Journal of Natural Sciences Research www.iiste.org
ISSN 2224-3186 (Paper) ISSN 2225-0921 (Online)
Vol.4, No.10, 2014

The Genetic Structure of the Population of the Koplik Municipality Calculated by the Isonomy

Ermira Hoxhaj [1]* Zyri Bajrami [1,2]
[1] Departament of Biology – Chemistry,University of Shkodra,Shkodra,Albania.
*E-mail:hoxhajermira@mail.com
[2] Departament of Biology,Universiyt of Tirana,Bulevardi Zogu I,Tirana,Albania
E-mail:zyribajrami@yahoo.com

Abstract
The main focus of this study is the human population of the Municipality of Koplik,in Albania. The city of Koplik is situated 17 km at the North of Shkodra city,in Albania. The genetic structure and connections among the inhabitants of Koplik, have been studied according to the distribution of the surnames of 18831 individuals, identified by 652 surnames, which resulted from the compilation of the genealogical trees.The coefficient of kinship was calculated, because of the isonomy of the case, along with evaluations of the α of Fisher as indicator of the surname variety, evaluations of Fst inbreeding coefficients and evaluations of the Karlin-McGregor`s coefficient. The surname distribution has been calculated based on the logarithm of the surname number and the number of the times such surname is identified.
All these parameters have been calculated by means of the data extracted from the genealogical trees in three different methods, based on the surnames inherited from the father, from the mother and a another one, based on casual distribution of the surnames of the same persons,but without considering the genealogical relationship.All these methods show a high value of kinship, but when the three methods are compared, the results have obvious differences. For instance, the value Fisher α was calculated by the Pedigree from the father α=98.27, the mother α=61.478, and from the casual distribution α=110.86. The value of the a of Fisher (α= 98.27) for this population, indicates high rates of kinship among the population of Koplik. A lower value of the α coefficient from the mother side, can be easily explained with the fact that the kiniship in the marriages starting from the mother, is much higher than the one started from the father.
Key Words: Koplik,Isonomy,surname,genealogical tree,inbreeding, genetic structure

1.Introduction

The study of the structure, dynamics and evolutional tendencies of the human population, nowadays has become an important object of the research work.
The surnames are unique bio-cultural features, used to study the structure and the microevolution in the human populations. In our society, the surname is inherited in a patrilinear way (from the father individual). This observation has led to the thesis that the surnames are monophyletic and can be considered as genes related to the chromosome Y and therefore, their heritage is analogue to the heritage of the polymorphic genes related to the chromosome Y (Jobling M A.et.al.2001). The unbiological nature of the distribution of the chromosome Y is expected to be independent from the fertility and mortality, and is expected to be subjected to the neutral model of the evolution, (Kimura 1983). In compliance with this model, it is expected that the logarithm frequency of the surnames follows a linear trend (Barrai et al 1987) .In spite of the fact that the latest studies have revealed a deviation from the expected linear trend (Barrai et al 1987). The deviations from the expected linear trend, come as a result of the excessive repetition of surnames (Barrai et al 2002). Most of the studies have been based on big samples, which comprise a whole nation or region, whereas concerning small populations like ours, such studies have been seldom conducted.
In this light, it would be interesting to examine the neutral distribution in the small countries, since the demographic events and the marriages, bring about major changes in the blood relations, by influencing the structure of surnames. Under these conditions, the surname neutrality might not be expected.A study in this field in Albania is done by,Scapoli,Barbujani,Mamolini,Sandri,Carrieri,Mikerenzi,,Xhina and Rodrigues-Larralde(2013) This is the first study of this kind in Koplik city.

