



LJMU Research Online

Irish, JD

Who were they really? model-free and model-bound dental nonmetric analyses to affirm documented population affiliations of seven South African "Bantu" samples.

<http://researchonline.ljmu.ac.uk/id/eprint/2568/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Irish, JD (2015) 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. ISSN 1096-8644

LJMU has developed [LJMU Research Online](#) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>



**Who were they really? Model-free and model-bound dental
nonmetric analyses to affirm documented population
affiliations of seven South African "Bantu" samples**

Journal:	<i>American Journal of Physical Anthropology</i>
Manuscript ID	AJPA-2015-00342.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	22-Nov-2015
Complete List of Authors:	Irish, Joel; Liverpool John Moores University, Research Centre in Evolutionary Anthropology and Palaeoecology
Key Words:	dental morphology, phenetic affinity, genetic structure, population history, isolation-by-distance
Subfield: Please select your first choice in the first field.:	Bioarchaeology [including forensics], Genetics [primate and human]

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Who were they really? Model-free and model-bound dental nonmetric analyses to affirm documented population affiliations of seven South African “Bantu” samples

Joel D. Irish

Research Centre in Evolutionary Anthropology and Palaeoecology,
Liverpool John Moores University, Liverpool L3 3AF, United Kingdom.

32 text pages

12 reference pages

1 footnote page

7 tables

5 figures

Running Title: South African “Bantu” dental affinities

KEY WORDS Raymond A. Dart Collection, South Africa, dental morphology, phenetic affinity, genetic structure, R-matrix, F_{st} , isolation-by-distance, population history

Send correspondence to: Prof. Joel D. Irish, Research Centre in Evolutionary Anthropology and Palaeoecology, School of Natural Sciences and Psychology, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, United Kingdom +44 (0)151 231 2387 (Office)
+44 (0)151 231 2338 (Fax)
J.D.Irish@ljmu.ac.uk

Grant sponsor: National Science Foundation (BCS-0840674)

1
2
3
4
5
6
7
8
9 **ABSTRACT Objectives:** For bioarchaeological biodistance analyses it is common to “assume”
10 that skeletal samples are representative of the populations to which they are attributed. Here,
11 alternatively, samples with “known” attribution in the Raymond A. Dart Collection are assessed
12 regarding their suitability for use in such analyses. Prior curation issues may call their ascribed
13 identities into question.
14

15
16
17 **Materials and Methods:** These 20th century samples ostensibly derive from South African
18 Ndebele, Sotho, Swazi, Tswana, Venda, Xosa, and Zulu populations. First, the mean measure of
19 divergence (MMD) is used to obtain among-sample dental phenetic distances for comparison
20 with documented population relationships. Second, the Mantel test evaluates fit of the isolation-
21 by-distance model between MMD and geographic distances, i.e., among the historic homelands.
22 Third, R-matrices and minimum and estimated F_{st} from MMD distances give an indication of
23 genetic micro-differentiation.
24
25

26
27
28
29 **Results:** Output from these model-free and model-bound analyses suggest that five and perhaps
30 six samples are representative of their attributed populations – presenting differences along
31 population lines and evidence of more ancient ancestry.
32
33

34
35 **Discussion:** Other than the Swazi and perhaps Ndebele, the among-sample variation: 1) mirrors
36 documented population history, 2) reveals a moderately positive correlation between phenetic
37 and geographic distances, and 3) although evidencing much homogeneity, provides measures of
38 genetic distance in support of the phenetic distances. Therefore, with the two noted exceptions –
39 perhaps from collection issues, swamping of past genetic structure, or both, most samples appear
40 suitable for bioarchaeological analyses. On this basis, results are offered to supplement published
41 findings concerning the biological relationships of these peoples.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Standard operating procedure for biodistance study is to record and compare data in skeletal samples that are then “assumed” to be representative of their respective populations. An alternative approach is taken here. Skeletal samples of individuals “documented” to derive from specific recent populations are analyzed to determine if they really are representative of those populations. The samples are from the Raymond A. Dart Collection of Human Skeletons, curated at the University of the Witwatersrand in Johannesburg. Specifically, the objective is to explore the legitimacy of an unsubstantiated view, by various researchers familiar with the material, that many individuals may not be correctly attributed. If this opinion is discounted for at least part of the collection, then certain samples may prove useful for subsequent bioarchaeological study. This work also serves as a focused follow-up to Dayal and colleagues’ (2009) overview of the full collection and the 2,605 skeletons therein.

Most of these skeletons came from cadavers of early to mid-20th century South Africans. Information on these individuals that is now available to researchers may vary from that recorded at the times of death; for example, the names are not provided. Otherwise data categories in these records include: ID and accession numbers, cause of death, notes of interest, sex, age, death date, population group (often to the level of “tribe” for Africans), skeletal inventory, and additional notes. As Dayal et al. (2009) relate, the percentage of individuals in the collection can be broken down into the country’s standard census categories (Stats SA, 2014; Jacobson, 1982): “Indian” (0.3%), “Coloured,” i.e., ancestry from two or more of the other categories (4%), “White” (15%), and “African” (76%).

The focus here is on the last category, specifically those assigned to one of seven “Bantu” ethnic groups: Ndebele, Sotho, Swazi, Tswana, Venda, Xosa, and Zulu. The dearth of equivalent

1
2
3
4
5
6
7
8
9 remains curated elsewhere compelled the author to select these samples for inclusion in a 2008
10 National Science Foundation grant proposal (BCS-0840674) on the “‘Bantu’ Expansion” (see
11 below). It was assumed that the samples are representative of their respective groups because the
12 Collection is “one of the best documented in the world” (Morris, 1992, p 76) and the period
13 when these remains were accessioned predates the most extensive detribalization, so individuals
14 are likely “unhybridised” (De Villiers, 1968, p 5). However, as detailed by Dayal et al. in 2009,
15
16
17
18
19
20 most had died near Johannesburg, not their ethnic homelands, so hospital identifications could be
21
22 made via physical appearance, the name, or other subjective factors (e.g., purported language)
23
24 (also see Tal and Tau, 1983). Other issues include flooding of the skeletons from a burst pipe in
25
26 1959 with potential subsequent mixing of some elements and, prior to that, specimen exchange
27
28 among institutions with sparse record keeping (Dayal et al., 2009).

29
30 Given these concerns, model-free and model-bound quantitative analyses are used to help
31 (re)confirm the validity of these ethnic identities. By necessity of the methods employed, “ethnic
32 group” is deemed synonymous with “population,” where members of the former share biological
33 features (from a common gene pool) that differentiate them from those of other ethnic groups.
34
35
36 Phenetic affinities and estimates of genetic structure from dental nonmetric data using the mean
37
38 measure of divergence (MMD) (following the approach in Irish, 2010) are compared with known
39
40 population relationships, and interpreted by means of a hypothesis-guided approach.
41
42

43 The MMD was chosen over other measures, e.g., Mahalanobis D^2 for nonmetric traits
44 (Konigsberg, 1990), for three reasons (detailed in Irish, 2010). First, it was established to be a
45
46 robust statistic, giving consistent results before and after the removal of problematic traits (i.e.,
47
48 highly correlated trait pairs, many missing data, fixed frequencies) (also Nikita, 2015). Second,
49
50 for model-bound analyses – those which are “derived directly from evolutionary models and
51
52

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

allow estimates of specific parameters” (Relethford and Blangero, 1990; Relethford and Harpending, 1994, p 251), MMD values are consistently highly correlated with geographic distances to, for example, assess the fit of the isolation-by-distance model. Third, although necessitating an additional step, MMD distances can be used to approximate a genetic relationship (or R-) matrix and Sewall Wright’s fixation index (F_{st}) like those from the D^2 (Konigsberg, 2006).¹

Finally, based on these analyses it is suggested that five, or perhaps six of the samples do appear representative of their assigned ethnic groups/populations. In these cases, results are presented to supplement published findings concerning the biological characterization and relationships of these peoples (among others, De Villiers, 1968; Jacobson, 1982; Excoffier et al., 1987; Lane et al., 2002; Abbé et al., 2006). As well, it appears that these five or six synchronic samples should prove useful as proxies for premodern populations to permit future diachronic study of “Bantu” origins and affinities on regional and broader levels; again, it was this prospect that prompted the recording of these data, as expanded upon below.

SOUTH AFRICAN “BANTU” POPULATION HISTORY

The term “Bantu” was coined by Wilhelm Bleek in 1862 to classify a group of over 400 languages spoken across sub-Saharan Africa (Greenberg, 1963; Lwango-Lunyiigo and Vansina, 1988; Schoenbrun, 2001). Since its inception the term was also commonly used to classify the speakers of these languages and their cultures (Lwango-Lunyiigo and Vansina, 1988). Though inexact, this practice is continued here in accordance with a number of prior publications for simplicity.

In brief, all Bantu populations are descendants of peoples who once lived in Nigeria and Cameroon near the Cross River Valley (July, 1992; Ruhlen, 1994; Newman, 1995; Vogel, 1997).

1
2
3
4
5
6
7
8
9 Around 4,000-3,000 BP these agriculturalist proto-Bantu began to expand outward from this
10 region to the south and east as a result of growth beyond local carrying capacity and resulting
11 social stress (Soper, 1982; Hiernaux, 1975; Collett, 1982; Barker, 2006). The eastward migrants,
12 after reaching the plains of east Africa and Lake Victoria region by the early 1st millennium BC
13 (Newman, 1995; Ehret, 2000), or later (July, 1992; Iliffe, 1997; Vogel, 1997), eventually turned
14 southward. These travelers into eastern South Africa are among the last Bantu populations to
15 have reached their historic homelands (Huffman, 1988). Here, “homeland” is intended to denote
16 the settlement location of these groups, not that established by the Apartheid Government (a.k.a.,
17 “Bantustans”) (Butler et al., 1978), although there is some correspondence.
18
19
20
21
22
23
24
25

26 Two general routes were used from east to South Africa. The first followed the coastal
27 plain (Maylam, 1986), and was taken by speakers attributed to the Nguni Branch of the Bantu
28 family of languages. Their entry into the country is once thought to be recent, perhaps in the 16th
29 century (Monnig, 1967; Maylam, 1986). However, archaeological evidence points to a much
30 earlier date. ca. 1100 AD (Huffman, 2007; see overview in Warren et al., 2014). The second
31 route, along the western shores of Lake Malawi (Monnig, 1967; Sutton, 1981; Nurse et al, 1985),
32 was associated with speakers of Sotho languages. They reached the northern Transvaal (now
33 Limpopo Province) around the 13th-14th centuries AD (Monnig, 1967), as supported by recovery
34 of diagnostic pottery dating to ca. 1300 in the Soutpansberg Mountains (Huffman, 1989; Hall,
35 2010; Warren et al., 2014). Upon arrival and movement around South Africa, both groups
36 encountered and differentially interbred with earlier Bantu migrants and indigenous Khoesan-
37 speaking peoples (Tobias, 1974; Denbow, 1981; Parkington, 1981; Nurse et al., 1985; Loubser,
38 1989; Soodyall, 1993 in Mitchell, 2010; Newman, 1995). Though ultimately having a common
39 origin, Nguni and Sotho populations differ in many respects, most notably language and social
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

organization (Van Warmelo, 1962). This biocultural history accounts for the backgrounds of all seven populations covered; some inhabit neighboring countries but only those in South Africa proper are of specific interest here. Additional details are provided in the sample descriptions below.

MATERIALS

Nonmetric traits were recorded in 408 of the total 1,390 South African “Africans” that are currently part of the Dart Collection (Dayal et al., 2009) using the Arizona State University Dental Anthropology System (described below). In cases of few ethnically-identified individuals, i.e., the Ndebele ($n=38$ individuals) and Venda ($n=51$), all relevant crania with permanent teeth were scored. For the other five ethnic groups with many more identified individuals, systematic random selection was conducted. That is, beginning with the first accession ID number for each ethnic group, every other, or second, specimen was recorded until a minimum representative number (i.e., $n=15$, see below) for each trait was obtained (refer to counts in Table 1).

Ndebele

This sample consists of 38 individuals assigned to the Ndebele ethnic group/population. The latter speak an Nguni language, isiNdebele (Byrnes, 1996), so are held to be descendants of the coastal route group; indeed, and importantly (below), they may have originally settled in far east-central South Africa, south of Swaziland in KwaZulu Natal province (Byrnes, 1996). The Ndebele are said to be offshoots of the Nguni-speaking Zulu (Van Warmelo, 1962; Seligman, 1967). However, their present location places them alongside many Sotho-speakers (Van Warmelo, 1962) and influence of the latter is evident, as their language may be classified as a form of seSotho (Byrnes, 1996). Overall, the Ndebele are understood to be an amalgamation of peoples brought together during the 19th century; the reason, in part, is related to the Mfecane (or

1
2
3
4
5
6
7
8
9 Difaqane), a time of major upheaval and forced migration when, among others, the Ndebele were
10 displaced back toward the country's northwest corner and beyond (July, 1992; Byrnes, 1996).

