



# Surnames in south-eastern France: structure of the rural population during the 19th century through isonymy

Célia Kamel, Bérengère Saliba-Serre, Marie-Hélène Lizée, Michel Signoli,  
Caroline Costedoat

## ► To cite this version:

Célia Kamel, Bérengère Saliba-Serre, Marie-Hélène Lizée, Michel Signoli, Caroline Costedoat. Surnames in south-eastern France: structure of the rural population during the 19th century through isonymy. *Journal of Biosocial Science*, inPress, 10.1017/S0021932021000699 . hal-03521816

**HAL Id: hal-03521816**

**<https://hal.science/hal-03521816>**


Submitted on 15 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

# Surnames in south-eastern France: structure of the rural population during the 19th century through isonymy

Célia Kamel<sup>1\*</sup> , Bérengère Saliba-Serre<sup>1</sup>, Marie-Hélène Lizée<sup>2</sup>, Michel Signoli<sup>1</sup> and Caroline Costedoat<sup>1</sup>

<sup>1</sup>Aix Marseille Univ, CNRS, EFS, ADES, Marseille, France and <sup>2</sup>RECOVER-INRAE, Aix-en-Provence, France

\*Corresponding author. Email: [celiakamel1@gmail.com](mailto:celiakamel1@gmail.com)

(Received 26 May 2021; revised 22 November 2021; accepted 22 November 2021)

## Abstract

An analysis of the distribution of surnames through time and space allows us to understand the structure of human groups, their exchanges or even their possible isolation. The French population has already been studied through surnames and it has been shown that the Sud-Provence-Alpes-Côte d'Azur region differed from the rest of France in both the 20th and 21st centuries (Mourrieras *et al.*, 1995; Scapoli *et al.*, 2005). The objective of this study was to understand the population evolution and particularities of the Sud-Provence-Alpes-Côte d'Azur region through an analysis of the distribution of surnames over an earlier period: the 19th century. For this work, 806,069 birth records from 521 communes between 1810 and 1890 were recorded and a total of 23,340 surnames were collected. The estimation of various isonymic parameters has allowed a description of this corpus never exploited before. In order to appreciate the population evolution, the data set was divided into three periods of 25 years. The canton was the geographical unit of this study, and similarities and differences between each of them were evaluated using Lasker distances, which allow the construction of dendrograms. A positive and significant correlation ( $p < 0.0001$ ) was found between Lasker distances and geographical distances using the Mantel test. The lowest inbreeding estimates were found in the Durance Valley. Migration, estimated from the  $\nu$ -index of Karlin and McGregor (1967), showed higher values in the south-western quarter of the region. The decrease in  $R_{st}$  values across the three periods is consistent with a homogenization of the patronymic between the cantons. This three-period approach showed a population evolution influenced by linguistic, cultural, historical and migratory phenomena since the Middle Ages, disrupted by the socioeconomic changes of the 19th century.

**Keywords:** Human population biology; Migration and social mobility; Demography

## Introduction

The distribution of surnames is not random across time and space (Darlu *et al.*, 2012). Naming practices reflect social norms and ethno-cultural customs. These practices have been developed and transmitted over generations and are often maintained after migrations (Mateos *et al.*, 2011). Naming is thus the product of cultural, linguistic and legislative processes that have enabled its normalization and systematic transmission (Cheshire, 2014). As a source of social, cultural and historical information, the study of surnames makes it possible to trace the history of populations over the centuries. The accessibility of surnames at various times and their reliability as a historical-demographic source (Boattini *et al.*, 2006) have made them very interesting for the reconstruction of human genetic structures. The analogy between genes and surnames, due to their patrilineal transmission since the Middle Ages in most European countries, has allowed the emergence of numerous methodologies (Mourrieras *et al.*, 1995).

As early as 1965, Crow and Mange developed the theory of isonymy in order to estimate the relationship between consanguinity and family names (Crow & Mange, 1965). Originally, their calculations were based on marriage registers because they believed that isonymous unions (same name) would often correspond to a marriage between first cousins (Crow & Mange, 1965). Since then, the use of isonymy theory has been extended to larger databases and allows for the estimation of diversity or isolation of human groups.

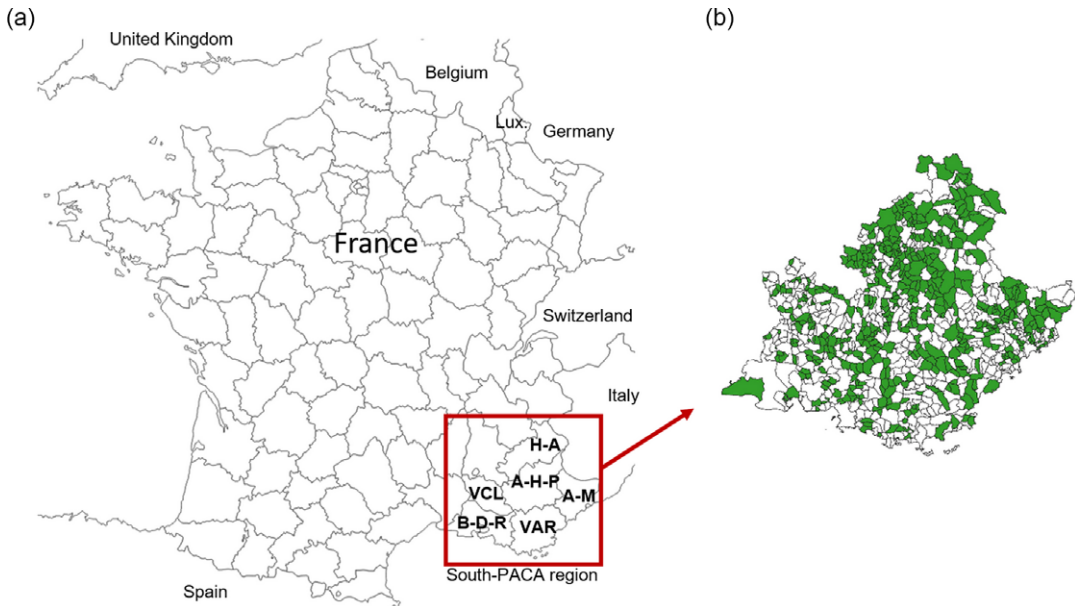
Generally, studies analysing the structure of the French population in the 20th century using surnames use the 'surname file' designed in the 1970s by INSEE (Mourrieras *et al.*, 1995; Darlu *et al.*, 1997; Vernay, 2001; Darlu & Oyharçabal, 2006; Gibert *et al.*, 2012). This file provides the names of natives, geographically referenced at the commune level, from the period 1891 to 1990. The studies concerning the 21st century are based on the telephone directory (Scapoli *et al.*, 2005). These studies have shown that south-eastern France differed from the rest of the country in the 20th century because the Massif Central seemed to act as a migration barrier (Mourrieras *et al.*, 1995). For the 21st century, Scapoli *et al.* (2005) showed that the French population is structured according to different cultural and linguistic phenomena.

This study looked at different levels of administrative division, among which the departmental analysis of France showed the isolation of the Alpes-Maritimes department (located in the south-east) from the French territory. The aim of the study, therefore, was to analyse the distribution of surnames of an earlier period, the 19th century, in order to grasp the micro-evolution and the particularities of the Sud-Provence-Alpes-Côte d'Azur region, and then to identify the isonymic structures of the population studied in order to evaluate the impact of geographical remoteness and historical, migratory, cultural and linguistic phenomena on these structures. There is a fairly extensive bibliography on the migration history of the southern region. These sources emphasize the importance of the rural exodus of the 20th century, which began in the middle of the 19th century and demographically strengthened the cities (Marseille, Nice, Toulon, etc.). At the same time, there are reflections on the extent of migratory movements in the old French village society (Dupâquier, 2002). This study contributes to the debate on rural society and migration by providing elements of a response through the use of an original and never-before-exploited database. Furthermore, the results of this new analysis, coupled with the analogy between isonymic and genetic structures, suggest several types of applications in the field of public health, which are developed in terms of perspectives.

