Genetic Affinities of Ukrainians From the Maternal Perspective

Andrey Pshenichnov,[1]* Oleg Balanovsky,[1,2] Olga Utevska,[1,3] Ene Metspalu,[4] Valery Zaporozhchenko,[1] Anastasia Agdzhoyan,[2] Mikhail Churnosov,[5] Lyubov Atramentova,[3] and Elena Balanovska[1]

[1]*Research Centre for Medical Genetics, RAMS, Moscow, Russia, 115478*
[2]*Vavilov Institute of General Genetics, RAS, Moscow, Russia, 119991*
[3]*Karazin Kharkiv National University, Kharkov, Ukraine, 61022*
[4]*Estonian Biocentre and Department of Evolutionary Biology, University of Tartu, Tartu, Estonia, 51010*
[5]*Belgorod State University, Belgorod, Russia, 308015*

KEY WORDS ethnogenetics; mtDNA; mitochondrial DNA; gene geography; haplotype co-occurrence analysis

ABSTRACT The area of what is now the Ukraine has been the arena of large-scale demographic processes that may have left their traces in the contemporary gene pool of Ukrainians. In this study, we present new mitochondrial DNA data for 607 Ukrainians (hypervariable segment I sequences and coding region polymorphisms). To study the maternal affinities of Ukrainians at the level of separate mitochondrial haplotypes, we apply an original technique, the *haplotype co-occurrence analysis*. About 20% of the Ukrainian maternal gene pool is represented by lineages highly specific to Ukrainians, but is scarcely found in other populations. About 9% of Ukrainian mtDNA lineages are typical for peoples of the Volga region. We also identified minor gene pool strata (1.6–3.3%), each of which is common in Lithuanians, Estonians, Saami, Nenets, Cornish, and the populations of the North Caucasus. Am J Phys Anthropol 000:000–000, 2013.

The area north of the Black Sea has been subject to various broad-scale demographic processes. Forty-five thousand years ago, when modern humans reached Europe, this area was one of the entering routes from interior western Asia (Torroni et al., 2006; Pala et al., 2009). After the last glacial maximum about 11.5 thousand years ago, Europe was repopulated from a range of southern refugia, among which was the East European Plain; the successive postglacial expansion from this area was marked by the dispersal of U4 and U5a mitochondrial lineages (Semino et al., 2000; Torroni et al., 2006; Malyarchuk et al., 2010; Soares et al., 2010; Pala et al., 2009, 2012). Since the Neolithic, until the Slavic expansion, and later on, this area had seen multiple migration waves, which resulted in admixture and replacement of numerous populations. These complex demographic processes led to the emerging pattern of genetic relatedness of Ukrainians—the modern inhabitants of this area—to other European populations.

The diversity of paternally inherited Y-chromosomal markers shows that Ukrainians are genetically close to other Eastern Slavs, Western Slavs, and Slovenes and western Croats (the westernmost Southern Slavic populations) (Rebala et al., 2007). In this study, we address the diversity of maternally inherited mitochondrial DNA (mtDNA) markers in Ukrainians. We present new mtDNA data for 607 Ukrainian individuals originating from the main historical subdivisions of the Ukrainian people. By using a set of published data on mtDNA variability in Europe and its surroundings (486 populations, 41,839 samples), we analyze the genetic affinities of Ukrainians with other Europeans. We introduce the original haplotype co-occurrence technique, which allows us to trace genetic relations of separate strata of the gene pool under study. In the analysis, we supplement our new data with the mtDNA sample of 240 Ukrainians of predominantly Eastern origin from Gusar (2006) and 159 Western Ukrainians from Mielnik-Sikorska et al. (2013), which gives us a total sample of 1,006 Ukrainian mtDNA samples.

METHODS

DNA collecting and genotyping

DNA samples were collected in 2002 in collaboration with Kharkov National University (Kharkov, Ukraine) and Belgorod State University (Belgorod, Russia). The sampled populations represent the main historical and ethnic subdivisions, i.e., Eastern, Dnieper, Podol, and Western Ukrainians (Dyachenko, 1965; Ponomaryov,

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Russian Foundation for Basic Research; Grant numbers: 13-04-90420; 13-04-01711; 12-06-12002; Grant sponsor: the Presidium of RAS programs "Molecular and cell biology" and "Gene pools dynamics"; Grant sponsor: EU European Regional Development Fund through the Centre of Excellence in Genomics to Estonian Biocentre and the University of Tartu and by the Estonian Basic Research (to E.M.); Grant number: SF 0270177Bs08.

*Correspondence to: Andrey Pshenichnov; 115478, Laboratory of Human Population Genetics, Research Centre for Medical Genetics, Moskvorechie-1, Moscow, Russia. E-mail: pxemon@gmail.com

Received 22 March 2013; accepted 23 August 2013

DOI: 10.1002/ajpa.22371
Published online 00 Month 2013 in Wiley Online Library (wileyonlinelibrary.com).

