Richard Villems

Marginalia on the Topic of Identity

1. The Significance and Insignificance of Genetic Identity

Thucydides, citing Pericles' speech in praise of the soldiers who fell at the beginning of the Peloponnesian War, presented an exceptionally revealing description of national identity, through the essence of Athenian social institutions. Now, nearly two and a half millennia later, it is difficult to be more eloquent and clear. My topic, however, is altogether a different matter – genes. I would like to begin with an observation from my long career as a university lecturer. In introducing students to contemporary human genetic variation, I always observe with great interest their reaction to the assertion that the DNA of two Estonian students, randomly selected from that very auditorium, would presumably exhibit differences six times greater than the differences between the DNA of ten thousand randomly selected Bantu and ten thousand randomly selected Estonians. As a rule, over half of the students are unwilling to accept this assertion. I then explain the quantitative evidence resulting from the research performed by Richard Lewontin *et al* a couple of decades ago and the real meaning and explanation of the assertion: humankind is relatively young and thus a very homogeneous species, much more homogeneous than gorillas or chimpanzees. At this point there remains in the lecture hall only a relatively significant minority for whom there remains a contradiction with this assertion: If genetics (or rather geneticists) do not wish to distinguish Estonians from the inhabitants of Tierra del Fuego or pygmies, and consider that difference to be significant, so much the worse for genetics (geneticists). There is also a small contingent that does not heed Lewontin's arguments – staunch Christians and the followers of the great ideals of the Enlightenment, for whom the unity of humankind is axiomatic, not to mention Buddhists, in whose opinion the unity of humankind is expressed in the fact that only those (re)born as human beings can break the chain of birth and rebirth: for them, all else is secondary in this context.

It may appear incongruous to begin a chapter in a book devoted to the reception of Estonian identity with a request to forget for a moment national identity while reading the following, and instead concentrate on the prehistory of humankind; or at least make a clear distinction between (the sentiment of) national identity and a people's genetic identity. I would, however, encourage the reader to do so; it would in any case be advisable. Is it, however, also

PDF created with pdfFactory trial version www.pdffactory.com

possible? We can (I can) declare in the strongest possible terms the need to separate these concepts, but does that mean that it is possible in psychological terms? Even in Europe? Here I am not referring to the gangs of skinheads that operate almost everywhere. Last year Berkeley socio-anthropologist Paul Rabinow, who achieved fame with his Moroccan research at the end of the 1970s, created havoc with the book *French DNA: Trouble in Purgatory*, which did not concern the physical identity of the French genome, but geneticists', and others' understanding of the "Frenchness of French DNA." This was not about the genes themselves, but the reception of the concept, the objects of the research being scientists, journalists and the government, with digressions into the sociology and history of the sense of identity and technocracy. Hence the moral: on the level of the social unconscious, the area of overlap of national and genetic identity is extensive, and we must unfortunately take this into consideration, despite clear examples from recent history, where the instruments of those manipulating the sense of identity have generally been, depending on the requirements of the time, either religious affiliation or close linguistic and genetic kinship. One need only briefly consider the three entities Serbia, Bosnia and Croatia.

However, the term "national genetic identity" is not one that a geneticist finds impossible to define with accuracy. For him it is an empirical term that can be formulated as an "all-encompassing table" of the frequencies of genetic variations characteristic of a specific human population (in this case a people). It appears, however, that in the case of identity there is a desire to reveal and emphasise not so much that table, but the features that distinguish the selected population from others; and not just others, but preferably one's close neighbours. Yet this is also data that can be presented empirically: the first of the two ancient questions *who are we* and *where do we come from*, seen from the geneticist's viewpoint. The second – *where do we come from* – is of equal importance to the researcher of genetic identity. Incidentally, the question of "coming from somewhere" is apparently easier for Northern Europeans to understand on a psychological level than it is for other peoples: as predominantly rational people who have received relevant instruction in school, it is accepted that we could not have always occupied the area that twenty thousand years ago was covered by glaciers and the Scandinavian ice cap. Thus the crux of the question for the Estonian general public is also not one of *arriving/not arriving*, but instead of *arrival* and *arrivals*. From how many directions did the genetic code contained in Estonians' present-day gene

PDF created with pdfFactory trial version www.pdffactory.com

pool come? Where did the "material bearers" of those strands live in ancient times? When and in which order did they arrive here?

