Principles Underlying the Determination of Population Affinity with Craniometric Data

David Bulbeck[1]
The Australian National University

This paper investigates the value for forensic anthropology of craniometric data in assessing population affinity. It finds that generally speaking cranial measurements do not contain the information to directly make a positive match for a skull's population affinity. Rather, cranial measurements should be thought of as containing information that allows for the elimination of any population affinity for the skull which would be a mismatch. A minimum of 13 measurements is required to capture enough information to be confident that the eliminated population affinities are indeed the mismatches. In addition, if a reasonably sized sample of crania from the same population is available for analysis, the affinities of the sampled population can be reliably assessed using the methodology outlined in this paper.

Key Words: Craniometrics; Population Affinity; Multivariate Analysis; Race; Geography.

Introduction

Craniometric analysis is a major tool in the branches of forensic anthropology which deal with osteological remains. Ideally, it would be able to produce a reliable assessment for every skull's population affinity, but there are grounds for believing this is not always the case. This study provides a rationale for why the perfectly correct classification of every skull would be an unrealizable holy grail, regardless of how many measurements are analyzed or how many populations are represented in the comparative database. However, this study also finds that when a sample of skulls is available for analysis, and certain other conditions are satisfied, we can expect correct identification of the affinities of the

[1] Address for correspondence: David Bulbeck, Department of Archaeology and Natural History, School of Culture, History and Language, College of Asia and the Pacific, The Australian National University, Canberra ACT 0200, Australia, mbulbeck@hn.ozemail.com.au

population from which the sample of skulls is drawn.

The present study employs the craniometric module which, as part of the Fordisc 2.0 computer program (Ousley & Jantz 1996), compares individually measured crania with the populations measured by W.W. Howells (1973, 1989). This particular Fordisc 2.0 functionality has been criticized by Williams et al. (2005) on the basis of their analysis of 42 ancient Nubian crania. In the light of previous studies which had found ancient Nubian and Egyptian crania to be metrically similar, Williams et al hypothesized that the Late Period Dynastic Egyptians measured by Howells should emerge as the closest match for most or all of their analyzed Nubian crania. Disappointingly, only a minority of the Nubian crania would have been classified with Howells's Egyptians. Accordingly, Williams et al concluded that factors such as intra-population variation and cranial plasticity (developmental variation) were responsible for the inability of Fordisc 2.0 to provide reliable 'racial' classifications from craniometric data.

Several aspects of the study by Williams et al. (2005) warrant scrutiny. First, as pointed out by Hubbe and Neves (2007), Williams et al. employed only 11 of the theoretical maximum of 21 measurements that could have been used in their analysis. Had they incorporated more information into their analysis by using more measurements, in all likelihood a larger proportion of Nubian crania would have been correctly classified. Secondly, from the point of view of classifying crania to their correct race, the criterion of success for Nubian crania should be to detect a 'Caucasoid' affinity rather than a specifically Egyptian affinity. This is because the Egyptian populations studied by Howells (1973, 1989) are consistently more similar to Europeans than to populations elsewhere in the world. Indeed, in nine cases a European population measured by Howells provided the closest match to one of the Nubian specimens studied by Williams et al. (2005), similar to the ten cases where Howells's Egyptian population made the closest match. Thirdly, Fordisc 2.0 provides considerably more statistical information than merely which is the closest Howells population, and Williams et al. made no use of this additional information.

In reviewing the issues outlined above, this study uses craniometric data recorded for a large sample of recent Thais (Figure 1). Thailand lies near the homelands of several other tropical 'Mongoloid' populations measured by Howells, specifically Hainan Chinese, the Atayal of Taiwan, and Filipinos. However, in terms of geographical proximity to Thailand, the closest of the Howells populations is the Andaman Islanders, who are of unclear 'racial' affinity (Bulbeck et al. 2006). Therefore, if geography were the main determinant of population affinities we would expect Andaman Islanders to be the population most similar to Thais. Conversely, if racial affinity were important but geography were not, we would expect the Thais to show broad affinities with Mongoloids, including those in the New World, but no particular similarity with Andamanese. If both racial affinity and geography were important we would expect other tropical East Asian Mongoloids, specifically the Hainan, Atayal and Filipinos, to be the Howells populations most similar to Thais. Finally, if neither racial affinity nor geography influenced craniometric similarities, we would expect the populations most similar to Thais to be distributed randomly across the globe. These four expectations, respectively labeled 'G', 'R', 'GR' and 'X', are presented in Table 1.

It may be objected that the Andaman Islands are separated from Thailand by sea, and therefore should be thought of as more isolated from Thailand than places on the Eurasian landmass even if their direct geographical distance from Thailand is somewhat greater. However, from the point of view of distinguishing between the G and GR expectations, this objection would be irrelevant, because the Hainan, Atayal and Filipinos are also separated from the Eurasian landmass by sea (Figure 1). Moreover, Andamanese traditional material culture includes outrigger canoes (Cooper 2002), which points to an Andamanese sea-going capacity and in all probability contacts in recent millennia with one or more surrounding maritime societies that introduced the outrigger canoe to the Andaman Islands.

Table 1. Four possible expectations for Thai Crania

Cause for craniometric similarity	Expectation for Thai crania	Label
Geography	Andaman Islanders closest to Thais	G
'Race' (Mongoloid for Thais)	Mongoloid populations across East Asia, the Pacific and New World closest to Thais	R
Both geography and race	Hainan, Atayal and Filipinos closest to Thais	GR
Neither geography nor race a cause for craniometric similarity	Populations other than Mongoloids and Andaman Islanders closest to Thais	X

Two other questions raised by the study of Williams et al. will be investigated here. The first question is how many measurements are required in order to obtain reliable results. Say for instance that race emerges as the crucial determinant for craniometric similarity, and so a successful analysis would be one where Howells's Mongoloid populations are found to be closest to Thais. The answer to our first question would then be: how many measurements should be used before the addition of another measurement would not significantly increase the proportion of Mongoloid classifications. The second question is whether there are more effective methods for interpreting the

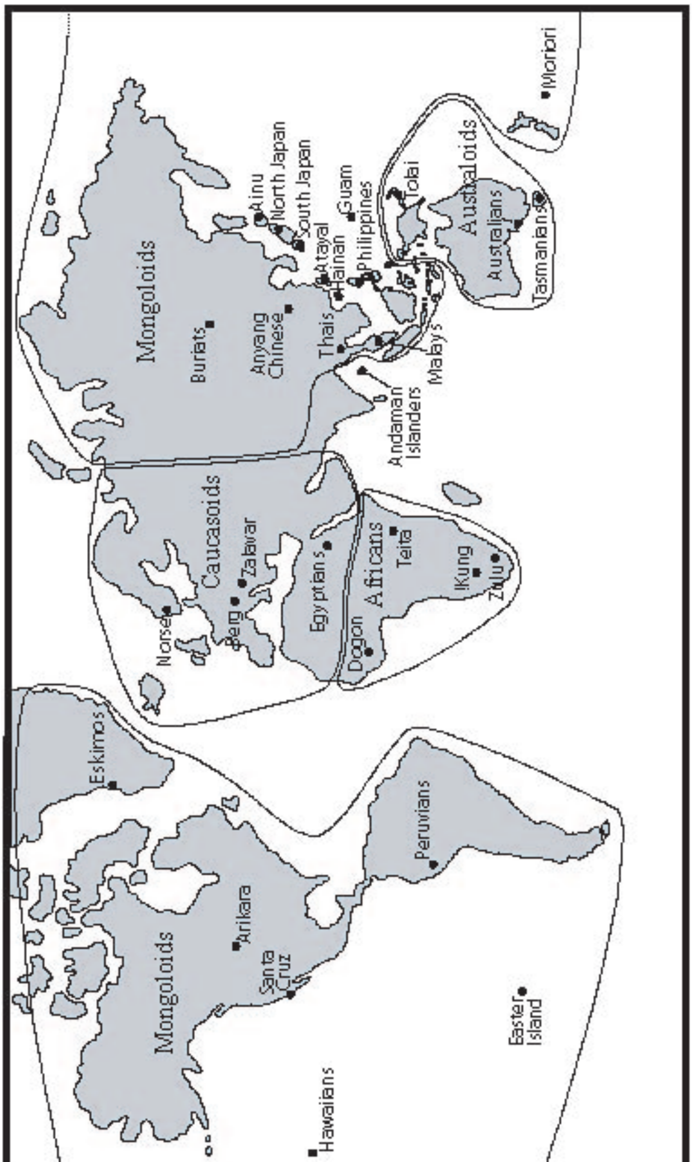

Figure 1. Location of Thais, Malays, and the Howells populations

Fordisc 2.0 results than to simply consider the 'classification' that would be made based on the closest Howells population. For instance, a Thai cranium might be classified as non-Mongoloid on the basis that the closest Howells population is not Mongoloid, but have Mongoloid affinities in the sense that all of the other Howells populations close to it are Mongoloid. If these secondary affinities could be incorporated into the analytical method then the analysis might be more diagnostic. Indeed, analytical methods that are not based simply on classifications might prove to be particularly robust in the sense that relatively few measurements might be required before obtaining a result that did not change significantly with the addition of further measurements.

The expectations of the multiple hypotheses investigated in this paper are summarized in Table 2.

One issue not addressed in this study is whether Fordisc 3 (Jantz & Ousley 2003) might be an improvement on Fordisc 2.0 in realizing the utility of craniometrics to detect population affinity. There are two main reasons for restricting this study to Fordisc 2.0. First, the Thai measurements (Saengvichien 1971) were taken using the main measurements in Martin's system (Martin & Saller 1957), and Fordisc 2.0 accommodates these measurements as well as Fordisc 3 does. Secondly, background information relevant to this study has already been generated using Fordisc 2.0 (Bulbeck et al. 2006).

