ISSN 2073-2244
www.alb-shkenca.org

AKTET
Journal of Institute Alb-Shkenca
Revistë Shkencore e Institutit Alb-Shkenca

MICROEVOLUTION PROCESSES IN PËRMETI POPULATION AS ESTIMATED BY ISONYMY TECHNIQUES

* DERVISHI A.[a], MIKEREZI I.[b]
[a] Universiteti i Tiranës, Fakulteti i Shkencave Natyrore, Departamenti i Bioteknologjisë
[b] Universiteti i Tiranës, Fakulteti i Shkencave Natyrore, Departamenti i Biologjisë

PËRMBLEDHJE

Struktura gjenetike e popullates së Përmetit është analizuar përmes shpërndarjes së mbiemrave nga një kampion prej 1057 martesash të regjistruara në periudhën kohore 1946-2006. Vlerësimet e α së Fisherit, si tregues i pasurisë së mbiemrave dhe ν së Karlin-McGregorit, si tregues i ritmeve të migrimit, janë fituar për vendasit dhe për të ardhurit në Përmet. Ato treguan vlera më të larta në të ardhurit. E njëjta gjë është demonstruar edhe nga studimi i vijës së regresionit të shpërndarjes lo_2k-log_2S. Dy vlerësime të ndryshëm midis të gjithë kombinimeve të mundshëm të individëve që i përkisnin nëntë dhjetëvjeçarëve u fituan përmes distancës Eukldiane dhe koefiçientit të ngjashmërisë sipas Laskerit, përkatësisht. Krahasimet e ecurisë së F_t, koefiçienti i inbridingut tërësor dhe α së Fisherit treguan se çdo zvogëlim i vlerave të F_t ndiqej nga rritja e vlerave të α si pasojë e migrimit të kohëve të fundit. Niveli i inbridingut në popullatë si dhe vlera e lartë e endogamisë janë ndikuar nga mënyra e migrimit të popullatës si dhe nga pozicioni gjeografik i qytetit të Përmetit.

SUMMARY

The genetic structure of Përmeti population was analyzed though the distribution of surnames obtained in a sample of 1057 marriages registered in the period 1946-2006. Estimates of Fisher's α, an indicator of richness of surnames and ν of Karlin-McGregor, and indicator of migration rates, were obtained for both natives and immigrants. They showed higher values in immigrants. The same was demonstrated even by log2k-log2S regression line. Two different estimates of relationships between all possible combinations of cohorts born in nine different decades were obtained through Euclidean distance and Lasker's coefficient of relationship, respectively. Comparisons of Ft, coefficient of total inbreeding, and Fisher are α indicated that any decrease of Ft was followed by an increase of α as the result of recent migration. The level of population inbreeding and high value of endogamy (61,9%) are affected by the migration pattern and geographic position of Përmeti town.

Keywords: isonymy, surname distribution, marital structure

INTRODUCTION

In recent decades surnames have been frequently used as genetic markers to evaluate different factors that influence the genetic structure of human populations. Giving that in most societies surnames are transmitted through the paternal line, they can be considered as neutral alleles of one locus on Y chromosome (1, 2, 4, 9). This property of surnames allows investigating the temporal trends in the variability of the population (4, 8).

The study of Përmeti population is part of an investigation aiming to estimate the genetic relationships among all different local populations in Albania by using different methods. By this way will be possible to cast more light on genetic history of Albanian population.

Përmeti is located in South-Eastern part of Albania in a rather isolated geographical position. In this framework, giving the absence of genetic investigations in this population, it seems interesting to study the frequency distribution of surnames in order to obtain population dynamics indicators especially related to migration. On the other side, this study aims to test whether the geographic position and migration patterns have affected the inbreeding level. However, the investigation of Përmeti population, as in other

ones, will continue in the future by using other alternative genetic methods.

MATERIALS AND METHODS

The records of 2114 marriages of individuals resident in Përmeti were obtained from the registers of the Municipality. All the marriages celebrated during one year were recorded for every five years starting from 1946 (than 1951, 1956 and so on) up to 2006. Unfortunately, the records before 1946 were not used because the registers for were either incomplete or absent. For each individual (subject) surname, gender, year and place of birth, and place of origin were recorded. A total of 1081 different surnames were counted. Among them, 710 surnames were subjects born in Përmeti, and 371 surnames were subjects who had immigrated from elsewhere. All the records were grouped in 9 classes according to the date of birth of the bearers. In order to analyze chronological trends in marital structure the data were grouped in 7 classes according to the period of marriage.

