Europeans Have Larger Testes than Sub-Saharan Africans but Lower Testosterone Levels

Article *in* The Mankind Quarterly · March 2022

DOI: 10.46469/mq.2022.62.3.4

Emil O. W. Kirkegaard
Ulster Institute for Social Research

MANKIND QUARTERLY 2022 62:3 484-497

Europeans Have Larger Testes than Sub-Saharan Africans but Lower Testosterone Levels

Edward Dutton*
Asbiro Univerisity, Łódź, Poland

Helmuth Nyborg**
Aarhus University, Denmark (1968-2007, retired)

Emil Kirkegaard***
Ulster Institute for Social Research, London, UK

* Asbiro University, ul. Milionowa 12, 93-193 Łódź, Poland;
Email: e.c.dutton@dunelm.org.uk

** Email: helmuthnyborg@hotmail.com

*** Email: emil@emilkirkegaard.dk

Species and subspecies differ substantially in the size of their testicles. A study has found differences in average testis size when comparing Europeans and Northeast Asians. Other studies have found differences in testosterone levels between Blacks, Whites and Northeast Asians. We sought to replicate and extend these findings in a dataset of 4,462 US soldiers (the Vietnam Experience Study). We find that those of Sub-Saharan African descent (Blacks, n = 525) average smaller testes than those of European descent (Whites, n = 3,654), *d* = 0.24, while Blacks average higher testosterone levels, *d* = 0.11. These patterns are not explained by differences in age, weight, or height. We discuss these findings in terms of evolutionary theory of sperm competition and mating systems. These differences are thought to relate to life history speed differences, especially regarding mating patterns. A more polygynous mating pattern is associated with smaller testes and high testosterone, whereas a more competitive/free mating pattern is associated with larger testes, helping with semen production and thus sperm competition. These findings are consistent with how individual level differences in ejaculate quality relate to life history speed.

Keywords: Testicles; Testis; Race; Testosterone

Comparing various mammalian species, it has been found that large testicle size, in relation to body size, positively correlates with high copulatory frequency and with the extent to which females are likely to mate with multiple males during any given ovulatory cycle (Kenagy & Trombulak, 1986). Among primates, those whose females will normally only mate with one male during each cycle, such as gorillas, copulate less and have smaller testes in relation to their body size than chimpanzees whose females, living in larger groups, will mate with many males during their cycle (Harcourt et al., 1981). This has also been found when comparing other primate species (Kappeler, 1997). It is widely accepted that larger testes, in relation to body size, develop in order to allow the male to produce more semen, elevating his chances in a system of sperm competition (Schillaci, 2006). The relationship between testis size, sperm competition and evolution has been explored in a number of studies. For example, among the horned beetle, males with smaller horns will compensate for this by developing larger testes (Simmons & Emlen, 2006). A number of studies have found that ejaculate quality in human males is negatively associated with social assertiveness and psychoticism, both markers of high testosterone (see DeLecce et al., 2020).

In humans, larger testes correlate with male hormone levels and with semen quantity (Takihara et al., 1987). It has also been found that a European sample averaged larger testes than samples of Hong Kong Chinese (Diamond, 1986). Many studies have recorded that there are consistent differences in average testosterone levels when comparing Blacks with Whites (Ellis & Nyborg, 1992) and with East Asian males. It has been noted that European-descended males have, on average, higher free testosterone levels than Chinese males when controlling for age (Xu et al., 2014). A meta-analysis of this issue found that after adjustment for age, Black men have a modest but significantly higher free testosterone level than White men of between 2.5% and 4.9% (Richard et al., 2014). Studies that have not found this have often involved unrepresentative samples, such as university students (e.g., Maestripieri et al., 2014), which are problematic due to race differences in higher education participation in Western countries (Herrnstein & Murray, 1994). Studies in other mammals have noted that free testosterone strongly correlates with testicle size (Preston et al., 2011).

