



Jadwiga Pstrusińska (University of Warsaw, Poland)

ON THE ORIGIN OF IRANIAN-SPEAKING NOMADS OF THE EURASIAN STEPPES IN THE LIGHT OF HUMAN POPULATION GENETICS

Keywords: Nomads of the Eurasian Steppes, history and geography of human genes, genetics

Many diverse ideas can be read about the early history of the so-called Iranian Nomads of the Eurasian Steppes. The aim of this paper is to present remarks on their origin in the light of the research undertaken, more or less within the last forty years, by molecular biologists¹ who have created such new disciplines as population genetics, history and geography of human genes or archaeogenetics, unveiling past migrations, diffusions and relationships among groups of populations living in Eurasia as well as the rest of the world. At present, scholars can support or exclude their ideas on the matter including data provided by geneticists researching mitochondrial DNA and Y chromosome. We already have, at our disposal, a vast amount of highly professional literature which is still developing. Numerous new publications on the subject appear constantly. Let us mention here the award-winning book published in 2003 by John H. Relethford, entitled *Reflections of Our Past: How Human History Is Revealed In Our Genes*. Of course, this does not mean that we already know everything and that numerous traps do not have to be avoided. Understandably, many geographical and historical regions and periods have not yet been fully examined by these relatively young branches of human genetics. Much still has to be done. Certainly, however, we should not ignore such data that already exists. It seems, in the end, that the field of humanities gained something crucial from what is perceived as hard

¹ This text avoids the very specialised terminology used by molecular biologists and will not describe their highly specialised methodology of research, instead it concentrates just on the final results relevant to this topic.

science and we would not have to guess the truth regarding remote human dispersion. Let us add that the recent speed of development within the field of molecular biology is enormous in comparison with developments within the so-called humanities.

The beginning of the new discipline, population genetics, was not very easy and the key scholar for its development, Luigi L. Cavalli-Sforza had many difficulties initially to convince academic gremia about his concepts, but now his ideas about the history and geography of human genes have been already broadly implemented, much improved and numerous teams of scholars work on such matters because, as we read, there is no doubt about the relationship between genetics and history.² We find that not only the MtDNA says much but the human Y chromosome has proven to be a valuable tool for the study of population history, allowing complex demographic events to be deconstructed despite extensive admixtures in some geographic regions. *Y chromosome provides anthropologists and geneticists with an extremely powerful tool for historical and demographic studies* (Wells et al. 2001, 10244). Let us add that this genetic research, which took into account the Y-chromosome, in selected Eurasian populations as well, shows understandable male migration.

Michael Witzel has already written several years ago: *It must be pointed out that genetic evidence, though still in its infancy, is often superior to (even multi-variate) paleontological evidence [...]. Genetic evidence frequently allows to pinpoint (sub-)branches in the cladistic tree at a particular point in time and space* (Witzel 2001, 9). Data provided by genetics for both historical and demographic studies are, of course, important for many linguistic studies as well and especially crucial for historical linguistics and sociolinguistics. Despite several exceptions, which can be easily shown, geneticists stress that generally both genes and languages are normally passed from parents to children through generations.³ Contemporary population genetics started to compare not only ethnic and geographical areas but linguistic regions as well.⁴ Studying the population genetics publications we can find, in fact, not only numerous texts confirming the correlation between genetic data and languages distribution but also texts which deny such a possibility in many cases (e.g., Rosser 2000). We can read about numerous correlations between some linguistic and ethnic boundaries in Eurasia, already discovered by scholars, using other tools and other methodologies, but simultaneously we find other data which does not confirm such relations in other examined cases. From a linguistics point of view, however, both results are informative and might be useful in writing the history and sociolin-

² E.g., Sykes 2001, 144; Zerjal 2002b.

³ E.g., Cavalli-Sforza et al 1994, 99; Wells 2001; Olson 2003, 163–164.

⁴ E.g., Nasidze 2004a, 214; Manni, Barrai 2000.

guistic history of a given area. It might perhaps help to understand better, for instance, the role of convergency in the creation of the so-called linguistic families and their branches and subbranches. Generally, there is hope that archeology, linguistics and population genetics will together eventually explain in a better way our history achieving a new kind of synthesis (Olson 2003, 160). We already have numerous publications, which are the results of such interdisciplinary and transdisciplinary insights.⁵

Thanks to this research within population genetics and a fresh insight into the demographic and migrational history of Eurasia, we learn that Asia has served as a focal point for human migration (Karafet 2001) and the importance of the area designated here as Central Asia (Wells et al. 2001, 10244–10249) in this respect gained in recent years special strength (e.g., Pstrusińska 2009, 282–290). As a result of this new scientific input into the humanities made by population genetics, we can read about many new views on our history, some of them rather surprising. What, however, can we learn from the history and geography of human genes regarding the group of so-called Iranian people, among whom scholars traditionally include Scythians, Sarmatians and Alans, placed among Iranian nomads of Eurasian Steppes.

