Hum Genet (2001) 109 : 659–674
DOI 10.1007/s00439-001-0627-9

ORIGINAL INVESTIGATION

Michael E. Weale · Levon Yepiskoposyan
Rolf F. Jager · Nelli Hovhannisyan · Armine Khudoyan
Oliver Burbage-Hall · Neil Bradman
Mark G. Thomas

Armenian Y chromosome haplotypes reveal strong regional structure within a single ethno-national group

Received: 14 September 2001 / Accepted: 14 September 2001 / Published online: 30 October 2001

Abstract Armenia has been little-studied genetically, even though it is situated in an important area with respect to theories of ancient Middle Eastern population expansion and the spread of Indo-European languages. We screened 734 Armenian males for 11 biallelic and 6 microsatellite Y chromosome markers, segregated them according to paternal grandparental region of birth within or close to Armenia, and compared them with data from other population samples. We found significant regional stratification, on a level greater than that found in some comparisons between different ethno-national identities. A diasporan Armenian sub-sample (collected in London) was not sufficient to describe this stratified haplotype distribution adequately, warning against the use of such samples as surrogates for the non-diasporan population in future studies. The haplotype distribution and pattern of genetic distances suggest a high degree of genetic isolation in the mountainous southern and eastern regions, while in the northern, central and western regions there has been greater admixture with populations from neighbouring Middle Eastern countries. Georgia, to the north of Armenia, also appears genetically more distinct, suggesting that in the past Trans-Caucasia may have acted as a genetic barrier. A Bayesian full-likelihood analysis of the Armenian sample yields a mean estimate for the start of population growth of 4.8 thousand years ago (95% credible interval: 2.0–11.1), consistent with the onset of Neolithic farming. The more isolated southern and eastern regions have high frequencies of a microsatellite defined cluster within haplogroup 1 that is centred on a modal haplotype one step removed from the Atlantic Modal Haplotype, the centre of a cluster found at high frequencies in England, Friesland and Atlantic populations, and which may represent a remnant paternal signal of a Paleolithic migration event.

M.E. Weale (✉) · R.F. Jager · O. Burbage-Hall · N. Bradman
M.G. Thomas
Centre for Genetic Anthropology,
Departments of Biology and Anthropology,
University College London, University of London,
Darwin Bdg, Gower St, London WC1E 6BT, UK
e-mail: m.weale@ucl.ac.uk,
Tel.: +44-207-4043040, Fax: +44-207-4042081

L. Yepiskoposyan · N. Hovhannisyan · A. Khudoyan
Institute of Man, Yerevan, Armenia

R.F. Jager
Faculteit Biologie, Vrije Universiteit Amsterdam,
Amsterdam, The Netherlands

Introduction

Armenians have a strong and distinct ethnic and cultural identity that unites them as an ethno-national group. The present-day country (size approx. 30,000 km^2, population approx. 3.7 million) is situated in southern Trans-Caucasia between the Black and Caspian Seas at the boundaries of the Middle East, Northern Asia and Central Asia, although many self-identified Armenians continue to live in neighbouring countries or did so until recently (Fig. 1). Armenia occupies an important location in the context of theories of early human population expansion and language development. Neolithic farming in Western Asia began between 8000 and 6000 BC in the Fertile Crescent some 500 km to the south, initiating a major but uneven population expansion that may have spread to other parts of Asia, including the Indian sub-continent and Europe (Cavalli-Sforza et al. 1994). Archaeological evidence suggests that farming may have started in Armenia within the same period (Kushnareva 1990), with an increase in the local density of settlements occurring primarily in the Early Bronze Age (Kuro-Araxian culture) c. 3500–2500 BC (Badalyan 1986). It has been suggested that cranial similarities between modern Armenians and Armenian inhabitants of 1600–700 BC indicate a genetic continuity with ancient populations (Movsessyan and Kotchar 2000).

The Armenian language is an isolated branch, with uncertain affiliation, of Indo-European, the language group spoken today in most of Europe and east of Armenia throughout Iran, Afghanistan, Pakistan and India (Djahukian 1987). The origins of the hypothesised Proto-Indo-

Fig. 1 Map of Armenia, including definition of the regions "Ararat", "North", "West", "Syunik" and "Karabakh". The "Iranian" region covers a wider area to the southeast of this map

European language remain controversial. While the first records of Indo-European languages appear in western Anatolia c. 1900–1700 BC (Hittite, Palaic, Luwian), the Proto-Indo-European homeland has been variously placed in the Ukraine (Mallory 1989), Anatolia (Renfrew 1987) and Armenia (Gamkrelidze and Ivanov 1984) among others. The relative role of the Balkans (west of the Black Sea) and Trans-Caucasia (east of the Black Sea) as routes for early migrations that would have spread Indo-European languages to the north or south remains uncertain (Mallory 1989).

The first evidence of Indo-European speaking people in the Armenian region dates to between 1300 and 700 BC. These people eventually replaced the non-Indo-European speaking Hurrians and later Urartians by 600 BC (Bournutian 1993; Hovannisian 1997; Redgate 1998). The Kingdom of Armenia reached its greatest extent by the first century BC, stretching southwest from present-day Armenia to the northeastern Mediterranean. In 301 AD Armenia became the first country to adopt Christianity as the state religion. For most of the period from the first century AD to the present day Armenia has been subject to the hegemony of more powerful neighbours, although a notable exception was the Armenian Bagratid dynasty of the ninth to eleventh centuries. External powers that have ruled or exerted dominant political influence over Armenia include the Romans, Parthians (and later Persians), Byzantium, Seljuk Turks, Mongols (thirteenth to early fifteenth centuries), the Ottoman and Russian Empires, and most recently (until 1991) the Soviet Union. Forced and voluntary dispersions over the years have led to a large worldwide Armenian diaspora.

The paternally inherited non-recombining portion of the Y chromosome has over the past few years become increasingly useful in the study of human prehistory (for example, Kayser et al. 2001; Malaspina et al. 2000; Rosser et al. 2000; Semino et al. 2000; Thomas et al. 2000; Underhill et al. 2000). It can be expected in time to provide the most accurately known human gene genealogy because it is the largest stable non-recombining portion of the genome (approx. 35 Mb of euchromatic DNA) with a large number of both slowly and rapidly evolving markers. Slowly evolving biallelic Unique Event Polymorphisms (UEP) allow almost unequivocal identification of descendants of single common ancestors. More rapidly evolving microsatellites allow more accurate inferences to be made on the timing of genetic and demographic events. Modern screening techniques allow rapid characterisation of both UEP and microsatellite markers in large population samples and can be performed on DNA obtained from mouth swabs rather than blood samples, facilitating data collection (Thomas et al. 1999; Underhill et al. 1997). There is evidence that Y chromosome population stratification may be found on a finer geographic scale than autosomal and mitochondrial variation, making it useful for local discrimination studies (Jorde et al. 2000; Pérez-Lezaun et al. 1999; Seielstad et al. 1998).

We typed DNA from 734 Armenian males, collected from four regional collection areas and one diasporan location (London, UK), for 11 biallelic and 6 microsatellite Y chromosome markers and compared the data with Y chromosome haplotypes from samples collected in neighbouring and more distant countries. Sufficiently large data sets were collected to ask the following questions relevant to Armenia's long recorded history and important geographic location: (a) Are Armenians regionally stratified, despite their ethnic unity and ancestral geographic proximity, and if so to what extent? (b) What are the implications of stratification for interpreting Armenian demographic history? (c) How do Armenian Y chromosome haplotype distributions compare with the distributions in samples from neighbouring populations, and what historical inferences can be made? (d) How do Armenian Y chromosome haplotype distributions compare with the distributions in samples from more distant populations, especially with regard to the ancient peopling of Europe? (e) Can signals of population growth be detected and dated? We also addressed an additional question: (f) Can a sample taken from a diasporan community (living in London) adequately describe Armenian Y chromosome diversity as a whole? Since samples from diasporan or displaced ethnic groups are sometimes easier to collect than samples from their original geographic locations, we wished to test whether this sampling strategy could be considered reliable in future anthropological or epidemiological genetic studies.

Subjects and methods

Subjects

Mouth swabs from 741 informed consenting self-identified ethnic Armenian males, unrelated at the paternal grandfather level, were collected anonymously between 1997 and 1999 at four regional

collection areas: Yerevan (the capital; *n*=150), northern Armenia (the towns of Gyumri and Vanadzor; *n*=150), southern Armenia (the town of Goris; *n*=150), the Karabakh region in Azerbaijan (an area of territorial dispute between Armenia and Azerbaijan; *n*=200), and also from diasporan Armenians living in London, UK (*n*=91). Seven samples were later discarded as Y chromosome screening was unsuccessful. Subjects were asked to name the birthplace and cultural identity of themselves and their immediate ancestors, including their paternal grandfather.

Samples were assigned to six regions according to paternal grandparental place of birth (Fig. 1): "Ararat", the valley region surrounding the capital, Yerevan, to the east of the Aras river; "North", northern Armenia plus three districts in southern Georgia (Bolnisi, Akhalkalaki and Akhaltsikhe) and one in northwestern Azerbaijan (around Gyanja); "Syunik", a mountainous region of southern Armenia; "Karabakh", a mountainous enclave within Azerbaijan; "Iranian", within present-day Iran, believed to be mainly descendants of Armenians removed to Isfahan (central Iran) from Julfa (see Fig. 1) by Shah Abbas I in 1604 AD; "West", the area of eastern Turkey historically part of Greater Armenia. All regions are or were historically areas with large ethnic Armenian populations.

Comparative data sets

The Armenian samples were compared with samples collected from other countries for the same Y chromosome markers: "Turkey", 173 students of Istanbul University; "Azerbaijan", 29 residents of Baku (the capital); "Syria", 44 students of Damascus University; "Georgia", 68 students resident in Tbilisi (the capital); "Greece", 132 residents of Athens (the capital); "Mongolia", 402 army recruits, primarily of Khalkh ethnic origin; "England", 310 residents of five market towns in the Midlands and East Anglia (Ashbourne, Southwell, Bourne, Fakenham and North Walsham) that form a rough east-west transect across central England with 210 km separating the outermost towns; "Friesland", 94 residents of Dutch Friesland. All comparative data were obtained from informed consenting volunteers and were collected anonymously. The data are unpublished and currently undergoing study. The location of paternal or grandpaternal birthplace was generally geographically extensive within a given country (restricted to mainland Anatolia in the case of "Turkey"). However, we recognise that the sample collection protocols for most of these comparative data sets were not as rigorous as that for the Armenians, that present-day political boundaries do not necessarily coincide with ethnic boundaries, and that the labels "Turkey", "Azerbaijan", etc., should therefore only be taken as convenient indicators of the general geographical location of these samples.

