A phylogenetic protein-coding genome-phenome map of complex traits across 224 primate species

Authors: Alejandro Valenzuela[1,2,*] , Fabio Barteri[1,4,9,*] , Lukas Kuderna[10] , Joseph Orkin[11] , Jean Boubli[12] , Amanda Melin[13] , Hafid Laayouni[1] , Kyle Farh[10] , Jeffrey Rogers[14] , Tomàs Marquès-Bonet[1] , Gerard Muntaner[5]† , David de Juan Sopeña[15]† , Arcadi Navarro[1,2,3,5,6,7,8,9,†]

Affiliations:

* Co-first authors

† Corresponding authors: Gerard Muntaner (gerard.muntane@upf.edu), David de Juan Sopeña (da.juan@cnb.csic.es), Arcadi Navarro (arcadi.navarro@upf.edu)

1. Institute of Evolutionary Biology (IBE, UPF–CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona (PRBB), C. Doctor Aiguader 88, 08003 Barcelona, Spain

2. Institució Catalana de Recerca i Estudis Avançats (ICREA) and Universitat Pompeu Fabra, Pg. Lluís Companys 23, 08010 Barcelona, Spain

3. Center for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Av. Doctor Aiguader 88, 08003 Barcelona, Spain

4. BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation, C. Wellington 30, 08005 Barcelona, Spain

5. Hospital Universitari Institut Pere Mata, Institut d'Investigació Sanitària Pere Virgili (IISPV) and Universitat Rovira i Virgili, Reus, Spain

6. Centro de Investigación Biomédica en Red en Salud Mental (CIBERSAM), Spain

7. Genomes for Life – GCAT Lab, Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain

8. European Molecular Biology Laboratory (EMBL), Meyerhofstraße 1, 69117 Heidelberg, Germany

9. Museu de Ciències Naturals de Barcelona, Barcelona, Spain

10. Illumina, San Diego, CA, USA

11. Université de Montréal, Montreal, Canada

12. University of Salford, Salford, UK

13. University of Calgary, Calgary, Canada

14. Baylor College of Medicine, Houston, TX, USA

15. Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

Abstract

Leveraging 60 million years of primate evolutionary diversity, we conducted a large-scale cross-species genome-phenome analysis of protein-coding genes across 224 primate species and 263 phenotypic traits spanning five domains: behavior, ecology, life history, morphology, and physiology. Using phylogenetic comparative methods, we generated the Primate Protein-coding Genome-Phenome Map (P3GMap), available in the Primate Genome-phenome Archive (PGA https://pgarchive.github.io), identifying over 29,000 convergent amino acid substitutions (CAASs) and 3,900 gene-trait associations. As case studies, we applied targeted analyses to insectivorous diet, white blood cell count, and maximum lifespan, uncovering lineage-specific adaptations and signatures of directional selection. We show that traits undergoing greater phylogenetic divergence were linked to more extensive genetic change, and that most CAASs were fixed in humans, suggesting a decoupling of macro- and microevolutionary processes. Notable specific findings include associations between mitochondrial metabolism genes and body mass, immune-related loci and lactation, and olfactory and xenobiotic-processing genes with white blood cell counts. Together, our results establish a scalable framework to explore the molecular basis of complex traits across evolutionary time and demonstrate the value of integrating intra- and interspecies data to understand the genetic foundations of human biology.

One-Sentence Summary: We conducted a comprehensive cross-species genome-phenome analysis of protein-coding genes across 224 primate species, uncovering thousands of candidate causal amino-acid changes linked with 263 phenotypes and demonstrating the value of integrating evolutionary perspectives to understand the genetic architecture of complex human traits.

Introduction.

For decades, our understanding of the genetic basis of complex traits has been largely derived from both experimentation with model organisms (1–3) and analyses of genomic and phenotypic variation within populations, primarily through GWAS in humans (4–7). Although these intra-species approaches are powerful and successful, they are limited by the variation present in extant populations, and so they cannot detect genetic variants fixed in the lineage of the species under study (8,9). In contrast, complementary approaches like Phylogenetic Comparative Methods (PCMs) (9–12), can help unveiling phenotypic associations across a broad spectrum of genomic variations accumulated over millions of years, including gene gain and loss, protein evolution rates, structural variations, and point mutations (13–16). Even when applied to a limited set of traits or focused on candidate genes and episodes of positive or purifying selection (17–22), *PCMs* have advanced our understanding of the genetic architecture of key traits in human evolution such as brain size (17,23,24) and maximum lifespan (25–28), to mention only two.

Scaling up PCMs, we introduce the Primate Genome-Phenome Archive (PGA), a comprehensive resource containing high-quality phenotypic data for 225 traits across 224 sequenced primate species (29), aimed to elucidate the genetic underpinnings of complex traits. Applying different PCMs to the PGA, our analysis identified thousands of genome-phenome relationships in protein-coding genes, providing unprecedented insights into the evolutionary dynamics and genetic underpinnings of many complex traits [...].

The Primate Phenomic Space

We selected representative trait values for 225 non-redundant primate species- from a pool of 263 high-quality quantitative and qualitative measurements curated from over 800 phenotypes available from public sources (30–33). We categorized these 225 traits into five primary biological domains: behavior (29 traits), ecology (7 traits), life history (29 traits), morphology (93 traits), and physiology (67 traits), and 15 secondary biological domains (See Fig. 1, Methods and Table S1) that constitute the the Primate Phenomic Space.

This phenotypic dataset was compiled from diverse studies with varying sampling depths and strategies, which resulted in heterogeneous data densities across species (see also Supplementary Material). To characterize the sampling structure and facilitate comparisons of traits with similar data coverage across clades, we clustered them into 13 groups highlighting differences in phylogenetic sampling (Figures S2-S22, see also SM). Some traits in clusters I and II, such as *Median Longevity* (N=24) and *Maximum Litter Size* (N=24), were more thoroughly sampled within *strepsirrhines*. Notably, *Strepsirrhine* families, such as *Lemuridae* and *Indriidae*, contributed, on average, more than 90% of the traits across all primary biological domains. In contrast, traits such as *Body Mass* (N=214), *Locomotion* (N=220), and *Diel Activity* (N=220) were densely represented across the phylogeny (clusters XI-XIII), including in under-studied families such as *Lepilemuridae* and *Tarsiidae*, which contributed data for fewer than 50% of traits on average. Mass-related traits were among the best represented in the most densely

populated clusters, explaining XX% of the variance in the first component of the Factor Analysis of Mixed Data (FAMD) for clusters VIII, IX, and X (Figure 1A).

Influence of allometry on trait variation is a well established phenomenon (34–38), thus we controlled allometric effects in 74 morphological, 14 life history, and two physiology traits known to scale with body size (Data S1). This resulted in a curated phenotypic dataset that is made publicly available to support further research on primate phenotypic evolution (https://pgarchive.github.io).

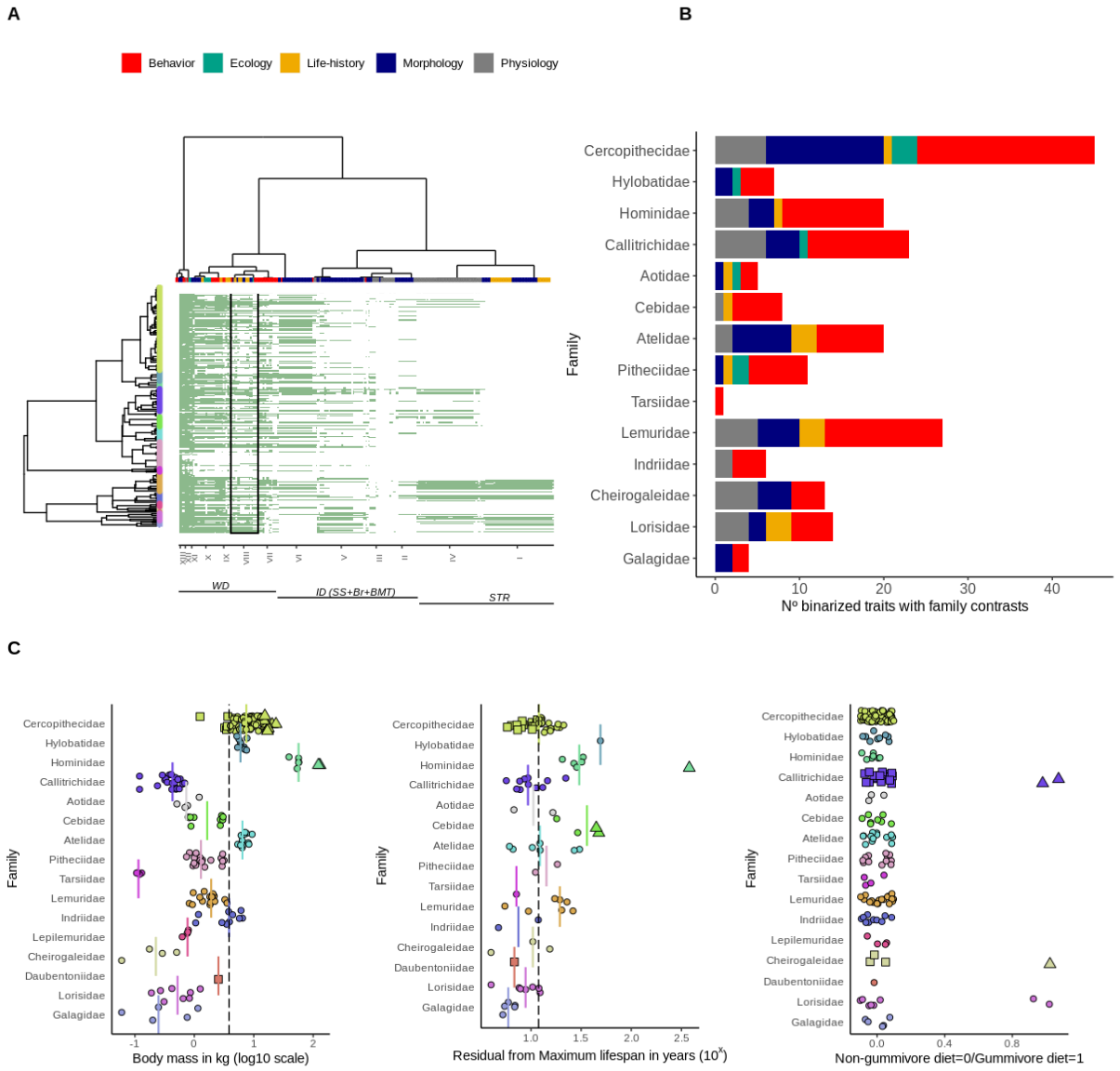


Figure 1. Overview of the *Primate Traits Database (PDB)* (A) Hierarchical clustering of phenomic data, mapped onto the primate phylogeny and grouped by primate family. Based on data availability, the traits were classified into 13 data clusters [see SM]~~, categorized as *WD* (widely distributed), *ID* (intermediately distributed), and *STR* (strepsirrhine-focused)~~. The area within the square highlights traits from one of the clusters, *Cluster VIII,* which includes a broad range of ecological, life history, and morphological traits. **(B)** Total number of traits (x-axis)

showing extreme phenotypic differences between species within the same family. Top and Bottom values refer to species with contrasting trait values (see text for definitions). **(C)** Examples illustrating the distribution of values for three traits: *Body Mass (left)*, *Maximum Lifespan (center)*, and *Gummivore, sap-eating, Diet (right)*. Each point represents a species. The vertical bars indicate median family values. Dashed lines correspond to the median global trait value. Species selected as *Top* and Bottom extremes (based on criteria explained in the text) are indicated by triangles and squares, respectively.

