J. Biol. Res. 2006 n. I - Vol. LXXXI - Rubbettino - Soveria Mannelli

Surnames as genetic markers in population studies

E. Lucchetti, P. Pizzetti, S. Deiasio

Dept. Genetics Anthropology Evolution, University of Parma

KEY WORDS: surnames, biodemography, population structure, migration, population similarity

Abstract

Surnames, information of cultural nature used to identify people and family groups, include simultaneously the genealogical and hereditary facet; the territorial (both geographic and socio-cultural) and temporal dimension. They represent an useful source for either specialistic or interdisciplinary analyses, and for the various disciplines which have the study of human groups as core issue of their researches (sociology, linguistics, history, demography, anthropology, genetics, etc.).

Family names, their typology and distribution can provide information about the characteristics of human groups, and specifically about their internal structure, degree of isolation or openness, relationships with other populations, and about their evolution.

When transmitted through the male line, the surname behaves as a genetic marker, thereby allowing the use of all the methods peculiar of population genetics and biology. This paper illustrates, by means of some examples, possible applications and opportunities offered by the use of surnames in the studies of human populations, along with the limits that they implicitly present.

Surnames, their typologies, number and frequencies, provide a source of valuable information for the characteristics of a population, its internal structure, its degree of isolation or mobility, its relationships with other populations, as well as the reconstruction of its history and evolution. (Lasker, 1985; De Felice, 1980; Lucchetti, 1990, 2004).

In a large number of populations, surnames are handed down by heredity, from father to son, and are preserved through the male line: surnames are comparable to genetic markers, and can be investigated using the numerous specific analytical methods applied in biological and genetic population studies (Zei et al, 1983; Lucchetti et al., 2004). Surnames, cultural entities transmitted as genetic markers, encompass the family and hereditary dimension, the territorial (geographical, socio-cultural or economic) dimension and the temporal dimension; these are all precious sources for interdisciplinary studies (Brunet et al., 2001), involving the contribution of various and diverse areas of research (linguistics, sociology, history, demography, genetics, anthropology, ...) where the common denominator is the study of human groups.

Surnames and characterisation of populations

Any item of information is valuable for the description and characterisation of a population, if it enables us to infer specific characteristics, both independently and in relation to other populations. Surnames are a source of information that meet this requirement: the typologies of surnames, their numbers and frequency distribution are appropriate for the description of the structural characteristics of a population; the comparison between surname distribution in various populations, for the investigation on the relations between these and on their evolution in time. This information is useful both from the qualitative point of view, if we take into account the typology of surname forms, and from the quantitative point of view, if we analyse their number and frequency distribution.

In the various areas, depending on the language or dialect spoken, and on the cultural heritage in general, surnames are distributed differentially, thus enabling us to characterise and distinguish one population or group from another. The descriptive and typifying potential of surnames can be observed not only in large areas, but also within limited areas, such as adjacent villages and hamlets (Lucchetti et al., 1989).

Let us examine, for instance, the distribution over the Italian peninsula of the two surnames with the highest frequency: *Rossi* and *Russo*. Their diffusion over the national

territory is clearly differentiated. The surname *Rossi* is widespread particularly in the central northern regions, whereas it is infrequent in the southern and insular regions. The surname *Russo* is widespread above all in the southern regions, with high frequency in Sicily, whereas it is infrequent in Sardinia. In the North of Italy, the frequency of the surname *Russo* is consistent in the more highly developed regions: the presence of the surname in these regions can be ascribed to the extensive migrations from the South to the North of Italy in the decades following the Second World War. This basic observation is a good example of how surnames can lend themselves to surveys on migrations: the surname *Russo*, clearly of southern origin, not only indicates the areas in northern Italy privileged by the migration flow from the South, but it can also shed light on the entity of the migratory phenomenon. Another example concerns the distribution of some surnames associated with the art of iron-working, which take on different forms according to the region in which they originated: *Fabbri, Ferrari, Ferrero, Magnani, Forgione, Frau*. The surname *Ferrari*, the third most widespread surname over the Italian territory, is diffused mainly in the northern regions; *Fabbri* is present above all in the Romagna area, in Tuscany and in the upper part of the Marche region; *Magnani* is found principally in the Emilia-Romagna region, as well as in the upper part of Tuscany; *Ferrero* is a typical Piedmont surname; *Forgione* can be found in the southern regions, especially in the Campania region; *Frau* is typical of Sardinia.