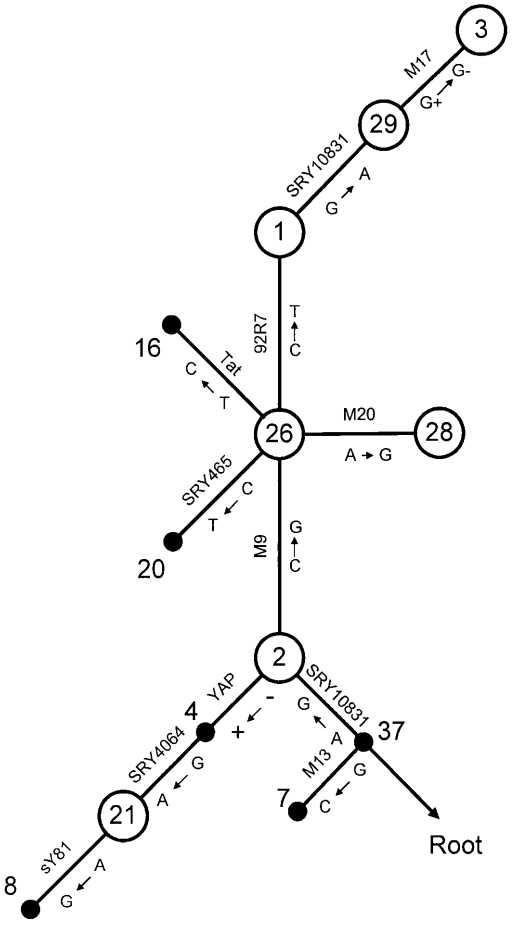

Fig. 2 Y chromosome haplogroup network defined by the 11 UEP markers used in this study. Haplogroup numbers follow a nomenclature expanded from that of Vogt et al. (1997) and Rosser et al. (2000). *Open circles* haplogroups found in the Armenian sample; *closed circles* those that were not; *branches* the inferred UEP mutation events and the mutational change involved (the back-mutation at SRY10831 has been resolved by maximum parsimony using other markers, including ones not shown here); *Koot* human Y chromosome common ancestor, deduced by comparison with other great ape species [Underhill et al. (2000) and P. Underhill (personal communication) for position of SRY10831]

Molecular analysis

All samples were screened for 11 biallelic UEP markers. These UEP markers comprise the following: nine base-pair substitutions – 92R7 (Mathias et al. 1994), M9, M13, M20 (Underhill et al. 1997), sY81 (Seielstad et al. 1994), SRY+465 (T. Shinka and Y. Nakahori, personal communication), SRY4064 (also termed SRY-8299), SRY10831 (also termed SRY-1532; Whitfield et al. 1995) and Tat (Zerjal et al. 1997); one single basepair deletion, M17 (Underhill et al. 1997); and one *Alu* insert, YAP (Hammer 1994). Samples were also screened for six microsatellite markers: four tetranucleotide repeats, DYS19, DYS390, DYS391 and DYS393; and two trinucleotide repeats, DYS388 and DYS392 (Jobling and Tyler-Smith 1995). This screening was carried out using three multiplex PCR kits and Genescan technology (Thomas et al. 1999). Microsatellite repeat sizes were assigned according to the nomenclature of Kayser et al. (1997). UEP-defined haplogroups (so-called because they allow the classification of microsatellite haplotypes within a UEP-defined genealogy) were assigned using a nomenclature expanded from that of Vogt et al. (1997) and Rosser et al. (2000), and detailed in Fig. 2. In this study we found that the marker M17 splits the old haplogroup 3 (hg3) into two further subgroups. We retain the name hg3 for M17"G–" individuals and assign the new name hg29 to M17"G+" individuals, as the latter are much rarer in the present study. Samples were re-screened to resolve missing, ambiguous, or unlikely data. Seven samples (0.9% of the 741 screened) failed to yield complete or consistent results and were discarded.

Statistical analysis

Genetic diversity, *h*, and Nei's genetic identity, *I*, were estimated from unbiased formulae given in Nei (1987). Genetic distances F_{ST} and R_{ST} were estimated from analysis of molecular variance (AMOVA) Φ_{ST} values with the aid of the Arlequin program (Excoffier et al. 1992; Michalakis and Excoffier 1996; URL: http://anthropologie.unige.ch/arlequin). Confidence intervals for these statistics were constructed using bootstrap estimates of standard errors, based on resampling haplotypes according to observed population frequencies. Tests for significant population differentiation were carried out using the exact test for population differentiation

of Raymond and Rousset (1995). Principal coordinates analysis was performed on similarity matrices calculated as one minus genetic distance (F_{ST} or R_{ST}). Values along the main diagonal, representing the similarity of each population sample to itself, were calculated from the estimated genetic distance between two copies of the same sample. For AMOVA-based F_{ST} and R_{ST} distances, the resulting similarity of a sample to itself simplifies to $n/(n–1)$, where n is the sample size.

We carried out full-likelihood Bayesian inference of genetic and demographic parameters under population splitting and growth using the BATWING program (Bayesian analysis of trees with internal node generation; URL: http://www.maths.abdn.ac.uk/~ijw), extended from the algorithm presented by Wilson and Balding (1998). BATWING uses a Markov chain Monte Carlo procedure to generate a sequence of genealogical and population trees, with associated model parameter values, consistent with the genetic data observed in a sample of individuals. At equilibrium the sequence of trees correctly samples from the posterior probability distribution of trees given the observed data and the assumed underlying genetic and demographic model. The extended BATWING version used here assumes an unbounded single stepwise mutation model (Moran 1975) for the microsatellite loci and a coalescent process under an exponential model of population growth from an initially constant-size population. UEP mutations were used only to condition the space of permissible trees, assuming the network in Fig. 2 and a known root. Population trees superimposed on the genealogy were modelled under a strict binary fission process of an initially panmictic population into a series of isolated sub-populations with no subsequent migration. Locus-specific priors for the microsatellite mutation rate per generation were based on data from three published studies (Bianchi et al. 1998; Heyer et al. 1997; Kayser et al. 2000), restricted to the same microsatellite loci as those we used (DYS19, DYS390, DYS391, DYS392 and DYS393). As no data exist for DYS388 in these studies, this locus was excluded from the BATWING analysis. The observed mutational events together with the number of observed meioses were combined with a standard exponential pre-prior. The priors for each locus were: DYS19 set as gamma(3, 1459) (2.5%, 50%, 97.5% quantiles: 0.00042, 0.0018, 0.0050); DYS390 set as gamma(5, 929) (2.5%, 50%, 97.5% quantiles: 0.0017, 0.0050, 0.011); DYS391 set as gamma(3, 878) (2.5%, 50%, 97.5% quantiles: 0.00070, 0.0030, 0.0082); DYS392 set as gamma(2, 878) (2.5%, 50%, 97.5% quantiles: 0.00028, 0.0019, 0.0063); and DYS393 set as gamma(1, 878) (2.5%, 50%, 97.5% quantiles: 0.000029, 0.00079, 0.0042). Weakly informative priors were also given to other parameters to aid in the convergence of the MCMC process. The initial effective population size was given a gamma(1, 0.0001) prior (2.5%, 50%, 97.5% quantiles: 371, 7903, 38960), which covers the values commonly assumed for the global Y chromosome effective population size as well as lower values to compensate for this being a regional sample and for representing the effective size before growth. The population growth rate r per generation was given a gamma(1.01, 1) prior (2.5%, 50%, 97.5% quantiles: 0.026, 0.703, 3.71), which is a very flat prior giving support to much lower r values than that implied by the 2.5% quantile, and covers estimates of real (census) population growth in various parts of the world over the past few thousand years (Cavalli-Sforza et al. 1994) as well as supporting lower values to allow for growth in effective population size being plausibly lower than real growth. All other parameters were given flat, uninformative priors. Generation time was set at 25 years.

For analysis of duplicated DYS19 samples (see "Results") in BATWING and calculation of R_{ST} and repeat size variance statistics, we used the lower of the two repeat sizes (Goldstein et al. 1996).

Results

Y chromosome haplogroups and haplotypes

The 11 UEP markers that were typed on all samples define seven haplogroups within the Armenian data set (Fig. 2, Table 1). Regional differences within Armenia in the common haplogroups (hg1 and hg2) are highly significant, with the frequency of hg1 in Syunik and Karabakh approximately twice that in Ararat, the North and West (exact test for regional differences at the haplogroup level: $P<0.001$).

The six microsatellite loci in addition define 253 compound UEP+microsatellite haplotypes (Table 2). Regional differences within Armenia at the UEP+microsatellite haplotype level are also highly significant (exact test: $P<0.001$). Homoplasy of microsatellite haplotypes across UEP haplogroups was low (5/253=2.0% of haplotypes, two or fewer chromosomes in one of the two haplogroups involved in each case), consistent with other studies using similar numbers of microsatellite loci (Bradman et al. 2000; Malaspina et al. 1998; Scozzari et al. 1999; Thomas et al. 2000).

