DIGITALCOMMONS
—@WAYNESTATE—

Wayne State University

Human Biology Open Access Pre-Prints

WSU Press

An Admixture Approach to Trihybrid Ancestry Variation in the Philippines with Implications for Forensic Anthropology

Bridget F. B. Algee-Hewitt
Center for Comparative Studies in Race & Ethnicity, Stanford University, Stanford, California, bridgeta@stanford.edu

Cris E. Hughes
Department of Anthropology, University of Illinois at Urbana–Champaign, Urbana, Illinois

Matthew C. Go
Department of Anthropology, University of Illinois at Urbana–Champaign, Urbana, Illinois

Beatrix Dudzik
DeBusk College of Osteopathic Medicine, Lincoln Memorial University, Harrogate, Tennessee, Beatrix.Dudzik@lmunet.edu

Recommended Citation

An Admixture Approach to Trihybrid Ancestry Variation in the Philippines with Implications for Forensic Anthropology

Bridget F. B. Algee-Hewitt,[1]* Cris E. Hughes,[2] Matthew C. Go,[2,3] and Beatrix Dudzik[4]

[1]Center for Comparative Studies in Race & Ethnicity, Stanford University, Stanford, California, USA.

[2]Department of Anthropology, University of Illinois at Urbana–Champaign, Urbana, Illinois, USA.

[3]SNA International, supporting the Defense POW/MIA Accounting Agency, Central Identification Laboratory, Joint Base Pearl Harbor-Hickam, Hawaii, USA.

[4]DeBusk College of Osteopathic Medicine, Lincoln Memorial University, Harrogate, Tennessee, USA.

*Correspondence to: Bridget F. B. Algee-Hewitt, PhD, Center for Comparative Studies in Race & Ethnicity, Stanford University, 450 Serra Mall, Building 360, Main Quad, Stanford, CA 94305 USA. E-mail: bridgeta@stanford.edu.

Short Title: Forensic Estimation of Ancestry in the Philippines

KEY WORDS: CRANIOMETRICS, FILIPINOS, POPULATION STRUCTURE, AFFINITY ANALYSIS, BIOLOGICAL PROFILE ESTIMATION

Abstract

In this study, we investigate, for the first time from a forensic anthropological perspective, the question of mixed ancestry estimation for modern Filipinos with geographic origins in the Philippines. We derive estimates of continental ancestry using craniometrics from four sources: a new documented collection of current forensic significance from the Manila North Cemetery; the Howells cranial series representing a sample of unclaimed individuals from Manila but said largely to originate from more remote areas, with dates of death before 1940; the Hanihara sample aggregated from various locations and time periods across the Philippines; and the Hanihara series capturing various local indigenous, ethnic groups that are together identified as Philippine Negrito. Parental craniometrics are selected from the Howells dataset and more recently collected samples from Europe and Asia. Using unsupervised clustering, we investigate the algorithmically defined three-cluster, or trihybrid admixture, model to infer continental ancestry for each individual, reporting their relative proportions of Asian, European, and African admixture. We employ similar clustering procedures to identify more complex models, with a larger number of clusters, to explore patterns of affinity between our four Philippine samples and the recently acquired samples from Vietnam, Thailand, China (Hong Kong), Japan, and Korea. These analyses give insight into the relationships between both macro and micro geographic regions, such that, at the country level, we reveal how different population dynamics – whether geo-political, -economic, -historical and/or -social – structure the ancestral makeup of Asian peoples, especially in the degree of European and African admixture. From these ancestry estimates, we find that population of origin explains 38-51% of the variation in each ancestry component and we detect significant differences among the Asian samples in their quantities of ancestry. Filipinos appear considerably admixed, as they appear to carry almost 20% less Asian

ancestry than the average quantity (90%) estimated for the other Asian groups. We also reveal substructure within our representation of modern Filipinos, such that differences in the patterns of three-way admixture exist between each of the four Philippine samples, finding that the Manila cemetery sample has the highest level of Asian ancestry and, as we might expect, that the Negrito sample has the greatest quantity of African ancestry. We perform additional analyses that introduce craniometrics from the Howells Australo-Melanesian series in order to more fully investigate their relationship to the Asian samples and to better understand the African contributions common to the Philippine Negritos especially, as well as the other Southeast Asians and the Spanish and Portuguese groups. By mapping the cluster patterns on a global scale, these analyses reveal, with craniometrics just as with genetic loci, patterns of affinity that are informative of the complex history of Southeast Asia, as they are suggestive of the vestiges of migration, trade, and colonialism, as well as more recent periods of isolation, marginalization, and occupation.

In this paper, we respond to the concerns for improved methodological and data sourcing approaches for analyses in biological anthropology that broadly motivate this special issue, "Thinking Computationally about Forensics." The goals of the present study are, therefore, twofold. We seek, first, to expand the way ancestry is typically conceptualized and/or commonly treated statistically in forensic anthropology, and, second, to expand our focus on underrepresented peoples, for whom patterns of cranial variation and standards for the probabilistic estimation of ancestry have not yet been established (Algee-Hewitt 2016; Algee-Hewitt et al. 2018a; Algee-Hewitt 2017a; Go 2018; Go et al. 2017b; Konigsberg and Frankenberg 2018). We subject craniofacial shape data, generated from a suite of traditional craniometrics, to an unsupervised model-bound clustering, in accordance with prior analyses of other craniometric (Algee-Hewitt 2016; Algee-Hewitt et al. 2018a; Algee-Hewitt 2017a; Algee-Hewitt 2017b) and genetic datasets (Alexander et al. 2009; Algee-Hewitt et al. 2016; Pemberton et al. 2013; Pritchard et al. 2000; Rosenberg 2011) in order to present new information on the ancestry composition and history of population interactions for contemporary Filipinos, who represent a largely understudied population of mixed ancestry (Delfin et al. 2011; Go et al. 2017a).

The unique population history of the Philippine archipelago suggests unique pattern(s) of admixture for modern Filipinos. These complex ancestry signatures may be structured by factors like geography, time, socio-economic status, and ethnic identity (Algee-Hewitt et al. 2018a; Algee-Hewitt 2017a; Algee-Hewitt 2017b; Go et al. 2017a; Hughes et al. 2017; Hughes et al. 2018). Accordingly, such within-population variation in ancestry estimates may be observable across differently ascertained skeletal samples. Owing to this diversity on various scales – from the level of the Filipino population down to specific aggregates – contemporary peoples from the

Philippines likely represent exceptional subsets of the variation within Asia and, therefore, warrant special consideration. Yet, the ancestry composition of contemporary Filipinos has received remarkably little attention in the bioanthropology literature and has not been addressed from an admixture perspective, with a concern for its implications for human identification in forensic anthropology (Go 2018; Go et al. 2017a). We argue that addressing ancestry in the forensic context for Filipinos is of immediate social justice concern in the Philippines and should be of paramount interest to forensic anthropologists, as the geo-political and socio-economic environment, past and present, of this country necessitates ancestry research that enables improved human identification practice not only for casework in this region but also in the United States and Canada, where many Filipinos reside (Citizenship and Immigration Canada 2015; Go 2018; Go et al. 2017; Lopez et al. 2017; Passel and Cohn 2016; Statistics Canada 2017).

The model-based clustering approach used in this study is especially appropriate for elucidating ancestry variation in the Philippines under an admixture model, as it allows for the detection of latent structure, the estimation of continental ancestry proportions, and the inference of population affinity at the level of the individual or group. As this procedure permits shared membership across multiple clusters, it generates multiple ancestries for admixed individuals and provides quantitative measures of the degree of similarity, or biological affinity, among different peoples and, when aggregated, among geographically, temporally, or sampling-based groups (Algee-Hewitt et al. 2018a; Algee-Hewitt 2017a; Algee-Hewitt 2017b). This approach is well-suited to advancing forensic casework methodology, as, in providing such probabilistic estimates of mixed membership, it responds to the recent demand for strengthening the rigor of the forensic sciences through improved statistics – a concern that should extend to forensic

anthropological approaches to cranial ancestry estimation (Algee-Hewitt et al. 2018b; Steadman 2018).

As a region for the study of mixed ancestry, the Philippines is particularly interesting. According to the 2015 census records, more than half of the total population resides in Luzon, the largest and among the most northern islands in the archipelago, which is composed of eight regions and includes the two adjacent and most populous cities, Quezon City and, the capital, Manila, that together support almost 5 million people (Philippine Statistics Authority 2017). Thus, we anticipate considerable biological and social diversity in its inhabitants, attributable to such factors as the early peopling of the archipelago as well as more recent periods of Western colonialism, trade and migration, and socio-economic stratification (Delfin et al. 2011; Go 2018; Go et al. 2017b; Griffin 1996; Hanihara 1989; Hanihara 1990; Headland 1984; Headland 1989; Headland and Early 1998; Kutschera and Pelayo III 2012; Kutschera et al. 2012; Kutschera et al. 2015; Larruga et al. 2017; Lipson et al. 2014; Matsumoto et al. 1979; Molnar 2017; Omoto et al. 1978; Omoto et al. 1981; Padilla Jr 2013; Padilla 2000; Peng et al. 2010; Phelan 2011; Philippine Statistics Authority 2017; Reid 1994; Reid 2013; Skoglund et al. 2016; Tenasas and Ramas 1974; Winkelmann 2017). Negrito ethnic groups are found across Southeast Asia, including Thailand and Malaysia, and may represent the earliest peoples in the region. The Negrito of the Philippines, including the Agta, Aeta, Ati, Batak, and Mamanwa, are, however, the most geographically dispersed (Padilla 2013). The Philippines experienced colonial rule, first, under Spain (1521-1898) and, then, the United States (1898-1946), and it was also linked to the Americas through trade, specifically to Mexico via the Manila-Acapulco Galleon route (1565-1815). Evidence of gene flow under Spanish rule is given by historical records that report intermarriages between Filipinos, Latin Americans, and the Chinese (Phelan 2011). Similarly,

under United States colonialism, intermarriages are documented between Filipinos and European Americans, with a sex bias likely in favor of foreign males (Molnar 2017; Winkelmann 2017). Similar trends are identified in even very recent times, owing to the U.S. military presence and its legacy. Over 50,000 admixed, Filipino "Amerasian" children were born to transient military personnel and contractors at the time of the American troop withdrawal in 1992 (Kutschera et al. 2012). That estimate grows to 250,000 when subsequent offspring are considered, and it does not take into account the number of children resulting from the new foreign presence in the Philippines that began in 2001 as America assumed its leading role in the global war on terrorism (Kutschera et al. 2012; Kutschera et al. 2015). Further, legacy systems of prostitution and sex tourism are present in post-military installation areas (Kutschera et al. 2012; Kutschera et al. 2015). Together, these cultural conditions have shaped and perpetuate an admixed and structurally vulnerable subgroup of socially and economically marginalized peoples within the larger Filipino population (Kutschera and Pelayo 2012).

