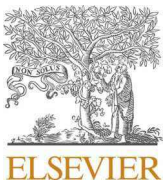
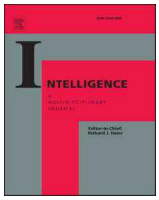
Intelligence 101 (2023) 101795



Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell



The ten-million-year explosion: Paleocognitive reconstructions of domain-general cognitive ability (*G*) in extinct primates

Mateo Peñaherrera-Aguirre [a,*], Matthew A. Sarraf [b], Michael A. Woodley of Menie [c], Geoffrey F. Miller [d]

[a] *University of Arizona, School of Animal and Comparative-Sciences Research, Tucson, AZ, USA*
[b] *Independent researcher, Boston, MA, USA*
[c] *Independent researcher, London, UK*
[d] *University of New Mexico, Department of Psychology, Albuquerque, NM, USA*

ARTICLE INFO

Keywords:
Big *G*
Paleocognitive reconstructions
Phylogenetic bracketing
Extinct primates
Extant primates
Punctuated dynamics

ABSTRACT

The correlation between primate "Big *G*" scores and brain volume in 68 *extant* species was employed to estimate probable *G* values for an additional 68 *extinct* and 1 *extant* species with endocranial volume data, employing phylogenetic bracketing. Three different methods were used to generate bracketed estimates, which all showed high convergence. The average of these *G* estimates (for the extinct primates) coupled with the values from the extant species were found to correlate strongly with neurocognitive measures of both extant and extinct primate taxa, specifically Transfer Index scores (an indicator of cognitive flexibility) and the neuroanatomical covariance ratio (a measure of neural integration). Ancestral character reconstruction incorporating *G* values was made possible with a phylogenetic tree containing data on the relationships among extant and extinct primates. Negative correlations were found between *G* and branch length, indicating that higher-*G* species do not persist as long as lower-*G* ones, consistent with the presence of the grey-ceiling effect (brain mass negatively predicts maximum population growth rate, and therefore a heightened vulnerability to extinction). Cladogenesis rates were also positively associated with *G*. Both associations were robust to models that controlled for false positive rates. Comparative models revealed that *G* evolved in extinct and extant primates in a punctuated pattern. The biggest increase in *G* occurred after the split between the members of the tribes *Hominini* and *Gorillini* 10 million years ago. Hence at the macroevolutionary scale, there can be said to have been a "ten-million-year explosion" in primate *G* leading up to modern humans.

1. Introduction

Across primate species, different measures of cognitive ability have been found to covary positively, indicating the presence of a common "Big *G*" (*G*) factor[1] (Burkart, Schubiger, & van Schaik, 2017a; Deaner, van Schaik, & Johnson, 2006; Fernandes, Woodley, & te Nijenhuis, 2014; Reader, Hager, & Laland, 2011). The meaning of this factor has proven controversial. Some researchers argue that it is simply the little "*g*" factor of intelligence—which captures individual-differences variance in general cognitive ability (or GCA)—scaled up to the level of species differences (Burkart et al., 2017a). Others argue that *g* and *G* result from different processes and so should not be conflated—perhaps *g* arises from variation in pleiotropic mutation load and other quality factors varying within populations, whereas *G* arises from differences in

* Corresponding author.
E-mail address: mpeaher@email.arizona.edu (M. Peñaherrera-Aguirre).

[1] Fernandes et al. (2014) introduced *G*, as distinct from g, into the comparative psychology literature to refer to general cognitive ability variance between species, where the "units of analysis are *aggregates* comprising multiple individual observations" (p. 315; emphasis added). By contrast, *g* represents "*individual differences*" in general cognitive ability (Fernandes et al., 2014, p. 315; emphasis added). To be more precise, as a statistical entity, *G* is an aggregation of within-species cognitive data from at least three species (Woodley of Menie & Peñaherrera-Aguirre, 2023), which can take the form of individual-differences data (such as from cognitive tests), but need not (for example, cognitive-behavioral count data from different species can be used to derive *G*; such counts reflect individual *observations* but do not necessarily provide individual *differences* data). It is important to understand that although data from at least three species are needed to derive a *G* factor, a given species can be assigned a *G* score, just as persons (or animals) each can be assigned a *g* score even though *g* cannot be estimated without data from at least three individuals.

https://doi.org/10.1016/j.intell.2023.101795
Received 22 May 2023; Received in revised form 25 September 2023; Accepted 26 September 2023
Available online 11 October 2023

species-level neuroanatomical constraints across populations (e.g., Arden & Zietsch, 2017; Arslan, von Borell, Ostner, & Penke, 2017; Lewis, Al-Shawaf, & Anderson, 2017). Recent evidence suggests that the first interpretation is correct – that Big *G* is just little *g* scaled up to the cross-species level – insofar as *G* and *g* estimates strongly correlate ($r = 0.918$) in mixed species and individual-differences data, but only when the latter is estimated with respect to subtests exhibiting sufficiently high coefficients of variance (indicating both high individual- and species-level variability[2] with floor or ceiling effects in cognitive task performance masking the underlying variation in general intelligence) (Woodley of Menie & Peñaherrera-Aguirre, 2023; see also Woodley of Menie, Fernandes, te Nijenhuis, Peñaherrera-Aguirre, & Figueredo, 2017; Woodley of Menie & Peñaherrera-Aguirre, 2022).

The availability of data on *G* for large numbers of primates (Reader et al., 2011) has enabled substantial analysis, allowing theories concerning the relationship between mean species differences and *G* loadings across abilities to be tested (Fernandes et al., 2014). Further, research in this area has established a basis for comparing the macroevolutionary modes of *G* to those for different neuroanatomical volume indicators (Fernandes, Peñaherrera-Aguirre, Woodley of Menie, & Figueredo, 2020) and the specific abilities constituting *G*, both in their residualized and unresidualized (for *G* variance) forms (Woodley of Menie, Peñaherrera-Aguirre, & Jurgensen, 2022). Building on this research program, comparative phylogenetic methods coupled with phylogenetic bracketing will be used here to estimate realistic *G*-factor scores for extinct primates. Fossil data on a correlate of *G* in living primates, specifically endocranial capacity (basically, brain volume), make this possible. After reviewing relevant literature, we describe these methods for estimating *G* factors. Then we compile a dataset containing *G*-factor scores for a large selection of both extant and extinct primates (including several members of the genus *Homo*), which allows us to conduct better tests of macroevolutionary models of the evolution of brain size and intelligence. We also use these new species-level measures of domain-general cognitive ability (*G*) to test other significant hypotheses, such as the "grey-ceiling" effect, which links increased brain mass to heightened vulnerability to extinction through the impact of reduced maximum population growth rate.

1.1. Previous macroevolutionary examinations of primate G

Deaner et al. (2006) conducted a meta-analysis of studies evaluating cognitive abilities in nonhuman primates. They used a Bayesian latent variable model and identified significant differences in cognitive performance across primate genera. A genus-by-genus comparison revealed the presence of a task-performance scale that consistently ranked genera in over 80% of cases. Deaner et al. concluded that Great Apes exhibited higher performance than other nonhuman primates. Reader et al. (2011) also explored the presence of (what is now called) *G* in 62 primate species. They reviewed the literature (over 4000 research articles) and estimated frequency counts for five cognitive indicators (tactical deception, extractive foraging, innovation, tool use, and social learning). Reader et al. restricted their analyses to publications featuring naturalistic observations in the field, without human intervention. Previous comparative studies with these cognitive indicators revealed high inter-rater reliability, indicating that the researchers did not introduce any (substantial) biases when coding the behaviors (Reader & Laland, 2002). A principal component analysis found that *G* explained 65% of the variance across behavioral indicators, with tactical deception exhibiting the smallest factor loading (0.74) and tool use the largest (0.88). According to the authors, *G* was positively and significantly correlated with neocortex ratio, log-transformed body mass, log-transformed brain volume, log-transformed neocortex volume, as well as the log-transformed brain volume residuals controlled for the effects of log-transformed body mass. Reader et al. (2011) also found a positive and sizable correlation between *G* and Deaner et al.'s (2006) measure of primate laboratory performance across multiple cognitive tasks ($r = 0.48$). A similar pattern emerged when exploring the association between *G* and Riddle & Corl's (1977) dataset on learning tasks ($\rho = 0.95$).

In a phylogenetic comparative investigation, Fernandes et al. (2014) gathered additional behavioral data to expand the dataset originally collected by Reader et al. (2011). Both unit-weighted and principal axis factor estimation revealed the existence of *G* across species of nonhuman primates (the factor structure remained relatively unaltered even after controlling for research effort aimed at different cognitive indicators). The latent *G* factor loaded positively and significantly onto the five cognitive abilities. The authors also identified a positive association between the *G* loadings and the between-species mean pair-wise difference across cognitive abilities and also with variance of cognitive indicators among species. According to Fernandes et al., score differences across primate taxa were more pronounced on more *G*-loaded indicators. The authors also reported a positive correlation between the rate of evolutionary change and the various *G* loadings, suggesting that selective pressures on *G* were especially pronounced during the macroevolution of primate cognition. Fernandes et al. (2020) compared the evolutionary lability (rate of change) of *G* relative to an assortment of neuroanatomical volume indicators (NVIs; brain volume, neocortex size, neocortex ratio, and cerebellum size; the authors used both absolute and body-size-residualized scores). Whereas brain size exhibited the lowest lability value (~ 0.00 standard deviations [SDs] per million years), *G* exhibited the highest evolutionary rate (~ 0.15 SDs per million years). Of the NVIs, only the absolute and residualized cerebellum estimates had noteworthy lability estimates (~0.10 SDs per million years). The authors also determined that most NVIs displayed greater phylogenetic inertia than *G*, indicating that similar NVI values are attributable to the species' shared macroevolutionary history (mean Pagel's λ across all NVIs = 0.85; Pagel's λ for *G* = 0.62).

1.2. Phylogenetic bracketing

Witmer (1995, 1997) proposed a clever method for reconstructing extinct phenotypes by integrating phylogenetic information. This procedure involves three steps: (1) estimating correlations between osteological and soft tissues (e.g. endocranial volumes and brain volume) based on data collected from extant taxa (living species); (2) using character optimization techniques that overlay the association between these variables onto a phylogenetic tree; (3) describing the relevant osteological attributes in extinct species based on fossil data (e.g. endocranial volumes). Witmer also proposed three possible levels of inference based on the presence of osteological and soft tissue correlations across an underlying phylogeny. The first, and strongest, level of inference involves reconstructing a fossil species' soft tissue (e.g. brain size) if the lineage possesses a specific osteological trait (e.g. a certain endocranial volume), and two or more extant outgroups (phylogenetically related to the fossil species) show both the osteological feature and the corresponding soft tissue. The second level of inference involves reconstructing a fossil species' soft tissue when only one of the extant outgroups possesses the osteological and soft tissue traits. The third, and weakest, level of inference involves reconstructing a fossil species' soft

[2] Burkart, Schubiger, and van Schaik (2017b) explicitly acknowledge that the work of Woodley of Menie et al. (2017a) "solved" the anomaly of an apparent discrepancy between *G* and *g*: "the anomaly of the lack of success of the mixed intraspecific and interspecific studies to generate a common main factor has been solved by Woodley of Menie, Fernandes, te Nijenhuis, Aguirre, & Figueredo. They suggested that variables with floor or ceiling effects may obscure differences in general intelligence across species because they cannot load on *g*. Their analysis supports this idea because species differences are especially striking for tests that load highly on *g*. Overall, then, the increasing plausibility of the idea that *g* and *G* can be equated automatically supports the argument that animals have something that closely resembles human *g*, and may even be homologous to it" (Burkart et al., 2017b, p. 53).

tissue even if none of the extant outgroups possesses the osteological attribute and its corresponding soft tissue. According to Ross, Lockwood, Fleagle, and Jungers (2002) all three levels of inference require the "causal association between soft tissue and osteological structure [to] be established in extant taxa (not necessarily the first and second outgroups of the fossil of interest)" (p. 25). The better our data on living species, the more confidently we can infer patterns of cognitive evolution in extinct species.

Although this *phylogenetic bracketing* approach was initially developed for reconstructing soft tissues, Ross et al. (2002) recommend extending this approach to behavioral and functional traits. Thus, phylogenetic bracketing might allow us to leverage data on *G* and endocranial volume in living primates, to reconstruct *G* in extinct primates—if their endocranial volumes are known from fossil skulls. These *paleocognitive* reconstructions of *G* could then be used to test various theories about the macroevolution of brain size and intelligence.

However, Nunn & van Schaik (2002) identified four major challenges for attempts to reconstruct the behavior of extinct primates using such methods. First, the apparent correlations between trait variables could be inflated due to shared macroevolutionary history among the species being examined. Without proper phylogenetic controls, for example, correlated residuals plus incorrect specification of degrees of freedom could increase the probability of Type I errors (false positives) (Nunn, 2011). Nunn and van Schaik recommend using phylogenetic comparative methods that control for these pseudoreplication problems. Second, hidden third variables (such as overall body size) might correlate with the examined traits, such that the apparent associations between phenotypic traits could disappear after controlling for a third variable (Nunn & van Schaik, 2002). Previous comparative studies have frequently controlled for estimated adult body mass, when analyzing the relationship between multiple neuroanatomical volume indicators and various cognitive abilities in nonhuman primates, and the overall evidence strongly suggests that uncorrected brain mass is a better predictor of cognitive phenotypes in nonhuman primates than body-mass-corrected brain mass (Deaner, Isler, Burkart, & van Schaik, 2007). Additionally relevant are many current neuroscientific studies across human participants, wherein *g* has been found to be associated with various non-neocortical regions. Goel and Dolan (2001) reported that logical reasoning was associated with activation within bilateral (Brodmann's Area; BA 19) and left (BAs 17 and 18) occipital regions. Similarly, Colom, Jung, and Haier (2006) identified significant correlations between *g* and neural activation of BAs 18 and 19 in the left occipital and BAs 17, 18, and 19 in the right occipital. Consequently, areas often associated with visual processing are also integral neuroanatomical correlates of *g*. In light of the current phylogenetic comparative and neurocognitive evidence, our paleocognitive reconstructions were not based on relative neuroanatomical measures such as the encephalization quotient or neuroanatomical estimates residualized for total brain size or body size.

Third, data on the trait of interest could be absent in the fossil record. Since the phylogenetic bracketing approach is correlational in nature, if any of the attributes are not described in a particular species, it is not feasible to proceed with the reconstruction. Hence, in the current study, extinct species were excluded if they lacked neuroanatomical data. The fourth and final problem concerns the estimation of adult body mass in the fossil record. Although Nunn & van Schaik, 2002 reconstructed multiple behavioral and socioecological traits in extinct primates based on adult body mass, as previously indicated, the current paper will not employ this variable as part of the analyses, thus circumventing this last problem.

1.3. Prior paleocognitive reconstructions of extinct primate abilities

Comparative and developmental researchers have proposed several methods for assessing the cognitive complexity of extant and extinct primates (Parker & Mckinney, 1999; Tomasello, 2010; Wynn, 1989, 2002). Employing a different approach, Beran, Gibson, and Rumbaugh (1999) reconstructed the cognitive ability of extinct hominins by examining the association between cranial capacity (estimated overall brain volume) and learning in extant nonhuman primates. Rumbaugh's (1970) experimental research on learning reversal inspired the authors to develop a Transfer Index (TI), a measure of cognitive flexibility wherein an individual must first reach a learning criterion (such as associating a cue with a reward) and is then tested under conditions where the initial baited stimuli are no longer rewarded (Rumbaugh, 1970). Basically, this procedure involves an inhibitory control task based on control of prepotent responses (although Beran et al., 1999 did not provide such a description in their publication) to assess cognitive flexibility.

Beran and collaborators presented 116 individuals from 12 nonhuman primate genera with a task-reversal procedure. Then they estimated a TI score based on subject performance. Beran and colleagues identified strong and positive associations between these TI scores and genus-level cranial capacity (*Spearman* $\rho = 0.83$) and body size (*Spearman* $\rho = 0.87$). Given these sizable correlations, the authors estimated the corresponding TI for ten extinct hominins: *Australopithecus afarensis*, *A. africanus*, *Paranthropus aethiopicus*, *P. boisei*, *P. robustus*, *Homo habilis*, *H. rudolfensis*, *H. erectus*, and *H. neanderthalensis*, using the following equation:

$$\mathrm{TI} = \left(0.05^{*}\text{cranial capacity in cm}^{3}\right)\text{–}7.2$$

TI scores, as a percent change, ranged from less than −20% to >60%. Their reconstruction indicated that the TI increased continuously, yet not necessarily evenly, during the evolution of extinct hominins. According to Beran and colleagues the TI scores for the genus *Australopithecus* and *Paranthropus* surpassed those of most extant primates in their dataset, but were similar to those of great apes (*Pan* sp., *Gorilla* sp., and *Pongo* sp.). Beran et al. also noticed that species within the genus *Paranthropus* had slightly higher estimated TI scores compared to the more gracile *Australopithecines*. Although the reconstruction suggested that *H. habilis* outscores earlier hominins, a considerable rise in TI scores occurred with the evolution of *H. rudolfensis*. The authors also identified two subsequent increases in TI, the first occurred in *H. erectus*, and the second with the evolution of *H. neanderthalensis* and extant *H. sapiens*. As *H. neanderthalensis* had a cranial capacity larger than *H. sapiens*, the authors concluded that Neanderthals could have had similar, or even higher, TI scores than *H. sapiens*.

