Eco-evo-devo implications and archaeobiological perspectives of wild and domesticated grapevines fruits covariating traits

Vincent Bonhomme[1,&,@] , Sandrine Picq [1,2,&], Sarah Ivorra[1], Allowen Evin[1], Thierry Pastor[1], Roberto Bacilieri[3], Thierry Lacombe[3], Isabel Figueiral[1], Jean-Frédéric Terral[1,†] and Laurent Bouby[1,†]

[1] ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France. Équipe « Dynamique de la biodiversité, anthropo-écologie », Place Eugène Bataillon - CC065 34095 Montpellier Cedex 5, France.
[2] Laurentian Forestry Centre, Natural Resources Canada, 1055 Du P.E.P.S. Street, P.O. Box 10380, Québec, Quebec, G1V 4C7
[3] UMR AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Equipe « Diversité, Adaptation et Amélioration de la Vigne », 2 Place Viala, 34060 Montpellier, France
[&,†] these authors contributed equally to this work
[@] corresponding author: bonhomme.vincent@gmail.com +33 (0)4 67 14 41 60

Short title : grapevine berry-pip traits covariation

Abstract

The phenotypic changes that occurred during the domestication and diversification of grapevine are well known, particularly changes in seed morphology, but the functional causes and consequences behind these variations are poorly understood. Wild and domesticate grapes differ, among others, in the form of their pips: wild grapes produce roundish pips with short stalks and cultivated varieties have more elongated pips with longer stalks. Such variations of form are of first importance for archaeobotany since the pip form is, most often, the only remaining information in archaeological settings. This study aims to enlight archaeobotanical record and grapevine pip development by better understanding how size and shape (co)variates between pip and berry in both wild and domesticated *Vitis vinifera*. The covariation of berry size, number of seeds per berry ("piposity"), pip size and pip shape were explored on 49 grapevine accessions sampled among Euro-Mediterranean traditional cultivars and wild grapevines. We show that for wild grapevine, the higher the piposity, the bigger the berry and the more elongated the pip. For both wild and domesticated grapevine, the longer is the pip, the more it has a "domesticated" shape. Consequences for archaeobotanical studies are tested and discussed, and these covariations allowed the inference of berry dimensions from archaeological pips from a Southern France Roman site. This systematic exploration sheds light on new aspects of pip-berry relationship, in both size and shape, on grapevine eco-evo-devo changes during domestication, and invites to explore further the functional ecology of grapevine pip and berry and notably the impact of cultivation practices and human selection on grapevine morphology.

Keywords: archeobotany, berry, pip, cultivation, eco-evo-devo, domestication, *Vitis vinifera*, grape cultivars.

Introduction

Grapevine (*Vitis vinifera* L.) is one of the most cultivated fruit species in the world [1], and has held a central economic and cultural role since ancient times, particularly in the Mediterranean area (Brun, 2003; McGovern, 2007) . The berries of grapevine are primarily used in wine production, but can be consumed fresh or dried (i.e. table grape). The wild progenitor of grapevine, *Vitis vinifera* subsp. *sylvestris*, was first domesticated in the South Caucasian area [4], which has yielded the oldest wine making evidence (McGovern et al., 2017) dated to early Neolithic period (~8000 BP). The existence of other domestication centres has also been argued [6,7]. Since the early times of domestication, grapevine varieties (or cultivars or "cépages") of *Vitis vinifera* subsp. *vinifera* have been selected and propagated; today there are several thousand varieties, identified by ampelography (i.e. grape morphology) and molecular markers [8,9]. *V. vinifera* subspecies differ in their reproductive biology and other phenotypic changes following domestication include larger bunches, larger berries, higher diversity in berry shape and skin colour, and higher sugar content [10,11].

The quantitative morphological description of archaeobotanical material has brought major insights into the intertwined relationships between humans and domesticated plants [12–20], including grapevine (Bouby et al., 2013; Karasik, Rahimi, David, Weiss, & Drori, 2018; Mangafa & Kotsakis, 1996; Orrù, Grillo, Lovicu, Venora, & Bacchetta, 2013; Pagnoux et al., 2015; Terral et al., 2010; Ucchesu et al., 2016). So far, molecular approaches on ancient grapevine have yielded limited information on domestication [28,29], with the study of ancient DNA hindered by its poor preservation in charred archaeobotanical material (but see Ramos-Madrigal et al., 2019) .

Wild and domesticate grape seeds differ in their form (i.e. size plus shape): wild grapes produce roundish pips with short stalks and cultivated varieties produce more elongated pips with longer stalks [31]. Such form variations have been identified on archaeological grapevine pips (Bouby et al., 2013; Pagnoux et al., 2015; Terral et al., 2010). Archaeological material is often charred which can cause domesticated pips to appear more similar to wild pips [32] yet experimental charring has demonstrated the robustness of identification (Bouby et al., 2018; Ucchesu et al., 2016) .

The functional causes and consequences, if any, behind the form variation of grapevine pip are still poorly understood. If size, shape, taste and colour of berries are phenotypic traits that have been selected by humans, pip shape was likely not a direct target of selective pressures but may possibly be affected by: the berry size; the number of pips per berry; the growing

environment and cultivation practices; the domestication status and the variety for domesticated grapevine; and developmental stochasticity. For instance, previous works suggested that pip size and the number of pips per berry are positively correlated to berry size [8,34,35] and that this correlation is stronger for wild grapevines (Bouby et al., 2013).

Direct selection for numerous and larger berries is likely as they are key yield factors, but to what extent the form changes observed in archaeological pips imply changes in the form of berries? How this relation could be affected by the cultivation and subsequent domestication of wild individuals? Pips with a form similar to that of modern wild grapevine pips have been repeatedly found in ancient viticultural sites, which is suggestive of cultivated "true" wild grapevines or "weakly" domesticated forms alongside "true" fully domesticated varieties (Bouby et al., 2013; Pagnoux et al., 2015).

This paper scrutinizes how the form of berries and pips they contain covariate. A dataset of domesticated and wild contemporary grapevines allowed to compare patterns of covariation between these two, wild and domestic, *Vitis vinifera* compartments. This article is divided into four questions: i) how does size (co)vary between pips and berries, and depending on the number of pips?; ii) how does shape (co)vary between pips and berries and depending on the number of pips per berry?; iii) how much pip shape depends on berry size, number of pips per berry, status, accession, and which practical consequences for archaeobotanical studies? iv) can we infer the dimensions of the berry dimensions from the (recovered archaeological) pips?

Results

Preliminary analyses on modern material

The average piposity is equivalent between domesticated and wild accessions (mean±sd: domesticated=2.01±0.891, wild=2.1±0.968; generalized linear model with Poisson error: df=1468, z=1.234, P=0.217 – Figure 2). The distribution of piposity, however, does differ (Fisher's exact test: P=0.004) due to a higher proportion of 4-pips berries in wild grapevines (Fisher's exact test with 4-pips berries removed: P=0.6272). No difference was observed between cultivated (domesticated and wild) accessions and those collected from wild (mean±sd: cultivated=2.06±0.931, non-cultivated=2.00±0.902; generalized linear model with Poisson error: df=1468, z=-0.676, P=0.499).

Covariation between pip and berry size in relation to the number of pips

Wild vs. domesticated. All berries and pips measurements were overall smaller for wild accessions compared to their domesticated counterparts (Wilcoxon one-tailed rank tests: all $P<10^{-8}$ – Figure A). The extent of the difference between wild and domesticated varied between the pip dimensions ($pip_{LengthStalk} > pip_{PositionChalaza} > pip_{Length} > pip_{Thickness} > pip_{Breadth}$).

Overall, the higher the piposity, the lower the contrast between domesticated and wild (Figure 3a). Pip dimensions of wild grapevines increase more substantially with increased piposity than their domesticated counterpart decrease. In wild grapevines, larger berries have more and larger pips (Figure 3a). No marked differences in berry dimensions/mass along increasing piposity were found for the cultivated grapevines, excepted between 1- and 2-pip for $pip_{Thickness}$ (Wilcoxon rank tests: P=0.006).

Table vs. wine cultivars. Table grapes have significantly higher dimensions than wine varieties (Figure 3b). With increasing piposity, table varieties tend to have bigger berries which is not the case for wine varieties. For pips, the only significant differences between low and high piposity were found for wine varieties and for $pip_{Breadth}$ and $pip_{Thickness}$ ($P<10^{-16}$).

Wild grown in collection vs. wild in natura. Wild grapevine pips and berries are bigger when in cultivation than their counterparts growing *in natura* (Figure 3c). Besides these global differences, trends of all measured variables are similar along increasing piposity. The berry mass ratios, relatively to wild collected *in natura*, were on average, 6.4 for wine varieties, 15.6 for table ones and 1.8 for cultivated wild.

Bivariate comparisons (Figure B, Supplementary information) indicate positive correlations between all measurements. The total pip_{Length} appears to be the most consistent variable, between domesticated and wild grapevines: indeed, only the correlation with the $pip_{LengthStalk}$ show a significant interaction. Inversely, the correlations implying $pip_{LengthStalk}$ always show a significant interaction. For pips dimensions, the best correlations were found between pip_{Length} and $pip_{PositionChalaza}$ (adj. $r^2=0.8$) among those with non-significant interactions, and between $pip_{LengthStalk}$ and $pip_{PositionChalaza}$ (adj. $r^2_{wild}=0.615$, adj. $r^2_{domesticated}=0.717$) among those with significant interactions. Compared to pips dimensions, correlations between berry dimensions were much better and the three possible interactions were all significant.

Covariation between pip and berry shape in relation to the number of pips

The PCA shows that the first two PCs (Figure 5) gathered 69% of the total shape variation, and higher rank components clearly levelled off (PC1=43.0%; PC2=26.2%; PC3=6.7%;

PC4=3.9%), only the first two PCs were used as synthetic shape variables. Shape differences between wild and domesticated grapevines are mostly captured on PC1 yet scores on both PC1 (Wilcoxon rank tests, $P<10^{-16}$) and on PC2 ($P<10^{-16}$) were found different. Here, PC1 represents how prominent is the stalk and how round is the pip; PC2 represents the circularity, a more global length/width ratio of pips, for the two views.

Regarding shape versus pip dimensions, pip_{Length}, correlated to all other measurements, is itself correlated with position on PC1. Two regressions were justified (Analysis of covariance: df=1, F=362.7, $P<10^{-16}$); their slope are identical (df=1, F=0.037, P=0.848) but their intercept differed between wild and domesticated. These two regressions were significant yet r^2 were low (wild: $P<10^{-16}$, adj. r^2=0.195; domesticated: $P<10^{-16}$, adj. r^2=0.240 – Figure 5a). When PC1 and PC2 are considered jointly, two regressions were not justified (P=0.04) and the r^2 was lower (P=0.04, adj. r^2=0.181 – Figure 5a). The longer the pip is, the more "domesticate" it looks, particularly in terms of stalk prominence.

As concerns shape versus piposity, the latter is associated with shape changes on PC1 between wild and domesticated both overall (see above) and within levels (Wilcoxon rank tests, all $P<10^{-10}$ – Figure 5b). Within domesticated accessions, differences were never significant. Within wild accessions, differences were not found between pairs of successive piposity levels but those between 1-3, 2-3 and 2-4 (all with $P<10^{-8}$ – not shown). For PC2, general differences observed between domesticated and wild vanished for high piposity (1-pip: $P<10^{-12}$; 2-pips: $P<10^{-9}$; 3-pips: P=0.016; 4-pips: P=0.035 – Figure 5b). No differences within wild/domesticated and between successive piposities were found significant.

Mean shapes (Figure 6) illustrate these results. The mean absolute difference (MD) confirms that larger changes between extreme piposities are observed within wild grapevines (particularly for cultivated ones) and reveals that most of these changes affect the dorsal side of the pips (Figure 6).

Pip shape and size in relation to status, accession and piposity; consequences for archaeobotanical inference

The respective contributions of berry height, accession and piposity on the shape of pips (Figure 7) show that the accession is the factor affecting the most the pip shape. Among the different subsets, the accession factor has a higher impact on domesticated grapevine than on wild, and on cultivated wild accessions than on those collected *in natura*. By contrast, its contributions for wine and table domesticated varieties were similar. Here again, piposity and berry height

both affect the pip shape of wild accessions but have a limited (piposity) and very limited (berry height) contribution for domesticated accessions.

Classification accuracies were compared using different training data and on different subsets (Figure 8). When different piposity levels were pooled, mirroring archaeobotanical admixtures, classification was very good at the status level (Figure 8a). Size + shape performed better (95%), than shape (93.7%) and size (92.5%) alone. When these models were evaluated on piposity subsets, they all have an accuracy above 91%, except for 4-pips berries. In all cases, accuracies were much higher than what could be obtained by chance alone. As expected, accuracies were lower at the accession level (Figure 8b) and when piposity levels were pooled, size + shape (89.8%) outperformed shape alone (81.3%) and size alone (46.3%). The same model ranking was observed on piposity subset, except for 4-pips berries. Overall, accuracies were nevertheless much better than chance alone.

Application to archaeological pips: can we infer the dimensions of the (vanished) berry dimensions from the (recovered) pips?

On modern material, we used the size of pips to predict berry heights and diameters. Both regressions show a significant interaction of the domestication status ($berry_{Diameter}$: df=1, F=8369, $P<10^{-16}$; $berry_{Diameter}$: df=1, F=7730, $P<10^{-16}$), and two regressions for berry diameter and two others for its height were obtained (Figure 9). All were significant (all $P<10^{-16}$) yet the adjusted r^2 were quite low ($berry_{Diameter}$ adj. r^2_{wild}=0.585, adj. $r^2_{domesticated}$=0.491; $berry_{Height}$ adj. r^2_{wild}=0.615, $r^2_{domesticated}$=0.511). Final models all used pip_{Length}, $pip_{Thickness}$, and at least one PC. Table 2). On unlogged (to compare "real" deviations obtained) berry diameter and height, the relative deviations were obtained (Figure C – ESM). Mean relative deviation per accession for $berry_{Diameter}$ ranged from -12.9% to +10.3% for wild, and from -22.9% to +17.7% for domesticated; for $berry_{Height}$ they ranged from -13.0% to +13.1% for wild, and from -29.4% to +29.4% for domesticated. The average predictions were all centred (on zero) ±1.6%.

