Table S1. Cranial measurements used in the study.

Code[a]	Measurement
GOL	Glabella-occipital length
NOL	Nasio-occipital length
BNL	Basion-nasion length
BBH	Basion-bregma height
XCB	Maximum cranial breadth
XFB	Maximum frontal breadth
BPL	Basion-prosthion length
AUB	Biauricular breadth
NPH	Nasion-prosthion height
NLH	Nasal height
NLB	Nasal breadth
OBH	Orbital height
OBB	Orbit breadth
MAB	Palate breadth
ZMB	Bimaxillary breadth
EKB	Biorbital breadth
DKB	Interorbital breadth
FMB	Bifrontal breadth
NAS	Nasio-frontal subtense
FRC	Nasion-bregma chord
FRS	Nasion-bregma subtense
FRF	Nasion-subtense fraction
VRR	Vertex radius
NAR	Nasion radius
PRR	Prosthion radius
ZMR	Zygomaxillare radius
AVR	Molar alveolus radius

[a]"Code" is Howells' [1-3] three-letter measurement designation.

Table S2. Common chimpanzee sample.

Taxon	Sample size	Repository institutions[a]
Pan troglodytes schweinfurthii	80	AMNH, NHM, NMNH, RMCA
Pan troglodytes troglodytes	75	AMNH, HMCZ, NHM, NMNH, PCM, RMCA
Pan troglodytes verus	82	AMNH, HMCZ, HPM, MPI, NHM, NMNH, RMCA

[a]AMNH: American Museum of Natural History (New York); HMCZ: Harvard Museum of Comparative Zoology (Cambridge); HPM: Harvard Peabody Museum (Cambridge); MPI: Max Planck Institute for Evolutionary Anthropology (Leipzig); NHM: Natural History Museum (London); NMNH: National Museum of Natural History (Washington); PCM: Powell-Cotton Museum (Birchington); RMCA: Royal Museum of Central Africa (Terveuren).

Appendices

A1. Morphological split time estimator.

For two groups (e.g., populations, subspecies, species) that shared a common ancestral group, Weaver and colleagues [4] introduced an estimator of the time in generations when the ancestral group split into the two daughter groups

$$PT_{\mathrm{D}} = \frac{\left(x_1 - x_2\right)^2 + 2V_1 + 2V_2 - 4V_0}{4V_m}, \qquad (1.1)$$

where x_1 and x_2 are the group means for a single morphological measurement (or scores along a single eigenvector) for the two daughter groups; V_1 and V_2 are the additive genetic variances for the two daughter groups; V_0 is the additive genetic variance for the ancestral group; and V_m is the additive genetic variance introduced by mutation per zygote per generation into all of the groups. PT_{D} is based on T_{D}, which is a split time estimator for microsatellite (short tandem repeat) DNA markers introduced by Zhivotovsky [5].

Note that the additive genetic variances for the two daughter groups depend on the (narrow-sense) heritability, h^2, according to equations of the form

$$V = h^2\sigma^2 \qquad (1.2)$$

where σ^2 is the morphological variance for the group; and the mutational variance depends on the heritability and a mutation constant, m, according to the relationship

$$V_m = m(1-h^2)\sigma_P^2 \qquad (1.3)$$

where σ_P^2 is the pooled within-group morphological variance [4].

If the group means for the morphological measurement have diverged neutrally (by genetic drift and mutation alone) and all of the quantities in Eq. (1.1) are known, then PT_{D} will be expected to give the split time, even if the groups are not at mutation drift equilibrium (balance between the addition of variation by mutation and removal of variation by genetic drift) [4, 5]. To see why PT_{D} has this property, first, let t_g be the split time in generations, τ_{wd} be the mean coalescence time of pairs of alleles chosen from the same daughter group, and τ_{wa} be the mean coalescence time of pairs of alleles chosen from the ancestral group. Then, using results in Slatkin [6, 7] and Whitlock [8]

$$\mathrm{E}\left\{\left(x_1 - x_2\right)^2\right\} = 4\left(t_g + \tau_{wa} - \tau_{wd}\right)V_m, \qquad (1.4a)$$

$$\mathrm{E}\left\{2\left(V_1 + V_2\right)\right\} = 4\tau_{wd}V_m, \text{ and} \qquad (1.4b)$$

$$\mathrm{E}\left\{4V_0\right\} = 4\tau_{wa}V_m, \qquad (1.4c)$$

where E is the mathematical expectation (average of the distribution of possible evolutionary outcomes). Finally, combining Eqs. (1.1) and (1.4) gives

$$\mathrm{E}\{PT_{\mathrm{D}}\} = \frac{4\left(t_g + \tau_{wa} - \tau_{wd}\right)V_m + 4\tau_{wd}V_m - 4\tau_{wa}V_m}{4V_m} \tag{1.5a}$$

$$= t_g + \tau_{wa} - \tau_{wd} + \tau_{wd} - \tau_{wa} \tag{1.5b}$$

$$= t_g\,. \tag{1.5c}$$

These results assume the classical quantitative genetics model of heredity [9, 10]; that the genetic basis of a measurement is a large number of loci that contribute equally and additively (i.e., no interactions among them) to the value of the measurement.

A2. Required mutational variance.

