Journal of Geodesy and Geoinformation Science
http://jggs.chinasmp.com

Article Open Access

The Distribution of Surnames in Xiantao City through Isonymy

ZHANG Jingjing[1,2], ZHU Yan[1,3], XIAO Hong[1]

1 College of Chinese Language and Literature, Wuhan University, Wuhan 430072, China;
2 College of Literature and Communication, China Three Gorges University, Yichang 443002, China;
3 School of Foreign Languages, China University of Geosciences, Wuhan 430072, China

ABSTRACT

The population migration in Hubei Province was frequent in history, accompanied by the migration of surnames, so it is important to study their population surnames. We take Xiantao City as an example to explore the isonymic structure of small and medium-sized cities in Hubei. The surname distributions of 223 327 residents registered in 2013 were analyzed in 5 towns and 105 villages of Xiantao. The number of different surnames found was 422. As for surnames, the α-value reflects the influence of ethnic composition on the abundance of surnames. The correlation between the isonymic distance and the geographic distance between villages was calculated and indicated that Euclidean distance was weakly correlated with the geographic distance (r =0.177± 0.012), and the isonymic distance increased with the geographical distance. Furthermore, the dendrogram and PCA built from the matrix of Euclidean distances between villages identified a main surname differentiation between the urban and rural areas.

Key words: Xiantao City; surname distribution; isonymy; isolation by distance; urban and rural areas

Citation: ZHANG Jingjing, ZHU Yan, XIAO Hong. The Distribution of Surnames in Xiantao City through Isonymy[J]. Journal of Geodesy and Geoinformation Science, 2024, 7(2): 52-65. DOI: 10.11947/j.JGGS. 2024. 0204.

1 Introduction

Surnames, as a human identifier, reflect the national historical, socioeconomic, and linguistic background of a particular population. In a system of surname attribution through paternal lineage, surnames simulate neutral alleles of genes transmitted only through the Y chromosome and thus satisfy the expectations of neutral evolutionary theory, which are completely described by random genetic drift, mutation and migration. In the process of inter-generational transmission, surnames are not influenced by disease, climate, and living area and environment, which are neutral in nature. This property and validity of surnames are important in population studies[1-2].

Chinese surnames are an important part of the Chinese national culture with a long history of nearly 4000 years, which are characterized by continuity and stability, settlement and migration in history. Since ancient times, Chinese surnames are generally inherited from generation to generation in a patrilineal manner, while the female descendants only retain their father's surname for life and will not pass it on to the next generation. Surnames are rarely changed in inter-generational transmission, and the surnames changed are usually the existing common surnames or locally well-known surnames, with little impact on the overall distribution of the

∗ Foundation support: National Social Science Foundation of China (No. 20BYY045).
First auther: ZHANG Jingjing, PhD, majors in Chinese philology.
E-mail: jing1030@whu.edu.cn.
Corresponding author: XIAO Hong.
E-mail: xiaohong@whu.edu.cn.

surnames in the group[2-3]. Population migration is the main factor causing the change in the surname proportion of the population[2], especially group migration is much more influential on surnames than other forms.

The study on the Chinese surnames mainly based on the collection, record and textual research[4], focusing on studying the origin of surnames from a cultural or social perspective, such as the dictionaries of Chinese surnames, which collected and analyzed the historical origins and approximate geographic distribution of Chinese surnames[5-6]. Also, the genealogical cultural study of a particular surname, which documented the origins and evolution of a particular surname. Since the 1980s, there have also been some quantitative studies of Chinese surnames from the perspectives of statistics and geography or human population genetics. For example, ZHANG[4] counted the frequency of surnames in seven regions of China, indicating that the regional distribution of surnames is uneven. YUAN & ZHANG[2] studied Chinese surnames based on population genetics and geographic distribution. LIU et al.[7] studied the distribution of Chinese surnames at the provincial, city and county levels through isonymy. The large-scale studies of Chinese surnames above are helpful to grasp the overall distribution of surnames, but to some extent, it covers up some detailed information on the distribution of surnames. However, the micro-study of surnames can precisely make up for this shortcoming, and can show the specific history, culture, and distribution details of Chinese surnames in detail. But there are fewer studies available on the microscopic study of surnames, for example, ZHU[8] preliminarily counted the number of surnames of Xiantao and summarized their characteristics, which is a good attempt of micro-study, but lacking research on the geographical distribution of surnames. And ZHAO[9] analyzed the spatial distribution of surnames in the hereditary ethnic villages in Pingguo County, Guangxi. In general, the micro-researches on the geographical distribution of surnames are very rare, which is not conducive to revealing the diversity and unique distribution pattern of surnames in different places in China, nor is it conducive to understanding the historical evolution of surnames, as well as their cultural connotations and migratory paths, etc. It is imperative to carry out in-depth research on surnames of a appropriate scale.

Given the current status of research, we have selected the surnames of mid or small-sized cities in Hubei Province as the object of our study, which is based on the following considerations: firstly, Hubei Province is in the middle of China, which is the junction between the north and south of China and its population migrated frequently throughout history. A large number of immigrants moved into Hubei during the three large-scale southward migrations of the northern population (i.e., the Yongjia Rebellion, the An-shi Rebellion, and the Jingkang Rebellion) and the "Jiangxi Filling of Hubei" migration movement. Moreover, frequent floods in the mid-lower reaches of the Yangtze River have aggravated population migration in Hubei Province. Generally speaking, before the Ming Dynasty, the immigrants in Hubei mainly came from the northern provinces, while after the Ming Dynasty, a large number of Jiangxi people moved to Hubei, where the convenient transportation and stable social environment are suitable for people to live. The origin of immigrants in Hubei changed mainly from the north to the Yangtze River basin, especially the mid-lower reaches[10]. YUAN & ZHANG[2] pointed out that the distribution of surnames in Hubei was closer to that in the north. Hubei surnames can be considered as the result of many population migrations in history, reflecting the characteristics of the population surnames in the mid-lower reaches of the Yangtze River, and it is particularly important to study the Hubei surnames. At present, the large provincial capital city like Wuhan, where the population migration and origin are very complicated, cannot well reveal the demographic situation in Hubei Province. In contrast, there are a large number of small and medium-sized cities, and their current population migration status is relatively sta-

ble and can better reveal the general population situation in Hubei Province, that is, the short-distance intra-provincial or intra-local migration is predominant, and the cross-provincial migration is less and mostly from neighboring provinces. In addition, modern commerce in Xiantao is relatively developed and the scale is moderate, which is very suitable as the object of micro-study. Importantly, our team has been deeply engaged in the study of Xiantao dialect and obtained first-hand information on Xiantao surnames and personal names, and has conducted in-depth research on Xiantao surnames and names for many years. Besides, we have also done many field investigations. So we are very familiar with Xiantao. Therefore, we take Xiantao City, a small and medium-sized city as an example to explore the distribution patterns of population surnames in small and medium-sized cities in Hubei Province.

