Surname Distributions and Their Association with Y-Chromosome Markers in the Aleutian Islands

Orion M Graf
Mark Zlojutro
Rohina Rubicz
Michael H Crawford

Human Biology, Volume 82, Numbers 5-6, October-December 2010, pp. 745-757 (Article)

Published by Wayne State University Press

For additional information about this article
http://muse.jhu.edu/journals/hub/summary/v082/82.5-6.graf.html
Surname Distributions and Their Association with Y-Chromosome Markers in the Aleutian Islands

ORION M. GRAF,1 MARK ZLOJUTRO,2 ROHINA RUBICZ,2 AND MICHAEL H. CRAWFORD1

Abstract We examine surname distribution, origin, and association with Y-chromosome haplogroups in native communities from the Aleutian archipelago. The underlying hypothesis is that surnames and Y-chromosome haplogroups should be associated because both are paternally inherited markers. We used Lasker’s coefficient of relationship through isonymy ($R_a$) to identify the distribution of 143 surnames in the Aleutian Islands. The geographic distribution of surnames was explored both through frequency distribution and through the use of Mantel tests. Multidimensional scaling, chi-square, and Mantel tests were used to examine the relationship between surname and Y-chromosome markers. Overall, we observed that the distribution of surnames in the Aleutian archipelago is culturally driven rather than being one of paternal inheritance. Surnames follow a gradient from east to west, with high frequencies of Russian surnames found in western Aleut communities and high levels of non-Russian surnames found in eastern Aleut communities. A nonsignificant correlation ($r = -0.0132; P = 0.436$) was found between distance matrices based on haplogroups of the nonrecombining portion of the Y chromosome and surnames, although an association was found between non-Russian surnames and the predominantly non-Russian haplogroups (R1b, I1a, and I).

We examine the distribution and origins of surnames and their association with Y-chromosome haplogroups in native communities from the Aleutian archipelago. As a proxy for genetic markers, surnames have been used as a nonmolecular means of characterizing the genetic structure of populations and their levels of inbreeding. The first such study was conducted by George Darwin, the son of Charles, who used surnames to estimate frequencies of consanguineous marriages in Victorian England (Darwin 1875). This simple approach was later elaborated and extended by Crow and Mange (1965) to include the calculation of inbreeding coefficients, a method referred to as isonymy, based on the frequency of marriage between individuals with the same surname (Crow 1983).
Advances in molecular biology have led to the increased use of genetic variation to better understand human migration and population structure. The field of anthropological genetics has experienced an increase in studies using mitochondrial DNA (mtDNA) and the nonrecombining portion of the Y chromosome (NRY) to examine population structure and human migration during the latter half of the 20th century (Crawford 2007a). Markers from mtDNA and the NRY are haploid and eliminate the confounding effects of recombination. This makes them ideal for population studies. Mitochondrial DNA is inherited maternally, meaning that it is passed from a mother to her offspring (Giles et al. 1980). This offers a method for identifying a maternal lineage. In the same way, Y chromosomes are inherited paternally, offering a male-specific lineage. Because both NRY markers and surnames are inherited paternally in most societies, they are assumed to be associated with one another (Jobling 2001).

Studies comparing Y-chromosome markers and surnames have yielded useful results. In a study conducted in Great Britain, the so-called Sykes Y chromosome has been shown to correspond to the surname of the same name with few mutational differences for the last 700 years (Sykes and Irven 2000). Of those sampled \( (n = 48) \), almost half shared the same Y-chromosome haplotype, suggesting a single surname founder for extant Sykes males (Jobling 2001; Sykes and Irven 2000). In similar research, performed for Irish subjects, surnames were categorized on the basis of kinship systems, such as clans, and their putative origins, both regionally within Ireland and in foreign countries. The results showed a significant association between Irish surnames and both Gaelic (Ulster, Muster, Leinster, and Connaught) and foreign (Scottish, Norman/Norse, and English) origins and specific NRY haplogroups (Hill et al. 2000).

Given the reported concordance between Y-chromosome variation and surnames, we investigated whether this relationship is evident among Aleut communities in southwestern Alaska, a subdivided island population whose genetic structure exhibits significant geographic patterning and high levels of admixture stemming from substantial male-mediated gene flow from Russian and Western European groups. The genetic structure of Aleut populations reflects three historic periods: the peopling of the Aleutian Islands, the Russian period, and the post-Russian period. An ancestral population crossed Beringia before dividing into two groups: Aleuts and Eskimos (Laughlin et al. 1979). Of these groups, it has been suggested that Eskimos traveled further into North America and that Aleuts diffused westerly through the Aleutian archipelago (Crawford 2007b; Rubicz 2007; Zlojutro et al. 2006).

