BJMG 11/2 (2008) 25-30
10.2478/v10034-008-0030-0

ORIGINAL ARTICLE

ALU INSERTION POLYMORPHISMS IN POPULATIONS OF THE SOUTH CAUCASUS

Litvinov S*, Kutuev I, Yunusbayev B, Khusainova R, Valiev R, Khusnutdinova E

***Corresponding Author:** Mr. Sergey Litvinov, Institute of Biochemistry and Genetics of Ufa Science Center of Russian Academy of Sciences, Prospekt Oktyabrya, 71, Ufa, 450054, Russia; Tel./Fax: +7-3472-356088; e-mail: litviss@mail.ru

ABSTRACT

An analysis of eight Alu insertion loci (ACE, ApoA1, PV92, TPA25, NBC27, NBC102, NBC148 and NBC182) was carried out in three ethnic groups of the South Caucasus (Abkhazians, Armenians and Georgians). Genetic differentiation between these groups (Gst = 1%) was half as much as in European populations and four time less than in populations of Dagestan. Although it was not possible to determine a contribution of Neolithic farmers to the Caucasian gene pool, the principal component analysis showed clear differences between these populations and those of Europe, Siberia and Asia. No evidence of correlation between genetic and linguistic data in our populations was disclosed. However, long-term habitation on common territory has led to common cultural, anthropological and detectable genetic traits.

Key words: Population genetics; South Caucasus; Alu insertion

INTRODUCTION

Because of the geographic peculiarities of the Caucasus, study of its human populations offers an opportunity to assess the influence of geographic barriers on their genetic structure. Previous publications, based on mitochondrial DNA (mtDNA) data, have shown that genetic relationships reflect geographic rather than linguistic relationships [1-3]. Despite the presence of the Caucasus Mountains as a potentially significant barrier to gene-flow, considerable correlation between pairs of geographically-distant populations has been observed [2,4]. Also, these mountains had no detectable influence on the genetic structure of the Caucasus populations [3]. These results led to the proposal of genetic drift as the major factor influencing the genetic structure of the Caucasus populations [5,6]. An analysis based on mtDNA and Alu-insertion data has also shown, that the Caucasus populations are more similar to European rather than to western Asian populations, whereas Y-chromosome analysis revealed a closer relationship to the latter populations [3,5]. Mitochondrial DNA and Y-chromosome-based analyses have also shown an affinity of Caucasus populations with Near East [5,7].

The Alu family of short interspersed elements is a relatively stable autosomal polymorphic marker with a unique mutational mechanism, for which the ancestral state is the absence of Alu insertion. Because it reflects both maternal and paternal history, an Alu insertion polymorphism is a highly informative tool for studying the genetic structure of human populations [8].

We have analyzed three populations from the South Caucasus: Abkhazians (North-Caucasian language family, Abkhaz-Adyghe language subgroup), Georgians (Kartvelian language family) and Armenians (Indo-European language family) to assess the genetic diversity of linguistically and historically different, but related populations in this specific geographic region [9]. Because the Georgian nation

includes multiple ethnic groups we examined the Mingrelian population inasmuch as it shares significant traits with Abkhazians from their centuries-old interaction. Armenians are a separate ethnic group, which originated from Neolithic tribes of the Armenian Uplands. In the 12th- 11th centuries BC, this group gained Hittite, Hurrite and partially Abkhaz-Adyghe and Kartvel elements. Later, in 8th-7th centuries BC, a Cimmerian-Scythian element was added to its gene pool. We also compared our results with those from other populations of Eurasia.

MATERIALS AND METHODS

Blood samples (212 in total) from unrelated individuals from the following three ethnic groups: Abkhazians (98), Georgians (Mingrelians) (72), Armenians (42) were collected in the Abkhazian Republic after informed consent was obtained from all participants. The origin of all subjects participating in this study was ascertained over three generations. Published data for additional ethnic groups has been taken for comparison in this study (Table 1).

