Population differences in the postcrania of modern South Africans and the implications for ancestry estimation

Leandi Liebenberg[1,2,*], Ericka N. L'Abbé[1] and Kyra E. Stull[3]

[1]Department of Anatomy, University of Pretoria, Arcadia, South Africa

[2]Department of Basic Medical Sciences, University of the Free State, Bloemfontein, South Africa

[3]Department of Anthropology, Idaho State University, Pocatello, ID, USA

*Corresponding author at: Department of Basic Medical Sciences, University of the Free State, IB G25: Basic Medical Sciences, P.O. Box 339, Bloemfontein, 9300, South Africa. Tel.: +27 82 072 0057/51 405 7249.

Highlights

- •Postcranial variation among modern South Africans is assessed.
- •Flexible and linear discriminant analysis (FDA and LDA) is compared.
- •Multivariate subsets achieved the greatest classification accuracies.
- •FDA yielded overall greater results than LDA.
- •A multivariate postcraniometric approach can accurately assign ancestry.

Abstract

The cranium is widely recognized as the most important skeletal element to use when evaluating population differences and estimating ancestry. However, the cranium is not always intact or available for analysis, which emphasizes the need for postcranial alternatives. The purpose of this study was to quantify postcraniometric differences among South Africans that can be used to estimate ancestry. Thirty-nine standard measurements from 11 postcranial bones were collected from 360 modern black, white and coloured South Africans; the sex and ancestry distribution were equal. Group differences were explored with analysis of variance (ANOVA) and Tukey's honestly significant difference (HSD) test. Linear and flexible discriminant analysis (LDA and FDA, respectively) were conducted with bone

models as well as numerous multivariate subsets to identify the model and method that yielded the highest correct classifications. Leave-one-out (LDA) and k-fold (k = 10; FDA) cross-validation with equal priors were used for all models. ANOVA and Tukey's HSD results reveal statistically significant differences between at least two of the three groups for the majority of the variables, with varying degrees of group overlap. Bone models, which consisted of all measurements per bone, resulted in low accuracies that ranged from 46% to 63% (LDA) and 41% to 66% (FDA). In contrast, the multivariate subsets, which consisted of different variable combinations from all elements, achieved accuracies as high as 85% (LDA) and 87% (FDA). Thus, when using a multivariate approach, the postcranial skeleton can distinguish among three modern South African groups with high accuracy.

Introduction

The estimation of ancestry is an essential aspect of the biological profile used to assist with the identification of unknown individuals from skeletal remains. Accurate estimates can be achieved through the quantification and interpretation of morphological variation among different groups. Differences in the general morphology of skeletal elements develop from the complex interaction between proximate and ultimate influences [1]. During the ontogeny of the skeleton, genetic information is translated into a specific phenotype through the processes of growth, development and allometry. Intra-specific skeletal differences among modern humans can arise through three inter-related aspects of ontogeny, namely (1) early onset of population-specific morphologies, (2) divergence in ontogenetic shape scaling trajectories (shape differences) and (3) ontogenetic scaling where populations have different adult end points on the same ontogenetic shape scaling trajectories (size differences) [2]. Essentially human variation is multifactorial, evolutionarily dynamic and population

2

historically contingent [3]. Specifically, each population has different origins with bouts of selection specific to their ancestral environment, ultimately producing the distinct morphology that we see among groups today. South Africa provides a testament to how geographical and social factors can influence population variation. The wide range of morphological variation among groups in South Africa is the cumulative result of varying geographic origins, unique population histories and structure, positive assortative mating and social and/or political barriers [4-7].

The South African population is diverse and consists of three major peer-reported groups, namely South African blacks (80.2%), whites (8.4%) and coloureds (8.8%). The remaining 2.5% of the population consists mainly of individuals classified as Asian and Indian [8]. Each group has a unique population history within the country. Black South Africans originally descended from Bantu-speaking groups from western and central Africa, who migrated throughout sub-Saharan Africa [9,10]. The linguistic term "Bantu" refers to the general assemblage of several closely-related African languages belonging to the Niger-Kordofanian phylum [10]. Further divisions among the southern Bantu-speaking groups have led to numerous ethnic groups currently inhabiting South Africa, including the Sotho, Nguni, Venda and Shangaan-Tsonga [11]. During the 17th century the establishment of the Cape colony along the trade route of the Dutch East India Company saw the influx of European settlers into South Africa. The colonists predominantly consisted of Dutch settlers, with later additions of British, French and German immigrants [12,13]. In South Africa the term "coloured" refers to a self-identified social group that exhibits highly diverse origins [14]. Because South African coloureds constitute such a complex combination of continental populations, the group typically presents with some of the highest degrees of genetic variation in the world [15,16]. Parental populations are inclusive of (1) indigenous Khoesan, (2) descendants of slaves brought from countries such as Malaysia, Indonesia and

India, (3) Europeans and (4) Bantu-speaking groups [14,17-19]. The different group histories shaped the unique and heterogeneous basis of modern variation among South Africans [6].

Numerous studies identify the cranium as the most accurate element to estimate ancestry; specifically, inclusion of midfacial variables in metric analyses can increase the accuracy of classification models [6,20–22]. Using standard linear cranio-metric variables the South African literature reports accuracies ranging from 98% between black and white South Africans [20] to 84% accuracy with a three group classification among black, white and coloured groups [23]. However, a holistic approach with a variety of techniques from numerous elements should be used rather than relying solely on one element. In South Africa, when the cranium is not recovered or is not present for the skeletal analysis, the alternative options are limited and the anthropologist is unable to accurately and reliably estimate ancestry. Even though the postcranial skeleton has not been studied as thoroughly as the cranium with regard to ancestry, numerous studies that assessed sex and stature affirm group differences [41–43]. Additionally, the a priori knowledge of ancestry can increase the accuracy of sex and stature estimates [44,45]. The presence of group differences when investigating other biological parameters inherently indicates the potential of the postcranial skeleton specifically for the estimation of ancestry. In contrast to crania, postcrania are frequently labelled as exceedingly variable and unreliable for assessing population affinity, owed primarily to the extreme plasticity exhibited by the appendicular skeleton [24–26]. For example, long bone lengths and epiphyseal breadths have been correlated to climate, specifically minimum average temperatures [27]. Similarly, stature is frequently correlated with socio-economic status (SES), where lower SES typically results in shorter stature [28,29]. Yet the genetic contribution to postcranial variation is often overlooked. Discernable population differences have been noted early in life among the postcrania of subadults, which is suggestive of a substantial genetic component to postcranial ontogeny [30–33]. Furthermore, stature, which is partially a function of long