The Russian period (1741–1867) was marked by an influx of paternal genetic material, great demographic change (depopulation), and forced relocation of Aleuts to the Pribilof and Commander islands. Surnames were first introduced to Aleut peoples during this period through baptism, adoption, and admixture (Black 2004). After the Russian period (1867–present), British and Scandinavian fishermen admixed with Aleuts in the eastern islands in part because of the hunting policies of the U.S. Treasury Department. A man was allowed to hunt fur-bearing
animals only if he was married to a native woman [Reedy-Maschner 2010 (this issue)].

Our primary focus in this study is to describe the distribution and origins of surnames in the Aleutian archipelago. The underlying hypothesis is that surnames and Y chromosomes should be linked because both are paternally inherited. Surname origins were assessed based on historical data, and these were compared to molecular data collected for the Aleut Research Program (1999–present).

**Materials and Methods**

**Population Samples.** One hundred forty-three surnames were used to explore the distribution and origin of surnames in the Aleutian archipelago. Three of us (Crawford, Rubicz, and Zlojutro) collected DNA samples and demographic information during field investigations, between 1999 and 2005, from 11 locations: Attu, Akutan, Bering Island, False Pass, King Cove, Nelson Lagoon, Nikolski, St. George and St. Paul islands, Sand Point, and Unalaska. Participant surnames were collected using demographic questionnaires; the spellings were verified using telephone directories during the sampling process. An expanded sample (n = 732), assembled using surnames from participants’ spouses, parents, and grandparents, was used in comparison with a geographic distance matrix in this study. Mitochondrial and NRY haplotype data from these samples have been previously analyzed and described (Crawford 2007b; Rubicz 2001, 2007; Rubicz et al. 2003; Zlojutro 2008; Zlojutro et al. 2006, 2009).

**Analytical Procedures.** Surnames were assigned to geographic location or population origins based on historical data and region-specific surname databases. The historical record indicates that most paternal gene flow into the Aleutian Islands came from Western Europe, Northern Europe, and Russia, so surnames were assigned as originating in one of these three geographic regions. Fishermen from Western Europe were largely from the United Kingdom; those from Northern Europe were mostly of Scandinavian origin. Russia was the first European power to colonize the Aleutian Islands. Surnames with minimal spelling differences were grouped together. This is standard practice in surname analysis (Fuster 1986; Küchemann et al. 1979; Madrigal and Ware 1997; Pettener 1990; Pollitzer et al. 1988; Rojas-Alvarado and Garza-Chapa 1994; Sanna et al. 2001).

Each NRY haplogroup was assigned to one of the three ethnic groups as well, with the addition of a category for haplogroups found at high levels in native North Americans. They were organized on the basis of current Y-chromosome haplogroup frequencies for each location (Alonso et al. 2005; Karlsson et al. 2006; Rootsi et al. 2004; Semino et al. 2000, 2004) and the historical record. For example, haplogroups I1a and I were classified as being Northern European in origin because they are found at their highest levels (30–40%) in this region. Haplogroup R1b was designated Western European because of its high incidence throughout the United Kingdom (60–80%). Haplogroups R1a and N were listed
as predominantly Russian because combined they make up nearly 70% of the haplotypic diversity in Russia. Historical records verify that most of the paternal admixture in the Aleutian Islands came from Russia, Western Europe, and Northern Europe. A fourth category included those with predominantly Native American haplogroups Q and Q3. These haplogroups represent approximately 80% of haplotypic diversity in Native American groups (Hammer et al. 2001). The remaining haplogroup data from these samples are at high levels in the Balkans, Central Asia, and the Mediterranean and were excluded from the surname/Y-chromosome cross-tabulation (Marjanovic et al. 2005; Semino et al. 2000, 2004). Of these, NRY haplogroup E3 (2.8%) occurred least frequently in this data set, followed by haplogroup J (4.9%). Native American haplogroups Q and Q3 constitute approximately 15% of the data.

The expected hypothesis for each group is that most of the individuals with NRY haplogroup R1a or N would have Russian surnames. It follows that most individuals with NRY haplogroup I1a or I would have Northern European surnames, and those with NRY haplogroup R1b would have Western European surnames. Any deviation from this prediction suggests the absence of association between the two variables. The significance of this deviation was examined using contingency tables and chi-square tests. A fourth category was designated as Native American in origin. These were individuals with NRY haplogroups Q or Q3. The prediction for this category was that all individuals would have Russian surnames. This was based on historical data. Aleuts had no formal surname system before Russian contact and took surnames through admixture, baptism, and adoption.