One of the most interesting publications relevant, among others, to the question of the origin of the Scythians, Sarmatians and Alans is the paper entitled ‘The Eurasian Heartland: A continental perspective on Y chromosome diversity’ which appeared in the series *Proceedings of National Academy of Sciences* (Wells et al. 2001, 10244–10249) showing research results obtained by an international team composed of about thirty scholars working in dozens of research centres. Numerous laboratories, starting from Oxford to Archangielsk through Yerevan, Tbilisi, Madurai in India to Chicago and so on have taken part in the project. Some data previously published have been also included. Y Chromosomes of men belonging to 49 Eurasian populations have been examined primarily during three expeditions organized in 1996, 1998 and 2000, with a particular focus on Central Asia. The research was undertaken by a large team of geneticists, headed by R. Spencer Wells, working in several maternal laboratories around the world, having simultaneously its main base at Oxford University, who published their research results at the end of August 2001 (Wells et al. 2001, 10244). We read here, as well, a most crucial piece of information, recently confirmed by geneticists, that Y chromosomes reveal traces of historical migrations, and provide an insight into the earliest patterns of settlement of anatomically modern humans on the Eurasian continent and Central Asia is revealed to be an important reservoir of genetic diversity, and the source of at least three major waves of migration leading into Europe, the Americas, and India (Wells et al. 2001, 10244).

⁵ E.g., Renfrew 2000a; Renfrew 2000b; Rosser 2000; Bellwood, Renfrew 2003.

It was discovered in the course of examination of the collected samples, mentioned above, that the Pakistani population of Central Asia shows the highest genetic diversity in Eurasia being among the oldest on the continent. We read: *In addition, Y chromosome microsatellites indicate that Central Asian (Pakistani) populations are the most diverse in Eurasia. [...] This pattern of high diversity is consistent with an early settlement of Central Asia...* (Wells et al. 2001, 10247). The region near the mountainous knot created by Hindukush, Pamir and Karakorum was discovered as a very important reservoir of genetic human diversity existing at present in Eurasia and that *these results are compared with data from other populations in an effort to reconstruct the history of early human migrations in Eurasia, as well as more recent events in the region of Central Asia* (Wells et al. 2001, 10247). Moreover, in the discussed project, the genetic results have been interpreted in the context of Eurasian linguistic patterns. An effort has been made to collect samples of several linguistic groups speaking languages belonging to the so-called Afro-Asiatic, Indo-European, Dravidian, South Caucasian, North Caucasian, Altaic, Uralic, and Sino-Tibetan language families. Some data taken from existing literature have been added (Wells et al. 2001, 10244).

We read that this pattern of high genetic diversity *is consistent with an early settlement of Central Asia by anatomically modern humans, perhaps 40,000 – 50,000 years ago, followed by subsequent migrations into Europe, America and India...* (Wells et al. 2001, 10244). Dispersed lineages reached the eastern and the western extremes of the continent. One such group of the Central Asian sub-population migrated westwards and gave rise to the population of the greater part of Europe (Wells et al. 2001, 10244). To present more clearly the results of their research, scholars involved in the project have prepared the so-called *neighbour-joining tree which shows several population clusters defined by branches from a central point* (Wells et al. 2001, 10244). Following this we can show here, for instance, how speakers of the so-called Indo-Iranian languages⁶ are placed in this publication within several separate waves of population, diverse genetically and originating in distant periods of time.

Cluster I

Greek, **Yagnobi**, Armenian, Turkmen, Czech/Slovak, Orkney, British, Basque.

Cluster II

Kurdish, Middle Eastern, **Ossetian**, **Shiraz**, **Tehran**, Lazgi, Svanetian, Lebanese, **Isfahan**, **Iranian Sam.**, Turkish, Kazbegi, Azeri, Tuvinian and Nenets.

Cluster IV

Mongolian, Kazakh, Cambodian, Dungan, Chinese, Taiwanese, Korean, Japanese.