bone length, has a rather large genetic component, with heritability estimates among populations ranging from 60% to 90% [32]. Although extrinsic factors may affect postcranial variance, it is likely not powerful enough to completely obliterate the inherent genetic variation present among populations. Previous attempts have been made to quantify postcranial variation with the specific intent to estimate ancestry. However, there are limitations associated with attempts to explore ancestry variation via the postcranium. Published techniques that use postcranial elements to estimate ancestry [e.g. 34–38], are not population-specific, not representative of the entire South African population and/or do not satisfy Daubert criteria [39]. While South African courts do not specifically use the Daubert guidelines, similar standards are expected to be on par with international expectations [40]. The availability of computers and personal software has made robust multivariate techniques a fairly popular choice regarding ancestry estimation. The inclusion of more measurements in an analysis has generally shown to better highlight group differences, thereby increasing the positive predictive accuracy of classification models [46]. To date a South African craniometric database exists for use with the personal computer software program, FORDISC (FD3), and is frequently used with medico-legal case analyses. However, no comparable postcranial database is available for South Africans. Postcranial data is currently available for black and white North Americans in the Forensic Databank (FDB) and analyses have demonstrated that accuracies as high as 94% can be achieved between black and white groups using a model inclusive of all postcranial measurements [47]. Hispanic postcranial data has also recently been incorporated into the FDB. Three-group classification models among black, white and Hispanic groups in the United States result with moderately high accuracy (76%) [48]. The classification rates achieved using the postcrania of US populations [47,48] demonstrate that postcrania can be used to confidently assign ancestry in heterogeneous populations and can be useful in a medico-legal setting. The purpose of the current study was to use robust statistical techniques to explore and quantify postcranial differences among modern black, white and coloured South Africans.

5

Materials and methods

Postcraniometric data were collected from 360 peer-reported black, white and coloured South Africans of known sex, age at death, and ancestry. The skeletal remains were obtained from two contemporary cadaver-based collections, namely the Pretoria Bone Collection (University of Pretoria) and the Kirsten Collection (Stellenbosch University). While each group consisted of an equal distribution of males and females, the sexes were pooled for all statistical analyses as the main focus of the study was to assess group differences rather than sex differences. Age at death for the sample ranged from 18 to 87 years, with mean ages of 47, 65 and 48 years for the black, white and coloured groups, respectively. Any skeletal material with visible pathology, ante-mortem trauma or post-mortem damage that prevented accurate measurement was excluded. A total of 39 standard linear measurements were taken to the nearest millimeter on 11 skeletal elements (clavicle, scapula, humerus, radius, ulna, sacrum, innominate, femur, tibia, fibula and calcaneus) [49] (Table 1).

Statistical analyses

Ten individuals were randomly selected from the sample to assess inter- and intra-observer agreement. The repeatability of the measurements was gauged using two methods. First the technical error of measurement (TEM) was calculated. Calculation of TEM provides an accuracy index that expresses error margins through the standard deviation between repeated measurements [50,51]. Additionally, absolute TEM was converted to relative TEM (%TEM), which provides an estimate of magnitude of the measurement error relative to the size of the measurement [51,52]. Second, Bland-

Table 1. The thirty-nine standard postcraniometric measurements and their associated abbreviations. See [49] for measurement definitions.

Measurement	Abbreviation	Measurement	Abbreviation
Clavicle maximum length	**claxln**	Innominate height	**innoht**
Clavicle sagittal midshaft diameter	**claapd**	Iliac breadth	**iliabr**
Clavicle vertical midshaft diameter	**clavrd**	Femur maximum length	**femxln**
Scapula height	**scapht**	Femur bicondylar length	**fembln**
Scapula breadth	**scapbr**	Femur epicondylar breadth	**femebr**
Humerus maximum length	**humxln**	Femoral head diameter	**femhdd**
Humerus epicondylar breadth	**humebr**	Femur A-P subtrochanteric diameter	**femsap**
Humeral head diameter	**humhdd**	Femur transverse subtrochanteric diameter	**femstv**
Humerus maximum diameter	**hummxd**	Femur A-P midshaft diameter	**femmap**
Humerus minimum diameter	**hummwd**	Femur transverse midshaft diameter	**femmtv**
Radius maximum length	**radxln**	Tibia condylo-malleolar length	**tibxln**
Radius A-P midshaft diameter	**radapd**	Tibia proximal epiphyseal breadth	**tibpeb**
Radius transverse midshaft diameter	**radtvd**	Tibia distal epiphyseal breadth	**tibdeb**
Ulna maximum length	**ulnxln**	Tibia maximum diameter at nutrient foramen	**tibnfx**
Ulna dorso-volar diameter	**ulndvd**	Tibia minimum diameter at nutrient foramen	**tibnft**
Ulna transverse diameter	**ulntvd**	Fibula maximum length	**fibxln**
Ulna physiological length	**ulnphl**	Fibula maximum diameter	**fibmdm**
Sacrum anterior height	**sacaht**	Calcaneus maximum length	**calcxl**
Sacrum anterior breadth	**sacabr**	Calcaneus middle breadth	**calcbr**
Transverse diameter of S1	**sacs1b**		

Altman plots were used to visualize measurement agreement. A Bland-Altman plot essentially provides a graphical representation of TEM [53,54].