11 The Ndebele live in other regions of South Africa, but have inhabited parts of today's Limpopo
12 Province, mostly in the vicinity of KwaNdebele (Fig. 1), for over a century (Byrnes, 1996).

13
14
15
16 **[FIGURE 1 HERE]**

17 18 **Sotho**

19
20 This sample comprises individuals said to be of Southern Sotho origin ($n=66$), who are
21 centered in Lesotho (Fig. 1); they are differentiated from the Northern Sotho, or Pedi, who like
22 the Ndebele mostly inhabit Limpopo Province. As the name implies, they are identified as the
23 descendants of Bantu immigrants who are believed to have entered South Africa using the inland
24 route (above). After arriving, those who would become the Southern Sotho eventually continued
25 south in the 15th century and after (Byrnes, 1996; Hall, 2010). Along the way they met up with
26 Khoesan peoples, as indicated by the incorporation of some click sounds in their language,
27 unlike that of Northern Sotho. By the 1830s this loosely associated group, and some Nguni
28 peoples displaced by the Mfecane, were united and established in their current homeland (Van
29 Warmelo, 1962; Byrnes, 1996).

30 31 32 33 **Swazi**

34 This sample ($n=58$) ostensibly contains Nguni-speaking individuals belonging to an
35 ethnic group of relatively recent origin (Van Warmelo, 1962). Their society, prior to the late 18th
36 century, consisted of related Nguni patrilineal descent groups in what is now southern Swaziland
37 (Fig. 1) (Seligman, 1967; Byrnes, 1996). After that, a distinct ethnic identity was formed in the
38 mid- to late 19th century by two Swazi leaders who subjugated and integrated neighboring Nguni
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

and, eventually, Sotho groups; the latter formerly inhabited much of northern Swaziland (Van Warmelo, 1962; Byrnes, 1996).

Tswana

The Tswana sample consists of 63 individuals. Members of this ethnic group, also known as the Western Sotho, are thought to be descended from several populations in north-central South Africa (Byrnes, 1996), in the present-day North West Province (Fig. 1). Their language, seTswana, is closely related to seSotho (Byrnes, 1996), so speakers of these languages may share a common origin from the second group of immigrants to South Africa. Some of the latter then moved westward several centuries ago (Van Warmelo, 1962) along the southern fringes of the Kalahari Desert to their present location (July, 1992). Like their Southern Sotho relatives, the Tswana encountered and interacted extensively with local Khoesan peoples, mostly San (Van Warmelo, 1962).

Venda

The Venda ethnic group/population, to which the individuals ($n=51$) in this sample are assigned, live in the Soutpansberg Mountains region of Limpopo Province (Fig. 1). They speak neither a Sotho nor Nguni language, though there is some similarity to the former. The Venda appear to be a regional amalgamation (Loubser, 1989). That is, like other populations, they are comprised of several Bantu groups. It is thought that Shona-speakers from Zimbabwe migrated south into the Soutpansberg during the 15th century AD, and interacted with Northern Sotho who lived there since the 14th century. Their integration resulted in a Venda population by the mid-16th century (Loubser, 1989). The Singo from Zimbabwe conquered this first Venda incarnation during the late 17th century; however, the former group adopted the language and customs of the

1
2
3
4
5
6
7
8
9 latter to maintain a Venda identity (Loubser, 1989). Given their relative isolation they have since
10 interacted minimally with other populations (Van Warmelo, 1962; Abbé et al., 2006).

11 12 **Xosa**

13
14 The Xosa sample ($n=65$) is assumed to be made up of individuals who spoke isiXosa, an Nguni
15 language somewhat like Zulu (Byrnes, 1996). Of all local Bantu languages it contains the
16
17 greatest number (i.e., 12) of click sounds; based on the long history between Xosa and Khoesan
18 the integration of these sounds is not surprising. Xosa peoples first reached their southern coastal
19 location, in present-day Eastern Cape Province (Fig. 1) (Seligman, 1967), from the Drakensberg
20 region to the east (Phillipson, 1994) prior to the 15th century; others followed during the 16th-17th
21 centuries (July, 1992; Byrnes, 1996). This location, like the whole of the country, was inhabited
22 by Khoesan, mostly Khoekhoe (Van Warmelo, 1962; Byrnes, 1996), with whom the Xosa lived
23 alongside and came to dominate (Maylam, 1986; Byrnes, 1996).
24
25
26
27
28
29
30

31 **Zulu**

32
33 The final sample ($n=67$) consists of peoples identified as Nguni-speakers who lived on
34 the far eastern coast of South Africa in present-day Kwazulu Natal Province (Fig. 1). In the 18th
35 century a number of “Natal” Nguni (Van Warmelo, 1962; Seligman, 1967) groups inhabited the
36 area’s Tugela River region that could be considered the first Zulu (Byrnes, 1996). However, they
37 did not unite into a truly cohesive group until the early 19th century under two powerful kings,
38 including Shaka (Davidson, 1974; Maylam, 1986). As major players in the Mfecane, the Zulu
39 basically changed Bantu population structure in South Africa – particularly affecting other Nguni
40 groups. They eliminated some, subjugated others, and forced thousands more to retreat north
41
42
43
44
45
46
47
48
49 (Davidson, 1974; Maylam, 1986; Byrnes, 1996).
50
51
52
53
54
55
56
57
58
59
60

METHODS

Dental trait recording

This comparative study is based on morphological variation of the permanent dentition. Up to 125 nonmetric crown, root, and osseous traits were recorded in each individual. Of these, 36 (see list in Table 1) that have proven useful in prior African studies (Irish, 1993, 1997, 2005, 2006, 2013; Irish et al., 2014) were employed here. With the exception of midline diastema, all are from the Arizona State University Dental Anthropology System (ASUDAS). Most traits are present in both antimeres, i.e., are mirror images of one another. As such, during recording, a decision regarding which antimeres to score is required. One method entails counting only one side in all specimens (Haeussler et al., 1988). A second method is to score both antimeres and, allowing for asymmetry, count the side with the highest expression (Turner and Scott, 1977). To maximize sample size if only one side is present, that side is scored and assumed to represent the highest expression. This standard protocol is used here; it assumes scoring for the individual's maximum genetic potential (Turner, 1985a).

As detailed in Turner et al. (1991) and Scott and Turner (1997) ASUDAS traits hold a number of advantages. First, many remain observable despite slight attrition. Of course to avoid potentially biased data (Burnett, 2016), proper scoring restraint must be exercised (Nichol and Turner, 1986; Turner et al., 1991; Burnett et al., 2013; Stojanowski and Johnson, 2015); this is especially important with near-occlusal traits that are more affected at early wear stages (Burnett, 2016). Second, rank-scale reference plaques comprising the ASUDAS promote intra- and inter-observer recording repeatability; however, additional measures, like dichotomization (below), are used to address concordance issues (Nichol and Turner, 1986; Turner et al., 1991), especially between observers (Stojanowski and Johnson, 2015). Third, all dental morphogenetic fields (or

1
2
3
4
5
6
7
8
9 regions) are represented (Butler, 1939, 2001; Dahlberg, 1945; Osborn, 1978; Townsend et al.,
10 2016. Fourth, these traits possess a high genetic component in expression (Scott, 1973; Larsen,
11 1997; Scott and Turner, 1997; Rightmire, 1999; Martinon-Torres et al., 2007; Hughes and
12 Townsend, 2011, 2013). Lastly, “the fossil record has shown that (whatever their adaptive value)
13 they evolve very slowly” (Turner et al., 1991 p 13). Therefore, “the conservative nature of their
14 evolution” makes these dental nonmetric traits valuable for biodistance analyses (Larsen, 1997).
15
16 In addition, these traits have been demonstrated to show no or little sexual dimorphism
17 (Scott, 1973, 1980; Smith and Shegev, 1988; Bermudez de Castro, 1989; Hanihara, 1992; Irish
18 1993). Significant dissimilarities by sex that may occur appear to be random, in that different
19 traits are affected among studies. As such, it is standard procedure to pool the sexes (Turner et al,
20 1991). The absence of dimorphism is supported in the present study. Chi-square tests for the 36
21 traits in 399 sex-identified individuals from the full sample ($n=408$), revealed only root number
22 UP1 ($p=0.015$, 1 df), root number UM2 ($p=.025$, 1 df), and premolar odontome ($p=0.017$, 1 df)
23 to differ significantly (again, refer to the trait list in Table 1). Females have higher frequencies of
24 root fusion and very rare odontomes. These traits account for 8.0% of the 36 traits, near the 0.05
25 alpha level for random association (i.e., $p=0.05$; $1.8/36$). Further, any potential bias is offset by
26 the 3.1:1 ratio of 302 males and 97 females that is, other than small Ndebele (7.3:1) and Venda
27 (9.2:1), roughly emulated across samples. Therefore, the sexes are combined for analyses.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 **Model-free analyses**

44
45 Rank-scale ASUDAS data were dichotomized into categories of presence and absence
46 (Turner et al., 1991; Scott and Turner, 1997; Irish, 1993, 1997) to calculate inter-sample phenetic
47 distances with the mean measure of divergence (MMD) (Sjøvold, 1977). The MMD yields inter-
48 sample phenetic distances, where small values indicate similitude and vice versa. In addition, a
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Khoesan sample is included in a final comparison to explore if the documented differential gene flow is detectable, especially in Sotho, Tswana, and Xosa, though such influence is evident in all South African Bantu (Jenkins et al., 1970; Nurse et al., 1985; Soodyall, 1993 in Mitchell, 2010).

The rationale for this comparison comes from Relethford and Crawford (1995), who note that finer grained assessments of phenetic (and genetic) patterning may be possible if “outside” samples with known relationships to the populations of interest are included. The Khoesan sample ($n=135$) includes 20th century San dentitions from Botswana and South Africa, most of which have been previously studied (Irish, 1993, 1997).

The present MMD formula contains the Freeman and Tukey angular transformation that corrects for very low or very high trait frequencies and small sample sizes (Sjøvold, 1973, 1977; Green and Suchey, 1976). To determine if samples are significantly different, the distance is compared with its standard deviation, where if the $MMD > 2 \times s$, the null hypothesis of $P_1 = P_2$ is rejected at the 0.025 alpha level (Sjøvold, 1977). Although a robust statistic (above) it is still recommended that problematic traits be edited out prior to analyses. First, those having many missing data are deleted, because the bias transformation is not intended to correct for trait observations of less than 10 (Green and Suchey, 1976; Green et al., 1979). Second, fixed or largely invariant traits are removed because they provide no useful information for identifying differences among samples, and can result in negative MMD distance values; the latter is a statistical artifact that has “no biological meaning” (Harris and Sjøvold, 2004, p 91). Traits that are minimally discriminatory can also be identified quantitatively using, for example, principal components analysis (PCA). In the current study, any variable nor receiving a PCA loading of at least |0.5| was eliminated from further analysis. Third, Kendall's tau- b is used to find correlated trait pairs. As many traits as desired may be used, but they should not be highly correlated ($\tau_b \geq$

1
2
3
4
5
6
7
8
9 | 0.5) with others or differential weighting of the underlying dimensions can lead to erroneous
10 | distances (Sjøvold, 1977).

11
12 | Lastly, the distance matrix is submitted to interval-level multidimensional scaling (MDS)
13 | to visualize inter-sample affinities (SPSS 21.0 Procedure Alscal). The sum of squared differences
14 | between Euclidean values derived from this matrix (i.e., d_{ij}) and those in the resulting (\hat{d}_{ij}) matrix
15 | are minimized, i.e., optimally scaled (Hintze, 2007). From this, plots of 1 to n dimensions can
16 | illustrate sample relationship (Kruskal and Wish, 1978; Cox and Cox, 1994; Borg and Groenen,
17 | 1997).