As emerges from studies conducted by various authors, Sardinia sets itself apart from the rest of Italy, by way of the surname typologies present on the island. (Lucchetti et al., 1996; Lisa et al., 2001); this also seems to correspond with our data regarding the genetic characteristics (Cavalli Sforza et al., 1994). Surnames of Sardinian origin are uncommon in Sicily; among the regions in continental Italy they can be traced in Lombardy and in Piedmont; they have a minor – though not insignificant – incidence in other regions of central northern Italy, whereas they are poorly represented in all the southern regions; their distribution shows the direction and the intensity of the migrations from Sardinia towards the other Italian regions.

The characterising potential of surnames is remarkable even in the case of smaller communities, such as municipalities, or even a group of houses within the same municipality. In the upper Parma Valley – an area comprising three municipalities, subdivided in turn into 40 hamlets or parishes – the surnames entered in the birth registers from the year 1800 to the first decades of 1900 reveal a certain amount of differentiation in their diffusion and distribution, both among the municipalities and among the villages (Soliani et al., 1988).

The study of several or all surnames clearly facilitates the characterization of populations or groups. The lists of the 15 most common surnames found in each of the provinces in the Emilia-Romagna region are visibly different and typical for each province. The surname *Rossi* is a common denominator in every province, but it is the only surname featuring among the 15 most frequent surnames in all provinces. The lists of the 15 most common surnames confirm that the provinces of the Romagna area differ from those in the Emilia area; on the whole, the most frequent surname typologies enable us to distinguish one province clearly from another.

Surnames as genetic markers

The interest in surnames on the part of anthropologists, biologists and human population geneticists stems from the fact that, in many societies, surnames are handed down by heredity, from father to son: they are transmitted and preserved from generation to generation, in the same way to that of the Y chromosome and the characters enclosed in it (Zei et al., 1983).

The use of surnames as genetic markers has certain advantages with respect to authentic genetic markers, even if it also inevitably entails a few incongruities which must be taken into consideration (Lucchetti et al., 1990, 2004).

The main advantage is associated with the high degree of variability – or elevated polymorphism – provided by surnames, which is undoubtedly far superior to that provided by genetic markers, even in the presence of studies conducted with the aid of the most up-to-date molecular biology techniques (Darlu, 2001). As has been observed, a high degree of polymorphism allows us to characterise populations with a certain amount of accuracy and to differentiate them even in limited contexts and situations, in which authentic genetic markers lack the power to identify differences and to discriminate among them.

Surnames are very easily obtainable, and at a low cost; the process does not require either the length of time or the highly specialised technology involved in the determination of biological characteristics. Lists which can be easily transferred from one computer to another and which contain surnames entered under several headings – residents, births, heads of families, subscribers, etc. – facilitate our task. The search can be easily extended to the whole of the population.

Surnames are reminiscent of their origins. Numerous surnames of southern Italian origin can very often be identified among the populations of northern Italy: even in subsequent generations, the number of descendants – regardless of their birthplace – can be ascertained and, albeit in probabilistic terms, their contribution to the development of the local communities can be assessed.

A further advantage lies in the possibility of extending the search to past generations: with the aid of ecclesiastical or civil archives, the process can be conducted with relative effortlessness and surnames can be traced at least a few centuries back in time (Lucchetti, 1990). Surnames constitute a sort of historical DNA of populations (Darlu, 2001): their evolution in time can be monitored without interruption from generation to generation.