Contemporary biology postulates that genotype forms the basis of phenotype. Whereas the deciphering of the typical human genotype nears completion, the investigation of the connection between genotype and phenotype is just beginning. Genetic identity is a concept about which rational discussion is completely possible, while the description of a national phenotype tends to be intrusively close to ideology. On the one hand this is unavoidable, since other origins (e.g. the natural environment, behavioural customs) have been able to add only very limited empirical supporting material to the discussion. Another reason is, of course, that generalised descriptions of a national phenotype have on several occasions been tailored to suit political aims, especially as concerns psychotype. The investigation of phenotype, which is based on knowledge of genotype, i.e. the ascertainment of the ties between genotype and phenotype, is presumably an altogether more long-term task than the nearly completed international programme for the ascertainment of the structure of the human genome, and will last for several more generations. This is all the more so, as the centre stage will be occupied not by the history of human phenotypes, but instead by complex problems connected with the genetic background of the main diseases. Thus we may entertain some small hope that the increase in the level of education of the world's population could ensure this new and nascent knowledge receives a response as is required for the survival of humankind as a species. As long as the sinister shadow of the national ideas of the Third Reich lingered on – to name just one prominent 20th-century attempt to promote national identity. If there is scope for hope, one may also doubt this.

2. The Genetic History of Humankind: The Background for the Understanding of Estonians' Genetic History

I will now turn from the concept of genetic identity to the more neutral-sounding genetic history of the species, which in this context is more accurately referred to as our demographic history. In order to understand the demographic history of humankind, including that of the Estonian people, one must first recall that for over 95 percent of the time that man as a species is considered to have existed, we have lived in the Palaeolithic period. The prehistoric period represents, however, only a tiny portion of the volume of

PDF created with pdfFactory trial version www.pdffactory.com

books devoted to general history: precisely 5 percent of J. M. Roberts' *World History*, which sits on my bookshelf. Thus there is an almost exactly inverse proportion between the number of pages and the amount of time covered. The above-mentioned comparison is not intended as a reproach to historians, on the contrary: this ratio is understandable and reflects the volume of our empirical knowledge. From this point of view, even the said 5% is a great exaggeration.

Acceptance of such a disproportion does not in the least infer an underestimation of the Palaeolithic. On the contrary, one may consider the assumption that the genetic drift caused by long-term Palaeolithic isolation has determined a great deal in the history of our species to be an altogether rational working hypothesis. The same has, of course, been performed by various trends in natural selection, in correspondence with the most diverse ecological conditions (and their continual alteration) in which the members of our species spent thousands of generations. At the same time, attempts to interpret predominantly social phenomena through genetics are premature, and in many cases most likely to be altogether irrational, for the same reason that it would be irrational to attempt to describe DNA replication on the basis of the String Theory.

Historians should not, however, be surprised, much less protest, if researchers of the genetic history of humankind define periods, about which even archaeologists of the prehistoric period possess little information, as important and decisive. Most geneticists and paleoanthropologists are in agreement that our species, *homo sapiens*, originates from a hominine primate that developed in Africa about 150 thousand years ago – perhaps from one branch among other similar branches. This branch is the only one to have survived to the present day. Since here and below I am not speaking of humankind in general, but am focusing on Estonians or at least Finno-Ugrians, I will begin from a time that, from the point of view of the species, is truly recent – the settlement of Eurasia by anatomically modern man about 40 thousand years ago. Furthermore, one of the largest black holes in both paleoanthropology and genetics is the 50- to 60-thousand-year period just before the colonisation of Eurasia.