## Methods

### *Study area and population*

The Sud-Provence-Alpes-Côte d'Azur region, located in the south-east of France, is one of the thirteen French regions and covers 31,400 km<sup>2</sup>. To the west, the Rhône River forms its regional boundary with Occitania, while to the north it borders on the Auvergne-Rhône-Alpes region. The Italian regions of Liguria and Piedmont border the region to the east, and to the south it is bordered by the Mediterranean Sea. Both the borders and the name of the Sud-Provence-Alpes-Côte d'Azur region have varied over the centuries. In 1972, the Provence-Alpes-Côte d'Azur region was administratively created by the French state and designated by the acronym PACA, but since 2018 it has been called the Sud-Provence-Alpes-Côte d'Azur region by the regional council. This territory includes six departments, three of which are coastal: Bouches-du-Rhône, Alpes-Maritimes and Var; the other three are located inland: Vaucluse, Alpes-de-Haute-Provence and Hautes-Alpes. These departments originate from the former provinces of Provence, the Dauphiné for the Hautes-Alpes, the Comtat Venaissin for the Vaucluse and the Comté de Nice for the Alpes-Maritimes. The southern region offers a wide variety of contrasting landscapes (plains, hills, mountains and coastline), whose cohesion is ensured by the routes traced by the Durance and the Rhône (Bouvier, 1979; Temime, 1997).



**Figure 1.** (a) Map of the French departments with the six departments of the Sud-Provence-Alpes-Côte d’Azur region framed: A-H-P: Alpes-de-Haute-Provence; H-A: Hautes-Alpes; A-M: Alpes-Maritimes; B-D-R: Bouches-du-Rhône; Var; VCL: Vaucluse. (b) Map showing the 521 communes studied (in green).

According to the National Institute of Statistics and Economic Studies (INSEE), the region had more than 5 million inhabitants in 2018 (Michaïlesco & Mora, 2020), whereas in the 1851 census its population did not exceed 1.5 million (Le Mée, 1999). The demographic growth of the region has therefore been spectacular due to socioeconomic changes and migration (Gastaut, 2009). It is now a highly urbanized region; 90% of its inhabitants are city dwellers, whereas at the beginning of the 19th century, 60% of its population lived from agriculture (Gastaut, 2009). The distribution of the population is now very unbalanced. The technical, industrial and tourist progress of the 20th century, located on the Rhône axis and on the coast (Marseille, Nice, Toulon), has attracted migrants to these new centres of attraction, which today accommodate three-quarters of the region’s inhabitants (Gastaut, 2009). As a result, the interior of the peasant region, which traditionally subsisted on olive trees, vines and wheat, was economically weakened and depopulated during the rural exodus (Temime, 1997).

### Data

A total of 806,069 birth certificates were recorded from 521 communes with a population of less than 15,000 in the 19th century (Figure 1). The municipalities surveyed were therefore mainly part of the rural area of that period. These records cover the period from 1810 to 1890. As the INSEE ‘surname files’ are made up of birth certificates, they were also chosen in order to facilitate comparison with studies using this corpus. A total of 23,340 surnames were collected. The geographical unit of the study was the canton in order to minimize the differences in the number of births between these cantons since in some communes the number of births could be very low. The communes were therefore grouped according to the canton to which they belonged in 2014, which represented 81 of the 126 cantons in the southern region. In order to assess the population evolution on a micro-evolutionary scale, the dataset was split into three periods of 25 years, each representing one generation (Darlu *et al.*, 1997): P1=1810–1835; P2=1836–1861; P3=1862–1890.

In France, as in most European countries, the patronymic system appeared during the Middle Ages, in around the 10th century. The use of patronymics was progressively regulated by the political power, and its transmission to descendants and its generalization were put in place from the 14th century onwards (Fordant, 1999). The legal standardization of patronymic names took place in several stages. After the French Revolution, the parish registers, drawn up at the parish where baptisms, marriages and burials were recorded, became the registers of civil status, drawn up at the town hall where births, marriages and deaths are recorded (Darlu *et al.*, 2012). From 1794 onwards, only the Conseil d'Etat could authorize a change of name and it was forbidden to use names other than those recorded in the civil status register. The appearance of the Family Record Book in 1870 durably consolidated the spelling and patrilineal transmission of surnames in France (Fordant, 1999).

Birth certificates were chosen for this study because they are doubly relevant. Firstly, the INSEE 'surname files' are made up of birth certificates and, as a result, working on the same type of certificate made it possible to compare and put into perspective the many existing studies that have used this corpus. Secondly, it is true that the use of surnames raises the issue of the inclusion or exclusion of women's surnames in the studies as this can change the results (Cheshire, 2014). However, it has been shown that variation is more likely at small geographical scales (Bowden *et al.*, 2008). At larger scales some studies have shown that excluding women's surnames does not have a significant impact. This was shown, for example, by Winney *et al.* (2012) in a 19th century UK-wide study. According to these authors, this is probably due to marriage remaining relatively local during the 19th century in Britain. Thus, the use of birth certificates allows the surnames of women and men before their marriages to be obtained. This was done for the present study because in the 19th century France was under a patriarchal system and marriage led to the change of women's names.

### Statistical analyses

All calculations and statistical analyses were carried out using RStudio® software Version 1.3.1093. First, the evolution of the surname stock was described using descriptive statistics: the number of surnames that remained or disappeared over the three periods and the proportion of individuals they represented. Particular interest was given to surnames registered at a single birth. These, conceptualized by Chareille and Darlu (2013) under the name of 'hapax', allow for a very interesting interpretation. Indeed, they can be considered as the last representative of a rare surname, the witness of a migration or a transcription error (Chareille & Darlu, 2013).

The evolution of differences between cantons was assessed using the  $R_{st}$  index, an estimator used in population genetics (Relethford, 1988; Boattini *et al.*, 2006). For this study it was adapted to the context of surnames and allowed for a measurement of the homogeneity between the subdivisions of the total population studied – here, the cantons (Relethford, 1988; Boattini *et al.*, 2006). This index was calculated for all three periods. A decrease of  $R_{st}$  values across a period reflects an increasing exchange of surnames between the cantons. Conversely, if there are fewer exchanges of surnames the  $R_{st}$  value will increase (Boattini *et al.*, 2006). The  $R_{st}$  index was calculated using the  $F_{st}$  function of the Biodem package of the R® software, developed by A. Boattini and colleagues.

A study of the distribution of surnames through the theory of isonymy is the main method used in patronymic studies. The evaluation of the isolation of each canton was made possible by the isonymy index  $I_{ii}$ , which is calculated as follows:

$$I_{ii} = \sum_k p_{ki}^2,$$

where  $p_{ki}$  is the frequency of patronymic  $k$  in canton  $i$ , and the total is equal to the sum of the squared frequencies of each patronymic  $k$  in canton  $i$  (Darlu *et al.*, 1997). This mathematical formula calculates the probability of randomly drawing two identical surnames in the same canton;

the higher this index, the more isolated the canton (Darlu & Oyharçabal, 2006; Carrieri *et al.*, 2020). Isonymy was calculated using the *uri* function of the Biodem package of the R<sup>®</sup> software.

In order to estimate migration on the basis of surnames, the Karlin–McGregor  $\nu$ -index was calculated in accordance with Carrieri *et al.* (2020). This index was developed by Karlin and McGregor (1967) based on the theory of neutral mutation behaviour in finite populations of constant size, in conjunction with the patrilineal transmission of patronyms that can be considered as alleles (Rossi, 2013). This index estimates the ‘mutation rate’ (Darlu & Ruffié, 1992). Since true surname mutations are relatively rare, a high value of  $\nu$  will indicate an influx of new surnames and thus a migratory phenomenon. This index is calculated for each canton  $i$  as follows:

$$\nu = \alpha / (N - \alpha),$$

where  $N$  is the number of people per canton and  $\alpha = 1/I_{ii}$  (Dipierri *et al.*, 2005; Carrieri *et al.*, 2020).