23 | **Model-bound analyses**

24
25 | R-matrices and F_{st} are approximated from the MMD distances using a modified method
26 | (Irish, 2010) from Konigsberg (2006), which is based on an approach to obtain this output using
27 | metric data (Williams-Blangero and Blangero, 1989; Relethford and Blangero, 1990). Nonmetric
28 | data may hold an advantage over metric because they are more likely to be selectively neutral to
29 | more closely “parallel molecular measures of genetic divergence” (Leigh et al., 2003, p 116).
30 | The off-diagonal r_{ij} values in the R-matrix provide a measure of genetic distance among samples
31 | (Relethford and Crawford, 1995; Leigh et al., 2003), where positive values indicate greater
32 | similarity and negative values lesser similarity than average (Relethford and Harpending, 1994;
33 | Relethford et al., 1997). On-diagonal r_{ii} values help to evaluate internal variation; samples near
34 | the centroid possess greater heterogeneity, or heterozygosity; those farther away are more
35 | homogeneous as a result of genetic drift and lower migration (Relethford and Crawford, 1995;
36 | Konigsberg, 2006). Finally F_{st} , the mean of r_{ii} values, provides a measure of population sample
37 | differentiation (Relethford and Harpending, 1994).
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9 Königsberg (2006) uses a squared Euclidean-based distance matrix, i.e., D^2 for nonmetric
10 traits (Königsberg, 1990), to calculate a C- or co-divergence matrix, and from that an R-matrix
11 and F_{st} . This same output from non-Euclidean MMD distances is obtained by substituting the
12 optimally-scaled matrix from MDS (above). Specifically, because: 1) D^2 and MMD distance
13 matrices are highly correlated when comparing the same samples, 2) MDS optimal values are
14 rescaled into Euclidean distances, and 3) the latter provide as near a match as possible to the
15 original distances, the R-matrix from the MMD is proportionate to that obtained directly from
16 the D^2 (Irish, 2010). Of course, results are related to the MDS solution so dimensionality may
17 need to be increased to obtain a stress value of ≤ 0.10 (Kruskal and Wish, 1978), which is
18 considered an excellent fit (Borgatti, 1997).
19
20
21
22
23
24
25
26

27 Weighting with relative population sizes (w) may be done to correct for the impact of
28 small groups that commonly lie farther from the regional centroid from genetic drift (Relethford
29 and Crawford, 1995; Leigh et al., 2003). Unfortunately, w is not known in this study because the
30 census categories do not differentiate among South African “Africans” (Christopher, 2011; Stats
31 SA, 2014). Recent approximations are available (e.g., Byrnes, 1996), but they vary from source
32 to source and may not be reliable for the early to mid-20th century date of the Dart Collection.
33
34
35
36
37
38

39 Therefore, following standard procedure w is equal across all samples (Relethford, 1994). Trait
40 heritability can also be included, where r_{ii} , r_{ij} , and F_{st} all decrease when h^2 increases (Relethford
41 and Blangero; 1990; Relethford, 1994; Relethford et al., 1997). When heritability is unknown,
42 $h^2=1$ is the default used to calculate minimum F_{st} ; when known, estimated F_{st} may be calculated
43 (Relethford, 1994). Both are presented here, with the latter conservatively estimated as $h^2=0.65$,
44 i.e., between 0.55 cranial measurements (Relethford, 1994) and 0.80 for dental nonmetric traits
45
46
47
48
49
50
51 of known high heritability, like Carabelli’s cusp (Hughes and Townsend, 2011, [2013](#)).
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

assuming such factors are suggestive of reproductive isolation, the Nguni Ndebele, Swazi, Xosa, and Zulu samples should exhibit close affinities to one another, relative to more distinct Sotho and Tswana who would be similar to each other, as well as the separately-originating Venda.

A second broad hypothesis accounts for the more recent documented interactions of these seven ethnic groups/populations, especially since the Mfecane. Namely, in accordance with the isolation-by-distance model, populations in spatial proximity, regardless of ancestry and current ethnic identity, will exhibit closer affinities to one another than to those living farther away. So, for example, the Venda sample should be more similar to Ndebele than to Xosa (Fig. 1).

And third, at a more specific level certain pairwise sample affinities would be expected, as laid out in the aforementioned hypotheses and documented population history, as follows:

Nguni samples

- Ndebele should appear least like the far southern fellow-Nguni-speaking Xosa (Fig. 1), closer to the Swazi whom they contacted in the Mfecane, and most like the Zulu from whom they likely diverged. Similarities to the Sotho as reflected by language, and now-neighborhood Venda are possible.
- Swazi will be closest to the Zulu sample as a result of shared ancestry, simultaneous founding, and geographic proximity. Of all the Nguni samples, they will be least like the Xosa. The historic contact in northern Swaziland may be indicated by some affinity with Sotho.
- Zulu will be like other Nguni samples (above) other than Xosa, and appear increasingly divergent from all others that are progressively farther away geographically from the Zululand region.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
- Xosa will appear somewhat distinct from other Nguni samples, despite shared ancestry, because of their location, plus long term Khoekhoe contact. The latter influence, obtained in parallel, may serve to push the sample toward the Sotho and Tswana samples.

15 *Sotho samples*

- 16
17
18
19
20
21
22
23
24
25
26
27
28
- Tswana will be dentally akin to the Sotho sample because of their common origin; the separate interactions with Khoesan, in this case San, around their respective homelands may further serve to link them to one another, and perhaps Xosa, though indirectly (as above).
 - Sotho will share a close affinity with the Tswana sample; however, because the Sotho of Lesotho are closest of all seven Bantu groups to the region's geographic center (Fig. 1), they should be most similar to all others.

29 *Venda sample*

- 30
31
32
33
34
35
36
- The Venda should be divergent. Although they were in contact with the Northern Sotho, they have neither a Sotho nor Nguni background. The Venda are somewhat isolated in the Soutpansberg Mountains, and geographically distant from other Bantu populations.

37 **RESULTS**

38
39
40
41
42

Percentages of individuals across samples that express each of the 36 traits are listed in Table 1. The ASUDAS presence/absence dichotomies are presented beneath each trait name.

43
44
45
46
47
48
49
50
51
52

Dichotomization is based upon each trait's appraised morphological threshold (Haeussler et al., 1988), as ascertained by Scott (1973), Nichol (1990), and others according to standard ASUDAS procedure (Turner, 1985b, 1987; Scott and Turner, 1997). Very small numbers of observations (i.e., <10), designated as "n" in the table, are not an issue; however, the small Ndebele (NDB) and Venda (VEN) samples have four traits between them with fewer observations than desired

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(<15). Some trait variation is evident, including a few that differ by >30% between samples of different origin, such as NDB and VEN for Bushman Canine and Tswana (TSW) and Zulu (ZUL) for UC distal accessory ridge, among others. Moreover, the ZUL, NDB and VEN samples have the most divergent, i.e., highest or lowest, percentages for 11, 12 and 15 traits, respectively. Otherwise values appear generally uniform across samples.

[TABLE 1 HERE]

Model-free analyses

To gain an initial impression of inter-sample affinities a full 36-trait MMD comparison was undertaken (Table 2). Some of the aforementioned trait variation is visualized, such as the divergence of NDB, VEN, and ZUL, but overall uniformity is evident. Only two sample pairs differ significantly; all 0.00s in the table were originally negative MMD values reset to specify no divergence (as above), which resulted from traits with no or minimal discriminatory value influencing the bias correction. The MDS solution for the MMD matrix (Fig. 2) yields an r^2 of 0.972 and Kruskal's stress formula 1 value of 0.062.

[TABLE 2 HERE]

[FIGURE 2 HERE]

The 36 traits were then edited in accordance with the steps outlined above. First, all observations are >10 so no traits required deletion. Second, eight traits having no or minimal discriminatory value were removed; following prior protocol (Irish, 2005, 2006, 2013; Irish et al., 2014). these included traits with no expression across samples, and those with no expression in some plus well under 10% across the remaining samples. These traits are: palatine torus (0.00-2.13%), UI1 double shoveling (0.00-6.67%), UM3 parastyle (0.00-2.94%), mandibular torus (0.00-2.13%), LM1 C1-C2 crest (0.00-6.90%), LM1 protostylid (0.00-7.17%), LC root number

1
2
3
4
5
6
7
8
9 (0.00%), and LM1 root number (0.00-5.56%). The remaining 28 percentages were submitted to
10 PCA to identify additional traits of minimal discriminatory value or, conversely, those most
11 important in driving inter-sample variation. Six unrotated components² accounting for 100% of
12 the total variance were obtained. However, the component matrix (not shown but available from
13 author) reveals a discernable drop-off of strong loadings, i.e., $>|0.5|$ in components 4-6; as such,
14 only the first three components (70.77% of the variance) are listed in Table 3. Group component
15 scores are plotted in Figure 3. On Comp 1 seven strongly positive loadings, particularly for UI2
16 interruption groove (0.894) and LM2 cusp number (0.846), push samples with high occurrences
17 of these traits (NDB and ZUL) toward the positive end of the x-axis (Fig. 3). Strong negative
18 loadings for seven others, notably UI1 winging (-0.816) and Bushman Canine (-0.864), drive the
19 others [VEN and Tswana (TSW)] toward the negative side. Key traits were likewise identified in
20 Comp 2 (y-axis), like LM2 groove pattern (0.794) and LM1 cusp 7 (-0.734), and in Comp 3 (z-
21 axis) [LM2 root number (-0.746)]. As a result four more traits were deleted: P1-P2 odontome
22 [also shown (above) to be significantly dimorphic], rocker jaw, LM1 deflecting wrinkle, and LP1
23 Tomes' root. In the third step, three-remaining trait-pairs are strongly correlated (i.e., >0.5) –
24 UI1 shoveling/UI2 interruption groove ($\tau_b=0.558$), UI2 interruption groove/UI2 tuberculum
25 dentale ($\tau_b=0.751$), and UM2 hypocone/UM1 cusp 5 ($\tau_b=0.518$). Given the very high positive
26 Comp 1 loading (above) for UI2 interruption groove, it was retained, so shoveling and
27 tuberculum dentale, along with hypocone, were dropped. Thus, 21 traits, as indicated by asterisks
28 in Table 3, are available for the final MMD comparison.

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46 [TABLE 3 HERE]

47 [FIGURE 3 HERE]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The MDS solution provides excellent representations, or fit, where $r^2=0.965$ and the Kruskal's stress value is 0.056 (i.e., ≤ 0.10 per Borgatti, 1997). Sample locations (Fig. 4) are unchanged from the 36-trait version (Fig. 2), supporting the claim for robusticity of the MMD ; however, the trait editing, including removal of minimally discriminatory traits, has succeeded in reducing the number of inter-sample 0.000-differences (Table 4) while, accordingly, most MMD values have increased, with six now indicating a significant difference. Greater discrimination will be of value in addressing the objectives set out in the introduction. Of interest, both MDS configuration (Figs. 2 and 4) appears somewhat reminiscent of the general population locations illustrated in the South Africa map (Fig. 1).

[FIGURE 4 HERE]

[TABLE 4 HERE]

Finally, results from the San/Bantu comparisons based on all 36 traits are presented in Table 5. As a non-Bantu outlier, the San (SAN) sample is divergent as indicated by the larger, significant inter-sample distances (compare to Table 4). The Bantu samples appear uniformly distinct from SAN, though with some variation in the expected directions.

[TABLE 5 HERE]

Model-bound analyses

Two R-matrices from the MDS optimally-scaled matrix of MMD distances based on 21 traits are provided in Table 6. For the top diagonal $h^2=1.0$ and for the bottom $h^2=0.65$. Minimum and estimated F_{st} are listed and, as is evident, magnitudes increase when h^2 decreases. The results are largely concordant with phenetic distances, e.g., focusing on the bottom diagonal ZUL ($R_{ii}=0.061$) is farthest from the centroid and SOT (0.020) closest. For illustrative purposes the lower

1
2
3
4
5
6
7
8
9 diagonal of Table 6 is approximated in Figure 5 for comparison to the phenetic variation (Fig. 4).
10 It is an approximation in that negative and positive R_{ij} values were submitted to the distances
11 (i.e., proximities) function in SPSS 21.0 to obtain Euclidean distances, with the latter submitted
12 to MDS. In any event, variation among samples (Fig. 5) results from the R_{ij} measures while R_{ii}
13 distances are represented as lines to the plot centroid. Beyond this, the F_{st} values imply that of the
14 total genetic variation, 2.5-3.8% results from among-group differences. The remaining 96.2-
15 97.5% of the variation (i.e., P or panmictic index) based on dental traits resides within them.
16
17
18
19
20
21

22 [TABLE 6 HERE]

23 [FIGURE 5 HERE]

24
25
26 Lastly, the matrix of among-homeland geographic distances (Table 7) is compared with
27 that from the 21-trait MMD. The Mantel correlation is positive but weak, i.e., $r=0.195$ ($p=0.100$),
28 despite some suggestion of correspondence in geographic and phenetic patterning as summarized
29 by the MDS graphs in Figure 6ab; the most divergent, or outlying, samples in common for both
30 graphs are, clockwise from top, VEN, ZUL, XOS, and TSW, with SOT in a central position. A
31 number of exceptions in location between geographic vs. phenetic distances are apparent,
32 including SWZ and NDB, which appear to have switched positions. When SWZ is removed,
33 $r=0.242$ ($p=0.097$). With SWZ and NDB both gone the value increases to 0.308 ($p=0.079$) – a
34 moderately positive correlation (per Cohen, 1988).
35
36
37
38
39
40
41
42

43 [TABLE 7 HERE]

44 [FIGURE 6 HERE]

45 46 47 DISCUSSION

48 49 Samples

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

So, do ethnic affiliations assigned to individuals comprising the samples, or at least most of them, appear credible based on dental nonmetric analyses? Sample locations in Figures 2-6b that somewhat reflect population provenience (Figs. 1 and 6a) may suggest a tentative “yes,” at least for VEN, ZUL, XOS, TSW, and SOT. However, the most systematic way to approach this problem is in relation to the expectations/hypotheses presented above; each is briefly restated and addressed by means of comparison to the quantitative findings.

