Irish, JD, Morez, A, Girdland Flink, EL, Phillips, E and Scott, GR

Do dental nonmetric traits actually work as proxies for neutral genomic data? Some answers from continental- and global-level analyses

Received: 21 January 2020 | Revised: 25 February 2020 | Accepted: 10 March 2020

DOI: 10.1002/ajpa.24052

American Journal of PHYSICAL ANTHROPOLOGY | WILEY

RESEARCH ARTICLE

Do dental nonmetric traits actually work as proxies for neutral genomic data? Some answers from continental- and global-level analyses

Joel D. Irish[1,2] | Adeline Morez[1] | Linus Girdland Flink[1,3] | Emma L.W. Phillips[1] | G. Richard Scott[4]

[1]School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, UK

[2]Evolutionary Studies Institute and Centre for Excellence in PaleoSciences, University of the Witwatersrand, South Africa

[3]Department of Archaeology, School of Geosciences, University of Aberdeen, Aberdeen, UK

[4]Anthropology Department, University of Nevada Reno, Reno, Nevada

Correspondence
Prof. Joel D. Irish, Research Centre in Evolutionary Anthropology, and Palaeoecology, School of Biological and Environmental Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK.
Email: j.d.irish@ljmu.ac.uk

Funding information
American Museum of Natural History; Arizona State University; National Science Foundation, Grant/Award Numbers: BCS-0840674, BNS-0104731, BNS-9013942

Abstract

Objectives: Crown and root traits, like those in the Arizona State University Dental Anthropology System (ASUDAS), are seemingly useful as genetic proxies. However, recent studies report mixed results concerning their heritability, and ability to assess variation to the level of genomic data. The aim is to test further if such traits can approximate genetic relatedness, among continental and global samples.

Materials and Methods: First, for 12 African populations, Mantel correlations were calculated between mean measure of divergence (MMD) distances from up to 36 ASUDAS traits, and F_{ST} distances from >350,000 single nucleotide polymorphisms (SNPs) among matched dental and genetic samples. Second, among 32 global samples, MMD and F_{ST} distances were again compared. Correlations were also calculated between them and inter-sample geographic distances to further evaluate correspondence.

Results: A close ASUDAS/SNP association, based on MMD and F_{ST} correlations, is evident, with r_m-values between .72 globally and .84 in Africa. The same is true concerning their association with geographic distances, from .68 for a 36-trait African MMD to .77 for F_{ST} globally; one exception is F_{ST} and African geographic distances, r_m = 0.49. Partial MMD/F_{ST} correlations controlling for geographic distances are strong for Africa (.78) and moderate globally (.4).

Discussion: Relative to prior studies, MMD/F_{ST} correlations imply greater dental and genetic correspondence; for studies allowing direct comparison, the present correlations are markedly stronger. The implication is that ASUDAS traits are reliable proxies for genetic data—a positive conclusion, meaning they can be used with or instead of genomic markers when the latter are unavailable.

KEYWORDS

Africa, Arizona State University Dental Anthropology System, geographic distance, population affinities, single nucleotide polymorphisms

1 | INTRODUCTION

Nonmetric traits of the human permanent dentition, as in those from the Arizona State University Dental Anthropology System (ASUDAS), have a significant genetic component in expression. At least this was suggested in earlier research (e.g., Sofaer, Niswander, MacLean, & Workman, 1972; Scott, 1973; Brewer-Carias, Le Blanc, & Neel, 1976; Scott, Yap Potter, Noss, Dahlberg, & Dahlberg, 1983; Turner II, 1985a; Sofaer, Smith, & Kaye, 1986; Scott & Turner II, 1988, 1997; Turner II, Nichol, & Scott, 1991; Pinkerton, Townsend, Richards, Schwerdt, & Dempsey, 1999; Rightmire, 1999; although see Harris, 1977). Therefore, phenetic affinities based on these traits approximate genetic relatedness, at levels of comparison ranging from local to global. At least this was assumed in previous work (Irish, 1997; Scott et al., 1983; Scott & Turner II, 1988, 1997; Turner II, 1985a; Turner II et al., 1991). These and other attributes, such as their ease of recording, observer replicability—especially if dichotomized (below), evolutionarily conservative nature, lack of sexual dimorphism, and a low likelihood of selection in expression (Scott & Irish, 2017; Scott & Turner II, 1997; Scott, Turner II, Townsend, & Martinón-Torres, 2018; Turner II et al., 1991), are part of a common refrain in ASUDAS studies, including most cited here. Recently, however, the degree of genetic contribution and concordance of dental and genetic data, in particular, have come under renewed scrutiny—several examples of which are summarized. Before proceeding, it of course goes without saying that non-ASUDAS traits are also used in dental research (e.g., Bailey & Hublin, 2013; Martinón-Torres et al., 2007); however, because few have been universally accepted or formally tested (below), such traits are not considered further.

Concerning genetic input, narrow-sense heritability (h^2) estimates in samples of Australian twins for several ASUDAS molar traits surpass 0.60, with UM1 Carabelli's reaching 0.80 and the UM1 and UM2 hypocone 0.87 and 0.93, respectively (Higgins, Hughes, James, & Townsend, 2009; Hughes & Townsend, 2011, 2013; Hughes, Townsend, & Bockmann, 2016). Yet, analyses of an African American Gullah sample (Stojanowski, Paul, Seidel, Duncan, & Guatelli-Steinberg, 2018, 2019) yielded some estimates that are more modest. For eight distinct ASUDAS crown traits recorded across incisor, canine, and premolar fields and between antimeres and isomeres, the average h^2 for traits with significant p-values is ≥.33. The range is .00–.82, depending on whether the rank-scale traits were treated as continuous data or dichotomized using alternate breakpoints (defined in Turner II et al., 1991; Scott & Turner II, 1997; Scott & Irish, 2017; Scott, Turner II, et al., 2018; below). In the same Gullah sample, Stojanowski et al. (2019) then looked at 14 distinct ASUDAS premolar and molar traits in maxillary and mandibular dental fields. Heritabilities are nearer the Australian findings, as UM1 Carabelli's reached 0.85 and, on average, all statistically significant trait estimates are higher than for Gullah anterior teeth. Still, several premolar and molar values are unexpectedly low. The practice of pooling sexes, in this sample at any rate, was also questioned contra one of the abovementioned ASUDAS attributes. This h^2 range is 0.00–1.00, depending on if the traits were treated as continuous or dichotomized; the latter provided much higher estimates. Socioeconomic stress and reproductive isolation in the Gullah, along with small samples, are acknowledged that may account for the low to moderate h^2 estimates (Stojanowski et al., 2018, 2019). In support, prior research also suggests stress affects expression of certain crown traits (Riga, Belcastro, & Moggi-Cecchi, 2014). Lastly, several of the above authors contributed to another study analyzing Australian twins (Paul, Stojanowski, Hughes, Brook, & Townsend, 2020). Results for permanent teeth reflect those of the Gullah studies, but with higher h^2 estimates (mostly 0.4–0.8), greater trait heritability when appropriately dichotomized, and less concern for sexual dimorphism in expression. Nevertheless, the above results offer mixed signals about genetic input to expression across a range of ASUDAS traits.

With regard to correspondence of dental and neutral genetic data in appraising relatedness, four recent studies are referenced. First, in a model-free analysis of four samples of living Kenyans (n = 295 individuals), correlations were calculated between pseudo-Mahalanobis D^2 distances based on nine ASUDAS crown traits, and delta-mu squared distances from 42 short tandem repeats (STRs) (Hubbard, 2012; Hubbard, Guatelli-Steinberg, & Irish, 2015). To date, this is the only research to compare phenotypic and genomic data in the same individuals. A moderate to strong (per Cohen, 1988) positive, though nonsignificant correlation, .50, resulted from Mantel (p = 0.21) and bivariate Pearson tests (p = 0.31) (Hubbard, Guatelli-Steinberg, & Irish, 2015). Second, 19 dental and up to 19 genetic global samples were matched for comparison by population or similar provenience and ethno-linguistic affinity (Rathmann et al., 2017). Most are recent—although the dates are not listed, two samples are from medieval to Victorian times, and one includes prehistoric material (>2000 BP). The model-bound R-matrix method, initially derived to compare allele frequencies, was used to produce pairwise population kinships from both dental (methods in Relethford, 1991; Konigsberg, 2006) and genomic data, all of which had been published previously: 12 ASUDAS crown traits (19 samples, 1872 individuals), 28 crown measurements (19 samples, 1,016 inds), 645 STR loci (13 samples, 265 inds), and 1,778 SNPs, that is, single-nucleotide polymorphisms (19 samples, 1,652 inds). Focusing just on ASUDAS results, Mantel tests yielded higher correlations than Hubbard et al. (2015), with an r_m-value for STRs in 13 samples of .55 (10,000 random permutations, $p < .001$) and for SNPs in 19 samples, .64 ($p < .001$). The dental-SNP correlation is strong (per Cohen, 1988), prompting Rathmann et al. (2017:3) to suggest dental data may be used as genetic proxies, although they "reason that a substantial portion of the variation can be explained by natural selection on dental morphology." Indeed, the presence and expression of several traits have been linked with selection (Bryk et al., 2008; Hlusko et al., 2018; Kimura et al., 2009; Park et al., 2012); this idea challenges yet another perceived attribute of the ASUDAS (see Scott & Irish, 2017; Turner II et al., 1991). Third, somewhat tangentially, Delgado et al. (2019) calculated phenetic distances from 16 ASUDAS traits in 477 living Colombians to those in dental samples of Europeans, Native Americans, and Africans (Irish, 1993, 1997; Scott & Turner II, 1997); these three are said to represent principal ancestors of admixed Latin Americans. The same Colombians had also been

genotyped, and 93,328 SNPs were compared with samples of the same populations to quantity admixture proportions. Dental-based affinities revealed the closest link with Europeans, and average ancestry estimates based on genetic and dental data generally concur. However, dental traits were not useful in assessing individual ancestries (Delgado et al., 2019). Lastly, to again explore whether dental and genetic data return similar information on admixture, Gross and Edgar (2019) employed Fisher Information in samples from West Africa, Europe, and North America. A model-bound clustering method, for multi-locus genotype data in the program STRUCTURE, was then applied to consider correspondence in estimating ancestry of individuals, that is, African, European/European American, and African American, with 53 unspecified crown traits (797 inds), up to 992,601 SNPs (271 inds), and 645 STRs (177 inds). Like most all dental studies many trait data are missing, which was suggested to affect the performance. Still, SNPs, followed by STRs delivered superior results in "detecting differences in admixture proportions between individuals within admixed populations;" dental data were, however, deemed to be useful to investigate population-level variation (Gross & Edgar, 2019:528). As above, outcomes of these four studies offer mixed support for ASUDAS traits, in this case pertaining to use as genetic proxies for analyses of populations and, in the latter two cases, individuals.

Today, it is patent that neutral genomic markers are the definitive choice in population (and individual) studies, and the standard to which all phenotypic data are and should be compared (see Rathmann et al., 2017). That said, on the above bases expression of the latter is minimally heritable for some, while analyses based on 9, 12, 16 and 53 ASUDAS traits failed to account fully for variation among population samples and/or individuals. Two other long held attributes—lack of sexual dimorphism and minimal selection—were also called into question. Nonetheless, given its long successful run, the aim here is to give the ASUDAS another chance to demonstrate its once-posited potential, through enhanced comparative analyses at both continental and global levels; heritability is not considered directly, but rather the capacity for these traits to approximate genetic relatedness.

All of these recent studies provide a foundation on which to build. The global approach of Rathmann et al. (2017), specifically, is used as a baseline. However, the first author's (JDI) African and other recent ASUDAS data (>100 samples, >6,000 individuals), and vast CG Turner II database (>300 samples, >23,000 inds; parts of which are presented in Scott & Irish, 2017) accessed here offer additional comparative choices. As such, many more dental samples could be matched with their genetic counterparts by population at a level of concordance not possible before (below). So, in 12 populations across the continent of Africa, correlations were calculated between the matrices of phenetic distances based on 36 and 25 ASUDAS traits, respectively, and genetic distances from >350,000 SNPs. The latter were selected over STRs because (a) published SNP data are available for more global populations, (b) they afford superior differentiation among such populations, and (c) these markers seemingly correspond more closely with dental nonmetric data (Gross & Edgar, 2019; Rathmann et al., 2017). Matrices calculated from the same dental traits and SNPs were then tested for correlation in an expanded analysis of 32 total global populations. Finally, it is assumed genetic and, by extension (as above), phenetic distances among populations increase exponentially as geographic distances increase (Wright, 1943; Relethford, 2004); thus, correlations between the latter distances among samples and those from dental and genetic data were calculated to explore the influence of geographic structure (e.g., extreme isolation) on the two datasets.

2 | MATERIALS AND METHODS

Given the large dataset and familiarity of JDI with the post-Pleistocene peopling of Africa, the latter was the clear choice for continental-level analyses; dental and genetic samples of three North and nine sub-Saharan African populations were compared. Other than one dental sample (Riet River San, Table 1) with a few earlier historic specimens, all data were recorded in recent, 19-20th century crania and hardstone casts to match close as possible the existing genomic data from living individuals (Table 2). Next, combined with Africans for global analyses were 20 dental samples from Europe, Asia, Australia, Melanesia, and the Americas. With two earlier historic exceptions (Table 1), the latter comprise similarly recent specimens to compare with 20 matched genetic samples (Table 2; Figure 1). In all following tables and figures, samples are abbreviated with a prefix of "D" (dental) or "G" (genetic), followed by sample number (1–32), and three letters for the name. For instance, the Bedouin dental sample D1_BED (Table 1) corresponds with genetic sample G1_MOR from Morocco, and so forth (Table 2). An extensive anthropological literature review facilitated sample matching based on (a) shared language and ethnic groups (e.g., Turner II, 1985a; Scott & Turner II, 1997; Irish, 1993, 1997, 2000; Irish et al., 2014; Irish, 2016; see below), and (b) similar geographic locations. The average distance in km between the matched African dental and genetic samples, as determined from latitude and longitude coordinates (Tables 1 and 2), is 286.6 with a low of 64.2 and high of 505.1. The mean for all 32 samples is 347.9 km with a range, excluding the African low, of 89.6 to 1,361.3 km. The latter is the distance between the dental and genetic Aleut samples—one from the American and the other from the Russian side of the island chain. These and the two Pima samples separated by 581.9 km were included to provide some level of New World coverage. That is, recent Native North American dental data are ample in Turner's database, but matching genetic data are not (Reich et al., 2012; Skoglund et al., 2015), with the reverse true for recent dental and genetic data in Meso- and South Americans.

Thirty-six crown, root, and intra-oral osseous nonmetric traits (refer to list in Table 3) used in previous affinity studies (Irish, 1993, 1997, 1998a, 1998b, 2000, 2005, 2006, 2010, 2016; Irish et al., 2014; Irish et al., 2017) were initially compared with the SNP data for the African analysis. Beyond the abovementioned ASUDAS attributes, the rationale for choosing, and the standard approach in recording these specific traits are detailed in the preceding references and elsewhere (Scott & Irish, 2017; Scott & Turner II, 1997; Scott,

TABLE 1 Dental samples used in the study with background information

Dental sample	Abbreviation	Region	Country/Area	Data source	n	Lat	Lon
Africa							
Bedouin (Arab)	D1_BED	North Africa	Morocco and Algeria	Irish, 1993, 1998a	49	34.8	−5.2
Kabyle (Berber)	D2_KAB	North Africa	Algeria	Irish, 1993, 1998a	32	36.6	3.7
Kikuyu	D3_KKU	East Africa	Kenya	Irish unpublished data	60	−0.3	36.1
Riet River (San; >12-19th Cent)[a]	D4_RRI	South Africa	South Africa	Irish, Black, Sealy, & Ackermann, 2014	66	−29.3	24.8
San	D5_SAN	South Africa	Botswana, South Africa	Irish, 1993, 1997	99	−22.4	24.6
Senegambia (Wolof)	D6_SEN	West Africa	Senegambia	Irish, 1993, 1997	42	15.2	−16.7
Shawia (Berber)	D7_SHA	North Africa	Algeria	Irish, 1993, 1998a	26	35.4	6.7
Somalia	D8_SOM	East Africa	Somalia	Irish, 2010	77	9.0	46.4
Sotho	D9_SOT	South Africa	South Africa	Irish, 2016	66	−29.4	28.3
Tswana	D10_TSW	South Africa	South Africa	Irish, 2016	63	−25.8	23.0
Yoruba	D11_YOR	West Africa	Benin (Dahomey)	Irish unpublished data	28	6.6	2.6
Zulu	D12_ZUL	South Africa	South Africa	Irish, 2016	67	−28.0	32.4
				Total	675		
America, Asia, Australia, Melanesia, Europe							
Pima 94	D13_PIM	North America	Salt River–Maricopa, Arizona	Turner unpublished data	165	33.3	−111.5
Aleut (Western US)	D14_ALE	North America	Attu, Atka plus Western Aleut Historic	Turner unpublished data; Scott & Irish, 2017	95	52.0	−174.0
Kazak 94 (17-19th Cent)[a]	D15_KAZ	Central Asia	East Kazakhstan	Turner unpublished data	204	47.0	76.0
Mongol 2 and 3 Pooled	D16_MON	Central Asia	Northeast Mongolia	Turner unpublished data; Turner 1990	82	48.0	110.0
Lower Ob Khanty	D17_LOK	Central Asia	Khant-Mansi (Ugrian), Central Russia	Turner unpublished data	49	63.0	70.0
Chukchi plus Eastern Siberia	D18_CHU	Central Asia	Northeast Russia	Turner unpublished data	126	67.5	170.0
Recent Thailand	D19_THA	East Asia	Central Thailand	Turner unpublished data; Turner 1990	189	13.0	101.0
Recent Tonkin, Historic Annam	D20_VIE	East Asia	North Vietnam	Turner unpublished data	76	20.0	107.0
Recent Japanese	D21_JAP	East Asia	Central Japan	Turner unpublished data; Scott & Irish, 2017	131	36.0	138.0
Malay Composite	D22_MAL	Southeast Asia	Central Malaysia	Turner unpublished data; Scott & Irish, 2017	58	1.0	102.5
Philippines no 2 Calatagan BP	D23_PHI	Southeast Asia	Central Philippines	Turner unpublished data; Scott & Irish, 2017	58	12.3	122.0
Borneo 94	D24_BOR	Southeast Asia	Central Borneo	Turner unpublished data; Scott & Irish, 2017	144	1.5	114.5
Australia—North BP	D25_AUN	Australia	Northeast Australia	Turner unpublished data; Scott & Irish, 2017	57	−20.8	139.5
New Britain 1_4 738 BP_ no 3	D26_NBR	Melanesia	New Britain	Turner unpublished data; Scott & Irish, 2017	238	−6.0	150.0
Nepal 94 BP	D27_NEP	South Asia	Central Nepal	Turner unpublished data	97	28.0	84.0
Greek Recent	D28_GRK	South Europe	South Greece	Irish, Lillios, Waterman, & Silva, 2017	70	37.5	22.3
Italy Modern	D29_ITY	South Europe	Central Italy	Irish et al., 2017	55	42.0	14.0

TABLE 1 (Continued)

Dental sample	Abbreviation	Region	Country/Area	Data source	n	Lat	Lon
Kaberla 1,2,3 (13th–17th Cent)[b]	D30_KBR	North Europe	North Estonia	Turner unpublished data	160	59.5	25.3
Ladoga Finns	D31_FIN	North Europe	Finland/Western Russia	Turner unpublished data	51	61.0	30.0
Lapps (Kola Peninsula)	D32_LAP	North Europe	Lapland/Northwest Russia	Turner unpublished data; Scott & Irish, 2017	64	67.0	40.0
				Total	2,169		
				Grand Total	2,844		

[a]Samples contain some pre-19th century specimens.
[b]Sample specimens are all pre-19th century.