Pursuing this untapped line of research is important from a human variation perspective just as it is timely for forensic practice in the United States and in the Philippines, as we expect the composition of forensic casework to reflect current demographic trends (Algee-Hewitt 2017a; Go 2018). Despite the significant representation of Asians, including Filipinos, in both global and regional communities, there is a clear underrepresentation of these populations in forensic literature and limited availability of formalized guidelines or resources for implementing standard case analysis methods. Only recently, for example, have we seen a comprehensive, quantitative study of the misclassification of Asians using the gold standard method for ancestry estimation via craniometrics (Go et al. 2017a). However, forensically relevant samples of Filipinos are not yet among the Asian populations included in the Fordisc 3.1 software used by

forensic anthropologists in the United States and abroad (Ousley and Jantz 2012). Asia contributes almost 60% of the global world population, and the Philippines ranks thirteenth worldwide and second in Southeast Asia (Worldometers.info 2018a; Worldometers.info 2018b). In the United States, demographic records indicate that the Asian American population increased by 72% in the 2000-2015 period and future projections identify Asians, originating from the Far East, Southeast Asia, and Indian subcontinent, as the fastest growing group, comprising 38% of all U.S. immigrants in the next 50 years (Lopez et al. 2017). For the last decade, the Philippines has ranked third among Asian countries as a source for undocumented immigrants in the United States (Passel and Cohn 2016). Presently, Filipinos account for 19% of all Asian Americans, making them the third largest group of Asian origin in the United States (Lopez et al. 2017). Filipinos have an even greater presence in Canada, where they represent the largest immigrant "permanent resident" group and the fourth largest visible minority (Citizenship and Immigration Canada 2015; Statistics Canada 2017). They also had the greatest growth rate in the 2006-2016 period, nearly doubling their numbers in the last 10 years (Statistics Canada 2017).

While the individuals who comprise a collection of forensic cases are a biased sampling of the population at large, prior work on the major U.S. populations has shown that the ancestral composition of the Forensic Anthropology Data Bank – an unprecedented source of morphometric and demographic data from actual forensic anthropology cases and documented skeletal collections – approximates current census demographics (Algee-Hewitt 2016; Algee-Hewitt 2017a; Algee-Hewitt 2017b). Further, recent research has shown how craniometrically-derived admixture estimates and measures of affinity computed for the dead, including admixed individuals, are concordant with those ancestry proportions generated from genotypic data sourced from the living (Algee-Hewitt 2016; Algee-Hewitt et al. 2018a; Algee-Hewitt 2017a).

The agreement in these patterns suggests for the present context that an increase in Filipinos in the population at large will lead to a concomitant increase in their representation among forensic cases and, as reliable ancestry estimates and measures of affinity for admixed populations are achievable from craniometrics, they are equally likely to be attainable for Filipinos.

Coming to an understanding of ancestry for modern Filipinos is a pressing forensic research concern, as computational standards are immediately needed for the estimation of the biological profile parameters including ancestry, for this, and any such emergent, population in the Americas and abroad. This study represents a first step towards improved theory and practice in ancestry estimation for Filipinos as we investigate the following research questions to elucidate patterns of admixture and affinity for samples from the Philippines and with respect to other global geographic, especially Asian, populations.

1) Using the methods of unsupervised clustering, can we produce a model that corresponds to the trihybrid ancestry patterns expected for modern Filipinos, whose ancestral history represents a mixture of Asian, African, and European (Iberian, specifically) parental contributors and reflects a lengthy history of Western contact and colonial rule?

2) Can we find population structure, defined here as variation in trihybrid ancestry estimates, among the Asian samples, such that inferred proportions of ancestry are informative of an individual's population of membership?

3) Can we detect substructure in the Philippines, by identifying specifically ancestry variation within the full Filipino sample, which represents an aggregate of peoples from four different skeletal collections that are known to differ in sampling time and place and whose individuals reflect different local origins, ethnic identities, and socio-economic statuses?

4) Using models with many clusters, can we reveal patterns of affinity among the four different samples from the Philippines, the parental reference samples from Asia, Africa, and Europe and, in select analyses, samples representing Australo-Melanesia/Micronesia?

5) Can we demonstrate that shifts in admixture affect craniofacial shape by identifying for the Philippine samples which of this project's inter-landmark distance measurements change in their dimensions under increased or decreased quantities of Asian, European, or African ancestry?

6) Finally, we ask what implications, if any, do these analyses and their results have on the forensic anthropological evaluation of ancestry for modern Filipino cases?

This study responds to the stated lacuna in bioanthropological and forensic research on the Philippines by using this investigative framework to explore ancestry variation in this region, and to identify best practice recommendations for forensic anthropological casework for persons with likely origins in the Philippines and/or of Filipino identity.

Materials and Methods

Skeletal Samples. For this study, we analyze a mixed-sex sample of 1,301 individuals. Tables 1a and 1b give the sample sizes per population by geographic region. As we sought to capture the widest possible range of diversity in the Philippines, given present data availability, and explore the effects of sample ascertainment with tests of substructure, we include craniometrics for four different samples: (1) a newly assembled skeletal collection from Manila North Cemetery, representing contemporary Filipinos of forensic significance and low socio-economic status, selected as available from low-cost niche burials (Go et al. 2017), (2) the Philippine series from the Howells (1989) worldwide craniometric dataset, which is said to

include crania from "convicts who died in prison in Manila" but who likely originated from "remoter parts" and which have "a combination of features, or lack of indigenous features, suggesting a possible European component", (3) the Philippine series from the Hanihara craniometric dataset (pers. comm. 2008), which is an aggregate of diverse ethnic/linguistic groups, from various geographic locations, representing different burial periods, and (4) the Negrito series, also from the Hanihara craniometric dataset, composed of indigenous peoples identified elsewhere as Aeta from west-central Luzon (Hanihara 1989; Hanihara 1990; Hanihara 1992a; Hanihara 1992b).

We also include craniometrics for parental reference samples from Africa, Europe, and Asia in the cluster analyses to capture the variation for the most likely sources of continental ancestry. The African samples – the Dogon, Teita, and Zulu – are sourced from the Howells (1989) dataset. The European sample is limited to the Iberian Peninsula, specifically published craniometric data from the Oloriz and Wamba collections in Spain and a Portuguese sample in Lisbon (Humphries and Ross 2011). The Asian sample is drawn from a newly reported dataset already used in discussions of the forensic determination of ancestry (Dudzik and Jantz 2016).

In secondary analyses, we studied an additional 355 individuals from the Australo-Melanesian/Micronesian series in the Howells (1989) dataset, using different configurations of the Australia (n=101), Tasmania (n=87), Guam (n=57) and Tolai (n=87) samples. These Southwestern Pacific series were incorporated in large-K cluster analyses to better reveal the affinities between the Southeast Asian populations.

Craniometric Data Selection and Treatment.   We merged the craniometric data available for our four Philippine and parental population samples and performed both case and variable-wise deletion to select only those inter-landmark distance measurements (ILDs) with complete

data. We selected the following suite of 12 standard ILDs: maximum cranial length (GOL), cranial base length (BNL), cranial vault height (BBH), maximum cranial breadth (XCB), biauricular breadth (AUB), nasal height (NLH), nasal breadth (NLB), mastoid height (MDH), orbital height (OBH), frontal chord (FRC), parietal chord (PAC), and occipital chord (OCC) (Moore-Jansen et al. 1994). We prescreened our final dataset for unrealistic measurements but not for atypical cases because model-based clustering is a robust tool for outlier detection (Evans et al. 2015; Tao and Pi 2009; Yamanishi et al. 2004). Prior to cluster analysis, we converted the ILDs to Mosimann shape variables by geometric mean transformation to account for the issue of size differences (Darroch and Mosimann 1985).

Theory and Implementation of the Multivariate Normal Mixture. We implemented the unsupervised model-based clustering methods of finite mixture analysis (Fraley and Raftery 2002; McLachlan and Basford 1988; McLachlan and Peel 2000) to discover latent population structure, reveal patterns of population affinity, and estimate proportions of ancestry. For the multivariate normal mixture, K is the number of components in the model, each component corresponds to a k-cluster, and the estimated membership coefficient, τ_k, represents the probability that an individual belongs to the kth component.

As the value of K is not prespecified, we identified the best clustering solution among all of the fitted models by applying a mixture decomposition and allocation routine that assigns the sampled individuals into k clusters. We let the value of K increase to a maximum of 15 components just as we test ten different parameterizations for the component covariance matrix, $\sum_k$ (Fraley and Raftery 2007; Fraley et al. 2012). We use the Bayesian Information Criterion (BIC) to identify statistically the optimal value of K and the parameterization of $\sum_k$. Accordingly, we identified nine major clusters with unconstrained covariance matrices. Because

the BIC selects the number of mixture components only to provide a good approximation to the density rather than the number of clusters and sometimes, when a multivariate mixture-model is used for Gaussian clustering, a non-normal cluster may be better fit by a mixture of multiple normal distributions (Baudry et al. 2010; Biernacki et al. 2000). Therefore, we performed a secondary analysis on this $K = 9$ solution, to determine if the value of K can be alternatively resolved, such that there are fewer clusters. We successively merged components according to an entropy criterion, fitted a piecewise linear regression to the rescaled entropy plot, and then selected the optimal number of clusters from the estimated breakpoint in the plot (Baudry et al. 2010). As displayed in Figure 2, this analysis yields a reduced component model of $K = 4$, with three major clusters and one spurious cluster containing a single outlier. We validated the reduced three major cluster model by rerunning the cluster analysis, with and without the outlier, specifying the value of $K < 9$. We obtain a comparable $K = 3$ solution.

We performed this unsupervised clustering without incorporating prior knowledge of the individuals' true population affiliation, geographic origin, or sampling location. We reassociated these "anonymous" cases with their respective identifiers after the optimal clustering solution was obtained in order to interpret the clustering results and to permit the downstream analysis of the membership coefficients by standard statistical methods.