This paper will conduct an in-depth study on the isonymic structure of Xiantao in order to reveal the variability of its surnames. Furthermore, the correlation between isonymic distance and geographical distance between villages was calculated to reveal the geographical distribution pattern of surnames. The dendrogram and PCA built from the matrix of Euclidean distances between villages will visualize the regional difference in the surname distribution.

2 Materials and Methods

2.1 Data and material

Xiantao City, formerly known as Mianyang, has been a county-level municipality in Hubei Province since 1994. Covering an area of 2538 km^2, Xiantao is located in the Jianghan Plain in central Hubei Province, adjacent to the Han River to the north, the Yangtze River to the south, Wuhan City to the east, Qianjiang City to the west and Jingzhou City to the south, shown in Fig. 1. As one of the birthplaces of Jingchu culture, it is located at the intersection of the "Liangjiang" (the Yangtze River and Han River) economic belt in Hubei, known as "the treasure land in Hubei" and "the Pearl of Jianghan".

In the present work, we obtained the data of 223 327 individuals of Xiantao in 2013 from Xiantao Public Security Bureau, excluding 15 Uyghurs from Xinjiang Uygur Autonomous Region, since they did not have surnames. The data contains the following information: household number, name, sex, year of birth, ancestral native place, and current residential address. These people were distributed in five towns and 105 villages, including Gongyeyuan and Shazui sub-district in the urban areas, and Tonghaikou Town, Miancheng Town, and Shahu Town in the rural areas, as shown in Fig. 1. Among them, there are the largest number of administrative villages in Shahu town (39), and the least in Gongyeyuan (7), shown in Tab. 1.

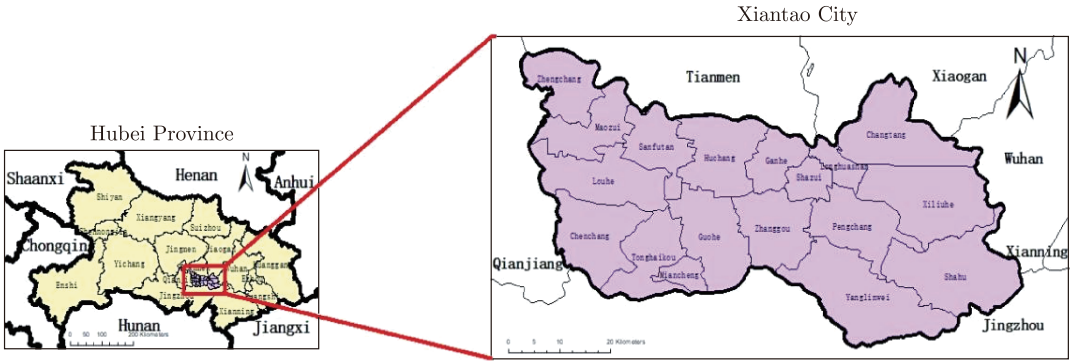


Figure 1 Location map of Xiantao City and Hubei Province

Table 1 Surname data in Xiantao

Town	Village	N	S	I	α	ID
Miancheng		24 475	233	0.027 464 668	36.410 415 84	
	Miancheng	6845	177	0.027 532 33	36.320 936 69	1
	Chengjiao	410	77	0.059 240 265	16.880 410 71	2
	Eryang	1937	118	0.049 980 374	20.007 853 68	3
	Gubaimen	1219	105	0.029 677 883	33.695 125 27	4
	Huangjin	1535	114	0.064 849 301	15.420 366 73	5
	Jiangbei	1121	104	0.049 851 854	20.059 434 41	6
	Nanqiao	1225	93	0.059 293 051	16.865 382 88	7
	Qihong	1354	110	0.042 426 644	23.570 094 18	8
	Shangguan	1132	111	0.042 218 494	23.686 302 08	9
	Shaoshendu	1871	126	0.034 843 102	28.700 085 31	10
	Wanghe	1398	112	0.042 675 752	23.432 510 26	11
	Yuanlou	1916	115	0.032 675 232	30.604 220 54	12
	Zhouling	2512	129	0.043 028 826	23.240 234 33	13
Tonghaikou		62 775	298	0.026 605 402	37.586 351 16	
	Tonghaikou	12 407	209	0.025 647 038	38.990 856 78	14
	Caisang	1964	109	0.043 662 647	22.902 871 67	15
	Chenjia	1928	116	0.050 664 611	19.737 642 91	16
	Chenzha	1861	126	0.038 359 536	26.069 136 92	17
	Dahe	2806	125	0.062 623 891	15.968 346 65	18
	Diwan	2901	134	0.040 151 672	24.905 563 19	19
	Dongdi	1809	113	0.036 560 071	27.352 244 6	20
	Gonghe	1955	113	0.045 110 168	22.167 951 07	21
	Guihuatai	1236	99	0.105 564 509	9.472 880 725	22
	Haifeng	98	43	0.033 452 556	29.893 081 76	23
	Liuli	2590	118	0.036 813 009	27.164 310 31	24
	Matao	2641	124	0.043 334 71	23.076 189 85	25
	Nianpan	3188	128	0.065 540 135	15.257 826 27	26
	Panba	3502	144	0.055 881 399	17.895 042 43	27
	Shiyuan	1933	107	0.067 090 171	14.905 313 07	28
	Sutan	1688	117	0.034 976 837	28.590 349 59	29
	Wangjiadu	1602	108	0.039 136 744	25.551 435 57	30
	Wuqi	864	75	0.085 170 701	11.741 126 82	31
	Wunao	1123	103	0.046 472 795	21.517 965 71	32
	Xiangyang	1009	90	0.041 226 187	24.256 427 38	33
	Xiewei	2448	120	0.038 279 165	26.123 871 58	34
	Xinjie	1044	98	0.074 517 95	13.419 585 42	35
	Xinghong	3081	130	0.057 329 379	17.443 063 38	36
	Xiongmiao	1445	110	0.046 252 72	21.620 350 22	37
	Yongchanghe	2274	110	0.050 097 876	19.960 926 22	38
	Paihuyuanzhongchang	1181	100	0.028 870 965	34.636 874 29	39