The Phylogenetic Protein-coding Genome-Phenome Map (P3GMap)

We conducted a genome-phenome analysis of protein-coding sequences to build the first phylogenetic protein-coding genome-phenome map (P3GMap) to be uploaded in the PGA. This resource catalogs the phylogenetic associations between protein-coding genes and a broad array of phenotypes across the primate lineage.

The P3GMap consists of multiple sequence alignments for 16,133 one-to-one orthologous protein-coding genes across the primate phylogeny. Over 75% of the alignments included more than 200 species (see Figure S1 and SM). To identify genomic factors associated with phenotype evolution, we applied two complementary strategies. First, we identified Convergent Amino Acid Substitutions (CAAS) between species with divergent phenotypes using CAASTools (40, see SM "CAAStools") . Second, we linked gene-specific relative evolutionary rates (RERs) to phenotypic rate of change across the phylogeny, using RERconverge (39) (see SM "RERconverge"). These methods capture genomic-phenotypic associations across a broad range of evolutionary scenarios bypassing the need to detect classical signatures of strong positive selection [REF paper Guojie].

To enable systematic cross-trait comparisons, we developed a *standard strategy* that prioritizes consistency and scalability across hundreds of traits, rather than tailoring parameters to optimize fine-tuned detection for individual traits or groups of traits. We must note that the P3GMap focuses only on single-copy protein-coding genes, and so it is not intended to fully resolve the genetic architecture of each trait. However, it offers for the first time a comprehensive panoramic view of the genome-phenome relationship in primates, and provides a framework for identifying genetic signatures underlying key evolutionary transitions.

Convergent amino acid substitutions linked to primate traits

For each trait we identified CAASs between groups of species with extremely different quantitative phenotypic values (or contrasting qualitative values). To do so, for each phenotype we classified species with high trait values (or arbitrary qualitative states) as *Top* species, and species with low values (or complementary qualitative states) as *Bottom* species (See Fig. 1B-C, Figure S24-S27, Data S1, and SM sections A.1-A.4). These pairwise comparisons are named *Contrasts*. The remaining species, which were not classified as *Top* or *Bottom*, were not used in the first phase of the analysis. For the *standard strategy* we focused on *contrasts* within the same family. This design increases the likelihood that phenotypic differences observed in different families represent independent cases of convergence.

Among our set of traits, 83 (37%) exhibited at least one intra-family phenotypic *contrast*, out of which 63 showed *contrasts* in two or more families, unveiling a considerable amount of

independent convergent evolution in primate traits. We applied CAAStools on these 83 traits and controlled for phylogenetic effects using a permulations (41). Then, we performed a per-trait multiple testing correction (see SM A.4). Altogether, 29,155 unique significant CAASs were identified in 8,780 genes across 76 traits ($p_{\text{permul-FDR}} < 0.05$), corresponding to a median of 49 genes and 51 positions across the primary biological domains (Fig. 2A, Figure S28-S31 and Data S2). To our knowledge, this represents the largest set of candidate causal amino acid substitutions ever identified in association with such a comprehensive and diverse array of phenotypic traits, providing an unprecedented resource for the study of convergent molecular evolution.

To evaluate the consistency of these associations, we validated the recovered CAASs in the species not included in the original groups (the *Top* and *Bottom* groups) by means of phylogenetic generalized least squares (PGLS) for quantitative traits and phylogenetic logistic regression (phyloglm) for qualitative traits (See SM Internal phylogenetic test in primates) (Fig. 2C and Figure S39). This a highly conservative sanity check, since power was drastically reduced by removing species with extreme values and since intermediate species do not necessarily share the *Top* or *Bottom* amino-acid residues. We obtained a median of 7,14% nominally significant positions ¿per trait? (Figure S40-S44 and Data S2).

For a more targeted investigation, we focused on 10 traits with <100 validated and significant phylogeny-wide associations ($p_{PGLS\text{-}FDR} < 0.05$, and $p_{phyloglm\text{-}FDR} < 0.05$). These analyses revealed both, previously known and novel gene-trait associations (Suppl. Table S3). Notable examples include associations between *Body Mass* and mitochondrial metabolism genes (50,51): *NDUFS7* (NADH:ubiquinone oxidoreductase core subunit S7) with substitutions Y223L ($p_{PGLS} = 1.37 \times 10^{-5}$), R230H ($p_{PGLS} = 3.57 \times 10^{-5}$), K231G ($p_{PGLS} = 1.37 \times 10^{-5}$), and I232A ($p_{PGLS} = 1.37 \times 10^{-5}$) and *NOXO1* T/V223A (NADPH Oxidase Organizer 1) ($p_{PGLS} = 2.06 \times 10^{-7}$); as well as genes in the immunoglobulin protein superfamily (52), including *ICAM1* R/V418M ($p_{PGLS} = 2.39 \times 10^{-5}$). For *Lactation Period Length*, we found associations with genes linked to familial Mediterranean fever (FMF)(53–55), such as *MEFV* K/R546Q ($p_{PGLS} = 1.66 \times 10^{-6}$), a key regulator of innate immunity.

To further validate our results, we gathered a set of 10 equivalent species-specific trait values from non primate mammals and tested whether the XX CAASs presented a consistent and significant direction of the effects. This test is even more conservative, since changes in the trait in distant species are not likely to be mediated by the same amino-acid change. Still, the test yielded a median of 7,28% nominally significant positions ($p_{PGLS} < 0.05$) (Figure S45-S47). Eight CAAS remained significant after FDR correction ($p_{PGLS\text{-FDR}} < 0.05$) (See also Supp. Table S4). Notable examples include position Q176RN *in ELL1* for *Lactation Period Length* ($p_{PGLS} = 1.09 \times 10^{-4}$), previously linked to late milk production in dairy goats (56) and to bone formation, regeneration, and craniosynostosis (57–59), and *MPDZ* H/P956Q for *Gestation Period Length* ($p_{PGLS} = 1.65 \times 10^{-5}$), a gene with germline mutations associated with congenital non-communicating hydrocephalus (60,61).

Correlations between rates of gene evolution and traits

While the CAAS approach can detect genome-phenome associations regardless of the underlying evolutionary mechanism, complementary insights can be gained by identifying genes whose rates of evolution correlate with trait changes across the phylogeny. Such associations may reflect episodes of accelerated or constrained protein evolution (potential signatures of positive or purifying selection). To carry-out such analysis, we used RERconverge in 16,109 alignments of orthologous protein-coding genes and 200 quantitative traits (See also SM: RERconverge). Gene-trait correlation significance was assessed via permulations (41) (see also SM B.2). Five qualitative traits (*Extinction Risk*, *Natal Dispersal, Female Ranging Patterns*, *Infanticide* and *Stable Groups*) were excluded because of the high computational cost of permulations (see Supplementary Materials and Methods (SM)).

We identified significant gene-trait correlations involving 3,911 genes and 194 traits ($p_{\text{permul-FDR}} < 0.05$), with a median of 19 associated genes per trait (Fig. 2A, Figure S28-S31 and data S2). RERconverge applies a more permissive correction than CAASTools (see SM), which may lead to inflated significance. To address this, in the selection of specific gene-trait associations, we additionally required significance based on FDR correction of alternative parametric p-values (p_{corr}). RERconverge provides p_{corr} as a high-resolution, computationally efficient, but phylogeny-agnostic alternative to $p_{\text{permul,}}$, with both statistics showing similar values across our dataset (see SM).

RERconverge allowed identifying genes with functions linked to the associated traits (Fig. 2B and Fig, S38; see also SM RERconverge B.1 and Suppl. Table S2). Examples include *NKX6-2* for *Percentage of Time Moving* (*Rho* = 0.62; $p_{\text{permul-FDR}} < 10^{-3}$), implicated in spastic ataxia in humans (42,43), *DEFA5* for *Percentage of Fauna Ingest* (*Rho* = -0.5, $p_{\text{permul-FDR}} < 10^{-3}$), a defensin family gene associated with gastritis (44), and *DESI2* for *Mesencephalon Relative Size* (*Rho* = 0.69, $p_{\text{permul-FDR}} < 10^{-3}$), linked to amyotrophic lateral sclerosis type 5 (47).

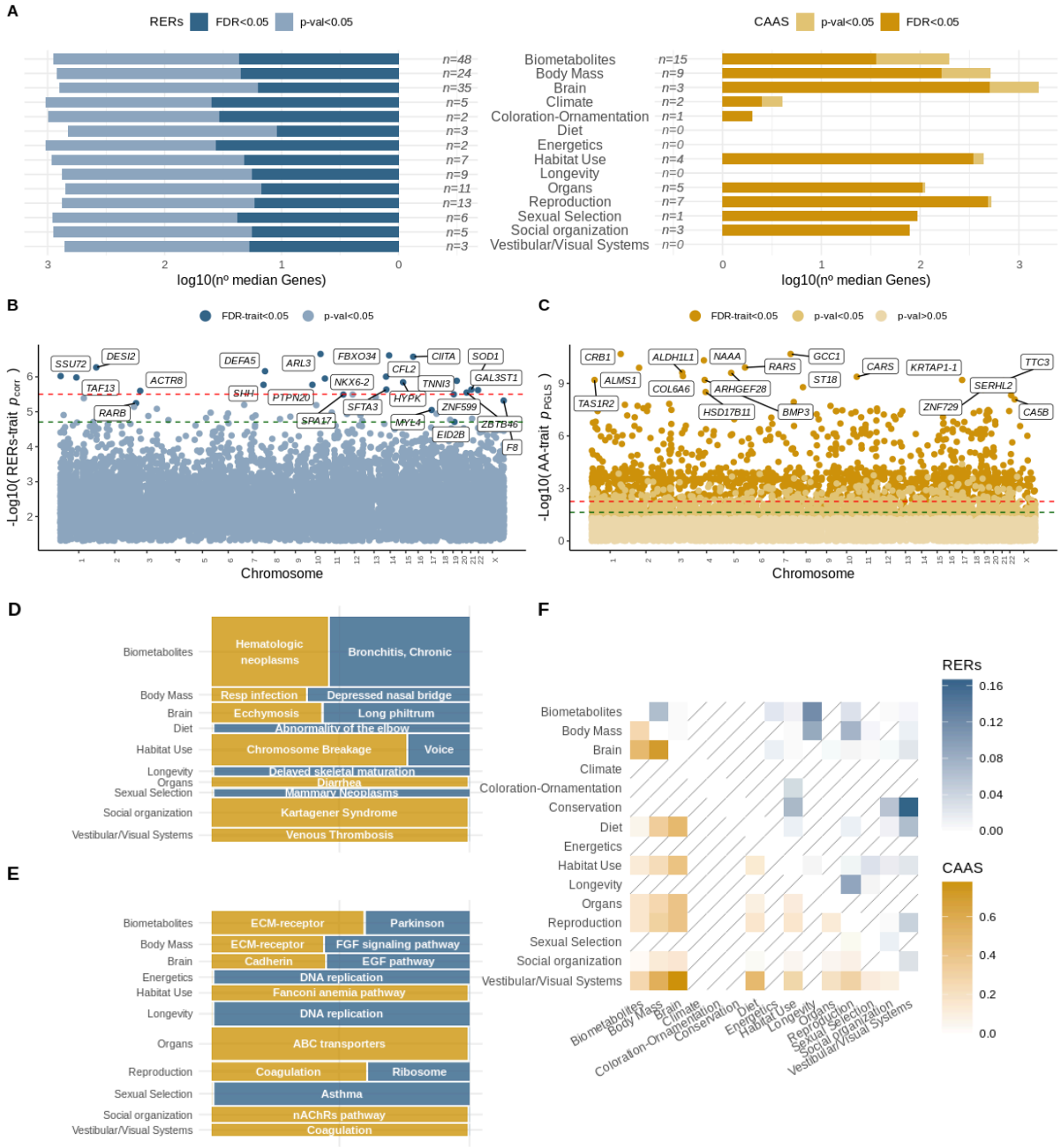


Figure 2. Phylogenetic Protein-coding Genome-Phenome Map (P3GMap) exploration. (A) Median number of genes after multiple-gene test correction ($p_{\text{Permul-FDR}} < 0.05$) for permulation of each quantitative trait, reporting the log10 scale of the median gene count on RERs (left) and CAAS (right) for each secondary biological domain. The *n* counts (middle) indicate the number of traits with associations per secondary biological domain. **(B and C)** Manhattan plot with correlation p-values for RERs (left) and PGLS p-values for CAAS (right) showing the lowest quantitative RERs-trait p-value and lowest AA-trait p-value for each gene, respectively. The dashed green line corresponds to the FDR threshold and the dashed red line corresponds to the strict Bonferroni threshold. Genes labeled with their corresponding gene names for RERs ($p_{\text{bonf-trait}} < 0.05$) and CAAS ($p_{\text{PGLS-trait}} < 1e\text{-}8$). **D)** Heatmap with the proportion of significant pairs of traits ($p_{\text{FDR}} < 0.05$) obtained from the hypergeometric test of gene set overlap between secondary biological domains. Boxes with strip lines represent pairs of secondary biological domains with no significant gene set overlap for each PCM methodology.