Table 1. Data used in the present study (ACE, ApoA1, PV92, TPA25, NBC27, NBC102, NBC148, NBC182 and He)

Population	*2n*	Reference	Population	*2n*	Reference
South Caucasus:			**North Caucasus:**		
Abkhazians	196	Present study	Karachays	162	17
Georgians	144	Present study	Kumyks	120	17
Armenians	84	Present study	Kuban Nogays	126	17
			Karanogays	150	17
			Lezgiz	96	18
			Agars	120	18
			Tabassarans	108	18
			Dargins	146	18
			Andis	124	18
			Bagvalals	64	18
			Chamalals	70	18
Central Asia:			**Europe:**		
Kazakhs	166	17	Finnish	40	16
Uighurs	126	17	French	40	16
Uzbeks	144	17	Northern European	136	16
			Polish	20	16
Central South Siberia:					
Kalmyks	90	17			
Yakuts	85	17			
Evenks	41	17			
Volga-Ural Region:					
Bashkirs	68	17, 23			
Tatars	152	17, 23			
Komis	140	17, 23			
Maris	98	17, 23			
Mordvinians	74	17, 23			
Udmurts	140	17, 23			
Total			**Total**	3270	

2n: number of chromosomes.

Table 2. Alu insertion frequencies and the average expected heterozygosity (He) in three studied groups from the South Caucasus

Population	*n*	ACE	ApoA1	pv92	TPA25	NBC27	NBC102	NBC148	NBC182	He
Abkhazians	196	0.490	0.954	0.260	0.490	0.316	0.495	0.168	0.515	0.390
Georgians	144	0.361	0.972	0.208	0.417	0.319	0.368	0.208	0.479	0.382
Armenians	84	0.369	1.000	0.226	0.429	0.310	0.381	0.119	0.643	0.359
Average He per locus		0.476	0.047	0.355	0.492	0.432	0.479	0.273	0.485	

DNA samples were extracted from peripheral blood lymphocytes using a phenol-chloroform extraction method [10]. All the samples were genotyped for eight Alu insertion markers (ACE, ApoA1, pv92, TPA25, Ya5NBC27, Ya5NBC102, Ya5NBC148, Ya5NBC182). Primer sequences have been described previously [11].

Allele frequencies and heterozygosity were determined by direct counting. A principal component analysis made use of the Statistica 6.0 package. The Gst value (the coefficient of gene differentiation) has been calculated as described by Nei and Chesser [13]. The Hardy-Weinberg equilibrium was evaluated by an exact test [14].

RESULTS

Allele frequencies of the eight Alu insertions in the three groups are shown in Table 2. In all the loci, two alleles were revealed with the exception of ApoA1 in Armenians, which is fixed in the presence of an Alu insertion. Two of 24 tests performed for Hardy-Weinberg equilibrium indicated significant departure from equilibrium (ApoA1 in Georgians and Armenians; $p < 0.05$). Thus, these deviations most likely represent random statistical fluctuation.

The average expected heterozygosity for each group varied from 0.359 to 0.398 (Table 2). The expected average heterozygosity for each marker in all populations was 0.047 for ApoA1 to 0.492 for TPA25.

The standard test for heterogeneity of allele frequencies between the groups showed substantial heterogeneity. While there are only two alleles in the locus, Gst value is equivalent to Wright's Gst [15]. The Gst values ranged from 0.001 for the NBC27 locus to 0.023 for NBC182. The average Gst for the whole dataset was 0.010, which is almost one half that for European populations (0.018) [16]. This result is most similar to that obtained for the North Caucasus populations by Kutuev *et al.* [17], whereas the level of genetic differentiation only in the populations of Dagestan, for the same loci (except B65, NBC123 and NBC51) was 0.045 [18]. The Gst value that was obtained for both North and South Caucasus populations and for different Alu markers (ACE, TPA25, PV92, APO, FXIIIB, D1, A25, B65) was 0.113 [4].

The level of genetic differentiation for each pair of our three groups has been estimated. The Gst value for Abkhazian and Georgian population was the same as for Georgian and Armenian populations, *i.e.,* 0.007. The differentiation between Abkhazian and Georgian populations was 0.009.

To assess the genetic relationships of the three groups, a principal component analysis was performed (Figure 1). The first two components together account for 55.4% of the observed variance. The first principal component accounts for 36.3%. It separates the studied Caucasian populations from populations of Central Asia and Siberia.