Inter-sample relationships should reflect known ancestral origins. As above, some of the trait percentages differ (Table 1), e.g., VEN appears the most divergent, but they are largely uniform across samples. Similarly, phenetic distances are relatively small in magnitude (Table 4), with only six of the 21 inter-sample MMD values differing significantly ($p \leq 0.025$). This across-Bantu homogeneity is demonstrated by the overall equidistant separation in individual MMD distances (Table 5) from SAN. The R-matrix presents similar evidence in the form of off- and on-diagonal values (Table 6), and F_{st} estimates suggesting minimal total genetic variation among the groups. These results, then, reflect previous suggestions of overall sample and population homogeneity based, in fact, on most of the same Dart Collection crania (see De Villiers, 1968; Jacobson, 1967, 1982).

That said, the variation that has been captured is informative. The four Nguni samples are not grouped together in the 21-trait MDS plot but, other than SWZ, they inhabit the same general side of Figure 4 (and 5), with some separation from both Sotho samples and VEN in particular. The SWZ sample is separated from the other Nguni groups based principally on its unexpectedly large MMD distance from ZUL, and the lack of a significant difference from all others (Table 4), as expanded on below. The remaining sample proximities are in the expected directions, with TSW and SOT in the same general vicinity and VEN divergent. The excellent fit of the MDS

1
2
3
4
5
6
7
8
9 solution indicates that sample locations closely correspond with the MMD matrix (Table 4).

10 Therefore, although “contemporary patterns of population structure [often] ‘swamp’ or ‘erase’
11 past history” (Relethford and Crawford, 1995, p 32; Relethford et al., 1997), especially on a
12 regional level, the results roughly mirror language and other pre-modern links (Greenberg, 1963;
13 Nurse et al., 1985; Maylam, 1986, July, 1992; among others). Lane et al. (2002) report similar
14 findings in their genetic study of living Bantu. Thus, at this broad level the ethnic/population
15 identities of these samples cannot be precluded, with the potential exception of SWZ.
16
17
18
19
20
21

22 ***Samples of populations that lived near one another will exhibit greater similarity than to those***

23 ***farther away.*** The among-sample plots illustrating phenetic and approximating genetic measures

24 (Figs. 2, 4-6) appear at least qualitatively, to be comparable to the historic population locations
25 (Fig. 1). Some variation is patent (Fig. 6) but the general dental-derived pattern noted above is
26 recurrent. Two obvious exceptions are SWZ and NDB, which appear most out of place relative
27 to Figure 1. Mantel correlations help corroborate these observations by testing the isolation-by-
28 distance model (Wright, 1943), which addresses the expectation/hypothesis of a link between
29 geographic and phenetic proximities. With straight-line distances in the unidimensional stepping-
30 stone variant, the assumption is that an infinite number of subpopulations live along each linear
31 habitat, where they exchange migrants at an equivalent rate with adjacent subpopulations. Some
32 minimal gene flow may also occur with nonadjacent subpopulations and an external source (e.g.,
33 Khoesan) of infinite size (Kimura and Weiss, 1964; Konigsberg, 1990; Schillaci et al., 2009).
34
35
36
37
38
39
40
41
42
43
44

45 The bottom line, though, is that greater phenetic (and by proxy genetic) and geographic distances
46 are linked.
47

48
49 The Mantel correlation among all seven Bantu samples is 0.195, which is a positive but
50 weak r (Cohen, 1988). However, r increases to 0.242 after SWZ is dropped from analysis and,
51
52

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

when NDB is removed, r increases further to 0.308 – a moderate positive correlation (Cohen, 1988). Lane et al. (2002) reported a similar correspondence ($r=0.399$, $p=0.04$) between their geographic and genetic distances based on Y-chromosome haplotypes among many of the same Bantu populations. Again, for SWZ the MMD distances (Table 4) seem counter to those expected, so the sample may not be representative of the Swazi population. On the other hand, the reasons for the NDB phenetic/-geographic discrepancy may have more to do with the complex origin of the Ndebele, in tandem with the current occupation of their post-Mfecane homeland, both of which are recent occurrences relative to other groups. Indeed, their phenetic position (Fig. 6b) appears more in line with their proposed original homeland location (see above) in far east-central South Africa near Zululand.

In any event, given the: 1) approximate geographic locations and vast size of homelands, 2) use of linear distances that likely do not reflect reality on the South African landscape, and 3) regional scope of study where gene flow beyond that envisioned by the model is documented [contra broader continental and global scales (Scott and Turner, 1997)], the correlation of 0.308 after removal of potential outliers does seem supportive of the isolation-by-distance model. That is, at this second broad level of examination the results do not contradict most assigned ethnic/population identities of these samples, with the potential exception of SWZ (and perhaps NDB).

The seven hypothesized among-sample relationships should be identified by the quantitative analyses. The Ndebele (NDB) sample is slightly more divergent from the geographically-distant fellow-Nguni XOS (Table 4, MMD=0.038) than neighboring ZUL (0.033) – from whom they purportedly branched (Van Warmelo, 1962; Seligman, 1967), and SWZ (0.000) – as illustrated by MDS (Fig. 4); none of these phenetic distances differ significantly. NDB appears similar to SOT (MMD=0.000), which is not unexpected given the influence of Sotho peoples living near

1
2
3
4
5
6
7
8
9 them [e.g., seSotho attributes in their isiNdebele language (Van Warmelo, 1962; Byrnes, 1996)],
10 though not to TSW who long ago emigrated westward (MMD=0.052 $p>0.025$). Off-diagonal
11 measures of genetic distance (Table 6; Fig. 5) provide analogous results, where positive numbers
12 indicate greater similarity [e.g., NDB/ZUL $r_{ij}=0.00034$ (rounded to 0.000) in lower diagonal of
13 Table 6], and negative numbers less similarity than on average (NDB/ TSW $r_{ij}=-0.023$). As a
14 likely result of the Mfecane-prompted displacement (July, 1992) NDB also shows some affinity
15 to the now-neighboring VEN (MMD=0.031). In fact, it is this 19th century in-country movement
16 that appears to prevent a “better” correlation between geographic and phenetic distances. So, in
17 accordance with hypothesized expectations, the ethnic/population identity of the NDB sample
18 may be credible.
19
20
21
22
23
24
25
26

27
28 The Swazi (SWZ) sample, contrary to expectations from shared ancestry and proximity
29 (Van Warmelo, 1962), is least akin to nearby Nguni ZUL based on phenetic (Table 4, MMD=
30 0.037 $p>0.025$) and genetic measures ($r_{ij}=-0.012$ in lower diagonal of Table 6). Further, it is most
31 like far-flung XOS (MMD=0.000; $r_{ij}=0.001$) according to individual distances and MDS plots
32 (Figs. 4-6). For that matter, SWZ appears similar to all samples *other* than ZUL (Tables 4 and 6).
33 The r_{ii} values are particularly instructive; SWZ is second closest (0.023, Table 6) of all samples
34 to the regional centroid, which is suggestive of much internal heterogeneity, or heterozygosity.
35 Finally, other than NDB, SWZ is the most divergent in geographic vs. phenetic location *relative*
36 to other samples (Fig. 6). All told, these findings imply that the SWZ sample is a heterogeneous
37 amalgamation not representative of the 20th century Swazi population – whether from recent
38 (undocumented) population changes that obscure relationships (Relethford and Crawford, 1995;
39 Relethford et al., 1997) or, simply, issues with the assigned ethnicity or other sampling problems.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9 The Zulu (ZUL) sample does show some resemblance to their purported offshoots, NDB
10 (MMD=0.033, r_{ij} =0.00034) (Van Warmelo, 1962; Seligman, 1967), and are more distinct from
11 Nguni XOS (MMD=0.054 $p>0.025$, r_{ij} =-0.013) (Tables 4 and 6, Figs. 4-5), and the ostensible
12 SWZ outlier that will not be discussed further. It is divergent from TSW and VEN as expected,
13 but like SOT (MMD=0.013, r_{ij} =0.001) as elaborated upon below. As well, ZUL is farthest from
14 the centroid (r_{ii} =0.061, Table 6; Fig. 5). The r_{ii} value suggests it is the most homogeneous of all
15 samples, plausibly due to higher outside- and lower local admixture (Konigsberg, 2006) and drift
16 (Relethford et al., 1997); unfortunately, any effects of the latter mechanism cannot be quantified
17 due to the shortcomings of regional census data that prevent estimates of relative population size.
18 Perhaps these results are explainable by the Zulu role in eliminating and driving out other eastern
19 Bantu groups during the Mfecane. In any event, the ethnic/population identity of the ZUL sample
20 cannot be rejected here.
21
22
23
24
25
26
27
28
29
30

31 The Xosa (XOS) sample differs to some extent from NDB, though not significantly (i.e.,
32 MMD=0.033 $p\leq 0.025$), and is more divergent from Nguni ZUL as noted (Tables 4 and 6; Figs.
33 4-5), in accordance with among-sample expectations. It is distinct from VEN (MMD=0.044
34 $p>0.025$, r_{ij} =-0.017). Also as expected an affinity is evident with SOT (MMD=0.008) and TSW
35 (MMD=0.006), perhaps in part due to the San admixture (Van Warmelo, 1962; Maylam, 1986;
36 Byrnes, 1996) posited to have occurred in parallel. Such influence is apparent (Table 1) by the
37 high incidence of Bushman Canine (31.58%), among others (e.g., root reduction), known to be
38 characteristic of the Khoesan (Haeussler et al., 1989; Irish, 1993, 1997); XOS also exhibits the
39 lowest MMD distance (0.042) from SAN (Table 5). All of this, plus the phenetic/geographic
40 correspondence (Fig. 6) suggests XOS likely is representative of the Xosa population.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9 The Tswana (TSW) sample is close to SOT phenetically (MMD=0.001) (Table 4), though
10 they share a small negative r_{ij} (-0.003) (Table 6 and Figs. 4-5). This affinity was expected from
11 common ancestry as supported by related languages (Byrnes, 1996), and perhaps from gene flow
12 with the Khoesan-speaking San (Van Warmelo, 1962; July, 1992) after reaching their respective
13 homelands (Fig. 1). Khoesan influence is evident in TSW, like with XOS (Table 1), based on the
14 second lowest MMD distance (0.046) (Table 5). In sum the results, including largely concordant
15 geographic and phenetic locations (Fig. 6), do not discount the Tswana identity of this sample.
16
17
18
19
20
21

22 The Sotho (SOT) sample is much like TSW, including a resemblance to SAN (Tables 1
23 and 5); and, as anticipated based on their interaction with many populations since reaching South
24 Africa (Van Warmelo, 1962; Byrnes, 1996; Hall, 2010) and centralized location (Fig. 1), they are
25 highly comparable to all remaining Bantu samples (Table 4; Fig. 4) including VEN. The latter
26 affinity may result from Venda contact with Northern Sotho in the region where present-day
27 Southern Sotho (SOT) were first established (Fig. 1). The resemblance of SOT to all Bantu
28 samples is sustained by the R-matrices (Table 6) that, despite mostly negative (though small) r_{ij}
29 values, show SOT to be closest (0.021; Table 6) of all samples to the regional centroid (also Fig.
30 5). Again, this proximity is indicative of marked heterogeneity likely resulting from gene flow
31 expected in a group geographically nearest all other populations. Therefore, indications are that
32 the sample is representative of the Sotho population.
33
34
35
36
37
38
39
40
41
42

43 The Venda (VEN) sample, finally, presents many extreme percentages that differentiate it
44 from others (Table 1). Relative to the three Nguni samples that remain under discussion, VEN is
45 phenetically similar to neighboring NDB (above) but distinct from ZUL (MMD= 0.062 $p>0.025$,
46 $r_{ij}=-0.013$) and XOS (MMD=0.044 $p>0.025$, $r_{ij}=-0.008$) (Tables 4 and 6; Figs. 4-5), as expected
47 based on population history (Loubser, 1989). Otherwise, it is closely akin to SOT (MMD =0.000,
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9 $r_{ij}=-0.003$) and TSW (MMD=0.005, $r_{ij}=0.003$), which is may be explainable by the Venda link
10 with neighboring Northern Sotho (Loubser, 1989). In sum, these results and phenetic/geographic
11 concordance (Fig. 6) cannot disprove the Venda identity of this sample.
12
13

14 15 16 17 **Populations**

18
19 Again, prior skeletal analyses of these Dart Collection samples, based on craniometrics
20 and qualitative comparisons of dental morphometric data (Jacobson, 1967, 1982; De Villiers,
21 1968, p. 201), found overall homogeneity – prompting the authors to conclude that “the [Bantu]
22 ‘tribal’ series may be regarded as samples of a single South African ‘Negro’ population.” Dental
23 nonmetric-based analyses similarly reveal: 1) general trait uniformity across the samples (Table
24 1), 2) low inter-sample MMD distances, many of which are not significant (Table 4), 3) largely
25 uniform significant distances of all Bantu samples from the San outlier (Table 5), and 4) minimal
26 inter-sample measures of genetic differentiation (Table 6). So migration, gene flow, and drift
27 since the arrival of the original Nguni and Sotho immigrant groups, particularly during recent
28 history, apparently did play roles in erasing past history (see Relethford and Crawford, 1995;
29 Relethford et al., 1997). Dart (1937) himself wrote of widespread admixture in South African
30 Bantu and others.
31
32

33
34 That said, the present study also identifies variation that the prior skeletal analyses did
35 not. In this way, the results parallel those of a genetic study that includes five of the seven, albeit
36 modern, groups examined here. Lane and colleagues (2002, p. 178) found “very little genetic
37 differentiation among . . . [these] southeastern Bantu-speakers,” yet could discern differences
38 along population lines. So at least in this case, measures of divergence from nonmetric traits do
39 appear comparable with those from molecular data (per Leigh et al., 2003). Specifically, model-
40
41
42
43
44
45
46
47
48
49
50
51
52

1
2
3
4
5
6
7
8
9 free analyses of autosomal and Y-haplotype data yielded genetic affinities that also link Zulu
10 with Xosa, relative to correspondingly related Sotho and Tswana, and distinct Venda (Lane et al.,
11 2002). As above, their model-bound results include an analogous correlation between geographic
12 and genetic distances based on Y-chromosome haplotypes. And, although their estimated F_{st} of
13 0.014 from Y-haplotypes is lower than the 0.025 in the current study (Table 6), both indicate
14 “little differentiation” using the qualitative guidelines of Wright (1969), i.e., 0.00-0.05. Thus, If
15 not sample- or data-related, this difference in F_{st} magnitude is plausibly an indicator of ongoing
16 detribalization (per De Villiers, 1968) and increasing gene flow among populations between the
17 early/mid-20th and early 21st centuries.

