

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

© Springer-Verlag 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.

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 km², 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-

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 Bldg, 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



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 ex-

ample, 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