The studied Abkhazian group is closer to the European rather than to Asian populations along the first axis. It is clearly separated from Siberian and Asian populations, but there are no visible correlations on plot with Volga-Ural and Dagestan populations according to their linguistic or geographic attributes. The Abkhazian population from the studied group appears closer to populations of Dagestan, who are placed together except Karanogays, positioned with Kazakhs and Uzbeks. This can be explained by an apparent Asian origin of Karanogays [19]. With the exception of that, Dagestan, Central Asia and Siberian populations are clearly separated

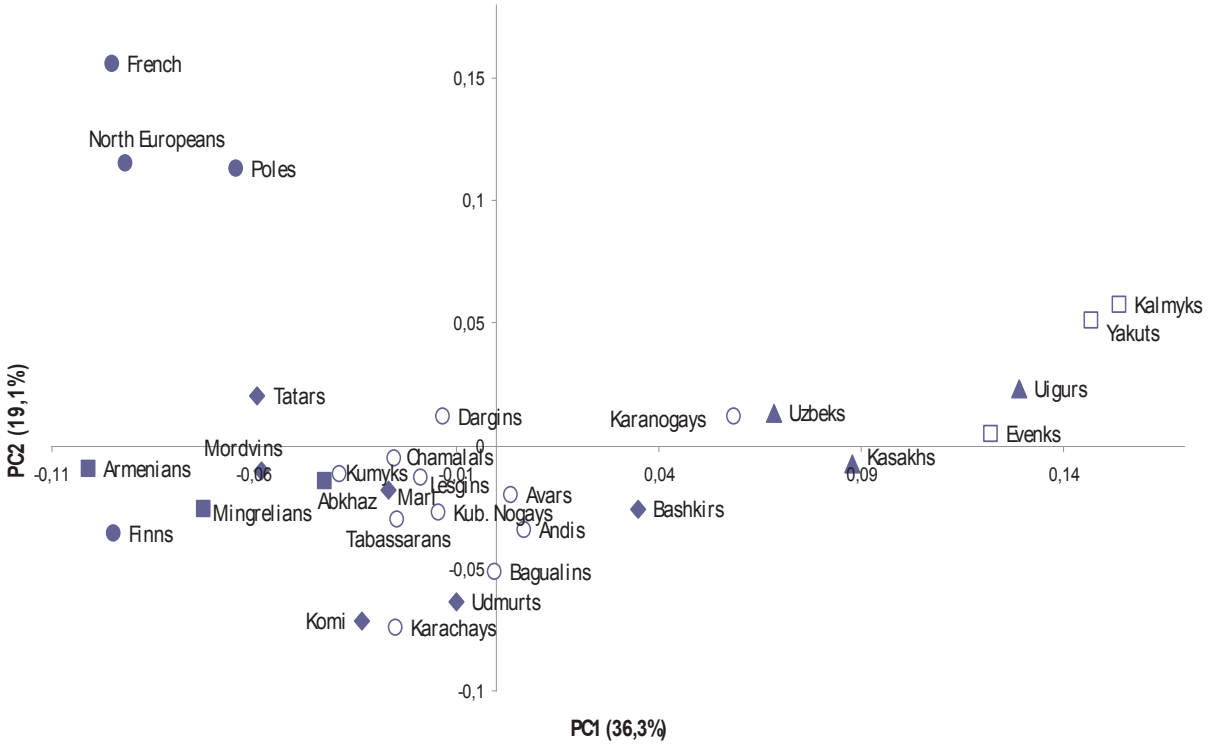

Figure 1. Principal component analysis of investigated populations based on eight Alu insertion polymorphisms (ACE, ApoA1, PV92, TPA25, NBC27, NBC102, NBC148 and NBC182). Studied populations of Abkhazia are marked as (■), European populations as (●), Dagestan populations as (○), Volga-Ural populations as (♦), Central Asia populations as (▲), Central South Siberia populations as (□).

from each other. Armenian and Georgian populations cluster together with Tatars, Mordvinians and Finns. Regarding the second principal component, the latter are located close to Uralic speaking groups such as Mari, Udmirts and Komis. This concurs with the idea of Finno-Ugric Turkish interaction during past millennia.