The distribution of surnames was studied by sorting the main files by place of birth and period of birth. For each of both groups that were formed by the above sorting, several parameters were estimated. The unbiased isonymy (I) and Fisher's α were estimated according to (8):

$I = \Sigma_i q_i^{\,2} - 1/n$ (q_i - relative frequency of the i-th surname in each cohort, n – total number of surnames in the same cohort) where the summation is over all surnames and

$\alpha = 1/I$

Karlin-McGregor's v was calculated according to (9):

$v = (1–B)/ [B (n-1)]$, where *B* is the biased isonymy estimated as:

$B = \Sigma_i q_i^{\,2}$ (q_i - as defined above)

The relationship between cohorts born in different decades was studied according to the following formula:

$R = \Sigma_k (q_{ik}q_{jk})$, where q_{ik} and q_{jk} are the relative frequencies of surname *k* in cohorts *i* and *j* respectively. This parameter will be considered here as Lasker's coefficient.

The Euclidean distance between cohorts was also estimated according to (4):

$D = \sqrt{(1 - \cos\theta)}$, where $\cos\theta = \Sigma_k \sqrt{(q_{ik}q_{jk})}$, with q_{ik} and q_{jk} as defined above.

Both groups of immigrants and natives to Përmeti were compared by the slope of regression lines fitted to $\log_2 k$-$\log_2 S$ distributions according to (1) where S is the number of surnames appeared k times.

In order to analyze the endogamy, the marriages were subdivided into three groups according to the birthplace of the spouses: a) both partners born in Përmeti, b) one partner born in Përmeti and the other born elsewhere, c) both partners born elsewhere. Changes in premarital migration were studied by measuring the linear geographic distance (in kilometers) between birthplaces of mates. The total inbreeding (Ft) and its random component (Fr) were estimated from marital isonymy according to (6) and (5) respectively.

RESULTS AND DISCUSSION

1. Surname distribution by place birth

The distribution frequency of surnames was given as a function of the number of individuals carrying the same surname. In Figure 1 the distribution of surnames was shown by place of birth.

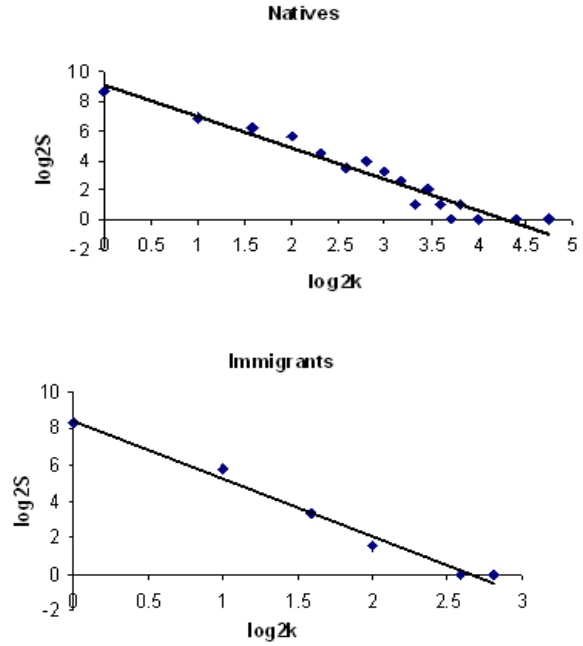


Figure 1: $\text{Log}_2 k$ – $\log_2 S$ distribution of surnames by place of birth.

It is shown that the slopes of the regression line change significantly with the origin: in immigrants it had the value b = - 3.159 whereas in natives the value of the slope was b = - 2.121. Indeed, from 710 different surnames belonging to subjects born in Përmeti, here considered as natives, 391 surnames appeared only once, whereas among 371 surnames belonging to immigrants 311 had the frequency one. The higher number of surnames with unique frequency or with a small one among immigrants, is a good indication that migration moves generally single individuals or small family groups.