Materials and Methods

The data employed in this study are the individual measurements provided by Saengvichien (1971) for 145 skulls of known Thai adults, curated in the Congden Anatomical Laboratory in Bangkok. Up to 20 of the measurements utilized by Fordisc 2.0 are provided by Saengvichien, but many of the crania lack some of these measurements. Three of those most frequently missing are palate breadth, nasion-prosthion length and basion-prosthion length, which suggests that necrosis of the dental arcade, probably through periodontal disease, had obliterated the anatomical landmarks required to take these measurements.

Table 2. Expectations* for Thai crania based on this paper's multiple hypotheses

Test conditions	Geography important	Race important	Race and geography both important	Neither race nor geography important
Most or all measurements available	G	R	GR	X
Classifications reliable even for small measurement subsets	G here and above	R here and above	GR here and above	X here and above
Only techniques other than classifications work for small measurement subsets	G here, but not above in row 2	R here, but not above in row 2	GR here, but not above in row 2	X here, but any result in row 2 above
Small measurement subsets unreliable with any technique	R, GR or X	G or X	G, R or X	Any result

* *For explanation of the G, R, GR and X labels, see Table 1.*

Fordisc 2.0 is used to compare the Thais craniometrically with the populations measured by Howells (1973, 1989). These populations are spread across the world excluding South Asia (Figure 1). Note that there are more Mongoloid than other populations in the Howells database, especially if the Ainu of Japan are considered Mongoloid, as they appear to be craniometrically (Howells 1989: Figure 3 and 4). Amongst the 28 Howells male populations (Table 3), 16

Table 3. Fordisc 2.0 results for male Thai crania S.243 and Sankas 24 (20 variables each)

Howells male population	S.243 Typicality probability	S.243 Posterior probability	Sankas 24 Typicality probability	Sankas 24 Posterior probability
Anyang Chinese (Mongoloid)	0.203	0.265*	0.000	0.000
South Japanese (Mongoloid)	0.203	0.265	0.000	0.002
Guam Micronesians (Mongoloid)	0.202	0.262	0.000	0.000
Filipinos (Mongoloid)	0.117	0.068	0.000	.809*
Hawaii Polynesians (Mongoloid)	0.105	0.053	0.000	0.002
Hainan Chinese (Mongoloid)	0.094	0.041	0.000	0.001
Tolai Melanesians (Australoid)	0.081	0.029	0.000	0.020
North Japanese (Mongoloid)	0.048	0.010	0.000	0.000
Zulu (African)	0.020	0.002	0.000	0.001
Taiwan Atayal (Mongoloid)	0.016	0.001	0.000	0.000
Easter Island Polynesians (Mongoloid)	0.015	0.001	0.000	0.000
Moriori Polynesians (Mongoloid)	0.008	0.000	0.000	0.000
Zalavár Hungarians (Caucasoid)	0.008	0.000	0.000	0.000

Tasmanians (Australoid)	0.007	0.000	0.000	0.027
Ainu (craniometrically Mongoloid)	0.007	0.000	0.000	0.000
Greenland Eskimos (Mongoloid)	0.007	0.000	0.000	0.000
Mali Dogon (African)	0.006	0.000	0.000	0.056
Arikara Amerinds (Mongoloid)	0.006	0.000	0.000	0.000
Santa Cruz Amerinds (Mongoloid)	0.003	0.000	0.000	0.009
Andaman Islanders (unassigned)	0.002	0.000	0.000	0.069
Peru Amerinds (Mongoloid)	0.002	0.000	0.000	0.002
1st Dynasty Egyptians (Caucasoid)	0.002	0.000	0.000	0.000
Swanport Australians (Australoid)	0.001	0.000	0.000	0.000
Kenyan Teita (African)	0.001	0.000	0.000	0.000
Mongolian Buriats (Mongoloid)	0.000	0.000	0.000	0.000
Berg Austrians (Caucasoid)	0.000	0.000	0.000	0.000
Oslo Norse (Caucasoid)	0.000	0.000	0.000	0.000
San Bushmen (African)	0.000	0.000	0.000	0.000
Sum of probabilities	1.059	1.000	0.000	1.000

** The posterior probability of the Howells population closest to the analyzed Thai skull.*

Table 4. Fordisc 2.0 results for female Thai crania S.108 (17 variables) and S.74 (20 variables)

Howells female population	S.108 Typicality probability	S.108 Posterior probability	S.74 Typicality probability	S.74 Posterior probability
Hawaii Polynesians (Mongoloid)	0.822	0.542*	0.002	0.199*
Zulu (African)	0.640	0.136	0.000	0.000
Ainu (craniometrically Mongoloid)	0.593	0.099	0.000	0.000
Zalavár Hungarians (Caucasoid)	0.505	0.053	0.000	0.009
Guam Micronesians (Mongoloid)	0.452	0.036	0.000	0.003
1st Dynasty Egyptians (Caucasoid)	0.424	0.029	0.000	0.003
Hainan Chinese (Mongoloid)	0.405	0.025	0.000	0.005
Oslo Norse (Caucasoid)	0.362	0.018	0.000	0.002
Mali Dogon (African)	0.350	0.016	0.000	0.001
Moriori Polynesians (Mongoloid)	0.328	0.013	0.000	0.006
Arikara Amerinds (Mongoloid)	0.262	0.007	0.002	0.149
Tasmanians (Australoid)	0.257	0.007	0.000	0.001
North Japanese (Mongoloid)	0.237	0.006	0.000	0.001
South Japanese (Mongoloid)	0.197	0.004	0.000	0.000
Taiwan Atayal (Mongoloid)	0.177	0.003	0.002	0.194
Berg Austrians (Caucasoid)	0.163	0.002	0.001	0.109
Mongolian Buriats (Mongoloid)	0.130	0.001	0.002	0.127
Tolai Melanesians (Australoid)	0.113	0.001	0.000	0.000
San Bushmen (African)	0.061	0.000	0.000	0.000
Peru Amerinds (Mongoloid)	0.055	0.000	0.002	0.125

Kenyan Teita (African)	0.051	0.000	0.000	0.000
Santa Cruz Amerinds (Mongoloid)	0.039	0.000	0.000	0.000
Greenland Eskimos (Mongoloid)	0.038	0.000	0.000	0.000
Easter Island Polynesians (Mongoloid)	0.034	0.000	0.000	0.000
Swanport Australians (Australoid)	0.027	0.000	0.000	0.000
Andaman Islanders (unassigned)	0.016	0.000	0.001	0.066
Sum of probabilities	6.738	1.000	0.012	1.000

* *The posterior probability of the Howells population closest to the analyzed Thai skull.*

(57%) are Mongoloid, including three close to Thailand (11%), four are Caucasoid (14%), four Sub-Saharan African (14%), and three southwest Pacific or 'Australoid' (11%), while the Andaman Islanders are unassigned (4%). Considering this composition of the Howells database, we would infer that the a priori probability of obtaining the 'G' expectation is 4%, the 'GR' expectation is 11%, the 'R' expectation is 57% and the 'X' expectation is 39%. Amongst the 26 female populations (Table 4), 14 are Mongoloid (54%), including two close to Thailand (8%), and the number is the same as for males with Caucasoids (15%), Africans (15%), Australoids (12%) and Andaman Islanders (4%). Accordingly, for a female skull the a priori probability of expectation 'G' is 4%, 'GR' 8%, 'R' 54% and 'X' 42%.

Craniometric analysis proceeded as follows. With each Fordisc 2.0 program run, the measurements of a 'target specimen' (e.g., S.243 in Table 3) are entered. Fordisc 2.0 uses canonical variate analysis to calculate the 'typicality' probability (TP) that a specimen with these measurements would belong to each Howells population included in the analysis. The computer program then uses linear discriminant analysis to maximize the correct classification of the populations measured by Howells, and calculates the relative or 'posterior' probabilities (PP) of the specimen's membership with every Howells population (see Tables 3 and 4).

The TP and PP carry different types of information. With each run, the target specimen's TP range between 0 and 1 with respect to each population in the analysis, whereas the target specimen's PP sum to 1 with respect to all populations in the analysis. The implications can be comprehended by sampling the variety of results. A skull can combine a (virtually) zero TP of belonging to a Howells population with a very high PP, approaching unity, that it would belong to that Howells population if it belonged to any of them (Sankas 24 and Filipinos, Table 3). Conversely, a skull could be over 50% 'typical' of a Howells population and yet have a low PP, barely 5%, of being assigned to that particular Howells population (S.108 and Zalavár, Table 4). The frequently arbitrary nature of classifying a specimen based on which particular Howells population is the closest is also evident. Anyang Chinese, South Japanese and Guam Micronesians are all, essentially, equally close to S.243 (Table 3), as are Hawaiians and Atayal with respect to S.74 (Table 4).

The Fordisc 2.0 program was run a total of 1,640 times to produce the data used in this study. Initially it was run 144 times (85 times for the males and 59 times for the females) using all of the Fordisc-compatible measurements provided by Saengvichien that did not generate a warning from Fordisc 2.0 of being too high or too low.[2] I then repeated the program runs for all of the Thai (male of female) skulls with all of the measurements in the 26 measurement suites listed in the Appendix to this paper (Table 5). These suites of measurements are the sets of three or more measurements (eligible for Fordisc 2.0 analysis), up to 19 measurements, published for skulls from the Neolithic sites of Ban Kao (Sangvichien et al. 1969) and Khok Phanom Di (Tayles 1999). They often included palate breadth, nasion-prosthion length or basion-prosthion length, in which case many of the Thai crania were ineligible for inclusion. While these measurement sets were not randomly generated, they tend

[2] Whenever any of Saengvichien's measurements produced either of these warnings, it was excluded from analysis, because of the risk of a misprint or measurement error.

to differ substantially from each other owing to the vagaries of archaeological preservation. They are here assumed to satisfactorily illustrate how the Fordisc 2.0 classifications, including the generated TP and PP, are affected by entering different numbers (3 to 19) of measurements.