⁶ Designations of population groups almost as found in Wells et al 2001. The so-called Aryans (Indo-Iranians) shown in bold.

Cluster V

Macedonian, Pomor, Saami, Russian North, Russian Tashkent, Ukrainian, Kyrgyz, **Tajik Khojant, Ishkashim.**

Cluster VI

Bartangi, Sinte Romani, Hunza

Cluster VII

Kallar, Sourashtra, Yadava, Dushanbe Tajik, Shughnan, Samarkand Tajik, Arab Buchar.

Cluster VIII

Uzbek, Tatar, Karakalpak.

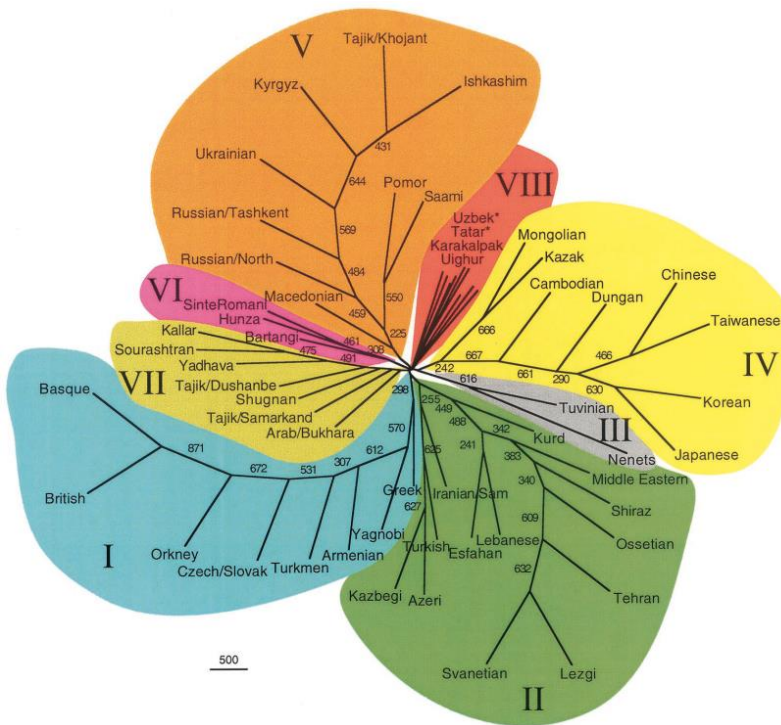


Fig. 1. Neighbor-joining tree of 61 Eurasian populations, based on Y-chromosome biallelic haplotype frequencies. Reproduced after: Wells et. al 2001, Fig. 2 (www.ncbi.nlm.nih.gov/pmc/articles/PMC56946)

The figure of neighbour-joining tree shown in the literature presents the examined population groups as so-called clusters defined by eight branches going from a central point. Thousands of years passed between the separate clusters, this means separate mutations and, as a consequence, later separate migratory

waves. This neighbour-joining tree could be discussed and commented almost endlessly taking into account numerous aspects. Cluster number II is very revealing, composed, as we see, of Kurds⁷, Middle Eastern population, Ossetian, people of Shiraz, Tehran and Isfahan, Lazgi, Svanetian, Lebanese, Iranian Sam, Turkish, Kazbegi and Azeri (Wells et al. 2001, 10247). It is visible from this neighbour-joining tree that the so-called Indo-Iranian people speaking Iranian languages do not exist as one genetically related group of a common origin. Simultaneously, it is shown that the languages grouped as the so-called Indo-Iranian family have been spoken and still are spoken by the population belonging, according to this genetic research, to six separate clusters originated in different times in Central Asia and each one having its own genetic specificity as far as Y chromosome is concerned. We read: *Intriguingly, the population of present day Iran [...] appears to have had little genetic influence from the wave carrying Indo-Iranians [...] population of Iran is mainly an eastern extension of the great civilizations of Mesopotamia and language replacement and the later usage of Persian language can be explained by elite-dominance model. [...] The sketch of Eurasian population movements outlined here is based entirely on Y chromosome evidence. The actual history of these populations presumably has included the migration of women [...] thus we await further study of their DNA* (Wells et al. 2001, 10244).

