

Overview of the Genetic STR Clustering among Worldwide Human Populations

Serkan Dogan¹, Adna Asic¹, Imer Muhovic¹, Larisa Besic¹ and Damir Marjanovic^{1,2}

¹*International Burch University, Department of Genetics and Bioengineering, Sarajevo, Bosnia and Herzegovina*

²*University of Sarajevo, Institute for Genetic Engineering and Biotechnology, Sarajevo, Bosnia and Herzegovina*

KEYWORDS STR Loci. Global Phylogeny. Population Data. Diversity. Clustering

ABSTRACT Although DNA genetic markers, including Y-chromosomal short tandem repeats (Y-STRs), are widely used in the analysis of population data, autosomal short tandem repeats (STRs) have a wide role in the investigation of human migration patterns throughout the history, genealogical research, and population genetics. In this review, allele frequencies of 13 autosomal STR loci (D3S1358, TH01, D21S11, D18S51, D5S818, D13S317, D7S820, D16S539, CSF1PO, vWA, D8S1179, TPOX, and FGA) have been reviewed in 64 different worldwide populations. Allele frequency data for 13 STR loci was collected from previously published scientific papers in the journal databases for each studied population and molecular genetic diversity among the 64 sample populations was compared. Further, a worldwide phylogenetic tree and genetic distance values were created using POPTREE2 software and UPGMA method. Results confirmed that the differences among local sub-populations are much smaller than the differences among geographically separated populations. The obtained results, as the researchers had expected, were in the compliance with previously published papers with the difference that the researchers used data on populations from all over the world and thus created a more detailed phylogenetic tree. In that way, the authors offer an insight into the global phylogenetic tree created on the basis of STR allele frequencies for the first time. The goal of this manuscript is to prove the usefulness of these 13 STR markers within the analysis of the genetic distance and its correlation with “geographically-based genetic clustering” among the worldwide populations.

INTRODUCTION

Since 99.7 percent of the human genome is the same, the differences between two individuals are found in 0.3 percent of the entire DNA in the human body (Butler 2011). Short tandem repeats (STRs), also called simple sequence repeats (SSRs) or microsatellites, have 2-7 bp long repetitive units and are sequences that can be up to 400 bp in length, found in the non-coding region of DNA. They occur on all 22 pairs of autosomal chromosomes, as well as on the X and Y chromosomes. They are especially valuable for the study of genetic relationships of closely related populations. Autosomal STRs are much more polymorphic than Y-STRs due to the lack of recombination on this chromosome (Gunn 2006).

STRs were first used in forensic laboratories in the 1990s. Today, the analysis of STR polymorphisms is an inevitable step in various laboratories (Goodwin et al. 2011; Chishti et al. 2014; Yuan et al. 2014). They are exceptionally useful DNA repeat markers since they can easily get amplified and analyzed within the multiplex systems. They represent approximately three percent of the human genome and are found on every 10000 nucleotides (Butler 2011; Press et al. 2014). There are more than 20000 tetranucleotide STR loci identified in the human genome (Butler and Hill 2012), and 13 most widely used ones are the so-called CODIS (Combined DNA Index System) loci selected by the FBI (Gunn 2006).

STRs are frequently used for studying the diversity of various populations as well as for the determination of similarity between closely related populations (Dogan et al. 2013; Hong et al. 2013). They can also be used to determine the genetic authenticity of a specific population, since the population having no similarities in STR allele frequencies with the populations from the other parts of the world can be considered as an autochthonous population that did not experience the gene flow from other gene pools

Address for correspondence:

Serkan Dogan
International Burch University,
Department of Genetics and Bioengineering,
Francuske revolucije bb, Ilidza, Sarajevo,
Bosnia and Herzegovina
Telephone: +387 33 944 494,
Fax: +387 33 944 500
E-mail: serkan.dogan@ibu.edu.ba

(Bentayebi et al. 2014). Statistical parameters related to specific populations or even subpopulations require the existence of population genetic databases (Talledo et al. 2010).

In the present paper, researchers have compared allele frequencies of 13 autosomal STR loci among 64 populations worldwide (Fig. 1). The goal of the paper was to test previously obtained results using an extended number of different populations and to confirm the usefulness of STR markers within the analysis of genetic relationships of an observed population through the “global picture”. Having viewed previous papers, it was expected that geographically close populations would show stronger clustering than the populations that are distant from each other.

MATERIAL AND METHODS

Compared Populations

The following 64 populations have been compared in the paper: Greek, Italian, Swiss, Portuguese, Serbian from Vojvodina (Serbia), Croatian, Lithuanian, Belgian, Belarusian, Bosnian, Andalusian from Spain, Ashkenazi Jewish from Hungary, Finnish, Macedonian, Montenegrin, Nayarit from Mexico, Kuwaiti, Egyptian,

Palestinian, Iraqi, Iranian, Nepalese, Uygur from China, Bangladeshi, Pakistani, Tatar from Poland, Taiwanese, Han Chinese, Japanese, Libyan, Bhutani, Tibetan, Indonesian, German, Equatorial Guinean, Angolan, Mozambican, Sudanese, Karamoja population from Uganda, Szekler and Csango populations from Romania, two Hungarian populations: from Hungary and Serbia, Caucasian and Romany populations from Slovakia, Albanian populations from Kosovo and from North Italy, Turkish populations from Turkey and from Bosnia and Herzegovina, two Berber populations from Morocco: from Asni and Bouhria, two populations from Bihar (India): Yadav and Kurmi, two populations from Tunisia: Kesra (Berber) and Zriba (Arab), three Polish populations: from Northern and Central Poland, from Northern Poland, and from all over Poland, three populations from USA: Caucasian, Hispanic, and African-American, and three Argentinian populations: from Buenos Aires, Santa Fe, and Mendoza.

Collection of Data

A total of 64 population STR data from all over the world has been obtained through the search of journal databases. The original research papers have brought population data in



Fig. 1. Geographical locations of the 64 compared populations

the form of frequencies of allelic variants on autosomal STR loci. All the papers that did not give brief information on experimental procedures have not been included. A minimum of 13 CODIS STR loci of interest have had to be typed in order for a paper to have been included in the research. A number of papers that have contained good population data have been excluded because they did not contain all of the loci of interest. Papers presenting data from several subpopulations within a single population have been included to serve as a quality control: subpopulations within a single population should be in close proximity within the phylogenetic tree if all analyses have been done correctly. The allele frequencies on all STR loci have been summed up to serve as a checkup for errors in the original publications, and the ones in which discrepancies were noticed have been discarded (Table 1).