Thus, all else being equal, we would expect humans of Sub-Saharan African descent to have larger testes than humans of European descent. This would be in line with Europeans having larger testes than Northeast Asians, and Sub-

Saharan Africans being highest in free testosterone and Northeast Asians being lowest. Accordingly, we set out to investigate human sub-population differences in testosterone levels and in testis size and the relationship between the two variables.

Method

Data source

We used archival data from the Vietnam Experience Study (VES; https://www.cdc.gov/nceh/veterans/default1c.htm; http://ves.emilkirkegaard.dk/). The VES is a US military dataset based on a sample of 4,462 enlisted men (3,654 Whites, 525 Blacks, 200 Hispanics, 49 Amerindian/Native Americans, and 34 Asians). They were inducted in the military between 1965 and 1971, and a follow-up wave was conducted in 1985-1986. The purpose of the study was to examine the possible health effects of exposure to the Agent Orange chemical weapon during the Vietnam War. About 60% of the sample served in the Vietnam War while the rest were controls who served elsewhere (such as in Korea).

A very wide variety of health measures were taken during the second wave due to varied reports in the media about the health effects of the war. These measurements included psychiatric interviews (DIS), self-rating (MMPI), medical history, general physical examination, blood and urine analyses, sensory acuity, and intelligence (19 tests). For a subsample, there were tests of semen and some measures of numbers of offspring fathered by the men, the purpose of which was to look for fertility issues and possible birth defects in the children. The physical examination included measures of body size, including testis sizes (variables GP01N05 and GP01N06), and the blood analyses included measurement of the testosterone level (variable LD6861).

Measurements

The testes were in the beginning examined using a caliper to measure the long axis of both testes. However, after 1st January 1986, by which time approximately half of the veterans had been examined, the caliper was replaced by Prader's orchidometer (Prader, 1966) in order to improve precision. The medical examiner stood in front of the subject and palpated (examined by touch) one testis with one of his hands. At the same time, he went through a string of testis-shaped beads with his other hand until he found one which was the closest in size to the testis he was palpating. The Prader orchidometer consists of a string of 12 1-25 ml beads (shown in Figure 1). This was then repeated for the second testis.

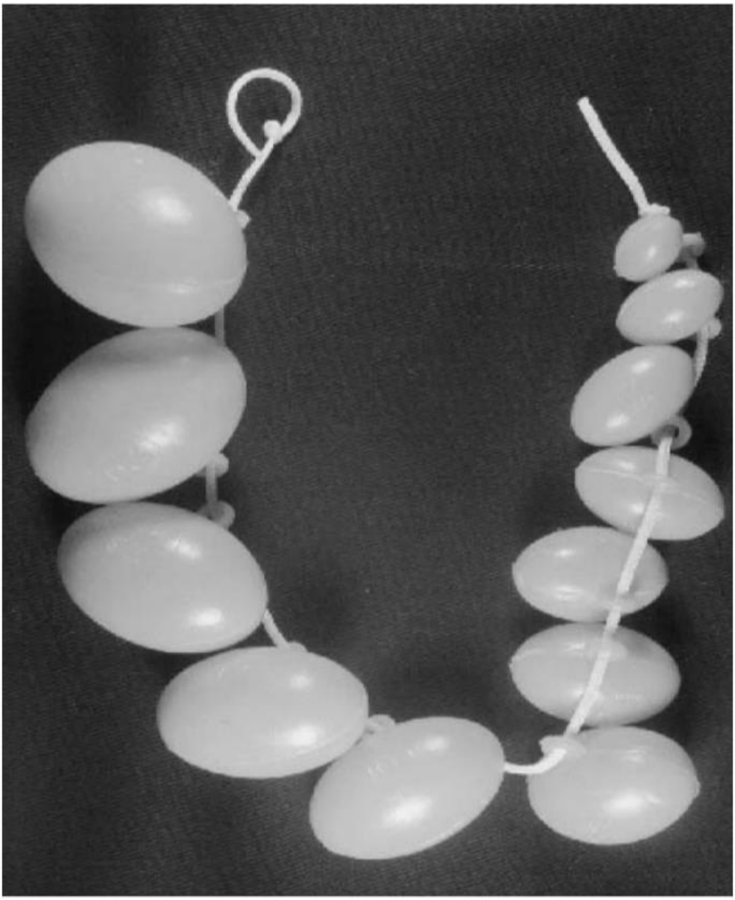

Figure 1. *The Prader orchidometer beads used to measure testis size (Karaman et al., 2005).*

