MUSICAL DIVERSITY AND HUMAN MIGRATION

CROSS-CULTURAL MUSICAL DIVERSITY AND IMPLICATIONS FOR ITS USE IN STUDYING HUMAN MIGRATION

By

TOM RZESZUTEK, H.B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2011)	McMaster University
(Psychology, Neuroscience and Behaviour)	Hamilton, Ontario

TITLE:	Cross-cultural musical diversity and implications for its use in studying human migration
AUTHOR:	Tom Rzeszutek, H.B.Sc.
SUPERVISOR:	Professor Steven Brown
NUMBER OF PAGES:	vi, 44

Abstract

Cross-cultural musical diversity and implications for its use in studying human migration

Tom Rzeszutek
Master of Science
Depratment of Psychology, Neuroscience and Behaviour
McMaster University
2011

Music is greatly underappreciated in the scope of cross-cultural analysis. This is due in part to methodological problems plaguing recent comparative approaches, and modern ethnomusicology's stance against cross-cultural analysis. Language, on the other hand, has a long history of cross-cultural study and recent advances in quantitative techniques, borrowed mostly from biology, have put language at the forefront of studying population prehistory from a cultural perspective. Chapter 2 of this thesis presents a novel quantitative approach to studying cross-cultural musical diversity based on the AMOVA methodology borrowed from population genetics. This method allows researchers to quantify the amount of variability found between as well as within populations, and gives us a measure of population-level divergence that accounts for intra-population variability. Our major finding is that the vast majority of musical variability (~98%) is found within populations rather than between. This approach solves many of the quantitative issues with the original Cantometrics approach, and is widely applicable to the analysis of many domains of culture. Aside from methodological issues a major open question is whether music has the requisite time-depth to answer questions about recent human pre-history. Chapter 3 focuses on addressing this question generally, and more specifically investigating which musical features trace population history most effectively. Using a corpus of songs from 9 Taiwanese aboriginal tribes and quantitative methods from chapter 2, we show that features related to song structure are correlated with mitochondrial DNA data from the same populations, while features of singing style are not. Both the quantitative methods and provisional support for music's time depth presented here will hopefully usher in a new era of comparative musicology and provide scholars of pre-history with an additional tool to answer unresolved questions.

Acknowledgements

I am eternally indebted to my supervisor Steven Brown, who shared his idea for this project with me early on, and encouraged me to take on a Master's degree to work on this fascinating and novel project. He has provided me not only with material resources but his own valuable time, giving me motivation and guidance, but also space and flexibility whenever required. I would also like to thank Patrick Savage, my colleague on this project, without whom the completion of this work would not have been possible. His constant excitement about our work and his openness to new ideas kept me stimulated and eager throughout this project. Additionally, I am grateful to my committee, Ben Evans, Laurel Trainor, and Susan Dudley, who bravely agreed to advise me on a project that some have referred to as "crazy" because of its sheer novelty. This project would also not have been possible without many of our international collaborators and people I have met while exploring this new field. They include (but are not limited to): Jean Trejaut, Marie Lin, Yingfen Wang, Albert Ko, Frederick Delfin, Mark Stoneking, and Paul Heggarty. The other members of the NeuroArts Lab have made it an amazing place to come into work everyday; I wish them the best of luck. Special thanks goes to Michel Belyk, who has been a very insightful sounding board for programming problems as well as a dear friend. Lastly, to many of the friends I've made during my time here in the Department of Psychology, Neuroscience and Behaviour, I couldn't ask for better a support base, you will truly be lifelong friends.

Contents

1 General Introduction 1

1.1 Comparative Musicology: Criticisms and Applications......... 1

1.2 Contributions and Chapter Contents...................................... 3

2 The structure of cross-cultural musical diversity 6

2.1 Abstract... 6

2.2 Introduction... 6

2.3 Materials and Methods.. 9

2.4 Results..10

2.5 Discussion..12

2.6 Supplementary Material...20

3 Coevolution of music and genes in aboriginal Taiwan 32

3.1 Abstract...32

3.2 Introduction...33

3.3 Methods.. 34

3.4 Results..36

3.5 Discussion..37

4 General Discussion 42

List of figures and tables

Fig. 2.1 Multidimensional scaling plot of 421 songs.............................11
Table 2.1 Musical AMOVA results..11
Fig. 2.2 NeighborNet Plot of pairwise Φ_{ST} 12
Fig. 2.S1 Calculating distance for an ordinal variable............................26
Fig. 2.S2 Calculating distance for an ordinal variable............................28
Table. 2.S3 Modal profiles..29
Fig. 2.S4 Neighbour-Net of modal profile distances..............................30
Fig. 3.1 Cantometric features of song-structure and singing-style......... 33
Fig. 3.2 Musical and genetic distances among 9 Taiwanese aboriginal tribes.. 36
Fig. 3.3 CantoCore analysis...37

Chapter 1

General Introduction

1.1 Comparative musicology: criticisms and applications

Human prehistory is studied through the lens of many different marker systems, encompassing genetics, archeology, linguistics, and culture. Many times these marker systems can be in conflict (Diamond, 2000; Oppenheimer & Richards, 2001) due to underlying differences in processes responsible for shaping diversity in these domains (Newson, Richerson, & Boyd, 2007). As a result, each domain can provide a unique perspective, and it is only through the synthesis of information across these domains that scholars can truly understand the demographics of recent human prehistory (McMahon & McMahon, 2008; Renfrew, 2010). Probably the most widely discussed relationship here, in the context of cross-cultural work, is that between genetics and languages, popularized originally by Cavalli-Sforza et al.'s publication on the congruence of a worldwide tree of linguistic data and a corresponding genetic tree (Cavalli-Sforza, Piazza, Menozzi, & Mountain, 1988). Later criticisms of this original approach highlighted problems with autocorrelation, associated with the definition of genetic populations on a linguistic basis (McMahon & McMahon, 2005). This, along with the often-discussed limited time depth of linguistic analysis (McMahon & McMahon, 2008), highlights the fact that each marker system does indeed have its limitations and drawbacks. These widely studied markers may therefore benefit from an interdisciplinary approach that incorporates novel information supplemented by other cultural systems, such as music.

Music provides a good alternative because it contains many of the necessary components of a novel marker system: it is universal (Nettl, 2000), cross-culturally variable and can be studied using tools available to contemporary musical scholars (Lomax, 1968). While linguistics has a long history of cross-cultural comparison that continues to today (McMahon & McMahon, 2005) comparison in musicology is not widespread. Erich von Hornbostel was an early twentieth century Austrian ethnomusicologist, largely credited with the early development of comparative musicology (Hornbostel, 1905) and an important classification scheme for instruments (Hornbostel & Sachs 1914). Hornbostel broke new ground at the time, but his work and that of other early comparative musicologists has since been highly criticized for its ethnocentric stance on non-European musics (Rehding, 2000). A contemporary attempt to revive comparative musicology was Alan Lomax's visionary Cantometrics project of the 1960's, which attempted to collect and classify a large cross-section of the world's folk music. Unfortunately, staunch criticisms of Lomax's misinterpretation of correlations between certain musical features and aspects of social structure plagued the project

(Erikson 1976). While this was probably the most vocal of critiques, there are many other valid musical and statistical grounds on which to criticize Cantometrics. For example, many ethnomusicologists have since been very critical of Lomax's use of 'modal profiles', an average set of features to represent a given cultures musical style (Feld, 1984; Henry, 1976). This kind of approach to musical sampling can seriously underestimate the rich internal diversity of musical repertoires.

Chapter 2 of this thesis attempts to address this criticism by providing a new quantitative approach to comparative musicology that involves extensive sampling of indigenous musics. Rather than taking one modal set of features as a profile for a population, we sample all available traditional adult group music available for a population, much like a sample of individuals is taken in a population genetic analysis. Following in the footsteps of historical linguistics, which has benefitted from methods adapted from evolutionary biology (Mace & Holden, 2005), we adapt the analysis of molecular variance (AMOVA) (Excoffier, Smouse, & Quattro, 1992) from population genetics in order to measure musical divergence at the population level from a sample of sub-population units of analysis (songs). In addition to this essential function, AMOVA provides estimates of the distribution of musical diversity, both between and within populations, which has not been previously studied from a quantitative standpoint. We also attempt to improve on the Cantometrics coding scheme by using one developed in our lab, called CantoCore (Savage, Merritt, Rzeszutek and Brown, in prep.). This scheme focuses exclusively on the structural features of vocal music, and solves many of the inherent problems with Cantometrics, such as coding ambiguity. Resolving these methodological issues is key to the revival of comparative musicology. Classically, ethnomusicology has taken a strong stance against comparison, in favour of ethnography (McLeod, 1974). However, the recent appearance of Journals such as *Analytical Approaches to World Music* as well as the advent of computational ethnomusicology (Tzanetakis et al. 2007) indicate that this view may be changing. Comparative musicology may also see its revival through use by anthropologists and pre-historians, seeking an alternative perspective on population prehistory.

To this end, studying the origin of Austronesian-speaking populations may be a good test case to apply music's potential to inform us about prehistory. This agriculturally driven expansion is widespread, and quite recent – within the last 6000 years by some estimates (Diamond, 2000; Diamond & Bellwood, 2003; Gray et al., 2009) – making it a good place to test music's time depth. Linguistically, the highest level of diversity, by far, is found in Taiwan (Diamond, 2000). This has led some linguists and archaeologists to argue that Austronesian-speaking people originated quite recently in Taiwan, and rapidly expanded to colonize the Pacific, as though on an "express train" (Diamond & Bellwood, 2003). Alternative models, such as the "slow boat" model, oppose the "express train" idea based on the dating of mitochondrial variants associated with the Austronesian expansion (Richards, Oppenheimer & Sykes, 1998). The timing suggested by these genetic studies is quite a bit earlier than the time frame suggested by linguistic studies, and this continues to be an unresolved problem in the field (Soares et al., 2011).

Tackling this open question using music is a daunting task, and the first step is to establish whether music has a robust time depth. Using music to study prehistoric human migration is not a completely novel idea (Grauer, 2006), but it has been criticized on the grounds that music changes too rapidly to contain information about prehistory (Stock, 2006). We hypothesized that if any features were slowly changing enough to trace the past, they would be the structural features, such as rhythm, melody and form, because multiple singers in a group must necessarily agree on these features. It is plausible that performance features dealing with singing style, such as nasality and rubato, would be more idiosyncratic and therefore more easily modifiable. Chapter 3 addresses this question empirically, using Cantometric codings (containing both structural and performance features) and music from the 9 major Austronesian-speaking Taiwanese aboriginal populations. We use the methodology outlined in Chapter 2, separately for structural and performance features, in order to create pairwise matrices of musical divergence. To test the relative time depth of these features we compare them to a marker of known time depth – namely the hypervariable region of the mitochondrial genome – that is widely used in population genetic research (Pakendorf & Stoneking, 2005). This provides one of the first comparisons of this kind, and although provisional, also lends some much needed empirical support to Grauer's idea (Grauer, 2006).

1.2 Contributions and Chapter Contents

Chapter 2 – The structure of cross-cultural musical diversity

Here I present a novel adaptation of a method for quantifying both the within- and between-population variation of cultural forms, specifically music. For this work, I came up with the idea for this analysis, developed the distance measure for music in conjunction with Patrick Savage (PS) and Steven Brown (SB), programmed that distance measure, performed the analysis and wrote the paper. PS did all of the musical coding and developed CantoCore with SB. SB helped with revisions on the paper and provided helpful feedback during the writing process. We plan to submit this manuscript for publication in the very near future to *Proceedings of the Royal Society B: Biological Sciences*.

Chapter 3 – Co-evolution of music and genes in aboriginal Taiwan

Here I present the application of methodology from Chapter 2 to investigate the time depth of musical features. PS came up with the specific concept for the study, performed the CantoCore codings and wrote the original manuscript. I helped with conceptual aspects of the study design, performed all of the musical and genetic analyses, and helped with re-writing and revisions. Victor Grauer performed all of the Cantometrics codings. SB helped with revisions on the paper, provided extensive feedback, and came up with the general concept for this kind of work. Despite its limitations, we hope to publish this work in some form in the near future.

References

Cavalli-Sforza, L. L., Piazza, A., Menozzi, P., & Mountain, J. L. (1988). Reconstruction of human evolution: Bringing together genetic, archaeological, and linguistic data. *Proceedings of the National Academy of Sciences*, *85*(16), 6002-6006.

Diamond, J. M. (2000). Taiwan's gift to the world. *Nature*, *403*(6771), 709-10.

Diamond, J., & Bellwood, P. (2003). Farmers and their languages: the first expansions. *Science (New York, N.Y.)*, *300*(5619), 597-603.

Erickson, E. E. (1976). Tradition and evolution in song style: A reanalysis of Cantometric data. *Cross-Cultural Research*, *11*(4), 277-308.

Excoffier, L., Smouse, P., & Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479.

Feld, S. (1984). Sound structure as social structure. *Ethnomusicology*, *28*(3), 383–409.

Grauer, V. (2006). Echoes of our forgotten ancestors. *World of Music*, *48*(2), 5-59.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, *323*(5913), 479-83.

Henry, E. (1976). The variety of music in a north Indian village: reassessing cantometrics. *Ethnomusicology*, *20*(1), 49-66.