Such research results can certainly influence our thinking on certain basic matters within Iranian studies, ancient history, historical linguistics, religious studies and many, many others. By the way, the origin and dispersal of the so-called Indo-Iranians has been rightly perceived as one of the greatest puzzles of the so-called Indo-European studies (e.g., Mallory, Mair 2000, 258). At present, we should take into account that it was written in the already quoted *Proceedings of National Academy of Sciences* (Wells et al. 2001, 10244–10249) that the population of present day Iran, speaking Farsi appears to have had little genetic influence from cluster number VII carrying most of the Indo-Iranian speakers and genetically was in fact the eastern extension of the great civilisations of Mesopotamia. Thus, it was rather only a language replacement and not a broad dispersal of Indo-Iranians. We learn, for instance, that the inhabitants of western Iran appear to be more similar genetically to Afro-Asiatic-speaking Middle Eastern populations than they are to Central Asians or Iranian populations. It seems that the Dasht-e Kavir and Dasht-e Lut deserts have acted as significant barriers to gene flow (Wells et al. 2001, 10248). It is relevant also to the so-called Iranian-Speaking Nomads of the Eurasian Steppes, including the Scyths, Sarmatians and Alans populations. Thus we cannot treat them any more as people of clearly Indo-Iranian provenance.

⁷ See also Pstrusińska 2004.

Moreover, we should expect further research regarding admixtures within genetic pictures of the so-called Iranian Nomads of the Eurasian Steppes. Latyshev has already written (Latyshev 1949, 303) that Alans gradually started to include other people into their own area (e.g., Oranskij 1988, 72). Such processes started certainly much earlier. Alans are, as we most often read, a large group of nomadic tribes first mentioned far in the East in the 2nd century BC in the Chinese Annals of the Han dynasty and in classical authors of the 1st century AD. They display close relations to the Scythians and Sarmatians and their language has been included in the Iranian group. We perceive Ossetic as a modern continuation of Alanic (e.g., Schmitt 2000, XIII). Scythians (often designated in English as Scytho-Sarmatian) originated from seminomadic tribes in the 1st millennium BC. Scythians resided north of the Black Sea, within the period of the 8th–7th century BC – 4th–5th century AD.⁸ Let us mention that according to genetic research we learn, among other things, that: *Putting together the archaeological and genetic data, and assuming a common origin of South and North Ossetians (which is supported by the mtDNA data) a plausible scenario is that “alteration” of the initial Osetian Y-chromosome gene pool took place in North Caucasus groups. ... Genetic studies of such complex and multiple migrations as the Ossetians can provide additional insights into the circumstances surrounding such migrations.* (Nasidze 2004b, 597, 598).

L. L. Cavalli-Sforza stated in his 1994 publication that it might be difficult to differentiate Sarmatians genetically due to the fact that their origin was similar to that of Scythians. Sakas and Massagetians have been most likely their eastern reflexion (Cavalli-Sforza 1994, 201, 295). Let us mention also that, interestingly, there is strong evidence for limited Greek contribution to the Pathan (Pashtuns) population of Pakistan, whose tribes are perceived as related to Scythians, Sarmatians and Alans. S. Firasat writes: *The genetic data alone do not tell us when the Balkan chromosomes arrived in Pakistan: it is necessary to turn to the historical record for this. There has been no known Greek admixture within the last few generations, but in addition to Alexander’s armies, the possibility of admixture between Greek slaves, who were brought to this region by Xerxes around one hundred and fifty years before Alexander’s arrival, and the local population cannot be discounted. At that time Afghanistan and present day Pakistan were part of the Persian Empire. Nevertheless, Alexander’s army of 25,000–30,000 mercenary foot soldiers from Persia and West Asia and 5000–7000 Macedonian cavalry perhaps provides a more likely explanation because of their elite status and substantial political impact on the region* (Firasat 2007, 125).

The Jasz (Jassy) people of Central Europe, giving one more example, are perceived as descendants of Alans, and according to the population genetics they

⁸ Isaev 2000a, 105; Isaev 2000b, 107.

together with many groups living in Hungary have Iranians as their parental ancestors (Guglielmino 2000, 145–159). The study of genetic admixture aims to increase and substantiate the historical knowledge of populations, especially in cases where history is not fully documented. This study also shows how varied the genetic composition of relatively small populations can be (Guglielmino 2000, 158).