Beran et al.'s (1999) method for estimating TI is similar to the phylogenetic bracketing approach used in this paper for estimating *G*, since it uses cranial capacity to infer the level of cognitive flexibility among extinct primates, based on the association between the two in living primates. A major drawback of the Beran et al. (1999) method is that it employed a single indicator of cognitive ability obtained under laboratory conditions for a limited number of primates aggregated at the level of genera ($n = 12$). Also, they seem not to have controlled for phylogeny in modeling these associations (an important potential limitation—see Nunn & van Schaik, 2002). The paleocognitive reconstructions of general intelligence in our study overcome these limitations by using (1) species-level data on a much larger sample of extant primates, (2) a multivariate and high-resolution cognition measure (*G*), the component data of which were obtained under naturalistic conditions, and (3) comparative phylogenetic methods and phylogenetic bracketing. However, the results of these earlier efforts can be used to establish the potential convergent validity of the *G* values estimated in the current study.

1.4. The grey-ceiling effect and extinction risk

Another potential use for the reconstructed *G* values estimated in this paper is in testing for the so-called grey-ceiling effect. This effect is based on the *Expensive Brain* hypothesis (Isler & van Schaik, 2009), which

maintains that the bioenergetic cost associated with increased brain mass requires either increased overall energy acquisition, or reduced allocation of energy to other domains (e.g., development and reproduction). Given these energetic trade-offs, larger-brained species generally reach various developmental milestones later and show lower fertility rates (Isler & van Schaik, 2009). Even though longer reproductive lifespan could compensate for lower fertility, prolonged interbirth interval and delayed maturation rates might also increase vulnerability to adverse environmental factors (e.g. famine, disease) and demographic factors (e.g. increased social competition under higher population densities), due to the bioenergetic costs of greater brain mass.

Isler & van Schaik (2009) collected brain mass, and other life history data for 536 species of eutherian (placental) mammals and estimated their corresponding maximum rate of population increase (r_{max}), based on the species' maximum reproductive lifespan and the annual young produced per female. These researchers argued that, since r_{max} predicts a species' capacity to bounce back from population collapses, it can be used as an index of (reduced) extinction risk. They found that brain mass and r_{max} were negatively related, suggesting that certain lineages are apparently constrained in terms of their maximum brain mass (i.e., they encounter a greyceiling), and that the evolutionary growth of brain mass beyond this neuroanatomical threshold increases a species' risk of extinction. Basically, selection for intelligence might favor larger brains most of the time (at least in some mammal taxa), but rare catastrophes might favor higher maximum reproductive rates sometimes (or else a lineage goes extinct), leading to a dynamic evolutionary balance between costly, slow-growing brains and cheaper, fast-growing brains. Thus, the greyceiling represents the increased extinction risk that might be imposed by extra grey matter, given that larger, costlier, slower-growing brains reduce the ability of species to bounce back from catastrophes.

Isler & van Schaik (2012) also reconstructed the length of the interbirth interval and the minimum population doubling time (DT_{min} = ln (2)/ r_{max}) in a sample of extinct hominins. The authors' estimates varied depending on whether each species was presumed to feature cooperative breeding (such as pair bonding or alloparenting). They argued that the evolution of cooperative breeding in early species of *Homo* allowed them to circumvent the macroevolutionary limitations of the grey ceiling. However, as indicated by the high extinction rate of hominins, including the genus *Homo* (except for contemporary *Homo sapiens*), it is feasible that paleoecological stochastic dynamics between the Late Miocene to the Late Pleistocene partially eroded the protective effects of cooperative breeding that initially facilitated the emergence and evolutionary persistence of large-brained taxa. Further studies should explore the macroevolutionary factors beyond cooperative breeding that allowed *Homo sapiens* to temporarily avoid extinction, given its sizable endocranial volume and high *G* score.

We will conduct a more direct test of the relationship between cognition and extinction vulnerability by analyzing the association between *G* and phylogenetic branch length (which measures a species' actual persistence in phylogenetic time), as opposed to estimated life-history characteristics such as r_{max}. We expect that, if the greyceiling is really a problem, there should be negative associations between *G* and branch length, indicating lower persistence (and heightened extinction vulnerability) for higher-*G* species.

2. Methods

2.1. Sample

G-factor-score data for nonhuman primates were sourced from Reader et al.'s (2011) meta-analysis (covering >4000 publications and encompassing 69 extant species of nonhuman primates) with supplemental data sourced from Fernandes et al. (2020). These databases contain ethological counts for five broad cognitive abilities including: (1) *Tool use*: Producing and using material artifacts to solve physical and social problems; (2) *Extractive foraging*: Obtaining food items that are enclosed or hidden (such as nuts, underground tubers, or bone marrow); (3) *Innovation*: Developing novel solutions to sophisticated, and often new, social or ecological challenges; (4) *Social learning*: Acquiring information and skills from conspecifics including parents, peers, and other group members; and (5) *Tactical deception*: Influencing the attention of conspecifics, gaining an advantage or imposing a cost by misleading them. Given that some primate species have been studied more extensively than others, it was essential to control for research effort (operationalized as the number of publications per species across the same scientific sources from which the behavioral counts were collected—these data were also available in Reader et al., 2011). Thus, the residuals were extracted from general linear models examining the influence of research effort on the various ethological counts. *G* values were computed using a unit-weighted factor (UWF) estimation procedure on the residualized behavioral counts. The UWF estimates are calculated by first standardizing the pertinent indicators and then computing an average across standardized values (Gorsuch, 2014).

Data on extinct primate endocranial volumes and the corresponding phylogeny for extant and extinct primate lineages were collected from Melchionna et al. (2020), who generated these by combining both morphological and (in the case of extant primates) genetic data. In total, usable *G* and cranial capacity data were available for 69 extant species (in the case of *Homo sapiens*, *G* was imputed based on brain size), and data on cranial capacity were available for 68 extinct species, from Melchionna et al.'s phylogeny. Neocortex volume (mm^3), brain mass (g), and body mass (g) were also collected from Lindenfors, Wartel, and Lind's (2021) database. Encephalization quotient (*EQ*) values were generated by first estimating a general linear model with Log body mass as a predictor of Log endocranial volume and extracting the corresponding intercept (an exponential transformation was used to obtain the raw intercept) and slope from a general linear model. Hence *EQ* = endocranial volume (cm^3) / (0.127* Body mass (g) $^{0.729}$). These data are further described in Appendix A. Previously, Sansalone et al. (2023) found that extinct and extant hominins show higher levels of neuroanatomical integration than other primates, operationalized as a covariance ratio (the covariation between neuroanatomical regions divided by the covariation within neuroanatomical regions, with values closer to 1.00 suggesting high covariation). So, we developed an additional phylogenetic comparative model by examining the association between this neuroanatomical covariance ratio (obtained from Sansalone et al., 2023) and *G* scores. Clearly, there are some limitations in using aggregate endocranial volume for these analyses. For example, measures of cranial capacity show considerable variability due to individual differences within species and measurement error by researchers (De Miguel & Henneberg, 2001). Thus, the cranial values of extinct primates often overlap across taxa due to these wide variances. This variability has been attributed to geographical and sex differences and ecological fluctuations associated with climate change (Ruff, Trinkaus, & Holliday, 1997; Stibel, 2023). Additional paleocognitive reconstructions are required to determine the relation between endocranial variability and *G* in extinct primate species due to these morphological and paleoecological factors.

2.2. Phylogenetic bracketing

A Reduced Major Axis model (RMA) was used to examine the association between brain volume and *G* scores in extant nonhuman primates. The predicted *G* values for extinct species were estimated using the RMA equation. According to Dunbar & Shultz (2023), RMA, instead of Ordinary Least Squares models, is recommend when reconstructing the phenotypes of extinct species because RMA accounts for the underlying measurement error. RMA models were computed with the R package *lmodel2* (Legendre & Oksanen, 2018). However, because RMA does not take phylogeny into account, it suffers from potential

pseudoreplication problems (i.e. correlations among the residuals arising from shared macroevolutionary history). Hence traditional statistical procedures (such as RMA) are more susceptible to type I errors, overestimating significant and sizable values. To address these limitations, we also evaluated the statistical relationship between brain volume and *G* scores while controlling for the underlying phylogenetic structure. In addition to using a Phylogenetic Generalized Least Squares (PGLS) model, we also developed a Phylogenetic Reduced Major Axis model (PRMA). This follows the recommendation by Schultz and Dunbar (2023) that RMA should take phylogenetic history into account. Thus, we used these PGLS and PRMA equations to estimate the extinct primate species' *G* values. The three separate scores (based on RMA, PGLS, and PRMA systems of equations) were also averaged to yield an inter-method mean score. All phylogenetic comparative models were computed using the R packages *caper* and *phytools* (Orme et al., 2013; Revell, 2012) in R version 4.0.1.

2.3. Macroevolutionary model comparison

Previous publications have estimated several macroevolutionary models to try to understand the selection regimes that drove the evolution of *G* and associated traits (Fernandes et al., 2014; Fernandes et al., 2020; Woodley of Menie et al., 2022). But these studies were restricted to data collected from extant nonhuman primates. Our study revisited these macroevolutionary models based on the combined extant and extinct *G* estimates. Eight macroevolutionary models were compared (using the *Geiger v2.0* package in R v.4.0.1; Pennell et al., 2014):

i) *Brownian Motion*: Under Brownian Motion the temporal extent of common ancestry, operationalized as branch length in the phylogenetic tree, moderates the likelihood of trait change, with older lineages being more prone to phenotypic shifts (Fernandes et al., 2020).
ii) *Ornstein-Uhlenbeck*: Although trait variation can occur at random, it is either positive or negative, which means that globally this alteration is oriented toward a central estimate (Nunn, 2011). It is not uncommon for researchers to argue that this model therefore provides an indication of stabilizing and adaptation-optimizing selection (e.g., Butler & King, 2004).
iii) *Lambda*: This model considers the degree to which phylogenetic history influences the magnitude of the covariation between traits. This model adjusts the tree according to its corresponding phylogenetic signal—a statistical parameter of trait conservation at the macroevolutionary level.
iv) *Early Burst*: This model considers whether the evolutionary rate increases or decreases exponentially across time. If the parameter estimate is close to zero then the selection regime is analogous to Brownian Motion; in contrast, if this value is smaller than zero, then the model evidences a process of rapid niche-filling followed by considerable reduction as ecological niches become saturated.
v) *Kappa*: Corresponds to a cladogenic (species-forming) mode of macroevolution wherein trait shifts occur as a function of the frequency of speciation events. This model can be taken to indicate the action of *punctuated* evolution, where rapid (as opposed to gradual) cladogenesis occurs when the Kappa value is <0.5 (Pagel, 1999).
vi) *Mean Trend*: An evolutionary model based on a trend element or directional drift, for example increasing values over time. This model is influenced by non-ultrametric phylogenies.
vii) *Rate Trend*: Corresponds to an evolutionary rate that is treated as a linear trend assuming a variable rate diffusion model.
viii) *White Noise*: This model presupposes that the data represent a normal distribution lacking any covariance among trait values across species due to shared phylogeny. This model can be taken to indicate the presence of true noise (or error) in the data, or it can be taken to reflect the action of non-phylogenetic (e.g., ontogenetic) transmission pathways conditioning the emergence and maintenance of traits (Pennell et al., 2014).

2.4. Branch length and G-score reconstruction

Given the *grey-ceiling hypothesis*, we predicted significant negative associations between the reconstructed *G* score, and the species' branch lengths extracted from a dated and non-ultrametric phylogeny (i.e. one in which the branches' lengths are not equidistant from the root). Branch lengths were calculated as the amount of time that passed since the taxa speciated until it either became extinct (e.g., in the case of fossil species), or persisted to the present (as extant lineages). Branch length is potentially a more direct metric for estimating a species' retrospective vulnerability to extinction than r_{max}.

2.5. Ancestral character reconstruction

We also generated an ancestral character reconstruction based on the phylogenetic tree containing both extant and extinct primate species (from Melchionna et al., 2020). This uses maximum likelihood estimation to determine the ancestral values associated with the various nodes in the phylogeny, then superimposes a phenotypic heat map onto the corresponding phylogeny, with colder colors (e.g., blue, purple) indicating lower *G* scores and warmer ones (e.g., red, orange) representing higher *G* scores. This visualization was generated using the *contMap* function, for continuous traits, found in the *phytools* package (Revell, 2012).

2.6. Phylogenetic convergent validity tests

To evaluate the convergent validity of the reconstructed *G* scores, we generated several different Phylogenetic Generalized Least Squares models to compare the standardized *G* scores to the TI values estimated for each species using the formula in Beran et al. (1999). Note that both the phylogenetic bracketing approach used here and Beran and colleagues' TI estimates use endocranial volume as the basis for inferring "missing" cognitively relevant information in extinct taxa, so there is a potential problem of pseudoindependence between the two measures. Despite this, we expected that the two sets of values will still show some degree of independence, as they reflect somewhat different sets of cognitive abilities (*G* in the case of the current study, and learning flexibility in the case of Beran et al., 1999). A second set of phylogenetic analyses were also conducted examining the association between *G* scores and the neuroanatomical covariance ratio, a measure of neuroanatomical integration. As *g* has been found to be distributed across the neural cortex, we predicted a positive association between *G* and the neuroanatomical covariance ratio.

3. Results

A basic RMA model revealed that brain volume was a positive and significant predictor of *G* ($b = 0.009$, $\beta = 0.982$, $p = .01$). The more phylogenetically sophisticated PGLS model also showed a positive association between these variables ($b = 0.005$, $\beta = 0.553$, $p < .0001$). This pattern persisted when computing a PRMA model using brain volume to predict *G* ($b = 0.028$, $p < .0001$; $N = 68$ extant nonhuman primate species for each model).[3] The results of this analysis are graphed in Fig. 1. A PGLS model revealed a significant positive association between log-transformed endocranial volume and *G* in extant species of primates. By contrast, neither the log-transformed brain mass residuals (adjusted for the influence of log-transformed body mass) nor the encephalization quotient significantly predicted *G*. The models also indicated that even though the raw and log-transformed neocortex volumes positively and

[3] It was not possible to estimate a corresponding β value for the PRMA model.

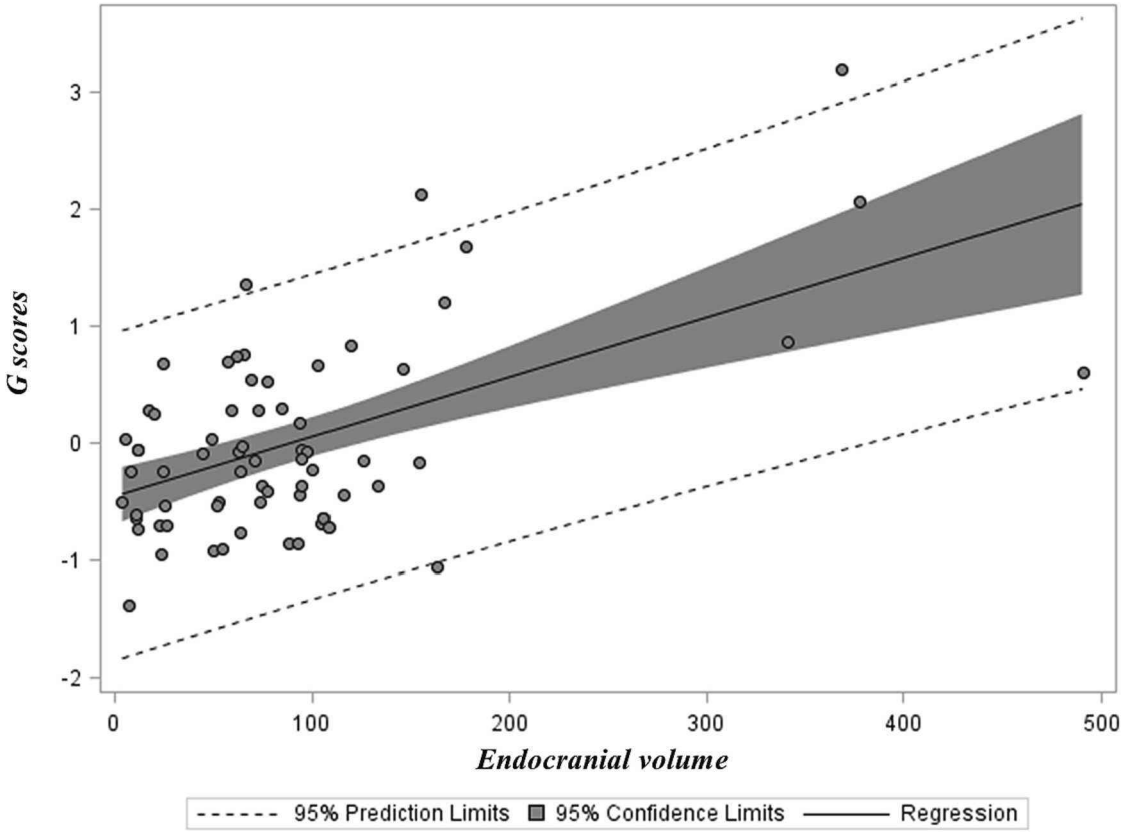


Fig. 1. Basic (non-phylogenetic) association between endocranial volume and *G* scores across a sample of 68 extant nonhuman primate species. The trend includes 95% confidence and prediction intervals.

significantly predicted *G*, neither the log-transformed neocortex residuals (adjusted for log-transformed brain volume) nor the neocortex ratio predicts *G* in extant nonhuman primates. Fisher *z*-tests also indicated that the effect sizes of raw and log-transformed neocortex values on *G* did not differ from those associated with the raw and log-transformed endocranial estimates. These results are described in more detail in Table 1.