2000; Deryabin, 2002). Samples were taken from individuals of Ukrainian descent unrelated for at least three generations. A total of 607 samples were collected. Among these, 98% originate maternally (in the third generation) from one of four locations: Western (Lvov and Ivano-Frankovsk regions, $N = 142$), Podol (Khmelnitsky region, $N = 180$), Dnieper (Cherkassy region, $N = 174$), and Eastern (Belgorod region, $N = 96$). The remaining 15 individuals trace their maternal ancestry to other parts of the Ukraine. In the following analysis, we attribute the samples not to their sampling location, but to the places of their maternal origin in the third generation.

The DNA was extracted from whole blood according to the procedures of Powell and Gannon (2002) at the Institute of Immunology of Federal Medical-Biological Agency (Moscow). Genotyping was performed at the Estonian Biocentre (Tartu, Estonia). For each sample, the mitochondrial hypervariable segment I region was sequenced between the 16024 and 16400 positions in both directions[1]. Haplogroups were confirmed by checking single-nucleotide polymorphisms (SNPs) in the mtDNA coding region or by hypervariable segment II sequencing (van Oven and Kayser, 2009). The details of SNP and hypervariable segment II analysis and haplogroup attribution are given in Supporting Information Appendix 1.

Haplotype co-occurrence analysis

A close phylogenetic relatedness of two lineages implies their common origin and shared history. But even far-related lineages of different phylogenetic branches may acquire common history. This happens if they occasionally occur in the same population before its demographic expansion. In this case, the haplotypes that co-occur in the ancestral population, although unrelated, will co-occur in its derivates. The co-occurrence of haplotypes in a set of current-day populations, regardless of these haplotypes' relatedness, may therefore represent a signal of some ancient population expansion. The search for such co-occurrence traces may complement the common ancestry-based phylogenetic analysis in studying ancient demographics. Our purpose here is to find such traces among the haplotypes of Europe and its surroundings.

The approach is based on dividing the set of haplotypes under study (in our case, the haplotypes present in the European region) into separate geographic co-occurrence clusters, each embracing mtDNA haplotypes with similar spatial distribution. To do this, we used a data set on mtDNA variability in populations of Europe and its surroundings (see the *Reference Data Sets* section below). As long as each resulting co-occurrence cluster consists of haplotypes, its frequency can be determined for any population as the sum of frequencies of its comprising haplotypes. We can then visualize each co-occurrence cluster's frequencies on a geographic map and study the cluster's spatial distribution. The areas of co-occurrence clusters' high frequency may mark the population process that took place in the past, and the cluster's share in the Ukrainian sample corresponds to the share of gene stratum touched by this process.

To find the co-occurrence clusters, we took the following steps:

1. For all haplotypes found in the European region, we calculated the inter-haplotype D_{NEI} distance (Nei, 1975) between each pair of haplotypes, a measure that reflects dissimilarity of spatial distributions of the two lineages (see Supporting Information Appendix 2). This analysis was performed only for the non-unique haplotypes, i.e., the haplotypes found in at least two different populations. As a result, we obtained an inter-haplotype D_{NEI} distance matrix.
2. We grouped the studied haplotypes by the use of clustering analysis, on the basis of the inter-haplotype D_{NEI} distance matrix. We used the complete linkage clustering algorithm, which, compared with other clustering algorithms, provides maximal similarity of objects within each cluster (that is, haplotypes in each resulting cluster are closer to each other than to any haplotype outside the cluster).
3. Clustering is an iterative process, as each step (a clustering event) unites the two closest existing clusters. With each step, the distance between the next two united clusters grows. By plotting the D_{NEI} distances of clustering to the number of clustering events that occur at each D_{NEI} value (Fig. 2), we determined the threshold D_{NEI} value, below which the largest share of clustering events has occurred, and above which the clustering decreases. The clusters formed below the threshold were taken into the next step as candidates to form co-occurrence clusters.
4. For further analysis, we selected from the cluster candidates formed in the previous step, those with frequencies exceeding 1.5% in Ukrainians and consisting of three or more mtDNA haplotypes.
5. Each cluster's frequency was plotted onto a geographical map for visual analysis.

The presence of a co-occurrence cluster in a population's gene pool does not imply this population's direct genetic link through recent gene flow with other populations that have a high frequency of this cluster. The co-occurrence analysis does not distinguish between recent gene flow and deep common ancestry, and its primary aim is to detect similarity.

Reference data sets

In the mitochondrial genomics era, the potential of phylogenetic analysis of complete mtDNAs has rendered older data, like HVS sequences with coding region haplogroup-diagnostic SNPs, obsolete. Yet, only about 15% of complete sequence data are coupled with lineages' population frequencies (Zaporozhchenko et al.). Therefore, for the haplotype co-occurrence analysis, which requires a comprehensive reference data set of population gene frequencies, we must reject using the complete mtDNA sequence data and stay with data on hypervariable segment I and SNPs. Hopefully, in a few years, the availability of new mtDNA data coupled with population frequencies will be substantial enough to test the haplotype co-occurrence technique on the data of a high phylogenetic resolution.

The other issue with reference data sets is the availability of population frequencies of sub-branches of the

[1]Nucleotides are numbered according to Behar et al. 2012. For most samples, sequencing between the 16024 and 16400 positions yielded a double read of each sample. However, in samples with 16189TC transition, both the forward and the reverse read usually ended within the poly-C track in the vicinity of this position. As a result, for most of the 16189TC samples, the total coverage was one read per nucleotide.