Continued

Town	Village	N	S	I	α	ID
	Paihuyuchang	2197	124	0.033 300 087	30.029 951 45	40
Shahu		63 152	301	0.030 622 32	32.655 918 52	
	Shahu	10 182	197	0.028 342 53	35.282 665 83	41
	Bayi	1544	100	0.089 797 145	11.136 211 51	42
	Chenhe	672	70	0.059 137 925	16.909 622 74	43
	Chiling	1315	92	0.055 036 431	18.169 782 75	44
	Chunliang	585	70	0.068 504 859	14.597 504 7	45
	Chunnan	1220	89	0.039 287 78	25.453 207 37	46
	Fengle	1845	97	0.065 207 014	15.335 773 46	47
	Hongjunba	1145	91	0.066 818 335	14.965 952 2	48
	Hongtuhu	1079	75	0.085 675 082	11.672 005 14	49
	Huangjinba	891	68	0.080 681 488	12.394 416 94	50
	Huanglou	1113	90	0.039 880 225	25.075 084 08	51
	Jiahe	1457	88	0.063 849 586	15.661 808 79	52
	Jingfeng	816	75	0.053 431 373	18.715 596 33	53
	Liangting	1217	91	0.043 433 486	23.023 710 25	54
	Zongshuwan	768	84	0.038 006 573	26.311 238 16	55
	Maling	801	76	0.152 662 297	6.550 405 823	56
	Mopan	1067	92	0.048 708 395	20.530 341 86	57
	Pengtai	661	69	0.062 361 894	16.035 433 36	58
	Pitiaozhou	1121	84	0.083 297 757	12.005 125 16	59
	Qunhe	678	62	0.156 525 187	6.388 748 156	60
	Qunxing	1478	110	0.030 949 984	32.310 194 78	61
	Sangou	1696	94	0.077 752 268	12.861 361 16	62
	Shuangfeng	601	65	0.052 640 044	18.996 944 47	63
	Shuangsheng	702	67	0.048 916 688	20.442 921 24	64
	Tanhu	696	68	0.051 562 888	19.393 793 6	65
	Tangtai	601	60	0.061 242 374	16.328 563 67	66
	Xinfa	682	62	0.089 948 799	11.117 435 85	67
	Xinhe	1208	91	0.069 961 648	14.293 545 6	68
	Xinkou	1854	113	0.034 334 829	29.124 944 89	69
	Yangtai	1373	90	0.043 070 334	23.217 837 16	70
	Yaobang	1179	93	0.039 620 927	25.239 187 32	71
	Youhe	1862	106	0.049 797 096	20.081 492 39	72
	Youhu	1199	91	0.042 837 59	23.343 983 62	73
	Yuchang	928	84	0.050 238 534	19.905 039 57	74
	Zhouhu	855	78	0.052 530 233	19.036 656 59	75
	Zhupaikou	1361	81	0.073 502 399	13.604 998 16	76
	Shahuyuanzhongchang	11 018	198	0.035 692 557	28.017 045 38	77
	Wuhuyuchang	2522	130	0.041 024 781	24.375 510 86	78
	Xuqinchang	1160	98	0.043 284 936	23.102 725 37	79

Continued

Town	Village	N	S	I	α	ID
Gongyeyuan		15 905	249	0.045 561 288	21.948 457 55	
	Chuangyelu	5538	195	0.029 828 75	33.524 703 01	80
	Duhuyuanzhongchang	1106	109	0.028 011 75	35.699 304 78	81
	Chuanwan	1801	93	0.051 087 667	19.574 195 73	82
	Guangou	1062	80	0.056 491 85	17.701 668 39	83
	Qingshuiwan	3024	118	0.177 364 85	5.638 095 708	84
	Tiejiangwan	2683	113	0.141 604 707	7.061 912 146	85
	Fangzhidadao	691	162	0.027 064 326	36.949 008 06	86
Shazui sub-district		57 020	297	0.034 343 698	29.117 423 49	
	Shazui sub-district	8084	230	0.025 554 607	39.131 886 55	87
	Duliu	5237	155	0.155 474 511	6.431 922 456	88
	Balouwan	1960	137	0.031 852 465	31.394 744 16	89
	Gaofeng	523	52	0.121 550 442	8.227 037 126	90
	Jintai	2856	113	0.100 410 357	9.959 132 026	91
	Jiushidun	1827	105	0.065 111 918	15.358 171 44	92
	Liukou	2526	131	0.040 871 413	24.466 979 18	93
	Lvwan	3266	112	0.098 180 427	10.185 329 52	94
	Meihu	2100	101	0.107 760 612	9.279 828 547	95
	Qiwei	2835	118	0.288 094 305	3.471 085 627	96
	Shazui	4049	155	0.081 491 355	12.271 240 24	97
	Shiyidun	6252	172	0.056 416 412	17.725 338 62	98
	Wanxiangyuan	1956	99	0.068 192 825	14.664 299 3	99
	Wangshikou	1129	79	0.059 353 975	16.848 071 12	100
	Xutai	811	71	0.063 530 773	15.740 403 51	101
	Yanggang	3258	109	0.270 276 062	3.699 920 711	102
	Yehe	4435	134	0.128 152 297	7.803 215 602	103
	Yuji	1066	76	0.105 097 376	9.514 985 417	104
	Zhoujiazha	2850	116	0.080 843 632	12.369 558	105

Note: Sample size (N), number of surnames (S), random isonymy(I), Fisher's α, and code.

2.2 Isonymy theory

Next, we will briefly introduce some parameters used in the study of the surname distribution, such as isonymy, namely $4F_{ST}$, where F_{ST} represents the inbreeding coefficient, Fisher's α and the surname distance.

2.2.1 Isonymy

Based on the surname distribution, the random isonomy refers to the probability that any two people who randomly meet in a crowd have identical surnames[2]. The isonymy can be used to describe the degree of differentiation and concentration of surnames in a population. If the concentration of population surnames is high, the value of the isonymy is large, and vice versa. The random isonymy between groups i and j was estimated as: $I_{ij} = \sum P_{ki} P_{kj}$.

Where P_{ki} and P_{kj} are the relative frequencies of the k_{th} surname in groups i and j, respec-

tively, and $\sum P_{ki}P_{kj}$ is the total sum of the products of P_{ki} and P_{kj}. When there are no identical surnames between two groups, the isonymy is 0. The random isonymy within the group was: $I_{ii} = \sum n_{ik}(n_{ik}-1)/N_i(N_i-1)$[11].

2.2.2 Fisher's α

Fisher's α was estimated according to BARRAI et al.[12] It is an important parameter for analyzing the frequency distribution of surnames and studying the degree of surname differentiation within a group[1, 13]. $\alpha = 1/I_{ii}$, where I_{ii} is the random isonymy of the ith sample. α is equivalent to the number of "effective alleles" in a genetic system, described as the number of "effective surnames". It can effectively estimate the number of surnames with the same frequency in a population. A small value of α indicates large inbreeding and drift, while large value indicates migration and low inbreeding.