One approach in this investigation employs the Fordisc 'classifications', i.e. the Howells population with the highest TP/PP on any run. This is the Fordisc functionality utilized by Williams et al (2005) in their analysis of Nubian crania. My second approach is to treat the generated TP and PP as the data for analysis. For instance, referring to the results in Tables 3, we have two TP (0.117, 0.000) and two PP (0.068, 0.809) documenting the similarity of Filipino males to Thai males, two TP (0.105, 0.000) and two PP (0.053, 0.002) for the similarity of Hawaiian males to Thai males, etc. The more similar a Howells population is to the Thais on any particular measurement suite, the higher its TP and PP should tend to be. In this second approach, all the Fordisc-generated information, not just the classifications, can be used in assessing the relative craniometric similarities of Thais to the various Howells populations.

One problem with analyzing the probabilities (both TP and PP) is that they are dominated by values of 0.000, at three decimal places (cf. Table 3). The distributions have a strong positive skew, which makes any reliance on mean values potentially misleading. Standard normalizing techniques such as log-transforms would have little effect in correcting the distributions' positive skew, because so many values are not distinguished from zero. Similarly, the median values would also be ineffective in distinguishing between populations because the median in most cases would be 0.000 or a similarly tiny fraction. Accordingly, populations are compared for their similarity to Thais based on benchmark percentile values above the median. Percentile analysis was performed using the Excel spreadsheet percentile function.

Table 5. Information on measurement suites (see Appendix) used in this study

Measurement suite	Sex	Number of measurements	Number of specimens
All available	M	Average 19.1	85
General #2	M	16	53
General #4	M	13	55
General #5	M	12	54
General #9	M	11	54
General #10	M	11	61
General #11	M	10	54
General #12	M	9	63
General #14	M	8	62
Facial #1	M	7	55
General #17	M	6	59
General #20	M	4	55
Facial #3	M	3	83
Facial #5	M	3	63
All available	F	Average 19.2	59
General #1	F	19	40
General #3	F	13	42
General #6	F	11	43
General #7	F	11	41

General #8	F	11	40
General #11	F	10	42
General #13	F	8	45
General #14	F	8	43
General #15	F	7	55
General #16	F	7	42
General #18	F	5	42
General #19	F	4	59
Facial #2	F	4	42
Cranial #1	F	3	58
Facial #4	F	3	48
Facial #5	F	3	43

In the presentation of the main results (Tables 10 to 19), the last column indicates which of the G, R, GR or X expectations is supported by the analysis. This is based on the highest 'observed-to-expected ratio' with regard to the proportional representation of populations in the Howells database. For instance, looking at classifications, we would expect just one of each 28 Thai male skulls to be classified as Andamanese ('G') because Andamanese make up just one of the 28 Howells male populations. Say a quarter of the male Thai skulls were classified as Andamanese, this would be seven times (700%) the expected number. More Thai skulls might be classified with some other population but this need not imply greater support than for the G expectation. For instance, with so many Mongoloid populations in the Howells database, by chance alone there might be more classifications to a particular Mongoloid population than to Andamanese. To find support for the 'R' expectation, we would expect that the proportion of Thai male skulls classified with a Mongoloid population, divided by the proportion of Howells male populations that are Mongoloid (57%), would exceed the observed-to-expected ratio for the GR and X populations as well as the G (Andamanese) population.

In addition, the main results present the five Howells populations most similar to Thais in each analysis. The number was set at five for two reasons. First, if two analyses are producing similar results, we would expect the closest five populations in one analysis to be much the same as the closest five in the other analysis (but not necessarily in the same order, from closest to fifth closest). Secondly, the number five casts a sufficiently wide net to capture the populations close to Thais for the analysis of percentile values. For instance, we would not expect Andamanese to be amongst the closest five by chance alone, as they make up only one of 28 Howells male populations and one of 26 Howells female populations. If they do occur amongst the closest five, their observed-to-expected ratio would be respectively 560% for males and 520% for females.[2]

In some cases, there is only a small difference between the supported expectation and the second-best supported expectation in terms of their observed-to-expected ratios. To determine whether a supported expectation is statistically significant at the 95% confidence level, the Wilson 95% confidence interval (Wilson 1927) was calculated for the numerator and denominator. If the observed-to-expected ratio for every value in the Wilson confidence interval exceeds the observed-to-expected ratio for any alternative expectation, then statistically significant support for the expectation in question is inferred (Table 20).

---

[2] These are high observed-to-expected ratios but they can be equaled by GR populations. If all three male GR populations (Filipinos, Hainan and Atayal) are amongst the closest five, the resulting observed-to-expected ratio is 560%, and if both female GR populations (Hainan and Atayal) are amongst the closest five, the resulting observed-to-expected ratio is 520%. If one of these results is obtained as well as Andamanese amongst the closest five, the GR expectation is deemed to be more strongly supported than the G expectation, because of the larger number of populations involved in its support.

The analytical methodology employed in this study is summarized in Table 6.

Table 6. Summary of analytical methodology for G, R, GR and X expectations*

	Classification analysis	**Percentile analysis**
Numerator	Number of Thai skulls with a G, R, GR or X classification	Number of G, R, GR or X populations amongst five denominator populations
Denominator	Number of Thai skulls in the analysis	Five populations with highest percentile values
Supported expectation	Whichever of G, R, GR or X has highest observed-to-expected ratio	Whichever of G, R, GR or X has highest observed-to-expected ratio
Calculation of observed-to-expected ratio	Divide by the proportion of Howells populations that are G, R, GR or X	Divide by the proportion of Howells populations that are G, R, GR or X
Confidence interval	Wilson interval based on numerator and denominator	Wilson interval based on numerator and denominator
Statistically significantly supported expectation	Observed-to-expected ratio for entire Wilson interval higher than any other observed-to-expected ratio	Observed-to-expected ratio for entire Wilson interval higher than any other observed-to-expected ratio

** For explanation of the G, R, GR and X labels, see Table 1.*

Results *Thai and Malay Comparisons: All Measurements Available per Specimen*

As an introduction to the main analysis, it is instructive to compare Thais with Malays, a Mongoloid population which overlaps geographically with Thais (Figure 1). Malays are not one of the populations measured by Howells, but they have been compared with the Howells populations, using Fordisc 2.0, by Bulbeck et al. (2006). If either geography or race is important for craniometric similarities, and especially if both are important, we would expect Thais and Malays to be very similar in how they compare to the Howells populations.

Table 7 shows that Thais and Malays are very similar in their racial classifications. In both cases, over 80% of crania would be classified as Mongoloid, on the basis of having a Mongoloid population as their closest Howells population. This proportion is higher than the expected c. 55% (see Materials and Methods). Caucasoid classifications are the second most common, and African classifications the least common, for both Thais and Malays. They both contrast strongly in these regards with Australian Aborigines, eastern Indonesians and Punjabis from India (Bulbeck et al. 2006).

Table 7. Thai and Malay classifications compared (sexes combined)

Classification	Thais (n = 144)	Malays (n = 92)
Mongoloid (including Ainu)	121 (84.0%)	74 (80.4%)
Caucasoid	13 (9.0%)	7 (7.6%)
Australoid	7 (4.9%)	6 (6.5%)
Andamanese	2 (1.4%)	3 (3.3%)
Africans	1 (0.7%)	2 (2.2%)

In addition to race, geography also plays a role in these classifications. Far more male Thais (23 cases) and male Malays (16 cases) would be classified as Filipinos than any other male Howells population. However, since the Howells database does not include female Filipinos, it would not be possible for Thai females to be classified as Filipino. Here we find that female Thais (22 and 11 cases respectively) and female Malays (8 and 9 cases respectively) would both be most frequently classified as Hawaiian or Buriat.

Table 8. Ninetieth percentile posterior probabilities for Thais and Malays (top three for either)

Howells population	Thais	Malays
Hawaiians	**0.700**	0.623
Filipinos	0.531	**0.687**
Buriats	0.314	0.516

Table 9. Ninetieth percentile typicality probabilities for Thais and Malays (top three for either)

Howells population	Thais	Malays
Filipinos	**0.415**	**0.330**
Hainan Chinese	0.248	0.270
Anyang Chinese	0.213	0.051 (20^{th})
Hawaiians	0.159 (5^{th})	0.290

More revealing of the tropical East Asian Mongoloid status of Thais and Malays is percentile analysis. This analysis additionally accommodates the lack of female Filipinos amongst the Howells populations, because the result obtained for male Filipinos stays the same even when sexes are combined (as done here for the Howells populations represented by both males and females). At the 90^{th} percentile benchmark (Tables 8 and 9), the strong affinities of both Thais and Malays to both Filipinos and Hawaiians are revealed by both the PP and TP. In addition, with the TP 90^{th} percentile scores, Hainan Chinese emerge as a strong match for both Thais and Malays.