DISCUSSION

The average Gst value for all markers in the three South Caucasus groups was 0.010. It is half as much as that for European populations (0.018) [16] and it is significantly lower than Gst obtained for mountain inhabitants of Dagestan (0.045) [18]. The Gst calculated for the Dagestan populations is based on the same markers that we used in this study. In spite of the absence in our study of three Alu loci used by Yunusbayev *et al.* [18] (NBC51, B65, A25), the difference in genetic differentiation between the South Caucasus and Dagestan populations is obvious. It was shown that pure drift was not the sole reason for such differentiation in Dagestan and could possibly be the result of migrations and isolation in steppe and mountain populations, respectively. Thus, the possible reason for the difference in the Gst value is that the isolation in Abkhazian settlements is not as strong as it is in Dagestan, where an aul (a remote mountain village) can be considered as an isolate [20]. Furthermore, the populations studied do not represent the aul inhabitants. The main reason for the expected genetic differences in Abkhazian populations, which can be seen on the PC plot, could be historical and genetic distinction. The Mingrelians are the descendants of Georgian tribes, which inhabited the central region and foothills of Colchis on the territory of present-day western Georgia and Abkhazia. Due to centuries-old neighborhoods and

contacts with western Georgian ethnic groups, the Abkhazians gained common anthropological and cultural traits from them. Although anthropologically Abkhazians are most close to western Georgian groups, which are intermediate between Pontic and Armenoid types (while signs of Caucasian type are not expressed), Abkhazians are more culturally and linguistically close to North-Caucasian Adygs. The Indo-European speaking Armenians are quite different from our two other groups, both linguistically and anthropologically. Yet the genetic difference between them and the two other groups (0.007) is slightly lower than that between the Abkhazian and Georgian populations (0.009). The main reason for this could be the penetration of Kartvel and Abkhaz-Adyghe elements to the Armenian gene pool because of dispersal settling of different Armenian groups in the Caucasus.

Although classical markers (blood groups, serum proteins and red cell enzymes) showed clinal frequency distribution from Anatolia to Europe, later studies did not show any such signs [21]. In a previous study [4], it was shown that Alu insertion analyses along with mtDNA data places the Caucasus populations alongside European populations. It was proposed then that the Caucasus populations represent an earlier "layer" of the European populations [22]. That European and Caucasus populations are close to each other has now been confirmed [17].

Our principal component analysis clearly distinguishes the populations of the Caucasus and Asia. However, we cannot exclude a Neolithic contribution to the contemporary gene pool. The possible reason for the absence of the frequency distribution gradient can be genetic drift, reinforced by isolation that could conceal the influence of Neolithic farmers on the Caucasus populations [1,21]. Our study can neither confirm nor disprove both these assumptions. Markers with high powers of resolution should be used to determine the extent of impact on the Caucasus populations from the Near East.

In summary, our analysis of eight Alu loci in three South Caucasian groups revealed a close relationship between them. The geographic neighborhood showed significant influence on genetic proximity in spite of linguistic and cultural differences between the groups. Our results confirm that geographic differentiation correlate with genetic diversity to a greater degree than linguistic differentiation. While an Alu insertion marker does not have enough power of resolution to assess the contribution of the influence of Neolithic farmers on the Caucasian gene pool, it clearly separates both South and North Caucasus populations (except Karanogays) from Siberian and Asian populations.

ACKNOWLEDGMENTS

This study was supported by grants of the Russian Foundation for Basic Research (07-04-01016-a) and the Presidium of the Russian Academy of Sciences Foundation "Biodiversity and Dynamics of Gene Pools" (both to E. Khusnutdinova) and Grant of the President of the Russian Federation (to I. Kutuev).Dynamics of Gene Pools."

REFERENCES

1. Barbujani G, Whitehead GN, Bertorelle G, Nasidze I. Testing hypotheses on processes of genetic and linguistic change in the Caucasus. Hum Biol 1994; 66(5): 843-864.

2. Nasidze I, Stoneking M. Mitochondrial DNA variation and language replacements in the Caucasus. Proc R Soc Lond (Biol) 2001; 268(1472): 1197-1206.

3. Nasidze I, Sarkisian T, Kerimov A, Stoneking M. Testing hypotheses of language replacement in the Caucasus: evidence from the Y chromosome. Hum Genet 2003; 112(3):255-261.

4. Nasidze I, Risch GM, Robichaux M, Sherry ST, Batzer MA, Stoneking M. Alu insertion polymorphisms and the genetic structure of human populations from the Caucasus. Eur J Hum Genet 2001; 9(4): 267-272.

