Journal of Genetics and Genetic Engineering
ISSN: 2637-5370 | Volume 6, Issue 1, 2024
https://doi.org/10.22259/2637-5370.0601002

RESEARCH ARTICLE

Distribution of 18 STR-Haplotypes of Y-Chromosome in Megacities in Connection with Creating Reference Databases

I. G. Udina[1], M.A. Gubina[2], A. S. Gracheva[3]

[1]*Doctor of biological sciences, leader of the group of population immunogenetics, N.I. Vavilov General Genetics Institute of the Russian Academy of Sciences, Moscow, Russia.*

[2]*Candidate of biological scientist, senior scientist, Federal Research Center Institute of Cytology and Genetics of the Siberian Department of the Russian Academy of Sciences, Novosibirsk, Russia.*

[3]*Candidate of biological scientist, senior scientist, N.I. Vavilov General Genetics Institute of the Russian Academy of Sciences, Moscow, Russia.*

Received: 21 November 2024 Accepted: 06 December 2024 Published: 26 December 2024

Corresponding Author: I. G. Udina, Doctor of biological sciences, leader of the group of population immunogenetics, N.I. Vavilov General Genetics Institute of the Russian Academy of Sciences, Moscow, Russia.

Abstract

For the senior generations of megacities - Moscow and Novosibirsk, no significant differences by the distribution of 18 STR of Y-chromosome were observed. By 18 STR-haplotypes (*DYS389I, DYS389II, DYS390, DYS19, DYS385A, DYS385B, DYS456, DYS437, DYS438, DYS447, DYS448, DYS449, DYS391, DYS392, DYS393, DYS439, DYS635* и *DYS576*) haplogroups of Y-chromosome were detected with high probability using predictors Whit Athey's and NEVGEN. ORG for prediction haplogroups of Y-chromosome. Significant differences were revealed between two senior generations by «Southern origin haplotypes», which are introduced into population gene pool with migrant flows. The youngest generation of Moscow, which is under intensive influence of modern migration from the regions of Middle Asia and Northern Caucasus, demonstrates high frequencies of «Southern origin haplogroups». *(C3,G2a, G2c, J1, J2, L, O2, O3, Q, R2* and *T*). Revealed peculiarities of the gene pool of megacities evidence necessity of development special reference databases by DNA-markers of Y-chromosome for each megacity using simultaneously typing of DNA-markers and collecting questionnaire data for detection genetic demographic parameters for inhabitants of megacities, especially, affiliation to particular generation.

Keywords: Megacities, Migration, Y-chromosome Haplogroups, Reference Database.

1. Introduction

Population reference databases serve the basis for interpretation of probability estimates of the results of personal identification in criminalistics for DNA-identification of victims of disasters, terroristic acts, remains of those who died on the war, historical remains, and for identifications intended on solving social oriented problems, establishing paternity, kinship etc. During the forensic examination reliability of personal identification depends on the correct choice of the reference database. There is necessity of exact knowledge with which frequency considered DNA-markers are found in the population. Frequencies of DNA-markers in the reference database might be obtained by the investigation of the representative sample from the population of an individual's origin. Affiliaion to specific haplogroup is determined by key SNP or by several additional SNP (for considering subhaplogroups) in specific positions of Y-chromosome [Human Y-chromosome DNA haplogroup]. During generations SNP are accumulated in Y-chromosome, and, also, linked STR-markers, because SNP- and STR-markers are localized in non-recombining area of Y-chromosome (NRY). The peculiarity of Y-chromosome as genetic marker provides an opportunity to detect affiliation to definite haplogroup of Y-chromosome by STR-haplotype using specially designed computer

Citation: I. G. Udina, M.A. Gubina, A. S. Gracheva. Salceda. Distribution of 18 STR-haplotypes of Y-chromosome in Megacities in Connection with Creating Reference Databases. Journal of Genetics and Genetic Engineering. 2024; 6(1):16-19.

programs- predictors of Y-chromosome haplogroups [Whit Athey's Haplogroup Predictor., Y-DNA Haplogroup Predictor – NEVGEN.ORG].

Population of megacity represents the population of mixed type with specific peculiarities of genetic demographic processes (high effective volume of population, complex infrastructure, centripetal migration, high density of population, multi-ethnic composition, outbreeding, unstable character of natural reproduction of population, artificial habitat, environmental pollution) (the main factor of gene pool dynamics being migration) [Kurbatova et al. 2013, Kurbatova O., Yankovsky N. 2016].

The goal of our investigation is considering the peculiarities of the frequency profile of Y-chromosome 18 STR-haplotypes in the senior generation of Moscow and Novosibirsk in comparative aspect aiming to study gene pools of megacities on the European territory and in Siberia before the modern migrant flows formed after collapse of the USSR, particularly, from North Caucasus and Middle Asia.