The time from this period of earliest settlement to the beginning of the last great Ice Age about 24 thousand years ago was presumably also the most important in the development of the genetic identity of the major human races. The next significant stage is the Last Great Ice Age itself, which lasted for millennia and had a direct and drastic influence on settlement density in

PDF created with pdfFactory trial version www.pdffactory.com

Europe, and also exerted an indirect influence (i.e. not through cold, but desertification) in the northern part of Africa, Arabia, present-day Iran, the Indus Valley and perhaps even on the people of more easterly regions. It is very important to know that the meticulous investigation of archaeological finds has now made it possible to identify two large refuges where proto-Europeans sheltered during the winter. These are the so-called Franco-Cantabrian region in northwest France and northern Iberia, and the middle reaches of the Dnieper-Don River north of the Black Sea. Archaeologists claim that instead of decreasing, the population density in both of these regions grew, even at the peak of the last Great Ice Age 22 thousand years ago. The third important stage was the re-settlement of northern Europe and other extensive areas, this being a long-term process that took place in parallel with the melting of layers of the inland ice sheet and increasing humidity. It is natural to presume that re-settlement took place on a large scale from the above-mentioned refuges, although not necessarily only from those areas. The settlement of the Americas by anatomically modern man is generally believed to fall within this period, as well as one other significant geophysical phenomenon, the precise influence of which remains to be assessed – the nearly hundred-metre rise in the world's oceans. The last undeniably momentous event in demographic history is, of course, the evolution to the Neolithic lifestyle, i.e. the transition from hunting and gathering to animal husbandry and agriculture, which originated from the Levant, Anatolia and northern Mesopotamia roughly 10 thousand years ago. This, on the one hand, favoured those populations that had the good fortune to live in the belt stretching from the Pacific Ocean to the Atlantic Ocean that was favourable to agriculture and animal husbandry. On the other hand, this process began very gradually to homogenise the genetic differences that had accumulated during the course of Palaeolithic isolation. In this context, increased prosperity gave rise to genetic effects: the food producers' gene pool selectively expanded at a rate much faster than that of the gatherers. There is no doubt that it predominantly expanded, but did not become deeper: genetic changes are so slow to develop that expansion only gradually leads to the enrichment of variants.

To return to discussion of the background: Palaeolithic is, like Mesolithic and Neolithic, a cultural term to which one cannot apply a precise time scale. When the Tasmanians, who were only recently destroyed or led to destruction by the white man, first encountered Europeans, they were technologically in an early stage of the upper Stone Age – an earlier stage than that of the Cro-

PDF created with pdfFactory trial version www.pdffactory.com

Magnons of Europe more than thirty thousand years ago. On the other hand, the Neolithic period in the Levant and Estonia is separated by thousands of years.

Almost all classical linguists are of the opinion that it is not possible to take reconstructions further into the past than 5000–7000 years. From the point of view of the history of Europe and the Levant, this corresponds to the beginning of the Bronze Age or the late Neolithic Age. Representatives of the Nostratian school (designating a hypothetical language family comprising Indo-European, Semitic, Altaic, and Dravidian) operate with linguistic macro-groups that are twenty thousand years old. Geneticists see no serious theoretical difficulties in reconstructing our genetic lines at least ten times further into the past. In practical terms such reconstructions are, however, simple in only two (albeit important and extremely intriguing) hereditary systems – in the observation of the genetic history of our maternal and paternal lines, and the demographic history that flows from that. In contrast to ordinary genes, these are handed down without merging: either from the maternal or the paternal side. In genetic terms, it is much easier to investigate the lineage of the genes of our mitochondrial and Y chromosomes than it is to reconstruct the history of the continually Mendel-type merging and also unpredictably recombining "ordinary" genes.

3. The History of the Finno-Ugrians' DNA and Identity: Who Are We and Where Do We Come from?

In the above subtitle I replaced the word "genetic" by the abbreviation DNA, in order to give a clear indication that what follows is not a paraphrase of the results of "classical" genetics and physical anthropology. The second limitation is that the statements below only concern the results of the investigation of the above-mentioned maternal and paternal lines. This work, especially the reconstruction of maternal lines, has now reached the stage of the first significant generalisations about humankind in general, but especially about Eurasia and Mediterranean Africa, and we can answer with much greater accuracy the question of the genetic identity of the Estonians (Armenians, Basques, etc.).