The kits that were found to be most widely used in research papers are PowerPlex 16 (Promega Corporation, Madison, WI, USA) and AmpF/STR Identifier (Applied Biosystems, Foster City, CA, USA), both of them giving DNA profiles along 13 CODIS and two additional autosomal loci. Old papers have mostly been relying on older kits manufactured by Applied Biosystems and Promega Corporation. These kits have given reliable results but on less than 13 autosomal loci, so two or more kits have been combined in these papers to obtain all loci needed to get satisfying information on a population.

Statistical Analysis

Phylogenetic comparison of the allelic frequencies over 13 STR loci in the populations tested has been performed using POPTREE2 software and UPGMA (unweighted pair group method with arithmetic mean) method. The genetic distances among populations have been obtained by using MEGA5.21 software and the lowest distance for each population has been determined. Multidimensional scaling (MDS) analysis, based on pairwise distance values calculated from the 13 STR loci in 64 populations, has been performed using the SPSS 20.0 software package (SPSS, Chicago, IL, USA).

RESULTS

The major findings of the present paper are presented in the rectangular (Fig. 2) and circular (Fig. 3) phylogenetic trees created by using the

UPGMA method. UPGMA method has been the method of choice in the paper since it uses the sequential clustering algorithm in which the highest level of similarity between the objects of analysis is recognized and the tree is built in a step-wise manner. For such reason, researchers could be able to show the results clearly and clear graphical and tabular representations have been enabled.

On the top of the rectangular phylogenetic tree (Fig. 2), a cluster of European populations has been observed. While high levels of differences have been observed among the populations, those that were closely related and belonged to the same branch have also been observed. Such examples are the Hungarian populations from Hungary and Serbia and the Argentinian populations from three different cities in Argentina that have also clustered with the European populations. This finding is not surprising if the influence of the gene pool from Spain on the Argentinian population is taken into consideration. Due to minimal genetic distances, the populations from South America can be considered as a part of the cluster of European populations. Following this cluster, the cluster of Asian populations has been obtained. Researchers found that Asian populations share much more similarities among each other when compared to European populations. This has been proved by the fact that many populations belong to the same branch in the phylogenetic tree. As an instance, four Arabic populations have clustered together, as well as several populations from The Far East. At the bottom of the rectangular phylogenetic tree, a cluster of Sub-Saharan populations from Africa has been detected. Once again, the researchers have noticed a very strong cluster of six populations being genetically highly similar to each other. Outside this cluster, Arabic populations from the Sub-Saharan region of Africa have been found at the end of this phylogenetic tree. The results presented in the circular phylogenetic tree (Fig. 3), demonstrated to be identical to those presented in the rectangular tree. These two types of graphical representations of the results have enabled the researchers to gain a deeper insight into the results of the paper.

An MDS plot has been made to serve as a way of visual representation of results with an emphasis put on the clustering of populations (Fig. 4). A two-dimensional plot has been constructed to show similarities and differences among world populations. The populations have

Table 1: Compared populations

<i>Population</i>	<i>Size (N)</i>	<i>Reference</i>
1-Szekler-Romania	257	Egyed et al. (2006)
2-Csango-Romania	220	Egyed et al. (2006)
3-Taiwanese-Taiwan	597	Wang et al. (2003)
4-Polish-North and Central Poland	412	Czarny et al. (2005)
5-Belgian-Belgium	100	Decorte et al. (2004)
6-Indonesian-Indonesia	105	Dobashi et al. (2005)
7-Andalusian-Spain	114	Coudray et al. (2007)
8-Kuwaiti-Kuwait	502	Alenizi et al. (2008)
9-Iranian-Iran	126	Hedjazi et al. (2013)
10-Hungarian-Serbia	291	Petriæ et al. (2012)
11-Nepalese-Nepal	233	Jha et al. (2012)
12-Mexican-Nayarit Mexico	231	Cubero et al. (2011)
13-Polish-Poland	870	Soltyszewski et al. (2006)
14-Ashkenazi-Hungary	178	Egyed et al. (2006)
15-Serbian-Vojvodina Serbia	100	Veselinoviæ et al. (2004)
16-Polish-North Poland	145	Szczerkowska et al. (2004)
17-Caucasian-USA	302	Butler et al. (2003)
18-African American-USA	258	Butler et al. (2003)
19-Hispanic-USA	140	Butler et al. (2003)
20-HanChinese-China	2211	Wang et al. (2003)
21-Turkish-Turkey	116	Cakir et al. (2003)
22-German-Germany	7373	Immel et al. (2004)
23-Finnish-Finland	1274	Sippel et al. (2000)
24-Portugese-Portugal	200	Pinheiro et al. (2005)
25-Italian-Italy	835	Carboni et al. (2007)
26-Swiss-Switzerland	511	Gehrig et al. (1999)
27-Slovakian Caucasian-Slovakia	138	Sotak et al. (2008)
28-Romany-Slovakia	138	Sotak et al. (2008)
29-African-Equatorial Guinea	134	Alves et al. (2005)
30-Argentinian-BuenosAires Argentina	516	Marino et al. (2006)
31-Argentinian-SantaFe Argentina	562	Marino et al. (2006)
32-Argentinian-Mendoza Argentina	290	Marino et al. (2006)
33-Bangladeshi-Bangladesh	148	Ferdous et al. (2009)
34-Palestinian-Palestine	125	Abu Halima et al. (2009)
35-Pakistani-Pakistan	328	Rakha et al. (2009)
36-Albanian-North West Italy	100	Robino et al. (2001)
37-Hungarian-Hungary	8426	Agnes Rak et al. (2010)
38-Uygur-China	500	Dong et al. (2010)
39-Bhutani-Bhutan	936	Kraaijenbrink et al. (2007)
40-Tibetan-Tibet China	196	Li et al. (2007)
41-Tatar-Poland	168	Pepinski et al. (2005)
42-Yadav-Bihar India	45	Ashma et al. (2002)
43-Kurmi-Bihar India	50	Ashma et al. (2002)
44-Egyptian-Egypt	265	Omran et al. (2009)
45-Sudanese-Sudan	491	Babiker et al. (2011)
46-Angolan-Angola	110	Beleza et al. (2004)
47-Ugandan-Karamoja Uganda	218	Gomes et al. (2009)
48-Berber-Asni Morocco	105	Coudray et al. (2006)
49-Berber-Bouhria Morocco	104	Coudray et al. (2006)
50-Libyan-Libya	99	Khodjet-el-Khil et al. (2012)
51-Mozambican-Mozambique	344	Alves et al. (2004)
52-Turkish-Bosnia and Herzegovina	100	Dogan et al. (2013)
53-Bosnian-Bosnia and Herzegovina	100	Marjanoviæ et al. (2006)
54-Croatian-Croatia	195	Projia et al. (2007)
55-Montenegrin-Montenegro	101	Jeran et al. (2007)
56-Macedonian-Macedonia	100	Jakovski et al. (2006)
57-Albanian-Kosovo	136	Kubat et al. (2004)
58-Greek-Greece	318	Kovatsi et al. (2006)
59-Japanese-Japan	164	Hara et al. (2004)
60-Lithuanian-Lithuania	300	Caplinskiene et al. (2011)
61-Iraqi-Iraq	103	Barni et al. (2007)
62-Belarusian-Belarus	176	Rebela et al. (2007)
63-Berber-Kesra Tunisia	44	Cherni et al. (2005)
64-Arab-Zriba Tunisia	45	Cherni et al. (2005)

