA, Taylor PG, Donnelly P, de Knijff P, Mieremet R, Zerjal T, et al (1998) Jefferson fathered slave's last child. Nature 396:27–28

Hammer MF (1995) A recent common ancestry for human Y chromosomes. Nature 378:376–378

Hammer MF, Horai S (1995) Y chromosomal DNA variation and the peopling of Japan. Am J Hum Genet 56:951–962

Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6:799–803

Hurles ME, Veitia R, Arroyo E, Armenteros M, Bertranpetit J, Pérez-Lezaun A, Bosch E, et al (1999) Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. Am J Hum Genet 65:1437–1448

Jobling MA, Bouzekri N, Taylor PG (1998) Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (DYF155S1). Hum Mol Genet 7:643–653

Jobling MA, Heyer E, Dieltjes P, de Knijff P (1999) Y-chromosome-specific microsatellite mutation rates re-examined using a minisatellite, MSY1. Hum Mol Genet 8:2117–2120

Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110:118–124

Karafet T, Zegura SL, Vuturo-Brady J, Posukh O, Osipova L, Wiebe V, Romero F, et al (1997) Y chromosome markers and trans-Bering Strait dispersals. Am J Phys Anthropol 102: 301–314

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133, 141–149

Kwok C, Tyler-Smith C, Mendonca BB, Berkovitz GD, Goodfellow PN, Hawkins JR (1996) Mutation analysis of the 2kb 5′ to SRY in XY females and XY intersex subjects. J Med Genet 33:465–468

Macaulay VA, Richards MB, Forster P, Bendall KE, Watson E, Sykes B, Bandelt H-J (1997) mtDNA mutation rates—no need to panic. Am J Hum Genet 61:983–990 (includes reply by Howell and Mackey)

Morral N, Bertranpetit J, Estivill X, Nunes V, Casals T, Gimenez J, Reis A, et al (1994) The origin of the major cystic fibrosis mutation (delta F508) in European populations. Nat Genet 7:169–175

Morell V (1990) Confusion in earliest America. Science 241: 439–441

Nata M, Brinkmann B, Rolf B (1999) Y chromosomal STR haplotypes in a population from north west Germany. Int J Legal Med 112:406–408

Pérez-Lezaun A, Calafell F, Seielstad M, Mateu E, Comas D, Bosch E, Bertranpetit J (1997) Population genetics of Y-chromosome short tandem repeats in humans. J Mol Evol 45:265–270

Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, et al (1996) Palaeolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203

Rolf B, Meyer E, Brinkmann B, de Knijff P (1998) Polymorphism at the tetranucleotide repeat locus DYS389 in 10 populations reveals strong geographic clustering. Eur J Hum Genet 6:583–588

Rolf B, Röhl A, Forster P, Brinkmann B (1999) On the genetic origins of the Turks: study of six Y-chromosomal short tandem repeats. In: Papiha SS, Deka R, Chakraborty R (eds) Genomic diversity: applications in human population genetics. Kluwer Academic/Plenum, New York, pp 75–82

Roosevelt AC, Lima da Costa M, Lopes Machado C, Michab M, Mercier N, Valladas H, Feathers J, et al (1996) Palaeoindian cave dwellers in the Amazon: the peopling of the Americas. Science 272:373–384

Santos FR, Pandya A, Tyler-Smith C, Pena SDJ, Schanfield M, Leonard WR, Osipova L, et al (1999) The central Siberian origin for Native American Y chromosomes. Am J Hum Genet 64:619–628

Saunders NJ (1998) Book review of T. D. Dillehay: Monte Verde—a late pleistocene settlement in Chile. Vol 2: The archaeological context and interpretation. Nature 392: 145–146

Scozzari R, Cruciani F, Santolamazza P, Malaspina P, Torroni A, Sellitto D, Arredi B, et al (1999) Combined use of biallelic and microsatellite Y-chromosome polymorphisms to infer affinities among African populations. Am J Hum Genet 65: 829–846

Seielstad M, Bekele E, Ibrahim M, Toure A, Traore M (1999) A view of modern human origins from Y chromosome microsatellite variation. Genome Res 9:558–567

Siffert W, Forster P, Jockel KH, Mvere DA, Brinkmann B, Naber C, Crookes R, et al (1999) Worldwide ethnic distribution of the G protein beta 3 subunit 825T allele and its association with obesity in Caucasian, Chinese, and Black African individuals. J Am Soc Nephrol 10:1921–1930

Torroni A, Bandelt H-J, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, et al (1998) mtDNA analysis reveals a major late palaeolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62: 1137–1152

Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, et al (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Res 7:996–1005

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507

Watson E, Forster P, Richards M, Bandelt H-J (1997) Mito-

chondrial footprints of human expansions in Africa. Am J Hum Genet 61:691–704

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128

Whitfield LS, Sulston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome. Nature 378:379–380

Wiegand P, Bajanowski T, Brinkmann B (1993) DNA typing of debris from fingernails. Int J Legal Med 106:81–83