Then, these four models were applied on the archaeobotanical material after being classified at the wild/domesticated level using LDA. 46 pips (22%) were classified with a posterior probability <0.8 and were filtered out. Among the remaining pips, 114 (72%) were classified as domesticated and 45 (28%) as wild. When compared to their modern analogues (Figure 10), the length of "domesticated pips" were closer to those of wine varieties than table varieties; the lengths of "wild pips" were intermediate between wild accessions collected in their habitat and those cultivated. For archaeological pips identified as domesticated, both inferred berry height and diameter were intermediate between wine and table modern varieties

yet closer to wine ones. Similarly, for wild archaeological material, inferred berry height and diameter were intermediate between wild collected in their habitat and those grown in collection but closer to the former.

Discussion

This study opens major fronts in our understanding of *Vitis vinifera* phenotypic changes under domestication and helps disentangle the interplay of the number of pips per berry, berry dimensions, domestication, pip shape, varietal diversity and cultivation practices in both wild and domesticated grapevines. We discuss implications for *Vitis vinifera* eco-evo-devo and perspectives for archaeobotanical studies for which a possible application is proposed.

Patterns of covariation between the form of the pip, the form of the berry and the piposity

With two ovules per ovary, and two ovaries per berry, the theoretical maximum number of pips per berry is four, yet one was observed with five. Such abnormal piposity have been reported, for berries having more than two ovaries [8,36]. Most berries had two pips, and more than 70% had only one or two. This is in accordance with previous publication [35]. There were no differences neither between domesticated or wild (Figure 2), nor between cultivated wild individuals and those collected in their habitat. Higher piposity could be expected for domesticated grapevine (since they have hermaphroditic flowers), and more generally for cultivated accessions (since the pollen rain is expected to be lower, or even limiting, in the natural habitat).

Overall, and this is no surprise, wild pips and berries are smaller than their domesticated counterparts. Similarly, pips and berries of wine varieties are smaller than those of table varieties, as pips and berries of wild grapevines collected in their habitat are smaller than those from cultivated wild individuals. This study details the effect of piposity on the pip form reported by previous studies (Bouby et al., 2013; Houel et al., 2013; Negrul, 1960; Olmo, 1995). Among vertebrate dispersed plants, the reward (the fruit pulp mass) associated with a given seed mass is commensurate with work required to move it, and is expected to scale relatively [37]. For wild grapevine, and vertebrate dispersed plant species in general, berry and pip dimensions/masses are expected to be constrained by their dispersers and by a general trade-off between pip size and number [38].

For all but wine varieties, the higher the piposity the longer the pip and the bigger the berry in which they develop (Figure 3). For these groups, it seems that more numerous pips are

not limited by space or nutrients but rather contribute the development of bigger berries. The stages of berry development are well known [39,40] and can be divided into two phases of enlargement. The first, prior to anthesis, is a period of rapid berry growth mostly due to cell division. After anthesis, berry growth is largely due to cell enlargement and it has been suggested that pip growth may also increases cell mitosis in the developing berry (Ojeda et al., 1999). Auxins, cytokinins and gibberelins, upregulated shortly after fertilisation in grapevine ovaries, are likely to trigger berry growth by cell expansion [35].

The absence of positive (or even negative) correlations between piposity, pip and berry dimensions for wine varieties remains unclear. For these varieties, the regulation, if any, may be at the bunch or stock scale, whether it has been selected (for example to concentrate sugars, aromas and flavours) or it is a by-product of another trait under selection. Since table varieties are larger than wine varieties, the berry dimensions of the latter cannot be argued to have reached a developmental limit.

Finally, bivariate correlations concerning berry dimensions and mass are the strongest observed. This indicate robust allometries between berry size and mass, in other words that berry largely remain ellipsoid in shape, independently of their dimensions.

Morphometrics and domestication as a wedge into grapevine eco-evo-devo

For grapevine and domesticated plants in general, domestication results in a change of desirable phenotypic patterns (bigger fruits for instance) but also releases many “natural” constraints such as dispersion [42]. Cultivation practices such as pruning may explain why wild individuals grown in collection have bigger berries for higher piposity: the number of bunches is reduced, leading to larger pips. Cultivation also reduces growth constraints such as competition for water and light, self-supporting and climbing costs, those related to dispersers, etc.

Evidence of plastic and canalized phenotypic expression may be fuel for further eco-evo-devo studies. The latter brings a conceptual and experimental framework that relies on environmentally mediated regulatory systems to better understand ecological and evolutionary changes [43]. Here, the norm of reaction of the pip size and shape, along increasing piposity and berry dimensions, is clearly different at the three investigated levels: between wild and domesticated, between wine and table, between wild individuals grown in collection and those collected in their natural habitat.

Consequences for archaeological inference

Taken independently or in pairwise comparisons, some pip lengths differences between wild and domesticated appear more "robust" to increasing piposity, notably $pip_{LengthStalk}$ and $pip_{PositionChalaza}$ (Figure 3a; Figure A Supplementary information). Their interest in discriminating domestication status has long been used, including when archaeological material is charred (Bouby et al., 2018; Mangafa & Kotsakis, 1996; Smith & Jones, 1990).

Overall, shape variability tends to distinguish the wild and domesticated grapevines (Bouby et al., 2013; Terral et al., 2010). Similar to the positive correlation between piposity and pip length of wild accessions, piposity is also correlated with shape changes for wild accessions (Figure 5). Most of these changes affect the dorsal side of wild pips, particularly when piposity is high. There is as much as ~70% difference, between 1- and 4-pips for wild accessions, than those observed between average wild and domesticated. Differences in extreme piposity are even larger than this "domestication gap", for the cultivated wild. This does not answer the question whether past vineyards cultivated "true" wild grapevines or "weakly" domesticated forms (Bouby et al., 2013; Pagnoux et al., 2015) but it points out how piposity and cultivation practices contribute to this confusion, further enhancing the continuum of pip forms.

Pip shape being largely used in archaeobotany, it was crucial to point out which factors contribute to its variability, or at least covary with it, and if they could preclude identification of archaeological remains. Here, the main factor associated to pip differences was, by far, the accession and it was even more important for domesticated accessions; in other words accession effect appears stronger than domestication (Figure 7). Relatively to accession, berry height and piposity poorly contributed to observed differences. This confirms the usefulness of shape and its robustness to identify morphotypes that are shape varietal archetypes. It may also indicate that domestication favours pip shape diversification whether this results from genetical linkage with selected loci or is the product of drift.

Here, we show the reliability of classification, independently of piposity. Indeed, classification accuracies at the status level were all high (Figure 8), even when the models trained on the pip admixture where evaluated on piposity subsets. Shape was nonetheless superior to size alone in discrimination power but, when considered jointly, the classification was improved. Whenever possible, size should thus be included along morphometric coefficients and used jointly in classification models. Accuracies at the accession levels evidenced even more clearly the latter conclusions.

Shape is overall more robust than sizes when models were evaluated on piposity subsets. The only exception, for both status and accession levels, were obtained on the 4-pips subset. Our experimental design reflects real-world admixtures, and sample sizes of all studied factors were not balanced. That being said, results here evidence that such bias in the piposity structure is very unlikely to affect archaeobotanical identification either at the status or accession levels.

An application on archaeological material: inferring berry size from pip

Berry is very likely home to the most selected traits, from the beginnings of domestication to varietal breeding and diversification times. Unfortunately, its dimensions cannot be quantified directly on archaeological material where fleshy parts are usually absent or too degraded. The only route to investigate changes in berry shape and size is through actualistic inference based on pip, and trained on modern material. Here, multiple regressions on pip dimensions show that berry diameter (Figure 9a) and height (Figure 9b) were not perfectly predicted but nevertheless centred on zero and overall in the ±25% range.

Our archaeological application used material from Sauvian - La Lesse, a Roman farming establishment involved in wine production, were an admixture of wild and domesticated type is attested (Figueiral et al., 2015). As in many cases in Southern France, the presence of numerous wild type pips in a vinicultural site let us consider that these vines were locally cultivated to make wine in Roman times. Berry dimensions inferred from pips of this site are intermediate between the wild growing in their habitat and those cultivated (Figure 10). This may suggest that wild, or weakly domesticated, individuals were cultivated in Roman vineyards. The berry dimensions inferred for domesticated varieties were closer to modern wine varieties than to table ones. This is congruent with the wine production attested at this period and in this region (Figueiral et al., 2010; Figueiral et al., 2015).

Conclusion

The main finding of this systematic exploration of berry and pip form covariation is that for wild grapevine, the higher the piposity, the bigger is the berry and the longer is the pip. For both wild and domesticated, the longer is the pip, the more its shape looks like “domesticated”. Further studies will clarify the contribution of cultivation practices contribution on pip shape, largely used in archaeobotanical studies to better understand viticulture history. These findings pave the way for dedicated studies to shed light on genetic, functional and evolutionary changes

that occurred in *Vitis vinifera* between the pip, its reproductive unit, and the berry, its dispersal reward and the main target of its domestication and varietal improvement.

Material and Methods

Statistical environment

Statistical analyses were performed using the R 3.6.2 environment [46], the package Momocs 1.3.0 for everything morphometrics (Bonhomme et al., 2014) and the tidyverse 1.2.1 packages for data manipulation and most graphics [48]. Alpha significance level was chosen equal to 10^{-5} all along analyses. This level both ensure marked differences for subsequent archaeobotanical application and an overall alpha level below 0.05 when repeated tests were done (i.e., a Bonferroni correction).

Nomenclature

Hereafter, *status* designates compartment (domesticated vs. wild); *accession* designates the variety (or cultivar, or *cépage*) for domesticated grapevine and the individual for wild grapevine; synecdochically, a domesticated/wild pip/berry refers to the accession they were collected from; *cultivation* designs whether wild individuals were *cultivated* (grown in field collection) or sampled *in natura*; *form* is used when *shape* and *size* are used in combination; "piposity" is short for "given a pip, the number of pips in the berry where it was sampled".

Modern and archaeological material

The modern reference material included 49 accessions (30 domesticated and 19 wild) from Euro-Mediterranean traditional cultivars and wild grapevines (Table 1). Fourteen wild grapevines were collected at ripeness in their habitat, and five were cultivated in the French central ampelographic collection (INRA, Vassal-Montpellier Grapevine Biological Resources Center; https://www6.montpellier.inra.fr/vassal), along with the domesticated accessions. Of the domesticated accessions, 21 were wine varieties and 9 table varieties. For each accession, 30 normally developed berries have been haphazardly collected from a single, fully ripe bunch.

Archaeobotanical material comes from two wells at the Roman farm of Sauvian - La Lesse, extensively described elsewhere (US3022, US3063, US3171 and US3183 in Figueiral et al., 2015). These archaeological layers were dated to 2025-1725 BP based on pottery and coins. The waterlogged conditions ensured very good preservation the pips used in this study (N=205).

Traditional measurements

355 On modern material, the berry diameter ($berry_{Diameter}$), height ($berry_{Height}$) and mass ($berry_{Mass}$) were obtained before dissection (Table 1). Mass was not available for 9 accessions that were removed from further analyses involving mass. Then, the number of pips (hereafter "piposity") was recorded and one pip was randomly chosen. A single berry from the variety "Kravi tzitzi" was found with 5 pips and was discarded from further analyses. The final dataset thus consisted 360 of 1469 pips (48 accessions × 30 pips + 1×29).

All pips, archaeological and modern, were photographed in dorsal and lateral views by the same operator (TP) using an Olympus SZ-ET stereomicroscope and an Olympus DP camera. On each pip, five length measurements were manually recorded by the same operator (LB) and using ImageJ (Rasband, 2008, Table 1, Figure 1): total length (pip_{Length}), length of stalk 365 ($pip_{LengthStalk}$), position of the chalaza ($pip_{PositionChalaza}$), breadth ($pip_{Breadth}$) and thickness ($pip_{Thickness}$). All length measurements were log-transformed to focus on relative changes and minimize size differences; the mass was log cubic-root transformed for the same reason (Bouby et al., 2013).

As preliminary analyses on modern material, differences between average piposity were 370 tested using generalized linear model with Poisson error; differences in their distributions were tested using two-sided Fisher's exact tests on count data.

Testing the covariation between pip and berry size in relation to the number of pips

On modern material, three sets of differences in pips and berries measurements were tested using multivariate analysis of covariance: i) the interaction between status and piposity; ii) if 375 the latter was significant, we also tested differences between status for a given piposity level; iii) whether the average piposity differs between status. These three possible sets of differences were tested between different subsets: domesticated and wild accessions; wine and table varieties for domesticated accessions; cultivated wild individuals and those collected *in natura*. Piposity was then discarded and sets were compared using Wilcoxon rank tests.

380 Bivariate comparisons were explored between the domesticated and wild accessions (discarding piposity), and tested with an analysis of covariance. When the domestication status was significant, separate regressions were tested and, if significant, the adjusted r^2 was obtained.

Testing the covariation between pip and berry shapes in relation to the piposity

For pips, shapes data were extracted from the dorsal and lateral outlines. 2D coordinates were 385 extracted from photographs, centred, scaled, aligned along their longer axis (using the variance-

covariance matrix of their coordinates) and normalized for the position of their first point before elliptical Fourier transforms (EFT). These preliminary steps removed positional, size, rotation and phasing differences between outlines and EFT could then be used without numerical normalization [50]. EFT were performed on the dorsal and lateral views separately, and the number of harmonics was chosen to gather 99% of the total harmonic power (8 for both views). This generated 64 coefficients for each pip (2 views × 8 harmonics per view × 4 coefficients per harmonic).

To explore the overall variability of shape, a principal component analysis (PCA) was calculated on the full matrix of coefficients. The first two PCs (see Results) were used as synthetic shape variables. To test the effect of piposity and pip dimension on pip shape, the same approach than for length measurements using PC1 and PC2 as the response variables. To test the relation between shape and pip length (only pip_{Length} was used), analyses of covariance first tested if separate regressions were justified. Then Wilcoxon tests were used to test for shape differences between and within piposity levels.

To visualize shape differences between extreme piposity levels (1 and 4), mean shapes for the dorsal and lateral views were calculated on the matrix of coefficients. These differences were quantified with the mean absolute difference (MD) between each sets of Fourier coefficients. To make these differences meaningful, they were divided by the mean difference of Fourier coefficients between cultivated and wild accessions with all piposity levels pooled. For each subset, MD was calculated as: $(|\text{coefficients}_{\text{subset, 1-pip}} - \text{coefficients}_{\text{subset, 4-pips}}|)/(|\text{coefficients}_{\text{domesticated, all pips}} - \text{coefficients}_{\text{wild, all pips}}|)$. For example, a MD equals to 0 would indicate no difference between pips with a piposity of 1 or 4; a MD greater than unit would indicate more differences relatively to differences that exist between domesticated and wild individuals.