Patronymic similarities and differences between cantons taken two-by-two were evaluated using Lasker distances. In 1977, Lasker proposed an extension of Crow and Mange’s (1965) formula to use isonymy as a measure of relatedness between two populations (Lasker, 1977). Lasker’s index  $R_{ij}$  calculates, from surnames, the probability that two human groups  $i$  and  $j$  have genes in common (Rossi, 2013):

$$R_{ij} = \frac{\sum_k S_{ki} S_{kj}}{2 \sum_i S_i \sum_j S_j},$$

where  $S_{ki}$  is the number of occurrences of surname  $k$  in canton  $i$  and  $S_{kj}$  is the number of occurrences of the same surname in canton  $j$  (Colantonio *et al.*, 2007; Roman-Busto & Fuster, 2015). The logarithmic transformation of the Lasker index yields the Lasker distance between the two cantons:  $L_{ij} = -\ln(2R_{ij})$  (Rodriguez-Larralde *et al.*, 1998; Colantonio *et al.*, 2007). The  $R_{ij}$  index is multiplied by two because the Lasker index is equal to twice the isonymy between two groups (Barrai *et al.*, 1999, 2000; Dipierri *et al.*, 2005). In order to detect isolation by distance, the correlation between Lasker distances and geographical distances was evaluated using the Mantel test, which uses the Monte Carlo method with 9999 permutations. This test was performed with the *mantel.randtest* function of the R<sup>®</sup> software package, *ade4*. The geographical distances between each pair of townships were calculated from the straight-line distance between the township capitals using QGIS<sup>®</sup> Desktop 3.4.7. A hierarchical ascending classification with Ward’s method was performed using the Lasker distance matrices to create a partition of the townships for each period. The *hclust* function, with the ‘ward.D2’ method of the R<sup>®</sup> software package *stats* was used to obtain the dendrograms. The analysis of the inertia jump of the dendrograms allowed the most homogeneous grouping of the townships possible. The use of cartography, with the software QGIS<sup>®</sup> Desktop 3.4.7, made it possible to visualize the groups obtained in geographical space.

## Results and Discussion

### Evolution of the surname stock in Sud-Provence-Alpes-Côte d’Azur

The distribution of surnames within each canton and the parameters calculated from the theory of isonymy are given in Table 1. The 25 most frequent names were studied; ten of them (Martin, Bernard, Robert, Michel, Roux, Bertrand, Bonnet, Blanc, Faure and Girard) were found in the list of 25 most frequent names at the national level of INSEE’s files of patronymic names for the period from 1891 to 1915. The surname Martin is the most frequent in France today and has been since the 19th century (Darlu *et al.*, 1997). But in the corpus studied, the family name Blanc was the most frequent during the three periods with 10,662 occurrences, followed by Roux with 8831 occurrences and Martin (8755 times), which was in third position. They are followed by the surnames Michel (8715), Giraud (6565), Arnaud (5702) and Bernard (4409). These 25 names



**Table 1.** Distribution of study sample surnames and parameters derived from isonymy theory by time period and Sud-Provence-Alpes-Côte d’Azur cantons

Canton	Period 1: 1810–1835				Period 2: 1836–1861				Period 3: 1862–1890			
	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>
Allauch	1821	229	0.02373	0.0237	2336	311	0.01777	0.02469	2717	536	0.00721	0.05382
Antibes-3	1068	180	0.02093	0.04685	855	182	0.01451	0.08768	798	238	0.00978	0.14697
Apt	8107	603	0.00714	0.0176	7737	688	0.00631	0.02092	6462	770	0.00585	0.02719
Argentiere-la-Bessee	3257	188	0.02533	0.01228	3068	242	0.02165	0.0153	3182	361	0.01814	0.01764
Arles	655	163	0.01894	0.08772	687	167	0.01953	0.08058	728	183	0.01533	0.09847
Barcelonnette	4567	366	0.0114	0.01959	4346	359	0.01095	0.02148	4438	401	0.00967	0.02388
Beausoleil	3016	514	0.01022	0.03356	2761	428	0.01081	0.0347	4742	1321	0.0048	0.04598
Berre-l’etang	2697	356	0.01382	0.02758	2487	516	0.00824	0.05135	1934	476	0.00801	0.06907
Bollene	1878	348	0.00659	0.08804	2617	416	0.00665	0.06098	2261	412	0.00604	0.07909
Briancon-1	5945	355	0.00995	0.01721	5192	322	0.01133	0.01731	4867	313	0.01249	0.01674
Briancon-2	1684	188	0.02311	0.02638	1392	186	0.01965	0.03796	1213	174	0.01859	0.04643
Brignoles	1638	193	0.02274	0.0276	1274	181	0.02084	0.03914	942	178	0.01858	0.06062
Cagnes-sur-Mer-2	622	137	0.04451	0.03748	437	130	0.03678	0.06637	304	112	0.02211	0.17485
Carpentras	413	93	0.02083	0.13155	479	100	0.02376	0.09638	412	106	0.02463	0.10934
Castellane	7871	839	0.00918	0.01405	6623	533	0.00976	0.01573	5528	520	0.0087	0.02126
Chateaurenard	3341	444	0.00788	0.03951	4178	534	0.00678	0.03661	3939	545	0.00587	0.04527
Chtx-Arnoux-St-Auban	917	141	0.02032	0.05672	777	160	0.01508	0.09332	718	173	0.01318	0.11826
Cheval-Blanc	3326	444	0.00699	0.04498	3197	500	0.00657	0.05006	3009	528	0.00591	0.05966
Chorges	4467	351	0.00904	0.02542	4158	379	0.00879	0.02814	4325	460	0.00744	0.03211
Ciotat-la	4603	745	0.00627	0.03595	5503	1152	0.00461	0.04108	9853	2250	0.00182	0.05931
Contes	9332	725	0.00673	0.01621	10996	728	0.00676	0.01365	8567	749	0.00714	0.01664
Crau-la	1373	214	0.01723	0.04416	1246	249	0.01442	0.05896	1145	316	0.00915	0.10554
Digne-les-Bains-1	2933	222	0.01567	0.02225	2405	229	0.0147	0.02912	2211	280	0.0142	0.03291
Digne-les-Bains-2	975	157	0.02336	0.04594	704	127	0.01951	0.07854	842	158	0.01687	0.07574
Draguignan	876	152	0.02439	0.04912	776	173	0.01965	0.0702	900	271	0.00897	0.14142
Embrun	5510	391	0.01096	0.01685	5094	384	0.01058	0.01892	4849	419	0.01033	0.02039
Flayosc	13462	823	0.00633	0.01189	10433	870	0.00618	0.01577	8641	966	0.00569	0.02079
Forcalquier	2769	316	0.01152	0.03238	2466	365	0.0109	0.03868	2328	458	0.00867	0.05218
Gardanne	4297	458	0.00973	0.02453	4085	560	0.00779	0.03247	4225	841	0.00499	0.04984
Garde-la	1868	375	0.00813	0.07058	1626	362	0.00741	0.09058	1954	589	0.00457	0.12616
Gareoult	2226	288	0.01407	0.03301	1904	262	0.01274	0.04303	1447	298	0.01123	0.06562
Grasse-1	1686	152	0.0199	0.03073	1273	161	0.01921	0.04266	1062	175	0.01431	0.07046
Grasse-2	557	110	0.02531	0.07638	477	119	0.02151	0.10804	495	204	0.01177	0.20722
Guillestre	8417	660	0.00692	0.01747	6944	629	0.00732	0.02009	6472	660	0.00778	0.02029

(Continued)

Table 1. (Continued)