2.2.3 Isolation by distance

To detect isolation by distance, we calculate the linear correlation of surname distances (Lasker's, Euclidean and Nei's) between villages i and j, with their geographic distance.

Lasker's distance[14] was defined as

$$L = -\log(I_{ij})$$

Euclidean distance[15] was defined as

$$E = \sqrt{1-\sum_k \sqrt{P}_{ki}P_{kj}}$$

Nei's distance[16] was defined as

$$N_d = -\log(I_{ij}/\sqrt{I_{ii}I_{jj}})$$

As for geographic distance, we use the longitude and latitude coordinates of each town (village) to calculate, which were obtained from Baidu Maps, and converted from the BD-09 coordinate system to the WGS1984 coordinate system by the equal offset method.

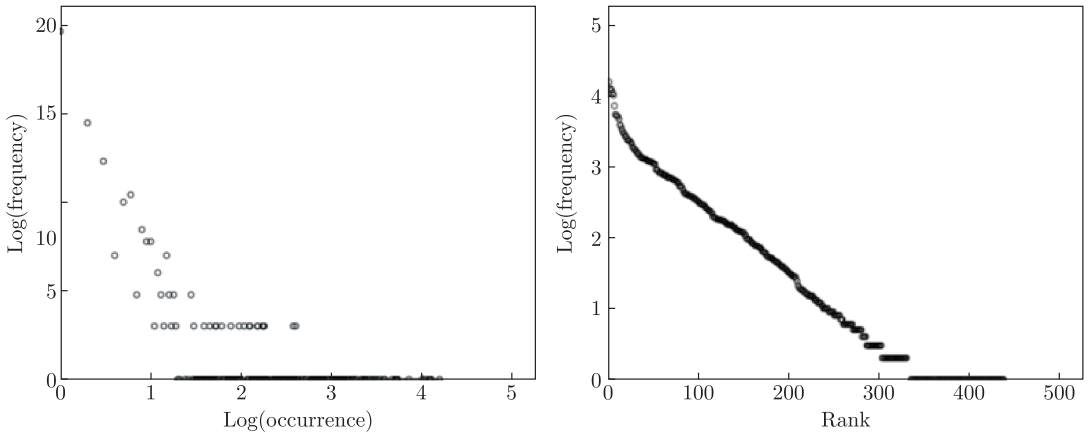
3 Results and Discussion

3.1 Distribution of individuals

The sample size, number of surnames, α and I in 5 towns analyzed are given in Tab. 1. The number of individuals per town ranged from 15 905 in Gongyeyuan to 63 152 in Shahu Town. The total number of different surnames obtained in Xiantao was 422, containing 418 monosyllable surnames and 4 polyphonic surnames, such as OUYANG/欧阳, SHANGGUAN/上官, ZHUGE/诸葛 and XIANYU/鲜于. Some of these polyphonic surnames have been simplified into monosyllable surnames. For example, OUYANG/欧阳 is simplified to OU/欧 (166) or YANG/阳 (83), SHANGGUAN/上官 to GUAN/官 (8), ZHUGE/诸葛 to ZHU/诸 (5), and XIANYU/鲜于 to XIAN/鲜 (6). The above data reflects the fact that Chinese surnames are mostly monosyllabic, but the number of common Chinese characters is small, and even fewer are used to express surnames, so the number of Chinese surnames is small. Specifically, the number of people in Shahu Town was the largest, with 301 surnames. While the number of people in Gongyeyuan was the least, with 249 surnames. The total population in the urban areas (Gongyeyuan and Shazui) was 72 925, with 343 surnames. The total population of the rural areas was 150 402, with 376 surnames.

3.2 The most frequent surnames

The log-log frequency distribution of surname occurrence is shown in Fig. 2(a). The graph is weakly linear, while the graphs of European countries are very linear[1, 17-18]. The distribution of points in the linear part is few and scattered, and the two truncations at the bottom become flat. At the second cut-off part, there is a horizontal long tail, which reflects the very concentrated characteristics of the Xiantao surnames, that is, the number of common surnames with a large population is small, while the number of rare surnames with a small population is large[7, 19-20]. Among all the surnames in Xiantao, there are 92 surnames that appear only once and 185 surnames that appear 10 or less, accounting for 0.23%. These rare surnames are used by individual or family group immigrants, including many married female immigrants. In other words, most surnames are used less frequently, and only a few common surnames are used more frequently, as



(a) Log-log distribution of frequency of occurrence of surnames in Xiantao (223 327 individuals, 2013)

(b) The distribution of the frequency of surnames, arranged by rank (the Zipf plot)

Figure 2 Surname distribution

shown in Fig. 2(b).

The list of the 20 most frequent surnames is shown in Tab. 2, and the number of people using these surnames is 132 201, 59% of the total population of the five towns. Among them, the most common surname is LI/李 (15 706, 7%), followed by WANG/王 (12 745, 5.7%), ZHANG/张 (12 279, 5.5%), LIU/刘 (11 990, 5.4%), ZHOU/周 (10 963, 4.9%) and CHEN/陈 (10 145, 4.5%) and so on. The 100 most frequent surnames comprise 209 635 individuals, about 94%, which is slightly higher than the proportion in China, 82.1%[3], but far exceeds the proportion in Western countries: 8.1% in France[21], 7.4% in western Europe[22] and 29.5% in Argentina[18], and so on. It further reflects that a small number of Chinese surnames are shared by the vast majority of people.

According to the statistics of the National Bureau of Statistics in 2010, the 10 most frequent surnames in China were WANG/王, LI/李, ZHANG/张, LIU/刘, CHEN/陈, YANG/杨, HUANG/黄, ZHAO/赵, ZHOU/周 and WU/吴[3]. Most of these surnames are in 10 most frequent surnames in Xiantao, except ZHAO/赵 and WU/吴, ranking 21st and 22nd, respectively in Xiantao. ZHAO/赵 is mainly distributed in the north of China, and WU/吴 is a common surname in Jiangsu, Zhejiang, Shanghai and southeast coastal areas of China. In Ming Dynasty, the ZHAO family and the WU family settled and multiplied in Xiantao(See *ZHAO's genealogy* and *WU's genealogy in Mianyang*). The 10 most frequent surnames in Hubei are LI/李, LIU/刘, ZHANG/张, CHEN/陈, YANG/杨, HU/胡, HUANG/黄, WANG/王, XU/徐 and ZHOU/周[23]. Only HU/胡 and XU/徐 are not among the 10 most frequent surnames in Xiantao, ranking 11th and 18th in Xiantao, respectively. Among the 10 most frequent surnames in Xiantao, XU/许 and DU/杜 are not among the 10 most frequent surnames in China and Hubei, but the common surnames in the neighboring provinces of Anhui, Hebei and Henan. It can be seen that the frequency distribution of the common surnames in Xiantao is roughly the same as that in China and Hubei Province, reflecting that the common surnames in Xiantao have both certain universality but also regional characteristics.