2.Methods of the Study

The material for this study has been taken from interviews with autochthon inhabitants and also from newcomers from the villages around the Municipality of Koplik. The genealogic trees for 18831 inhabitants of the Municipality of Koplik were built by using the computer program SQL-SERVER, by exploiting the advantages of genealogic codification for each inhabitant. The value of casual isonomy has been calculated through three different methods, using the outcome data of the genealogic trees. Afterwards, based on the value of the casual isonomy, the values of the other parameters have been calculated as well. In the beginning the isonomy value for 18831 inhabitants has been calculated based on the distribution of surnames deriving from the father identified with 652 surnames, and on the surname of the mother of 15659 individuals identified with 592 surnames. There

are less individuals from the mother side because the father surname has been defined for the difference, but it was not possible to define the surname of the mother.
The third way of isonomy calculation was grounded on the casual distribution of surnames of the same persons , without considering the genealogic relations at all.
18831 individuals have been idientified with 590 surnames. There are less surnames, than in the case of the calculation through genealogic trees, because there are 15 different surnames which identify 5120 individuals, who have only four descendants. There are also 82 surnames which are repeated from the pedigrees, which bear no genealogic relation but which are identified with 28 different surnames based on the method of casual distribution. Based on the distribution of surnames among the population, the casual isonomy within the population with regard to the same surname has been calculated. It is very likely that they belong to the same family. Thus, isonomy is an indicator of biological connection.
Different authors use different methods to calculate isonomy. In this study, the casual isonomy was calculated by the formula:
Iíí = ∑ Pik 2 – 1/ Ní based on the surname of the men, referred to Rodrigues-Larralde et al, (1993) where Pik – the relative frequence of each surname and Ní the number of individuals of the population.
A high value of the isonomy is probable where we are dealing with few surnames while a low value of the isonomy is probable where the number of the surnames is big. The low value of the isonomy indicates recent migration.
Iíí = ∑ qik 2 – 1/ Ní based on the surname of the women
Iíí=∑píqí based on the surname of men and women where Pí is the relative frequency of the surnames in paternal line and q is the frequency of the same surname in maternal line.
The inbreeding component within the population, which shows the value of kinship is calculate with the formula: F st = Iíí/4 where Iíí is the casual isonomy as referred to Relethford (1998). Inbreeding which is based on the values of Isonomy allows conclusions concerning the rate of genetic differentiation.
The variety of surnames has been calculated through the coefficient Fisher`s alpha which is used to evaluate surname diversity and as a consequence genetic diversity, is an indicator of genetic isolation and is calculated with the formula:
a= 1/ Iíí where I íí is the casual isonomy evaluated according to Barrai et al. (1992)
A high value of α indicates high diversity of surnames and high level of migration, whereas its low value, indicates genetic isolation and high inbreeding.
Karlin-MacGregor`s, which is used to evaluate the migration rate v, is calculated with the formula:
V=a/ (Ní+a) Zei et al. (1983)
The high value of the coefficient v indicates a high genetic diversity, because human migration affects genetic diversity for through the migration, new alleles can join in or leave the population. Thus, the coefficient v is in direct proportion with the coefficient alpha.
The surname distribution for all the population has been studied through the model Log-log (8). A chart has been built, where in the X values there is log2, S where S is the number of surnames and in the Y values there is log2 K, where K is the number of the cases a surname is identified.

3.The Resultats and the Discussin

The parameters of isonomy in the Municipality of Koplik

3.1The comparison among the three different methods used for the calculation of these parameters

The distribution of the number of the surnames of Koplik used in this analysis, and the main parameters deriving from the theory of isonomy calculated with the above-mentioned methods, have been presented in Table 1.
Tab 1. No. of individuals, no. of surnames S, Fishers` α, Fst, and Karlin-McGregor v, calculated with the three above-mentioned methods in Koplik.