Hornbostel, E. M. von, & Sachs, C. (1914). Classification of musical instruments [in German]. *Zeitschrift fur Ethnologie*, *14*, 3-29.

Hornbostel, E. V. (1905). The problems of comparative musicology [In German]. *Zeitschrift der Internationalen Musikgesellschaft*, *7*(3), 85-97.

Lomax, A. (1968). *Folk song style and culture* (p. 363). New Brunswick, N.J. Transaction Books.

Mace, R., & Holden, C. J. (2005). A phylogenetic approach to cultural evolution. *Trends in Ecology and Evolution*, *20*(3), 116-121.

McLeod, N. (1974). Ethnomusicological research and anthropology. *Annual Review of Anthropology*, *3*(1), 99-115.

McMahon, A. M. S., & McMahon, R. (2005). *Language Classification by Numbers* (p. 265). Oxford, U.K. Oxford University Press.

McMahon, A., & McMahon, R. (2008). Genetics, historical linguistics and language variation. *Language and Linguistics Compass*, *2*(2), 264-288.

Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (p. 463–472). Massachusetts: The MIT Press.

Newson, L., Richerson, P. J., & Boyd, R. (2007). Cultural Evolution and the Shaping of Cultural Diversity. In S. Kitayama & D. Cohen (Eds.), *Handbook of Cultural Psychology* (pp. 454-476). New York, N.Y. The Guildford Press.

Pakendorf, B., & Stoneking, M. (2005). Mitochondrial DNA and human evolution. *Annual review of genomics and human genetics*, *6*, 165-83.

Rehding, A. (2000). The quest for the origins of music in Germany circa 1900. *Journal of the American Musicological Society*, *53*(2), 345-385.

Renfrew, C. (2010). Archaeogenetics--towards a "new synthesis"? *Current Biology*, *20*(4), R162-5.

Richards, M., Oppenheimer, S, & Sykes, B. (1998). mtDNA suggests Polynesian origins in Eastern Indonesia. *American journal of human genetics*, *63*(4), 1234-6.

Soares, P., Rito, T., Trejaut, J., Mormina, M., Hill, C., Tinkler-Hundal, E., et al. (2011). Ancient Voyaging and Polynesian Origins. *The American Journal of Human Genetics*, *88*(2), 239-247.

Stock, J. P. J. (2006). Clues from our present peers?: A response to Victor Grauer. *World of Music*, *48*(2), 73–91.

Tzanetakis, G., Kapur, A., Schloss, W. A., & Wright, M. (2007). Computational ethnomusicology. *Journal of Interdisciplinary Music Studies*, *1*(2), 1-24.

Chapter 2

The structure of cross-cultural musical diversity

Tom Rzeszutek, Patrick E. Savage, and Steven Brown

Department of Psychology, Neuroscience and Behaviour, McMaster University,
Hamilton, ON, Canada

Keywords: AMOVA; cultural diversity; music; population structure; Austronesian language family; Taiwan

2.1 Abstract

Human cultural traits, such as languages, musics, rituals, and material objects, vary widely across cultures. However, the majority of comparative analyses of human cultural diversity focus on between-culture variation without consideration for within-culture variation. In contrast, biological approaches to genetic diversity, such as the Analysis of Molecular Variance (AMOVA) framework, partition genetic diversity into both within- and between-population components. We attempt here for the first time to quantify both components of cultural diversity by applying the AMOVA model to music. By employing this approach with 421 traditional songs from 16 Austronesian-speaking populations, we show that the vast majority of musical variability is due to differences within populations rather than differences between. This demonstrates a striking parallel between the structures of cultural and genetic diversity in humans. A Neighbor-Net analysis of pairwise population musical divergence shows a large amount of reticulation, indicating the pervasive occurrence of borrowing and/or convergent evolution of musical features across populations.

2.2 Introduction

Human cultural traits exhibit an astounding myriad of forms, perhaps best exemplified by the approximately 6900 known languages currently spoken across the world [1]. Any approach to characterising this cross-cultural diversity depends on the creation of a reliable classification of forms for a given domain of culture. There are many important examples of cultural classification, spanning from the seminal work of Murdock on the classification of over 100 categories of cultural behaviour across 1100 world populations [2] to contemporary examples in linguistics such as the World Atlas of Language Structures [3] and the Austronesian Basic Vocabulary Database [4]. The primary goal of these kinds of classification systems is the identification of salient differences between populations, as these differences can aid in reconstructing the history of human population movements and cultural interactions [5-7]. A major criticism of these approaches, though, is that they place an exclusive emphasis on the diversity *between* cultures, downplaying or ignoring the internal diversity present within each culture. Overall, there is a dichotomy between comparative approaches – whose goal is to characterize differences between cultures – and ethnographic approaches, whose goal is to rigorously catalogue the richness of forms that exist within single cultures. Here, we propose a compromise solution that allows for the simultaneous consideration of between-culture and within-culture facets of cultural diversity.

The hierarchical structure of human cultural diversity is reminiscent of the structure of human genetic diversity in that this diversity can be compartmentalized into within- and between-population components. Population geneticists, starting with Lewontin [8], have repeatedly observed that the vast majority of the genetic diversity in human populations is found within populations rather than between them [9]. Some cultural scholars have argued that human cultures exhibit a much lower level of internal diversity than that seen in the genetic domain due to processes such as conformity or frequency-dependent selection [10] that homogenize behaviours within populations and thereby push particular cultural variants to fixation [11]. While this is a plausible argument, no one, to the best of our knowledge, has done a rigorous quantification of the hierarchical structure of cultural diversity. Perhaps the closest study is that of Bell et al. [12] that used internal behavioural variation to calculate cultural variation among populations using a population genetic model. However, this work did not explicitly quantify the degree of internal variation.

One requirement in applying population genetic models to cultural forms is the necessity that there be quantifiable features that vary among individuals or entities both within and between populations. For example, Bell et al. [12] used questions from the World Values Survey, administered to a sample of individuals from each focal culture. This is comparable to looking at variation among individuals at a particular genetic locus. Alternatively, if one wanted to investigate variation in some aspect of material culture, such as ceramics, one would need a number of exemplars from each culture, appropriate features to describe these exemplars, and a suitable quantitative measure of differences among entities. Clearly, there is a difference between studying variation among individuals in terms of behaviour and variation among entities of material culture. What is most important for the study of cultural diversity is that the unit of analysis and the means of measuring difference between cultural variants have domain-specific validity, and this must be worked out on a case-by-case basis for each domain of culture.

Music seems to satisfy these important requirements and thereby affords a novel opportunity to study the structure of cultural diversity. Not only is music a human universal [13] but its form also varies quite prominently both between [14] and within [15] cultures. Musical features are also quite amenable to comparative analysis [14]. Most importantly for our purposes, the “song” provides a reliable unit for the cultural analysis of music. Ethnographic analyses of human cultures have clearly shown that the song represents the fundamental unit of both structure and function [13]. In addition, the song was adopted as the unit of analysis in the most ambitious comparative attempt to classify the world's musics, namely Lomax's Cantometrics project of the 1960's [14], in which more than 4000 songs from over 200 cultures were analysed and compared.

In order to make such a global project feasible, Lomax employed a small sample of only ten songs per culture, and these were averaged into a 'modal profile' that represented the 'typical' song-style for each culture [14]. While Lomax believed that his modal profiles were representative of the cultures he was sampling, ethnographers studying musics from those same cultures questioned Lomax's findings, because his approach strongly underestimated the degree of internal musical diversity in those cultures [15,16]. To date, there has been no quantitative method applied to music that retains the cross-cultural scope of Lomax's global framework while at the same time taking internal variation into account.

Exactly such a method is used in the study of genetic diversity in population genetics, and this method provides a promising approach for thinking about the hierarchical structure of cultural diversity as well. The Analysis of Molecular Variance (AMOVA) is a method closely related to the Analysis of Variance (ANOVA) that allows the hierarchical partitioning of genetic variance into components [17]. These components generally include: variability within populations; variability between populations; and variability between regional groups. The population structure being tested is defined a priori by the researcher, and can include divisions based on geographic region or language [17]. In its original application, AMOVA was designed to investigate molecular diversity based on haplotype restriction polymorphism data, but the generalizability of the method was recognized early on [17] and has since been applied to many different kinds of genetic loci [18]. The flexibility of this method rests on the fact that variability is calculated as a measure of distance between haplotypes. The distance measure itself is defined by the user and can incorporate information about sequence evolution such as mutation rate [17]. Consequently, given an appropriate unit of analysis and distance measurement, this method can be extended to quantify the hierarchical structure of cultural diversity.

We attempt here for the first time to quantify both the within- and between-population components of cultural diversity by applying AMOVA to the analysis of musical diversity using the song as the unit of analysis. An important distinction here is that we are looking at populations of *songs* rather than populations of individuals. To this end, we focus on a rigorous sampling of tribal musics from Austronesian-speaking populations in Taiwan and the Philippines, itself part of a larger project devoted to prehistoric migrations in the region. To quantify musical variability, we calculate the distance between songs using a musical-classification system we developed that is inspired by Cantometrics. The AMOVA framework is then applied to this data in order to apportion musical variability into within- and between-population components. We also measure pairwise population musical divergence with Φ_{ST} and use it in a 'Neighbor-Net' analysis [19] to explore the degree of reticulation in the data due to borrowing and/or convergence. Distances based on Φ_{ST} are also compared with the corresponding modal profiles to test the accuracy of Lomax's modal-profile approach for distinguishing differences between populations. Our novel application of AMOVA to cultural forms provides a general means of performing population-level cultural analyses while simultaneously addressing the internal diversity of cultural forms.

2.3 Materials and Methods

a) Sample

The musical sample consists of 421 traditional group (choral) songs from 16 Austronesian-speaking aboriginal populations from Taiwan and the northern Philippines, including the Amis (30 songs), Atayal (10), Bunun (30), Paiwan (30), Puyuma (30), Rukai (30), Saisiyat (30), Tao (30), Tsou (22), Plains (Siraya) (24), Kavalan (18), Thao (30), Ibaloi (30), Ifugao (30), Kankanai (17), and Ayta (30). Songs were obtained from commercial ethnomusicology recordings as well as from the Taiwan National Music Archive in Taipei [20] and the Centre for Ethnomusicology at the University of the Philippines in Quezon City. Thirty songs were randomly sampled from each population. For populations with less than 30 available songs, all recordings meeting our inclusion criteria were used.

b) Classifying songs

P.E.S. coded all the songs using the “CantoCore” song-classification scheme developed in our lab (Savage, Merritt, Rzeszutek and Brown, submitted). This comprehensive scheme, modelled after Lomax and Grauer’s (1968) original Cantometric scheme [14], codes 26 characters related to song structure, including rhythm, pitch, syllable, texture, and form (see electronic supplementary material, S1).

c) Quantifying musical distance

Either phylogenetic distances based on sequence evolution or phenetic distances based on sequence similarity can be used in genetic analyses [17]. Since we currently lack information about song evolution, we attempted to develop a simple phenetic measure of distance between songs, based on our codings, that is both musically and statistically valid. Leroi and Swire [21] and Busby [22] identified a number of methodological solutions to issues related to converting Cantometric song-codings into distances, and these issues apply equally well to CantoCore. These include: the presence of both ordinal and nominal characters; simultaneous coding of multiple states for a number of characters (multi-coding); the redundancy of some codings when certain states are absent; and equal weighting of all characters. We built on their work to program an algorithm that takes these issues into account while at the same time being flexible enough to handle a variety of coding schemes. The algorithm was programmed in R version 2.12.2 [23] by T.R. and is available upon request. Details of the algorithm are found in the electronic supplementary material (S2).

d) Visualizing song relationships

In order to visualize songs in two dimensions, we performed nonmetric multidimensional scaling on the song-level distances obtained from our algorithm using isoMDS in R, with 50 iterations and metric scaling as an initial configuration.

e) AMOVA analysis

Distances were prepared for the AMOVA analysis by a Euclidean transform of the data using Lingoes' method [24], as implemented in the ade4 package for R [25]. The distances were then squared, as recommended by Excoffier et al. [17]. AMOVA was performed in Arlequin 3.11 using the prepared distance matrix and standard settings [26]. Musical variability was apportioned "between" and "within" ethno-linguistically defined populations of songs [1]. The parameter Φ_{ST} is the proportion of total variability due to differences between populations [17], and was calculated pairwise as a measure of musical divergence between populations. To test the significance of the between-population component of musical variance, we permuted songs randomly between populations using 1000 permutations.

f) Neighbor-Net analysis

Pairwise Φ_{ST} was used in a 'Neighbor-Net' analysis [19] to determine the level of reticulation in the data due to borrowing and convergence. The analysis was performed in SplitsTree4 using standard settings [27]. All negative Φ_{ST} values were set to zero before performing the analysis [28].

g) Modal Profile Analysis

In order to test the efficacy of Lomax's modal profile approach at distinguishing differences between populations, we created a modal song coding for each population, consisting of the most common coding in its musical repertoire for each of the 26 CantoCore characters. This resulted in some combinations not present in any single song, but nonetheless best approximated Lomax's creation of modal profiles. These modal profiles are available in the electronic supplementary material (S3). Distances between modal profiles were calculated using the same algorithm applied to the original song data, giving us a population-level distance devoid of any information about internal diversity. These modal distances were then compared to the population pairwise Φ_{ST} measures using Spearman's rho (r_s) and a Mantel test with 20,000 permutations.