Canton	Period 1: 1810–1835				Period 2: 1836–1861				Period 3: 1862–1890			
	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>
Istres	861	291	0.00969	0.13625	1839	548	0.00612	0.09766	1942	628	0.0047	0.12314
L'isle-sur-la-Sorgue	3255	403	0.00863	0.03695	3847	568	0.00645	0.04205	3079	609	0.00491	0.07087
Laragne-Monteglin	3505	347	0.0101	0.02908	3201	383	0.00908	0.03567	3317	460	0.00875	0.03572
Luc-le	1233	228	0.01487	0.05773	1050	218	0.01666	0.06067	1144	306	0.01175	0.08044
Manosque-1	1001	152	0.02308	0.04525	801	170	0.01656	0.08157	712	190	0.01108	0.14531
Manosque-2	134	30	0.05275	0.16481	102	32	0.05572	0.21354	119	42	0.05213	0.19219
Manosque-3	1416	232	0.00931	0.08216	1172	240	0.00751	0.12836	867	239	0.00696	0.19872
Marignane	2487	247	0.04606	0.00881	1721	267	0.01987	0.03013	1576	296	0.01586	0.04169
Menton	1134	103	0.03193	0.02841	1147	84	0.03199	0.02802	1198	111	0.03307	0.0259
Monteux	3230	403	0.00623	0.05237	3903	523	0.00523	0.05157	2942	547	0.00454	0.08096
Nice-3	573	108	0.03513	0.05229	459	128	0.02482	0.09626	312	118	0.0214	0.17621
Nice-7	519	76	0.04365	0.04619	577	71	0.0652	0.02731	590	110	0.04789	0.0367
Ollioules	1536	273	0.01334	0.05134	1899	372	0.00931	0.06002	1538	334	0.00862	0.08165
Oraison	669	135	0.01741	0.09396	532	130	0.01844	0.11354	510	142	0.01381	0.16553
Orange	1810	245	0.01034	0.05646	1750	275	0.00901	0.06774	1407	265	0.0099	0.07737
Pelissanne	3425	389	0.00772	0.03933	3375	455	0.00632	0.04926	2516	483	0.00549	0.07813
Pernes-les-Fontaines	7008	660	0.00492	0.02992	6727	768	0.00511	0.02999	5212	739	0.00468	0.04279
Pertuis	5115	585	0.00589	0.03439	4038	570	0.00555	0.04672	3383	644	0.0045	0.07033
Pontet	1213	180	0.01383	0.06342	1608	309	0.00903	0.07404	1227	268	0.00913	0.0981
Reillanne	5676	474	0.01052	0.01705	5135	534	0.00927	0.02148	4599	650	0.00707	0.03174
Riez	5136	460	0.01086	0.01828	4390	423	0.01107	0.02102	3762	417	0.00979	0.02794
Roquebrune-sur-Argens	3387	423	0.00821	0.03735	2754	401	0.00897	0.04224	2741	553	0.00643	0.06022
Salon-de-Provence-1	5298	643	0.005	0.03929	5309	653	0.00526	0.03719	4350	677	0.00455	0.05329
Salon-de-Provence-2	1753	295	0.00848	0.07216	1759	349	0.0079	0.07759	2182	515	0.00577	0.08632
Serres	10747	616	0.00572	0.01656	9371	632	0.00581	0.01873	8403	751	0.00552	0.02207
Seyne-les-Alpes	12373	726	0.00705	0.01161	10613	699	0.00732	0.01305	9911	746	0.0065	0.01577
Sisteron	2525	248	0.01305	0.03131	2027	239	0.01253	0.041	1691	257	0.01126	0.05548
Sollies-Pont	1574	243	0.02412	0.02706	1116	198	0.02244	0.0416	837	208	0.01545	0.08386
Sorgues	3350	365	0.00808	0.03838	3474	408	0.00829	0.036	2695	441	0.00684	0.05745
St-bonnet-en-Champsaur	13706	678	0.00641	0.01153	13514	669	0.0064	0.01171	14334	736	0.00664	0.01063
St-cyr-sur-mer	1019	146	0.01956	0.05283	931	137	0.01947	0.05842	815	165	0.0161	0.08251
St-Maximin-la-ste-Baume	6382	625	0.00721	0.02224	5452	595	0.00703	0.0268	4342	609	0.00747	0.03183

(Continued)



Table 1. (Continued)

Canton	Period 1: 1810–1835				Period 2: 1836–1861				Period 3: 1862–1890			
	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>	<i>B</i>	<i>N</i>	<i>I</i>	<i>v</i>
St-Raphael	737	211	0.01089	0.14236	790	264	0.00868	0.17075	1714	693	0.003	0.24213
Ste-maxime	2190	382	0.01785	0.02626	1859	364	0.01464	0.03817	1781	484	0.00796	0.0759
Tallard	4884	457	0.00785	0.02681	4920	467	0.00735	0.02847	4538	501	0.00681	0.03345
Toulon-3	630	114	0.04656	0.0353	565	109	0.03252	0.05758	441	114	0.03151	0.07755
Tourrette-Levens	13531	732	0.00639	0.01171	14611	843	0.00649	0.01067	13499	1149	0.006	0.01252
Trets	6150	749	0.00554	0.03028	5319	761	0.00563	0.03456	4298	836	0.00494	0.04946
Vaison-la-Romaine	6349	665	0.00721	0.02234	5842	706	0.00632	0.02788	4660	724	0.00595	0.03743
Valbonne	1958	254	0.01883	0.02789	1538	285	0.01493	0.04556	1197	344	0.01061	0.08548
Valensole	2217	275	0.0116	0.04048	1864	306	0.00978	0.05808	1540	326	0.00897	0.07811
Valreas	1007	172	0.01435	0.07436	929	203	0.01036	0.11603	796	202	0.00824	0.18007
Vence	6688	591	0.0075	0.02035	6190	617	0.00774	0.02134	5855	680	0.00798	0.0219
Veynes	4577	260	0.01176	0.01894	4129	255	0.01161	0.02133	4464	310	0.01092	0.02095
Vidauban	2214	251	0.01779	0.02606	2094	292	0.0153	0.03224	1963	452	0.00998	0.05381
Villeneuve-loubet	1480	241	0.02431	0.0286	1000	215	0.02407	0.04337	1072	362	0.00849	0.12361
Vitrolles	1078	203	0.01494	0.06622	960	216	0.01452	0.07732	838	225	0.0112	0.11935

*B*=number of registered births; *N*=number of different surnames; *I*=isonymy; *v*=Karlin–McGregor’s *v*-index.

represent 13.7% of the births registered between 1810 and 1890. They remain essentially the same throughout the three periods, which suggests a certain stability of the population. It is only from the period 1916 to 1940 that the Italian surname Rossi enters the ranking of the 25 most frequent names in the region, according to INSEE’s files of patronymic names.

Studying the distribution of frequent and rare surnames is particularly relevant for determining local surnames for genetic studies (Darlu *et al.*, 2012; Rossi, 2013); this point is discussed further in the Conclusion section. In France, between 1891 and 1940 the surname Martin was the most frequent surname, and it was present in 81 French departments; Durand was found in 65 departments, Richard in 59 departments and the surname Roux in 44 departments (Darlu *et al.*, 1997). These surnames are defined as polyphyletic because they cannot be considered as specific to one region as they probably have several geographical origins.

In France between 1891 and 1940, 96% of the surnames were present in fewer than ten departments at the same time and two-thirds of them were strictly localized in one department. These surnames are therefore regionally specific markers, and are defined as monophyletic, specific to one region and having only one central origin, as opposed to polyphyletic surnames (Darlu *et al.*, 1997; Rossi, 2013). The existence of these monophyletic surnames can be explained by the old presence of surnames on the French territory and the regional linguistic richness because the surnames are derived from words (Manni *et al.*, 2005; Rossi, 2013). These patronyms are notably used in population genetics studies (Winney *et al.*, 2012; Leslie *et al.*, 2015) in order to improve sampling strategies. They are also used in forensic DNA analysis in court cases (King *et al.*, 2006). This methodology hypothesizes a relationship between genetic characteristics of the Y chromosome, specific to male lineages, and a monophyletic patronymic that has been patrilineally transmitted for centuries and has a single geographical founder (King *et al.*, 2006). This hypothesis is not shared by all authors, as there are some very important factors to take into account, such as illegitimate births and abandoned children (Rossi, 2013). However, this type of bias may be limited depending on the temporal and spatial scale of the study or on the initial problem.