The correlation between surnames and genetic markers is certainly not a categorical one; the use of surnames to represent genetic markers occurs by analogy: the

assumptions underlying this analogy must always be scrupulously verified (Lucchetti, 1990; Zei, 2001). There are theoretical problems involved in the polyphyletic origin of surnames; there are practical problems caused by situations that are likely to upset the surname-biological heredity correlation, such as illegitimacy, adoptions, transcription errors and changes connected with the evolution of the language. These factors certainly introduce variations on an individual level but, as a general rule, they do not affect the basic characterising potential of the surname system, especially when the survey is carried out on the population as a whole.

If, on the one hand, surnames lend themselves to studies on past populations, the temporal depth that can be explored is considerably reduced compared with biological evolutionary times. At a given historical moment, surnames are conditioned by synchronic cultural and linguistic factors and they superimpose themselves, regardless of the biological characteristics previously acquired by the same populations over a long period of time; surnames prove to be more appropriate in microevolutionary studies and for the description and analysis of the events taking place in the last few centuries.

For their use in the study of human populations, as a general rule, surnames have an extremely advantageous *quality/price ratio*, (Darlu, 2001); they are characterised by their availability, by the low cost of the search process, which can be extended to past generations, and above all by their rich store of information, high characterising potential and high degree of polymorphism.

Examples and applications

Thanks to their hereditary characteristics, surnames are used as 'data' in many biodemographic studies; in other studies, they are used to describe the structure of a population or the relations among various populations, regardless of gene analogies. They can be used in connection with geographical distribution or with the mobility of their bearers (territorial dimension), or with their distribution over the various generations (temporal dimension) or both, depending on the objectives of the study, the sources available and the search methods, or the analytical methods adopted.

a. Evaluation of consanguinity and analysis of marriage patterns and behaviour (choice of husband or wife)

From a historical point of view, consanguinity represents one of the primary areas of research in the field of biodemography: it is important from the social, historical and economic point of view, as well as from the genetic and medical point of view (Conterio et al., 1974; Cavalli Sforza et al., 2004).

The first person to use surnames as an instrument to evaluate consanguinity is considered to be George Darwin (1875), who made an attempt to estimate consanguinity using the number of isonymic marriages. Crow and Mange (1965) observed that, in all types of consanguineous marriages, the relation between the degree of consanguinity F and the probability of isonymy I between husband and wife is constant, with $F/1 = 1/4$. Consanguinity F can be estimated considering the frequency I of isonymic marriages.

This method is straightforward and easy to apply, but it yields consanguinity estimations that are usually higher than those deriving from the direct consultation of parish registers (Cavalli Sforza et al., 1971, 2004; Zei 2001). In actual fact, the method entails assumptions that do not always fully correspond to reality. It implies a monophyletic surname origin, that is to say that the bearers of the same surname are in some way related, albeit remotely; that obviously surname transmission is linked to biological transmission, without taking into consideration deviations due to illegitimacy, adoptions, transcription errors; that the various types of consanguineous marriages of the same degree (for example, marriages between first cousins can be divided into four types, depending on the sex of the intermediate ancestors) all take place with the same probability, whereas it is a well-known fact that certain types of marriage take place more frequently than others (Cavalli Sforza et al., 1971, 2004).

Marriage behaviour patterns and couple formation can be assessed on the basis of the surname combinations of both partners. Lasker and Kaplan (1985) have proposed the Rp (Repeated pairs) index, which evaluates the number of repetitions in surname combinations in the couples.