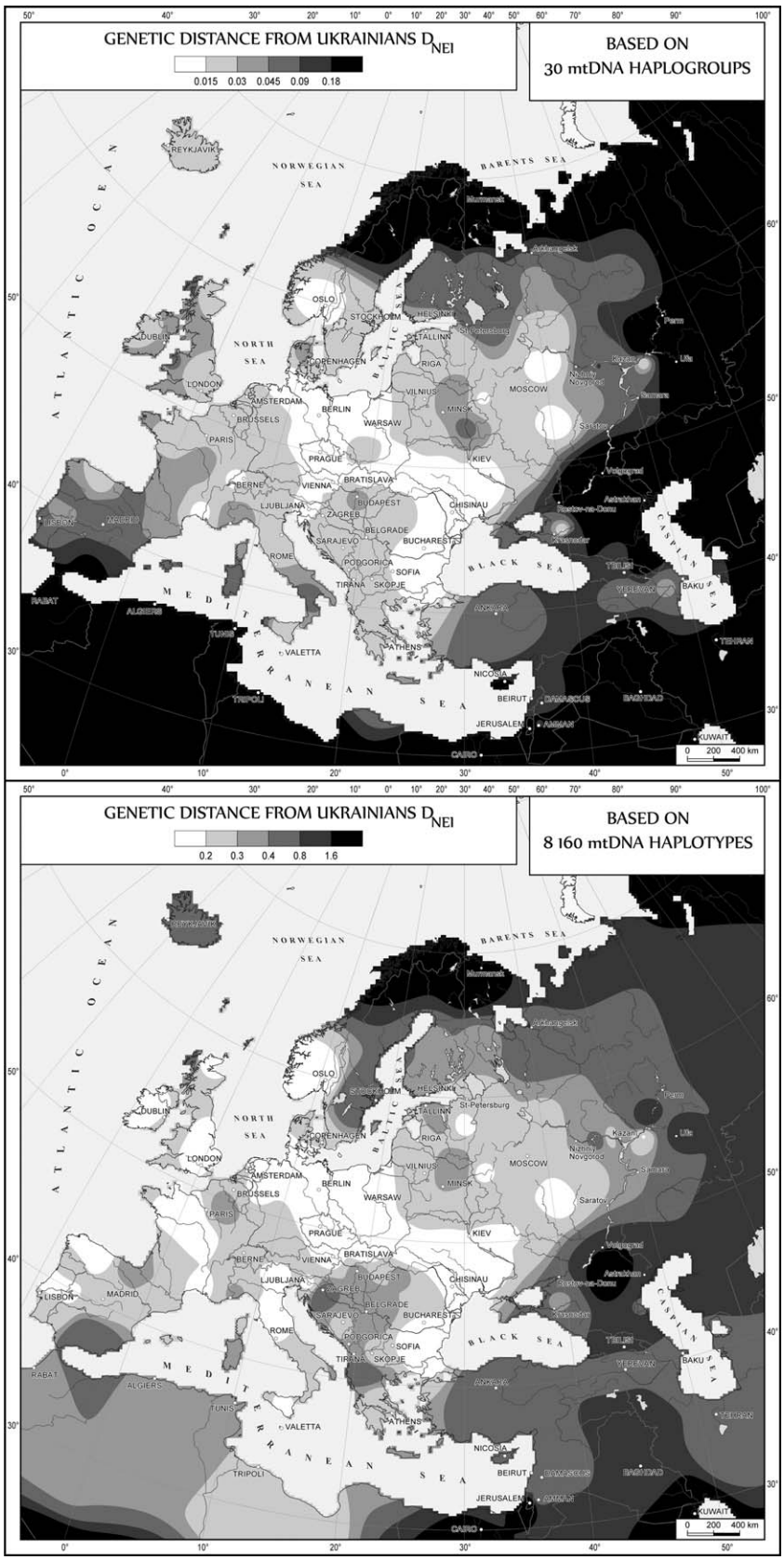

Fig. 1. D_{NEI} genetic distances from Ukrainians in Western Eurasia. **(A)** Based on mtDNA frequencies of 30 mtDNA haplogroups. (**B**) Based on frequencies of 8 160 mtDNA haplotypes.

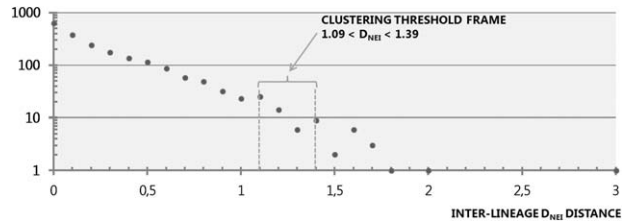

Fig. 2. The distribution of interlineage D_{NEI} distances in the clustering analysis. Selection of any D_{NEI} within the range $1.09 < D_{NEI} < 1.39$ and application of the cluster selection criteria (see the *Methods* section) yields a set of 13 co-occurrence clusters. $D_{NEI} = 1.09$ threshold corresponds to 1,948, and $D_{NEI} = 1.39$ corresponds to 1,989, of the total 2,097 clustering events between 2,098 non-unique haplotypes (92.9 and 94.8%, respectively). The 41 clustering events that occur between these two thresholds neither generate any clusters that would pass the criteria nor modify any of the 13 clusters that pass these criteria. The last event before the 1.09 threshold is the formation of the cluster IX; the next event after the 1.39 threshold is the merging of clusters I and XIII.