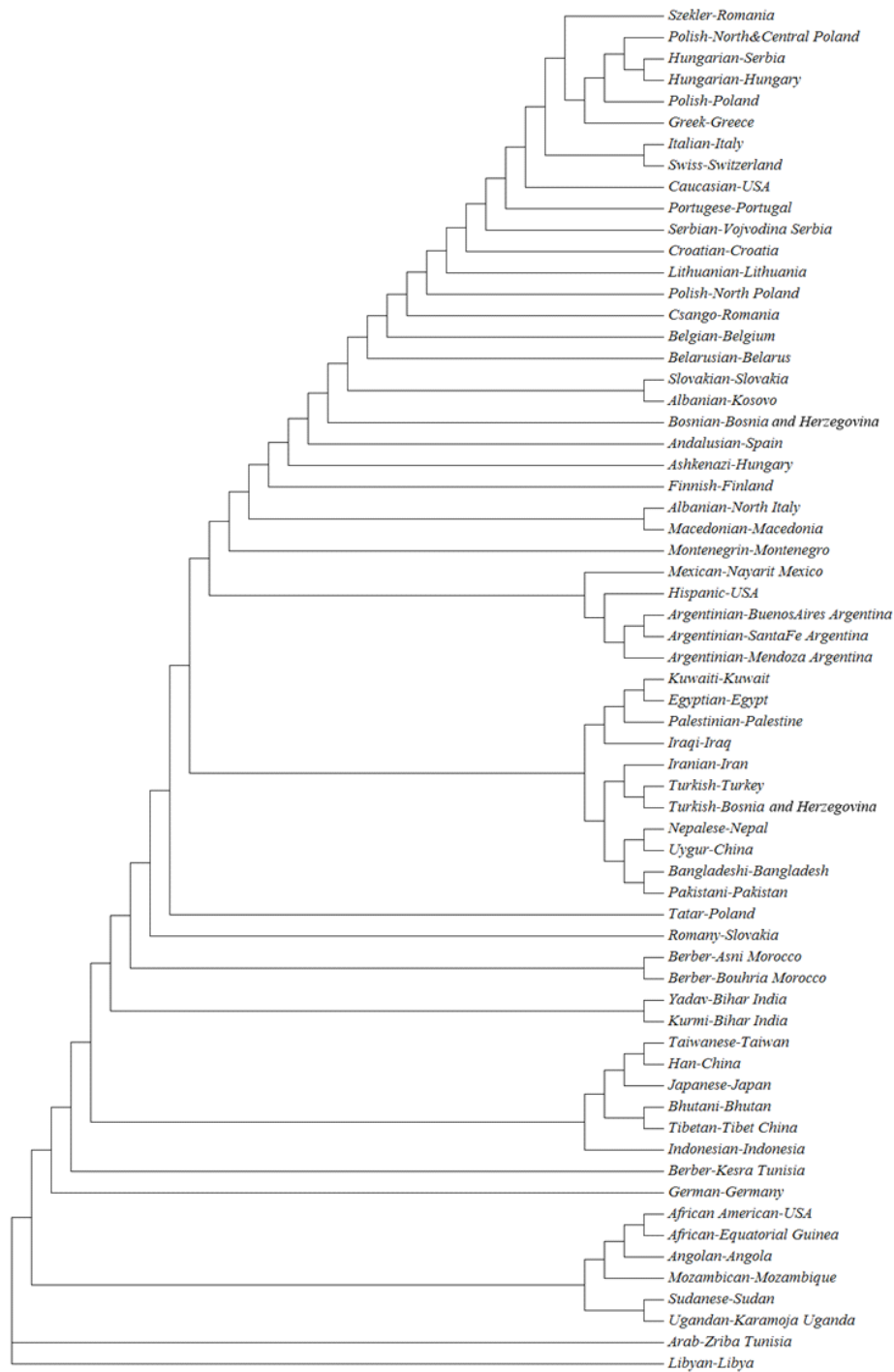


Fig. 2. UPGMA rectangular phylogenetic tree of 13 STRs for compared populations

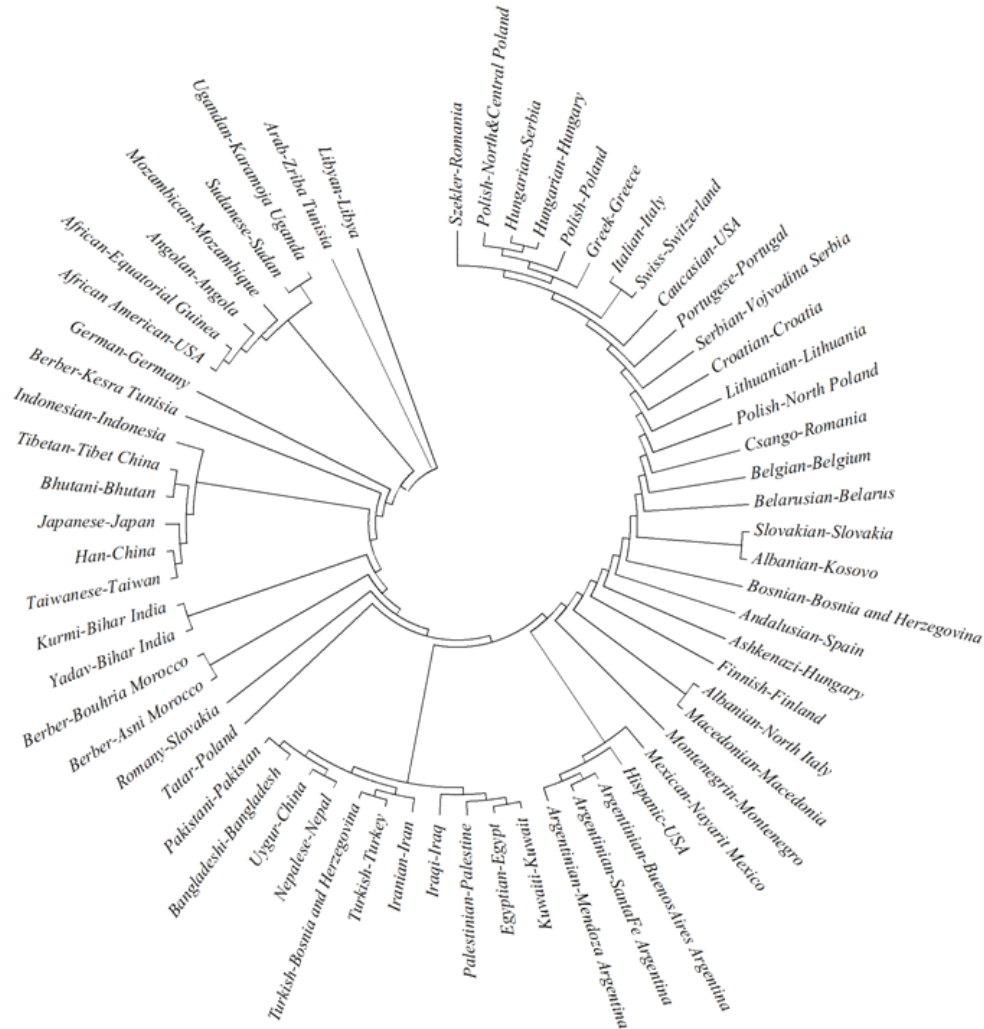


Fig. 3. UPGMA circular phylogenetic tree of 13 STRs for compared populations

clustered into three groups, mostly on the basis of the continent they belong to. Following are the observed clusters: European, Asian, and Sub-Saharan African clusters. Populations from South America have clustered closely to European populations, thus confirming the results obtained by the phylogenetic trees. The cluster of the European population has shown to be the largest one since most of the populations analyzed have been those from Europe and South America (European cluster).

As a final result of the paper, it could be concluded that the hypothesis defined at the very beginning of the research has proven to be correct, according to which the populations belonging to the same continent, or having a similar origin, cluster together in the phylogenetic tree and/or MDS plot. Minor exceptions from this rule are observed in the structure of the phylogenetic tree and will be further discussed on a case-by-case basis.