Eight cases of duplication at the DYS19 locus were detected (8/734=1.1% of sample), all within hg2 (Table 2), and further "silent" cases are likely to be present where the repeat sizes at the two DYS19 loci are not differentiated. Duplication at the DYS19 locus has been reported previously (Kayser et al. 1997, 2000; Santos et al. 1996). The eight cases form a connected single-step microsatel-

Table 1 Y chromosome haplogroup frequencies in six Armenian regions (determined by paternal grandparental birthplace) and eight comparative data sets

	Ararat (n=44)	North (n=189)	Syunik (n=140)	K'bakh (n=215)	Iranian (n=56)	West (n=90)	Turkey (n=173)	Azer. (n=29)	Syria (n=44)	Georgia (n=68)	Greece (n=132)	Mongol. (n=402)	England (n=310)	Friesl. (n=94)
hg1	0.2273	0.2222	0.4000	0.4279	0.3214	0.2222	0.1908	0.1034	0.0909	0.0882	0.1515	0.0547	0.6323	0.5532
hg2	0.5909	0.5873	0.4286	0.4233	0.5179	0.5778	0.4913	0.5862	0.6136	0.8088	0.4848	0.6244	0.3000	0.3511
hg3	–	0.0423	0.0929	0.0558	0.0179	0.0333	0.1040	0.1034	0.0227	0.0441	0.0606	0.0249	0.0290	0.0745
hg4	–	–	–	–	–	–	–	–	–	–	–	0.0149	–	–
hg16	–	–	–	–	–	–	0.0116	0.0345	–	–	–	0.0622	–	–
hg20	–	–	–	–	–	–	–	–	–	–	–	0.0075	–	–
hg21	0.0909	0.0688	0.0286	0.0279	0.1429	0.0556	0.0983	0.0690	0.2273	0.0147	0.2652	0.0025	0.0323	0.0213
hg26	0.0455	0.0529	0.0500	0.0512	–	0.0667	0.0578	0.1034	0.0455	0.0441	0.0379	0.2040	0.0032	–
hg28	0.0455	0.0212	–	0.0140	–	0.0333	0.0462	–	–	–	–	0.0050	0.0032	–
hg29	–	0.0053	–	–	–	0.0111	–	–	–	–	–	–	–	–

Table 2 Y chromosome UEP+microsatellite haplotype frequencies in six Armenian regions

Hap. no.	Microsatellite haplotype[a]	Ararat (*n*=44)	North (*n*=189)	Syunik (*n*=140)	K'bakh (*n*=215)	Iranian (*n*=56)	West (*n*=90)	Total (*n*=734)	Total counts
(Haplogroup 1)									
1	14-12-24-11-13-12	0.0455	0.0476	0.1429	0.1116	0.0714	0.0444	0.0858	63
2	14-12-24-11-14-12	0.0227	0.0106	0.0286	0.0326	0.0357	0.0111	0.0232	17
3	14-12-24-11-13-13	–	–	0.0786	0.0279	–	–	0.0232	17
4	14-12-24-10-13-12	0.0455	0.0159	0.0071	0.0279	0.0179	–	0.0177	13
5	14-12-23-11-13-12	–	0.0106	0.0143	0.0279	0.0179	0.0111	0.0163	12
6	14-12-24-10-14-12	0.0227	0.0159	–	0.0233	–	0.0111	0.0136	10
7	14-12-23-11-14-12	–	0.0212	–	–	0.0179	0.0333	0.0109	8
8	16-12-22-10-11-13	–	–	0.0071	0.0233	–	0.0111	0.0095	7
9	14-12-24-10-13-11	–	0.0053	0.0071	0.0186	–	–	0.0082	6
10	14-12-25-11-14-12	–	–	0.0071	0.0186	–	–	0.0068	5
11	14-12-25-11-13-12	–	–	–	0.0140	–	0.0111	0.0054	4
12	14-12-25-10-13-12	–	–	0.0071	0.0093	0.0179	–	0.0054	4
13	14-12-23-10-14-12	–	0.0053	–	0.0093	0.0179	–	0.0054	4
14	14-12-23-10-13-12	–	0.0053	–	–	0.0357	0.0111	0.0054	4
15	15-12-25-11-13-12	–	–	0.0071	0.0047	0.0179	–	0.0041	3
16	15-12-24-11-14-12	–	0.0106	–	0.0047	–	–	0.0041	3
17	14-13-23-10-10-14	0.0227	–	0.0143	–	–	–	0.0041	3
18	13-12-25-10-13-12	–	–	0.0071	0.0093	–	–	0.0041	3
19	14-12-23-10-10-14	0.0227	–	–	0.0047	–	–	0.0027	2
20	15-12-24-11-13-12	–	–	0.0071	0.0047	–	–	0.0027	2
21	15-12-23-11-13-12	–	0.0106	–	–	–	–	0.0027	2
22	13-12-25-11-13-12	0.0227	–	–	0.0047	–	–	0.0027	2
23	14-12-21-10-12-12	–	0.0106	–	–	–	–	0.0027	2
24	13-13-25-11-14-12	–	0.0106	–	–	–	–	0.0027	2
25	14-12-23-11-14-13	–	–	–	0.0047	–	0.0111	0.0027	2
26	14-12-24-12-14-12	–	–	–	0.0093	–	–	0.0027	2
27	14-12-24-12-13-12	–	–	–	0.0093	–	–	0.0027	2
28	14-12-24-10-13-13	–	–	0.0071	–	0.0179	–	0.0027	2
29	14-12-24-10-13-14	–	–	0.0071	–	0.0179	–	0.0027	2
30	14-12-24-11-15-12	–	0.0106	–	–	–	–	0.0027	2
31	14-12-24-10-15-12	–	–	–	0.0047	–	0.0111	0.0027	2
32	13-12-24-10-14-12	–	0.0053	0.0071	–	–	–	0.0027	2
33	14-11-23-11-13-12	–	–	0.0071	–	–	–	0.0014	1
34	13-12-24-11-13-12	–	–	0.0071	–	–	–	0.0014	1
35	13-12-23-10-15-14	–	0.0053	–	–	–	–	0.0014	1
36	14-12-24-11-13-11	–	–	0.0071	–	–	–	0.0014	1
37	13-12-26-11-14-12	–	–	–	–	0.0179	–	0.0014	1
38	14-12-24-12-13-13	–	–	0.0071	–	–	–	0.0014	1
39	14-12-22-11-13-12	–	–	–	0.0047	–	–	0.0014	1
40	14-11-24-10-13-12	–	0.0053	–	–	–	–	0.0014	1
41	14-11-24-11-13-12	–	–	–	0.0047	–	–	0.0014	1
42	14-12-25-11-13-13	–	–	0.0071	–	–	–	0.0014	1
43	14-11-24-10-15-12	–	–	–	0.0047	–	–	0.0014	1
44	14-12-23-10-13-11	0.0227	–	–	–	–	–	0.0014	1
45	14-13-23-11-13-12	–	–	–	0.0047	–	–	0.0014	1
46	14-13-23-11-13-13	–	–	–	0.0047	–	–	0.0014	1
47	14-13-24-11-13-12	–	–	–	–	–	0.0111	0.0014	1
48	14-14-24-10-13-12	–	0.0053	–	–	–	–	0.0014	1
49	15-12-23-10-10-15	–	–	–	–	0.0179	–	0.0014	1
50	14-12-23-11-13-13	–	–	–	–	–	0.0111	0.0014	1
51	14-12-23-11-12-12	–	–	–	–	–	0.0111	0.0014	1
52	15-12-24-11-13-13	–	0.0053	–	–	–	–	0.0014	1
53	14-12-22-10-13-12	–	–	–	–	–	0.0111	0.0014	1
54	14-12-23-09-10-13	–	–	–	–	–	0.0111	0.0014	1
55	15-12-26-11-13-12	–	0.0053	–	–	–	–	0.0014	1

Table 2 (continued)

Hap. no.	Microsatellite haplotype[a]	Ararat (*n*=44)	North (*n*=189)	Syunik (*n*=140)	K'bakh (*n*=215)	Iranian (*n*=56)	West (*n*=90)	Total (*n*=734)	Total counts
56	15-13-23-10-13-13	–	–	0.0071	–	–	–	0.0014	1
	Total (hg1)	0.2273	0.2222	0.4000	0.4279	0.3214	0.2222	0.3243	238
(Haplogroup 2)									
57	14-15-23-10-11-12	0.0909	0.0423	0.0214	0.0744	0.0893	0.0556	0.0559	41
58	14-16-23-10-11-12	0.0227	0.0212	0.0143	0.0140	–	0.0444	0.0191	14
59	14-14-23-10-11-12	–	0.0212	0.0214	0.0140	0.0179	0.0111	0.0163	12
60	14-17-23-10-11-12	0.0227	0.0265	–	0.0093	–	0.0222	0.0136	10
61	14-14-24-10-11-12	0.0227	0.0159	0.0071	0.0186	–	0.0111	0.0136	10
62	15-12-21-10-11-14	–	0.0159	–	0.0093	–	0.0444	0.0123	9
63	14-17-24-10-11-12	–	0.0106	0.0429	–	–	0.0111	0.0123	9
64	14-17-23-11-11-12	–	0.0053	0.0071	0.0140	0.0714	–	0.0123	9
65	15-15-23-10-11-12	0.0227	0.0159	–	0.0093	–	0.0111	0.0095	7
66	14-15-22-10-11-12	–	0.0106	–	–	0.0536	0.0222	0.0095	7
67	15-12-22-10-11-13	–	0.0159	0.0071	0.0047	–	0.0111	0.0082	6
68	15-13-24-10-11-14	0.0227	0.0159	–	0.0047	–	–	0.0068	5
69	15-13-21-10-11-14	–	–	0.0214	0.0093	–	–	0.0068	5
70	14-15-24-10-11-12	–	0.0106	0.0071	0.0047	0.0179	–	0.0068	5
71	14-13-23-10-11-12	–	0.0053	0.0071	0.0093	–	0.0111	0.0068	5
72	16-13-24-10-11-14	–	0.0159	–	0.0047	–	–	0.0054	4
73	15-15-24-11-11-12	0.0227	–	–	–	0.0179	0.0222	0.0054	4
74	15-15-24-10-11-12	–	0.0106	0.0071	0.0047	–	–	0.0054	4
75	15-15-22-10-11-12	–	–	0.0214	–	0.0179	–	0.0054	4
76	15-14-24-10-11-12	–	0.0053	–	0.0140	–	–	0.0054	4
77	15-13-23-10-11-14	0.0227	0.0053	–	0.0093	–	–	0.0054	4
78	15-12-22-10-11-14	–	0.0106	–	0.0047	–	0.0111	0.0054	4
79	15-12-21-10-11-15	–	0.0159	–	–	–	0.0111	0.0054	4
80	14-16-23-11-11-13	–	–	0.0286	–	–	–	0.0054	4
81	14-15-23-10-11-13	–	0.0159	–	0.0047	–	–	0.0054	4
82	14-14-24-10-11-13	–	–	–	0.0093	0.0357	–	0.0054	4
83	15-16-24-10-11-12	0.0227	–	–	–	0.0357	–	0.0041	3
84	15-14-22-10-11-12	–	0.0159	–	–	–	–	0.0041	3
85	14-13-24-10-11-12	–	–	–	0.0093	–	0.0111	0.0041	3
86	13-15-23-10-11-12	–	0.0053	0.0143	–	–	–	0.0041	3
87	15-12-22-11-11-14	–	–	–	0.0093	0.0179	–	0.0041	3
88	14-16-24-10-11-12	–	0.0106	–	0.0047	–	–	0.0041	3
89	14-15-23-11-11-12	–	–	0.0214	–	–	–	0.0041	3
90	14-15-22-11-11-12	–	0.0053	–	0.0093	–	–	0.0041	3
91	14-15-22-10-11-13	0.0227	0.0053	–	–	–	0.0111	0.0041	3
92	15-12-23-10-11-14	–	–	–	0.0093	–	–	0.0027	2
93	16-13-24-11-11-13	–	–	–	0.0047	0.0179	–	0.0027	2
94	14-15-23-09-11-12	–	0.0106	–	–	–	–	0.0027	2
95	16-12-23-10-12-13	–	0.0053	–	–	–	0.0111	0.0027	2
96	16-12-21-10-11-13	–	0.0053	–	–	–	0.0111	0.0027	2
97	14-15-23-10-11-14	–	–	–	0.0093	–	–	0.0027	2
98	15-17-23-10-11-12	0.0227	–	–	0.0047	–	–	0.0027	2
99	15-16-24-11-11-12	–	0.0106	–	–	–	–	0.0027	2
100	15-16-23-10-11-12	–	–	–	0.0047	–	0.0111	0.0027	2
101	15-15-25-10-11-12	–	–	0.0071	–	0.0179	–	0.0027	2
102	15-15-24-11-11-13	0.0227	–	0.0071	–	–	–	0.0027	2
103	15-15-23-11-11-12	–	–	0.0071	–	–	0.0111	0.0027	2
104	14-14-24-11-11-12	–	–	–	0.0093	–	–	0.0027	2
105	15-13-23-10-11-12	–	–	0.0071	–	–	0.0111	0.0027	2
106	14-16-22-10-12-12	–	–	0.0071	–	0.0179	–	0.0027	2
107	14-12-23-10-12-13	–	0.0053	–	–	0.0179	–	0.0027	2
108	15-12-23-10-11-13	–	0.0053	–	0.0047	–	–	0.0027	2
109	15-12-22-11-11-15	–	–	–	0.0093	–	–	0.0027	2