In addition, the distribution of the 20 most common surnames between the urban and rural areas is also slightly different from the overall distribution, shown in Tab. 2. The common surnames YIN/尹, ZHAO/赵 and CHANG/昌 in urban areas and DENG/邓, GUO/郭, YU/余 and WU/吴 in ru-

Table 2 The 20 most frequent surnames in Xiantao

Total		Rural areas		Urban areas	
LI/李	15 706	LI/李	11 546	ZHOU/周	6942
WANG/王	12 745	WANG/王	10 574	DU/杜	4765
ZHANG/张	12 279	ZHANG/张	9142	LIU/刘	4215
LIU/刘	11 990	CHEN/陈	7788	LI/李	4160
ZHOU/周	10 963	LIU/刘	7775	YANG/杨	3451
CHEN/陈	10 145	HUANG/黄	4453	ZHANG/张	3137
YANG/杨	7201	ZHOU/周	4021	HU/胡	2689
HUANG/黄	5431	YANG/杨	3750	XU/许	2567
DU/杜	5348	ZHU/朱	3108	CHEN/陈	2357
XU/许	5189	PENG/彭	2818	JIN/金	2274
HU/胡	5033	XIAO/肖	2690	YE/叶	2207
XIAO/肖	4686	XU/许	2622	WANG/王	2171
PENG/彭	3905	XIONG/熊	2434	XIAO/肖	1996
ZHU/朱	3752	HU/胡	2344	ZENG/曾	1544
JIN/金	3431	XU/徐	2292	YIN/尹	1179
ZENG/曾	3189	LUO/罗	1815	PENG/彭	1087
XIONG/熊	3003	DENG/邓	1802	ZHAO/赵	1073
XU/徐	2857	GUO/郭	1716	CHANG/昌	1048
LUO/罗	2677	YU/余	1690	HUANG/黄	978
YE/叶	2671	WU/吴	1651	LUO/罗	862

Note: Since there are many homophones in Chinese, Chinese characters are marked after English surnames in order to distinguish them.

ral areas are not among the 20 most common surnames in Xiantao. Further comparison shows that the population sample in rural areas is much larger than that in urban areas, so the distribution of surnames in rural areas has a greater impact on the overall distribution than that in urban areas. The distribution of surnames in urban areas was more different from the overall distribution of Xiantao than in rural areas. On the one hand, more importantly, it does reflect the difference in the distribution of surnames between urban and rural areas. For one thing, surnames with high frequency are more concentrated in urban areas than in rural areas. In urban areas, the 20 most common surnames comprised 50 702 individuals, or 70%, while in rural areas there were 86 031 individuals, or 57%. On the other hand, the distribution of the 20 most common surnames is significantly different between urban and rural areas. ZHOU/周 and DU/杜, these 2 most common surnames in urban areas, ranked 7th and 64th in rural areas, respectively. However, LI/李 and WANG/王, the most common surnames in rural areas, ranked 4th and 12th in the urban areas, respectively. YIN/尹, ZHAO/赵 and CHANG/昌 in urban areas, ranked 65th, 26th and 95th respectively in rural areas. DENG/邓, GUO/郭, YU/余 and WU/吴 in rural areas, ranked 28th, 24th, 29th and 22nd, respectively, in urban areas. Most of the above surnames are brought by immigrants who migrated from Jiangxi to Xiantao during the movement of "Jiangxi Filling Huguang" in the late Yuan and early Ming dynasties. After about six hundred years of reproduction, these surnames have developed into common surnames in Xiantao, especially

ZHOU/周, which surpassed LI/李, the most common surname of Xiantao, to become the most common surname in urban areas.

3.3 Isonymy parameters in Xiantao

3.3.1 Fisher's α

The value of α is 38.49 for the five towns of Xiantao as a whole, higher than the average α value of 19.89 for 105 villages, which is attributed to the "Prefecture Effect", named by SCAPOLI et al.[22] which means that for the same area and population, F_{ST} is smaller in larger areas and populations as compared to smaller subdivisions. Compared with the average α of 25 for Chinese counties calculated according to LIU et al.[7] the value of Xiantao is larger, consistent with the conclusion of LIU et al.[7] that the isonomy of the mid-lower reaches of the Yangtze River is very low. It is closely related to the fact that Hubei Province used to be the transit station of large-scale immigrations and the frequent migration phenomenon of flood disasters. The frequent population migration accompanied by the flow of surnames resulted in a higher degree of differentiation of surnames in this region.

The α values of each town and village are shown in Tab. 1 and Fig. 3. Among the five towns, the highest α value is in Tonghaikou, 37.59, followed by Miancheng, 36.41. The value of Gongyeyuan is the lowest, 21.95. The results reflect that the degree of differentiation of surnames in Tonghaikou and Miancheng is higher than that in Shahu and the urban areas, which is mainly influenced by ethnic composition. In Tonghaikou and Miancheng, there are a large number of Hui people and other minority populations, such as the Tu family and the Manchu. Miancheng is the only district-level Hui town in Hubei Province at present, and the Hui people account for one third of the town's total population, mainly distributed in Minzulu, Qilicheng, Honghuadi, Qihong, Jiangbei and other places in Miancheng. From the end of Yuan Dynasty to Qing Dynasty, Hui immigrants moved in Miancheng, resulting in many surnames with distinct Hui characteristics, such as "WEI/魏, DA/答, DING/定,