Estimating Ancestry and Revealing Population Affinities. As the mixture analysis allows for overlapping clusters, each individual holds some proportion of membership in all of the k clusters. We characterized every individual, therefore, by a vector of posterior probabilities $(\tau_{i1}, \dots, \tau_{in})$ that must sum to 1 for the specified value of K. Following Algee-Hewitt (2016), we (1) determine population affinities by the degree to which individuals affiliated with the same population have similar membership coefficients, τ_k, across the specified number of k-clusters

and (2) we use these same membership coefficients, when K corresponds to continental groupings, to give the relative proportions of ancestry. These membership values can be represented as percent estimates, whether, by row, to characterize the individual, or, as a matrix, to generalize the patterns for one sample or the multi-sample dataset.

We found, by the post-hoc review of individual parental reference assignment, that the clusters of the reduced $K = 3$ model correspond to the three continental ancestry groups of African, Asian, and European. The clusters of the full $K = 9$ model represent subsets of these continental groupings, thereby revealing latent structure within the parental samples. We used, therefore, the reduced $K = 3$ model to estimate trihybrid ancestry. As the goal of this study is to elucidate ancestry variation, the $K = 3$ analysis is the main focus of this paper. However, we also used the large $K = 9$ model to investigate population affinities and performed an additional analysis that incorporated samples from the Australo-Melanesian and Micronesian series in the Howells (1989) dataset, for which we obtained a $K = 11$ model with unconstrained covariance matrices.

For the three-cluster model, we produced the population-specific means for each of the ancestry components by partitioning the data by its population identifier and averaging the individual estimates across the samples with respect to each cluster. Normality was assessed with the Shapiro-Wilk test and each of the ancestry components were treated as needed by $\log_{10}$-transformation. Structure plots were used to display the cluster results, for both the reduced and large K models, as proportions of ancestry or patterns of affinity at various, i.e., individual, sample, population and continental, levels.

Testing Population Differences in Ancestry. We used Analysis of Variance (ANOVA) to test for structured differences in the calculated admixture estimates for the $K = 3$ solution.

Means were evaluated separately for each of the three admixture components. *Post-hoc* evaluation with the Tukey-Kramer test was performed to assess mean differences between each of the sampled populations. We assessed within-sample variation in proportions of ancestry for significant differences using one-sided *t*-tests, adjusting the alpha level as appropriate for multiple comparisons.

Identifying the Morphological Response to Admixture. We performed Spearman rank correlation analysis to evaluate the magnitude and direction of the association between the 12 ILDs that make up this paper's dataset of craniometrics and the membership coefficients by ancestry component. The statistical significance of all resulting correlations was determined by testing if the observed value of the correlation coefficient, ρ, differs from 0 at the specified threshold, $a = 0.05$. To visualize these patterns, we plotted the correlation coefficients from the analysis of the ancestry estimates and ILDs for each of the Philippine samples.

Software. The *R* statistical computing environment was used for all statistical analysis. *Python* scripts were written for data formatting.

Results

Estimating Ancestry. For the trihybrid ancestry analysis, we calculated mean ancestry estimates for each sampled population. These results are presented in Tables 1a and 1b.

To visualize these estimates, we produced a structure plot that displays mixed ancestry at the level of the individual for the four Philippine samples. Figure 3 shows for each individual their percentage of membership across the Asian, European, and African ancestry components.

Testing Population Differences in Ancestry. To determine if the samples differ in their relative proportions of ancestry, we performed a one-way ANOVA on each of the three vectors

of posterior probabilities of cluster membership. As Table 2 shows, these tests are significant, with population of origin explaining on average 43% of the variation in ancestry. Assessments with Tukey-Kramer tests further identify population-specific differences among the mean values calculated for each admixture component. Here, we have placed emphasis on the comparisons between the Asian populations and the Philippine samples; the European and African samples are left pooled as continental groups. The results of these *post-hoc* analyses are displayed graphically in Tables 3a-c.

Revealing Population Affinities. We investigated larger component models to discover latent substructure within the continental groups and elucidate affinities between the [micro]regional levels, e.g. within Asia, between and within the two Iberian parental populations, or among the four Philippine samples. We defined, via *BIC,* a nine major cluster solution as the optimal model. From this clustering, we identified visually the patterns of relationships among the African, Iberian, and Asian, including the four Philippine, populations. We displayed, using population-level structure plots, the mean estimates of membership in each of the clusters, after eliminating from consideration any trivial mean estimates (<5%). In Figure 4, we demonstrate how the cluster membership patterns map onto world-wide geography to allow for the detection of spatial relationships. To confirm substructure, we performed one-sided *t*-tests to assess the perceived differences in quantities of European and African Ancestry among the Iberian samples. We find significant differences ($a = 0.025$) in mean proportions between the 16th Century Wamba sample from Spain and both the 20th Century Oloriz sample from Spain and the 20th Century Lisbon sample from Portugal, such that individuals from the Wamba sample carry on average 12% less European and 17% more African ancestry.

To better explain the African component observed in the Philippine samples and elucidate affinities between the Southeast Asian populations, we added Australo-Melanesian samples. We performed an identical cluster analysis and produced a model with 11 major clusters. In this case, the African component previously observed in the Southeast Asian samples is split into three clusters. An African cluster is retained while two new clusters are produced, which capture, individually, the Australo-Melanesian component in the Asian samples and variation unique to the Southeast Asian groups. In Figure 5, we display this new cluster solution with a population-level structure plot.

Identifying the Morphological Response to Admixture. We used Spearman rank correlation analysis to evaluate the magnitude and direction of the relationship between the ILDs that make up the craniometric dataset and the ancestry estimates inferred from the model-based clustering. In Figure 6, the significant correlation coefficients, those for which $\rho \neq 0$ at $a = 0.05$, are plotted for each of the Philippine samples by ancestry component, displaying how aspects of the craniofacial morphology change as the quantity of a given ancestry increases or decreases.

Discussion

For this project, we applied unsupervised model-based clustering to detect population structure and infer proportions of trihybrid – Asian, African, and European – ancestry from cranial shape data for four different samples from the Philippines, including a modern skeletal collection of forensic significance from the Manila North Cemetery, the Howells and Hanihara Philippine series, and the Hanihara Philippine Negrito series. To produce mixed ancestry estimates, we included in our analysis craniometrics for parental reference samples representing most likely sources of continental ancestry. To investigate differences in ancestral makeup by population, we

aggregated the individual-level ancestry estimates by population identifier (Tables 1a and b) and we compared the mean estimates among our samples. We found that population of origin explains between 38% and 51% of the variation in each ancestry component (Table 2) and we detected significant differences in quantities of ancestry for some samples from the Philippines and for the rest of Asia (Tables 3a-c). Filipinos appear considerably admixed with respect to the other Asian population samples, carrying, on average, less Asian ancestry (71%) than our Korean (99%), Japanese (96%), Thai (93%), and Vietnamese (84%) reference samples. We also revealed substructure in our Filipino sample, showing that the patterns of ancestry vary within the Philippines – that is, between the four differently sourced Filipino samples. Mean estimates of Asian (76%) and European (7%) ancestry are greatest for the cemetery sample of forensic significance from Manila. The Hanihara and Howells samples have equal Asian ancestry (71%) and greater African ancestry (24%) than the Manila cemetery sample (17%). The Negrito sample has equal Asian and African (47%) ancestry.

Using models with more clusters, we identified patterns of relationships among the Filipino samples relative to the other geographic groups. The Manila cemetery sample tends to align with low-admixture Asian groups, even holding membership in an Asian cluster that is not represented among the other Philippine samples. The Howells and Hanihara series show greater affinity with the European and African groups, displaying memberships in two such clusters that are not shared with the Manila cemetery sample. These findings are quite interesting for the Howells series, as it is noted how some individuals have combined "Tagalog and Hispanized" names while others are labeled as "Moro" (Howells 1989). The Hanihara sample appears the most cosmopolitan, with coverage across all but one of the continental clusters, perhaps owing to a more geographically diverse sampling strategy. Only the Negritos have membership in all three

of the African clusters. We also discovered similar kinds of substructure in some of the parental reference samples, noting in particular variation in African ancestry between the two Spanish samples (Wamba and Oloriz) and good agreement in ancestry proportions between the contemporary Spanish (Oloriz) and Portuguese (Lisbon) samples.

We repeated our analyses after including Australo-Melanesian data, which changed the allocation patterns in two key ways: two new clusters emerged, each capturing Australo-Melanesian and Southeast Asian variation, and the most common African component in the Asian populations was reduced such that only the Negrito, Vietnamese, and Hanihara Philippine samples retained a non-trivial proportion of membership (5%-15%) in this cluster. Overall, these trends reflect different levels of admixture in this region and variation in the source populations for the Southeast Asian samples.

With respect to forensic casework, these discoveries attest to the importance of taking into consideration the unique population histories for peoples of Asian origin. For the Philippines in particular, the differences in the magnitude and patterning of admixture among the four samples observed here indicate that Filipino reference samples should be evaluated critically: at present, the Manila cemetery likely represents the most appropriate "forensic" reference sample given the number of available individuals (n=127), the recent dates of death, and the comparability in the marginalized, low socio-economic status of this cemetery's burials and those structurally vulnerable persons who typically comprise the population of forensic cases in this region (Go 2018; Go et al. 2017).

Population Structure. Overall, the trends discovered from our trihybrid continental ancestry analysis are remarkably concordant with our expectations for admixture in not only modern Filipinos but also more broadly in our sample of Southeast Asian populations. Results reflect the

history of population interactions within Asia, revealing the vestiges of European colonialism, and even the extent of European presence, as well as the substantial impact of early peopling from Africa. We find that for the mean Asian ancestry contributions, only the Hanihara Negrito sample exhibit significantly less Asian ancestry than the other Asian reference samples (Korean, Thai, Vietnamese, Chinese, Japanese), while the remaining Philippine samples are generally comparable (Table 3a). However, the agreement among Philippine and other Asian samples deviates when comparing the African and European contributions. The majority of the Asian reference samples have significantly lower mean African contributions than the Philippine samples (Table 3b). The differences observed for this Negrito sample may be a result of genetic contributions from indigenous Australians (Delfin et al. 2011). The large African component observed for the Chinese and the non-Negrito samples may reflect shared genetic contributions from the Taiwanese peoples related to the Austronesian Expansion (Lipson et al. 2014; Peng et al. 2010; Reich et al. 2011; Skoglund et al. 2016)

While European contributions are also typically greater in the Philippine samples (with the exception of the Negrito sample), the mean estimates do not significantly differ from Chinese or Vietnamese, as seen in Table 3. Furthermore, the Howells Philippine sample does not differ from the Japanese, leaving only the Korean (1%) and Thai (<1%) as samples with significantly lower mean European contributions. Thailand was never colonized by Europeans but was a target of Western missionary work and surrounded by Malaysian and Myanmar neighbors when under British and French rule (Chaiwan 1984). Thailand is also bounded by Indochina in addition to Myanmar and Malaysia. The low percentage of European admixture observed in Korean samples agrees with Korea's historically low admixture percentages, with over 95 % of the Korean population exhibiting no admixture with other ethnic groups (Korean Statistical

Information Service 2016). Such homogeneity is said to stem from cultural attitudes towards, and the social stigma associated with, mixed-race offspring (Yoo 2017). However, in the last decade, admixture between Korean males and females from surrounding Asian countries such as China and Vietnam has become more common (Fackler 2009). Only our sample from China, which is comprised of individuals from Hong Kong, displays a greater quantity of European admixture (15%) than the Philippines, reflecting, perhaps, the effect of a century of British Crown rule over this region. Most notably, our mixture clustering results capture the Philippines' unique and extended history of colonialism and genetic exchange with the West, including the special case of introgression via trade-driven contact with Latin America – population dynamics already observed biologically by the reported rates of misclassification of Filipinos as groups from Latin America using traditional craniometrics and a set of global reference samples (Go et al. 2017a). In short, the continental ancestry patterns that are distinctive for the Philippine samples, when compared to all other Asian samples, include increased African and European ancestry.