Table 10. Classification results for male Thai crania using different measurement suites

Suite	Most classifications	%	Second most classifications	%	Third most classifications	%	Fourth most classifications	%	Fifth most classifications	%	Supported expectation
All available	**Filipinos**	27.1	**Arikara**	12.9	**Hawaiians**	10.6	**Anyang**	8.2	**Zalavár**	7.1	GR
General #2	**Filipinos**	17.0	Buriats	13.2	**Anyang**	11.3	Guam	9.4	**Arikara**	7.5	GR
General #4	**Arikara**	25.5	Buriats	16.4	Hainan	9.1	**Filipinos**	5.5	Peru	5.5	R
General #5	Tasmanians	18.5	**Arikara**	7.4	**Hawaiians**	7.4	Guam	7.4	**Anyang**	5.6	R
General #9	**Arikara**	22.2	Buriats	13.0	Andamanese	11.1	Guam	7.4	**Anyang**	7.4	G
General #10	**Arikara**	21.3	Berg	18.0	**Filipinos**	11.5	Buriats	11.5	**Zalavár**	6.3	GR
General #11	**Filipinos**	18.5	Hainan	16.7	Andamanese	11.1	**Anyang**	9.3	**Hawaiians**	5.6	GR
General #12	**Arikara**	11.1	Tasmanians	9.5	Australians	9.5	Guam	7.9	**Zalavár**	6.3	X
General #14	**Filipinos**	19.4	Hainan	16.1	Andamanese	12.9	Arikara	8.1	Anyang	6.5	GR

Facial #1	Moriori	20.0	Australians	12.7	Egypt	12.7	Tolai	7.3	South Japan	5.5	X
General #17	**Arikara**	33.9	Buriats	15.3	Berg	15.3	Andamanese	6.8	**Zalavár**	3.4	G
General #20	Santa Cruz	12.7	Norse	12.7	**Anyang**	10.9	Egypt	9.1	Buriats	7.3	G*
Facial #3	Buriats	15.3	**Anyang**	10.8	**Filipinos**	8.4	Peru	8.4	Berg	7.2	GR
Facial #5	Egypt	11.1	Easter Island	11.1	Moriori	9.5	Buriats	6.3	Tasmanians	6.3	G*

** Although there were not enough Andamanese classifications for them to be in the closest five, there were enough for the ratio of observed to expected Andamanese classifications to be higher than for any other expectation (see Materials and Methods).*

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 11. Classification results for female Thai crania using different measurement suites

Suite	Most classifications	%	Second most classifications	%	Third most classifications	%	Fourth most classifications	%	Fifth most classifications	%	Supported expectation
All available	**Hawaiians**	39.0	**Buriats**	16.9	**Hainan**	10.2	**Guam**	6.8	**North Japan**	5.1	GR
General #1	**Hawaiians**	25.0	**Buriats**	12.5	**Hainan**	12.5	**Guam**	10.0	Atayal	7.5	GR
General #3	**Buriats**	23.8	**Hawaiians**	21.4	**Guam**	7.1	Atayal	7.1	Peru	7.1	R
General #6	**Buriats**	30.2	Berg	20.9	**Hawaiians**	9.3	**Hainan**	9.3	Atayal	7.0	GR
General #7	**Hawaiians**	22.0	Peru	17.1	**Buriats**	14.6	Berg	7.3	**Guam**	4.9	R
General #8	**Buriats**	25.0	**Hawaiians**	15.0	Dogon	15.0	**Guam**	12.5	**Hainan**	7.5	G*
General #11	Andamanese	21.4	**Hawaiians**	16.7	**Hainan**	16.7	**Buriats**	9.5	**Guam**	9.5	G
General #13	**Hawaiians**	24.4	Andamanese	24.4	**Hainan**	13.3	**Guam**	8.9	**Buriats**	6.7	G
General #14	Andamanese	23.3	**Hawaiians**	20.9	**Hainan**	18.6	**Buriats**	9.3	Dogon	9.3	G

General #15	**Hawaiians**	21.8	Andamanese	21.8	**Hainan**	18.2	**Buriats**	9.1	Dogon	9.1	G
General #16	**Buriats**	31.0	Andamanese	21.4	Berg	14.3	Dogon	7.1	**Hainan**	4.8	G
General #18	**Buriats**	16.7	Egypt	14.3	**Hawaiians**	11.9	Berg	11.9	Dogon	11.9	G*
General #19	**Buriats**	25.4	Andamanese	23.7	Berg	10.2	Eskimos	8.5	Arikara	6.8	G
Facial #2	Berg	29.3	**Buriats**	9.5	Zalavár	9.5	Tolai	9.5	Australians	7.1	X
Cranial #1	Berg	29.3	**Buriats**	22.4	Andamanese	20.7	Arikara	12.1	Dogon	3.4	G
Facial #4	**Buriats**	8.3	Dogon	8.3	Moriori	8.3	Atayal	6.3	Zalavár	6.3	X
Facial #5	Tasmanians	14.0	**Buriats**	11.6	**Guam**	9.3	Berg	9.3	Zulu	7.0	X

** Although there were not enough Andamanese classifications for them to be in the closest five, there were enough for the ratio of observed to expected Andamanese classifications to be higher than for any other expectation (see Materials and Methods).*

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 12. Ninetieth percentile posterior probability (PP) results for male Thai crania using different measurement suites

Suite	Highest PP	PP	Second highest PP	PP	Third highest PP	PP	Fourth highest PP	PP	Fifth highest PP	PP	Supported expectation
All available	**Filipinos**	.493	**Arikara**	.416	**Hawaiians**	.329	**Anyang**	.247	**Guam**	.230	GR
General #2	**Filipinos**	.476	**Guam**	.407	Buriats	.387	**Arikara**	.270	**Anyang**	.266	GR
General #4	**Arikara**	.542	Buriats	.458	**Hawaiians**	.189	**Filipinos**	.187	Hainan	.169	GR
General #5	Tasmanians	.620	**Guam**	.233	**Arikara**	.222	Zalavár	.194	Moriori	.158	R
General #9	**Arikara**	.476	Buriats	.318	Andamanese	.293	**Guam**	.198	**Filipinos**	.189	G
General #10	**Arikara**	.433	Berg	.362	Buriats	.344	**Filipinos**	.249	Tasmanians	.226	GR
General #11	Andamanese	.389	**Filipinos**	.388	Hainan	.344	**Anyang**	.251	**Hawaiians**	.163	G
General #12	Tasmanians	.254	Australians	.199	**Arikara**	.185	Moriori	.169	Zalavár	.168	X

General #14	Andamanese	.412	**Filipinos**	.407	Hainan	.361	**Anyang**	.160	Buriats	.153	G
Facial #1	Egypt	.290	Moriori	.210	Australians	.187	Ainu	.149	**Hawaiians**	.143	R
General #17	**Arikara**	.473	Buriats	.424	Berg	.292	**Guam**	.124	**Filipinos**	.113	GR
General #20	Egypt	.143	Norse	.130	Santa Cruz	.128	**Anyang**	.120	Peru	.111	R
Facial #3	Buriats	.225	Peru	.135	**Filipinos**	.135	Santa Cruz	.125	**Anyang**	.122	GR
Facial #5	Moriori	.163	Easter Island	.120	Buriats	.111	**Guam**	.107	Egypt	.100	R

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 13. Sixtieth percentile posterior probability (PP) results for male Thai crania using different measurement suites

Suite	Highest PP	PP	Second highest PP	PP	Third highest PP	PP	Fourth highest PP	PP	Fifth highest PP	PP	Supported expectation
All available	**Filipinos**	.189	**Arikara**	.048	**Hainan**	.036	**South Japan**	.024	**Hawaiians**	.022	GR
General #2	**Filipinos**	.118	**Arikara**	.049	**Guam**	.052	**South Japan**	.028	**Hainan**	.019	GR
General #4	**Arikara**	.117	**Filipinos**	.055	**Hainan**	.041	**Guam**	.021	**South Japan**	.018	GR
General #5	**Arikara**	.102	Zalavár	.075	**Guam**	.045	Egypt	.041	Norse	.029	X
General #9	**Arikara**	.147	**Filipinos**	.072	Buriats	.051	Berg	.038	**Hainan**	.034	GR
General #10	**Arikara**	.084	**Filipinos**	.081	Berg	.065	**Hainan**	.036	Zalavár	.026	GR
General #11	**Filipinos**	.139	**Hainan**	.134	**Anyang**	.057	**South Japan**	.045	North Japan	.022	GR
General #12	**Arikara**	.057	Australians	.047	Guam	.045	Zalavár	.043	Egypt	.043	X

General #14	**Hainan**	.176	**Filipinos**	.120	Anyang	.058	**South Japan**	.036	Atayal	.026	GR
Facial #1	Moriori	.077	Ainu	.067	**Hawaiians**	.067	Australians	.051	Zalavár	.043	R
General #17	**Arikara**	.184	Berg	.065	Santa Cruz	.059	**Filipinos**	.050	**Hainan**	.045	GR
General #20	Peru	.064	**Anyang**	.058	Berg	.058	Zalavár	.058	**Hainan**	.055	GR
Facial #3	**Hainan**	.068	**Filipinos**	.064	**Anyang**	.058	Atayal	.052	Berg	.048	GR
Facial #5	**Hawaiians**	.051	**Hainan**	.049	**Arikara**	.048	Norse	.045	Ainu	.044	GR

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 14. Ninetieth percentile typicality probability (TP) results for male Thai crania using different measurement suites

Suite	Highest TP	TP	Second highest TP	TP	Third highest TP	TP	Fourth highest TP	TP	Fifth highest TP	TP	Supported expectation
All available	**Filipinos**	.403	**South Japan**	.305	**Hainan**	.300	**Anyang**	.201	**North Japan**	.175	GR
General #2	**Filipinos**	.351	**South Japan**	.339	**Hainan**	.278	**Anyang**	.258	**Guam**	.212	GR
General #4	Arikara	.468	**Filipinos**	.436	**South Japan**	.385	**Hainan**	.373	**Anyang**	.345	GR
General #5	**Guam**	.544	Arikara	.512	Zalavár	.460	Egypt	.440	Norse	.407	X
General #9	**Arikara**	.428	Berg	.365	**Filipinos**	.351	**Hainan**	.317	Andamanese	.279	G
General #10	Berg	.367	Zalavár	.330	**Arikara**	.322	**Filipinos**	.301	**Hainan**	.272	GR
General #11	**Hainan**	.698	**Filipinos**	.681	**South Japan**	.497	**North Japan**	.487	Peru	.391	GR
General #12	Egypt	.589	**Arikara**	.572	Zalavár	.539	**South Japan**	.492	Moriori	.468	R