18
19
20
21
22
23
24
25
26 The likelihood that inter-group genetic variation of 2.5% (with a corresponding intra-
27 group P of 97.5%), based on 21 dental nonmetric traits, is representative of 20th century South
28 African Bantu populations is supported by a range of studies. Craniometric data ($h^2=0.55$) used
29 by Relethford (2001) obtained an estimate of 9% [“moderate differentiation” between 0.05-0.15
30 according to Wright (1969)] among groups distributed across the entire sub-continent. Similarly,
31 the present minimum F_{st} of 0.038 is less than the 0.059 attained from nine dispersed groups from
32 western, eastern, and southern Africa based on 13 dental nonmetric traits (Irish, 2010). Such
33 differences in F_{st} magnitude would be expected when comparing regional- vs. continental-scale
34 populations. For example, Lane et al. (2002, p- 178) state their Bantu groups have
35 “approximately one-tenth of the between-group autosomal variance and about half of the
36 between-group Y-chromosome variance found among African populations from widely
37 separated locations.” On the other hand, by way of methodological comparison, the present
38 minimum F_{st} (0.025) is five times higher than those (0.005-0.007) for three periods from a much
39 smaller geographic region in Ireland (Relethford et al., 1997). Thus, again, F_{st} values based on
40
41
42
43
44
45
46
47
48
49
50
51
52

1
2
3
4
5
6
7
8
9 dental nonmetric data may indeed provide a realistic indication of population differentiation
10 among these Bantu groups – individual sample representativeness (e.g., SWZ) notwithstanding.

11
12 Finally, with this wide range of information in mind, the question concerning use of these
13 seven samples as proxies for premodern populations in bioarchaeological research is revisited. In
14 other words, does this “. . . synchronic snapshot” of “among-group variation . . . [reflect *enough*]
15 past population history” (Relethford and Crawford, 1995, p. 25, 32) to permit diachronic studies
16 of Bantu origins and affinities? Depending on the level of analysis, i.e., regional vs. continental,
17 the answers are “conditionally yes” and “yes.” Except for SWZ, the samples seem representative
18 *enough* of their respective ethnic groups/populations to reconstruct historic change in genetic and
19 demographic structure. If NDB is removed, more ancient links in eastern South Africa, at least
20 before the 19th century, may be explored. At a continental level all samples, being representative
21 of a single South African Bantu population, should be useful for comparison with other regional
22 pooled groups; studies at this level are likely to reflect longer-term patterns consistent with major
23 events in population history (Relethford et al., 1997), including the “‘Bantu’ expansion.”
24
25
26
27
28
29
30
31
32
33
34

35 SUMMARY

36 Model-free and model-bound quantitative analyses of dental nonmetric traits were used
37 to help (re)confirm the validity of ethnic group identities attributed to individuals from seven
38 “Bantu” samples in the Raymond A. Dart Collection: Ndebele, Sotho, Swazi, Tswana, Venda,
39 Xosa, and Zulu. Information was also obtained concerning whether the samples that do appear
40 representative best reflect past or recent patterns of population structure. The goal was to assess
41 whether these synchronic samples (n=408 individuals) can be used to yield credible diachronic
42 estimates of population affinity and history in bioarchaeological research. It appears that five to
43 six samples are largely representative of their attributed populations from early- to mid-20th
44 century South Africa, in that they: 1) display phenetic variation in line with documented
45
46
47
48
49
50
51
52
53

1
2
3
4
5
6
7
8
9 population history including evidence of ancestry with initial immigrants from eastern Africa, 2)
10 indicate a moderately positive correlation between phenetic and geographic distances relative to
11 the isolation-by-distance model, and 3) though evidencing minimal among-group differentiation
12 do provide measures of genetic distance in support of the phenetic distances. Whether related to
13 collection issues, recent swamping of past genetic structure, or both – it appears that only the
14 Swazi (SWZ) and perhaps Ndebele (NDB) samples may not be suitable for population history
15 study.
16
17
18
19
20
21

22 ACKNOWLEDGEMENTS

23
24 I thank Marin Pilloud and Joseph Hefner for inviting me to their 2015 AAPA symposium
25 in which much of this information was first presented. Thanks also go to Brendan Billings, the
26 Curator of the Dart Collection at the University of the Witwatersrand, and everybody else there
27 who helped me out. Funding was provided by the National Science Foundation (BCS-0840674).
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

LITERATURE CITED

Abbé EN, Ribot I, Steyn M. 2006. A craniometric study of the 20th century Venda. South Afr Archaeol Bull 61:19–25.

Barker G. 2006. The agricultural revolution in prehistory: Why did foragers become farmers? New York: Oxford University Press.

Bermudez de Castro JM. 1989. The Carabelli trait in human prehistoric populations of the Canary Islands. Hum Biol 61:117-131.

Borg I, Groenen P. 1997. Modern multidimensional scaling. New York: Springer-Verlag.

Borgatti SP. 1997. Multidimensional scaling. www.analytictech.com/borgatti/mds.htm.

Burnett SE, Irish JD, Fong MR. 2013. Wears the problem? Examining the effect of dental wear on studies of crown morphology. In: Scott GR, Irish JD, editors. Anthropological perspectives on tooth morphology: Genetics, evolution, variation. Cambridge: Cambridge University Press. p 535-553.

Burnett SE. 2016. Crown wear: Identification and categorization. In: Irish JD, Scott GR, editors. A companion to dental anthropology. London: Wiley Blackwell p 415-432.

Butler J, Rotberg RI, Adams J. 1978. The Blacks homelands of South Africa: The political and economic development of Bophuthatswana and Kwa-Zulu. Berkeley: University of California Press.

Butler PM. 1939. Studies of the mammalian dentition. Differentiation of the post-canine dentition. Proc Zool Soc Lon B, 109:1–36.

- 1
2
3
4
5
6
7
8
9 Butler PM. 2001. What happened to the field theory? In: Brook AH, editor. Dental
10 morphology. Sheffield: Sheffield Academic Press Ltd. p 3-12.
11
12
13 Byrnes RM. 1996. South Africa: A country study. Washington: GPO for the Library of
14 Congress.
15
16
17 Christopher AJ. 2011. The Union of South Africa censuses 1911-1960: An incomplete record.
18 Historia 56:1-18.
19
20
21 Cohen J. 1988. Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale:
22 Erlbaum.
23
24
25 Collett DP. 1982. Models of the spread of the Early Iron Age. In: Ehret C, Posnansky M,
26 editors. The archaeological and linguistic reconstruction of African history. Berkeley:
27 University of California Press. p 182-198.
28
29
30
31 Cox TF, Cox MAA. 1994. Multidimensional scaling. London: Chapman and Hall.
32
33
34 Dahlberg AA. 1945. The changing dentition of man. J Am Dent Assoc 32:676-690.
35
36
37 Dart RA. 1937. Racial origins in the Bantu-speaking tribes of South Africa. In: Schapera L,
38 editor. The Bantu-speaking tribes of South Africa: an ethnographical survey. London: George
39 Routledge & Sons, Ltd. p 1-31.
40
41
42
43 Davidson B. 1974. Africa in history. New York: Macmillan Publishing Company.
44
45
46 Dayal MR, Kegley ADT, Strkalj G, Bidmos MA, Kuykendall KL. 2009. The history and
47 composition of the Raymond A. Dart Collection of Human Skeletons at the University of the
48 Witwatersrand, Johannesburg, South Africa. Am J Phys Anthropol. 140:324-335.
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9 De Villiers H. 1968. The skull of the South African Negro. Johannesburg: Witwatersrand
10 University Press.
11
12
13 Denbow JR. 1981. Broadhurst: A 14th century AD expression of the Early Iron Age in south-
14 eastern Botswana. S Afr Archaeol Bull 66-74.
15
16
17 Ehret C. 2000. Language and history. In: Heine B, Nurse D, editors. African languages: An
18 introduction. Cambridge: Cambridge University Press. p 272-297.
19
20
21 Ersts PJ. Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural
22 History, Center for Biodiversity and Conservation. Available from
23 http://biodiversityinformatics.armh.org/open_source/gdrrg. Accessed on 2014-7-9.
24
25
26
27 Excoffier L, Pellegrini B, Sanchez-Mazas A, Simon C, Langaney A. 1987. Genetics and
28 history of sub-Saharan Africa. Yrbk Phys Anthropol 30:151-194.
29
30
31
32 Green R, Suchey J. 1976. The use of inverse sine transformation in the analysis of non-
33 metrical data. Am J Phys Anthropol 45:61-68.
34
35
36
37 Greene R, Suchey J, Gokhale D. 1979. The statistical treatment of correlated bilateral traits in
38 the analysis of cranial material. Am J Phys Anthropol 50:629-634.
39
40
41
42 Greenberg JH. 1963. The languages of Africa. The Hague: Mouton & Co.
43
44
45 Haeussler AM, Turner CG II, Irish JD. 1988. Concordance of American and Soviet methods in
46 dental anthropology. Am J Phys Anthropol 75:218.
47
48
49 Hall S. 2010. Farming communities of the second millennium: Internal frontiers, identity,
50 continuity and change. In: Hamilton C, Mbenga BK, Ross R, editors. Cambridge history of
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9 South Africa. Cambridge: Cambridge University Press. p 112-167.

10
11 Hanihara T. 1992. Dental and cranial affinities among populations of East Asia and the
12 Pacific: The basic populations in East Asia, IV. *Am J Phys Anthropol* 88:163-182.

13
14
15 Harris EF, Sjøvold T. 2004. Calculation of Smith's mean measure of divergence for intergroup
16 comparisons using nonmetric data. *Dent Anthropol* 17:83-93.

17
18
19
20 Hiernaux J. 1975. *The people of Africa*. New York: Charles Scribner's Sons.

21
22 Hintze JL. 2007. *NCSS help system*. Kaysville, Utah: NCSS, LLC.

23
24
25 Huffman TN. 1988. Southern Africa to the south of the Zambezi. In: Hrbek I, editor. *General*
26 *history of Africa III: Africa from the seventh to the eleventh century*. Berkeley: University of
27 California Press. p 664-680.

28
29
30
31 Huffman TN. 1989. Ceramics, settlements and late Iron Age migrations. *Afr Arch Review*
32 7:155-182.

33
34
35 Huffman TN. 2007. *Handbook to the Iron Age: The archaeology of pre-colonial farming*
36 *societies in southern Africa*. Pietermaritzburg: University of Natal Press

37
38
39
40 Hughes TE, Townsend GC. 2011. Twin studies of dental crown morphology: Genetic and
41 environmental determinants of the Cusp of Carabelli. In *Program 15th International*
42 *Symposium on Dental Morphology*, 37. Newcastle, UK.

43
44
45
46 Hughes TE, Townsend GC. 2013. Twin and family studies of human dental crown
47 morphology: genetic, epigenetic, and environmental determinants of the modern human
48 dentition. In: Scott GR, Irish JD, editors. *Anthropological perspectives on tooth morphology:*
49

1
2
3
4
5
6
7
8
9 genetics, evolution, variation. Cambridge: Cambridge University Press. p 31-68.

10
11 Iliffe J. 1997. Africans: The history of a continent. Cambridge: Cambridge University Press.

12
13 Irish JD, Black W, Sealy J, Ackermann R. 2014. Questions of Khoesan continuity: Dental
14 affinities among the indigenous Holocene peoples of South Africa. *Am J Phys Anthropol*
15 155:33-44.
16
17

18
19
20 Irish JD. 1993. Biological affinities of late Pleistocene through modern African aboriginal
21 populations: The dental evidence. Ph.D. Dissertation, Arizona State University, Tempe.