An analysis of variance (ANOVA) was conducted in conjunction with Tukey's honestly significant difference (HSD) test to identify mean differences among the groups for each measurement and to assess if the differences were statistically significant. Following this, linear discriminant analysis (LDA) and flexible discriminant analysis (FDA) were applied to numerous multivariate models. While LDA is the traditionally used classification method, FDA was assessed as an alternative classification technique. Both LDA and FDA operate to maximize overall differences among groups, thereby identifying boundaries between groups while taking the effect of individual variation within each group into account. Some key differences exist between LDA and FDA. LDA makes use of linear boundaries for group separation and requires a sufficient sample size, multivariate normal data and equal variance-covariance matrices for optimal results [46,55]. FDA is considered a non-parametric technique as it is not restricted by the assumptions associated with LDA [56,57]. Furthermore, FDA differs from LDA in that curved boundaries rather than linear boundaries are used to separate groups [58]. Specifically, when executed with the mda package in R [59], FDA uses multivariate adaptive regression splines (MARS) models in contrast to the linear regression used by LDA [56]. Subsequently, FDA demonstrates enhanced flexibility and is thought to be superior to LDA when more than two response categories are present [60].

Variable selection techniques were employed to remove any superfluous variables that do not contribute to classification, as unnecessary variables may negatively affect the accuracy of classification models [55,61]. Backward stepwise selection was conducted with LDA and a forward and backward stepwise selection with automatic variable selection is integral to FDA [62]. Cross-validation provides the prediction estimates of a model while preventing optimistic bias and overfitting

8

Table 2. Variables and sample size for each multivariate subset. See Table 1 for measurement abbreviations.

	n	Variables
Upper limb	316	humxln, humebr, humhdd, hummxd, humwd, radxln, radapd, radtvd, ulnxln, ulndvd, ulntvd, ulnphl
Upper limb + Shoulder girdle	183	claxln, clavrd, claapd, scapht, scapbr, humxln, humebr, humhdd, hummxd, humwd, radxln, radapd, radtvd, ulnxln, ulndvd, ulntvd, ulnphl
Lower limb	225	femxln, fembln,femebr, femhdd, femsap, femstv, femmap, femmtv, tibxln, tibpeb, tibdeb, tibnfx, tibnft, fibxln, fibmdm
Lower limb + Pelvic girdle	177	femxln, fembln, femebr, femhdd, femsap, femstv, femmap, femmtv, tibxln, tibpeb, tibdeb, tibnfx, tibnft, fibxln, fibmdm, sacabr, innoht, iliabr
Pelvis + Femur	265	femxln, fembln, femebr, femhdd, femsap, femstv, femmap, femmtv, innoht, iliabr
Proximal elements	280	humxln, humebr, humhdd, hummxd, humwd, femxln, fembln,femebr, femhdd, femsap, femstv, femmap, femmtv
Distal elements	247	radxln, radapd, radtvd, ulnxln, ulndvd, ulntvd, ulnphl, tibxln, tibpeb, tibdeb, tibnfx, tibnft, fibxln, fibmdm
All-measurement	140	claapd, scapht, humxln, humhdd, ulnxln, ulnphl, sacabr, iliabr, femstv, tibpeb, calcbr

of the data, thus presenting more realistic results [46]. Leave-one-out cross-validation (LOOCV) was used with the LDA while *k*-fold cross validation (K = 10) was used with the FDA. LOOCV involves

the consecutive removal of one individual from the sample, after which a discriminant function is created using all remaining individuals to attempt to classify the removed individual [46,56]. *K*-fold cross-validation is different in that the sample is randomly divided into *k* folds, or groups, of equal size. In turn, each fold is treated as a validation set (i.e. a hold-out group) while the model is fit to the remaining folds similar to LOOCV [63]. All statistical analyses were performed with R version 3.0.1 [64].

To maximize the practical applicability of the data several different models were created, namely bone models and multivariate subsets. Bone models consisted of all measurements collected per bone and the multivariate subsets were made up of numerous measurement combinations from several different bones (Table 2). Every possible measurement combination could not be explored, as the large number of variables included in the study would result in too many model permutations to assess.

Results

Technical error of measurement (TEM) was used to demonstrate the degree of measurement repeatability between three observers. The intra-observer TEM and %TEM ranged between 0 and 0.71mm and 0% and 4.52%, respectively. The inter-observer error was slightly higher, with TEM and %TEM ranging between 0 and 1.10mm and 0% and 5.36% respectively (Table 3). The measurements with the highest TEM for both intra- and inter-observer agreement were the midshaft measures of the clavicle, humerus and ulna as well as the subtrochanteric measurements of the femur. Midshaft measurements of the radius also presented higher levels of measurement error for the inter-observer analysis.

Table 3. Technical error of measurement (TEM) and relative technical error of measurement (%TEM) demonstrating inter- and intra-observer agreement.

	Intra-observer error		Inter-observer error 1		Inter-observer error 2	
	TEM	%TEM	TEM	%TEM	TEM	%TEM
claxln	0.00	0.00	0.35	0.24	0.35	0.24
claapd	0.35	3.58	0.35	3.67	0.00	0.00
clavrd	0.00	0.00	0.50	4.17	0.00	0.00
scapht	0.00	0.00	0.71	0.50	0.00	0.00
scapbr	0.35	0.37	0.35	0.37	0.61	0.64
humxln	0.00	0.00	0.55	0.18	0.32	0.10
humebr	0.00	0.00	0.32	0.56	0.00	0.00
humhdd	0.00	0.00	0.55	1.36	0.45	1.10
hummxd	0.00	0.00	0.00	0.00	0.00	0.00
hummwd	0.32	2.07	0.45	2.90	0.32	2.07
radxln	0.35	0.15	0.35	0.15	0.61	0.26
radapd	0.00	0.00	0.45	4.07	0.45	3.99
radtvd	0.00	0.00	0.32	2.34	0.45	3.29
ulnxln	0.35	0.14	0.00	0.00	0.35	0.14
ulndvd	0.63	4.52	0.71	5.36	0.55	3.94