The methods	Ni	Ns	Iii	α	Fst	v
By genealogical trees						
• **From the father's surrname**	**18831**	**652**	**0.010176**	**98.27**	**0.0025**	**0.0025**
• **From the mother'surname**	**15659**	**592**	**0.016266**	**61.478**	**0.004**	**0.00391**
By casual distribution of surnames	**18831**	**590**	**0.009072**	**110.86**	**0.00227**	**0.00585**

Fisher`s alpha is used to evaluate the diversity of the surnames and as a consequence, the genetic difference among the population. If we compare the value of α with the three calculated ways, the one calculated from the mother side (α=61.478), is comparatively smaller and indicates a higher genetic closeness among individuals based on mother`s side, than on father`s side (α=92.27). This is quite reasonable to a certain extent, because not so much attention is paid to avoid the kinship resulting from the mother side, if compared to the father`s side. This is mostly due to the fact, that female line relations are not as known as the male ones. The value of α=110.86, calculated from the casual distribution of the surnames of the same individuals, (18831), identified with the same surnames (590), without considering the genealogic relations resulting from the genealogic trees,

changes in value especially considering the one calculated from the mother and the father as well.
The reasons of these changes can be as follows:
When the casual distribution is calculated, there are fewer surnames if compared to the tree calculation because 5120 individuals identified with 15 different surnames, have only 4 descendants if we follow the genealogic connections deriving from the genealogic trees.
Tab.2 The surnames which have the same descendant extracted from the genealogic trees, but which actually have different surnames.

Common ancestor	Actual surname	Nr of inhabitants
Culi		
	• Curri	694
	• Culaj	
	• Fili	
Bajrami		
	• Bajramaj	1749
	• Rrustemaj	
	• Ramçaj	
Hoti		
	• Hoti	787
	• Isufi	
Mehi		
	• Bajraktari	1890
	• Semanaj	
	• Avdyli	
	• Kurti	
	• Hysaj	
	• Duli	
	• Shabani	
Total		5120

There are also 82 surnames which are repeated, that if we trace the connections in the genealogic trees, they have no connection at all, which means they have different origins. Only 28 surnames are identified with the method of casual distribution.

Tab 3 Origins of surnames which appear more than once in different familes.

Nr	Surname	Origin1	Nr.	Origin 2	Nr.	Origin 3	Nr	Origin 4	Nr	Origin 5	Nr	Total
1	Curri	Koplik	117	Lepurosh	65	Zagor	13	Bratosh	8	Goraj	6	207
2	Hasaj	Koplik	24	Repisht	216	Flake	128	Dodç	24	Rrjoll	10	402
3	Hysaj	Koplik	271	Polvar	107	Reç	20	Bajz	20	Grizhe	17	435
4	Ujkaj	Koplik	78	Sterbeq	181	Potgor	52	Vajush	42	Bajz	36	389
5	Bjaramaj	Koplik	246	Grizhe	8	Laç	4	Velipoj	8			266
6	Hoxhaj	Kaldrun	189	Repisht	50	Vajush	43	Zagor	14			296
7	Kurti	Koplik	275	Polvar	175	Linaj	29	Grizhe	19			498
8	Smakaj	Koplik	230	Lopç	55	Lohe	42	Grizhe	73			400
9	Çokaj	Dober	20	Koplik	246	Grud	72					338
10	Isufaj	Koplik	166	Linaj	73	Vorf	17					256
11	Kukaj	Koplik	219	Grizhe	100	Marshej	100					419
12	Rrukaj	Kamic	85	Marshej	40	Sterbeq	20					145
13	Zekaj	Koplik	149	Kaldrun	175	Grizhe	73					397
14	Zenelaj	Koplik	182	Kcar	34	Grizhe	34					250
15	Aliaj	Flake	59	Lopç	19							78
16	Bajraktari	Koplik	208	Grizhe	11							219
17	Çelaj	Reç	92	Postrib	4							96
18	Gjetaj	Koplik	30	Jubic	21							51
19	Hasanaj	Kaldrun	101	Grud	46							147
20	Hoti	Koplik	283	Polvar	26							309
21	Lekaj	Jubic	152	Reç	52							204
22	Luli	Kukes	3	Jubic	1							4
23	Metaj	Reç	24	Marshej	46							70
24	Memaj	Zagor	62	Reç	48							110
25	Mustafaj	Kçar	54	Kaldrun	38							92
26	Ramaj	Koplik	199	Reç	31							230
27	Rexhaj	Koplik	82	Grizhe	27							109
28	Zaraj	Kaldrun	153	Moçovil	77							230

3.2The distribution of surnames and the most frequent surnames in Koplik

There are 18831 individuals which go by 590 surnames based on the casual distribution.There are 49 surnames which have a higher frequency than 100, held by 10296 individuals, and that means 56% of the individuals. There are 329 surnames, which have an absolute frequency smaller than 10 and 5% of the population is identified with them. There are 133 surnames with an absolute frequency of 1. All the surnames belong to the Muslim faith besides the surname “Smakaj-Llazan” which belongs to the Catholic faith.