General #14	**Hainan**	.705	**Filipinos**	.624	**North Japan**	.469	**Anyang**	.457	**South Japan**	.406	GR
Facial #1	Moriori	.641	Hawaiians	.607	Ainu	.583	Egypt	.564	Australians	.516	R
General #17	**Arikara**	.656	Santa Cruz	.492	Berg	.465	Zalavár	.484	**Filipinos**	.410	GR
General #20	Zalavár	.902	Egypt	.891	**Hainan**	.884	Berg	.882	Norse	.881	X
Facial #3	**Anyang**	.894	Berg	.874	**Hainan**	.870	Santa Cruz	.839	Peru	.832	R
Facial #5	Zalavár	.919	Berg	.902	Norse	.891	**North Japan**	.869	Peru	.860	X

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 15. Seventieth percentile typicality probability (TP) results for male Thai crania using different measurement suites

Suite	Highest TP	TP	Second highest TP	TP	Third highest TP	TP	Fourth highest TP	TP	Fifth highest TP	TP	Supported expectation
All available	**Filipinos**	.078	**Arikara**	.037	**Hainan**	.036	**South Japan**	.029	**Hawaiians**	.028	GR
General #2	**Filipinos**	.096	**Hainan**	.061	Guam	.058	**South Japan**	.050	**Arikara**	.043	GR
General #4	**Filipinos**	.138	**Arikara**	.138	**Hainan**	.114	Guam	.099	**South Japan**	.085	GR
General #5	**Arikara**	.276	Zalavár	.202	**Filipinos**	.162	Egypt	.150	Guam	.147	GR
General #9	**Arikara**	.218	Berg	.178	**South Japan**	.153	Guam	.130	**Filipinos**	.123	R
General #10	**Filipinos**	.172	**Arikara**	.088	**South Japan**	.088	**Hawaiians**	.071	Berg	.070	R
General #11	**Filipinos**	.301	**Hainan**	.262	**Anyang**	.172	**South Japan**	.153	Atayal	.117	GR
General #12	Guam	.260	Zalavár	.239	**Arikara**	.221	Moriori	220	Egypt	.200	R

General #14	**Hainan**	.357	**Filipinos**	.317	Anyang	.189	**South Japan**	.160	North Japan	.110	GR
Facial #1	Moriori	.352	**Hawaiians**	.327	Guam	.280	Ainu	.239	**Arikara**	.227	R
General #17	**Arikara**	.341	Santa Cruz	.249	Guam	.192	North Japan	.180	Berg	.175	R
General #20	North Japan	.692	Egypt	.684	**South Japan**	.676	Santa Cruz	.651	Norse	.649	R
Facial #3	**Hainan**	.677	**Anyang**	.623	Atayal	.606	**Filipinos**	.596	Berg	.593	GR
Facial #5	**Hainan**	.637	**Hawaiians**	.604	North Japan	.604	Zalavár	.591	Norse	.585	GR

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 16. Ninetieth percentile posterior probability (PP) results for female Thai crania using different measurement suites

Suite	Highest PP	PP	Second highest PP	PP	Third highest PP	PP	Fourth highest PP	PP	Fifth highest PP	PP	Supported expectation
All available	**Hawaiians**	.881	**Buriats**	.574	**Hainan**	.359	**Guam**	.203	**Atayal**	.129	GR
General #1	**Hawaiians**	.706	**Buriats**	.369	**Hainan**	.339	**Guam**	.316	**Atayal**	.161	GR
General #3	**Buriats**	.545	**Hawaiians**	.522	**Atayal**	.248	Arikara	.234	Peru	.226	GR
General #6	**Buriats**	.774	Peru	.266	Berg	.484	**Hawaiians**	.265	**Atayal**	.221	GR
General #7	**Buriats**	.609	**Hawaiians**	.326	Peru	.293	Tasmanians	.258	Arikara	.135	R
General #8	**Buriats**	.802	Dogon	.365	**Hawaiians**	.358	**Guam**	.282	**Hainan**	.215	GR
General #11	Andamanese	.578	**Hawaiians**	.497	**Hainan**	.336	**Buriats**	.302	Dogon	.222	G
General #13	**Hawaiians**	.699	Andamanese	.637	**Hainan**	.370	**Buriats**	.245	Dogon	.141	G

General #14	Andamanese	.563	**Hawaiians**	.491	**Hainan**	.392	Dogon	.271	**Buriats**	.271	G
General #15	**Hawaiians**	.660	Andamanese	.484	**Hainan**	.324	Dogon	.263	**Buriats**	.237	G
General #16	**Buriats**	.895	Andamanese	.445	Berg	.347	Dogon	.169	**Hainan**	.131	G
General #18	**Buriats**	.421	Egypt	.245	Berg	.209	**Hawaiians**	.171	Dogon	.163	X
General #19	Andamanese	.496	**Buriats**	.310	Arikara	.194	Dogon	.151	Berg	.136	G
Facial #2	**Buriats**	.174	Tasmanians	.161	**Guam**	.150	Australians	.132	Dogon	.121	X
Cranial #1	**Buriats**	.595	Andamanese	.594	Berg	.434	Arikara	.215	Peru	.134	G
Facial #4	**Buriats**	.135	Tasmanians	.125	Teita	.115	Moriori	.111	**Atayal**	.110	GR
Facial #5	**Buriats**	.242	Tasmanians	.166	**Guam**	.144	Zulu	.123	Australians	.110	X

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 17. Sixtieth percentile posterior probability (PP) results for female Thai crania using different measurement suites

Suite	Highest PP	PP	Second highest PP	PP	Third highest PP	PP	Fourth highest PP	PP	Fifth highest PP	PP	Supported expectation
All available	**Hawaiians**	.248	**Hainan**	.048	**Guam**	.035	**Buriats**	.030	**Atayal**	.014	GR
General #1	**Hawaiians**	.100	**Hainan**	.087	**Guam**	.056	**Buriats**	.048	**Atayal**	.022	GR
General #3	**Buriats**	.087	**Hawaiians**	.054	**Hainan**	.054	Arikara	.051	**Atayal**	.043	GR
General #6	**Buriats**	.102	Berg	.058	**Hainan**	.051	Peru	.046	**Atayal**	.039	GR
General #7	**Hawaiians**	.116	Zalavár	.054	Peru	.037	**Hainan**	.033	Norse	.031	GR
General #8	**Buriats**	.096	**Hawaiians**	.094	**Hainan**	.061	**Guam**	.046	South Japan	.016	GR
General #11	**Hainan**	.113	**Hawaiians**	.085	Andamanese	.074	South Japan	.039	Dogon	.032	G
General #13	**Hainan**	.107	**Hawaiians**	.106	Andamanese	.084	South Japan	.035	Dogon	.025	G

General #14	**Hainan**	.122	**Hawaiians**	.073	Andamanese	.050	North Japan	.037	**Atayal**	.033	G
General #15	**Hawaiians**	.104	**Hainan**	.095	**Atayal**	.041	Andamanese	.040	North Japan	.030	GR
General #16	Berg	.098	**Hainan**	.065	Andamanese	.045	**Buriats**	.044	Peru	.040	G
General #18	**Hawaiians**	.101	Berg	.066	Ainu	.062	Zalavár	.053	**Hainan**	.047	GR
General #19	**Buriats**	.085	Arikara	.081	Berg	.054	Dogon	.049	**Hainan**	.043	GR
Facial #2	**Guam**	.055	Ainu	.054	**Hawaiians**	.052	North Japan	.052	**Hainan**	.049	GR
Cranial #1	Berg	.210	Peru	.084	Andamanese	.080	**Buriats**	.053	**Hainan**	.051	G
Facial #4	**Hawaiians**	.054	North Japan	.051	Australians	.050	Ainu	.049	Teita	.045	R
Facial #5	**Hainan**	.053	Ainu	.053	South Japan	.052	North Japan	.051	Tolai	.051	GR

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 18. Ninetieth percentile typicality probability (TP) results for female Thai crania using different measurement suites

Suite	Highest TP	TP	Second highest TP	TP	Third highest TP	TP	Fourth highest TP	TP	Fifth highest TP	TP	Supported expectation
All available	**Hawaiians**	.148	**Hainan**	.127	**South Japan**	.117	**Atayal**	.067	**North Japan**	.057	GR
General #1	**Hainan**	.247	**Guam**	.141	**Hawaiians**	.126	**South Japan**	.121	**North Japan**	.058	GR
General #3	**South Japan**	.317	**Atayal**	.312	**Hainan**	.270	**Hawaiians**	.220	**North Japan**	.217	GR
General #6	Berg	.460	**Atayal**	.433	Peru	.292	**Hainan**	.291	**South Japan**	.231	GR
General #7	**Hainan**	.607	**South Japan**	.525	**Hawaiians**	.496	Peru	.448	Norse	.409	GR
General #8	**Guam**	.408	**Hainan**	.384	Arikara	.375	**Hawaiians**	.359	**North Japan**	.332	GR
General #11	Dogon	.608	**South Japan**	.600	**Hainan**	.578	**North Japan**	.545	**Atayal**	.466	GR
General #13	Dogon	.590	**South Japan**	.587	**Hainan**	.533	**North Japan**	.527	**Hawaiians**	.481	GR