Table 2 (continued)

Hap. no.	Microsatellite haplotype[a]	Ararat (*n*=44)	North (*n*=189)	Syunik (*n*=140)	K'bakh (*n*=215)	Iranian (*n*=56)	West (*n*=90)	Total (*n*=734)	Total counts
110	15-12-22-09-11-13	–	0.0106	–	–	–	–	0.0027	2
111	15-12-21-11-11-14	–	0.0106	–	–	–	–	0.0027	2
112	14-19-24-10-11-12	–	0.0106	–	–	–	–	0.0027	2
113	14-16-24-08-11-12	–	–	0.0071	0.0047	–	–	0.0027	2
114	13-18-24-10-11-12	–	–	0.0143	–	–	–	0.0027	2
115	14-17-23-10-11-13	–	0.0053	–	–	–	0.0111	0.0027	2
116	16-17-23-10-11-12	–	–	–	–	–	0.0111	0.0014	1
117	14-17-22-10-10-12	–	0.0053	–	–	–	–	0.0014	1
118	14-17-23-10-10-12	0.0227	–	–	–	–	–	0.0014	1
119	13-15-23-11-11-12	0.0227	–	–	–	–	–	0.0014	1
120	14-17-21-11-11-12	0.0227	–	–	–	–	–	0.0014	1
121	14-12-22-10-11-13	–	0.0053	–	–	–	–	0.0014	1
122	14-17-23-11-11-13	–	0.0053	–	–	–	–	0.0014	1
123	14-16-25-10-11-12	–	–	–	–	0.0179	–	0.0014	1
124	14-17-24-10-12-12	–	–	0.0071	–	–	–	0.0014	1
125	14-18-23-10-10-12	–	0.0053	–	–	–	–	0.0014	1
126	14-18-23-10-11-12	–	–	–	–	–	0.0111	0.0014	1
127	14-14-26-11-11-12	–	–	–	0.0047	–	–	0.0014	1
128	15-12-21-09-11-14	–	–	–	–	–	0.0111	0.0014	1
129	13-16-24-10-11-12	–	–	–	0.0047	–	–	0.0014	1
130	14-13-24-11-11-13	–	0.0053	–	–	–	–	0.0014	1
131	14-13-25-11-11-12	–	0.0053	–	–	–	–	0.0014	1
132	15-12-21-12-11-16	–	0.0053	–	–	–	–	0.0014	1
133	14-16-23-11-11-12	–	0.0053	–	–	–	–	0.0014	1
134	15-12-22-10-10-14	–	–	–	0.0047	–	–	0.0014	1
135	14-12-22-11-11-13	0.0227	–	–	–	–	–	0.0014	1
136	14-13-24-10-12-14	–	0.0053	–	–	–	–	0.0014	1
137	15-12-22-11-10-14	–	–	–	–	–	0.0111	0.0014	1
138	14-14-25-10-11-12	–	–	–	–	–	0.0111	0.0014	1
139	14-16-23-10-11-14	–	–	–	0.0047	–	–	0.0014	1
140	13-12-24-10-11-14	–	–	–	–	–	0.0111	0.0014	1
141	14-16-23-09-11-12	–	–	0.0071	–	–	–	0.0014	1
142	15-12-23-11-11-14	–	0.0053	–	–	–	–	0.0014	1
143	15-12-24-10-11-13	–	–	–	–	–	0.0111	0.0014	1
144	15-12-25-10-11-13	–	–	–	–	–	0.0111	0.0014	1
145	15-13-21-10-11-12	0.0227	–	–	–	–	–	0.0014	1
146	15-13-21-10-11-13	0.0227	–	–	–	–	–	0.0014	1
147	14-12-24-11-13-12	0.0227	–	–	–	–	–	0.0014	1
148	15-13-22-10-11-13	–	–	–	–	–	0.0111	0.0014	1
149	15-13-22-10-11-14	–	–	–	0.0047	–	–	0.0014	1
150	14-16-22-11-11-12	–	–	–	0.0047	–	–	0.0014	1
151	14-15-26-10-11-12	–	0.0053	–	–	–	–	0.0014	1
152	15-13-23-10-12-12	–	–	–	0.0047	–	–	0.0014	1
153	15-13-23-10-12-14	–	0.0053	–	–	–	–	0.0014	1
154	15-13-23-11-12-14	–	–	–	0.0047	–	–	0.0014	1
155	15-13-24-10-11-13	–	–	0.0071	–	–	–	0.0014	1
156	14-12-24-10-12-13	–	0.0053	–	–	–	–	0.0014	1
157	15-13-24-10-12-12	–	–	–	–	–	0.0111	0.0014	1
158	15-13-24-11-11-13	–	0.0053	–	–	–	–	0.0014	1
159	15-13-24-11-11-14	–	0.0053	–	–	–	–	0.0014	1
160	15-13-25-10-11-14	–	–	0.0071	–	–	–	0.0014	1
161	13-16-23-10-11-12	–	0.0053	–	–	–	–	0.0014	1
162	15-14-22-11-11-12	–	–	–	0.0047	–	–	0.0014	1
163	15-14-23-10-08-13	–	0.0053	–	–	–	–	0.0014	1
164	15-14-23-10-11-12	0.0227	–	–	–	–	–	0.0014	1
165	14-13-23-11-11-12	–	–	–	0.0047	–	–	0.0014	1

Table 2 (continued)

Hap. no.	Microsatellite haplotype[a]	Ararat (*n*=44)	North (*n*=189)	Syunik (*n*=140)	K'bakh (*n*=215)	Iranian (*n*=56)	West (*n*=90)	Total (*n*=734)	Total counts
166	15-14-24-11-11-12	0.0227	–	–	–	–	–	0.0014	1
167	15-14-25-11-11-12	–	–	–	0.0047	–	–	0.0014	1
168	14-13-23-10-12-12	–	–	–	0.0047	–	–	0.0014	1
169	15-15-23-09-11-12	–	–	0.0071	–	–	–	0.0014	1
170	15-15-23-10-11-13	–	–	0.0071	–	–	–	0.0014	1
171	14-15-25-11-11-12	–	–	0.0071	–	–	–	0.0014	1
172	15-15-24-09-11-12	–	–	0.0071	–	–	–	0.0014	1
173	14-13-23-10-11-13	–	–	–	–	–	0.0111	0.0014	1
174	14-15-25-10-11-12	–	–	0.0071	–	–	–	0.0014	1
175	14-15-25-10-10-12	–	–	–	–	–	0.0111	0.0014	1
176	15-15-25-11-11-12	–	0.0053	–	–	–	–	0.0014	1
177	15-15-25-11-11-13	–	0.0053	–	–	–	–	0.0014	1
178	15-16-21-10-11-14	–	–	–	0.0047	–	–	0.0014	1
179	15-16-22-10-11-13	–	–	–	–	0.0179	–	0.0014	1
180	15-16-23-09-11-12	–	0.0053	–	–	–	–	0.0014	1
181	15-16-23-09-11-13	–	0.0053	–	–	–	–	0.0014	1
182	14-15-24-10-11-13	–	–	–	0.0047	–	–	0.0014	1
183	14-14-23-10-11-13	–	–	–	0.0047	–	–	0.0014	1
184	14-13-22-10-11-12	–	–	0.0071	–	–	–	0.0014	1
185	15-16-25-10-11-12	–	0.0053	–	–	–	–	0.0014	1
186	14-15-24-10-11-11	–	–	–	–	–	0.0111	0.0014	1
187	14-12-23-10-10-13	–	0.0053	–	–	–	–	0.0014	1
188	16-12-22-10-10-14	–	0.0053	–	–	–	–	0.0014	1
189	16-12-23-10-10-13	–	–	0.0071	–	–	–	0.0014	1
190	16-12-23-10-11-14	–	–	–	–	–	0.0111	0.0014	1
191	14-14-22-10-11-13	–	0.0053	–	–	–	–	0.0014	1
192	16-12-23-11-11-14	–	0.0053	–	–	–	–	0.0014	1
193	16-12-24-11-11-12	–	0.0053	–	–	–	–	0.0014	1
194	16-13-23-10-11-14	–	–	–	–	–	0.0111	0.0014	1
195	14-13-23-09-11-12	–	0.0053	–	–	–	–	0.0014	1
196	16-14-22-10-11-12	–	–	0.0071	–	–	–	0.0014	1
197	16-15-23-09-11-13	–	–	–	–	–	0.0111	0.0014	1
198	16-15-23-10-11-12	–	–	0.0071	–	–	–	0.0014	1
199	17-12-22-10-11-13	–	–	0.0071	–	–	–	0.0014	1
200	17-13-24-11-11-12	–	–	–	–	0.0179	–	0.0014	1
201	17-13-25-10-11-14	–	–	–	0.0047	–	–	0.0014	1
202	17-15-23-09-11-13	–	–	–	–	–	0.0111	0.0014	1
203	18-13-24-10-11-14	–	–	–	0.0047	–	–	0.0014	1
(Haplogroup 2, duplicated DYS 19 locus)									
204	15/16-12-23-10-11-14	0.0227	–	0.0143	0.0047	–	–	0.0054	4
205	15/16-12-22-10-11-13	–	0.0053	–	–	–	–	0.0014	1
206	14/15-12-22-10-11-14	–	–	–	–	0.0179	–	0.0014	1
207	14/15-13-22-10-11-13	0.0227	–	–	–	–	–	0.0014	1
208	12/15-12-22-10-11-14	–	–	0.0071	–	–	–	0.0014	1
	Total (hg2)	0.5909	0.5873	0.4286	0.4233	0.5179	0.5778	0.5027	369
(Haplogroup 3)									
209	15-12-25-11-11-13	–	0.0265	0.0857	0.0140	–	–	0.0272	20
210	16-12-25-11-11-13	–	0.0053	0.0071	0.0140	–	0.0111	0.0082	6
211	17-12-25-10-11-13	–	0.0053	–	0.0047	–	0.0111	0.0041	3
212	17-12-23-11-11-13	–	–	–	0.0093	–	–	0.0027	2
213	16-12-25-10-11-13	–	0.0053	–	–	0.0179	–	0.0027	2
214	15-13-24-10-11-13	–	–	–	0.0047	–	–	0.0014	1
215	15-12-25-10-11-14	–	–	–	–	–	0.0111	0.0014	1
216	16-12-26-11-11-13	–	–	–	0.0047	–	–	0.0014	1
217	16-12-24-11-11-13	–	–	–	0.0047	–	–	0.0014	1
	Total (hg3)	–	0.0423	0.0929	0.0558	0.0179	0.0333	0.0504	37

