White Rose Research Online
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/138020/

Version: Accepted Version

**Article:**

Milla, R., Bastida, J.M., Turcotte, M.M. et al. (2018) Phylogenetic patterns and phenotypic profiles of the species of plants and mammals farmed for food. Nature Ecology and Evolution, 2 (11). pp. 1808-1817.

https://doi.org/10.1038/s41559-018-0690-4

**Reuse**
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

# Phylogenetic patterns and phenotypic profiles of the species of plants and mammals farmed for food

Rubén Milla[1*;] Jesús M. Bastida[1]; Martin M. Turcotte[2]; Glynis Jones[3]; Cyrille Violle[4], Colin P. Osborne[5]; Julia Chacón-Labella[1]; Ênio E.Sosinski Jr.[6]; Jens Kattge[7,8]; Daniel C. Laughlin[9]; Estelle Forey[10]; Vanessa Minden[11,12]; Johannes H. C. Cornelissen[13]; Bernard Amiaud[14]; Koen Kramer[15]; Gerhard Boenisch[7]; Tianhua He[16]; Valerio D. Pillar[17]; and Chaeho Byun[18]

[1]Departamento de Biologia, Geología, Física y Química Inorgánica. Universidad Rey Juan Carlos. Tulipán s/n, 28933 Móstoles, Spain

[2]Department of Biological Sciences, University of Pittsburgh, Pittsburgh PA 15260, USA

[3]Department of Archaeology, Minalloy House. 18-22 Regent Street, University of Sheffield, Sheffield S1 3NJ, United Kingdom.

[4]CEFE UMR 5175, CNRS – Université de Montpellier – Université Paul-Valéry Montpellier – EPHE – 1919 route de Mende, FR-34293 Montpellier, Cedex 5, France.

[5]Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, United Kingdom.

[6]Embrapa Clima Temperado, Pelotas, Rio Grande do Sul, Brazil.

[7]Max Planck Institute for Biogeochemistry, Hans Knöll Str. 10, 07745 Jena, Germany

[8]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany.

[9]Department of Botany, University of Wyoming, 1000 E. University Ave., Laramie, Wyoming 82071 USA.

[10]Laboratoire Ecodiv URA/EA 1293, IRSTEA, FR CNRS 3730 SCALE, UFR Sciences et Techniques, Normandieuniversite, Universite de Rouen, 76821 Mont Saint AignanCedex, France.

[11]Department of Biology, Ecology and Biodiversity, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium.

[12]Landscape Ecology Group, Institute of Biology and Environmental Sciences, Carl von Óssietzky-Strasse 9-11, 26111 Oldenburg, Germany.

[13]Systems Ecology, Departament of Ecological Science, Faculty of Earth and Life Sciences, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.

[14]UMR Ecologie et Ecophysiologie Forestières, Faculté des Sciences et Technologies, Université de Lorraine, INRA, F-54506, Vandoeuvre-lès-Nancy Cedex, France.

[15]Wageningen Environmental Research (Alterra) - Wageningen University, Droevendaalsesteeg 4, 6700AA Wageningen, The Netherlands.

[16]School of Molecular and Life Sciences, Curtin University, GPBox U1987, WA 6845, Australia.

[17]Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 91501-970, Brazil.

[18]School of Civil & Environmental Engineering, Yonsei University A460 121dong Yonsei-ro 50, Seodaemun-gu, Seoul 03722, Korea.

* Corresponding author: (+34) 914 884 645; E-mail: ruben.milla@gmail.com

**The origins of agriculture were key events in human history, during which people came to depend for their food upon small numbers of animal and plant species. However, the biological traits determining which species were domesticated for food provision and which were not, are unclear. Here, we investigate the phylogenetic distribution of livestock and crops, and compare their phenotypic traits with those of wild species. Our results indicate that phylogenetic clustering is modest for crop species but more intense for livestock. Domesticated species explore a reduced portion of the phenotypic space occupied by their wild counterparts and have particular traits in common. For example, herbaceous crops are globally characterized by traits including high leaf nitrogen concentration and tall canopies, which make them fast growing and proficient competitors. Livestock species are relatively large mammals with low basal metabolic rates, which indicate moderate to slow life histories. Our study therefore reveals ecological differences in domestication potential between plants and mammals. Domesticated plants belong to clades with traits advantageous in intensively-managed high-resource habitats whereas domesticated mammals are from clades adapted to moderately productive environments. Combining comparative phylogenetic methods with ecologically relevant traits has proven useful to unravel the causes and consequences of domestication.**

The plant and animal species domesticated for human food supply represent only a small fraction of global biodiversity. Of around 370,000 extant flowering plants[1], only 1,000-2,000 have undergone some form of domestication for that purpose [2–4]. Similarly, humans have domesticated 20-31 species of mammals for food [5,6], from ~ 5,400 species contemporary to late Paleolithic people[7]. The taxonomic distribution of species used for farming seems non-random[5,8], such that certain families include numerous domesticated species (*e.g.* grasses and legumes among flowering plants, and bovids and camelids among mammals), while many others contain none. An uneven phylogenetic distribution of the species that became domesticated would imply that certain combinations of phenotypic traits are more adaptive for husbandry, if these traits are phylogenetically conserved[9]. However, global comparative analyses between domesticates and wild relatives are rare[10–12] or consider taxonomically and/or geographically restricted groups of species[13,14]. Filling that gap would direct agricultural sciences towards the phylogenetic groups and traits that could be pursued for new food sources. Additionally, investigating such patterns at a global scale, while explicitly linking phylogenetic and trait distributions, would highlight the usefulness of the tools and concepts of evolutionary ecology to address questions at its interface with agricultural sciences and archaeology.

The phenotypes of current livestock and crops are the result of early domestication processes and millennia of unconscious and deliberate selection under farming[15]. Evolution under farming has caused the traits of domesticated species to change under shifting selective forces[13]. For example, local breeding preserves mutants that would otherwise be eliminated by natural selection and thereby offset the sampling effects of early farming[4]. In fact, diversifying selection has promoted remarkable variance in the size of crop seeds or in animal coat colors[5,16]. Conversely, directional selection for productivity has resulted in the convergence of a number of livestock and crop traits, *i.e.* the domestication syndrome[5,9,17](but see [18,19]). Domestication syndrome traits include increased docility or reductions in brain size in livestock[5,6,20] and increased palatability or the loss of seed dispersal mechanisms in crop plants[4,9].

Adaptations to the farming environment might also promote convergent evolution *via* natural selection[13]. For instance, evolution in high-resource croplands should select for suites of traits that enable fast resource acquisition and rapid canopy closure, according to trait-based ecology[21]. A few recent studies have partially supported this view, showing that several crop plants display traits indicative of high competitive ability[14,22–24]. Therefore, if directional and stabilizing selection are strong, the phenotypic diversity of domesticated species will be low, adding to early domestication bottlenecks. In contrast, diversifying selection, associated with centuries-long geographic expansion under farming, is expected to promote the widening of phenotypic spaces[16]. The net effects of the early filtering of wild species, of subsequent domestication processes, and of later crop and livestock evolution, on the phenotypic spaces explored by domesticated species remain unknown. In this paper, we show the results of the first broad-scale phylogenetic analyses addressing whether domesticates are a limited phylogenetic and phenotypic sample of wild plants and animals, and uncovering traits that distinguish domesticated species from wild species.

We used phylogenetic comparative methods to investigate the phylogenetic patterns of domestication events, and to ask whether domesticates are phenotypically distinct from their wild relatives. We did this by compiling and analyzing two large datasets. First, we compiled a database on the distribution of species domesticated for food across mammal and angiosperm families and genera. With that dataset we investigated evolutionary patterns of the relative abundance of domesticated species (proportion of all domesticated species within a given genus or family), and of domestication frequencies (proportion of all species in a genus or family that were domesticated) across mammal and angiosperm phylogenies. Second, we put together a large-scale database of three key phenotypic traits for farm mammals (size-corrected basal metabolic rate, adult body mass and neonate body mass) and crop plants (leaf nitrogen content, plant canopy height and seed dry mass) across a broad sample of domesticated (23 mammals, 181 angiosperms) and wild species (885 mammals, 2,943 angiosperms). Traits were selected based on their key functional relevance for resource-use-acquisition, life history and ecological strategies, both for domesticated and for wild species[21,25–28]. Using this second dataset, we compared the phenotypic spaces of domesticates to those of their wild relatives. We set out to address three questions: 1) How are food domesticates distributed across the phylogenies of mammals and angiosperms? 2) Do livestock and crop species have particular phenotypic profiles, when compared to their wild counterparts? And, if so, 3) Do the phenotypic traits of domesticated species fall within the trait space exhibited by wild species, or do they extend their phenotypic space beyond the boundaries set by wild plants and animals?

## Results

Some families and genera contain more livestock or crop species than others. Livestock species were found in only ten families. In particular, Bovidae harbour ~40 % of domesticated species (Supplementary Table 1), and only 22 genera of mammals contain domesticated species (Supplementary Table 1). In contrast, crop species were distributed across 120 families and 453 genera of angiosperms (Supplementary Data 1). Fabaceae, *Solanum*, and Poaceae are examples of taxa yielding high proportions of crops. The abundances of domesticated species were far from randomly distributed across families and growth forms (plants) or dietary types (mammals), both for mammals and angiosperms (Supplementary Table 2). We next investigated whether the above deviations from a random distribution were phylogenetically structured.

In general, we found that the phylogenetic structure of domestication abundances and frequencies was modest for plants, though stronger for mammals. First, Local Indicators of Phylogenetic Affinity (LIPAs) indicated that ca. 90% of the plant families hosting domesticated species were randomly distributed with respect to the domestication status of their phylogenetic neighbourhood (Supplementary Data 2). Approximately 10% of angiosperm families departed from a non-significant LIPA score (Supplementary Data 2). However, such departure signalled over-dispersion (*i.e.* focal family surrounded by families without domesticated species more than the random expectation), rather than clustering (Supplementary Data 2). Interestingly, that ca. 10% of angiosperm families include Fabaceae, Poaceae, Rosaceae, Solanaceae or Asteraceae, which host crops of high agricultural relevance. For mammals, four (Suidae, Cervidae, Caviidae, and Cunilidae) out of ten families with domesticated species had at least one LIPA score indicating phylogenetic clustering, either for abundances or frequencies. Second, we investigated phylogenetic clustering at the scale of the whole phylogenetic tree. When two contrasting evolutionary models were compared (*i.e.* a Brownian motion model of evolution, representing strong phylogenetic structure, *vs.* a star phylogeny, representing full phylogenetic independence), phylogenetically independent models showed better fit to the data than Brownian motion models, both for mammals and angiosperms (Supplementary Tables 3 and 4). Finally, global phylogenetic clustering was investigated with a gradual approach (phylogenetic signal), which complemented the binary (non-phylo vs phylo) comparison of phylogenetic models above. The frequency of domestication events generally showed a phylogenetic signal (Figure 1 and Supplementary Figure 1, but see angiosperm families in Figure 1). Domestication abundance, instead, showed low or no phylogenetic signal in angiosperms, low signal in mammals at the family scale, and high signal when mammals were examined separately for each dietary type (Figure 1 and Supplementary Figure 1).