Pip shape and size in relation to status, accession and piposity

To quantify the respective contribution of berry dimensions, accession and piposity onto pip shape, a multivariate analysis of variance used the following model: all Fourier coefficients ~ $berry_{Height}$ + accessions + piposity within accession. Since it was highly correlated to other berry measurements (see Results), only $berry_{Height}$ was used to describe berry dimensions. The contribution of each variable is the ratio of its sum of squares over the total sum of squares (including residuals). Again, this is tested on the different subsets of interest (Figure 7).

Linear discriminant analyses (LDA) were used to evaluate whether piposity could preclude status and accession classification accuracies. Different combinations of predictors (sizes; shape; sizes + shape) were evaluated to benchmark their performance to classify the pips to their correct status and accession (Figure 8).

Given a combination of (status, accession) × (sizes+shape, sizes, shape), a leave-one-out cross-validation was used to assess classification accuracies, evaluated on all pips, to mirror archaeological admixtures where piposity is unknown (Figure 8). To cope with unbalanced group structures, we calculated a baseline for each subset that estimates the mean and maximum accuracy one can obtain by chance, using 10^4 permutations (see (Evin *et al.*, 2013). If the accuracy observed is higher than the maximum value obtained using permutations, the LDA can be considered to perform better than random, with an estimated alpha below 10^{-4}.

Predicting the dimensions of the archaeological berry dimensions

Separate multivariate regressions were calculated on the modern material, for berry height and diameter (using the five length measurements on pips). As predictor variable, we used length measurements (for dimensions) and the first two principal components (for shape). The difference between domesticated and wild grapevines regressions was first tested using an analysis of covariance: two regressions (one for cultivated, one for wild) were obtained for the berry height and two others for its diameter (Figure 9). These four regressions were fitted using stepwise regression with backward elimination based on the AIC (Venables & Ripley, 2002), and started with full models: $berry_{Height/Diameter}$ for wild/domesticated ~ pip_{Length} + $pip_{LengthStalk}$ + $pip_{PositionChalaza}$ + $pip_{Breadth}$ + $pip_{Thickness}$ + PC1 + PC2, all but PCs were log-transformed). Then, archaeological pips were classified into domesticated or wild using an LDA trained using the same variables but of modern pips. Pips assigned to wild/domesticated with a posterior probability <0.8 were filtered out. Finally, the berry height and diameter of this archaeological material were inferred using the corresponding models (Figure 10).

Acknowledgements

This study is funded by ANR project “Vignes et vins en France du Néolithique au Moyen Âge. Approche intégrée en archéosciences” (PI: Laurent Bouby) and supported by the OSU-OREME (https://oreme.org/) that helped to the constitution of the wild grape pip collection. We are grateful to the INRA Vassal-Montpellier grapevine collection (Marseillan-Plage, France) that

provided all the pips from cultivated varieties and cultivated wild grapes. We warmly acknowledge Michael Wallace for his help with English.

Author contribution

LB, SP, SI and JFT conceived the ideas and designed methodology; SP, SI, TP, IF and LB collected the data; VB, led analysed the data with important contributions from SP, AE, JFT, LB, TL and RB; VB led writing of the manuscript with the help of all co-authors, and chiefly SP.

References

1. FAO. FAOSTAT. 2017.
2. Brun J-P. Le vin et l'huile dans la Méditerranée antique : Viticulture, oléiculture et procédés de fabrication. Editions Errance; 2003.
3. McGovern PE. Ancient Wine The Search for the Origins of Viniculture. Princeton University Press; 2007.
4. Imazio S, Maghradze D, De Lorenzis G, Bacilieri R, Laucou V, This P, et al. From the cradle of grapevine domestication: molecular overview and description of Georgian grapevine (Vitis vinifera L.) germplasm. Tree Genetics & Genomes. 2013;9: 641-658. doi:10.1007/s11295-013-0597-9
5. McGovern PE, Jalabadze M, Batiuk S, Callahan MP, Smith KE, Hall GR, et al. Early Neolithic wine of Georgia in the South Caucasus. Proceedings of the National Academy of Sciences. 2017;114: E10309-E10318. doi:10.1073/pnas.1714728114
6. Grassi F, Labra M, Imazio S, Spada A, Sgorbati S, Scienza A, et al. Evidence of a secondary grapevine domestication centre detected by SSR analysis. 2003; 1315-1320. doi:10.1007/s00122-003-1321-1
7. Arroyo-Garcia R, Ruiz-Garcia L, Bolling L, Ocete R, Lopez MA, Arnold C, et al. Multiple origins of cultivated grapevine (Vitis vinifera L. ssp. sativa) based on chloroplast DNA polymorphisms. Molecular Ecology. 2006;15: 3707-3714. doi:10.1111/j.1365-294X.2006.03049.x
8. Galet P. Cépages et vignobles de France. Dehan, éditeur. Montpellier; 1988.
9. This P, Jung A, Boccacci P, Borrego J, Botta R, Costantini L, et al. Development of a standard set of microsatellite reference alleles for identification of grape cultivars.

Theoretical and Applied Genetics. 2004; 1448-1458. doi:10.1007/s00122-004-1760-3

10. Olmo H. The Origins and Ancient History of Wine. Dans: McGovern P, éditeur. The Origins and Ancient History of Wine. Amsterdam: Gordon and Breach; 1995. p. 31-43.

11. This P, Lacombe T, Thomas M. Historical origins and genetic diversity of wine grapes. Trends in Genetics. 2006;22: 511-519. doi:10.1016/j.tig.2006.07.008

12. Zohary D, Hopf M. Domestication of plants in the world. The origin and spread of cultivated plants in west Asia, Europe and the Nile valley. Clarendon Press, Oxford; 1993.

13. Jones G, Valamoti S, Charles M. Early crop diversity: A « new » glume wheat from northern Greece. Vegetation History and Archaeobotany. 2000;9: 133-146. doi:10.1007/BF01299798

14. Terral J-F, Alonso N, Chatti N, Capdevila RB i, Fabre L, Fiorentino G, et al. Historical biogeography of olive domestication (Olea europaea L.) as revealed by geometrical morphometry applied to biological and archaeological material. Journal of Biogeography. 2004;31: 63-77. doi:10.1046/j.0305-0270.2003.01019.x

15. Terral J-F, Newton C, Ivorra S, Gros-Balthazard M, de Morais CT, Picq S, et al. Insights into the historical biogeography of the date palm (*Phoenix dactylifera* L.) using geometric morphometry of modern and ancient seeds. Journal of Biogeography. 2012;39: 929-941. doi:10.1111/j.1365-2699.2011.02649.x

16. Willcox G. Measuring grain size and identifying Near Eastern cereal domestication: evidence from the Euphrates valley. Journal of Archaeological Science. 2004;31: 145-150. doi:10.1016/j.jas.2003.07.003

17. Fuller DQ. Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical Insights from the Old World. Annals of Botany. 2007;100: 903-924. doi:10.1093/aob/mcm048

18. Ros J, Evin A, Bouby L, Ruas M-P. Geometric morphometric analysis of grain shape and the identification of two-rowed barley (Hordeum vulgare subsp. distichum L.) in southern France. Journal of Archaeological Science. 2014;41: 568-575. doi:10.1016/j.jas.2013.09.015

19. Bonhomme V, Forster E, Wallace M, Stillman E, Charles M, Jones G. The first shoots of a modern morphometrics approach to the origins of agriculture. Web Ecology. 2015;16: 1-2. doi:10.5194/we-16-1-2016

20. Wallace M, Bonhomme V, Russell J, Stillman E, George TS, Ramsay L, et al. Searching for the Origins of Bere Barley: A Geometric Morphometric Approach to Cereal Landrace

Recognition in Archaeology. Journal of Archaeological Method and Theory. 2018. doi:10.1007/s10816-018-9402-2

21. Mangafa M, Kotsakis K. A New Method for the Identification of Wild and Cultivated Charred Grape Seeds. 1996; 409-418.

22. Terral J-F, Tabard E, Bouby L, Ivorra S, Pastor T, Figueiral I, et al. Evolution and history of grapevine (Vitis vinifera) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars. Annals of Botany. 2010;105: 443-455. doi:10.1093/aob/mcp298

23. Orrù M, Grillo O, Lovicu G, Venora G, Bacchetta G. Morphological characterisation of *Vitis vinifera* L. seeds by image analysis and comparison with archaeological remains. Vegetation History and Archaeobotany. 2013;22: 231-242. doi:10.1007/s00334-012-0362-2

24. Bouby L, Figueiral I, Bouchette A, Rovira N, Ivorra S, Lacombe T, et al. Bioarchaeological Insights into the Process of Domestication of Grapevine (Vitis vinifera L.) during Roman Times in Southern France. Degryse P, éditeur. PLoS ONE. 2013;8: e63195. doi:10.1371/journal.pone.0063195

25. Pagnoux C, Bouby L, Ivorra S, Petit C, Valamoti S-M, Pastor T, et al. Inferring the agrobiodiversity of Vitis vinifera L. (grapevine) in ancient Greece by comparative shape analysis of archaeological and modern seeds. Vegetation History and Archaeobotany. 2015;24: 75-84. doi:10.1007/s00334-014-0482-y

26. Ucchesu M, Orrù M, Grillo O, Venora G, Paglietti G, Ardu A, et al. Predictive method for correct identification of archaeological charred grape seeds: Support for advances in knowledge of grape domestication process. PLoS ONE. 2016;11: 1-18. doi:10.1371/journal.pone.0149814

27. Karasik A, Rahimi O, David M, Weiss E, Drori E. Development of a 3D seed morphological tool for grapevine variety identification, and its comparison with SSR analysis. Scientific Reports. 2018;8: 6545. doi:10.1038/s41598-018-24738-9

28. Wales N, Ramos Madrigal J, Cappellini E, Carmona Baez A, Samaniego Castruita JA, Romero-Navarro JA, et al. The limits and potential of paleogenomic techniques for reconstructing grapevine domestication. Journal of Archaeological Science. 2016;72: 57-70. doi:10.1016/j.jas.2016.05.014

29. Bacilieri R, Bouby L, Figueiral I, Schaal C, Terral J-F, Breton C, et al. Potential of combining morphometry and ancient DNA information to investigate grapevine domestication. Vegetation History and Archaeobotany. 2017;26: 345-356.

doi:10.1007/s00334-016-0597-4

30. Ramos-Madrigal J, Runge AKW, Bouby L, Lacombe T, Samaniego Castruita JA, Adam-Blondon A-F, et al. Palaeogenomic insights into the origins of French grapevine diversity. Nature Plants. 2019;5: 595-603. doi:10.1038/s41477-019-0437-5

31. Levadoux L. Les populations sauvages et cultivées des Vitis vinifera L. Annales de l'amélioration des plantes. 1956;6: 59-118.

32. Smith H, Jones G. Experiments on the effects of charring on cultivated grape seeds. Journal of Archaeological Science. 1990;17: 317-327. doi:10.1016/0305-4403(90)90026-2

33. Bouby L, Bonhomme V, Ivorra S, Pastor T, Rovira N, Tillier M, et al. Back from burn out: are experimentally charred grapevine pips too distorted to be characterized using morphometrics? Archaeological and Anthropological Sciences. 2018;10: 943-954. doi:10.1007/s12520-016-0425-x

34. Negrul A. Evolucija razmera semjan i jagod u vinograda. Izvaet TSHA. 1960; 167–176.

35. Houel C, Martin-Magniette M-L, Nicolas SD, Lacombe T, Le Cunff L, Franck D, et al. Genetic variability of berry size in the grapevine (Vitis vinifera L.). Australian Journal of Grape and Wine Research. 2013;19: 208-220. doi:10.1111/ajgw.12021

36. Bonnet A. L'hybridité chez la vigne. Etude de la graine. Annales de l'Ecole Nationale d'Agriculture de Montpellier. 1912;2: 5-134.

37. Edwards W. Plants reward seed dispersers in proportion to their effort: The relationship between pulp mass and seed mass in vertebrate dispersed plants. Evolutionary Ecology. 2006;20: 365-376. doi:10.1007/s10682-006-0006-z

38. Leishman MR. Does the seed size/number trade-off model determine plant community structure? An assessment of the model mechanisms and their generality. Oikos. 2001;93: 294-302. doi:10.1034/j.1600-0706.2001.930212.x

39. Dokoozlian N. Grape berry growth and development. Raisin Production Manual. Oakland: University of California, Agricultural and Natural Resources; 2000. p. 30-37.

40. Ristic R, Iland PG. Relationships between seed and berry development of Vitis Vinifera L. cv Shiraz: Developmental changes in seed morphology and phenolic composition. Australian Journal of Grape and Wine Research. 2005;11: 43-58. doi:10.1111/j.1755-0238.2005.tb00278.x

41. Ojeda H, Deloire A, Carbonneau A, Ageorges A, Romieu C. Berry development of grapevines: Relations between the growth of berries and their DNA content indicate cell multiplication and enlargement. Vitis. 1999;38: 145-150.

42. Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M, et al. Current perspectives and the future of domestication studies. Proceedings of the National Academy of Sciences. 2014;111: 6139-6146. doi:10.1073/pnas.1323964111
43. Sultan SE. Development in context: the timely emergence of eco-devo. Trends in Ecology & Evolution. 2007;22: 575-582. doi:10.1016/j.tree.2007.06.014
44. Figueiral I, Pomarèdes H, Court-Picon M, Bouby L, Tardy C, Terral J-F. New insights into Mediterranean Gallo-Roman farming: a closer look at archaeological wells in Southern France. Archaeological and Anthropological Sciences. 2015;7: 201-233. doi:10.1007/s12520-014-0181-8
45. Figueiral I, Bouby L, Buffat L, Petitot H, Terral JF. Archaeobotany, vine growing and wine producing in Roman Southern France: the site of Gasquinoy (Béziers, Hérault). Journal of Archaeological Science. 2010;37: 139-149. doi:10.1016/j.jas.2009.09.024
46. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Vienna, Austria; 2020.
47. Bonhomme V, Picq S, Gaucherel C, Claude J. Momocs: Outline Analysis Using R. Journal of Statistical Software. 2014;56. doi:10.18637/jss.v056.i13
48. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. Journal of Open Source Software. 2019;4: 1686. doi:10.21105/joss.01686
49. Rasband W. ImageJ. 2008.
50. Claude J. Morphometrics with R. New York, NY: Springer New York; 2008. doi:10.1007/978-0-387-77789-4
51. Evin A, Cucchi T, Cardini A, Strand Vidarsdottir U, Larson G, Dobney K. The long and winding road: identifying pig domestication through molar size and shape. Journal of Archaeological Science. 2013;40: 735-743. doi:10.1016/j.jas.2012.08.005

Figures legends

Figure 1 Dorsal and lateral views of a grapevine pip, here from the *Vitis vinifera* subsp. *sylvestris* wild" individual "Pic Saint-Loup 13", with indications of morphometric measurements: pip_{length} (total length), pip_{stalk} (length of the stalk), $pip_{chalaza}$ (position of the chalaza), pip_{width} and $pip_{thickness}$.