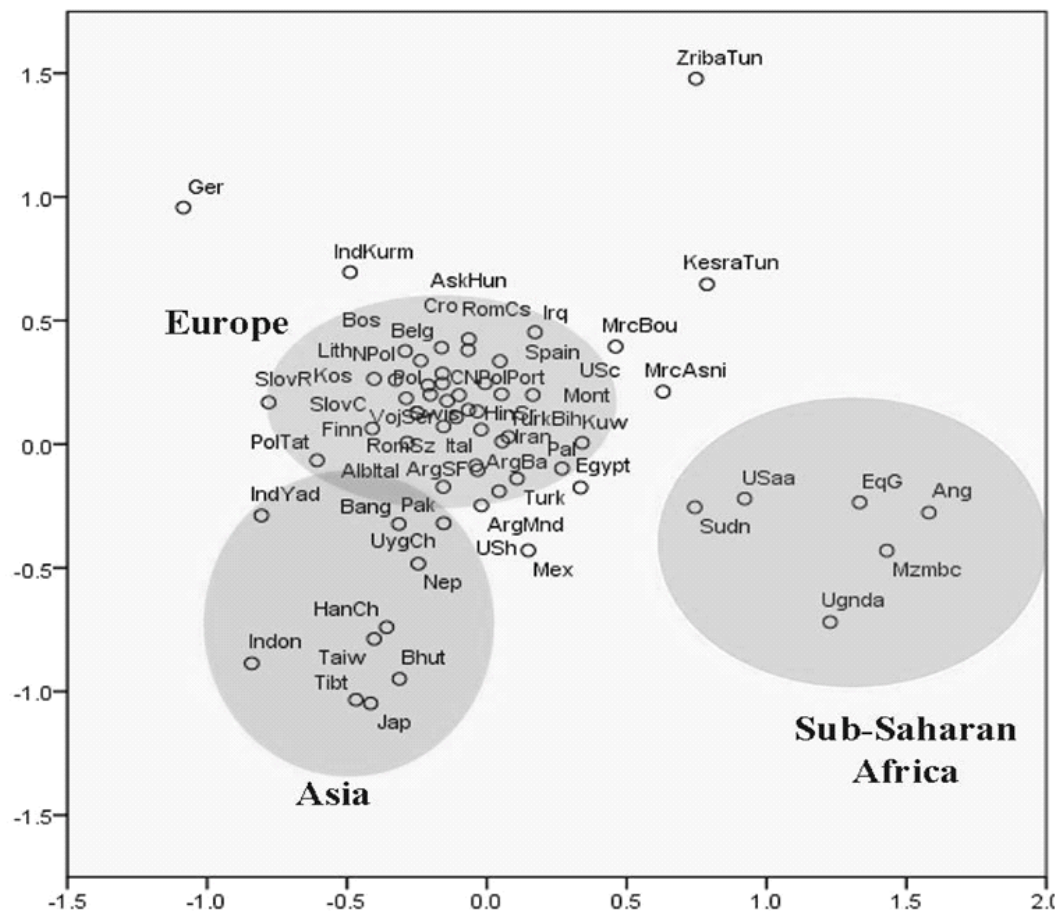


Fig. 4. MDS plot of the data from the study shows that the populations can be divided into three major clusters according to the geographic origin of the sample populations

DISCUSSION

The clustering of worldwide populations according to the continent of origin has been expected, since in the paper done by Hong and colleagues (Hong et al. 2013) similar results have been obtained. The authors of that paper have obtained three clusters of populations: the European cluster, in which Argentinian and Mexican populations have been placed as well, the cluster of African populations (containing populations from Mozambique, Angola, and Equatorial Guinea), and the cluster of Asian populations (this cluster consisted mainly of the populations from The Far East and from India). The

difference between the mentioned reference paper and the present one is that the present paper has used a larger number of populations, thus enabling the creation of a global phylogenetic tree with a more detailed structure.