H haplogroup, which comprises a large share of the European mitochondrial gene pool. For the haplogroups H1 and H2, population frequency data are available for just a dozen European peoples, which is insufficient for haplotype co-occurrence analysis. This forced us to discard haplogroup H subdivisions in our analysis.

Given these margins, we searched through the available mtDNA data (literature and our unpublished sources) and picked the populations with the following analyzed markers: (1) hypervariable segment I sequence within the range 16,069–16,365 (filtering for a broader read range would decrease the set of the available populations); and (2) haplogroup-diagnostic SNPs studied sufficiently enough to attribute each sample—by combining the SNP and hypervariable segment I data—to one of the following 30 haplogroups: A; B; C; D; G; H; HV; I; J; K; L(xM,N); M(xC,D,G,Z); N(xI,N1b,N9a,R,W); N1b; N9a; R(xHV,J,T,U); T; U(xK,U1,U2e,U3,U4,U5a,U5b,U7,U8a); U1; U2e; U3; U4; U5a; U5b; U7; U8a; V; W; X; and Z (haplogroup attribution according to Van Oven and Kayser, 2009). This gave us a data set of 41,839 samples from 486 populations of Europe and its surroundings. For details and references, see Supporting Information Appendix 3. The attribution of samples to haplogroups was performed automatically by MURKA software (Zaporozhchenko et al.).

In the haplotype co-occurrence analysis, we considered two samples identical by their haplotype if they both belonged to the same one of the 30 mtDNA haplogroups listed above and if their hypervariable segment I sequence coincided within the range of 16,069–16,365.

In the data set, some ethnic groups (like Italians, Spanish, Germans, and some others) were represented by a great number of populations, whereas others were underrepresented (like Adygs, Latvians, and some other peoples). To smooth out this discrepancy, which could affect the co-occurrence analysis, we merged the 486 source populations, which yielded 94 large populations of the *co-occurrence analysis data set*. When performing the last step of the co-occurrence analysis, the plotting of co-occurrence clusters to a geographic map, we grouped the same 486 source populations into 129 populations of the *mapping data set* (see Supporting Information Appendix 3). This allowed us to represent well-studied peoples by more than two populations on the map, as long as the mapping analysis was not as sensitive to sample discrepancy as the co-occurrence analysis.

D_{NEI} genetic distances estimation

We estimated genetic distances D_{NEI} (Nei, 1975) between the average Ukrainian sample and the 129 populations of the mapping data set (1) based on frequencies of 30 mtDNA haplogroups found in 129 populations and (2) based on frequencies of 8,160 mtDNA haplotypes found in 129 populations.

The frequency of each haplogroup or haplotype in the average Ukrainian sample was calculated as the average unweighted frequency of Western ($N = 301$; Mielnik-Sikorska, 2013; our data), Podol ($N = 180$; our data), Dnieper ($N = 174$; our data), and Eastern ($N = 96$; our data) locations. We used the unweighted average, because, if weighted, the overrepresentation of the Western Ukrainian sample would bias the average frequencies. We also had to discard the data from Gusar 2006 from calculating average frequencies because of its ambiguous geographic affiliation.

The resulting D_{NEI} values for the 129 mapping populations were plotted onto the geographic map of European regions. Among these, we also plotted the four local Ukrainian samples to form the area of the lowest genetic distance from the average Ukrainian sample.

Software

D_{NEI} was calculated in Phylip v3.69 (Felsenstein, 2005). Maps were built with original GeneGeo software as described in Balanovsky et al. (2011). Clustering was performed in Microsoft Excel 2007 using VBA for applications and confirmed in Statistica 6.0.

RESULTS AND DISCUSSION

Ukrainian mtDNA haplotypes

The mtDNA haplotypes of the sample studied ($N = 607$) are presented in Supporting Information Appendix 1 and are available in GenBank (accession numbers JX895515–JX896122). For further analysis, we united our new 607 samples, the 240 samples from Gusar (2006), and 159 samples from Mielnik-Sikorska et al. (2013), which produced the total sample of $N = 1,006$ Ukrainian mtDNAs. To make this set compatible with our 129-population mapping reference dataset, in the further analysis, we reduced the read of hypervariable segment I to the 16,069–16,365 frame and assigned Ukrainian samples to one of 30 haplogroups of the reference data set. This gave us 416 Ukrainian mtDNA haplotypes belonging to 28 haplogroups. Of these haplotypes, 130 are not found in any other ethnic groups (93 non-Ukrainian populations of the co-occurrence analysis data set) and hereafter are referred to as *unique Ukrainian haplotypes*. Haplotypes of 14.5% of 1,006 Ukrainians belong to one of the uniques, whereas 286 haplotypes are shared with other ethnic groups.