Distance matrices were calculated from Y-chromosome STRs, geographic data, and a correlation matrix of Lasker’s coefficients ($R_{ib}$). Lasker’s coefficient of relationship by isonymy ($R_{ib}$) was used to evaluate the relationship between populations. This is based on the assumption that individuals with shared surnames are more closely related than individuals without a shared surname (Biondi et al. 1990; Colantonio et al. 2003; Holloway and Sofaer 1989; Küchemann et al. 1979; Raspe and Lasker 1980). The elements are given as

$$R_{ib} = \frac{\sum S_{i1}S_{i2}}{2n_1n_2},$$

where $S_{i1}$ and $S_{i2}$ are the numbers of the $i$th surname in population 1 and population 2, respectively, and $n_1$ and $n_2$ are the total numbers of people in each population, respectively. By using this method, we were able to formulate a matrix of relationship by isonymy between Aleut communities based on surnames.

Nonmetric multidimensional scaling was performed using Statistica (Statistica-Statsoft, version 4.0) on a correlation matrix of Lasker’s coefficient ($R_{ib}$). Mantel tests were used to test the significance of the correlation between two or more matrices. In this case, it was used to evaluate the relationship between genetics, surnames, and geography in the Aleutian Islands. A Mantel test
(Mantel 1967; Smouse et al. 1986) was used to test the significance of the relationship between matrices using Y-chromosome STRs, surnames, and geographic data. The relationship between surnames and geography was analyzed using both the original and expanded sample sets. An \( R_{ST} \) matrix was derived from Y-chromosome STR data using the program Arlequin, version 3.1 (Schneider et al. 2000). A geographic distance matrix (km) was computed using GEOG, version 2.1 (Relethford 2000). Mantel tests were performed using Mantel, version 3.1 (Relethford 1990).

**Results**

We examined the relationship of surname and NRY haplogroup origins using cross-tabulation and chi-square tests. Surnames were ordered in lists based on a shared haplogroup. The initial list included surnames found in conjunction with Native American haplogroups Q and Q3 (Hammer et al. 2001). Most of these surnames (85.7%) were Russian in origin. Three non-Russian surnames were highlighted as differing from the expected group. The second list consisted of surnames found in conjunction with haplogroups I and I1a. In this case, surnames were expected to be non-Russian European, so Russian surnames deviated from expectation. This was also the case when Russian surnames were found in conjunction with the predominantly Western European haplogroup R1b. The last category included haplogroups R1a and N. These haplogroups are found at high frequencies in Russia (Alonso et al. 2005; Karlsson et al. 2006; Rootsi et al. 2004; Semino et al. 2000, 2004). Non-Russian European surnames were highlighted as deviating from expectation. The highest percentage of Russian surnames was found in the Native American pairing, followed by the Russian group (66.7%). Both European groups were composed of approximately 40% Russian surnames. Two additional groups were organized to represent the total haplogroup diversity for Russian and non-Russian European surnames. The Russian total included haplogroups R1a, N, Q, and Q3, and the European total included haplogroups R1b, I1a, and I. Non-Russian European surnames were deviations from expectation for the Russian group, and Russian surnames were deviations from expectation for the European total. Haplogroups Q and Q3 were included in the Russian total because of their association with predominantly Russian surnames.

We used contingency tables and chi-square tests to test the association between each surname and haplogroup pairing. A significant association \((P < 0.005)\) was found between the Russian total and Russian surnames. The same association \((P < 0.001)\) was also found between the European total and non-Russian European surnames. This association held up between Russian surnames and Native American haplogroups Q and Q3 at \( P < 0.05 \) but was found to be nonsignificant between the remaining groupings. Surprisingly, the weakest association was between Russian surnames and haplogroups R1a and N. This is most likely the result of a large number of Russian surnames that were paired with non-Russian haplogroups.
Multidimensional scaling based on a distance matrix constructed from Lasker’s coefficient \( (R_{gb}) \) indicates that Bering Island and False Pass are the most divergent communities. Bering Island shared surnames with no other community in this study. False Pass shared surnames with King Cove. In an expanded sample set \( (n = 732) \), Bering Island shared surnames with the community of Unalaska; False Pass was not nearly as isolated. Unalaska, St. George, and St. Paul cluster together. Akutan is slightly divergent but notably closer to this cluster than the remaining communities. The eastern communities of King Cove and Nelson Lagoon are grouped closely together with Sand Point in close proximity (Figure 1). Nikolski lies as an intermediate community between east and west.