Both in consanguinity estimation and in the study of couple formation, a random component and a non-random component can be distinguished. (Chakraborty, 1985); the former is related to the expected value in the event of total panmixia, the latter shows the occurrence of phenomena (alliance strategies, social or geographical subdivisions) which favour, or hinder, various types of marriage. Applications of couple formation studies demonstrate that the non-random component is negatively linked to population size and to exogamy (Mascie-Taylor et al., 1987); its positive excess seems, however, to benefit positively from population size when we consider large areas or large-sized populations, revealing the presence of socio-cultural or economic stratifications. (Relethford, 1992).

b. Analysis of the internal structure of a population

The number of surnames, related to the size of the population, enables us to infer structural characteristics of the population. The difference is obvious in extreme cases; on the one hand, a population with a limited number of surnames, each of which has a high degree of repetitiveness or, on the other hand, a population with a high number of different surnames, each of which has low repetitiveness. In the first case, an isolated population, with no inflow of surnames (and thus individuals) from outside, can be logically hypothesised; the continuation of the inhabitants is entrusted to a handful of local families, the number of surnames tends to become reduced, with a consequent

increase in the frequency of the remaining surnames. In the second case, however, the population is an open community: the greater wealth of surnames presumes an inflow of people from outside (immigration). The probability that two people taken at random may bear the same surname is indicated by Lasker (1985) as an index of isonymy within the population.
The S/N ratio (number of surnames/population size) is a useful index; it can, however, take on similar values in different situations: for example, it does not discriminate if the surnames (S) are all characterised by similar repetitiveness or if some of these have a high degree of repetitiveness and others have low or even unitary frequency. The analysis of frequency distribution according to class enables us to discriminate in situations of this type. Barrai et al. (1989) have demonstrated that, by using a log-log transformation, the points tend to arrange themselves in a linear way. The curve provides us with an index to evaluate the immigration impact on the population.
By analogy with biological indices measuring the wealth of the species and the biodiversity present in a given environment (the number of species and the number of members per species, or the number of alleles for a given character and the frequency distribution) the surname wealth of the population and the inflow rate of new surnames can be evaluated. Fisher's · index (1943), obtained from the assessment of the abundance of species in a given ecosystem, measures the surname wealth of a population; Karlin McGregor's ni index (1967) evaluates the frequency with which new allelic forms appear in a population: by analogy with surnames, it assesses the inflow rate of new surnames, by evaluating the immigration impact.
Applications on historical data from the Parma Valley (Yasuda et al., 1974) from Sardinia (Zei et al., 1983) and on recent data regarding the Italian regions (Piazza et al., 1987) and Sicily (Scapoli et al., 1997) and on data regarding French surnames (Darlu and Ruffie, 1992) confirm the validity of the surname utilisation.