Our analysis also uncovered hundreds of novel gene-trait associations potentially relevant to both biomedicine and evolutionary biology (Fig. 2). Notably, traits that exhibited greater evolutionary divergence across the primate phylogeny were more likely to be linked across with a higher number of genomic associations, consistent with the expectation that extensive phenotypic change is often accompanied by widespread molecular adaptation. These findings reinforce the utility of phylogenetic rate analyses for capturing evolutionary signatures that may escape detection in traditional within-species approaches.

A first functional exploration analysis of the standard P3GMap

We evaluated the biological terms linked to our results using over-representation analysis (ORA) (Suppl. Tables XXX). For significant CAAS genes, we identified 199 unique enriched terms across the 27 traits analyzed (Suppl. Data S3). Body mass-related traits were enriched with DNA repair-related categories (Fig. XX or Suppl. Table XX), consistent with recent findings for the order *Carnivora* (87). Additionally, DNA repair genes have been highlighted in several studies as targets of positive selection for anti-cancer mechanisms in long-lived species (28,48,88–90). Two other notable cases involved the trait *Percentage of Time with Social Activity*, with associated genes enriched with the terms *chromosome breakage* (*ER* = 7.21; $p = 4.53 \times 10^{-7}$) and *Fanconi anemia* (ER = 4.73; $p = 3.03 \times 10^{-5}$). These genes are associated with developmental and congenital disorders (91,92).

Focusing on significant genes from RERconverge (P<0.05), we found 94 functionally enriched terms across XXX traits (*p*FDR < 0.05; hypergeometric test). Among these, we identified enrichments of the term *Long philtrum* (defined as a distance >2SD between the nasal base and the midline of the upper lip), associated with the relative sizes of the brain regions *Pallidum* (ER = 3.35; $p = 2.97\times10^{-6}$) and *Subthalamus* (ER = 3.34; $p = 3.12\times10^{-6}$). These results highlight nose morphology as a key phenotypic distinction among primates, reflecting one of the most characteristic differences between *strepsirrhines* and *haplorhines* (62,63), and a trait that has been implicated in numerous human syndromes involving cognitive, sensory, and communication disabilities, as well as facial abnormalities (64–68). Two additional significant enrichments were associated with traits under sexual selection. *Hair Dimorphism* was linked with the MHC protein complex ($p = 7.76\times10^{-6}$), a relationship proposed in several studies as relevant for mate choice in mammals, birds, and fish (73–77), although some studies have disputed this association in specific taxa (78–82). Likewise, *Mass Dimorphism* was associated with *mammary neoplasms* (*ER* = 2.16; $p = 4.65 \times 10^{-6}$), which is supported by studies linking sex-based dimorphisms to differential cancer incidence in humans (83–85).

Functional interpretation of the gene sets obtained for each biological domain may also be enhanced by examining their relationships with other biological categories (97–99). Gene set overlap was evaluated through hypergeometric tests between all pairs of traits across the primary and secondary biological domains (Fig. 2D). First, focusing on gene sets containing significant CAAS-trait associations, we obtained gene overlaps for 16 trait pairs ($p < 1 \times 10^{-50}$),. These overlaps were concentrated in *Trichromatic Color Vision* and brain and body mass domains, suggesting that gene sets involved in sensory perception and morphological development may contribute pleiotropically to multiple complex traits across evolutionary contexts. In addition, we found five overlaps involving life-history traits above this threshold: *Age of Female At First*

Reproduction with both *Trichromatic Color Vision* and with *Maximum Young Adult Weight*; *Maximum Possible Reproductive Success Under Optimal Conditions* with *Trichromatic Color Vision*; and *Gestation Period Length* with both *Maximum Young Adult Weight* and *Maximum Young Adult Weight in Males*. Several of these reproductive traits are shared among families that have undergone recent phenotypic changes in color vision, particularly within the *Atelidae* family. Notably, genera within the *Atelinae* subfamily, such as *Ateles* and *Lagotricha*, exhibit polymorphic color vision, whereas the *genus Alouatta* displays routine trichromatic vision (117). These findings suggest that recent evolutionary shifts in visual perception may have co-occurred with changes in reproductive and somatic traits, pointing to coordinated molecular adaptations linked to ecological or social pressures.

Second, focusing on the RERconverge results, we identified robust gene overlaps for 11 trait pair combinations ($p < 10^{-30}$; hypergeometric test). Notably, four of these associations involved longevity-related traits. For example, genes associated with *Percentage of Infant Mortality* overlapped with *Maximum Male Longevity* and *Gamma-Glutamyl Transferase (GGT) levels* ($p = 4{,}46 \times 10^{-74}$ and $p = 7{,}13 \times 10^{-39}$, respectively). *GGT* is an important antioxidant enzyme involved in glutathione metabolism, previously associated with cardiovascular and neurological diseases (100–106), lung inflammation (107), and various cancers (108–110) in humans. It has also been proposed as a predictor of mortality in the general population (111). We also found a robust significant overlap between genes linked to *Cholesterol Levels* and *Median Male Longevity* ($p = 6{,}74 \times 10^{-48}$), although the relationship between cholesterol, lipid metabolites, and human mortality remains complex and sometimes contradictory (112). Additionally, genes associated with *Serum Oxaloacetic Transaminase Levels*, also known as aspartate transaminase (AST), overlapped significantly with those linked to *Median Female Longevity* ($p = 2{,}18 \times 10^{-38}$). Elevated AST levels have been associated with mortality in patients with liver, neurological, and cardiovascular diseases (113–116).

Focusing on traits associated with specific primate families provides an opportunity to detect lineage-specific evolutionary patterns that may be masked in broader analyses. Such comparisons can reveal coordinated molecular adaptations shaped by shared ecological niches, social structures, or historical constraints within particular clades. In this context, we found that *Cathemeral Activity* shared enriched gene sets with both *Dichromatic Color Vision* ($p = 7{,}72 \times 10^{-49}$) and *Seasonal Breeding* ($p = 2{,}34 \times 10^{-200}$) associations mainly driven by members of the *Lemuridae* family. Similarly, *Solitary Social System* showed robust overlap with genes linked to *Monochromatic Color Vision* ($p = 2{,}98 \times 10^{-41}$), involving *Lorisidae* and *Galagidae*, and with *Dichromatic Color Vision* ($p = 9{,}85 \times 10^{-48}$), associated with *Lepilemuridae, Microcebus*, and *Nycticebus*. These family-level patterns highlight potential co-evolution between visual systems, reproductive timing, and social behavior, traits that may have jointly responded to nocturnal lifestyles, seasonal environments, or mating strategies in these lineages.

Finally, we assessed whether the genes or pathways associated with traits across the primate phylogeny were enriched for heritability in corresponding human traits based on GWAS data. No significant trend was detected, suggesting limited overlap between phylogenetically inferred associations and loci contributing to current within-population variation in humans. Furthermore, among the full set of genomic positions underlying CAASs, the vast majority (86,16%) were

fixed in human populations for traits included in the P3GMap. This highlights a potential decoupling between macroevolutionary changes, which shape long-term phenotypic divergence across species, and microevolutionary processes, which govern variation within species. Such divergence reinforces the central notion of our finding: the complementary nature of cross-species and population-based approaches for understanding trait architecture

Three case studies: the genomic architecture of traits with complex phylogenetic histories

The *standard strategy* for the selection of species to be compared to detect CAAS applied above is designed to provide a consistent and scalable framework for constructing the P3GMap across a highly heterogeneous set of phenotypes, enabling broad cross-trait comparisons while preserving phylogenetic representativeness. Beyond this general strategy, more targeted selections of contrasting phenotypes within specific clades can uncover lineage-specific evolutionary mechanisms. To demonstrate this potential, we examined three case studies spanning distinct biological domains—Insectivorous Diet, White Blood Cell Count, and Maximum Lifespan. By refining species selection based on the particular phylogenetic distribution of each trait, we identified additional gene sets that improved downstream functional interpretation and revealed shared molecular pathways driving trait evolution. Importantly, the structure of the PGA enables researchers to perform similar clade-specific or trait-focused analyses, expanding the utility of the resource beyond the standardized genome-phenome map.

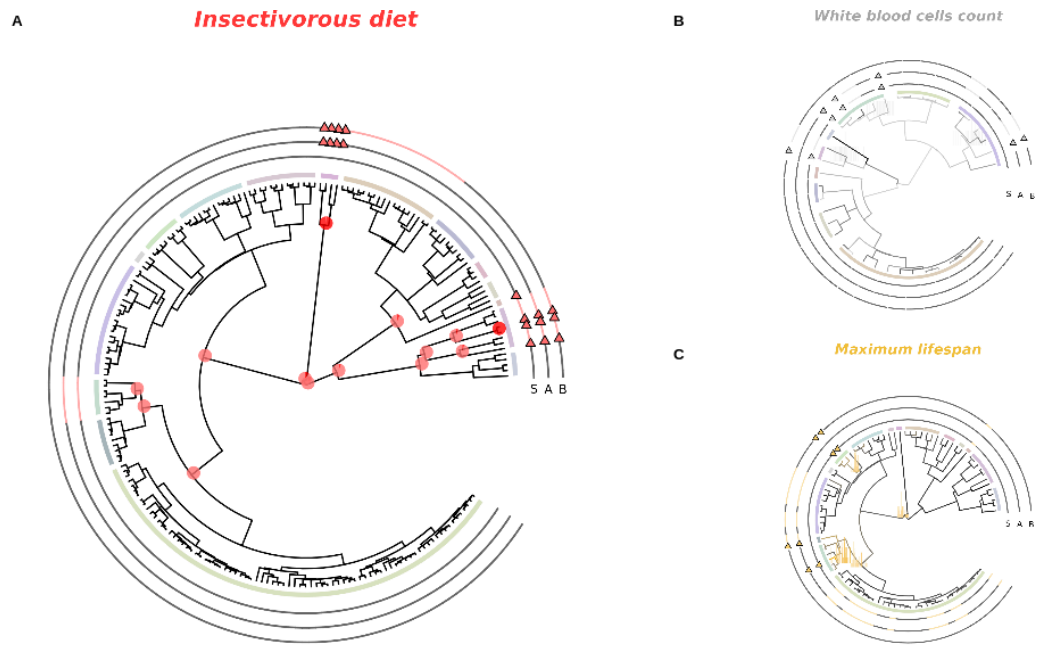
The case studies of *Insectivorous Diet* and *White Blood Cell Count* (Fig. 3A & 3B) illustrate how targeted analyses can uncover deeper biological insights. We also examined Maximum Lifespan (Fig. 3C), a trait whose phenotypic contrasts occur primarily across families rather than within them, often aligning with common species groupings in the alignments and leading to more permissive associations. To better understand the mechanisms behind these associations, we performed a focused analysis of the evolutionary relevance of the implicated amino acid changes, including natural selection analysis. By identifying sites under positive selection, we were able to assess whether these substitutions reflect adaptive pressures. This combined approach demonstrates the added value of the PGA/P3GMap when integrated with selection-based methods to dissect the molecular basis of complex traits.