We converted all data to Z scores to put them on the same scale without having to deal with the conversion between the units from the two measurement methods. In doing so, we found a small number of data errors, which we re-coded as missing data due to uncertainty about how to correct them. We verified that our data conversion worked as intended by comparing the correlation between the right and left testis. In the unadjusted data, these correlated about .80 and that was also the case in the corrected and merged data. The appendix provides more information about the re-coding. Although Prader's orchidometer method may seem bizarre, it has been found by comparison to ultrasound measurement to be quite precise (Mbaeri et al., 2013; Oehme et al., 2018; Sakamoto et al., 2008). Accordingly, it is accepted as a precise instrument and it cannot, thus, be averred that there may be subjective differences between the results obtained by different medical examiners.

Blood specimens were taken in the morning before breakfast, following an overnight fast commencing at 7pm. Plasma testosterone concentration was determined (in nanograms/deciliter, or ng/100ml) using a standard double antibody radioimmunoassay system (Leeco Diagnostics, Inc), and monitored with bench and blind repeat quality control procedures. The coefficient of variation for these repeats was between 4.49 and 9.72 with a median of about 5 (4 tests). This value is close and possibly superior (Owen et al., 2013) to values from modern equipment (Xu et al., 2018). The documentation for the measurement of testosterone can be found on page 158ff of the *Laboratory Methods and Quality Control* report (supplement A).

Results

Descriptive statistics for the main variables are shown in Table 1. Table 2 shows the correlations between the numerical variables.

The sample was on average borderline overweight (mean BMI = 26) with a fairly representative mean age of 38. The two testis measurements were very strongly correlated (r = .80), indicating reliability of the measurements. Average testis size was weakly related to height and weight (r = .10 and .09), and less so to BMI (r = .05). It was not related to testosterone level (r = .03, $p > .01$). Table 3 shows summary statistics for average testicle size and testosterone levels by race.

Table 1. *Descriptive statistics for the main variables; MAD = median absolute deviation, SD = standard deviation.*

	N	Mean	Median	SD	MAD	Min	Max	Skew	Kurtosis
Testis mean	4413	0.00	-0.36	1.00	0.92	-4.1	4.6	0.19	0.73
Testosterone (ng/dl)	4462	679.6	657	234.6	215.0	53	1950	0.82	1.49
Age (years)	4462	38.3	38	2.5	3.0	31	49	0.14	-0.01
Height (cm)	4462	176.2	176	6.75	5.9	150	202	-0.01	0.13
Weight (kg)	4461	83.6	82	15.4	13.3	47	190	1.22	3.77
BMI	4460	25.9	25.5	3.74	3.3	16.2	51.2	1.02	2.34

Table 2. *Correlations between numerical variables*; *Testost. = testosterone.*

	Testis right	Testis left	Testis avg.	Testost.	Age	Height	Weight
Testis left	0.80***	1					
Testis avg.	0.95***	0.95***	1				
Testost.	0.02	0.03	0.03	1			
Age	0.00	0.00	0.00	-0.19***	1		
Height	0.10***	0.09***	0.10***	-0.02	0.01	1	
Weight	0.09***	0.09***	0.09***	-0.32***	0.06***	0.43***	1
BMI	0.05***	0.05***	0.06***	-0.34***	0.06***	0.04*	0.81***

* p < .01, ** p < .005, *** p < .001

Table 3. *Summary statistics for testis size and testosterone level by race; SE = standard error, MAD = median absolute deviation, SD = standard deviation.*