c. Analysis of the inter-relations among populations

The surnames from two populations are more and more similar to each other, in terms of typology and frequency, the more the two populations originate from a common cultural area and/or the more numerous the migration exchanges between the two. Clearly, if there are shared exchanges between the two the farther back in time these relations extend, the greater the similarity among surnames; inversely, if two populations have very little contact and there are no mutual exchanges, they are most likely to have a low number of surnames in common. From the resemblances among surnames, it is therefore possible to make inferences on migrations, as well as on the genetic exchanges that have taken place in the course of time (Darlu, 2001). The probability of two individuals, taken at random from two different populations, being bearers of the same surname, affords an estimation of the similarity of the populations considered (Lasker, 1985); the index can be standardised to take on value 1 when the comparison is made within the population itself and value 0 when the two populations compared have no surnames in common (Chen et al., 1983). This approach involves methodologies taken from population genetics, where the similarity between populations is defined as the probability of two alleles of a given locus, taken from each of the genic pools under comparison, being identical.
Similarity matrices are usually interpreted using "tree" representations or topological representations (Lalouel, 1980; Lucchetti et al., 2004).
Similarities among surnames produce representations in which the geographical locations of the populations are often reflected, but these can also show discrepancies, due to the presence of easily accessible ways of communication, or vice-versa of obstacles to mobility, of intense migratory flowes responding to particular events and situations, and also due to the sharing of a common cultural identity. For instance, the application of the method to the distribution of surnames in the 20 Italian regions produces a topological representation showing a neat correspondence with the geographical location of the regions (Lucchetti et al., 1996). Sardinia is a separate case, sharply differentiated from the rest of Italy, also from the point of view of the frequency of genetic markers; the Aosta Valley, the Trentine region, regions in which numerous surnames of (respectively) French and German origin can be found, are also differentiated, albeit to a lesser extent.
The effect of the South-North migrations that have characterised Italy in the last decades is clearly visible: Lombardy, Piedmont and Liguria, the three regions where the migratory routes from the South have been mainly directed, are detached from the other northern regions and are situated in an intermediate position, closer to the southern regions. When we separate the main cities (Milan, Turin and Genoa), where the highest concentration of immigrants is to be found, from the rest of their respective regions, the intermediate position closest to the southern regions is still occupied by the cities, whereas the rest of the region is situated in the cluster grouping together the other northern regions.
The application of the method to the popultions situated on both slopes of the Apennine ridge (Lucchetti et al., 1989), belonging to the provinces of Parma, Piacenza, La Spezia, Genoa and Massa Carrara, demonstrates, alongside the important role played by the geographical proximity of the municipalities (corresponding also to their political unity in the centuries prior to the unification of Italy), the dividing effect of the ridge. Similarly, a study of the municipal populations living on the banks of Lake Como and those enclosed within the two southern branches of the lake has demonstrated that the lake does not constitute a separating element, but an easily accessible way of communication used in exchanges among the populations living on the opposite banks of the lake (Lucchetti et al., 2004). However, the mountain chain situated in the southern part, enclosed between the two branches of the lake (one towards Como and the other towards Lecco) acts as a divider: the populations adjacent

to the ridge prove to be more similar to the populations living on the lake banks, even to the populations living on the opposite bank, as opposed to the bordering populations on the other side of the ridge. Thus, exchanges among populations are more likely to take place by crossing the lake, rather than crossing the mountain ridge, which is more impervious, especially in the winter season, even if the distance is shorter.
The effect of geographical proximity, expressed in concrete terms by their residence in the same valley, can be clearly observed in the study of Alpine populations (Caravello et al., 2002, 2004). The fact that they are situated in the same valley is more influential than their affiliation, in terms of dialect and tradition, to a common cultural identity. Generally speaking, these communities show strong resemblances when they are geographically close, but when they belong to distinct geographical areas, the sharing of a common language or cultural identity is less influential than geographical proximity in determining the intensity of shared exchanges.

Discussion

Surnames undoubtedly constitute a rich source of information for population studies, especially considering the relative facility with which they can be collected; they can be studied from several perspectives and respond to the interests of numerous areas of research. In biodemographic studies, surnames are analysed as genetic markers or, at least, in relation to their inseparability from the persons who bear them. The assumptions underlying this analogy are not always totally verifiable: therefore considerable caution is required in the assessment of the results. However, it must be specified that, even in the case of polyphyletism, the same surname tends to confirm the descent from a common cultural matrix: therefore, the relations among populations have a certain validity, even if not in the strictly biological sense.
The diversity of sources from which surnames can be obtained is another issue which needs to be taken into careful consideration, because different sources can provide different types of information and can be suitable for the description and study of different phenomena. Surnames can be obtained from vital registrations (birth, death and marriage registers): in this case, the range of time over which the search is conducted is important, but also the type of source used. For example, surnames of newborn children and surnames of deads can describe the same population in different ways. In particular, the marriage registers found in communities where marriage is traditionally celebrated in the bride's parish are representative of the parish in terms of the brides' surnames, but not in terms of the bridegrooms' surnames. Census describe a population at a given moment in time: the surnames can be extracted from all residents or from particular categories of residents. Surnames of heads of family are often used, which can be obtained from readily available lists: in this case, there is no internal sub-classification according to age and there is disparity in terms of sex. By using the complete list of residents, these drawbacks can be overcome and the surname structure can be analysed according to age; however, a complete study involves a higher degree of repetitiveness which – depending on the type of analysis we aim to conduct – may be a disturbing element but, on the other hand, it may be a source of additional information.
A final issue briefly concerns the possibility of recording Christian names: as opposed to surnames, Christian names do not constitute a hereditary character, thus attracting the interest of anthropologists, biologists and geneticists. Names are markers influenced by cultural changes, for the very reason that they are chosen autonomously, without any predetermined conventions, by the child's parents. The repetition of certain names (when, for instance, a newborn child is given the same name as his grandfather) is a sign of attachment to traditions, including marriage and reproduction traditions. Considerable variations in names, either for the sake of following fashions or particular events, making a break with the past, distinguishing oneself from other people or from past generations, or establishing values and ideals to be reaffirmed or introduced, show that people's cultural perspectives are changing. Cultural changes are the presupposition or evidence of changes in behaviour: breaking away from tradition, even in the case of names given to children, is tantamount to breaking away from a world with which the community previously identified itself. Also the use of names afford new possibilities of integrated analyses, in which behaviour patterns, cultural perspectives, social, but also biological, consequences, can be studied in combination.