Turner II, et al., 2018; Turner II et al., 1991). As well, these traits were found to be largely independent of one another (Nichol, 1990), a key factor in avoiding data redundancy that adds little additional discriminatory value in biodistance analyses (see below); independence of these 36 traits was subsequently supported with a range of pairwise Kendall's tau-*b* correlations among the rank-scale data in 1625 Africans of just |0.00–0.33| (Irish, 1993). A succession of later studies did identify some strong correlations, $\tau_b \geq |0.5|$, but the lack of patterning suggests the affected trait pairs vary across populations (Irish, 2005, 2006, 2010, 2016; Irish et al., 2014). Next, as mentioned, the 12 dental and 12 genetic samples were compared again after dropping 11 traits: UI1 labial curvature, palatine torus, UC distal accessory ridge, UI2 peg-reduced, UI1 midline diastema, LM1 anterior fovea, mandibular torus, rocker jaw, LM1 deflecting wrinkle, LM1 C1-C2 crest, and LM2 torsomolar angle. This reduction allowed direct comparison of the Africans with 20 additional global samples (Tables 1 and 2), 18 of which are in Turner's dental database; he regularly recorded just 25 traits (e.g., Turner II, 1985a; and below). Any dissimilarities in the African results relating to different dental trait numbers were then quantified.

All SNP data, except whole genome sequences of the West African Wolof sample (Table 2), were genotyped with the Affymetrix Human Origins Array (AHOA) (in Patterson et al., 2012; Pickrell & Pritchard, 2012; Lazaridis et al., 2014; Pickrell et al., 2014; Skoglund et al., 2016; or by request from these authors via signed letters). These data are high density, that is, 593,124 SNPs, and ascertained for all modern populations. Specifically, because the AHOA was built from 13 different global samples, ascertainment bias (i.e., systematic distortion of true allele frequencies) is limited, to reliably represent demographic history. The five low coverage Wolof sequences (9X) were produced by the Gambian Genome Variation Project. To merge them with the AHOA dataset, all reads were mapped against the reference genome (Hg19/GRCh37, 1,000 Genome release) by the second author (AM), using the Burrow-Wheeler Aligner-MEM with –M option (BWA-MEM) (Li & Durbin, 2009). Any duplicates reads were removed with markdup, in Samtools Release 1.9 (http://www.htslib.org/) (Li et al., 2009). The genotypes were called only on SNPs in the AHOA. Samtools 1.9 (Li et al., 2009) mpileup was used to generate genotype likelihoods for each SNP, and post-genotype filtering was employed to remove bases with a phred score (base quality) of <30, reads with a mapping quality of <30 (probability of correctly mapping a read >.999), and for sequences with mismatches of >50% (when this percentage of the read's bases differ from the reference). Then, genotypes were called using bcftools default method (–m option) (Danecek, Schiffels, & Durbin, 2014), omitting insertions-deletions (indels). Filtered out with the bcftools filter were all variants having a calling quality of <20% and depth of >118—where a general rule for maximum coverage depth is to filter position depth as DP > 2*DP (for the Wolof genomes DP = 59). SNPs (n = 144) that did not match Human Origins allele codes (i.e., a "new" allele was discerned in a sequence) or had minor allele frequencies (<.05) were removed. This progression yielded a final total of 353,091 SNPs for the African and global comparative analyses.

At both geographic levels a model-free approach (Hubbard et al., 2015), standard in most phenotype affinity studies, was conducted (though see below). For example the R-matrix method, to estimate between-population kinship coefficients (Rathmann et al., 2017), may not be the best suited for SNP data—particularly the large numbers, and correlation of results (below) based on dental and genomic data is unlikely to be dependent on the distance measures (Relethford, personal communication, 2019). An advantage of the R-matrix method is that it can correct for genetic drift among samples, but this weighting procedure is only possible when effective population sizes are known (Leigh, Relethford, Park, & Konigsberg, 2003; Relethford & Crawford, 1995), a difficult proposition with premodern peoples (Irish, 2016). So, to evaluate correspondence, common field-specific measures of divergence based on ASUDAS and SNP data among respective dental and genetic samples were obtained here, where low distance values indicate similitude and vice versa between samples.

For dental data, the mean measure of divergence (MMD) was chosen relative to others, for example, pseudo-Mahalanobis D^2 (Konigsberg, 1990). It is a robust statistic that yields reliable results even with problematic traits, such as those, that are highly intercorrelated or invariant across samples; it is also less affected by missing data that characterize most dental studies and, while not necessary for the present comparative analyses, has a significance test (Irish, 2010; Nikita, 2015; Sjøvold, 1973, 1977). Finally, it was

TABLE 2 Genetic samples used in the study with background information

Sample name	Abbreviation	Region	Country/area	Data source	n	Lat	Lon
Africa							
Moroccan	G1_MOR	North Africa	Morocco, Casablanca	Lazaridis et al., 2014	10	33.5	−7.6
Algerian	G2_ALG	North Africa	Algeria	Lazaridis et al., 2014	7	36.8	3.0
Kikuyu	G3_KKU	East Africa	Kenya	Lazaridis et al., 2014	4	−0.4	36.9
Khomani (San)	G4_KHO	South Africa	South Africa	Lazaridis et al., 2014	11	−27.8	21.1
Ju_hoan_North (San)	G5_JUH	South Africa	Namibia	Patterson et al., 2012; Pickrell & Pritchard, 2012	21	−18.9	21.5
Wolof	G6_WOL	West Africa	Gambia	Gambian Genome Variation Project	5	13.4	−16.7
Mozabite	G7_MOZ	North Africa	Algeria	Patterson et al., 2012	21	32.0	3.0
Somalia	G8_SOM	East Africa	Somalia	Lazaridis et al., 2014	13	5.6	48.3
Sotho	G9_SOT	South Africa	South Africa	Patterson et al., 2012	1	−29.0	29.0
Tswana	G10_TSW	South Africa	South Africa/ Botswana/ Namibia	Patterson et al., 2012; Pickrell & Pritchard, 2012	7	−28.0	24.0
Yoruba	G11_YOR	West Africa	Nigeria	Lazaridis et al., 2014	70	7.4	3.9
Zulu	G12_ZUL	South Africa	South Africa	Patterson et al., 2012	1	−28.0	31.0
				Total	171		
America, Asia, Australia, Melanesia, Europe							
Pima	G13_PIM	Mesoamerica	Chihuahua, Mexico	Patterson et al., 2012	14	29.0	−108.0
Aleut (Nikolskoye)	G14_ALE	East Russia	Bering Island, Russia	Lazaridis et al., 2014	2	55.2	166.0
Kyrgyz	G15_KRG	Central Asia	North Kyrgyzstan	Lazaridis et al., 2014	9	42.9	74.6
Mongola	G16_MON	Central Asia	East Mongolia	Patterson et al., 2012	6	45.0	111.0
Mansi	G17_MAN	Central Asia	Central Russia (Konda River)	Lazaridis et al., 2014	3	62.5	63.3
Chukchi	G18_CHU	Central Asia	Northeast Russia	Lazaridis et al., 2014	10	69.5	168.8
Thai	G19_THA	East Asia	Central Thailand	Lazaridis et al., 2014	10	13.8	100.5
Kinh_Vietnam_KHV	G20_KIN	East Asia	North Vietnam	Lazaridis et al., 2014	8	21.0	105.9
Japanese	G21_JAP	East Asia	Central Japan	Patterson et al., 2012	29	38.0	138.0
Malays	G22_MAL	Southeast Asia	Central Malaysia	Skoglund et al., 2016	9	4.2	102.0
Visayan, Kankanaey, Ilocano, Tagalog	G23_PHI	Southeast Asia	Central Philippines	Skoglund et al., 2016	21	9.8	125.5
Lebbo	G24_LEB	Southeast Asia	Central Borneo	Skoglund et al., 2016 (signed letter)	8	0.0	115.0
CAI - North Australia/ Queensland, WPA - North Australia/Queensland, Australian_ECCAC	G25_AUN	Australia	Northeast Australia	Lazaridis et al., 2014	3	−16.9	145.0
All HO New Britain from Skoglund 2016	G26_NBR	Melanesia	New Britain	Skoglund et al., 2016 (signed letter)	156	−5.8	150.8
Kusunda	G27_KUS	South Asia	Central Nepal	Lazaridis et al., 2014	10	28.1	82.5
Greek_Coriell	G28_GRK	South Europe	East Greece	Lazaridis et al., 2014	20	38.0	23.7
Italian_Tuscan	G29_ITY	South Europe	Central Italy	Patterson et al., 2012	20	43.0	11.0
Estonian	G30_EST	North Europe	West Estonia	Lazaridis et al., 2014	10	58.5	24.9
Finnish_FIN	G31_FIN	North Europe	South Finland	Lazaridis et al., 2014	8	60.2	24.9

TABLE 2 (Continued)

Sample name	Abbreviation	Region	Country/area	Data source	n	Lat	Lon
Saami_WGA	G32_SAM	North Europe	North Finland	Lazaridis et al., 2014; Mallick et al., 2016	3	68.4	23.6
				Total	359		
				Grand Total	530		

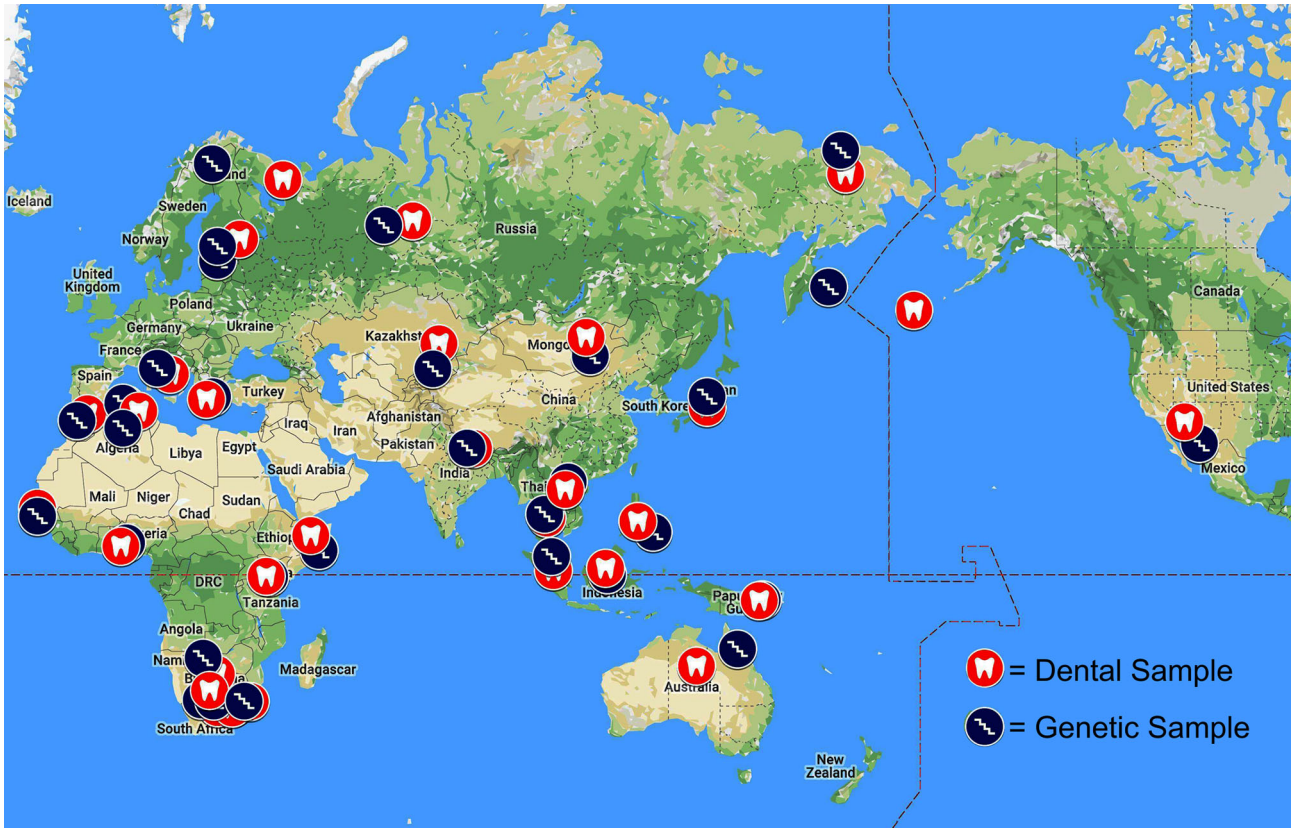

FIGURE 1 Origin locations of the 32 dental and matched 32 genetic samples. Modified version originally created using Google My Maps (https://www.google.com)

found that MMD values are more highly correlated with geographic distances (Irish, 2010, 2016; Schillaci, Irish, & Wood, 2009). The formula used here has a bias correction, the Freeman and Tukey angular transformation, to correct for very low or high trait frequencies and small sample sizes (Green & Suchey, 1976; Sjøvold, 1973, 1977). As required by the MMD and to simplify presentation of dental trait frequencies, rank-scale ASUDAS data were dichotomized into categories of present and absent using standard breakpoints (refer to Table 3) (Irish, 1993, 1997, 2005, 2006; Scott & Irish, 2017; Scott & Turner II, 1997; Scott, Turner II, et al., 2018). A few workers suggest rank-scale data would give better results (Gross & Edgar, 2019; Nikita, 2015; Rathmann et al., 2017)—a concept that is not novel (see Sjøvold, 1977; Turner II, 1985b). However, suitably dichotomized trait data, perhaps surprisingly, hold several advantages (a) importantly, h^2 estimates were demonstrated to be higher (see Stojanowski et al., 2019 for details), (b) weighting bias from different grade numbers across ASUDAS traits is avoided, (c) proven distance statistics like the MMD (and D^2) can be applied, and (d) residual intra- and inter-observer error is reduced further. That said, the latter should be negligible, at least relative to the above studies, because data were recorded by Turner—the ASUDAS designer, and JDI, who was directly instructed by and calibrated with him (Haeussler, Turner II, & Irish, 1988; Irish & Turner II, 1990).

Many measures of divergence are available for genomic data. However, only f_2, outgroup-f_3, and F_{ST} were considered due to their ubiquitous application, the availability of online programs and, importantly, the capability of these programs to process large numbers of markers (Holsinger & Weir, 2009; Patterson et al., 2012; Peter, 2016;

TABLE 3 Percentages (%) of 36 dental traits considered present and number of individuals scored (n) in the 12 North and sub-Saharan African samples

Trait/grades present		D1_BED[a]	D2_KAB	D3_KKU	D4_RRI	D5_SAN	D6_SEN	D7_SHA	D8_SOM	D9_SOT	D10_TSW	D11_YOR	D12_ZUL
Winging UI1	%	5.4	0.0	0.0	8.0	13.8	4.0	0.0	6.0	1.9	4.2	0.0	0.0
(+ = ASU 1)[b]	n	37	29	53	50	116	25	26	67	52	48	21	57
Labial curvature UI1	%	37.5	50.0	50.0	53.1	63.5	50.0	28.6	27.8	60.6	68.6	42.9	77.8
(+ = ASU 2–4)	n	24	8	12	32	85	12	7	18	33	35	7	36
Palatine torus	%	2.4	3.5	0.0	8.1	0.9	10.5	0.0	0.0	0.0	0.0	0.0	0.0
(+ = ASU 2–3)	n	41	29	59	62	117	38	25	70	63	60	28	61
Shoveling UI1	%	0.0	0.0	0.0	0.0	7.5	10.0	0.0	0.0	0.0	0.0	0.0	0.0
(+ = ASU 3–6)	n	25	7	16	26	80	10	7	18	33	29	9	37
Double shoveling UI1	%	12.5	12.5	0.0	0.0	0.0	8.3	25.0	5.9	2.7	0.0	0.0	0.0
(+ = ASU 2–6)	n	24	8	16	31	83	12	8	17	37	34	8	36
Interruption groove UI2	%	37.5	21.4	14.3	3.7	14.4	21.4	46.2	0.0	4.7	0.0	10.0	15.8
(+ = ASU +)	n	24	14	14	27	90	14	13	24	43	38	10	38
Tuberculum dentale UI2	%	43.5	50.0	33.3	42.3	44.2	58.3	25.0	39.1	36.6	37.5	50.0	29.7
(+ = ASU 2–6)	n	23	12	15	26	86	12	12	23	41	32	10	37
Bushman canine UC	%	0.0	0.0	11.5	31.0	38.8	10.5	0.0	3.1	24.4	37.8	22.2	23.4
(+ = ASU 1–3)	n	29	16	26	29	85	19	14	32	45	45	18	47
Distal accessory ridge UC	%	12.0	27.3	29.6	16.7	21.8	41.7	22.2	34.5	39.0	60.6	14.3	26.3
(+ = ASU 2–5)	n	25	11	27	24	78	12	9	29	41	33	14	38
Hypocone UM2	%	58.8	63.6	79.0	96.0	86.4	64.5	68.4	71.0	93.7	77.6	73.9	91.1
(+ = ASU 3–5)	n	34	22	57	50	110	31	19	69	63	49	23	56
Cusp 5 UM1	%	8.8	11.8	25.0	24.1	31.8	11.1	10.0	19.4	22.0	17.0	4.8	18.4
(+ = ASU 2–5)	n	34	17	56	29	85	27	20	62	59	47	21	49
Carabelli's trait UM1	%	39.4	52.6	44.6	12.5	22.1	31.0	27.8	51.6	32.2	29.4	43.5	44.0
(+ = ASU 3–7)	n	33	19	56	24	95	29	18	62	59	51	23	50
Parastyle UM3	%	0.0	0.0	4.4	2.7	0.0	0.0	7.7	2.1	0.0	2.3	4.0	0.0
(+ = ASU 3–5)	n	20	22	45	37	69	28	13	48	52	43	25	42
Enamel extension UM1	%	5.6	0.0	0.0	0.0	2.3	5.9	4.8	4.4	4.8	3.9	10.7	1.7
(+ = ASU 1–3)	n	36	23	51	61	43	34	21	68	63	51	28	58
Root number UP1	%	50.0	52.2	69.2	44.7	39.5	69.0	52.2	75.0	74.4	57.1	53.9	69.0
(+ = ASU 2+)	n	32	23	39	38	38	29	23	44	39	35	13	29
Root number UM2	%	69.0	68.4	87.5	80.8	84.9	78.3	72.2	82.9	83.3	65.0	100.0	84.6
(+ = ASU 3+)	n	29	19	16	26	33	23	18	41	24	20	8	26
Peg-reduced UI2	%	0.0	6.3	0.0	1.6	7.0	0.0	0.0	0.0	3.3	1.9	4.2	3.2