ulntvd	0.00	0.00	0.55	3.78	0.63	4.65
ulnphl	0.00	0.00	0.61	0.27	0.61	0.27
sacaht	0.00	0.00	0.00	0.00	0.50	0.45
sacabr	0.71	0.74	0.61	0.65	0.94	0.99
sacs1b	0.00	0.00	1.00	2.33	0.41	0.94
iliabr	0.00	0.00	1.10	0.58	0.55	0.29
innoht	0.32	0.22	0.55	0.38	0.63	0.44
femxln	0.32	0.07	0.45	0.10	0.45	0.10
fembln	0.32	0.07	0.45	0.10	0.45	0.10
femebr	0.00	0.00	0.41	0.54	0.63	0.87
femhdd	0.00	0.00	0.35	0.84	0.32	0.76
femsap	0.00	0.00	0.95	3.87	0.55	2.18
femstv	0.45	1.55	0.77	2.71	0.71	2.41
femmap	0.00	0.00	0.32	1.14	0.32	1.15
femmtv	0.00	0.00	0.00	0.00	0.55	2.27
tibxln	0.00	0.00	0.41	0.11	0.58	0.15
tibpeb	0.00	0.00	0.82	1.14	0.50	0.70
tibdeb	0.00	0.00	1.06	2.34	0.35	0.79
tibnfx	0.00	0.00	0.55	1.70	0.32	0.97

|						
tibnft	0.32	1.38	0.77	3.40	0.45	1.96
fibxln	0.00	0.00	0.41	0.11	0.58	0.16
fibmdm	0.00	0.00	0.32	2.34	0.00	0.00
calcxl	0.00	0.00	0.00	0.00	0.71	0.90
calcbr	0.35	0.87	0.41	0.99	0.50	1.24
Mean	0.13	0.40	0.48	1.42	0.43	1.04
Min	0.00	0.00	0.00	0.00	0.00	0.00
Max	0.71	4.52	1.10	5.36	0.94	4.65

Bland-Altman plots were used to illustrate measurement agreement among the three observers. Sufficient agreement was noted among the three observers, with only a few measurements presenting with more than 2mm difference (Figure 1). The intra-observer plot presented with high precision, as the majority of the measurements did not have more than 1mm difference and had an overall narrower distance between the upper and lower agreement levels (Figure 1). For both the inter- and intra-observer agreement more inconsistencies were noted in smaller measurements (i.e. midshaft measurements) rather than the larger measurements (i.e. maximum lengths). Because measurement error was minimal, the inter- and intra-observer agreement was considered satisfactory, and all measurements were retained in the analyses.

The ANOVA detected statistically significant differences for the majority of the variables. However, when further explored with Tukey's HSD, numerous measurements had similar means for two of the three groups. Black South Africans were generally intermediate in size between coloured
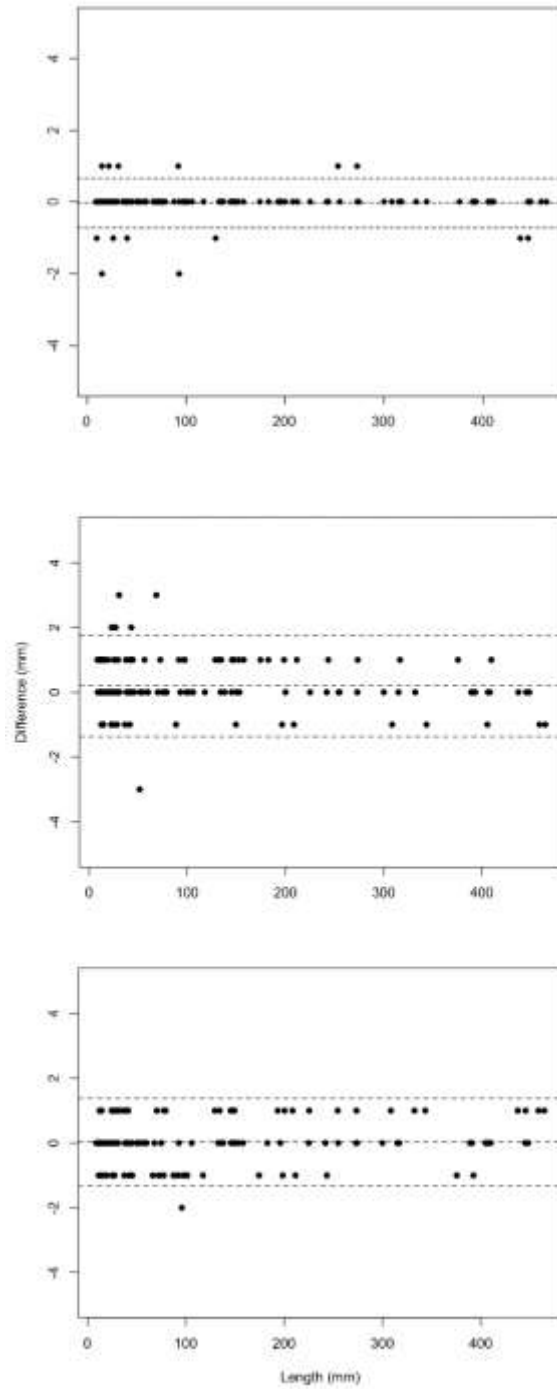
Fig. 1. Bland-Altman plots illustrating the inter- and intra-observer agreement of measurements compared among three observers. The *x*-axis represents the mean size (mm) of the measurements and the *y*-axis represents the difference in a pair of measurements compared between two observers.

and white South Africans and demonstrated substantial overlap with the other groups. White South

Africans typically had larger dimensions than the other groups, while coloured South Africans typically