References

Barrai I, Formica G, Barale R, Beretta M. 1989. Isonymy and migration distance. Annals of Human Biology, 17, 1: 7-18.

Brunet G, Darlu P, Zei G., Bideau A. 2001. Introduction générale. In: Brunet Guy, Darlu P., Zei G. eds. Le Patronyme histoire, anthropologie, société. CNRS Editions: 1-3.

Caravello GU, Tasso M, Lucchetti E. 2002.. Distribution of surnames and identities in the Cimbro-Mòcheno communities of Italy. Anthropologischer Anzeiger 60, 3: 241-253.

Caravello G.U, Tasso M, Lucchetti E., Pizzetti P. 2004. Distribution of surnames and linguistic-cultural identities of the Slovenian and German minorities of northeastern Italy. Anthropologischer Anzeiger 62, 2: 185-202.

Cavalli-Sforza LL and Bodmer WF. 1971. The genetics of human populations. San Francisco (Ca.), WH Freeman, XVI + 965 p.

Cavalli Sforza LL., Menozzi P, Piazza A. 1994. The Histhory and Geography of Human Genes. Princeton University Press 541+518p.

Cavalli Sforza LL., Moroni A, Zei G. 2004. Consanguinity, Inbreeding, and Genetic Drift in Italy. Princeton Paperbacks 314p.

Chen K, Cavalli-Sforza LL 1983. Surnames in Taiwan: interpretation based on geography and history, Human Biology, 55, 2: 367-374.

Chakraborty R. 1985. A note on the calculation of random RP and its sampling variance, Human Biolology, 57, 4: 713-777.

Conterio F, Moroni A. 1974. Démographie et génétique, Population et famille, 31, 1: 64-115.

Crow JF, Mange AP. 1965. Measurement of inbreeding from the frequency of marriages between persons of the same surname, Eugenics Quarterly, 12, 4: 199-203.

Darlu P. and Ruffie J. 1992. L'immigration dans les départements français étudiée par la méthode des patronymes, Population, 47, 3: 719-734.

Darlu P. 2001. Espace, temps, migration. In: Brunet G., Darlu P., Zei G. eds. Le Patronyme histoire, anthropologie, société. CNRS Editions: 155-158.

Darwin G. 1875. Marriages between first cousins in England and their effects, Journal of the Statistical Society, 38, 2, p. 153-184.

De Felice E. 1980. I cognomi italiani. Il Mulino, Bologna.

Degioanni A, Lisa A, Zei G. and Darlu P. 1996. Patronymes italiens et migrations italiennes en France entre 1891 et 1940, Population, 51, 6: 1153-1180.