TABLE 3 (Continued)

Trait/grades present		D1_BED[a]	D2_KAB	D3_KKU	D4_RRI	D5_SAN	D6_SEN	D7_SHA	D8_SOM	D9_SOT	D10_TSW	D11_YOR	D12_ZUL
(+ = ASU P or R)	n	27	16	55	64	115	20	13	73	61	54	24	63
Odontome P1-P2	%	0.0	0.0	4.1	0.0	0.0	0.0	0.0	0.0	1.6	1.6	4.2	0.0
(+ = ASU +)	n	40	22	49	39	105	32	23	68	64	61	24	63
Congenital absence UM3	%	21.1	3.5	5.1	7.3	2.0	13.5	23.1	4.1	4.6	7.1	3.7	6.8
(+ = ASU -)	n	38	29	59	69	98	37	26	74	66	56	27	59
Midline diastema UI1	%	8.8	12.0	7.6	5.1	9.7	7.7	0.0	3.0	8.5	2.1	15.8	10.2
(+ ≥ 0.5 mm)	n	34	25	53	59	114	26	23	67	59	47	19	59
Lingual cusp LP2	%	64.3	69.2	80.8	74.2	64.6	58.3	92.3	63.2	66.7	68.6	50.0	54.7
(+ = ASU 2–9)	n	28	13	26	31	96	12	13	38	57	51	8	53
Anterior fovea LM1	%	37.5	60.0	84.8	70.0	70.6	50.0	29.4	50.0	68.8	74.4	70.0	54.0
(+ = ASU 2–4)	n	24	10	46	20	68	12	17	52	48	43	10	50
Mandibular torus	%	2.9	0.0	0.0	1.8	0.0	0.0	4.2	0.0	1.5	0.0	0.0	0.0
(+ = ASU 2–3)	n	35	19	55	55	116	15	24	64	65	61	14	63
Groove pattern LM2	%	46.9	27.8	75.0	67.4	74.1	58.3	36.8	58.3	64.8	75.6	42.9	72.0
(+ = ASU Y)	n	32	18	48	46	112	12	19	60	54	45	14	50
Rocker jaw	%	9.4	10.5	11.1	5.7	8.9	6.7	8.3	9.4	1.5	5.2	14.3	3.2
(+ = ASU 1–2)	n	32	19	54	53	45	15	24	64	66	58	14	62
Cusp number LM1	%	12.5	31.3	13.0	8.0	4.7	7.1	9.5	14.0	13.8	6.3	0.0	1.9
(+ = ASU 6+)	n	32	16	46	25	107	14	21	57	58	48	11	54
Cusp number LM2	%	42.9	33.3	82.0	96.7	92.6	54.6	31.6	57.9	83.6	76.6	83.3	90.0
(+ = ASU 5+)	n	28	18	50	30	108	11	19	57	55	47	12	50
Deflecting wrinkle LM1	%	15.6	6.7	34.0	34.6	25.6	27.3	5.0	27.1	30.0	26.7	20.0	31.4
(+ = ASU 2–3)	n	32	15	47	26	78	11	20	59	50	45	10	51
C1-C2 crest LM1	%	3.0	0.0	0.0	0.0	3.1	8.3	0.0	0.0	1.9	2.2	0.0	2.0
(+ = ASU +)	n	33	14	47	25	65	12	20	60	53	45	11	51
Protostylid LM1	%	0.0	0.0	0.0	5.6	7.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
(+ = ASU 3–6)	n	33	16	46	18	103	13	21	58	59	49	13	56
Cusp 7 LM1	%	5.9	5.9	31.9	22.0	27.9	28.6	4.8	28.8	37.9	37.5	18.2	49.1
(+ = ASU 2–4)	n	34	17	47	41	111	14	21	59	58	48	11	55
Tome's root LP1	%	6.3	5.3	14.9	5.3	2.6	14.3	10.5	18.4	10.5	5.7	33.3	7.4
(+ = ASU 3–5)	n	32	19	47	38	39	14	19	49	38	35	12	27
Root number LC	%	0.0	20.0	0.0	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0
(+ = ASU 2+)	n	26	10	50	37	42	12	16	53	32	29	13	33
Root number LM1	%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

(Continues)

TABLE 3 (Continued)

Trait/grades present		D1_BED[a]	D2_KAB	D3_KKU	D4_RRI	D5_SAN	D6_SEN	D7_SHA	D8_SOM	D9_SOT	D10_TSW	D11_YOR	D12_ZUL
(+ = ASU 3+)	n	33	17	25	29	32	15	22	51	26	22	14	28
Root number LM2	%	88.9	88.9	96.4	83.3	85.7	92.3	95.5	93.8	92.9	89.3	100.0	72.7
(+ = ASU 2+)	n	27	18	28	24	28	13	22	48	28	28	10	22
Torsomolar angle LM3	%	20.0	21.4	6.3	9.1	20.2	18.2	23.5	10.0	8.0	8.1	0.0	8.3
(+ = ASU +)	n	25	14	48	44	84	11	17	50	50	37	14	48

[a]D1_BED = Bedouin, D2_KAB=Kabyle, D3_KKU=Kikuyu, D4_RRI = Riet River San, D5_SAN=San, D6_SEN=Senegambia, D7_SHA = Shawia, D8_SOM = Somalia, D9_SOT = Sotho, D10_TSW = Tswana, D11_YOR = Yoruba, D12_ZUL = Zulu. Details in Table 1.
[b]ASUDAS rank-scale trait breakpoint information in Irish (1993, 1997, 2005, 2006), Scott & Turner II (1997), Scott & Irish (2017) and Scott, Turner II, et al., 2018.

Reich, Thangaraj, Patterson, Price, & Singh, 2009; Skoglund et al., 2015). Of these, F_{ST} was selected as most appropriate for the SNP data, samples, and overall approach (Supporting information Text S1.1). Calculated under the Hardy–Weinberg model, it is theoretically not model-free; however, it is used in that capacity here, like the MMD, to "describe overall patterns of variation that can be interpreted in light of population history and structure" (Relethford & Harpending, 1994:251). The motivation is that, beyond its descriptive attributes and use for identifying genomic regions under selection, F_{ST} yields a more generalized estimate of genetic differentiation among population pairs (but see Séré, Thévenon, Belem, & De Meeûs, 2017). Further, it (a) works with small samples if, as here, many loci are included, (b) does well on a broad geographic scale, (c) is not associated with population divergence time, (d) is reliable regardless of the population structure model, and (e) like the MMD, maintains constant pairwise values if any samples are added (Diniz-Filho et al., 2013; Holsinger & Weir, 2009; Nelis et al., 2009; Ortega-Del Vecchyo & Slatkin, 2019; Peter, 2016; Tian et al., 2009; Weir, Cardon, Anderson, Nielsen, & Hill, 2005; Willing, Dreyer, & van Oosterhout, 2012). Both the Weir and Cockerham (1984) and Hudson F_{ST} estimators (Hudson, Slatkin, & Maddison, 1992) were used, with the results of the latter detailed below. The Hudson estimator is stated to return more accurate distances with genomic data, and is less affected by very small sizes, that is, $n < 4$, that affect several genetic samples (see Table 2) (Bhatia, Patterson, Sankararaman, & Price, 2013; Ortega-Del Vecchyo & Slatkin, 2019).

Geographic distances among the 32 dental and 32 genetic samples were determined based on associated latitudes and longitudes in decimal degrees submitted to the Similarity and Distances Indices module in PAST 3.23 (http://folk.uio.no/ohammer/past; Hammer, Harper, & Ryan, 2001; Hammer, 2019). The default PAST output is meters, which were used for all quantitative analyses; however, for convention they were converted to km in the corresponding tables and figures (below and Supporting information). Without reference to hypothesized migration routes they measure, simply, straight-line distances along a great circle over the surface of the earth between coordinates of sample pairs in the WGS84 datum. This method also accounts for variance between longitudes at high and low latitudes (Hammer, 2019, personal communication, 2019). Equivalent results were returned with the Geographic Distance Matrix Generator (version 1.2.3) (Ersts, 2014).

Finally, as mentioned, correlations between dental, genetic, and geographic distances were calculated. Like prior studies (Hubbard et al., 2015; Rathmann et al., 2017) Mantel tests were used, with a null hypothesis of no association between matrices (Mantel, 1967; Smouse & Long, 1992; Smouse, Long, & Sokal, 1986; Sokal & Rohlf, 1995). To explore dental vs. genetic correspondence irrespective of geographic separation, partial Mantel tests were also conducted; for these, the third variable consists of "midpoint" distances calculated from mean latitude and longitude coordinates between the matching pairs of dental and genetic samples. Of course, Mantel tests are not without criticism. It was said that r_m-values "obtained by permutations do not display enough variability,"

TABLE 4 MMD distance matrix (bottom diagonal) for the three North and nine sub-Saharan African dental samples based on 36 ASUDAS traits, and F_{ST} distance matrix (top diagonal) for the three North and nine sub-Saharan African genetic samples based on 353,091 SNPs

	G1_MOR[a]	G2_ALG	G3_KKU	G4_KHO	G5_JUH	G6_WOL	G7_MOZ	G8_SOM	G9_SOT	G10_TSW	G11_YOR	G12_ZUL	
D1_BED[b]		0.016	0.053	0.122	0.171	0.082	0.009	0.031	0.091	0.094	0.086	0.095	G1_MOR
D2_KAB	0.000		0.070	0.138	0.188	0.099	0.024	0.046	0.110	0.112	0.103	0.112	G2_ALG
D3_KKU	0.095	0.078		0.063	0.101	0.026	0.066	0.020	0.016	0.020	0.015	0.022	G3_KKU
D4_RRI	0.136	0.145	0.025		0.028	0.078	0.136	0.087	0.039	0.038	0.070	0.055	G4_KHO
D5_SAN	0.149	0.150	0.050	0.000		0.111	0.186	0.131	0.072	0.068	0.105	0.089	G5_JUH
D6_SEN	0.000	0.000	0.015	0.050	0.046		0.096	0.053	0.028	0.032	0.021	0.031	G6_WOL
D7_SHA	0.000	0.000	0.113	0.177	0.205	0.006		0.043	0.106	0.110	0.100	0.110	G7_MOZ
D8_SOM	0.048	0.031	0.018	0.087	0.115	0.000	0.067		0.051	0.057	0.049	0.057	G8_SOM
D9_SOT	0.117	0.096	0.000	0.009	0.032	0.008	0.146	0.025		0.002	0.011	0.004	G9_SOT
D10_TSW	0.150	0.127	0.017	0.026	0.038	0.023	0.178	0.048	0.000		0.016	0.008	G10_TSW
D11_YOR	0.049	0.045	0.000	0.032	0.041	0.000	0.085	0.007	0.001	0.043		0.014	G11_YOR
D12_ZUL	0.133	0.139	0.020	0.029	0.033	0.033	0.187	0.075	0.000	0.024	0.024		G12_ZUL
	D1_BED	D2_KAB	D3_KKU	D4_RRI	D5_SAN	D6_SEN	D7_SHA	D8_SOM	D9_SOT	D10_TSW	D11_YOR	D12_ZUL	

[a]G1_MOR = Moroccan, G2_ALG = Algerian, G3_KKU=Kikuyu, G4_KHO=Khomani San, G5_JUH = Ju-hoan North San, G6_WOL = Wolof, G7_MOZ = Mozabite, G8_SOM = Somalia, G9_SOT = Sotho, G10_TSW = Tswana, G11_YOR = Yoruba, G12_ZUL = Zulu. Details in Table 2.

[b]D1_BED = Bedouin, D2_KAB=Kabyle, D3_KKU=Kikuyu, D4_RRI = Riet River San, D5_SAN=San, D6_SEN=Senegambia, D7_SHA = Shawia, D8_SOM = Somalia, D9_SOT = Sotho, D10_TSW = Tswana, D11_YOR = Yoruba, D12_ZUL = Zulu. Details in Table 1.

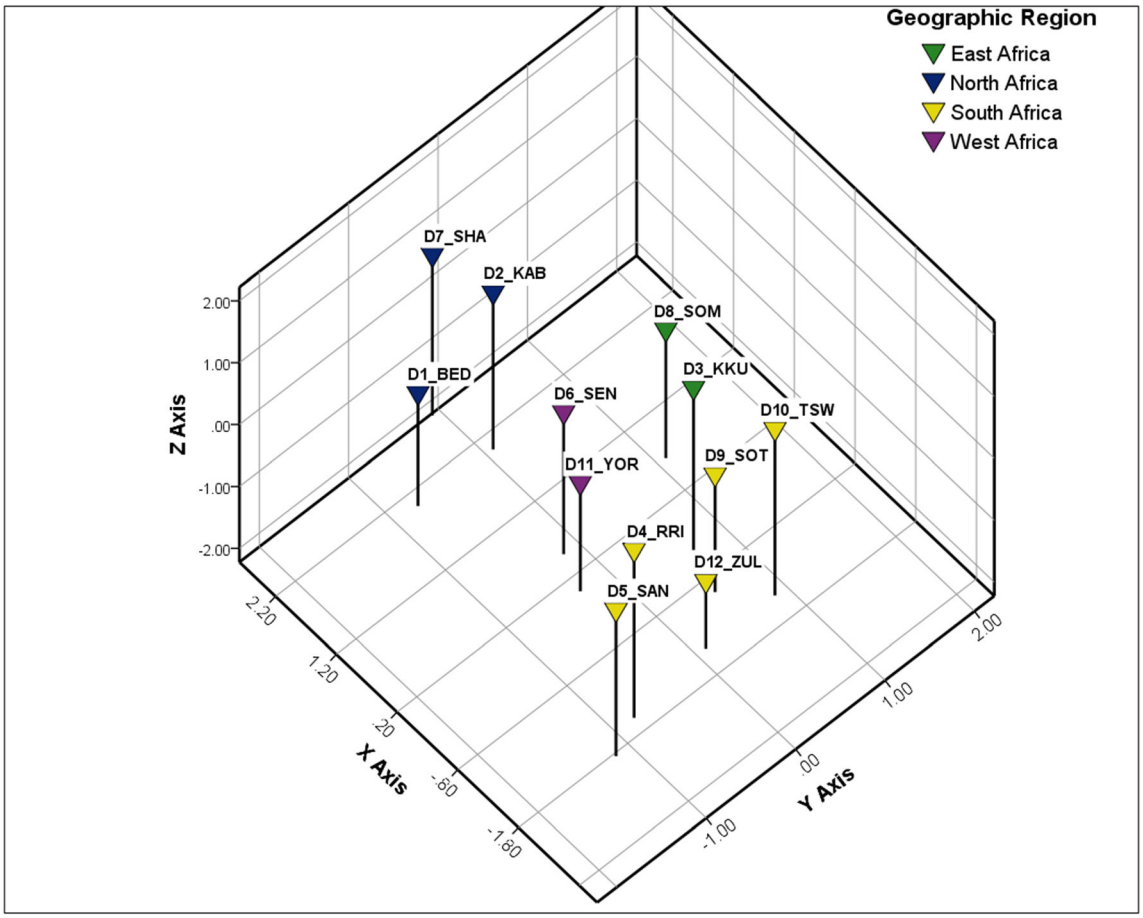

FIGURE 2 Three-dimensional MDS plot of the 36-trait MMD distances among the 12 African dental samples. The sample abbreviations are defined in Tables 1, 3-6

and spatial autocorrelation, which is often inherent with biological data, can lead to inaccurately low *p*-values (Legendre & Fortin, 2010; Guillot & Rousset, 2013:341). Nevertheless, Mantel tests have been shown to be robust (Séré et al., 2017), are easy to interpret, remain widely used (above) to facilitate between-study comparisons, and various alternative methods are neither unreservedly accepted nor, in this case, applicable, for example, comparing directly the non-dichotomized ASUDAS and SNP data. In any event, results based on applications of Mantel and alternative methods were demonstrated to largely converge (Diniz-Filho et al., 2013). Therefore, distance matrices were submitted to the Mantel Test module of PAST 3.23. The resulting r_m-values are Pearson's correlation coefficients, with one-tailed *p*-values from 10,000 random permutations (Hammer, 2019). A *p*-value adjustment such as the Bonferroni procedure is often used to address Type I error from multiple testing (see Rathmann et al., 2017) and/or spatial autocorrelation. Obviously, a lower alpha level can increase "a Type II error [some say] to unacceptable levels," and because of the magnitude of the resulting r_m-values in the present study (see below) all significant *p*-values are exceptionally low, that is, p = 0.001 to 9.999E-05. Thus, correction was not to be deemed necessary (Nakagawa, 2004:1045; Gelman, Hill, & Yajima, 2012). As above, identical results were achieved with an alternate program, in this case Function Mantel in Vegan R-package 2.5–6 (https://cran.r-project.org; https://github.com/vegandevs/vegan; Oksanen et al., 2013).