A Mantel test showed a nonsignificant negative correlation between matrices based on Y-chromosome and surname data (Lasker’s \( R_{gb} \)) \( (r = -0.0132; P = 0.436) \). This is a surprising result, because a stronger correlation was expected between these two paternally inherited markers. A nonsignificant correlation \( (r = 0.470; P = 0.104) \) was found between Lasker’s \( R_{gb} \) and geography; however, this correlation was found to be significant in the expanded sample set \( (r = 0.471; P = 0.032) \). These results are more comparable to the correlation between the maternally inherited mtDNA and geography \( (r = 0.717; P < 0.001) \) than the correlation between Y-chromosome markers and geography \( (r = -0.149; P = 0.796) \) (Zlojutro 2008). A Mantel test run between Lasker’s \( R_{gb} \) and an \( F_{ST} \) distance matrix from mtDNA data resulted in a nonsignificant correlation of 0.388 \( (P = 0.150) \).

Figure 1. Multidimensional-scaling plot based on a distance matrix from Lasker’s coefficient of relationship \( (R_{gb}) \). Stress = 0.0981013.
Unlike the findings of Sykes and Irven (2000) and Hill et al. (2000), Aleut surnames and their ethnic origins are not significantly associated with NRY haplogroups, despite the fact that both systems are mostly transmitted paternally between generations. Mantel tests showed a nonsignificant negative correlation between surnames and Y chromosomes, suggesting a lack of association between the two markers. Multidimensional scaling revealed Bering Island to be a significant outlier when using surname data (Figure 1); however, Bering Island clusters with Akutan and Sand Point in a multidimensional scaling derived using Y-chromosome markers (Figure 2). Further discordance is evident in the fact that Nelson Lagoon and Nikolski are significant outliers using Y-chromosome data (Figure 2) but not when using surname data (Figure 1). Chi-square tests showed an association between surnames and haplogroups, but only when they were compiled as two large groups (European total and Russian total), and with the Native American category. When separated into more specific arrangements (Western European, Northern European, and Russian), this association broke down.

Based on the multidimensional-scaling plot for $R_{ST}$, there is evidence for geographic patterning. Along the first dimension the eastern communities are separated out from the more westerly ones, with the exception of Bering Island, which is an
outlier. This pattern likely reflects the higher frequency of non-Russian surnames in this region versus the higher frequency of Russian surnames in the western communities and in the Pribilof Islands. This is also reflected by a significant Mantel correlation between geography and $R_g$ in the extended sample set of 732 surnames. The clustering of Unalaska, St. Paul, and St. George is probably the result of the Pribilof Islands being established through forced relocation from Unalaska (Black 2004; Elliott 1886). Aleuts conscripted from Atka and Attu populated Bering Island (Lantis 1984; Rychkov and Sheremetyeva 1972). These communities would cluster together, but there was insufficient surname data to include Atka and Attu in this study.

Surname frequency distribution followed a geographic pattern from west to east with a prevalence of European surnames in the eastern islands and Russian surnames in the west. Russian surnames grew in frequency until they were ultimately fixed in three of the western communities (Figure 3). This is a trend that more closely resembles maternal heritage, as shown in mtDNA studies (Crawford 2007b; Rubicz 2007) rather than in Y-chromosome studies (Zlojutro 2008; Zlojutro et al. 2009). Zlojutro et al. (2009) found significant genetic (NRY) discontinuity between the western and eastern parts of the Aleutian island chain, but a geographic distribution based on haplogroup frequencies was not as clear. These results further indicate that surnames have not been inherited along with Y chromosomes in the Aleutian archipelago. Instead, there appears to be a significant cultural aspect to their inheritance.

In fact, most Aleut surnames were obtained through Native baptism into Russian Orthodoxy, not direct paternal inheritance. Those baptized received a baptismal name, usually the name of a saint whose feast fell on or near the day of baptism. Baptismal names were often adopted as first names, whereas the baptismal sponsor’s surname was given to the newly baptized Native as their surname (Black 1997, 2004). The baptismal relationship between Russians and Natives resembled adoptions. Both Natives and Russians took the relationship seriously. This bond was second only to parent-child relationships. Russians sponsoring Natives in baptism were taking on a great deal of responsibility, including paying any costs associated with the event, and they would usually pay for the Natives’ education afterward (Black 1997, 2004).