Despite geographical proximity, the Slovakian Caucasian population has not clustered with Slovakian Romani population. The significant differences between the two populations have been found on all of the loci tested. These results are found to be consistent with those obtained by Sotak and colleagues (Sotak et al. 2008). In the aforementioned paper, Romanies from Slovakia have clustered with the Romanies from Hungary. In the researchers' phylogenetic tree,

Slovakian Romanies has not shown any clustering, which is related to the fact that this is the only population presenting Roma people in the present paper. The second exception from the rule, that geographical proximity determines similar allele frequencies on STR loci has been observed in the case of Poland. While Polish Caucasian populations tested have grouped closely, Polish Tatar population could have been observed on a completely different part of the phylogenetic tree, thus, being more similar to populations from The Far East. This has been found to be consistent with the findings by Pepinski and colleagues (Pepinski et al. 2005), where it was explained that Tatars differ significantly from the autochthonous Polish population due to their different origin and preservation of distinct ethnic and religious character although they live in Poland for centuries. Finally, US Caucasian population has clustered with European populations rather than with US Afro-American and US Hispanic populations which has been expected since the members of Caucasian population living in the USA are known to originate from Europe, while Afro-American and Hispanic populations in the USA originate from Africa and South America, respectively, thus clustering with those populations in the researchers' phylogenetic tree.

In addition, geographical proximity has proven to be the most crucial factor causing clustering of populations in a great majority of cases. The researchers' results have shown to be consistent with those published before and this has also confirmed the hypothesis. As an evidence, in a paper done by Projic and colleagues (Projic et al. 2007), no significant differences between Croatian and Serbian populations have been found on any of the tested loci after Bonferroni correction has been applied. Researchers compared Croatian and Serbian populations from Vojvodina (a region located in the northern part of Serbia), and found no significant differences which has shown to be in compliance with the literature published before.

The investigation of Asian populations in this research has given results similar to those obtained by Rakha and colleagues (Rakha et al. 2009). While populations from Bhutan, Nepal, Tibet, and Uyghur ethnic group have shown clustering, similar behavior has been shown by Pakistani and Bangladeshi populations. The most distant one from these populations has been shown to be the one from Indonesia. Significant

differences between the phylogenetic tree that has been made by Rakha and colleagues (Rakha et al. 2009) and the researchers' phylogenetic tree have been found due to a vastly different number of populations analyzed in the two papers (the paper from 2009 has brought the analysis of 11 populations, all of them are from Asia). More recent papers have confirmed these findings since Zhai and colleagues (Zhai et al. 2014), as well as Yin and co-workers (Yin et al. 2014), have found that Thai, Chinese, and Tibetan populations exhibit strong clustering among each other.

Comparison of Uyghur ethnic group with neighboring populations; Tibetan, Han Chinese, Indian, Iraqi, Pakistani, Palestinian, Turkish, and Nepalese has been done in a paper by Dong and colleagues (Dong et al. 2010). Having incomplete data that has not been explained in detail, researchers could have only concluded that the Uyghur group represents a distinct population which is closely related to the neighboring ones. According to the present paper, the Uyghur population has shown to be most closely related to the Nepalese population, which has also been found in Dong's work (Dong et al. 2010).

Another paper (Kraaijenbrink et al. 2007) comparing Asian populations, those from Bhutan, Nepal, Bangladesh, and India (Yadav population), has obtained results similar to those of the present paper. This paper (Kraaijenbrink et al. 2007) has implied that the differences among these four populations are minor, although they exist, and that observation has been confirmed by the researchers' phylogenetic tree, where these populations have shown to be located in the same region. Additionally, Hossain and co-workers (Hossain et al. 2014) compared Indian and Bangladeshi populations with those from The Far East. The results suggested that the Bangladeshi population strongly clusters with the populations from The Far East (Tibetan, Chinese, Thai, and Nepalese), while Indian subpopulations are more distant. The present study has completely confirmed these results.

The clustering of Asian populations that has been presented in the paper has proved to be correct in a work that is done by Garcia-Bertrand and colleagues (Garcia-Bertrand et al. 2014). They found that the populations of Kuwait, Egypt, and Iraq have clustered together and found to be relatively distant from Turkish population and even more distant from the Pakistani

population. The same results are visible from the phylogenetic tree that is presented in this paper.

In a paper by Alves and colleagues (Alves et al. 2005), the comparison of African populations from Equatorial Guinea, Mozambique, and Angola has been done. Although different data sources have been used, the same model of clustering has been obtained in the researchers' paper showing that Cabinda population from Angola is very similar to the population from Equatorial Guinea and that Mozambican population differs more from these populations. However, statistically significant differences have not been observed on any of the 13 loci in any of these three populations, making the researchers' results further consistent with those from the paper by Alves and colleagues (Alves et al. 2005).

Finally, quality control has been done by using several subpopulations from a single population that has been expected to have similar allelic variants on autosomal STR loci. Across, researchers have checked the credibility of the papers used in this research, as well as the precision of the researchers' methods. Data generated has been expected: two Indian populations (Yadav and Kurmi) clustered together, as well as two Berber populations from Morocco (Asni and Bouhria), Hungarians living in Hungary and in Serbia, and three Polish Caucasian populations. Three Argentinian populations have also been compared in this paper. No significant differences between them are found, which is consistent with the data reported by Marino and colleagues (Marino et al. 2006).

CONCLUSION

Researchers have compared available autosomal STR data on a significant number of human populations, described their genetic (dis)similarity, and clustered them. These studies, once again, have shown that autosomal STR data could be a great descriptor of the geographical clustering for the global human population. By finding the genetic similarities and differences between populations, their geographical location and/or origin could be more precisely determined. In fact, even if geographical location of examined population is not identified, according to the autosomal STR data based on its comparison with same data of some known starting population, geographical position of this

population in most cases could be successfully approximated.