General #14	**Atayal**	.635	**Hainan**	.569	Dogon	.564	**North Japan**	.524	**South Japan**	.483	GR
General #15	**Hainan**	.535	**North Japan**	.484	**Atayal**	.471	**Hawaiians**	.467	**South Japan**	.463	GR
General #16	Berg	.617	**South Japan**	.483	Dogon	.482	Buriats	.466	**North Japan**	.465	R
General #18	**Hainan**	.870	**South Japan**	.844	Egypt	.817	**North Japan**	.806	Dogon	.761	GR
General #19	Berg	.666	Tolai	.652	Buriats	.618	Arikara	.613	**Hainan**	.519	GR
Facial #2	Zalavár	.957	Berg	.938	**Hawaiians**	.919	Norse	.902	Ainu	.902	X
Cranial #1	Berg	.783	**Hainan**	.618	Santa Cruz	.577	Arikara	.567	**Hawaiians**	.553	GR
Facial #4	Zalavár	.938	Norse	.937	Egypt	.914	Berg	.911	**North Japan**	.876	X
Facial #5	Zalavár	.976	Berg	.966	Norse	.914	Tasmanians	.902	**Hainan**	.861	GR

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

Table 19. Seventieth percentile typicality probability (TP) results for female Thai crania using different measurement suites

Suite	Highest TP	TP	Second highest TP	TP	Third highest TP	TP	Fourth highest TP	TP	Fifth highest TP	TP	Supported expectation
All available	**Hawaiians**	.031	**Hainan**	.024	**Guam**	.022	**Atayal**	.014	**North Japan**	.009	GR
General #1	**Guam**	.049	**Hawaiians**	.036	**Hainan**	.032	**Atayal**	.026	Arikara	.020	GR
General #3	Arikara	.089	**Hainan**	.067	Berg	.058	**Atayal**	.055	**Guam**	.052	GR
General #6	Arikara	.064	Berg	.063	**Hainan**	.058	Peru	.048	Buriats	.048	GR
General #7	**Hawaiians**	.259	Zalavár	.190	Norse	.145	Berg	.125	**Hainan**	.124	GR
General #8	**Hainan**	.125	**Hawaiians**	.115	**North Japan**	.103	**Guam**	.092	South Japan	.071	GR
General #11	**Hainan**	.220	**Hawaiians**	.168	**Atayal**	.155	South Japan	.149	**North Japan**	.137	GR
General #13	**Hainan**	.175	**South Japan**	.107	**Hawaiians**	.102	Andamanese	.098	**Atayal**	.086	GR

General #14	**Hainan**	.261	**Hawaiians**	.220	Andamanese	.178	**North Japan**	.168	South Japan	.156	G
General #15	**Hainan**	.220	**Hawaiians**	.176	**Atayal**	.141	Andamanese	.103	**North Japan**	.099	GR
General #16	**Hainan**	.266	**Hawaiians**	.191	**North Japan**	.186	Berg	.185	**Atayal**	.175	R
General #18	**Hawaiians**	.542	Berg	.506	Egypt	.461	Zalavár	.456	Norse	.453	X
General #19	Buriats	.254	Arikara	.219	**Hawaiians**	.166	Tolai	.140	**Hainan**	.134	GR
Facial #2	Ainu	.635	Australians	.628	**Hainan**	.616	Tolai	.606	**Atayal**	.604	GR
Cranial #1	Berg	.328	Peru	.265	**Hainan**	.234	Arikara	.203	**Atayal**	.161	GR
Facial #4	**North Japan**	.636	Ainu	.609	**Hawaiians**	.587	**Hainan**	.587	South Japan	.559	GR
Facial #5	**Hawaiians**	.668	Ainu	.646	Tolai	.644	**Hainan**	.640	**North Japan**	.610	GR

N.B.: populations in bold face are the closest five based on the analysis of all available measurements.

In summary, percentile analysis confirms the similarity of Thais and Malays and the tropical East Asian Mongoloid status of both of them. The TP analysis moreover detected an affinity with Hainan Chinese which had not been evident from classification analysis.

Thai Comparisons: Various Measurement Suites

When all available measurements are utilized, the GR expectation is supported, regardless of whether classification results or percentile analysis is performed (first row in Tables 10 to 19). This result indicates that both race and geography are important for craniometric affinity, and that this correct finding is robust whatever analytical approach is employed.

The GR expectation is also supported whenever at least 16 measurements are included in the analysis (General #1 and General #2 suites in Tables 10 to 19). Therefore, 16 measurements would appear to be a sufficient number to make a correct finding (here, GR) probable. Note that percentile analysis appears superior to classifications in reproducing the finding that would be found utilizing all available measurements. For females, in terms of percentile analysis, with few exceptions the same five populations are the five closest to Thais for both the General #1 suite and all available measurements (Tables 16 to 18), whereas this is not the case with classifications (Table 11). For males, comparing the five closest populations for all available measurements and the General #2 suite (percentile analysis), there is at most a mismatch of one Howells population (Tables 12 to 15), but in the case of classifications there is a mismatch of two populations (Table 10).[3]

[3] Note that 16 is not a large enough number of measurements to fix a skull's classification. This study's original data include many cases where one suite of 16 or more measurements would strongly imply one classification, but an overlapping suite of 16 or more measurements would strongly imply a quite different classification. The point being made here is that the conclusion to be drawn from the skulls considered together is stabilized once a minimum of 16 measurements is used. For instance, adding one more measurement to the analysis may well make some skulls switch from a GR to an X classification, but they would be counterbalanced by other skulls switching from an X to a GR classification.

Percentile analysis also appears better than classifications in detecting a GR 'signature' when at least 16 measurements are available. For males, looking at classifications, Filipinos are the sole GR population to occur in the closest five to Thais, but looking at percentiles, Hainan join Filipinos in the sixtieth PP and both TP analyses (Tables 10 and 13 to 15). For females, both Hainan and Atayal are amongst the five populations closest to Thais in all but one of the percentile analyses, but in only one of the two classification analyses (Tables 11 and 16 to 19).

Percentile analysis also supports a GR expectation whenever at least 13 measurements are available whereas classifications analysis fails to. The measurement suites of relevance here are General #3 for females and General #4 for males. The GR expectation is supported in every percentile analysis (Tables 12 to 19) but with the classification results, the most strongly supported expectation is 'R' (Tables 10 and 11). R is not an incorrect expectation but it has less specificity than GR.

Results become less predictable once the number of utilized measurements falls below 13. For males, the R, G and X expectations are often as strongly supported as GR (Tables 10 and 12 to 14, rows headed General #5 to Facial #5). However, the seventieth percentile typicality probability analysis has the virtue that either the GR or R expectation is supported even when the number of measurements decreases to three (Table 15). For females, the General #6 suite (11 measurements) correctly supported the GR expectation regardless of the analytical approach, but otherwise the G expectation is the most often supported, at least for classifications and posterior probability analysis. Interestingly, however, typicality probability analysis supports the GR expectation in the great majority of cases even when the number of measurements is as few as three (Tables 11 and 16 to 19, rows headed General #7 to Facial #5).

A scan down the columns in Tables 10 to 19 reveals several interrelated trends associated with a decrease in the number of utilized measurements. First, there is a general decline in how often the five closest Howells populations are the same five as found through analysis of all available measurements. For instance, in the male classifications based on measurement suites Facial #1 (seven measurements) and Facial #5 (three measurements), the five closest Howells populations are completely different from the closest five based on all available measurements (Table 10). Secondly, classifications tend to become spread more evenly amongst the Howells populations. Compare the 27% of Filipino classifications for all available measurements with the 11% of Egyptian classifications for the Facial #5 suite (Table 10, male Howells populations with the most classifications), or the 39% of Hawaiian classifications for all available measurements with the 14% of Tasmanian classifications for the Facial #5 suite (Table 11, female Howells populations with the most classifications). Thirdly, as the number of measurements decreases, so the posterior probabilities of the closest Howells populations plummet while their typicality probabilities steeply rise (Tables 12 to 19). Further exploration of this point will be deferred to the Discussion. For the time being, we may observe that the generation of low typicality probabilities is frequently part and parcel of the process of using a sufficiently large battery of measurements to capture enough information to produce a useful diagnosis of craniometric affinity.

Table 20 presents the expectations with statistically significant support (see Table 6), ordered by the number of measurements used in the different analyses. The GR and R expectations are not mutually exclusive since the GR populations are also R populations. 'GR/R' is generally found to be statistically significant when Mongoloid populations including one or more located close to Thailand dominate the results in Tables 10 to 19; 'GR' when Mongoloid populations close to Thailand account for half or more of a strong representation by Mongoloid populations; and 'R' when Mongoloid populations dominate the results but those close to Thailand make up at most a minor component.

Table 20. Expectations with statistically significant support from the different analyses

Measurement suite	Classifications	90th PP	60th PP	90th TP	70th TP
All available ♀	GR/R	GR/R	GR/R	GR/R	GR/R
All available ♂	GR	GR/R	GR/R	GR/R	GR/R
General #1 ♀	GR/R	GR/R	GR/R	GR/R	GR/R
General #2 ♂	GR/R	GR/R	GR/R	GR/R	GR/R
General #3 ♀	GR/R	GR/R	GR/R	GR/R	GR
General #4 ♂	GR/R	GR/R	GR/R	GR/R	GR/R
General #5 ♂	—	—	—	—	—
General #6 ♀	GR/R	GR/R	GR	GR	R
General #7 ♀	R	R	—	R	—
General #8 ♀	—	R	GR/R	GR/R	GR/R
General #9 ♂	G	—	GR/R	—	R
General #10 ♂	—	—	GR	GR	R
General #11 ♂	GR	—	GR/R	GR/R	GR
General #11 ♀	G	—	—	GR	GR/R

General #12 ♂	—	—	—	—	—
General #13 ♀	G	—	—	R	—
General #14 ♂	—	—	GR	GR/R	GR/R
General #14 ♀	G	—	—	GR	—
General #15 ♀	G	—	—	GR/R	—
General #16 ♀	G	—	—	—	GR
Facial #1 ♂	—	—	—	—	R
General #17 ♂	—	R	GR/R	R	R
General #18 ♀	—	—	—	—	X
General #19 ♀	G	G	—	—	R
General #20 ♂	—	—	—	—	—
Facial #2 ♀	X	—	GR/R	—	GR
Cranial #1 ♀	G	—	—	GR/R	GR
Facial #3 ♂	GR/R	GR/R	GR	R	GR
Facial #4 ♀	—	—	—	X	GR/R
Facial #5 ♂	—	R	R	—	—
Facial #5 ♀	X	—	R	—	R

Considering first the analyses with at least 13 measurements, in the top six rows, we find that statistically significant support always emerged for a similarity between Thai crania and Mongoloid populations. The most frequently supported expectation was 'GR/R', which means that an affinity with Andamanese and other non-Mongoloid populations can be ruled out, but it cannot be decided that Thais are more similar to Mongoloid populations close to Thailand compared to Mongoloid populations in general (of which, Hawaiians and Arikara frequently featured as similar to Thais in Tables 10 to 19). Occasionally, however, statistically significant support emerged for a specific affinity with Mongoloid populations close to Thailand ('GR') as opposed to Mongoloid populations elsewhere in the world.