Table 2 (continued)

Hap. no.	Microsatellite haplotype[a]	Ararat (*n*=44)	North (*n*=189)	Syunik (*n*=140)	K'bakh (*n*=215)	Iranian (*n*=56)	West (*n*=90)	Total (*n*=734)	Total counts
(Haplogroup 21)									
218	13-12-24-10-11-13	0.0227	0.0106	0.0143	0.0047	0.0179	0.0222	0.0123	9
219	14-12-24-10-11-13	0.0227	0.0106	–	–	0.0179	–	0.0054	4
220	13-12-26-10-12-13	–	–	–	–	0.0714	–	0.0054	4
221	13-12-23-10-11-13	0.0227	0.0053	0.0071	0.0047	–	–	0.0054	4
222	12-12-24-10-11-13	–	0.0053	0.0071	0.0047	0.0179	–	0.0054	4
223	13-12-23-09-11-14	–	0.0159	–	–	–	–	0.0041	3
224	13-12-25-10-11-13	–	0.0106	–	–	–	–	0.0027	2
225	15-12-25-07-11-13	–	–	–	0.0047	–	–	0.0014	1
226	13-12-24-11-11-13	–	0.0053	–	–	–	–	0.0014	1
227	13-12-25-10-12-13	–	–	–	–	0.0179	–	0.0014	1
228	13-12-24-09-11-13	–	–	–	0.0047	–	–	0.0014	1
229	14-12-23-10-11-14	–	–	–	–	–	0.0111	0.0014	1
230	13-12-24-09-11-14	0.0227	–	–	–	–	–	0.0014	1
231	14-12-25-10-11-13	–	0.0053	–	–	–	–	0.0014	1
232	14-12-25-10-11-14	–	–	–	–	–	0.0111	0.0014	1
233	14-12-25-10-12-13	–	–	–	0.0047	–	–	0.0014	1
234	15-12-23-10-11-14	–	–	–	–	–	0.0111	0.0014	1
	Total (hg21)	0.0909	0.0688	0.0286	0.0279	0.1429	0.0556	0.0545	40
(Haplogroup 26)									
235	14-12-23-10-13-13	–	0.0212	0.0143	0.0279	–	–	0.0163	12
236	14-12-23-10-15-14	–	0.0106	0.0214	–	–	–	0.0068	5
237	14-12-24-10-13-13	–	0.0053	–	–	–	0.0222	0.0041	3
238	14-12-23-11-13-13	0.0227	–	–	–	–	0.0222	0.0041	3
239	15-12-23-10-15-14	0.0227	–	–	–	–	0.0111	0.0027	2
240	15-12-23-10-13-13	–	0.0053	–	0.0047	–	–	0.0027	2
241	13-12-23-11-13-13	–	0.0053	–	0.0047	–	–	0.0027	2
242	14-12-23-10-13-14	–	–	–	0.0093	–	–	0.0027	2
243	16-12-23-10-13-13	–	–	0.0071	–	–	–	0.0014	1
244	15-12-23-10-13-12	–	–	–	–	–	0.0111	0.0014	1
245	13-12-23-10-13-12	–	0.0053	–	–	–	–	0.0014	1
246	14-12-23-10-14-13	–	–	0.0071	–	–	–	0.0014	1
247	15-12-24-10-13-14	–	–	–	0.0047	–	–	0.0014	1
	Total (hg26)	0.0455	0.0529	0.0500	0.0512	–	0.0667	0.0490	36
(Haplogroup 28)									
248	15-12-23-10-13-11	0.0227	0.0212	–	–	–	0.0333	0.0109	8
249	17-12-23-10-13-11	–	–	–	0.0047	–	–	0.0014	1
250	14-12-23-10-15-13	–	–	–	0.0047	–	–	0.0014	1
251	15-12-23-10-15-12	–	–	–	0.0047	–	–	0.0014	1
252	16-12-23-11-14-11	0.0227	–	–	–	–	–	0.0014	1
	Total (hg28)	0.0455	0.0212	–	0.0140	–	0.0333	0.0163	12
(Haplogroup 29)									
253	15-13-25-11-14-13	–	0.0053	–	–	–	0.0111	0.0027	2

Microsatellite haplotypes within haplogroups are arranged in decreasing order of total frequency. This table is also available in electronic format on request from the corresponding author

[a] Microsatellite haplotypes are defined by a string of 6 numbers giving the repeat size at loci DYS19, DYS388, DYS390, DYS391, DYS392 and DYS393 respectively

lite network, supporting the hypothesis that they are the result of a single duplication event. Duplicated DYS19 loci were found in two other comparative populations, also within hg2: Turkey (2/173=1.2%) and Mongolia (39/402=9.7%). Apart from one Turkish chromosome, none of these other haplotypes fit into the Armenian network.

Genetic diversity values (*h* based on UEP+microsatellite haplotype frequencies) are lower in the Syunik, Karabakh and Iranian regions than in Ararat, the North and West, in some cases significantly so (Fig. 3, Table 3). This suggests a pattern of genetic differentiation within Armenia similar to that found with haplogroup frequencies. With the exception of England and Friesland, where

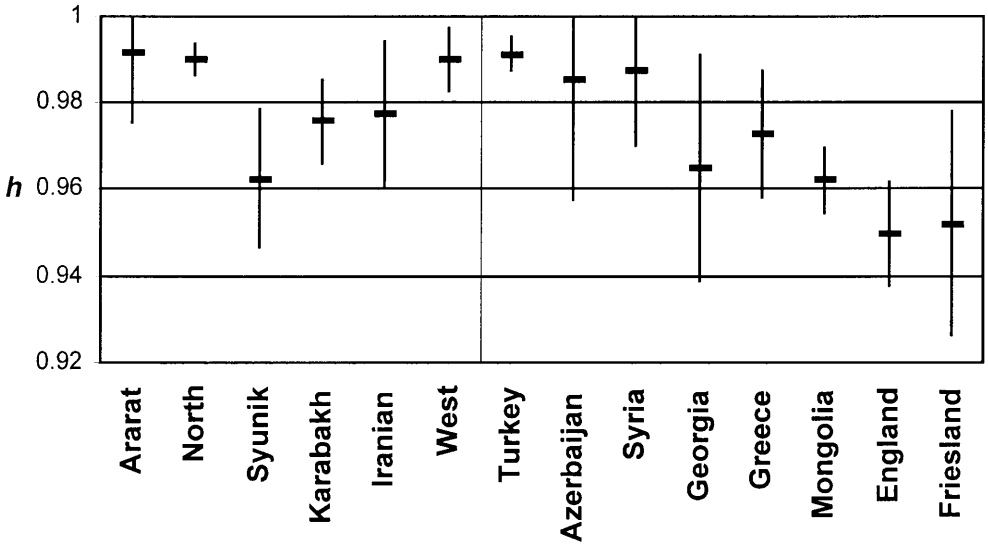

Fig. 3 Genetic diversity, *h*, with bootstrap 95% confidence intervals across six Armenian regions and eight comparative data sets

Table 3 Pairwise differences in *h* values (*lower left table*, row value minus column value) based on UEP+microsatellite haplotype frequencies, and *P* values (two-tailed) for difference in *h* (*upper right table*) based on bootstrap standard errors

	Ararat (n=44)	North (n=189)	Syunik (n=140)	K'bakh (n=215)	Iranian (n=56)	West (n=90)
Ararat	–	0.721	0.023	0.109	0.196	0.740
North	–0.002	–	<0.001*	0.002*	0.075	0.976
Syunik	–0.029*	–0.027*	–	0.146	0.208	<0.001*
K'bakh	–0.016	–0.014*	0.013	–	0.812	0.018
Iranian	–0.015	–0.013	0.014	0.001	–	0.134
West	–0.002	0.000	0.027	0.014	0.013	–

*P<0.05

the genetic diversity is lower, the range of *h* values within Armenia is comparable to that found among the comparative data sets. Genetic diversity is particularly high in Ararat, North, West and the neighbouring data sets from Turkey, Azerbaijan and Syria.