Insectivorous diet

Diet interacts closely with other morphological traits, including brain (118–120) and dental adaptations (121,122), both across the entire evolutionary history of primates and within specific clades such as strepsirrhines (120). We focused on the *Insectivorous Diet*, for which our *standard strategy* selected contrasting insectivorous and non-insectivorous species within the family *Lorisidae* (slender lorises show greater insect dependence than slow lorises) and included *Daubentonia madagascariensis* (the only extant *Daubentiidae*) as an insectivorous species, along with non-insectivorous species from different families. For this targeted analysis, we included *Tarsiidae* as an additional insectivorous lineage using alternative selections of non-insectivorous species (Fig. 3A).

In doing so, we identified 619 amino acid positions in 537 genes, all of them showing significant permulation results (Fig. 3D) and were enriched in four biological categories (pFDR < 0.05, hypergeometric test). Of course, many of these genes were also detected by the *standard strategy*, including 19 genes associated with the apical part of the cell (GO:0045177), eight linked to cargo receptor activity (GO:0038024), seven to lipase activity (GO:0016298), and 14 to endopeptidase activity (GO:0004175). However, none of these enrichments were detected without the selection of alternative species. The role of the epithelial cell surface in nutrition and metabolism has been extensively studied, particularly in the context of small intestine function (123–125). Structural changes such as villus blunting or flattening of the epithelial surface are known to contribute to inflammatory and nutritional disorders (126–128).

A notable gene linked to this category is Sucrase-Isomaltase (SI), for which we identified CAASs A/R/T917K across all species selections and I627M in those including *Tarsiidae,* with directional selection detected. Sucrase-Isomaltase is an essential gene for carbohydrate digestion, and loss-of-function (LoF) mutations in Arctic human populations have been associated with gastrointestinal symptoms upon sucrose introduction in childhood diets (129–131).



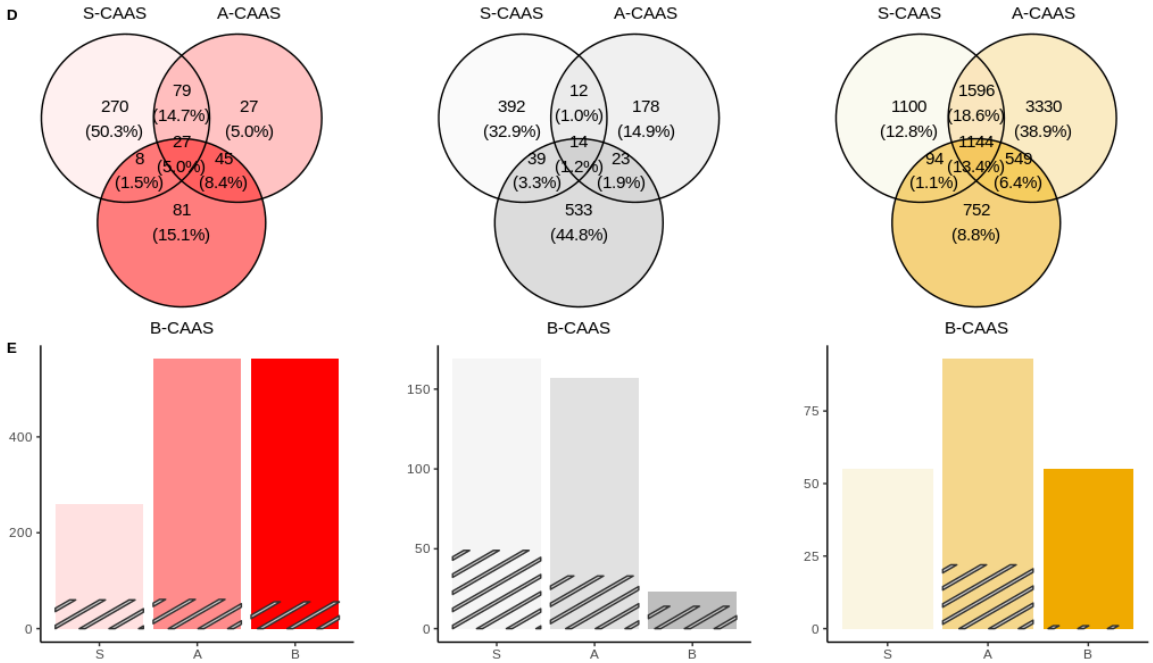


Figure 3. Study cases: The potential of the CAAS selection approach applied to *Insectivorous Diet*, *Maximum Lifespan* and *White Blood Cells Count*. A, B, and C) Selection of *Top* (triangle) and *Bottom* (colored line) species for each trait using the *standard strategy* for the complete genome-phenome map (S), and two alternative selection approaches (A and B). Pie charts represent ancestral states for the qualitative trait, whereas internal nodes with strong trait shifts are represented by continuous posterior distributions in the shape of histograms for the quantitative traits. **D)** Venn diagram representing gene overlap (*p-perm* < 0.05) between alternative family selections for each trait studied **E)** Directional selection counts for each phenotype studied. Stray lines indicate the number of positions detected with CAAS, which show positive directional selection.

Lipases also play a critical role in fat digestion and intestinal absorption (132,133). Consistent with this category, we detected CAASs in *ASPG* p.G374D (Asparaginase), *HPSE* A/F/L44V (Heparanase), *PLA2G3* L432P (Phospholipase A2 Group III), and *PNLIP* K394R (Pancreatic lipase) in the species selection, including *Tarsiidae*, with directional selection also being detected.

Interestingly, among the 27 genes and 14 amino acid positions shared across all species selections, *DEFB113* V52I (Defensin Beta 113) was the only CAAS with directional selection detected in all comparisons. A member of the β-defensin family mainly involved in host defense, *DEFB113* has undergone multiple gene gain and loss events throughout mammalian evolution (134–136), with diet proposed as an external force influencing gene family diversification (137–139).

White Blood Cells

Although white blood cell (WBC) metrics have not been extensively studied from a macroevolutionary perspective, several studies in primates have explored the associations between immune and mating systems, as well as correlations with sexually dimorphic traits, such as testis size (140,141). For this analysis, we included species with contrasting WBC from the

two families used in our *standard strategy* (*Hominidae and Lorisidae*) and added *Callitrichidae*, since they present extreme values. We identified 2,829 positions in 1,659 genes, including 1,901 positions in 1,191 genes with significant permulation results. These were enriched for olfactory receptor activity (GO:0004984), which was also observed in the *standard* P3GMap analysis, and for two extracellular matrix (ECM)-related categories, ECM constituent (GO:0005201) and ECM-receptor interaction (hsa04512). These ECM-related functions showed partial gene overlap with olfactory receptor activity (nine and six, respectively), but were not detected using the standard P3GMap. Olfactory receptors (ORs), a highly conserved sensory system in mammals, are also expressed in immune cells such as macrophages (142) and in human leukemia cell lines (143–145). Across all species selections, we identified four OR genes of particular interest. OR4D5 (I48T) showed directional selection in Hominidae. OR10G9 displayed CAASs (R89K and F/Y114L) in Callitrichidae and Lorisidae, with consistent associations across the primate phylogeny (pPGLS $< 1 \times 10^{-3}$); four additional positions in this gene also showed signs of directional selection. OR52B6 and OR10AD1 carried CAASs as well, with OR10AD1 presenting a substitution (L/M102I) in *Hominidae* and *Lorisidae* under directional selection. These findings suggest that OR genes may have been recurrently co-opted during primate evolution to mediate not only chemosensory functions but also immune-related processes, potentially linking environmental sensing with host defense.

We also identified a cluster of genes involved in xenobiotic metabolism, highlighting potential evolutionary adaptations to environmental exposures. SULT1C3 (Sulfotransferase Family 1C Member 3) showed four CAASs (K45R, H/R196Q, E227D, Q243Y) under directional selection in *Callitrichidae* and *Hominidae*, along with 19 additional significant associations (pPGLS $< 1 \times 10^{-3}$) in *Callitrichidae*-inclusive sets of species. Similarly, CES2 (Carboxylesterase 2) exhibited a CAAS (L599F) with directional selection and a consistent association across primates (pPGLS = 5.46×10^{-4}). These genes play key roles in detoxification and are part of neuro-immune pathways implicated in xenobiotic avoidance behaviors, allergy susceptibility, and immune dysregulation (146-148).

In addition, we detected directional selection in key immune-related genes, particularly within the Leukocyte Immunoglobulin-Like Receptor family, LILRB3 in *Hominidae*, and LILRA6 under selection despite lacking CAASs. Within the Major Histocompatibility Complex (MHC), HLA-DQA1 showed directional selection in *Hominidae*, while HLA-DPB1 exhibited a CAAS (Q75R) in *Callitrichidae* and *Lorisidae*, with significant phylogeny-wide association (pPGLS = 1.67×10^{-4}). Again, these findings illustrate the role of immune system genes in lineage-specific adaptations and point to the evolutionary shaping of immune recognition mechanisms in response to environmental and pathogenic pressures.

Maximum lifespan

Longevity-related traits are key targets for identifying molecular convergence across primate and mammalian phylogenies, and for understanding their implications in human biology. For instance, within the great apes, humans have evolved markedly longer lifespans than their closest relatives (48,149–151). In the case of Maximum Lifespan (after correcting for Body Mass), in addition to the species included in the *standard strategy* (humans within *Hominidae*, other long-lived *Cebidae*, and several short-lived species from different families), we specifically

targeted *Hylobatidae* as another long-lived lineage and included additional short-lived species across the phylogeny (Fig. 3C).

We identified 30,886 positions in 8,565 genes showing significant permulation results (Fig. 3D). As explained above, the phylogenetic distribution of *Maximum Lifespan* in primates does not meet our standard criteria for recovering intra-family groups of contrasting species, leading to noisier and potentially overestimated CAAS detection. To address this, we assessed association consistency in both primates and mammals. In primates, we evaluated 62,72% of the positions where more than five additional species carried amino-acids found in both the *Top* and *Bottom* groups. This refinement yielded 5,225 positions in 3,338 genes with significantly consistent associations across the primate phylogeny. We then applied the same criteria to the mammalian alignments, testing the association consistency in mammals for 30,71% of the positions. This analysis identified 1,061 positions in 921 genes, with significantly consistent associations across both primates and non-primate mammals. This bespoken phylogenetic curation enabled us to derive a more informative gene set for investigating the genetic basis of *Maximum Lifespan*.