2. Indicators of temporal changes in the group structure: In order to investigate changes in the population structure, different indicators based on surname distribution are used. Fisher's α is an indicator of the richness of surnames as a consequence of the accumulated migration while, Karlin-McGregor's v is an indicator of the immigration since it is correlated to the surnames that appear only once. Figure 2 gives the values of Fisher's α in different decades for both natives and immigrants (All) It is shown that the highest value of α corresponded to the second (1911-1920) and to the eighth (1971-1980) periods, while the lowest value is reached during the fifth period (1941-1950).

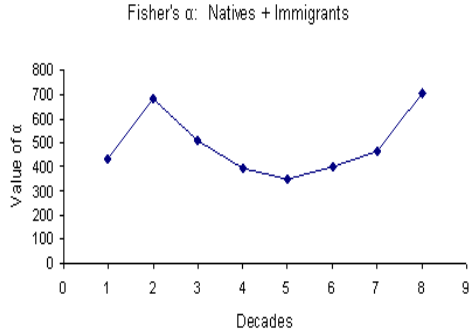


Figure 2 Fisher's α values for Natives and Immigrants (All)

Figure 3 gives the comparison between Fisher's α among natives and immigrants according to the decade of birth.

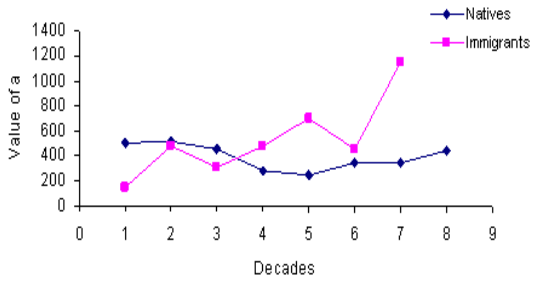
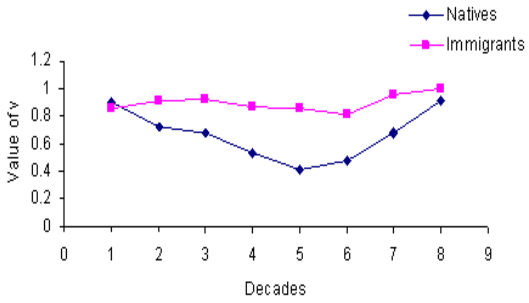


Figure 3: Comparison of Fisher's α between immigrants and natives in Përmet

Similar values of α for both residents and immigrants groups were obtained for the first three periods of time while the values of α were higher in immigrants in the consecutive periods. Since migration generally moves single persons or small family groups, the richness of surnames and consequently surname variability was higher among immigrants than in natives.



In Figure 4 the comparison of Karlin-McGregor's v between immigrants and natives in Përmeti by the decade of birth are shown.

Figure 4: Comparison of Karlin-McGregor's v between immigrants and natives in Përmeti. The values of Karlin-McGregor's v were higher among immigrants compared to natives. This is a consequence of the fact that v is correlated to the surnames of unique frequency.

3. Relationship between cohorts as measured by Lasker's coefficient and Euclidean distances: In Table 1 the matrix of Lasker's and Euclidean distances between all the possible combinations of cohorts are given. The distances were calculated separately for natives, immigrants and for all together. If immigration through different decades was random regarding the surname, Lasker's distance between cohorts was expected to be nil, while Euclidean distance should be one. The tendency toward low values for the Lasker's distances and the tendency for high values of Euclidean distance were obvious in Table 1. Lasker's distance values tended to be lower among immigrants compared to natives while for Euclidean distance the relationship is inverted. It was shown that the relationship between the cohorts is decreasing in the following decades for both Lasker's and Euclidean distance.

Rodriguez-Larralde *et al*. (1993) estimated Lasker's and Euclidean distances in relationship with time periods using data from Municipality of Perugia, Italy, for the period 1900-1990. They observed that the relationship among the cohorts decreased with time, showing that the population became more heterogeneous with time. A similar trend is observed in our study as well (Table 1).