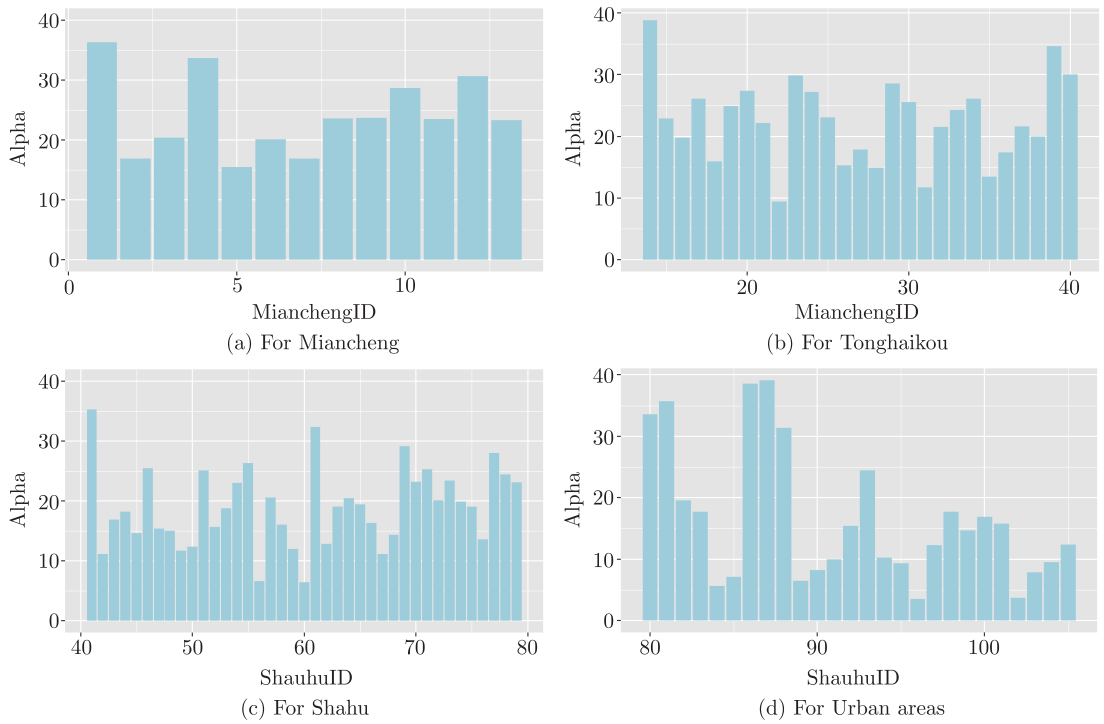


Figure 3 Histogram of α (Gongyeyuan and Shazui sub-district)

HA/哈, MI/米 and BU/卜", etc, which are very rare among Han surnames. Among them, WEI/魏, the surname of Hui immigrants from Shuntianfu (now Beijing) is the most frequent surname among Hui people in Xiantao, accounting for about a quarter of Xiantao Hui people. LI/李, originally from Shanxi (or Shaanxi in other words), is the second most common surname among Hui people in Xiantao. In addition, there are some Manchu surnames, such as "JIN/金, WANG/王, FEI/费, MA/马, FU/傅, ZHANG/章, LANG/郎, GUAN/关 and SHU/舒"[24-30]. These ethnic minority surnames are different from Han surnames, increasing the heterogeneity of surnames in the two towns. In the other three towns, the minority population is small and the degree of differentiation of surnames is lower.

3.3.2 Isolation by distance

Isolation by distance was studied through the correlation of geographic with surname distances at the village level. We found that Euclidean, Lasker's and Nei's distances were weakly correlated with the linear geographical distances, with $r = 0.177 \pm 0.012$, $r = 0.054 \pm 0.013$, $r = -0.054 \pm 0.011$, respectively. Given the high correlation between the three measures of distance ($r_{[\mathrm{Euclidean-Nei}]} = 0.798 \pm 0.005$, $r_{[\mathrm{Nei-Lasker}]} = 0.646 \pm 0.008$, and $r_{[\mathrm{Euclidean-Lasker}]} = 0.532 \pm 0.01$), we used Euclidean distance for the subsequent analysis,which had the largest correlation. The scatter diagram of Euclidean distance and geographical distance between villages is given in Fig. 4. The minimum Euclidean distance is obtained between Chuangyelu in Gongyeyuan and Shazui sub-district, with 0.182 and 3.94 km apart. The 10 most common surnames in Chuangyelu are LIU/刘, LI/李, ZHANG/张, CHEN/陈, WANG/王, ZHOU/周, HU/胡, XU/许, YANG/杨 and PENG/彭, while the 10 most common surnames in Shazui are LIU/刘, ZHANG/张, LI/李, WANG/王, CHEN/陈, HU/胡, YANG/杨, ZHOU/周, XU/许 and XIAO/肖. There is only one difference among the 10 most common surnames in the two places, indicating the high similarity of surnames in these two places. The maximum Euclidean distance is between Gaofeng Village of Shazui and Haifeng Village of Tonghaikou, with 0.798 and 28.75 km apart. The population and surnames of these two villages are very small, with great differences in the distribution of surnames. For example, the most common surnames in Gaofeng Village are XIANG/向 (144) and YAN/鄢 (82), not common in Xiantao, while the most frequent surnames in Haifeng Village are WANG/王 (10) and HUANG/黄 (8), which are common surnames in Xiantao. Due to the clear difference in the distribution of surnames, there is a large isonymic distance between Gaofeng Village and other villages.

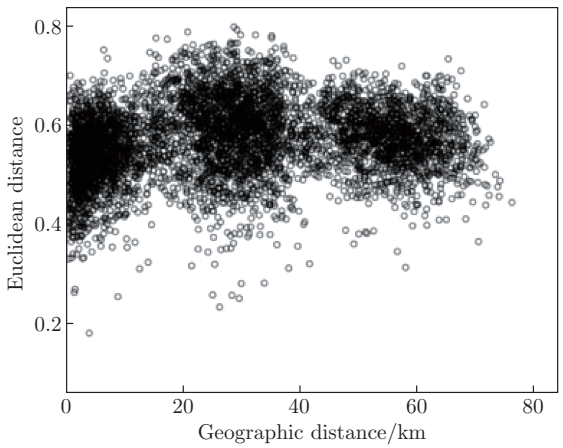


Figure 4 Variation of Euclidean distance between 105 villages with the geographic distance in Xiantao

The signal extracted from the scatter diagram of Euclidean distance over kilometers for villages is given in Fig. 5. Euclidean distance between adjacent or close villages is low, and it increases with the increase of geographical distance, that is, kinship decreases with the increase of distance[1]. Especially, the isonymic distance between urban and rural areas reaches the maximum, reflecting the large difference of surnames between urban and rural areas, and then decreases slightly with the increase of distance, indicating the flow of population between villages. Generally speaking, the distance of surnames within a town is relatively low, but the Euclidean distance between villages located in Shahu Town varies greatly in the east-west direction due

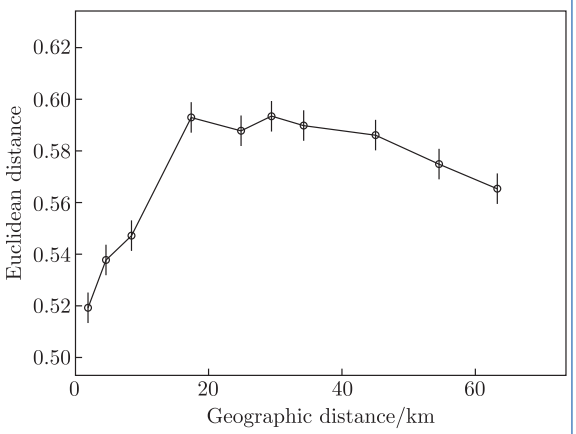


Figure 5 Signal extraction from the variation of Euclidean distance (±SD) over kilometers between 105 villages

to the large east-west span of Shahu.