Population Substructure. Filipino population substructure, defined as ancestral differences among the four Philippine samples, was also detected. The most significant example of divergence in ancestry estimates is seen for the Hanihara Negrito sample, which is not surprising given what is known about the ethnic identities and potential source populations for Philippine Negritos (Delfin et al. 2011; Hanihara 1989; Hanihara 1990; Hanihara 1992a; Hanihara 1992b; Headland 1984; Jinam et al. 2012; Lipson et al. 2014; Padilla Jr 2013; Padilla 2000; Reid 2013). Here, the Negritos display the greatest mean proportions of African ancestry of all the Philippine samples, though only significantly greater than the Manila Philippine study sample (Table 3b). The remaining two Philippine samples bridge this gap, having African proportions that do not

significantly differ from the Negrito or the Manila sample, likely a result of the broad geographic and temporal origins documented for the Hanihara Philippine sample, and the stated "general" representation of cranial variation by the Howells sample, with individuals carrying Hispanic and Moorish names (Howells 1989). While the Manila sample shows the highest quantities of European and Asian ancestry, these contributions do not significantly differ for any of the Philippine samples. The mean Asian ancestry proportion of the Negrito sample is significantly lower than all other Asian study samples, including the three additional Philippine samples. These patterns may be understood in light of arguments for migration events and more recent gene flow between indigenous Australians and Filipino Negritos (Aghakhanian et al. 2015; Delfin et al. 2011; Jinam et al. 2012; Wang and Li 2013).

When examining the more complex analysis, with many clusters, that includes the Australo-Melanesian series from the Howells dataset, we observed further Filipino population substructure, again with respect to the Philippine Negrito sample. All four Philippine samples have contributions from both the African and the Australo-Melanesian clusters, but the proportions of these two clusters differ. Specifically, the Manila, Hanihara and Howells Philippine samples all have larger contributions from the Australo-Melanesian than the African cluster, yet the Negrito sample exhibits almost five times greater contributions from the African cluster than the Australo-Melanesian cluster. Such a pattern suggests that not only does the Negrito sample have a generally greater amount of African ancestry (as demonstrated in the trihybrid analysis above), but its source of African ancestry substantially differs from the remaining three Philippine samples. Interpretations for these genetic patterns range from the Negrito populations exhibiting a separate migratory group into East and Southeast Asia that was distinct from other migration events out of Africa, as well as gene flow during and following the

initial peopling of Australia (Aghakhanian et al. 2015; Jinam et al. 2012; Wang and Li 2013).

When comparing the Philippine samples to the other Southeast Asian samples included in this study, we see that, in general, the Southeast Asian samples conform to the patterns exhibited by the majority of the Philippine samples such that the Australo-Melansian contributions are greater than African cluster contributions. Only the Vietnamese sample has an inflated contribution from the African cluster, similar to the Negrito sample. These findings indicate that there is great diversity in the Philippines that reflects complex patterns of gene flow and secular change across the archipelago. Modern Filipinos represent, therefore, a unique component of the variation within Asia – a point of considerable importance for forensic research and applied casework (Go 2018; Go et al. 2017a; Go et al. 2017b).

Interestingly, some genetic studies found no differences between Negrito and non-Negrito in the Philippines and between Filipinos and other Austronesian-speaking groups (Abdulla et al. 2009). While other studies agree with our current findings (Delfin et al. 2011): that the nature of the variation observed suggests that an additional source of African ancestry is present in Filipinos – found here most disproportionately in the Hanihara Negrito sample. While our analyses suggest that the Philippine samples are more similar to each other than to the other Asian populations, there appears here to be substantial enough substructure to caution us against any oversimplification of diversity by treating all "Filipinos" or peoples originating from the Philippine archipelago as representing a single, homogenous population. For forensic anthropological purposes, therefore, ancestry estimation methods using cranial data should use at the very least two populations to represent the ancestral variation in the Philippines.

It is worthwhile to note that we also discovered substructure in the Spanish reference samples when investigating the larger values of K. Significantly more African and less European

ancestry is identified for the 16th Century Wamba Ossuary relative to the more recent 20th Oloriz sample. In fact, the Oloriz collection better aligns in its proportions with the sample of 20th Century individuals from Portugal. Our results are, therefore, in good agreement with previous suggestions that the morphological characteristics observed for the Wamba sample reflect Moorish admixture and that the expulsion of the Moors between the 15th and 17th centuries may have obliterated any evidence of African admixture in later Iberian samples (Fox et al. 1996; Humphries and Ross 2011). These findings underscore the importance of selecting parental reference samples with caution, understanding that the unique population histories of peoples or subsets within a "single geographic" population sample, if unaccounted for, may influence the estimates of ancestry for the target individuals.

Morphology and Ancestry. To understand the morphological effect of differences in the magnitude and direction of admixture on modern Filipinos, we used the trihybrid ancestry proportions to test for significant correlations between percent ancestry and inter-landmark distances. In doing so, we revealed how increased or decreased quantities of Asian, African, and European ancestry produce change in craniofacial shape. Increased African ancestry is characterized by an elongated cranium and greater prognathism for the non-Negrito samples and an enlarged nasal breadth for all samples. Decreased African ancestry is defined by a reduction to nasal height and cranial breath for the non-Negrito samples. Correlations for the Negrito and Hanihara samples suggest a decrease in mastoid height and the Manila sample displays a unique reduction in frontal cord with decreased African proportions. High correlations with European ancestry are observed for cranial length for the non-Negrito samples as well as frontal cord for the Hanihara sample. Mastoid height is also increased for all samples but the Hanihara, while the Negrito sample alone presents a significant increase in nasal height. Decreased European

ancestry is associated with a reduction in nasal width for all groups, in cranial breadth for the non-Negritos, and in parietal cord and cranial height for the Howells sample alone. Finally, greater quantities of Asian ancestry are related to enlarged cranial breadth dimensions for all non-Negritos samples, increases to nasal and mastoid heights for the Hanihara and Negrito samples, and increased cranial height for the Howells sample. For all samples, decreased Asian ancestry is associated with a shortening of the cranium; further, we see a reduction in parietal chord and basion-nasion length for the Manila and Hanihara samples. Lastly, nasal breadth appears significantly reduced for the Howells and Negrito samples. These changes generally agree with the morphological features that tend to differentiate between biogeographic groups when performing ancestry classification, especially dimensions of the cranial vault, which have been shown to be the least plastic and, therefore, the most accurate at predicting population affinity (Carson 2006; Holló et al. 2010; Hubbe et al. 2009; Martínez-Abadías et al. 2009).

Forensic Implications. The African and European admixture trends for the Philippine samples have significant implications in forensic anthropological analysis of ancestry. In a recent study, Go et al. (2017a) examined the group classifications of the Manila sample in the population affinity program Fordisc 3.1 (Ousley and Jantz 2012). The authors found, for example, that the most common Fordisc reference groups to which the male Filipino cases classified by cranial form were Vietnamese (36%) and Chinese (24%): two groups which do not significantly differ from the Philippine samples in mean African ancestry estimates, respectively, in the present study. They also reported how, when averaged across all analyses, classification into the American Black reference group was twice as common as classification into the American White group – again, consistent with the important role that African ancestry proportions appear to play in the study of our Filipino samples. Furthermore, Dudzik and Jantz's

(2016) analysis of Fordisc classification trends for the Chinese, Japanese, Thai, Korean and Vietnamese samples (also used in the present study as the "Asian" parental reference samples) can be similarly compared to the Fordisc classifications of the Manila sample by Go et al. (2017a). The Fordisc classification trends are a forensic testament to how the ancestry patterns found in the present study play out in the casework setting using conventional ancestry estimation tools. That is, the increased African, and to a lesser extent European, ancestry contributions will influence the patterns of population affinity for potential forensic cases, such that the Philippine samples are more likely to deviate from an Asian population classification more frequently than other Asian samples. Accordingly, this information should be of great value for interpreting misclassifications using traditional methods for single-ancestry estimation.

Owing to the uniqueness of the Philippines, in both its history and its peoples, making determinations of ancestry in this context poses a significant challenge for forensic anthropology. We argue that this is true for both the estimation of ancestry at the level of the individual and the identification of ancestry signatures among local/regional or ethnic groups in the archipelago. It is further complicated by whether the persons of interest are multi-generational Filipinos, representing a subset of Asians in Asia, or peoples of Filipino descent, who trace their biogeographic origins to the Philippines but reside or were born in North America or elsewhere and who may reflect additional or other admixture events and exposure to different environments. To this point, understanding the nature of this variation is critical for casework in the United States and Canada.

Our results suggest that forensic anthropologists, who rely upon the conventional metric approaches to the craniofacial determination of ancestry, which implement single-ancestry, or hard classification, models, would benefit from adopting a mixture, specifically trihybrid,

estimation methodology. An admixture approach can offer the forensic anthropologist several computational and practical advantages. First, it does not require access to a predefined reference sample, whose craniometric diversity must capture enough of the range of morphological variation in modern Filipinos to produce classification with low error and the desired posterior probabilities and typicalities (Algee-Hewitt 2016; Ousley and Jantz 2012). Second, it allows the analyst to take into account the complex population and genetic history of Southeast Asia, and specifically, the Philippines, and delivers quantitative measures of the relative contributions of ancestral sources for modern Filipinos, which are grounded in evolutionary and population-specific history (Abdulla et al. 2009; Delfin et al. 2011; Lipson et al. 2014; Peng et al. 2010; Stoneking and Delfin 2010) and the robust theory of probability statistics (Algee-Hewitt 2016; McLachlan and Peel 2000). Lastly, it permits a more satisfying comparison between craniometrically-derived estimates of admixture, as trihybrid ancestry, and patterns of affinities and the genetic marker-derived estimates that are generated using similar mathematical methods (Alexander et al. 2009; Algee-Hewitt et al. 2018a; Pritchard et al. 2000).