		Per 1	Per 2	Per 3	Per 4	Per 5	Per 6	Per 7	Per 8	Per 9
	All		0.0024	0.0029	0.0014	0.0008	0.0011	0.0018	0.0008	0.0019
Per 1	PR		0.0019	0.0035	0.0015	0.0010	0.0014	0.0014	0.0009	0.0012
	ESW		0.0208	0.0113	0.0000	0.0000	0.0000	0.0000	0.0104	0.0000
	All	0.9656		0.0012	0.0010	0.0006	0.0008	0.0013	0.0010	0.0025
Per 2	PR	0.9778		0.0017	0.0013	0.0005	0.0012	0.0012	0.0014	0.0034
	ESW	0.9250		0.0000	0.0016	0.0018	0.0007	0.0020	0.0000	0.0018
	All	0.9386	0.9213		0.0013	0.0011	0.0010	0.0012	0.0010	0.0018
Per 3	PR	0.9315	0.9205		0.0015	0.0012	0.0012	0.0012	0.0012	0.0023
	ESW	1.0000	0.9451		0.0027	0.0016	0.0003	0.0010	0.0014	0.0010
	All	0.8665	0.9431	0.9573		0.0020	0.0016	0.0019	0.0015	0.0013
Per 4	PR	0.8724	0.9401	0.9608		0.0024	0.0020	0.0021	0.0021	0.0019
	ESW	0.9629	0.9793	1.0000		0.0023	0.0013	0.0015	0.0000	0.0000
	All	0.8138	0.8901	0.9548	0.9627		0.0025	0.0024	0.0015	0.0013
Per 5	PR	0.8116	0.8973	0.9676	0.9628		0.0033	0.0030	0.0022	0.0017
	ESW	0.9665	0.9710	0.9624	1.0000		0.0014	0.0016	0.0003	0.0013
	All	0.7485	0.8384	0.8852	0.9354	0.9569		0.0025	0.0017	0.0009
Per 6	PR	0.7511	0.8379	0.8861	0.9285	0.9538		0.0032	0.0025	0.0016
	ESW	0.9415	0.9624	0.9860	0.9866	1.0000		0.0008	0.0007	0.0000
	All	0.7350	0.7587	0.8016	0.8707	0.9328	0.9426		0.0017	0.0016
Per 7	PR	0.7363	0.7652	0.8049	0.8821	0.9418	0.9547		0.0023	0.0022
	ESW	0.9593	0.9507	0.9659	0.9668	0.9695	1.0000		0.0006	0.0021
	All	0.8307	0.8452	0.8659	0.8817	0.9018	0.9438	0.9696		0.0010
Per 8	PR	0.8206	0.8274	0.8419	0.8635	0.9081	0.9393	0.9729		0.0014
	ESW	0.9785	0.9770	0.9912	1.0000	0.9668	1.0000	0.9475		0.0018
Per 9	All	0.9465	0.9218	0.9411	0.9189	0.9356	0.9133	0.9400	0.9667	
	PR	0.9419	0.9114	0.9331	0.9358	0.9295	0.9186	0.9327	0.9821	
	ESW	0.9687	0.9681	1.0000	0.9739	1.0000	0.9837	0.9779	1.0000	

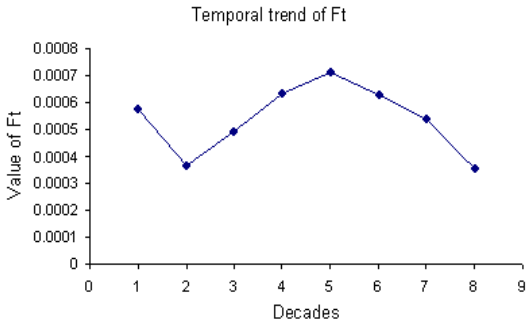
Table 1**:** Matrix of Lasker's coefficient (above the diagonal) and Euclidean distance (below the diagonal) between 9 cohorts for natives (PR), immigrants (ESW) and for the entire group (All).

		Group 1		Group 2		Group 3	
Period	**Marriages**	**Total**	**%**	**Total**	**%**	**Total**	**%**
1946	118	73	61.9	43	36.4	2	1.70
1951-1956	84	72	85.7	12	14.3	0	0.00
1961-1966	150	99	66.0	47	31.3	4	2.70
1971-1976	136	83	61.0	43	31.6	10	7.40
1981-1986	248	131	52.8	103	41.5	14	5.60
1991-1996	217	150	69.1	64	29.5	3	1.40
2001-2006	104	46	44.2	49	47.1	9	8.70
1946-2006	1057	654	61.9	361	34.2	42	3.90

Table 2**:** Temporal trends in the frequency of endogamous and exogamous marriages: group 1 – both mates born in Përmet, group 2 – one born in Përmet and the other born elsewhere, group 3 – both born elsewhere.