The subset of domesticated species used for phenotypic space analyses covered a wide range of phylogenetic and geographic origins (Figure 2). In spite of this taxonomic and geographic diversity, domesticated species were distributed across a portion of the phenotypic space generally occupied by wild species, and rarely fell beyond the bounds set by wild mammals and plants (Figures 3 and 5; and Supplementary Tables 5 and 6, but see ruminant livestock in Figure 3). Livestock occupied a small subset of the phenotypic spaces of wild mammals (Figure 3 and Supplementary Table 5). Within the common phenotypic boundary occupied by wild and domesticated mammals, livestock species were, on average, larger as adults and gave birth to larger neonates, but had lower basal metabolic rates, compared to their wild counterparts (Figure 4 and Supplementary Tables 7 and 9). Those phenotypic biases were upheld when investigated separately for ruminants and non-ruminants, though domestic ruminants lied mostly outside the phenotypic boundaries of wild ruminants. In contrast, domesticated crops have been selected from a wide range of botanical diversity in the three focal traits (Figure 5 and Supplementary Table 6). An exception to that pattern is the small phenotypic space occupied by domesticated graminoids (grass-like monocot plants), in comparison with their wild analogues (Figure 5 and Supplementary Table 6). However, although crops were phenotypically diverse, they occupied some regions of the phenotypic space more than others, which lead to phenotypic differentiation when compared to wild plants. Specifically, herbaceous crops, both graminoid and non-graminoid, were generally larger plants with larger seeds, and with leaves with higher nitrogen content, than their wild counterparts (Figure 6 and Supplementary Tables 8 and 10). Woody crops were more similar to wild woody plants, though they consistently had larger seeds (Figure 6 and Supplementary Tables 8 and 10). The phenotypic departure of domesticated species from the trait medians of their wild counterparts was generally unrelated to the differences in geographic origin, climate at geographic origin, or antiquity of domestication of crop and livestock species (Supplementary Table 11).

## Discussion

Our results provide the first quantitative global test of the long-standing hypothesis that domestication events have a strong phylogenetic structure. We found only weak evidence for phylogenetic clustering in crops, but stronger evidence of such clustering in livestock species. Interestingly, the non-random phylogenetic distribution of species that became domesticated was associated with non-random phenotypic spaces of crops and farm mammals. Livestock species had moderate to low basal metabolic rates, gave birth to large offspring and were large adults. Herbaceous crops had high leaf nitrogen content (an indicator of photosynthetic rates), were large as adults, and produced large seeds. These results show that domesticated mammals and plants occupy specific portions of the spectra of phenotypic variation[21,25–28]. Despite such phenotypic differentiation, and in spite of substantial trait evolution during domestication[16], domesticated species were rarely positioned outside the bounds of the phenotypic spaces set by the wild species of their kind, excluding ruminant livestock. These findings have important implications for understanding the potential of wild species for farming, the patterns of phenotypic convergence under domestication, and the adaptation of wild species to the environmental conditions of farming habitats[13,18,20,29,30]. In addition, we demonstrate that a macroevolutionary approach, scarcely embraced in the domestication literature, has the strong potential to advance this field.

Our analyses showed that domesticated mammals represent a small sub-sample of the total phenotypic variation displayed by wild animals. This was expected, since livestock species are a very small fraction of all contemporary mammals. More unexpectedly, farmed mammals had mid-to-low basal metabolic rates and were mid-to-large adults and neonates. These results portray domesticated mammals as moderate to slow life history strategists, *i.e.* species with intermediate body temperatures, with moderately long juvenile periods, giving birth to few but relatively large offspring, and living for reasonably long time spans, accordingly to the fast-slow life-history framework[26,27]. Low basal metabolic rates, which might entail slow relative growth rates[31], are adaptive in unproductive and unpredictable environments in ruminants[32], rodents[33] and mammals in general[27,28]. In addition, a moderate-slow lifestyle might genetically associate with behavioural traits that are critical to animal domestication, such as boldness, tameness, or sociality[33–35], which remains to be investigated. Interestingly, many domesticated mammals evolved body size reduction after domestication[29,36]. This suggests that the moderate-slow lifestyle of livestock identified in our current work is largely the result of early selection of wild animals, rather than of further evolution under farming. In livestock species that show such body size reductions during domestication, decreases in sexual dimorphism were also reported, which, following Rensch´s rule[37], might account for their overall smaller adult size [5,29,38]. Further studies advancing this line of enquiry should consider the wild progenitors of livestock species, account for sexual dimorphism, and address species domesticated for purposes beyond food provision, which clearly display body size reduction after domestication (*e.g.* dogs or cats).

For crops, our results for non-woody plants are compatible with hypotheses positing that early human selection favoured traits advantageous in the fertile, disturbed habitats surrounding human settlements and early agricultural fields[5,9,30]. Herbaceous crops occupied only a portion of the phenotypic space of their wild counterparts, suggesting habitat filtering[39]. This is in line with previous case studies reporting that crops are a subsample of the phenotypic variation found in nature, and have not surpassed the biological limits observed for wild plants[13,23,40,41], which suggests limitations of artificial selection to move phenotypes beyond what is observed in the wild. Additionally, the phenotypic profiles described here indicate that herbaceous crops are fast growing species (high leaf nitrogen content) and proficient competitors in resource abundant environments (tall plants and large seeds)[21,25], which would suit the ecological requirements of early agricultural habitats[42]. Such phenotypic differentiation passed unnoticed in the previous literature addressing smaller sets of crops, where crop-specific contrasting results were commonly reported[14,22–24,41]. This highlights the relevance of a global approach and the usage of the comparative method in this field. Woody crops yielded large seeds but, in contrast to herbs, were neither tall species nor species with high nitrogen content in their leaves. Multiple explanations might account for such discrepancy between growth forms, including trade-offs in resource allocation to fruit tissue, to vegetative growth, and to the maintenance of woody tissues, the clonal mode of propagation common to woody crops, or crop uses[4,9,18,43].

A direct comparison of the phenotypic spaces of plants and mammals yielded an additional relevant insight. In plants, metabolism and size are largely decoupled[25]. In contrast, the evolution of metabolic rates and body size are coordinated in mammals[44]. Our trait analyses were consistent with these patterns, both for wild and domesticated species

(Figures 3 and 5). The phenotypes of domesticates therefore evolve under different constraints in plants and mammals. As a consequence, we found wider combinations of traits and larger phenotypic spaces for crops than for livestock. Endothermy, design of the vascular system, and size-metabolism constraints might impede the simultaneous maximization of mass-specific growth rates and body size in mammals[45], both of which are desirable traits for productive husbandry. Within these constraints, human selection for farming favoured animals with intermediate-high sizes, although at the cost of low metabolic rates, and thus probably modest relative growth rates. Breeding livestock that overcome size-metabolism constraints are therefore expected to be challenging. On the other hand, crop plants occupied a wider part of the tradeoff-free spectrum of metabolism (leaf nitrogen) *vs.* size (plant and seed sizes), in line with their wild counterparts. Plant modularity and nitrogen transfer among modules, which underpin such uncoupling between metabolism and size[46], might thus promote the greater phenotypic diversity of crops than livestock mammals, even when considered within plant growth-forms. Breeding to simultaneously optimize variation in plant and organ sizes, and variation in metabolic rates, might be tradeoff-free in plants.

Finally, we highlight two limitations of the current work. First, trait data come from plants and animals sampled in their typical habitats, which are different among species, and are undoubltley so among organisms living in farm vs wild habitats. Thus, the phenotypic patterns encountered here came from a mixture of genotypic and environmental drivers, whose relative importance is plainly indistinguisible for large scale macroecological studies. However, the few experiments that grew sets of crops and of their wild progenitors in common gardens, and phenotyped the types of traits that we measured here, tended to concur with our results, which suggests a strong genetic component at play [22,40,47].

Therefore, it will be necessary to take these experimental approaches one step further in terms of number of species and phylogenetic breadth, both for wild and domesticated organisms. Second, we note that, for both domesticated and wild species, intraspecific variation was not considered here. Such variation, even if unmanageable to explore systematically at the macroevolutionary scale of the current study, might expand trait spaces greatly[48]. Leaf trait intraspecific variation in sunflower, wheat, maize and coffee was recently reported, and occupies a fair portion of wild phenospaces, though very rarely spreading outside wild boundaries [40,49,50]. We supplemented those leaf trait patterns available in the literature with *ad hoc* analyses for seed and plant sizes (Supplementary Figure 2). Similarly to leaf traits, intraspecific variation in size traits is constrained within wild envelopes in sunflower, soybean, and barley (Supplementary Figure 2). However, maize, as a crop species in which the seed-plant size centroid is outside the phenosphace of its corresponding wild analogs (Figure 5), expands most of its intraspecific variation outside wild boundaries (Supplementary Figure 2). In our view, these analyses, and the available literature, are still too scant to reach solid solid conclusions on the role of intraspecific variation in trait differences between wild and domesticated organisms. Thus, investigating how and to what extent diversifying breeding of domesticates expands phenospaces is needed to bridge the macro- and the micro-evolutionary scales.