Armenian regions share "Frequently Encountered Haplotypes", defined as those occurring in any one population at a frequency greater than or equal to 10%, with all comparative populations, although the level and pattern of sharing varies (Table 4). The Armenian modal hg2 haplotype (haplotype 57 in Table 2) is shared with neighbouring or Middle Eastern countries, and is also the modal hg2 haplotype in Turkey and Azerbaijan. This haplotype has a DYS388 repeat size of 15, and Table 4 supports the hypothesis that haplotypes with high DYS388 repeat such sizes as this one are signatures of Near Eastern/Southeast Asian origin (Bradman et al. 2000). In contrast, the Georgian modal hg2 haplotype (haplotype 78) has a DYS388 repeat size of 12. The Armenian modal hg1 haplotype (haplotype 1), prevalent particularly in the Syunik, Karabakh and Iranian regions, is also the Turkish modal hg1 haplotype and is a one-step neighbour of both the English modal (haplotype 3) and the Frisian modal (haplotype 50) hg1 haplotypes. The English modal haplotype is also found at high frequency in Syunik.

Genetic distances

We visualised the patterns of genetic affinities within Armenia and among comparative data sets using principal coordinates analysis applied to both F_{ST} values (UEP+microsatellite haplotype frequency data) and R_{ST} values (microsatellite data only) (Fig. 4, Table 5). Mongolia, England and Friesland have been removed from the plot based on F_{ST} values to aid in the separation of more closely related samples, as the two-dimensional projection based on F_{ST} values explains less of the overall variance than the projection based on R_{ST} values.

Figure 4 reveals several interesting features further supported by significance tests. Firstly, the degree of genetic differentiation or structure among different Armenian regions is comparable to that found for many comparisons across ethno-national boundaries. Pairwise exact tests for population differentiation reveal significant differences for Syunik vs. all other regions, and also for Karabakh vs. all other regions. The Iranian region is also significantly different from all regions except Ararat, and thus there is also evidence for genetic isolation of this region despite the impression given in Fig. 4a. In both plots in Fig. 4, Syunik and Karabakh appear as genetically isolated groups, although F_{ST} values place Syunik as the most isolated while R_{ST} values place Karabakh as the most isolated among the Armenian regions.

A second feature is that Syunik and Karabakh appear more isolated from the data sets collected in Turkey, Azerbaijan, Syria and Greece than do other Armenian regions, even though Karabakh is located wholly within Azerbaijan (Fig. 1). This is backed up by bootstrap tests on the two regions with the largest sample sizes – Syunik and the North. Bootstrap tests based on F_{ST} values show that Azerbaijan, Syria, Turkey and Greece are signifi-

Table 4 Frequently Encountered Haplotype table for six Armenian regions and eight comparative data sets

	Hap. no.	Microsatellite haplotype[a]	Ararat (*n*=44)	North (*n*=189)	Syunik (*n*=140)	K'bakh (*n*=215)	Iranian (*n*=56)	West (*n*=90)	Turkey (*n*=173)	Azer (*n*=29)	Syria (*n*=44)	Georgia (*n*=68)	Greece (*n*=132)	Mongol (*n*=402)	Engl. (*n*=310)	Friesl. (*n*=94)
hg1	1	14–12–24–11–13–12	0.0455	0.0476	0.1429*	0.1116*	0.0714	0.0444	0.0462	0.0345	–	0.0294	0.0076	–	0.0032	0.0106
hg1	3	14-12-24-11-13-13	–	–	0.0786	0.0279	–	–	0.0289	0.0345	–	–	0.0379	–	0.1516*	0.1277
hg1	28	14-12-24-10-13-13	–	–	0.0071	–	0.0179	–	0.0116	–	–	–	0.0227	0.0025	0.1000*	0.0213
hg1	50	14-12-23-11-13-13	–	–	–	–	–	0.0111	0.0058	–	–	–	0.0076	–	0.0968	0.1702*
hg2	57	14-15-23-10-11-12	0.0909	0.0423	0.0214	0.0744	0.0893	0.0556	0.0405	0.1034*	0.0227	0.0147	–	–	–	–
hg2	92	15-12-23-10-11-14	0.0227	–	0.0143	0.0140	–	–	–	–	–	–	–	0.1020*	–	–
hg2	78	15-12-22-10-11-14	–	0.0106	–	0.0047	–	0.0111	0.0231	–	–	0.1618*	0.0076	–	–	–
hg21	218	13-12-24-10-11-13	0.0227	0.0106	0.0143	0.0047	0.0179	0.0222	0.0289	0.0345	0.0227	0.0147	0.1288*	–	0.0129	0.0213

[a] A haplotype is included if observed at a frequency ≥10% in at least one region or country
*frequency ≥10%

cantly closer to the North than they are to Syunik ($P<0.01$ in each case). Significant differences were also found using R_{ST} values, although R_{ST} bootstrap variances were larger, and therefore fewer results were significant.

Thirdly, the Armenian regions all appear only distantly related to Georgia, despite its geographical proximity. Bootstrap tests based on mean F_{ST} values from all six Armenian regions show that, overall, Armenia is significantly closer to Azerbaijan ($P=0.037$) and Turkey ($P<0.001$) than it is to Georgia. Finally, Syunik and Karabakh appear more closely related to England and Friesland than do the other Armenian regions. This is due to the sharing of similar hg1 haplotypes as discussed in the previous section.

We repeated these affinity analyses using other genetic distance measures, including F_{ST} based on haplogroup frequencies; Nei's genetic identity *I* based on haplogroup and haplotype frequencies; F'_{ST} based on microsatellite data (where inter-haplotype distance is defined as the number of microsatellite loci with different alleles); and all of these measures repeated within hg1 and hg2 only. While differences existed in the distance matrices produced by these other analyses, general features were consistent with those found above.

Representativeness of the London Armenian sample

The regional differences within Armenia raise the question of the extent to which a diasporan sample (in this case the London Armenians) adequately reflects the Y chromosomes of its source population (Armenia). In our survey, 97% of London Armenians ($n=89$) had a paternal descent from only two of the six Armenian regions defined in our study: West ($n=46$) and Iranian ($n=40$). This bias in regional representation reflects greater historical emigration of ethnic Armenians from these areas. Not surprisingly, given the high regional differentiation reported above, the biases in regions sampled caused the UEP+microsatellite haplotype frequency distributions to differ significantly between London and non-London Armenians when compared as a whole ($P=0.012$). However, even when London Armenians from Iran were compared to non-London Armenians from Iran a significant difference still remained ($P=0.005$). West Armenians, the only other group with a large London sample size, were not significantly different when compared in this way ($P=0.662$). Examination of UEP+microsatellite haplotypes reveals that the incongruence of the Iranian samples is because haplotype 64 (in hg2) and haplotype 220 (in hg21) are at high frequency in the non-London Iranian Armenian subsample (4/16=25% in both cases) but absent from the London Iranian Armenian sub-sample.

Our results show that (a) due to high regional structuring, a single sample labelled "Armenian" would not adequately represent any one region; (b) the sample of London Armenians was, for historical reasons of population migration, regionally highly biased, with most regions not being adequately represented; and (c) even when grandpa-

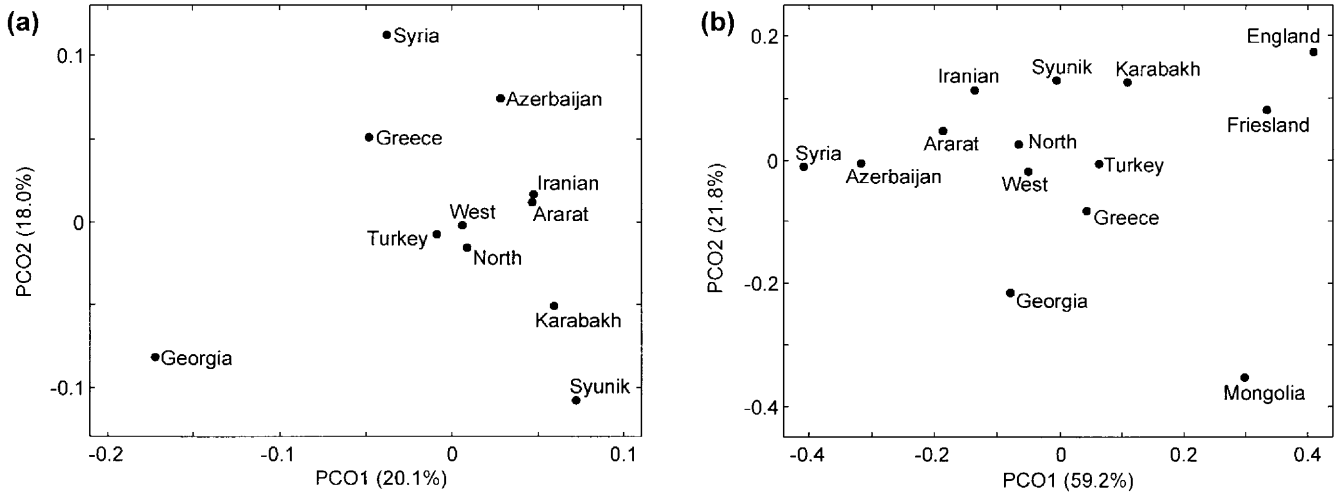

Fig. 4 Principal coordinates plots (first two axes) based on F_{ST} values (**a**), calculated from UEP+microsatellite haplotype frequency data, and R_{ST} values (**b**), calculated from microsatellite data only. *PCO1* first axis, *PCO2* second axis. *Numbers in brackets* percentage of total variation explained by each axis

ternal region of origin was taken into account, differences still existed between London Armenians and non-London Armenians.

BATWING

Using BATWING, we fitted a model to the Armenian data that included: (a) exponential growth from an initially constant-size population and (b) binary fission (without subsequent migration) of an initially panmictic population into the six regions that we defined in and around Armenia today. Posterior distributions of parameters of interest are presented in Fig. 5. Despite wide credible intervals, the dates for the time to most recent common ancestor, start of growth and the oldest (deepest) population split are clearly separated. The posterior growth rate (mean= 0.026, 95% credible interval: 0.010–0.053) indicates a strong growth signal in the data, but with a recent start date under the assumed exponential growth model (mean=4.8 thousand years ago, 95% credible interval: 2.0–11.1), most likely within the last 10,000 years (95.7% support).

In support of our other genetic distance analyses, the majority (81%) of posterior population trees contain either Syunik/Karabakh/Iranian vs. Ararat/North/West regions as the deepest split (56.2%), or Syunik/Karabakh vs. other regions as the deepest split (24.8%). All other configurations occur at a frequency of less than 5%.

