The origins of the Irish Travellers and the genetic structure of Ireland

KARI E. NORTH†‡, LISA J. MARTIN†‡ and MICHAEL H. CRAWFORD†
†Department of Anthropology, University of Kansas, USA
‡Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, USA

Summary. Ireland’s unique and well-documented history provides insight into the formation and origins of population subdivisions. Of particular interest, is the controversial ethnogenesis of an itinerant population of Ireland: the Travellers. The objectives of this study were: (1) to determine the genetic affinity of the Travellers to the general Irish population based on gene frequency data, subdivided by county, and (2) to explore the relationship between subpopulations of Ireland, given its turbulent history.

The gene frequencies of standard genetic markers collected from populations residing in counties of Ireland and the Travellers were calculated and analysed using several multivariate methods. First, a relationship (R) matrix was used to ascertain the scaled variance–covariance matrix of population similarity. Second, mean per locus heterozygosity (H) was regressed on distance of the region from the gene frequency centroid (r_i). The results of this study include: (1) the confirmation of Crawford’s (1975, in Biosocial Interrelations in Population Adaptations, E. S. Watts et al. (eds), pp. 93–103) conclusions concerning the origins and genetic affinity of the Travellers; (2) based on several multivariate analyses, the major influence on population structure was unique historical events; and (3) Relethford and Crawford’s (1995, American Journal of Physical Anthropology, 96, 25–38) hypothesis concerning the distinctiveness of the midland counties was verified by this study.

1. Introduction

Ireland’s unique and well-documented history provides insight into the formation and origins of population subdivisions. A large amount of gene flow has been documented to and from the island of Ireland, including invasions and successive emigrations of the Celtic, English, Scottish, Norse and Anglo-Normans (Sunderland, Tills, Bouloux et al. 1973, Relethford and Crawford 1995). Although gene flow normally reduces the formation of population subdivisions, in Ireland, cultural differences of the migrants (i.e. religion, ethnicity and economic factors) have increased this potential.

Irish population structure has been analysed using different hierarchical levels of population subdivisions. First, subdivisions of the local population within a single small region (usually a county) has been studied through the use of blood genetics, surname analyses and quantitative traits (e.g. Casey, Hale and Casey 1969, Relethford, Lees and Crawford 1980, 1981, Relethford and Blangero 1990, Bittles and Smith 1994). Second, regional population subdivisions have been documented by examining relationships among major provinces (or regions) using blood genetics and quantitative traits (e.g. Sunderland et al. 1973, Tills 1977, Tills, Teesdale and Mourant 1977). Third, studies of countrywide variation have compared counties using quantitative traits (e.g. Hooton, Dupertuis and Dawson 1955, Relethford 1983, Relethford and Crawford 1995). However, only recent research by Relethford and colleagues examined population structure using multivariate analytical analyses.
Previous research has been inconclusive concerning the determinants of population structure of Ireland. Depending on the level of analysis (i.e. one small region or the entire island), different results have been reported. Relethford and colleagues have suggested that these inconsistencies concerning the population structure of Ireland may not be inconsistencies *per se*. Indeed, they have suggested that the level of analysis in these studies of human variation can explain differences in results (Relethford and Crawford 1995, Relethford et al. 1997, North et al. 1999).

Most research of Irish population structure has not considered the Travellers, an itinerant population in Ireland previously known as the Tinkers. The ethogenesis of the Travellers has been a topic of considerable debate with several different hypotheses proposed to explain their origins. According to MacMahon (1971), the Travellers could have originated from several different groups, including outcasts of prehistoric Irish populations, native chieftains, displaced farmers and labourers, and Romany Gypsies. A study by Crawford (1975), based on genetic distance, suggested an Irish origin of the Travellers. However, Crawford’s study was limited to univariate analysis of gene frequency data. Recent multivariate techniques now enable a more detailed analysis of the ethogenesis of the Travellers and their genetic affinity to Irish populations.

The objectives of this study were (1) to determine the genetic affinity of the Travellers to Irish counties based on gene frequency data, and (2) to explore the relationships among subpopulations of Ireland, given its turbulent history. These goals will be accomplished through the use of several multivariate methods.

2. Population

Archaeological evidence reveals that Ireland was inhabited by successive waves of migration beginning in the immediate postglacial period (Harbison 1976). After this initial influx of people, little immigration to Ireland occurred until the Celtic invasions, beginning in the 6th century BC. Each successive wave of Celtic invaders pushed the earlier inhabitants westward, and by 300 AD the Celtic occupation of Ireland was well established.

In the late 8th century, the Norse Vikings began to settle in Ireland and its previous cultural uniformity was disrupted (Chauvire 1956, Sunderland et al. 1973). The Norse Viking invasions forced the Celts inland, whereas the Vikings established permanent coastal settlements. Over the next several hundred years, the Gaelic Irish for the most part, culturally assimilated the Norse. Those Norse who were not assimilated became genetically isolated in their original coastal settlements by 1014 AD (Hackett, Dawson and Dawson 1956).

The next wave of migrants, the Anglo-Normans, invaded Ireland between 1169 and 1300 (Pringle 1985). Although their migration was fairly scattered, the Anglo-Norse were able to spread throughout the entire country. The Norman conquest was at its peak in the 13th century with 75% of the island controlled by Anglo-Normans. By the mid-15th century, Norman colonization was sparse and the Gaelic Irish had culturally assimilated most Normans.

The next wave of migrants came from Scotland and England, from the 15th to the 17th centuries (Pringle 1985). During this period, the English tried to gain political control of Ireland. Because of political conflicts, the 17th century was characterized by large-scale population movements (Pringle 1985).
The historical pattern of migration to the east coast of Ireland over the centuries, with successive groups displacing previous inhabitants to some extent towards the west, should have resulted in a tendency for an east–west cline in biological variation (Pringle 1985). Moreover, the last group of migrants from England should have the greatest influence in the north and east. Indeed, previous research on the population structure of Ireland supports the presence of an east–west cline (e.g. Relethford and Crawford 1995).

The Travellers are a unique historical group in Ireland, wanderers that travelled in patrilineally related groups of two to four families (Crawford 1975). Traditionally, Travellers travelled within one or two counties, repairing pots and pans in exchange for food or money. Socially, the Travellers were outcasts (Ní Shúinéar 1996), and often not welcome in settled communities for more than the duration of their work (about 2 weeks). This forced social isolation caused the Travellers to be relatively endogamous, and to possibly develop their own language, Gammin. However, the origin of Gammin is the subject of much debate (Binchy 1996).

Beginning in the 1950s, the need for the Travellers’ skills in rural areas declined (Crawford and Gmelch 1974). Specifically, the development of plastics and other metallic substitutes eliminated the traditional occupations of the Travellers, forcing the majority to migrate into urban population centres. In the 1960s, the Irish government implemented a settlement program by constructing several serviced campsites and housing for the Travellers. Nonetheless, in the 1970s, the majority of Traveller families continued to live by the roadside with varying degrees of mobility. Although the Traveller community has modernized, they remain a social isolate having little contact with members of the settled Irish community (Ní Shúinéar 1996). Indeed, the Travellers have maintained their ethnic separateness until the present day, through language, endogamy and broader cultural differences.

The ethnogenesis of the Travellers is a source of great controversy. Ó Síocháin and colleagues summarized several different theories proposed to explain their origins (Ó Síocháin, Ruane and McCann 1996). First, the Travellers could be descendants of Prehistoric Irish populations (MacMahon 1971) or of the early historical period, either from a pre-Celtic group, from one of a number of Celtic groups, or from indigenous itinerant crafts-workers of the early Christian period (Ní Shúinéar 1996). If this is the origin of the Travellers, they should be genetically distinct from other Irish populations, because of the duration of social isolation. Second, the Travellers could be descendants of native chieftains and their families, who were displaced by the English immigration (MacMahon 1971). Again, if this is the appropriate origin of the Travellers, they would be distinct from other Irish subpopulations and from England. Third, the Travellers could be displaced farmers and labourers who were driven from their land by stochastic processes, such as the potato famine (Crawford 1975, Gmelch and Gmelch 1976). If this is the Travellers’ origin, they would show genetic affinity to other Irish populations, because of a reduced period of isolation from other Irish groups. Fourth, the Travellers could have originated very recently, as a challenge to the status quo (Binchy 1996, Kenny 1996). Indeed, according to Kenny (1996), the entire concept of ethnic separateness or ethnic groups is an emergent phenomenon. However, this hypothesis is largely a socio-cultural construct, an argument about the definition of ethnic groups, not about the biological affinity of Traveller populations. Fifth, the Travellers could be a hybrid population, which experienced gene flow with another Gypsy population, the Romany Gypsies (e.g. Acton 1996). If the Travellers
originated from Romany Gypsies, as is the consensus, then their gene frequencies should resemble those of Northern India (Gypsies originated in India). However, Crawford (1975) demonstrated that the Travellers have minimal genetic affinity with Gypsies or Indian populations, thus this theory is unlikely.

3. Materials and methods

3.1. Data

The late Don Tills of the British Museum of Natural History provided the gene frequencies of standard genetic markers from counties of Ireland. Gene frequencies for the Travellers were compiled from Crawford (1975). The Travellers were sampled from individuals on the road in Dublin, County Wicklow and County Wexford and from the government settlements of LaBrea Park, Finglas and Church Town. However, the majority of the sample included Travellers from the government settlements surrounding Dublin. Thus, this sample is representative of Travellers who travel through several counties. Ten blood systems for the Irish and the Travellers were used including: ABO, Rhesus, Duffy, Kell, MNS, P, transferrin, phosphoglucomutase, adenylate kinase and haptoglobin.

With the exception of the Travellers and a sample from the City of Dublin, the basic unit of analysis for this study was the county. The Republic of Ireland has 26 counties (figure 1). Tills separated Dublin City because it is a major population

Figure 1. Map of the Republic of Ireland displaying counties sampled.
centre and the largest city in Ireland (Sunderland et al. 1973). Moreover, there has been extensive in- and out-migration from the City of Dublin (Pringle 1985). The Travellers were treated as a separate population because of their suggested cultural isolation and genetic divergence (Crawford and Gmelch 1974).

In addition to these 28 Irish groupings, several separate European outlier populations were considered for comparative reasons, these were Northern Ireland, Iceland, Norway, Scotland and Tilting, Newfoundland. With the exception of Tilting, these gene frequency data were compiled from Mourant, Kopec, and Domaniewska-Sobezak (1976). The gene frequencies for Tilting (an Irish fishing village from Newfoundland) were compiled by Crawford. Overall, the sample sizes from all populations for all loci considered were greater than 50.

3.2. Analytical methods

A relationship matrix (\( R \)-matrix), was used to ascertain the scaled variance–covariance matrix of population similarity. \( R \)-matrix is a multivariate statistical method of graphically demonstrating relationships between populations by reducing large sets of data onto a two dimensional plot of the eigenvalue scaled by the square of the eigenvector (Harpending and Jenkins 1973). The \( R \)-matrix was reduced through least squares and all axes were examined to determine their information content. Informative axes were then plotted to display population affinities. All groups received equal weight in the \( R \)-matrix computation (Harpending and Jenkins 1973, Workman, Harpending and Laluel 1973). We also experimentally weighted groups by census population sizes and sample size to correct for bias introduced by sampling error. However, these matrices did not significantly modify the results. Therefore, only the results using unweighted matrices are reported.

As an aid to interpretation of the distribution of populations in the \( R \)-matrix plots, plots of the distribution of alleles in reduced space were produced. These \( A \)-matrix plots display the specific alleles responsible for population dispersion.

\( R_{ST} \) values were calculated, from the average value of the diagonal elements \( r_{ii} \) of the \( R \)-matrix. The \( R_{ST} \), also known as the Wahlund \( F \), measures the variation in genetic distances from the subpopulations to the gene frequency centroid of the contemporary populations (Workman et al. 1973). While \( R_{ST} \) is not technically equivalent to Wright’s \( F_{ST} \) (which measures variation relative to the gene frequency centroid in the founding population), \( R_{ST} \) values closely approximate \( F_{ST} \) values.

Mean per locus heterozygosity (\( \bar{H} \)) was regressed on distance from the gene frequency centroid (\( r_{ii} \)). This method was used to evaluate the extent to which local genetic structure results from the local social structure and other influences (Harpending and Ward 1981). This method assumes that the relationship between the genetic distance of a subpopulation from the centroid’s gene frequencies and the relative homozygosities are linear if the exchange with external populations is equal between subpopulations. Moreover, this method assumes that the genetic distance between groups should be small if the gene frequencies are similar. In this method, the interpretation of outliers is the focus, not the fit of the regression line. Therefore, subpopulations that experience above average external gene flow will exhibit greater than expected heterozygosity and fall well above the regression line. In contrast, subpopulations that are more isolated and more susceptible to genetic drift will display lower heterozygosities than expected and fall below the regression line.
4. Results and discussion

4.1. R-matrix analysis

Three R-matrix analyses were completed on these data. The first R-matrix analysis was limited to 28 Irish groups using the 10 available blood systems. Only 35% of the total variation was explained by the first two eigenvectors, suggesting extreme homogeneity and little genetic micro-differentiation among the subdivisions.

The second R-matrix analysis was completed using only the five blood groups (ABO, MNS, Duffy, Kell and P) that were most informative in the initial analysis. Subpopulation relationships were examined through the R-matrix plot (figure 2). Each axis was scaled by the square root of its corresponding eigenvalue to ascertain an accurate assessment of relative dispersion (Harpending and Jenkins 1973). The first two scaled eigenvectors accounted for 51% of the total variation. The first axis separated two major clusters, with no clear geographic pattern. Two main outlier groups were separated by the second axis. One of the outlier groups consisted of Clare, Dublin and the Travellers. Although the Travellers were sampled from three counties (Wicklow, Wexford and Dublin), the fact that they do not cluster with non-Travellers from the same counties, demonstrates the representativeness of the Traveller sample. Overall, this R-matrix analysis demonstrated no association between geography and genetics in Ireland. This finding is in concordance with previous research of the entire country of Ireland, although based on anthropometrics, which identified population history as the main determinant of population structure (e.g. Relethford and Crawford 1995, North et al. 1999).

Figure 2. Relationships between subpopulations in Ireland plotted using the first two scaled eigenvectors of the R-matrix.
Figure 3 is an $A$-matrix, illustrating the relationship between alleles for the Irish subpopulations of figure 2. The Rhesus, Duffy, Kell and P systems are located on the periphery of the matrix and carry the most discriminating power. Moreover, these blood groups provide the most information about the genetic relationships between Irish populations.

In the third $R$-matrix analysis, five outlier subpopulations (Northern Ireland, Iceland, New Foundland, Norway and Scotland) were added. In this analysis, 44% of the total variation was explained by the first two eigenvectors (figure 4). The first axis separates three main clusters of subpopulations. The left cluster was composed of counties that are coastal or slightly inland, the Scottish outlier group and the Travellers (table 1). This cluster may have been affected by the most recent waves of Irish migration, the Scottish and English. Moreover, the Travellers were clustering with heterogeneous counties, suggesting the influence of various Irish subpopulations in the origin of the Travellers. The central cluster includes several midland counties (Roscommon, Longford, Westmeath and Cavan), Norway and Tilting (table 1). The clustering of Tilting with various Irish groups supports Crawford, Koertvelyessy, Collins et al.’s hypothesis (1996) that Tilting is an isolated group of Irish Catholics. In addition, the midland counties (with the exception of Cavan) clustering with Norway supports Relethford (1983) and Relethford and Crawford’s (1995) hypothesis suggesting Viking admixture. The right cluster does not display any clear patterns. The second axis separates Norway and Tilting outlier populations from all other groups.

Figure 5 illustrates the relationship between alleles for Irish subpopulations and other European populations of figure 4. The Rhesus, Duffy, ABO and P systems are
located on the periphery of the matrix; thus carrying the most discriminating power. These blood groups provide the most information about genetic relationships between Irish populations and other outlier populations.

The $R_{ST}$ values obtained from all three analyses, along with $R_{ST}$ values reported for Ireland in previous research were reported in table 2. These $R_{ST}$ values were consistent with those found in relatively homogeneous populations. In addition, the $R_{ST}$ values from all three matrices were in good agreement with the values reported

Table 1. Separation of first eigenvector in $R$-matrix analysis with outliers.

<table>
<thead>
<tr>
<th>Outliers</th>
<th>Right</th>
<th>Middle</th>
<th>Outlier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scotland</td>
<td>England</td>
<td>Iceland</td>
</tr>
<tr>
<td></td>
<td>Northern Ireland</td>
<td>Norway</td>
<td>Tílting, Newfoundland</td>
</tr>
<tr>
<td>Counties</td>
<td>Carlow</td>
<td>Cavan</td>
<td>Galway</td>
</tr>
<tr>
<td></td>
<td>Clare</td>
<td>Cork</td>
<td>Kildare</td>
</tr>
<tr>
<td></td>
<td>Donegal</td>
<td>Kerry</td>
<td>Kilkenny</td>
</tr>
<tr>
<td></td>
<td>Dublin</td>
<td>Laois</td>
<td>Leitrim</td>
</tr>
<tr>
<td></td>
<td>Dublin City</td>
<td>Longford</td>
<td>Limerick</td>
</tr>
<tr>
<td></td>
<td>Louth</td>
<td>Meath</td>
<td>Offaly</td>
</tr>
<tr>
<td></td>
<td>Mayo</td>
<td>Roscommon</td>
<td>Sligo</td>
</tr>
<tr>
<td></td>
<td>Monaghan</td>
<td>Westmeath</td>
<td>Tipperary</td>
</tr>
<tr>
<td></td>
<td>Travellers</td>
<td></td>
<td>Wicklow</td>
</tr>
<tr>
<td></td>
<td>Waterford</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wexford</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Relationships between subpopulations of Ireland and other European populations plotted using the first two scaled eigenvectors of the $R$-matrix.
by Tills (1977) in his analysis of gene frequency variation using the ABO and Rh systems. In contrast, the values were higher than reported by Relethford (1983) using the ABO and Rh systems from a different data set. Sampling variation could have caused the variation in these results.

4.2. **Heterozygosity versus \( r_{ii} \) analysis**

Mean per locus heterozygosity (\( \bar{H} \)) was regressed on distance from the gene frequency centroid (\( r_{ii} \)) (figure 6). The expected and the observed regression line have been indicated in the plot. The Travellers and Longford County had the two highest \( r_{ii} \)'s, 0.0233 and 0.0245, respectively. The high \( r_{ii} \) value for the Travellers demonstrated their relative isolation from the rest of Ireland. In contrast, Louth and Dublin City displayed the lowest \( r_{ii} \)'s, 0.0072 and 0.0063, respectively. The low \( r_{ii} \) value for Dublin City is expected and most likely is due to the large amount of population movement in this urban centre.
Many of the subpopulations deviated significantly from the theoretical and observed regression line (figure 6). For example, Wexford, Westmeath, Roscommon and Cork fell significantly above the regression lines and have high heterozygosity and low $r_{ii}$ values. This suggests that these county populations experienced above average external gene flow. It is possible that these groups maintained the remnants of historical invasions. In fact, Cork and Wexford are coastal communities which have been subjected to continuous migration; beginning with the Vikings and continuing through the Anglo-Norse and English (Pringle 1985). Therefore, Cork and Wexford could be heterozygous as a result of any of these migrations.

Kerry, Monaghan, Cavan and Mayo displayed lower than expected heterozygosities and fell significantly below the regression lines. Monaghan and Cavan may have been more isolated because of their proximity to Northern Ireland. According to Pringle (1985), the conflict between Northern Ireland and the Republic of Ireland has a long history. Therefore, it is possible that living on the border of these two countries has been an isolating mechanism. However, as other border counties (e.g. Donegal, Leitrim and Louth) do not display lower than expected heterozygosities, this may not be the isolating mechanism.

The positioning of the City of Dublin in this plot also deserves attention. Dublin fell on the observed regression line and had a relatively low heterozygosity (figure 4). This finding is interesting because Dublin is the largest city in Ireland and has experienced a high level of gene flow (Pringle 1985). While there is a high level of migration to Dublin City, Humphreys (1966) and Kammeyer (1976) have suggested that few migrant genes are being left behind. In fact, the 1970 census reported a high
rate of single unmarried individuals living in Dublin City (Kammeyer 1976). Moreover, according to Humphreys (1966), fertility in unmarried individuals was low. Therefore, this may explain why Dublin City does not have the characteristic high rate of heterozygosity often found in urban centres.

In the $\bar{H}$ versus $r_{ii}$ analysis, the Travellers had high heterozygosity and low $r_{ii}$, suggesting that the Travellers were a heterogeneous group that has been relatively isolated through time. This finding is in concordance with the general consensus that the Travellers are an Irish social isolate that has differentiated through time (e.g. Ó Síocháin et al. 1996).

5. Conclusion

The turbulent history of Ireland, with its waves of invasions and immigrations, offers an unique opportunity to examine population structure in relation to biological variation and population history. In this study, three main conclusions can be made. First, these analyses support Crawford’s (1975) hypotheses concerning the Irish origins and genetic affinity of the Travellers. Judging from the $R$-matrix analyses, the Travellers are undoubtedly of Irish ancestry, due to their proximity to the centroid. Furthermore, the Travellers clustered with several heterogeneous counties in Ireland, including Wexford and Westmeath. Therefore, these data support that the origin of the Travellers was not a sudden event; rather a gradual formation of populations. Indeed, the Travellers probably originated with craftsmen and artisans forced to leave their monasteries (Crawford 1975). Later, their population grew as they were joined by various Irish groups that were forced to leave their homes because of various calamities and political upheavals (i.e. the potato famine and the repression of British occupation) (Crawford 1975). However, the timing of the Traveller origin is not certain and may have predated the historical period (e.g. Ní Shúinéar 1996).

Second, our analyses suggest that at the level of the country, the major influence on population structure was population history. Although our analyses suggest that Ireland is genetically homogeneous, several distinct subpopulations were identified that may have been formed because of historical gene flow. These findings further confirm that the hierarchical level of analysis has an important impact on the results of studies of the population structure of Ireland (Relethford et al. 1997, North et al. 1999). Indeed, recent research by North and colleagues identified different spatial patterns of biological variation, depending on the level of analysis (North et al. 1999).

Third, the $R$-matrix analyses support Relethford and Crawford’s (1995) hypothesis concerning the distinctiveness of the midland counties due to Viking influence. In this analysis, the midland counties clustered with Norway.

Acknowledgements

We would like to first acknowledge the people in Ireland. Without their participation, this study would not have been possible. Additionally, we would like to recognize Don Tills for providing the gene frequencies of Ireland. Grant no. 2647 from Wenner-Gren Foundation for Anthropological Research funded in part the Traveller portion of this study.
References


Address for correspondence: K. E. North, Department of Genetics, Southwest Foundation for Biomedical Research, P.O. Box 760549, San Antonio, TX 78245–0549, USA.


Résumé. L’histoire particulière et bien connue de l’Irlande offre des aperçus sur les origines et la manière dont se forment les subdivisions de populations. En particulier, est intéressante l’ethnogénèse controversée d’une population itinérante d’Irlande : les « Travellers ». Les objectifs de cette étude sont (1) de déterminer les affinités génétiques des Travellers par rapport à la population irlandaise en général à partir de fréquences géniques subdivisées par comtés et (2) d’explorer les relations entre sous populations irlandaises, compte tenu de leur histoire turbulente. Les fréquences géniques de marqueurs génétiques standard collectées à partir de populations résidant dans les comtés d’Irlande et chez les Travellers, ont été calculées et analysées au moyen de diverses méthodes multivariées. Tout d’abord, une matrice de relation (R) a été utilisée pour établir la matrice variance-covariance de similarité ordonnée. Puis, la moyenne d’hétérozygotie par locus (H) a été régressée sur la distance de la région à partir du centre (r ii) de fréquence génique. Les résultats indiquent : (1) la confirmation des conclusions de Crawford (1975, dans Biosocial Interrelations and Population Adaptations, E. S. Watts et al. (eds), pp 93–103) à propos des origines et des affinités génétiques des Travellers ; (2) l’influence majeure d’événements historiques particuliers sur la structure des populations, observée via diverses analyses multivariées ; et (3) que l’hypothèse de Relethford et Crawford sur la particularité des comtés des Midlands (1995, American Journal of Physical Anthropology, 96, 25–38) est confirmée par cette étude.