This study placed domesticates within their broader botanical and zoological context, which facilitates appreciation of the qualities and potentials of the species that support our food system, and could help in the search for suitable future domesticates. Suitable phenotypes among mammals include moderate-slow life histories, while fast growth traits and large size dominate among herbaceous crops. Further, the usage of a phylogenetic comparative approach, which was seldom embraced in the domestication literature (but see[12,14,51]), provided unique insights, that are unattainable at smaller scales. Overall, our work indicates that certain phylogenetic clades and phenotypic profiles have been more exploited than others for provisioning human food, and that such filtering was based on strong, conscious or unconscious, early selection at human settlements. Future work should investigate biogeographic and historical determinants, disentangle genotypic from environmental drivers, and address the microevolutionary scale, of the broad phylogenetic and phenotypic patterns of differences between domesticated and wild kins revealed here.

## Methods

### Study system

We explored macroevolutionary patterns of the distributions of species domesticated for human food, and compared their phenotypic trait space occupancy with that of wild species. We included the broadest possible diversity of mammals and angiosperms farmed for human food provision, with distinct domestication histories and intensities, phylogenetic affinities, and phenotypic profiles (see Supplementary Table 1, Supplementary Data 1, and Figures 1-6).

### Collection of data on the distribution and abundance of angiosperms and mammals farmed for food

In order to assess patterns of distribution and abundance of food domesticates across phylogenies, we compiled the abundances and identities of domesticated species within mammals and angiosperms, at the family and genus levels. We aimed to build a comprehensive working list of all putative species domesticated for food. Therefore we included all species for which farming could be substantiated, as an indicator of some degree, even if incipient, of domestication.

We used the literature to build a preliminary list of mammals farmed for food[5,6,52,53]. We excluded most anectodal evidence of deer and antelopes farming, but for the sake of inclusivity we considered species like reindeer, sika deer, moose, bison, muskox or common eland, which are regularly farmed regionally, and thus should include incipiently domesticated populations. We also included recent incipient domesticated species for food like the African giant rat, or the South American paca. The set of mammals comprised 27 domesticated species, distributed across 22 genera and 10 families. Taxonomy was checked using the *taxize* R package[54].

For plants, we compiled a list of all putative domesticated species from the literature[2,4,18,55]. From that list, we extracted taxa for which cultivation could be demonstrated, and filtered that extract by species used for human food or forage. We include forage species because human food supply depends on livestock feeding, and a substantial part of the agricultural land is devoted to that usage. To assign usage for food or forage we primarily used two studies[2,56], supplemented with http://www.pfaf.org and with other sources when needed. Plant taxonomy was standardized using the *Taxonstand* R package[57], synonymous names were cleaned, and binomials were attributed to families using The Plant List (http://www.theplantlist.org/). Thirty-five species were taxonomically unresolved, and their provisional binomials were used. Our final 944 species list should include the vast majority of angiosperms known to have been cultivated for human food. The species of our list belonged to 453 genera and 120 angiosperm families.

Domesticated mammals were grouped into two dietary types, *i.e.* ruminants and non-ruminants, based on the MammalDIET database[58], and on information at http://www.ultimateungulate.com/ungulates.html - Ruminantia. The ruminant category also included pseudo-ruminants, *i.e.* Camelidae and Hippopotamidae. Domesticated plants were grouped by growth form into herbaceous, graminoids (Poaceae, Cyperaceae and Juncaceae), and woody plants (shrubs, trees, woody vines and tree-like species). Growth forms were assigned using the TRY database[59], the Global Woodiness Database[60], and were supplemented species-wise with primary literature when not available in those sources.

**Selection and compilation of phenotypic traits**

We selected a set of three traits for mammals and three traits for angiosperms that are functionally analogous. Selection of traits used the following criteria: 1) previous evidence of domestication effects on those traits[14,22,29,30,36]; 2) functional relevance for basic metabolism, resource-use, competition and reproductive strategies; and 3) availability of data, both for domesticated and for wild species. By functional analogy, the selected traits can be grouped into: 1) plant canopy height (m) and mammal adult body mass (g) as proxies for adult size and competitive ability for resources[61,62]; 2) leaf nitrogen content (mg N $mg^{-1}$), and size-corrected mammalian basal metabolic rate (ml $O_2$ $h^{-1}$ $g^{-1}$) as proxies of photosynthetic and metabolic rate, respectively[26,63]; and 3) seed dry mass (mg) and neonate body mass (g) as proxies for offspring size and likelihood of survival [64,65].

The two trait datasets were assembled separately for mammals and angiosperms. The dataset on mammal traits was compiled from the PanTHERIA database for adult and neonate body mass[66], and from [67] for basal metabolic rate (supplemented with PanTHERIA if basal metabolic rate was unavailable at [67]). The dataset on mammal traits comprised 480 species, including 23 domesticated species (see Supplementary Table 1 for the identity of domesticated species). For domesticated plants, trait data were not available for all 944 crops species. Thus, we started by retrieving data from the 203 species list of domesticated plants published in Meyer et al.[18], which was built to maximize crop diversity. This helped to avoid biases in growth form, taxonomy, or agricultural relevance. That list of crop taxa was further supplemented, again using criteria to maximize diversity and filtering by the availability of trait data. Then, data for wild angiosperms were added. Overall, the sources for angiosperm trait data were: 1) the TRY database[59] (https://www.try-db.org, accessed 2016-11-13; 2) literature searches for wild species incompletely recorded or not present in the TRY database; 3) literature searches for trait data of crop species, which are mostly absent from TRY[68]; 4) our own data already collected on crops and other wild species (see Data availability section). The final angiosperm trait dataset comprised 3,124 species, including 181 domesticated species (see Supplementary Data 1 for the identity of domesticated species).

Plant species names were standardized using the *Taxonstand* R package[57], and were attributed to families according to The Plant List (http://www.theplantlist.org/). A majority of crop binomials are synonymous to the wild genotypes of their wild progenitors. Therefore, to decide whether a given observation of a crop related binomial was attributable to a crop or a synonymous wild species we used the following criteria. First, we looked for explicit statements in the original publication or database on whether the studied taxa were crop or wild. If uncertain, an observation was assigned to “wild” if the study was observational and was conducted under natural field conditions, or if the seeds for an experiment were collected from wild stocks. In contrast, an observation was assigned to "crop" if the seeds came from commercial companies, cultivars or varieties or if studies had been conducted in an agricultural setting, and no explicit mention to wild status was found.

### Data handling prior to analyses

Our angiosperm dataset had 1.51 % missing data (leaf nitrogen 1.70 %, plant height 1.27%, and seed mass 1.58 %). Since several plots and analyses involved the joint use of two or more traits, we adopted a multiple imputation approach to deal with missing data, following recommended procedures[69,70]. We generated ten complete datasets using the Predictive Mean Matching method (PMM) of the *MICE* package[71]. Phylogenetic relatedness (built as described below) was incorporated into the imputation procedure as phylogenetic orthogonal eigenvectors[72]. Results reported in the main body of the paper are from averaged imputed data of those ten complete datasets. A dataset without imputed data, and thus with a slightly reduced sample size, was used to test for robustness and sensitivity to our data handling procedures. The results of sensitivity analyses were consistent with those shown in the main text (Supplementary Note, and Supplementary Tables 12-14). The mammalian traits dataset lacked basal metabolic rate data for six species, which were estimated using the phylogenetically-corrected allometric scaling of adult body mass to basal metabolic rate available at [67] ($BMR = 2.382m^{0.729}$). All continuous variables were log-transformed prior to analyses. An exception was seed mass, which was log-generalized transformed because a few crops do not yield seeds (*Musa acuminata, Vaccinium corymbosum and Allium sativum*). This latter procedure is recommended when data contain zeros, and the smallest positive value is not close to one.

### Macroevolutionary patterns in the abundance and frequency of food domesticates

We performed separate analyses on the relative abundance of domesticated species, and domestication frequency, at the family and genus levels. Relative abundance, at family and genus level, was calculated as the number of domesticated species in a particular family or genus divided by the total number of domesticated species. Because this metric does not account for species richness within clades, we also calculated a domestication frequency metric as the number of domesticated species in a focal family or genus per total number of species in that same family or genus. These two metrics inform about different features of the distribution of domesticated species. Species richness at family and genus level, needed to compute domestication frequencies, was retrieved from lists of accepted names from Wilson and Reeder (2005) for mammals, and from The Plant List v1.1 (http://www.theplantlist.org/), making use of the *taxonlookup* R package (https://github.com/traitecoevo/taxonlookup), for plants.

Phylogenetic hypotheses were built at the family and genus levels for mammals and for angiosperms separately. Mammal phylogenies were based on Bininda-Emonds et al's[73] megaphylogeny as a backbone. The mammalian family-level tree included 142 families (91% of total mammalian families), while the genus-level tree included 498 genera (39.6 % of total mammalian genera). Seventy-two genera were ruminants and 337 were non-ruminant herbivores (http://tolweb.org). Angiosperm phylogenies were based on the PhytoPhylo megaphylogeny[60,74]. The angiosperm family-level tree included 404 families (97% of total angiosperm families), while the genus-level tree included 7,233 genera (ca. 56 % of total angiosperm genera) (http://www.theplantlist.org/). All families, and all but seven genera with domesticated species, were present in the megaphylogeny. Those seven genera (*Gigantochloa, Nopalea, Parmentiera, Polianthes, Sphenostylis*, *Stizolobium,* and *Vitellaria*) were bound to the PhytoPhylo megaphylogeny based on published phylogenies of their respective families. The genus level tree included 2,745 genera of herbs, 421 of graminoids, 3,500 of woody plants and 567 genera hosting both herbs and woody plants. Genera containing herbs and woody plants were included in both of their respective growth form analyses. Angiosperm phylogenies were fully resolved, but mammalian phylogenies contained some internal polytomies (12 % of nodes in the family-level tree, 24 % in ruminants and 38 % in non-ruminant herbivores). Therefore, analyses were run across 100 randomly resolved family and genus-level mammalian trees.