The General #5 suite (males), involving 12 measurements, did not produce statistically significant support for any expectation, regardless of the analytical method employed. Where to to 11 measurements were used (General #6 to General #11 rows), statistically significant support often emerged for the GR, R or GR/R expectation, but there were also two cases of statistically significant support for the G expectation. This last result is known to be wrong in the sense that it was never in contention when 13 or more measurements were used. The present results therefore suggest that with 12 or less measurements, there is no guarantee for finding statistically significant support for the 'correct' expectation, and there is even a risk of finding statistically significant support for an incorrect conclusion.

To rely on classifications appears to entail a risk of obtaining a spurious result when less than 13 measurements are used (see the second column in Table 20). Expectation G featured in over half of the cases wherever statistically significant support emerged for any expectation, and expectation X also featured a couple of times. However, the risk of obtaining statistically significant support for an incorrect expectation appears to be much less when percentile analysis is applied to the results. There was only one instance of statistically significant support for the G or X expectation in the 90th percentile posterior probability analysis, 90^{th} percentile typicality probability analysis and 70^{th} percentile typicality probability analysis, and no instances at all in the 60^{th} percentile posterior probability analysis. Even when as few as three measurements are used, percentile analysis evidently involves minimal risk of obtaining statistically significant support for an incorrect expectation. This is true even though there is considerable scope for obtaining support for an expectation that is less specific than desirable (the R rather than the GR/R expectation), or for not obtaining statistically significant support for any expectation.

Discussion

Two of the multiple hypotheses investigated in this study (Table 2) were clearly falsified in this study's application of the Fordisc 2.0 Howells craniometric module to Thai crania. One falsified hypothesis is the importance of geography (independent of race) for craniometric affinities, and the other is that neither race nor geography is important. Instead, the Mongoloid status of Thais is revealed through Fordisc 2.0 as long as 13 or more measurements are employed. In these cases, the importance of geography as well as race is evident in that statistically significant support regularly emerged for both the GR and R expectations or, occasionally, for the GR expectation in preference to the R expectation (Table 20).

With 12 or less measurements there is no reason to expect a correct result. When percentile analysis is used, statistically significant support for the GR or GR/R expectation may emerge, but it also may emerge for just the R expectation (showing that the geographical specificity potential of Fordisc 2.0 analysis has been lost), or there may be no statistical support for any expectation. However, percentile analysis appears to involve minimal risk of statistically significant support for an incorrect expectation, even when as few as three measurements are used. This is not the case when analysis is based on classifications, when there may be a greater probability of statistically significant support for an incorrect than for a correct expectation.

The results from the current analysis thus explain that the 'unsatisfactory' label Williams et al. (2005) placed on Fordisc 2.0 analysis is due to an unsatisfactory application of the technique in their study of ancient Nubian crania. The 11 measurements they used would be too few to expect Fordisc 2.0 analysis to be diagnostic. Moreover, by relying on Fordisc 2.0 classification results, rather than analyzing the PP or TP percentiles, they increased the likelihood of obtaining spurious results for their small measurement suite.

In fact, taking into account the number of populations in the Howells database, the study by Williams et al. would actually support the 'GR' and R' expectations. Allowing for 27 Howells populations in the analysis (i.e., the average of the male and female numbers of populations), the expected-to-observed ratio for Egyptian classifications across the entire Wilson confidence interval is 364% to 1,040%, a range that falls above the next highest expected-to-observed ratio (317%, for Caucasoids). The expected-to-observed ratio for Caucasoid classifications across the entire Wilson confidence interval is 211% to 405%, a range that falls well above the expected-to-observed ratio for non-Caucasoid classifications (62%). Therefore, despite the risky approach adopted by Williams et al., they obtained results contrary to their own 'X' conclusion. This point usefully indicates that the results from the present study are not particularly affected by the fact that Mongoloid populations make up over half of the Howells populations, and instead a GR/R expectation should be statistically supported whichever population is analyzed.

This study's results would not however support the view that craniometric analysis of a single skull would be likely to identify the Howells population that is the most closely related to the population of the analyzed skull. Overlap in craniometric variability between populations, even those that are only distantly related, is a fact. For instance, considering the analysis of Thai and Malay male skulls (all measurements available per specimen), we found that a Filipino classification accounted for a minority of specimens in both cases even though it was the single most common classification. We also found that around 20% (sexes combined) would have been classified to a wrong race, with every race represented in the classifications (Table 7). It is possible that those results could have been improved with the introduction of additional measurements, especially facial subtenses so as to take account of Mongoloid 'facial flatness'. However, overlap in craniometric variability between populations would seem to place a limit on how far craniometric classification of single skulls can be optimized. This is witnessed by the lack of any study that can demonstrate any result approaching perfect classification of every skull in a population through craniometrics. There may be occasions when an incorrect result is unlikely – for instance, a skull in southeast Australia in a circa 200 year old context should be correctly classified as either Aboriginal or European (leaving aside the possibility of mixed ancestry) because of the high proportion of Aboriginal skulls that would be classified as Australoid (Bulbeck et al. 2006) – but when populations are so distinct from each other, visual inspection by an expert would be just as efficacious as craniometric analysis.

As noted under Results, typicality probabilities appear to fall and posterior probabilities appear to rise as the number of analyzed measurements increases. Figure 2 illustrates how that would be the case if we conceptualize an analysis with 18 measurements as the product of six separate analyses of three measurements each (although this is obviously not quite how multivariate analysis works). Imagine we have six populations – G1, GR1, R1, R2, X1 and X2 – that are amongst the closest three populations to the target specimen on at least two of the suites of three measurements. Typicality probabilities tend to be high when only three measurements are used, so treat the typicality probabilities of the closest three populations per analysis as high and the typicality probabilities of the other three as medium. However, with so many medium to high typicality probabilities, discrimination between the populations on which is the closest to the target specimen is difficult, and so all posterior probabilities are either low or very low (Figure 2). When we start to link the small suites of measurements

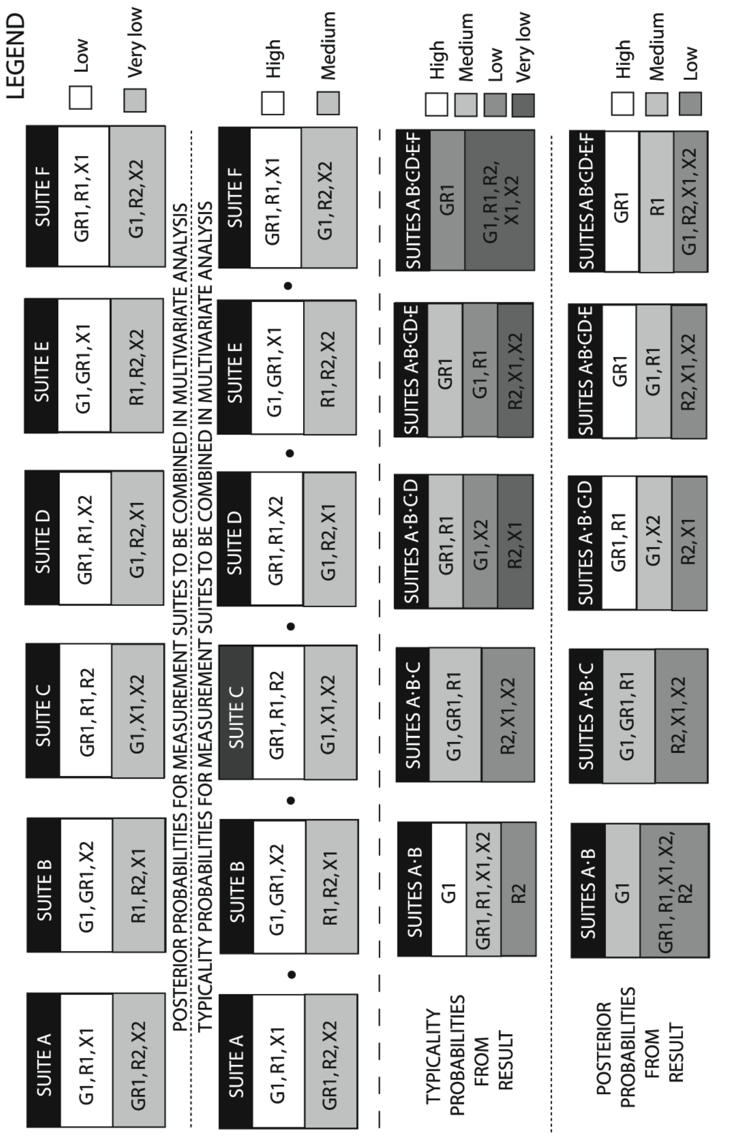

Figure 2. Model for changes to typicality and posterior probabilities with increased number of measurements

into larger suites, the typicality probabilities tend to decrease. This would be expected in the same way that the product of probabilities will result in a lower probability – for instance, an 80% probability for six independent events would be expected to result for every one of the events only 26% of the time. However, this decrease in the typicality probabilities is much more marked for the populations least often amongst the closest three to the target specimen than the populations which are most often amongst the closest three. With the resulting differentiation between populations in their typicality probabilities, a high posterior probability can be found for the closest population overall, compared to medium and low posterior probabilities for the other populations (Figure 2).