We identified four enriched functional categories among the genes carrying CAASs with significantly consistent associations across primates, including chromosome segregation (GO:0007059), which was also detected in our standard species selection, and the complement and coagulation cascades (hsa04610). Among the genes with significantly consistent associations across mammalian phylogeny, we found eight enriched categories, including interleukin-10 production (GO:0032613) and regulation of multi-organism processes (GO:0043900). Additionally, we identified 139 amino acid positions in 136 genes that showed consistent associations in both primates and mammals. These genes were significantly enriched in two human phenotypes: cranial nerve paralysis (HP:0006824) and abnormal cranial nerve physiology (HP:0031910). The human brain contains 12 cranial nerves, and paralysis typically arises from congenital, traumatic, or vascular disorders, including life-shortening conditions, such as diabetes (152,153), ischemic strokes (154), or aneurysms (155). Seven overlapping genes were also found in the standard species selection, and six of them correspond to CAASs in all species selections. A notable example is *IL1B* (Interleukin 1 Beta), where we identified the CAAS L152M, a variant in a gene that plays a central role in cytokine-induced immune and inflammatory responses (96,156,157). We also identified CAASs in two other interleukins from the same gene set, *IL12B* and *IL4I1*, and detected directional selection in *IL23R* and *IL5*.

Another example is *KLK4* Q227P (Kallikrein-related peptidase 4), which promotes tumorigenesis by activating growth factors, including insulin-like growth factor (*IGF*). The insulin-like growth factor 1 (*IGF-1*) signaling pathway has been repeatedly highlighted in comparative studies of longevity across species (28,158). In addition, we detected both directional selection and CAASs in *KLK1* (Kallikrein 1) P57R, in the selection of *Hylobatidae* and *Hominidae*. *KLK1* is known to degrade insulin-like growth factor-binding protein-3, although this particular amino acid position is not consistently associated across primates or mammals.

Concluding remarks and limitations

In this study, we leveraged the most comprehensive catalog of primate genomic and phenotypic variation to date to develop a standardized, cross-trait comparative framework for mapping

protein-coding genomic variation onto phenotypic diversity. By integrating this framework with phylogenetic comparative methods, we generated the Primate Protein-coding Genome-Phenome Map (P3GMap), made available through the Primate Genome-Phenome Archive (PGA). This resource revealed widespread associations between protein-coding variation and 263 complex traits across 224 primate species. Our results highlight how macroevolutionary analyses can complement within-species approaches to better understand the genetic architecture of traits with both evolutionary and biomedical relevance. The primate gene-trait associations uncovered here provide a foundation for exploring their functional implications in human health, especially through future comparative multi-omics studies that can illuminate the conservation and divergence of relevant genes and pathways.

While this broad strategy offers a panoramic view of phenotypic evolution, it also enables trait-focused exploration, as illustrated in our case studies on Insectivorous Diet, White Blood Cells, and Maximum Lifespan. These examples show how targeted species selections, informed by phylogenetic trait distributions, can uncover lineage-specific signatures and deepen our understanding of convergent molecular evolution (11,159,160).

Nonetheless, our approach has limitations. The use of permulation-based methods grounded in Brownian motion models offers a rigorous control for phylogenetic structure but may not fully capture the complexity of evolutionary dynamics. Additionally, differences between the two analytical methods (RERconverge and CAASTools), reflect distinct conceptual and statistical foundations, which introduce some variability in the gene-trait associations recovered. Rather than being a drawback, this highlights the value of complementary approaches for capturing a broader spectrum of evolutionary relationships between genomes and phenome. Moreover, by focusing on protein-coding sequences, we did not account for regulatory variants, structural variation, or gene expression dynamics—key contributors to trait architecture (164). Finally, most CAAS positions were found to be fixed in human populations, which expands our understanding of the molecular basis of these traits and is a substantial addition to lists of actionable genes and individual variants. Importantly, while our analyses identified thousands of such candidate protein-altering variants, many with potential functional relevance, experimental validation will be essential to confirm causality.

Despite these limitations, the scope and resolution of our analysis compare favorably with large-scale GWAS, which typically identify far fewer loci with limited causal interpretability. In contrast, our phylogenetic approach implicates thousands of protein-coding variants directly linked to phenotypic changes. These findings lay the groundwork for future efforts to functionally characterize the evolutionary and biomedical roles of the implicated loci. As phenotypic datasets become richer, more classes of genomic variants are explored and statistical models more refined, PGA can serve as a platform for expanding genome-phenome inference beyond coding regions, ultimately helping to bridge comparative genomics and human medical genetics.

BIBLIOGRAPHY

1. Ogienko AA, Omelina ES, Bylino OV, Batin MA, Georgiev PG, Pindyurin AV. Drosophila as a Model Organism to Study Basic Mechanisms of Longevity. Int J Mol Sci. 2022 Sep 24;23(19):11244.
2. Zhang S, Li F, Zhou T, Wang G, Li Z. Caenorhabditis elegans as a Useful Model for Studying Aging Mutations. Front Endocrinol [Internet]. 2020 [cited 2024 Jan 17];11. Available from: https://www.frontiersin.org/articles/10.3389/fendo.2020.554994
3. Irion U, Nüsslein-Volhard C. Developmental genetics with model organisms. Proc Natl Acad Sci. 2022 Jul 26;119(30):e2122148119.
4. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet. 2021 Oct;53(10):1415–24.
5. Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, et al. A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet. 2019 Sep;51(9):1339–48.
6. Liu X, Tian D, Li C, Tang B, Wang Z, Zhang R, et al. GWAS Atlas: an updated knowledgebase integrating more curated associations in plants and animals. Nucleic Acids Res. 2023 Jan 6;51(D1):D969–76.
7. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nat Genet. 2018 Nov;50(11):1593–9.
8. de Magalhães JP, Wang J. The fog of genetics: what is known, unknown and unknowable in the genetics of complex traits and diseases. EMBO Rep. 2019 Nov 5;20(11):e48054.
9. Smith SD, Pennell MW, Dunn CW, Edwards SV. Phylogenetics is the New Genetics (for Most of Biodiversity). Trends Ecol Evol. 2020 May 1;35(5):415–25.
10. Garamszegi LZ, editor. Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice [Internet]. Berlin, Heidelberg: Springer; 2014 [cited 2023 Sep 25]. Available from: http://link.springer.com/10.1007/978-3-662-43550-2
11. Nagy LG, Merényi Z, Hegedüs B, Bálint B. Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing. Nucleic Acids Res. 2020 Mar 18;48(5):2209–19.
12. Harmon L. Phylogenetic Comparative Methods: Learning From Trees. 2019 May 20 [cited 2023 Sep 27]; Available from: https://ecoevorxiv.org/repository/view/4486/
13. Hiller M, Schaar BT, Indjeian VB, Kingsley DM, Hagey LR, Bejerano G. A “forward genomics” approach links genotype to phenotype using independent phenotypic losses among related species. Cell Rep. 2012 Oct 25;2(4):817–23.
14. De Bie T, Cristianini N, Demuth JP, Hahn MW. CAFE: a computational tool for the study of gene family evolution. Bioinformatics. 2006 May 15;22(10):1269–71.
15. Yan H, Hui Z, Thomas GWC, Edwards SV, Sackton TB, Liu JS. PhyloAcc-GT: A Bayesian method for inferring patterns of substitution rate shifts on targeted lineages accounting for gene tree discordance. Mol Biol Evol. 2023 Sep 4;msad195.
16. Rey C, Guéguen L, Sémon M, Boussau B. Accurate Detection of Convergent Amino-Acid Evolution with PCOC. Mol Biol Evol. 2018 Sep 1;35(9):2296–306.
17. DeCasien AR, Trujillo AE, Janiak MC, Harshaw EP, Caes ZN, Galindo GA, et al. Equivocal evidence for a link between megalencephaly-related genes and primate brain size evolution. Sci Rep. 2022 Jun 28;12(1):10902.

18. Sun Y, Liu Y, Sun X, Lin Y, Yin D, Xu S, et al. Insights into body size variation in cetaceans from the evolution of body-size-related genes. BMC Evol Biol. 2019 Jul 27;19(1):157.
19. Chen L, Qiu Q, Jiang Y, Wang K, Lin Z, Li Z, et al. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science. 2019 Jun 21;364(6446):eaav6202.
20. Silva FA, Souza ÉMS, Ramos E, Freitas L, Nery MF. The molecular evolution of genes previously associated with large sizes reveals possible pathways to cetacean gigantism. Sci Rep. 2023 Jan 19;13(1):67.
21. Xu S, Sun X, Niu X, Zhang Z, Tian R, Ren W, et al. Genetic basis of brain size evolution in cetaceans: insights from adaptive evolution of seven primary microcephaly (MCPH) genes. BMC Evol Biol. 2017 Aug 29;17(1):206.
22. Shao Y, Zhou L, Li F, Zhao L, Zhang BL, Shao F, et al. Phylogenomic analyses provide insights into primate evolution. Science. 2023 Jun 2;380(6648):913–24.
23. Montgomery SH, Capellini I, Venditti C, Barton RA, Mundy NI. Adaptive Evolution of Four Microcephaly Genes and the Evolution of Brain Size in Anthropoid Primates. Mol Biol Evol. 2011 Jan 1;28(1):625–38.
24. Montgomery SH, Mundy NI. Microcephaly genes evolved adaptively throughout the evolution of eutherian mammals. BMC Evol Biol. 2014 Jun 5;14(1):120.
25. Farré X, Molina R, Barteri F, Timmers PRHJ, Joshi PK, Oliva B, et al. Comparative Analysis of Mammal Genomes Unveils Key Genomic Variability for Human Life Span. Mol Biol Evol. 2021 Oct 27;38(11):4948–61.
26. Muntané G, Farré X, Rodríguez JA, Pegueroles C, Hughes DA, de Magalhães JP, et al. Biological Processes Modulating Longevity across Primates: A Phylogenetic Genome-Phenome Analysis. Mol Biol Evol. 2018 Aug 1;35(8):1990–2004.
27. Kowalczyk A, Partha R, Clark NL, Chikina M. Pan-mammalian analysis of molecular constraints underlying extended lifespan. Kaeberlein M, Tautz D, editors. eLife. 2020 Feb 11;9:e51089.
28. Yu Z, Seim I, Yin M, Tian R, Sun D, Ren W, et al. Comparative analyses of aging-related genes in long-lived mammals provide insights into natural longevity. The Innovation. 2021 May 28;2(2):100108.
29. Kuderna LFK, Gao H, Janiak MC, Kuhlwilm M, Orkin JD, Bataillon T, et al. A global catalog of whole-genome diversity from 233 primate species. Science. 2023 Jun 2;380(6648):906–13.
30. DeCasien AR, Higham JP. Primate mosaic brain evolution reflects selection on sensory and cognitive specialization. Nat Ecol Evol. 2019 Oct;3(10):1483–93.
31. Kemnitz JW. DATABASE FOR INDICES OF AGING IN NONHUMAN PRIMATES. Innov Aging. 2019 Nov 8;3(Suppl 1):S957.
32. Galán-Acedo C, Arroyo-Rodríguez V, Andresen E, Arasa-Gisbert R. Ecological traits of the world's primates. Sci Data. 2019 May 13;6(1):55.
33. Zehr SM, Roach RG, Haring D, Taylor J, Cameron FH, Yoder AD. Life history profiles for 27 strepsirrhine primate taxa generated using captive data from the Duke Lemur Center. Sci Data. 2014;1:140019.
34. Simmen B, Morino L, Blanc S, Garcia C. The energy allocation trade-offs underlying life history traits in hypometabolic strepsirhines and other primates. Sci Rep. 2021 Jul 9;11(1):14196.
35. Isler K, van Schaik CP. Allomaternal care, life history and brain size evolution in mammals. J Hum Evol. 2012 Jul 1;63(1):52–63.