2.4 Results

a) Multidimensional Scaling

Figure 1 shows a multidimensional scaling plot for the 421 songs used in our sample colour-coded for the 16 tribes. The high level of stress (34.3) in this two-dimensional ordination indicates the complex multidimensional nature of the musical data. A scree plot did not reveal a clear elbow, and showed that instead our data would require more than 8 dimensions to achieve an acceptable level of stress under 10. Despite this, the multidimensional scaling plot clearly demonstrates the high level of internal heterogeneity in each population's musical repertoire, and the high degree of overlap between populations.

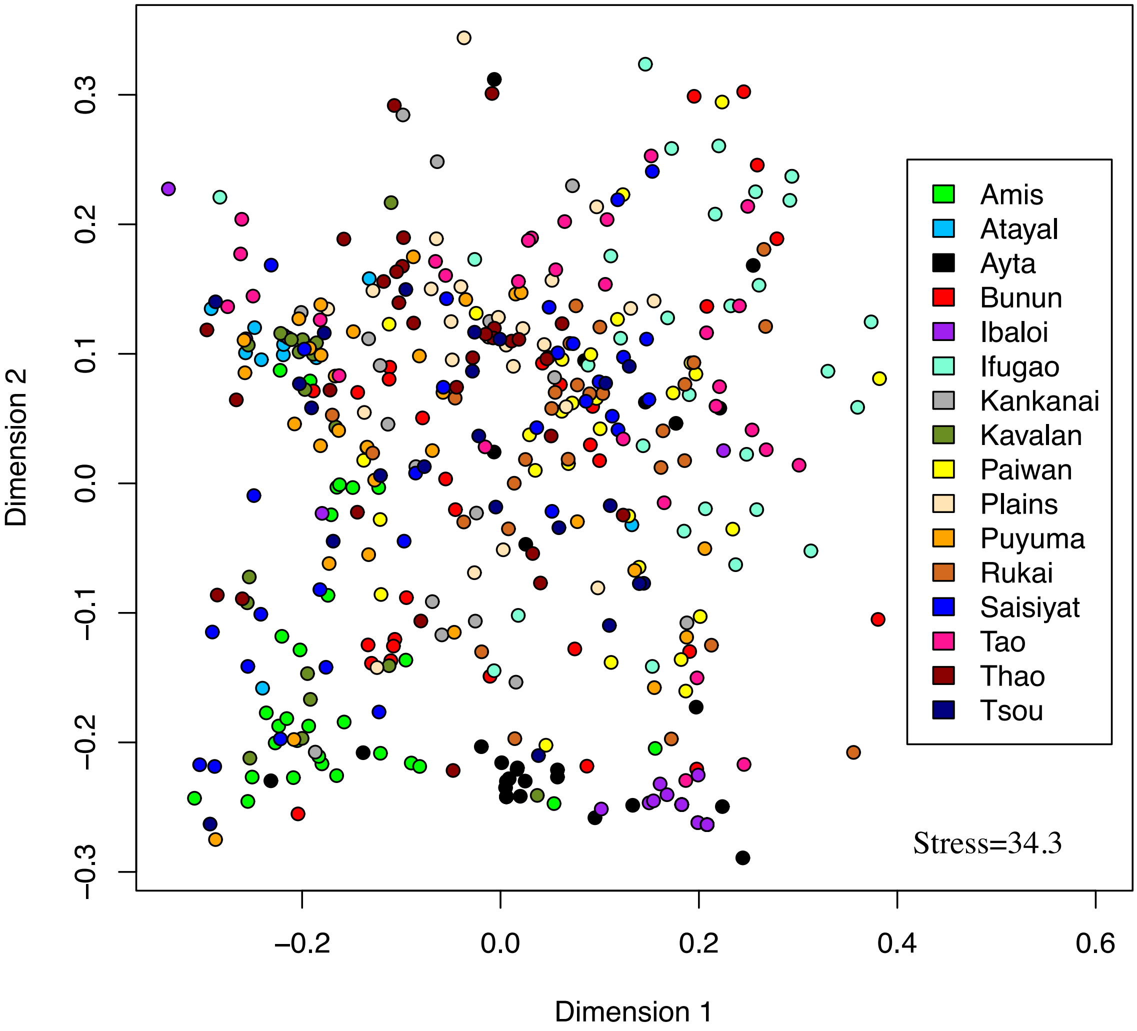


Figure 2.1 - Multidimensional scaling plot of distances between 421 songs from Austronesian-speaking populations. There is a large amount of overlap between populations and spread within populations. Each point represents a song and is coloured according to population of origin.

Table 2.1.
Musical AMOVA Results

Source of Variation	Percentage of Variation	Degrees of Freedom
Between Populations	2.06	15
Within Populations	97.94	405
Total	100	420

b) Song-level AMOVA analysis

The AMOVA analysis confirms the multidimensional scaling result (table 1), with a majority of the variance in our sample (~98%) being accounted for by differences within populations and a smaller portion (~2%) accounting for differences between populations. Despite accounting for a much smaller proportion of the variance, musical diversity between populations was statistically significant (Φ_{ST}=0.021, p<0.001).

c) Neighbor-Net analysis

The Neighbor-Net analysis (figure 2) demonstrated that our musical data did not appear tree-like and instead contained a fair amount of reticulation. This reticulation is reflective of borrowing and/or convergence of musical features between populations.

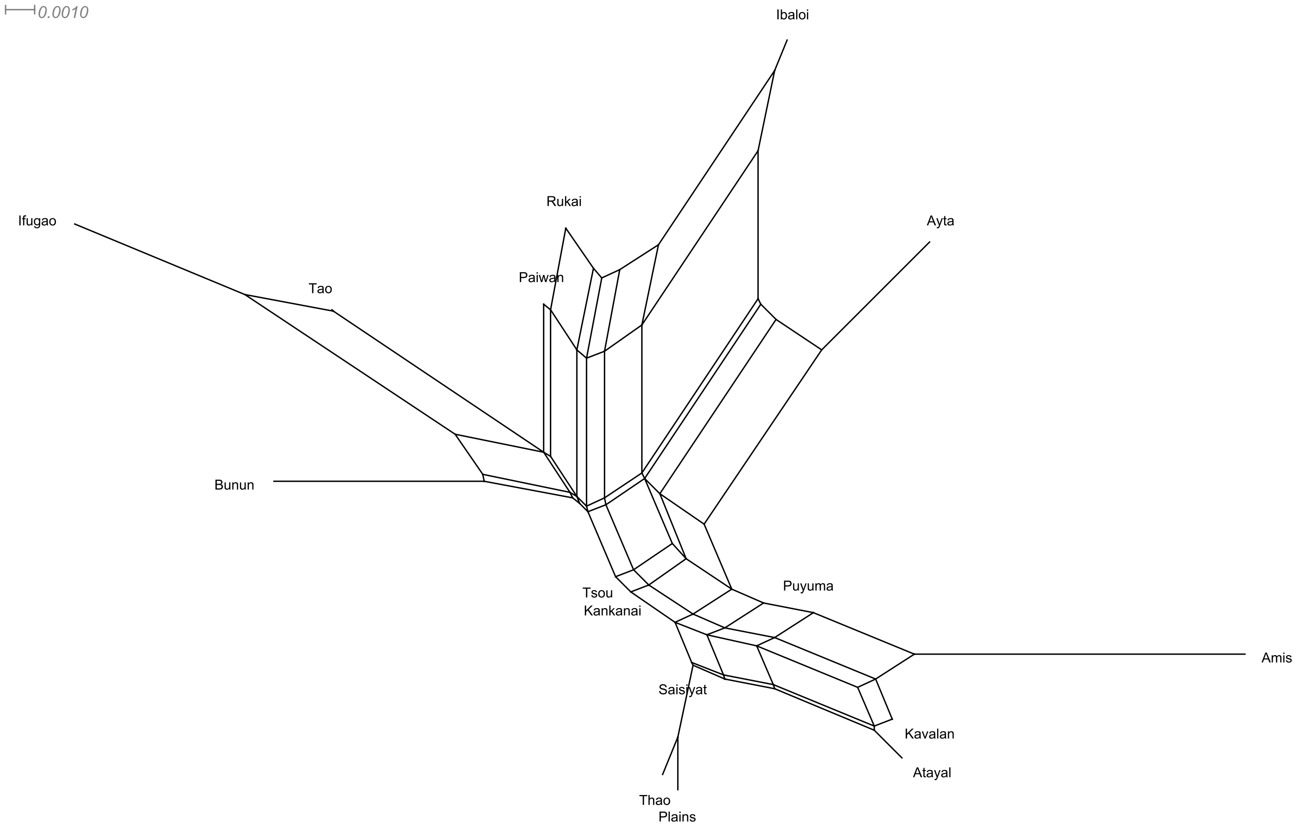


Figure 2.2 - A Neighbor-Net plot of population-level musical divergence between 16 Austronesian-speaking populations, based on pairwise Φ_{ST} from and AMOVA analysis of 421 traditional group songs. This plot shows a high degree reticulation in our dataset, indicating the presence of borrowing and/or convergence.

d) Modal profile analysis

The pairwise population-level distances based on the modal profiles (ignoring internal diversity) were highly correlated with pairwise Φ_{ST} distances (r_s=0.730, p<0.001), which take into account the internal variation in musical repertoires. This indicates that, although it cannot capture information about internal diversity within cultures, the modal profile approach may still adequately approximate overall patterns of variation between populations (see electronic supplementary material, S4).

2.5 Discussion

We have applied the Analysis of Molecular Variance (AMOVA) framework to a cultural dataset, allowing us for the first time to quantify the hierarchical structure of cultural diversity. Our application of this approach to a sample of aboriginal Austronesian songs demonstrated that the vast majority of musical variation in this sample (~98%) was found within populations, while a far smaller proportion of this variation (~2%) occurred between populations. This validates and quantifies the critiques of ethnomusicologists that Cantometrics' cross-cultural approach underestimated the diversity of musical repertoires within each culture [15,16]. Next, a Neighbor-Net analysis of population pairwise Φ_{ST} distances showed that our musical data was not very tree-like, providing some preliminary insight into the evolution of musical repertoires and the presence of forces that diversify musics within cultures.

a) How much diversity is sufficient?

The high level of internal musical diversity found in this study parallels general findings on the structure of human genetic diversity, with some estimates of this diversity being as high as 93-95% globally, and as high as 99% within some regions [9]. However, as in genetic domain, this raises the important question of how much diversity is sufficient for describing differences between populations. This has been extensively addressed in population genetics. Lewontin's 1972 analysis of human genetic variation led him to argue that the small proportion of variation found between populations in his study (14.6%) meant that differences between populations were not informative [8]. Some scholars [29,30], most prominently Edwards [31], have noted that this conclusion is statistically inaccurate, as it ignores information contained in the correlation of allele frequencies across many loci. Modern clustering approaches use the correlated nature of genetic data to distinguish between major human groups that coincide with their geographic distribution, despite the small amount of variation (3-5%) accounting for these differences [9].

This situation is qualitatively the same in the study of musical diversity, since the correlation between different musical features in songs reveals much more about the unique musical repertoires of populations than the frequency of the features themselves. Therefore, our observation that between-population musical variance is a very small proportion of the total variance in no way precludes using this component for taxonomic and comparative analyses of world musics, as Lomax did [14]. This kind of comparative methodology should not be applied recklessly but in consultation with expert ethnomusicologists, who can attest to the validity of the sample. The between-population component should be sufficient to distinguish populations musically, and this is validated by our modal-profile analysis. That analysis demonstrated that a methodology that ignores internal diversity might still be successful at detecting the overall pattern of differences between populations, even though it might fail to detect more fine-grained population relationships.

b) Cultural evolution of music

The transmission of cultural traits is distinct from that of biological traits in that there are many more possible modes of transmission. Unlike the human-genetic domain, where variants are passed vertically across generations, features of culture can also pass horizontally between members of the same cohort, as well as obliquely from unrelated elder members of a focal individual's group [32]. The presence of alternative modes of transmission has been a central issue in the application of phylogenetic models to cultural traits [33]. Our preliminary attempts to apply such models to our song sample support Leroi and Swire's [21] claim that musical evolution is much less "tree-like" than genetic evolution, with much independent invention (convergence) as well as borrowing (horizontal transmission) of individual musical features and even entire songs between populations.

This brings up the more general issue of the dynamics of musical evolutionary change. There are cultural forces that both diversify and homogenize musical repertoires, and some of them are conceptually analogous to forces that influence the dynamics of genetic change [34]. As with genes, cultural forms like songs can undergo random changes over time, a kind of musical "drift" [35]. Small population sizes may enhance the effects of genetic drift, although it is unclear as of yet how population sizes affect musical diversity and change over time. Another major force that can diversify repertoires is admixture through cultural contact, a kind of musical "flow". Recent contact situations, such as that between the Paiwan and Rukai of Taiwan in our sample [36], can lead to high levels of acculturation, despite the maintenance of distinct languages. This particular contact situation is well reflected musically, with Paiwan and Rukai producing the only negative pairwise Φ_{ST} value in our analysis. This is unsurprising as music actually provides an excellent model for "hybridization" in the cultural domain, since it is composed of a series of modular components (mainly pitch and rhythm) that can undergo "syncretisms" or blendings of features. A good example of this is found in African-American music, which contains a novel fusion of European tonal features and African rhythmic features [37]. Other cultural forces that can affect the frequency of cultural variants within and between populations include convergence, borrowing, innovation, conformity, extinction, and replacement (through imposition, as in situations of conquest or economic globalization).