22
23 Irish JD. 1997. Characteristic high- and low-frequency dental traits in sub-Saharan African
24 populations. *Am J Phys Anthropol* 102:455–467.
25
26

27
28 Irish JD. 2005. Population continuity versus discontinuity revisited: Dental affinities among
29 Late Paleolithic through Christian era Nubians. *Am J Phys Anthropol* 128:520-535.
30
31

32
33 Irish JD. 2006. Who were the ancient Egyptians? Dental affinities among Neolithic through
34 post- dynastic peoples. *Am J Phys Anthropol* 129:529-543.
35
36

37
38 Irish JD. 2010. The mean measure of divergence (MMD): its utility in model-free and model-
39 bound analyses relative to the Mahalanobis D^2 distance for nonmetric traits. *Am J Hum Biol*
40 22:378–395.
41
42

43
44 Irish JD. 2013. Afridonty: the "Sub-Saharan African Dental Complex" revisited. In: Scott
45 GR, Irish JD, editors. *Anthropological perspectives on tooth morphology: genetics, evolution,*
46 *variation.* Cambridge: Cambridge University Press. p 278–295.
47
48

49
50 Jacobson A. 1967. The Bantu dentition: A morphological and metrical study of the teeth, the
51
52

1
2
3
4
5
6
7
8
9 jaws and the bony palate of several large groups of South African Bantu-speaking Negroids.

10 Ph.D. Thesis, University of the Witwatersrand, Johannesburg.

11
12
13 Jacobson A. 1982. The dentition of the South African Negro. Anniston, AL: Higgenbotham,
14
15 Inc.

16
17 Jenkins T, Zoutendyk A, Steinberg AG. 1970. Gammoglobulin groups (Gm and Inv) of
18
19 various southern African populations. Am J Phys Anthropol 32:197-218.

20
21 July RW. 1992. A history of the African people. Prospect Heights, IL: Waveland Press.

22
23 Kruskal JB, Wish M. 1978. Multidimensional scaling. Beverly Hills, CA: Sage Publications.

24
25
26
27 Konigsberg LW. 2006. A post-Neumann history of biological and genetic distance studies in
28
29 bioarchaeology. In: Buikstra JE, Beck LA, editors. Bioarchaeology: the contextual analysis of
30
31 human remains. New York: Academic Press. p 263–279.

32
33 Lane AB, Soodyall H, Arndt S, Ratshikhopha ME, Jonker E, Freeman C, Young L, Morar B,
34
35 Toffie L. 2002. Genetic substructure in South African Bantu-speakers: Evidence from
36
37 autosomal DNA and Y-chromosome studies. Am J Phys Anthropol 119:175-185.

38
39 Larsen CS. 1997. Bioarchaeology. Cambridge: Cambridge University Press.

40
41 Leigh SR, Relethford JH, Park PB, Konigsberg LW. 2003. Morphological differentiation of
42
43 *Gorilla* subspecies. In: Taylor AB, Goldsmith ML, editors. Gorilla biology: a multidisciplinary
44
45 perspective. Cambridge: Cambridge University Press. p 104–131.

46
47 Loubser JHN. 1989. Archaeology and early Venda history. S Afr Archaeol Soc Goodwin
48
49 Series 6:54-61.

1
2
3
4
5
6
7
8
9 Lwango-Lunyiigo SL, Vansina J. 1988. The Bantu-speaking peoples and their expansion. In:
10 Hrbek I, editor. General history of Africa III: Africa from the seventh to the eleventh century.
11 Berkeley: University of California Press. p 140-162.
12

13
14 Martinon-Torres M, Bermudez de Castro J, Gomez-Robles A, Arsuaga J, Carbonell E,
15 Lordkipanidze D, Manzi G, Margvelashvili A. 2007. Dental evidence on the hominin
16 dispersals during the Pleistocene. Proc Nat Acad Sci 104:13279–13282.
17
18

19
20
21 Maylam P. 1986. A history of the African people of South Africa: From the early iron age to
22 the 1970's. New York: St. Martin's Press.
23

24
25 Mitchell P. 2010. Genetics and southern African prehistory: An archaeological view. J
26 Anthropol Sci 88:73-92.
27

28
29 Monnig HO. 1967. The Pedi. Pretoria: JL Van Schaik LTD.
30

31
32 Morris AG. 1992. The skeletons of contact: A study of protohistoric burials from the Lower
33 Orange River valley, South Africa. Johannesburg, Witwatersrand University Press.
34

35
36 Newman JL. 1995. The peopling of Africa: A geographic interpretation. New Haven: Yale
37 University Press.
38

39
40 Nichol CR, Turner II CG. 1986. Intra- and interobserver concordance in classifying dental
41 morphology. Am J Phys Anthropol 69:299-315.
42

43
44 Nichol CR. 1990. Dental genetics and biological relationships of the Pima Indians of Arizona.
45
46 Ph.D. Dissertation, Arizona State University, Tempe, AZ.
47

48
49 Nikita E. 2015. A critical review of the mean measure of divergence and Mahalanobis
50
51

1
2
3
4
5
6
7
8
9 distances using artificial data and new approaches to the estimation of biodistances employing
10 nonmetric traits. *Am J Phys Anthropol* 157:284-294.

11
12 Nurse GT, Weiner JS, Jenkins T. 1985. *The peoples of southern Africa and their affinities*.
13 Oxford: Clarendon Press.

14
15
16
17 Osborn JW. 1978. Morphogenetic gradients: Fields versus clones. In: Butler PM, Joysey KA,
18 editors. *Development, function and evolution of teeth*. London: Academic Press. p 171-201.

19
20
21
22 Parkington JE. 1981. Southern Africa: Hunters and food-gatherers. In: Mokhtar G, editor.
23 *General history of Africa II: Ancient civilizations of Africa*. Berkeley: University of California
24 Press. p 639-670.

25
26
27
28 Phillipson DW. 1994. *African archaeology*. Cambridge: Cambridge University Press.

29
30
31
32 Relethford J, Blangero J. 1990. Detection of differential gene flow from patterns of
33 quantitative variation. *Hum Biol* 62:5-25.

34
35
36
37 Relethford JH, Crawford MH. 1995. Anthropometric variation and the population history of
38 Ireland. *Am J Phys Anthropol*. 96:25-38.

39
40
41
42 Relethford JH, Harpending HC. 1994. Craniometric variation, genetic theory, and modern
43 human origins. *Am J Phys Anthropol*. 95:249-270.

44
45
46
47 Relethford JH, Crawford MH, Blangero J. 1997. Genetic drift and gene flow in post-famine
48 Ireland. *Hum Biol* 69:443-465.

49
50
51
52 Relethford JH. 1994. Craniometric variation among modern human populations. *Am J Phys*
53 *Anthropol* 95:53-62.

- 1
2
3
4
5
6
7
8
9 Relethford J. 2001. Global analysis of regional differences in craniometric diversity and
10 population substructure. *Hum Biol* 73:629–636.
11
12
13 Relethford J. 2004. Global patterns of isolation by distance based on genetic and
14 morphological data. *Hum Biol* 76:499-513.
15
16
17 Rightmire GP. 1999. Dental variation and human history. *Rev Archaeol* 20:1-3.
18
19
20 Ruhlen M. 1994. *The origin of language: Tracing the evolution of the mother tongue*. New
21 York: John Wiley & Sons, Inc.
22
23
24 Schillaci MA, Irish JD, Wood C. 2009. Further analysis of the population history of ancient
25 Egyptians. *Am J Phys Anthropol* 139:235-243.
26
27
28 Schoenbrun D. 2001. Representing the Bantu expansions: What's at stake? *Int J Afr Hist*
29 *Studies* 34:1-4.
30
31
32
33 Scott GR, Turner CG II. 1997. *The anthropology of modern human teeth: Dental morphology*
34 *and its variation in recent human populations*. Cambridge: Cambridge University Press.
35
36
37 Scott GR. 1973. *Dental morphology: A genetic study of American white families and variation*
38 *in living Southwest Indians*. Ph.D. Dissertation. Arizona State University, Tempe, AZ.
39
40
41 Scott GR. 1980. Population variation of Carabelli's trait. *Hum Biol* 52:63-78.
42
43
44 Seligman CG. 1967. *Races of Africa*, 4th ed. Oxford: Oxford University Press.
45
46
47 Sjøvold T. 1973. Occurrence of minor non-metrical variants in the skeleton and their
48 quantitative treatment for population comparisons. *Homo* 24:204-233.
49
50
51 Sjøvold T. 1977. Non-metrical divergence between skeletal populations: The theoretical
52

1
2
3
4
5
6
7
8
9 foundation and biological importance of C.A.B. Smith's mean measure of divergence. *Ossa* 4,
10 Suppl 1:1-133.

11
12
13 Smith P, Shegev M. 1988. The dentition of Nubians from Wadi Halfa, Sudan: An evolutionary
14 perspective. *J Dent Assoc S Afr* 43:539-541.

15
16
17 Smouse PE, Long JC, Sokal RR. 1986. Multiple regression and correlation extensions of the
18 Mantel test of matrix correspondence. *Syst Zool* 35:627-632.

19
20
21 Soper RC. 1982. Bantu expansion into eastern Africa. In: Ehret C, Posnansky M, editors. The
22 archaeological and linguistic reconstruction of African history. Berkeley: University of
23 California Press. p 223-238.

24
25
26
27
28 Statistics South Africa. 2014. Mid-year population estimates. Statistical release P0302.
29 www.statssa.gov.za.

30
31
32 Stojanowski CM, Johnson KM. 2015. Observer error, dental wear, and the inference of New
33 World sundadonty. *Am J Phys Anthropol* 156(3):349-362.

34
35
36 Sutton JEG. 1981. East Africa before the seventh century. In: Mokhtar G, editor. General
37 history of Africa II: Ancient civilizations of Africa. Berkeley: University of California Press.
38 p 568-592.

39
40
41 Tal H, Tau S. 1983. Statistical survey of the human skulls in the Raymond Dart Collection of
42 Skeletons. *S Afr J Sci* 79:215-217.

43
44
45 Tobias PV. 1974. *Biology of the South African Negro*. Cape Town: University of the
46 Witwatersrand.

- 1
2
3
4
5
6
7
8
9 Townsend G, Brook A, Yong R, Hughes T. 2016. Tooth classes, field concepts, and
10 symmetry. In: Irish JD, Scott GR, editors. A companion to dental anthropology. London:
11 Wiley Blackwell p 171-201.
12
13
14
15 Turner CG II. 1985a. The dental search for Native American origins. In: Kirk R, Szathmary E,
16 editors. Out of Asia. Canberra: The Journal of Pacific History. p 31-78.
17
18
19 Turner CG II. 1985b. Expression count: A method for calculating morphological dental trait
20 frequencies by using adjustable weighting coefficients. Am J Phys Anthropol 68:263-268.
21
22
23 Turner CG II. 1987. Late Pleistocene and Holocene population history of East Asia based on
24 dental variation. Am J Phys Anthropol 73:305-322.
25
26
27
28 Turner CG, II, Nichol CR, Scott GR. 1991. Scoring procedures for key morphological traits of
29 the permanent dentition: the Arizona State University dental anthropology system. In: Kelley
30 MA, Larsen CS, editors. Advances in dental anthropology. New York: Wiley-Liss. p 13-32.
31
32
33
34 Turner CG II, Scott GR. 1977. Dentition of Easter Islanders. In: Dahlberg AA, Graber TM,
35 editors. Orofacial growth and development. The Hague: Mouton. p 229-249.
36
37
38
39 Van Warmelo NJ. 1962. Grouping and ethnic history. In: Schapera I, editor. Bantu-speaking
40 tribes of South Africa: An ethnographical survey. London: Routledge & Kegan Paul, Ltd. p
41 43-66.
42
43
44
45 Vogel JO. 1997. Bantu expansion. In: Vogel JO, editor. Encyclopedia of precolonial Africa:
46 Archaeology, history, languages, cultures, and environments. Walnut Creek: Alta Mira Press.
47 p 435-438.
48
49
50
51 Warren KA, Hall S, Ackermann RR. 2014. Craniodental continuity and change between Iron
52

1
2
3
4
5
6
7
8
9 Age peoples and their descendants. S Afr J Sci. 110, Art. #'2013-0275.

10
11 Williams-Blangero S, Blangero J. 1989. Anthropometric variation and the genetic structure of
12 the Jirels of Nepal. Hum Biol. 61:1-12.

13
14
15 Wright S. 1943. Isolation by distance. Genetics 28:114–138.

16
17
18 Wright S. 1969. Evolution and the genetics of populations. Vol. 2. The theory of gene
19 frequencies. Chicago: University of Chicago Press.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FOOTNOTES

¹All analyses were also conducted with the latter statistic to further corroborate these inferences (Irish, 2010); this highly concordant output is available from the author and is planned for presentation elsewhere.

²Varimax rotation delivers analogous results (output available from the author).