Coming back to the neighbour-joining tree, we see groups related genetically but speaking languages placed by scholars into several language families: Kurds and some others like Ossetians and inhabitants of Shiraz or Tehran speak the so-called Iranian ones. Ossetians are of key importance for our discussion as they are perceived as contemporary descendants of Alans. (People of cluster II, designated here as Middle Eastern, speak Afro-Asiatic languages namely Arabic, Hebrew etc. Others speak Turkic languages and Dene-Caucasian ones). It is clear that following the picture given by the neighbour-joining tree it is rather impossible to place all Iranian-speaking Nomads genetically only within a so-called Indo-Iranian group of people. The neighbour-joining tree discussed here suggests that many so-called Indo-Iranian speaking groups are on a biological level not related to Iranians living west of Dasht-e Kavir and Dasht-e Lut and placed in cluster II, thus they do not have a common Indo-Iranian origin. What is crucial for our discussion is that Ossets, perceived as descendants of Alans, belong also to cluster II. Interestingly, the well-known Russian linguist, D.I. Edelman, concluded some years ago that the contemporary spoken Iranian languages are not autochthonous even in one case (Edelman 2002, 11). Nasidze wrote that in the region of Caucasus we find geographical and not linguistic correlations. It is relevant to the languages of the Ossets and some others from the so-called Iranian group of languages. Mountains, in this case the Caucasus, as we see, had not been a strong barrier in the flow of genes (Nasidze 2004a, 218). However, deserts have created such barriers. Just recently an Iranian scholar working at Portsmouth University, among other researchers, adds to the discussion stressing the non-Aryan origin of “Iranians”. *The Aryan genetic markers that exist in central Asia and Caucasus are found very few in the Iranian Plateau*. Over 2600 Iranian DNA-sequences belonging to 26 diverse contemporary Iranian groups and those belonging to ancient Iranian bones have been researched.⁹ Some scholars suggest that simultaneously we should take into account the Proto-Turkic and then Turkic (in broad sense) people and their languages while discussing origin of Iranians.¹⁰ It should be applied to the so-called Iranian-speaking Nomads of the Eurasian Steppes and their languages as well.

⁹ Ashrafian 2013; Chikisheva 2013.

¹⁰ Chikisheva 2013; Klyosov 2010; Comas 2004; Zerjal 2002a; Calafell 2000.

To sum up it was shown by population genetic research of recent decades that the populations grouped in the past as Indo-Iranians cannot be put together, as Indians and Iranians are of common origin, because they belong to several diverse genetic clusters originated in Central Asia separately and having their own genetic specificity. This fact adds to the accuracy of our picture of the origin of the so-called Iranian-speaking Nomads of the Eurasian Steppes as well. Molecular genetics, without any doubt, will gradually further clarify our knowledge of the subject.

Bibliography

- Ashrafian, B. 2013: 'The non-Aryan origin of "Iranians"'. Genetic evidence R1a', <http://youtube.com/watch?v=fin-leETI>, accessed: 10.06.2014.
- Bellwood, P., Renfrew, C. (eds.) 2003: *Examining the farming/language dispersal hypothesis*, Cambridge.
- Calafell, F. 2000: 'Genetics and Population History of Central Asia' in Renfrew C., Boyle K. (eds), *Archeogenetics: DNA and the population prehistory of Europe*, Cambridge, 259–266.
- Cavalli-Sforza, L.L. 1994 : *The History and Geography of Human Genes*, Princeton.
- Chikisheva T. 2013: 'Mongoloid Caucasoids of Mesolithic' [lecture online, <http://youtube.com/watch?v=fjn-leETI>, accessed 11.12.2014].
- Comas, D. 2004: 'Admixture, migrations, and dispersals in central Asia: evidence from maternal DNA lineages' *European Journal of Human Genetics* 12, 495–504.
- Edelman, D.I. 2002: *Iranskie i slavjanskie jazyki. Istoricheskie otnosheniia*, Moskva.
- Firasat, S. et al. 2007: 'Y-chromosomal evidence for a limited Greek contribution to the Pathan population of Pakistan' *European Journal of Human Genetics* 15, 121–126.
- Guglielmino, C.R. et al. 2000: 'Probable ancestors of Hungarian ethnic groups: an admixture analysis' *Annals of Human Genetics* 64, 145–159.
- Isaev, M.I. 2000a: 'Skifskii jazyk' in *Jazyki mira, iranskie jazyki 3. Vostochnye jazyki*, Moskva, 107–110.
- Isaev, M. I. 2000b: 'Alanski jazyk' in *Jazyki mira, iranskie jazyki III. Vostochnye jazyki*, Moskva, 105–107.
- Karafet, T. et al. 2001: 'Paternal population history of East Asia: sources, patterns, and microevolutionary processes' *The American Journal of Human Genetics*, 69(3), 615–628.
- Klyosov, A.A. 2010: 'The principal mystery in the relationship of Indo-European and Turkic linguistic families and an attempt to solve it with the help of DNA genealogy; reflections of a nonlinguist' *Journal of Russian Academy of DNA Genealogy* 3/1, 3–58.
- Latyshev, V.V. 1949: 'Izvestia drevnikh pisatelej o Skifii i Kavkaze', *VDI*, 1949/3.
- Malaspina, P. et al. 2000: 'Human Y-chromosomal Networks and Patterns of Gene Flow in Europe, West Asia and North Africa' in Renfrew, C., Boyle, K. (eds.), *Archeogenetics: DNA and the population prehistory of Europe*, Cambridge, 163.
- Mallory, J.P., Mair, V.H. 2000: *The Tarim Mummies. Ancient China and the Mystery of the Earliest Peoples from the West*, London.
- Manni, F., Barraï, I. 2000: 'Patterns of Genetic and Linguistic Variation in Italy: a Case Study' in Renfrew, C., . Boyle, K. (eds.), *Archeogenetics: DNA and the population prehistory of Europe*, Cambridge, 333–338.