One means by which musical repertoires diversify internally is through a fissioning into an increasing numbers of genres or functional song-types, a universal feature of musical repertoires. A classic example of genre-based variation in song structure is found in Arom's work on the music of the Pygmies of the Central African Republic [38], which qualitatively describes systematic differences in the musical features of songs performed in different social contexts, comprising roughly two dozen distinct musical genres (e.g., music for the hunting of elephants, music for the birth of twins). This is the same case with our Austronesian musical sample, with genres such as wedding songs and headhunting songs appearing in the repertoires of multiple populations. Unfortunately, the limited number of songs in the current study prevented us from doing any sort of meaningful genre-level analysis. It is plausible that some genres of song are less malleable or prone to borrowing, which could affect our results. Given a larger, more comprehensive dataset, the AMOVA approach could be used to explore how variability in genres is structured within and between populations.

Our work on the cultural evolution of music has important limitations, especially as related to our use of archival material. The reliance of our work on archival recordings highlights the difficulty in sampling the musical variation of indigenous populations in the modern age. One concern for the current work is that the kinds of songs represented in the archives that we used did not cover all of the genres of a population's musical repertoire due to ascertainment bias. This, however, does not negate our major finding, as the inclusion of unrecorded music of other genres in our analyses would most likely have increased, not reduced, the internal diversity of the musical repertoires.

Archival recordings are essential in a world where globalization and the associated expansion of Western culture threaten to extinguish much of the rich cultural diversity seen in human populations across the globe [39]. This decline is reflected in the sheer proportion of living languages classified as vulnerable, endangered, or critical, which is at least 27%, according to a conservative recent analysis [40]. The dominant influence of western music has led to non-traditional (western) musical features being incorporated into indigenous musical repertoires through a kind of imposed hybridization. Archival recordings reduce the potential of encountering this form of unwanted admixture but are problematic in other ways.

In addition to the possible sampling bias discussed above, some recordings may be poorly documented, misclassified, non-traditional, or of poor recording quality. We were fortunate enough to work with a very well-documented archive and to have received advice from an ethnomusicologist with expertise in the traditional musics of the Taiwan aborigines. This kind of work may be substantially more difficult in regions with less-organized archives and where ethnomusicological expertise on these traditional musics is lacking. Despite the inherent difficulty in doing this kind of work, the task of characterising and comparing worldwide musical diversity, as other scholars have done with languages [4], is an extremely important endeavour, not least considering the current rapid rate of cultural extinction [40].

c) How generalizable are these results to other aspects of culture?

Many useful parallels have been drawn between cultural and biological evolution [41], but the forces shaping cultural diversity can differ markedly from those that drive the structure of genetic diversity [42]. For example, some have argued that cultural variants will necessarily always display less intra-population variation than will genetic variants [11]. Language is one of the best-cited examples of a cultural trait that is mostly variable between speech communities (rather than within), due to strong constraints that ensure that members of a speech community can communicate with one another [10]. The relative strength of processes that reduce internal diversity and those that increase it is likely to differ across cultural domains. It is plausible that music, for example, may be subject to lesser constraints than a system like language, and that innovation in this domain may be more highly valued in some cultures. The current work only covers musical variation in a small number of populations within the same language family. Populations in other regions of the world may have much more homogeneous musical repertoires. However, our results demonstrate that a high degree of internal heterogeneity in a population's musical repertoire is a possibility, in at least some cases.

d) Conclusion

While the present-day structure of human genetic diversity has been rigorously quantified, we lack the same kind of quantitative information for most aspects of culture. The AMOVA framework provides cross-cultural researchers a means of quantifying variability for a number of cultural forms, and of exploring the forces responsible for balancing diversity and conformity. The current work is by no means meant as a comprehensive sampling of worldwide musical diversity, and indeed the partitioning of musical variance may differ substantially in other regions of the world. We do, however, present a crucial tool that can be applied to many other aspects of culture, a tool that can be useful for the study of human migrations and associated histories of cultural contact.

ACKNOWLEDGEMENTS

This work was supported by a grant to S.B. from the Social Sciences and Humanities Research Council of Canada and by an Amherst College Roland Wood Fellowship to P.E.S. We thank Yingfen Wang for providing expert ethnomusicological assessments of the aboriginal Taiwanese musics. We would like to thank Tom Currie for helpful comments on an earlier version of this manuscript. We would also like to thank Jean Trejaut and Victor Grauer for advice and support while conducting this research, as well as Michel Belyk for advice on programming.

REFERENCES

1 Lewis, M. (ed.) 2009 *Ethnologue: Languages of the world*. 16th ed. Dallas, TX: SIL International.

2 Murdock, G. 1967 Ethnographic atlas: A summary. *Ethnology* **6**, 109-236.

3 Haspelmath, M., Dryer, M., Gil, D. & Comrie, B., eds. 2005 *The world atlas of language structures*. Oxford: Oxford University Press.

4 Greenhill, S. J., Blust, R. & Gray, R. D. 2008 The Austronesian Basic Vocabulary Database: From bioinformatics to lexomics. *Evol. Bioinform.* **4**, 271-283.

5 Currie, T. E., Greenhill, S. J., Gray, R. D., Hasegawa, T. & Mace, R. 2010 Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* **467**, 801-804. (doi:10.1038/nature09461)

6 Gray, R. D., Drummond, A. J. & Greenhill, S. J. 2009 Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479-483. (doi:10.1126/science.1166858)

7 Reesink, G., Singer, R. & Dunn, M. 2009 Explaining the linguistic diversity of Sahul using population models. *PLoS. Biol.* **7**, e1000241. (doi:10.1371/journal.pbio.1000241)

8 Lewontin, R. C. 1972 The apportionment of human diversity. In *Evolutionary biology* (eds. T. Dobzhansky M. K. Hecht & W. Steere), pp. 391-398. New York: Appleton-Century-Crofts.

9 Rosenberg, N., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. & Feldman, M. W. 2002 Genetic structure of human populations. *Science* **298**, 2381-2385. (doi:10.1126/science.1078311)

10 Newson, L., Richerson, P. J. & Boyd, R. 2007 Cultural evolution and the shaping of cultural diversity. In *Handbook of cultural psychology* (eds. S. Kitayama & D. Cohen), pp. 454-476. New York: Guildford Press.

11 Foley, R. & Mirazón Lahr, M. 2011 The evolution of the diversity of cultures. *Philos. Trans. R. Soc. B* **366**, 1080-1087. (doi:10.1098/rstb.2010.0370)

12 Bell, A. V., Richerson, P. J. & McElreath, R. 2009 Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proc. Natl. Acad. Sci. USA.* **106**, 17671-17674. (doi:10.1073/pnas.0903232106)

13 Nettl, B. 2000 An ethnomusicologist contemplates universals in musical sound and musical culture. In *The origins of music* (eds. N. L. Wallin, B. Merker & S. Brown), pp. 463-472. Cambridge, MA: MIT Press.

14 Lomax, A. 1968 *Folk song style and culture*. New Brunswick, NJ: Transaction Books.

15 Henry, E. 1976 The variety of music in a north Indian village: Reassessing Cantometrics. *Ethnomusicology* **20**, 49-66.

16 Feld, S. 1984 Sound structure as social structure. *Ethnomusicology* **28**, 383-409.

17 Excoffier, L., Smouse, P. & Quattro, J. 1992 Analysis of molecular variance

inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

18 Michalakis, Y. & Excoffier, L. 1996 A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**, 5-8.

19 Bryant, D. & Moulton, V. 2004 Neighbor-net: An agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* **21**, 255-265. (doi:10.1093/molbev/msh018)

20 National Taiwan Preparatory Office of Traditional Arts 2010 Taiwan National Music Archive. http://music.ncfta.gov.tw [in Chinese]

21 Leroi, A. M. & Swire, J. 2006 The recovery of the past. *World of Music* **48**, 43-54.

22 Busby, G. 2006 Finding the blues: An investigation into the origins and evolution of African-American music. [Unpublished Master's Thesis]

23 R Development Core Team 2011 R: A Language and Environment for Statistical Computing.

24 Lingoes, J. 1971 Some boundary conditions for a monotone analysis of symmetric matrices. *Psychometrika* **36**, 195-203.

25 Dray, S. & Dufour, A. B. 2007 The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* **22**, 1-20.

26 Excoffier, L., Laval, G. & Schneider, S. 2005 Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**, 47-50.

27 Huson, D. H. & Bryant, D. 2006 Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254-267. (doi:10.1093/molbev/msj030)

28 Meirmans, P. G. 2006 Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* **60**, 2399-2402.

29 Mitton, J. B. 1977 Genetic differentiation of races of man as judged by single-locus and multilocus Analyses. *Am. Nat.* **111**, 203-212.

30 Mitton, J. B. 1978 Measurement of differentiation: Reply to Lewontin, Powell, and Taylor. *Am. Nat.* **112**, 1142-1144.

31 Edwards, W. F. 2003 Human genetic diversity: Lewontin's fallacy. *BioEssays* **25**, 798-801. (doi:10.1002/bies.10315)

32 Cavalli-Sforza, L. L., Feldman, M. W., Chen, K. H. & Dornbusch, S. M. 1982 Theory and observation in cultural transmission. *Science* **218**, 19-27.

33 Greenhill, S. J., Currie, T. E. & Gray, R. D. 2009 Does horizontal transmission invalidate cultural phylogenies? *Proc. R. Soc. B.* **276**, 2299-2306. (doi:10.1098/rspb.2008.1944)

34 Cavalli-Sforza, L. L. & Feldman, M. W. 1981 *Cultural transmission and evolution: A quantitative approach.* Princeton, NJ: Princeton University Press.

35 Bentley, R. A., Hahn, M. W. & Shennan, S. J. 2004 Random drift and culture

change. *Proc. R. Soc. B.* **271**, 1443-1450. (doi:10.1098/rspb.2004.2746)

36 Huteson, G. 2003 Sociolinguistic survey report for the Tona and Maga dialects of the Rukai language. *SIL Electronic Survey Reports* **12,** 1-46.

37 Kubik, G. 1999 *Africa and the blues*. Jackson, MS: University Press of Mississippi.

38 Arom, S. 1991 *African polyphony & polyrhythm: Musical structure and methodology*. Cambridge, UK: Cambridge University Press.

39 Lomax, A. 1977 Appeal for cultural equity. *J. Commun.* **27**, 125-138.

40 Sutherland, W. J. 2003 Parallel extinction risk and global distribution of languages and species. *Nature* **423**, 276-279. (doi:10.1029/2002GL015811)

41 Henrich, J., Boyd, R. & Richerson, P. J. 2008 Five misunderstandings about cultural evolution. *Hum. Nature* **19**, 119-137. (doi:10.1007/s12110-008-9037-1)

42 Mesoudi, A., Whiten, A. & Laland, K. N. 2006 Towards a unified science of cultural evolution. *Behav. Brain. Sci.* **29**, 329-347; discussion 347-383. (doi:10.1017/S0140525X06009083)

2.6 Supplementary Material

S1 – “CantoCore” Musical Classification Scheme

I) “MELODY” (between-note)

A) Rhythm

1) **METER**
   Cyclic groupings of strong and weak beats into bars
   - (a) *A-metric:* No consistent beat
   - (b) *Hetero-metric:* There is a consistent beat, but strong and weak beats occur without a consistent pattern
   - *(c) Poly-metric:* Multiple cyclic patterns of strong and weak beats coexist simultaneously (e.g., 6/8 against 3/4)
   - *(d) Iso-metric:* There is a single, consistent pattern of strong and weak beats (e.g., 3/4, 6/8, 5/4, 2+2+3/8)

   N.B. Songs not classified as “*iso-metric*” must be coded “?” for parameters **(2-5)**.

2) **NUMBER OF BEATS**
   The number of beats in a bar
   - (i) *Duple:* Multiples of 2 (e.g., 4/4, 6/8, 2+3/8)
   - (ii) *Triple:* Multiples of 3 (e.g., 3/4, 9/8, 2+2+3/8)
   - (iii) *Complex:* Multiples of prime numbers greater than 3 (e.g., 5/4, 5/8, 2+2+3+2+3/8)
   - (?) *A-/hetero-/poly-metric:* See **(1)**

3) **BEAT SUB-DIVISION**
   Division of beats into sub-beat-level metric groupings
   - (a) *A-divisive:* Beats are not sub-divided (e.g., a 4/4 piece containing only ♩ and 𝅗𝅥 notes)
   - (b) *Hetero-divisive:* Beats are sub-divided, but the number of sub-beats per beat changes (e.g., 2+2+3/8)
   - (c) *Iso-divisive:* Beats sub-divided into a consistent number of sub-beats (e.g., 6/8, a 4/4 piece containing ♪ notes)
   - (?) *A-/hetero-/poly-metric:* See **(1)**

   N.B. Songs not classified as “*iso-divisive*” must be coded “?” for parameter **(4)**.