Overall maternal affinities of Ukrainians

By estimating D_{NEI} genetic distance between Ukrainians and other ethnic groups of Europe and its surroundings, we aimed to study the maternal genetic relatedness of these populations with Ukrainians. The maps on Figure 1A,B represent genetic distances from Ukrainians based on two levels of genetic resolution,

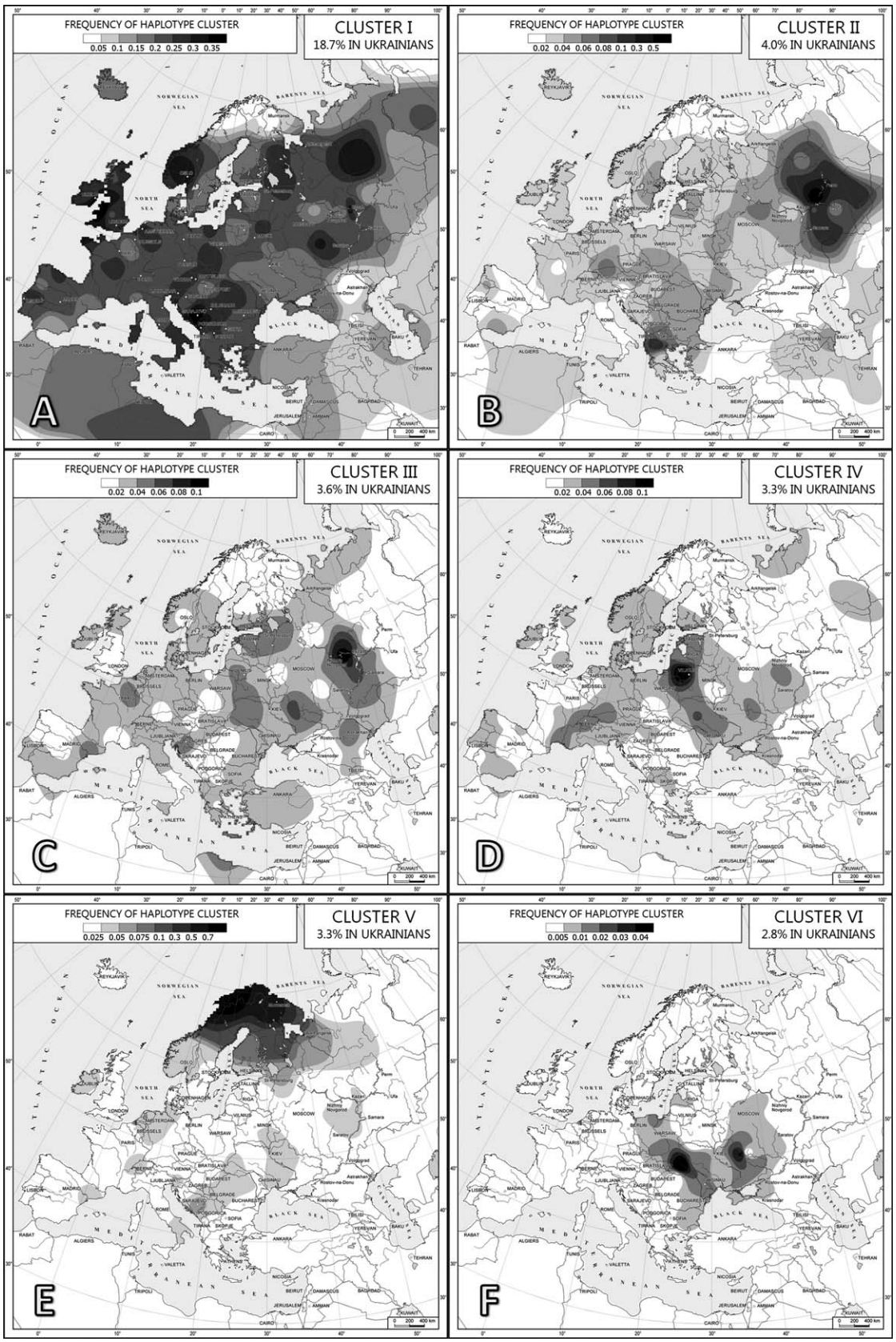

Fig. 3. Frequencies of haplotype geographic co-occurrence clusters in Europe and its surroundings. See Supporting Information Appendix 4 for details of cluster composition and spatial distribution.