3.3.3 Relations between the villages

The dendrogram and PCA constructed from the matrix of Euclidean distance between villages are given in Fig. 6 and Fig. 7, respectively, which both indicate two main clusters: A and B. A indicates villages located in the urban area, and B mainly indicates villages in the rural areas, indicating the significant difference in the distribution of surnames between the urban and rural areas. The results of the dendrogram and PCA identified a main surname differentiation between the urban and rural areas. The clustering results show that the villages in the same cluster are almost conterminous, indicating that Euclidean distance is correlated with the geographic distance. There are many outliers in the figure, such as Gaofeng Village, Yanggang Village, Yuji Village and Qunhe Village, etc. The distribution of surnames in these villages is significantly different from that in other villages. In addition, the outliers may be due to the small sample size of some villages.

4 Conclusion and Future Work

Surnames in Xiantao in general reflect the characteristics of Chinese surnames. The distribution of the 10 most common surnames in Xiantao is basically the same as that in Hubei Province and even in China, only individual surnames have differences. The most common surname in Xiantao is LI/李. The distribution of high-frequency common surnames is roughly the same in different areas, with slight differences. For example, the most frequent surname in rural and urban areas is LI/李

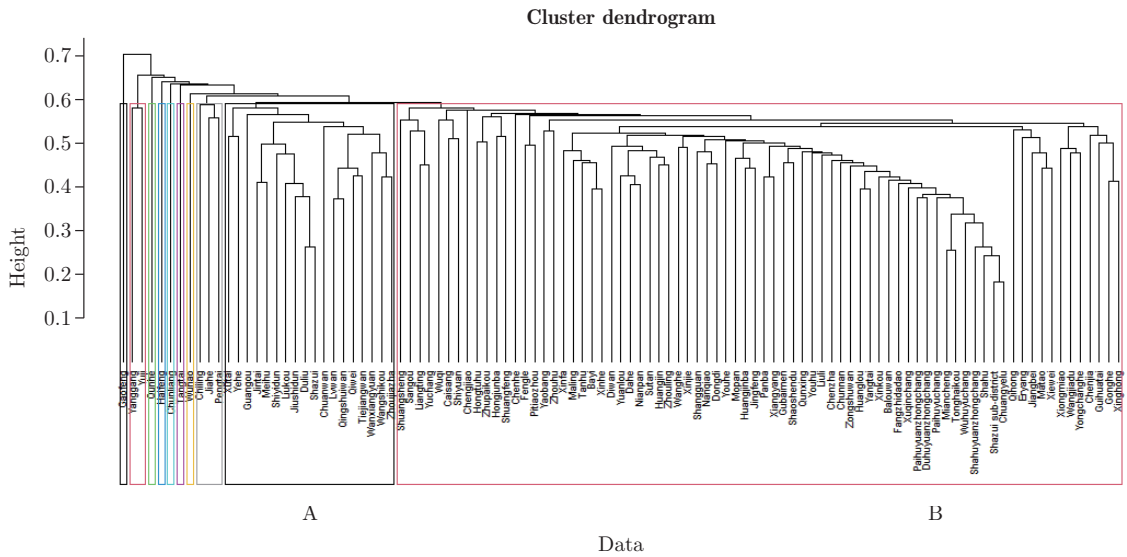


Figure 6 Dendrogram obtained from the matrix of Euclidean distance

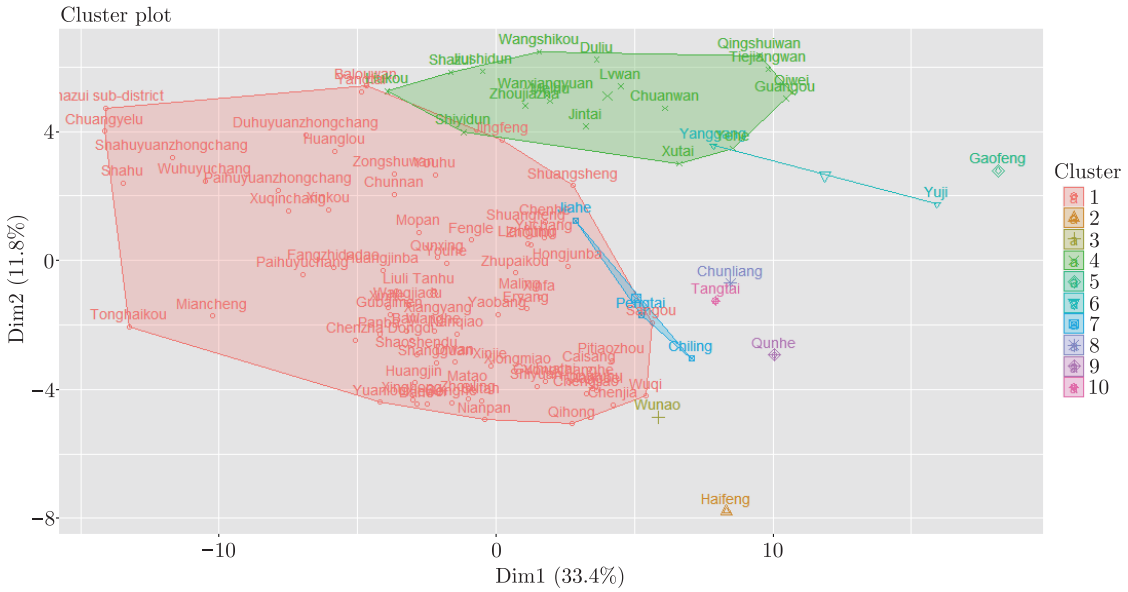


Figure 7 Projection of the Euclidean distance matrix between villages on the first two factors of PCA

and ZHOU/周, respectively.

In addition, the ethnic composition of the population is also an important factor affecting the distribution of surnames. Miancheng Town and Tonghaikou Town, which have a large Hui population, have significantly higher α-values and internal differentiation of surnames than the other three towns.