- Nasidze, Y.S. 2004a: 'Mitochondrial DNA and Y-Chromosome Variation in the Caucasus' *Annals of Human Genetics* 68/3, 205–221.
- Nasidze, Y.S. 2004b: 'Genetic Evidence Concerning the Origins of South and North Ossetians' *Annals of Human Genetics* 68/6, 597–598.
- Olson, S. 2003: *Mapowanie historii ludzkości. Przeszłość ukryta w naszych genach*, Poznań.
- Oranskiĭ, I. M. 1988: *Vvedenie v iranskuĭu filologiiu*, Moskva.
- Pstrusińska, J. 2004: 'Remarks on the origin of Kurds in the light of Human population genetics' in *XXX International Conference on Kurdish Studies, May 2004 (Polish Academy of Sciences, Cracow Branch, publications of the Oriental Committee 25)*, Cracow, 47–51.
- Pstrusińska, J. 2009: 'On the Origin of Central Asia for Biosociolinguistic Research' in Gacek, T., Pstrusińska, J. (eds.) *Proceedings of the Ninth Conference of the European Society for Central Asian Studies*, Cambridge, 282–290.
- Relethford, J.H. 2003: *Reflection of Our past: How Human History Is Revealed in Our Genes*, Boulder CO – Oxford.
- Renfrew, C. (ed.) 2000a: *America Past, America Present: Genes and Languages in the Americas and Beyond*, Cambridge.
- Renfrew, C. et al. (2000b) 'Foreword' in Renfrew, C. et al. (eds.), *Time Depth in Historical Linguistics, vol. 1, Papers in the Prehistory of Languages*, Oxford, vii–ix.
- Rosser, Z.M. et al. 2000: 'Phylogeography of European Y-chromosome Diversity' *The American Journal of Human Genetics* 67/3, 1526–1543.
- Schmitt, R. 2000: 'Foreword' in A. Alemany *Sources on the Alans. A Critical Compilation (Handbook of Oriental Studies)*, Leiden – Boston – Köln, xiii – xv.
- Sykes, B. 2001: *The seven daughters of Eve*, London.
- Wells, R.S., Yuldasheva, N., Ruzibakiev, R., Underhill, P.A., Evseeva, I., Blue-Smith, J. et al. 2001: 'The Eurasian Heartland: A continental perspective on Y chromosome diversity' *Proceedings of the National Academy of Sciences of the USA* 98 (18), 10244–10249 (doi:10.1073/pnas.171305098).
- Witzel, M. 2001: 'Autochthonous Aryans? The Evidence from Old Indian and Iranian Texts' *Electronic Journal of Vedic Studies* 7(3), 9.
- Zerjal, T. et al. 2002a: 'A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia' *The American Journal of Human Genetics* 71(3), 466–482.
- Zerjal, T. et al. 2002b: 'What can Y-chromosomal DNA Analysis Contribute to the Understanding of Prehistory?' in Boyle, K., Renfrew, C., Levin, M. (eds.), *Ancient interactions: east and west in Eurasia*, Cambridge, 315–325.

Abstract

We learn from the history and geography of human genes regarding the origin of the so-called Iranian people, the Iranian-speaking Nomads of the Eurasian Steppes inclusive, among whom scholars traditionally include Scythians, Sarmatians and Alans, that they cannot necessarily be treated as belonging in the past to an Indo-Iranian community speaking an Indo-Iranian language.