Overall, these findings support the results of Deaner et al. (2007) and Reader et al. (2011), who found that both log-transformed neocortex volume and log-transformed brain mass positively correlate with an aggregate measure of general cognitive performance. Similar to our results, Deaner and colleagues determined that relative neuroanatomical measures such as encephalization quotient, log-transformed brain mass residuals (adjusted for log-transformed body mass), neocortex residuals (adjusted for the influence of log-transformed brain size), and the neocortex ratio did not predict the global cognitive performance of extant nonhuman primates. Consequently, due to the better statistical fit of endocranial volume and its statistical equivalence with overall neocortex volume, our ancestral reconstructions of *G* in extinct primates did not rely on relative measures of neuroanatomy (e.g. brain size corrected for body size).

G scores broken out by species and estimation method are presented in Table 2. These include the values imputed by phylogenetic bracketing on the basis of the association between endocranial volume and *G* scores in extant primates. Convergence between the estimates was extremely high (Cronbach's $\alpha = 0.99$).

Mean standardized *G* scores across phylogenetic reconstructions indicated that *Dryopithecus brancoi*, *Dryopithecus fontani*, and

Table 1
Phylogenetic Generalized Least Squares models examining the influence of various morphological predictors on *G* in a sample of extant nonhuman primates.

Predictors	*β*	*Std. Error*	*t-value*	*p-value*	*N*
z-Log Body mass (g)*	0.378	0.113	3.34	0.0014	68
z-Endocranial volume (cm^3)	0.553	0.102	5.43	<0.0001	68
z-Log Endocranial volume (cm^3)	0.442	0.110	4.04	0.0001	68
z-Encephalization quotient	0.257	0.130	1.98	0.0513	68
z-Log Brain mass (g) residuals adjusting for Log Body mass* (g)*	0.227	0.149	1.52	0.1340	61
z-Neocortex volume (mm^3)	0.675	0.118	5.71	<0.0001	40
z-Log Neocortex volume (mm^3)	0.501	0.139	3.62	0.0008	40
z-Log Neocortex (mm^3) residuals adjusting for Log Brain volume (mm^3)	0.044	0.155	0.28	0.7772	40
z-Neocortex ratio	0.225	0.161	1.39	0.1716	40
Fisher z Effect Size Comparison				***z-value***	***p-value***
z-Log Neocortex volume (mm^3) v. z-Log Endocranial volume (cm^3)				0.37	0.711
z-Neocortex volume (mm^3) v. z-Endocranial volume (cm^3)				0.96	0.337

Note. Symbols inside the parentheses indicate the original scale used to measure each morphological phenotype prior to standardization. *Z*-scores were estimated after computing the various adjusted values, as in the cases of Log Neocortex ratio and the various unstandardized neuroanatomical residuals. * Lindenfors et al.'s (2021) database classifies these variables as weights. Although weight and mass are often erroneously used as if they refer to the same concept, obviously they do not.

Table 2
Standardized *G* scores for extant and extinct primates based on three different phylogenetic and traditional reconstruction procedures.

Binomial	*ECV*	*z-G PRMA*	*z-G PGLS*	*z-G RMA*	*Mean z-G*	*Status*
Homo neanderthalensis	1404	5.542	5.000	5.389	5.310	Extinct
Homo sapiens (Modern human)	1349	5.302	4.781	5.152	5.079	Extant
Homo heidelbergensis	1242	4.835	4.355	4.693	4.628	Extinct
Homo erectus	991	3.739	3.355	3.616	3.570	Extinct
Homo habilis	624.3	2.137	1.895	2.042	2.024	Extinct
Paranthropus boisei	515	1.660	1.459	1.573	1.564	Extinct
Paranthropus robustus	493.33	1.565	1.373	1.480	1.472	Extinct
Australopithecus africanus	460	1.419	1.240	1.337	1.332	Extinct
Australopithecus garhi	450	1.376	1.200	1.294	1.290	Extinct
Australopithecus afarensis	446	1.358	1.184	1.276	1.273	Extinct
Pan troglodytes (Chimpanzee)	368.35	0.184	2.265	1.264	1.238	Extant
Paranthropus aethiopicus	431.75	1.296	1.128	1.215	1.213	Extinct
Homo floresiensis	425.7	1.270	1.104	1.189	1.187	Extinct
Kenyanthropus platyops	425	1.267	1.101	1.186	1.185	Extinct
Sahelanthropus tchadensis	350	0.939	0.802	0.864	0.868	Extinct
Ardipithecus ramidus	325	0.830	0.702	0.757	0.763	Extinct
Papio papio (Guinea baboon)	155.44	0.019	1.435	0.761	0.739	Extant
Dryopithecus brancoi	317.5	0.797	0.673	0.725	0.731	Extinct
Pongo pygmaeus (Bornean orangutan)	377.38	0.008	1.383	0.729	0.707	Extant
Dryopithecus fontani	289	0.673	0.559	0.603	0.611	Extinct
Dryopithecus wuduensis	289	0.673	0.559	0.603	0.611	Extinct
Papio ursinus (Chacma baboon)	178	−0.052	1.079	0.545	0.524	Extant
Oreopithecus bambolii	251.83	0.510	0.411	0.443	0.455	Extinct
Cebus apella (Tufted capuchin)	66.63	−0.102	0.828	0.393	0.373	Extant
Papio anubis (Olive baboon)	167.42	−0.127	0.705	0.319	0.299	Extant
Theropithecus brumpti	187	0.227	0.153	0.165	0.182	Extinct
Pan paniscus (Bonobo)	341.29	−0.180	0.439	0.157	0.139	Extant
Nomascus gabriellae (Yellow-cheeked gibbon)	119.38	−0.183	0.425	0.149	0.130	Extant
Theropithecus oswaldi	168	0.144	0.077	0.083	0.101	Extinct
Proconsul heseloni	167	0.140	0.073	0.079	0.097	Extinct
Proconsul nyanzae	167	0.140	0.073	0.079	0.097	Extinct
Cebus albifrons (White-fronted capuchin)	65.45	−0.196	0.357	0.108	0.090	Extant
Cercopithecus mona (Mona monkey)	61.84	−0.198	0.348	0.102	0.084	Extant
Piliocolobus kirkii (Zanzibar red colobus)	57.25	−0.206	0.311	0.080	0.062	Extant
Saimiri oerstedii (Central American squirrel monkey)	25.07	−0.209	0.295	0.070	0.052	Extant
Macaca fuscata (Japanese macaque)	102.92	−0.209	0.294	0.069	0.051	Extant
Cercopithecoides williamsi	156	0.092	0.029	0.032	0.051	Extinct
Theropithecus darti	152	0.074	0.013	0.014	0.034	Extinct
Papio hamadryas (Hamadryas baboon)	146.17	−0.215	0.264	0.051	0.034	Extant
Parapapio whitei	150	0.065	0.005	0.006	0.026	Extinct
Gorilla gorilla (Western gorilla)	490.41	−0.219	0.242	0.038	0.020	Extant
Cebus olivaceus (Wedge-capped capuchin)	69.84	−0.230	0.190	0.006	−0.011	Extant
Trachypithecus johnii (Nilgiri langur)	78	−0.231	0.181	0.001	−0.016	Extant
Megaladapis edwardsi	137	0.009	−0.046	−0.050	−0.029	Extinct
Theropithecus baringensis	129	−0.026	−0.078	−0.084	−0.063	Extinct
Megaladapis madagascariensis	118	−0.074	−0.122	−0.131	−0.109	Extinct
Papio izodi	118	−0.074	−0.122	−0.131	−0.109	Extinct
Hadropithecus stenognathus	115	−0.087	−0.134	−0.144	−0.122	Extinct
Hylobates pileatus (Pileated gibbon)	84.69	−0.268	0.001	−0.109	−0.125	Extant
Parapapio jonesi	114	−0.092	−0.138	−0.149	−0.126	Extinct
Cercopithecus ascanius (Red-tailed monkey)	59.58	−0.269	−0.007	−0.113	−0.129	Extant
Aotus vociferans (Spix's night monkey)	17.7	−0.269	−0.007	−0.113	−0.130	Extant
Cebus capucinus (Colombian white-faced capuchin)	72.93	−0.269	−0.007	−0.113	−0.130	Extant
Eulemur mongoz (Mongoose lemur)	20.17	−0.276	−0.041	−0.134	−0.150	Extant
Anapithecus hernyaki	107.116	−0.122	−0.165	−0.178	−0.155	Extinct
Archaeolemur edwardsi	104	−0.135	−0.178	−0.192	−0.168	Extinct
Macaca tonkeana (Tonkean macaque)	93.79091	−0.286	−0.093	−0.166	−0.182	Extant
Palaeopropithecus maximus	99	−0.157	−0.198	−0.213	−0.189	Extinct
Archaeolemur majori	93	−0.184	−0.222	−0.239	−0.215	Extinct
Macaca anderssoni	87.896	−0.206	−0.242	−0.261	−0.236	Extinct
Chiropotes satanas (Black bearded saki)	49.40678	−0.308	−0.205	−0.233	−0.249	Extant
Loris tardigradus (Red slender loris)	5.87	−0.309	−0.206	−0.234	−0.250	Extant
Turkanapithecus kalakolensis	84.3	−0.222	−0.256	−0.276	−0.251	Extinct
Palaeopropithecus ingens	80	−0.240	−0.273	−0.295	−0.269	Extinct
Chlorocebus aethiops (Grivet)	65	−0.318	−0.251	−0.261	−0.277	Extant
Otolemur crassicaudatus (Brown greater galago)	11.78	−0.323	−0.276	−0.276	−0.291	Extant
Cercocebus galeritus (Tana River mangabey)	95.33714	−0.323	−0.280	−0.279	−0.294	Extant
Otolemur garnettii (Northern greater galago)	11.5	−0.324	−0.281	−0.279	−0.295	Extant
Erythrocebus patas (Patas monkey)	97.73	−0.325	−0.288	−0.283	−0.299	Extant
Cercopithecus diana (Diana monkey)	62.61	−0.326	−0.291	−0.285	−0.301	Extant
Mesopithecus pentelicus	72.5	−0.273	−0.303	−0.327	−0.301	Extinct
Daubentonia madagascariensis (Aye-aye)	44.85	−0.329	−0.305	−0.294	−0.309	Extant
Semnopithecus entellus (Northern plains grey langur)	95.23685	−0.334	−0.332	−0.310	−0.326	Extant

(*continued on next page*)

Table 2 (*continued*)

Binomial	*ECV*	*z-G PRMA*	*z-G PGLS*	*z-G RMA*	*Mean z-G*	*Status*
Cercopithecus mitis (Blue monkey)	71.33	−0.338	−0.351	−0.322	−0.337	Extant
Nomascus concolor (Black Crested Gibbon)	125.6714	−0.338	−0.352	−0.322	−0.337	Extant
Mandrillus sphinx (Mandrill)	153.88	−0.340	−0.365	−0.330	−0.345	Extant
Macaca arctoides (Stump-tailed macaque)	100.7	−0.349	−0.410	−0.358	−0.372	Extant
Macaca fascicularis (Crab-eating macaque)	63.98	−0.351	−0.416	−0.361	−0.376	Extant
Eulemur macaco (Black lemur)	24.51	−0.351	−0.416	−0.361	−0.376	Extant
Saguinus fuscicollis (Brown-mantled tamarin)	7.94	−0.351	−0.420	−0.364	−0.378	Extant
Victoriapithecus macinnesi	54	−0.354	−0.377	−0.406	−0.379	Extinct
Babakotia radofilai	48	−0.380	−0.401	−0.432	−0.404	Extinct
Pachylemur jullyi	46	−0.389	−0.409	−0.441	−0.413	Extinct
Xenothrix mcgregory	45	−0.393	−0.413	−0.445	−0.417	Extinct
Mesopropithecus globiceps	41	−0.411	−0.429	−0.462	−0.434	Extinct
Paralouatta varonai	41	−0.411	−0.429	−0.462	−0.434	Extinct
Theropithecus gelada (Gelada)	133.33	−0.370	−0.514	−0.421	−0.435	Extant
Antillothrix bernensis	40.58	−0.412	−0.430	−0.464	−0.436	Extinct
Macaca radiata (Bonnet macaque)	74.87	−0.371	−0.519	−0.424	−0.438	Extant
Macaca nigra (Celebes crested macaque)	94.9	−0.372	−0.521	−0.425	−0.439	Extant
Macaca silenus (Lion-tailed macaque)	78	−0.378	−0.551	−0.443	−0.457	Extant
Lophocebus albigena (Grey-cheeked mangabey)	93.97	−0.383	−0.576	−0.458	−0.472	Extant
Rhinopithecus roxellana (Golden snub-nosed monkey)	115.5352	−0.384	−0.583	−0.462	−0.476	Extant
Aegyptopithecus Zeuxis	30	−0.459	−0.472	−0.509	−0.480	Extinct
Callithrix pygmaea (Pygmy marmoset)	4.17	−0.392	−0.624	−0.487	−0.501	Extant
Alouatta caraya (Black howler monkey)	52.63	−0.393	−0.631	−0.491	−0.505	Extant
Colobus guereza (Mantled guereza)	74.39	−0.393	−0.631	−0.491	−0.505	Extant
Dolichocebus gaimanensis	22.14	−0.493	−0.504	−0.543	−0.513	Extinct
Leptadapis magnus	21.7	−0.495	−0.506	−0.545	−0.515	Extinct
Alouatta guariba (Brown howler monkey)	51.7	−0.398	−0.653	−0.505	−0.518	Extant
Eulemur fulvus (Common brown lemur)	25.77	−0.398	−0.653	−0.505	−0.518	Extant
Homunculus patagonicus	19.85	−0.503	−0.513	−0.553	−0.523	Extinct
Saguinus mystax (Moustached tamarin)	11.09	−0.409	−0.710	−0.540	−0.553	Extant
Parapithecus grangeri	11.4	−0.540	−0.547	−0.589	−0.559	Extinct
Notharctus tenebrosus	10.43	−0.544	−0.550	−0.593	−0.563	Extinct
Notharctus osborni	10.4	−0.544	−0.551	−0.593	−0.563	Extinct
Macaca nemestrina (Pig-tailed macaque)	105.59	−0.413	−0.729	−0.551	−0.564	Extant
Cercocebus torquatus (Collared mangabey)	105.99	−0.414	−0.733	−0.554	−0.567	Extant
Smilodectes gracilis	9.31	−0.549	−0.555	−0.598	−0.567	Extinct
Callimico goeldii (Goeldi's monkey)	11.43	−0.415	−0.737	−0.556	−0.569	Extant
Adapis parisiensis	8.63	−0.552	−0.558	−0.601	−0.570	Extinct
Chilecebus carrascoensis	7.46	−0.557	−0.562	−0.606	−0.575	Extinct
Rooneyia viejaensis	7.23	−0.558	−0.563	−0.607	−0.576	Extinct
Microsyops annectens	5.9	−0.564	−0.568	−0.613	−0.582	Extinct
Plesiadapis tricuspidens	5.21	−0.567	−0.571	−0.616	−0.585	Extinct
Plesiadapis cookei	5	−0.568	−0.572	−0.616	−0.585	Extinct
Pronycticebus gaudryi	4.8	−0.569	−0.573	−0.617	−0.586	Extinct
Ateles geoffroyi (Black-handed spider monkey)	105.09	−0.421	−0.766	−0.574	−0.587	Extant
Microchoerus erinaceus	4.26	−0.571	−0.575	−0.620	−0.589	Extinct
Necrolemur antiquus	3.8	−0.573	−0.577	−0.622	−0.591	Extinct
Necrolemur zitteli	3.8	−0.573	−0.577	−0.622	−0.591	Extinct
Catopithecus browni	3.1	−0.576	−0.580	−0.625	−0.593	Extinct
Propithecus verreauxi (Verreaux's sifaka)	26.21	−0.423	−0.780	−0.582	−0.595	Extant
Lemur catta (Ring-tailed lemur)	22.9	−0.424	−0.781	−0.583	−0.596	Extant
Ignacius frugivorus	2.14	−0.580	−0.583	−0.629	−0.598	Extinct
Ignacius graybullianus	2.14	−0.580	−0.583	−0.629	−0.598	Extinct
Tetonius homunculus	1.5	−0.583	−0.586	−0.632	−0.600	Extinct
Macaca thibetana (Tibetan macaque)	108.9	−0.425	−0.791	−0.588	−0.602	Extant
Leontopithecus chrysomelas (Golden-headed lion tamarin)	11.83577	−0.427	−0.800	−0.594	−0.607	Extant
Piliocolobus badius (Western red colobus monkey)	63.59	−0.434	−0.834	−0.615	−0.628	Extant
Macaca mulatta (Rhesus macaque)	88.98	−0.447	−0.900	−0.654	−0.667	Extant
Macaca sylvanus (Barbary macaque)	93.2	−0.448	−0.905	−0.658	−0.671	Extant
Alouatta seniculus (Colombian red howler monkey)	55.22	−0.455	−0.938	−0.678	−0.690	Extant
Alouatta palliata (Mantled howler monkey)	49.88	−0.457	−0.949	−0.684	−0.697	Extant
Saimiri sciureus (Common squirrel monkey)	24.14	−0.462	−0.976	−0.700	−0.713	Extant
Papio cynocephalus (Yellow baboon)	163.19	−0.479	−1.060	−0.752	−0.764	Extant
Callithrix jacchus (Common marmoset)	7.24	−0.530	−1.316	−0.907	−0.918	Extant