TABLE 1. Dental trait percentages (%) and number of individuals scored (n) for the seven South African "Bantu" samples

	Trait ²	Sample ¹	NDB	SOT	SWZ	TSW	VEN	XOS	ZUL
1)	Winging UI1 (+=ASU 1)	% n	0.00 27	1.92 52	2.94 34	4.17 48	10.53 38	3.45 58	0.00 57
2)	Labial Curvature UI1 (+=ASU 2-4)	% n	71.43 14	60.61 33	61.11 18	68.57 35	60.71 28	61.54 39	77.78 36
3)	Palatine Torus (+=ASU 2-3)	% n	0.00 35	0.00 63	0.00 53	0.00 60	2.13 47	0.00 60	0.00 61
4)	Shoveling UI1 (+=ASU 2-6)	% n	0.00 17	12.12 33	11.11 18	20.69 29	5.00 20	8.11 37	2.70 37
5)	Double Shoveling UI1 (+=ASU 2-6)	% n	6.67 15	2.70 37	0.00 19	0.00 34	0.00 25	0.00 43	0.00 36
6)	Interruption Groove UI2 (+=ASU +)	% n	9.09 22	4.65 43	4.00 25	0.00 38	3.70 27	2.63 38	15.79 38
7)	Tuberc. Dentale UI2 (+=ASU 2-6)	% n	30.43 23	36.59 41	34.48 29	37.50 32	34.62 26	44.44 36	29.73 37
8)	Bushman Canine UC (+=ASU 1-3)	% n	8.00 25	24.44 45	35.00 40	37.78 45	45.16 31	31.58 38	23.40 47
9)	Distal Acc. Ridge UC (+=ASU 2-5)	% n	38.10 21	39.02 41	44.12 34	60.61 33	37.04 27	31.43 35	26.32 38

1										
2										
3										
4										
5	10)	Hypocone UM2	%	77.42	93.65	74.47	77.55	90.48	93.44	91.07
6		(+=ASU 3-5)	n	31	63	47	49	42	61	56
7										
8										
9	11)	Cusp 5 UM1	%	22.22	22.03	17.39	17.02	14.63	21.15	18.37
10		(+=ASU 2-5)	n	27	59	46	47	41	52	49
11										
12										
13	12)	Carabelli's Trait UM1	%	31.03	49.15	51.06	47.06	34.15	52.63	58.00
14		(+=ASU 2-7)	n	29	59	47	51	41	57	50
15										
16										
17	13)	Parastyle UM3	%	0.00	0.00	0.00	2.33	2.94	1.92	2.38
18		(+=ASU 1-5)	n	25	52	41	43	34	52	42
19										
20										
21	14)	Enamel Extension UM1	%	3.45	4.76	2.08	3.92	5.00	1.85	1.72
22		(+=ASU 1-3)	n	29	63	48	51	40	54	58
23										
24										
25	15)	Root Number UP1	%	64.29	74.36	64.52	57.14	76.19	60.61	68.97
26		(+=ASU 2+)	n	14	39	31	35	21	33	29
27										
28										
29	16)	Root Number UM2	%	92.86	83.33	77.27	65.00	83.33	60.87	84.62
30		(+=ASU 3+)	n	14	24	22	20	12	23	26
31										
32										
33	17)	Peg-Reduced UI2	%	0.00	3.28	2.13	1.85	7.69	3.17	3.17
34		(+=ASU P or R)	n	30	61	47	54	39	63	63
35										
36										
37	18)	Odontome P1-P2	%	0.00	1.56	0.00	1.64	0.00	0.00	0.00
38		(+=ASU +)	n	36	64	57	61	44	64	63
39										
40										
41	19)	Congenital Abs. UM3	%	5.56	4.55	3.77	7.14	13.04	6.35	6.78
42		(+=ASU -)	n	36	66	53	56	46	63	59
43										
44										
45										
46										
47										
48										
49										

1										
2										
3										
4	20)	Midline Diastema UI1	%	20.69	8.47	11.11	2.13	5.88	4.92	10.17
5		(+ 0.5 mm)	n	29	59	36	47	34	61	59
6										
7										
8	21)	Lingual Cusp LP2	%	65.52	66.67	68.75	68.63	54.29	75.00	54.72
9		(+=ASU 2-9)	n	29	57	48	51	35	48	53
10										
11	22)	Anterior Fovea LM1	%	68.97	68.75	70.27	74.42	66.67	73.33	54.00
12		(+=ASU 2-4)	n	29	48	37	43	33	45	50
13										
14										
15	23)	Mandibular Torus	%	0.00	1.54	1.79	0.00	2.13	0.00	0.00
16		(+=ASU 2-3)	n	34	65	56	61	47	63	63
17										
18										
19	24)	Groove Pattern LM2	%	75.86	64.81	65.85	75.56	80.56	64.29	72.00
20		(+=ASU Y)	n	29	54	41	45	36	56	50
21										
22										
23	25)	Rocker Jaw	%	2.94	1.52	0.00	5.17	2.13	1.59	3.23
24		(+=ASU 1-2)	n	34	66	57	58	47	63	62
25										
26										
27	26)	Cusp Number LM1	%	6.45	13.79	2.27	6.25	12.50	0.00	1.85
28		(+=ASU 6+)	n	31	58	44	48	40	56	54
29										
30										
31	27)	Cusp Number LM2	%	88.46	83.64	77.50	76.60	71.43	87.50	90.00
32		(+=ASU 5+)	n	26	55	40	47	35	56	50
33										
34										
35	28)	Deflecting Wrinkle LM1	%	10.34	30.00	42.50	26.67	26.47	31.25	31.37
36		(+=ASU 2-3)	n	29	50	40	45	34	48	51
37										
38	29)	C1-C2 Crest LM1	%	6.90	1.89	0.00	2.22	5.56	2.08	1.96
39		(+=ASU +)	n	29	53	40	45	36	48	51
40										
41										
42										
43										
44										
45										
46										
47										
48										
49										

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

30)	Protostylid LM1	%	0.00	0.00	2.27	0.00	2.50	3.64	7.14
	(+=ASU 1-6)	n	31	59	44	49	40	55	56
31)	Cusp 7 LM1	%	31.25	37.93	33.33	37.50	18.42	42.84	49.09
	(+=ASU 2-4)	n	32	58	45	48	38	55	55
32)	Tomes' Root LP1	%	15.79	10.53	25.00	5.71	14.29	8.57	7.41
	(+=ASU 3-5)	n	19	38	36	35	21	35	27
33)	Root Number LC	%	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(+=ASU 2+)	n	21	32	31	29	23	40	33
34)	Root Number LM1	%	0.00	0.00	0.00	0.00	5.56	0.00	0.00
	(+=ASU 3+)	n	17	26	25	22	18	25	28
35)	Root Number LM2	%	100.00	92.86	100.00	89.29	94.12	100.00	72.73
	(+=ASU 2+)	n	16	28	21	28	17	26	22
36)	Torsomolar Angle LM3	%	0.00	8.00	6.82	8.11	5.71	11.54	8.33
	(+=ASU +)	n	23	50	44	37	35	52	48

¹NDB=Ndebele, SOT=Sotho, SWZ=Swazi, TSW=Tswana, VEN=Venda, XOS=Xhosa, ZUL=Zulu (see text for sample details).

²ASU rank-scale trait breakpoints from Irish (1993, 1997, 1998a,b, 2005, 2006) and Scott and Turner (1997).

TABLE 2. MMD distance matrix for the seven South African "Bantu" samples based on 36 dental traits

	NDB	SOT	SWZ	TSW	VEN	XOS	ZUL
NDB	0						
SOT	0.000	0					
SWZ	0.003	0.000	0				
TSW	<u>0.035</u>	0.000	0.000	0			
VEN	0.011	0.000	0.000	0.000	0		
XOS	0.027	0.000	0.000	0.000	0.014	0	
ZUL	0.019	0.004	0.023	<u>0.036</u>	0.024	0.018	0

NDB=Ndebele, SOT=Sotho, SWZ=Swazi, TSW=Tswana, VEN=Venda, XOS=Xhosa, ZUL=Zulu (see text for sample details).

Underlined MMD distances indicate significant difference at the 0.025 level.

TABLE 3. Component loadings, eigenvalues, and variance explained for 28 traits in seven South African "Bantu" samples

Trait	Comp 1	Comp 2	Comp 3
Winging UI1*	-.816¹	.481	.181
Labial Curvature UI1*	.741	.068	.194
Shoveling UI1	-.681	-.492	-.067
Interruption Groove UI2*	.894	.241	.363
Tuberc. Dentale UI2	-.613	-.605	.047
Bushman Canine UC*	-.864	.060	.395
Distal Acc. Ridge UC*	-.573	-.144	-.487
Hypocone UM2	-.016	.041	.597
Cusp 5 UM1*	.535	-.416	-.369
Carabelli's Trait UM1*	.115	-.720	.611
Enamel Extension UM1*	-.493	.601	-.232
Root Number UP1*	.007	.685	.344
Root Number UM2*	.576	.715	-.177
Peg-Reduced UI2*	-.530	.515	.602
Odontome P1-P2	-.390	-.246	-.112
Congenital Abs. UM3*	-.425	.701	.371
Midline Diastema UI1*	.763	.287	-.531
Lingual Cusp LP2*	-.212	-.807	-.499
Anterior Fovea LM1*	-.637	-.359	-.647
Groove Pattern LM2*	-.093	.794	-.048
Rocker Jaw	.031	.087	.056
Cusp Number LM1*	-.356	.628	-.131
Cusp Number LM2*	.846	-.389	.065
Deflecting Wrinkle LM1	-.329	-.442	.464
Cusp 7 LM1*	.500	-.734	.391
Tomes' Root LP1	.007	.200	-.471
Root Number LM2*	-.362	-.094	-.746
Torsomolar Angle LM3*	-.377	-.636	.661
Eigenvalue	7.931	7.009	4.875
Variance (%)	28.327	25.033	17.410
Total Variance	28.327	53.360	70.770

¹Values in bold-face indicate strong loadings (i.e., > |.5|) as detailed in text.

TABLE 4. MMD distance matrix for the seven South African "Bantu" samples based on 21 dental traits after editing

	NDB	SOT	SWZ	TSW	VEN	XOS	ZUL
NDB	0						
SOT	0.000	0					
SWZ	0.000	0.000	0				
TSW	<u>0.052</u>	0.001	0.000	0			
VEN	0.031	0.000	0.000	0.005	0		
XOS	0.038	0.008	0.000	0.006	<u>0.044</u>	0	
ZUL	0.033	0.013	<u>0.037</u>	<u>0.060</u>	<u>0.062</u>	<u>0.054</u>	0

NDB=Ndebele, SOT=Sotho, SWZ=Swazi, TSW=Tswana, VEN=Venda, XOS=Xhosa, ZUL=Zulu (see text for sample details).

Underlined MMD distances indicate significant difference at the 0.025 level.

TABLE 5. MMD distances between the San and seven South African "Bantu" samples based on 36 dental traits

	SAN
NDB	<u>0.089</u>
SOT	<u>0.048</u>
SWZ	<u>0.054</u>
TSW	<u>0.046</u>
VEN	<u>0.050</u>
XOS	<u>0.042</u>
ZUL	<u>0.050</u>

SAN=San, NDB=Ndebele, SOT=Sotho, SWZ=Swazi, TSW=Tswana, VEN=Venda, XOS=Xhosa, ZUL=Zulu (see text for sample details).

Underlined MD distances indicate significant difference at the 0.025 level.