	White	*Black*	*Hispanic*	*Asian*	*Native*
Testis size average					
N	3616	518	197	34	48
Mean	0.04	-0.21	-0.03	-0.75	0.11
SD	0.99	0.98	1.07	1.04	0.95
Median	-0.36	-0.36	-0.36	-0.67	-0.05
MAD	0.92	0.92	0.92	0.84	0.92
Min	-4.10	-4.10	-2.66	-2.75	-2.85
Max	4.62	3.37	4.62	2.13	1.51
SE of mean	0.02	0.04	0.08	0.18	0.14
Testosterone (ng/dL)					
N	3654	525	200	34	49
Mean	676.0	701.4	680.7	727.7	678.1
SD	230.5	248.5	260.0	255.1	256.3
Median	654.5	679	635.5	686	618
MAD	215.7	228.3	197.9	210.5	220.9
Min	53	122	83	362	175
Max	1915	1846	1950	1450	1405
SE of mean	3.8	10.8	18.4	43.8	36.6

The sample sizes for Asians and Native Americans are very small, and so of limited interest here. They are included for the possible use in future meta-analyses. The results show that Whites average slightly larger testes than Blacks and Hispanics. For testosterone, the relationship is reversed. Blacks have the highest level and Hispanics are again intermediate. However, the relationships are quite small and for the Hispanic comparisons, not statistically reliable. Thus, we focus only on the White-Black comparisons. Standardized effect sizes [with 95% confidence intervals] for testis size are: White-Black 0.25 *d* [0.16, 0.34], White-Hispanic 0.068 *d* [-0.075, 0.21]. For testosterone levels: White-Black -0.11 *d* [-0.20, -0.017], White-Hispanic -0.02 *d* [-0.16, 0.12]. Thus, the effect sizes can be described as small to medium in strength. Because testosterone is on a ratio scale, we can also compute the percentage differences, which are 3.9% for White-Black, and 2.9% for White-Hispanic comparisons, respectively. Our finding for the White-Black difference is closely in line with that of a 2014 meta-analysis which found Blacks had 2.5 to 4.9% higher levels (Richard et al., 2014).

The relationships found above, however, could be due to some confounding with age, height, and weight. To examine this, we used linear regression to include these as covariates. Results are shown in Table 4.

Table 4. *Regression results for testis size. Z = standardized to mean = 0, standard deviation = 1. Race reference is White.*

	Model 1			**Model 2**		
	B	SE	p	B	SE	p
Intercept	0.04	0.017	0.025	0.07	0.227	0.750
Race: Black	-0.25	0.047	<0.001	-0.24	0.046	<0.001
Race: Hispanic	-0.07	0.073	0.351	0.01	0.074	0.877
Race: Asian	-0.78	0.171	<0.001	-0.67	0.171	<0.001
Race: Native	0.07	0.145	0.606	0.04	0.145	0.787
Age				0.00	0.006	0.877
height_z				0.10	0.015	<0.001
BMI_z				0.05	0.015	<0.001
R^2 adj.	0.010			0.022		
N	4413			4411		

* p < .01, ** p < .005, *** p < .001.

The model results show that the Black-White testis size difference cannot be explained with age, height or BMI differences (0.25 *d* before, 0.24 *d* with covariates). Overall, however, testis size was very poorly predicted by the model, with 2.2% of the variance explained with the full model. Table 5 shows the comparable model for testosterone level.

Table 5. *Regression results for testosterone level. z = standardized to mean = 0, standard deviation = 1. Race reference is White*.

	Model 1			**Model 2**		
	B	SE	p	B	SE	p
Intercept	0.00	0.017	1	2.60	0.215	<0.001
Race: Black	0.11	0.047	0.02	0.11	0.044	<0.01
Race: Hispanic	0.02	0.074	0.78	0.09	0.070	0.22
Race: Asian	0.22	0.175	0.20	0.09	0.163	0.58
Race: Native	0.01	0.146	0.95	0.10	0.137	0.48
Age				-0.07	0.006	<0.001
height_z				-0.02	0.014	0.23
BMI_z				-0.33	0.014	<0.001
R^2 adj.		0.001			0.146	
N		4462			4460	

* $p < .01$, ** $p < .005$, *** $p < .001$.