REFERENCES

- Abu Halima MSH, Pineda Bernal L, Sharif FA 2009. Genetic variation of 15 autosomal short tandem repeat (STR) loci in the Palestinian population of Gaza Strip. *Legal Medicine*, 11(4): 203-204.
- Agnes Rak S, Zalan A, Szabados G, Pamjav H 2010. Population genetic data on 15 STR loci in the Hungarian population. *Forensic Sci Int-Gen*, 5(5): 543-544.
- Alenizi M, Goodwin W, Ismael S, Hadi S 2008. STR data for the AmpF/STR identifier loci in Kuwaiti population. *Legal Medicine*, 10(6): 321-325.
- Alves C, Gusmao L, Damasceno A, Soares B, Amorim A 2004. Contribution for an African autosomic STR database (AmpF/STR Identifier and Power Plex 16 System) and a report on genotypic variations. *Forensic Sci Int*, 139(2-3): 201-205.
- Alves C, Gusmao L, Lopez-Parra AM, Mesa MS, Amorim A, Arroyo-Pardo E 2005. STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpF/STR Identifier and Powerplex 16 Kits. *Forensic Sci Int*, 148(2-3): 239-242.
- Ashma R, Kashyap VK 2002. Genetic polymorphism at 15 STR loci among three important sub-populations of Bihar, India. *Forensic Sci Int*, 130: 58-62.
- Avise JC 1989. Gene trees and organismal histories: A phylogenetic approach to population Biology. *Evolution*, 43(6): 1192-1208.
- Babiker HMA, Schlebusch CM, Hassan HY, Jakobsson M 2011. Genetic variation and population structure of Sudanese populations as indicated by 15 identifier STR loci. *Investig Genet*, 2(1): 12.
- Barni F, Berti A, Pianese A, Boccellino A, Miller MP, Caperna A, Lago G 2007. Allele frequencies of 15 autosomal STR loci in the Iraq population with comparisons to other populations from the Middle-Eastern region. *Forensic Sci Int*, 167(1): 87-92.
- Beleza S, Alves C, Reis F, Amorim A, Carracedo A, Gusmao L 2004. 17 STR data (AmpF/STR Identifier and PowerPlex 16 System) from Cabinda (Angola). *Forensic Sci Int*, 141: 193-196.
- Bentayebi K, Abada F, Ihzmad H, Amzazi S 2014. Genetic ancestry of a Moroccan population as inferred from autosomal STRs. *Meta Gene*, 2: 427-438.
- Butler JM 2007. Short tandem repeat typing technologies used in human identity testing. *BioTechniques*, 43(4): 2-5.
- Butler JM 2010. *Fundamentals of Forensic DNA Typing*. Burlington, MA, USA: Elsevier Inc.
- Butler JM 2011. *Advanced Topics in Forensic DNA Typing: Methodology*. Waltham, MA, USA: Elsevier.
- Butler JM, Hill CR 2012. Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Sci Rev*, 24(1): 15-26.
- Butler JM, Schoske R, Vallone PM, Redman JW, Kline MC 2003. Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. *J Forensic Sci*, 48(4): 908-911.
- Cakir AH, Celebioglu A, Altunbas S, Yardimci E 2003. Allele frequencies for 15 STR loci in Van-Aöry dis-