TABLE 6. *R*-matrices from MDS optimally-scaled matrix of MMD distances using overall trait heritability of $h^2 = 0.65$ for calculation of estimated F_{st} (bottom diagonal) and $h^2 = 1.0$ for minimum F_{st} (top diagonal) for the seven South African "Bantu" samples based on 21 dental traits

	NDB	SOT	SWZ	TSW	VEN	XOS	ZUL	
	0.028	0.000	0.001	-0.015	-0.006	-0.008	0.000	NDB
		0.013	-0.007	-0.002	-0.001	-0.004	0.000	SOT
NDB	0.043		0.015	-0.001	0.000	0.000	-0.008	SWZ
SOT	0.000	0.020		0.025	0.003	0.003	-0.012	TSW
SWZ	0.001	-0.010	0.023		0.027	-0.011	-0.012	VEN
TSW	-0.023	-0.003	-0.002	0.038		0.028	-0.008	XOS
VEN	-0.009	-0.001	0.000	0.004	0.041		0.040	ZUL
XOS	-0.012	-0.006	0.001	0.005	-0.017	0.042		
ZUL	0.000	0.001	-0.012	-0.019	-0.018	-0.013	0.061	
	NDB	SOT	SWZ	TSW	VEN	XOS	ZUL	

Minimum $F_{st} = 0.0252$

Estimated $F_{st} = 0.0382$

TABLE 7. Geographic straight-line distances in km among the seven South African "Bantu" populations from the approximated centers of their historic homelands

	NDB	SOT	SWZ	TSW	VEN	XOS	ZUL
NDB	0						
SOT	473.83	0					
SWZ	294.14	453.94	0				
TSW	599.72	650.86	849.96	0			
VEN	314.36	766.41	409.42	848.45	0		
XOS	875.28	403.53	830.24	874.47	1169.58	0	
ZUL	474.00	430.36	201.14	965.21	605.61	739.22	0

NDB=Ndebele, SOT=Sotho, SWZ=Swazi, TSW=Tswana, VEN=Venda, XOS=Xhosa, ZUL=Zulu (see text for sample details).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Figure 1. Historic homeland locations of the seven South African "Bantu" populations.
138x125mm (300 x 300 DPI)

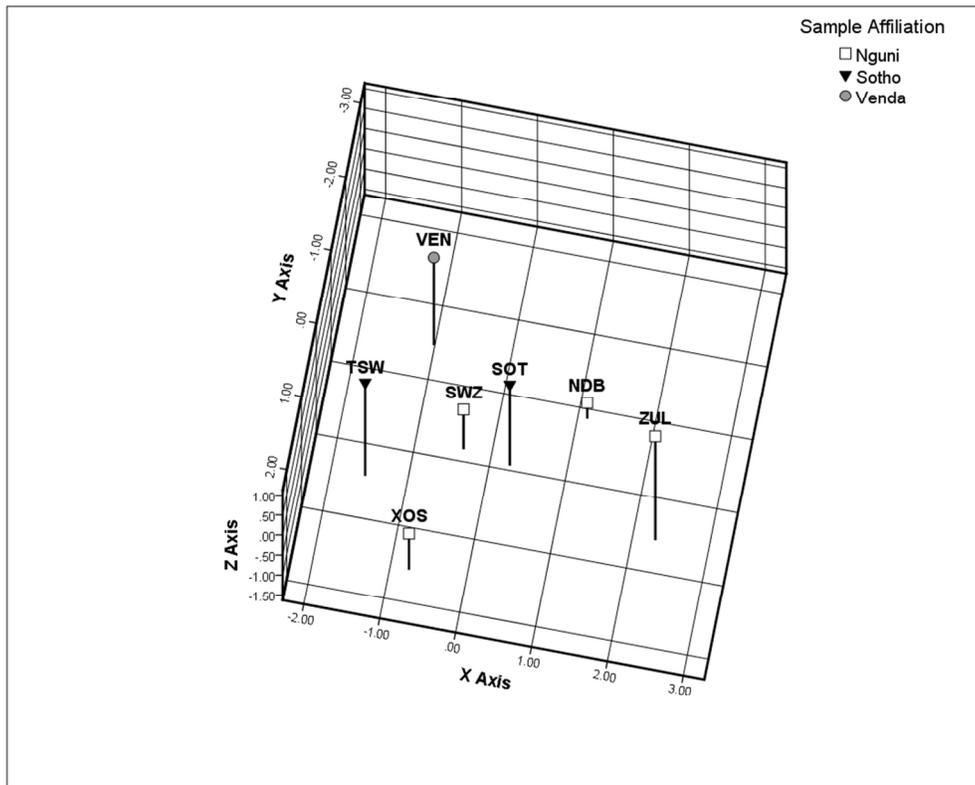


Figure 2. Three-dimensional MDS of 36-trait MMD distances among the seven "Bantu" samples. The three-letter sample abbreviations are defined in Table 1 and the text.
121x97mm (300 x 300 DPI)

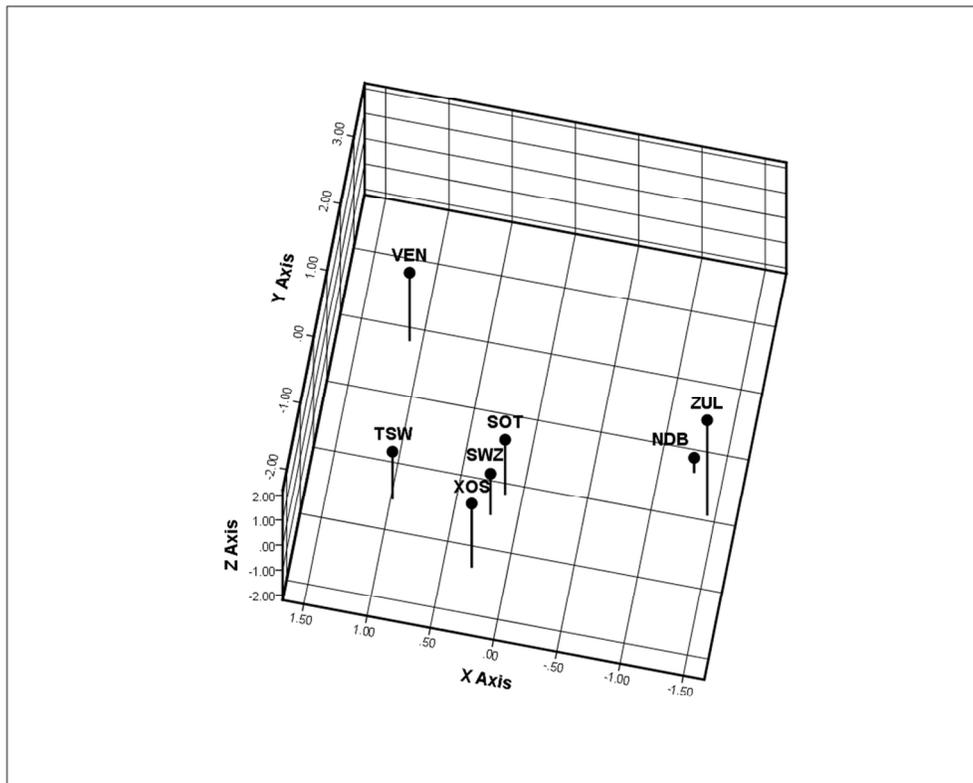


Figure 3. Three-dimensional scatterplot of the first three components among the seven "Bantu" samples for 28 dental traits from Table 3. Accounts for 70.77% of the total variance (28.33% on x-axis, 25.03% on y-axis, and 17.77% on z-axis). Sample abbreviations defined in Table 1 and the text.
121x97mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

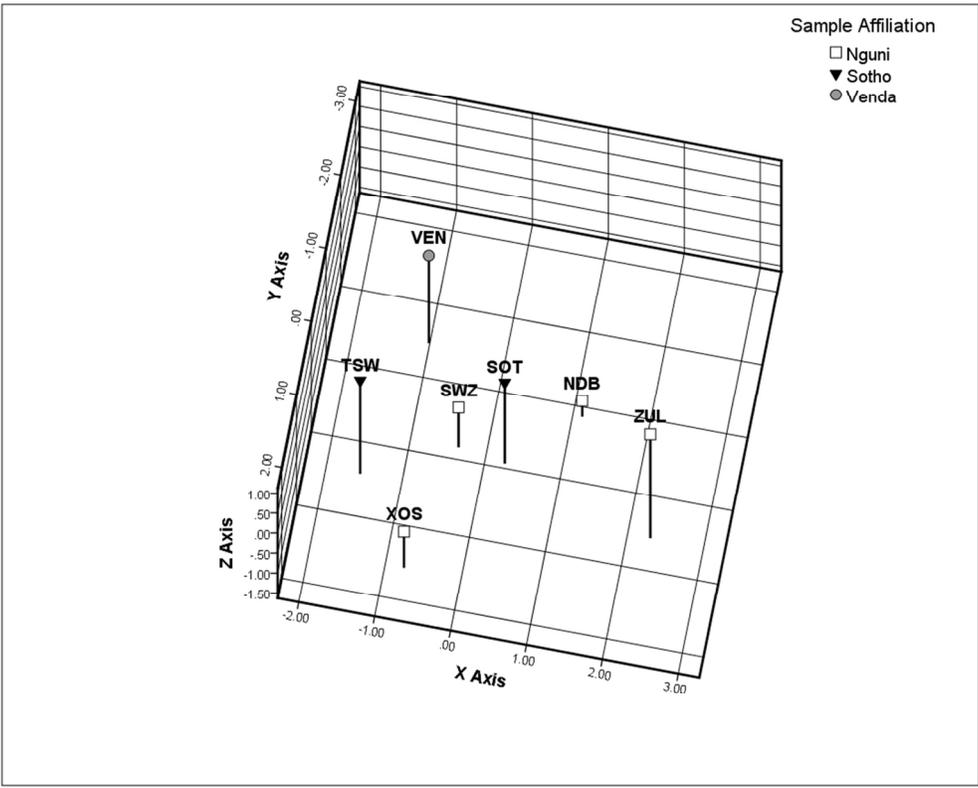


Figure 4. Three-dimensional MDS of 21-trait MMD distances among the seven "Bantu" samples. Sample abbreviations defined in Table 1 and the text.
121x97mm (300 x 300 DPI)

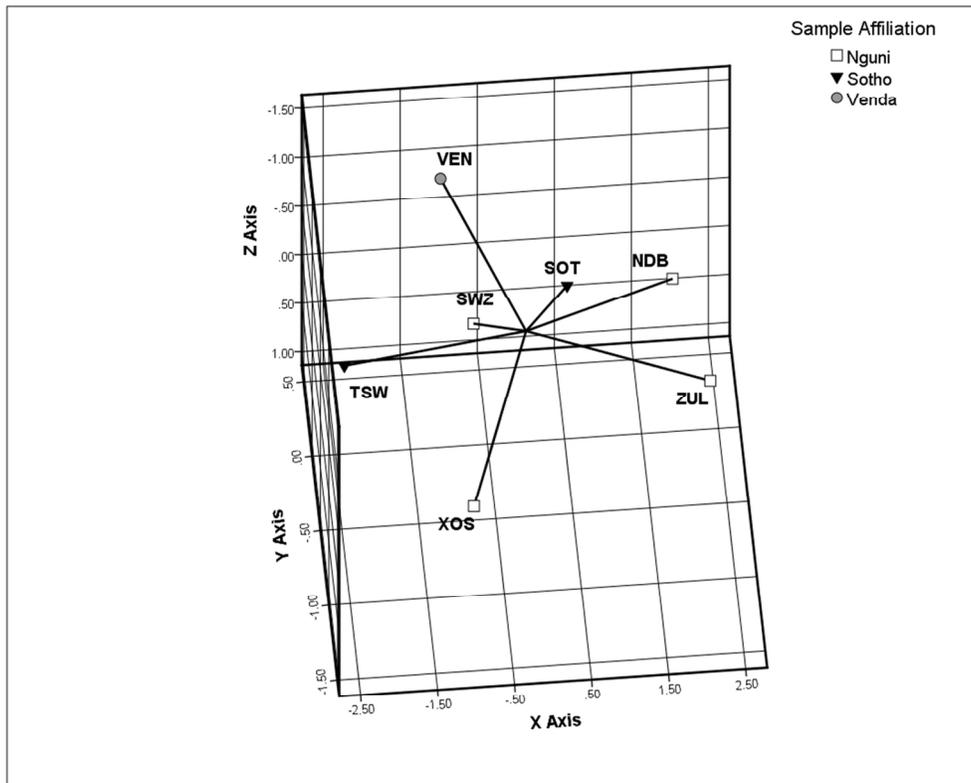


Figure 5. Three-dimensional MDS approximation of the MMD-based R-matrix for the seven "Bantu" samples. Sample abbreviations defined in Table 1 and the text. 121x97mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

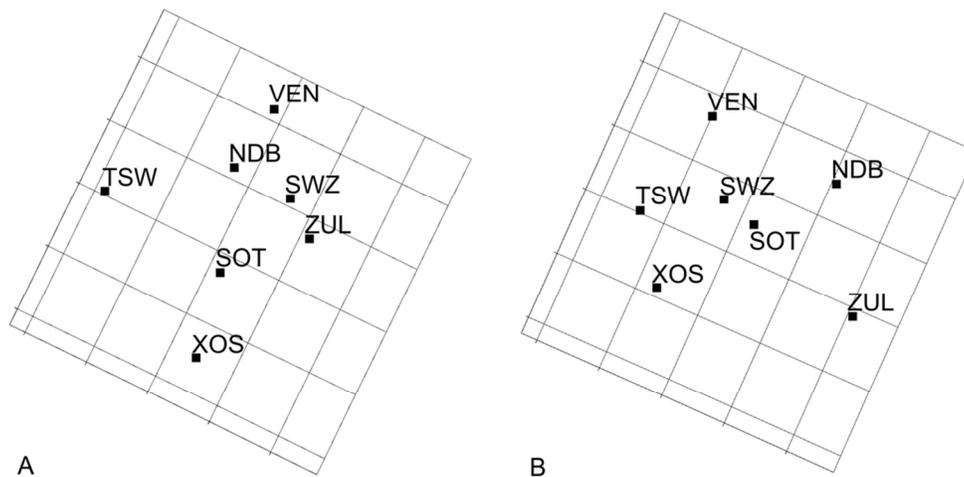


Figure 6. Two-dimensional MDS graphs of inter-sample (a) geographic distances in Km and (b) 21-trait MMD distances among the seven "Bantu" samples. Sample abbreviations defined in Table 1 and the text. 101x50mm (300 x 300 DPI)