had the smallest measurement means (Figure 2).
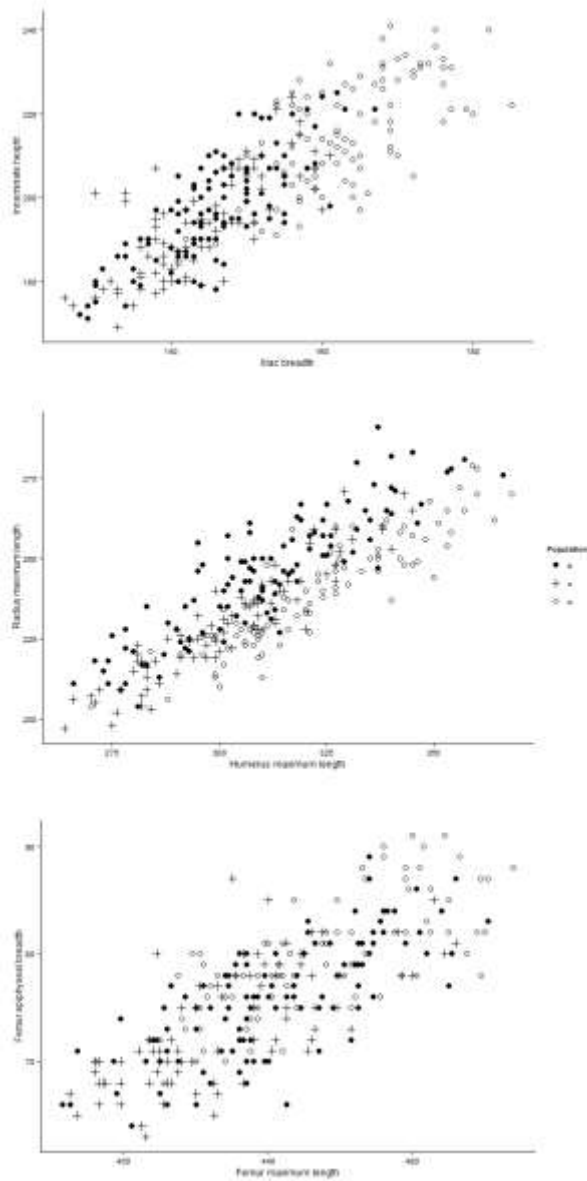
Fig. 2. Scatter plots demonstrating group variation for three variable combinations representing different patterns and

degrees of overlap among South African groups (black = b, coloured = c and white = w).

Table 4. Cross-validated classification accuracies (%) for each multivariate bone model using LDA, LDA with stepwise variable selection and FDA. Bold indicates the highest correct classification accuracy for each bone model.

	n	LDA	Stepwise LDA	FDA
Clavicle	253	46	**47**	46
Scapula	294	46	**47**	42
Humerus	343	57	58	**62**
Radius	344	51	**53**	**53**
Ulna	339	53	53	**54**
Sacrum	160	61	63	**66**
Pelvis	331	**62**	61	61
Femur	288	57	**60**	**60**

Using LDA, cross-validated classification accuracies for the bone models ranged from 46% to 63% (Table 4). Similar results were noted using FDA, with accuracies ranging from 41% to 66%. The sacrum and innominate models achieved the highest correct classifications while the fibula and scapula achieved the lowest. The multivariate subsets demonstrated greater accuracies than the bone models, with 63% to 85% using LDA and 62% to 87% using FDA (Table 5). The all-measurement model, showcasing a collection of variables from several different bones, achieved the highest classification
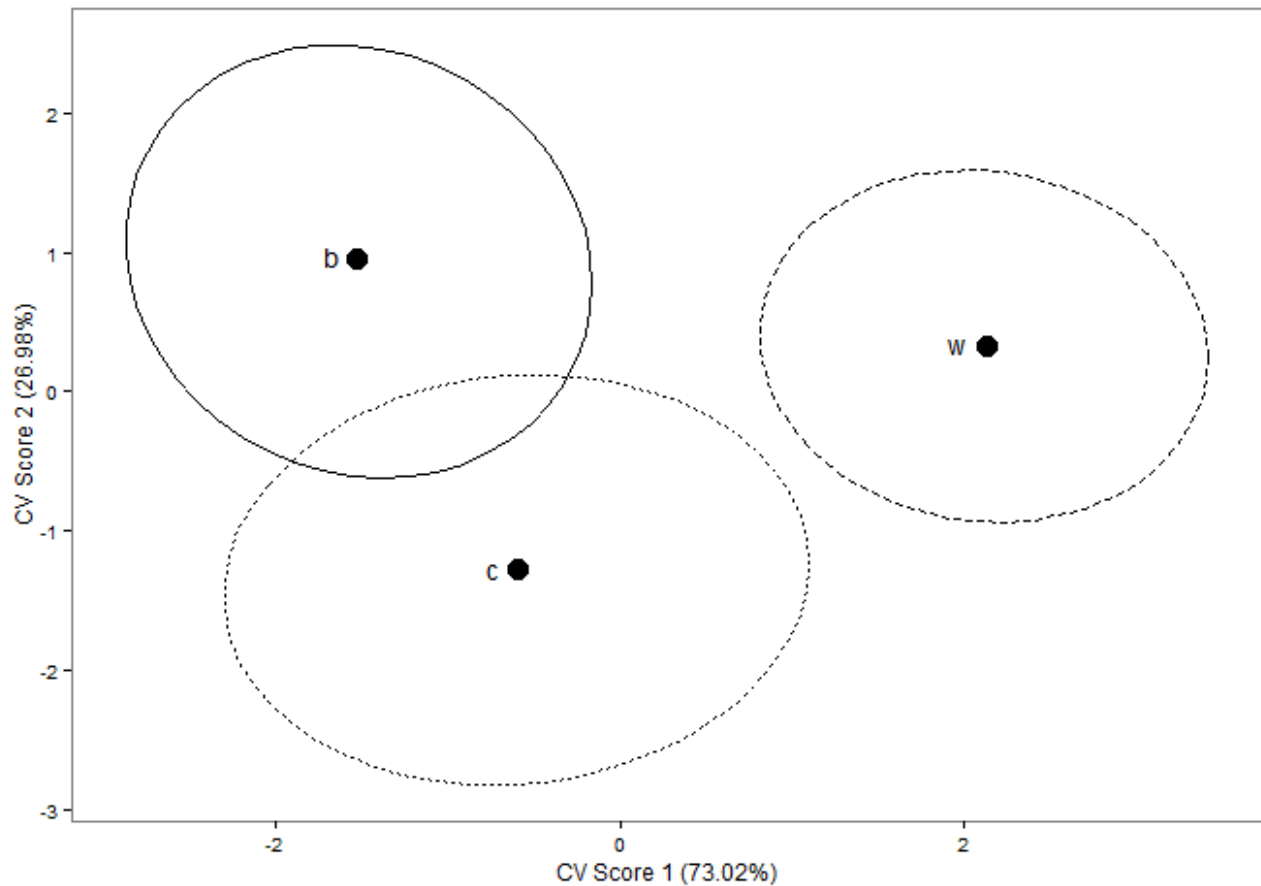
Table 5. Cross-validated classification accuracies (%) for each multivariate subset using LDA, LDA with stepwise variable selection, and FDA. Bold indicates the highest correct classification accuracy for each multivariate subset.