Note. ECV: Endocranial estimates collected from Melchionna et al. (2020); PRMA: Phylogenetic Reduced Major Axis model reconstruction; PGLS: Phylogenetic Generalized Least Squares model reconstruction; RMA: Reduced Major Axis model reconstruction. The values for humans were based on extant endocranial estimates sourced from Melchionna et al. (2020).

Dryopithecus wuduensis had *G* values higher than that of Chacma baboons (*Papio ursinus*), but lower than Guinea baboons (*Papio papio*). Extinct hominins such as *Ardipithecus ramidus*, *Kenyanthropus platyops*, *Sahelanthropus tchadensis*, *Homo floresiensis*, and *Paranthropus aethiopicus* had higher *G* scores compared to Guinea baboons but remained below the chimpanzee scores (*Pan troglodytes*). By contrast, *Australopithecus afarensis*, *Australopithecus garhi*, *Australopithecus africanus*, *Paranthropus robustus*, *Paranthropus boisei*, and *Homo habilis* exhibited *G* scores between 1.25 and 2.02 standard deviations above the primate mean. *Homo erectus* and *Homo heidelbergensis* respectively reached 3.57 and 4.63

standard deviations above the mean. Both *Homo sapiens* and *Homo neanderthalensis* exhibited *G* scores that were more than five standard deviations above the primate mean, with Neanderthals showing slightly higher scores than modern humans.

3.1. Testing for grey-ceiling-like effects

Our study also analyzed possible grey-ceiling effects, i.e. potential negative associations between *G* and phylogenetic branch length, where high-*G* species would be expected to feature shorter branches due to higher speciation and extinction rates. The Phylogenetic Comparative model revealed a significant negative influence of the reconstructed PRMA values for *G* for the extant and extinct nonhuman primates on the standardized log-transformed length of branches (henceforth length of branches).

Furthermore, a Phylogenetic Comparative model indicated a significant negative association between the reconstructed PGLS values and the length of the branches in the phylogenetic tree. A Phylogenetic Comparative model evaluated the influence of the Reduced Major Axis reconstructed *G* values (estimated without phylogenetic controls) on the length of the branches of the phylogenetic tree. The model identified a negative and significant effect alongside a sizable phylogenetic signal. Finally, the Phylogenetic Generalized Least Squares model examining the mean value of standardized reconstructed *G* scores indicated that this variable negatively and significantly predicted the length of branches (Fig. 2). Results are shown in Table 3.

As a robustness test, we re-ran a PGLS evaluating the association between standardized mean *G* scores on Log branch lengths in a subsample of extinct primate species, after removing species outside a range of −3.7 to +3.7, and implementing a winsorization procedure on the data. The PGLS reached statistical significance and revealed that mean *z*-*G* negatively and significantly predicted the criterion variable ($\beta = -0.424$, $p = .04182$). Consequently, the negative association between mean *G* scores and branch length is not just reflecting the influence of unusually large-brained extinct or extant species such as *H. sapiens* or *H. neanderthalensis*.

3.2. Quantitative trait speciation and extinction model

Four rival macroevolutionary models were computed to analyze the relations between Log mean *G* scores, speciation rates (λ), and extinction rates (μ) in a subsample of 68 extinct primates. The results are shown in Table 4. A model comparison revealed that the model with variable speciation and extinction rates had a better statistical fit (*AIC weight* = 1.000). As shown in Fig. 3, greater mean *G* scores tend to increase both speciation rates and extinction rates. Recently, Rabosky & Goldberg, 2015 provided empirical evidence suggesting that these kinds of speciation and extinction models can incur Type I errors incorrectly identifying an association between the trait of interest and the corresponding rates of speciation and extinction. To check for this, we simulated 100 models based on a random attribute (modeling a trait that has no influence over the speciation or extinction rates) estimated with the *sim. character* function found in the *diversitree* package. We computed *ΔAIC* values between the constant speciation and extinction model, and the variable speciation and extinction model, for each of the 100 simulations. This leads to a distribution of *ΔAIC* values, and we ran a *z*-significance test to determine whether the *ΔAIC* estimate computed with the actual data significantly differed from these simulated values. We found a statistically significant difference between the actual *ΔAIC* value and the simulated estimates, suggesting the current results were not false positives (see Fig. 4). We also conducted several QuaSSE models with the raw *G* scores. The results remained consistent with the previous findings. The full model was not only statistically significant ($p < .0001$), but also featured the strongest statistical fit (*AIC* weight = 0.978). We compared the full model's *AIC* (408.71) against the simulated *AIC* values: a *z*-test indicated a significant difference between the actual *AIC* and the simulated scores ($p < .0001$).

3.3. Phylogenetic tests of convergent validity

Our Multiple Phylogenetic Generalized Least Squares (PGLS) models showed positive and significant associations between the Transfer Index (TI) values and the reconstructed *G* values for all extant and extinct primate species in the dataset (using Beran et al.'s equation described previously). The analysis estimated a positive and significant association between the mean standardized *G* scores and the reconstructed standardized TI values ($\beta = 0.923$, $p < .0001$) across all extinct and extant primates in the sample (Fig. 5). (See Table 5.)

We also ran a robustness test, using PGLS to evaluate the association between standardized mean *G* scores and standardized TI scores in a subsample of extinct primate species, after removing species outside a range of −3.7 to +3.7, and applying a winsorization procedure to the data. The PGLS reached statistical significance and revealed that mean *z*-*G* positively and significantly predicted the TI scores ($\beta = 0.949$, $p < .0001$). Thus, the positive association between *G* and the TI scores is not a by-product of the presence of unusually large-brained extinct or extant species such as *H. sapiens* or *H. neanderthalensis*.

Multiple PGLS evaluating the association between neuroanatomical covariance ratio estimates (the total covariation between neuroanatomical regions divided by the total covariation within neuroanatomical regions; a measure of neuroanatomical integration) and *G* scores revealed the presence of positive and significant associations (β values ranging from 0.369 to 0.388, see Table 6). Consequently, these results strongly suggest that *G* in extant and extinct primates is associated with the degree of neuroanatomical integration (Fig. 6). As restricting the analyses to extinct species featuring neuroanatomical covariance ratio estimates considerably reduced the statistical power due to its corresponding small sample size, no further statistical examinations were conducted.

3.4. Ancestral character reconstruction

The Ancestral Character Reconstruction with the mean *z*-*G* scores (Fig. 7) suggests that ever since the emergence of the primate order during the upper Cretaceous (c. 66 million years ago), *G* scores remained relatively stable across most primate species. The model also detected a slight increase in *G* values in the last common ancestor of Old World Monkeys (catarrhines) and Apes. *G* scores showed a second substantial increase around 23–16 million years ago with the diversification of Hominoidea (apes). The paleontological record suggests this clade evolved during the Oligocene-Miocene transition (c. 23 million years ago) in tropical Africa (Prothero, 2006). Then, around 14 million years ago, a decline in global temperatures led to a reduction of tropical forests in East Africa and to the expansion of more open savanna-like environments (Prothero, 2006). Even though initially Eurasia experienced a decline in temperature and humidity, a subsequent increase in global temperatures during the Middle Miocene (c. 16–11 million years ago) enabled the return of forests in Europe and Asia (Prothero, 2006). Novel ecological niches associated with these favorable environmental conditions facilitated the radiation and spread of Hominoids in Eurasia (Begun, 2010, 2013).

Temperatures declined again during the Upper Miocene, around 11 million years ago. Close to 1.5 million years later, the paleoecological record indicates temperate forests replaced subtropical forests in Europe (an ecological disruption often referred to as the Vallesian crisis; Agustí, Cabrera, & Garcés, 2013). Africa and certain parts of Asia remained

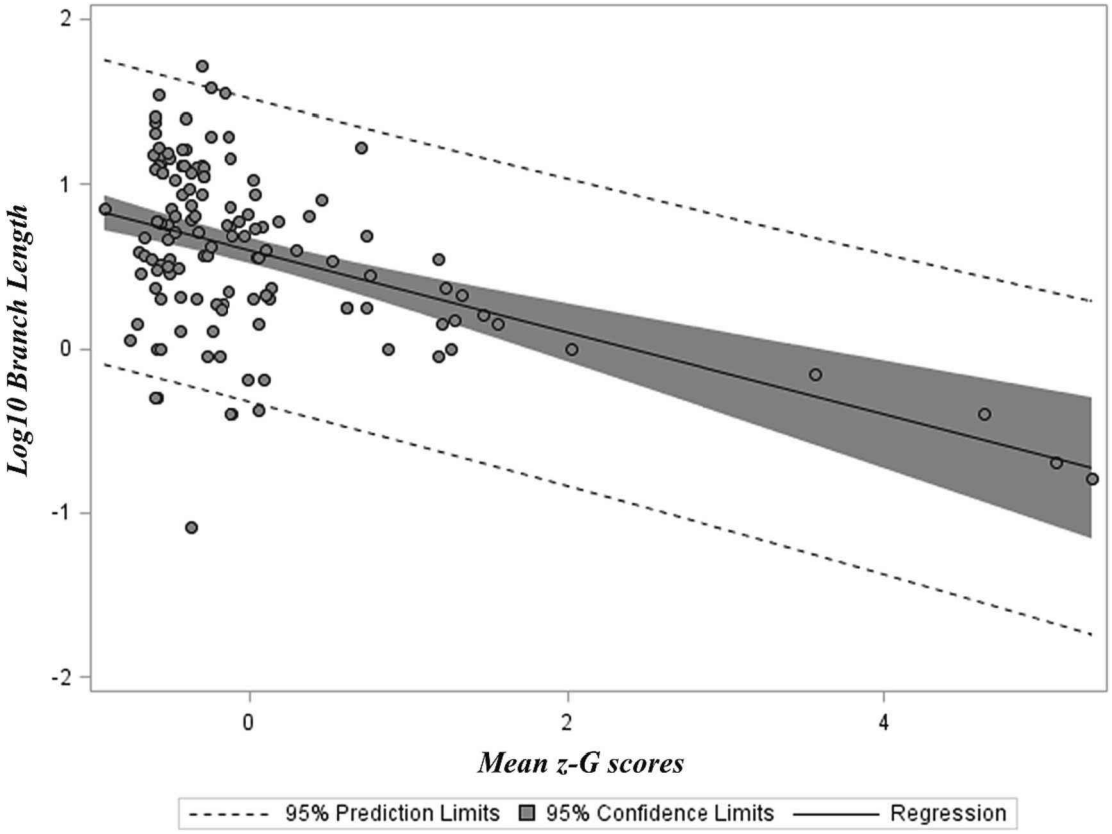


Fig. 2. Non-phylogenetic association between mean standardized reconstructed *G* scores and log-transformed branch lengths in a sample of extant and extinct primate species. The trend includes 95% confidence and prediction intervals.

ecological safe havens for Miocene hominoids.[4] This rapid increase in *G* scores continued with the emergence of the family Hominidae (Great Apes). Extinct lineages, such as *Oreopithecus bambolii*, *Dryopithecus fontani*, and *Dryopithecus wuduensis* attained *G* values slightly lower than those of Bornean orangutans (*Pongo pygmaeus*). The considerable morphological similarities between extinct Dryopiths and contemporary hominines (African apes) provide additional evidence supporting a European origin of this clade. Hence, the more parsimonious hypothesis posits that these morphological features evolved once in Dryopiths and persisted in their hominine descendants, rather than evolving as homoplasies twice: first in Middle-to-Late Miocene European apes and later independently in Late Miocene African hominines (Begun, 2010, 2013).

By the end of the Miocene and the early Pliocene, the common ancestor of *Pan* and extinct hominins showed a clear increase in *G* scores, as evidenced in descendant taxa such as *Sahelanthropus tchadensis*, *Ardipithecus ramidus*, and *Kenyanthropus platyops*. The evolution of *Australopithecus afarensis*, *Australopithecus garhi*, *Paranthropus aethiopicus*, *Paranthropus robustus*, *Paranthropus boisei*, and *Australopithecus africanus* also showed a progressive increase in *G* scores. These values continue to rise with the evolution of the genus *Homo,* first in the last common ancestor of *Homo erectus* and *Homo heidelbergensis*, and then in the common ancestor of *Homo neanderthalensis* and *Homo sapiens*, where the values are highest.

3.5. Macroevolutionary model comparison

Based on our AIC estimates, the Kappa model yielded a considerably better fit to the data (*AIC weight* = 1.000) than alternative macroevolutionary models did (see Table 7). A similar pattern emerged when considering AICc values, where the Kappa model featured a better fit (*AICc weight* = 1.000) than the other models. Hence *G* values across extinct and extant primate species vary as a function of the number of speciation events. This dynamic is generally described as cladogenic (meaning that speciation involves changes in trait levels). Moreover, when the Kappa statistic is <0.5 (it is <0.0001 in the current model), the macroevolutionary mode can be described as punctuational (Pagel, 1999). Punctuation was once believed to be quite rare in fossil taxa, but is now thought to be relatively common (Erwin & Anstey, 1995); and it is typified not just by rapid cladogenesis, but can also occur with long periods of relatively little evolutionary change (termed *stasis*). The

[4] Three species of Late Miocene apes have been recovered from East African sediments, with some researchers claiming that these lineages led to the speciation of extant African apes and extinct hominins. Ishida and Pickford (1997) argue that the partial maxilla of the Kenyan ape *Samburupithecus* (~9.5 million years ago) shares several features with the genus *Gorilla*. However, these similarities have been called into question in that this taxon shares several morphological characteristics with more primal African apes (e.g., *Proconsul*; Begun, 2010, 2013) and thus could be considered more similar to them than extant genera such as *Gorilla* and *Pan*. There is a similar controversy with respect to the evolutionary history of *Nakalipithecus*, an extinct hominoid taxon recovered from Kenyan sediments dated approximately 9.8 million years ago. Although some paleoanthropologists claim that the invasion of *Nakalipithecus* into Southeastern Europe led to the subsequent speciation of specific European apes, such as *Ouranopithecus*, other researchers argue that since *Nakalipithecus* evolved more than two million years after the rise of European apes, this taxon probably descended from European ancestors rather than vice versa (Begun, 2010, 2013). Controversy of the same sort surrounds other taxa, such as *Chororapithecus*, a genus of late Miocene Ethiopian ape dated around 10.5 million years ago. Even though Suwa, Kono, Katoh, Asfaw, and Beyene (2007) proposed that *Chororapithecus* dental attributes were morphologically similar to gorillas, these conclusions have been disputed (Begun, 2010, 2013). Moreover, the fact that *Chororapithecus* speciated ~2 million years after the rise of European apes suggests this taxon descended from Dryopiths rather than African ancestors (Begun, 2010, 2013).