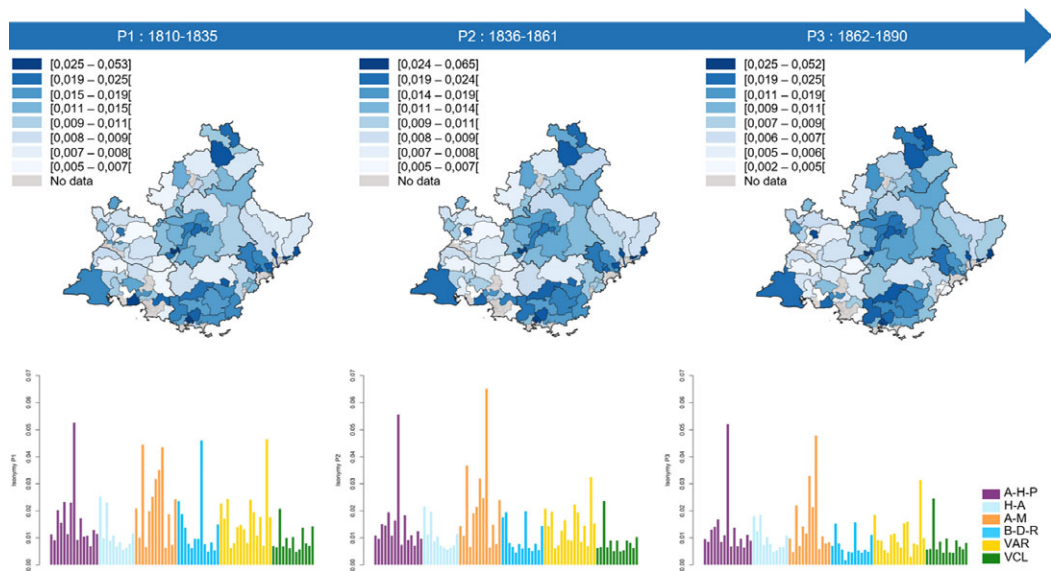
For this work, between 1810 and 1890, 23,340 surnames were collected: 11,578 surnames for P1, 12,269 for P2 and 15,469 for P3. This difference in registration is due to the appearance of new surnames in P2 and P3, coupled with the disappearance of other surnames. Only 26.3% of the surnames are common to the three periods but they represent a large part of the births (93.4%) over the entire study of 1810–1890. During the third period, 6898 new surnames were recorded, i.e. 29.5% of the surnames, representing 1.87% of the births. The stock of surnames thus seems to be variable; however, the majority of individuals have names that cross the three periods, suggesting relative stability of the population.

A ‘hapax’ is a name that appears only once in a corpus (Chareille & Darlu, 2013). For the three periods, 8299 ‘hapax’ were recorded, i.e. 35.5% of the surnames counted, but they represent only 1% of the registered births. This observation is comparable on a national scale: ‘hapax’ represented 36% of the surnames collected by INSEE between 1891 and 1915 (Darlu *et al.*, 1997). For the first period, 3386 ‘hapax’ were recorded, compared with 3604 for P2 and 5125 for P3. The number of ‘hapax’ thus increased significantly for the third period. Given that ‘hapaxes’ are likely to represent a migration (Chareille & Darlu, 2013), their increase in this corpus, particularly in the third period, seems to indicate an increase in migrations. Moreover, these ‘hapax’ had a strong impact on the proportion of names collected, but not on the population as they represented few births. It is therefore possible to say that migration tended to increase, but that its impact on the population’s gene pool was still anecdotal. The estimated  $R_{st}$  values, respectively for P1  $R_{st}=0.0022$ , P2  $R_{st}=0.0019$  and for P3  $R_{st}=0.0016$ , decreased across the three periods and especially in the third period. This indicates a homogenization between cantons over time (Boattini *et al.*, 2006). This phenomenon can be explained by the increase in migration, and therefore in the exchange of surnames between cantons, which led to a greater sharing of surnames between cantons. Indeed, the end of the 19th century was marked by great upheavals such as the arrival of the railway in the south-east of France in 1849, which favoured the mobility of individuals. The construction of this means of transport, as well as the development of industry, required a great deal of labour (Gastaut, 2009). This result is consistent with the increase in ‘hapax’.

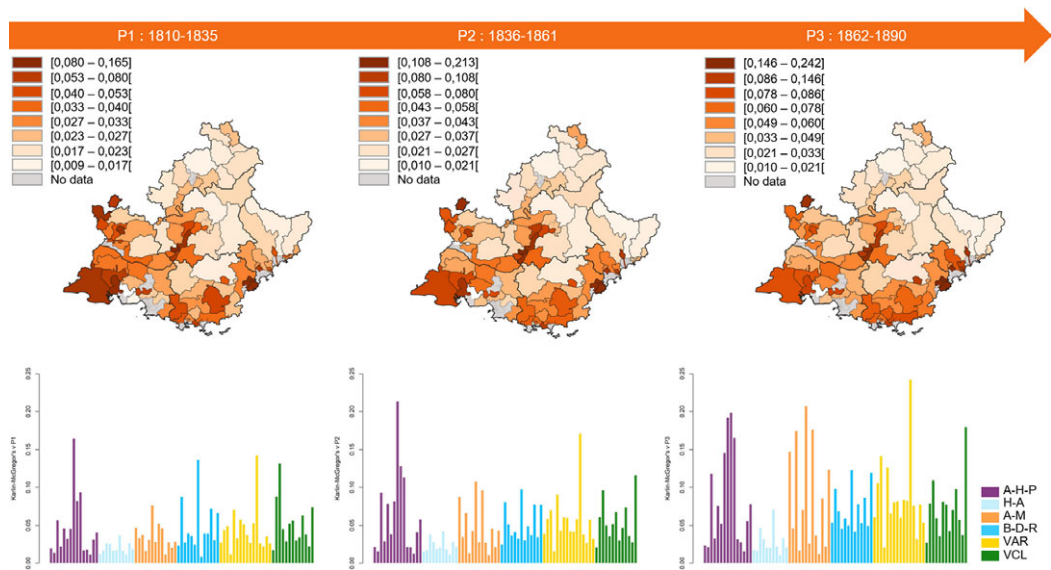
### Isolation and migration across the century

In general, the estimates of inbreeding, presented in Table 1, calculated from the isonymy index, tend to decrease across the three periods. Indeed, according to the cantonal distribution of the estimated isonymy values, presented by canton and by period in the bar charts of Figure 2, the values tend to decrease across the three periods. Thus, the townships became less and less isolated across the three periods. The lowest inbreeding estimates were found in the Durance Valley on the borders of the Vaucluse and Bouches-du-Rhône departments (Figure 2). This area corresponded to a communication corridor known as the Domitian Way, which since its creation during Roman domination has linked Italy to Spain (Temime, 1997). The economic role of this valley (as well as the Rhône Valley) is fundamental; it seals the age-old exchanges between north and south through transhumance and various commercial exchanges (Bouvier, 1979; Gastaut, 2009).

Migration, estimated from the Karlin–McGregor  $\nu$ -index, is presented in Table 1; higher values were found in the south-western quarter of the region (Figure 3). These observations are in line with the unbalanced distribution of today’s population within the southern region (Gastaut, 2009). In fact, the migratory flows directed towards the coastal cities, rich in industrial centres, have demographically enriched the south-western coastal quarter of the region during the rural exodus. According to the cantonal distribution of the estimated values of the Karlin–McGregor  $\nu$ -index, presented by canton and by period in the bar charts of Figure 3, an increase in values is noticeable across the three periods. Thus, the migration rate increased over the 19th century in the population sample of the southern region studied in this work.