4) **NUMBER OF SUB-BEATS**
   The number of sub-beats in a beat

(i) *Simple:* Multiples of 2 (e.g., ♩ beat divided into ♪ note sub-beats; includes 3/4, 4/4, etc.)
(ii) *Compound:* Multiples of 3 (e.g., ♩. beat divided into ♪ note sub-beats; includes 6/8, 9/8, "swing", etc.)
(iii) *Complex:* Multiples of prime numbers greater than 3 (e.g., ♩ beat divided into 5 sub-beats)
(?) *A-/hetero-/poly-metric or a-/hetero-/poly-divisive:* See **(1/3)**

N.B. Songs not classified as "*iso-divisive*" must be coded "?" for parameter **(4)**.

5) SYNCOPATION

The percentage of notes that are accented but in a metrically weak position
(i) *Un-syncopated:* <5%
(ii) *Moderately syncopated: 5-20%*
(iii) *Highly syncopated:* >20%
(?) *A-/hetero-/poly-metric:* See **(1)**

6) MOTIVIC REDUNDANCY

The percentage of all notes that are constructed from a single recurring rhythmic pattern
(i) *Non-motivic:* **<**20%
(ii) *Moderately motivic: 20-50%*
(iii) *Highly motivic*: >50%

7) DURATIONAL VARIABILITY

Maximum number of different types of duration values in a song

(i) *Low durational variability:* <3 duration values (e.g., only ♪ and ♩)

(ii) *Moderate durational variability:* 3-4 duration values (e.g., ♪, ♩ and 𝅗𝅥)

(iii) *High durational variability:* >4 duration values (e.g., 𝅘𝅥𝅯, ♪, ♩, 𝅗𝅥 and 𝅝)

B) Pitch

8) TONALITY

Organization of discrete pitches around one or more tonic notes
(a) *Indeterminate a-tonal:* No discrete pitches (e.g., exclamations, heightened speech)
(b) *Discrete a-tonal:* Discrete pitches, but no tonic
(c) *Hetero-tonal:* Tonic modulates/shifts between phrases
(d) *Poly-tonal:* Multiple, simultaneous tonics in different vocal parts
(e) *Iso-tonal:* Single tonic throughout

N.B. Songs not classified as *"iso-tonal"* must be coded "?" for parameters **(9-10)**.

9) MODE

Presence of scale degrees at a minor 3rd (250-350 cents) or major 3rd (350-450 cents) above the tonic

- (a) *A-modal:* No 3rd present
- (b) *Hetero-modal:* Both major and minor 3rd appear in separate phrases
- (c) *Poly-modal:* Both major and minor 3rd appear in the same phrase
- (d) *Minor iso-modal:* Minor 3rd only
- (e) *Major iso-modal:* Major 3rd only
- (?) *A-/hetero-/poly-tonal:* See **(8)**

10) NUMBER OF SCALE DEGREES

Number of scale degrees found in the scale

- (i) *Sparse scale:* <4 scale degrees
- (ii) *Moderately dense scale:* 4-5 scale degrees
- (iii) *Dense scale:* >5 scale degrees
- (?) *A-/hetero-/poly-tonal:* See **(8)**

11) HEMITONICITY

Percentage of melodic intervals that are semitones (50-150 cent intervals)

- (i) *Anhemitonic: <5%*
- (ii) *Moderately hemitonic: 5-20%*
- (iii) *Highly hemitonic:* >20%

12) MELODIC INTERVAL SIZE

Maximum pitch distance between successive notes within any vocal part

- (i) *Small intervals:* <350 cents (i.e., minor 3rd or less; formerly divided into "*monotone*", "*narrow*", and "*diatonic*" intervals)
- (ii) *Medium intervals:* 350-750 cents (i.e., major 3rd - perfect 5th; formerly divided into "*wide*" and "*very wide*" intervals)
- (iii) *Large intervals:* >750 cents (i.e., minor 6th or greater)

13) MELODIC RANGE

Maximum pitch distance between the highest and lowest notes within any vocal part

- (i) *Small range:* <750 cents (i.e., perfect 5th or less)
- (ii) *Medium range:* 750-1250 cents (i.e., perfect 5th - octave)
- (iii) *Large range:* >1250 cents (i.e., more than an octave)

14) MELODIC CONTOUR

Shape resulting from all changes in interval direction within a vocal part

- (a) *Horizontal:* No ascending or descending intervals
- (b) *Ascending:* Ascending intervals only
- (c) *Descending:* Descending intervals only

- (d) *U-shaped:* First descending, then ascending intervals
- (e) *Arched:* First ascending, then descending intervals
- (f) *Undulating:* Multiple changes of interval direction

C) Syllable

15) MELISMA

Maximum number of consecutive notes without articulating a new syllable

- (i) *Syllabic:* 1-2 notes
- (ii) *Mildly melismatic:* 3-5 notes
- (iii) *Strongly melismatic:* >5 notes

16) VOCABLES

The percentage of syllables containing only vowels or semi-vowels (e.g., "y", "h", "w")

- (i) *Few vocables:* <20%
- (ii) *Some vocables:* 20-50%
- (iii) *Many vocables:* >50%

II) "TEXTURE" (between-part)

17) NUMBER OF VOCAL PARTS

Maximum number of simultaneous vocal parts

- (i) *One-part:* 1 (formerly divided into "*solo*" and "*unison*")
- (ii) *Two-part:* 2
- (iii) *Many-part:* >2

N.B. Songs classified as "*one-part*" (including both unison and solo songs) must be coded "?" for parameters **(18-20)**.

18) RHYTHMIC TEXTURE (Cantometrics Line 12)

Temporal asynchrony in the relative onsets of different vocal parts (in seconds)

- (a) *Hetero-rhythmic (heterophonic):* 0.1–1s
- (b) *Poly-rhythmic (polyphonic):* >1s
- (c) *Iso-rhythmic (homophonic):* **<**0.1s
- (?) *One-part (monophonic):* See **(17)**

N.B. Songs not classified as "*iso-rhythmic*" must be coded "?" for parameter **(19)**.

19) HARMONIC TEXTURE

Minimum harmonic interval (octave-equalized) between simultaneous vocal parts that is sustained for at least 1 second

- (i) *Dissonant:* 50-250 cents (e.g., $2^{nds}/7^{ths}$)
- (ii) *Consonant:*>250 cents (e.g., 3^{rds}-6^{ths})
- (?) *One-part, or poly-/hetero-rhythmic:* See **(17/18)**

20) RELATIVE MOTION

Relationship of the melodic contours (see **13**) of two simultaneous parts
- (a) *Hetero-contour (drone):* One part is horizontal, the other changes direction
- (b) *Poly-contour (independent motion):* Both parts have different, non-horizontal contours
- (c) *Iso-contour (parallel motion):* Both parts have the same contour
- (?) *One-part: S*ee **(17)**

III) "FORM" (between-phrase)

21) PHRASE REPETITION

Maximum number of successive phrases before a phrase is repeated
- (i) *Non-repetitive*: >8 phrases, or no repeat at all
- (ii) *Moderately repetitive:* 3-8 phrases
- (iii) *Repetitive:* 1-2 phrases

22) PHRASE LENGTH

Maximum phrase length, in seconds
- (i) *Short phrases:* **<**5 s
- (ii) *Medium-length phrases:* 5-9 s
- (iii) *Long phrases*: >9 s

23) PHRASE SYMMETRY

Ratio of the length of the longest phrase in a song relative to the shortest phrase
- (i) *Symmetric:* **<**1.5 times the length of the shortest phrase
- (ii) *Mildly asymmetric:* 1.5-2.5 times the length of the shortest phrase
- (iii) *Very asymmetric*: >2.5 times the length of the shortest phrase

24) SOLO/GROUP ARRANGEMENT

Number of singers in each phrase
- (a) *Solo:* Only solo phrases throughout
- (b) *Mixed:* Individual phrases contain both group and solo sub-sections
- (c) *Alternating:* Alternation between distinct solo and group phrases
- (d) *Group:* Only group phrases throughout

25) RESPONSORIAL ARRANGEMENT

Alternation of phrases between different vocal parts
- (a) *A-responsorial:* No alternation between parts
- (b) *Hetero-responsorial:* Irregular alternation between parts
- (c) *Iso-responsorial:* Consistent alternation between parts

N.B. Songs classified as "*a-responsorial*" must be coded "?" for parameter **(26)**.
Comments: See comments in **(24)**.

26) PHRASE OVERLAP

Maximum overlap between a "call" phrase and the "response" phrase that alternates with it (as the percentage of time in which the latter phrase overlaps with the former)

(i) *Non-overlapping:* 0%
(ii) *Mildly overlapping:* 1–25%
(iii) *Highly overlapping:* >25%
(?) *A-responsorial:* See **(25)**

S2 – Song-to-song distance matrix algorithm description

a) Preparation of data

The non-independent nature of some of the characters in the CantoCore scheme results in the potential for redundancy to be introduced into the codings. For example, if the character “meter” is coded as “a-metric” (i.e., no recognisable beat), then the codings for the characters “number of beats”, “sub-beat division”, and “number of sub-beats” carry the same information, and this can overemphasise the importance of the absence of a beat. In order to reduce this potential redundancy, a “?” is put in place of redundant codings and is treated as missing data. Uncoded characters are also denoted with a “?”.

b) Ordinal and nominal variables

CantoCore contains both ordinal and nominal characters. The first part of the distance-matrix algorithm separates ordinal and nominal variables, which are defined a priori by the user. Ordinal characters are coded using lower-case Roman numerals (i-iii in the case of CantoCore) and are never multi-coded. Nominal characters are coded using lower-case letters (a-f in the case of CantoCore), and can be multi-coded where appropriate.

c) Ordinal characters

Ordinal characters can have a number of possible character states, and this can vary across characters. In order to keep this consistent across ordinal characters, as well as for ordinal comparisons to be equivalent to nominal comparisons, the raw codings of the ordinal character states are converted to *scaled* values from 0 to 1 such that the minimum ordinal value is coded as 0 and the maximum ordinal value is coded as 1, with intermediate states taking intermediate values (see figure S1 a-b). In CantoCore, all ordinal characters have three possible states (i, ii, or iii). As a result, i becomes 0, ii becomes 0.5, and iii becomes 1. For a character with five states (i, ii, iii, iv, and v), the converted values would become 0, 0.25, 0.50, 0.75, and 1, respectively.

The algorithm then creates a separate pairwise distance matrix for each ordinal character by taking the absolute difference between the scaled codings for each pair of songs (Figure S1 b-c):

$$|\text{Song X} - \text{SongY}|$$

The maximum possible difference is 1 (if the codings are maximally different) and the minimum value is 0 (if the codings are identical). For pairs of songs where one or both songs lack codings for that character (“?”), the distance is listed as NA (See Figure S1 c). The end result is a separate pairwise distance matrix for each ordinal character, from 1-j.

(a)

Song	Ordinal variable j
Song A	iii
Song B	i
Song C	?
Song D	ii

i = 0
ii = 0.5
iii = 1
? = ?

Codings scaled from 0-1

(b)

Song	Ordinal variable j
Song A	1
Song B	0
Song C	?
Song D	0.5

Distance calculated as
| Song X – Song Y |

(c)

j	Song A	Song B	Song C	Song D
Song A	0	1	NA	0.5
Song B	1	0	NA	0.5
Song C	NA	NA	NA	NA
Song D	0.5	0.5	NA	0

Figure S1. A sample calculation of pairwise distance across 4 songs (A-D) for a single ordinal variable (j). (a) The raw CantoCore codings. (b) The same codings after they have been converted into a scale from 0 to 1. (c) A sample distance matrix based on the absolute difference of the scaled codings for each pair of songs. Any pairwise distance involving a redundant or uncoded character (?) is denoted as NA. This overall process is repeated for each ordinal character from 1-j, resulting in j distance matrices for the song set.

d) Nominal characters

In order to accommodate the potential for multi-coded characters, the matrix of raw codings (comprised of lower-case letters) is converted into a "presence-absence" matrix, as in Busby (2006), where Y denotes the presence of a character state, and N denotes its absence (see Figure S2 a-b). Character states coded as "?" (representing redundant or missing codings) are denoted as NA when they are converted into the presence-absence matrix and are thus treated as missing data.

The conversion of scores in the presence-absence matrix into distances is based on pairwise matching of songs across all possible character states. For example, if one song contains an “a” coding (Y for character state a) and another song lacks it (N for character state a), then the program scores this as a 1, implying maximum distance between the two songs. If both songs contain a “b” coding (i.e., both are scored as Y for character state b), then the program scores this as a 0, implying minimum distance between them.

(a)

Song	Nominal variable k
Song A	ad
Song B	c
Song C	?
Song D	d

Conversion of raw codings into a presence-absence matrix

Y = Present
N = Absent
NA = ?

(b)

	Character states			
Var. k	a	b	c	d
Song A	Y	N	N	Y
Song B	N	N	Y	N
Song C	NA	NA	NA	NA
Song D	N	N	N	Y

Calculation of mean pairwise distance across all character states

If both Y, then = 0
If one Y and other N = 1
If both N, then don’t include in mean
If one or both NA, then denote NA

(c)

k	Song A	Song B	Song C	Song D
Song A	0	(1+1+1)/3=1	NA	(1+0)/2=0.5
Song B	(1+1+1)/3=1	0	NA	(1+1)/2=1
Song C	NA	NA	NA	NA
Song D	(1+0)/2=0.5	(1+1)/2=1	NA	0

Figure S2. A sample calculation of mean distance for a nominal variable. Raw CantoCore codings (a) are first converted into a “presence-absence” matrix (b), where Y denotes the presence of a character state, N denotes its absence, and NA represents a redundant coding for each song. The mean pairwise distance is calculated by taking the pairwise distance between songs across all character states, except those involving mutual absence, which are ignored. This process is repeated for each nominal variable from 1-k.