Fisher Ronald A, Corbet Steven A. and Williams CB. 1943. The relation between the number of species and the number of individuals in a random sample of animal population, Journal of Animal Ecology, 12, 1: 42-58.

Karlin S, McGregor J. 1967. The number of mutant forms maintained in a population, in: Fifth Berkeley Symposium of Mathematics, Statistics and Probabilities. Vol. 4: Biology and problems of health: 415-438. Berkeley, University of California Press, xix + 934 p.

Lalouel JM. 1980. Distance analysis and multidimensional scaling, in: James H. Mielke and Michael H. Crawford (eds.), Current Developments in Anthropological Genetics. New York, Plenum: 209-250.

Lasker GW. 1985. Surnames and genetic structure. Cambridge, Cambridge University Press, p148.

Lasker GW, Kaplan BA. 1985. Surmames and genetic structure: repetition of the same pairs of names of married couples. A measure of subdivision of the population, Human Biology, 57: 431-440.

Lisa A, Fiorani O, Siri E, Zei G. 2001. Migrations récentes in Italie et distribution géographiques des patronymes. In: Brunet G., Darlu P., Zei G. eds. Le Patronyme histoire, anthropologie, société. CNRS Editions: 231-244.

Lucchetti E, Battisti C N, Ghisolfi G, Soliani L. 1989. Delimitation and aggregation in human populations examined by surname Structure. Int. J. Anthrop. 5, 1: 49-62.

Lucchetti E. 1990. Uso dei cognomi in analisi sulla struttura ed evoluzione di popolazioni umane. Antropologia Contemporanea, 13, 2/3: 107-119.

Lucchetti E, Siri E, Soliani L. 1996. Internal migrations and relationships between italian regional populations: an analysis based on surname distributions. Rivista di Antropologia, 74: 171-185.

Lucchetti E, Soliani L. 2004. Génétique et Démographie. In: Caselli G., Vallin J.S, Wunsch G. eds. Démographie: analyse et synthèse VI Population et Société: 517 – 554.

Mascie-Taylor NC, Lasker GW and Boyce AJ. 1987. Repetition of the same surnames in different marriages as an indication of the structure of the population of Sanday Island, Orkney Islands, Human Biology, 59, 1: 97-102.

Piazza A, Rendine S, Zei G, Moroni A. and Cavalli-Sforza LL. 1987. Migration rates of human populations from surname distributions, Nature, 329: 714-716.

Relethford JH. 1992. Analysis of marital structure in Massachusetts using repeating pairs of surnames, Human Biology, 64, 1: 25-33.

Scapoli C, Rodriguez-Larralde A, Beretta Ma, Nesti C, Lucchetti E. and Barrai I. 1997. Correlations between isonymy parameters. International Journal of Anthropology, 12, 1: 17-37.

Soliani L.o, Anelli A.o, Anghinetti W., Siri E. 1988. Analisi con cognomi e nomi delle variazioni demografiche, biologiche, culturali di comunità e popolazioni. Genus, 92, 3-4: 55-88.

Yasuda N, Cavalli-Sforza LL, Skolnick M, Moroni A. 1974. The evolution of surnames: an analysis of their distribution and extinction, Theoretical Population Biology, 5, 1: 123-142.

Zei G, Guglielmino C, Siri E, Moroni A, Cavalli-Sforza LL. 1983. Surnames as neutral alleles: Observations in Sardinia. Human Biology, 55: 357-365.

Zei G, Guglielmino C, Siri E, Moroni A, Cavalli-Sforza LL. 1983. Surnames in Sardinia. I. Fit of frequency distribution for neutral alleles and genetic population structure, Annals of Human Genetics, 47, 4: 329-352.

Zei G. 2001. Alliance, parenté, structure génétique. In: Brunet G., Darlu P., Zei G. eds. Le Patronyme histoire, anthropologie, société. CNRS Editions: 265- 268.