As before, the Black-White gaps remain after adjustment for covariates (0.11 *d* before and after). We find here an inverse relation with age, a -0.07 standard deviation lower testosterone level for each year of age (standardized beta = -0.17). The effect size of BMI is surprisingly large at $\beta = -0.33$, and this is the chief reason this model explains variation much better than the testis model (14.6% explained vs. 2.2%).

Discussion

Our key finding is that Sub-Saharan Africans have slightly higher testosterone levels than Europeans together with smaller testes. This may potentially make sense in terms of differences in breeding patterns between these subgroups. Anatomically modern humans are descended from approximately twice as many females as males, with approximately 80% of females breeding but only around 40% of males doing so, these being the more dominant males selected by the females as part of polygamous unions (Baumeister, 2010; Seielstad et al., 1998; Wilder et al., 2004). There are, however, racial differences in the extent of this imbalance. X chromosome diversity has been shown to be lower in European and East Asian populations than in African populations (Arbiza et al., 2014). This would imply that polygamous mating patterns were more pronounced in Sub-Saharan Africa than they were in Europe or Northeast Asia. Outside of Africa, it appears, individual males monopolized the females to a lesser extent, engaged in fewer sexual partnerships and, in essence, moved closer to monogamy. Congruous with this difference, a large longitudinal study has found that Black males score much higher on socio-sexuality than White or East Asian males even within countries: 'Black men were generally more permissive than White, Hispanic, and Asian men' (Sprecher et al., 2013). This is one of many studies that have found higher levels of socio-sexuality when comparing Blacks to other groups (see Dutton, 2020, pp.125-128; Johnson et al., 1994; Kogan et al., 2015; Rushton, 1995).

Testosterone is associated with high levels of mating effort and thus with socio-sexuality (Edelstein et al., 2011). Accordingly, finding that Blacks have higher average testosterone levels than Whites would be congruous with evidence of higher levels of socio-sexuality among Blacks, as well as with evidence that Blacks are adapted to an ecology which is easier, in terms of basic needs being met, but also more unstable, resulting in greater effort being put into copulation as opposed to bonding and nurture. This direction of effort towards mating, rather than nurture, in an unstable ecology is known as a relatively fast life history strategy, and many studies have presented evidence for group differences in average life history speed (see Dutton, 2020, Ch. 6; Figueredo et al., 2021).

On the other hand, our finding that Blacks have smaller testes than Whites, with Whites also having larger testes than Northeast Asians, would appear to defy a simple explanation. One possibility is that Whites have larger testes than Northeast Asians simply due to being higher in socio-sexuality, meaning that females will be more likely to be unfaithful, resulting in the development of larger testes in order to triumph in sperm competition in a context of relatively large, cooperative social groups in which females are able to have contact with males other than their main sexual partner. However, it may be that among Sub-Saharan Africans, there was less of a pressure to develop large and cooperative social groups, due to the easier ecology, with such cooperation anyway less likely to be reciprocated due to the relative instability of the ecology. In these smaller social groups, there would be fewer males around to attempt to cuckold any given dominant male, with the result that sperm selection pressure would be weaker, leading to the development of smaller testes among Sub-Saharan Africans despite their having the highest testosterone levels. This is congruous with individual-level data which found that males who pursue a slower life history strategy concomitantly reflect the cuckoldry-risk hypothesis. They produce higher quality ejaculates, allowing them to triumph in the battle of sperm selection (Barbaro et al., 2019). Similarly, Blacks pursue a faster life history strategy than Whites but their smaller testicles imply lower ejaculate quality, with sperm quality and quantity robustly correlating (Liao et al., 2019).