5. Nasidze I, Ling EYS, Quinque D, Dupanloup I, Cordaux R, Rychkov S, Naumova O, Zhukova O, Sarraf-Zadegan N, Naderi GA, Asgary S, Sardas S, Farhud DD, Sarkisian T, Asadov C, Kerimov A, Stoneking M. Mitochondrial DNA and Y-chromosome variation in the Caucasus. Ann Hum Genet 2004; 68(Pt. 3): 205-221.

6. Bulayeva KB, Davudov OM, Pavlova TA, Kurbanov RM, Bulayev OA, Harpending H. Genetic subdivision of Daghestan ethnic populations. Russ J Genet 2003; 39(1): 68-76. Translated from Genetika 2003; 39(1): 83-92.

7. Reguero M, Cadenas AM, Gayden T, Underhill PA, Herrera R.J. Iran: tricontinental nexus for Y-chromosome driven migration. Hum Hered 2006; 61(3): 132-143.

8. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. Nat Rev Genet 2002; 3(5): 370-379.

9. People of the Caucasus. Gardanov BA, Ed. Narody Kavkaza. Moscow: Izdatelstvo Akademii Nauk SSSR, 1962 (in Russian).

10. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1989.

11. Watkins WS, Ricker CE, Bamshad MJ, Carroll ML, Nguyen SV, Batzer MA, Harpending HC, Rogers AR, Jorde LB. Patterns of ancestral human diversity: An analysis of Alu-insertion and restriction-site polymorphisms. Am J Hum Genet 2001; 68(3): 738-752.

12. StatSoft Inc: STATISTICA (data analysis software system), version 6, 2001; www.statsoft.com.

13. Nei M, Chesser RK. Estimation of fixation indices and gene diversities. Ann Hum.Genet 1983; 47(Pt. 3): 253-259.

14. Guo S, Thompson E. Performing the exact test of Hardy-Weinberg proportion for multiple alleles.Biometrics 1992; 48(2): 361-372.

15. Nei M. F-statistics and analysis of gene diversity in subdivided populations. Ann Hum Genet 1977; 41(2): 225-233.

16. Watkins WS, Rogers AR, Ostler CT, Wooding S, Bamshad MJ, Brassington A-ME, Carroll ML, Nguyen SV, Walker JA, Prasad BVR, Reddy PG, Das PK, Batzer MA, Jorde LB. Genetic variation among world populations: Inferences from 100 Alu insertion polymorphisms. Genome Res 2003; 13(7): 1607-1618.

17. Kutuev I, Khusainova R, Karunas A, Yunusbayev B, Fedorova S, Lebedev Y, Hunsmann G, Khusnutdinova E. From East to West: patterns of genetic diversity of populations living in four Eurasian regions. Hum Hered 2006; 61(1): 1-9.

18. Yunusbayev B, Kutuev I, Khusainova R, Guseinov G, Khusnutdinova E. Genetic structure of Dagestan populations: a study of 11 Alu insertion polymorphisms. Hum Biol 2006; 78(4): 465-476.

19. Osmanov MO. Narody Dagestana (nationalities of Dagestan). Tishkov VA, Cheshko SV, Eds. Moscow: Nauka, 2002; 492-493.

20. Bulayeva KB, Jorde L, Watkins S, Ostler C, Pavlova TA, Bulayev OA, Tofanelli S, Paoli G, Harpending H. Ethnogenomic diversity of Caucasus, Daghestan. Am J Hum Biol 2006; 18(5): 610-620.

21. Barbujani G, Nasidze IS, Whitehead GN. Genetic diversity in the Caucasus. Hum Biol 1994; 66(4): 639-668.

22. Bulayeva K, Jorde LB, Ostler C, Watkins S, Bulayev O, Harpending H. Genetics and population history of Caucasus populations. Hum Biol 2004; 75(6): 837-853.

23. Khusainova R, Akhmetova V, Kutuev I, Salimova A, Korshunova T, Lebedev Y, Khusnutdinova E. Genetic structure of the Volga-Ural and Central Asia populations inferred form the data on Alu polymorphism Russ J Genet 2004; 40(4): 443-449.