Table 3
Phylogenetic Generalized Least Squares models using *G* scores (including the reconstructed values for extinct nonhuman primates) to predict *z*-Log branch lengths.

Measure	*β*	*Std. error*	*p-value*	*κ*	*Pagel's λ (95%CI)*	*LB*	*UB*
z-G PRMA	−0.407	0.100	0.0001	1	0.910 (0.792, 0.965)	<0.0001	<0.0001
z-G PGLS	−0.255	0.096	0.0086	1	0.902 (0.768, 0.963)	<0.0001	<0.0001
z-G RMA	−0.346	0.100	0.0007	1	0.900 (0.768, 0.961)	<0.0001	<0.0001
Mean z-G	−0.350	0.100	0.0007	1	0.900 (0.768, 0.961)	<0.0001	<0.0001

Note. PRMA: Phylogenetic Reduced Major Axis model reconstruction; PGLS: Phylogenetic Generalized Least Squares model reconstruction; RMA: Reduced Major Axis model reconstruction.

Table 4
QuaSSE model comparison of four macroevolutionary models exploring the association between Log mean *G* scores, speciation rates, and extinction rates in a sample of 68 extinct primate species.

Models	*Df*	*lnLik*	*AIC*	*ΔAIC*	*AIC weight*	x^2	*p-value*
Constant speciation, constant extinction	3	−145.29	296.57	60.83	0.000		
Variable speciation, constant extinction	4	−124.97	257.94	22.20	0.000	40.629	0.0000
Variable extinction, constant speciation	4	−145.29	298.58	62.84	0.000	−0.002	1.0000
Variable speciation, variable extinction	5	−112.87	235.74	0.00	1.000	64.837	0.0000

combination of these two macroevolutionary modes was termed *Punctuated Equilibria* by Eldredge and Gould (1972).[5] Stasis is also clearly evident in these data, as *G* scores remained relatively stable across most primate lineages until the rise of catarrhines and the subsequent evolution of the superfamily Hominoidea. Within the ape superfamily, *G* values increased in a fairly step-wise fashion (except in *Homo floresiensis*), compared to other primate superfamilies—this shift occurred precipitously in the last common ancestor of hominoids.

4. Discussion

The relationship between brain volume and *G* in 68 extant primate species was used to infer the levels of *G* of 68 extinct primate species (and one extant species: *H. sapiens*), based on endocranial volume data. Three different approaches to phylogenetic bracketing, each based on slightly different assumptions, yielded highly convergent estimates of *G*. These scores also strongly converged with the Transfer Index—a measure of cognitive flexibility inferred based on the association between learning flexibility and brain size in extant primates, and also with a measure of neural integration. *G* also negatively predicted branch length, suggesting that higher-*G* species tend to go extinct faster.

Fitting different macroevolutionary models to the phylogeny of *G* supported a Kappa model, which shows change consistent with a punctuated equilibria pattern. The apparent stasis of *G* in many of the more ancient primate lineages also suggests punctuated equilibria, with *G* having undergone rapid and cladogenic evolution in a (relatively small) number of primate species. Based on the ancestral character reconstruction, the biggest punctuation period followed the separation of the tribes *Gorillini* (gorillas) and *Hominini* (genera *Pan*, *Sahelanthropus*, *Orrorin*, *Kenyanthropus*, *Ardipithecus*, *Australopithecus*, *Paranthropus*, and *Homo*) around 10 million years ago. This is the split that eventually led to the emergence of anatomically modern *H. sapiens* around 300,000 years ago. These findings are consistent with Melchionna et al.'s (2020) observations on the evolution of brain volume across extinct and extant primates. Specifically, they note that "[t]he increases in both speciation rate and encephalization begin in the Oligocene (c. 34-23 million years ago), suggesting the two variables are causally associated. The substitution of early, stem Primates belonging to plesiadapiforms with crown Primates seems to be responsible for these macroevolutionary trends. *However, our findings also suggest that cognitive capacities favoured speciation in hominins*" (p. 14; emphasis added).

The period following the split between these two cladistic tribes (gorillas vs. other African great apes) is what we call the "ten-million-year explosion." This rapid evolution of *G* during this period seems loosely analogous to the rapid period of human adaptive evolution that occurred during the Holocene epoch (around 11,700 years ago to the present), with the frequency of novel haplogroup formation suggesting that rates of adaptive evolution during this period were substantially greater than those occurring in human populations during the preceding Pleistocene epoch (Cochran & Harpending, 2009; Frost, 2011; Hawks, Wang, Cochran, Harpending, & Moyzis, 2007).

Given its larger brain volume, *H. neanderthalensis* was assigned a slightly greater *G* value in all reconstructions. The possibility that this species may have had higher *G/g* than *H. sapiens* due to its larger brain volume, or its having acquired some local cognitive adaptations, has led to speculation that gene flow between Neanderthals and humans might have contributed *g*-relevant alleles to the modern human genome (Cochran & Harpending, 2009; Evans, Mekel-Bobrov, Vallender, Hudson, & Lahn, 2006; Lari et al., 2010).These alleles may have then flowed out to the rest of the human range due to the fitness benefits that they conferred.[6]

The theory that Neanderthals had higher *G* is nevertheless controversial. Several paleoneurological publications have emphasized the unique contribution of certain neural regions to the evolution of cognitive abilities in extinct hominins. For example, Balzeau, Holloway, and Grimaud-Hervé (2012) estimated that relative to *Homo sapiens*,

[5] The irony that punctuated equilibria, a model aggressively championed by Steven Jay Gould (1941–2002) over most of his working life, best describes the macroevolution of *G*, *g* being a trait that Gould vociferously argued was "chimerical" (Gould, 1996), is not lost on the authors. On the other hand, there are reasonable doubts about whether Gould and Eldredge should be credited as having actually originated punctuated equilibria in light of remarks from Ernst Mayr (1904–2005), who maintained that he first developed the idea and that both researchers were aware of this (a relevant early publication is Mayr, 1954). In an interview with *Skeptic* (Shermer & Sulloway, 2000), Mayr stated that "Gould was my course assistant at Harvard where I presented [punctuated equilibria] again and again for three years. So he knew it thoroughly. So did Eldredge. In fact, in his 1971 paper Eldredge credited me with it. But that was lost over time" (p. 79).

[6] Cochran and Harpending (2009) note that "a tiny bit of Neanderthal ancestry thrown into the mix tens of thousands of years ago could have resulted in many people today, possibly even all modern humans, carrying the advantageous Neanderthal version of some genes" (p. 44). Eswaran, Harpending, and Rogers (2005) suggested on the basis of simulations that among modern humans "as much as 80% of nuclear loci have assimilated genetic material from non-African archaic humans" (p. 1).

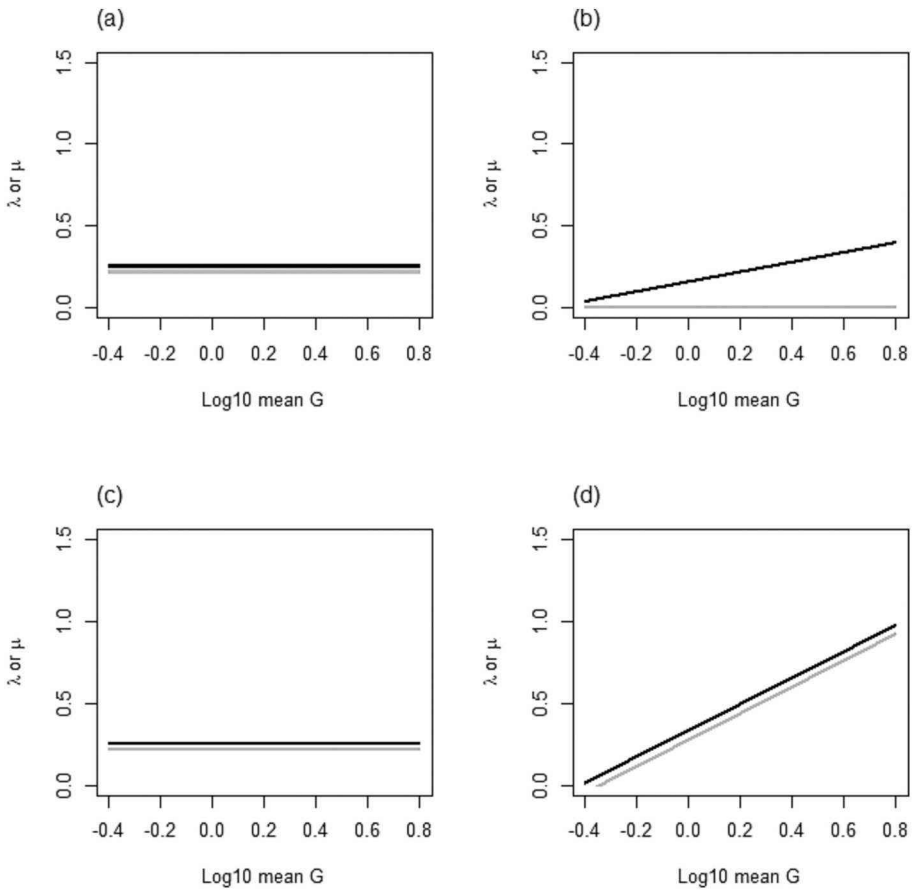


Fig. 3. QuaSSE models featuring the association between Log mean *G* scores, speciation rates, and extinction rates in extinct primates. The images represent four macroevolutionary models: (a) Constant speciation and extinction; (b) Variable speciation and constant extinction; (c) Variable extinction and constant speciation; and (d) Variable speciation and extinction. Speciation rates represented as dark lines (λ) and extinction rates represented as grey lines (μ).

Neanderthals had larger frontal and occipital lobes and smaller parietal-temporal regions. Kochiyama et al. (2018) also found that Neanderthals and early *Homo sapiens* differed in other neuroanatomical features. Their reconstruction revealed that early *Homo sapiens* had larger cerebellar hemispheres and a reduced occipital region compared to Neanderthals —a result that echoes Balzeau and colleagues' conclusions. Moreover, the positive associations between cerebellar volume and working and episodic memory, attention, language comprehension, and cognitive flexibility, above and beyond motor coordination, suggest that further paleocognitive and phylogenetic research would be required to estimate *G* in extinct primate taxa based on cerebellar data. Such data, alas, are not yet available for many prehistoric primates. Similarly, a debate persists concerning the mosaic or global reorganization of the brain during hominin evolution. For example, Bruner and Holloway (2010) suggested that the frontal lobe (including Broca's region) in Neanderthals and *Homo sapiens* enlarged independently from other neuroanatomical areas. In the same way, Bruner (2008, 2010) reported evidence of mosaic evolution in Neanderthals and *Homo sapiens*, with additional support for expanding and globularization of the temporal lobes in the latter species. Consistent with this is the study of Gunz et al. (2019), who found that "[introgressed] Neanderthal alleles on chromosomes 1 and 18 are associated with reduced endocranial globularity. These alleles influence expression of two nearby genes, *UBR4* and *PHLPP1*, which are involved in neurogenesis and myelination, respectively" (p. 120).

Even though some of these publications emphasize the unique contribution of certain neuroanatomical regions to potential cognitive differences between *Neanderthals* and *Homo sapiens*, the current neuroanatomical evidence indicates that *g* is distributed across the brain, and is not limited to specific neural areas.

Colom and collaborators (2006) identified multiple regions in the frontal cortex associated with intelligence, including BAs 9, 10, and 46, locations often associated with cognitive abilities such as working memory and attention. Other regions include BA 11, recruited during decision-making processes, executive functioning, planning, reasoning, and memory retrieval. Broca's area itself (BA 45, traditionally associated with speech production), is also a correlate of general intelligence through certain cognitive abilities, including semantic working memory, semantic decision process, and context-dependent information evaluation. Similarly, according to Colom and collaborators (2006), other regions involved in linguistic, and in particular syntactical, abilities (such as BA 47), are also associated with general intelligence. Beyond the frontal cortex, current neuroanatomical studies strongly suggest parietal regions, such as BA 3 and BA 5 (somatosensory areas), BA 7 (part of the somatosensory cortex and involved in object location), and BA 40 (a subregion of Wernicke's area), play crucial roles in intelligence. Temporal regions also contribute to general intelligence. For example, BAs 20, 21, and 22 (the inferior, middle, and temporal gyrus, respectively), are involved in cognitive abilities such as recognition memory, production and comprehension of words and general language use, visual processing, and auditory understanding. Lastly, Colom, Jung and Haier (2006) reported that occipital regions involved in shape recognition, integration functions, and feature extraction (BAs 18 and 19, components of the visual association cortex) provide essential input for higher-order cognitive processes. Consequently, the neuroscientific evidence reveals that *g* is distributed across the cerebral cortex, rather than being limited to frontal cortical regions.

A recent molecular genetic study, which examined the effects of introgressed Neanderthal variants on cognitive phenotypes, paints a

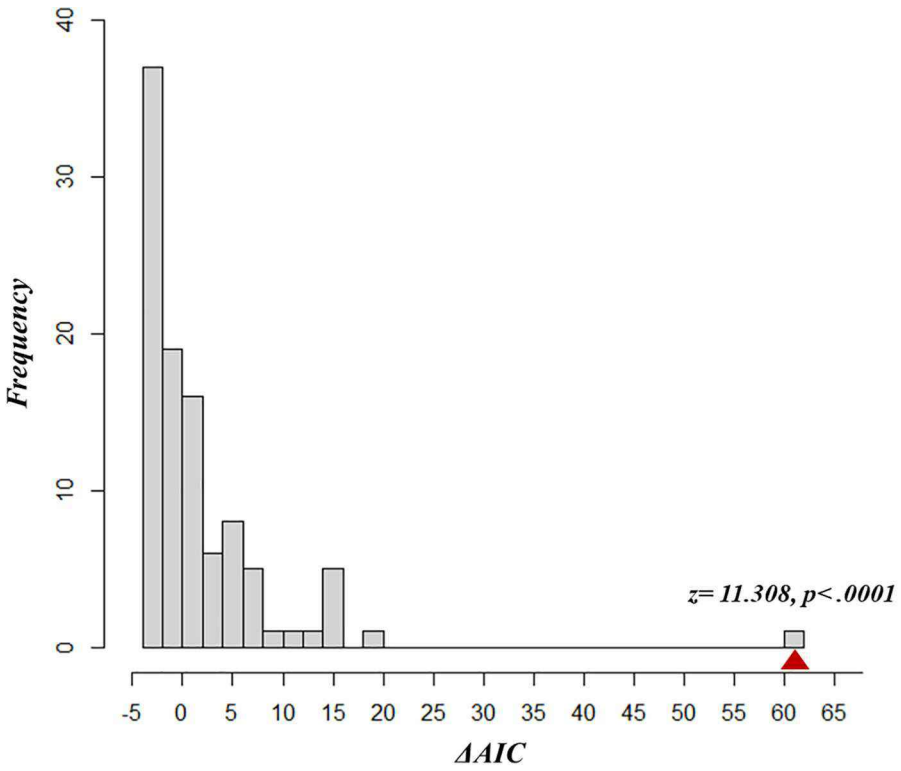


Fig. 4. Distribution of QuaSSE *ΔAIC* values comparing the best-fitting model (variable speciation and extinction) to the null model (constant speciation and extinction). The sample involved 100 simulations that modelled no influence of a trait on speciation and extinction rates. A z-significance test compared the observed *ΔAIC* value, obtained with the actual data, to those estimated using model simulations.

mixed picture on whether the effects of these alleles promote or inhibit cognitive ability in modern humans. Koller et al. (2022) examined the effects of various Neanderthal- and Denisovan-derived alleles on a large array of phenotypes (based on merging two very large Biobanks). These researchers identified nine SNPs that had either introgressed from Neanderthals alone, or in conjunction with Denisovan variants, and which

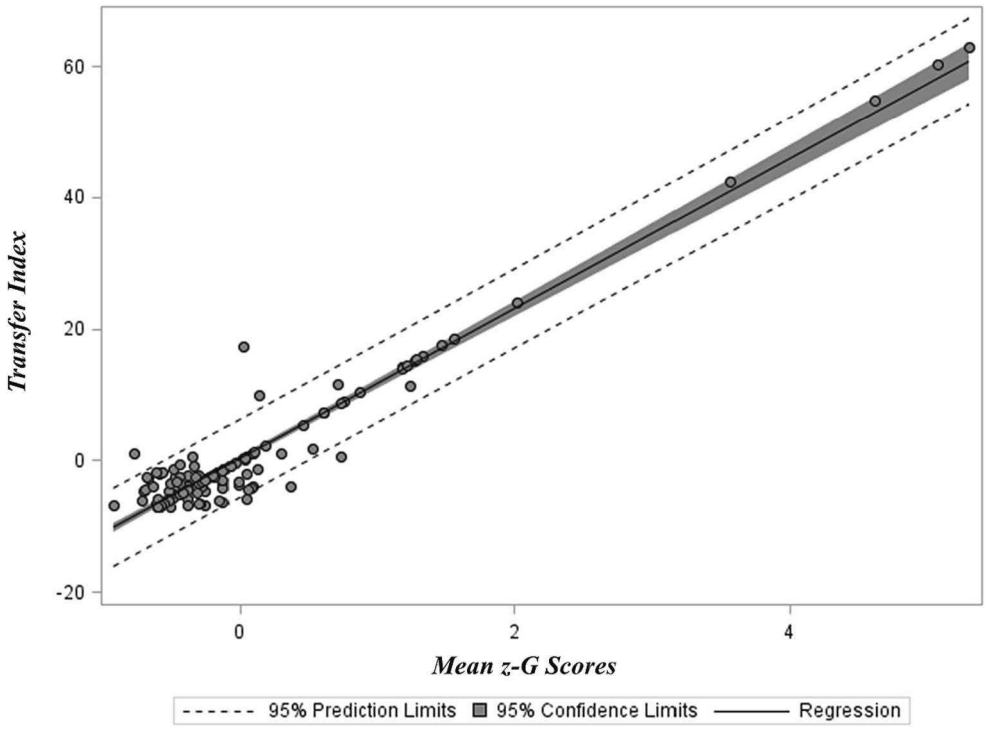


Fig. 5. Non-phylogenetic association between mean standardized reconstructed *G* scores and reconstructed Transfer Index values in extant and extinct primates. The trend includes 95% confidence and prediction intervals.