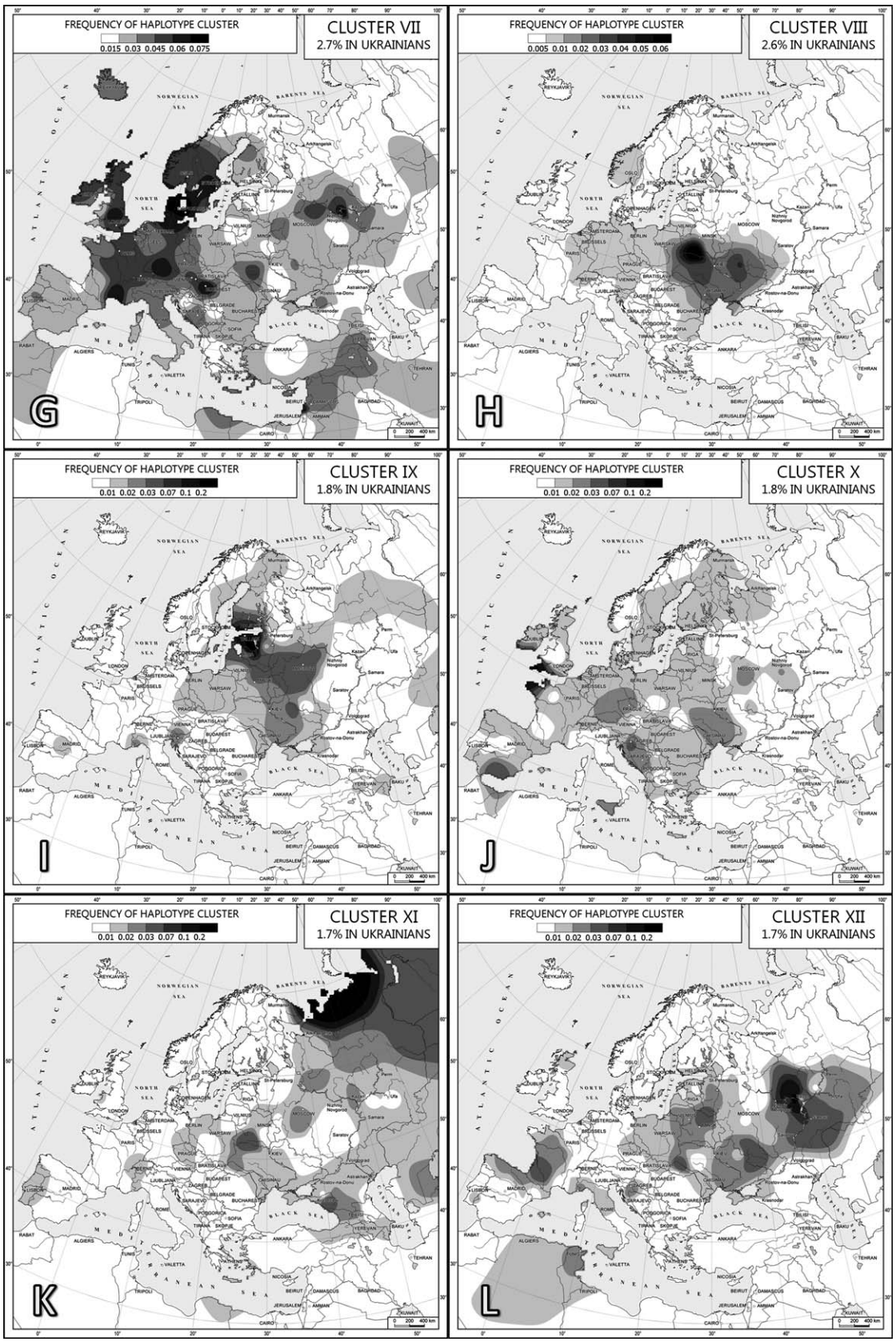

Fig. 3. (Continued)

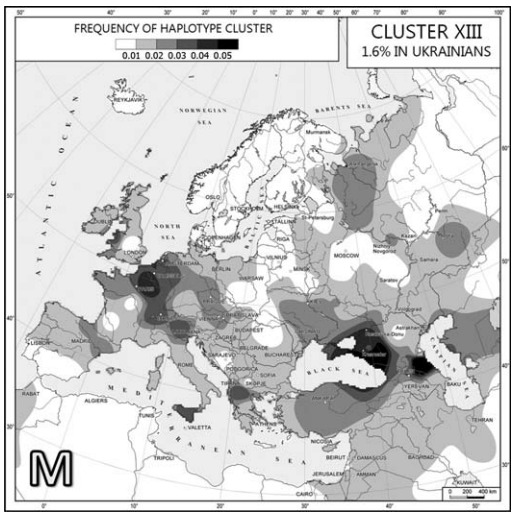

Fig. 3. (Continued)

30 haplogroup and 8,160 haplotype frequencies. On both levels of resolution, the maximal genetic similarity corresponds to a similar area: the Ukraine; Poland; the western coast of the Black Sea, including Bulgaria, Central Russian populations, and Norway. At the haplotype level (Fig. 1B), some other Western European populations are genetically close to Ukrainians, too.

Haplotype-based relatedness estimates (Fig. 1B) are more sensitive to exact matches of separate haplotypes than haplogroup-based estimates. An exact haplotype match brings up the haplotypic similarity estimate, but does not noticeably increase the haplogroup-based estimate. Separate lineages, which coincide with Ukrainian lineages in Western European populations, are detectable at the level of haplotypes, but do not make an input substantial enough to show such a similarity at the level of the 30 studied haplogroups. This makes mtDNA haplotype-based estimates more sensitive to recent gene flow.

To distinguish populations that show maternal genetic affinities of separate strata of the Ukrainian gene pool, we undertook the haplotype co-occurrence analysis.