To assess whether the abundance and frequency of domestication are randomly distributed across mammal and angiosperm families and genera, we performed four complementary analyses. First, we conducted randomisation analyses to test whether the distribution of the abundances of domesticated species per family, and of the frequencies of domestication events, differed from random expectations. Observed kurtosis and skewness of the distribution of abundances were compared to that of 1,000 randomised distributions at each respective level. Second, we computed Local Indicators of Phylogenetic Association (LIPA) based on Local Moran's *I* [75], to detect families surrounded by phylogenetic neighbourhoods with similar or distinctive (positive or negative autocorrelation, respectively) relative abundances of domesticated species or domestication frequencies. For each LIPA score, statistical significance was assigned by performing non-parametric two-sided tests with 999 randomisations. For mammals, LIPAs were averaged across the 100 randomly resolved trees. Third, we calculated the phylogenetic signal of the relative abundances of domesticated species, and of the frequencies of domestication, at the genus and family levels, and separately for mammals and angiosperms. Provided that our data followed either zero-inflated log-normal (abundances), or binomial (frequencies), distributions, we did not calculate standard Pagel's $\lambda$ or Blomberg's *K* metrics, which are meant for continuous data with normal or log-normal distributions. Instead, we computed phylogenetic signal as the phylogenetic heritability parameter (λ) of phylogenetic mixed models (PMMs), where our metrics of domestication were the response, an intercept was set as the sole fixed-effects predictor, and inverse matrices of the phylogenetic distances matrices were the covariance structure terms[76,77]. PMMs allow the specification of family distributions of data deviating from Gaussian. The lambda parameter of such models, specified without meaningful fixed-effect predictors, and without additional covariance structures in the random term, is the phylogenetic signal of the response variable, analogously to a null Phylogenetic Generalized Least Square (PGLS) model[76]. Fourth, we fitted and compared two evolutionary models to test whether relative abundances and domestication frequencies were phylogenetically structured or phylogenetically independent. We used a Brownian motion (BM) model to approximate neutral drift evolution or randomly fluctuating selection[78]. Under BM, relative abundances and frequencies evolve as a random walk through trait space along the branches of the phylogeny, and thus represents strong phylogenetic structuring. BM was compared to a non-phylogenetic model (a star phylogeny), which was used to approximate a phylogenetic independent distribution. To compare both models we used the bias-corrected Akaike Information Criterion ($AIC_c$), and calculated the difference between the $AIC_c$ of the best (smallest $AIC_c$) and the alternative model[79]. In addition, for each model we calculated the $AIC_c$ weights ($AIC_c$-w), with a high $AIC_c$-w indicating a low relative $AIC_c$ for that model and hence higher support[79]. For mammals, all parameters were averaged across 100 randomly resolved trees and the percentage of preferred models was calculated. Phylogenetic signals were computed using the *MCMCglmm* function of the R package *MCMCglmm*[80], setting family distribution as zero inflated poisson for domestication abundances, and as binomial for frequencies of domestication. Evolutionary model fitting was performed with the *FitContinuous* function of the R package *geiger*[81]. Local Moran's *I* was calculated using the *lipaMoran* function of the R package *phylosignal*[75].

## Comparative analyses of phenotypic trait space occupancy of wild and domesticated species

New sets of phylogenetic trees were built at the species level for those mammals and angiosperms included in our traits datasets. The mammal phylogeny for the 480 species with trait data was built from Bininda-Emonds et al's[73] megaphylogeny using Phylomatic v.3[82] (http://phylodiversity.net/phylomatic/pmws). The angiosperm phylogeny for the 3,124 species with trait data was based on the PhytoPhylo megaphylogeny[60,74], and was built using scenario three of the R package *S.PhyloMaker*[74]. To account for phylogenetic uncertainty (20.3 % of unresolved nodes for mammals and 15.3 % for angiosperms), all analyses were performed on 100 randomly resolved trees by using the *multi2di* function of the *ape* R package[83].

To visualize the phenotypic spaces explored by wild and domesticated species we used bivariate phenospaces. Additionally, we used convex hulls to draw the minimum convex envelope for each pair of traits, domestication status and growth form or dietary type[84]. In addition, for each growth form or dietary type, we calculated the area and volume of each three-traits convex hull. To test for significant differentiation in trait space between domesticated and wild species, we performed phylogenetic-corrected MANOVAs and ANOVAs, separately for each growth form or dietary type. Convex hull calculations were performed using function *convhulln* of the R package *geometry*[85]. Phylogenetic MANOVAs/ANOVAs were run with the function *aov.phylo* in the R package geiger[81]. See

Supplementary Methods for tests on whether the phenotypic departure of domesticated species from the trait medians of their wild counterparts was related to differences between domesticates in geographic origin, climate at geographic origin, or antiquity of domestication. All analyses of the paper were conducted in R v3.4.3[86].

## Data availability

All phenotypic traits of mammalian species included in this study are available from the literature (see Methods section). For plants, most data are available from the database TRY[59] (https://www.try-db.org), and all original sources of TRY data are listed in Supplementary References 1. All references for data not included in TRY are available in the Supplementary References 2. Unpublished data owned by R.M. and J.M.B. are available from Supplementary Data 3. Unpublished data from the University of Sheffield database of weed functional attributes can be requested from G.J. Lists of livestock and crop taxa are available at Supplementary Table 1 and Supplementary Data 1, respectively. Phylogenetic trees used in this study are available from Supplementary Data 4. Data on geography and climate at domestication sites are available as Supplementary Data 5.

## References

1. RBG Kew. *The State of the World's Plants Report - 2016*. (Royal Botanic Gardens, Kew, 2016).
2. Zeven, A. C. & De Wet, J. M. J. *Dictionary of cultivated plants and their regions of diversity: excluding most ornamentals, forest trees and lower plants*. (Pudoc, 1982). at <http://agris.fao.org/agris-search/search.do?recordID=XE8282385>
3. Dirzo, R. & Raven, P. H. Global state of biodiversity and loss. *Annu. Rev. Environ. Resour.* **28,** 137–67 (2003).
4. Meyer, R. S. & Purugganan, M. D. Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14,** 840–852 (2013).
5. Zeder, M. A. in *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability* (eds. Gepts, P., Famula, T. R., Bettinger, R. L. & Al, E.) (Cambridge University Press, 2012).

6. Clutton-Brock, J. *A natural history of domesticated mammals*. (Cambridge University Press, 1999).
7. Wilson, D. & Reeder, D. A. *Mammal Species of the World*. (Johns Hopkins University Press, 2005).
8. Gepts, P. *Crop domestication as a long-term selection experiment*. *Plant breeding reviews* **24,** (John Wiley & Sons, Inc, 2010).
9. Blomberg, S. P. & Garland Jr, T. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J. Evol. Biol.* **15,** 899–910 (2002).
10. Diamond, J. Evolution, consequences and future of plant and animal domestication. *Nature* **418,** 700–707 (2002).
11. Salman-Minkov, A., Sabath, N. & Mayrose, I. Whole-genome duplication as a key factor in crop domestication. *Nat. Plants* **2,** 1–4 (2016).
12. Whitehead, S. R., Turcotte, M. M. & Poveda, K. Domestication impacts on plant–herbivore interactions: a meta-analysis. *Philos. Trans. R. Soc. B Biol. Sci.* **372,** 20160034 (2017).
13. Milla, R., Osborne, C. P., Turcotte, M. M. & Violle, C. Plant domestication through an ecological lens. *Trends Ecol. Evol.* **30,** 463–469 (2015).
14. Preece, C. *et al.* Were Fertile Crescent crop progenitors higher yielding than other wild species that were never domesticated? *New Phytol.* **207,** 905–913 (2015).
15. Darwin, C. *The variation of animals and plants under domestication*. **2,** (O. Judd, 1868).
16. Abbo, S. *et al.* Plant domestication versus crop evolution: a conceptual framework for cereals and grain legumes. *Trends Plant Sci.* **19,** 351–360 (2014).
17. Harlan, J. R. *Crops and man*. (American Society of Agronomy, 1992).
18. Meyer, R. S., DuVal, A. E. & Jensen, H. R. Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytol.* **196,** 29–48 (2012).
19. Sánchez-Villagra, M. R., Geiger, M. & Schneider, R. A. The taming of the neural crest: a developmental perspective on the origins of morphological covariation in domesticated mammals. *R. Soc. open Sci.* **3,** 160107 (2016).
20. Larson, G. & Fuller, D. Q. The evolution of animal domestication. *Annu. Rev. Ecol. Evol. Syst.* **45,** 115–136 (2014).
21. Reich, P. B. The world-wide 'fast–slow'plant economics spectrum: a traits manifesto. *J. Ecol.* **102,** 275–301 (2014).
22. Milla, R., Morente-López, J., Alonso-Rodrigo, J. M., Martín-Robles, N. & Chapin III, F. S. Shifts and disruptions in resource-use trait syndromes during the evolution of herbaceous crops. *Proceedings. Biol. Sci.* **281,** 20141429 (2014).
23. Tribouillois, H. *et al.* A functional characterisation of a wide range of cover crop species: growth and nitrogen acquisition rates, leaf traits and ecological strategies. *PLoS One* **10,** e0122156 (2015).
24. Milla, R. & Matesanz, S. Growing larger with domestication: a matter of physiology, morphology or allocation? *Plant Biol.* **19,** 475–483 (2017).
25. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529,** 167–171 (2015).
26. Ricklefs, R. E. & Wikelski, M. The physiology/life-history nexus. *Trends Ecol. Evol.* **17,** 462–468 (2002).
27. Lovegrove, B. G. The zoogeography of mammalian basal metabolic rate. *Am. Nat.*