3 | RESULTS

3.1 | African analyses

Table 3 lists percentages of expression for the 36 traits in the North and sub-Saharan African dental samples, and the total number of individuals recorded. All 36 ASUDAS breakpoints are provided as well. As mentioned, SNP data in the corresponding African genetic samples are available from the literature (Lazaridis et al., 2014; Patterson et al., 2012; Pickrell & Pritchard, 2012) and via the Gambian Genome Variation Project.

To maximize the comparison, all 36 dental traits were submitted first to the MMD. The resulting inter-sample distances are listed in the bottom diagonal of Table 4. Those from Hudson F_{ST} for the genetic samples are in the top diagonal. To visualize the inter-sample variation, multi-dimensional scaling (MDS) in SPSS 25.0 Procedure Alscal was used to generate three-dimensional spatial representations

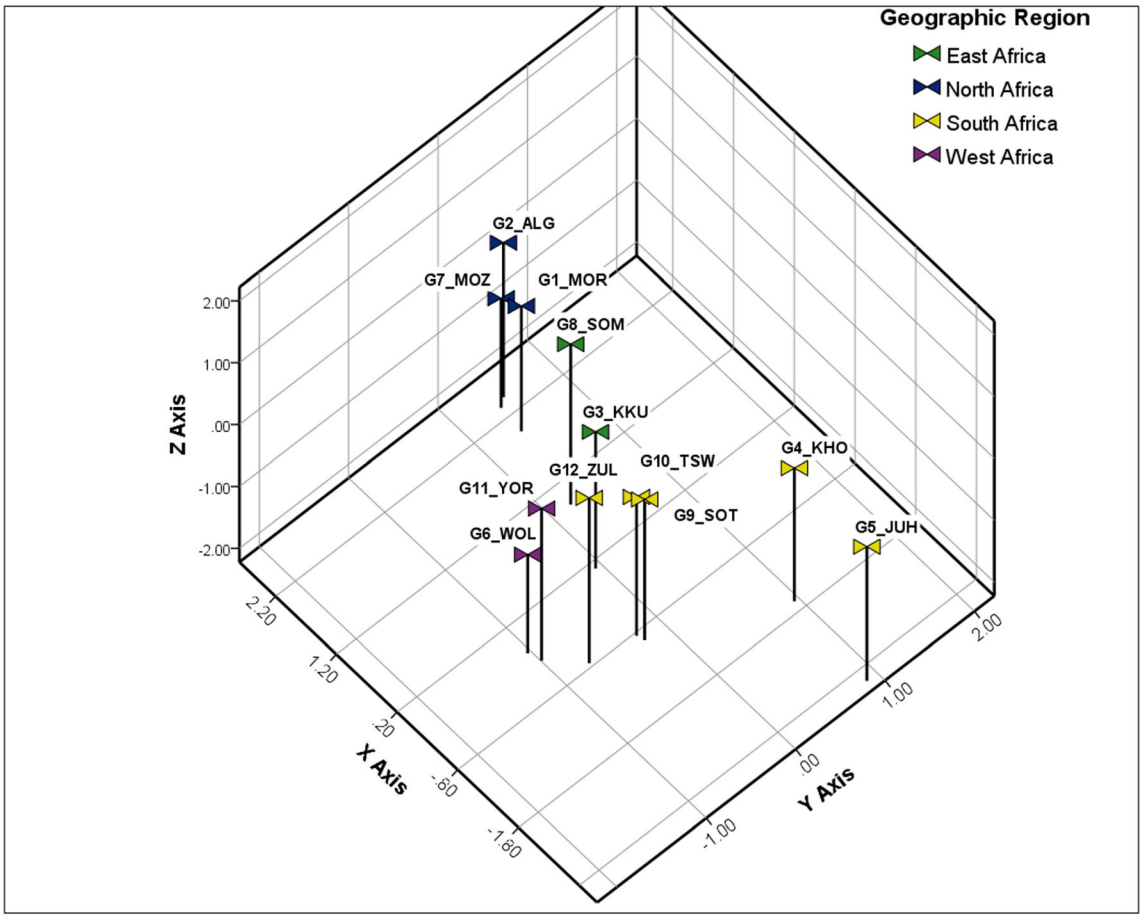

FIGURE 3 Three-dimensional MDS plot of the Hudson F_{ST} distances among the 12 African genetic samples. The sample abbreviations are defined in Tables 2, 4-5

of each matrix (Figures 2 and 3). Interval-level MDS solutions provide good representations of the MMD, with a Kruskal's stress formula 1 value of .099 and r^2 of .943 and, particularly, F_{ST} distances with a stress value of .044 and r^2 of .992. It is apparent that the dental sample locations correspond roughly with geographic origins, where north-to-south is along the X- and west-to-east on the Y-axis (Figure 2). A similar, yet less marked geographic distribution can be seen among the genetic samples (Figure 3).

These qualitative observations are sustained by Mantel test results. The correlation between 36-trait MMD and geographic distances (Table 5) for the dental samples is strongly positive, r_m = 0.682 (p = 0.000). Between F_{ST} and geographic matrices for the 12 matching genetic samples (Table 5) the r_m-value is moderately positive, .486 (p = 0.001). Yet, the r_m-value between the 36-trait MMD and F_{ST} matrices indicates a very strong correlation, .786 (p = 0.000). This correspondence is supported further in controlling for the geographic midpoint distances (Supporting information Table S1); that is, between MMD and F_{ST} residuals the partial Mantel correlation remains strong, r_m = 0.699 (p = 0.000). Scatterplots depicting these correlations are presented in Figure 4.

Next, the 36 ASUDAS traits were reduced to 25, matching those in Turner's database (refer to list in Table 7, below). However, before that, the 12 dental and 12 genetic samples were again compared to quantify any effect that reduced trait number has on the results. So as before, this trait set was submitted to the MMD. The new values are listed in Table 6 and depicted via MDS in Figure 5; for this solution the stress increased slightly (.107), and r^2 decreased (.939). Minor differences are evident between the 36- and 25-trait matrices (compare Tables 5 and 6) and in relative sample locations (Figures 2 and 5). Yet, not unexpectedly, the correlation between the two MMD matrices is almost perfect, r_m = 0.977 (p = 0.000), as seen in Figure 6a. The remaining correlations based on 25 traits all increased: (a) r_m = 0.696 (p = 0.000) for the MMD and geographic distances among dental samples (Figure 6b), (b) r_m = 0.838 (p = 0.000) for MMD and F_{ST} (Figure 6c), and (c) r_m = 0.782 (p = 0.000) for the partial correlation between MMD and F_{ST} residuals, relative to the geographic midpoint distances (Figure 6d).

3.2 | Global analyses

The 25 ASUDAS percentages for the 20 non-African dental samples using the same breakpoints as before are listed in Table 7. To include the maximum number of traits, plus dental samples—to match with

TABLE 5 Geographic distance matrix in km for the three North and nine sub-Saharan African dental samples (bottom diagonal), and for the three North and nine sub-Saharan African genetic samples (top diagonal)

	G1_MOR[a]	G2_ALG	G3_KKU	G4_KHO	G5_JUH	G6_WOL	G7_MOZ	G8_SOM	G9_SOT	G10_TSW	G11_YOR	G12_ZUL	
D1_BED[b]		1,029.17	5,974.47	7,432.69	6,572.97	2,413.47	1,005.29	6,530.46	7,922.45	7,582.08	3,125.61	7,934.43	G1_MOR
D2_KAB	825.32		5,401.40	7,397.40	6,464.35	3,251.16	532.46	5,752.43	7,774.22	7,501.96	3,257.30	7,748.19	G2_ALG
D3_KKU	5,792.78	5,280.23		3,469.45	2,648.97	6,104.37	5,057.71	1,430.34	3,274.61	3,349.88	3,763.84	3,118.40	G3_KKU
D4_RRI	7,769.95	7,621.70	3,423.22		986.52	6,115.56	6,890.49	4,711.68	785.35	286.36	4,312.25	974.67	G4_KHO
D5_SAN	7,073.09	6,886.77	2,738.92	768.10		5,506.25	5,968.31	3,997.92	1,353.41	1,039.61	3,491.84	1,398.06	G5_JUH
D6_SEN	2,460.48	3,111.13	6,051.85	6,645.64	6,128.67		2,877.33	7,171.23	6,792.78	6,342.60	2,349.89	6,885.24	G6_WOL
D7_SHA	1,082.38	301.77	4,993.02	7,407.45	6,664.88	3,225.73		5,537.32	7,289.91	7,003.08	2,725.16	7,272.64	G7_MOZ
D8_SOM	5,970.89	5,282.83	1,543.28	4,830.98	4,209.69	6,886.66	4,981.15		4,350.88	4,542.35	4,913.17	4,156.31	G8_SOM
D9_SOT	7,927.51	7,738.39	3,319.39	332.20	855.29	6,900.96	7,513.18	4,671.07		501.86	4,844.40	224.99	G9_SOT
D10_TSW	7,343.69	7,199.10	3,152.83	426.27	415.69	6,252.34	6,987.20	4,614.04	650.00		4,473.94	688.44	G10_TSW
D11_YOR	3,229.50	3,324.14	3,800.64	4,631.90	4,004.04	2,309.22	3,214.19	4,844.75	4,834.91	4,211.20		4,882.28	G11_YOR
D12_ZUL	8,005.21	7,759.48	3,092.51	752.07	1,004.21	7,132.66	7,520.31	4,368.11	430.16	964.81	4,994.72		G12_ZUL
	D1_BED	D2_KAB	D3_KKU	D4_RRI	D5_SAN	D6_SEN	D7_SHA	D8_SOM	D9_SOT	D10_TSW	D11_YOR	D12_ZUL	

[a]G1_MOR = Moroccan, G2_ALG = Algerian, G3_KKU=Kikuyu, G4_KHO=Khomani San, G5_JUH = Ju-hoan North San, G6_WOL = Wolof, G7_MOZ = Mozabite, G8_SOM = Somalia, G9_SOT = Sotho, G10_TSW = Tswana, G11_YOR = Yoruba, G12_ZUL = Zulu. Details in Table 2.

[b]D1_BED = Bedouin, D2_KAB=Kabyle, D3_KKU=Kikuyu, D4_RRI = Riet River San, D5_SAN=San, D6_SEN=Senegambia, D7_SHA = Shawia, D8_SOM = Somalia, D9_SOT = Sotho, D10_TSW = Tswana, D11_YOR = Yoruba, D12_ZUL = Zulu. Details in Table 1.

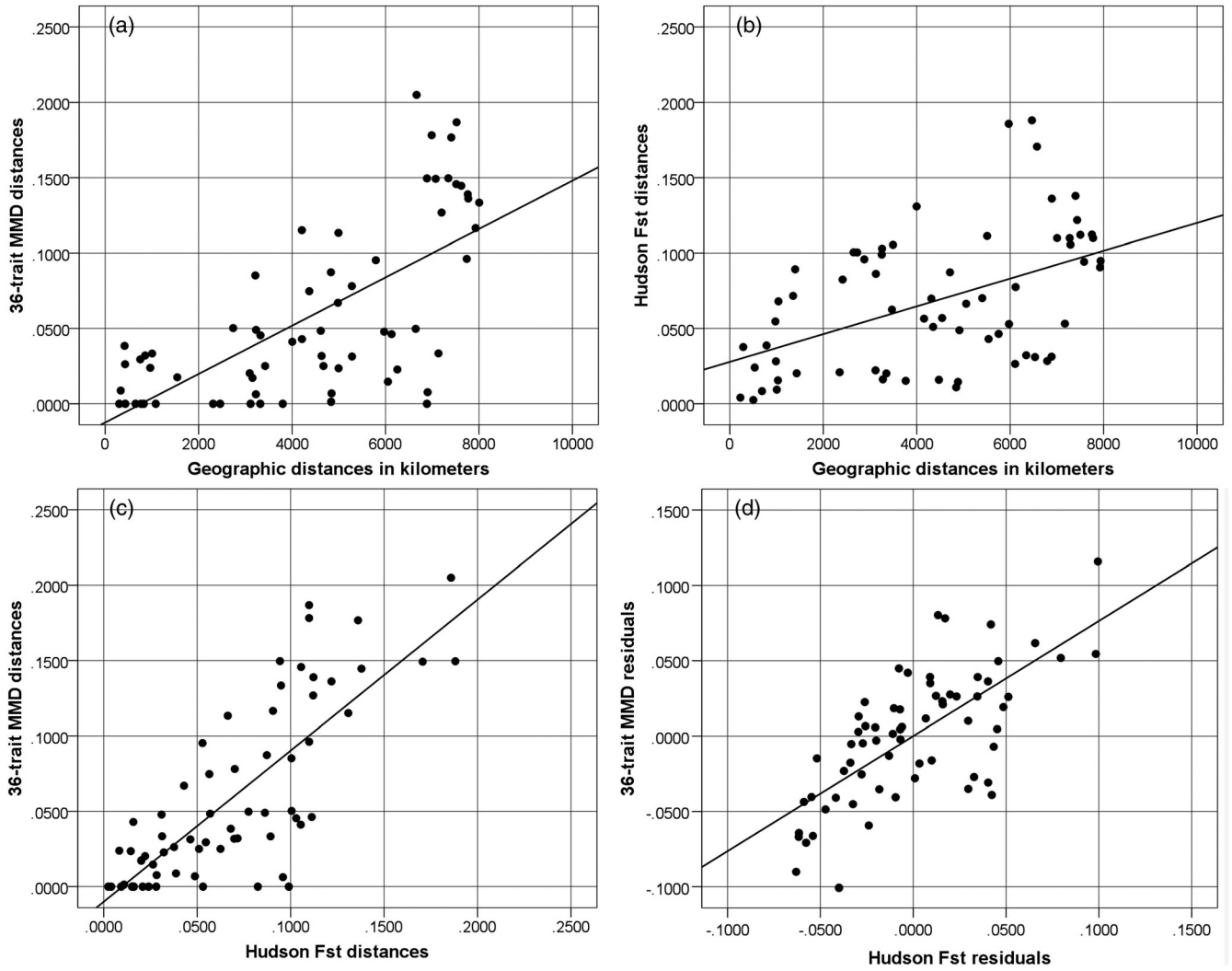

FIGURE 4 Scatterplots of Mantel correlations between matrices from the first African analysis for (a) 36-trait MMD and dental sample geographic distances, (b) Hudson F_{ST} and genetic sample geographic distances, (c) 36-trait MMD and Hudson F_{ST} distances, and (d) 36-trait MMD and F_{ST} residuals, controlling for geographic midpoint distances

the most available genetic sample counterparts, a few individual counts ($n \leq 6$), and/or samples (D17_LOK, D18_CHU, D25_AUN) are unavoidably small. In these cases, results should be interpreted with caution. SNP data for the 20 genetic samples are again from published sources (Lazaridis et al., 2014; Mallick et al., 2016; Patterson et al., 2012; Pickrell & Pritchard, 2012; Skoglund et al., 2016) or were obtained by request.

The 32 x 32 25-trait MMD and Hudson F_{ST} matrices are too large to fit with the main text, so are included in Supplemental Information Tables S2-S3. However, the resulting MDS plots are presented in Figures 7 and 8, respectively. The three-dimensional solutions yielded good resolution for the MMD (stress = 0.088; r^2 = 0.960) and F_{ST} distances (stress = 0.071; r^2 = 0.978). The dental and genetic samples cluster relative to their geographic origins.

The Mantel correlation between MMD and geographic distances for the 32 dental samples (Supporting information Table S4) is strongly positive, r_m = 0.710 (p = 0.000), as it is now also for F_{ST} and genetic sample geographic distances, r_m = 0.768 (p = 0.000) (Supporting information Table S5). The r_m-value between the MMD and F_{ST} matrices again indicates a very strong correlation, .720 (p = 0.000). Unlike the African findings, when controlling for the global geographic midpoint distances (Supporting information Table S6), the partial Mantel correlation between MMD and F_{ST} residuals is only moderately positive, r_m = 0.400 (p = 0.000). Scatterplots depicting these various correlations are provided in Figure 9. Lastly, because the abovementioned dental samples affected most by small numbers are from Turner's database, it was decided to rerun these same Mantel tests with only his 18 samples to explore whether correlations alter substantially. Doing so also serves to quantify if the relatively large number of African samples artificially inflated the r_m-value, while identifying indirectly major inter-observer error with JDI. Although slightly lower, as would be expected because of fewer more geographically limited samples, the Mantel correlations

TABLE 6 MMD distance matrix for the three North and nine sub-Saharan African dental samples based on 25 ASUDAS traits

Sample	D1_BED	D2_KAB	D3_KKU	D4_RRI	D5_SAN	D6_SEN	D7_SHA	D8_SOM	D9_SOT	D10_TSW	D11_YOR	D12_ZUL
D1_BED[a]	0.000											
D2_KAB	0.000	0.000										
D3_KKU	0.093	0.101	0.000									
D4_RRI	0.185	0.213	0.037	0.000								
D5_SAN	0.193	0.232	0.063	0.000	0.000							
D6_SEN	0.000	0.000	0.005	0.080	0.066	0.000						
D7_SHA	0.000	0.000	0.100	0.220	0.254	0.005	0.000					
D8_SOM	0.060	0.040	0.012	0.111	0.134	0.000	0.098	0.000				
D9_SOT	0.133	0.139	0.000	0.010	0.042	0.013	0.165	0.018	0.000			
D10_TSW	0.142	0.164	0.014	0.009	0.033	0.025	0.189	0.037	0.000	0.000		
D11_YOR	0.072	0.090	0.000	0.062	0.066	0.000	0.099	0.010	0.007	0.035	0.000	
D12_ZUL	0.160	0.194	0.010	0.033	0.045	0.043	0.216	0.073	0.000	0.020	0.032	0.000

[a]D1_BED = Bedouin, D2_KAB=Kabyle, D3_KKU=Kikuyu, D4_RRI = Riet River San, D5_SAN=San, D6_SEN=Senegambia, D7_SHA = Shawia, D8_SOM = Somalia, D9_SOT = Sotho, D10_TSW = Tswana, D11_YOR = Yoruba, D12_ZUL = Zulu. Details in Table 1.

remain similar in magnitude to those of the 32-sample analyses (Supporting information Text S1.2).