Components of maternal gene pool

We performed the haplotype co-occurrence analysis (see the *Methods* section) for the mtDNA haplotypes found in 94 populations of the co-occurrence analysis dataset. Of these 8,331 haplotypes, we performed this estimation for the 2,098 non-unique haplotypes (ones that occur in at least two populations). At the haplotype clustering stage, selecting any clustering threshold within the range $1.09 < D_{NEI} < 1.39$ (Fig. 2) and applying the cluster selection criteria (exceeding 1.5% in Ukrainians and consisting of three or more haplotypes) gives a similar set of 13 co-occurrence clusters. These 13 clusters were plotted onto a geographical map (Fig. 3A–M). Their description, including composition, is given in Table S4 (Supporting Information Appendix 4).

Almost one-fifth of the Ukrainian maternal gene pool is formed by cluster I (Fig. 3A), which is spread all over Europe, with lower frequencies in Central Europe and the highest frequencies at the margins of the region. Haplotypes most frequent in Ukrainians—the Cambridge Reference Sequence (see Andrews et al., 1999) H_16,129-16,187-16,189-16,223-16,230-16,278-16,311 (9.5%) and the lineage J_16,069-16,126-16,129-16,187-16,189-16,223-16,230-16,278-16,311 (3.4%)—build up more than one-half of this cluster in Ukrainians. The three other H haplogroup haplotypes differ from the Cambridge Reference Sequence in just one mutation (16,189, 16,304, or 16,311). As long as such lineages emerged more than once in the haplogroup H phylogeny, each of them is a paraphyletic composite. To separate the major cluster I, a more comprehensive dataset on H sub-branch frequencies in European local populations is needed. The latest fraction to join the cluster (at $D_{NEI} = 0.75$) is the U4 and the U5a lineages (whereas the rest of the cluster is formed at $D_{NEI} = 0.48$). It is hard to predict the position of these U4 and U5a lineages in clustering if the H haplogroup is dissected.

Clusters VI and VIII (Fig. 3F,H; 5.4% total in Ukrainians) are most spread in Ukrainians, and in case of cluster VIII, southern Byelorussians. Together with the 14.5% unique Ukrainian haplotypes (ones which are not present in any non-Ukrainian populations of the reference data set), this shows that 19.9% of the Ukrainian gene pool belongs to haplotypes highly specific to Ukrainians. Among these 148 specific lineages, only 44 belong to haplogroup H; 26 lineages belong to J or T haplogroups; 19 belong to U4 or U5a; and 23 belong to other branches of U. Ten lineages belong to various branches of HV(xH,V). Six lineages belong to haplogroups of eastern origin (A, B, C, D, and G); among them, the C lineage belongs to C1c5 sub-branch, typical for Eastern Europe (Mielnik-Sikorska et al., 2013). The rest belong to haplogroups I, N1b, W, X, or V.

Clusters II, III, and XII (Fig. 3B,C,L; 9.2% total in Ukrainians) have their highest frequencies in Finnic and Turkic speakers of Volga (Udmurts, Mari, Komi, Chuvash, Tatars, and Bashkirs). Yet, cluster III, just like cluster I, includes H lineages that appeared more than once in the haplogroup H phylogeny, and applying the H dissection would seemingly result in reformation of the III cluster. Among the rest, clusters II and XII include D4b1 and U4a1 lineages typical for the Volga region.

Cluster VII (Fig. 3G) displays an irregular pattern, with high frequencies in various parts of Europe. The remaining clusters capture the relation of 1.6–3.3% of Ukrainian samples with one of the following populations: cluster IV is most frequent in Lithuanians (Fig. 3D); cluster V includes U5b1b, V, D5a3, and a few HV0(xV) lineages typical for Saami (Fig. 3E); and clusters IX, X, XI, and XIII include lineages most frequent in Estonians, Cornish, Nenets, and peoples of Northern Caucasus, respectively (Fig. 3I,J,K,M).

CONCLUSIONS

Although the haplotype co-occurrence analysis was performed with low-resolution data, it allowed us to detect genetic similarities between the Ukrainians and various European populations:

1. Twenty percent of the samples studied belong to lineages highly specific to Ukrainians.
2. Nine percent of the samples belong to lineages typical for the Volga region.
3. About 13% of the sample shows genetic affinities with one of the following populations: Lithuanians, Saami, Estonians, Nenets, Cornish, and peoples of the North Caucasus.

The estimates of genetic similarity between Ukrainians and other populations do not count the direct input of these populations' gene pools into the Ukrainian gene pool. The haplotype co-occurrence clusters may mark the longstanding common ancestry of Ukrainians and corresponding groups: modern populations that show mutual genetic affinity may display the patterns of genetic similarity of their predecessors who inhabited corresponding territories. Methods for reliable dating of geographic mtDNA haplotype clusters are yet to be developed. They may be based on lineage diversity accumulated within clusters, but this will work only for clusters consisting of many lineages, at least parts of which are closely related.

About 20% of the Ukrainian sample belong to the major haplotype cluster I, comprising paraphyletic haplogroup H lineages. Its interpretation is impossible without haplogroup H dissection. In future studies, it would be promising to adapt an mtDNA genome sequence-based haplotype frequency database for studying haplotype co-occurrence in Europe.