Who then are the Estonians, in genetic terms, in comparison with the people of Eurasia and the world? More precisely, where do our maternal and paternal ancestors come from, and how distant are Estonians from the

PDF created with pdfFactory trial version www.pdffactory.com

European, Asian or African "average," or indeed from their neighbours? Does the present level of development of genetics permit the identification of the elements that unite Finno-Ugrians, compared to Indo-Europeans, for instance? In other words: are the classifications "Finno-Ugrian" and "Indo-European" based on a strictly and purely linguistic distinction? In more impersonal terms: in the spread of peoples' genetic variants, is linguistic or geographical origin of primary importance? And is the answer to these questions the same when viewed comparatively through our maternal and paternal lines of descent?

Most of the languages spoken in Europe belong to the Indo-European language family, apart from exceptions such as the Basques and Finno-Ugrians, of which the Finns, the Estonians and the Hungarians have independent national states. At the same time, the Finno-Ugric languages are a sub-group of the Uralic language family. In speaking of the Finno-Ugrians I also refer to those Finno-Ugric peoples lacking independent national states but about whom there exist DNA data or fragments thereof. These are the Karelians and the Samis and a series of Finno-Ugric groups from Eastern Europe and the Volga basin: Maris, Mokshas, Udmurtians, and to a lesser extent Komis. Fragmentary data, mainly characterising only paternal lines, are available for several Samoyedic peoples of the Uralic language family. The same applies to the Hungarians, for whom only half of the DNA genetics have been investigated.

On the basis of the principle of "global reach," which in the pre-DNA era was applied to genetic features in their entirety (blood groups, etc.), in the past decade it has become clear that the worldwide distribution of variants of the gender-specific parts of our genome – the Y-chromosome and mitochondrial DNA – is far from homogeneous one-dimensionality. Thus the general rule that states that roughly 15 percent of genetic variation can be explained by continental differences (i.e. those comprising large-scale geographical divisions) does not apply in their case. Or rather that rule should be modified, although further explanation would require delving into details of population genetics that are superfluous in this context. Concerning humankind's maternal lines, and presumably as a result of Palaeolithic isolation, the following global regions of "primeval mothers" can be distinguished: a) sub-Saharan Africa; b) western Eurasia and North Africa; c) India; d) East Asia as well as the indigenous populations of Siberia, Polynesia and America. It should be noted that all of the Eurasian, American and Australian maternal lines that have so far been investigated are described in

PDF created with pdfFactory trial version www.pdffactory.com

the genetic tree as one branch of the sub-Saharan Africa variations, once again confirming the hypothesis of the (recent) African origin of modern man.

On this scale, the maternal lines of Estonians and other Finno-Ugric peoples may be identified as typically western European, with a minimal mongoloid component. Differences in detail within Europe can be explained as originating from genetic drift and other stochastic demographic processes. We can paint a complete picture from the Iberian Peninsula to the Volga River and the Ural Mountains. Furthermore, we can scrutinise the details and distinguish numerous maternal lines which are found in Cornwall, Crete and among the Finno-Ugrians of the Volga, and about which we can confidently assume that these are generations that originate from one progenitress. On the other hand, one can also find unique lines that, for instance, unite Albanians, Estonians and Syrians or Lapps and the French. In other words, one can see a spider's web of common progenitresses that unites peoples regardless of their membership in a linguistic group, and it is only against the background of this predominant homogeneity that individual emphases appear, for instance the greater or lesser frequency of certain maternal lines in one region or another, because the maternal lines of Estonians and Finns do indeed coincide in those features that are not necessarily characteristic of the inhabitants of Crete.

I would not refer to this pan-European similarity as a crisis of identity, but as a fundamental characteristic of the identity of Finno-Ugric mothers, yet in general terms, the reality of the relatedness of the maternal lines of Indo-Europeans, Hamito-Semitic and Finno-Ugric maternal lines. Basque maternal lines also belong to the same group.

And yet, modern genetics is nevertheless able to please those who yearn for a Finno-Ugric genetic identity (or rather, distinctiveness). Indeed, by contrasting with other groups, it became clear that most Finno-Ugrians share a massive gene variant that covers roughly half of all paternal lines; a unique mutation that is lacking in Africa and the western part of Europe, as well as in India, China and Japan. The border of the western spread of this mutation is so sharp that only one-twelfth of Swedes possess it, and it is almost absent among Germans, Dutch, French, English, Spanish, Italians, Greeks, Poles, Byelorussians, Czechs, Slovaks, Slovenians and Croatians, hence from most of the Indo-Europeans of Europe. Of the Slavs, the only exceptions are the Russians and Ukrainians of Eastern Europe, among whom the above-mentioned paternal line variant makes up about one-seventh of all Y chromosomes, quite clearly reflecting the so-called Finno-Ugric substrate in the eastern Slavs' gene pool – a phenomenon with an analogy well known to

PDF created with pdfFactory trial version www.pdffactory.com

researchers of toponyms and hydronyms. Moskva [Moscow] is, after all, a typical Finno-Ugric toponym.