Table 5
Phylogenetic Generalized Least Squares models using *G* scores (including the reconstructed values for extinct nonhuman primates) to predict *z*-transfer index.

Measure	*β*	*Std. error*	*p-value*	*κ*	*Pagel's λ (95%CI)*	*LB*	*UB*
z-G PRMA	0.904	0.025	<0.0001	1	0.462 (0.146, 0.999)	<0.0001	0.0392
z-G PGLS	0.822	0.040	<0.0001	1	0.414 (0.019, 0.707)	0.0403	<0.0001
z-G RMA	0.916	0.026	<0.0001	1	0.276 (0.012, 0.596)	0.0323	<0.0001
Mean z-G	0.923	0.026	<0.0001	1	0.242 (0.000, 0.575)	0.0539	<0.0001

Note. PRMA: Phylogenetic Reduced Major Axis model reconstruction; PGLS: Phylogenetic Generalized Least Squares model reconstruction; RMA: Reduced Major Axis model reconstruction.

Table 6
Phylogenetic Generalized Least Squares models with log-transformed neuroanatomical covariance ratio estimates predicting *G* scores (including the reconstructed values for extinct nonhuman primates).

Measure	*β*	*Std. error*	*p-value*	*κ*	*Pagel's λ (95%CI)*	*LB*	*UB*
z-G PRMA	0.369	0.081	<0.0001	1	0.992 (0.950, 1.000)	<0.0001	0.0407
z-G PGLS	0.377	0.092	0.0002	1	0.958 (0.805, 0.997)	<0.0001	0.0085
z-G RMA	0.388	0.083	<0.0001	1	0.984 (0.906, 0.999)	<0.0001	0.0166
Mean z-G	0.384	0.083	<0.0001	1	0.984 (0.908, 0.999)	<0.0001	0.0176

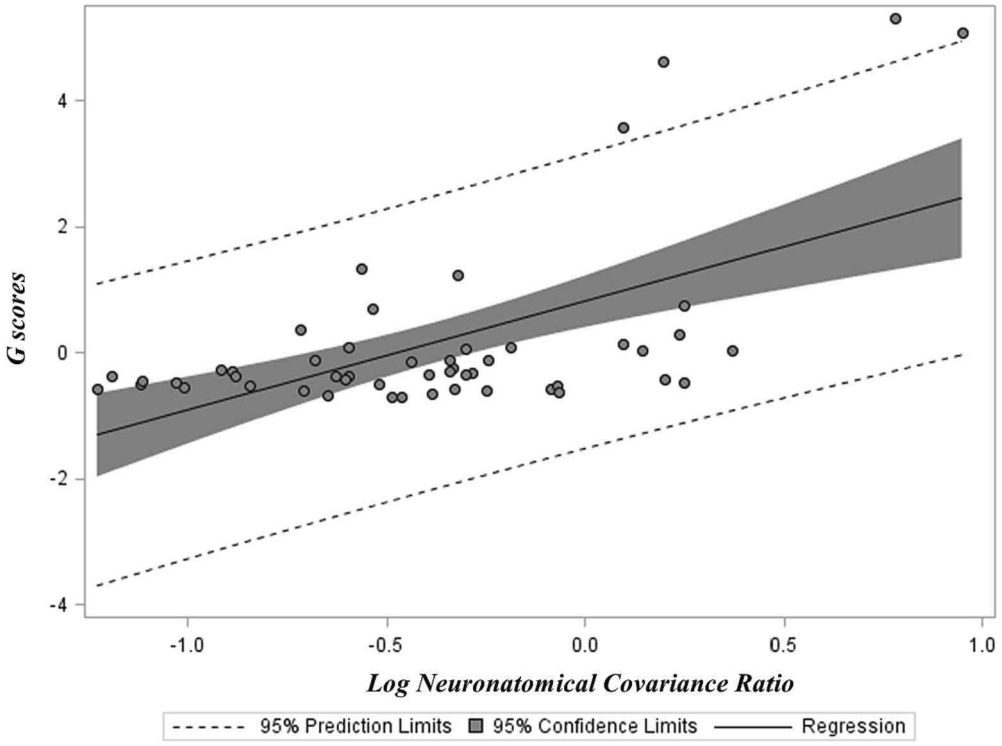


Fig. 6. Non-phylogenetic association between mean reconstructed *G* scores and log-transformed neuroanatomical covariance ratio in extant and extinct primates. The trend includes 95% confidence and prediction intervals.

were significantly associated with cognitive or (relevant) neurological performance measures. These cognition measures include number of correct matches in a matching round, number of times the snap-button was pressed (snap in this context implies that two alike images appear in the screen making this an associative task), three fluid intelligence measures (family relationship calculation, chained arithmetic, and conditional arithmetic), final attempt correct (a memory measure), forward digit-span, and mild cognitive impairment. *b* values estimated for each SNP ranged from −1.591 (in the case of the final attempt correct memory item) to 1.220 (in the case of the conditional arithmetic fluid intelligence measure). Using 1-*q* as the basis for weighting these effect sizes (*q* denotes the false-positive rate), a weighted-mean *b* of 0.061 is estimated (after reversing the sign of the *b* value associated with mild cognitive impairment). This hints at the possibility that across human populations, Neanderthal introgression may have enhanced *g*, perhaps through increased fluid rather than crystallized intelligence. This intriguing result should be followed up with a global admixture analysis using percent Neanderthal ancestry estimated via genotyping a large sample of individuals of homogeneous ancestry to predict *g* (ideally estimated using a broad ability battery) and educational attainment—-after very carefully controlling for population stratification and confounding environmental factors (such as parental socioeconomic status and other basic demographic factors such as age, sex, etc.). If percent Neanderthal ancestry remains a significant positive predictor of *g* after these controls, this would yield much higher- quality molecular genetic evidence for (a polygenic variant of) the cognitive archaic introgression model.

The observation that positive directional selection for *G* (and related

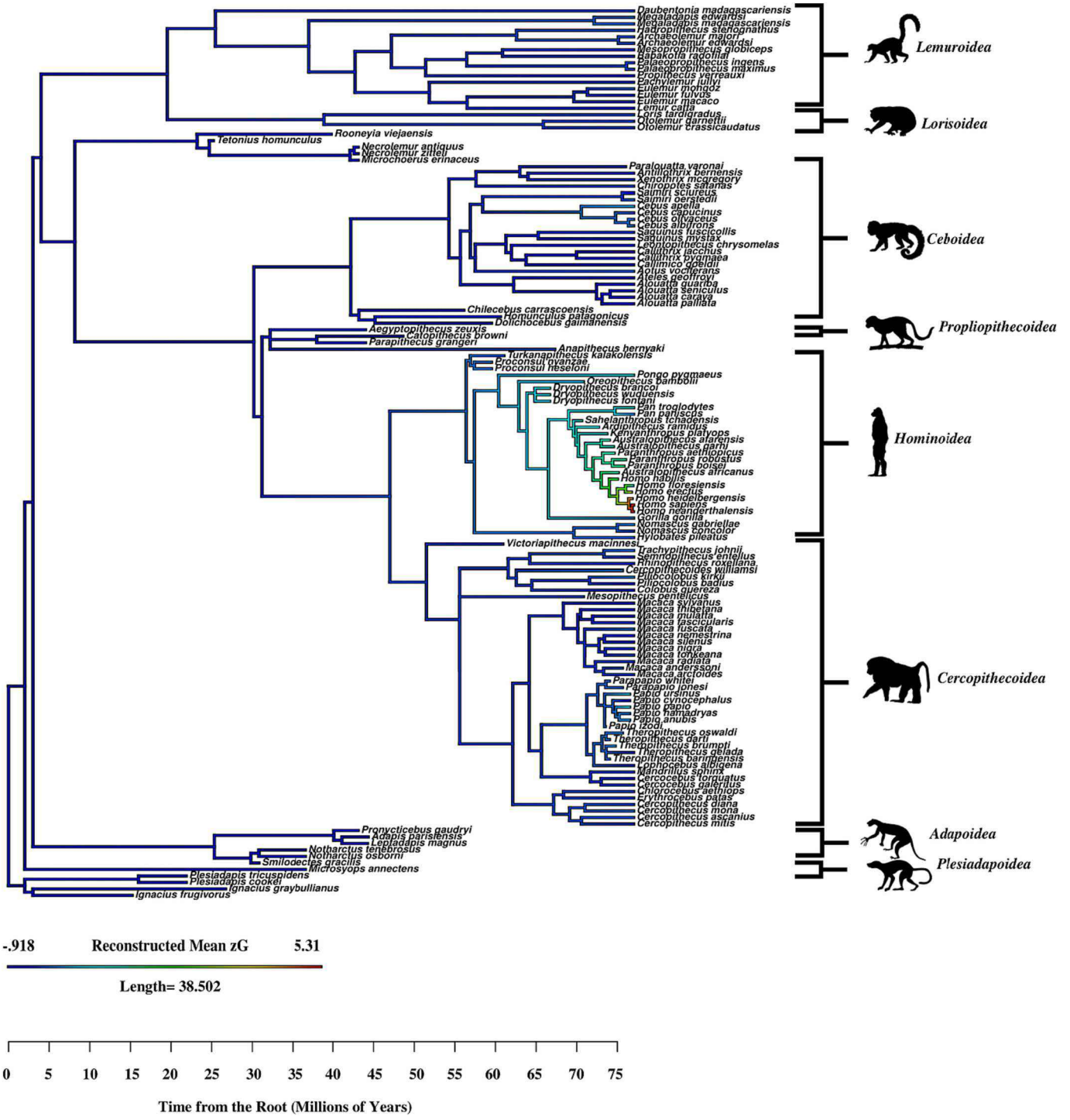


Fig. 7. Ancestral character reconstruction of mean standardized *G* scores for extant and extinct primate species. Silhouettes from http://www.phylopic.org. Images licensed for use under Public Domain, except *Plesiadapis* (Zica and modified by Keesey), *Notharctus tenebrosus* (Hartman), and *Aegyptopithecus zeuxis* (Tamura), licensed under Creative Commons 3.0 license http://creativecommons.org/licenses/by/3.0. The phylogenetic tree also features the superfamilies *Parapithecoidea*, *Paromomyoidea*, and *Omomyoidea*.

traits such as brain volume) has been very strong over the last ten million years suggests the existence of certain special conditions that provided fitness benefits sufficient to outweigh the costs of having higher *G/g*, these costs being the cause of the grey-ceiling effect.

No previous macroevolutionary comparative study has examined the connection between cognitive abilities and the risk of extinction estimated using direct measures of species persistence. This distinction is critical as there are several factors that are potentially problematic for the grey-ceiling hypothesis as currently construed. First, the use of r_{max} as a proxy for species vulnerability to extinction is based on some potentially dubious assumptions, such as that species with small effective population sizes and bioenergetically expensive tissue (such as brain mass) are necessarily more vulnerable to various mortality hazards and the action of *F*-type extinction vortices (positive feedback loops between shrinking population sizes and the rate at which deleterious mutations become stochastically fixed in such populations through increased

Table 7
Macroevolutionary model comparison evaluating the statistical fit of various modes on the mean standardized *G* scores.

Mean z-G Model AIC							
BM (AIC)	***OU (AIC)***	***LB (AIC)***	***EB (AIC)***	***KP (AIC)***	***MT (AIC)***	***RT (AIC)***	***WT (AIC)***
336.408	252.645	277.093	338.403	181.064	337.820	304.388	388.420
BM (Δ AIC)	***OU (Δ AIC)***	***LB (Δ AIC)***	***EB (Δ AIC)***	***KP (Δ AIC)***	***MT (Δ AIC)***	***RT (Δ AIC)***	***WT (Δ AIC)***
155.344	71.581	96.029	157.339	0.000	156.756	123.323	207.355
BM (w AIC)	***OU (w AIC)***	***LB (w AIC)***	***EB (w AIC)***	***KP (w AIC)***	***MT (w AIC)***	***RT (w AIC)***	***WT (w AIC)***
0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
Mean z-G Model AICc							
BM (AICc)	***OU (AICc)***	***LB (AICc)***	***EB (AICc)***	***KP (AICc)***	***MT (AICc)***	***RT (AICc)***	***WT (AICc)***
332.408	246.645	271.093	332.403	175.064	331.820	298.388	384.420
BM (Δ AICc)	***OU (Δ AICc)***	***LB (Δ AICc)***	***EB (Δ AICc)***	***KP (Δ AICc)***	***MT (Δ AICc)***	***RT (Δ AICc)***	***WT (Δ AICc)***
157.344	71.581	96.029	157.339	0.000	156.756	123.323	209.355
BM (w AICc)	***OU (w AICc)***	***LB (w AICc)***	***EB (w AICc)***	***KP (w AICc)***	***MT (w AICc)***	***RT (w AICc)***	***WT (w AICc)***
0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000

Note. BM: Brownian motion; OU: Ornstein–Uhlenbeck; LB: Lambda; EB: Early burst; KP: Kappa; MT: Mean trend; RT: Rate trend; and WT: White noise.

autozygosity; Gilpin & Soulé, 1986). Species with lower r_{max} values are likely to show a suite of traits that *adapt* them to living at the environmental carrying capacity (or even to increasing that carrying capacity). Such taxa are typically described as being *K*-selected (where *K* is the carrying capacity of a particular environment), or as having slow life history speeds (Woodley of Menie, Luoto, Peñaherrera-Aguirre, & Sarraf, 2021). Such *K*-selected adaptations can include increased parental solicitude, longer developmental periods (altricial development), greater investment of communitarian effort into the formation of inclusive-fitness-enhancing social structures (Ellis, Figueredo, Brumbach, & Schlomer, 2009), increased capacity to modify the parameters of the environment (via inceptive niche construction), and (potentially) increased genetic robustness to the action of new deleterious mutations (Woodley of Menie et al., 2021)—all of which may have the effect of buffering against mortality and extinction risks. Greater control over the environment can even lead *extrinsic* mortality to become *intrinsic* mortality (i.e., sources of mortality can become more controllable; this interpretation is consistent with Alexander's views on the evolution of human behavior and cognition (Alexander, 1989), further elaborated by Geary (2005)). So, it is important to decouple extinction vulnerability from r_{max}, which is a life history parameter that may, in many instances, contribute to increased phylogenetic persistence of specific lineages when low, as it will be associated with *K*-selected adaptations that can buffer species against sources of mortality. The solution here is to use persistence in time as a more neutral and direct indicator of species survival, as we have done.