	n	LDA	Stepwise LDA	FDA
Upper limb	316	70	70	**72**
Upper limb + Shoulder girdle	183	68	72	**78**
Lower limb	225	65	68	**72**
Lower limb + Pelvic girdle	177	76	80	**82**
Pelvis + Femur	265	72	73	**75**
Proximal elements	280	65	68	**73**
Distal elements	247	56	**63**	62
All-measurement	140	84	85	**87**

accuracy (85% to 87%) (Figure 3). For model creation, both LDA and FDA generally retained maximum length variables and epiphyseal breadths with the respective variable reduction techniques. The pelvic variables (iliabr and innoht) were retained in all models utilizing the innominate. Two variables that were repeatedly weighted as key discriminators in the multivariate subsets were iliac breadth (iliabr) and humeral head diameter (humhdd).

In Figure 3, the larger ellipses associated with black and coloured South Africans indicates higher levels of within-group variation compared to the smaller ellipse associated with white South Africans. Further exploration of the classification matrices revealed that black and coloured South

Fig. 3. Plot of canonical variate South African group means (black = b, white = w, and coloured = c) and 95% confidence ellipses for the all-measurement model that presented with an 85% cross-validated correct classification accuracy.

Africans were more likely to misclassify as one another, while white South Africans had the tendency to misclassify as either black or coloured. However, the pattern of misclassification depended largely on the variables included in the model. For example, models including many breadth variables demonstrated substantial misclassification between black and coloured South Africans, whereas inclusion of the lengths of the distal limb elements demonstrated more frequent misclassification between black and white South Africans. The assessment of the bone models and multivariate subsets revealed that white South Africans achieved the highest overall positive predictive values.

Discussion and Conclusions

Based on the high correct classifications associated with the multivariate subsets, postcraniometric data can be used to confidently estimate ancestry in the contemporary South African population. The results offer a significant advancement in the field considering this is the first research that offers a thorough and statistically sound methodology to explore population differences in the postcrania of South Africans. Furthermore, the 85% accuracies obtained using the all-measurement model is comparable to the 84% accuracy obtained using the cranium in the same population [stull 2014]. Thus, the results indicate that postcrania can be used in future analyses to supplement and/or substantiate results based on craniometrics. However, the postcraniometric approach is also robust enough to provide an accurate assignment of ancestry when the cranium is not available.

The accuracies observed with the single bone models were too low to be considered practical. Simply, the specific combinations of variables included in each bone model are not necessarily the best variable combinations for group separation. Rather, the inclusion of highly discriminatory variable combinations, such as with multivariate subsets, is necessary to separate out the size and proportional differences among the groups. The all-measurement subset applied several variables that were identified as important during model creation. During model creation the long bone lengths, epiphyseal breadths and pelvic variables were considered more important for separating the groups with moderately high accuracy. For example, black and coloured South Africans have fairly narrow diaphyses and pelves. However, coloured South Africans typically have shorter limbs, which distinguish them from black South Africans. Both of these groups can be distinguished from white South Africans who have a combination of long limbs, robust epiphyseal and diaphyseal dimensions, and large pelves.

Variation observed among the postcrania of peer-reported South African groups may be attributed to different histories, paired with variable population experiences and selective influences. Decades of positive assortative mating and segregation limited gene flow among the groups and enabled morphological differences to be maintained among modern South Africans [6]. Higher levels of within-group variation was noted among black and coloured South Africans, ultimately resulting in substantial group overlap. Because of socio-political regimes and the colonization history of South Africa, the black and coloured groups would be expected to have more similarities to one another than either group would have to white South Africans. For instance, numerous examples of intermarriage between the historic Khoesan and Bantu-speaking groups are noted within the literature [19,65]. Cranial morphology corroborates a Khoesan influence among some black subgroups, particularly the Xhosa and Zulu [19,66]. Furthermore, the myriad of laws associated with apartheid (e.g. Prohibition of Mixed Marriages Act of 1949 and Immorality Act of 1950) were not quite as stringent in preventing intermixture between black and coloured groups as with white and "non-white" groups [67]. As would be expected from the population history, the white South Africans presented with more homogeneity, as there was far less within-group variation and more between-group variation. Essentially white South Africans were easier to distinguish from the other groups, thus, achieving higher classification accuracies. Yet, prior to the implementation of segregation laws some intermixture occurred between white and "non-white" groups. As the European colonists that came to South Africa were mainly males, unions between European men and indigenous and/or slave women were not uncommon [17,18].

Comparison of the different discriminant analyses revealed that FDA marginally outperformed LDA. When conducted on the subsets, FDA yielded accuracies between 2% and 6% higher than LDA. Although not a particularly large difference in accuracy, FDA has other benefits that justify its

inclusion in anthropological analyses. The major advantage of FDA is that fewer assumptions about the data are involved. Second, by applying less rigid boundaries FDA can be particularly useful with classification problems where considerable variable overlap exists, such as ancestry estimation. And third, *k-fold* cross-validation, which is integral to FDA, offers a better bias-variance trade off in comparison to LOOCV [56]. Thus, when the data does not meet the assumptions associated with LDA, which is often the case with biological data, FDA can be used without compromising the classification accuracy or violating model assumptions. Overall FDA should be considered as an alternative to LDA. The models presented in the paper serve as examples of possible accuracies that can be achieved when estimating ancestry with postcrania. The data is available on request and can be used in either FD3 or R. If the data is imported into FD3 as a custom database, LDA can be applied. Or the data can be imported into R, and using the packages specified above, FDA can be applied.

References

[1] O.M. Pearson, Activity, Climate, and Postcranial Robusticity: Implications for Modern Human Origins and scenarios of Adaptive Change, Curr. Anthropol. 41 (2000) 569-607.

[2] U.S. Strand Viðarsdóttir, P. O'Higgins, Developmental variation in the facial skeleton of anatomically modern Homo sapiens, in: J.L. Thompson, G.E. Krovitz, A.J. Nelson (Eds.), Patterns of growth and development in the genus homo, 2003, Cambridge: Cambridge University Press. pp. 114-143.