This paternal line is, however, very common among the peoples of Siberia, even among the indigenous peoples of Kamchatka, the Koryak and Itelmen, and the Chukchis, who live adjacent to Alaska, and thus also among those Siberian peoples who speak Altai-Tungus tongues, and languages belonging to unclassified language families. This leads us back to the age-old problem of the ancient home of the Finno-Ugrians, which is almost as old and indistinct as the question of the ancient home of the Indo-Europeans. The classical view that the latter lies in Siberia has persisted for over a century. And indeed, among the Yakut the above-mentioned paternal line is more numerous than among even the Estonians or Finns. It now appears, however, that this circum-arctic paternal line more likely migrated from Eastern Europe to Siberia than *vice versa*. In determining the original home of gene lines, internal divergence is of greater importance than frequency of occurrence. This parameter is, however, considerably higher among the Finno-Ugrians of Europe. This leads to a very important generalisation: the extensive overlapping of the paternal lines of Siberian Ugrians and Finno-Ugrians, to which there is no parallel among maternal lines, is proof of an extensive eastward flow of genes caused predominantly by men, presumably in the late Upper Palaeolithic period. The Dnieper-Don refuge during the Ice Age, more precisely the process of resettlement following the retreat of the glaciers, would have been a natural source of this flow. The said process was presumably connected with the movement of the main animals of prey (during the time of the refuge mammoths and woolly rhinoceroses) northwards and to Siberia. In any case, in contrast to our pan-European (Western Eurasia-like) maternal lines, Estonians, Finns and most other Northern European Finno-Ugrians have extensive connections with the peoples of Siberia through their paternal lines, yet not with the main Mongol tribes, which lack the corresponding paternal line.

A few interesting details remain to be mentioned, firstly – Hungarians almost completely lack the paternal line that is characteristic of the other European Finno-Ugrians. Secondly, this same line, although absent in western and southern Slavic populations, is well represented among both Latvians and Lithuanians. Here we can see a clear contrast between linguistic and geographical influence: the two above-mentioned Baltic peoples are linguistically Indo-Europeans, and in the tree of languages, their languages are sisters to the Slavic languages. The question of whether this offers

PDF created with pdfFactory trial version www.pdffactory.com

evidence of the ancient linguistic exchange between proto-Latvian and proto-Lithuanian, perhaps in connection with the transition to the Neolithic lifestyle, as hypothesised by Finnish linguist Kalevi Wiik, remains an unresolved question. It is, however, already possible for geneticists to argue that the genetic identity borne by the Finno-Ugric maternal and paternal lines also includes the Latvians and Lithuanians.

Who then are the Estonians, in genetic terms, and where did they come from? Based on our present knowledge, we have no reason to devise complex patterns; the great majority of the Estonians' genes presumably originate from post-Ice Age Europeans, those who arrived about 40 thousand years ago and survived the 8000-year Ice Age in refuges Furthermore, there are no influential arguments that would lead one to presume the people who lived here before the Ice Age (and there is also no reason to believe that they did not live here) were not predominantly of the same genetic makeup as those who established the refuges 24–25 centuries ago, as the ice boundary moved southward and moved with it, i.e. northward, as the ice retreated. This of course concerns only our genetic history, and does not purport to explain the details of ethnogenesis.

I would like to conclude this article with a quotation from Steven Shapin's review of the above-mentioned book by Paul Rabinow:

> … DNA is a Post-Modernist molecule, since fragments of our contemporary expert culture insist that the reflexive condition for believing these things about DNA, or indeed disbelieving them, is ultimately ascribable to the workings of DNA itself, while the knowledge of these workings is an authentic item of our culture.

PDF created with pdfFactory trial version www.pdffactory.com