If both songs lack a particular character state, for example c for Songs A and D (where both are scored as N), then this mutual absence is ignored and is not incorporated into the mean distance calculation. This is done because the mutual absence of a character state is uninformative.

Finally, to calculate the mean pairwise distance for a particular character, we take the pairwise distances between songs across all character states for that character, except those involving mutual absence, which are ignored. The occurrence of mutual absence results in some distances (means) containing fewer comparisons than others. As a result, the denominator in the mean calculation is variable. For example, the AB distance contains three comparisons while the AD distance contains only two, since the latter pair has two mutual absences compared to only one for the former pair. As for the ordinal characters, 1 is the maximum possible mean distance, and 0 is the minimum. If one or both of the songs of a pair contain NA's anywhere in their fields (because the raw coding was a "?", as with Song C), the pairwise distance is denoted as NA. A separate distance matrix is created this way for each nominal variable from 1-k.

e) Combining ordinal and nominal characters into a final distance measure

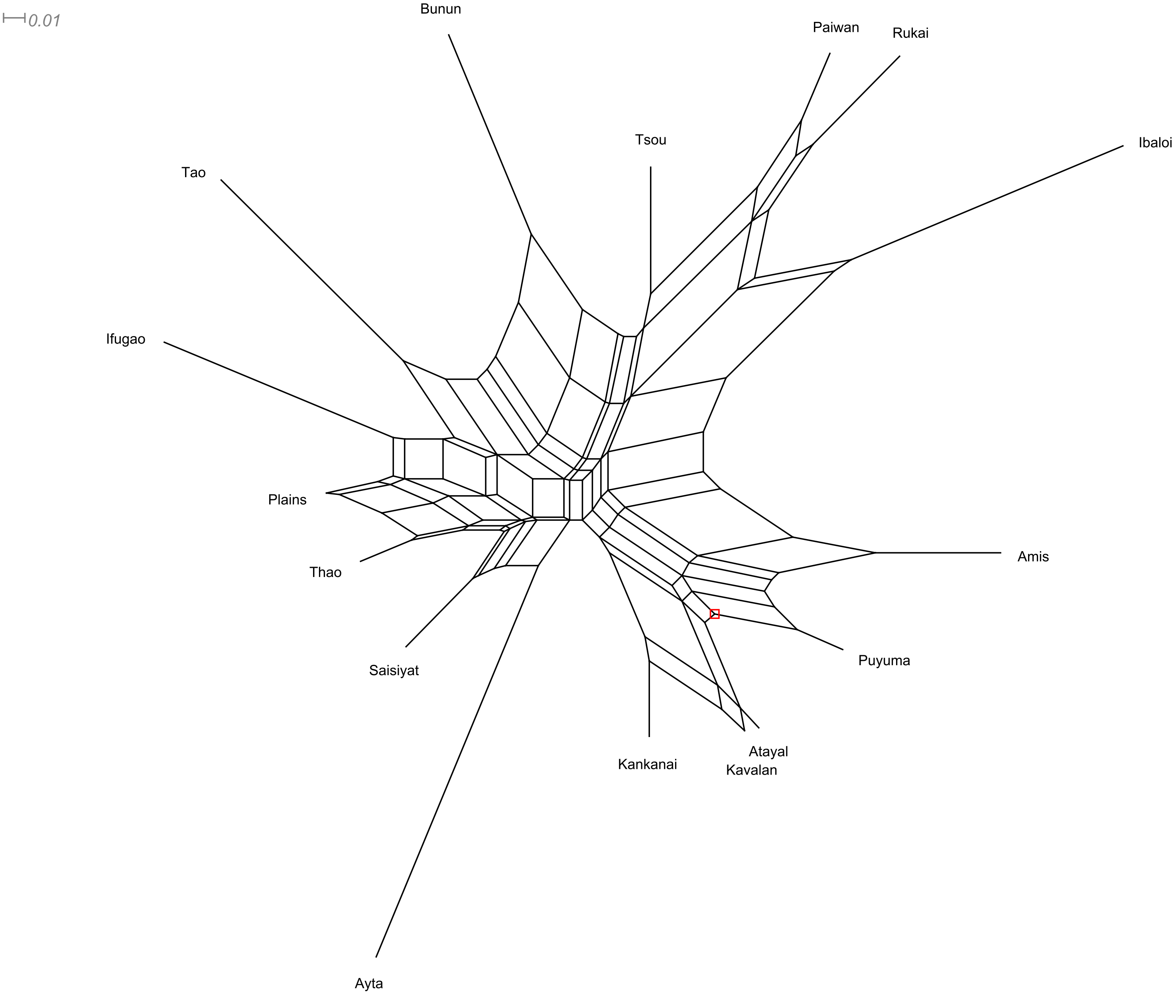
The final step of the algorithm combines information from the j ordinal variables with the k nominal variables to obtain an overall measure of distance between songs. For each pair of songs, the mean distance across all characters is taken, ignoring any distances denoted as NA. As a result, the final measure of distance incorporates information from each character equally, ignoring only redundant codings or uncoded characters.

Table 2.S3 – Modal Profiles

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Amis	d	a	c	a	ii	iii	i	e	d	ii	b	ii	iii	f	i	iii	i	b	?	b	iii	iii	ii	c	c	ii
Atayal	d	a	c	a	i	iii	i	e	d	ii	b	ii	ii	f	i	i	i	?	?	?	ii	ii	i	d	a	?
Ayta	a	?	?	?	?	ii	i	e	a	ii	b	ii	ii	c	i	ii	i	?	?	?	iii	ii	i	c	c	ii
Bunun	a	a	?	?	i	iii	ii	e	e	ii	b	ii	iii	a	ii	i	iii	c	ii	ca	iii	iii	i	b	a	?
Ibaloi	a	?	?	?	?	ii	iii	e	e	ii	a	ii	ii	fe	ii	ii	i	?	?	?	iii	iii	iii	c	c	i
Ifugao	a	?	?	?	?	ii	ii	e	?	?	?	ii	i	f	iii	ii	iii	ac	i	cb	iii	ii	ii	d	a	?
Kankanai	d	?	?	?	?	iii	i	e	e	ii	b	ii	ii	f	i	i	i	?	?	?	iii	ii	ii	d	a	?
Kavalan	d	a	c	a	i	iii	i	e	d	ii	b	ii	ii	f	i	i	i	?	?	?	ii	ii	ii	d	a	?
Paiwan	b	?	?	?	?	iii	i	e	e	ii	a	ii	ii	f	ii	i	ii	c	ii	a	iii	iii	ii	c	a	?
Plains	a	?	?	?	?	ii	ii	e	d	ii	b	ii	iii	f	ii	ii	i	?	?	?	iii	ii	i	d	a	?
Puyuma	d	a	c	a	i	iii	i	e	d	ii	b	ii	iii	f	ii	iii	i	?	?	?	ii	iii	ii	d	a	?
Rukai	b	?	?	?	?	iii	ii	e	e	ii	a	ii	ii	fa	ii	i	i	?	?	?	iii	iii	ii	cb	a	?
Saisiyat	a	?	?	?	?	iii	i	e	d	i	b	ii	ii	f	ii	ii	i	?	?	?	iii	ii	i	c	a	?
Tao	a	?	?	?	?	ii	iii	e	d	i	b	ii	i	a	i	i	i	?	?	?	iii	iii	i	d	a	?
Thao	a	?	?	?	?	iii	i	e	d	ii	b	ii	iii	f	ii	ii	i	?	?	?	ii	ii	i	d	a	?
Tsou	d	?	?	?	?	iii	i	e	e	ii	b	ii	iii	f	ii	ii	ii	c	ii	c	iii	iii	ii	c	a	?

We created modal profiles by taking the most commonly coded character state for every CantoCore character in each population's musical repertoire. Numbers across the top correspond to the 26 CantoCore categories found in S1. Nominal character states are coded as letters and multiple states are permitted. Ordinal characters are coded as numbers. "?" codings are treated as missing data, because their inclusion would carry information redunhdant with the coding of another character.

Figure 2.S4 – Neighbor-Net of modal profile distances



Chapter 3

Co-evolution of music and genes in aboriginal Taiwan

Patrick E. Savage[1], Tom Rzeszutek[1], Victor Grauer[2], Ying-fen Wang[3], Jean Trejaut[4], Marie Lin[4], and Steven Brown[1]

[1]Department of Psychology, Neuroscience & Behaviour, McMaster University, Canada
[2] Independent scholar, Pittsburgh, USA
[3]Graduate Institute of Musicology, National Taiwan University, Taiwan
[4]Transfusion Medicine Laboratory, Mackay Memorial Hospital, Taiwan

Keywords: music; genetics; evolution; migration; Cantometrics

3.1 Abstract

A recent claim that music can act as a marker of ancient population migrations has been criticized on the grounds that music's time-depth seems too shallow (i.e., music changes too rapidly to preserve deep relationships). We predicted that, if any musical features were to have the necessary time-depth, they would be the structural features – rather than performance features – of group songs. To test this prediction, we used Cantometric classifications of 234 traditional group songs from 9 Taiwanese aboriginal tribes to create separate distance matrices for music based on either structural or performance features. Both distance matrices demonstrated a positive correlation with distances based on mitochondrial DNA, a migration marker with well-established time-depth. However, this correlation was only statistically significant for the distance matrix based on song structure, which accounted for more than twice as much variance in the genetic data as did performance style. These results, although preliminary, provide the first quantitative evidence suggesting that songs and genes may have followed parallel evolutionary trajectories on the order of several thousand years. Follow-up analyses using different coders, different classification schemes, and different genetic data-sets, while partially confirming these patterns, highlight substantial challenges in obtaining and interpreting reliable samples for both music and genes.

3.2 Introduction

The discovery that our genes can be used to trace the migration of all anatomically modern humans back to a single African "mitochondrial Eve" has had an enormous impact on our understanding of human pre-history (Cann, Stoneking, & Wilson, 1987). Recently, Grauer (2006) proposed that music, too, traces these same migrations back to Africa, but critics argued that music changes too rapidly to preserve ancient relationships (e.g., Stock, 2006). These opposing claims present an excellent opportunity to bring new quantitative techniques and data from genetic anthropology to old questions of how and why music varies cross-culturally.

Grauer's claim drew on data from the landmark Cantometrics Project (Lomax, 1968), which remains the only global scientific study of human song. At the time, Lomax's causal interpretation of the correlation between culture and music – for example, male dominance causing nasal singing – was highly criticized even by other members of the Cantometrics Project (e.g., Erickson, 1976). While Grauer's recent migratory interpretation avoids Lomax's pitfall, Stock's rebuttal still echoed many of the original criticisms of the Cantometrics Project (e.g., Blacking, 1977; Downey, 1970). Could the acoustic surface of music really reflect ancient connections between cultures? If so, are these reflected in performance features, or in the structural features traditionally emphasized in Western musicology?

Lomax himself was highly critical of the use of Western musical notation in ethnomusicology, which he saw as emphasizing surface structural features at the expense of deeper performance features. He spent his life developing a performance-oriented approach that was concerned "not with songs abstracted from the stream of vocalizing we encountered on the tapes, but with the stream itself, with 'singing'" (Lomax, 1980). Nevertheless, the Cantometric classification scheme that Lomax and Grauer (1968) developed contained roughly equal numbers of features devoted to song structure and singing style (see Fig. 3.1).

Our own view differs from both Lomax's and his critics' in that we propose that the structural features of song should have the greatest time-depth to track migrations, especially when applied to group songs. Our reasoning is that structural features such as melody, texture and form require greater consensus among singers than the more idiosyncratic variation that goes into performance, such as timbre or ornamentation. Hence, features like scales and rhythms should be more stable over time than features like nasality or rubato.

These claims are testable. As a case-study to examine music's time-depth in the context of human migrations, we have examined the traditional group songs of the aboriginal tribes of Taiwan, who have been well-studied in terms of music, genetics, linguistics, and migration (Diamond, 2000; Kurosawa, 1973; Trejaut et al., 2005). Our primary aim, therefore, was to use existing information about the relative patterns of genetic and musical similarity among the Taiwanese aboriginal tribes to empirically test for the first time whether song structure or singing style has the time-depth required for studying human migrations. Our basic method was to compare music – a marker of unknown time-depth – against the best available marker with a well-established time-depth, namely mitochondrial DNA (mtDNA) (Oppenheimer, 2004).