36. Antoł A, Kozłowski J. Scaling of organ masses in mammals and birds: phylogenetic signal and implications for metabolic rate scaling. ZooKeys. 2020 Nov 2;982:149–59.
37. Burger JR, George MA Jr, Leadbetter C, Shaikh F. The allometry of brain size in mammals. J Mammal. 2019 Apr 24;100(2):276–83.
38. Sansalone G, Allen K, Ledogar JA, Ledogar S, Mitchell DR, Profico A, et al. Variation in the strength of allometry drives rates of evolution in primate brain shape. Proc R Soc B Biol Sci. 2020 Jul 8;287(1930):20200807.
39. Kowalczyk A, Meyer WK, Partha R, Mao W, Clark NL, Chikina M. RERconverge: an R package for associating evolutionary rates with convergent traits. Bioinformatics. 2019 Nov 15;35(22):4815–7.
40. Barteri F, Valenzuela A, Farré X, Juan D de, Muntané G, Esteve-Altava B, et al. CAAStools: a toolbox to identify and test Convergent Amino Acid Substitutions [Internet]. bioRxiv; 2023 [cited 2023 Sep 25]. p. 2022.12.14.520422. Available from: https://www.biorxiv.org/content/10.1101/2022.12.14.520422v3
41. Saputra E, Kowalczyk A, Cusick L, Clark N, Chikina M. Phylogenetic Permulations: A Statistically Rigorous Approach to Measure Confidence in Associations in a Phylogenetic Context. Mol Biol Evol. 2021 Jul 1;38(7):3004–21.
42. Chelban V, Patel N, Vandrovcova J, Zanetti MN, Lynch DS, Ryten M, et al. Mutations in NKX6-2 Cause Progressive Spastic Ataxia and Hypomyelination. Am J Hum Genet. 2017 Jun 1;100(6):969–77.
43. Toch M, Harris A, Schakman O, Kondratskaya E, Boulland JL, Dauguet N, et al. Onecut-dependent Nkx6.2 transcription factor expression is required for proper formation and activity of spinal locomotor circuits. Sci Rep. 2020 Jan 22;10(1):996.
44. Wu E, Zhu J, Ma Z, Tuo B, Terai S, Mizuno K, et al. Gastric alarmin release: A warning signal in the development of gastric mucosal diseases. Front Immunol [Internet]. 2022 [cited 2023 Oct 17];13. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2022.1008047
45. Trieschmann G, Wilhelm C, Berweck S, Zech M. De novo retinoic acid receptor beta (RARB) variant associated with microphthalmia and dystonia. Eur J Med Genet. 2023 Aug 1;66(8):104802.
46. Srour M, Chitayat D, Caron V, Chassaing N, Bitoun P, Patry L, et al. Recessive and Dominant Mutations in Retinoic Acid Receptor Beta in Cases with Microphthalmia and Diaphragmatic Hernia. Am J Hum Genet. 2013 Oct 3;93(4):765–72.
47. Foran E, Rosenblum L, Bogush AI, Trotti D. Sumoylation of critical proteins in amyotrophic lateral sclerosis: emerging pathways of pathogenesis. Neuromolecular Med. 2013 Dec;15(4):10.1007/s12017-013-8262–x.
48. Tejada-Martinez D, Avelar RA, Lopes I, Zhang B, Novoa G, de Magalhães JP, et al. Positive Selection and Enhancer Evolution Shaped Lifespan and Body Mass in Great Apes. Mol Biol Evol. 2022 Feb 3;39(2):msab369.
49. Orkin JD, Montague MJ, Tejada-Martinez D, de Manuel M, Del Campo J, Cheves Hernandez S, et al. The genomics of ecological flexibility, large brains, and long lives in capuchin monkeys revealed with fecalFACS. Proc Natl Acad Sci U S A. 2021 Feb 16;118(7):e2010632118.
50. Benssouina FZ, Parat F, Villard C, Leloup L, Garrouste F, Sabatier J marc, et al. Overexpression of a Novel Noxo1 Mutant Increases Ros Production and Noxo1 Relocalisation. Int J Mol Sci. 2023 Feb 28;24(5):4663.

51. Burch TC, Rhim JS, Nyalwidhe JO. Mitochondria Biogenesis and Bioenergetics Gene Profiles in Isogenic Prostate Cells with Different Malignant Phenotypes. BioMed Res Int. 2016;2016:1785201.
52. Bui TM, Wiesolek HL, Sumagin R. ICAM-1: A master regulator of cellular responses in inflammation, injury resolution, and tumorigenesis. J Leukoc Biol. 2020 Sep 1;108(3):787–99.
53. Makay B, Unsal E. Does breast-feeding affect severity of familial Mediterranean fever? Clin Rheumatol. 2009 Dec;28(12):1389–93.
54. Özen S, Batu ED, Demir S. Familial Mediterranean Fever: Recent Developments in Pathogenesis and New Recommendations for Management. Front Immunol [Internet]. 2017 [cited 2023 Sep 26];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5362626/
55. Yanmaz MN, Özcan AJ, Savan K. The impact of familial Mediterranean fever on reproductive system. Clin Rheumatol. 2014 Oct;33(10):1385–8.
56. Massender E, Oliveira HR, Brito LF, Maignel L, Jafarikia M, Baes CF, et al. Genome-wide association study for milk production and conformation traits in Canadian Alpine and Saanen dairy goats. J Dairy Sci. 2023 Feb 1;106(2):1168–89.
57. Zhang X, Zara J, Siu RK, Ting K, Soo C. The Role of NELL-1, a Growth Factor Associated with Craniosynostosis, in Promoting Bone Regeneration. J Dent Res. 2010 Sep;89(9):865–78.
58. Lai K, Xi Y, Du X, Jiang Z, Li Y, Huang T, et al. Activation of Nell-1 in BMSC Sheet Promotes Implant Osseointegration Through Regulating Runx2/Osterix Axis. Front Cell Dev Biol [Internet]. 2020 [cited 2023 Sep 26];8. Available from: https://www.frontiersin.org/articles/10.3389/fcell.2020.00868
59. Chen W, Zhang X, Siu RK, Chen F, Shen J, Zara JN, et al. Nfatc2 Is a Primary Response Gene of Nell-1 Regulating Chondrogenesis in ATDC5 Cells. J Bone Miner Res. 2011 Jun;26(6):1230–41.
60. Abebe MS, Seyoum G, Emamu B, Teshome D. Congenital Hydrocephalus and Associated Risk Factors: An Institution-Based Case–Control Study, Dessie Town, North East Ethiopia. Pediatr Health Med Ther. 2022 May 11;13:175–82.
61. Tully HM, Dobyns WB. Infantile hydrocephalus: a review of epidemiology, classification and causes. Eur J Med Genet. 2014 Aug;57(8):359–68.
62. Rossie JB, Smith TD. Ontogeny of the nasolacrimal duct in primates: functional and phylogenetic implications. J Anat. 2007 Feb;210(2):195–208.
63. Rossie JB, Smith TD, Beard KC, Godinot M, Rowe TB. Nasolacrimal anatomy and haplorhine origins. J Hum Evol. 2018 Jan 1;114:176–83.
64. Low KJ, Ansari M, Abou Jamra R, Clarke A, El Chehadeh S, FitzPatrick DR, et al. PUF60 variants cause a syndrome of ID, short stature, microcephaly, coloboma, craniofacial, cardiac, renal and spinal features. Eur J Hum Genet EJHG. 2017 May;25(5):552–9.
65. Popova S, Charness ME, Burd L, Crawford A, Hoyme HE, Mukherjee RAS, et al. Fetal alcohol spectrum disorders. Nat Rev Dis Primer. 2023 Feb 23;9(1):1–21.
66. Pehlivan D, Karaca E, Aydin H, Beck CR, Gambin T, Muzny DM, et al. Whole-exome sequencing links TMCO1 defect syndrome with cerebro-facio-thoracic dysplasia. Eur J Hum Genet EJHG. 2014 Sep;22(9):1145–8.
67. Verloes A, Di Donato N, Masliah-Planchon J, Jongmans M, Abdul-Raman OA, Albrecht B, et al. Baraitser–Winter cerebrofrontofacial syndrome: delineation of the spectrum in 42 cases. Eur J Hum Genet. 2015 Mar;23(3):292–301.

68. Kline AD, Moss JF, Selicorni A, Bisgaard AM, Deardorff MA, Gillett PM, et al. Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. Nat Rev Genet. 2018 Oct;19(10):649–66.
69. Ockeloen CW, Willemsen MH, de Munnik S, van Bon BWM, de Leeuw N, Verrips A, et al. Further delineation of the KBG syndrome phenotype caused by ANKRD11 aberrations. Eur J Hum Genet EJHG. 2015 Sep;23(9):1176–85.
70. Guo L, Elcioglu NH, Wang Z, Demirkol YK, Isguven P, Matsumoto N, et al. Novel and recurrent COL11A1 and COL2A1 mutations in the Marshall-Stickler syndrome spectrum. Hum Genome Var. 2017;4:17040.
71. Panarotto M, Davidson IF, Litos G, Schleiffer A, Peters JM. Cornelia de Lange syndrome mutations in NIPBL can impair cohesin-mediated DNA loop extrusion. Proc Natl Acad Sci U S A. 2022 May 3;119(18):e2201029119.
72. Fusco C, Micale L, Augello B, Teresa Pellico M, Menghini D, Alfieri P, et al. Smaller and larger deletions of the Williams Beuren syndrome region implicate genes involved in mild facial phenotype, epilepsy and autistic traits. Eur J Hum Genet EJHG. 2014 Jan;22(1):64–70.
73. E M, Song X, Wang L, Yang Y, Wei X, Yu J, et al. Mate choice for major histocompatibility complex (MHC) complementarity in the Yellow-rumped Flycatcher (Ficedula zanthopygia). Avian Res. 2021 May 29;12(1):27.
74. Havlíček J, Winternitz J, Roberts SC. Major histocompatibility complex-associated odour preferences and human mate choice: near and far horizons. Philos Trans R Soc Lond B Biol Sci. 2020 Jun 8;375(1800):20190260.
75. Roved J, Hansson B, Tarka M, Hasselquist D, Westerdahl H. Evidence for sexual conflict over major histocompatibility complex diversity in a wild songbird. Proc R Soc B Biol Sci. 2018 Aug;285(1884):20180841.
76. Han QH, Sun RN, Yang HQ, Wang ZW, Wan QH, Fang SG. MHC class I diversity predicts non-random mating in Chinese alligators (Alligator sinensis). Heredity. 2019 Jun;122(6):809–18.
77. Sommer S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front Zool. 2005 Oct 20;2:16.
78. Auer KE, Primus J, Istel S, Dahlhoff M, Rülicke T. Mate genetic similarity affects mating behaviour but not maternal investment in mice. Sci Rep. 2023 Jun 29;13(1):10536.
79. Achorn AM, Rosenthal GG. It's Not about Him: Mismeasuring 'Good Genes' in Sexual Selection. Trends Ecol Evol. 2020 Mar 1;35(3):206–19.
80. Thonhauser KE, Raveh S, Penn DJ. Multiple paternity does not depend on male genetic diversity. Anim Behav. 2014 Jul;93(100):135–41.
81. Amo L, Paz GA de, Kabbert J, Machordom A. House sparrows do not exhibit a preference for the scent of potential partners with different MHC-I diversity and genetic distances. PLOS ONE. 2022 Dec 21;17(12):e0278892.
82. Westerdahl H. No evidence of an MHC-based female mating preference in great reed warblers. Mol Ecol. 2004 Aug;13(8):2465–70.
83. He F, Furones AR, Landegren N, Fuxe J, Sarhan D. Sex dimorphism in the tumor microenvironment - From bench to bedside and back. Semin Cancer Biol. 2022 Nov;86(Pt 3):166–79.