Figure 2 Distribution of the number of pips per berry for wild and domesticated grapevines.

Figure 3 Comparisons of (logged) lengths and (log cubic-rooted) mass measurements. Each row represents a different comparison: a) wild and domesticated grapevines, b) table and wine varieties for domesticated accessions only, and c) cultivated and collected from wild for wild grapevine only. For each measurement, boxplots are displayed for each piposity level. Differences are tested using multivariate analyses of covariances, and differences of $P<10^{-5}$, are indicated by stars in the facet title (interaction), on the right (overall difference) and above each piposity level (difference within a given piposity).

Figure 4 Principal component analysis on the joint matrices of Fourier coefficients obtained for the two views. The first two principal component gathered 69% of the total variance. The component of shape variation they capture are illustrated with reconstructed shapes at each extreme of their range. Colour of markers and convex hulls indicate pips from wild (green) and domesticated (blue) grapevines.

Figure 5 a) Regressions PC1 and PC2 versus pip length for domesticated pips (blue) and cultivation (green); b) boxplots for each piposity level for domesticated pips (blue) and cultivation (green). Differences are tested using Wilcoxon rank tests, and differences of $P<10^{-5}$, are indicated by brackets above graphs.

Figure 6 Mean shapes calculated for all pips sampled from berry with 1 (blue) or 4 pips (red), for different subsets (in rows) and for the dorsal and lateral views (in columns). Between the two views, an index of shape differences between these two extreme piposity levels (0=identical shapes; unit=as much differences as between wild and domesticated average shapes).

Figure 7 Relative contribution of berry height, accession and number of pips per berry (coloured bars) onto the shape of pips for different subsets.

Figure 8 Classification accuracy (LDA leave-one-out) at the a) status and b) accessions levels. Models are trained (and evaluated) on the admixture of pips, then evaluated on piposity subsets. Different combinations of training data are used (Fourier coefficients of shape, lengths/mass measurements, both). Lines provide a random baseline and summarise 10000 permutations: solid line correspond to mean accuracy; dashed line to the maximal values obtained.

Figure 9 Regressions for berry dimensions from pip dimensions, obtained on modern material: predicted versus actual (logged) berry dimensions at the domestication status level (all accessions). Columns are for berry diameter and height, respectively.

645 **Figure 10** Distribution of pip lengths (observed for all), berry height and diameter inferred from pip dimensions for the archaeological material from Sauvian – La Lesse. Rows distinguish domesticated and wild grapes since separate regressions were required.

Figure A (ESM) Comparisons of (logged) lengths and (cubic-rooted) mass measurements. On rows are displayed different subsets: a) wild and domesticated grapevines,
650 b) for domesticated accessions, table and wine varieties and, c) for wild accessions, those collected *in natura* and others cultivated as domesticated varieties. Different piposity levels are pooled (see Figure 3 for the detail). Differences are tested using Wilcoxon rank tests and all of them have a $P<10^{-5}$.

Figure B (ESM) Bivariate pairwise plot between (logged) lengths and (cubic-rooted)
655 mass measurements. For the sake of readability, only the wild versus domesticated status are displayed using different colours (green for wild; blue for domesticated). If two regressions are justified, then they are shown using the corresponding colours; otherwise a single regression line is showed in black. Then, for each regression, the correlations are tested and, if significant, the adjusted R^2 is displayed on the regression lines.

660 **Figure C (ESM)** Predictions obtained from regressions for berry dimensions from pip dimensions on modern material. The relative deviation, at the accession level, and for unlogged measurements. Columns are for berry diameter and height, respectively.

Tables

Table 1 Accessions used in this study. D: domesticated; W: wild; Wn: wine grape; Tb: table grape. For domesticated grapevines, names correspond to the variety names. Dimensions are reported with mean±sd and given in mm, except for $berry_{mass}$ which is expressed in g.

id	accession	status	dest./cult.	pip_{Length}	$pip_{LengthStalk}$	$pip_{PositionChalaza}$	$pip_{Breadth}$	$pip_{Thickness}$	$berry_{Height}$	$berry_{Diameter}$	$berry_{Mass}$
cAlvarB	Alvarinho	dom.	wine	5.18 ± 0.28	1.35 ± 0.19	2.66 ± 0.23	3.5 ± 0.21	2.5 ± 0.15	10.6 ± 0.81	10.93 ± 0.98	0.82 ± 0.21
cBarbeN	Barbera	dom.	wine	6.73 ± 0.33	1.95 ± 0.14	3.58 ± 0.2	3.67 ± 0.23	2.7 ± 0.14	16.38 ± 1.07	15 ± 1.07	2.44 ± 0.49
cCabSaN	Cabernet-Sauvignon	dom.	wine	5.93 ± 0.34	1.78 ± 0.16	3.09 ± 0.2	3.57 ± 0.22	2.65 ± 0.17	14.05 ± 0.89	13.47 ± 0.84	1.57 ± 0.27
cCarigN	Carignan	dom.	wine	6.85 ± 0.26	2.14 ± 0.21	3.85 ± 0.24	3.54 ± 0.22	2.73 ± 0.09	18.1 ± 0.83	16.85 ± 0.84	3.16 ± 0.41
cChardB	Chardonnay	dom.	wine	5.28 ± 0.41	1.38 ± 0.2	3.06 ± 0.28	3.72 ± 0.28	2.78 ± 0.27	13.42 ± 1.24	12.9 ± 1.16	1.69 ± 0.41
cChevkN	Chevka	dom.	wine	5.97 ± 0.33	1.48 ± 0.13	3.05 ± 0.19	3.73 ± 0.19	2.53 ± 0.1	18.08 ± 1.17	16.51 ± 1.11	3.09 ± 0.58
cDebinB	Debina	dom.	wine	7.01 ± 0.48	1.73 ± 0.21	3.7 ± 0.3	4.21 ± 0.28	3.41 ± 0.19	19.61 ± 1.9	17.67 ± 1.58	3.89 ± 0.87
cFesAlB	Feteasca alba	dom.	wine	5.19 ± 0.23	1.4 ± 0.13	2.65 ± 0.19	3.46 ± 0.19	2.5 ± 0.13	12.49 ± 0.79	12.99 ± 0.69	1.62 ± 0.26
cGaidoB	Gaïdouria	dom.	wine	5.55 ± 0.37	1.61 ± 0.19	2.92 ± 0.23	3.8 ± 0.23	2.94 ± 0.14	16.51 ± 1.36	17.1 ± 1.3	3.46 ± 0.73
cGrenaN	Grenache	dom.	wine	5.32 ± 0.29	1.46 ± 0.1	2.84 ± 0.17	3.38 ± 0.2	2.45 ± 0.14	15.91 ± 0.78	15.12 ± 0.66	2.3 ± 0.27
cKypreN	Kypreiko	dom.	wine	6.79 ± 0.39	2.05 ± 0.23	3.81 ± 0.26	3.99 ± 0.27	3.07 ± 0.12	21.38 ± 1.35	17.98 ± 1.16	4.26 ± 0.79
cMavruN	Mavrud	dom.	wine	6.9 ± 0.26	2.35 ± 0.18	3.93 ± 0.2	4.3 ± 0.19	3.19 ± 0.14	14.91 ± 0.9	14.53 ± 0.84	2.04 ± 0.31
cMerloN	Merlot	dom.	wine	6.02 ± 0.32	1.5 ± 0.15	3.01 ± 0.24	3.77 ± 0.24	2.75 ± 0.18	12.62 ± 1.04	12.83 ± 1.16	1.41 ± 0.31
cMesFrB	Meslier Saint François	dom.	wine	5.51 ± 0.59	1.53 ± 0.2	3.05 ± 0.34	3.27 ± 0.31	2.61 ± 0.29	16.59 ± 1.64	15.54 ± 1.55	2.86 ± 0.77
cMourvN	Mourvèdre	dom.	wine	6.17 ± 0.41	1.51 ± 0.18	3.11 ± 0.27	3.73 ± 0.28	2.89 ± 0.2	15.71 ± 1.04	15.27 ± 1.16	2.28 ± 0.49
cMusPGR	Muscat à petits grains roses	dom.	wine	5.23 ± 0.3	1.33 ± 0.19	2.95 ± 0.23	3.12 ± 0.18	2.39 ± 0.13	12.93 ± 1.16	13.82 ± 0.9	1.97 ± 0.28
cMusPGN	Muscat noir à petits grains	dom.	wine	7.31 ± 0.29	2.27 ± 0.19	4.18 ± 0.26	4.2 ± 0.25	3.3 ± 0.16	20.07 ± 1.3	18.99 ± 1.39	
cPinNoN	Pinot noir	dom.	wine	6.35 ± 0.31	1.77 ± 0.21	3.54 ± 0.2	3.93 ± 0.4	2.78 ± 0.19	13.87 ± 1.05	13.42 ± 1.19	1.83 ± 0.39
cRoussB	Roussanne	dom.	wine	6.57 ± 0.23	1.46 ± 0.15	3.39 ± 0.19	4.16 ± 0.23	3.09 ± 0.2	15.37 ± 0.77	14.91 ± 0.83	2.13 ± 0.32
cSauviB	Sauvignon	dom.	wine	6.53 ± 0.39	1.91 ± 0.19	3.49 ± 0.28	3.74 ± 0.18	2.68 ± 0.16	15.46 ± 1.07	13.45 ± 0.87	1.78 ± 0.3
cSyrahN	Syrah	dom.	wine	5.39 ± 0.35	1.64 ± 0.18	3.14 ± 0.28	3.21 ± 0.17	2.54 ± 0.15	15.37 ± 0.91	13.51 ± 0.85	1.99 ± 0.29
cAinBoB	Ain el Bouma	dom.	table	5.56 ± 0.41	1.9 ± 0.2	3.42 ± 0.23	3.53 ± 0.29	2.63 ± 0.14	21.11 ± 2.06	17.96 ± 1.8	4.51 ± 1.1
cChaBlB	Chaouch blanc	dom.	table	6.95 ± 0.26	1.74 ± 0.16	3.86 ± 0.31	4.08 ± 0.25	3.44 ± 0.21	24.67 ± 1.85	22.55 ± 1.82	7.76 ± 1.59
cFerTaR	Ferral tamara	dom.	table	6.79 ± 0.25	1.95 ± 0.18	3.52 ± 0.17	4.21 ± 0.18	3.29 ± 0.16	23.45 ± 1.22	19.73 ± 1.33	5.79 ± 0.83
cHadarB	Hadari	dom.	table	7.42 ± 0.38	1.89 ± 0.19	3.49 ± 0.28	4.2 ± 0.19	3.24 ± 0.15	21.54 ± 1.2	19.37 ± 1.03	4.73 ± 0.7
cHunisN	Hunisa	dom.	table	8.32 ± 0.39	2.27 ± 0.23	4.41 ± 0.27	4.94 ± 0.25	3.4 ± 0.2	30.28 ± 2.53	23.92 ± 1.68	9.87 ± 1.48
cKarPaRs	Kara papigi	dom.	table	6.58 ± 0.36	1.92 ± 0.2	3.21 ± 0.29	3.65 ± 0.22	2.8 ± 0.18	18.45 ± 1.67	16.29 ± 1.33	3 ± 0.73
cKraTzN	Kravi tzitzi	dom.	table	7 ± 0.31	1.92 ± 0.21	3.51 ± 0.22	3.96 ± 0.17	3.09 ± 0.14	28.02 ± 2.11	19.37 ± 1.43	6.43 ± 1.19
cPemGeR	Pembe Gemre	dom.	table	6.98 ± 0.33	2.11 ± 0.18	3.93 ± 0.26	3.84 ± 0.18	2.91 ± 0.09	16.46 ± 0.83	16.16 ± 1.02	3.07 ± 0.46
cSlivaN	Sliva	dom.	table	6.98 ± 0.47	1.95 ± 0.29	3.67 ± 0.35	4.02 ± 0.2	3.23 ± 0.11	23.41 ± 1.65	19.8 ± 1.74	5.59 ± 1.28
wCamSa4	Camp Saure	wild	in natura	5.23 ± 0.26	1.04 ± 0.15	2.46 ± 0.17	3.38 ± 0.17	2.5 ± 0.2	8.29 ± 0.83	8.15 ± 1.02	
wChala7	Chalabre	wild	in natura	4.83 ± 0.28	0.92 ± 0.14	2.29 ± 0.17	3.59 ± 0.16	2.52 ± 0.25	7.68 ± 0.75	8.21 ± 0.92	
wCoBab2	Col de la Babourade	wild	in natura	4.8 ± 0.56	0.91 ± 0.22	2.41 ± 0.31	3.24 ± 0.32	2.31 ± 0.27	8.26 ± 0.96	8.56 ± 0.93	
wEscal13	L'Escale (13)	wild	in natura	5.51 ± 0.38	1.07 ± 0.18	2.84 ± 0.22	3.58 ± 0.2	2.76 ± 0.2	9.44 ± 0.97	9.14 ± 1.06	
wEscal14	L'Escale (14)	wild	in natura	4.84 ± 0.32	0.82 ± 0.12	2.24 ± 0.18	3.32 ± 0.15	2.4 ± 0.17	7.72 ± 0.81	7.81 ± 0.95	0.32 ± 0.1
wEscal16	L'Escale (16)	wild	in natura	5.34 ± 0.28	1.18 ± 0.14	2.69 ± 0.17	3.6 ± 0.26	2.56 ± 0.23	7.74 ± 0.7	7.76 ± 0.92	0.35 ± 0.09
wEscal17	L'Escale (17)	wild	in natura	4.94 ± 0.27	1.02 ± 0.16	2.38 ± 0.16	3.61 ± 0.18	2.77 ± 0.19	7.37 ± 0.68	7.41 ± 0.85	0.33 ± 0.09
wEscal18	L'Escale (18)	wild	in natura	5.17 ± 0.46	0.72 ± 0.16	2.34 ± 0.24	3.95 ± 0.37	2.78 ± 0.33	7.45 ± 0.78	7.84 ± 0.8	
wEscal20	L'Escale (20)	wild	in natura	5.05 ± 0.24	0.89 ± 0.14	2.39 ± 0.17	3.48 ± 0.2	2.53 ± 0.21	8.05 ± 0.67	8.23 ± 0.87	0.37 ± 0.1
wCalme1(	La Calmette (10)	wild	in natura	5.28 ± 0.31	0.89 ± 0.15	2.58 ± 0.23	3.7 ± 0.24	2.77 ± 0.27	7.98 ± 0.74	8.14 ± 0.82	
wCalme11	La Calmette (11)	wild	in natura	4.93 ± 0.42	1.15 ± 0.25	2.66 ± 0.31	3.33 ± 0.24	2.59 ± 0.24	7.66 ± 1.09	7.57 ± 1.14	
wPSL13	Pic Saint Loup (13)	wild	in natura	5.35 ± 0.29	0.98 ± 0.12	2.85 ± 0.21	3.75 ± 0.22	2.8 ± 0.3	8.75 ± 0.63	8.55 ± 0.73	0.42 ± 0.1
wPSLH	Pic Saint Loup (H)	wild	in natura	5.07 ± 0.44	0.75 ± 0.19	2.21 ± 0.27	3.78 ± 0.29	2.69 ± 0.26	7.67 ± 0.76	8.14 ± 0.88	
wRivel1	Rivel	wild	in natura	5.83 ± 0.55	0.97 ± 0.27	2.77 ± 0.28	4.08 ± 0.28	3.24 ± 0.26	7.8 ± 0.74	7.88 ± 0.95	0.38 ± 0.11
cKetsc27	Ile de Ketch (27)	wild	cultivated	5.68 ± 0.27	1.08 ± 0.15	2.63 ± 0.2	3.86 ± 0.27	2.84 ± 0.19	8.88 ± 0.72	8.75 ± 0.71	0.59 ± 0.11
cPalmA	Palma	wild	cultivated	5.78 ± 0.32	1.51 ± 0.19	3.06 ± 0.2	3.72 ± 0.14	2.66 ± 0.23	11.27 ± 1.3	10.78 ± 1.44	0.91 ± 0.33
cPSL13	Pic Saint Loup (13)	wild	cultivated	5.8 ± 0.34	0.91 ± 0.19	2.74 ± 0.25	4.12 ± 0.2	3.14 ± 0.25	8.79 ± 0.73	8.72 ± 0.96	0.52 ± 0.14
cPSL5	Pic Saint Loup (5)	wild	cultivated	6.05 ± 0.17	1.14 ± 0.09	2.82 ± 0.13	3.9 ± 0.18	2.79 ± 0.13	10.67 ± 0.57	10.9 ± 0.65	0.77 ± 0.12
wLambrN	wLambrN	wild	cultivated	5.48 ± 0.3	1.06 ± 0.1	2.71 ± 0.2	3.85 ± 0.18	2.94 ± 0.25	9.12 ± 0.84	8.53 ± 0.96	0.54 ± 0.15