4 | DISCUSSION

4.1 | Interpretations

Some unpublished data were included (D3_KKU and D11_YOR), but the 36-trait MMD distances (Table 4) and MDS plot of 12 African dental samples (Figure 2) parallel prior results (Irish, 1993, 1997, 1998a, 2010, 2016; Irish et al., 2014). That is, based on documented population history these ASUDAS traits provided reliable phenetic affinities among samples, and patterning indicative of geographic provenience. Reliability is why the ASUDAS has had such a long run in population studies worldwide, as evidenced by hundreds of publications (https://scholar.google.com/) and as summarized in several compendia (Scott & Irish, 2013, 2017; Scott & Turner II, 1997; Scott, Turner II, et al., 2018). Now, however, what had only been assumed from earlier research about the genetic component of trait expression (Scott et al., 1983; Sofaer et al., 1972; Turner II, 1985a), but recently queried (Hughes et al., 2016; Stojanowski et al., 2019; etc.)—that dental affinities approximate genetic relatedness, is readily testable empirically with genomic data (Delgado et al., 2019; Gross & Edgar, 2019; Hubbard et al., 2015; Rathmann et al., 2017). In the present analysis, the MDS plot of F_{ST} distances (Figure 3) is somewhat akin to that from dental data, but the correlation between 36-trait MMD and F_{ST} matrices (Table 4) is most telling (Figure 4; Table 8). Stronger than r_m-values in Hubbard et al. (2015) and Rathmann et al. (2017), it affords additional support for using ASUDAS traits as genetic proxies. So too does the relation between MMD and dental geographic distances (Table 5), and the partial MMD-F_{ST} correlation controlling for geographic midpoint distances (Supplemental Information Table S1). The r_m-value of the latter infers that geographic separation is not an overriding factor in the African dental-genetic correspondence. The lower, yet still moderately positive correlation between F_{ST} and geographic distances (Table 5) may, on this continental level, indicate that F_{ST} does not specifically detect geographic variation/isolation by distance to the same degree of some genetic distance measures (Séré et al., 2017). Conversely, it more likely indicates that F_{ST} is better able to detect increased gene flow between geographically remote populations—particularly since the 19th–20th century dates of the present dental samples, along with extreme reproductive isolation, as in those populations who may be geographically proximate but genetically divergent (Jay, Sjödin, Jakobsson, & Blum, 2012; Ramachandran et al., 2005).

Seemingly contrary to purpose, reducing the number of dental traits from 36 to 25 (Table 6) increased the r_m-values between MMD and geographic distances, and MMD and F_{ST} distances among the African samples. The partial MMD-F_{ST} correlation controlling for geographic midpoint distances increased most (Figure 6; Table 8). This is all despite the minimal change in 36- and 25-trait distances, as indicated by a correlation near 1.0 and similitude in MDS configurations

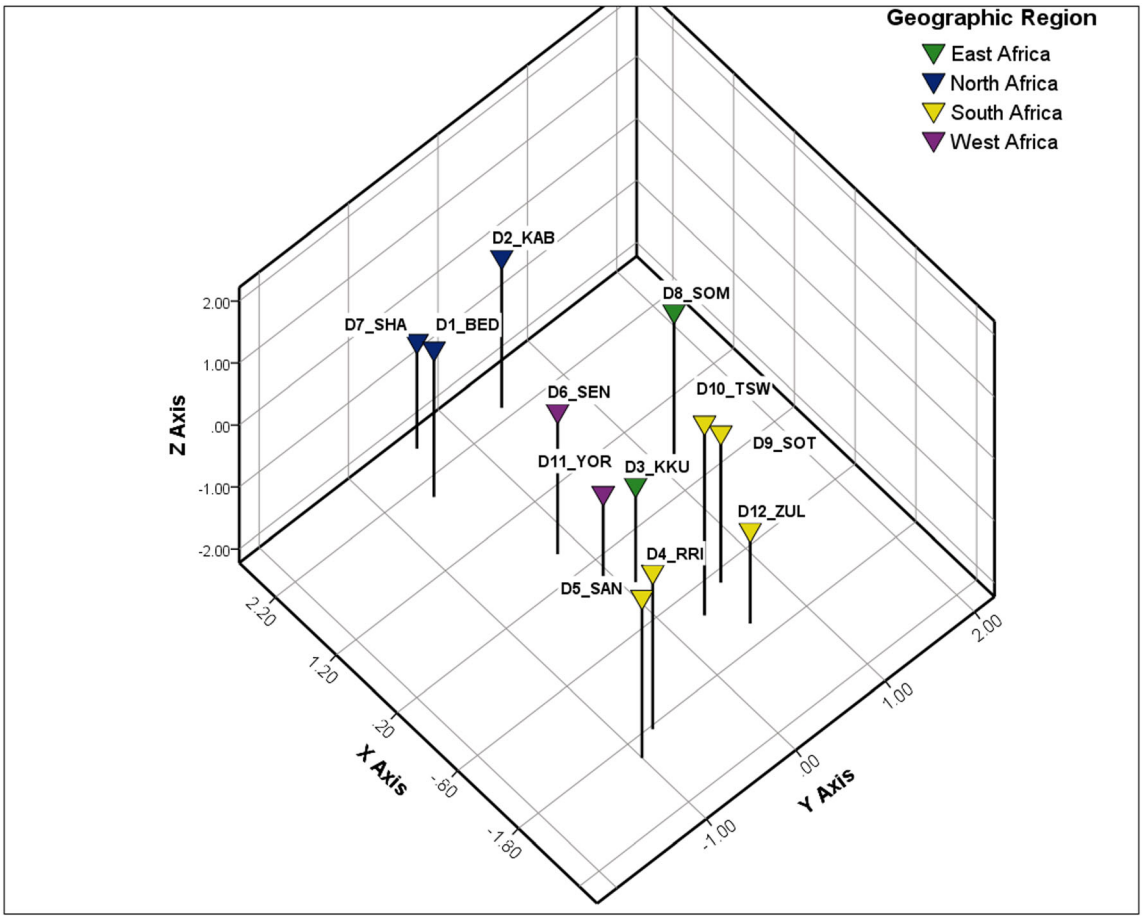

FIGURE 5 Three-dimensional MDS plot of the 25-trait MMD distances among the 12 African dental samples. The sample abbreviations are defined in Tables 1, 3-6

(Figure 5). Deleting the 11 traits essentially functioned to emulate the editing process typically used prior to submitting data to the MMD and other similar statistics (Irish, 2010). That is, reliable results are attainable with problematic traits as noted, but it is prudent to delete the same. If this study focused only on ASUDAS-based affinities (Irish, 2005, 2006, 2016; etc.), then nine of these traits would have been deleted in any event as standard practice, for being: 1) mostly invariant (palatine torus, UI2 peg-reduced, mandibular torus), 2) otherwise unimportant for driving inter-sample variation based on low loadings (<.5) in principal component analysis (UC distal accessory ridge, rocker jaw), and 3) highly inter-correlated ($\tau_b \geq |0.5|$) with other traits (UI1 labial curvature; LM1 anterior fovea, LM1 deflecting wrinkle, LM1 C1-C2 crest). Thus, while one idea was to include more traits than similar studies to maximize comparative analyses, it is apparent here that "more" is not always "better."

Finally, for the global-level analyses, the MDS plot of 32 dental samples (Figure 7) based on 25-trait MMD distances (Supporting information Table S2), and the matching plot (Figure 8) of the F_{ST} matrix (Supporting information Table S3) are comparable. Explicitly, from sub-Saharan Africa to the Americas and Melanesia, the samples cluster by regional origin and evidence overall geographic patterning. The latter is quantified by correlations >.7 (Table 8; Figure 9a,b) for both MMD and F_{ST} with dental and genetic sample geographic distances (Supporting information Table S4-S5). As first suggested >20 years ago (Irish, 1997:463), the MMD variation depicted in Figure 7 may be detecting the vestiges of "an expansive dental morphological cline." Recent sub-Saharan Africans were revealed to possess high frequencies of ancestral dental nonmetric traits, while other world populations transition toward higher frequencies of derived traits with increasing geographic separation (Irish, 1998b; Irish & Guatelli-Steinberg, 2003). The idea was later revisited using other dental data (Hanihara, 2013; Reyes-Centeno, Rathmann, Hanihara, & Harvati, 2017). Therefore, if beginning with the two samples of southern African San (D4_RRI, D5_SAN) in the plot, this presumed cline can be seen to extend from sub-Saharan to North Africa and into the Mediterranean region and farther north in Europe. Toward what would be east on the right side of the X-axis, are Northeast and Southeast Asians (also see Jay et al., 2012). The New World samples (D13_PIM, D14_ALU) are then farther to the north and east, and Australians (D25_AUN) and Melanesians (D26_NBR) south, near the bottom of the Y-axis. Naturally, neutral genomic data are increasingly being used to investigate ancient migrations, including routes out of Africa (e.g., (Jay et al., 2012; Kanitz, Guillot, Antoniazza, Neuenschwander, & Goudet, 2018;

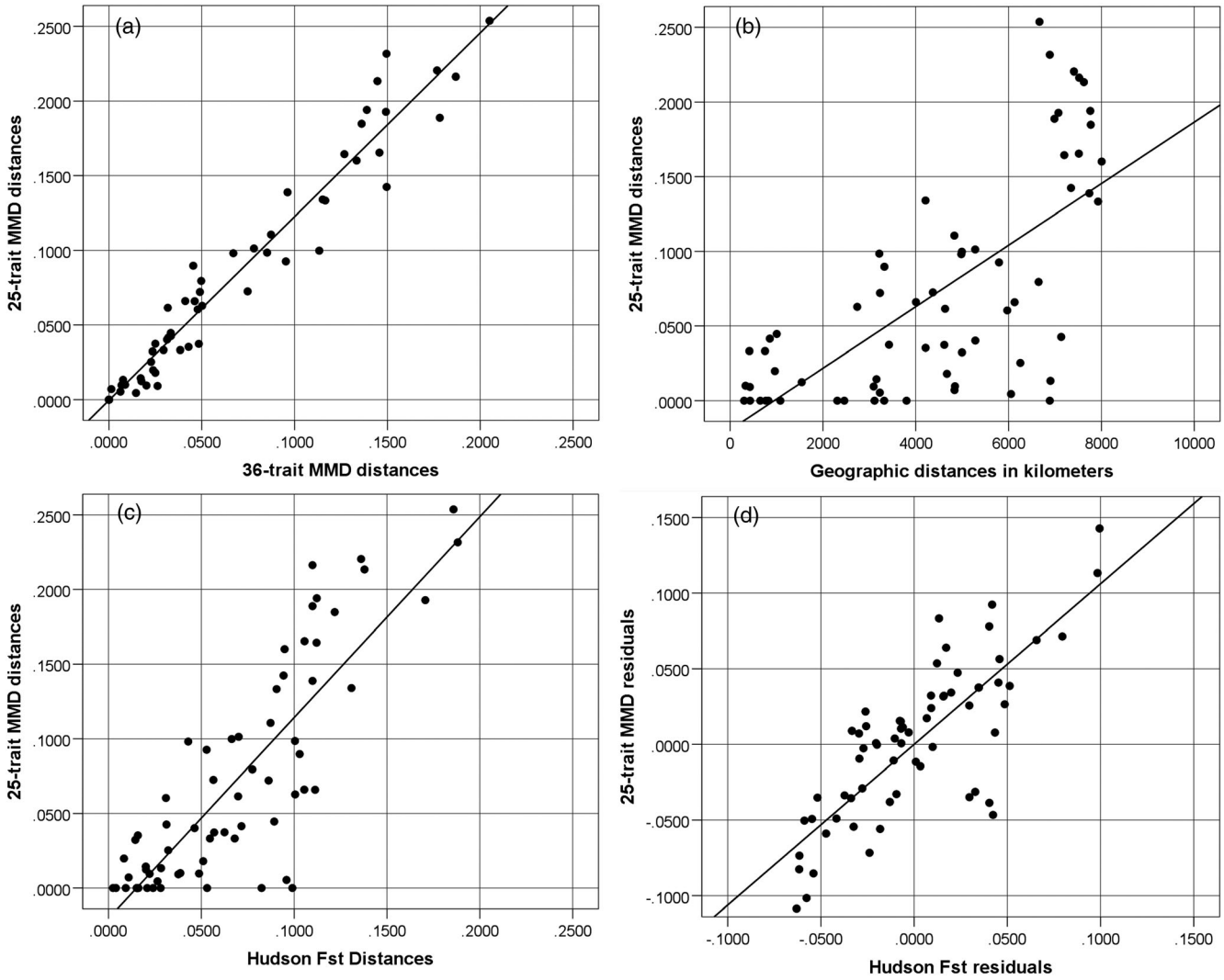

FIGURE 6 Scatterplots of Mantel correlations between matrices from the second African analysis for (a) 25- and 36-trait MMD distances, (b) 25-trait MMD and dental sample geographic distances, (c) 25-trait MMD and Hudson F_{ST} distances, and (d) 25-trait MMD and F_{ST} residuals, controlling for geographic midpoint distances

Ramachandran et al., 2005). And here, though a more linear trajectory is illustrated from F_{ST} distances (Figure 8), signs of such a cline are sustained—again starting with the southern African San samples (G4_KHO, G5_JUH) and ending with Australians (G25_AUN) and Melanesians (G26_NBR).

Considerably expanded from continental to global in scale with an additional 20 samples, the r_m-value between 25-trait MMD and F_{ST} matrices is still >.70 (Table 8; Figure 9c)—despite the caution that results may be influenced by low counts or small samples. Thus, for these 32 dental samples and MMD distances from ASUDAS data, and these 32 genetic samples and F_{ST} distances from SNP data, the correlation further supports use of dental traits for population affinity research, with or instead of neutral genomic data if the latter are unavailable. The higher correlation between F_{ST} and geographic distances relative to the African results is likely linked to the expanded scale. Irrespective of whether F_{ST} is or is not unequivocally suited to detect isolation by distance (above), it is perhaps picking up on ancient among-region affinities rather than that of recent within-region (or -continent) population movements and genetic exchange. This possibility is also likely why the global partial MMD-F_{ST} correlation controlling for geographic midpoint distances (Supplemental Information Table S6; Figure 9d), though positive (Table 8), is much lower than the African r_m-value; geographic separation does appear to be a contributing factor to the overall dental-genetic association in this broad-scale example.

4.2 | Implications

The results cannot be equated directly with ancestry estimates in individuals (Delgado et al., 2019; Gross & Edgar, 2019) but anecdotally, and at least relative to some population-level analyses also presented in these two articles, the present MMD-F_{ST} correlations suggest similar if not greater correspondence of dental and genomic data. More direct comparisons are possible, as above, with articles describing

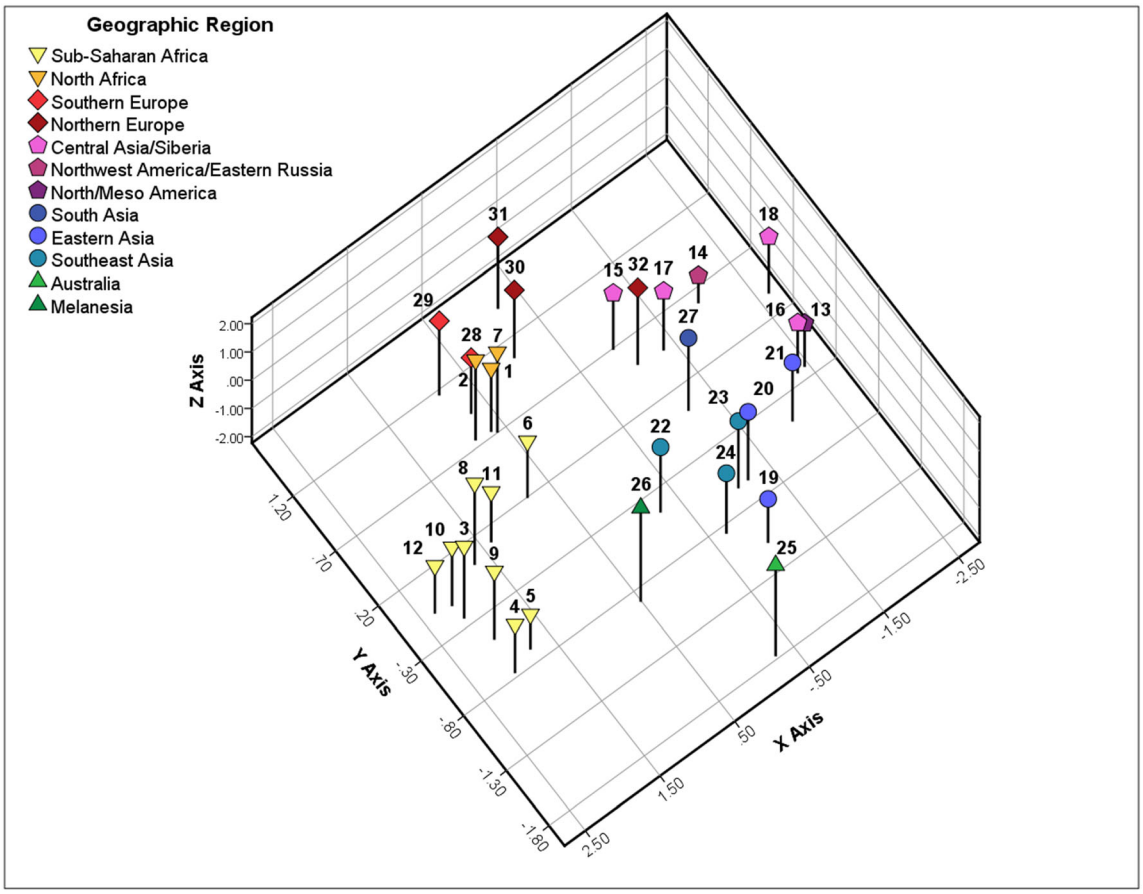

FIGURE 7 Three-dimensional MDS plot of the 25-trait MMD distances among the 32 global dental samples. The sample abbreviations are defined in Tables 1, 3-7

outcomes on regional and global levels (Hubbard et al., 2015; Rathmann et al., 2017); the current Mantel correlations are decidedly stronger. Why? And what is the significance?