- tricts of the Eastern Anatolia region of Turkey. *Forensic Sci Int*, 135: 60-63.
- Caplinskiene M, Pauliukevicius A, Baranoviene R, Jan-kauskienė J, Kukienė J, Savanevskyte K, Ruzgaite G 2011. Forensic application of autosomal STR analysis in Lithuanian population. *Forensic Sci Int-Gen Supplement Series*, 3(1): 562-564.
- Carboni I, Nutini AL, Porfiriio B, Genuardi M, Ricci U 2007. Genetic STRs variation in a large population from Tuscany (Italy). *Forensic Sci Int-Gen*, 1(3-4): 10-11.
- Cherni L, Loueslati Yaacoubi B, Pereira L, Alve C, Khodjet-El-Khil H, Ben Ammar El Gaaied A, Amorim A 2005. Data for 15 autosomal STR markers (PowerPlex 16 System) from two Tunisian populations: Kesra (Berber) and Zriba (Arab). *Forensic Sci Int*, 147(1): 101-106.
- Chishti HM, Ansar M, Ajmal M, Hameed A 2014. Application of short tandem repeat markers in diagnosis of chromosomal aneuploidies and forensic DNA investigation in Pakistan. *Gene*, 548(2): 217-222.
- Coudray C, Calderon R, Guitard E, Ambrosio B, González-Martín A, Dugoujon JM 2007. Allele frequencies of 15 tetrameric short tandem repeats (STRs) in Andalusians from Huelva (Spain). *Forensic Sci Int*, 168(2): 21-24.
- Coudray C, Guitard E, Keyser-Tracqui C, Melhaoui M, Cherkaoui M, Larrouy G, Dugoujon JM 2006. Population genetic data of 15 tetrameric short tandem repeats (STRs) in Berbers from Morocco. *Forensic Sci Int*, 167(1): 81-86.
- Cubero MJA, Gonzale LJM, Ponce DV, Najär EAD, Caloca GHM, Alvarez JC, Lorente JA 2011. Genetic variation of 15 autosomal microsatellite loci in a Nayarit population (Mexico). *Legal Medicine*, 13: 323-327.
- Czarny J, Grzybowski T, Derenko MV, Malyarchuk BA, Sliwka DM 2005. Genetic variation of 15 STR loci (D3S1358, vWA, FGA, TH01, TPOX, CSF1PO, D5S18, D13S317, D7S820, D16S539, D2S1338, D8S1179, D21S11, D18S51, and D19S433) in populations of north and central Poland. *Forensic Sci Int*, 147(1): 97-100.
- Decorte R, Engelen M, Larno L, Nelissen K, Gilissen A, Cassiman JJ 2004. Belgian population data for 15 STR loci (AmpF/STR SGM Plus and AmpF/STR Profiler PCR Amplification kit). *Forensic Sci Int*, 139(2): 211-213.
- Dobashi Y, Kido A, Fujitani N, Hara M, Susukida R, Oya M 2005. STR data for the AmpF/STR identifier loci in Bangladeshi and Indonesian populations. *Legal Medicine*, 7(4): 222-226.
- Dogan S, Kovacevic L, Marjanovic D 2013. Genetic polymorphisms of 15 STR loci within Turkish student population living in Sarajevo, Bosnia and Herzegovina. *Coll Antropol*, 37(4): 1313-1319.
- Dong L, Liu XX, Zhang HH, Bo R, Wei G, Zhang LS 2010. Genetic data for 15 STR loci in Uyghur ethnic group of North-West China. *Rom J Leg Med*, 1: 59-62.
- Egyed B, Füredi S, Angyal M, Balogh I, Kalmar L, Padar Z 2006. Analysis of the population heterogeneity in Hungary using fifteen forensically informative STR markers. *Forensic Sci Int*, 158: 244-249.
- Egyed B, Füredi S, Padar Z 2006. Population genetic study in two Transylvanian populations using forensically informative autosomal and Y-chromosomal STR markers. *Forensic Sci Int*, 164(2): 257-265.
- Ferdous A, Ali ME, Alam S, Hasan M, Hossain T, Akhteruzzaman S 2009. Forensic evaluation of STR data for the PowerPlex 16 System loci in a Bangladeshi population. *Legal Medicine*, 11: 198-199.
- Garcia-Bertrand R, Simms TM, Cadenas AM, Herrera RJ 2014. United Arab Emirates: Phylogenetic relationships and ancestral populations. *Gene*, 533(2014): 411-419.
- Gehrig C, Hochmeister M, Borer UV, Dirnhofer R, Budowle B 1999. Swiss Caucasian population data for 13 STR loci using AmpF/STR Profiler Plus and Cofiler PCR Amplification Kits. *J Forensic Sci*, 44(5): 1035-1038.
- Gomes V, Sanches-Diz P, Alves C, Gomes I, Amorim A, Carracedo A, Gusmao L 2009. Population data defined by 15 autosomal STR loci in Karamoja population (Uganda) using AmpF/STR Identifier Kit. *Forensic Sci Int-Gen*, 3(2): e55-e58.
- Goodwin W, Linacre A, Hadi S 2011. *An Introduction to Forensic Genetics*. Chichester, UK: John Wiley and Sons Ltd.
- Gunn A 2006. *Essential Forensic Biology*. Chichester, West Sussex, England: John Wiley and Sons Ltd.
- Hara M, Yamamoto Y, Takada A, Saito K, Kido A, Oya M, Kameyama H 2004. Population data for 15 STR loci D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penat D, ChvWA, D8S1179, TPOX and FGA in Japanese. In: Dautremepuich, Morling N (Eds.): *International Congress Series: Progress in Forensic Genetics 10*. Arcachon: Elsevier pp. 204-206.
- Havas D, Jeran N, Efremovska LJ, Dordevic D, Rudan P 2007. Population genetics of 15 AmpF/STR identifier loci in Macedonians and Macedonian Romani (Gypsy). *Forensic Sci Int*, 173: 220-224.
- Hedjazi A, Nikbakht A, Hosseini M, Hoseinzadeh A, Hosseini SMV 2013. Allele frequencies for 15 autosomal STR loci in Fars Province population, Southwest of Iran. *Leg Med*, 15(4): 226-228.
- Hong SB, Kim SH, Kim KC, Park MH, Lee JY, Song JM, Han MS, Kim W 2013. Korean population genetic data and concordance for the PowerPlex® ESX 17, AmpF/STR Identifier®, and PowerPlex® 16 systems. *Forensic Sci Int-Gen*, 7(3): e47-e51.
- Hossain T, Hasan M, Mazumder AK, Momtaz P, Sharmin T, Sufian A, Das SA, Akhteruzzaman S 2014. Population genetic data on 15 autosomal STR loci in Bangladeshi population. *Forensic Sci Int-Gen*, 13: 4-5.
- Immel UD, Kleiber M, Klitschar M 2004. Population genetic analysis in a German population from Saxony-Anhalt using the PowerPlex16 System. In: Ch Dautremepuich, Morling (Eds.): *International Congress Series: Progress in Forensic Genetics 10*. Arcachon: Elsevier pp. 216-218.
- Jakovski Z, Nikolova K, Furac I, Masic M, Janeska B, Kubat M 2006. Allele frequencies for 15 STR loci in a population from the Republic of Macedonia. *Int J Legal Med*, 120(1): 53-55.
- Jeran N, Havas D, Ivanovic V, Rudan P 2007. Genetic diversity of 15 STR loci in a population of Montenegro. *Coll Antropol*, 31(3): 847-852.
- Jha DK, Gonzalez LJM, Rijal JP, Tuladhar BS, Chhetri NT 2012. Allele frequencies of 15 AMPF/STR Identifier loci in the Nepalese population. *Scientific World*, 10(10): 20-23.