In our hypothetical example, the 'correct' result of the target skull's classification with the GR1 population first emerged with the analysis of 15 measurements, and was confirmed by the analysis with 18 measurements. This was so even though the GR1 population was not amongst the closest three populations on one of the six suites of three measurements, and even though the eventual typicality probability of the skull with respect to GR1 was low. Hence, even with this outcome of a correct result, it should be thought of as having been obtained through the elimination of less plausible classifications. It should not be thought of as the result dictated by how typical the skull's measurements are of the population with which it has been correctly classified.

Note that, had we stopped the analysis at 12 measurements ('Suites A·B·C·D' in Figure 2), reliance on classifications would have involved an arbitrary choice between GR1 and R1 as the closest population, whereas reference to the PP/TP results would correctly show that both GR1 and R1 are close on these 12 measurements. In addition, the 'correct' GR1 classification is to be expected for only a minority of the analyzed skulls, owing to the overlap in craniometric variability amongst the world's populations. Other skulls would be expected to be classified as G1, R1, R2, X1 or X2, albeit at lower frequencies than the

GR1 classification. However, with these 'incorrect' classifications, some other classification (e.g., GR1) would have been the preferred classification on a smaller suite of measurements, just as it took 15 measurements before GR1 emerged as the clearly best classification in Figure 2. The upshot of all this would appear to be that percentile analysis is a more reliable analytical method than classifications for inferring craniometric affinity, especially when relatively few measurements are used. Of course, percentile analysis would be nonsensical unless a reasonable sample of skulls from the same population were being analyzed, say a minimum of ten and preferably 30 or more. But if this paper's analysis points to any single conclusion, it would be that reliable craniometric classification need not be expected for single skulls, and the effective detection of craniometric affinities requires a population-based approach.

Conclusion

Taking a population-based approach to the analysis by Williams et al. (2005), we would conclude that their analysis supports the value of craniometric analysis using Fordisc 2.0. The proportion of ancient Nubian skulls classified as Egyptian is much higher than would be expected by chance, and the proportion classified as Caucasoid (the Egyptians' 'racial' group) is also much higher than expected. However, support of the kind to be found in the Williams et al. study (correctly interpreted) need not always be expected when repeating that study's test conditions, which limited the number of measurements to 11 and restricted analysis to the classification results. The same conditions repeated here found one instance where statistically significant support emerged for an incorrect conclusion, viz. an Andamanese affinity for Thai skulls using the classification results from the 11 measurements in the General #9 suite (Table 20). This instance of spurious support was neutralized through percentile analysis of the typicality probability and posterior probability data. In fact, there appears to be minimal risk of finding statistically significant support for an incorrect conclusion from percentile analysis of the typicality and posterior probability data generated by Fordisc 2.0, even when as few as three measurements are used. On the other hand, to have reasonable confidence in finding statistically significant support for the correct conclusion, a minimum of 13 measurements should be used. In that case the classification results should accord with the percentile analysis in pointing to the correct conclusion.

Populations overlap in their craniometric variability, and many skulls are not typical of their population. These points do not work against the validity of Fordisc 2.0 analysis as long as a population-based approach is taken and at least 13 measurements are used. In addition, percentile analysis of typicality and posterior probability data generated by Fordisc 2.0 should be used in preference to, or in combination with, the Fordisc 2.0 classification results.

Acknowledgments

The author thanks Helmut Loofs-Wissowa for providing me with copies of Sangvichien et al. (1969) and Saengvichien (1971), and Maneerat Jeerapaet for translating the title of Saengvichien (1971) into English.

Appendix:

Analyzed measurement sets

GOL = maximum glabella-occipital cranial length, XCB = maximum cranial breadth, BBH = basion-bregma cranial height, BNL = basion-nasion length, BPL = basion-prosthion length, MAB = external palate breadth, AUB = bi-auricular cranial breadth, UFHT = upper facial height, UFBR = upper facial breadth (across the anterior frontals), NLH = nasal height, NLB = nasal breadth, OBB = orbital breadth, OBH = orbital height, EKB = bi-orbital breadth (across the ectoconchia), DKB = interorbital breadth (between dacrya), FRC = frontal chord length, PAC = parietal chord length, OCC = occipital chord length, FOL = foramen magnum length, ZYB = bizygomatic breadth.

General #1 (females): GOL, XCB, BBH, BNL, BPL, MAB, AUB, UFHT, UFBR, NLH, NLB, OBB, OBH, EKB, DKB, FRC, PAC, OCC, FOL (19 variables).

General #2 (males): GOL, ZYB, BBH, BPL, MAB, AUB, UFHT, UFBR, NLH, NLB, OBB, EKB, DKB, FRC, PAC, OCC (16 variables).

General #3 (females): GOL, ZYB, BPL, MAB, AUB, UFHT, UFBR, NLH, NLB, EKB, DKB, FRC, FOL (13 variables).

General #4 (males): GOL, XCB, BNL, MAB, AUB, UFHT, UFBR, NLH, NLB, OBB, OBH, DKB, FOL (13 variables).

General #5 (males): XCB, BNL, BPL, MAB, UFHT, UFBR, NLH, NLB, OBB, OBH, EKB, DKB (12 variables).

General #6 (females): GOL, XCB, ZYB, MAB, AUB, UFBR, NLH, NLB, EKB, DKB, FRC (11 variables).

General #7 (females): XCB, ZYB, MAB, AUB, UFHT, UFBR, NLH, NLB, EKB, DKB, FOL (11 variables).

General #8 (females): GOL, XCB, MAB, UFHT, NLH, NLB, OBB, DKB, FRC, PAC, OCC (11 variables).

General #9 (males): GOL, BNL, BPL, MAB, UFHT, NLH, NLB, OBB, OBH, EKB, FRC (11 variables).

General #10 (males): GOL, XCB, ZYB, MAB, AUB, UFBR, NLB, OBB, OBH, EKB, DKB (11 variables).

General #11 (males and females): GOL, XCB, ZYB, BBH, BNL, BPL, MAB, UFHT, NLH, NLB (10 variables).

General # 12 (males): MAB, UFBR, NLH, NLB, OBB, OBH, EKB, DKB, FRC (9 variables).

General #13 (females): GOL, XCB, ZYB, BBH, BNL, BPL, UFHT, NLH (8 variables).

General #14 (males and females): GOL, XCB, ZYB, BBH, BNL, MAB, NLH, NLB (8 variables).

General #15 (females): GOL, XCB, ZYB, BBH, BNL, NLH, NLB (7 variables).

General #16 (females): GOL, XCB, ZYB, MAB, UFHT, NLH, NLB (7 variables).

Facial #1 (males): ZYB, MAB, UFHT, NLH, NLB, OBB, DKB (7 variables).

General #17 (males): GOL, XCB, UFHT, OBH, OBB, DKB (6 variables).

General #18 (females): XCB, ZYB, MAB, UFHT, NLB (5 variables).

General #19 (females): GOL, OBH, OBB, DKB (4 variables).

General #20 (males): BPL, MAB, UFHT, FRC (4 variables).

Facial #2 (females): MAB, UFHT, NLH, NLB (4 variables).

Cranial #1 (females): GOL, XCB, ZYB (3 variables).

Facial #3 (males): BNL, NLH, NLB (3 variables).

Facial #4 (females): UFHT, NLH, NLB (3 variables).

Facial #5 (males and females): MAB, NLH, NLB (3 variables).

References

Bulbeck, D., Rayner, D., Raghavan, P.
(2006). Races of *Homo sapiens*: if not in the southwest Pacific, then nowhere. *World Archaeology* 38: 109-132.

Cooper, Z.
(2002). *Archaeology and history: Early settlement in the Andaman Islands.* New Delhi: Oxford University Press.

Howells, W.W.
(1973). *Cranial variation in man.* Cambridge, Mass.: Peabody Museum.

Howells, W.W.
(1989). *Skull shapes and the map.* Cambridge, Mass.: Peabody Museum.

Hubbe, M., Neves, W.A.
(2007). On the misclassification of human crania. *Current Anthropology* 48: 285-286.

Jantz, R.L., Ousley, S.D.
(2003). *FORDISC 3.* https://web.dii.utk.edu/fordisc/ . Accessed May 28, 2009.

Martin, R. Saller, K.
(1957). *Lehrbuch der Anthropologie* (2^{nd} ed.). Volume I. Stuttgart: Gustav Fischer.

Ousley, S.D., Jantz, R.L.
(1996). *FORDISC 2.0: Personal computer forensic discriminant functions.* Knoxville: University of Tennessee.

Sangvichien, S., Sirigaroon, P., Jørgensen, J. B.
(1969). *Ban Kao Neolithic cemeteries in the Kanchanaburi Province. Part Two: The prehistoric Thai skeletons. Archaeological excavations in Thailand Volume III.* Copenhagen: Munksgaard.

Saengvichien, S.
(1971). *Thai skulls – a study in physical anthropology.* Ph.D. thesis. Bangkok: Mahidol University.

Tayles, N.G.
(1999). *The excavation of Khok Phanom Di, a prehistoric site in central Thailand. Volume V: The people.* London: The Society of Antiquaries.

Williams, F.L., Belcher, R.L., Armelagos, G.J.
(2005). Forensic misclassification of ancient Nubian crania: implications for assumptions about human variation. *Current Anthropology* 46: 340-346.

Wilson, E.B.
(1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 22: 209-212.