**156,** 201–219 (2000).

28. Lovegrove, B. G. The influence of climate on the basal metabolic rate of small mammals: a slow-fast metabolic continuum. *J. Comp. Physiol. B* **173,** 87–112 (2003).
29. Zohary, D., Tchernov, E. & Horwitz, L. K. The role of unconscious selection in the domestication of sheep and goats. *J. Zool.* **245,** 129–135 (1998).
30. Cunniff, J. *et al.* Functional traits differ between cereal crop progenitors and other wild grasses gathered in the Neolithic Fertile Crescent. *PLoS One* **9,** e87586 (2014).
31. Biro, P. & Stamps, J. Are animal personality traits linked to life-history productivity? *Trends Ecol. Evol.* **23,** 361–368 (2008).
32. Dittmann, M. T. *et al.* Characterising an artiodactyl family inhabiting arid habitats by its metabolism: low metabolism and maintenance requirements in camelids. *J. Arid Environ.* **107,** 41–48 (2014).
33. Careau, V., Bininda-Emonds, O. R. P., Thomas, D. W., Réale, D. & Humphries, M. M. Exploration strategies map along fast–slow metabolic and life-history continua in muroid rodents. *Funct. Ecol.* **23,** 150–156 (2009).
34. Réale, D. *et al.* Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **365,** 4051–4063 (2010).
35. Found, R. & St. Clair, C. C. Ambidextrous ungulates have more flexible behaviour, bolder personalities and migrate less. *Open Sci.* **4,** 160958 (2017).
36. Tchernov, E. & Horwitz, L. K. Body size diminution under domestication: unconscious selection in primeval domesticates. *J. Anthropol. Archaeol.* **10,** 54–75 (1991).
37. Rensch, B. *Evolution above the species level*. (Columbia University Press, 1960).
38. Vigne, J.-D. The origins of animal domestication and husbandry: a major change in the history of humanity and the biosphere. *C. R. Biol.* **334,** 171–181 (2011).
39. Li, Y. *et al.* Habitat filtering determines the functional niche occupancy of plant communities worldwide. *J. Ecol.* 1–9 (2017). doi:10.1111/1365-2745.12802
40. Donovan, L. A., Mason, C. M., Bowsher, A. W., Goolsby, E. W. & Ishibashi, C. D. A. Ecological and evolutionary lability of plant traits affecting carbon and nutrient cycling. *J. Ecol.* **102,** 302–314 (2014).
41. Rotundo, J. L. & Cipriotti, P. A. Biological limits on nitrogen use for plant photosynthesis: a quantitative revision comparing cultivated and wild species. *New Phytol.* **214,** 120–131 (2017).
42. Denison, R. F. *Darwinian agriculture: how understanding evolution can improve agriculture*. (Princeton University Press, 2012).
43. Miller, A. J. & Gross, B. L. From forest to field: perennial fruit crop domestication. *Am. J. Bot.* **98,** 1389–1414 (2011).
44. Brown, J. & West, G. *Scaling in Biology*. (Oxford Univ Press, 2000).
45. West, G. B., Brown, J. H. & Enquist, B. J. General model for the origin of allometric scaling laws in biology. *Science (80-. ).* **276,** 122 (1997).
46. Reich, P., Tjoelker, M., Machado, J.-L. & Oleksyn, J. Universal scaling of respiratory metabolism, size and nitrogen in plants. *Nature* **439,** 457–461 (2006).
47. Preece, C. *et al.* How did the domestication of Fertile Crescent grain crops increase their yields? *Funct. Ecol.* **31,** 387–397 (2017).
48. Vasseur, F., Violle, C., Enquist, B. J., Granier, C. & Vile, D. A common genetic

basis to the origin of the leaf economics spectrum and metabolic scaling allometry. *Ecol. Lett.* **15,** 1149–1157 (2012).

49. Martin, A. R. *et al.* Inter- and intraspecific variation in leaf economic traits in wheat and maize. *AoB Plants* **10,** 1–14 (2018).
50. Martin, A. *et al.* Intraspecific trait variation across multiple scales: the leaf economics spectrum in coffee. *Funct. Ecol.* **31,** 604–612 (2017).
51. Diamond, J. *Guns, Germs, and Steel: the Fates of Human Societies*. (W.W. Norton, 1997).
52. Smythe, N. & Brown de Guanti, O. *Domestication and husbandry of the paca (Agouti paca) (FAO conservation guide: 26)*. (Food and Agriculture Organization of the United Nations, 1995).
53. Moreira, J. R. & Pinheiro, M. S. in *Capybara: Biology, Use and Conservation of an Exceptional Neotropical Species* (ed. Moreira, J. R. et al.) 333–344 (Springer, 2013).
54. Chamberlain, S. *et al.* taxize: Taxonomic information from around the web. (2016).
55. Vincent, H. *et al.* A prioritized crop wild relative inventory to help underpin global food security. *Biol. Conserv.* **167,** 265–275 (2013).
56. Mansfeld, R. & Hanelt, P. *Mansfeld's Encyclopedia of Agricultural and Horticultural Crops.* (Springer, 2001).
57. Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. & Golicher, D. J. Taxonstand: An R package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* **3,** 1078–1083 (2012).
58. Kissling, W. D. *et al.* Establishing macroecological trait datasets: digitalization, extrapolation, and validation of diet preferences in terrestrial mammals worldwide. *Ecol. Evol.* **4,** 2913–2930 (2014).
59. Kattge, J. *et al.* TRY–a global database of plant traits. *Glob. Chang. Biol.* **17,** 2905–2935 (2011).
60. Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing environments. *Nature* **506,** 89–92 (2014).
61. Moles, A. T. *et al.* Global patterns in plant height. *J. Ecol.* **97,** 923–932 (2009).
62. Glazier, D. S. & Eckert, S. E. Competitive ability, body size and geographical range size in small mammals. *J. Biogeogr.* **29,** 81–92 (2002).
63. Wright, I. J. *et al.* The worldwide leaf economics spectrum. *Nature* **428,** 821–827 (2004).
64. Moles, A. T. *et al.* A brief history of seed size. *Science (80-. ).* **307,** 576–580 (2005).
65. Riesch, R., Martin, R. A., Lerp, H., Plath, M. & Wronski, T. Size and sex matter: reproductive biology and determinants of offspring survival in Gazella marica. *Biol. J. Linn. Soc.* **110,** 116–127 (2013).
66. Jones, K. E. *et al.* PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90,** 2648 (2009).
67. Genoud, M., Isler, K. & Martin, R. Comparative analyses of basal rate of metabolism in mammals: data selection does matter. *Biol. Rev.* **93,** 404–438 (2018).
68. Martin, A. R. & Isaac, M. E. Plant functional traits in agroecosystems: a blueprint for research. *J. Appl. Ecol.* **52,** 1425–1435 (2015).
69. Nakagawa, S. & Freckleton, R. P. Missing inaction: the dangers of ignoring missing data. *Trends Ecol. Evol.* **23,** 592–596 (2008).
70. Penone, C. *et al.* Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods Ecol. Evol.* **5,** 961–970 (2014).

71. van Buuren, S. & Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **45,** 1–67 (2011).
72. Santos, T., Diniz-Filho, J. A. F. & Rangel, T. F. PVR: Computes phylogenetic eigenvectors regression (PVR) and phylogenetic signal-representation curve (PSR) (with null and Brownian expectations). (2013).
73. Bininda-Emonds, O. R. P. *et al.* The delayed rise of present-day mammals. *Nature* **446,** 507–512 (2007).
74. Qian, H. & Jin, Y. An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.* **9,** 233–239 (2016).
75. Keck, F., Rimet, F., Bouchez, A. & Franc, A. phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecol. Evol.* **6,** 2774–2780 (2016).
76. de Villemereuil, P. & Nakagawa, S. in *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology* (ed. Garamszegi, L.) 287–304 (Springer-Verlag, 2014).
77. Hadfield, J. & Nakagawa, S. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* **23,** 494–508 (2010).
78. Felsenstein, J. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* **19,** 445–471 (1988).
79. Burnham, K. P. & Anderson, D. R. *Model selection and multimodel inference: a practical information-theoretic approach*. (Springer-Verlag, 2002). doi:10.1007/b97636
80. Hadfield, J. MCMC methods for multi-response generalised linear mixed models: The MCMCglmm R package. *J. Stat. Softw.* **33,** 1–22 (2010).
81. Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E. & Challenger, W. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24,** 129–131 (2007).
82. Webb, C. O. & Donoghue, M. J. Phylomatic: tree assembly for applied phylogenetics. *Mol. Ecol. Resour.* **5,** 181–183 (2005).
83. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20,** 289–290 (2004).
84. Cornwell, W. K., Schwilk, D. W. & Ackerly, D. D. A trait-based test for habitat filtering: Convex hull volume. *Ecology* **87,** 1465–1471 (2006).
85. Barber, C. B. *et al.* geometry: mesh generation and surface tessellation. Version 0.3-4. (2014).
86. Core Team, R. R: A language and environment for statistical computingNo Title. (2016). at <https://www.r-project.org/>

## Acknowledgements

R.M., J. C-L. and J.M.B. were funded by grants CGL2014-56567-R, CGL2017-83855-R and PCIN-2014-053 (Ministerio de Economia y Competitividad, MINECO, Spain), and

Eco-serve project (Biodiversa–FACCE, Horizon 2020, European Union). The study has been supported by the TRY initiative on plant traits (http://www.try-db.org). The TRY initiative and database is hosted, developed and maintained by J. Kattge and G. Bönisch (Max Planck Institute for Biogeochemistry, Jena, Germany). TRY is currently supported by DIVERSITAS/Future Earth and the German Centre for Integrative Biodiversity Research (iDiv) Halle-JenaLeipzig. R.M. thanks to Carlos F. Ingala from Universidad Rey Juan Carlos. E.S. thanks FAPESP/BIOTA program for financial support. J.K. thanks BACI (EU grant ID 640176). T.H. thanks the support of Australian Research Council (DP130013029). V.D.P. was supported by CNPq, Brazil, grant nº 307689/2014-0.

**Author contributions**

R.M. and J.M.B. designed the study and compiled the data, R.M., J.M.B., J. C-L. and M.M.T. performed statistical analyses. R.M. and J.M.B. wrote a first draft of the paper. M.M.T., G.J., C.P.O. and C.V. extensively revised drafts. All authors contributed to the writing of, and approved, the final version.

**Competing interests**

The authors declare no competing financial interests.

## FIGURE LEGENDS

**Figure 1. Distribution of the abundance of food domesticates and frequency of domestication events across mammalian and angiosperm families.** Length of blue bars are relative abundance of domesticated species (proportion of all domesticated species that are found within a given family), and of domestication frequencies (proportion of all species in a family that were domesticated). Raw data can be found in Supplementary Table 1 and Supplementary Data 1, and family identities in the phylogeny can be browsed online in the high resolution version of the Figure. Colors of tree edges correspond to domestication abundances or frequencies, according to a gradient of increasing rates from zero (red) to one (blue). Within each phylogeny, the inset indicates the posterior mode of the phylogenetic signal ($\lambda$), together with its 95% credible interval.

**Figure 2: Phylogeographic distribution of the putative place of origin of food domesticates included in phenotypic space analyses.** Locations in the map are medians of GBIF coordinate records for the putative wild progenitor of each domesticate. Data were retrievable for 168 wild progenitors of crops, out of 181 crop species, and for all of the 23 wild progenitors of mammal domesticates. Insets: Mantel test statistics for the correlation between phylogenetic and geographic distance matrices. See Supplementary Methods for further details.