2. Materials and Methods

In 2001-2004, the investigation of population of Novosibirsk (N=117) on the basis of medical institutions including genetic demographic questionnaire and blood collection, was performed. Earlier, senior sample of male population from Moscow (N=73) was studied, collected in 2001 [Udina et al. 2022]. Samples from two megacities consist from representatives of senior generation who were born before the beginning of modern intensive migration from North Caucasus and Middle Asia after collapse of the USSR (1991). Earlier studied sample from the youngest generation from Moscow is considered as comparative sample [Udina et al. 2022].

DNA from biological samples was isolated by standard method using reagents of «Isogen» (Moscow). Genotyping of DNA samples by STR Y-хромосомы (DYS389I, DYS389II, DYS390, DYS19, DYS385A, DYS385B, DYS456, DYS437, DYS438, DYS447, DYS448, DYS449, DYS391, DYS392, DYS393, DYS439, DYS635 и DYS576) was performed in the laboratory of OOO «Gordiz» (Moscow) with subsequent establishing of affiliation of revealed STR-genotype to haplogroup of Y-chromosome by the main predictor - Whit Athey's Haplogroup Predictor and additional NEVGEN.ORG predictor. Statistical analysis was performed using computer programs «WINPEPI» [Abramson J. 2011], «GenAlEx 6.5» [Peakall R., Smouse P. 2012],«Arlequin 3.5» [Excoffier, L., Lischer H. 2010], and MS Excel.

3. Results

18 STR-haplotypes in the samples compared are studied by R_{st}. Over 90% of haplogroups of Y-chromosome detected by STR-haplotype using predictor are predicted with probability 100%. For the other cases, additional detection of haplogroups was performed using other predictors available in the Internet. With single exclusions, the probability level of Y-chromosome haplogroup detection reached 99%. In population of Moscow and Novosibirsk, the distribution of Y-chromosome haplogroups with the most widely spread haplogroups R1a, R1b, E1b1b1, N, I1, I2and prevailing R1a was revealed (Table1) [Balanovskaya E., Balanovskii O. 2007].

For the studied sample from male population of Novosibirsk, the absence in the genepool migrant flows from the North Caucasus and Middle Asia, which are marked in our study by «Southern origin haplotypes»(C3,G2a, G2c, J1, J2, L, O2, O3, Q, R2 and T) [Udina et al. 2022] was found. From the total spectrum of the haplogroups, only haplogroups J1 and J2 with the frequency (2.7 %) were observed.

Table 1. *Frequencies of large haplogroups of Y-chromosome in the studied samples of the senior generation of Novosibirsk and two generations of Moscow.*

Haplogroup	Novosibirsk (senior generation)		Moscow (senior generation)		Moscow (the youngest generation)	
	N	*f*	*N*	*f*	*N*	*f*
E1b1b	5	0.043	1	0.014	17	0.043
I1	9	0.077	4	0.055	21	0.052
I2	14	0.120	6	0.082	42	0.104
J1	1	0.009	1	0.014	16	0.040
J2	2	0.018	4	0.055	29	0.073
N	15	0.129	13	0.178	29	0.073
R1a	63	0.543	38	0.520	181	0.452
R1b	6	0.052	3	0.041	26	0.065
«Rare»*	1	0.009	3	0.041	39	0.098
Total:	**116**	**1.000**	**73**	**1.000**	**400**	**1.000**

**- with frequency lower, than 2% in all the samples.*

Estimate of Rst=0,00 for the profiles of 18 STR-haplotypes of Y-chromosome for comparison the samples from the senior generation from Novosibirsk and Moscow. For the sample from Moscow, the frequency of «Southern by origin haplogroups» is 11%, four haplogroups (*J1, J2, T* and *Q*) were found. Between senior generations for two megacities, significant differences were observed by the frequency of «Southern origin haplogroups» (*p=0.024, OR=0.22 CI95% 0.06-0.81*). Observed peculiarities for the sample from Moscow compared with the sample from Novosibirsk indicate the earlier penetration of migration flows into megalopolis which bring some «Southern by origin haplogroups»». It is obvious that Moscow like the capital is more attractive for migrants. In the youngest generation of Moscow, 21,1% «Southern by origin haplogroups». Obtained estimates are in good agreement with the questionnaire data of origin and ethnic composition of migration.

The senior generation of megacities on the European territory (Moscow) and in Siberia (Novosibirsk) are similar by Rst estimates for STR-haplotypes of Y-chromosome. Corresponding estimates or the senior generation of Novosibirsk and the youngest generation of Moscow - Rst=0.005, for the senior and the youngest generations of Moscow - Rst=0,006.