Table 2 Estimates for pip lengths used to infer berry dimensions. Variables were all logged; so that $berry_{Diameter}$ (in mm and for wild) can be obtained with $\exp[0.65163 + 1.19914 \times \log(pip_{Length}) + 0.10617 \times \log(pip_{PositionChalaza}) - 0.13263 \times \log(pip_{Breadth}) - 0.45449 \times \log(pip_{Thickness})]$.

			Estimates of predictors						
log Predicted:	subset	Intercept	log pip_{Length}	log $pip_{LengthStalk}$	log $pip_{PositionChalaza}$	log $pip_{Breadth}$	log $pip_{Thickness}$	PC1	PC2
$berry_{Diameter}$	wild	0.689	1.136	–	–	–	-0.438	-0.223	0.209
	domesticated	1.298	0.478	–	0.149	-0.150	0.520	–	0.430
$berry_{Height}$	wild	0.838	0.732	–	0.216	0.171	-0.318	-0.524	0.258
	domesticated	0.760	1.312	0.279	–	-0.783	0.584	1.021	–

Figure 1

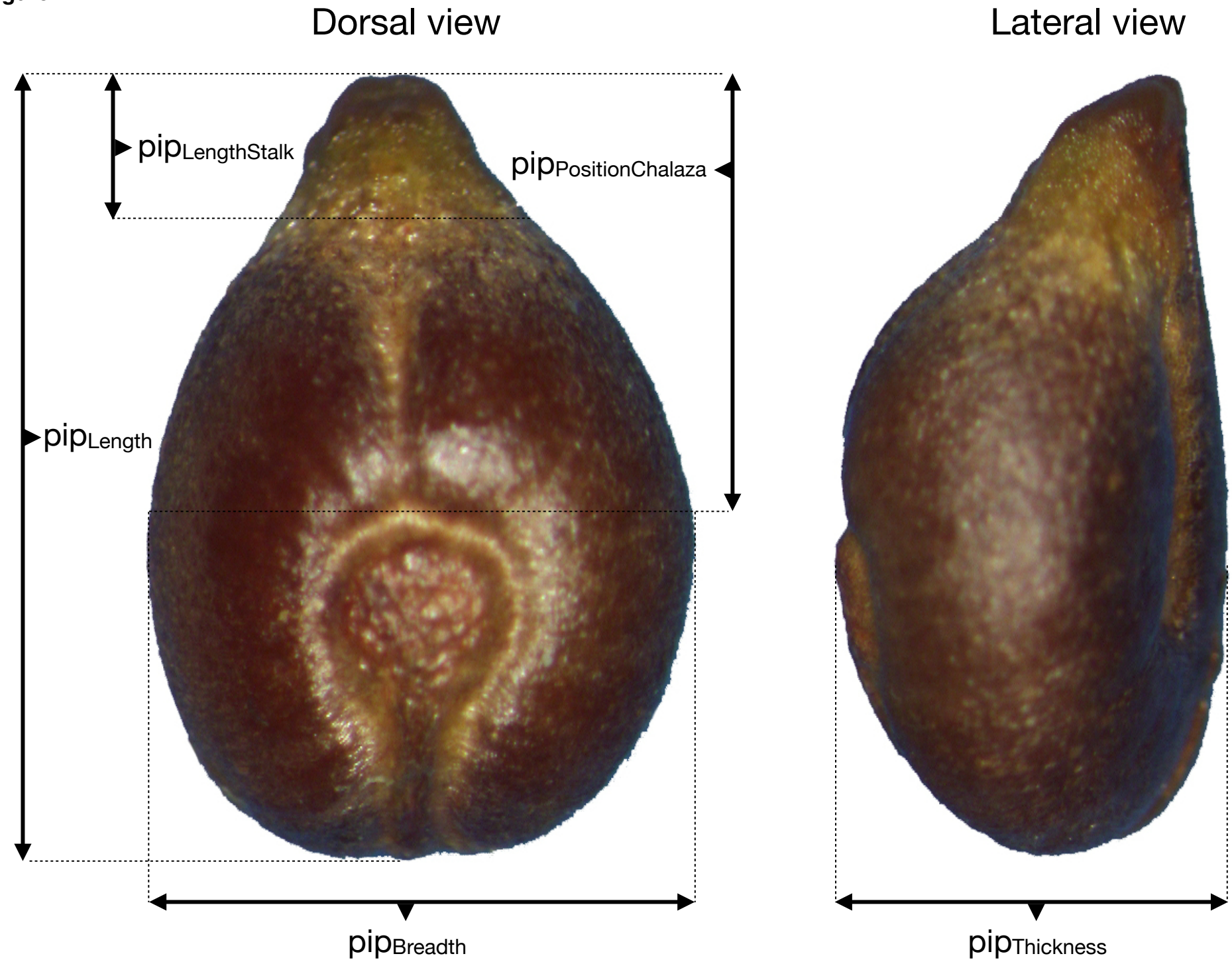

Dorsal view
Lateral view
pipLengthStalk
pipPositionChalaza
pipLength
pipBreadth
pipThickness

Figure 2

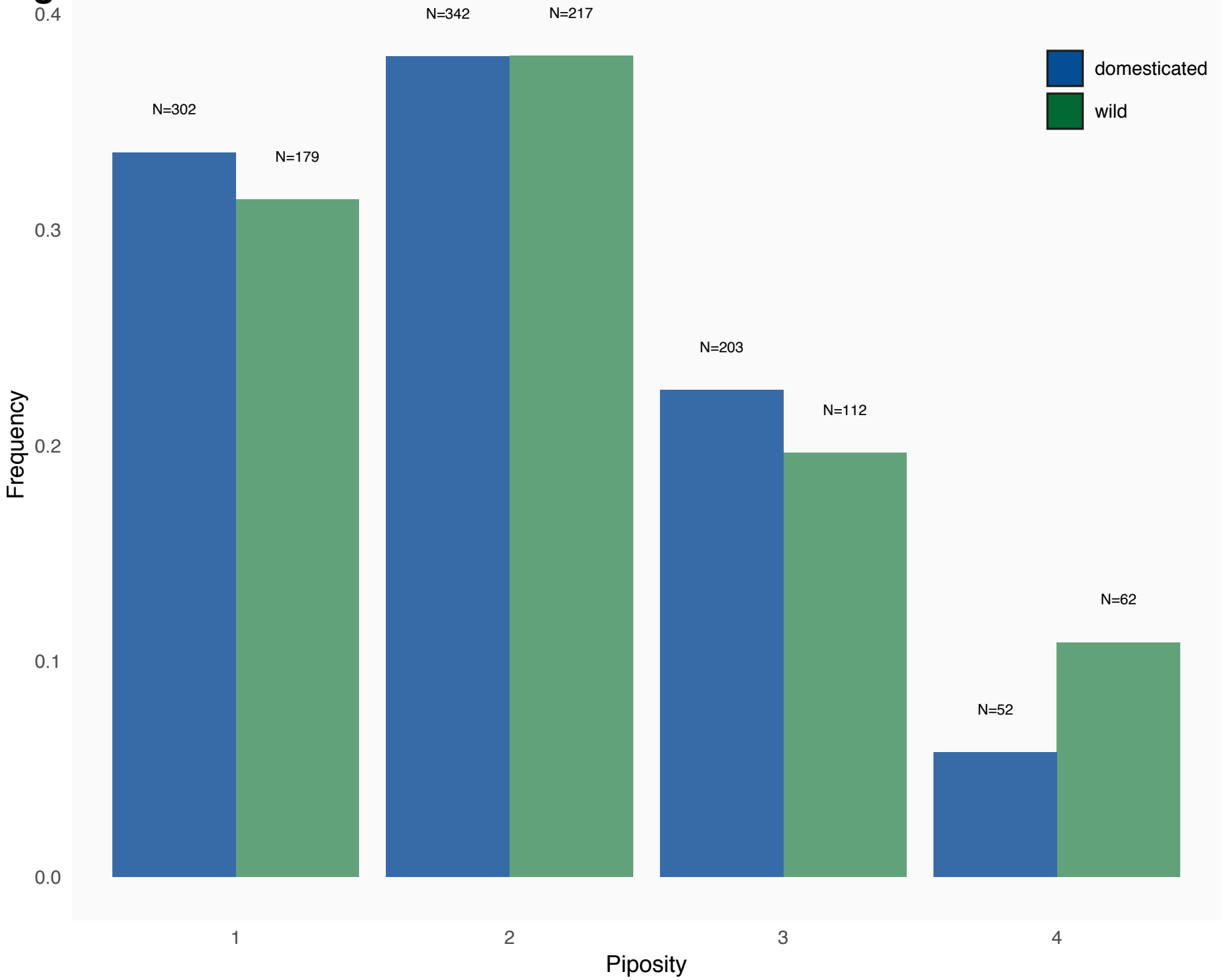

0.4
0.3
0.2
0.1
0.0
Frequency
N=302
N=179
N=342
N=217
N=203
N=112
N=52
N=62
domesticated
wild
1
2
3
4
Piposity

Figure 3

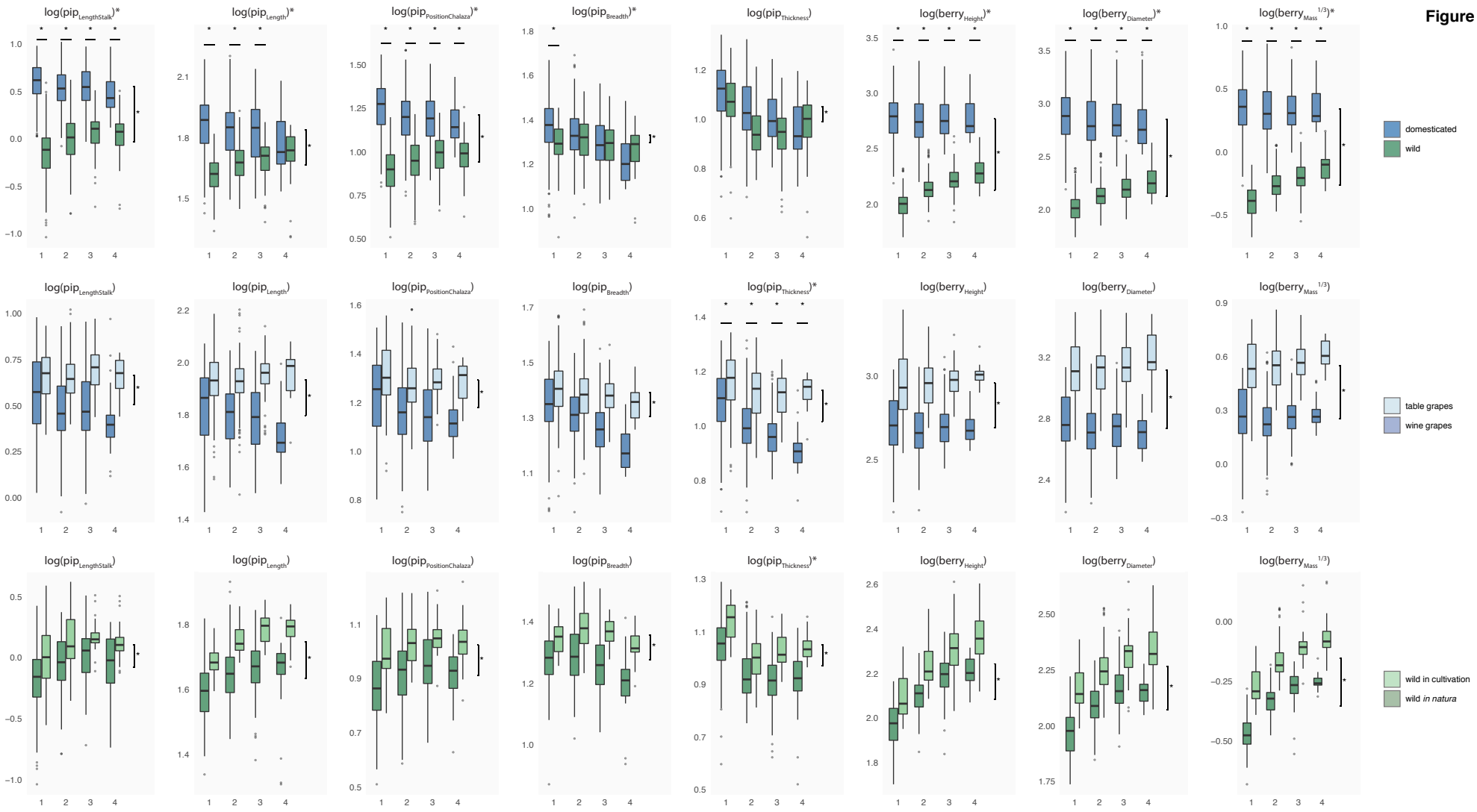

log(pip_LengthStalk)*
log(pip_Length)*
log(pip_PositionChalaza)*
log(pip_Breadth)*
log(pip_Thickness)
log(berry_Height)*
log(berry_Diameter)*
log(berry_Mass^1/3)*
log(pip_LengthStalk)
log(pip_Length)
log(pip_PositionChalaza)
log(pip_Breadth)
log(pip_Thickness)*
log(berry_Height)
log(berry_Diameter)
log(berry_Mass^1/3)
log(pip_LengthStalk)
log(pip_Length)
log(pip_PositionChalaza)
log(pip_Breadth)
log(pip_Thickness)*
log(berry_Height)
log(berry_Diameter)
log(berry_Mass^1/3)
domesticated
wild
table grapes
wine grapes
wild in cultivation
wild *in natura*