Thus the harsh, stable, and highly competitive ecology of Northeast Asia would result in the necessity of high levels of nurture and of stable, strongly cooperative groups, reducing both testosterone levels and testicle size. However, the plentiful but unstable ecology of Africa would result in high mating effort but in only small and fission-prone groups, meaning relatively few males in any given group. This fission tendency in Sub-Saharan African groups has been widely documented (Norbeck, 1961, p.238; Turnbull, 1968, p.132). This tendency to not form large polities would reduce the intensity of sperm competition, reducing testis size. Therefore Europeans, between these two extremes, have the largest average testis size.

Limitations

One issue with these data is that testes shrink due to various conditions, specifically malnutrition, alcoholism, and chronic, terminal illness (Handelsman & Staraj, 1985). However, this would not appear to be relevant to our sample. Another potential issue is differences in the method by which testes are measured. However, an analysis of the five main methods has found that they are all roughly equally reliable, so this is just as reliable as any other method and there is no reason to be prejudiced against any of the procedures because it appears bizarre (Chipkevitch et al., 1996).

Finally, it might be averred that there might have been differences between the White and Black men in our sample, which could have influenced our findings: for example, differences in physical, physiological, reproductive and psychological health; in war experience and trauma; in exposure to chemical agents during the war; in drug use; in military rank; in personal background or in reproductive history. However, in that our findings are congruous with meta-analyses on race differences in androgens it seems most improbable that these confounds are germane. Moreover, with the exception of exposure to chemical agents, these could be regarded as reflections of testosterone as well as inducers or suppressors of it.

Conflict of Interest: We declare that there is no conflict of interest.

Data Availability: The data that support the findings of this study are available from the American Government but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the American Government.

References

Arbiza, L., Gottipati, S., Siepel, A. & Keinan, A. (2014). Contrasting X-linked and autosomal diversity across 14 human populations. *American Journal of Human Genetics* 6: 827-844. https://doi.org/10.1016/j.ajhg.2014.04.011

Barbaro, N., Shackelford, T.K., Jeffery, A.J., Lopes, G.S., Holub, A.M. & Zeigler-Hill, V. (2019). Life history correlates of human (*Homo sapiens*) ejaculate quality. *Journal of Comparative Psychology* 133: 294-300.

Baumeister, R. (2010). *Is there Anything Good about Men? How Cultures Flourish by Exploiting Men*. Oxford: Oxford Univ. Press.

Chipkevitch, E., Nishimura, R., Tu, D. & Galea-Rojas, M. (1996). Clinical measurement of testicular volume in adolescents: Comparison of the reliability of 5 methods. *Journal of Urology* 6: 2050-2053.

DeLecce, T., Lopes, G.S., Zeigler-Hill, V., Welling, L.M., Shackelford, T.K. & Abed, M.G. (2020). A preliminary but methodologically improved investigation of the relationships between major personality dimensions and human ejaculate quality. *Personality and Individual Differences* 153: 109614.

Diamond, J. (1986). Ethnic differences: Variation in human testis size. *Nature* 320: 488-489.

Dutton, E. (2020). *Making Sense of Race.* Whitefish, MT: Washington Summit.

Edelstein, R., Chopnik, W. & Kean, E. (2011). Sociosexuality moderates the association between testosterone and relationship status in men and women. *Hormones and Behavior* 60: 248-255.

Ellis, L. & Nyborg. H. (1992). Racial/ethnic variations in male testosterone levels: A probable contributor to group differences in health. *Steroids* 57: 72-75.

Figueredo, A.J., Hertler, S.C. & Peñaherrera-Aguirre, M. (2021). The biogeography of human diversity in life history strategy. *Evolutionary Behavioral Sciences* 15: 10-26.

Handelsman, D. & Staraj, S. (1985). Testicular size: The effects of aging, malnutrition, and illness. *Journal of Andrology* 6: 144-151.

Harcourt, A.H., Harvey, P.H., Larson, S.G. & Short, R.V. (1981). Testis weight, body weight and breeding system in primates. *Nature* 293: 55-57.

Herrnstein, R. & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Free Press.