For two groups at mutation drift equilibrium (balance between the addition of variation by mutation and removal of variation by genetic drift), their split time in years can be estimated with the relationship

$$t_y = \frac{\left(x_1 - x_2\right)^2 g}{4V_m}\,, \tag{2.1}$$

where t_y is the split time in years; x_1 and x_2 are the group means (for a single morphological measurement) for the two groups; g is the generation length; and V_m is the additive genetic variance introduced by mutation per zygote per generation into the groups [4, 11-13]. Note that Eq. (2.1) is equivalent to $t_y = g \times PT_{\mathrm{D}}$ (see Appendix A1), because the two groups are at mutation drift equilibrium [at mutation drift equilibrium $V_0 = (V_1 + V_2)/2$, so the variance terms in Eq. (1.1), Appendix A1, cancel out]. Solving for the mutational variance in Eq. (2.1) gives

$$V_m = \frac{(x_1 - x_2)^2 g}{4t_y}. \qquad (2.2)$$

Additionally, the expected additive genetic variance for each of the groups can be estimated with the relationship

$$V = 2N_e V_m (1 - F_{\mathrm{ST}}), \qquad (2.3)$$

where V is the additive genetic variance for each of the groups; N_e is the (long-term) effective population size of the two groups (considered together); V_m is the additive genetic variance introduced by mutation per zygote per generation into the groups; and F_{ST} measures the degree of departure of the two groups from a

single randomly-mating group [8]. Solving for the mutational variance in Eq. (2.3) gives

$$V_m = \frac{V}{2N_e\left(1-F_{\mathrm{ST}}\right)}, \qquad (2.4)$$

and replacing the within-group additive genetic variance in Eq. (2.4) with the (narrow-sense) heritability times the within-group morphological variance, $h^2\sigma^2$, gives

$$V_m = \frac{h^2\sigma^2}{2N_e\left(1-F_{\mathrm{ST}}\right)} \ . \qquad (2.5)$$

Eqs. (2.2) and (2.5) give the mutational variances that are consistent with the between-group [Eq. (2.2)] and the within-group [Eq. (2.5)] patterns of variation, if the morphological measurements are evolving neutrally (by genetic drift and mutation alone). Eqs. (2.1)-(2.5) assume the classical quantitative genetics model of heredity [9, 10]; that the genetic basis of a measurement is a large number of loci that contribute equally and additively (i.e., no interactions among them) to the value of the measurement.

A3. Numerical example of morphological split time calculations.

At mutation drift equilibrium, Eq. (1.1), Appendix A1, reduces to

$$PT_{\mathrm{D_{MDE}}} = \frac{\left(x_1 - x_2\right)^2}{4V_m}\,. \qquad (3.1)$$

Taking the 3rd eigenvector of the human within-group covariance matrix as an example, substituting the mean scores for Neandertals and modern humans along this eigenvector, $x_1 = -28.0761$ and $x_2 = -16.1341$ respectively, gives the numerator $\left(-28.0761 + 16.1341\right)^2 = 142.6114$. Using Eq. (1.3), Appendix A1, with the values for the mutation constant ($m = 1.20 \times 10^{-4}$), (narrow-sense) heritability ($h^2 = 0.37$), and the pooled within-group morphological variance along the 3rd eigenvector ($\sigma_P^2 = 34.1396$) gives the denominator $4 \times \left[1.20 \times 10^{-4}\right] \times \left[1 - 0.37\right] \times 34.1396 = 0.0103$. The quotient is $PT_{\mathrm{D_{MDE}}} = 13846$, which is an estimate of the split time in generations. Multiplying by generation length ($g = 25$) gives 346,150 years ago. Finally, adding 25,000 years to account for the fact (averaging dates) that Neandertals lived about 50,000 years ago gives ~370,000 years ago as the split time estimate based on the 3rd eigenvector, assuming mutation drift equilibrium, for Neandertals and modern humans.

If instead of mutation drift equilibrium [$V_0 = \left(V_1 + V_2\right)/2$] we assume no additive genetic variance in the last common ancestor [$V_0 = 0$] then, assuming that the within-group additive genetic variance in Neandertals is the same as in modern humans, we need to add $4h^2\sigma_P^2$ to the numerator, which gives

$(-28.0761-16.1341)^2+4\times0.37\times34.1396=193.1380$. Dividing by the denominator, multiplying by generation length, and adding 25,000 years, gives ~490,000 as the split time estimate based on the 3rd eigenvector, assuming $V_0=0$, for Neandertals and modern humans.

References

1. Howells W.W. 1973 Cranial variation in man: a study by multivariate analysis of patterns of difference among recent human populations. Cambridge, Harvard University; 259 p.

2. Howells W.W. 1989 Skull shapes and the map: craniometric analyses in the dispersion of modernHomo. Cambridge, Harvard University; 189 p.

3. Howells W.W. 1995 Who's who in skulls: ethnic identification of crania from measurements. Cambridge, Harvard University; 108 p.

4. Weaver T.D., Roseman C.C., Stringer C.B. 2008 Close correspondence between quantitative- and molecular-genetic divergence times for Neandertals and modern humans. Proc Natl Acad Sci U S A 105, 4645-4649.

5. Zhivotovsky L.A. 2001 Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. Mol Biol Evol 18(5), 700-709.

6. Slatkin M. 1993 Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47, 264-279.

7. Slatkin M. 1995 A measure of population subdivision based on microsatellite allele frequencies. Genetics 139, 457-462.

8. Whitlock M.C. 1999 Neutral additive genetic variance in a metapopulation. Genet Res 74, 215-221.

9. Fisher R.A. 1918 The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society of Edinburgh 52, 399-433.

10. Falconer D.S., Mackay T.F.C. 1996 Introduction to quantitative genetics. Harlow, Pearson.

11. Lande R. 1979 Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33(1), 402-416.

12. Lynch M., Hill W.G. 1986 Phenotypic evolution by neutral mutation. Evolution 40(5), 915-935.

13. Turelli M., Gillespie J.H., Lande R. 1988 Rate tests for selection on quantitative characters during macroevolution and microevolution. Evolution 42(5), 1085-1089.