Why? In answer, it is stressed that the intent is to explore potential reasons for the enhanced dental-genetic correspondence, not to critique prior studies (Delgado et al., 2019; Gross & Edgar, 2019; Hubbard et al., 2015; Rathmann et al., 2017) that provide the foundation for this research. Multiple explanations are possible vis-à-vis differences in data, samples, and methods. First, the numbers of traits, 36 and 25, are larger than in three of the four studies, and unlike all four include root and root-related traits that are highly diagnostic in characterizing populations (Scott & Irish, 2017; Scott & Turner II, 1997; Scott, Turner II, et al., 2018; Turner II et al., 1991). Moreover, when the 11 traits were dropped from analyses, testing revealed that nine are problematic for the specific 32 samples under study, including four that are highly inter-correlated; this "editing" further enhanced the dental data for comparative purposes. And, dental data were recorded by Turner and JDI. Though standardized, the ASUDAS is not intuitive to the point where trait recording can be undertaken without requisite training, including quantification of inter-observer error (Scott & Irish, 2017); examples of suspect affinities and misidentified traits illustrate this potential issue (e.g., Irish & Morris, 1996). With regard to the >350,000 SNPs, among other attributes these genomic data have been shown to better differentiate among populations, and they appear to correspond more closely with dental traits than, for example, STRs (Gross & Edgar, 2019; Rathmann et al., 2017). These markers are also substantially greater in number than in three of the recent studies, including 1,718 SNPs in Rathmann et al. (2017).

Second, 32 dental samples consisting of 2,844 individuals were compared with 32 genetic samples comprised of 530 individuals—all considerably more than the previous studies. But most importantly, though not the same individuals (Hubbard et al., 2015), creation of the dental samples and matching them with their genetic counterparts were less subjective. To illustrate, given the focus of one admixture study, recent African Americans, the authors were obliged to use casts of individuals whose "race was assigned by the orthodontists who assembled the collections" (Gross & Edgar, 2019:522). For global population comparisons, Rathmann et al. (2017) did not have the present luxury of choice among dental samples, so several of their matched pairs are characterized by ethnic, as well as significant linguistic differences, including (a) dental sample Kikuyu of Niger-Congo Language Superfamily versus a genetic sample with some Masai and Luo of Nilo-Saharan Superfamily, (b) dental Zulu of Nguni Branch vs. genetic

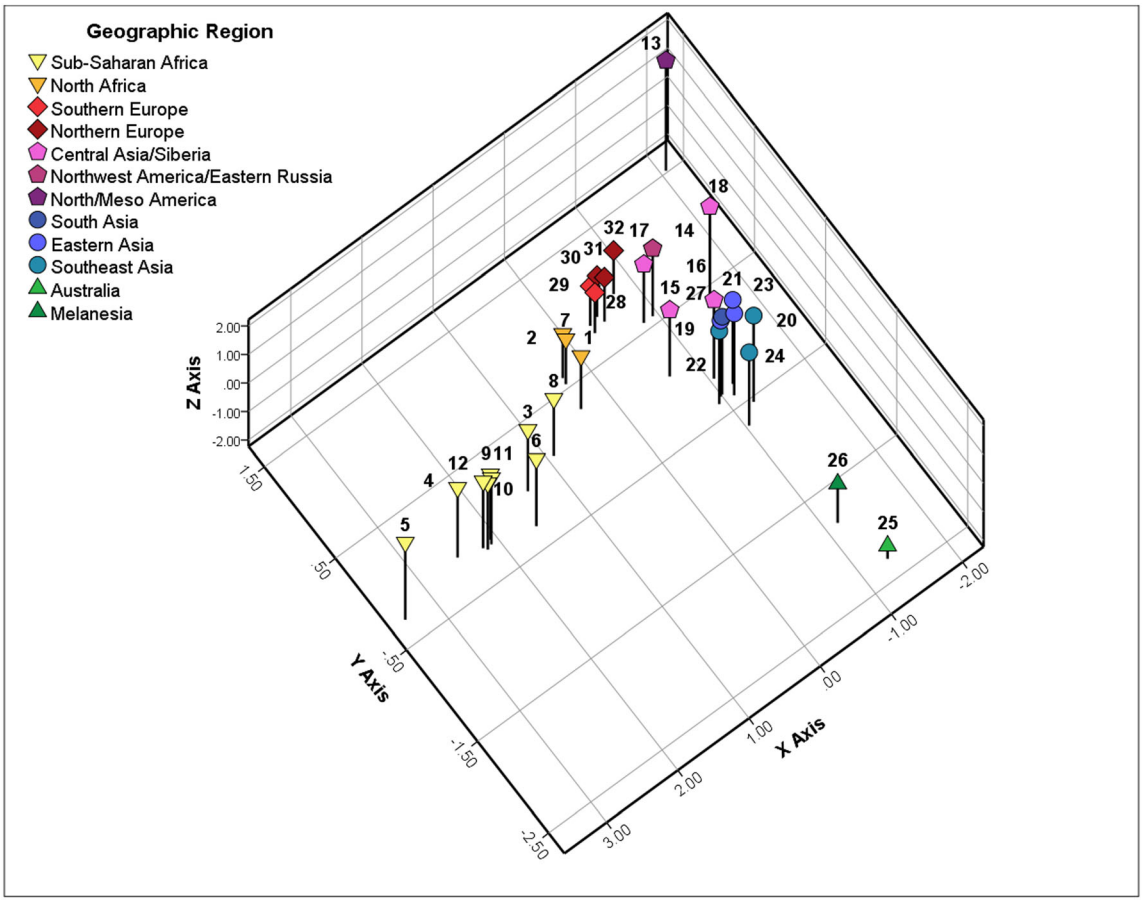

FIGURE 8 Three-dimensional MDS plot of the Hudson F_{ST} distances among the 32 global genetic samples. The sample abbreviations are defined in Tables 2, 4-5

Pedi, Sotho, and Tswana in the Sotho Branch of the Bantu Language Family, and (c) dental Haya of Niger-Congo Superfamily vs. genetic sample with Hadza who, if not a language isolate belong to the Khoisan Superfamily (Greenberg, 1963). Further, the overall lack of matches for the Americas necessitated their pairing of a recent dental sample of varied Mexican ethnicities with a genetic sample of archeological specimens, including 2,500 year-old Zapotec (Rathmann et al., 2017).

Third, methodological differences are likely a key factor in correspondence of dental and genomic data, especially their performance in estimating individual ancestry (Delgado et al., 2019; Gross & Edgar, 2019) relative to population affinities (Hubbard et al., 2015; Rathmann et al., 2017; this study). Again, comparing results between these two types of study is impractical, but another matter is the nature of ASUDAS data. While progress has been made using dental nonmetric traits to estimate individual ancestry and affinities (Edgar, 2013, 2015; Irish, 2015; Scott, Pilloud, et al., 2018; Stojanowski & Paul, 2015; Stojanowski & Schillaci, 2006), the ASUDAS was designed to analyze samples and variation therein (Scott & Turner II, 1997; Turner II et al., 1991); thus, the "low predictive power for genetic ancestry of individuals" compared with SNPs is not surprising (Delgado et al., 2019:439). In the Gross and Edgar (2019) study, as recognized, a contributing factor is the genetic program used to estimate ancestry, which is affected by missing data; thus, the poorer performance of dental traits is also related to this issue relative to more complete genomic datasets. Concerning population studies (Hubbard et al., 2015; Rathmann et al., 2017), the justification for instead using MMD and F_{ST} distances in a model-free approach, of course with much larger datasets, has already been discussed. This too may have played a role in stronger Mantel correlations calculated here.

So, what are the implications of this study? It did not address directly questions concerning such perceived attributes of the ASUDAS as low sexual dimorphism, minimal selection, and high heritability in trait expression. These matters have been dealt with elsewhere. Sexual dimorphism could relate simply to tooth size where, for example, larger crowns in males may promote the formation of additional, later-forming cusps, all other developmental factors being equal (Jernvall & Jung, 2000). That said, prior dental nonmetric studies found few or no statistically significant differences for cusp number or other trait by sex (Bermudez de Castro, 1989; Hanihara, 1992; Irish, 1993; Smith & Shegev, 1988). Significant differences that may occur appear random, in that different traits are affected among studies depending on the population (Irish, 2016); for example, it was a factor in the Gullah

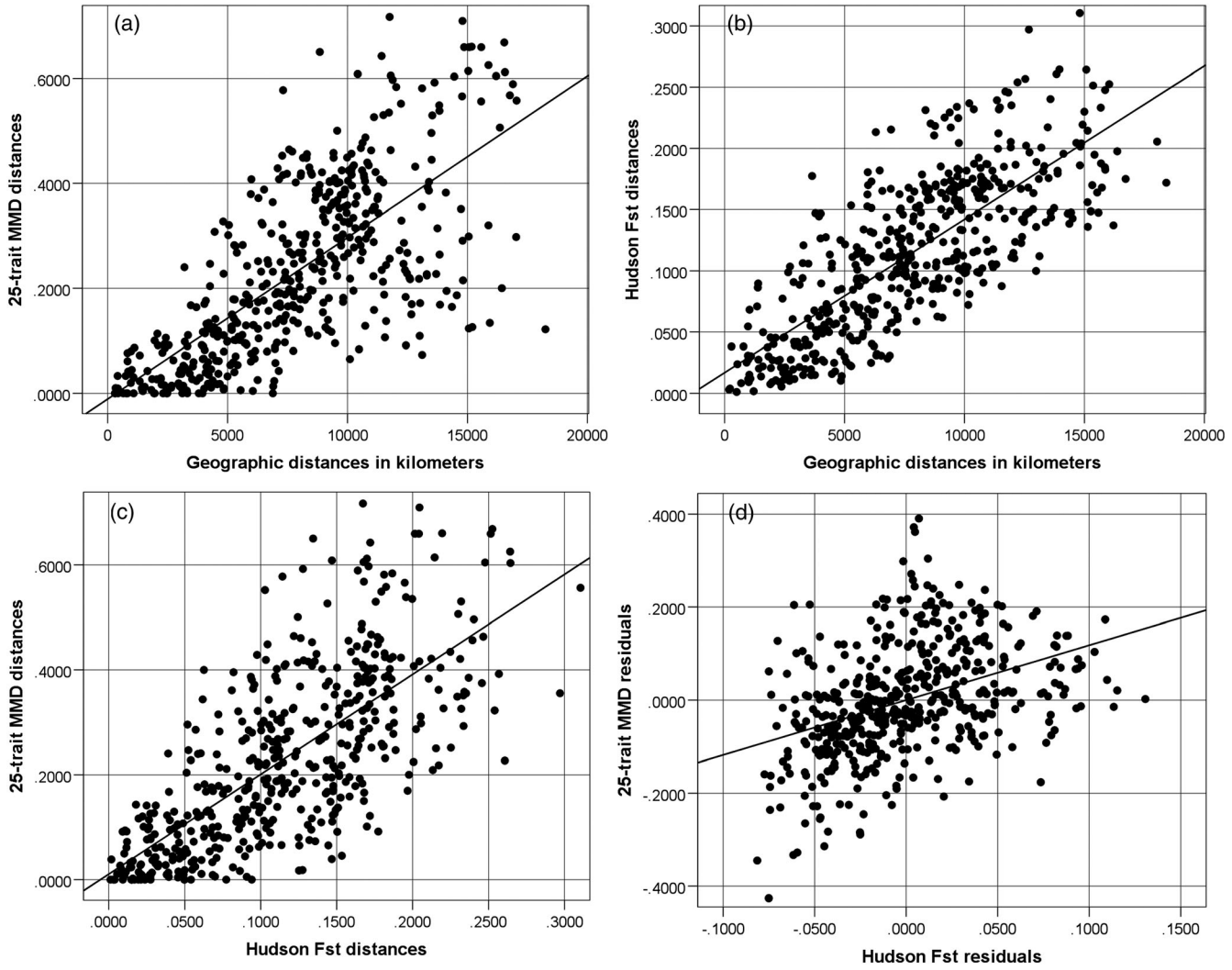

FIGURE 9 Scatterplots of Mantel correlations between matrices from the global analysis for (a) 25-trait MMD and dental sample geographic distances, (b) Hudson F_{ST} and genetic sample geographic distances, (c) 25-trait MMD and Hudson F_{ST} distances, and (d) 25-trait MMD and F_{ST} residuals, controlling for geographic midpoint distances. See text for details

heritability paper (Stojanowski et al., 2019), though not so much in the most recent Australian study (Paul et al., 2020). Concerning selection, it was established to impact some traits (Bryk et al., 2008; Kimura et al., 2009; Park et al., 2012; Hlusko et al., 2018), but indirectly as a consequence of pleiotropy. Recent trait variation, at any rate, is seemingly more "a product of random processes (i.e., genetic drift and founder effect) rather than genetic adaptation" (Scott, Turner II, et al., 2018: 223). Lastly, trait heritability was shown to vary across studies and dental fields. The effects of various stressors and other issues on the populations under study may play a role, along with methodological factors like appropriate dichotomization of traits (Higgins et al., 2009; Hughes et al., 2016; Hughes & Townsend, 2011, 2013; Paul et al., 2020; Riga et al., 2014; Stojanowski et al., 2018, 2019). However, the exact answers will require additional research, which is beyond the present scope of study.

Of course, dental and genetic data correspondence was addressed, and they do correspond to a much greater degree than before, based on comparing distances calculated from them. The r_m-values, however, do not approach 1.0 (Table 8). Consequently, under the assumption that neutral genomic markers are indeed the definitive choice in affinity study, dental nonmetric traits are not about to supplant them. That said, these correlations may be considered minimum values (also see Rathmann et al., 2017). The data were not collected from the same individuals, and though able to match at a high level of concordance, the paired dental and genetic samples are of different ages and several come from similar, not identical, populations. Some sample size issues, inter-observer error in the ASUDAS data, and other stochastic and nonstochastic factors (Rathmann et al., 2017), also cannot be ruled out totally. In any event, the results speak for themselves. If sufficient attention is paid to the data, samples, and methods, it does appear ASUDAS traits can serve as highly reliable proxies for neutral genomic markers—regardless of potential issues mentioned, including sexual dimorphism, selection, and heritability. If the latter could be identified on individual trait and/or sample bases, which was not the case here, data correspondence may conceivably be even higher.

TABLE 7 Percentages (%) of 25 dental traits considered present and number of individuals scored (n) in the 20 additional (non-African) global samples

Trait/grades present		D13 _PIM[a]	D14 _ALE	D15 _KAZ	D16 _MON	D17 _LOK	D18 _CHU	D19 _THA	D20 _VIE	D21 _JAP	D22 _MAL	D23 _PHI	D24 _BOR	D25 _AUN	D26 _NBR	D27 _NEP	D28 _GRK	D29 _ITY	D30 _KBR	D31 _FIN	D32 _LAP
Winging UI1	%	25.5	21.8	14.6	25.0	28.6	37.0	27.6	36.4	23.3	4.2	18.2	28.6	28.6	23.2	33.3	1.5	2.1	4.3	4.3	54.5
(+ = ASU 1)[b]	n	161	78	48	28	7	27	76	11	86	24	11	14	7	155	12	68	48	70	23	11
Shoveling UI1	%	84.2	53.3	27.9	71.5	40.0	61.5	32.5	55.6	71.9	28.6	47.4	29.4	12.5	8.9	18.8	0.0	5.9	8.7	9.5	16.7
(+ = ASU 3–6)	n	159	75	43	21	5	13	80	9	89	14	19	17	8	124	16	5	17	69	21	6
Double shoveling UI1	%	38.5	14.5	9.4	7.7	33.3	25.0	7.7	22.2	13.2	8.3	14.3	11.8	0.0	3.3	5.9	0.0	0.0	19.4	23.8	16.7
(+ = ASU 2–6)	n	156	76	43	13	3	8	65	9	83	12	7	17	5	123	17	5	17	67	21	6
Interruption groove UI2	%	54.7	46.5	32.8	36.4	33.3	46.2	33.7	25.0	44.6	35.3	27.8	30.0	0.0	19.2	25.0	35.0	17.7	31.0	45.8	45.0
(+ = ASU +)	n	117	71	67	22	6	26	86	8	83	17	18	20	5	146	12	20	17	71	24	20
Tuberculum dentale UI2	%	38.2	31.5	22.1	24.3	33.3	33.3	21.2	11.1	9.6	33.4	38.9	22.6	20.0	36.8	16.7	5.3	50.0	17.6	12.5	33.5
(+ = ASU 2–6)	n	141	73	68	33	6	21	85	9	83	18	18	22	5	144	12	19	18	74	24	21
Bushman canine UC	%	0.0	0.0	2.4	3.7	0.0	0.0	9.3	0.0	4.7	4.3	0.0	11.1	0.0	1.9	0.0	8.7	0.0	2.7	0.0	4.2
(+ = ASU 1–3)	n	96	69	85	27	9	24	97	19	86	23	24	27	7	154	24	23	23	73	29	24
Hypocone UM2	%	82.1	38.4	73.4	68.5	73.0	63.6	87.3	80.8	72.7	80.9	84.4	89.8	93.1	87.4	75.4	50.0	59.6	66.0	65.7	54.0
(+ = ASU 3–5)	n	112	73	124	54	26	55	157	52	99	42	45	79	29	214	73	54	47	94	35	37
Cusp 5 UM1	%	5.4	1.2	3.0	15.0	12.9	0.0	23.2	37.5	12.2	35.0	23.9	6.6	77.0	59.0	6.3	5.7	23.3	3.2	0.0	12.9
(+ = ASU 2–5)	n	94	82	101	40	31	21	112	40	90	40	46	60	13	183	64	53	43	94	22	31
Carabelli's trait UM1	%	47.8	11.5	31.4	24.0	23.1	10.4	39.9	30.6	27.0	41.8	40.4	44.2	33.4	37.8	27.4	58.3	51.2	40.2	40.9	14.7
(+ = ASU 3–7)	n	157	78	118	54	39	48	143	49	100	43	47	70	15	207	66	48	43	87	22	34
Parastyle UM3	%	0.0	0.0	0.0	0.0	0.0	0.0	1.7	2.7	2.1	2.9	0.0	5.2	3.3	2.0	0.0	0.0	0.0	0.0	4.5	3.7

TABLE 7 (Continued)