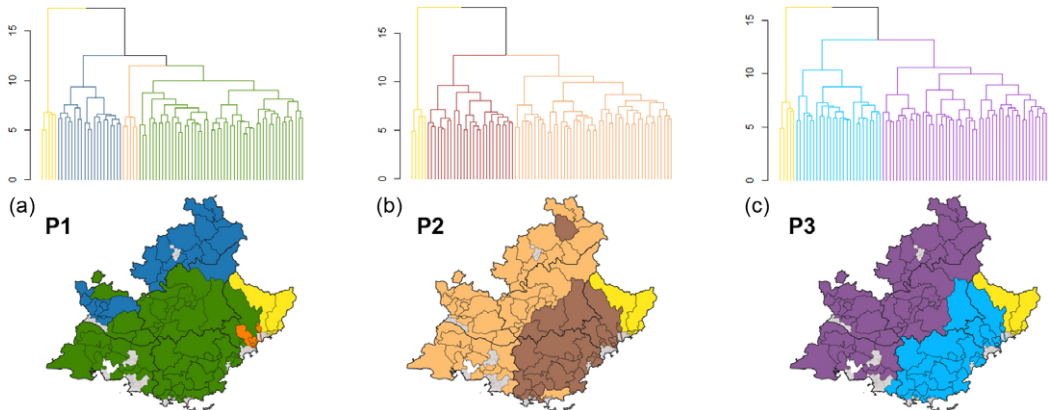
**Figure 2.** Maps of the cantonal distribution of the isonymy index by period. The bar graphs represent, for the three periods, the estimated values of the isonymy by canton and coloured according to their department. A-H-P: Alpes-de-Haute-Provence; H-A: Hautes-Alpes; A-M: Alpes-Maritimes; B-D-R: Bouches-du-Rhône; Var; VCL: Vaucluse.



**Figure 3.** Maps of the cantonal distribution of the Karlin-McGregor  $v$ -index by period. The bar graphs represent, for the three periods, the estimated values of the  $v$ -index by canton and coloured according to their department. A-H-P: Alpes-de-Haute-Provence; H-A: Hautes-Alpes; A-M: Alpes-Maritimes; B-D-R: Bouches-du-Rhône; Var; VCL: Vaucluse.

**Population structure based on Lasker distances**

The similarities and differences in surnames between the townships were studied using Lasker distances. These analyses showed that some cantons did not share any surnames. For the first period, the canton of Arles (Bouches-du-Rhône department) and the canton of Menton



**Figure 4.** Dendrograms obtained from a Hierarchical Ascending Classification with Ward's method using Lasker distances, for each period: (a) Period 1; (b) Period 2; (c) Period 3. The obtained groups were projected on the geographical space.

(Alpes-Maritimes department) did not share any patronymic. For the second period, this was the case for the canton of Menton with the canton of Briançon-2 (department of Hautes-Alpes) and the canton of Digne-les-Bains-2 (department of Alpes-de-Haute-Provence). These cantons were therefore the most different.

The geographical projection of the groups obtained from the dendrograms (Figure 4) has allowed for a synthesis of the relationships between the cantons according to the periods. The group in yellow in Figure 4a, b and c was maintained throughout the three periods and included five cantons in the eastern-most part of the Alpes-Maritimes: the cantons of Beausoleil, Contes, Menton, Nice-7 and Tourrette-Levens. The singularity of this group and its continuity through the three periods can be explained according to different historical and linguistic phenomena. On the one hand, from a linguistic point of view, the Roya Valley, whose territory extends within the cantons of Contes, Beausoleil and Menton, is closer to the Ligurian dialect of northern Italy (Bouvier, 1979; Caserio, 2015). It should not be forgotten that language has a particular influence on patronymics, as all patronymics are based on words derived from everyday language (Manni *et al.*, 2005; Rossi, 2013). On the other hand, the definitive annexation of this geographical area to France was late; 1860 for the County of Nice by referendum (Agulhon & Coulet, 1987; Temime, 1997) and 1947 for certain communes located in the Roya Valley following the Second World War and the Treaty of Paris (Baratier *et al.*, 1969; Agulhon & Coulet, 1987).

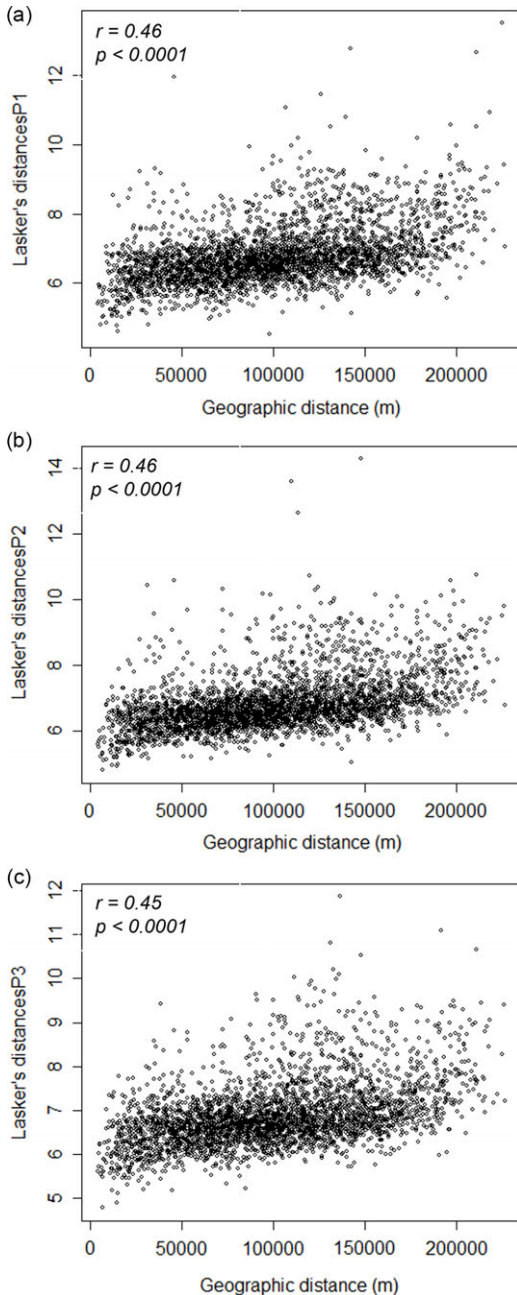
For the first period, four groups were defined from the dendrogram and are presented in Figure 4a. The largest group, in green, seems to reflect the politics of the County of Provence under the second Capetian house of Anjou-Provence (Baratier *et al.*, 1969; Agulhon & Coulet, 1987), this coinciding with the adoption of patronymics during the Middle Ages. In the north, the department of Hautes-Alpes, formerly Dauphiné, was united with the western territory of Vaucluse, formerly Comté Venaissin, and the canton of Barcelonnette, covering the Ubaye Valley in the north-east of the department of Alpes-de-Haute-Provence. This canton was lost by Provence in 1388 to join the States of Savoy and it was only in 1713, by the Treaty of Utrecht, that it was again attached to Provence (Baratier *et al.*, 1969; Agulhon & Coulet, 1987). The smallest group found around Nice, in orange (Figure 4a) corresponded linguistically to the Nissart dialect (Bouvier, 1979; Caserio, 2015) and also passed in 1388 under Savoyard suzerainty (Baratier *et al.*, 1969; Agulhon & Coulet, 1987). These findings highlight the multiple identities and histories of the Sud-Provence-Alpes-Côte d'Azur region that do not strictly coincide with current departmental boundaries (Gastaut, 2009). Indeed, the boundaries of the region and its constituent departments were created on economic factors rather than socio-cultural criteria (Bromberger & Meyer, 2003).

For the second period, three groups were defined from the dendrogram and are presented in Figure 4b. The group located in the south-east, in brown, gathered 27 cantons of the departments of Var, Alpes-Maritimes, Alpes-de-Haute-Provence and one canton of Hautes-Alpes: l'Argentière-la-Bessée. This group seems to be the witness of areas where migratory movements from outside began during the second period of this study (Gastaut, 2009). Indeed, in the 19th century, tourism and the labour force required for it, experienced a considerable boom on the Côte d'Azur. This was initiated in the 18th century with the first winter resorts of Hyères and Nice attracting English, Germans and Russians (Gastaut, 2009). In the countryside, particularly in the Var and Alpes-Maritimes, waves of immigrant workers came to supplement the lack of labour left by the rural exodus from the 1830s to 1840s (Gastaut *et al.*, 2008). The labour force, essentially Italian, was much sought after for viticultural or agricultural work because of their low wages (Gastaut *et al.*, 2008). The attachment of the canton of Argentière-la-Bessée in the Hautes-Alpes to the south-east group was, however, surprising at first glance. But it can be explained by the activity of its silver mine, which became notable and prosperous from 1855 onwards, employing up to 500 workers (Ancel, 1997). For the group located to the west, in beige (Figure 4b), it gathered 49 cantons of Bouches-du-Rhône, Vaucluse, Alpes-de-Haute-Provence, Hautes-Alpes and the canton of La Crau in Var. This grouping seems to be the witness of an intra-regional migratory movement (Alpine, peasants) directed towards the south-western coast (Baratier *et al.*, 1969; Agulhon & Coulet, 1987; Temime, 1997; Gastaut *et al.*, 2008; Gastaut 2009). Indeed, there has always been a co-dependence between the north, providing seasonal workers, and the south, providing blue-collar jobs thanks to the development of industry, within the Sud-Provence-Alpes-Côte d'Azur region (Baratier *et al.*, 1969; Agulhon & Coulet, 1987; Temime, 1997; Gastaut, 2009).