Figure 4

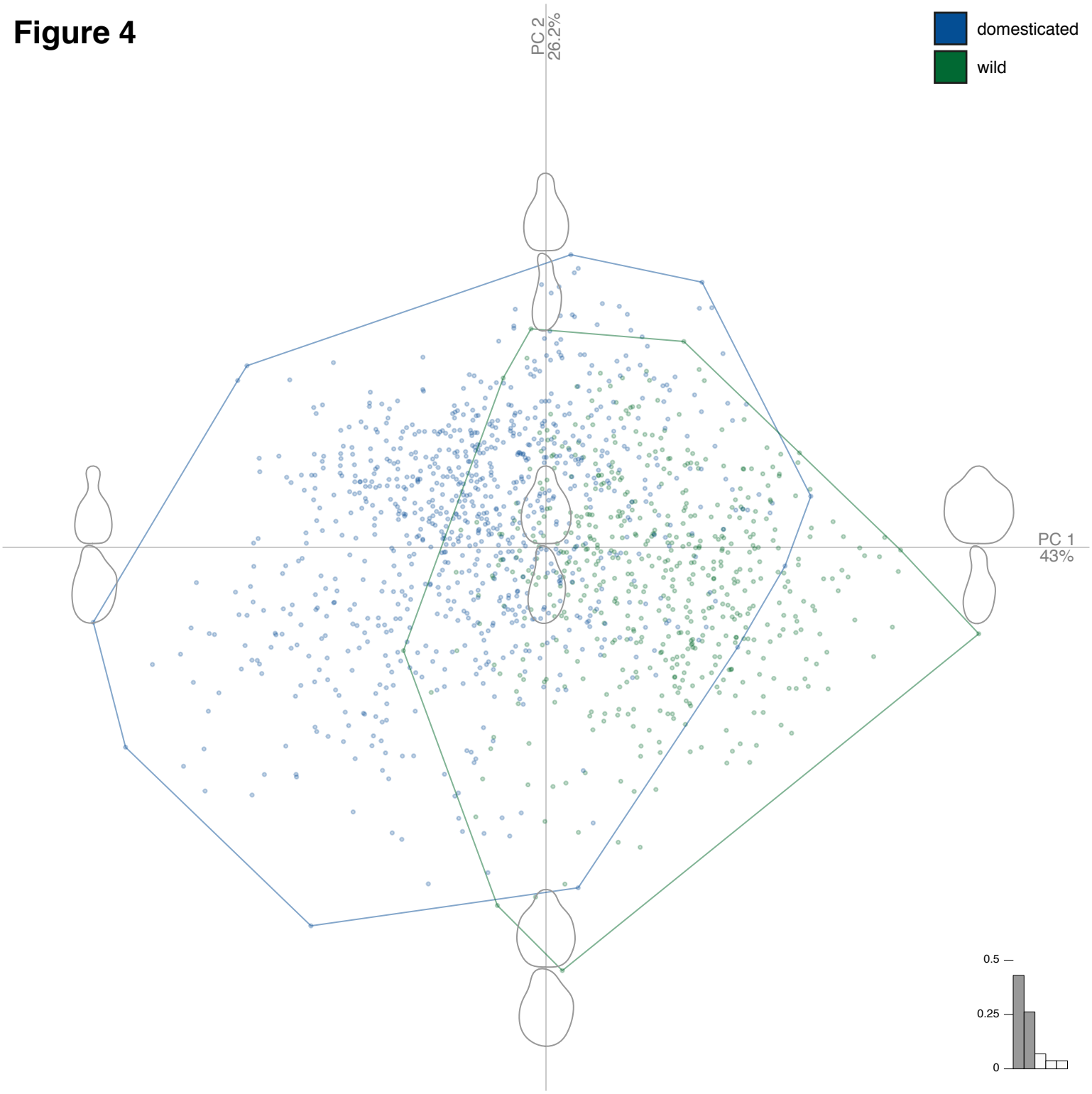


Figure 5

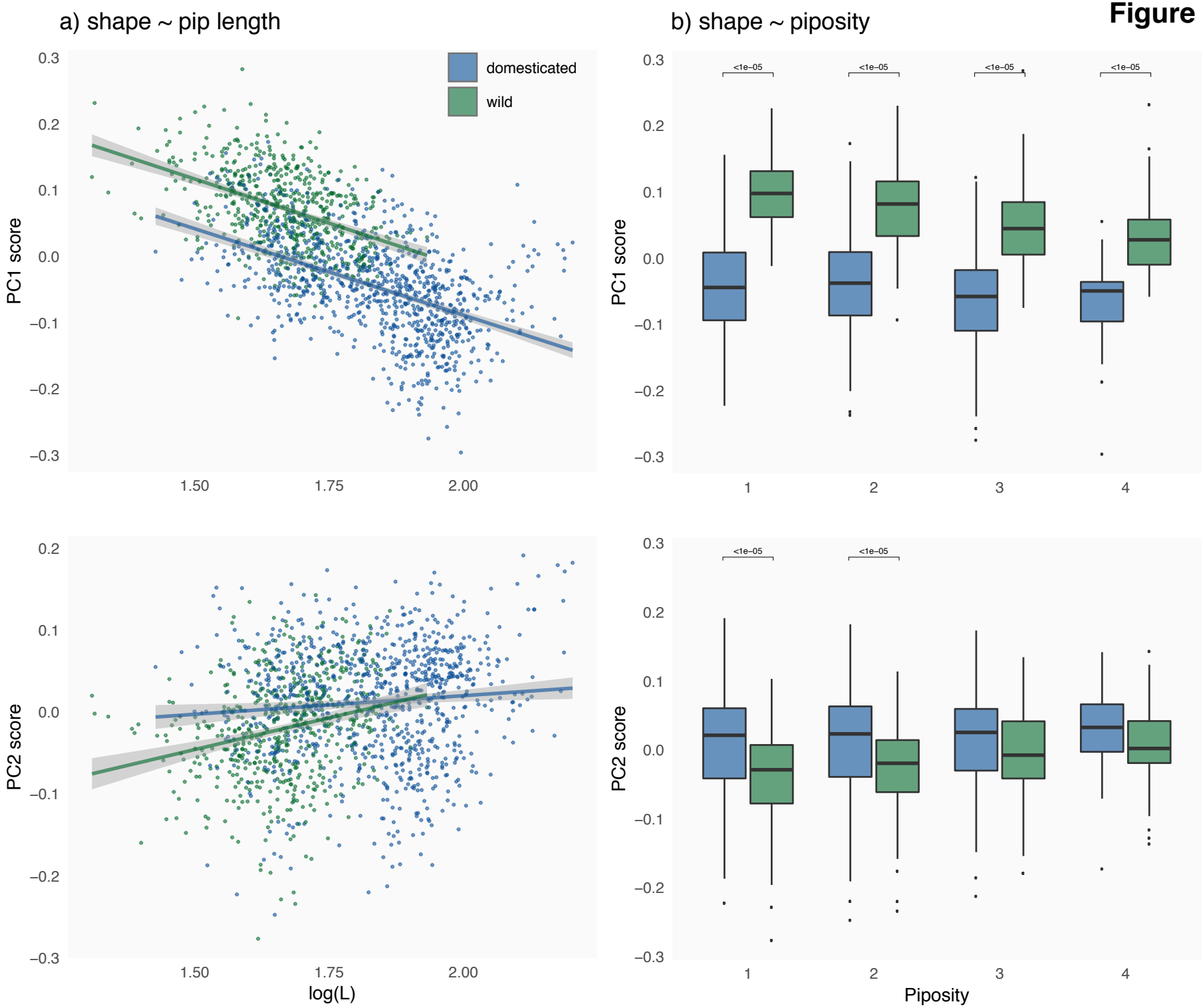

a) shape ~ pip length
b) shape ~ piposity
domesticated
wild
PC1 score
PC2 score
log(L)
Piposity
<1e−05
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
1.50
1.75
2.00
1
2
3
4

Figure 6

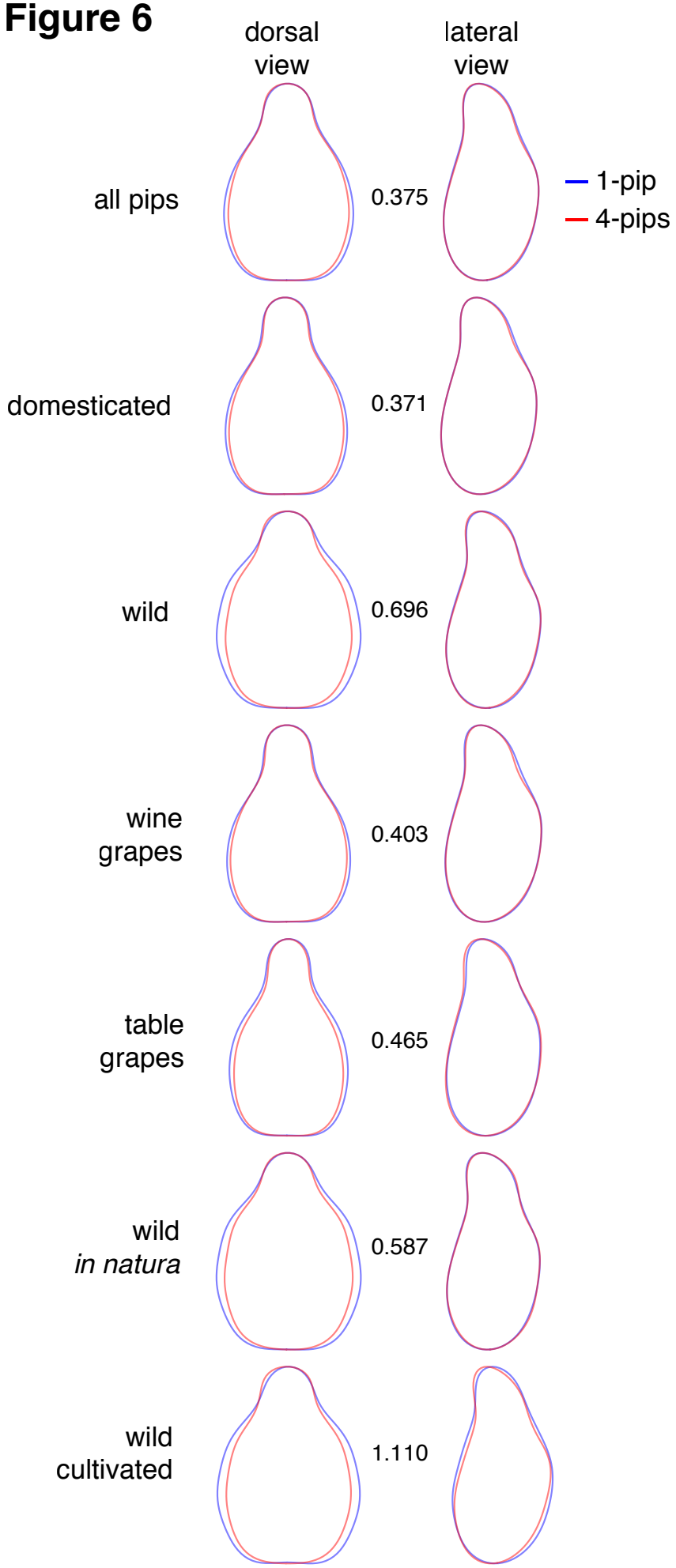

dorsal view
lateral view
1-pip
4-pips
all pips
0.375
domesticated
0.371
wild
0.696
wine grapes
0.403
table grapes
0.465
wild *in natura*
0.587
wild cultivated
1.110