We demonstrate, here, the advantages of such an integrative approach for this and future work. When historical, biological (both neutral phenotypic and genetic), geographic, and sociodemographic factors are considered in tandem, we can speak more holistically to the forensic case as an individual or to a population, as a group of once living people. This approach also helps us to define the pathway forward for the new area of computational anthropology (Algee-Hewitt 2016; Algee-Hewitt and Goldberg 2016; Algee-Hewitt et al. 2018b; Steadman 2018).

Literature Cited

Abdulla, M. A., I. Ahmed, A. Assawamakin et al. 2009. Mapping human genetic diversity in Asia. *Science* 326:1,541–1,545.

Aghakhanian, F., Y. Yunus, R. Naidu et al. 2015. Unravelling the genetic history of Negritos and indigenous populations of Southeast Asia. *Genome Biol. Evol.* 7:1,206–1,215.

Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19:1,655–1,664.

Algee-Hewitt, B. F. B. 2016. Population inference from contemporary American craniometrics. *Am. J. Phys. Anthropol.* 160:604–624.

Algee-Hewitt, B. F. B. 2017a. Geographic substructure in craniometric estimates of admixture for contemporary American populations. *Am. J. Phys. Anthropol.* 164:260–280.

Algee-Hewitt, B. F. B. 2017b. Temporal trends in craniometric estimates of admixture for a modern American sample. *Am. J. Phys. Anthropol.* 163:729–740.

Algee-Hewitt, B. F. B., M. D. Edge, J. Kim et al. 2016. Individual identifiability predicts population identifiability in forensic microsatellite markers. *Curr. Biol.* 26:935–942.

Algee-Hewitt, B. F. B., and A. Goldberg. 2016. Better together: Thinking anthropologically about genetics. *Am. J. Phys. Anthropol.* 160:557–560.

Algee-Hewitt, B. F. B., C. E. Hughes, and B. E. Anderson. 2018a. Temporal, geographic and identification trends in craniometric estimates of ancestry for persons of Latin American origin. *Forensic Anthropol.* 1:4–17.

Algee-Hewitt, B. F. B., J. Kim, and C. E. Hughes. 2018b. Thinking computationally about forensics: Anthropological perspectives on advancements in technologies, data, and algorithms. *Hum. Biol.* 90:5–10 (in press).

Baudry, J.-P., A. E. Raftery, G. Celeux et al. 2010. Combining mixture components for clustering. *J. Comput. Graph. Stat.* 19:332–353.

Biernacki, C., G. Celeux, and G. Govaert. 2000. Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Trans. Pattern Anal. Mach. Intell.* 22:719–725.

Carson, E. A. 2006. Maximum likelihood estimation of human craniometric heritabilities. *Am. J. Phys. Anthropol.* 131:169–180.

Chaiwan, S. 1984. A study of Christian mission in Thailand. *East Asia J. Theol.* 2:62–74.

Citizenship and Immigration Canada. 2015. *Facts and Figures 2014: Immigration Overview—Permanent Residents*. http://www.cic.gc.ca/english/pdf/2014-Facts-Permanent.pdf (Accessed September 24, 2018).

Darroch, J. N., and J. E. Mosimann. 1985. Canonical and principal components of shape. *Biometrika* 72:241–252.

Delfin, F., J. M. Salvador, G. C. Calacal et al. 2011. The Y-chromosome landscape of the Philippines: Extensive heterogeneity and varying genetic affinities of Negrito and non-Negrito groups. *Eur. J. Hum. Genet.* 19:224–230.

Dudzik, B., and R. L. Jantz. 2016. Misclassifications of Hispanics using Fordisc 3.1: Comparing cranial morphology in Asian and Hispanic populations. *J. Forensic Sci.* 61:1,311–1,318.

Early, J. D., and T. N. Headland. 1998. *Population Dynamics of a Philippine Rain Forest People: The San Ildefonso Agta*. Gainesville, FL: University Press of Florida.

Evans, K., T. Love, and S. Thurston. 2015. Outlier identification in model-based cluster analysis. *J. Classif.* 32:63–84

Fackler, M. 2009 November 28. Baby boom of mixed children tests South Korea. *The New York Times*. https://www.nytimes.com/2009/11/29/world/asia/29babies.html (Accessed September 24, 2018).

Fox, C. L., A. G. Martín, and S. V. Civit. 1996. Cranial variation in the Iberian Peninsula and the Balearic Islands: Inferences about the history of the population. *Am. J. Phys. Anthropol.* 99:413–428.

Fraley, C., and A. E. Raftery. 2002. Model-based clustering, discriminant analysis and density estimation. *J. Am. Stat. Assoc.* 97:611–631.

Fraley, C., and A. E. Raftery. 2007. Model-based methods of classification: Using the mclust software in chemometrics. *J. Stat. Softw.* 18:1–13.

Fraley, C., A. E. Raftery, T. B. Murphy et al. 2012. *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation*. Technical Report No. 597. Seattle, WA: Department of Statistics, University of Washington.

Go, M. C. 2018. Appraising forensic anthropology in the Philippines: Current status and future directions. *Forensic Sci. Int.* 288:329.e1–329.e9.

Go, M. C., A. R. Jones, B. F. B. Algee-Hewitt et al. 2018. Classification trends among contemporary Filipino crania using Fordisc 3.1. *Forensic Anthropol*. (in press).

Go, M. C., A. B. Lee, J. A. D. Santos et al. 2017. A newly assembled human skeletal reference collection of modern and identified Filipinos. *Forensic Sci. Int.* 271:128.e1–128.e5.

Griffin, M. B. 1996. The cultural identity of foragers and the Agta of Palanan, Isabela, the Philippines. *Anthropos* 91:111–123.

Hanihara, T. 1989. Affinities of the Philippine Negritos as viewed from dental characters: A preliminary report. *J. Anthropol. Soc. Nippon* 97:327–339.

Hanihara, T. 1990. Affinities of the Philippine Negritos with Japanese and the Pacific

populations based on dental measurements: The basic populations in East Asia, I. *J. Anthropol. Soc. Nippon* 98:13–27.

Hanihara, T. 1992a. Dental and cranial affinities among populations of East Asia and the pacific: The basic populations in East Asia, IV. *Am. J. Phys. Anthropol.* 88:163–182.

Hanihara, T. 1992b. Negritos, Australian Aborigines, and the "proto-sundadont" dental pattern: The basic populations in East Asia, V. *Am. J. Phys. Anthropol.* 88:183–196.

Headland, T. N. 1984. Agta Negritos of the Philippines. *Cult. Surv. Q.* 8:29–31.

Headland, T. N. 1989. Population decline in a Philippine Negrito hunter-gatherer society. *Am. J. Hum. Biol.* 1:59–72.

Holló, G., L. Szathmáry, A. Marcsik et al. 2010. Linear measurements of the neurocranium are better indicators of population differences than those of the facial skeleton: Comparative study of 1,961 skulls. *Hum. Biol.* 82:29–46.

Howells, W. W. 1989. *Skull Shapes and the Map: Craniometric Analyses in the Dispersion of Modern Homo*. Cambridge, MA: Harvard University Press.

Hubbe, M., T. Hanihara, and K. Harvati. 2009. Climate signatures in the morphological differentiation of worldwide modern human populations. *Anat. Rec.* 292:1,720–1,733.

Hughes, C. E., B. F. B. Algee-Hewitt, R. Reineke et al. 2017. Temporal patterns of Mexican migrant genetic ancestry: Implications for identification. *Am. Anthropol.* 119:193–208.

Hughes, C. E., B. Dudzik, B. F. B. Algee-Hewitt et al. 2018. Understanding (mis)classification trends of Latin Americans in Fordisc 3.1: Incorporating cranial morphology, microgeographic origin, and admixture proportions for interpretation. *J. Forensic Sci.* (in press).

Humphries, A. L., and A. H. Ross. 2011. Craniofacial sexual dimorphism in two Portuguese
    skeletal samples. *Anthropologie* 49:13–20.

Jinam, T. A., L. C. Hong, M. E. Phipps et al. 2012. Evolutionary history of continental Southeast
    Asians: "Early train" hypothesis based on genetic analysis of mitochondrial and
    autosomal DNA data. *Mol. Biol. Evol.* 29:3,513–3,527.

Konigsberg, L. W., and S. R. Frankenberg. 2018. Of typicality and predictive distributions in
    discriminant function analysis. *Hum. Biol.* 90:31–44 (in press).

Korean Statistical Information Service. 2016. *Population by Census*. http://kosis.kr/eng/
    (Accessed September 24, 2018).

Kutschera, P. C., and J. M. G. Pelayo III. 2012. The Amerasian paradox. Online Conference on
    Multidisciplinary Social Sciences — March 29–31, 2012. Australian International
    Cultural & Educational Institute.

Kutschera P. C., J. M. G. Pelayo III, and M. G. Talamera-Sandico. 2012. The continuing
    conundrum of Southeast Asia's 50,000 Filipino military 'Amerasians'. *Int. Proc. Econ.
    Dev. Res.* 48:62–67.

Kutschera, P. C., E. C. Tesoro, M. G. Talamera-Sandico et al. 2015. Modern day second
    generation military Filipino Amerasians and ghosts of the U.S. military prostitution
    system in west central Luzon's 'AMO Amerasian Triangle'. *Int. J. Soc. Behav. Educ.
    Econ. Bus. Ind. Eng.* 9:3,356–3,366.

Larruga, J. M., P. Marrero, K. K. Abu-Amero et al. 2017. Carriers of mitochondrial DNA
    macrohaplogroup R colonized Eurasia and Australasia from a Southeast Asia core area.
    BMC Evol. Biol. 17:115.

Lipson, M., P.-R. Loh, N. Patterson et al. 2014. Reconstructing Austronesian population history in Island Southeast Asia. *Nat. Commun.* 5:4689.

López, G., N. G. Ruiz, and E. Patten. 2017 September 8. Key facts about Asian Americans, a diverse and growing population. *Pew Research Center*. http://www.pewresearch.org/fact-tank/2017/09/08/key-facts-about-asian-americans/ (Accessed September 24, 2018).

Martínez-Abadías, N., M. Esparza, T. Sjøvold et al. 2009. Heritability of human cranial dimensions: Comparing the evolvability of different cranial regions. *J. Anat.* 214:19–35.