Johnson, E.H., Jackson, L.A., Hinkle, Y., Gilbert, D., Hoopwood, T. et al. (1994). What is the significance of black-white differences in risky sexual behavior? *Journal of the National Medical Association* 86: 745-759. PMCID: PMC2607704.

Kappeler, P.M. (1997). Intrasexual selection and testis size in strepsirrhine primates. *Behavioral Ecology* 8: 10-19.

Karaman, M.I., Kaya, C., Caskurlu, T., Guney, S. & Ergenekon, E. (2005). Measurement of pediatric testicular volume with Prader orchidometer: Comparison of different hands. *Pediatric Surgery International* 21: 517-520.

Kenagy, G.J. & Trombulak, S. (1986). Size and function of mammalian testes in relation to body size. *Journal of Mammalogy* 67: 1-22.

Kogan, S.M., Yu, T., Allen, K.A., Pocock, A.M. & Brody, G.H. (2015). Pathways from racial discrimination to multiple sexual partners among male African American adolescents. *Psychology of Men and Masculinity* 16: 218-228.

Liao, W.B., Zhong, M.J. & Lüpold, S. (2019). Sperm quality and quantity evolve through different selective processes in the Phasianidae. *Scientific Reports* 9: 19278.

Maestripieri, D., Klimczuk, A., Traficonte, D. & Wilson, M.C. (2014). Ethnicity-related variation in sexual promiscuity, relationship status and testosterone levels in men. *Evolutionary Behavioral Sciences* 8: 96-108.

Mbaeri, T.U., Orakwe, J.C., Nwofor, A.M., Oranusi, K.C. & Mbonu, O.O. (2013). Accuracy of Prader orchidometer in measuring testicular volume. *Nigerian Journal of Clinical Practice* 16: 348-351.

Norbeck, E. (1961). *Religion in Primitive Society.* New York: Harper.

Oehme, N.H.B., Roelants, M., Bruserud, I.S., Eide, G.E., Bjerknes, R., Rosendahl, K. & Juliusson, P.B. (2018). Ultrasound-based measurements of testicular volume in 6- to 16-year-old boys—intra- and inter-observer agreement and comparison with Prader orchidometry. *Pediatric Radiology* 48: 1771-1778.

Owen, L.J., MacDonald, P.R. & Keevil, B.G. (2013). Is calibration the cause of variation in liquid chromatography tandem mass spectrometry testosterone measurement? *Annals of Clinical Biochemistry* 50: 368-370.

Prader, A. (1966). Testicular size: Assessment and clinical importance. *Triangle* 7: 240-243.

Preston, B., Stevenson, I., Lincoln, G., Montfort, S., Pilkington, J. & Wilson, K. (2012). Testes size, testosterone production and reproductive behaviour in a natural mammalian mating system. *Journal of Animal Ecology* 81: 296-305.

Richard, A., Rohrmann, S., Zhang, L., Eichholzer, M., Bazaria, S. et al. (2014). Racial variation in sex steroid hormone concentration in black and white men: A meta-analysis. *Andrology* 2: 428-435.

Rushton, J.P. (1995). *Race, Evolution and Behavior: A Life History Perspective.* New Brunswick, NJ: Transaction Publishing.

Sakamoto, H., Ogawa, Y. & Yoshida, H. (2008). Relationship between testicular volume and testicular function: Comparison of the Prader orchidometric and ultrasonographic measurements in patients with infertility. *Asian Journal of Andrology* 10: 319-324.

Schillaci, M. (2006). Sexual selection and the evolution of brain size in primates. *PLoS ONE* 1(1): e62. https://doi.org/10.1371/journal.pone.0000062

Seielstad, M.T., Minch, E. & Cavalli-Sforza, L.L. (1998). Genetic evidence for a higher female migration rate in humans. *Nature Genetics* 20: 278.

Simmons, L.W. & Emlen, D.J. (2006). Evolutionary trade-off between weapons and testes. *Proceedings of the National Academy of Sciences USA* 103: 16346-16351.