Figure 7

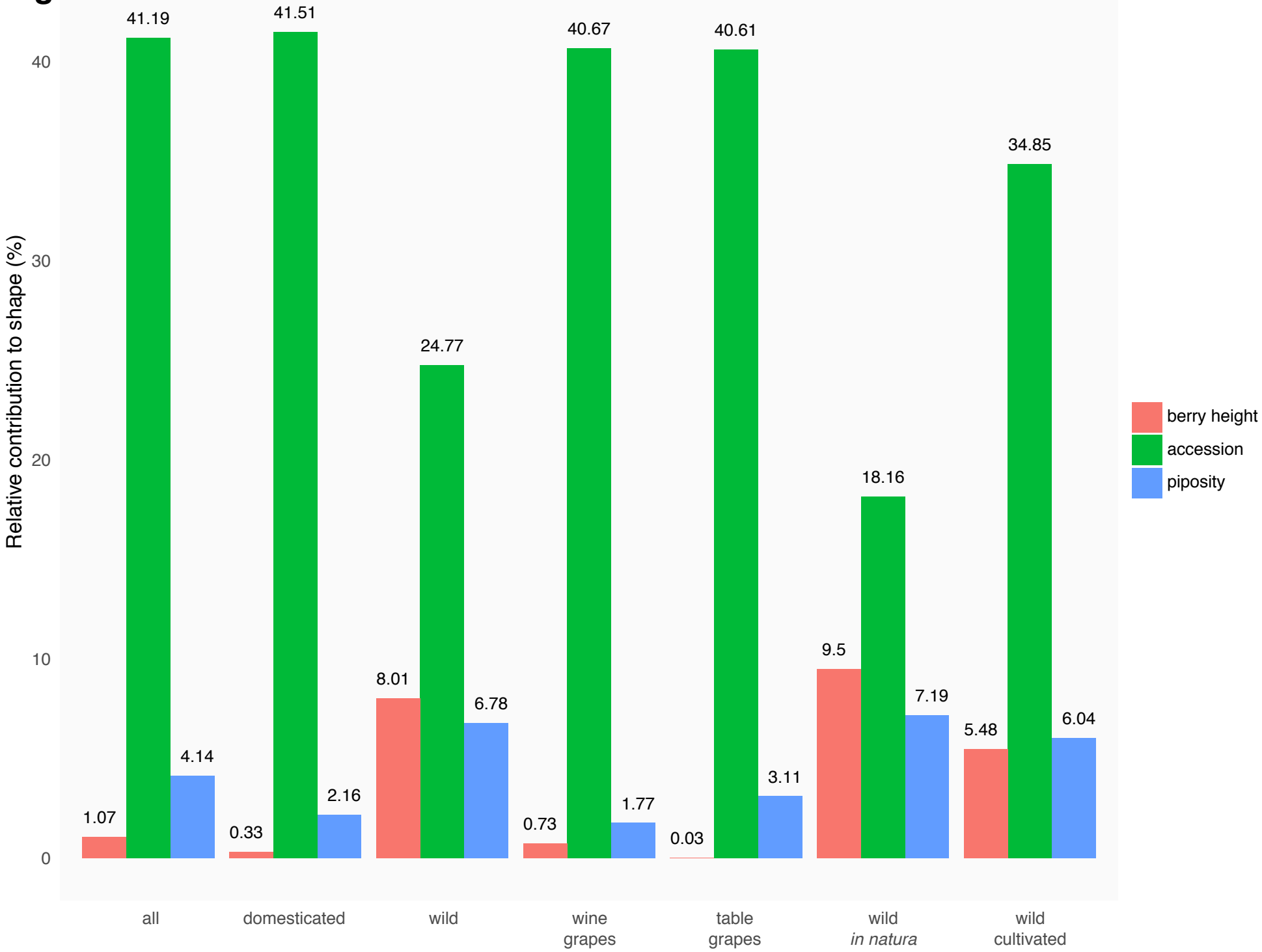

Relative contribution to shape (%)
0
10
20
30
40
41.19
41.51
40.67
40.61
34.85
24.77
18.16
9.5
8.01
7.19
6.78
6.04
5.48
4.14
3.11
2.16
1.77
1.07
0.73
0.33
0.03
all
domesticated
wild
wine grapes
table grapes
wild *in natura*
wild cultivated
berry height
accession
piposity

Figure 8

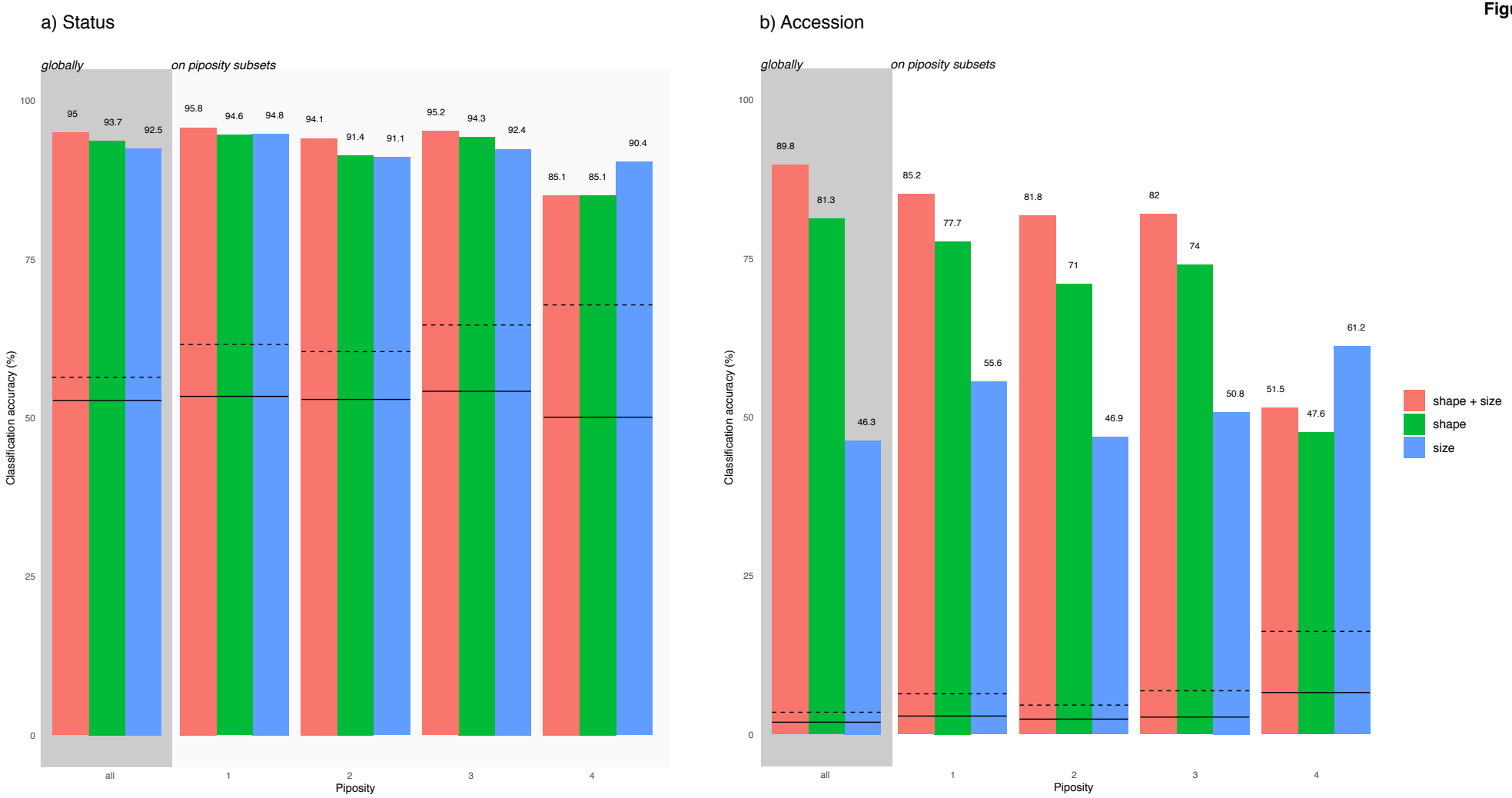

a) Status
globally
on piposity subsets
Classification accuracy (%)
100
75
50
25
0
95
93.7
92.5
95.8
94.6
94.8
94.1
91.4
91.1
95.2
94.3
92.4
85.1
85.1
90.4
all
1
2
3
4
Piposity
b) Accession
globally
on piposity subsets
Classification accuracy (%)
100
75
50
25
0
89.8
81.3
46.3
85.2
77.7
55.6
81.8
71
46.9
82
74
50.8
51.5
47.6
61.2
all
1
2
3
4
Piposity
shape + size
shape
size

Figure 9

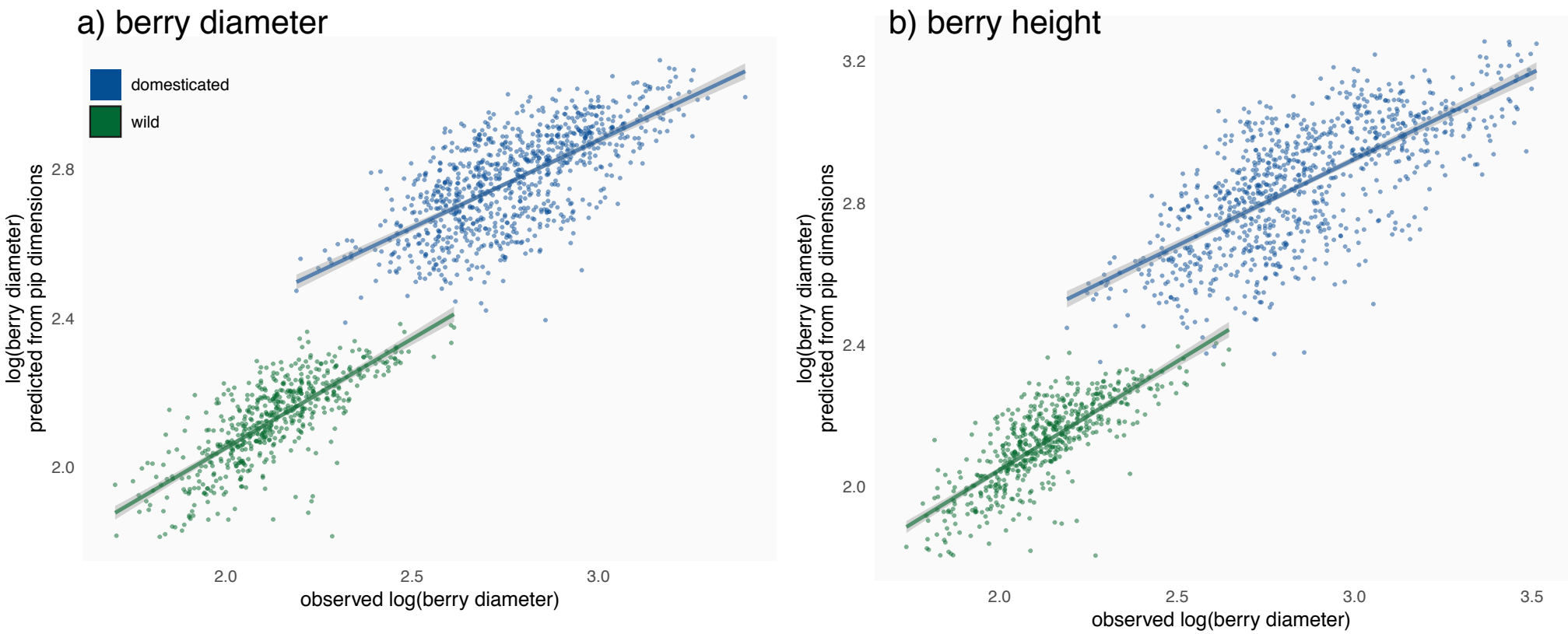
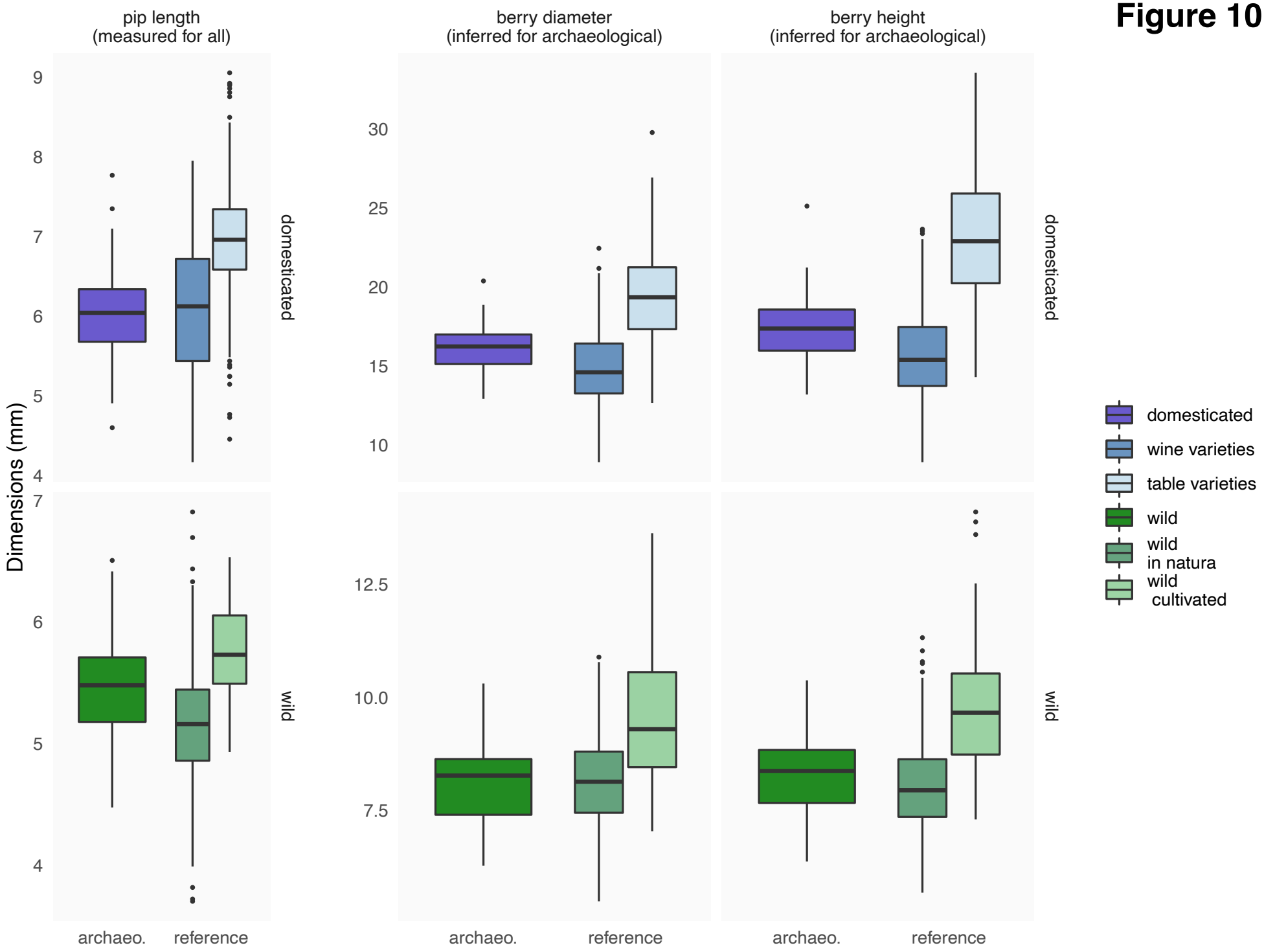

a) berry diameter
domesticated
wild
log(berry diameter)
predicted from pip dimensions
2.0
2.4
2.8
2.0
2.5
3.0
observed log(berry diameter)
b) berry height
log(berry diameter)
predicted from pip dimensions
2.0
2.4
2.8
3.2
2.0
2.5
3.0
3.5
observed log(berry diameter)

Figure 10


pip length
(measured for all)
berry diameter
(inferred for archaeological)
berry height
(inferred for archaeological)
Dimensions (mm)
domesticated
wild
archaeo.
reference
domesticated
wine varieties
table varieties
wild
wild
in natura
wild
cultivated