Trait/grades present		D13 _PIM[a]	D14 _ALE	D15 _KAZ	D16 _MON	D17 _LOK	D18 _CHU	D19 _THA	D20 _VIE	D21 _JAP	D22 _MAL	D23 _PHI	D24 _BOR	D25 _AUN	D26 _NBR	D27 _NEP	D28 _GRK	D29 _ITY	D30 _KBR	D31 _FIN	D32 _LAP
(+ = ASU 3–5)	n	26	35	79	38	8	27	116	37	48	35	31	39	30	151	33	33	28	66	22	27
Enamel extension UM1	%	59.2	55.1	40.6	83.3	69.1	78.1	57.8	66.1	71.3	52.1	74.5	63.6	62.1	21.5	62.7	16.7	9.3	24.1	16.7	47.3
(+ = ASU 1–3)	n	27	49	165	66	42	96	166	62	108	48	51	85	37	195	91	54	43	116	30	55
Root number UP1	%	14.3	10.6	21.2	17.2	12.2	13.1	33.9	32.3	24.7	62.0	32.7	46.6	50.0	49.2	31.9	61.9	59.5	37.0	48.6	27.8
(+ = ASU 2 +)	n	28	47	170	58	41	115	168	62	93	50	49	133	38	175	94	63	37	100	35	61
Root number UM2	%	54.5	63.3	36.9	40.7	38.2	55.2	79.9	81.0	71.3	70.5	72.0	81.9	70.6	81.0	61.6	58.3	81.8	58.8	52.0	44.9
(+ = ASU 3 +)	n	22	30	141	64	34	58	144	58	94	44	50	105	34	184	86	36	22	85	25	49
Odontome P1-P2	%	8.6	3.2	1.9	2.5	11.1	0.0	1.3	2.9	7.8	0.0	5.3	2.7	0.0	1.8	1.9	0.0	2.2	3.3	0.0	0.0
(+ = ASU +)	n	128	94	107	41	9	21	152	34	103	45	38	37	13	168	54	44	46	61	36	27
Congenital absence UM3	%	10.3	30.6	29.5	33.4	44.2	18.3	15.9	13.6	39.1	18.4	8.5	25.0	8.1	7.7	30.0	17.7	20.0	19.8	11.4	13.0
(+ = ASU -)	n	29	62	156	72	43	104	182	66	110	49	47	104	37	207	90	68	50	111	44	54
Lingual cusp LP2	%	34.2	40.3	71.4	40.0	50.0	25.0	69.7	54.5	63.3	81.8	71.4	83.8	100.0	90.7	58.3	60.0	27.6	65.1	48.5	66.6
(+ = ASU 2–9)	n	120	72	105	15	2	8	119	22	75	33	28	37	11	172	24	10	29	63	33	45
Groove pattern LM2	%	27.2	24.1	20.7	16.0	19.2	15.4	19.6	26.1	38.7	38.5	31.8	25.5	30.4	43.1	15.4	43.5	27.9	28.4	13.8	22.2
(+ = ASU Y)	n	82	54	140	25	26	13	148	46	75	39	44	55	23	211	52	23	43	88	29	45
Cusp number LM1	%	46.0	26.2	12.9	29.4	3.2	40.0	36.5	29.4	34.8	34.8	35.1	34.9	77.8	52.9	33.3	0.0	2.6	5.9	5.6	20.0
(+ = ASU 6 +)	n	150	65	116	17	31	5	96	34	69	23	37	43	9	170	39	19	38	85	18	30

(Continues)

TABLE 7 (Continued)

Trait/grades present		D13_PIM[a]	D14_ALE	D15_KAZ	D16_MON	D17_LOK	D18_CHU	D19_THA	D20_VIE	D21_JAP	D22_MAL	D23_PHI	D24_BOR	D25_AUN	D26_NBR	D27_NEP	D28_GRK	D29_ITY	D30_KBR	D31_FIN	D32_LAP
Cusp number LM2	%	84.1	82.4	41.4	76.2	47.8	62.2	78.4	66.7	68.1	66.7	66.7	75.0	81.2	50.3	50.1	47.6	37.1	27.6	17.9	38.3
(+ = ASU 5 +)	n	107	68	121	21	23	8	134	45	72	39	42	52	16	193	48	21	35	87	28	39
Protostylid LM1	%	24.4	7.6	1.6	19.1	0.0	0.0	0.7	0.0	13.0	0.0	2.4	8.1	0.0	0.0	1.9	0.0	0.0	0.0	0.0	2.4
(+ = ASU 3–6)	n	148	66	125	21	38	10	135	39	77	31	42	49	18	45	52	19	38	68	21	38
Cusp 7 LM1	%	6.8	3.8	2.0	0.0	10.5	0.0	2.0	4.5	1.2	3.1	7.0	11.6	0.0	9.1	3.6	5.6	2.6	2.0	0.0	0.0
(+ = ASU 2–4)	n	162	78	149	21	38	12	147	44	82	32	43	60	12	221	56	18	38	99	22	44
Tome's root LP1	%	20.0	0.0	14.1	14.3	33.3	0.0	9.9	37.5	34.9	25.8	42.9	45.1	26.1	14.3	36.9	7.1	13.2	13.9	10.0	8.8
(+ = ASU 3–5)	n	30	27	163	7	6	6	91	16	43	31	21	31	23	126	19	28	38	101	40	45
Root number LC	%	0.0	0.0	3.4	0.0	0.0	0.0	0.6	0.0	1.7	0.0	2.3	1.7	0.0	0.0	3.4	3.5	5.1	10.6	2.3	3.2
(+ = ASU 2 +)	n	38	20	175	14	41	17	157	49	58	30	44	58	29	130	59	29	39	104	44	62
Root number LM1	%	19.2	53.1	13.9	25.0	2.4	50.0	5.9	15.4	24.2	10.0	21.6	14.1	9.1	3.6	27.9	0.0	0.0	0.9	0.0	0.0
(+ = ASU 3 +)	n	26	49	180	28	42	20	186	52	95	40	51	85	33	193	61	22	25	116	30	60
Root number LM2	%	89.5	72.5	57.8	52.0	65.9	55.6	18.9	67.9	69.4	73.2	68.0	79.5	90.6	95.6	72.9	91.3	100.0	79.4	84.0	63.3
(+ = ASU 2 +)	n	19	40	159	25	41	18	180	53	85	41	50	73	32	184	59	23	38	92	25	60

[a]D13_PIM = Pima, D14_ALE = Aleut, D15_KAZ = Kazak, D16_MON = Mongol, D17_LOK = Lower Ob Khanty, D18_CHU=Chukchi, D19_THA = Thailand, D20_VIE = Tonkin-Annam Viet Nam, D21_JAP = Japanese, D22_MAL = Malay, D23_PHI=Philippines, D24_BOR = Borneo, D25_AUN = Australia North, D26_NBR = New Britain, D27_NEP=Nepal, D28_GRK = Greek, D29_ITY=Italy, D30_KBR = Kaberla, D31_FIN = Ladoga Finns, D32_LAP = Lapps. Details in Table 1.

[b]ASUDAS rank-scale trait breakpoint information in Irish (1993, 1997, 2005, 2006), Scott & Turner II (1997), Scott & Irish (2017) and Scott, Turner II, et al., 2018.

TABLE 8 Mantel and partial Mantel results summary

Analysis:		Africa 1 36 traits	Africa 2 25 traits	Global 25 traits
Matrices compared:				
MMD + DGeo	r_m=	.682	.696	.710
	p=	.000	.000	.000
F_{ST} + GGeo	r_m=	.486	.486	.768
	p=	.001	.001	.000
MMD + F_{ST}	r_m=	.786	.838	.720
	p=	.000	.000	.000
MMD + F_{ST} – DGGeo	r_m=	.700	.782	.400
	p=	.000	.000	.000

Abbreviations: DGeo, matrix of geographic distances among dental samples; DGGeo, matrix of geographic midpoint distances for matching dental and genetic samples; F_{ST}, Hudson F_{ST} distance matrix; GGeo, matrix of geographic distances among genetic samples; MMD, mean measure of divergence matrix.

The capability of using dental traits as proxies is not insignificant given the cost, destructive sampling, and processing time of genetic analyses. More critically, the traits can be substituted if DNA and, more likely, ancient DNA is not recoverable. Specifically, degradation is of particular concern in tropical and sub-tropical environments, like Africa, Southeast Asia, and other equatorial regions; in more temperate climates time is also a factor, though to a lesser degree, for example, negatively affecting recovery in specimens of Middle Pleistocene age and older (Kistler, Ware, Smith, Collins, & Allaby, 2017; Pinhasi et al., 2015; Smith, Chamberlain, Riley, Stringer, & Collins, 2003). In the latter instance, "paleoanthropologists [currently do] consider teeth the "safe box" of the genetic code" (Martinón-Torres et al., 2007:7), as evidenced by many studies using dental nonmetric traits (Bailey & Hublin, 2013; Bailey, Weaver, & Hublin, 2017; Irish & Guatelli-Steinberg, 2003; Irish, Guatelli-Steinberg, Legge, de Ruiter, & Berger, 2013; Martinón-Torres et al., 2007, 2008, 2013; among others). So, while there is no assurance heritability estimates in recent humans apply to our Plio-Pleistocene ancestors, these dental traits are likely as close as possible to genomic data for determining hominin origins and affinities (also see Irish, Bailey, Guatelli-Steinberg, Delezene, & Berger, 2018).

Lastly, while genomic markers are largely seen as the definitive data for population studies some caution, like that with the dental traits, should be exercised in choice and interpretation. For example, one limitation with the present SNP data is that each locus can only be represented by two alleles, as genetic software programs (see above) generally do not permit input of multi-allelic states. The result is a decrease in among-population variability, especially on a broad global basis. Another potential limitation is bias introduced when using coding positions potentially influenced by selection, that is, not all markers are neutral. Such is the case for rare variant positions like those, that code for diseases. In the present study, such bias is limited because of the large SNP number. Specifically, drift, unlike selection, influences the whole genome so selection effects are relatively few (Kimura, 1968). Moreover, the Human Origins ascertainment used in this study, unlike some other arrays, is not biased by the inclusion of SNPs with medical/clinical interest.

5 | CONCLUSION

In sum, the present study is the most comprehensive to date comparing dental nonmetric traits and neutral genomic markers, in terms of the amount of data and number of samples at continental and global levels. The correspondence of these datasets based on comparison of MMD and F_{ST} distance matrices is greater in than any prior studies, likely because of the data and samples, as well as the methods used. Mantel correlations between these distances are all strongly positive, ranging from .72 globally to .84 within Africa. These and inter-sample geographic distance matrices are also strongly correlated, ranging from .68 for the 36-trait MMD in Africa to .77 for F_{ST} globally; the only exception is between F_{ST} and African geographic distances, which though moderate, remains positive and significant, r_m = 0.49. This, and partial correlations between MMD and F_{ST} controlling for geographic distances, namely, high in Africa (.78) and moderate globally (.4), suggest that the genetic distance measure is better able to pick up on recent within-continent variation, while recognizing ancient relationships on a global level.

That said, as mentioned, all correlations may be seen as minimum values in light of several recognized limitations, notably, ASUDAS and SNP data are not from the same individuals across samples. Though prohibitive concerning cost and time, future researchers could follow and expand upon the approach of Hubbard et al. (2015); ideally, both sets of data would be collected from the same skeletal remains in local- through global-level samples. Among other potential advantages over living individuals, actual teeth instead of casts would promote more detailed trait recording and include key root data. The bottom-line then, in conjunction with the recent heritability studies and population analyses, and despite potential concerns (sex dimorphism, selection, etc.), is that dental nonmetric traits actually can and do work well as proxies for neutral genomic data.

ACKNOWLEDGMENTS

Gareth Weedall, Liverpool John Moores University, provided invaluable assistance to AM during the bioinformatics processing for the

genomic data, particularly realignment and processing of the Wolof raw data. Thanks are extended to individuals at institutions—past and present—from which dental data were collected by JDI, including: Douglas Ubelaker and David Hunt, National Museum of Natural History; Ian Tattersall, Jaymie Brauer, Gary Sawyer, and Ken Mowbray, American Museum of Natural History; Andre Langaney, Frances Roville-Sausse, Miya Awazu Periera da Silva, Alain Froment, and Phillippe Mennecier, Museé de l'Homme; Rob Foley, Marta Lahr, and Maggie Bellatti, Duckworth Laboratory, University of Cambridge; Brendan Billings, the Dart Collection at University of the Witwatersrand; James S. Brink and Sharon Holt, National Museum in Bloemfontein; David Morris, McGregor Museum in Kimberly; Sven Ouzman, Iziko Museums of South Africa in Cape Town; and Alan Morris, Department of Human Biology, University of Cape Town. Thanks also go to John Relethford, State University of New York College at Oneonta, for his input, and providing a copy of his RMAT 1.2 program for us to determine if it would work with the large SNP dataset. Finally, we acknowledge the enormous efforts of Christy G. Turner II who worked for decades developing the fundamental methods for using dental morphology in population studies while amassing the world's largest database on tooth crown and root variation on a global scale. Funding for data collection by JDI came from the National Science Foundation (BCS-0840674, BNS-0104731, BNS-9013942), Arizona State University Research Development Program, and American Museum of Natural History.

DATA AVAILABILITY STATEMENT

The ASUDAS frequency data used in the analyses are available in Tables 3 and 7 of this article, as well as the relevant references cited in the main text. Many ASUDAS nondichotomized data are available in Scott & Irish (2017). The remaining ASUDAS data from the study are in preparation for publication, and/or are available upon reasonable request from the first or fifth authors. As noted, all SNP data, except the whole genome sequences of the West African Wolof sample, were genotyped with the Affymetrix Human Origins Array (AHOA), available in the relevant references cited in the text.

ORCID

Joel D. Irish https://orcid.org/0000-0001-7857-8847

REFERENCES

Bailey, S. E., & Hublin, J. J. (2013). What does it mean to be dentally "modern"? In G. R. Scott & J. D. Irish (Eds.), *Anthropological Perspectives on Tooth Morphology: Genetics, Evolution, Variation* (pp. 222–249). Cambridge: Cambridge University Press.

Bailey, S. E., Weaver, T. D., & Hublin, J. J. (2017). The dentition of the earliest modern humans: How 'modern' are they? In A. Marom & E. Hovers (Eds.), *Human Paleontology and Prehistory* (pp. 215–232). Cham: Springer.

Bermudez de Castro, J. M. (1989). The Carabelli trait in human prehistoric populations of the Canary Islands. *Human Biology*, *61*, 117–131.

Bhatia, G., Patterson, N., Sankararaman, S., & Price, A. L. (2013). Estimating and interpreting F_{ST}: The impact of rare variants. *Genome Research*, *23*, 1514–1521.

Brewer-Carias, C. A., Le Blanc, S., & Neel, J. V. (1976). Genetic structure of a tribal population, the Yanomama Indians. XIII. Dental microdifferentiation. *American Journal of Physical Anthropology*, *44*, 5–14.

Bryk, J., Hardouin, E., Pugach, I., Hughes, D., Strotmann, R., Stoneking, M., & Myles, S. (2008). Positive selection in East Asians for an EDAR allele that enhances NF-κB activation. *PLoS One*, *3*, e2209.

Cohen, J. (1988). Set correlation and contingency tables. *Applied Psychological Measurement*, *4*, 425–434.

Danecek, P., Schiffels, S., & Durbin, R. (2014). Multiallelic calling model in bcftools (–m).

Delgado, M., Ramírez, L. M., Adhikari, K., Fuentes-Guajardo, M., Zanolli, C., Gonzalez-José, R., … Rothhammer, F. (2019). Variation in dental morphology and inference of continental ancestry in admixed Latin Americans. *American Journal of Physical Anthropology*, *168*, 438–447.

Diniz-Filho, J. A. F., Soares, T. N., Lima, J. S., Dobrovolski, R., Landeiro, V. L., Telles, M. P. D. C., … Bini, L. M. (2013). Mantel test in population genetics. *Genetics and Molecular Biology*, *36*, 475–485.

Edgar, H. J. H. (2013). Estimation of ancestry using dental morphological characteristics. *Journal of Forensic Sciences*, *58*, S3–S8.

Edgar, H. J. H. (2015). Dental morphological estimation of ancestry in forensic contexts. In G. E. Berg & S. C. Ta'ala (Eds.), *Biological Affinity in Forensic Identification of Human Skeletal Remains* (pp. 191–208). Boca Raton, FL: CRC Press.

Ersts, P.J. (2014). *Geographic Distance Matrix Generator (version 1.2.3)*. New York: American Museum of Natural History, Center for Biodiversity and Conservation. Available at: http://biodiversityinfonratics.armh.org/open_source/gdrrg, accessed on September 7, 2019.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, *5*, 189–211.

Green, R. F., & Suchey, J. M. (1976). The use of inverse sine transformations in the analysis of non-metric cranial data. *American Journal of Physical Anthropology*, *45*, 61–68.

Greenberg, J. H. (1963). *The languages of Africa*. The Hague: Mouton & Co.

Gross, J. M., & Edgar, H. J. (2019). Informativeness of dental morphology in ancestry estimation in African Americans. *American Journal of Physical Anthropology*, *168*, 521–529.

Guillot, G., & Rousset, F. (2013). Dismantling the Mantel tests. *Methods in Ecology and Evolution*, *4*, 336–344.

Haeussler, A. M., Turner, C. G., II, & Irish, J. D. (1988). Concordance of American and Soviet methods in dental anthropology. *American Journal of Physical Anthropology*, *75*, 218.

Hammer, Ø. (2019). *PAST: PAleontological STatistics, Version 3.23 Reference Manual*. Oslo: Natural History Museum, University of Oslo.

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, *4*, 1–9.

Hanihara, T. (1992). Dental and cranial affinities among populations of East Asia and the Pacific: the basic populations in East Asia, IV. *American Journal of Physical Anthropology*, *88*, 163–182.

Hanihara, T. (2013). Geographic structure of dental variation in the major human populations of the world. In G. R. Scott & J. D. Irish (Eds.), *Anthropological perspectives on tooth morphology: genetics, evolution, variation* (pp. 479–509). Cambridge: Cambridge University Press.

Harris, E. F. (1977). *Anthropologic and genetic aspects of the dental morphology of Solomon Islanders, Melanesia*. PhD dissertation, Arizona State University, Tempe.

Higgins, D., Hughes, T. E., James, H., & Townsend, G. C. (2009). Strong genetic influence on hypocone expression of permanent maxillary molars in South Australian twins. *Dental Anthropology Journal*, *22*, 1–7.

Hlusko, L. J., Carlson, J., Chaplin, G., Elias, S. A., Hoffecker, J. F., Huffman, M., … Scott, G. R. (2018). Evidence of environmental selection on the mother-to-infant transmission of vitamin D and fatty acids during the last ice age in Beringia. *Proceedings of the National Academy of Sciences*, *115*, e4426–e4432.

Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}. *Nature Reviews Genetics*, *10*, 639.

Hubbard, A.R. (2012). *An examination of population history, population structure, and biological distance among regional populations of the Kenyan coast using genetic and dental data*. Ph.D. Dissertation. Ohio State University, Columbus.