Sprecher, S., Treger, S. & Sakaluk, J. (2013). Premarital sexual standards and sociosexuality: Gender, ethnicity, and cohort differences. *Archives of Sexual Behavior* 42: 1395-1405.

Takihara, H., Cosentino, M.J., Sakatoku, J. & Cockett, A. (1987). Significance of testicular size measurement in andrology II: Correlation of testicular size with testicular function. *Journal of Urology* 137: 416-419.

Turnbull, C. (1968). The importance of flux in two hunting societies. In: R. Lee & I. DeVore (eds.), *Man the Hunter.* New Brunswick, NJ: Transaction Publishing.

Wilder, J.A., Mobasher, Z. & Hammer, M.F. (2004). Genetic evidence for unequal effective population sizes of human females and males. *Molecular Biology and Evolution* 21: 2047-2057.

Xu, L., Yeung, S.A., Kavikondala, S., Leung, G.M. & Schooling, C.M. (2014). Testosterone concentrations in young healthy US versus Chinese men. *American Journal of Human Biology* 26: 99-102.

Xu, X., Zhang, X., Zhong, Y., Saad, F., Perez-Patron, M. et al. (2018). Dynamic patterns of testosterone levels in individuals and risk of prostate cancer among hypogonadal men: A longitudinal study. *Journal of Urology* 199: 465-473.

Appendix: Testis data recoding

Some values are clearly wrong using the date breakpoint given, i.e., 1st January 1986. The documentation about the measurements can be found on page 187 and 268-271 of the medical examination book. Unfortunately, they don't go into more details about the methods. However, we note they have the same errors in their results, i.e., some data that supposedly concern sizes that are far beyond the possible values. There is at least one White subject with left testis size given as 20, which is an impossibly high value and which must instead indicate the volume. This subject was tested on 1985-12-17, i.e., before their breakpoint data. There is similarly one subject tested on 1985-01-22 with size 15, which must be volume 15, and one with size 10 on 1985-12-09 which must be volume 10. There are 4 people tested before June 1985, which seems to be before data collection started at all. They were likely tested in 1986 and we set this as the new test date. Other data errors include values that only occur once, such as one person whose value was given as 17. This might be size 7 (an unlikely value). To avoid introducing any more errors, these impossible values were recoded to missing (NA). For more detail, see the output of the R notebook. Figure A1 below shows the correlations between the left and right testis after data conversion.

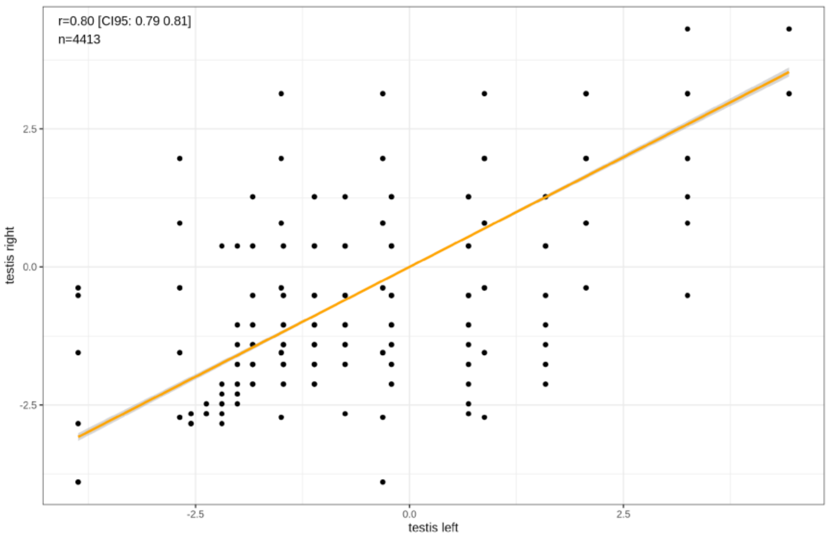

Figure A1. *Scatterplot of correlation between left and right testis after data correction and merging.*