ACKNOWLEDGMENTS

We thank Richard Villems (Estonian Biocentre and Department of Evolutionary Biology, University of Tartu, Tartu, Estonia) for valuable comments on this manuscript. We thank Irina Gus'kova (Institute of Immunology of Federal Medical-Biological Agency, Moscow, Russia) for help with DNA extraction. Mitochondrial DNA data in GenBank: accession numbers JX895515–JX896122.

LITERATURE CITED

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 23:147.

Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S, Pocheshkhova E, Haber M, Platt D, Schurr T, Haak W, Kuznetsova M, Radzhabov M, Balaganskaya O, Romanov A, Zakharova T, Soria Hernanz DF, Zalloua P, Koshel S, Ruhlen M, Renfrew C, Wells RS, Tyler-Smith C, Balanovska E; Genographic Consortium. 2011. Parallel evolution of genes and languages in the Caucasus region. Mol Biol Evol 28:2905–2920.

Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli EL, Silva NM, Kivisild T, Torroni A, Villems R. 2012. A "Copernican" reassessment of the human mitochondrial DNA tree from its root. Am J Hum Genet 90:675–684.

Deryabin VE. 2002. Ethnic anthropology of modern Slavic peoples of Eastern Europe. Multidimensional quantitative research. Moscow. Manuscript deposited in Russian Institute of Scientific and Technical Information. (In Russian).

Dyachenko VD. 1965. Anthropological composition of Ukrainian people. Comparative research of peoples of Ukrainian SSR and neighboring territories. Kyiv: Naukova Dumka. (In Ukrainian)

Felsenstein J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Gusar VA. 2006. Diversity and polymorphism of mitochondrial DNA of Ukrainian population. J Acad Med Sci 12:739–749. (In Ukrainian)

Malyarchuk B, Derenko M, Grzybowski T, Perkova M, Rogalla U, Vanecek T, Tsybovsky I. 2010. The peopling of Europe from the mitochondrial haplogroup U5 perspective. PLoS One 21:e10285.

Mielnik-Sikorska M, Daca P, Malyarchuk B, Derenko M, Skonieczna K, Perkova M, Dobosz T, Grzybowski T. 2013. The history of Slavs inferred from complete mitochondrial genome sequences. PLoS One 8:e54360.

Nei M. 1975. Molecular population genetics and evolution. Amsterdam: North-Holland Pub. Co.

Pala M, Achilli A, Olivieri A, Hooshiar Kashani B, Perego UA, Sanna D, Metspalu E, Tambets K, Tamm E, Accetturo M, Carossa V, Lancioni H, Panara F, Zimmermann B, Huber G, Al-Zahery N, Brisighelli F, Woodward SR, Francalacci P, Parson W, Salas A, Behar DM, Villems R, Semino O, Bandelt HJ, Torroni A. 2009. Mitochondrial haplogroup U5b3: a distant echo of the epipaleolithic in Italy and the legacy of the early Sardinians. Am J Hum Genet 84:814–821.

Pala M, Olivieri A, Achilli A, Accetturo M, Metspalu E, Reidla M, Tamm E, Karmin M, Reisberg T, Hooshiar Kashani B, Perego UA, Carossa V, Gandini F, Pereira JB, Soares P, Angerhofer N, Rychkov S, Al-Zahery N, Carelli V, Sanati MH, Houshmand M, Hatina J, Macaulay V, Pereira L, Woodward SR, Davies W, Gamble C, Baird D, Semino O, Villems R, Torroni A, Richards MB. 2012. Mitochondrial DNA signals of late glacial recolonization of Europe from Near Eastern refugia. Am J Hum Genet 90:915–924.

Ponomaryov AP. 2000. Historical-ethnographic zoning. In: Polischuk NS, Ponomaryov AP, editors. Ukrainians. Moscow: Nauka. p 27–43. (In Russian)

Powell R, Gannon F. 2002. Purification of DNA by phenol extraction and ethanol precipitation. Oxford: Oxford University Press.

Rebała K, Mikulich A, Tsybovsky I, Siváková D, Dzupinková Z, Szczerkowska-Dobosz A, Szczerkowska Z. 2007. Y-STR variation among Slavs: evidence for the Slavic homeland in the middle Dnieper basin. J Hum Genet 52:406–414.

Soares P, Achilli A, Semino O, Davies W, Macaulay V, Bandelt HJ, Torroni A, Richards MB. 2010. The archaeogenetics of Europe. Curr Biol 20:174–183.

Semino O, Passarino G, Quintana-Murci L, Liu A, Béres J, Czeizel A, Santachiara-Benerecetti AS. 2000. MtDNA and Y chromosome polymorphisms in Hungary: inferences from the Palaeolithic, Neolithic, and Uralic influences on the modern Hungarian gene pool. Eur J Hum Genet 8:339–346.

Torroni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. 2006. Harvesting the fruit of the human mtDNA tree. Trends Genet 22:339–345.

van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30:386–394.

Zaporozhchenko, Balanovsky, Pshenichnov, Balanovska. MURKA–global mitochondrial database and integrated software.