Tab 4. The surnames with an absolute frequencu higher than 100

Nr	Mb	f	Nr	Mb	f	Nr	Mb	f
1	Kurtaj	498	18	Culaj	214	34	Hebaj	138
2	Hysaj	435	19	Curri	207	35	Selgjekaj	133
3	Kukaj	419	20	Hasmujaj	205	36	Smajlaj	133
4	Hasaj	402	21	Lekaj	204	37	Delaj	124
5	Smakaj	400	22	Mataj	202	38	Bushati	123
6	Zekaj	397	23	Kurtulaj	201	39	Balaj	122
7	Ujkaj	389	24	Lulaj	198	40	Bercaj	121
8	Çokaj	338	25	Bajrakurtaj	186	41	Semanaj	121
9	Hoti	309	26	Dulaj	181	42	Gjonaj	119
10	Hoxhaj	296	27	Ramçaj	154	43	Çukaj	114
11	Bajramaj	266	28	Hasanaj	147	44	Ramekaj	112
12	Isufaj	256	29	Haxhiaj	147	45	Memaj	110
13	Delishaj	250	30	Rrukaj	145	46	Mustalaj	109
14	Zenelaj	250	31	Brojaj	143	47	Rexhaj	109
15	Ramaj	230	32	Senaj	143	48	Avdyli	106
16	Zaraj	230	33	Rrustemaj	139	49	Dukaj	102
17	Bajraktari	219						

3.3Distribution log 2 K-Log 2 S

Fig 1 and 3 shows the distribution of the logarithm of the number of surnames and of the number of the times they appear for husbands and wives in Koplik. Fig 2 and 4 show the graphic presentation between Lof 2 S-Log 2 S of husbands and wives in Koplik.

The logarithmic distribution of men's surnames in Koplik

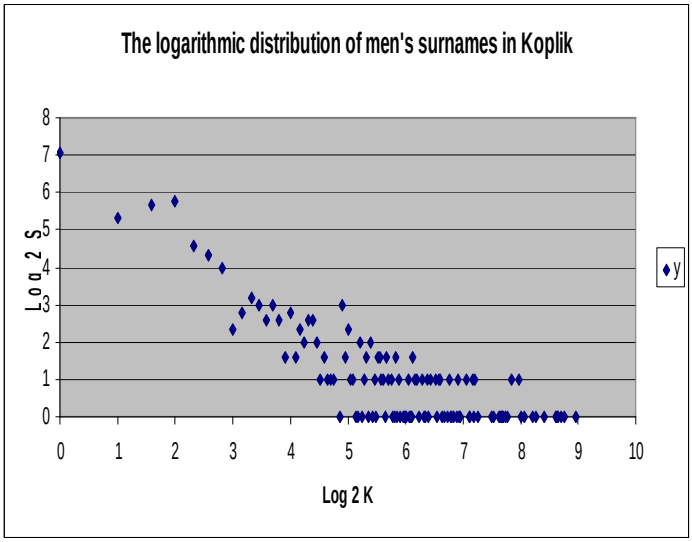

Fig 1.The logarithmic distribution of men'surname S and the number of the surname which appear K.