Another potential problem is in the focus on brain mass as expensive tissue. While brain volume (which is almost perfectly correlated with mass across species) is an excellent proxy for *G* at the species level (as has been demonstrated here), and is the main basis for estimating *G* in the case of extinct species, brain volume does not perfectly capture *G*. For example, overall brain mass appears to be more phylogenetically conserved than *G* among extant primates (Fernandes et al., 2020; see also Miller & Penke, 2007 for similar arguments). This suggests that species-level variation in *G* is, in addition to brain volume, also a function of other partially independent factors, such as variation with respect to specific neuroanatomical volume indicators and factors such as white matter integrity. Even within modern human lineages, brain size, and polygenic scores for both educational attainment and fluid intelligence appear to show different trajectories over the last 30,000 years, with brain size having decreased (Stibel, 2021), but intelligence having increased based on archaeogenetic analysis (Kuijpers et al., 2022; Woodley of Menie, Younuskunja, Balan, & Piffer, 2017). This suggests that smaller brains with underlying genotypes predisposing toward greater *g* may have emerged through factors such as increased efficiency of cellular metabolism, which may have come at the expense of more volumetrically demanding, and potentially cognitively less efficient, brain features. Similarly, current neurophysiological examinations indicate primate species differ in white matter connectivity. For example, chimpanzees and extant humans feature noticeable differences in prefrontal regions involved in default mode and fronto-pariental networks (Barks, Parr, & Rilling, 2015; Garin et al., 2022; Wei et al., 2019). Thus, to better understand the apparent fact that higher *G* species go extinct more frequently, further research could shift away from focusing on the brain as an expensive organ, and toward a more multidimensional understanding of the evolution and basis of *G/g*, as a trait with a more nuanced mixture of macroevolutionary costs, risks, and benefits.

So, we think a different theory is needed to account for our findings, beyond the standard grey-ceiling hypothesis that explains the extinction risks of large brains and high intelligence mostly in terms of the high bioenergetic costs of brain tissue. We propose instead that *G/g* is indeed costly, but for different reasons: since *g* involves the integration of multiple lower-order cognitive systems, giving rise to domain-*general* problem-solving ability—whereby solutions to problems derived in one fitness domain can be applied to other, unrelated, fitness domains—*g* should be an unusually large target for harmful mutations. Contemporary genomic studies of the genetic architecture of *g* in humans support this idea, finding that about 10,000 genes "modulate" cognitive ability (Huguet et al., 2021). Considering that there may be only 19,000 protein-coding genes in the human genome (Ezkurdia et al., 2014), this suggests that perhaps more than half of all human protein-coding genes underlie intelligence. Huguet et al. (2021) specifically claim that "half of the coding genome affects intelligence" (p. 2672). Moreover, genetically informed familial studies have found evidence that individual differences in the burdens of rare variants (mutations with a minor allele frequency of <0.01) contribute significantly to the genetic architecture of *g*, a finding that is consistent with the trait being (at least in part) under mutation-selection balance (Hill et al., 2018).

By contrast, high cognitive modularization—as would be expected in species with minimal *G*—is likely to provide robustness against deleterious mutations. Modularization organizes biological networks into relatively independent subunits, such that, for example, a particular subunit (i.e., a module) will be strongly affected by a certain set of genes that are crucial to its growth and functioning, but will be little affected by other modules' genes (and vice versa). Hence, Tran and Kwon (2013), in discussing widely held views on modularity and robustness in gene regulatory networks, note that "gene regulatory networks … are typically modular and this enhances the mutational robustness by reducing pleiotropy or allowing for the malfunctioning of one module without producing failure in other modules" (p. 1). We predict that high-*G* species, such as humans, have brains that are more vulnerable to mutations because the relative lack of modularity in a domain-general cognitive system should substantially *increase* the probability that a

defect in any given gene will (slightly) disrupt the general functioning of that system.

Intriguingly, Perote-Peña (2019) has proposed a theory for the evolution of domain-general problem solving that describes how fitness benefits accruing to the solvers of complex social problems may have entailed selection for mutations that "merge" independent cognitive modules, giving rise to individual variation with respect to domain-general problem-solving capacity through reduction in modularization. Fernandes (2020) found that higher *G* and more integrated *G* are positively associated in extant primates, which would appear to support Perote-Peñas' model insofar as it may indicate that selection for higher *G* has, in fact, occurred through a process of demodularization increasing the total cognitive variance for which *G* accounts. It should also be noted that much of the brain's modularization involves small local networks that deal with narrow forms of information (e.g., visual angle). Moreover, a reviewer proposed that the aforementioned demodularization observed at the brain level could be a product of increasing integration among multiple localized networks dedicated to processing specific information. This neurocognitive specialism is accompanied by the emergence of structures associated with increasing top-down regulation of these lower-level networks, a possibility supported by the current comparative evidence (Barks et al., 2015; Garin et al., 2022; Wei et al., 2019). A reviewer also suggested that the change in the proportion of variance associated with *G* as a function of brain volume might be related to fundamental reorganization coupled with the reallocation of bioenergetic resources away from these lower-level functions to structures that subserve domain-general problem-solving abilities such as the energy-demanding frontal-parietal network (Jung & Haier, 2007).

One interesting concrete example of *G*'s cross-domain utility in hominin evolution could be *H. erectus*' discovery of fire. Wrangham (2017) has argued that this may have directly contributed to brain size increases in that the use of fire, denaturing proteins in food prior to ingestion, could have freed up bioenergetic resources once needed for digestion for brain development, and thus enhanced selection for greater brain size in evolutionary time. However, additional evidence is required to determine whether cooking food acted as the main evolutionary driver for *G* or reduced several evolutionary constraints in the presence of other selection pressures. Fire, in fact, is also used to solve problems across multiple domains beyond food preparation— predator avoidance and heat and light regulation, for example—and higher and more integrated *G* would have facilitated such use.

The fitness costs associated with this higher vulnerability to harmful mutations would accordingly require that *G* bring substantial adaptive benefits to compensate for this effect. The benefits would have to be especially large then for a species to rapidly evolve high *G*, i.e., for *G* to be under strong directional selection *and* exhibit punctuational dynamics for a sustained phylogenetic period. Given that the best explanation of the evolution of *g* posits that it is an adaptation for the solving of *novel* problems (i.e., those for which a species' phylogeny has not supplied prepared solutions) (Geary, 2009),[7] high *G* is expected to have large fitness benefits only when species are dealing with unfamiliar ecological and environmental challenges.

Consistent with this expectation, there is strong evidence that higher *G* (proxied by brain size) is positively associated with successful adaptation to novel environments by mammalian species (Sol, Bacher, Reader, & Lefebvre, 2008). Pronounced climatic changes with noted ecological and environmental effects, such as the Vallesian crisis discussed earlier, shaped migration patterns and other factors influencing primate phylogeny in ways that would have greatly increased exposure to novelty. Primates expanding into novel environments and ecologies likely encountered substantial unfamiliar adaptive problems in the form of inter-species and inter-group competition and conflict, including lethal aggression and warfare (see Gintis, van Schaik, & Boehm, 2019; Peñaherrera-Aguirre, Fernandes, & Figueredo, 2021a, 2021b). Woodley of Menie et al. (2017) argue for the crucial role of inter-group competition and conflict in the evolution of high levels of *g* in humans, which is in line with a recent publication emphasizing that "[c]ompetition for vital resources [with] conspecific outsiders present[s] myriad threats and opportunities in all animal taxa across the social spectrum (from individuals to groups)" such as to substantially cause evolutionary changes in cognition (Ashton, Kennedy, & Radford, 2020). Indeed, the dynamics posited in Harpending's *starburst model* of frequent contact and conflict between different expansionistic human groups as a major feature of the species' evolutionary history (Harpending & Harris, 2016; see also Peñaherrera-Aguirre, Figueredo, & Hertler, 2020a, 2020b) may apply to high-*G* hominins generally—but with the addition of *heterospecific*, alongside outside conspecific, contact and conflict in the latter case. The 10,000-year-explosion model of human evolution in the Holocene, where occupation of novel niches massively accelerated genetic change via culture-gene coevolution, would in a sense apply to our own data, which indicate that multiple migration events both out of and into Africa over 10 million years facilitated rapid evolution of *G* through, especially, novel inter-group and inter-species contact and conflict, and admixture of hominins.

The work of Rushton and Rushton (2004) might suggest that the focal point of primate macroevolution is the brain, with changes in many other traits driven by neurological evolution. This seems consonant with our analysis finding that changes in *G* occur with speciation. A reasonable interpretation is that higher *G* enables species to persist in highly novel environments, which present these species with additional opportunities for further evolutionary change that might be sufficiently radical as to eventuate in speciation. But it might be argued, similarly to Rushton and Rushton (2004), that substantial increases in brain traits related to *G* will adaptively necessitate changes in various other phenotypes, such that large gains in *G* render speciation more or less inevitable. If, as we have argued, the evolution of higher *G* increases the trait's mutational target size, then the evolvability of *G* will also increase, because the probability that any beneficial mutation occurring in the genome will be one positively affecting *G* will rise as a higher proportion of the genome becomes related to *G*, which would enable more rapid speciation. Since these factors/processes are not incompatible, all may be involved to some extent, and help to make sense of the "burst" aspect of the Kappa model pattern, where speciation and changes in *G* happen surprisingly fast in evolutionary time.

The theory so far developed can further explain why *G* is also, perhaps seemingly paradoxically, associated with extinction risk. It might be expected that the process argued here to have given rise to high hominin *G* is *unstable* precisely because it depends on conditions related to the infiltration of novel ecologies and environments. Certain hard limits on any given species' adaptability—such as those imposed by its basic morphology and physiology—ensure that there is only a certain range of ecologies and environments into which it can expand, in which there will only be some competitor species. In the extreme case, a primate species that invades all environments it can adapt to and outcompetes all relevant species for niches will be unlikely to further experience the high evolutionary novelty that it encountered when first invading those territories and competing with other species. Such a primate species may have evolved very high and broad *G* to adapt to those earlier problems, but if such novel problems no longer occur or occur at a vastly reduced rate, *G* loses its fitness advantage insofar as the associated evolutionary novelty to contend with has greatly diminished. Under such conditions, the costliness of deleterious mutations affecting

[7] Koenigshofer (2017) has strongly challenged the idea that general intelligence is an adaptation to evolutionary novelty. His argument cannot be dealt with in detail here. In brief, the fundamental problem with it is that it relies on an erroneous conceptualization of evolutionary novelty, and offers an alternative model of the evolution of *g* that understands this trait as an adaptation to virtually *universal* features of environments and ecologies. This model, therefore, ultimately cannot adequately explain the patterns of enormous inter- and intra-species *variation* in *g*.

G could become severe enough to drive a species into extinction, especially in the case of a shrinking population vulnerable to an extinction vortex. This could neatly explain the *negative* effect of *G* on species persistence and therefore the disappearance of all high *G* hominins—*H. neanderthalensis*,[8] *H. erectus*, Denisovans, etc.—with the exception of *H. sapiens*, alongside the strong selection for and explosive evolution of *G* in hominin lines over the past 10 million years.

One alternative theory of the evolution of general cognitive ability that must be considered posits that high species *G*, such as in humans, is a result of sexual selection through mutual mate choice (e.g., Miller, 2000a, 2000b). Since sexual selection is reasonably understood as one form of social selection (Lyon & Montgomerie, 2012), the theory we have advanced in this Discussion is not incompatible with the notion that sexual selection played an important role in the evolution of *g*. On the other hand, there is a growing body of evidence that *g* is at best a weak criterion for mate choice in humans (at least for short-term mating), and perhaps is not a direct target at all (Driebe et al., 2021; Gignac, Darbyshire, & Ooi, 2018). Overall the relevant evidence seems to suggest that non-sexual social selection, related to in-group cooperation and inter-group conflict and competition, was a central factor driving the evolution of high cognitive ability in humans (Gintis et al., 2019; Woodley of Menie, Figueredo, et al., 2017), and by extension other high-*G* hominins: "The mating success of high cognition males was … grounded in their contribution to the mean fitness of band members, and hence in the long run, to the evolutionary success of ancestral humans. In a sense, hominins evolved to fill a *cognitive niche* that was relatively unexploited in the early Pleistocene" (Gintis et al., 2019). Sexual selection may nonetheless have had a complementary role, in amplifying, speeding up, and reinforcing other selection pressures that favored high *G* and high *g* in our ancestors.

Funding for this project

None.

Code availability

The R code will be made available upon request.

Ethics approval

Due to the archival nature of this study, using data taken from online sources, human subjects' approval from the Institutional Review Board at the University of Arizona was not needed.

Credit statement

MPA conceived the overall study and the data collection and analytic procedures; carried out the data collection and analyses; contributed to the refinement of the theory intended to explain the results; and contributed to writing and editing the manuscript. MAS developed the theory intended to explain the results, contributed to writing and editing the manuscript, and checked the analyses and results. MAW contributed to writing and editing the manuscript and the refinement of the theory intended to explain the results, and checked the analyses and results. GFM contributed to writing and editing the manuscript, and the refinement of the theory intended to explain the results.

[8] Consistent with this prediction, it has been claimed that the genome of *H. neanderthalensis*, relative to *H. sapiens*, contained far more deleterious mutations, and that the fitness of the former may have been 40% lower than that of the latter, contributing to the former's extinction (Harris & Nielsen, 2016).

Declaration of Competing Interest

None.

Data availability

The authors provided sources to the original data sources in the Methods section.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.intell.2023.101795.

References

Agustí, J., Cabrera, L., & Garcés, M. (2013). The Vallesian mammal turnover: A late Miocene record of decoupled land-ocean evolution. *Geobios, 46*(1–2), 151–157.
Alexander, R. D. (1989). Evolution of the human psyche. In P. Mellars, & C. Stringer (Eds.), *The human revolution: Behavioural and biological perspectives on the origins of modern humans* (pp. 455–513). Princeton: Princeton University Press.
Arden, R., & Zietsch, B. P. (2017). An all-positive correlation matrix is not evidence of domain-general intelligence. *Behavioral and Brain Sciences, 40*, Article e197.
Arslan, R. C., von Borell, C. J., Ostner, J., & Penke, L. (2017). Negative results are needed to show the specific value of a cultural explanation for *g*. *Behavioral and Brain Sciences, 40*, Article e198.
Ashton, B. J., Kennedy, P., & Radford, A. N. (2020). Interactions with conspecific outsiders as drivers of cognitive evolution. *Nature Communications, 11*, 4937.
Balzeau, A., Holloway, R. L., & Grimaud-Hervé, D. (2012). Variations and asymmetries in regional brain surface in the genus Homo. *Journal of Human Evolution, 62*, 696–706.
Barks, S. K., Parr, L. A., & Rilling, J. K. (2015). The default mode network in chimpanzees (*Pan troglodytes*) is similar to that of humans. *Cerebral Cortex, 25*, 538–544.
Begun, D. R. (2010). Miocene hominids and the origins of the African apes and humans. *Annual Review of Anthropology, 39*, 67–84.
Begun, D. R. (2013). The Miocene hominoid radiations. In D. R. Begun, & D. R. (Eds.), *A companion to paleoanthropology* (pp. 397–416). New York: John Wiley & Sons.
Beran, M. J., Gibson, K. R., & Rumbaugh, D. M. (1999). Predicting hominid intelligence from brain size. In M. C. Corballis, & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 88–97). Oxford: Oxford University Press.
Bruner, E. (2008). Comparing Endocranial form and shape differences in modern humans and Neandertals: A geometric approach. *PaleoAnthropology, 2008*, 93–106.
Bruner, E. (2010). Morphological differences in the parietal lobes within the human genus: A neurofunctional perspective. *Current Anthropology, 51*, S77–S88.
Bruner, E., & Holloway, R. L. (2010). A bivariate approach to the widening of the frontal lobes in the genus Homo. *Journal of Human Evolution, 58*, 138–146.
Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017a). The evolution of general intelligence. *Behavioral and Brain Sciences, 40*, Article e192.
Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017b). Future directions for studying the evolution of general intelligence. *Behavioral and Brain Sciences, 40*, Article e192.
Butler, M. A., & King, A. A. (2004). Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *American Naturalist, 164*, 683–695.
Cochran, G., & Harpending, H. (2009). *The 10,000 year explosion. How civilization accelerated human evolution*. New York: Basic Books.
Colom, R., Jung, R. E., & Haier, R. J. (2006). Distributed brain sites for the *g*-factor of intelligence. *Neuroimage, 31*, 1359–1365.
De Miguel, C., & Henneberg, M. (2001). Variation in hominid brain size: how much is due to method? *Homo, 52*, 3–58.
Deaner, R. O., Isler, K., Burkart, J., & van Schaik, C. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain, Behavior and Evolution, 70*, 115–124.
Deaner, R. O., van Schaik, C. P., & Johnson, V. (2006). Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evolutionary Psychology, 4*, 149–196.
Driebe, J. C., Sidari, M. J., Dufner, M., von der Heiden, J. M., Bürkner, P. C., Penke, L., … Arslan, R. C. (2021). Intelligence can be detected but is not found attractive in videos and live interactions. *Evolution and Human Behavior, 42*, 507–516.
Dunbar, R. I., & Shultz, S. (2023). Four errors and a fallacy: Pitfalls for the unwary in comparative brain analyses. *Biological Reviews, 98*, 1278–1309.
Eldredge, N., & Gould, S. J. (1972). Punctuated equilibria: An alternative to phyetic gradualism. In T. J. M. Schopf (Ed.), *Models in paleobiology* (pp. 82–115). San Francisco: Freeman Cooper.
Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk. *Human Nature, 20*, 204–268.
Erwin, D. H., & Anstey, R. L. (1995). *New approaches to speciation in the fossil record*. New York: Columbia University Press.
Eswaran, V., Harpending, H., & Rogers, A. R. (2005). Genomics refutes an exclusively African origin of humans. *Journal of Human Evolution, 49*, 1–18.
Evans, P. D., Mekel-Bobrov, N., Vallender, E. J., Hudson, R. R., & Lahn, B. T. (2006). Evidence that the adaptive allele of the brain size gene microcephalin introgressed