Matsumoto, H., T. Miyazaki, and K. Omoto. 1979. Population genetic studies of the Philippine Negritos. II Gm and Km allotypes of three population groups. *Am. J. Hum. Genet.* 31:70–76.

McLachlan, G. J., and K. E. Basford. 1988. *Mixture Models: Inference and Applications to Clustering*. New York: Marcel Dekker.

McLachlan, G. J., and D. Peel. 2000. *Finite Mixture Models*. New York: John Wiley & Sons.

Molnar, N. 2017. *American Mestizos, the Philippines, and the Malleability of Race: 1898–1961*. Columbia, MO: The University of Missouri Press.

Moore-Jansen, P. M., S. D. Ousley, and R. L. Jantz. 1994. *Data Collection Procedures for Forensic Skeletal Material*. Knoxville, TN: Department of Anthropology, The University of Tennessee.

Omoto, K., S. Misawa, and S. Harada. 1978. Population genetic studies of the Philippine Negritos. I. A pilot survey of red cell enzyme and serum protein groups. *Am. J. Hum. Genet.* 30:190–201.

Omoto, K., S. Ueda, and K. Goriki. 1981. Population genetic studies of the Philippine Negritos. III. Identification of the carbonic anhydrase-1 variant with CA_1 Guam. *Am. J. Hum. Genet.* 33:105–111.

Ousley, S. D., and R. L. Jantz. 2012. Fordisc 3 and statistical methods for estimating sex and ancestry. In *A Companion to Forensic Anthropology*, D. C. Dirkmaat, ed. Chichester, UK: Blackwell Publishing, 311–329.

Padilla Jr., S. G. 2013. Anthropology and GIS: Temporal and spatial distribution of the Philippine Negrito groups. *Hum. Biol.* 85:209–230.

Padilla, S. G. 2000. Negrito hunters and gatherers in the Philippines today. *Indig. Aff.* 2:52–55.

Passel, J. S., and D. Cohn. 2016. Overall number of U.S. unauthorized immigrants holds steady since 2009. *Pew Research Center*. http://www.pewhispanic.org/2016/09/20/overall-number-of-u-s-unauthorized-immigrants-holds-steady-since-2009/ (Accessed September 24, 2018).

Pemberton, T. J., M. DeGiorgio, and N. A. Rosenberg. 2013. Population structure in a comprehensive genomic data set on human microsatellite variation. *G3 (Bethesda)* 3:891–907.

Peng, M. S., H. H. Quang, K. P. Dang et al. 2010. Tracing the Austronesian footprint in Mainland Southeast Asia: A perspective from mitochondrial DNA. *Mol. Biol. Evol.* 27:2,417–2,430.

Phelan, J. L. 2011. *The Hispanization of the Philippines: Spanish Aims and Filipino Responses, 1565–1700*. Madison, WI: University of Wisconsin Press.

Philippine Statistics Authority. 2017. *Philippine Population Surpassed the 100 Million Mark (Results from the 2015 Census of Population)*. https://psa.gov.ph/content/philippine-

population-surpassed-100-million-mark-results-2015-census-population (Accessed September 24, 2018).

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

Reich, D., N. Patterson, M. Kircher et al. 2011. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am. J. Hum. Genet.* 89:516–528.

Reid, L. A. 1994. Possible non-Austronesian lexical elements in Philippine Negrito languages. *Ocean. Linguist.* 33:37–72.

Reid, L. A. 2013. Who are the Philippine Negritos? Evidence from language. *Hum. Biol.* 85:329–358.

Rosenberg, N. A. 2011. A population-genetic perspective on the similarities and differences among worldwide human populations. *Hum. Biol.* 83:659–684.

Skoglund, P., C. Posth, K. Sirak et al. 2016. Genomic insights into the peopling of the southwest Pacific. *Nature* 538:510–513.

Statistics Canada. 2017. *Immigration and Ethnocultural Diversity: Key Results from the 2016 Census*. https://www150.statcan.gc.ca/n1/daily-quotidien/171025/dq171025b-info-eng.htm (Accessed September 24, 2018).

Steadman, D. W. 2018. Who needs data? I've got experience! *Hum. Biol.* 90:77–82 (in press).

Stoneking, M., and F. Delfin. 2010. The human genetic history of East Asia: Weaving a complex tapestry. *Curr. Biol.* 20:R188–R193.

Tao, Y., and D. Pi. 2009. Unifying density-based clustering and outlier detection. In *Proceedings of the 2009 Second International Workshop on Knowledge Discovery and Data Mining*, Q. Luo and M. Gong, eds. Los Alamitos, CA: IEEE Computer Society Press, 644–647.

Tenasas, R. C. P., and L. L. Ramas. 1974. A map of the better-known cultural minorities of the Philippines. *Philipp. Q. Cult. Soc.* 2:3.

Wang, C.-C., and H. Li. 2013. Inferring human history in East Asia from Y chromosomes. *Investig. Genet.* 4:11.

Winkelmann, T. O. 2017. Rethinking the sexual geography of American Empire in the Philippines. In *Gendering the Trans-Pacific World*, C. C. Choy and J.-T. C. Wu, eds. Leiden, The Netherlands: Koninklijke-Brill NV, 39–76.

Worldometers. 2018a. Asia Population (Live). http://www.worldometers.info/world-population/asia-population/ (Accessed September 24, 2018).

Worldometers. 2018b. Philippines Population (Live). http://www.worldometers.info/world-population/philippines-population/ (Accessed September 24, 2018).

Yamanishi, K., J.-i. Takeuchi, G. Williams et al. 2004. On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms. *Data Min. Knowl. Discov.* 8:275–300.

Yoo, S. 2017. Legal reform related to interracial Koreans. *Wash. U. Global Stud. L. Rev.* 16:365–385.

Tables 1A and 1B. Counts of Individuals and Mean Ancestry Estimates (%) for (A) the Four Samples from the Philippines and (B) the Continental Reference "Parental" Populations

Table 1A.

Population sample	*n*	Ancestry Estimates: mean %		
		Asian	African	European
PHILIPPINES	328	71	23	6
Manila	109	76	17	7
Howells	50	71	24	5
Hanihara	141	71	23	5
Negrito Hanihara	28	47	47	6

Table 1B.

Population Sample	*n*	Ancestry Estimates: mean %		
		Asian	African	European
ASIAN	417	90	6	4
Chinese	69	72	13	15
Japanese	162	96	1	3

Korean	32	99	<1	1
Thai	107	93	7	<1
Vietnamese	47	84	12	4
AFRICAN	283	9	84	7
Dogon	99	13	82	5
Teita	83	6	85	9
Zulu	101	9	85	6
EUROPEAN	273	24	7	70
Portuguese	49	20	8	72
Spanish	224	24	7	70

Table 2. ANOVA Results by Ancestry Component

Ancestry Component	One-Way ANOVA			
	R^2	df	F-Ratio	Prob>F
Asian	0.41	10,1290	87.97	<0.0001
African	0.51	10,1290	135.92	<0.0001
European	0.38	10,1290	79.07	<0.0001

Tables 3A–C. A Graphical Display of the Results of the Turkey-Kramer Post-Hoc Test for Differences in (A) Asian, (B) African and (C) European Ancestry

Comparisons are made between the European and African macrogeographic groups and the Asian populations, including the four different Philippine samples. Populations not connected by the same letter are found to be significantly different in mean ancestry estimates at $a = 0.05$.

Table 3A.

Population Sample	% Asian			
Korean	A	B		
Japanese	A			
Thai	A	B		
Vietnamese	A	B		
Manila Philippines	A	B		
Chinese	A	B		
Hanihara Philippines		B		
Howells Philippines	A	B		
Hanihara Negrito Philippines			C	
EUROPEAN			C	
AFRICAN				D

Table 3B.

Population	% African					
AFRICAN	A					
Hanihara Negrito Philippines	A	B				
Howells Philippines		B	C			
Hanihara Philippines		B	C			
Manila Philippines			C	D		
Chinese			C	D		
EUROPEAN				D		
Vietnamese				D		
Thai					E	
Japanese					E	
Korean						F

Table 3C.

Population Sample	% European					
EUROPEAN	A					
Chinese		B				

Manila Philippines		B	C			
Hanihara Philippines		B	C	D		
Howells Philippines		B	C	D	E	
Vietnamese		B	C	D	E	
AFRICAN				D	E	
Japanese					E	
Hanihara Negrito Philippines			C	D	E	
Thai						F
Korean						F