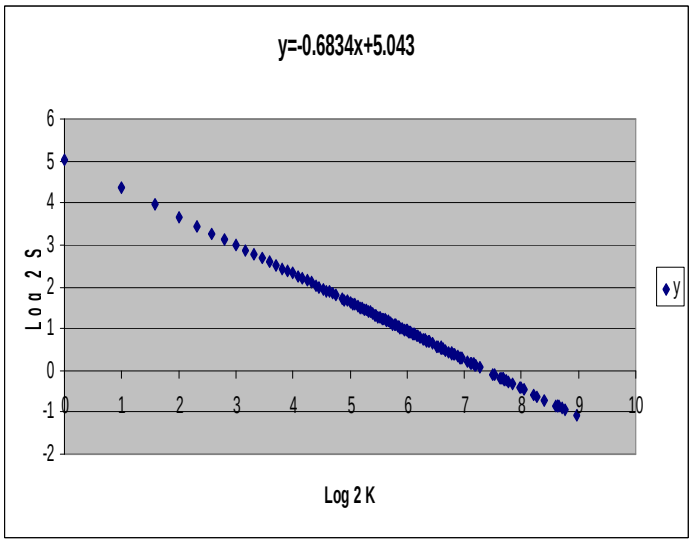

Fig 2.Graphic presentation of the regression between Log 2 S-Log 2 K for men.

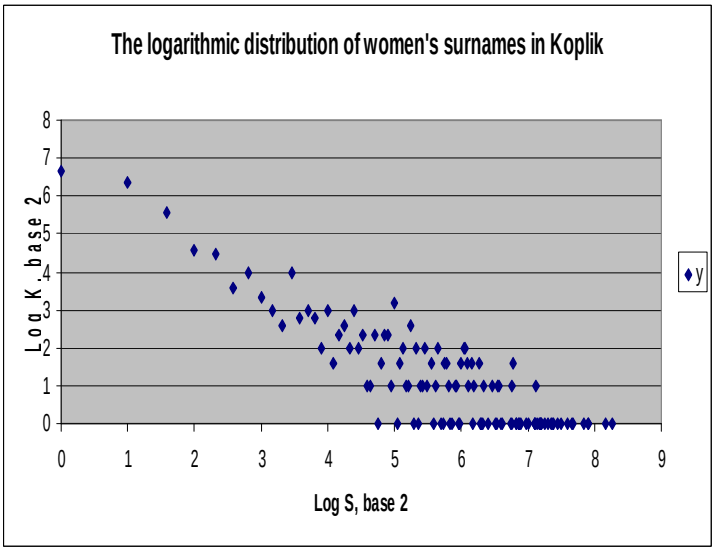

Fig 3.The logarithmic distribution of women'surname S and the number of the surname which appear K

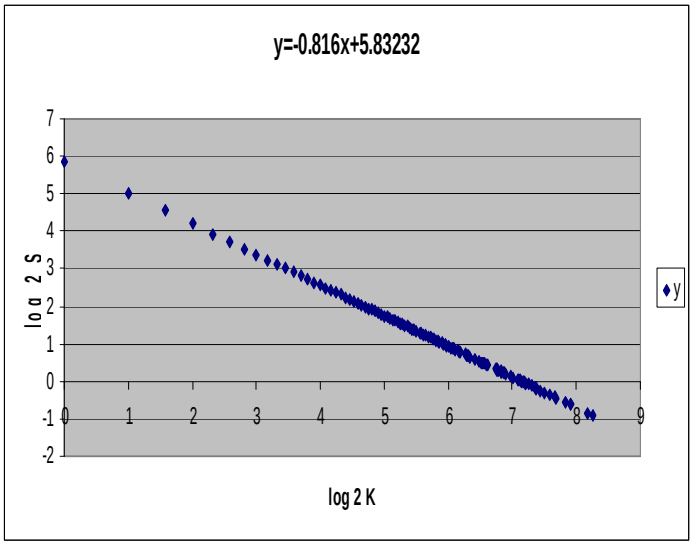

Fig 4.Graphic presentation of the regression between Log2 S-Log2 K for women

The graphic regression between these parameters clearly shows the linear trend in the distribution of surnames of husbands and wifes. This indicates a compliance with the model of the neutral alleles (Kimura 1983).
The values of Fishe alpha, casual isonomy, inbreeding coefficient and the Karlin-MacGregor`s v which result from this paper on Koplik, compared to the results of the study done by Scapoli,Barbujani,Mamolini,Sandri,Carrieri,Mikerenzi,,Xhina and Rodrigues-Larralde(2013) for Tirana and the District of Shkoder.