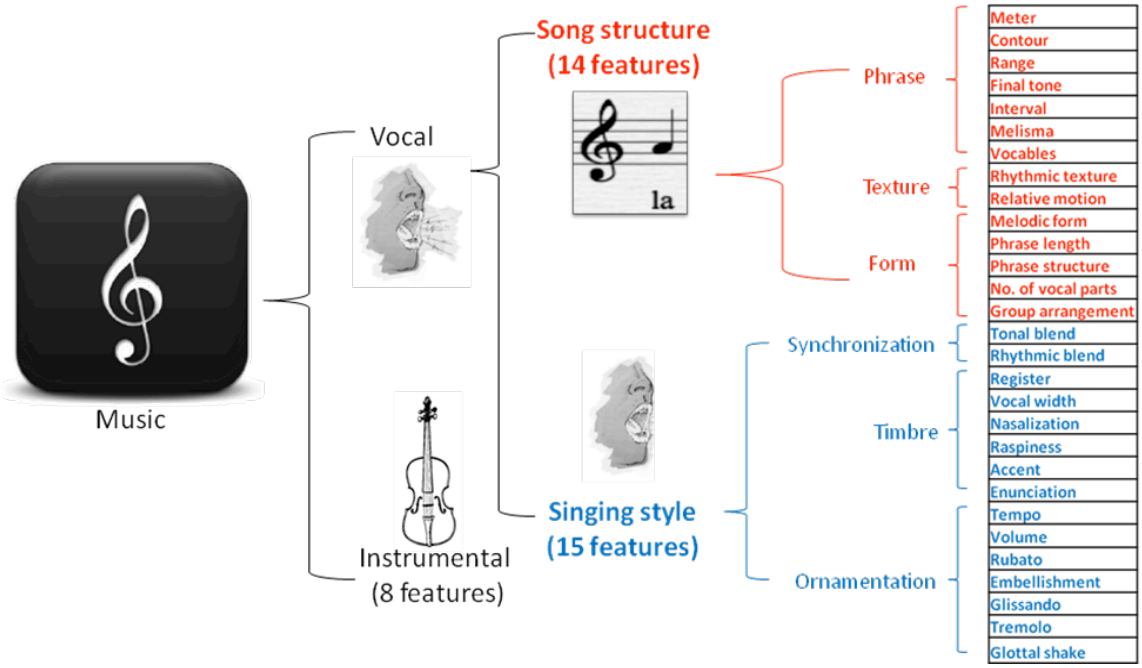


***Fig. 3.1:* Organization of the 14 song-structure (red) and 15 singing-style (blue) Cantometric classification features used in this analysis. Note that our method focuses on the vocal component of the music and therefore ignores 8 classification features related to instruments.**

3.3 Methods

Genetic sample: Genetic samples were obtained from JT and ML's previously published data (Trejaut et al., 2005). These samples consist of hypervariable segments 1 and 2 in the control region of the maternally inherited, haploid mitochondrial genome from 640 individuals. After editing, a 717-base-pair (bp) nucleotide string was used in the current analysis, with the same sample sizes previously reported.

Musical sample: Using the Taiwan National Music Archive[1], we obtained 234 traditional group songs from all 9 of the tribes whose mtDNA data was published in Trejaut et al. (2005). For populations with 30 or fewer eligible songs, all eligible songs in the archive were used. For populations with over 30 eligible songs, we randomly sampled 30 songs per population. Sample sizes were: Amis=30, Atayal=8, Bunun=30, Paiwan=30, Puyuma=30, Rukai=30, Saisiyat=26, Tao (Yami)=29, Tsou=21.

Distances between samples: Pairwise distances between individual genetic samples were calculated based on the number of pairwise differences between nucleotide sequences. This is the simplest possible measure of genetic distance, as it is based on phenetic (similarity) relationships rather than phylogenetic (evolutionary) relationships that require

---

[1] http://music.ncfta.gov.tw (in Chinese)

additional assumptions about how differences arose. Building on the work of Leroi and Swire (2006) and Busby (2006), we developed a comparable phenetic distance measure for Cantometrics that accounts for various statistical complications inherent in converting musical classification schemes into quantitative distances (for details, see Rzeszutek, Savage and Brown, in prep.; see Chapter 2 of this thesis). Cantometric classifications of the songs were coded by VG. Two separate musical distance-matrices were calculated: one using the 14 song-structure characters from Cantometrics, the other using the 15 singing-style characters (see Fig. 3.1 for details about these features). Eight Cantometric characters related to instruments alone were excluded from this analysis.

Distances between populations: For both the genetic and musical data, the 36 possible pairwise distances among the 9 tribes were calculated using the Analysis of Molecular Variance (AMOVA) framework (Excoffier, Smouse, & Quattro, 1992), as in Rzeszutek et al. (in prep.; see Chapter 2 of this thesis). These distances were measured using a statistic called ϕ_{ST}, which represents the proportion of variability among individual samples that is due to between-population differences. Thus, it explicitly incorporates within-population heterogeneity, avoiding the assumptions of within-population homogeneity that plagued Lomax's original statistical methodology (e.g., Henry, 1976; Leroi & Swire, 2006).

Correlations: The statistical significance of the correlations between musical and genetic distances was tested using the permutation-based Mantel test using 20,000 permutations and Spearman's rho, with the threshold for significance set at $p < 0.05$ (one-tailed). This test controls for the fact that the 36 pairwise distances among the nine tribes are not independent of one another. We used a rank-order Spearman correlation due to the distance nature of our musical and genetic datasets.

Follow-up analyses: In order to explore the effects of using different samples, different classification schemes, and different coders, we also applied the above methodology to the same set of songs coded by PES using an alternative classification scheme, CantoCore (Savage, Merritt, Rzeszutek, and Brown, in prep.; see Chapter 2 of this thesis), which is based on Cantometrics, but only contains features related to song structure. Second, we compared both VG's and PES's codings with those of a third coder, Emily Merritt, who was blind to the tribal affiliations of the songs. Inter-rater reliability values were calculated using $\varkappa$ as in Savage et al. (in prep.) for a pseudo-randomly selected subset of 45 songs (5 songs randomly selected per tribe). Finally, we obtained an unpublished high-resolution dataset from Albert Ko, Fred Delfin, and Mark Stoneking, containing whole-genome mtDNA samples from 407 individuals from the same nine Taiwanese populations and from 90 individuals from three aboriginal Philippines tribes (Ibaloi, Ifugao, and Kankanai) whose songs are analyzed in Rzeszutek et al. (in prep.; see Chapter 3 of this thesis). We performed the same analyses as above with these new data, once using the same 717-bp sub-section of the mtDNA genome as we used for Trejaut et al.'s data, and again using the whole 16,569-bp mtDNA genome.

3.4 Results

Population-level musical distances based on VG's Cantometric codings were almost, but not quite, significantly correlated with those based on Trejaut et al.'s published genetic data (r_s=0.30, p=0.07). In accordance with our predictions, this correlation appeared to be driven by the songs' structural features, as the correlation based on separate distance matrices for structural vs. performance features became significant only for feature of song structure, which accounted for more than twice as much variance in genetic distance as did features of singing style (structure: r_s=0.35, p=0.03; performance: r_s=0.22, p=0.14) (Figure 3.2). Perfomance features were themselves highly correlated with features of song structure (r_s=0.73, $p < 0.001$). Further analysis using the CantoCore coding scheme, based exclusively on features of song structure that were coded by a different rater, was also significantly correlated with Trejaut's data (r_s=0.32, p=0.045), replicating the above result with the structural features of Cantometrics (see Fig 3.3). Both classification schemes demonstrated similar reliability (Cantometrics: mean $\varkappa$ = 0.28; CantoCore: mean $\varkappa$ = 0.30) that was significantly greater than chance ($p<1x10^{-5}$) but not significantly different from one another (p=0.68).

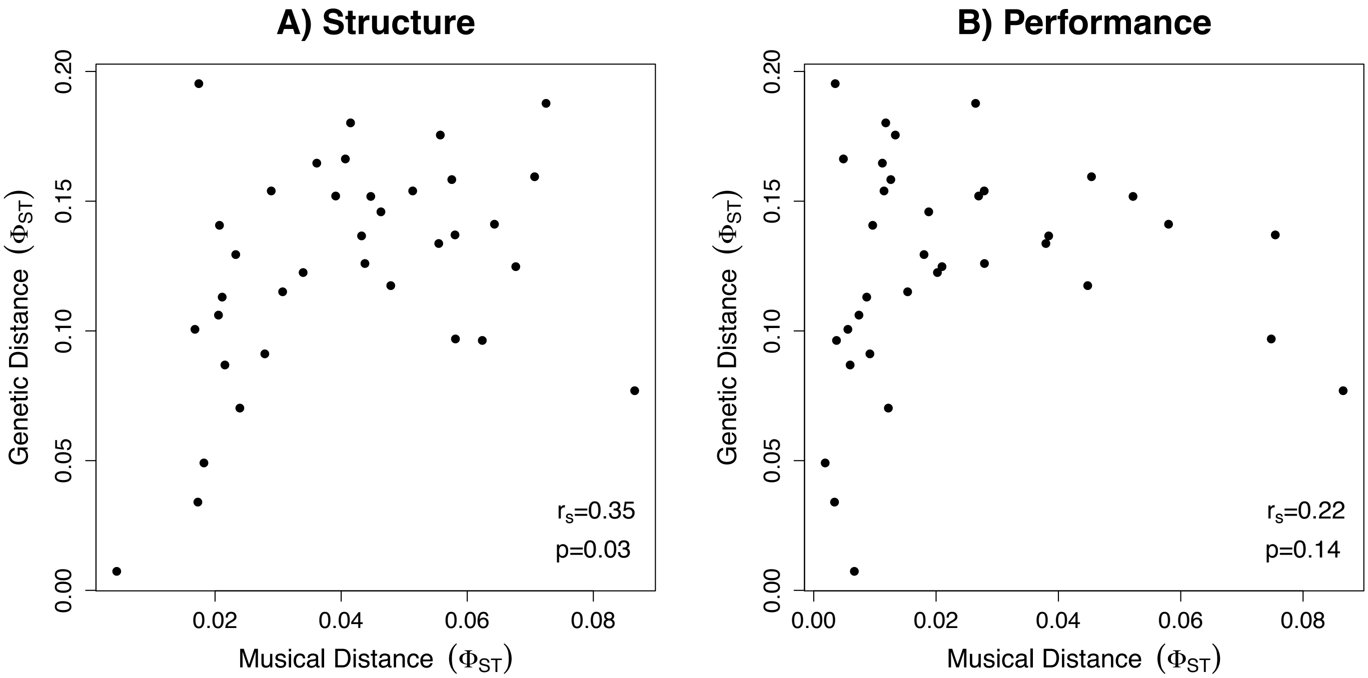


Fig. 3.2: **Scatterplots of the 36 pairwise musical and genetic distances among 9 Taiwanese aboriginal tribes. Genetic distances (y-axis) are based on an Analysis of Molecular Variance (AMOVA) of 640 mitochondrial DNA haplotypes from Trejaut et al. (2005). Analogous musical distances (x-axis) were calculated from 234 traditional choral songs using Cantometric characters of either A) song structure or B) singing style (i.e., performance). Statistical significance of distance-matrix correlations is based on a Mantel test and Spearman's rho.**

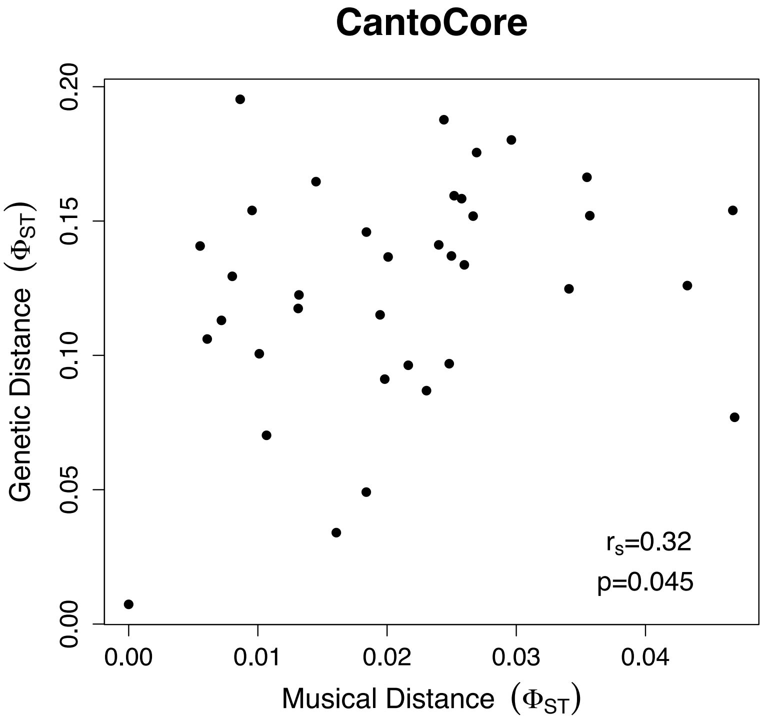


Fig. 3.3: **Follow-up analysis using the CantoCore coding scheme, focused exclusively on aspects of song structure confirms the correlation shown in Fig. 3.2. Genetic distances (y-axis) are based on an Analysis of Molecular Variance (AMOVA) of 640 mitochondrial DNA haplotypes from Trejaut et al. (2005). Musical distances were calculated in the same manner as in the previous analysis.**

Follow-up analyses using the new Taiwanese genetic data-sets were less conclusive. The comparison between VG's Cantometric codings and Ko et al.'s alternative genetic sample gave almost identical r_s values, although slight changes in the p-values pushed the correlation with structural features just barely beyond the threshold for significance (structure: r_s=0.35, p=0.051; performance: r_s=0.17, p=0.22). Surprisingly, however, this replication only held for Ko et al.'s data when we restricted the analysis to the same 717-bp sub-section of the mtDNA genome as was published by Trejaut et al. (2005). When we reran the analysis of Ko et al.'s data using the entire mtDNA genome, the correlations between musical and genetic similarities dropped dramatically (structure: r_s=0.2, p=0.21; performance: r_s=0.03, p=0.53). Although the same pattern of structural features being more strongly correlated with performance features was maintained, both correlations were now nowhere near statistical significance. Also surprising, the correlation between PES's CantoCore codings and the new genetic data-set were substantially reduced when using the new partial-genome data as well as the whole-genome data (partial: r_s=0.22, p=0.16; full: r_s=0.17, p=0.24).