- Khodjet-el-Khil H, Fadhlouli-Zid K, Gusmao L, Alves C, Benammar-Elgaai A, Amorim A 2012. Allele frequencies for 15 autosomal STR markers in the Libyan population. *Annals of Human Biology*, 39(1): 80-83.
- Kovatsi L, Parsons TJ, Just RS, Irwin JA 2006. Genetic variation for 15 autosomal STR loci (PowerPlex 16) in a population sample from Northern Greece. *Forensic Sci Int*, 159(1): 61-63.
- Kraaijenbrink T, Van Driem GL, Tshering of Gaselo K, De Knijff P 2007. Allele frequency distribution for 21 autosomal STR loci in Bhutan. *Forensic Sci Int*, 170(1): 68-72.
- Kubat M, Skavic J, Behluli I, Nuraj B, Bekteshi T, Behluli M, Klaric Martinovic I, Pericic M 2004. Population genetics of the 15 AmpF/STR identifier loci in Kosovo Albanians. *Int J Legal Med*, 118(2): 115-118.
- Li KL, Jianmin Z, Kai L, Xiaosong L 2007. Allele frequencies of 15 STR loci of Tibetan lived in Tibet Lassa. *Forensic Sci Int*, 168(2): 236-240.
- Marino M, Sala A, Corach D 2006. Population genetic analysis of 15 autosomal STRs loci in the Central Region of Argentina. *Forensic Sci Int*, 161: 72-77.
- Marjanovic D, Bakal N, Pojskic N, Kapur L, Drobnic K, Primorac D, Bajrovic K, Hadziselimovic R 2006. Allele frequencies for 15 short tandem repeat loci in a representative sample of Bosnians and Herzegovinians. *Forensic Sci Int*, 156: 79-81.
- Okamoto O, Yamamoto Y, Inagaki S, Yoshitome K, Ishikawa T, Imabayashi K, Miyashi S, Ishizu H 2003. Analysis of short tandem repeat (STR) polymorphisms by the PowerPlex 16 System and capillary electrophoresis: Application to forensic practice. *Acta Med Okayama*, 57(2): 59-71.
- Omran GA, Ruttly GN, Jobling MA 2009. Genetic variation of 15 autosomal STR loci in upper (Southern) Egyptians. *Forensic Sci Int-Gen*, 3(2): 39-44.
- Pepinski W, Niemcunowicz-Janica A, Skawronska M, Janica J, Koc-Zorawska E, Aleksandrowicz-Bukin M, Soltyszewski I 2005. Genetic data on 15 STR loci in the ethnic group of Polish Tatars residing in the area of Podlasie (North-Eastern Poland). *Forensic Sci Int*, 149(2): 263-265.
- Petric G, Draskovic D, Bosic DZ, Budakov B, Veselinovic I 2012. Genetic variation at 15 autosomal STR loci in the Hungarian population of Vojvodina Province, Republic of Serbia. *Forensic Sci Int-Gen*, 6(6): e163-e165.
- Pinheiro MF, Caine L, Pontes L, Abrantes D, Lima G, Pereira MJ, Rezende P 2005. Allele frequencies of sixteen STRs in the population of Northern Portugal. *Forensic Sci Int*, 148(2-3): 221-223.
- Press MO, Carlson KD, Queitsch, C 2014. The overdue promise of short tandem repeat variation for heritability. *Trends in Genetics*, 30(11): 504-512.
- Projic P, Skaro V, Samija I, Pojskic N, Durmic-Pasic A, Kovacevic L, Bakal N, Primorac D, Marjanovic D 2007. Allele frequencies for 15 short tandem repeat loci in representative sample of Croatian population. *Croat Med J*, 48(4): 473-477.
- Rakha A, Yu B, Hadi S, Sheng-bin L 2009. Population genetic data on 15 autosomal STRs in a Pakistani population sample. *Legal Medicine*, 11(6): 305-307.
- Rebala K, Wysocka J, Kapinska E, Cybulska L, Mikulich AI, Tsybovsky IS, Szczerkowska Z 2007. Belarusian population genetic database for 15 autosomal STR loci. *Forensic Sci Int*, 173: 235-237.
- Robino C, Gino S, Torre C 2001. Allele frequencies for the PowerPlex 16 STR loci in an Albanian population sample from Northern Italy. *J Forensic Sci*, 46(4): 998-999.
- Rout PK, Joshi MB, Mandal A, Laloe D, Singh L, Thangaraj K 2008. Microsatellite-based phylogeny of Indian domestic goats. *BMC Genetics*, 9(11): 1-11.
- Sippel H, Hedman M, Sajantila A 2000. DNA database in the Finnish population using AmpF/STR profiler and AmpF/STR SGM plus multiplex analysis by capillary electrophoresis. In: GF Sensabaug, PJ Lincoln B Olaisen (Eds.): *Progress in Forensic Genetics 8*. Amsterdam: Elsevier, pp. 132-135.
- Simkova H, Faltus V, Marvan R, Pexa T, Stenzl V, Broucek J, Horinek A, Mazura I, Zvarova J 2009. Allele frequency data for 17 short tandem repeats in a Czech population sample. *Forensic Sci Int-Gen*, 4(1): e15-e17.
- Soltyszewski I, Spolnicka M, Kartasinska E, Konarzewska M, Pepinski W, Janica J 2006. Genetic variation of STR loci D3S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX and FGA by GenePrint PowerPlex 16 in a Polish population. *Forensic Sci Int*, 159: 241-243.
- Sotak M, Petrejckikova E, Bernasovska J, Bernasovsky I, Sovicova A, Boronova I, Svickova P, Bozikova A, Gabrikova D 2008. Genetic variation analysis of 15 autosomal STR loci in Eastern Slovak Caucasian and Romany (Gypsy) population. *Forensic Sci Int-Gen*, 3: 21-25.
- Szczerkowska Z, Kapinska E, Wysocka J, Cybulska L 2004. Northern Polish population data and forensic usefulness of 15 autosomal STR loci. *Forensic Sci Int*, 144(1): 69-71.
- Takezaki N, Nei M, Tamura K 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with windows interface. *Molecular Biology and Evolution*, 27(4): 747-752.
- Talledo M, Gavilan M, Choque C, Aiquipa L, Arevalo J, Montoya Y 2010. Comparative allele distribution at 16 STR loci between the Andean and Coastal population from Peru. *Forensic Sci Int-Gen*, 4(4): e109-e117.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28(10): 2731-2739.
- Veselinovic I, Kubat M, Furac I, Skavic J, Klaric IM, Tasic M 2004. Allele frequencies of the 15 AmpF/STR identifier loci in the population of Vojvodina Province, Serbia and Montenegro. *Int J Legal Med*, 118: 184-186.
- Wang CW, Chen DP, Chen CY, Lu SC, Sun CF 2003. STR data for the AmpF/STR SGM plus and profiler loci from Taiwan. *Forensic Sci Int*, 138(1): 119-122.
- Wang W, Jia H, Wang Q, Cai Z, Wei L, Wang D, Bittles AH 2003. STR polymorphisms of "Forensic Loci" in the Northern Han Chinese population. *J Hum Genet*, 48: 337-341.

- Yin C, Ji Q, Li K, Mu H, Zhu B, Yan J, Yu Y, Wang J, Chen F 2014. Analysis of 19 STR loci reveals genetic characteristics of Eastern Chinese Han population. *Forensic Sci Int-Gen*, 14: 108-109.
- Yuan J-Y, Wang X-Y, Shen C-M, Liu WJ, Yan J-W, Wang H-D, Pu H-W, Wang Y-L, Yang G, Zhang Y-D, Meng H-T, Jing H, Zhu B-F 2014. Genetic profile characterization and population study of 21 autosomal STR in Chinese Kazak ethnic minority group. *Electrophoresis*, 35(4): 503-510.
- Zhai D, Yang J, Huang Y, Chen L, Wu J 2014. The allele frequency of 15 STRs among three Tibeto-Burman-Speaking populations from the southwest region of Mainland China. *Forensic Sci Int-Gen*, 13: 22-24.