We repeated the BATWING analysis using a model that included exponential growth but excluded binary fission of sub-populations (instead, the model assumed a single panmictic Armenian population). We found that the posterior distributions of both the growth rate (mean 0.023, 95% credible interval: 0.011–0.043) and the start-

Table 5 Pairwise F_{ST} values (*lower left table*) based on UEP+microsatellite haplotype frequencies, and R_{ST} values (*upper right table*) based on microsatellite data only

	Ararat (*n*=44)	North (*n*=189)	Syunik (*n*=140)	K'bakh (*n*=215)	Iranian (*n*=56)	West (*n*=90)	Turkey (*n*=173)	Azer. (*n*=29)	Syria (*n*=44)	Georgia (*n*=68)	Greece (*n*=132)	Mongol. (*n*=402)	England (*n*=310)	Friesl. (*n*=94)
Ararat	–	–0.009	0.016	0.040*	0.001	–0.005	0.015	–0.015	0.033	0.045*	0.051*	0.207*	0.179*	0.133*
North	–0.002	–	0.011*	0.024	0.006	–0.004	0.005	0.006	0.059*	0.040*	0.027*	0.152*	0.113*	0.079*
Syunik	0.012*	0.011*	–	0.005	–0.003	0.021*	0.016*	0.038*	0.098*	0.095*	0.031*	0.167*	0.091*	0.074*
K'bakh	0.000*	0.004*	0.006*	–	0.026*	0.027*	0.013*	0.083*	0.162*	0.110*	0.041*	0.142*	0.055*	0.049*
Iranian	0.001	0.005*	0.014*	0.004*	–	0.022*	0.028*	0.004	0.048*	0.093*	0.038*	0.202*	0.157*	0.126*
West	–0.003	–0.001	0.014*	0.004*	0.004*	–	0.000	0.016	0.079*	0.030*	0.034*	0.139*	0.125*	0.081*
Turkey	–0.001	0.002*	0.010*	0.005*	0.005*	0.001*	–	0.046*	0.113*	0.036*	0.016*	0.106*	0.065*	0.037*
Azer.	–0.005	0.000	0.015*	0.004*	–0.001	–0.001	0.000	–	–0.011	0.059*	0.073*	0.241*	0.262*	0.204*
Syria	0.005*	0.007*	0.022*	0.013*	0.006*	0.007*	0.006*	0.004	–	0.108*	0.139*	0.316*	0.344*	0.278*
Georgia	0.017*	0.015*	0.029*	0.022*	0.020*	0.015*	0.013*	0.015	0.020*	–	0.053*	0.151*	0.206*	0.133*
Greece	0.013*	0.014*	0.025*	0.021*	0.019*	0.014*	0.009*	0.012*	0.013*	0.026*	–	0.106*	0.124*	0.092*
Mongol.	0.021*	0.024*	0.035*	0.029*	0.031*	0.024*	0.022*	0.027*	0.026*	0.036*	0.032*	–	0.163*	0.127*
England	0.029*	0.029*	0.031*	0.032*	0.035*	0.028*	0.022*	0.027*	0.031*	0.042*	0.026*	0.044*	–	0.002
Friesl.	0.026*	0.027*	0.030*	0.030*	0.033*	0.026*	0.021*	0.026*	0.030*	0.041*	0.027*	0.043*	0.005*	–

*$P<0.05$ (*lower left table*, using exact tests for population differentiation; *upper right table*, using AMOVA permutation tests on R_{ST} values)

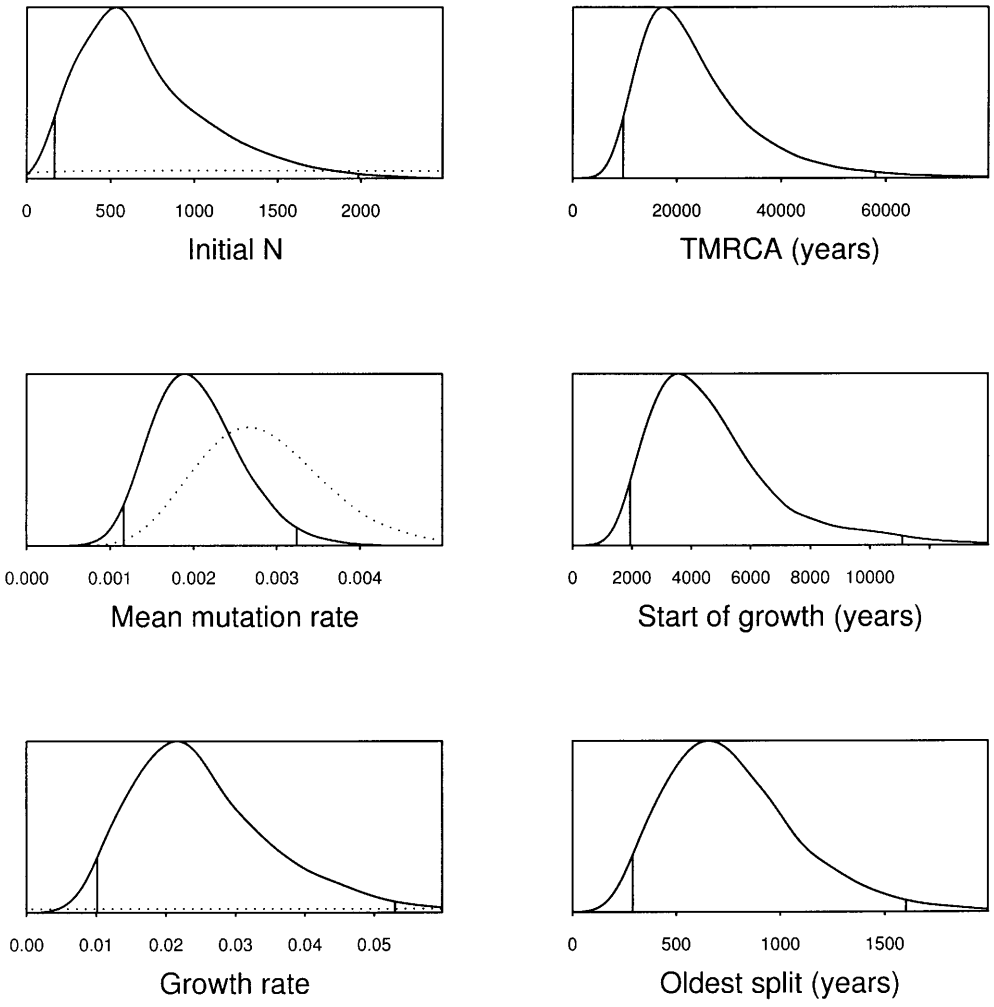

Fig. 5 Posterior marginal distributions (*solid lines*) of key parameters from the BATWING analysis described in the text. *Vertical lines* 95% credible intervals. Where applicable, prior distributions are also shown (*dotted lines*). *TMRCA* time to most recent common ancestor

of-growth date (mean 4.5 thousand years ago, 95% credible interval: 2.1–8.5) were little changed from those given above, the only noticeable differences being in the slightly reduced upper credible limits. This suggests that the assumptions relating to population structure had little influence on the inferred growth signal from these data.

Finally, we repeated the BATWING analysis using only the data from the London Armenian sample, restricted to those who were originally either from the West or Iranian regions. We fitted a model including both exponential growth and binary fission into the two regional groups. The inferred growth signal from these data was very different from that obtained using the entire Armenian data set. The posterior starting date for growth (mean=27.5 thousand years ago, 95% credible interval: 11.0–61.4) coincided with an initial effective population size close to one, indicating that BATWING was converging on a model of continuous population growth. The posterior growth rate (mean: 0.010, 95% credible interval: 0.0033–0.024) was smaller than that obtained using the entire Armenian data set, although the 95% credible intervals overlapped considerably. Overall, these results again indicate that in terms of genetic inference the diasporan population sample was not a good surrogate for the Armenian population as a whole.

Discussion

This study demonstrates that a high degree of Y chromosome genetic structuring can exist within a single geographically restricted ethno-national group with a strong ethnic identity, a finding which complicates inferences of migration patterns into and around Armenia. The high degree of within-Armenia structure highlights the importance of appropriate Y chromosome sampling schemes. An apparently attractive strategy for sampling relatively inaccessible populations, in both genetic anthropology and genetic epidemiology studies, is to select ethnically self-identified individuals living in major conurbations. However, as this study demonstrates (London Armenians vs. other Armenians) great care must be exercised in identifying appropriate sampling procedures, and it cannot be assumed that the target group is appropriately represented by its displaced subset, nor can it be assumed that the target group is genetically homogeneous.

The high degree of within-Armenia structure reveals a complex pattern of affinities. The Syunik and Karabakh regions are significantly differentiated from each other and from the Ararat/North/West group and also appear more distant from neighbouring comparative data sets. A

plausible explanation for these patterns is mountainous isolation. Historically, both Syunik and Karabakh were Armenian "melikhoods", or principalities, until the seventeenth century AD (Hovannisian 1997; Redgate 1998), and the two share similar Armenian dialects (Djahukian 1987). Research on palmar dermatoglyphics also reveals significant differences between Syunik and other regions of current and historical Armenia (N. Kotchar, personal communication).

In contrast to the Syunik, Karabakh and (to a lesser extent) Iranian regions, the Ararat, North and West regions are not significantly differentiated from each other and have higher affinities with samples collected in neighbouring countries (notwithstanding the less thorough collection protocols employed in the comparative data sets). This suggests a history of greater admixture both between these regions and with neighbouring countries. These regions are less mountainous and offer lower barriers to migration. Furthermore, unfavourable treatment at various times by foreign powers exerting political dominance over Armenia has led to periods of relocation of Armenian peoples and to greater mixing. In the most recent example, violence by officials of the Ottoman Empire directed against Armenians living in the West region led to many Armenians migrating to Ararat and the North (Hovannisian 1997; Redgate 1998).

Georgia appears genetically isolated not only from Armenia but also from all other data sets in our comparison (Fig. 4). It has two high-frequency haplotypes in hg2 (haplotype 78, 16.2%; and haplotype 134, 8.8%) found at high frequency in no other data set in this study. In contrast, Turkey has much higher genetic affinities with other data sets. This supports the hypothesis that patterns of migration into or out of the Middle East occurred to a much larger extent via Anatolia and to the west of the Black Sea than via Georgia to the east of the Black Sea. Further resolution of these migration patterns will require more extensive sampling of populations to the north and east of Armenia.