For the third period, three groups were defined from the dendrogram and are presented in Figure 4c. The blue group in the east included 26 cantons in the Var and Alpes-Maritimes, one canton in Alpes-de-Haute-Provence and the canton of La Ciotat in Bouches-du-Rhône, which in the previous period was part of the western group. The western group in blue (Figure 4c) included 50 cantons from Bouches-du-Rhône, Vaucluse, Alpes-de Haute-Provence and Hautes-Alpes. As for the canton of Argentière-la-Bessée, for the third period, it was no longer attached to the eastern group, probably because its silver mining activity declined from 1870 and stopped temporarily in 1881 (Ancel, 1997). At the same time, the canton of La Crau in the Var and La Ciotat in the Bouches-du-Rhône were attached to the western group in the third period. These two coastal cantons gradually welcomed Italian workers. Indeed, the commune of La Ciotat, which is the chief town of the canton of La Ciotat, is known for the development of its shipyards at the end of the 19th century (Gastaut, 2009).

These observations seem to depict an Italian migratory influence from east to west that will become a so-called 'mass' migration from the end of the 19th century (Temime, 1997). Indeed, immigrant workers settled everywhere in the region and particularly in the highly industrialized urban areas from the end of the 19th century (Gastaut, 2009). Of course, it is not possible to say, with the method used, whether the groups obtained in the third period are strictly the result of the evolution of the groups of the second period. But the change in population structures across the three periods indicates a fairly significant modification of the patronymic stock to the point of changing the relationships between the cantons (Chareille & Darlu, 2013). Obviously, the fact that cities with a population of more than 15,000 inhabitants, such as Marseille and Nice, were not taken into account in this study conditions these results. The population mix present in these urban centres would have disrupted the study of relations and structures in rural areas, as these cities were marked by the presence of Italian labour from the mid-19th century (Baratier *et al.*, 1969; Agulhon & Coulet, 1987; Temime, 1997; Gastaut *et al.*, 2008; Gastaut, 2009).





**Figure 5.** Variation of the Lasker distance with the geographical distance (in metres) between the 81 cantons studied in the Sud-Provence-Alpes-Côte d'Azur region by period: (a) Period 1; (b) Period 2; (c) Period 3.

For the Mantel test, the  $r$  correlations between Lasker distances and geographical distances ( $r=0.46$  for P1 and P2 and  $r=0.45$  for P3) were positive and significant ( $p<0.0001$ ) for all periods. The scatterplot in Figure 5 shows a linear and positive relationship between the geographical distance matrix and the patronymic distance matrix for all three periods. Thus, the greater the geographical distance between the cantons, the greater the patronymic distance. The geographical distance seems to have had an influence on the differentiations of the patronymic stocks between the cantons.



## Conclusion

With a multidisciplinary approach, this study has provided access to the population structures of the Sud-Provence-Alpes-Côte d'Azur region before the upheavals brought about by the migratory movements of the 20th century. Of course, this region has always welcomed migration, but in the 19th and 20th centuries, it became a major axis in France (Gastaut, 2009). This phenomenon began in the middle of the 19th century and had an impact on the structures of the population, which was mainly rural, that is studied here. The three-period approach was able to show this evolution; the structures obtained from groups linked to geographical areas impregnated with different historical, linguistic and geographical processes. The first period reflects an Ancien Régime Provence, while the structures found for P2 and P3 strongly echo the descriptions provided by the historian Gastaut (Gastaut *et al.*, 2008; Gastaut 2009) on the different types of migrations existing within the region.

The socioeconomic changes of the 19th century completely modified the patronymic structures of the southern region. The decrease in inbreeding estimates ( $R_{st}$  values), the increase in Karlin and McGregor's 'mutation rate'  $\nu$  and 'hapax', are consistent with increasing migration and population mixing over the chronological period studied. These trends are most pronounced in the third period, which coincides with the onset of the rural exodus, the Industrial Revolution and international migration in the South.

This study makes it possible to understand the beginnings of a so-called 'mass' migration wave described by historians (Temime, 1997; Gastaut, 2009). Thanks to the analogy between isonymic and genetic structures, it provides biological confirmation of historical facts and has made it possible to evaluate the importance of migratory flows within the region. This work highlights the singularity of the eastern part of the Alpes-Maritimes compared with the rest of the region, which is maintained throughout the three periods. This singularity will continue into the 21st century according to a study by Scapoli *et al.* (2005).

The study of this geographical area and the historical depth proposed can be relevant to a wide range of research applications in public health and population genetics. Initially, collaboration with the Etablissement Français du Sang (EFS) could allow the identification of rare genetic variants (blood systems, HLA) in this geographical area. This knowledge is essential in a public health strategy (Kristiansson *et al.*, 2008) that aims to improve blood transfusions by identifying geographical areas capable of supplying rare blood groups. The Sud-Provence-Alpes-Côte d'Azur region is a very good candidate for rare blood group research because of its geographical environment which is conducive to isolation (Chiaroni *et al.*, 2016). More than a century after the discovery of the ABO blood group, blood transfusion remains a major therapy that saves tens of thousands of patients per year (Bodmer, 2015) and the anthropological interest of erythrocyte blood groups can provide valuable information in this field (Marchini *et al.*, 2004). Secondly, the relevance of the use of surnames as a geodemographic indicator is undeniable (Darlu *et al.*, 2012, Cheshire, 2014). Indeed, the identification of locally rooted surnames makes it possible to improve sampling strategies for population genetics by targeting specific geographical areas (Darlu *et al.*, 2012). In particular, in the collaborative work of Winney *et al.* (2012) and Leslie *et al.* (2015), this methodology has enabled the identification of genetic clusters within Britain which they have been able to link to historical events dating back to the Middle Ages. This approach allows the identification of genetic isolates from the past and population structures that have been altered or disappeared today (Darlu *et al.*, 2012). In some cases, the correlation between Y-chromosome characteristics and surname has even proved fruitful in a forensic context (King *et al.*, 2006). It therefore seems that although we are in the era of the study of complete genomes, these genetic approaches through patronyms allow for indispensable interdisciplinary exchanges in the understanding of human populations.

**Funding.** The study was part of the doctoral project of the first author, who obtained doctoral funding from the Sud-Provence-Alpes-Côte d'Azur region and the French Blood Establishment (EFS Sud-Provence-Alpes-Côte d'Azur -Corse).