Figure 1.
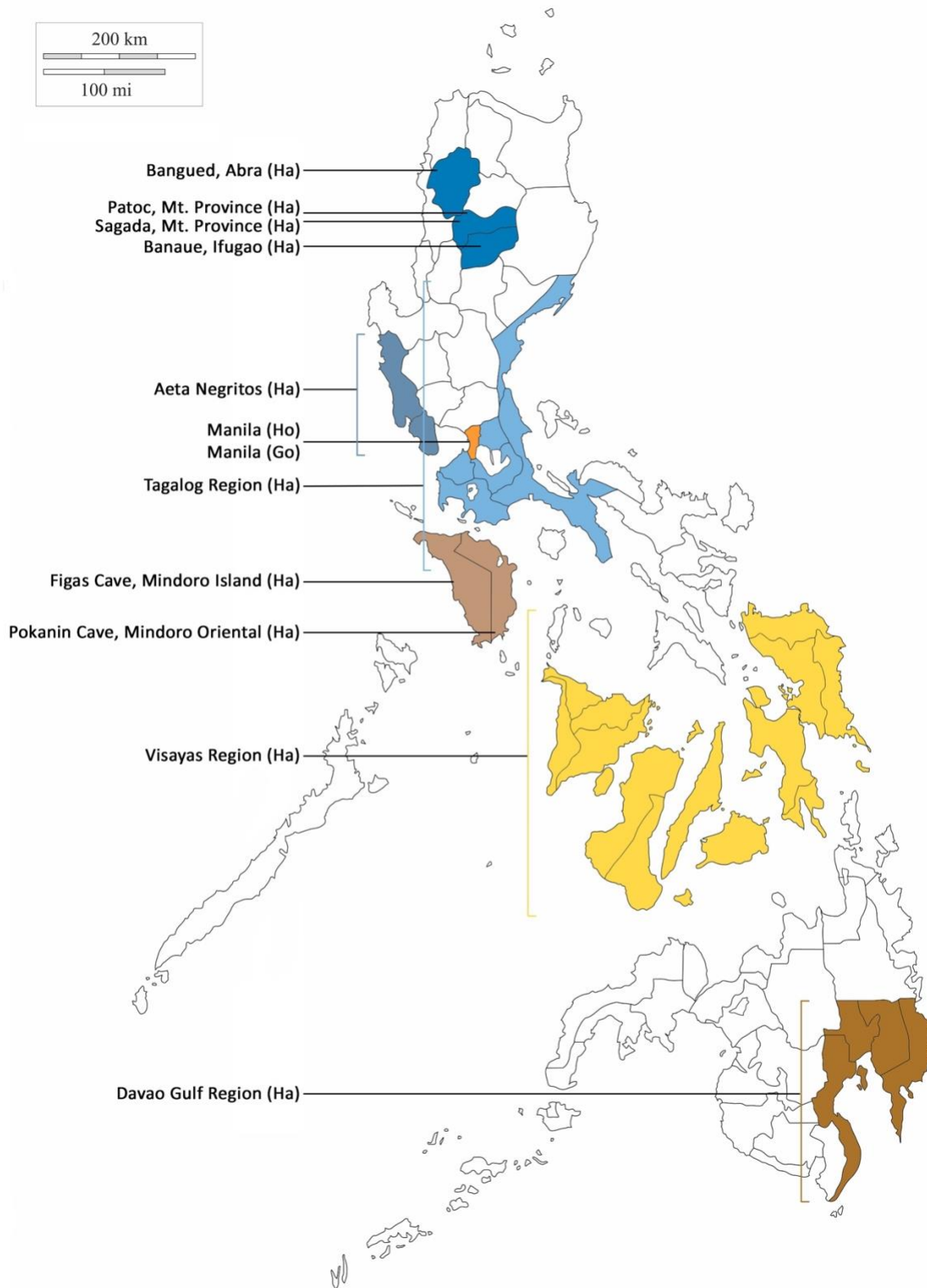
Figure 2.
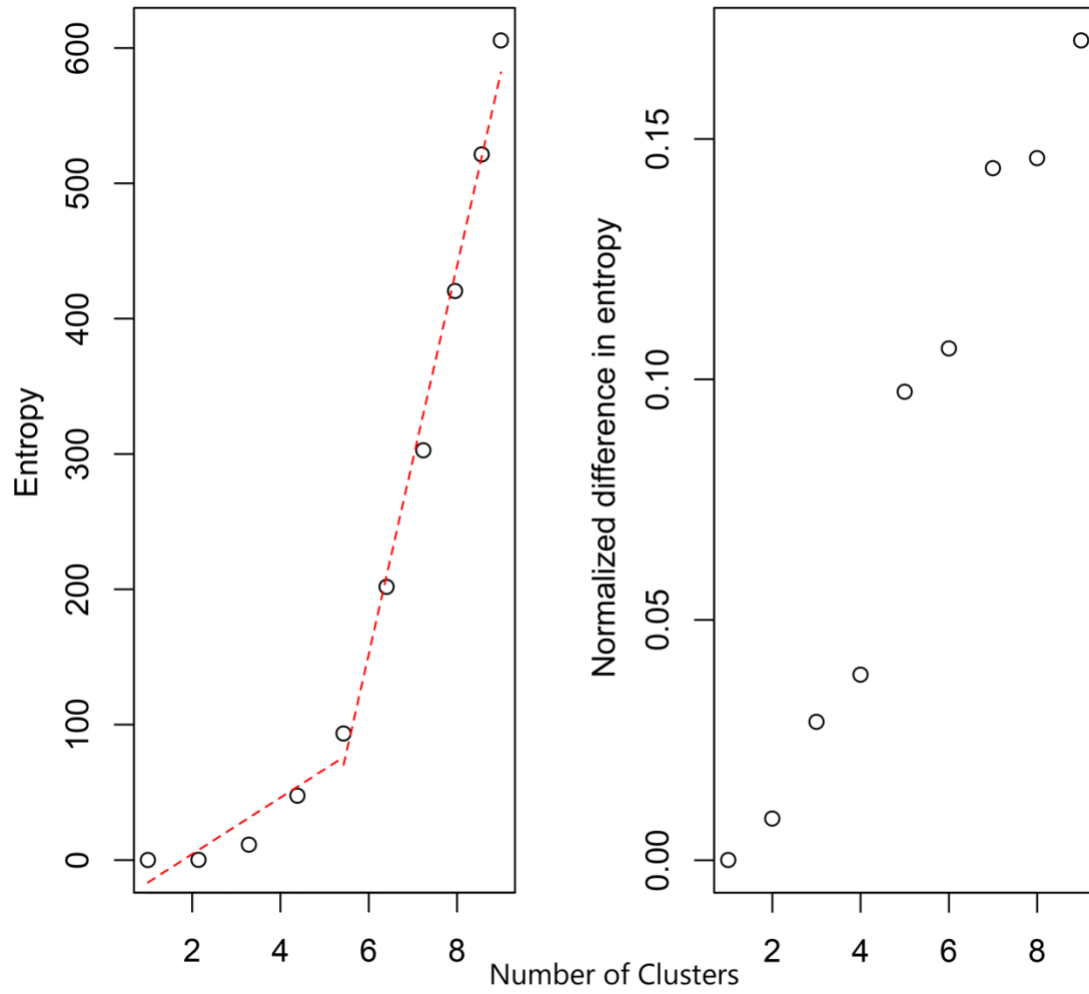
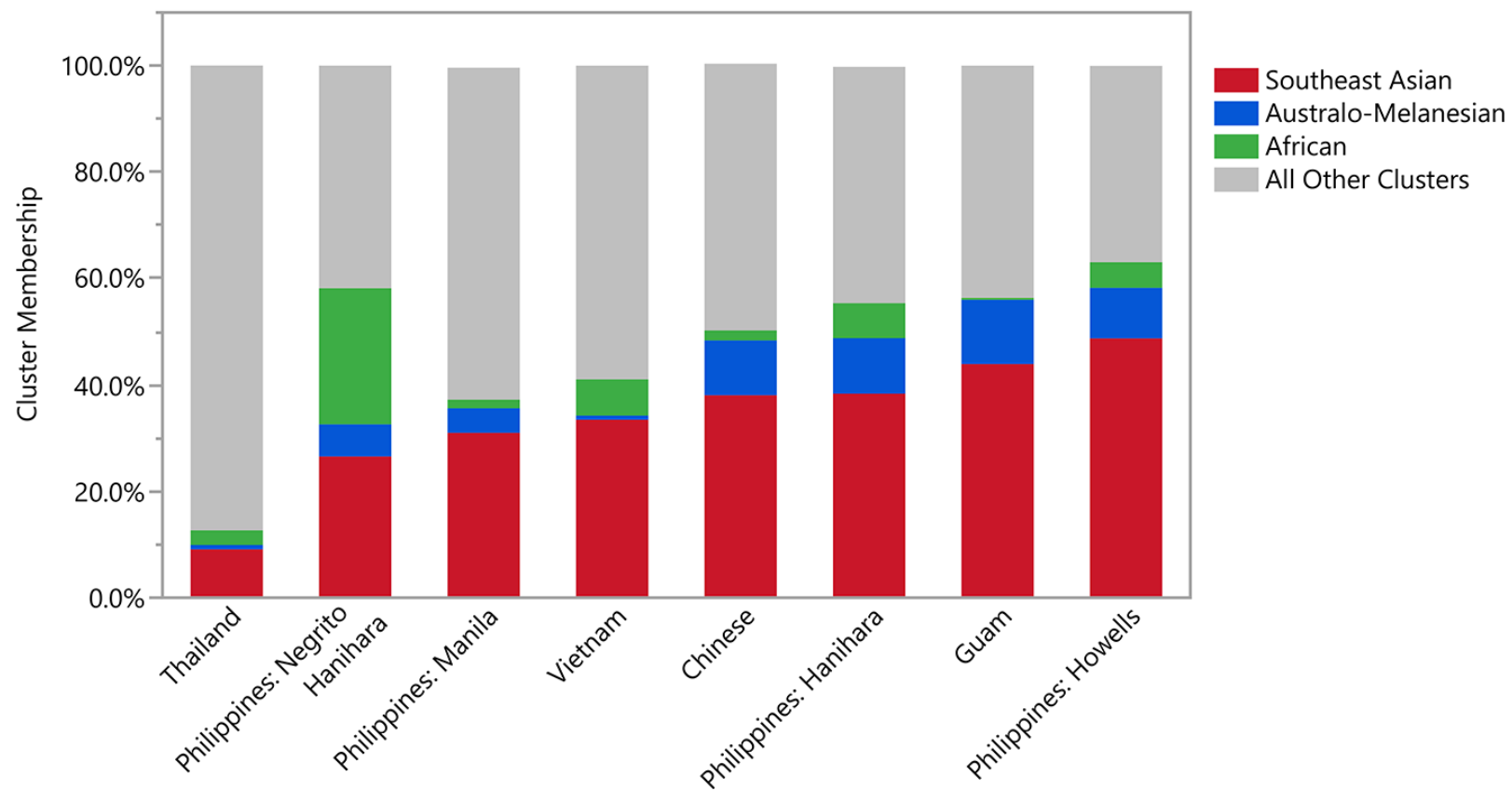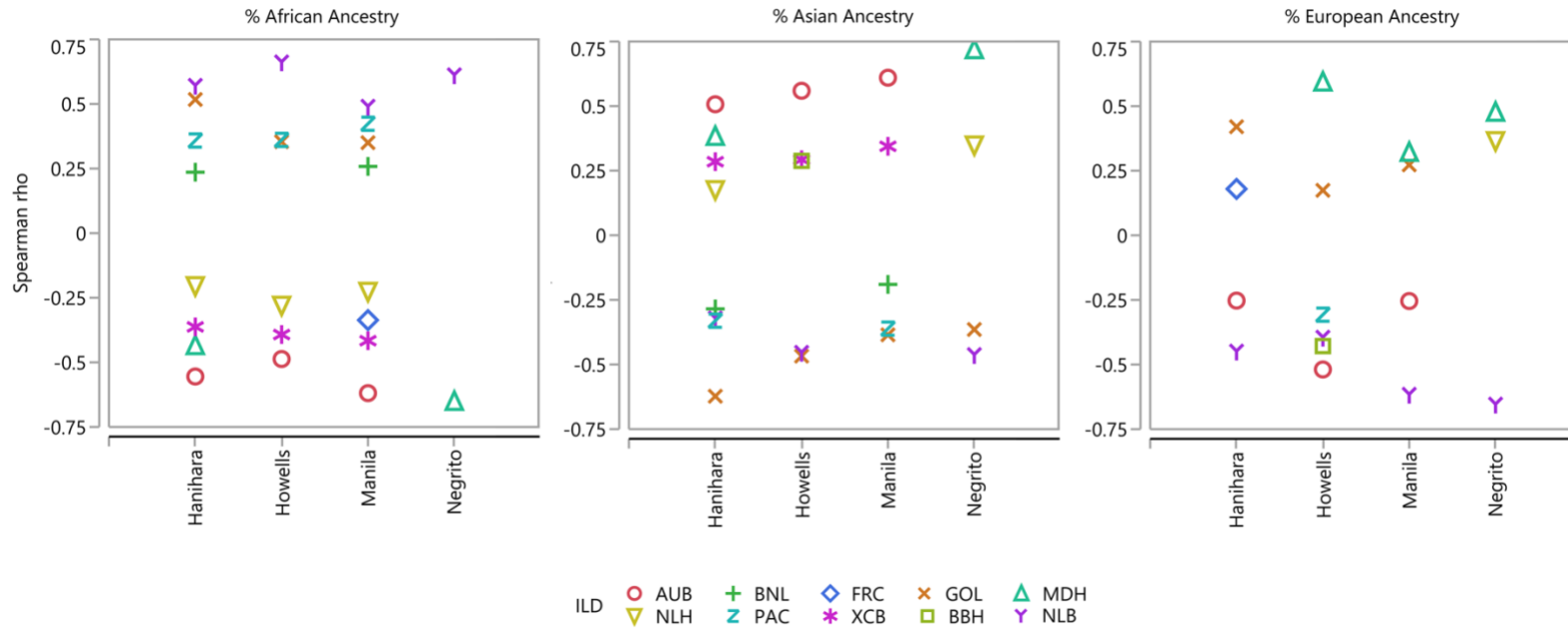**Figure 3.**
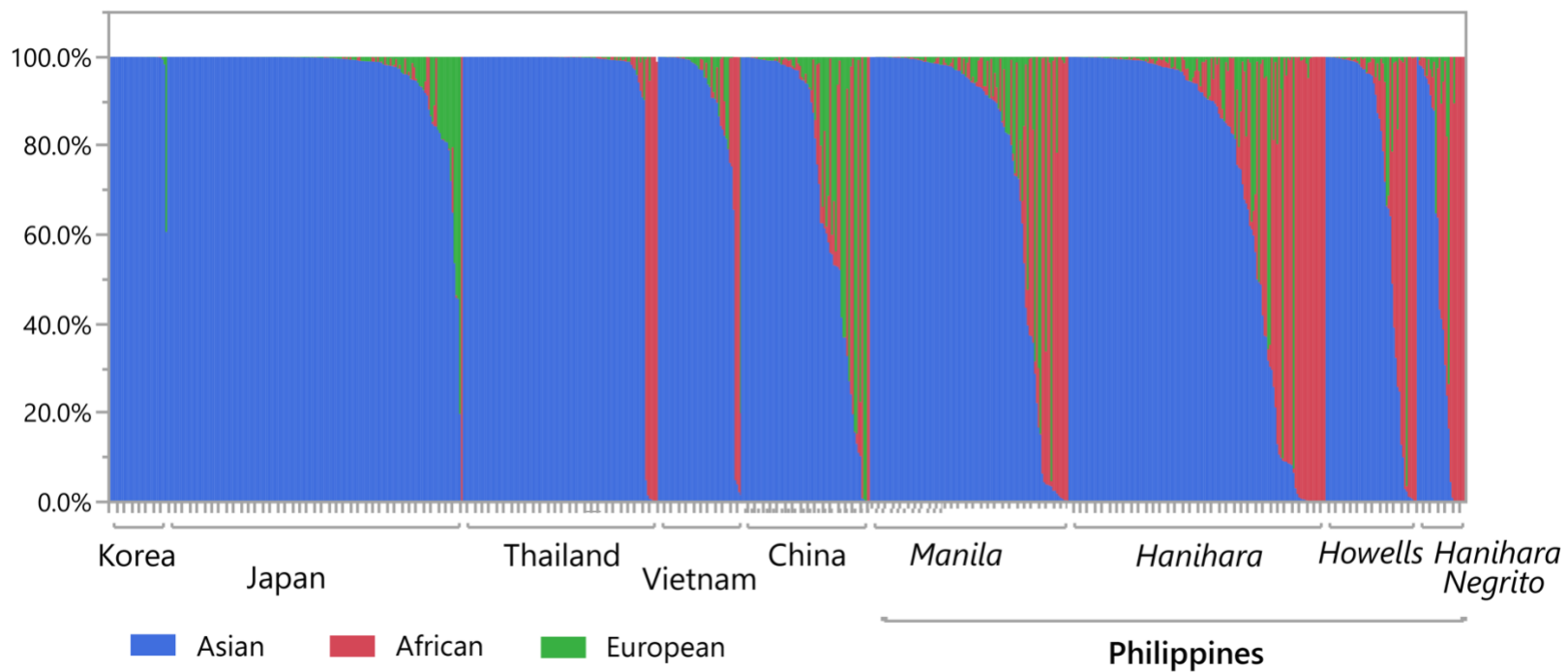
Figure 4.
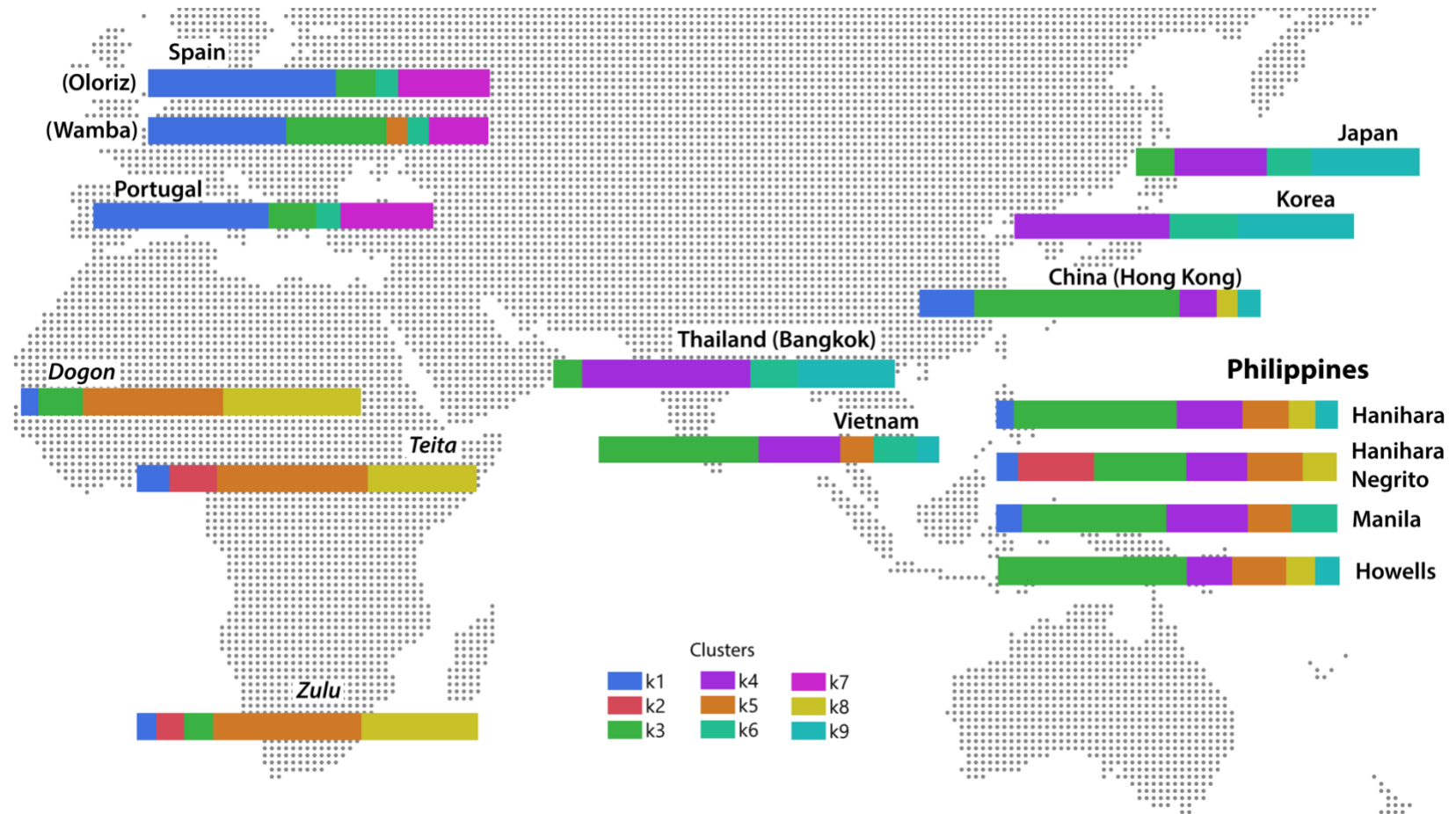
Figure 5.