[3] C.C. Roseman, B.M. Auerbach, Ecogeography, genetics, and the evolution of human body form, J. Hum. Evol. 78 (2015) 80-90.

[4] C.C. Roseman, T.D. Weaver, Multivariate apportionment of global human craniometric diversity, Am. J. Phys. Anthropol. 125 (2004) 257-263.

[5] J.H. Relethford, Race and global patterns of phenotypic variation, Am. J. Phys. Anthropol. 139 (2009) 16-22.

[6] K.E. Stull, M.W. Kenyhercz, E.N. L'Abbé, Ancestry estimation in South Africa using craniometrics and geometric morphometrics, Forensic Sci. Int. 245 (2014)206e1.

[7] J.L McDowell, M.W. Kenyhercz, E.N. L'Abbé, An evaluation of nasal bone and aperture shape among three South African populations, Forensic Sci. Int. (2015): In Press.

[8] Statistics South Africa, 2014, Mid-year population estimates: Statistical Release.

[9] S.A. Tishkoff, S.M. Williams, Genetic analysis of African populations: Human Evolution and complex disease, Nat. Rev. Genet. 3 (2002) 611-621.

[10] I. Ribot, Differentiation of Modern Sub-Saharan African Populations: Craniometric interpretations in relation to geography and history, Bull. Mém. Soc. Anthropol. Paris 16 (2004) 143-65.

[11] D. Franklin, L. Freedman, N. Milne, C.E. Oxnard CE, Geometric morphometric study of population variation in indigenous southern African crania, Am. J. Hum. Biol. 19 (2007) 20-33.

[12] L.M. Thompson, A history of South Africa, 2001, New Haven: Yale University Press.

[13] J.M. Greeff, Deconstructing Jaco: Genetic heritage of an Afrikaner, Ann. Hum. Genet. 71 (2007) 674-688.

[14] M. Adhikari, Contending approaches to coloured identity and the history of the coloured people of South Africa, History Compass 3 (2005) 1-6.

[15] L. Quintana-Murci, C. Harmant, H. Quach, O. Balanovsky, V. Zaporozhchenko, C. Bormans et al., Strong Maternal Khoisan contribution to the South African coloured population: A case of gender-biased admixture, Am. J. Hum. Genet. 86 (2010) 611–620.

[16] M. Daya, L. van der Merwe, U. Galal, M. Möller, M. Salie, E.R. Chimusa et al., A panel of ancestry informative markers for the five-way admixed South African Coloured Population, PLOS one 8 (2013) e82224.

[17] N. Patterson, D.C. Petersen, R.E. van der Ross, H. Sudoyo, R.H. Glashoff, S. Marzuki, D. Reich, V.M. Hayes, Genetic structure of a unique admixed population: Implications for medical research, Hum. Mol. Gen. 19 (2010) 411-419.

[18] E. de Wit, W. Delport, C.E. Rugamika, A. Meintjes, M. Möller, P.D. van Helden, C. Seoighe, E.G. Hoal, Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape, Hum. Genet. 128 (2010) 145-153.

[19] D.C. Petersen, O. Libiger, E.A. Tindall1, R. Hardie, L.I. Hannick, R.H. Glashoff et al., Complex patterns of genomic admixture within southern Africa, PLOS Genet. 9 (2013) e1003309.

[20] M.Y. İşcan, M. Steyn, Craniometric determination of population affinity in South Africans, Int. J. Legal Med. 112 (1999) 91-97.

[21] E.N. L'Abbé, C. van Rooyen, S.P. Nawrocki, P.J. Becker, An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample, Forensic Sci. Int. 209 (2011) 195-e1.

[22] J.L. McDowell, E.N. L'Abbé, M.W. Kenyhercz, Nasal aperture shape evaluation between black and white South Africans, Forensic Sci. Int. 222 (2012) 397.e1-397.e6.

[23] E.N. L'Abbé, M.W. Kenyhercz, K.E. Stull, N. Keough, S. Nawrocki, Application of Fordisc 3.0 to explore differences among crania of North American and South African blacks and whites, J. Forensic Sci. 6 (2013) 1579-1583.

[24] T.D. Stewart, Essentials of Forensic Anthropology: Especially as developed in the United States, 1979, Charles C Thomas. pp.227-238.

[25] E. St Hoyme, M.Y. İşcan, Determination of sex and race, in: M.Y. İşcan and K.A.R Kennedy (Eds), Reconstruction of life from the skeleton, 1989, Alan R. Liss, Inc. pp. 53-93.

[26] J. Albanese, S.R. Saunders, Is it possible to escape racial typology in forensic identification?, in: A. Schmitt, E. Cunha, J Pinheiro (Eds), Forensic Anthropology and Medicine: Complementary sciences from recovery to cause of death, 2006, Humana Press. pp.281-315.

[27] L. Betti, N. von Cramon-Taubadel, S.J. Lycett, Human pelvis and long bones reveal differential preservation of ancient population history and migration out of Africa, Hum. Biol. 84 (2012) 139-152.

[28] B. Bogin, J. Loucky, Plasticity, political economy, and physical growth status of Guatemala Maya children living in the United States, Am. J. Phys. Anthropol. 102(1997) 17–32.

[29] M. Henneberg, G.A. Harrison, G. Brush, The small child: Anthropometric and physical performance characterisctics of short-for-age children growing in good and in poor socio-economic conditions, Eur. J. Clin. Nutr. 52 (1998) 286-291.

[30] P. Thomas , J. Peabody, V. Tunier, R. Clark, A new look at intrauterine growth and the impact of race, altitude and gender, Pediatrics 106 (2000) e21-e21.

[31] C. Ruff, Ontogenetic adaptation to bipedalism: age changes in femoral to humeral length and strength proportions in humans, with a comparison to baboons, J. Hum. Evol. 45 (2003) 317-349.

[32] L.W. Cowgill, L.D Hager, Variation in the development of postcranial robusticity: an example from Çatalhöyük, Turkey, Int. J. Osteoarchaeol. 17 (2007) 235-252.