**Figure 3. Phenosphospace occupancy of livestock and wild mammals.** Separate plots are shown for all mammals, ruminants, and non-ruminant herbivores. Grey dots and red dots are wild and domesticated mammals, respectively. Black and red polygons are convex hulls for wild and domesticated mammals, respectively. Numbers in the insets are % of

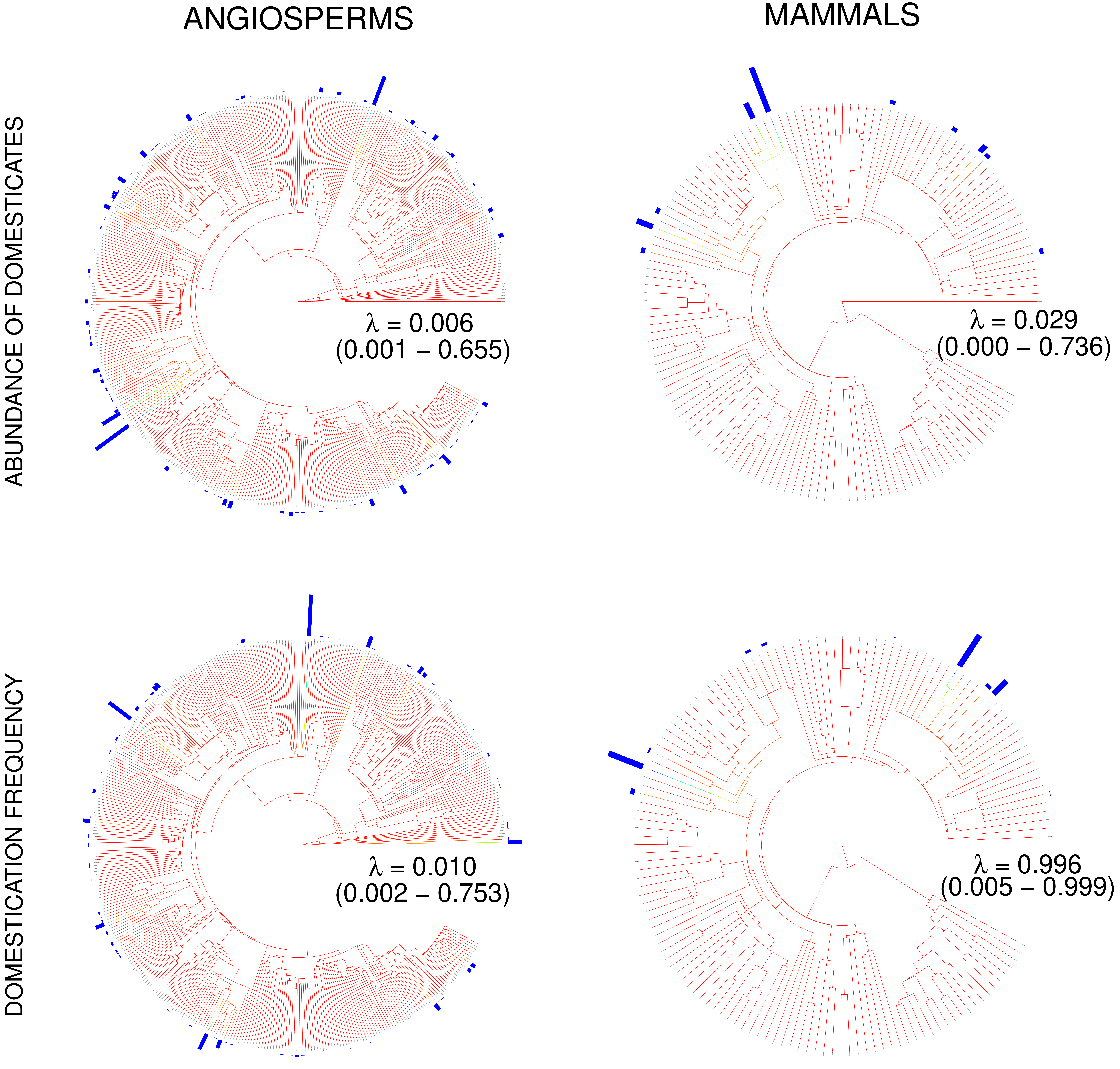
convex hull area of domesticates outside the wild boundary (light red), of domesticates inside the wild boundary (strong red-grey), and of wild space occupied by domesticates (grey).

**Figure 4. Phenotypic differentiation between livestock and wild mammals.** Separate panels are shown for all mammals, non-ruminants, and ruminants. Asterisks indicate statistically significant differences at phy-P ≤ 0.05 between domestication statuses according to phylogenetic Anovas (Supplementary Table 7). Center line, median; box limits, upper and lower quartiles; whiskers, lowest/highest datum still within 1.5*IQR; points, data exceeding whisker bounds.

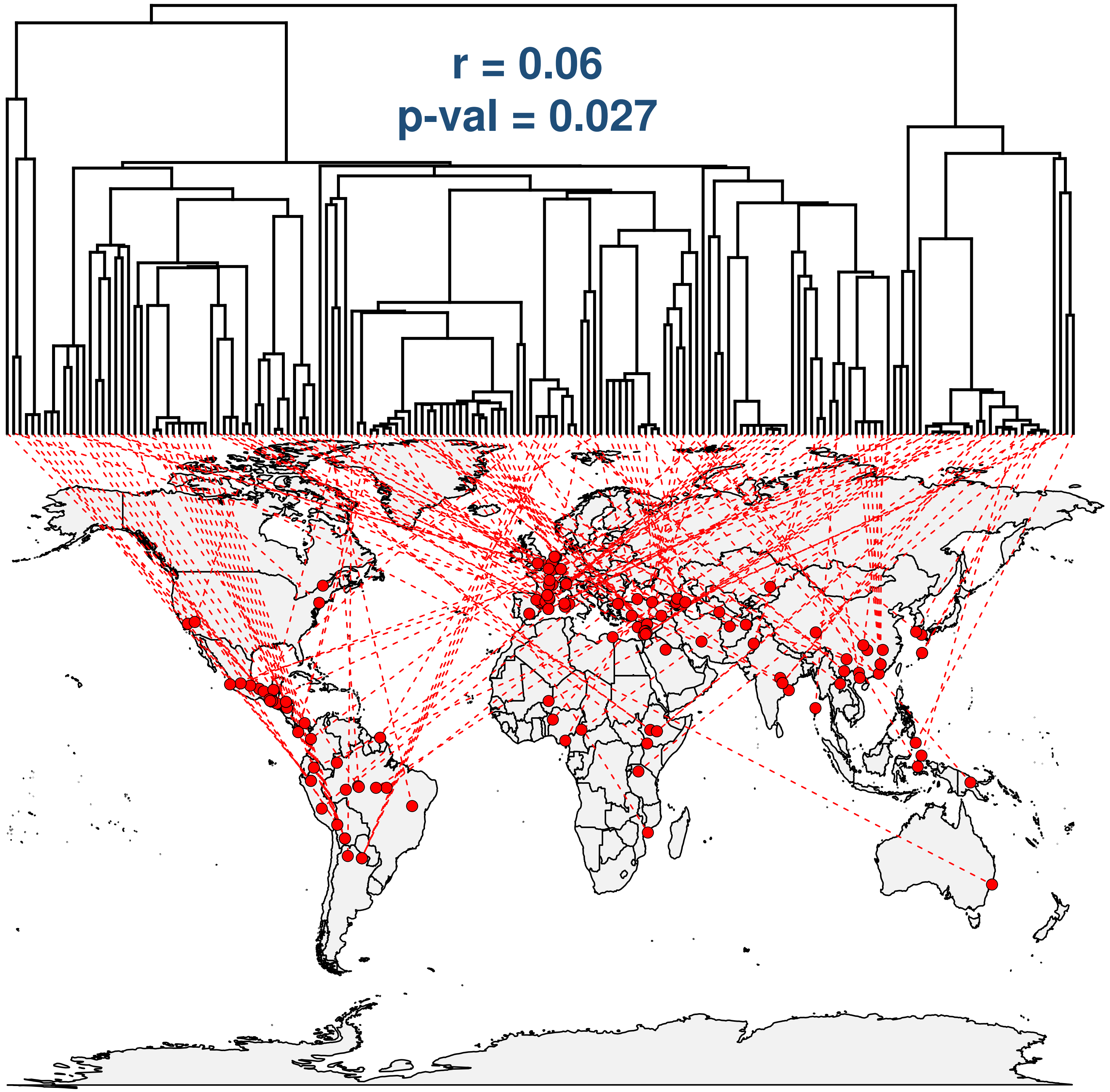
**Figure 5. Phenospace occupancy of crops and wild angiosperms.** Separate plots are shown for all angiosperms and for each growth form (*i.e.* herbaceous, graminoids and woody). Grey dots and red dots are wild and domesticated angiosperms, respectively. Black and red polygons are convex hulls for wild and domesticated angiosperms, respectively. Numbers in the insets are % of convex hull area of domesticates outside the wild boundary (light red), of domesticates inside the wild boundary (strong red-grey), and of wild space occupied by domesticates (grey).