Figure 6.

Figures Captions

Figure 1. Map of the Philippines showing the geographic distribution of some of the sampling locations for the craniometric data, identified by source according to the labeling system: the Manila North Cemetery (Go) and the samples collected by Howells (Ho) and Hanihara (Ha).

Figure 2. Entropy plots for the clustering of the Philippine and parental (Asian, European, and African) population data. To reduce the large value of K, a piecewise linear regression model is fit to the values in the entropy plot and the number of clusters are selected by the estimated breakpoint in the plot. The dotted regression line shows one elbow and indicates a reduced model with four clusters: three major clusters and one outlier cluster. The optimal solution has, therefore, three clusters, which corresponds directly to a trihybrid continental ancestry model.

Figure 3. Structure plot displaying the relative proportions of ancestry when plotted as percent estimates for the three Asian, European and African components. Each individual is represented by a single line that is partitioned into three differently colored segments that correspond to the ancestry components. The length of the colored line segment represents the estimated quantity of ancestry. The samples, and their individuals, are sorted in descending order on their quantity of Asian ancestry.

Figure 4. For the $K = 9$ model, the population-level (mean) estimates of cluster membership are plotted and the membership patterns are mapped onto geography by population. The length of the colored line segment represents the estimated quantity of membership in the given cluster.

Figure 5. For the $K = 11$ model with Australo-Melanesian samples included, the population-level (mean) estimates of membership are plotted for the one African, the two newly inferred clusters, and all other clusters combined. The length of the colored line segment represents the estimated quantity of membership in the given cluster.

Figure 6. Plots of the significant correlation coefficients for each of the Philippine samples by ancestry component, identifying which ILDs shift in accordance with changes in the proportions of ancestry.