Therefore, in the senior generations of two megacities compared with the youngest generation from Moscow relatively more low frequency of «Southern origin haplogroups» of Y-chromosome, characteristic for the modern migrant flows, is observed, which is 11% in Moscow and 2.6% in Novosibirsk due to different migration flows. There are some peculiarities in the distribution of haplogroups of Y-chromosome in the studied samples due to their geographical position: in Novosibirsk subhaplogroup *N1a2 (CTS6380),* characteristic for Siberian native peoples [Haplogroup N1a1 (Y-DNA), Haplogroup N1a2b (Y-DNA)], was found with frequency 2,6%, while in Moscow only subhaplogroup *N1a1 (M46)*was present.

4. Conclusion

Significant differences between senior generations of two megacities by the presence of «Southern origin haplotypes» of Y-chromosome and revealed peculiarities of dynamics of gene pool for Moscow population confirm necessity of creating reference databases for each megacity Thus, gene pool of megacity change under action of migration, migration and parameters of migration are specific for each megacity and representatives of distinct generations. Therefore, in the process of developing reference databases by DNA-markers, it is necessary, also, to take in consideration geographic peculiarities of haplogroups of Y- chromosome [Map of the geographic distribution of haplogroups of Y-chromosome on the territory of Eurasia]. DNA-markers of Y-chromosome are quite informative for considering dynamics of gene pool of megalopolis under action of migration [Udina et al, 2023].

Acknowledgement

Authors are grateful to the inhabitants of Novosibirsk who took part in survey and provided biological samples for the study. From each person in the sample informed approval for the using of biological sample (blood) and genetic demographic data from questionnaire in anonymous way in the population study.

The study is approved by ethical committee of IOGenRAS. All the procedures performed in the study involving people are corresponding ethical standards of institutional and/or national committee and Helsinki declaration 1964 year and its subsequent changes or comparable norms of ethics. The study is approved by Ethical committee of IOGen RAS. Authors declare the absence of conflict of interest.

The investigation is performed in the frames of the theme of IOGenRAS Government Assignment «The study of polymorphism on the cellular, organism and population levels as the basis for creating genetic technologies» (№ 0092-2022-0001).

5. References

1. Abramson JH. (2011). WINPEPI updated: computer programs for epidemiologists, and their teaching potential. *Epidemiol Perspect Innov*..8(1),1. doi: 10.1186/1742-5573-8
2. Balanovskaya E.V., Balanovskii O.P Russian gene pool on the Russian plain. – Moscow: Luch, 2007.– 416 p.
3. Excoffier, L. and H.E. L. Lischer (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564-567. (URL: http://cmpg.unibe.ch/software/arlequin35)
4. Haplogroup N1a1 (Y-DNA) [Electronic resource]: https://ru.wikipedia.org/wiki/haplogroup_N1a1_(Y-DNA)
5. Haplogroup N1a2b (Y-DNA) [Electronic resource]: https://ru.wikipedia.org/wiki/haplogroup_N1a2b_(Y-DNA)

6. Human Y-chromosome DNA haplogroup. URL: https://en.wikipedia.org/wiki/Human_Y-chromosome_DNA_haplogroup
7. Kurbatova, O. L., Pobedonostseva, E. Y., Veremeichyk, V. M., Prudnikova, A. S., Atramentova, L. A., Tsybovsky, I. S., & Udina, I. G. (2013). Genetic demography of populations of three megalopolises in relation to the problem of creating genetic databases. *Russian Journal of Genetics*, *49*, 448-456.
8. Kurbatova O.L., Yankovsky N.K. (2016) Migration as the Main Factor of the Russia's Urban Population Dynamics // *Russian Journal Genetics*, 52(7), 726–745.
9. Map of the geographic distribution of haplogroups of Y-chromosome on the territory of Eurasia. [Electronic resource]: URL: https://i.imgur.com/4rsFqdk.jpg
10. Peakall R. and Smouse P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics,*28, 2537-2539.
11. Udina, I. G., Gracheva, A. S., & Kurbatova, O. L. (2022). Frequencies of Y-chromosome haplogroups and migration processes in three generations of Moscow residents. *Russian Journal of Genetics*, *58*(11), 1365-1372.
12. Udina I.G., Gracheva A.S., Kurbatova O.L. (2023) Scientific and applied significance of the analysis of Y-chromosome variation for criminalistics, ethnodemographic studies and historical reconstructions, *Social Area,*9 (1) DOI: 10.15838/sa.2023.1.37.7
13. Whit Athey's Haplogroup Predictor. Haplogroup Predictor[Electronic resource]: (hprg.com)
14. Y-DNA Haplogroup Predictor – NEVGEN.ORG [Electronic resource]: http://nevgen.org.