**Figure 6. Phenotypic differentiation between crops and wild angiosperms.** Separate plots are shown for all angiosperms and each growth form (*i.e.* herbaceous, graminoids and woody). Asterisks indicate statistically significant differences at phy-P ≤ 0.05 between domestication statuses according to phylogenetic Anovas (Table 2). Center line, median;

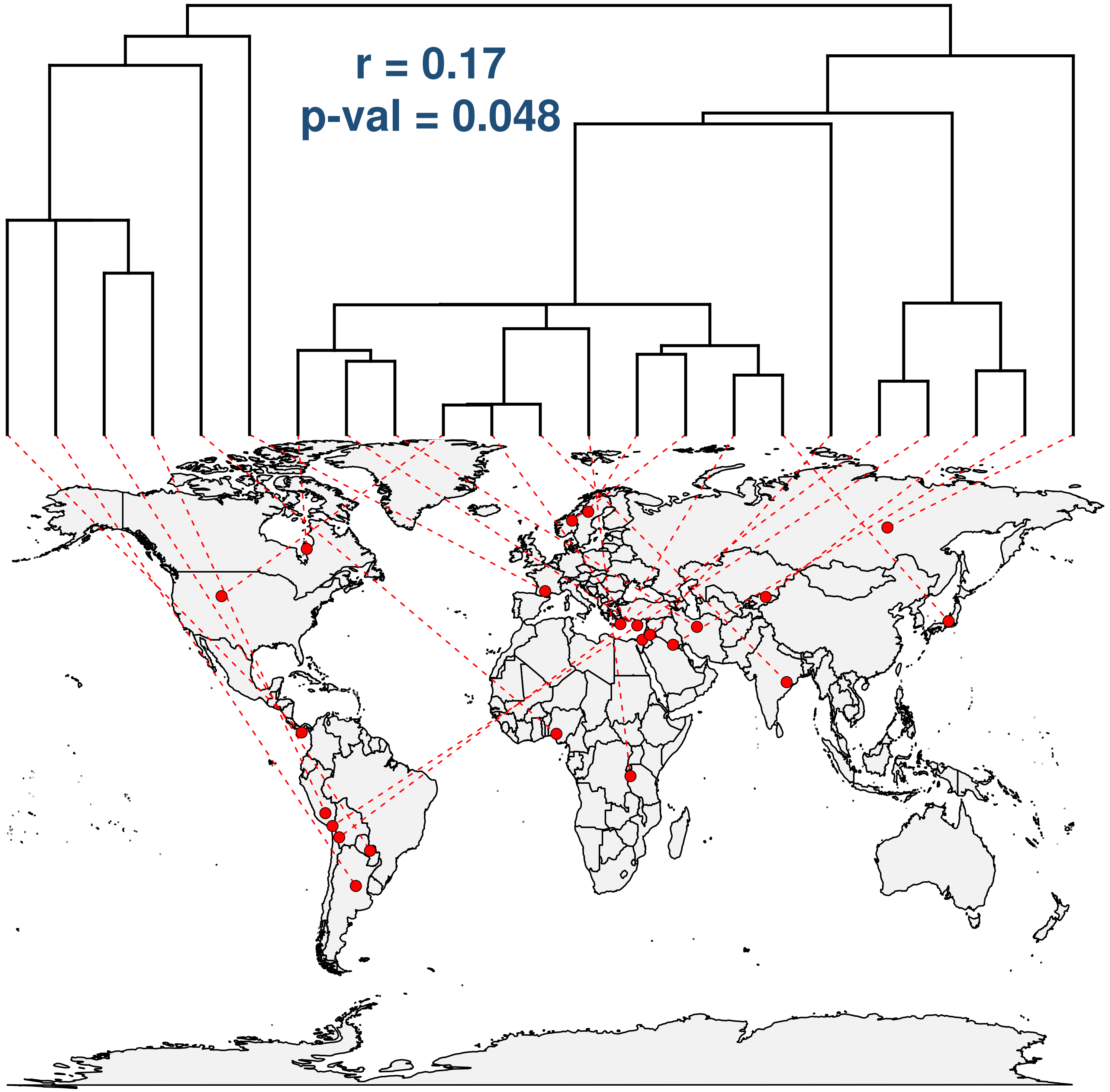
box limits, upper and lower quartiles; whiskers, lowest/highest datum still within 1.5*IQR; points, data exceeding whisker bounds.

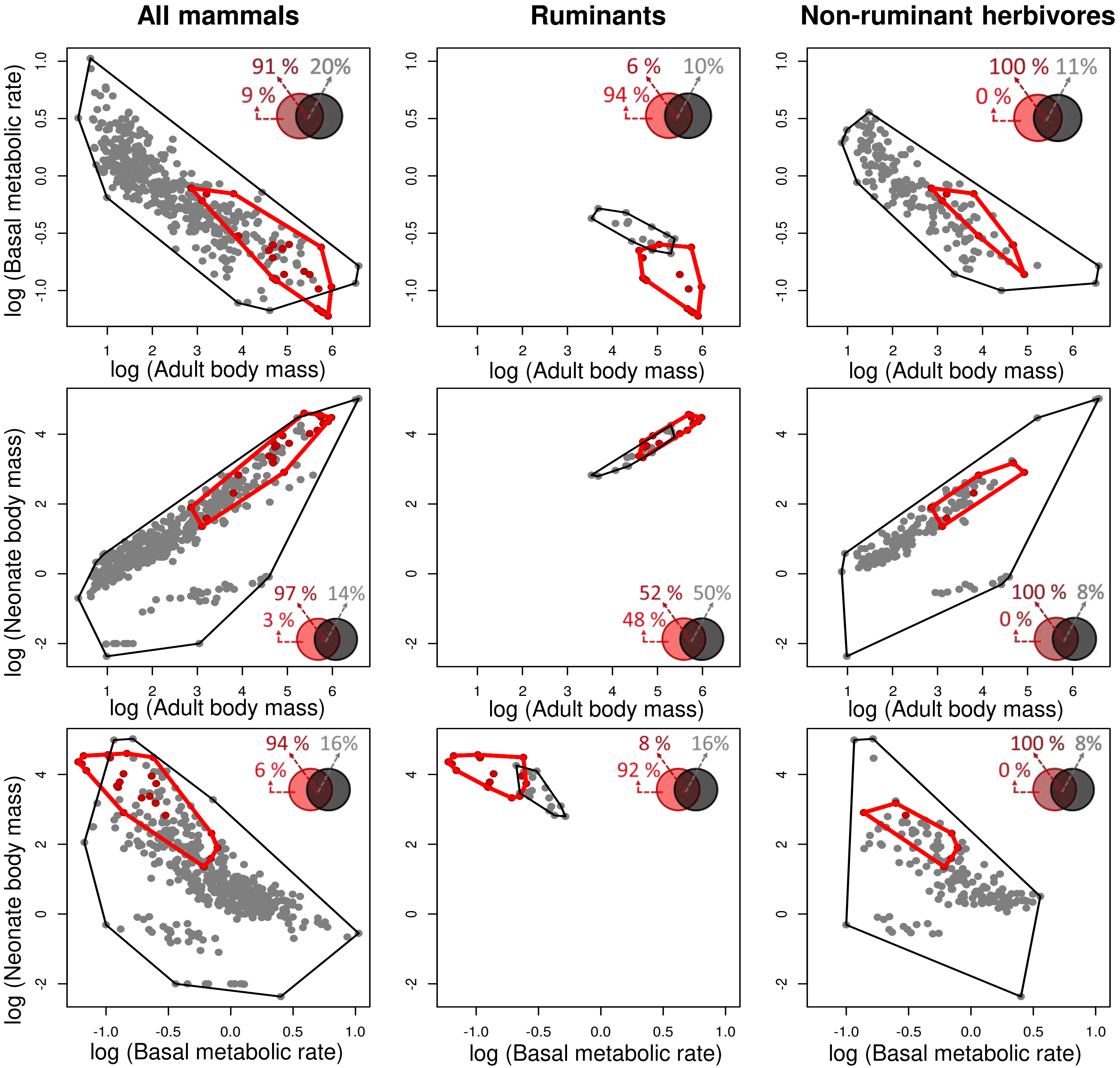

ANGIOSPERMS
MAMMALS
ABUNDANCE OF DOMESTICATES
DOMESTICATION FREQUENCY
λ = 0.006
(0.001 – 0.655)
λ = 0.029
(0.000 – 0.736)
λ = 0.010
(0.002 – 0.753)
λ = 0.996
(0.005 – 0.999)