Tab 5. The results of this study compared to the result of the study for Tirana, the District of Shkoder and other cities.

Rajoni	S	Iii	α	Fst	v
Tiranë	18415	0.00095	1048	0.000239	0.00167
Shkodra-Region	7350	0.00152	658	0.00381	0.00275
Shkodra -City	6642	0.00157	637	0.000394	0.00355
Pukë	892	0.0081	123	0.0020036	0.00562
M.Madhe	1235	0.00384	260	0.000965	0.00671
Koplik	590	0.009072	110.86	0.00227	0.00585

If we compare the values of casual isonomy with the values of this coefficient calculated by Mikerenzi for the other cities of the District of Shkoder, it is seen that the isonomy is higher in Koplik, than in the other cities. This means, that Koplik is genetically more isolated that Shkodra or Puka.
The value α of Fisher in Koplik is lower, than the other cities, which shows a lower diversity of surnames, low migration, and as a consequence, higher genetic isolation than the other cities.
The inbreeding coefficient is higher if compared to the other countries and this is a feature of the small countries. The reason is maybe the cultural impediments concerning marriages,which in Koplik ore more oriented towards endogamy.

References

Jobling,M,A (2001):In the name of the father.Surnames and genetics.Trends in Genetics nr.17.355-357
Kimura M (1983)The neutral Theory of Molekular Evolution.Cambidge University Press,Cambridge. Barrai I,Barbujani G,Beretta M,Maestri I,Russo A,Formica G,Pinto-Cisternas(1987)Surnames in Ferrara:distribution,isonymy,and levels of inbreeding.Ann Hum Biol,14:415-423.
Barrai I,Rorigues-Larralde A,Manni F,et al(2002).Isonomy and isolation by distance in Netherland.Hum Biol,74:263-283
Barrai I,Rodrigues- Larralde A,Manni F,Ruggiero V,Tartari D,Scapoli C,(2003).Isolation by distance in Belgium.Ann Hum Genet,68:1-16

BARRAI,I., C. SCAPOLI, M. BERETTA, C. NESTI, A. RODRIGUEZ-LARRALDE, (1996) Isonymy and the genetics structure of Sëitzerland I.The distribution of surnames. Ann. Hum. Biol., 23(1996)431-455.

Crow,JF,Manage AP(1965).Measures of Inbreeding of the frequencies of marriages between persons of the same surnames.Eugenics Quarterly,12:199-203.

3-Lasker Gw.1991.Cultural Factor in the Geographic Distribution of Personal Names.Pseudogenetic Analysis of First Names Used to Estimate the Cultural Component of Coefficients of Relationship by Isonomy.Human Biology 63(2):197-202.

Rodriguez-Larralde A,Gonzales Martin A,Scapoli C,Barrai I(2003).The names of Spain:A study of the isonymy structure of Spain.

Rodriguez –Larralde,Barrai I,Alfonzo JC(1993) Isonymy structure of four Venezuelan states.Ann Hum Biol ,20:131-145

Relethford JH (1988) Estimation of kinship and genetics surname from surnames.Hum boil,60:475-492.

Barrai,I.,C. Scapoli,M.Beretta, C.Nesti, A.Rodriguez-Larralde,(1996) Isonymy and the genetics structure of Switzerland I.The distribution of surnames. Ann. Hum. Biol., 23(1996)431-455.

Brownberg RA,Dipierri JF,Alfaro EL,Barrai I,Rodriguez-Larralde A,Castilla EE,Colonna V,Rodriguez –Arroyo G,and Bailliet G.2009.Isonymy Structure of Buenos Aires City.Human Biology 81(4):447-461.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Recent conferences: http://www.iiste.org/conference/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar