Correlations in the population structure of music, genes and language

Steven Brown1, Patrick E. Savage2, Albert Min-Shan Ko3, Mark Stoneking3, Ying-Chin Ko4, Jun-Hun Loo5 and Jean A. Trejaut5

1Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1
2Department of Musicology, Tokyo University of the Arts, Tokyo, Japan
3Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
4Graduate Institute of Clinical Medical Science, China Medical University, Taichung, Taiwan, Republic of China
5Mackay Memorial Hospital, New Taipei City, Taiwan, Republic of China

We present, to our knowledge, the first quantitative evidence that music and genes may have coevolved by demonstrating significant correlations between traditional group-level folk songs and mitochondrial DNA variation among nine indigenous populations of Taiwan. These correlations were of comparable magnitude to those between language and genes for the same populations, although music and language were not significantly correlated with one another. An examination of population structure for genetics showed stronger parallels to music than to language. Overall, the results suggest that music might have a sufficient time-depth to retrace ancient population movements and, additionally, that it might be capturing different aspects of population history than language. Music may therefore have the potential to serve as a novel marker of human migrations to complement genes, language and other markers.

1. Introduction

As human populations migrate to new regions of the world, their evolutionary divergence leaves its mark on both genes and culture. Correlations between cultural markers such as language and genetic markers such as mitochondrial, Y chromosome or autosomal DNA demonstrate that these features can co-migrate and coevolve over the course of thousands of years [1–3]. The same coevolutionary process has been proposed for music and genes [4–7], not least because music is a universal feature of human cultures [8,9] and because it shows quantifiable diversity both within and between populations [10,11]. However, this hypothesis has been criticized on the grounds that music evolution might occur at too rapid a rate [12] and therefore that music’s time depth might be too shallow to be correlated with something as ancient and slowly evolving as genes. Although a few studies have found suggestive parallels between music and genes [5,13–15], none have demonstrated a statistically significant correlation between them. We wanted to examine this relationship quantitatively for the first time, to our knowledge, and explore whether music might have the potential to serve as a new type of marker for the study of human population history.

We decided to examine Taiwan as a test case as it has several clear advantages for such an analysis. Taiwan has a small number of well-characterized indigenous populations that are located in geographically distinct regions of the island [16]. These populations have been well studied musically, linguistically and genetically such that there exists ample material for performing correlational analyses. The indigenous musics have been extensively recorded and archived by ethnomusicologists since the 1920s [17–21], and genetic analyses of mitochondrial DNA (mtDNA) haplotypes for most of the indigenous groups have been published [22]. Therefore, the degree of musical and genetic sampling makes Taiwan an ideal case for analysis. In addition to this, Taiwan has been the focal point of theories about one of the most significant migrational events in human history, namely the expansion of the Austronesian-speaking peoples
2. Material and methods

(a) Musical sample
We restricted our analysis to group-level (choral) vocal songs—excluding solo songs and purely instrumental music—because we predicted that the constraints involved in coordinating musical parts among multiple singers would make this repertoire the most resistant to change over time, and hence the most stable. The songs comprised a mixture of traditional genres with a focus on ritual songs. We excluded children’s songs and songs with explicit signs of borrowing, for example Christian missionary songs.

The musical sample consisted of 220 traditional, group-level vocal songs from nine indigenous populations from Taiwan (see figure 1 for the geographical location of these populations): the Amis (30 songs), Atayal (8), Bunun (30), Paiwan (30), Puyuma (24), Rukai (30), Saisiat (22), Tao/Yami (28) and Tsou (18). These songs were obtained in consultation with an ethnomusicologist with an expertise in these musics, Ying-fen Wang. Most of the songs are available from the Taiwan National Music Archive (http://music.ncfta.gov.tw). This archive contains a variety of ethnomusicological recordings, some of them published commercially (notably [19,21]). Our sample represents all of the populations whose genetic data were published by Trejaut et al. [22], which comprises the nine populations that have been officially recognized by the government for many decades. However, it does not include five groups that were only recognized in the twenty-first century or any of the groups that are not officially recognized. Musicological characterization of a larger dataset that included musics from these former five groups has been reported in other publications from our laboratory [11,34].

The complete corpus of 220 songs was coded using two different methods of classification. P.E.S. coded all the songs acoustically using the CantoCore song-classification scheme developed in our laboratory [35], while Victor Grauer coded all of the same songs using the Cantometric coding system [36,37]. In order to avoid unreliable characters as well as character duplication between the two schemes, we chose 41 characters a priori for the analysis: all 26 structural characters from CantoCore (related to rhythm, pitch, text, texture and form) and the 15 performance-style characters from Cantometrics (related to vocal style, ornamentation and dynamics; see Savage et al. [35] for descriptions of specific characters and their coding reliability). Distances between songs were calculated using these 41 characters, accounting for ordinal, nominal and missing characters, as described in Rzeszutek et al. [11].

(b) Genetic sample
Genetic samples were obtained from the published data of Trejaut et al. [22]. The samples consisted of hypervariable segments 1 and 2 of the control region of the mitochondrial genome from 640 individuals. While this analysis was being performed, a newly obtained cohort of 410 mtDNA samples from these same nine Taiwanese populations became available at the Max Planck Institute in Leipzig (Ko et al. [38]). The Mantel correlation between the genetic datasets was \( r_M = 0.45, p = 0.005 \). A combined dataset was prepared such that the same populations were pooled across the two datasets. The final sample sizes were: Amis (148 individuals), Atayal (159), Bunun (139), Paiwan (105), Puyuma (91), Rukai (100), Saisiat (87), Tao (113) and Tsou (108). After all sequences were aligned and edited, a 744 base-pair haplotype that encompassed the entirety of hypervariable segments 1 and 2 across a total of 1050 individuals was used for the correlational analyses.

(c) Distances between populations
For both the genetic and musical data, pairwise distances among the populations were calculated using the analysis of molecular variance (AMOVA) framework [39] in ARLEQUIN v. 3.5.1.2, as described in detail for music in Rzeszutek et al. [11]. These distances were measured using the statistic \( \Phi_{ST} \) (uncorrected for heterogeneous sites), which represents the proportion of variability among individual songs or genetic sequences that is owing to between-population differences.
It should be noted that the Mantel analyses that also use one-tailed tests by default [47–52].

make our analysis comparable with published gene–language a standard correlation not based on distances, and in order to test already has lower power to reject the null hypothesis than correlations between the various distances, because the Mantel Fortin [46], because coevolutionary hypotheses predict positive...

...tance matrices to construct an empirical null distribution that is another by randomly permuting the rows and columns of the distances among the nine populations are not independent of one...

Using 10,000 permutations, with the threshold for significance set at \( p < 0.05 \). This test controls for the fact that the 36 pairwise distances among the nine populations are not independent of one another by randomly permuting the rows and columns of the distance matrices to construct an empirical null distribution that is used to assess the significance of the observed correlation. Partial Mantel tests [45] were used to assess correlations between two distance matrices while controlling for a third distance matrix.

One-tailed tests were used, as recommended by Legendre & Fortin [46], because coevolutionary hypotheses predict positive correlations between the various distances, because the Mantel test already has lower power to reject the null hypothesis than a standard correlation not based on distances, and in order to make our analysis comparable with published gene–language analyses that also use one-tailed tests by default [47–52]. It should be noted that the Mantel \( r^2 \) is always smaller than an \( R^2 \)-value based on rectangular data tables in cases (unlike this one) when both types of data can be compared, and so it is not appropriate to interpret the Mantel \( r \)-value in terms of the percentage of variance accounted for [46].

3. Results

AMOVA analyses were performed for music and genes in order to examine the partitioning of variances into between-culture and within-culture components; such an analysis was not possible for language because there are no data available on intracultural linguistic diversity in Taiwan. The \( \Phi_{ST} \) value for music was 0.047 (4.7% between-culture variance), whereas that for genes was 0.127 (12.7%). These between-culture components for both musical and genetic diversity were highly significant (\( p < 0.00001 \)) despite their relatively small absolute magnitudes.

Next, we used NeighborNets to analyse population structure for music, genes and languages for the nine indigenous Taiwanese populations for which all three types of data were available (figure 2; see figure 1 for the geographical locations of the populations). In keeping with the extreme linguistic diversity of Taiwan, the NeighborNet for language was more or less star-shaped (figure 2c), implying that each language was nearly equidistant from every other language in the set. The Q-residual value for language was essentially zero (\( 2 \times 10^{-16} \)), suggesting that its branching pattern was almost completely tree-like.

In contrast to this, the NeighborNets for both music and genes showed specific structures to them such that clustering was seen among populations. For example, neighbouring groups such as Paiwan and Rukai or Bunun and Tsou showed proximity in both networks. In addition, both of these NeighborNets showed extensive amounts of reticulation, as reflected in Q-residual values that were far greater than that...
versus genes. The Mantel statistic are significant at the 0.05 level.)

Correlations between music, genes and language. (Values in italic Table 1. presents the correlations for all analyses, and figure 3 shows the associated regression plots, with the values for music and genes.

The magnitude of the music/gene correlation was quite high ($r = 0.417$). It is important to note that the value of this first-reported music/gene correlation is comparable to the language/gene correlation measured in this study as well as to most published language/gene correlations, which generally report maximum $r$-values in the range of 0.3–0.5 [47–52]. The observation that both music and language were significantly correlated with genes but not with one another suggests that these two cultural markers might be capturing partially distinct components of human population history, a contention that is supported by the fact that the musical and genetic data were much less tree-like than the linguistic data. This strengthens the case for language (0.211 for music and 0.104 for genes), suggesting that the population structures for music and genes are much less tree-like than that for language.

Because patristic distances are, by definition, calculated from a tree and may thus overestimate how tree-like the linguistic data really are, we also calculated the Q-residual value for language using Hamming distances. The obtained value was 0.001, which is greater than that calculated using patristic distances but still more than 100 times smaller than the values for music and genes.

In order to look for potential coevolutionary relationships, we analysed correlations between musical, genetic, linguistic and geographical distance matrices. Table 1 presents the correlations for all analyses, and figure 3 shows the associated regression plots, with the $r$- and $p$-values presented inside each plot. The correlation between music and genes was statistically significant ($r = 0.417$, $p = 0.015$). This correlation remained significant even when geographical distance was controlled for ($r = 0.385$, $p = 0.032$). This finding suggests that these correlations reflect a branching coevolution of music and genes through shared ancestry [53] rather than a process of ‘isolation by distance’ [54], in which case the correlations would simply result from recent diffusion to geographical neighbours.

The correlation between languages and genes was also significant ($r = 0.492$, $p = 0.006$), although this correlation became non-significant when geographical distance was taken into account ($r = 0.321$, $p = 0.071$). Interestingly, despite the fact that both music and language showed significant correlations with genes, the correlation between the two of them was not statistically significant ($r = 0.411$, $p = 0.085$), suggesting that these two cultural markers might be capturing, at least in part, distinct facets of genetic population history. Finally, correlations with geographical distance were significant for genes ($r = 0.468$, $p = 0.003$) and language ($r = 0.540$, $p = 0.014$) but not for music ($r = 0.174$, $p = 0.248$).

### 4. Discussion

These results provide, to our knowledge, the first quantitative evidence that music might be useful as a novel marker to study human population history by demonstrating statistically significant correlations between musical and genetic diversity for nine indigenous populations of Taiwan. Our results provide to the best of our knowledge, the first empirical support for the proposals of Lomax [10], Grauer [4,5] and Jordania [6] that music—particularly polyphonic group singing—might serve as a useful marker to study human migrations and human origins more generally. The fact that music can be shown to be correlated with a robust genetic marker like mtDNA suggests that it might have a sufficient time-depth to track population movements, although it is impossible at the present time to determine whether the correlations between musics, genes and languages date back to the initial peopling of Taiwan thousands of years ago or to more recent population movements within the last few centuries.

![Figure 3](image-url) Regression plots for music, genes and language. Regression plots are shown for (a) genes versus music, (b) language versus music and (c) language versus genes. The Mantel statistic $r$ and the associated $p$-value are shown inside each plot. The significance of the $p$-values is calculated using a Mantel permutation test to account for the fact that the 36 pairwise distances among the nine populations are not independent of one another.

<table>
<thead>
<tr>
<th>full correlations</th>
<th>$r$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>music–genes</td>
<td>0.417</td>
<td>0.015</td>
</tr>
<tr>
<td>music–language</td>
<td>0.411</td>
<td>0.085</td>
</tr>
<tr>
<td>music–geography</td>
<td>0.174</td>
<td>0.248</td>
</tr>
<tr>
<td>language–genes</td>
<td>0.492</td>
<td>0.006</td>
</tr>
<tr>
<td>language–geography</td>
<td>0.540</td>
<td>0.014</td>
</tr>
<tr>
<td>genes–geography</td>
<td>0.468</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>partial correlations</th>
<th>$r$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>music–genes (geography)</td>
<td>0.385</td>
<td>0.032</td>
</tr>
<tr>
<td>music–genes (language)</td>
<td>0.271</td>
<td>0.054</td>
</tr>
<tr>
<td>music–language (geography)</td>
<td>0.382</td>
<td>0.018</td>
</tr>
<tr>
<td>language–genes (geography)</td>
<td>0.321</td>
<td>0.071</td>
</tr>
<tr>
<td>language–genes (music)</td>
<td>0.387</td>
<td>0.031</td>
</tr>
</tbody>
</table>
using music as a complementary and informative marker for the study of population history.

We compared AMOVA analyses for music and genes, something made possible by our ability to measure both within- and between-cultural diversity for music—just as is routinely done for genes—but something not readily possible for languages [55]. We recently performed, to our knowledge, the first AMOVA analysis of a cultural trait, namely music [11]. Using a subset of the songs from that study, the current analysis found that the vast majority of musical diversity was accounted for by the within-cultural component. However, while the between-cultural component had a small absolute magnitude of 4.7%, this component was highly statistically significant ($p < 0.00001$). This suggests that there is ample between-culture musical diversity for performing cross-cultural comparisons for music.

Ross et al. [56] recently performed an AMOVA analysis of another cultural trait, namely 700 variants of a single folktales across 31 populations in Europe. Their measured $\Phi_{ST}$ value of 0.091 is comparable to but somewhat higher than our own value for music. However, these authors argued—along with Bell et al. [57] in their analysis of the World Values Survey—that cultural differences between populations are far greater than genetic differences (the latter measured in Europe as a $\Phi_{ST}$ value less than 0.01 [58,59]), whereas we found the opposite relationship. Also, whereas Ross et al. found geographical distance to be a strong predictor of variation in folktales across cultures, our correlational analyses showed far less of an effect of geographical distance on music. mtDNA $\Phi_{ST}$ values for Taiwan (0.127) are quite large for a region of this size and are more comparable to values obtained from world surveys, which tend to range from 0.05 to 0.30 [60–63]. Therefore, we argue that it is premature to make generalizations about the relative sizes of cultural versus genetic $\Phi_{ST}$, because each one might show substantial variation across world regions and because cultural $\Phi_{ST}$ might vary strongly as a function of the trait being measured and the classification tool used to measure it.

Recently, Pamjav et al. [15] performed a similar analysis to our own, examining the relationship between genetic distance and musical distance across Eurasia for 42 cultures for Y-chromosome analysis and 56 cultures for mtDNA analysis. Their automated computer analyses of the music were based on one-line song notations (as opposed to recordings in our case) and relied on a single musical feature related to melody. While clustering methods were employed for both the genes and the music, no significance testing or correlations were performed. However, their analysis lent support for a relationship between musical distance and genetic distance. Other qualitative findings have suggested similar relationships that span large distances in both geographical extent and in the amount of time that the populations are thought to have been isolated from one another, including that between central African Pygmies and southern African San ‘Bushmen’ [4,64], Bantu-speaking populations throughout sub-Saharan Africa [14] and Arctic cultures on both sides of the Bering Strait [13]. The fact that independent analyses of a variety of different regions using a variety of different methods and cultural samples coalesce on a similar conclusion lends support to the idea that music truly has a substantial time-depth that can aid in the study of human history. Our own findings are also, to the best of our knowledge, the first to demonstrate a relationship between music and genes over such a geographically restricted area, which is essential for understanding the mechanisms by which large-scale correlations might arise [49].

Unlike Pamjav et al. [15], we focused our analysis on choral songs based on the idea that the coordination involved in group singing should provide important constraints on song evolution. The observation of extensive polyphonic singing among most of the Taiwanese indigenous peoples creates an important cultural link between Taiwan and a majority of the extent Austronesian-speaking populations, including most ethnic groups throughout Island Southeast Asia, Oceania, coastal Papua New Guinea and Madagascar. It also creates a link to the southern region of China from which the proto-Austronesian peoples supposedly emanated [65,66], as polyphonic singing is ubiquitous among the ethnic minorities of southern China [67,68]. It therefore may be possible to use music to investigate the status of Taiwan versus other areas as the source-population for the expansion of the Austronesian-speaking people [3,23–31,40].

Overall, music showed greater similarities to genes than did languages with regard to population structure and degree of reticulation. This might reflect intrinsic similarities between music and genes in their mechanisms of evolution, migration and cross-cultural contact compared with languages. This work therefore opens the door to using music—a universal yet highly diverse feature of human culture—as a novel marker of human history, one that provides complementary information to the more established cultural marker of language. The correlations we observed between musical and genetic diversity support the contention that music and genes may have been coevolving for a significant time period and that music might possess the capacity to track population changes occurring on the time scale of perhaps thousands of years. To the extent that migrational models are validated by the concordance of results across multiple markers [1–3,69], music may well contribute to a richer understanding of human evolution.

Acknowledgements. We thank Tom Rzeszutek for his work on the distance matrices and correlational analyses, and Ying-fen Wang for ethnographic and musicological advice on song selection for the Taiwanese musical sample. We thank Victor Grauer for the Canto-metric coding of all of the songs in this study, Emily Merritt for the inter-rater reliability codings, Simon Greenhill for providing linguistic data and Marie Lin for her support of the collaborative arrangement with the Taipei group.

Funding statement. This work was supported by a grant to S.B. from the Social Sciences and Humanities Research Council of Canada, by an Amherst College Roland Wood Fellowship and a Japanese Ministry of Education, Culture, Sports, Science, and Technology Scholarship to P.E.S., and by the Max Planck Society for the research of A.M.-S.K. and M.S. The GenBank accession numbers for the Max Planck Taiwan sample are KF540506-KF541055.

References