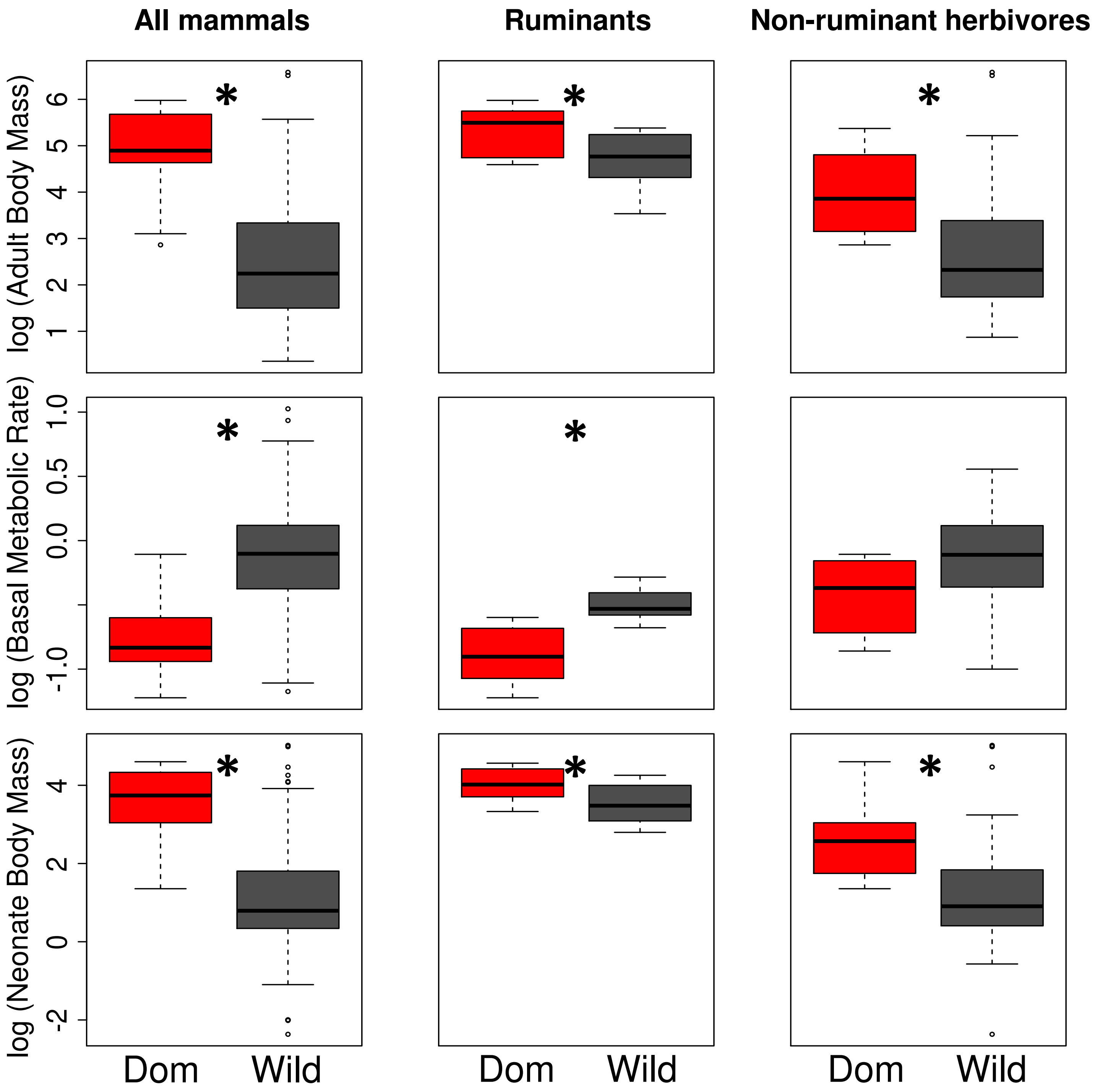
**ANGIOSPERMS**

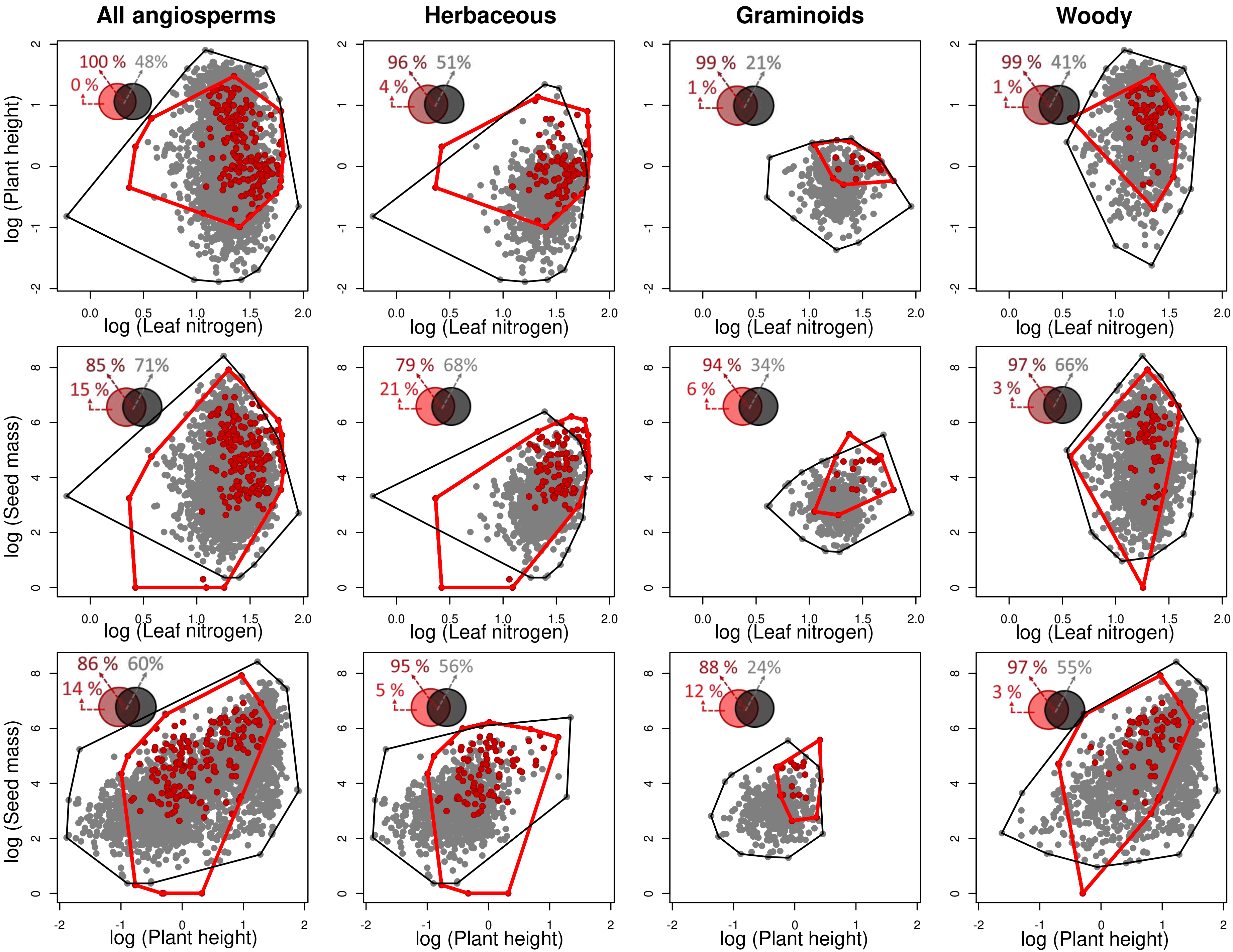


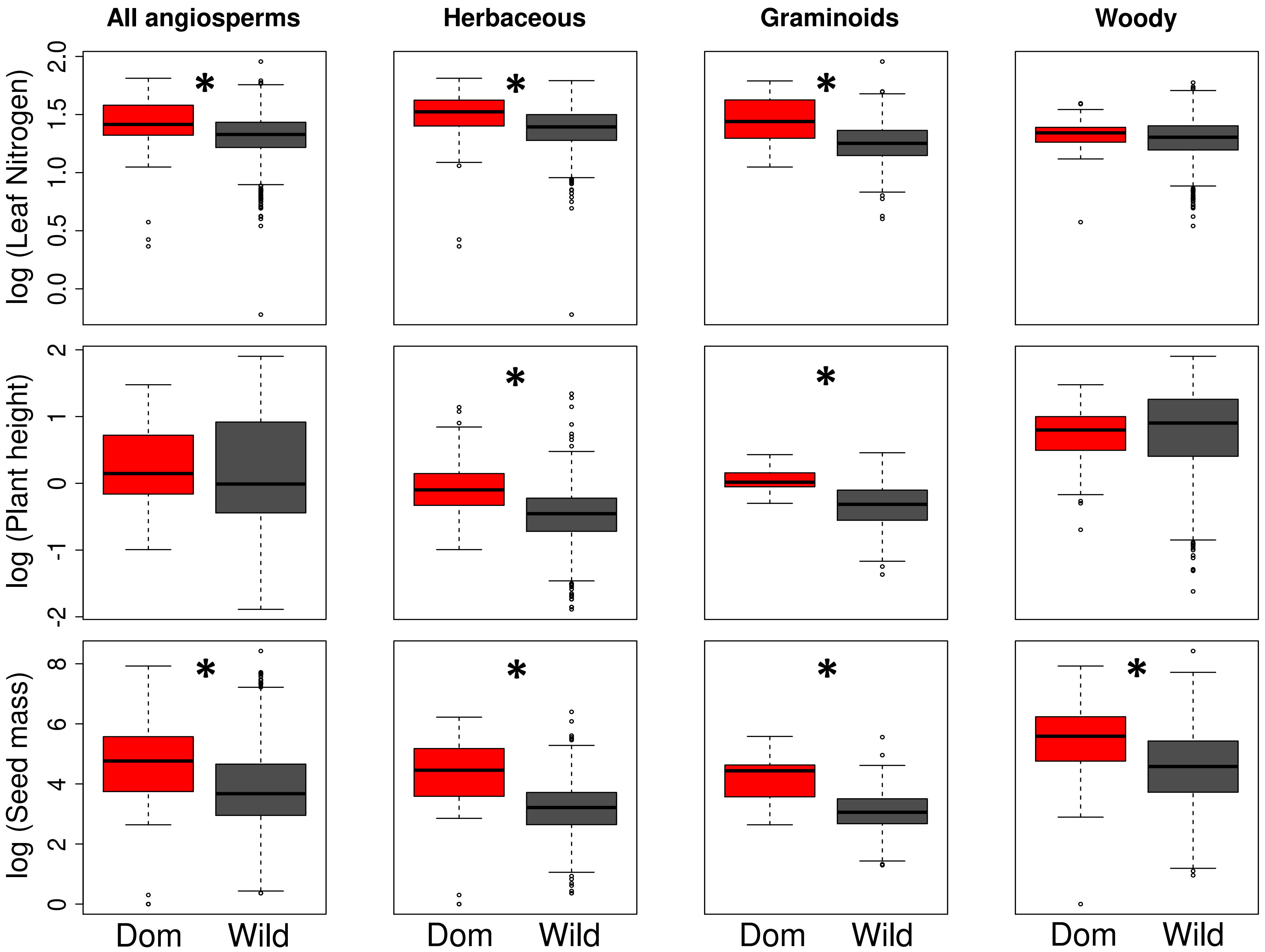
**MAMMALS**




All mammals
Ruminants
Non-ruminant herbivores
log (Basal metabolic rate)
log (Adult body mass)
log (Neonate body mass)
log (Basal metabolic rate)
91 %
20%
9 %
6 %
10%
94 %
100 %
11%
0 %
97 %
14%
3 %
52 %
50%
48 %
100 %
8%
0 %
94 %
16%
6 %
8 %
16%
92 %
100 %
8%
0 %
1.0
0.5
0.0
-0.5
-1.0
1
2
3
4
5
6
4
2
0
-2
-1.0
-0.5
0.0
0.5
1.0


All mammals
Ruminants
Non-ruminant herbivores
log (Adult Body Mass)
1
2
3
4
5
6
log (Basal Metabolic Rate)
-1.0
0.0
0.5
1.0
log (Neonate Body Mass)
-2
0
2
4
*
Dom
Wild


All angiosperms
Herbaceous
Graminoids
Woody
100 %
48%
0 %
96 %
51%
4 %
99 %
21%
1 %
99 %
41%
1 %
log (Plant height)
log (Leaf nitrogen)
85 %
71%
15 %
79 %
68%
21 %
94 %
34%
6 %
97 %
66%
3 %
log (Seed mass)
log (Leaf nitrogen)
86 %
60%
14 %
95 %
56%
5 %
88 %
24%
12 %
97 %
55%
3 %
log (Seed mass)
log (Plant height)
-2
-1
0
1
2
0.0
0.5
1.0
1.5
2.0
0
2
4
6
8


All angiosperms
Herbaceous
Graminoids
Woody
log (Leaf Nitrogen)
log (Plant height)
log (Seed mass)
0.0
0.5
1.0
1.5
2.0
-2
-1
0
1
2
0
2
4
6
8
*
Dom
Wild