into *Homo sapiens* from an archaic Homo lineage. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 18178–18183.
Ezkurdia, I., Juan, D., Rodriguez, J. M., Frankish, A., Diekhans, M., Harrow, J., … Tress, M. L. (2014). Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Human Molecular Genetics, 23*, 5866–5878.
Fernandes, H. B. F. (2020). *Evolutionary trajectories of cognitive abilities and of their putative neuroanatomical and allometric correlates: Testing novel hypotheses of cognitive evolution and cognitive integration with phylogenetic comparative methods* (Unpublished doctoral dissertation). Tucson, AZ: The University of Arizona.
Fernandes, H. B. F., Peñaherrera-Aguirre, M., Woodley of Menie, M. A., & Figueredo, A. J. (2020). Macroevolutionary patterns and selection modes for general intelligence (*G*) and for commonly used neuroanatomical volume measures in primates. *Intelligence, 80*, 101456.
Fernandes, H. B. F., Woodley, M. A., & te Nijenhuis, J. (2014). Differences in cognitive abilities among primates are concentrated on G: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence, 46*, 311–322.
Frost, P. (2011). Human nature or human natures? *Futures, 43*, 740–748.
Garin, C. M., Hori, Y., Everling, S., Whitlow, C. T., Calabro, F. J., Luna, B., … Constantinidis, C. (2022). An evolutionary gap in primate default mode network organization. *Cell Reports, 39*(2).
Geary, D. C. (2005). *The origin of mind: Evolution of brain, cognition, and general intelligence*. Washington, D.C.: American Psychological Association.
Geary, D. C. (2009). The evolution of general fluid intelligence. In S. M. Platek, & T. K. Shakelford (Eds.), *Foundations in evolutionary cognitive neuroscience* (pp. 22–56). New York: Cambridge.
Gignac, G. E., Darbyshire, J., & Ooi, M. (2018). Some people are attracted sexually to intelligence: A psychometric evaluation of sapiosexuality. *Intelligence, 66*, 98–111.
Gilpin, M. E., & Soulé, M. E. (1986). Minimum viable populations: Processes of species extinction. In M. E. Soulé (Ed.), *Conservation biology: The science of scarcity and diversity* (pp. 19–34). Sunderland: Sinauer.
Gintis, H., van Schaik, C., & Boehm, C. (2019). Zoon politikon: The evolutionary origins of human socio-political systems. *Behavioural Processes, 161*, 17–30.
Goel, V., & Dolan, R. J. (2001). Functional neuroanatomy of three-term relational reasoning. *Neuropsychologia, 39*, 901–909.
Gorsuch, R. L. (2014). *Factor analysis: Classic* (2nd ed.). New York: Routledge.
Gould, S. J. (1996). *The mismeasure of man (revised and expanded)*. New York: W.W. Norton.
Gunz, P., Tilot, A. K., Wittfeld, K., Teumer, A., Shapland, C. Y., van Erp, T. G. M., … Fisher, S. E. (2019). Neanderthal introgression sheds light on modern human endocranial globularity. *Current Biology, 29*, 120–127.
Harpending, H., & Harris, N. (2016). Human kinship as a greenbeard. In J. Carroll, D. P. AcAdams, & E. O. Wilson (Eds.), *Darwin's bridge: Uniting the humanities and sciences* (pp. 55–68). New York: Oxford University Press.
Harris, K., & Nielsen, R. (2016). The genetic cost of Neanderthal introgression. *Genetics, 203*, 881–891.
Hawks, J., Wang, E. T., Cochran, G. M., Harpending, H. C., & Moyzis, R. K. (2007). Recent acceleration of human adaptive evolution. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 20753–20758.
Hill, W. D., Arslan, R. C., Xia, C., Luciano, M., Amador, C., Navarro, P., & Penke, L. (2018). Genomic analysis of family data reveals additional genetic effects on intelligence and personality. *Molecular Psychiatry, 23*, 2347–2362.
Huguet, G., Schramm, C., Douard, E., Tamer, P., Main, A., Monin, P., … Jacquemont, S. (2021). Genome-wide analysis of gene dosage in 24,092 individuals estimates that 10,000 genes modulate cognitive ability. *Molecular Psychiatry, 26*, 2663–2676.
Ishida, H., & Pickford, M. (1997). A new Late Miocene hominoid from Kenya: Samburupithecus kiptalami gen. et sp. nov. *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science, 325*, 823–829.
Isler, K., & van Schaik, C. P. (2009). Why are there so few smart mammals (but so many smart birds)? *Biology Letters, 5*, 125–129.
Isler, K., & van Schaik, C. P. (2012). How our ancestors broke through the gray ceiling: Comparative evidence for cooperative breeding in early homo. *Current Anthropology, 53*, S453–S465.
Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *The Behavioral and Brain Sciences, 30*, 135–187.
Kochiyama, T., Ogihara, N., Tanabe, H. C., Kondo, O., Amano, H., Hasegawa, K., … Akazawa, T. (2018). Reconstructing the Neanderthal brain using computational anatomy. *Scientific Reports, 8*, 6296.
Koenigshofer, K. A. (2017). General intelligence: Adaptation to evolutionarily familiar abstract relational invariants, not to environmental or evolutionary novelty. *Journal of Mind & Behavior, 38*, 119–154.
Koller, D., Wendt, F. R., Pathak, G. A., De Lillo, A., De Angelis, F., Cabrera-Mendoza, B., … Polimanti, R. (2022). Denisovan and Neanderthal archaic introgression differentially impacted the genetics of complex traits in modern populations. *BMC Biology, 20*, 249.
Kuijpers, Y., Domínguez-Andrés, J., Bakker, O. B., Gupta, M. K., Grasshoff, M., Xu, C. J., … Li, Y. (2022). Evolutionary trajectories of complex traits in European populations of modern humans. *Frontiers in Genetics, 699*.
Lari, M., Rizzi, E., Milani, L., Corti, G., Balsamo, C., Vai, S., … Caramelli, D. (2010). The *Microcephalin* ancestral allele in a Neanderthal individual. *PLoS One, 5*, Article e10648.
Legendre, P., & Oksanen, M. J. (2018). Package 'lmodel2'. Available at: https://CRAN.R-Project.Org/Package=Lmodel2.
Lewis, D. M. G., Al-Shawaf, L., & Anderson, M. (2017). Contemporary evolutionary psychology and the evolution of intelligence. *Behavioral and Brain Sciences, 40*, Article e210.
Lindenfors, P., Wartel, A., & Lind, J. (2021). Dunbar's number'deconstructed. *Biology Letters, 17*, 20210158.
Lyon, B. E., & Montgomerie, R. (2012). Sexual selection is a form of social selection. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*, 2266–2273.
Mayr, E. (1954). Change of genetic environment and evolution. In J. Huxley, A. C. Hardy, & E. B. Ford (Eds.), *Evolution as a process* (pp. 157–180). London: Allen & Unwin, London.
Melchionna, M., Mondanaro, A., Serio, C., Castiglione, S., Di Febbraro, M., Rook, L., … Raia, P. (2020). Macroevolutionary trends of brain mass in Primates. *Biological Journal of the Linnean Society, 129*, 14–25.
Miller, G. F. (2000a). Sexual selection for indicators of intelligence. In G. R. Bock, J. A. Goode, & K. Webb (Eds.), *The nature of intelligence. Novartis Foundation symposium 233* (pp. 260–275). Chichester: Wiley Ltd.
Miller, G. F. (2000b). Mental traits as fitness indicators: Expanding evolutionary psychology's adaptationism. *Annals of the New York Academy of Sciences, 907*, 62–74.
Miller, G. F., & Penke, L. (2007). The evolution of human intelligence and the coefficient of additive genetic variance in human brain size. *Intelligence, 35*, 97–114.
Nunn, C. L. (2011). *The comparative approach in evolutionary anthropology and biology*. Chicago: University of Chicago Press.
Nunn, C. L., & van Schaik, C. P. (2002). A comparative approach to reconstructing the socioecology of extinct primates. In J. M. Plavcan, R. F. Kay, W. L. Jungers, & C. P. van Schaik (Eds.), *Reconstructing behavior in the primate fossil record. Advances in primatology* (pp. 159–215). Boston: Springer.
Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., & Pearse, W. (2013). *The caper package: comparative analysis of phylogenetics and evolution in R. R package version, 5*(2), 1–36.
Pagel, M. (1999). The maximum likelihood approach to reconstructing ancestral character states of discrete characters of phylogenies. *Systematic Biology, 48*, 612–622.
Parker, S. T., & Mckinney, M. L. (1999). *Origins of intelligence: The evolution of cognitive development in monkeys, apes, and humans*. Baltimore: Johns Hopkins University Press.
Peñaherrera-Aguirre, M., Fernandes, H. B., & Figueredo, A. J. (2021a). Nonhuman primates: Within-group conflicts. In T. K. Shackelford, & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of evolutionary psychological science*. Cham: Springer.
Peñaherrera-Aguirre, M., Fernandes, H. B., & Figueredo, A. J. (2021b). Nonhuman primates: Between-group conflicts. In T. K. Shackelford, & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of evolutionary psychological science*. Cham: Springer.
Peñaherrera-Aguirre, M., Figueredo, A. J., & Hertler, S. C. (2020a). Chimpanzee intercommunity conflict: Fitness outcomes, power imbalances, and multilevel selection. In S. C. Hertler, A. J. Figueredo, & M. Peñaherrera-Aguirre (Eds.), *Multilevel selection: Theoretical foundations, historical examples, and empirical evidence* (pp. 225–249). New York, NY: Palgrave Macmillan.
Peñaherrera-Aguirre, M., Figueredo, A. J., & Hertler, S. C. (2020b). Lethal intergroup competition in non-state societies: From small-scale raids to large-scale battles. In S. C. Hertler, A. J. Figueredo, & M. Peñaherrera-Aguirre (Eds.), *Multilevel selection: Theoretical foundations, historical examples, and empirical evidence* (pp. 251–273). Palgrave Macmillan.
Pennell, M. W., Eastmen, J. M., Slater, G., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., … Harmon, L. J. (2014). Geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics, 30*, 2216–2218.
Perote-Peña, J. (May 2019). Selection pressures against domain-specific modules in human brains. In *Oral presentation given at the 31st annual meeting of the Human Behavior and Evolution Society, Boston*.
Prothero, D. R. (2006). *After the dinosaurs: The age of mammals*. Indiana: Indiana University Press.
Rabosky, D. L., & Goldberg, E. E. (2015). Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology, 64*, 340–355.
Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 366*, 1017–1027.
Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences, 99*, 4436–4441.
Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution, 3*, 217–223.
Riddle, W. I., & Corl, K. G. (1977). Comparative investigation of the relationship between cerebral indices and learning abilities. *Brain, Behavior and Evolution, 14*, 385–398.
Ross, C. F., Lockwood, C. A., Fleagle, J. G., & Jungers, W. L. (2002). Adaptation and behavior in the primate fossil record. In J. M. Plavcan, R. F. Kay, W. L. Jungers, & C. P. van Schaik (Eds.), *Reconstructing behavior in the primate fossil record. Advances in primatology* (pp. 1–41). Boston: Springer.
Ruff, C. B., Trinkaus, E., & Holliday, T. W. (1997). Body mass and encephalization in Pleistocene Homo. *Nature, 387*, 173–176.
Rumbaugh, D. M. (1970). Learning skills of anthropoids. In L. Rosenblum (Ed.), *Primate behavior: Developments in field and laboratory research* (pp. 1–70). New York: Academic Press.
Rushton, J. P., & Rushton, E. W. (2004). Progressive changes in brain size and musculo-skeletal traits in seven hominoid populations. *Human Evolution, 19*, 173–196.
Sansalone, G., Profico, A., Wroe, S., Allen, K., Ledogar, J., Ledogar, S., … Raia, P. (2023). Homo sapiens and Neanderthals share high cerebral cortex integration into adulthood. *Nature Ecology & Evolution, 7*(1), 42–50.
Shermer, M., & Sulloway, F. J. (2000). The grand old man of evolution: An interview with evolutionary biologist Ernst Mayr. *Skeptic, 8*, 76–82.

Sol, D., Bacher, S., Reader, S. M., & Lefebvre, L. (2008). Brain size predicts the success of mammal species introduced into novel environments. *The American Naturalist, 172*, S63–S71.
Stibel, J. M. (2021). Decreases in brain size and encephalization in anatomically modern humans. *Brain, Behavior and Evolution, 96*, 64–77.
Stibel, J. M. (2023). Climate Change Influences Brain Size in Humans. *Brain Behavior and Evolution, 98*, 93–106.
Suwa, G., Kono, R. T., Katoh, S., Asfaw, B., & Beyene, Y. (2007). A new species of great ape from the late Miocene epoch in Ethiopia. *Nature, 448*, 921–924.
Tomasello, M. (2010). *Origins of human communication*. Cambridge: MIT press.
Tran, T.-D., & Kwon, Y.-K. (2013). The relationship between modularity and robustness in signalling networks. *Interface, 10*(88).
Wei, Y., de Lange, S. C., Scholtens, L. H., Watanabe, K., Ardesch, D. J., Jansen, P. R., … van den Heuvel, M. P. (2019). Genetic mapping and evolutionary analysis of human-expanded cognitive networks. *Nature Communications, 10*, 4839.
Witmer, L. M. (1995). The extant phylogenetic bracket and the importance of reconstructing soft tissues in fossils. In J. J. Thomason (Ed.), *Functional morphology in vertebrate paleontology* (pp. 19–33). New York: Cambridge University Press.
Witmer, L. M. (1997). The evolution of the antorbital cavity of archosaurs: A study in soft-tissue reconstruction in the fossil record with an analysis of the function of pneumaticity. *Journal of Vertebrate Paleontology, 17*, 1–76.
Woodley of Menie, M. A., Fernandes, H. B. F., te Nijenhuis, J., Peñaherrera-Aguirre, M., & Figueredo, A. J. (2017). General intelligence is a source of individual differences *between* species: Solving an anomaly. *Behavioral and Brain Sciences, 40*, Article e192.
Woodley of Menie, M. A., Figueredo, A. J., Sarraf, M. A., Hertler, S., Fernandes, H. B. F., & Peñaherrera-Aguirre, M. (2017). *The rhythm of the west: A biohistory of the modern era, AD 1600 to present. Journal of social, political and economic studies. Monograph series 37*. Washington, DC: Council for Social and Economic Studies.
Woodley of Menie, M. A., Luoto, S., Peñaherrera-Aguirre, M., & Sarraf, M. A. (2021). Life history is a major source of adaptive individual and species differences: A critical commentary on Zietsch and Sidari (2020). *Evolutionary Psychological Science, 7*, 213–231.
Woodley of Menie, M. A., & Peñaherrera-Aguirre, M. (2022). General intelligence as a major source of cognitive variation among individuals of three species of lemur, uniting *g* with *G*. *Evolutionary Psychological Science, 8*, 241–453.
Woodley of Menie, M. A., & Peñaherrera-Aguirre, M. (2023). Convergence between *G* and *g* in three monkey species (*Sapajus* spp, *Ateles geoffroyi*, and *Macaca fascicularis*). *Journal of Comparative Psychology, 137*, 62–73.
Woodley of Menie, M. A., Peñaherrera-Aguirre, M., & Jurgensen, J. (2022). Using macroevolutionary patterns to distinguish primary from secondary cognitive modules in primate cross-species performance data on five cognitive ability measures. *Intelligence, 92*, 101645.
Woodley of Menie, M. A., Younuskunja, S., Balan, B., & Piffer, D. (2017). Holocene selection for variants associated with cognitive ability: Comparing ancient and modern genomes. *Twin Research and Human Genetics, 20*, 271–280.
Wrangham, R. (2017). Control of fire in the Paleolithic: Evaluating the cooking hypothesis. *Current Anthropology, 58*, S303–S313.
Wynn, T. (1989). *The evolution of spatial competence*. Urbana: University of Illinois Press.
Wynn, T. (2002). Archeology and cognitive evolution. *Behavioral and Brain Sciences, 25*, 389–438.