Finally, introducing the new data from the three Philippines groups seemed to completely remove any trace of a significant correlation. When these groups were added, all possible permutations of musical features, classification scheme, genetic sample and mtDNA section gave similar, non-significant results (r_s<0.17, p>0.18).

3.5 Discussion

Our main finding was that musical similarities based on structural features among the 9 Taiwanese aboriginal tribes were significantly correlated with genetic similarities. This provides the first empirical support for Grauer's (2006) claim that music may have the time-depth required for use as a marker in studying prehistoric human migrations. Consistent with our predictions, performance features did not reflect the genetic relationships between populations as strongly as structural features did. The simplest interpretation is that the way a song is performed may be a more rapidly changing or malleable that the underlying structure of the song, which may be relatively more constrained. Similar dissociations in evolutionary rates among different features is also found in linguistics and genetics, where some words are more resistant to borrowing than others (McMahon & McMahon, 2005) and mutation rates vary across the human genome (Nachman & Cromwell 2000).

This does not mean that performance style is not important. The strongest correlations that Lomax originally found were between performance features and social structure (such as vocal tension correlating with sexual restrictiveness), leading him to conclude that structural features such as melody, meter and harmony "…are not connected to the prime characterizers of social and cultural structures" (Lomax, 1980:52). Lomax was fully aware of the complicated interrelationship between music, social structure, and migration, but was never able to resolve debates about the causality of these relationships. Perhaps the differential transmission of structural and performance features is such that song structure better reflects ancient population migrations, while singing style is a better marker of social structure and/or more recent patterns of movement and cultural contact that occurred in the absence of substantial gene flow. This has important implications for researchers wanting to use music to study human history.

Our follow-up analyses highlight many methodological challenges in quantifying and interpreting relationships between music and culture. Concerns about reliability, sampling, and the coarse nature of the classification scheme long dogged Cantometrics (Dowling & Harwood, 1986; Maranda, 1970; Nettl, 1970). Although our partial replications using a different classifying scheme, different coder and different genetic sample suggest that some of these effects on analyses are minor, the surprising dissociation between the correlations with the whole mtDNA genome and the partial mtDNA genome suggest that even in population genetics, a coarse analysis can substantially affect the data. Unfortunately, to our knowledge, there is currently no higher-resolution alternative to Cantometrics or CantoCore that is cross-culturally appropriate. This may change soon, however, if the new field of "computational ethnomusicology" (Tzanetakis, Kapur, Schloss & Wright, 2007) begins using these systems as a theoretical framework to build automated algorithms for cross-cultural classification and analysis.

The complete disappearance of correlations when we included the Philippines samples is difficult to interpret. On the one hand, it may imply that music does not function as a marker of prehistoric population relationships in the Philippines, perhaps due to different colonial and political histories. However, it may also reflect differences in the musical samples themselves. We were fortunate to be able to collaborate with a Taiwanese ethnomusicologist (YW) with expertise in traditional aboriginal music and work with a well-documented Taiwanese archive containing extensive liner notes, many of which were also published commercially. On the other hand, most of the recordings in the Philippines archive are unpublished field recordings containing long, continuous recordings of performances interspersed with talking, laughter, eating, etc. We had no ethnographic notes or ethnomusicological expertise to rely on regarding which songs were representative, which songs were repeat performances of the same song, etc., and thus had to simply use whatever recordings we could find. Our inability to interpret our Philippines data thus highlights the importance of musical sampling and the need for collaboration with expert ethnomusicologists.

Further work is needed to determine the extent to which the genetic and musical connections between the tribes are due to recent admixture as opposed to parallel co-evolutionary isolation and drift after diverging from one or more founding populations. For example, while the Paiwan and Rukai languages are mutually unintelligible due to differences that have evolved in the ~6,000 years since their ancestral language, proto-Austronesian, first arrived on Taiwan (Gray, Drummond, & Greenhill, 2009), our analyses found that they are not significantly different from each other either musically or genetically. It appears that musical and genetic similarities between the two tribes are a recent development due to extensive intermarriage and musical exchange, with linguistic differences maintained through bilingualism (Huteson, 2003). For example, the type of drone polyphony that is now so distinctive of both tribes was performed only by the Rukai before World War II (Kurosawa, 1973).

Regardless of how and when the similarities arose, however, our findings in Taiwan lend provisional support for the ability of musical structure to track population movements in the same manner as do genes. Whether the co-evolution and co-migration of music and genes extends as far back as Grauer's Out-of-Africa claim, however, remains an open empirical question.

ACKNOWLEDGMENTS

We would like to thank Albert Ko, Fred Delfin, and Mark Stoneking for sharing their unpublished genetic data; Ramon Santos and the University of the Philippines Center for Ethnomusicology for sharing their unpublished archival recordings; and Emily Merritt for coding the songs for the reliability analysis. This work was supported by the Social Sciences and Humanities Research Council (SSRHC) of Canada, grant No 90–2320-B-195–002 from the National Science Council of Taiwan, grant NHRIEX93-9218BI from the National Health Research Institute, and by an Amherst College Roland Wood Fellowship to P.E.S.

REFERENCES

Blacking, J. (1977). Some problems of theory and method in the study of musical change. *Yearbook of the International Folk Music Council*, *9*, 1–26.

Busby, G. (2006). Finding the blues: An investigation into the origins and evolution of African-American music. 1-55. (Unpublished Master's Thesis)

Cann, R. L., Stoneking, M., & Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature*, *325*, 31-36.

Diamond, J. (2000). Taiwan's gift to the world. *Nature*, *403*(6771), 709-710.

Dowling, W. J., & Harwood, D. L. (1986). Cultural contexts of musical experience. In W. J. Dowling & D. L. Harwood (Eds.), *Music Cognition* (pp. 225-239). Orlando: Academic Press.

Downey, J. C. (1970). Review of A. Lomax, Folk song style and culture. *Ethnomusicology*, *14*(1), 63-67.

Erickson, E. E. (1976). Tradition and evolution in song style: A reanalysis of Cantometric data. *Cross-Cultural Research*, *11*(4), 277-308.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479-491.

Grauer, V. (2006). Echoes of our forgotten ancestors. *The World of Music*, *48*(2), 5-59.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, *323*, 479-483.

Henry, E. O. (1976). The variety of music in a North Indian village: Reassessing Cantometrics. *Ethnomusicology*, *20*(1), 49-66.

Huteson, G. (2003). Sociolinguistic survey report for the Tona and Maga dialects of the Rukai language. *SIL Electronic Survey Reports*, *12*.

Kurosawa, T. (1973). *The music of Takasago tribe in Formosa (in Japanese)*. Tokyo: Oyamakaku.

Leroi, A. M., & Swire, J. (2006). The recovery of the past. *The World of Music*, *48*(3), 43-54.

Lomax, A. (1980). Factors of musical style. In S. Diamond (Ed.), *Theory & practice: Essays presented to Gene Weltfish* (pp. 29-58). The Hague: Mouton.

Lomax, A. (Ed.). (1968). *Folk song style and culture*. Washington, DC: American Association for the Advancement of Science.

Lomax, A., & Grauer, V. (1968). The Cantometric coding book. In A. Lomax (Ed.), *Folk song style and culture* (pp. 34-74). Washington, DC: American Association for the Advancement of Science.

Maranda, E. K. (1970). Deep significance and surface significance: Is Cantometrics possible? *Semiotica*, *2*(2), 173-184.

McMahon, A. M. S., & McMahon, R. (2005). *Language classification by numbers*. Oxford: Oxford University Press.

Nachman, M. W., & Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*, *156*(1), 297-304.

Nettl, B. (1970). Review of A. Lomax, Folk song style and culture. *American Anthropologist*, *72*(2), 438-441.

Oppenheimer, S. (2004). The "express train from Taiwan to Polynesia": On the congruence of proxy lines of evidence. *World Archaeology*, *36*(4), 591-600.

Stock, J. P. J. (2006). Clues from our present peers? A response to Victor Grauer. *The World of Music*, *48*(2), 73-91.

Trejaut, J. A., Kivisild, T., Loo, J. H., Lee, C. L., He, C. L., Hsu, C. J., et al. (2005). Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biology*, *3*(8), 1362-1372.

Tzanetakis, G., Kapur, A., Schloss, W.A., & Wright, M. (2007). Computational ethnomusicology. *Journal of Interdisciplinary Music Studies 1*(2), 1-24.

Chapter 4

General Discussion

The study of comparative musicology has historically shown some promise (Lomax, 1968), but is neglected by present-day ethnomusicologists. Chapter 2 establishes a new approach to studying and quantifying music cross-culturally. Previous approaches to establish worldwide musical comparisons, such as the Cantometrics project, have looked at differences across populations at the cost of ignoring the inherent internal diversity. The application of a new rigorous attitude towards sampling, along with AMOVA and Φ_{ST}, allows researchers to look at differences across populations, while taking into account internal variation in indigenous musical repertoires. In addition to providing a novel application of methodology, Chapter 2 also quantifies the proportion of diversity found within versus between populations. Internal diversity has been a thorn in the side of ethnomusicologists criticizing Lomax's 'modal profile' approach (Feld, 1984; Henry, 1976), and we have bolstered their qualitative analyses by showing that the vast majority of variance in our data (~98%) is found within populations. This preliminary analysis, of course, works only with a small number of closely related populations, and a broader sample would likely increase the proportion of diversity found between populations, but the overall pattern is likely to stay the same.

Given the appropriate data, this methodology could also be applied to other aspects of culture, such as variation in the beliefs or behaviours of individuals, both within and between cultures. For example, a recent analysis of the 'tightness' and 'looseness' of various cultures in terms of social norms and tolerance of deviant behaviour used information from a questionnaire given to a sample of individuals in each country (Gelfand et al., 2011). Unfortunately, these researchers lacked any measure of variation in responses within each country and instead reported only an overall mean 'tightness' score. The AMOVA approach could easily be applied in this sort of situation to investigate how variability in scores is apportioned within and between countries. Additionally, one could easily imagine this being applied to other aspects of material culture, such as textiles or ceramics.

Building on the important methodological contributions presented in Chapter 2, Chapter 3 broaches the topic of whether information contained in musical diversity can be useful to studying human migration. Central to this is the issue of time depth, and we provide some useful (albeit preliminary) evidence indicating that at least the structural features of vocal music reflect population relationships in a similar way to the hyper-variable segment of the maternally inherited mitochondrial DNA. The fact that we were unable to replicate the same results using the whole mitochondrial genome or with additional data from the Philippines reflects the fact that these correlations use only a small number of populations, and can be quite fragile. It is also unclear how these same musical data were to be compared against loci from the paternally inherited Y-chromosome, or autosomal loci. In order to have a robust finding, broader sampling beyond the scope of the current thesis would definitely be required. Performance features were consistently less correlated than were structural features in the Chapter 3 analysis. This does not necessarily mean that performance features are uninformative, just that they may be reflecting processes with a more shallow time depth. Performance features may therefore be useful to investigating recent contact situations or admixture.

The work presented in this thesis hopefully breathes some new life into the dormant field of comparative musicology. Rather than establishing a status quo that most expert ethnomusicologists find unappealing, as Lomax did (Feld, 1984; Henry, 1976), we have established a novel quantitative methodology that addresses some (but not all) of the major criticisms of the Cantometrics project. Unfortunately, our approach will never appease those who oppose comparison due to an ideological stance against applying western musical concepts to indigenous music (McLeod, 1974), nor will it placate those who dispute the use of quantification generally in anthropology (Linnekin, 1987). It should be stressed that our aim was not to oppose mainstream ethnomusicology, but to bring a new approach into the fold. The methods presented here should not be applied haphazardly without expert ethnomusicologist consultation, as our use of archives relies so much on the representativeness of the sample we obtain. Scholars of prehistory can work with open-minded musicologists to test the validity of our results reported here and employ music as an additional tool to complement the current array of available markers. Then, hopefully, music can be used to bring a novel perspective to open questions about recent human prehistory.

REFERENCES

Feld, S. (1984). Sound structure as social structure. *Ethnomusicology*, *28*(3), 383–409.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., et al. (2011). Differences Between Tight and Loose Cultures: A 33-Nation Study. *Science*, *332*(6033), 1100-1104. doi: 10.1126/science.1197754.

Henry, E. (1976). The Variety of Music in a North Indian Village: Reassessing Cantometrics. *Ethnomusicology*, *20*(1), 49-66.

Linnekin, J. (1987). Categorize, Cannibalize? Humanistic Quantification in Anthropological Research. *American Anthropologist*, *89*(4), 920-926. doi: 10.1525/aa.1987.89.4.02a00110.

Lomax, A. (1968). *Folk song style and culture* (p. 363). New Brunswick, N.J. Transaction Books.

McLeod, N. (1974). Ethnomusicological Research and Anthropology. *Annual Review of Anthropology*, *3*(1), 99-115. doi: 10.1146/annurev.an.03.100174.000531.