84. Zheng D, Trynda J, Williams C, Vold JA, Nguyen JH, Harnois DM, et al. Sexual dimorphism in the incidence of human cancers. BMC Cancer. 2019 Jul 12;19(1):684.
85. Rubin JB. The spectrum of sex differences in cancer. Trends Cancer. 2022 Apr;8(4):303–15.
86. Ly D, Forman D, Ferlay J, Brinton LA, Cook MB. An International Comparison of Male and Female Breast Cancer Incidence Rates. Int J Cancer J Int Cancer. 2013 Apr 15;132(8):1918–26.
87. Huang X, Sun D, Wu T, Liu X, Xu S, Yang G. Genomic insights into body size evolution in Carnivora support Peto's paradox. BMC Genomics. 2021 Jun 9;22(1):429.
88. Ma S, Upneja A, Galecki A, Tsai YM, Burant CF, Raskind S, et al. Cell culture-based profiling across mammals reveals DNA repair and metabolism as determinants of species longevity. Wagers AJ, editor. eLife. 2016 Nov 22;5:e19130.
89. Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, et al. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science. 2013 Jan 25;339(6118):456–60.
90. Keane M, Semeiks J, Webb AE, Li YI, Quesada V, Craig T, et al. Insights into the evolution of longevity from the bowhead whale genome. Cell Rep. 2015 Jan 6;10(1):112–22.
91. Bogliolo M, Surrallés J. Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics. Curr Opin Genet Dev. 2015 Aug 1;33:32–40.
92. Helbling-Leclerc A, Garcin C, Rosselli F. Beyond DNA repair and chromosome instability-Fanconi anaemia as a cellular senescence-associated syndrome. Cell Death Differ. 2021 Apr;28(4):1159–73.
93. Geremia A, Biancheri P, Allan P, Corazza GR, Di Sabatino A. Innate and adaptive immunity in inflammatory bowel disease. Autoimmun Rev. 2014 Jan;13(1):3–10.
94. Nie S, Yuan Y. The Role of Gastric Mucosal Immunity in Gastric Diseases. J Immunol Res. 2020 Jul 24;2020:7927054.
95. Kelley N, Jeltema D, Duan Y, He Y. The NLRP3 Inflammasome: An Overview of Mechanisms of Activation and Regulation. Int J Mol Sci. 2019 Jul 6;20(13):3328.
96. Boraschi D, Italiani P, Weil S, Martin MU. The family of the interleukin-1 receptors. Immunol Rev. 2018 Jan;281(1):197–232.
97. Huang J, Huffman JE, Huang Y, Do Valle Í, Assimes TL, Raghavan S, et al. Genomics and phenomics of body mass index reveals a complex disease network. Nat Commun. 2022 Dec 29;13(1):7973.
98. Kuningas M, Altmäe S, Uitterlinden AG, Hofman A, van Duijn CM, Tiemeier H. The relationship between fertility and lifespan in humans. Age. 2011 Dec;33(4):615–22.
99. Gorbunova V, Seluanov A, Zhang Z, Gladyshev VN, Vijg J. Comparative genetics of longevity and cancer: insights from long-lived rodents. Nat Rev Genet. 2014 Aug;15(8):531–40.
100. Cho EJ, Jeong SM, Chung GE, Yoo JJ, Cho Y, Lee KN, et al. Gamma-glutamyl transferase and risk of all-cause and disease-specific mortality: a nationwide cohort study. Sci Rep. 2023 Jan 31;13(1):1751.
101. Koenig G, Seneff S. Gamma-Glutamyltransferase: A Predictive Biomarker of Cellular Antioxidant Inadequacy and Disease Risk. Dis Markers. 2015;2015:818570.
102. Kalirawna TR, Rohilla J, Bairwa SS, Gothwal SK, Tak P, Jain R. Increased concentration of serum gamma-glutamyl transferase in ischemic stroke patients. Brain Circ. 2021;7(2):71–6.
103. Ndrepepa G, Kastrati A. Gamma-glutamyl transferase and cardiovascular disease. Ann Transl Med. 2016 Dec;4(24):481.

104. Zhou Y, Zhao S, Chen K, Hua W, Zhang S. Predictive value of gamma-glutamyltransferase for ventricular arrhythmias and cardiovascular mortality in implantable cardioverter-defibrillator patients. BMC Cardiovasc Disord. 2019 May 30;19(1):129.
105. Li S, Liao X, Pan Y, Xiang X, Zhang Y. Gamma-glutamyl transferase levels are associated with the occurrence of post-stroke cognitive impairment: a multicenter cohort study. BMC Neurol. 2022 Feb 23;22(1):65.
106. Ruttmann E, Brant LJ, Concin H, Diem G, Rapp K, Ulmer H, et al. Gamma-glutamyltransferase as a risk factor for cardiovascular disease mortality: an epidemiological investigation in a cohort of 163,944 Austrian adults. Circulation. 2005 Oct 4;112(14):2130–7.
107. Sun D, Liu H, Ouyang Y, Liu X, Xu Y. Serum Levels of Gamma-Glutamyltransferase During Stable and Acute Exacerbations of Chronic Obstructive Pulmonary Disease. Med Sci Monit Int Med J Exp Clin Res. 2020 Oct 22;26:e927771-1-e927771-7.
108. Akaydın SY, Salihoğlu EM, Güngör DG, Karanlık H, Demokan S. Correlation Between Gamma-Glutamyl Transferase Activity and Glutathione Levels in Molecular Subgroups of Breast Cancer. Eur J Breast Health. 2019 Dec 5;16(1):72–6.
109. Xiao Y, Yang H, Lu J, Li D, Xu C, Risch HA. Serum gamma-glutamyltransferase and the overall survival of metastatic pancreatic cancer. BMC Cancer. 2019 Oct 29;19(1):1020.
110. Hong SW, Lee HJ, Han K, Moon JM, Park S, Soh H, et al. Risk of gastrointestinal cancer in patients with an elevated level of gamma-glutamyltransferase: A nationwide population-based study. PloS One. 2021;16(2):e0245052.
111. Du G, Song Z, Zhang Q. Gamma-glutamyltransferase is associated with cardiovascular and all-cause mortality: A meta-analysis of prospective cohort studies. Prev Med. 2013 Jul 1;57(1):31–7.
112. Ridker PM. LDL cholesterol: controversies and future therapeutic directions. Lancet Lond Engl. 2014 Aug 16;384(9943):607–17.
113. Ndrepepa G. Aspartate aminotransferase and cardiovascular disease—a narrative review. J Lab Precis Med [Internet]. 2021 Jan 30 [cited 2024 Jan 18];6(0). Available from: https://jlpm.amegroups.org/article/view/5898
114. Ke P, Zhong L, Peng W, Xu M, Feng J, Tian Q, et al. Association of the serum transaminase with mortality among the US elderly population. J Gastroenterol Hepatol. 2022 May;37(5):946–53.
115. Xie K, Chen CH, Tsai SP, Lu PJ, Wu H, Zeng Y, et al. Loss of Life Expectancy by 10 Years or More From Elevated Aspartate Aminotransferase: Finding Aspartate Aminotransferase a Better Mortality Predictor for All-Cause and Liver-Related than Alanine Aminotransferase. Off J Am Coll Gastroenterol ACG. 2019 Sep;114(9):1478.
116. Ndrepepa G, Holdenrieder S, Cassese S, Xhepa E, Fusaro M, Laugwitz KL, et al. Aspartate aminotransferase and mortality in patients with ischemic heart disease. Nutr Metab Cardiovasc Dis. 2020 Nov 27;30(12):2335–42.
117. Henriques LD, Hauzman E, Bonci DMO, Chang BSW, Muniz JAPC, da Silva Souza G, et al. Uniform trichromacy in Alouatta caraya and Alouatta seniculus: behavioural and genetic colour vision evaluation. Front Zool. 2021 Jul 8;18(1):36.
118. DeCasien AR, Williams SA, Higham JP. Primate brain size is predicted by diet but not sociality. Nat Ecol Evol. 2017 Mar 27;1(5):1–7.

119. Grabowski M, Kopperud BT, Tsuboi M, Hansen TF. Both Diet and Sociality Affect Primate Brain-Size Evolution. Syst Biol. 2023 Mar 1;72(2):404–18.
120. López-Aguirre C, Lang MM, Silcox MT. Diet drove brain and dental morphological coevolution in strepsirrhine primates. PLOS ONE. 2022 Jun 6;17(6):e0269041.
121. Marcé-Nogué J, Püschel TA, Kaiser TM. A biomechanical approach to understand the ecomorphological relationship between primate mandibles and diet. Sci Rep. 2017 Aug 21;7(1):8364.
122. Berthaume MA, Lazzari V, Guy F. The landscape of tooth shape: Over 20 years of dental topography in primates. Evol Anthropol. 2020;29(5):245–62.
123. Sallee MD, Pickett MA, Feldman JL. Apical PAR complex proteins protect against programmed epithelial assaults to create a continuous and functional intestinal lumen. eLife. 10:e64437.
124. Qi M, Tan B, Wang J, Li J, Liao S, Yan J, et al. Small intestinal transcriptome analysis revealed changes of genes involved in nutrition metabolism and immune responses in growth retardation piglets. J Anim Sci. 2019 Sep;97(9):3795–808.
125. Stojanović O, Altirriba J, Rigo D, Spiljar M, Evrard E, Roska B, et al. Dietary excess regulates absorption and surface of gut epithelium through intestinal PPARα. Nat Commun. 2021 Dec 2;12(1):7031.
126. Hossain MS, Begum SMKN, Rahman MM, Mazumder RN, Parvez M, Gazi MA, et al. Alterations in the histological features of the intestinal mucosa in malnourished adults of Bangladesh. Sci Rep. 2021 Jan 27;11(1):2355.
127. Farràs M, Chandwe K, Mayneris-Perxachs J, Amadi B, Louis-Auguste J, Besa E, et al. Characterizing the metabolic phenotype of intestinal villus blunting in Zambian children with severe acute malnutrition and persistent diarrhea. PloS One. 2018;13(3):e0192092.
128. Bein A, Fadel CW, Swenor B, Cao W, Powers RK, Camacho DM, et al. Nutritional deficiency in an intestine-on-a-chip recapitulates injury hallmarks associated with environmental enteric dysfunction. Nat Biomed Eng. 2022 Nov;6(11):1236–47.
129. Andersen MK, Skotte L, Jørsboe E, Polito R, Stæger FF, Aldiss P, et al. Loss of Sucrase-Isomaltase Function Increases Acetate Levels and Improves Metabolic Health in Greenlandic Cohorts. Gastroenterology. 2022 Apr 1;162(4):1171-1182.e3.
130. Senftleber NK, Ramne S, Moltke I, Jørgensen ME, Albrechtsen A, Hansen T, et al. Genetic Loss of Sucrase-Isomaltase Function: Mechanisms, Implications, and Future Perspectives. Appl Clin Genet. 2023 Mar 23;16:31–9.
131. Senftleber NK, Skøtt Pedersen K, Schnoor Jørgensen C, Pedersen H, Bjerg Christensen MM, Kabel Madsen E, et al. The effect of sucrase-isomaltase deficiency on metabolism, food intake and preferences: protocol for a dietary intervention study. Int J Circumpolar Health. 82(1):2178067.
132. Pirahanchi Y, Sharma S. Biochemistry, Lipase. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 [cited 2024 Feb 9]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK537346/
133. Zhu G, Fang Q, Zhu F, Huang D, Yang C. Structure and Function of Pancreatic Lipase-Related Protein 2 and Its Relationship With Pathological States. Front Genet [Internet]. 2021 [cited 2024 Feb 9];12. Available from: https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2021.693538