4. Temporal trends in marital structure: Results on temporal changes in marital structure could give us the possibility to estimate the trends of genetic isolation of the population under investigation. Changes in the pattern of endogamy and exogamy for 10 year periods of time are shown in Table 2. The values of endogamy are especially high in the period of time 1951 -1991 when the mobility of the population was under total control according to the laws of the past regime. Lower values of endogamy are obtained after the year 1996 following the social and political changes when the population mobility was much higher compared to the past regime. However, for the entire period of investigation still high values of endogamy (61.9%) were observed. This is related to the isolation due to the geographic location of Përmeti and the limited mobility of the population compared to the other regions of the country.

a)



b)

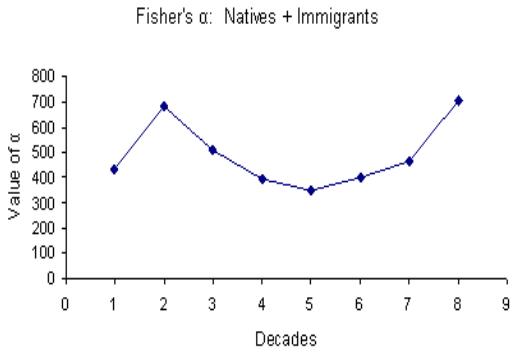


Figure 5: a) Temporal trend of total inbreeding Ft; b) Temporal trend of Fisher's α.

5. Temporal trends in Isonymy and Inbreeding: Inbreeding has been widely considered as a good indicator of genetic isolation. Fisher's α, that is an indicator of the richness of surnames, could give additional information on the trends of the genetic isolation. Trends of total inbreeding (Ft) estimated by isonymy and Fisher's α are shown in Figures 5a and 5b.

In Figure 5a is shown the continuous increase of the inbreeding level in the population due to both geographic location and the gradual decrease of the population's mobility, especially during the period of last political regime. The highest value was reached in the 5-th decade. It follows a slight decrease of the inbreeding level starting from the 6-th decade, due to the economic development of the region in this period of the past regime. However, in the last decade the decrease of inbreeding was higher because of free movement of the population following the political and social changes.

As the result of the recent migration any change of Ft values (Fig. 5a) is followed by an inverted change trend of Fisher's α value. For comparison reasons, these changes are shown in Figure 5b.

CONCLUSIONS

The population of Përmeti experienced interesting changes. However, it presents still a high degree of genetic isolation due to the pattern of migration or structure of marriages. The value of endogamous marriages is still high (61,9%) despite the social and economic factors that have influenced the gradual decrease of inbreeding coefficient measured by isonymy.

BIBLIOGRAPHY

1.Barrai I, Barbujani G, Beretta G. et al. (1987) Surnames in Ferrara: Distribution, isonymy and levels of inbreeding. Anaquo;. Hum. Biol. 14:415-423.

2.Barrai I, FormicaG, Barale R, Beretta M (1989) Isonymy and migration distance. Ann. Hum. Genet. 53:249-262.

3.Caravelle G, Tasso M (1999) An analysis of the spatial distribution of surnames in the Lecco area (Lombardy, Italy). Am. J. Hum. Biol. 11:305-315.

4.Cavalli-Sforza LL., Edwards A.W.F. (1967) Phylogenetic analysis models and estimation procedures. American Journal of Human Genetics, 19: 233-257.

5.Crow JF (1980): The estimation of inbreeding from isonymy. Human Biol., 52: 1-4.

6.Crow JF, Mange AP (1965) Measurements of inbreeding from the frequency of marriages between persons of the same surnames. Eugenics Quarterly ,12: 199-2003.

7.Chakraborty R., Barton SA, Ferrel RE et al. (1989) Ethnicity determination by surnames among the Aymara of Chile and Bolivia. Hum. Biol. 61:159-177.

8.Rodriguez-Larralde A, Formica G, Scapoli C, Beretta M, Mamolini E, Barrai I (1993) Microevolution in Perugia: isonymy 1890-1990. An. Hum. Biol., 20: 261-274.

9.Zei G, Matessi RG, Siri E, Moroni A, Cavalli-Sforza LL (1983) Surnames in Sardinia, I. Fit of frequency distributions for neutral alleles and genetic population structure. Annals of Human Genetics, 47: 329-352.