**Conflicts of Interest.** The authors have no conflicts of interests to declare.

**Ethical Approval.** The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

## References

- Agulhon M and Coulet N** (1987) *Histoire de la Provence*. PUF, Paris.
- Ancel B** (1997) *Les mines d'argent du Fournel. L'extraction souterraine au XIX<sup>ème</sup> siècle*. CCSTI, L'Argentière-la-Bessée.
- Baratier É, Duby G and Hildesheimer E** (1969) *Atlas historique. Provence, Comtat Venaissin, Principauté d'Orange, comté de Nice, principauté de Monaco*. Armand Colin, Paris.
- Barrai I, Rodriguez-Larralde A, Mamolini E, Manni F and Scapoli C** (2000) Elements of the surname structure of Austria. *Annals of Human Biology* 27(6), 607–622.
- Barrai I, Rodriguez-Larralde A, Mamolini E and Scapoli C** (1999) Isonymy and isolation by distance in Italy. *Human Biology* 71(6), 947–961.
- Boattini A, Calboli FCF, Villegas MJB, Guerres P, Franceschi MG, Paoli G et al.** (2006) Migration matrices and surnames in populations with different isolation patterns: Val di Lima (Italian Apennines), Val di Sole (Italian Alps), and La Cabrera (Spain). *American Journal of Human Biology* 18(5), 676–690.
- Bodmer W** (2015) Genetic characterization of human populations: from ABO to a genetic map of the British people. *Genetics* 199(2), 267–279.
- Bouvier JC** (1979) L'occitan en Provence: le dialecte provençal: ses limites et ses variétés. *Revue de linguistique romane* 43, 46–62.
- Bowden G, Balaesque P, King T, Hansen Z, Lee A, Pergl-Wilson G et al.** (2008) Excavation des structures de population passées par échantillonnage basé sur le nom de famille: l'héritage génétique des Vikings dans le nord-ouest de l'Angleterre. *Biologie moléculaire et évolution* 25(2), 301–309.
- Bromberger C and Meyer M** (2003) Cultures régionales en débat. *Ethnologie Française* 33, 357–361.
- Carriari A, Sans M, Dipierri JE, Alfaro E, Mamolini E, Sandri M et al.** (2020) The structure and migration patterns of the population of Uruguay through isonymy. *Journal of Biosocial Science* 52(2), 300–314.
- Casario JL** (2015) Entre provençal (nissart) et ligurien: le parler mentounasc, rameau de la langue d'oc. *Lengua et Fablas* 19, 61–65.
- Chaille P and Darlu P** (2013) Définition, identification et localisation de géophax dans le corpus des patronymes français: comparaison de sources anciennes et récentes. In *L'Onomastique dans l'espace public actuel*, pp. 484–503.
- Cheshire J** (2014) Analysing surnames as geographic data. *Journal of Anthropological Sciences* 92, 99–117.
- Chiaroni J, Chevè D, Berland-Benhaïm C and LeCoz P** (2016) *Le sang. Donner et recevoir*. CNRS Editions.
- Colantonio S, Fuster V and Kuffer C** (2007) Isonymous structure in the white population of Cordoba, Argentina, in 1813. *Human Biology* 79(5), 491–500.
- Crow J and Mange A** (1965) Measurement of inbreeding from the frequency of marriages between persons of the same surname. *Eugenics Quarterly* 12(4), 199–203.
- Darlu P, Bloothoof G, Boattini A, Brouwer L, Brouwer M, Brunet G et al.** (2012) The family name as socio-cultural feature and genetic metaphor: from concepts to methods. *Human Biology* 84(2), 169–214.
- Darlu P, Degioanni A and Ruffie J** (1997) Quelques statistiques sur la distribution des patronymes en France. *Population* 52(3), 607–634.
- Darlu P and Oyharçabal B** (2006) Distribution des patronymes et structure des populations dans les Pyrénées occidentales. *Lapurdum* 11, 137–165.
- Darlu P and Ruffie J** (1992) L'immigration dans les départements français étudiée par la méthode des patronymes. *Population* 47(3), 719–734.
- Dipierri JE, Alfaro EL, Scapoli C, Mamolini E, Rodriguez-Larralde A and Barrai I** (2005) Surnames in Argentina: a population study through isonymy. *American Journal of Physical Anthropology* 128(1), 199–209.
- Dupâquier J** (2002) Sédentarité et mobilité dans l'ancienne société rurale. Enracinement et ouverture: faut-il vraiment choisir ? *Histoire et Sociétés Rurales* 18, 121–135.
- Fordant L** (1999) *Atlas des noms de famille en France*. Archives and Cultures, Paris.
- Gastaut Y** (2009) Histoire de l'immigration en PACA: XIX<sup>e</sup> et XX<sup>e</sup> siècles. *Hommes et migrations* 1278, doi: 10.4000/hommesmigrations.226.
- Gastaut Y, Rinaudo C, Dahhan R, Véglia P, Folliet D, Sagatni T et al.** (2008) *Histoire et mémoire des migrations en PACA*. ACSE, Nice.
- Gibert M, Roumieu A, Telmon N, Sangoi JC and Sevin A** (2012) Structure des populations du Haut Viedessos: distribution spatiale des patronymes et des activités économiques au début du XX<sup>e</sup> siècle. *Revue géographique des Pyrénées et du Sud-Ouest* 33, 117–126.

- Karlin S and McGregor J** (1967) The number of mutant forms maintained in a population. *Proceedings of the Fifth Berkeley Symposium on Mathematics, Statistics and Probability* **4**, 415–438.
- King TE, Ballereau SJ, Schurer KE and Jobling MA** (2006) Genetic signatures of coancestry within surnames. *Current Biology* **16** (4), 384–388.
- Kristiansson K, Naukkarinen J and Peltonen L** (2008) Isolated populations and complex disease gene identification. *Genome Biology* **9**(8), 9.
- Lasker GW** (1977) Coefficient of relationship by isonymy – method for estimating genetic relationship between populations. *Human Biology* **49**(3), 489–493.
- Le Mée R** (1999) *Dénombrements, espaces et sociétés*. Société de Démographie Historique, Paris.
- Leslie S, Winney B, Hellenthal G, Davison D, Boumertit A, Day T et al.** (2015) The fine-scale genetic structure of the British population. *Nature* **519** (7543), 309, doi: [10.1038/nature14230](https://doi.org/10.1038/nature14230).
- Manni F, Toupance B, Sabbagh A and Heyer E** (2005) New method for surname studies of ancient patrilineal population structures, and possible application to improvement of Y-chromosome sampling. *American Journal of Physical Anthropology* **126**(2), 214–228.
- Marchini J, Cardon LR, Phillips MS and Donnelly P** (2004) The effects of human population structure on large genetic association studies. *Nature Genetics* **36**(5), 512–517.
- Mateos P, Longley P and O'Sullivan D** (2011) Ethnicity and population structure in personal naming networks. *Plos One* **6**(9), 12.
- Michaïlesco F and Mora V** (2020) Populations légales – Entre 2013 et 2018, la région gagne près de 20000 habitants par an. *INSEE Flash Provence-Alpes-Cte D'azur* 69.
- Mourrieras B, Darlu P, Hochez J and Hazout S** (1995) Surname distribution in France – a distance analysis by a distorted geographical map. *Annals of Human Biology* **22**(3), 183–198.
- Relethford JH** (1988) Estimation of kinship and genetic-distance from surnames. *Human Biology* **60**(3), 475–492.
- Rodriguez-Larralde A, Scapoli C, Beretta M, Nesti C, Mamolini E and Barrai I** (1998) Isonymy and the genetic structure of Switzerland. II. Isolation by distance. *Annals of Human Biology* **25**(6), 533–540.
- Roman-Busto J and Fuster V** (2015) Secular trends in the relationship between surnames in a population: study of a border town. *Homo-Journal of Comparative Human Biology* **66**(2), 158–164.
- Rossi P** (2013) Surname distribution in population genetics and in statistical physics. *Physics of Life Reviews* **10**(4), 395–415.
- Scapoli C, Goebel H, Sobota S, Mamolini E, Rodriguez-Larralde A and Barrai I** (2005) Surnames and dialects in France: population structure and cultural evolution. *Journal of Theoretical Biology* **237**(1), 75–86.
- Temime E** (1997) *Voyages en Provence, Alpes, Côte d'Azur*. Galimard, Paris.
- Vernay M** (2001) Répartition géographique des patronymes et structure génétique: le département de l'Ardèche au début du XX<sup>e</sup> siècle. *Comptes Rendus de l'Académie des Sciences-Serie III-Sciences De La Vie-Life Sciences* **324**(7), 589–599.
- Winney B, Boumertit A, Day T, Davison D, Echeta C, Evseeva I et al.** (2012) People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. *European Journal of Human Genetics* **20** (2), 203–210.