134. Fu J, Zong X, Jin M, Min J, Wang F, Wang Y. Mechanisms and regulation of defensins in host defense. Signal Transduct Target Ther. 2023 Aug 14;8(1):1–30.
135. Islam S, Akhand MRN, Hasan M. Evolutionary trend of bovine β-defensin proteins toward functionality prediction: A domain-based bioinformatics study. Heliyon. 2023 Mar 4;9(3):e14158.
136. Zhang ZY, Zhang HM, Li DS, Xiong TY, Fang SG. Characterization of the β-defensin genes in giant panda. Sci Rep. 2018 Aug 17;8(1):12308.
137. Robinson K, Ma X, Liu Y, Qiao S, Hou Y, Zhang G. Dietary modulation of endogenous host defense peptide synthesis as an alternative approach to in-feed antibiotics. Anim Nutr. 2018 Jun;4(2):160–9.
138. Chen J, Zhai Z, Long H, Yang G, Deng B, Deng J. Inducible expression of defensins and cathelicidins by nutrients and associated regulatory mechanisms. Peptides. 2020 Jan 1;123:170177.
139. Castellanos FX, Moreno-Santillán D, Hughes GM, Paulat NS, Sipperly N, Brown AM, et al. The evolution of antimicrobial peptides in Chiroptera. Front Immunol [Internet]. 2023 [cited 2024 Jan 18];14. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2023.1250229
140. Nunn CL, Gittleman JL, Antonovics J. Promiscuity and the primate immune system. Science. 2000 Nov 10;290(5494):1168–70.
141. Anderson MJ, Hessel JK, Dixson AF. Primate mating systems and the evolution of immune response. J Reprod Immunol. 2004 Feb 1;61(1):31–8.
142. Orecchioni M, Matsunami H, Ley K. Olfactory receptors in macrophages and inflammation. Front Immunol. 2022 Oct 13;13:1029244.
143. Manteniotis S, Wojcik S, Göthert JR, Dürig J, Dührsen U, Gisselmann G, et al. Deorphanization and characterization of the ectopically expressed olfactory receptor OR51B5 in myelogenous leukemia cells. Cell Death Discov. 2016 May 9;2(1):1–9.
144. Guardia GDA, Naressi RG, Buzzato VC, da Costa JB, Zalcberg I, Ramires J, et al. Acute Myeloid Leukemia Expresses a Specific Group of Olfactory Receptors. Cancers. 2023 Jun 6;15(12):3073.
145. Manteniotis S, Wojcik S, Brauhoff P, Möllmann M, Petersen L, Göthert JR, et al. Functional characterization of the ectopically expressed olfactory receptor 2AT4 in human myelogenous leukemia. Cell Death Discov. 2016 Jan 25;2(1):1–10.
146. Florsheim EB, Bachtel ND, Cullen JL, Lima BGC, Godazgar M, Carvalho F, et al. Immune sensing of food allergens promotes avoidance behaviour. Nature. 2023 Aug;620(7974):643–50.
147. Pollard KM, Christy JM, Cauvi DM, Kono DH. Environmental Xenobiotic Exposure and Autoimmunity. Curr Opin Toxicol. 2018 Aug;10:15–22.
148. Talbot S, Foster SL, Woolf CJ. Neuroimmunity: Physiology and Pathology. Annu Rev Immunol. 2016 May 20;34:421–47.
149. Colchero F, Aburto JM, Archie EA, Boesch C, Breuer T, Campos FA, et al. The long lives of primates and the 'invariant rate of ageing' hypothesis. Nat Commun. 2021 Jun 16;12(1):3666.
150. Byrne H, Webster TH, Brosnan SF, Izar P, Lynch JW. Signatures of adaptive evolution in platyrrhine primate genomes. Proc Natl Acad Sci U S A. 2022 Aug 30;119(35):e2116681119.
151. Finch CE. Evolution in health and medicine Sackler colloquium: Evolution of the human lifespan and diseases of aging: roles of infection, inflammation, and nutrition. Proc Natl Acad Sci U S A. 2010 Jan 26;107 Suppl 1(Suppl 1):1718–24.

152. Billerot E, Nguyen TH, Sedira N, Espinoza S, Vende B, Heron E, et al. Ocular motor nerve palsy in patients with diabetes: High-resolution MR imaging of nerve enhancement. J Fr Ophtalmol. 2023 Sep 1;46(7):726–36.
153. Choi DD, Han K, Oh SY, Park KA. Association between metabolic syndrome and incidence of ocular motor nerve palsy. Sci Rep. 2021 Nov 29;11(1):23033.
154. Kang DW, Ha SY, Sung JJ, Nam H. Cerebral small vessel disease as imaging biomarker predicting ocular cranial nerve palsy of presumed ischemic origin at admission. Sci Rep. 2022 Jul 18;12(1):12251.
155. Suda T, Matsushita E, Minamide H. Oculomotor nerve palsy owing to cerebral aneurysm. QJM Int J Med. 2022 Jun 1;115(6):395–6.
156. Boraschi D. What Is IL-1 for? The Functions of Interleukin-1 Across Evolution. Front Immunol. 2022 Apr 6;13:872155.
157. Kaneko N, Kurata M, Yamamoto T, Morikawa S, Masumoto J. The role of interleukin-1 in general pathology. Inflamm Regen. 2019 Jun 6;39(1):12.
158. Brown-Borg HM. Growth Hormone, Not IGF-1 Is the Key Longevity Regulator in Mammals. J Gerontol Ser A. 2022 Sep 1;77(9):1719–23.
159. Scossa F, Fernie AR. Ancestral sequence reconstruction - An underused approach to understand the evolution of gene function in plants? Comput Struct Biotechnol J. 2021;19:1579–94.
160. Uyeda JC, Bone N, McHugh S, Rolland J, Pennell MW. How should functional relationships be evaluated using phylogenetic comparative methods? A case study using metabolic rate and body temperature. Evolution. 2021;75(5):1097–105.
161. Sander PM, Griebeler EM, Klein N, Juarbe JV, Wintrich T, Revell LJ, et al. Early giant reveals faster evolution of large body size in ichthyosaurs than in cetaceans. Science. 2021 Dec 24;374(6575):eabf5787.
162. Benson RBJ, Hunt G, Carrano MT, Campione N. Cope's rule and the adaptive landscape of dinosaur body size evolution. Palaeontology. 2018;61(1):13–48.
163. Ksepka DT, Balanoff AM, Smith NA, Bever GS, Bhullar BAS, Bourdon E, et al. Tempo and Pattern of Avian Brain Size Evolution. Curr Biol. 2020 Jun 8;30(11):2026-2036.e3.
164. Fukushima K, Pollock DD. Detecting macroevolutionary genotype–phenotype associations using error-corrected rates of protein convergence. Nat Ecol Evol. 2023 Jan;7(1):155–70.
165. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5;215(3):403–10.
166. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7.
167. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009 Aug 1;25(15):1972–3.
168. Ranwez V, Douzery EJP, Cambon C, Chantret N, Delsuc F. MACSE v2: Toolkit for the Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. Mol Biol Evol. 2018 Oct 1;35(10):2582–4.
169. Ranwez V, Harispe S, Delsuc F, Douzery EJP. MACSE: Multiple Alignment of Coding SEquences Accounting for Frameshifts and Stop Codons. PLOS ONE. 2011 Sep 16;6(9):e22594.
170. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010 Jul 13;10(1):210.

171. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012;3(2):217–23.
172. Revell LJ, Chamberlain SA. Rphylip: an R interface for PHYLIP. Methods Ecol Evol. 2014;5(9):976–81.
173. Clavel J, Escarguel G, Merceron G. mvmorph: an r package for fitting multivariate evolutionary models to morphometric data. Methods Ecol Evol. 2015;6(11):1311–9.
174. Stefen C, Wagner F, Asztalos M, Giere P, Grobe P, Hiller M, et al. Phenotyping in the era of genomics: MaTrics—a digital character matrix to document mammalian phenotypic traits. Mamm Biol. 2022 Feb 1;102(1):235–49.
175. Soria CD, Pacifici M, Di Marco M, Stephen SM, Rondinini C. COMBINE: a coalesced mammal database of intrinsic and extrinsic traits. Ecology. 2021;102(6):e03344.
176. Devine J, Kurki HK, Epp JR, Gonzalez PN, Claes P, Hallgrímsson B. Classifying high-dimensional phenotypes with ensemble learning. BioRxiv Prepr Serv Biol. 2023 May 29;2023.05.29.542750.
177. Alghamdi SM, Hoehndorf R. Improving the classification of cardinality phenotypes using collections. J Biomed Semant. 2023 Aug 7;14(1):9.
178. Kwak HS, Im HG, Shim EB. A model for allometric scaling of mammalian metabolism with ambient heat loss. Integr Med Res. 2016 Mar 1;5(1):30–6.
179. Jones JH. Primates and the Evolution of Long-Slow Life Histories. Curr Biol CB. 2011 Sep 27;21(18):R708–17.
180. Freckleton RP, Harvey PH, Pagel M. Phylogenetic analysis and comparative data: a test and review of evidence. Am Nat. 2002 Dec;160(6):712–26.
181. Junier T, Zdobnov EM. The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. Bioinformatics. 2010 Jul 1;26(13):1669–70.
182. Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci CABIOS. 1997 Oct;13(5):555–6.
183. Collyer ML, Adams DC. RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. Methods Ecol Evol. 2018;9(7):1772–9.
184. Ives AR, Garland T Jr. Phylogenetic Logistic Regression for Binary Dependent Variables. Syst Biol. 2010 Jan 1;59(1):9–26.
185. Hussain T, Asghar S. EVALUATION OF SIMILARITY MEASURES FOR CATEGORICAL DATA. The Nucleus. 2013 Nov 27;50(4):387–94.
186. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res. 2019 Jul 2;47(W1):W199–205.
187. Herrendörfer G. JOHNSON, N. L., KOTZ, S.; KEMP, A. W.: Univariate discrete distributions. Second Edition — John Wiley & Sons 1992, 565 pp., £58.00. Biom J. 1994;36(2):164–164.
188. Pond SLK, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 2005 Mar 1;21(5):676–9.
189. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol. 2013 May;30(5):1196–205.
190. Tong C, Avilés L, Rayor LS, Mikheyev AS, Linksvayer TA. Genomic signatures of recent convergent transitions to social life in spiders. Nat Commun. 2022 Nov 22;13(1):6967.

191. Shank SD, Weaver S, Kosakovsky Pond SL. phylotree.js - a JavaScript library for application development and interactive data visualization in phylogenetics. BMC Bioinformatics. 2018 Jul 25;19(1):276.
192. Zhou W, Chen T, Chong Z, Rohrdanz MA, Melott JM, Wakefield C, et al. TransVar: a multilevel variant annotator for precision genomics. Nat Methods. 2015 Nov;12(11):1002–3.
193. Christmas MJ, Kaplow IM, Genereux DP, Dong MX, Hughes GM, Li X, et al. Evolutionary constraint and innovation across hundreds of placental mammals. Science. 2023 Apr 28;380(6643):eabn3943.
194. Kirilenko BM, Munegowda C, Osipova E, Jebb D, Sharma V, Blumer M, et al. Integrating gene annotation with orthology inference at scale. Science. 2023 Apr 28;380(6643):eabn3107.
195. Nature [Internet]. 2020 [cited 2023 Sep 28]. gnomAD. Available from: https://www.nature.com/collections/afbgiddede
196. Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, et al. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology. 2009;90(9):2648–2648.
197. Leeuw CA de, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Comput Biol. 2015 Apr 17;11(4):e1004219.
198. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet. 2015 Nov;47(11):1228–35.
199. Lee BT, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2022 update. Nucleic Acids Res. 2022 Jan 7;50(D1):D1115–22.
200. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 2005 Jan 1;33(suppl_1):D514–7.