According to Black (2004), the acceptance of the Orthodox faith was not culturally destructive in Alaska. Native lifestyles were left relatively unchanged. Russian Orthodoxy was not forced on Aleuts, and there was no pressure to use the dominant language or to completely change lifestyles (Black 2004; Dauenhauer 1990; Smith 1980). Native healers were not rejected, and even polygynous marriages were accepted during the early stages of Native conversion (Mousalimas 1989, 1990, 1992, 2003; Oleksa 1987). For these reasons, it was easier for Natives to accept Orthodoxy rather than many other religions. Today, many outsiders even consider the Orthodox church in Alaska to be a Native religion (Davis 1984; Hosley 1966; Rathburn 1981).

Although most Aleut surnames were acquired through Russian Orthodox baptism, a large minority of Aleuts gained surnames through Russian-Native
Figure 3. Surname frequencies in the Aleutian Islands displayed as pie charts.
intermarriage. This was initially encouraged for means of control and practicality but later became a means for increasing social mobility after the establishment of the creole estate. Akin to the burgher estate in the Russian economic system, the creole estate had the added privileges of being charged no taxes and it offered free education in Russia. Therefore marrying Russian and creole men became an avenue that Native women could use to increase their class status by acquiring wealth and education (Black 2004).

Members of the creole class achieved positions of management and leadership throughout the colony. They were teachers, clergy, navigators, cartographers, ship commanders, and artists. By 1863 approximately 2,000 people ascribed to this exclusively Alaskan estate (Black 2004). By the end of Russian-America, this class had been in existence for 46 years. It had encouraged more than a generation of Natives to adopt Russian surnames, serve Russian interests, and seek Russian education with the promise of social advancement.

When Alaska was sold to the United States in 1867, the creole class was dissolved. Rather than being treated with privilege by the United States, creoles were labeled “half-breeds” and “depraved” by military and civilian administrators (Black 2004; Dall 1870). Natives and creoles alike were classified as “Indians” and not given the same rights as U.S. citizens (Black 1990). After 1867, many creoles decided to move to Russia or the American mainland instead of accepting their new status in Alaska. Those who stayed had few employment options, and most were accepted back into the Native community (Black 2004). The exodus of Aleuts with mixed heritage after the sale of Alaska may explain the lack of congruity between Russian surnames and Y-chromosome haplogroups R1a and N that we find today in the archipelago.

After the Russian sale of Alaska in 1867, there was an influx of Scandinavian and Western European fishermen and fur traders. Like their Russian predecessors, these Europeans were encouraged to marry and admix with the Native population. However, unlike Russian-Native admixture, which was promoted by proximity, superstition, and for control, European admixture was a product of regulation. Because of the shortage of sea otters, the U.S. Treasury Department allowed only individuals married to Native women to hunt fur-bearing animals during the late 19th century. In an effort to take advantage of this regulation, Western European and Scandinavian men actively sought and wedded Native Aleut women [Reedy-Maschner 2010 (this issue)]. This produced a Y-chromosome signature in the eastern Aleutians, where this marital trend was most common, resulting in elevated frequencies of non-Russian haplogroups I1a and R1b, with the prevalence of haplogroups R1a and N in the western Aleutian Islands representing the remnants of earlier Russian conquest and occupation (Zlojutro et al. 2009).

Overall, surname distributions appear to follow a culturally derived path rather than one of biological inheritance in the Aleutian Islands. Most Russian surnames occur in the west, whereas non-Russian surnames are prominent in the east. There is a significant association between non-Russian surnames and non-Russian Y chromosomes in the Aleutians as a result of recent historical developments.
Surnames and Y-Chromosome Markers / 755

(i.e., hunting regulations and the corresponding increase in admixed marriages during the 19th century). Second, Native American NRY haplogroups Q and Q3 are significantly associated with Russian surnames. Natives adopting Russian surnames through the baptismal process of the Russian Orthodox faith explains this relationship. The lack of a significant association between Russian surnames and the predominantly Russian NRY haplogroups R1a and N has a more complicated explanation. When creoles were forced out of their prominent role in society after the sale of Alaska in 1867, many chose to leave. The paternal lineage of the creole estate would have been composed of individuals characterized by haplogroups found at their highest frequencies in Russia or North America along with an obligatory Russian surname. By removing these individuals, it is likely that the association between Russian surnames and the predominantly Russian NRY haplogroups R1a and N was weakened. Combined with Orthodox baptism, this offers an explanation for the lack of an association between these two markers.

Acknowledgments  We would like to thank the Aleut Corporation, the Aleutian/Pribilof Association, and the Aleut community. This research was funded by the National Science Foundation through grants OPP-990690 and OPP-0327676.

Received 1 June 2010; accepted for publication 29 June 2010.

Literature Cited