Hubbard, A. R., Guatelli-Steinberg, D., & Irish, J. D. (2015). Do nuclear DNA and dental nonmetric data produce similar reconstructions of regional population history? An example from modern coastal Kenya. *American Journal of Physical Anthropology*, *157*, 295–304.

Hudson, R. R., Slatkin, M., & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics*, *132*, 583–589.

Hughes, T.E., & Townsend, G.C. (2011). Twin studies of dental crown morphology: genetic and environmental determinants of the Cusp of Carabelli. In: Program 15th International Symposium on Dental Morphology, 37, Newcastle, UK. 25 August.

Hughes, T. E., & Townsend, G. C. (2013). Twin and family studies of human dental crown morphology: genetic, epigenetic, and environmental determinants of the modern human dentition. In G. R. Scott & J. D. Irish (Eds.), *Anthropological perspectives on tooth morphology: genetics, evolution, variation* (pp. 31–68). Cambridge: Cambridge University Press.

Hughes, T., Townsend, G., & Bockmann, M. (2016). An overview of dental genetics. In J. D. Irish & G. R. Scott (Eds.), *A companion to dental anthropology* (pp. 123–141). New York: Wiley-Blackwell.

Irish, J.D. (1993). *Biological affinities of late Pleistocene through modern African aboriginal populations: the dental evidence*. Ph.D. Dissertation. Arizona State University, Tempe.

Irish, J. D. (1997). Characteristic high- and low-frequency dental traits in sub-Saharan African populations. *American Journal of Physical Anthropology*, *102*, 455–467.

Irish, J. D. (1998a). Diachronic and synchronic dental trait affinities of Late and post-Pleistocene peoples from North Africa. *Homo*, *49*, 138–155.

Irish, J. D. (1998b). Ancestral dental traits in recent sub-Saharan Africans and the origins of modern humans. *Journal of Human Evolution*, *34*, 81–98.

Irish, J. D. (2000). The Iberomaurusian enigma: North African progenitor or dead end? *Journal of Human Evolution*, *39*, 393–410.

Irish, J. D. (2005). Population continuity versus discontinuity revisited: dental affinities among Late Paleolithic through Christian era Nubians. *American Journal of Physical Anthropology*, *128*, 520–535.

Irish, J. D. (2006). Who were the ancient Egyptians? Dental affinities among Neolithic through post- dynastic peoples. *American Journal of Physical Anthropology*, *129*, 529–543.

Irish, J. D. (2010). The mean measure of divergence (MMD): its utility in model-free and model-bound analyses relative to the Mahalanobis D2 distance for nonmetric traits. *American Journal of Human Biology*, *22*, 378–395.

Irish, J. D. (2015). Dental nonmetric variation around the world: Using key traits in populations to estimate ancestry in individuals. In G. E. Berg & S. C. Ta'ala (Eds.), *Biological Affinity in Forensic Identification of Human Skeletal Remains* (pp. 165–190). Boca Raton, FL: CRC Press.

Irish, J. D. (2016). Who were they really? Model-free and model-bound dental nonmetric analyses to affirm documented population affiliations of seven South African "Bantu" samples. *American Journal of Physical Anthropology*, *159*, 655–670.

Irish, J. D., Bailey, S. E., Guatelli-Steinberg, D., Delezene, L. K., & Berger, L. R. (2018). Ancient teeth, phenetic affinities, and African hominins: Another look at where *Homo naledi* fits in. *Journal of Human Evolution*, *122*, 108–123.

Irish, J. D., Black, W., Sealy, J., & Ackermann, R. R. (2014). Questions of Khoesan continuity: dental affinities among the indigenous Holocene peoples of South Africa. *American Journal of Physical Anthropology*, *155*, 33–44.

Irish, J. D., & Guatelli-Steinberg, D. (2003). Ancient teeth and modern human origins: an expanded comparison of African Plio-Pleistocene and recent world dental samples. *Journal of Human Evolution*, *45*, 113–144.

Irish, J. D., Guatelli-Steinberg, D., Legge, S. S., de Ruiter, D. J., & Berger, L. R. (2013). Dental morphology and the phylogenetic "place" of Australopithecus sediba. *Science*, *340*, 1233062.

Irish, J. D., Lillios, K. T., Waterman, A. J., & Silva, A. M. (2017). "Other" possibilities? Assessing regional and extra-regional dental affinities of populations in the Portuguese Estremadura to explore the roots of Iberia's Late Neolithic-Copper Age. *Journal of Archaeological Science: Reports*, *11*, 224–236.

Irish, J. D., & Morris, D. H. (1996). A supplemental description of the Bushman maxillary canine polymorphism. *South African Journal of Science*, *92*, 351–353.

Irish, J. D., & Turner, C. G., II. (1990). West African dental affinity of late Pleistocene Nubians: peopling of the Eurafrican-South Asian triangle II. *Homo*, *41*, 42–53.

Jay, F., Sjödin, P., Jakobsson, M., & Blum, M. G. (2012). Anisotropic isolation by distance: the main orientations of human genetic differentiation. *Molecular Biology and Evolution*, *30*, 513–525.

Jernvall, J., & Jung, J.-S. (2000). Genotype, phenotype, and developmental biology of molar tooth characters. *Yearbook of Physical Anthropology*, *48*, 171–190.

Kanitz, R., Guillot, E. G., Antoniazza, S., Neuenschwander, S., & Goudet, J. (2018). Complex genetic patterns in human arise from a simple range-expansion model over continental landmasses. *PLoS One*, *13*, e0192460.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, *217*, 624–626.

Kimura, R., Yamaguchi, T., Takeda, M., Kondo, O., Toma, T., Haneji, K., … Osawa, M. (2009). A common variation in EDAR is a genetic determinant of shovel-shaped incisors. *The American Journal of Human Genetics*, *85*, 528–535.

Kistler, L., Ware, R., Smith, O., Collins, M., & Allaby, R. G. (2017). A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Research*, *45*, 6310–6320.

Konigsberg, L.W. (2006). A post-Neumann history of biological and genetic distance studies in bioarchaeology. In J.E. Buikstra & L.A. Beck LA (Eds.). *Bioarchaeology: the contextual analysis of human remains* (pp. 263–279). New York: Academic Press.

Konigsberg, L. W. (1990). Analysis of prehistoric biological variation under a model of isolation by geographic and temporal distance. *Human Biology*, *62*, 49–70.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., … Berger, B. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, *513*, 409–413.

Legendre, P., & Fortin, M.-J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, *10*, 831–844.

Leigh, S. R., Relethford, J. H., Park, P. B., & Konigsberg, L. W. (2003). Morphological differentiation of *Gorilla* subspecies. In A. B. Taylor & M. L. Goldsmith (Eds.), *Gorilla biology: a multidisciplinary perspective* (pp. 104–131). Cambridge: Cambridge University Press.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Mantel, N. A. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., … Skoglund, P. (2016). The Simons genome diversity project: 300 genomes from 142 diverse populations. *Nature*, *538*, 201.

Martinón-Torres, M., Bermúdez de Castro, J., Gómez-Robles, A., Arsuaga, J., Carbonell, E., Lordkipanidze, D., … Margvelashvili, A. (2007). Dental evidence on the hominin dispersals during the Pleistocene. *Proceedings of the National Academy of Sciences*, *104*, 13279–13282.

Martinón-Torres, M., Bermúdez de Castro, J., Gómez-Robles, A., Margvelashvili, A., Prado, L., Lordkipanidze, D., & Vekua, A. (2008). Dental remains from Dmanisi (Republic of Georgia): morphological analysis and comparative study. *Journal of Human Evolution*, *55*, 249–273.

Martinón-Torres, M., Bermúdez de Castro, J. M., Martín-Frances, L., Gracia-Téllez, A., Martinez, I., & Arsuaga, J. L. (2013). Dental morphology of European Middle Pleistocene populations. In G. R. Scott & J. D. Irish (Eds.), *Anthropological Perspectives on Tooth Morphology: Genetics, Evolution, Variation* (pp. 201–222). Cambridge: Cambridge University Press.

Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology*, *15*, 1044–1045.

Nelis, M., Esko, T., Mägi, R., Zimprich, F., Zimprich, A., Toncheva, D., … Jakkula, E. (2009). Genetic structure of Europeans: a view from the North–East. *PLoS One*, *4*, e5472.

Nichol, C.R. (1990). *Dental genetics and biological relationships of the Pima Indians of Arizona*. Ph.D. Dissertation. Arizona State University, Tempe.

Nikita, E. (2015). A critical review of the mean measure of divergence and Mahalanobis distances using artificial data and new approaches to the estimation of biodistances employing nonmetric traits. *American Journal of Physical Anthropology*, *157*, 284–294.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., … Oksanen, M. J. (2013). Package 'vegan'. *Community Ecology Package*, *Version*, *2*, 1–295.

Ortega-Del Vecchyo, D., & Slatkin, M. (2019). F_{ST} between archaic and present-day samples. *Heredity*, *122*, 711–718.

Park, J. H., Yamaguchi, T., Watanabe, C., Kawaguchi, A., Haneji, K., Takeda, M., … Hanihara, T. (2012). Effects of an Asian-specific nonsynonymous EDAR variant on multiple dental traits. *Journal of Human Genetics*, *57*, 508.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., … Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*, 1065–1093.

Paul, K.S., Stojanowski, C.M., Hughes, T., Brook, A., & Townsend, G.C. (2020). Patterns of heritability across the human diphyodont dental complex: crown morphology of Australian twins and families. *American Journal of Physical Anthropology*. 1-15. Published online. https://onlinelibrary.wiley.com/doi/abs/10.1002/ajpa.24019

Peter, B. M. (2016). Admixture, population structure, and *F*-statistics. *Genetics*, *202*, 1485–1501.

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*, e1002967.

Pickrell, J. K., Patterson, N., Loh, P. R., Lipson, M., Berger, B., Stoneking, M., … Reich, D. (2014). Ancient west Eurasian ancestry in southern and eastern Africa. *Proceedings of the National Academy of Sciences*, *111*, 2632–2637.

Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., … Anders, A. (2015). Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One*, *10*, e0129102.

Pinkerton, S., Townsend, G., Richards, L., Schwerdt, W., & Dempsey, P. (1999). Expression of Carabelli trait in both dentitions of Australian twins. *Perspectives on Human Biology*, *4*, 19–28.

Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., & Cavalli-Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences*, *102*, 15942–15947.

Rathmann, H., Reyes-Centeno, H., Ghirotto, S., Creanza, N., Hanihara, T., & Harvati, K. (2017). Reconstructing human population history from dental phenotypes. *Scientific Reports*, *7*, 12495.

Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., … García, L. F. (2012). Reconstructing Native American population history. *Nature*, *488*, 370–375.

Reich, D., Thangaraj, K., Patterson, N., Price, A. L., & Singh, L. (2009). Reconstructing Indian population history. *Nature*, *461*, 489–494.

Relethford, J. (2004). Global patterns of isolation by distance basedon genetic and morphological data. *Human Biology*, *76*, 499–513.

Relethford, J. H. (1991). Effect of changes in population size on genetic microdifferentiation. *Human Biology*, *63*, 629–641.

Relethford, J. H., & Crawford, M. H. (1995). Anthropometric variation and the population history of Ireland. *American Journal of Physical Anthropology*, *96*, 25–38.

Relethford, J. H., & Harpending, H. C. (1994). Craniometric variation, genetic theory, and modern human origins. *American Journal of Physical Anthropology*, *95*, 249–270.

Reyes-Centeno, H., Rathmann, H., Hanihara, T., & Harvati, K. (2017). Testing modern human out-of-Africa dispersal models using dental nonmetric data. *Current Anthropology*, *58*(Supplement 17), S406–S417.

Riga, A., Belcastro, M. G., & Moggi-Cecchi, J. (2014). Environmental stress increases variability in the expression of dental cusps. *American Journal of Physical Anthropology*, *153*, 397–407.

Rightmire, G. P. (1999). Dental variation and human history. *Revista Archaeologica*, *20*, 1–3.

Schillaci, M. A., Irish, J. D., & Wood, C. (2009). Further analysis of the population history of ancient Egyptians. *American Journal of Physical Anthropology*, *139*, 235–243.

Scott, G. R. (1973). *Dental morphology: A genetic study of American White families and variation in living Southwest Indians*. PhD dissertation, Arizona State University, Tempe.

Scott, G. R., & Irish, J. D. (2017). *Human tooth crown and root morphology: The Arizona State University Dental Anthropology System*. Cambridge: Cambridge University Press.

Scott, G. R., Turner, C. G., II, Townsend, G. C., & Martinón-Torres, M. (2018). *The anthropology of modern human teeth: Dental morphology and its variation*. Cambridge: Cambridge University Press.

Scott, G. R., Pilloud, M., Navega, D., d'Oliveira Coelho, J., Cunha, E., & Irish, J. D. (2018). rASUDAS: A new web-based application for estimating ancestry from tooth morphology. *Forensic Anthropology*, *1*, 18–31.

Scott, G. R., Yap Potter, R. H., Noss, J. F., Dahlberg, A. A., & Dahlberg, T. (1983). The dental morphology of Pima Indians. *American Journal of Physical Anthropology*, *61*, 13–31.

Scott, G. R., & Turner, C. G., II. (1997). *The anthropology of modern human teeth: Dental morphology and its variation in recent human populations*. Cambridge: Cambridge University Press.

Scott, G. R., & Turner, C. G., II. (1988). Dental anthropology. *Annual Review of Anthropology*, *17*, 99–126.

Séré, M., Thévenon, S., Belem, A., & De Meeûs, T. (2017). Comparison of different genetic distances to test isolation by distance between populations. *Heredity*, *119*, 55–63.

Sjøvold, T. (1973). Occurrence of minor non-metrical variants in the skeleton and their quantitative treatment for population comparisons. *Homo*, *24*, 204–233.

Sjøvold, T. (1977). Non-metrical divergence between skeletal populations: The theoretical foundation and biological importance of C.A.B. Smith's mean measure of divergence. *Ossa*, *4*(Supplement 1), 1–133.

Skoglund, P., Mallick, S., Bortolini, M. C., Chennagiri, N., Hünemeier, T., Petzl-Erler, M. L., … Reich, D. (2015). Genetic evidence for two founding populations of the Americas. *Nature*, *525*, 104–108.

Skoglund, P., Posth, C., Sirak, K., Spriggs, M., Valentin, F., Bedford, S., … Fu, Q. (2016). Genomic insights into the peopling of the Southwest Pacific. *Nature*, *538*, 510.

Smith, C. I., Chamberlain, A. T., Riley, M. S., Stringer, C., & Collins, M. J. (2003). The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution*, *45*, 203–217.

Smith, P., & Shegev, M. (1988). The dentition of Nubians from Wadi Halfa, Sudan: an evolutionary perspective. *Journal of the Dental Association of South Africa*, *43*, 539–541.

Smouse, P. E., & Long, J. C. (1992). Matrix correlation analysis in anthropology and genetics. *American Journal of Physical Anthropology*, *35* (S15), 187–213.

Smouse, P. E., Long, J. C., & Sokal, R. R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, *35*, 627–632.

Sofaer, J. A., Niswander, J. D., MacLean, C. J., & Workman, P. L. (1972). Population studies on Southwestern Indian tribes V. Tooth morphology as an indicator of biological distance. *American Journal of Physical Anthropology*, *37*, 357–366.

Sofaer, J. A., Smith, P., & Kaye, E. (1986). Affinities between contemporary and skeletal Jewish and non-Jewish groups based on tooth morphology. *American Journal of Physical Anthropology*, *70*, 265–275.

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry*. New York: Freeman and Company.

Stojanowski, C. M., & Paul, K. S. (2015). Performance analysis of deciduous morphology for detecting biological siblings. *American Journal of Physical Anthropology*, *157*, 615–629.

Stojanowski, C. M., Paul, K. S., Seidel, A. C., Duncan, W. N., & Guatelli-Steinberg, D. (2018). Heritability and genetic integration of anterior tooth crown variants in the South Carolina Gullah. *American Journal of Physical Anthropology*, *167*, 124–143.

Stojanowski, C. M., Paul, K. S., Seidel, A. C., Duncan, W. N., & Guatelli-Steinberg, D. (2019). Quantitative genetic analyses of postcanine morphological crown variation. *American Journal of Physical Anthropology*, *168*, 606–631.

Stojanowski, C. M., & Schillaci, M. A. (2006). Phenotypic approaches for understanding patterns of intracemetery biological variation. *Yearbook of Physical Anthropology*, *131*(suppl. 43), 49–88.

Tian, C., Kosoy, R., Nassir, R., Lee, A., Villoslada, P., Klareskog, L., … Gregersen, P. K. (2009). European population genetic substructure: further definition of ancestry informative markers for distinguishing among diverse European ethnic groups. *Molecular Medicine*, *15*, 371–383.

Turner, C. G., II (1985a). The dental search for Native American origins. In R. Kirk & E. Szathmary (Eds.), *Out of Asia* (pp. 31–78). Canberra: The Journal of Pacific History.

Turner, C. G., II. (1985b). Expression count: a method for calculating morphological dental trait frequencies by using adjustable weighting coefficients. *American Journal of Physical Anthropology*, *68*, 263–268.

Turner, C. G. (1990). Major features of Sundadonty and Sinodonty, including suggestions about East Asian microevolution, population history, and late Pleistocene relationships with Australian aboriginals. *American Journal of Physical Anthropology*, *82*, 295–317.

Turner, C. G., II, Nichol, C. R., & Scott, G. R. (1991). Scoring procedures for key morphological traits of the permanent dentition: The Arizona State University Dental Anthropology System. In M. A. Kelley & C. S. Larsen (Eds.), *Advances in dental anthropology* (pp. 13–31). New York: Wiley-Liss.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

Weir, B. S., Cardon, L. R., Anderson, A. D., Nielsen, D. M., & Hill, W. G. (2005). Measures of human population structure show heterogeneity among genomic regions. *Genome Research*, *15*, 1468–1476.

Willing, E.-M., Dreyer, C., & van Oosterhout, C. (2012). Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PLoS One*, *7*, e42649.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Irish JD, Morez A, Girdland Flink L, Phillips ELW, Scott GR. Do dental nonmetric traits actually work as proxies for neutral genomic data? Some answers from continental- and global-level analyses. *Am J Phys Anthropol*. 2020;1–29. https://doi.org/10.1002/ajpa.24052