On the whole, Xiantao is located in Jianghan Plain, with small geographical barriers, the isonymic distance between towns is relatively close and the homogeneity of surnames is high. Moreover, the distance between the surnames of the five towns in Xiantao increases with the increase of geographical distance in general. The isonymic distance between urban and rural areas is greater than that within urban areas or within rural areas. The dendrogram and PCA built from the Euclidean distance matrix showed high heterogeneity of surnames between urban and rural areas.

The present work has only obtained surname data for five towns, but not for other towns in Xiantao or other cities in Hubei, which is not conducive to comparative research. In the future, we will continue to carry out field surveys, add more surname data, and use ArcGIS and other softwares to draw surname distribution maps, in order to reflect the geospatial distribution characteristics of surnames in medium- or small-sized towns in Hubei Province more comprehensively, and to provide ideas for the study of surnames in other regions.

Acknowledgments

The authors are grateful to the Xiantao Public Security Bureau for providing and allowing the use of these data for this study. We would like to thank Professor TANG Xingchun from the School of Life Sciences of Hubei University for his critical revision of the first draft. Thanks to ZHANG Xiaofei, Master of the School of Resource and Environmental Sciences of Wuhan University for helping to calculate the latitude and longitude conversion and the geographical distance between villages.

References

[1] BARRAI I, RODRIGUEZ-LARRALDE A, MAMOLINI E, et al. Isonymy and isolation by distance in Italy[J]. Human Biology, 1999, 71(6): 947-961.

[2] YUAN Yida, ZHANG Cheng. Chinese surnames:

population genetics and population distribution[M]. Shanghai: East China Normal University Press, 2002.

[3] WU Jie, YANG Jianchun. Zhang, Wang, Li, Zhao who is the most: the 2010 census surname structure and distribution characteristics[J]. China Statistics, 2014(6): 21-22.

[4] ZHANG S Y. An analysis and statistics of surname and name[M]//YUAN C. Quantitative Analysis of Modern Chinese. Shanghai Education Press, 1989.

[5] DOU Xuetian. Dictionary of Chinese ancient and modern surnames[M]. Jiangxi People's Press, 1997.

[6] YUAN Yida, QIU Jiaru. The dictionary of Chinese surnames[M]. Nanchang: Jiangxi People's Press, 2010.

[7] LIU Yan, CHEN Liujun, YUAN Yida, et al. A study of surnames in China through isonymy[J]. American Journal of Physical Anthropology, 2012, 148(3): 341-350.

[8] ZHU Y. A study of personal names in Xiantao[D]. Wuhan: Wuhan University, 2018.

[9] ZHAO Yuyao. Spatial distribution and influencing factors of native ethnic villages and surnames in Pingguo county[D]. Changsha: Hunan Normal University, 2019.

[10] SONG Chuanyin. An analysis of the migration characteristics of Hubei from Qin to Qing Dynasty[J]. Wuhan University Journal (Humanity Sciences), 2013, 66(5): 107-113.

[11] RELETHFORD J H. Estimation of kinship and genetic distance from surnames[J]. Human Biology, 1988, 60(3): 475-492.

[12] BARRAI I, SCAPOLI C, BERETTA M, et al. Isonymy and the genetic structure of Switzerland I. The distributions of surnames[J]. Annals of Human Biology, 1996, 23(6): 431-455. DOI: 10.1080/03014469600004672.

[13] BARRAI I, RODRIGUEZ-LARRALDE A, MAMOLINI E, et al. Elements of the surname structure of Austria[J]. Annals of Human Biology, 2000, 27(6): 607-622. DOI: 10.1080/03014460050178696.

[14] RODRIGUEZ-LARRALDE A, SCAPOLI C, BERETTA M, et al. Isonymy and the genetic structure of Switzerland. II. Isolation by distance[J]. Annals of Human Biology, 1998, 25(6): 533-540. DOI: 10.1080/03014469800006772.

[15] CAVALLI-SFORZA L L, EDWARDS A W F. Phylogenetic analysis: models and estimation procedures[J]. Evolution, 1967, 21(3): 550-570. DOI: 10.1111/j.1558-5646.1967.tb03411.x.

[16] NEI M. The theory and estimation of genetic distance[M]//MORTON N E. Genetic Structure of Populations. Honolulu: University of Hawaii Press, 1973: 45-54.

[17] RODRIGUEZ-LARRALDE A, BARRAI I, NESTI C, et al. Isonymy and isolation by distance in Germany[J]. Human Biology, 1998, 70(6): 1041-1056.

[18] DIPIERRI J E, ALFARO E L, SCAPOLI C, et al. Surnames in Argentina: a population study through isonymy[J]. American Journal of Biological Anthropology, 2005, 128(1): 199-209.

[19] FOX W R, LASKER G W. The distribution of surname frequencies[J]. International Statistical Review, 1983, 51(1): 81-87. DOI: 10.2307/1402733.

[20] CHESHIRE J A. Population structure and the spatial analysis of surnames[D]. London: University College London, 2011.

[21] SCAPOLI C, GOEBL H, SOBOTA S, et al. Surnames and dialects in France: population structure and cultural evolution[J]. Journal of Theoretical Biology, 2005, 237(1): 75-86.

[22] SCAPOLI C, MAMOLINI E, CARRIERI A, et al. Surnames in Western Europe: a comparison of the subcontinental populations through isonymy[J]. Theoretical Population Biology, 2007, 71(1): 37-48.

[23] ZHANG Shuyan. Name, Chinese characters, norms[M]. Beijing: Broadcasting Institute Press, 2004.

[24] DA Zhenyi. A preliminary study on the origin and development of the Hui nationality in Hubei[J]. Social Sciences in Ningxia, 1986(1): 17-24.

[25] United Front Work Department of Mianyang County CPC Committee. National records of Mianyang county[M]. [S.l.]:[s.n.], 1987.

[26] Xiantao Local Chronicles Compilation Committee. Mianyang county chronicles[M]. Wuhan: Central China Normal University Press, 1989.

[27] People's Government of Miancheng Hui Town, Xiantao City. Miancheng chronicle, Xiantao city[M]. [S.l.]:[s.n.], 1990.

[28] Mianyang City Annals Compilation Committee. Mianyang city annals (Updated Edition)[M]. Wuhan: Hubei Science and Technology Press, 2000.

[29] LI Anhui, WANG Shengyun. The study of the change of surname Da in Hui ethnic group in Zhong-xiang from the perspective of ethnology[J]. Journal of Hui Muslim Minority Studies, 2010, 20(2): 89-94.

[30] LIU Jinzeng. Population distribution and economic characteristics of Hui nationality in Hubei province in the Qing dynasty[J]. Hui Studies, 2019, 29(2): 37-42.