The BATWING results support a signal of population growth starting in the Neolithic (95% credible interval: 2.0–11.1 thousand years ago), and this signal appears robust to the model of population stratification used. This signal does not necessarily reflect the demographic history of the Armenian population per se. It may, for example, be present in other populations that at some time shared common demographic and genealogical histories with the Armenian sample. However, the date suggested for the start of population growth is consistent with the region's archaeological evidence that suggests an increase in local settlement density dating to the Kuro-Araxian culture c. 3500–2500 BC (Badalyan 1986). It will be interesting to compare the signal observed in this study with dates derived from data sets collected from other geographically localised regions in Southwest Asia as and when they become available. Pritchard et al. (1999), using a Bayesian summary-statistic rejection sampling method under the same demographic and mutation models as used here, found a similar signal in Y chromosome data (eight microsatellite markers) grouped into eight broad geographic regions around the world including 60 chromosomes from "West Asia" (mean: 17 thousand years ago, 95% credible interval: 5–43). While both methods support a Neolithic growth signal, credible intervals are wide, illustrating the difficulty in achieving firm inferences on growth using genetic data alone. Furthermore, uncertainty in the underlying demographic and mutational model, and of the male intergeneration time (Tremblay and Vézina 2000), means that these credible intervals should be stretched even further. Given this current uncertainty, absolute dating with confidence is difficult. However, in future the BATWING method should prove useful as a tool for studying comparative growth patterns between different human populations.

Wilson et al. (2001) have observed haplotype 3 (which they have called the Atlantic Modal Haplotype) to be modal in the Welsh, Basques and Irish. They suggest that it is a signature haplotype of the Palaeolithic peopling of Europe. It is interesting to observe that the Atlantic Modal Haplotype was found in the separate isolated regional samples of Syunik (7.9%) and Karabakh (2.8%) but not in the other four Armenian regions. Furthermore, the modal haplotype in these two regions (haplotype 1) is a one-step neighbour of the Atlantic Modal Haplotype that is also found at highest frequencies in Syunik and Karabakh (14.3% and 11.2%, respectively). The frequencies of the Atlantic Modal Cluster (defined as the Atlantic Modal Haplotype plus its one-step neighbours) are 24.3% in Syunik, 14.0% in Karabakh, and less than 10.0% in all other regions and data sets in our study apart from England (41.0%) and Friesland (36.2%). The frequency in Syunik is significantly greater than in all other non-Western European data sets included in this study. While it is not possible to discount convergent drift as an explanation for these results, it is worth noting that the more geographically isolated regions of Armenia differ from those areas that are more accessible by displaying a closer genealogical affinity to the Atlantic populations. If it is not the consequence of drift, the Atlantic Modal Cluster may represent a remnant paternal signal of an ancient, possibly pre-Neolithic population that spread from Southeast Asia into Europe. It will be interesting to determine whether the Atlantic Modal Haplotype and Cluster are detected at high frequencies in other isolated locations in future surveys of Europe and the Near East.

Acknowledgements We thank all those who agreed to donate DNA for this study and the following persons who helped in its collection: Haim Ben-Ami, Anna Bridges, Christian Cormack, Mohammed Issam, Steve Jones, Theologos Loukidis, Pieta Nasannen, P. Nymadawa, Fekret Osman, Hicram Pala, Ruslan Ruzibakiev, Claire Swetman and Deborah Weiss. M.G.T. is supported by a Nuffield Foundation grant (NUF-NAL).

References

Badalyan RS (1986) Early Bronze culture in the territory of Shirak. Thesis, Yerevan (in Russian)

Bianchi NO, Catanesi CI, Bailliet G, Martinez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera R, Lopez-Camelo JS (1998) Characterization of ancestral and derived Y-chromosomal haplotypes of New World populations. Am J Hum Genet 63: 1862–1871

Bournutian GA (1993) A history of the Armenian people. Mazda, Costa Mesa

Bradman N, Thomas MG, Goldstein DB (2000) The genetic origins of Old Testament priests. In: Renfrew C (ed) America past, America present: genes and languages in the Americas and beyond. Oxbow, Oxford, pp 31–44

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton

Djahukian GB (1987) The history of the Armenian language. Academy of Science, Yerevan

Djahukian GB (1990) Did Armenians live in Asia Anterior before the twelfth century BC? In: When words collide. Ann Arbor, pp 25–31

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Gamkrelidze TV, Ivanov VV (1984) Indo-European and Indo-Europeans: a reconstruction and historical typological analysis of a protolanguage and protoculture. Tbilisi (in Russian)

Goldstein DB, Zhivotovsky LA, Nayar K, Linares AR, Cavalli-Sforza LL, Feldman MW (1996) Statistical properties of the variation at linked microsatellite loci: implications for the history of human Y chromosomes. Mol Biol Evol 13:1213–1218

Hammer MF (1994) A recent insertion of an *Alu* element on the Y chromosome is a useful marker for human population studies. Mol Biol Evol 11:749–761

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

Hovannisian RG (1997) (ed) The Armenian people. St Martin's, New York

Jobling MA, Tyler-Smith C (1995) Fathers and sons – the Y-chromosome and human evolution. Trends Genet 11:449–456

Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, Seielstad MT, Batzer MA (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal and Y-chromosome data. Am J Hum Genet 66:979–988

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, PerezLezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel Greding J, Weichold G, de Knijff P, Roewer L (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Krüger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, Sajantila A (2000) Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet 66:1580–1588

Kayser M, Brauer S, Weiss G, Schiefenhövel W, Underhill PA, Stoneking M (2001) Independent histories of human Y chromosomes from Melanesia and Australia. Am J Hum Genet 68: 173–190

Kushnareva KK (1990) Southern Caucasus in the 9–7th millennium BC. Sankt, St. Petersburg (in Russian)

Malaspina P, Cruciani F, Ciminelli BM, Terrenato L, Santolamazza P, Alonso A, Banyko J, Brdicka R, García O, Gaudiano C, Guanti G, Kidd KK, Lavinha J, Avila M, Mandich P, Moral P, Qamar R, Mehdi SQ, Ragusa A, Stefanescu G, Caraghin M, Tyler-Smith C, Scozzari R, Novelletto A (1998) Network analyses of Y-chromosomal types in Europe, northern Africa, and western Asia reveal specific patterns of geographic distribution. Am J Hum Genet 63:847–860

Malaspina P, Cruciani F, Santolamazza P, Torroni A, Pangrazio A, Akar N, Bakalli V, Brdicka R, Jaruzelska J, Kozlov A, Malyarchuk B, Mehdi SQ, Michalodimitrakis E, Varesi L, Memmi MM, Vona G, Villems R, Parik J, Romano V, Stefan M, Stenico M, Terrenato L, Novelletto A, Scozzari R (2000) Patterns of male-specific inter-population divergence in Europe, West Asia and North Africa. Ann Hum Genet 64:395–412

Mallory JP (1989) In search of the Indo-Europeans: language, archaeology and myth. Thames and Hudson, London

Mathias N, Bayés M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. Hum Mol Genet 3:115–123

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142:1061–1064

Moran PAP (1975) Wandering distributions and the electrophoretic profile. Theor Popul Biol 8:318–330

Movsessyan A, Kotchar N (2001) Ancient populations of Armenia and their role in the genetic formation of modern Armenians (by non-metric cranial traits data). Vestnik antropologii (in press) (in Russian)

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Pérez-Lezuan A, Calafell F, Mateu E, Comas D, Bosch E, Martínez-Arias R, Clarimón J, Fiori G, Luiselli D, Facchini F, Pettener D, Bertranpetit J (1999) Sex-specific patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. Am J Hum Genet 65:208–209

Pritchard JK, Seielstad MT, Pérez-Lezuan A, Feldman MW (1999) Population growth of human Y chromosomes: a study of Y chromosome microsatellites. Mol Biol Evol 16:1791–1798

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49:1280–1283

Redgate AE (1998) The Armenians. Blackwell, Oxford

Renfrew C (1987) Archaeology and language: the puzzle of Indo-European origins. Pimlico, London

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, CÔrte-Real HBSM, de Knijff P, Decorte R, Dubrova TE, Evgrafov O, Gilissen A, Glisic S, Gölge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar A-V, Mitchell RJ, Nafa K, Nicholson J, Nørby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previderé C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67:1526–1543

Santos FR, Gerelsaikhan T, Munkhtuja B, Oyunsuren T, Epplen JT, Pena SDJ (1996) Geographic differences in the allelic frequencies of the human Y-linked tetranucleotide polymorphism DYS19. Hum Genet 97:39–313

Scozzari R, Cruciani F, Santolamazza P, Malaspina P, Torroni A, Sellitto D, Arredi B, Destro-Bisol G, De Stefano G, Rickards O, Martinez-Labarga C, Modiano D, Biondi G, Moral P, Olckers A, Wallace DC, Novelletto A (1999) Combined use of biallelic and microsatellite Y-chromosome polymorphisms of infer affinities among African populations. Am J Hum Genet 65:829–846

Seielstad MT, Herbert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D, Cavalli-Sforza LL (1994) Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. Hum Mol Genet 3:2159–2161

Seielstad MT, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. Nat Genet 20:278–280

Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiæ M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA (2000) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. Science 290:1155–1159

Thomas GM, Bradman N, Flinn HM (1999) High throughput analysis of 10 microsatellite and 11 binary polymorphisms on the human Y-chromosome. Hum Genet 105:577–581

Thomas GM, Parfitt T, Weiss DA, Skorecki K, Wilson JF, le Roux M, Bradman N, Goldstein DB (2000) Y chromosomes traveling south: the Cohen Modal Haplotype and the origin of the Lemba – the "Black Jews of Southern Africa." Am J Hum Genet 66:674–686

Tremblay M, Vézina H (2000) New estimates of intergenerational time intervals for the calculation of age and origins of mutations. Am J Hum Genet 66:651–658

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

Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonné-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ (2000) Y chromosome sequence variation and the history of human populations. Nat Genet 26:358–361

Vogt PH, Affara N, Davey P, Hammer MF, Jobling M, Lau Y-FC, Mitchell M, Schempp W, Tyler-Smith C, Williams G, Yen P, Rappold GA (1997) Report of the third international workshop on Y chromosome mapping 1997. Cytogenet Cell Genet 79: 1–20

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

Wilson IJ, Balding DJ (1998) Genealogical inference from microsatellite data. Genetics 150:499–510

Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB (2001) Genetic evidence for different male and female roles during cultural transitions in the British Isles. Proc Natl Acad Sci USA 98:5078–5083

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C (1997) Genetic relationships of Asians and northern Europeans, revealed by Y-chromosomal DNA analysis. Am J Hum Genet 60:1174–1183