[33] M.A. Frelat, P. Mittereocker, Postnatal ontogeny of tibia and femur form in two human populations: a multivariate morphometric analysis, Am. J. Hum. Biol. 23 (2011) 796-804.

[34] N.G. Walensky, A study of the anterior femoral curvature in man, Anat. Rec. 151(1965) 559-570.

[35] R. DiBennardo, T.V. Taylor, Multiple discriminant function analysis of sex and race in the postcranial skeleton, Am. J. Phys. Anthrop. 61 (1983) 305-314.

[36] M.L. Patriquin, M. Steyn, S.R. Loth, Metric assessment of race from the pelvis in South Africans, Forensic Sci. Int. 127 (2002) 104-113.

[37] D.J. Wescott, Population variation in femur subtrochanteric shape, J. Forensic Sci. 50 (2005) 1-8.

[38] M.A. Bidmos, Metrical and non-metrical assessment of population affinity from the calcaneus, Forensic Sci. Int. 159 (2006) 6-13.

[39] *Daubert v. Merrell Dow Pharmaceuticals, Inc*., 509 U.S. 579 (1993).

[40] L. Meintjies-Van der Walt, The proof of the pudding: The presentation and proof of expert evidence in South Africa, J. Afr. Law. 47 (2003) 88-106.

[41] M. Steyn, M.Y. İşcan, Sex determination from the femur and tibia in South African whites, Forensic Sci. Int. 90 (1997) 111-119.

[42] M. Steyn, M.Y. İşcan, Osteometric variation in the humerus: Sexual dimorphism in South Africans, Forensic Sci. Int. 106 (1999) 77-85.

[43] M. Steyn, J.R. Smith, Interpretation of ante-mortem stature estimates in South Africans, Forensic Sci. Int. 171 (2007) 97-102.

[44] C. Small, D.M. Brits, J. Hemingway, Quantification of the subpubic angle in South Africans, Forensic Sci. Int. 222 (2012) 395-e1.

[45] G.C. Krüger, E.N. L'Abbé, K.E. Stull, M.W. Kenyhercz, Sexual dimorphism in cranial morphology among modern South Africans, Int. J. Legal Med. 129 (2015) 869-875.

[46] S.D. Ousley, R.L. Jantz, FORDISC 3 and Statistical Methods for Estimating Sex and Ancestry, in: D.C. Dirkmaat (Ed), A companion to Forensic Anthropology, 2012, Blackwell Publishing LTD. pp.311-329.

[47] R.L. Jantz, S.D. Ousley, FORDISC 3.0: Personal Computer Forensic Discriminant Functions, (2005), Knoxville: University of Tennessee.

[48] M.K. Spradley, Metric ancestry estimation from the postcranial skeleton, in: G.E. Berg and S.C. Ta'ala (Eds), Biological affinity in forensic identification of human skeletal remains: Beyond black and white, 2014, CRC Press. pp. 83 – 94.

[49] P.M. Moore-Jansen, S.D. Ousley, R.L. Jantz, Data collection procedures for forensic skeleton material, 1994, The University of Tennessee, Knoxville: Department of Anthropology. pp. 70-82.

[50] T.A. Perini, G.L. de Oliveira, J.D.S. Ornellas, F.P. de Oliveira, Technical error of measurement in anthropometry, Rev. Bras. Med. Esporte. 11 (2005) 81-85.

[51] R. Goto, C.G.N. Mascie-Taylor, Precision of measurement as a component of human variation. J. Physiol. Anthropol. 26 (2007) 256-256.

[52] S. Stomfai, W. Ahrens, K. Bammann, É. Kovács, S. Mårild, N. Michels, L.A. Moreno, H. Pohlabeln, A. Siani, M. Tornaritis, T. Veidebaum, D. Molnár, Intra- and inter-observer reliability in anthropometric measurements in children, Int. J. Obesity 35 (2011) 45-51.

[53] J.M. Bland, D.G. Altman, Measuring agreement in method comparison studies, Stat. Methods Med. Res. 8 (1999) 135-160.

[54] E.F. Harris, R.N. Smith, Accounting for measurement error: A critical but overlooked process, Arch. Oral. Biol. 54 (2009) 107-117.

[55] S.K. Kachigan, Multivariate statistical analysis: A conceptual introduction, 1991, Radius Press. pp. 216-235.1991

[56] T. Hastie, R. Tibshirani, J. Friedman, The elements of statistical learning: Data mining, inference and prediction, 2009, New York: Springer-Verlag.

[57] K.E. Stull, E.N. L'Abbé, S.D. Ousley, Using multivariate adaptive regression splines to estimate subadult age from diaphyseal dimensions, Am. J. Phys. Anthropol. 154 (2014) 376-386.

[58] S. Milborrow, Earth: Multivariate Adaptive regression spline models, 2011.

[59] T. Hastie, R. Tibshirani, mda: Mixture and flexible discriminant analysis, 2013.

[60] T. Hastie, R. Tibshirani, A. Buja, Flexible Discriminant Analysis by optimal scoring, J. Am. Statist. Assoc. 89 (1994) 1255-1270.

[61] B.G. Tabachnick, L.S Fidell, Using multivariate statistics, 5th ed., 2007, Boston: Pearson Education.

[62] M. Kuhn, Caret: Classification and Regression Training, 2013.

[63] G. James, D. Witten, T. Hastie, R. Tibshirani, An introduction to statistical learning, 2013, New York: Springer.

[64] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[65] E.T. Wood, D.A. Stover, C. Ehret, G. Destro-Bisol, G. Spedini, H. McLeod et al., Contrasting patterns of Y chromosome and mtDNA variation in Africa: Evidence for sex-biased demographic processes, Eur. J. Hum. Genet. 13 (2005) 1-10.

[66] D. Franklin, A. Cardini, C.E.Oxnard, A geometric morphometric approach to the quantification of population variation in sub-Saharan African crania, Am. J. Hum. Biol. 22 (2010) 23-35.

[67] C.K. Jacobson, A.Y. Acheampong, T.B. Heaton, Inter-racial marriages in South Africa, J. Comp. Fam. Stud. 35 (2004) 443-458.