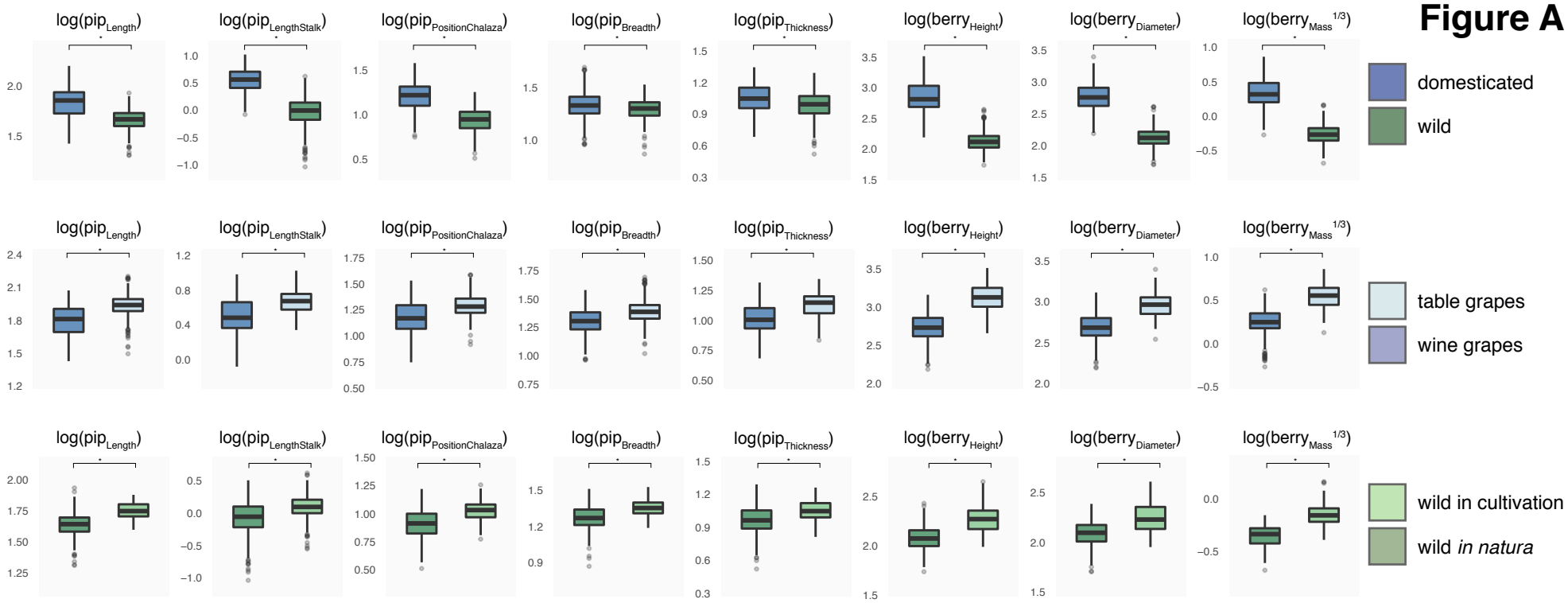


Figure A
log(pip_Length)
log(pip_LengthStalk)
log(pip_PositionChalaza)
log(pip_Breadth)
log(pip_Thickness)
log(berry_Height)
log(berry_Diameter)
log(berry_Mass^1/3)
domesticated
wild
table grapes
wine grapes
wild in cultivation
wild *in natura*

Figure B

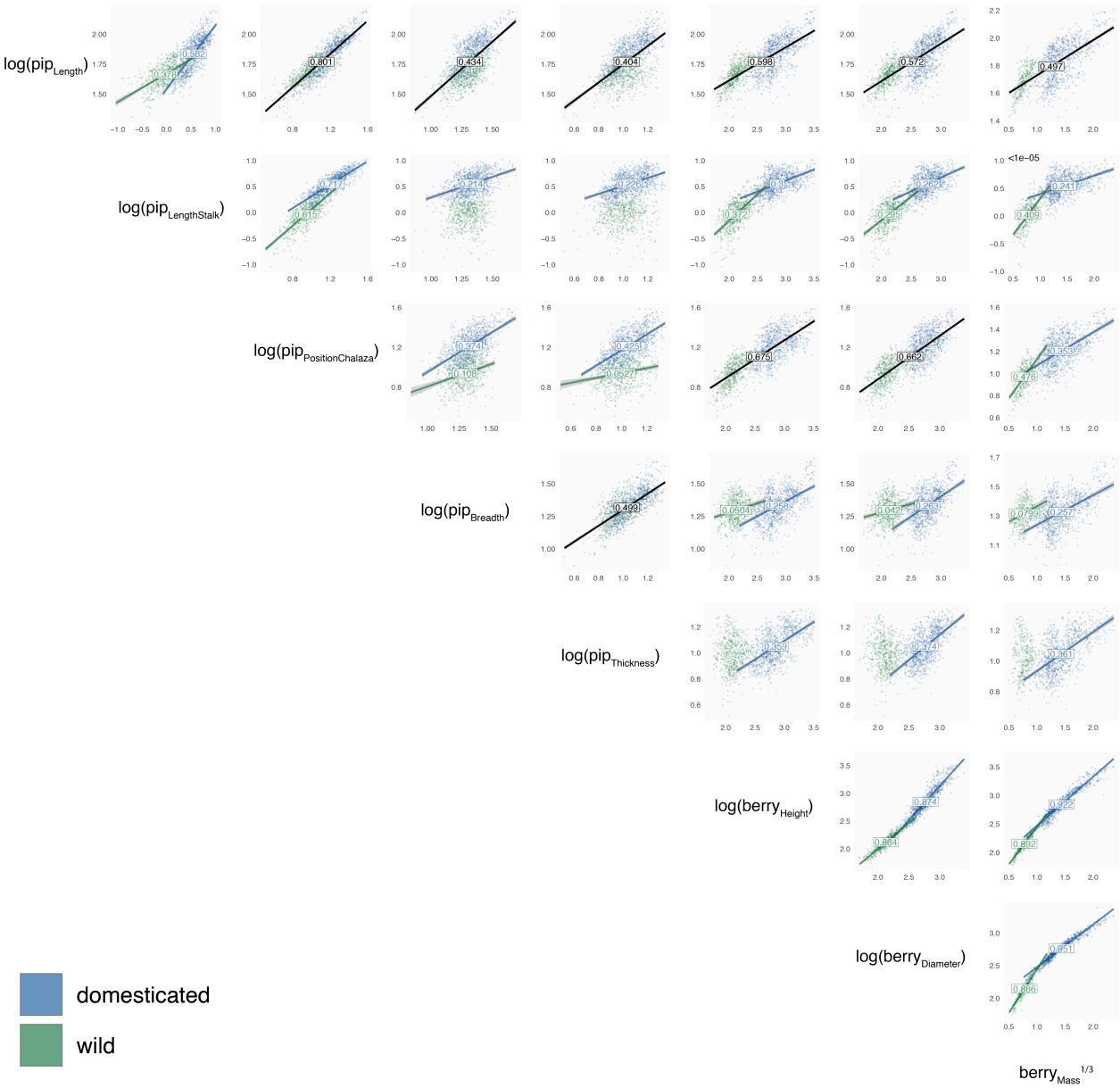
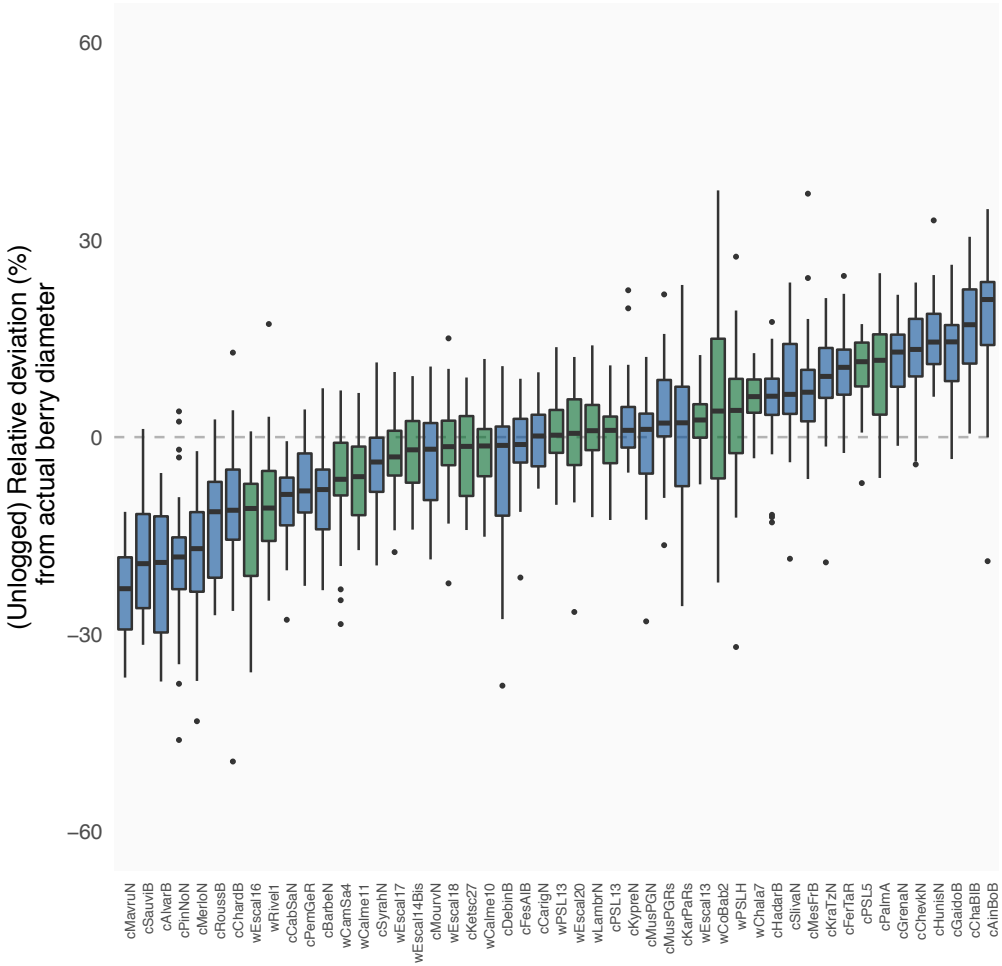
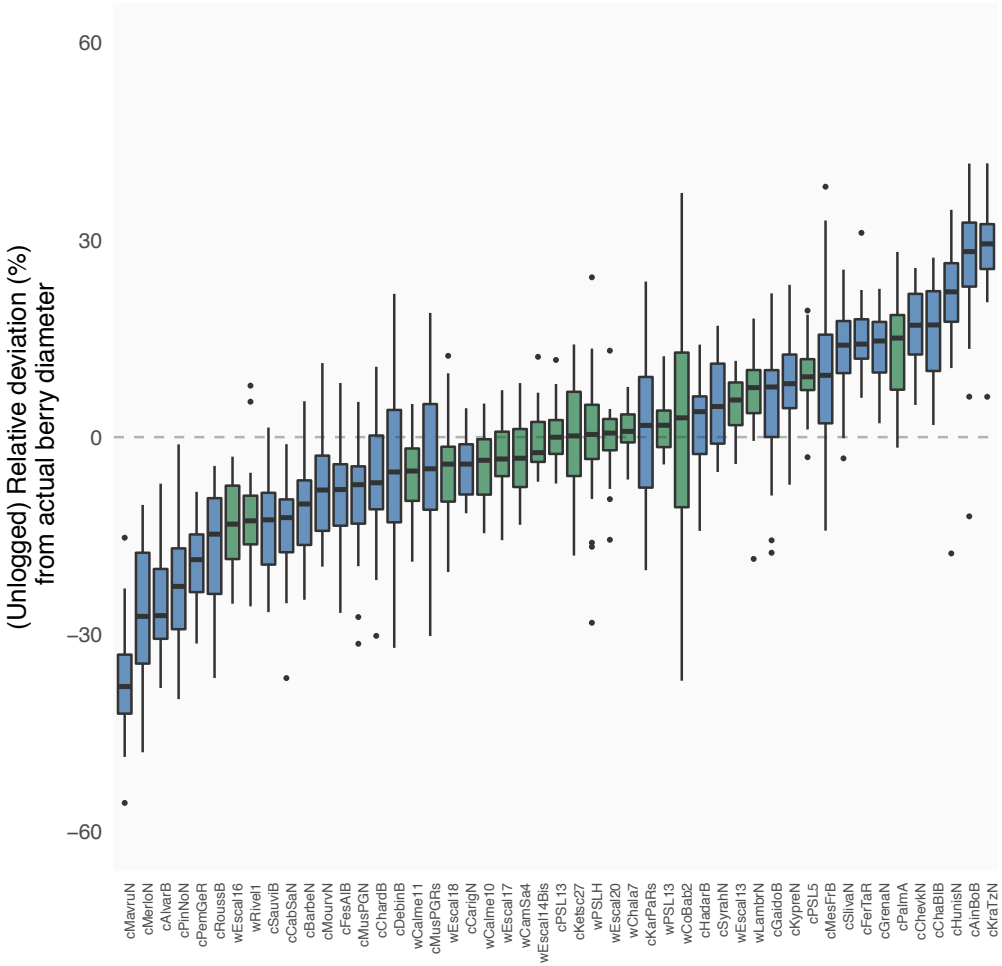

log(pip_Length)
log(pip_LengthStalk)
log(pip_PositionChalaza)
log(pip_Breadth)
log(pip_Thickness)
log(berry_Height)
log(berry_Diameter)
berry_Mass^1/3
<1e−05
domesticated
wild

Figure C


(Unlogged) Relative deviation (%) from actual berry diameter
60
30
0
−30
−60
cMavruN
cSauviB
cAlvarB
cPinNoN
cMerloN
cRoussB
cChardB
wEscal16
wRivel1
cCabSaN
cPemGeR
cBarbeN
wCamSa4
wCalme11
cSyrahN
wEscal17
wEscal14Bis
cMourvN
wEscal18
cKetsc27
wCalme10
cDebinB
cFesAlB
cCarigN
wPSL13
wEscal20
wLambrN
cPSL13
cKypreN
cMusPGN
cMusPGRs
cKarPaRs
wEscal13
wCoBab2
wPSLH
wChala7
cHadarB
cSilvaN
cMesFrB
cKraTzN
cFerTaR
cPSL5
cPalmA
cGrenaN
cChevkN
cHunisN
cGaidoB
cChaBlB
cAinBoB


(Unlogged) Relative deviation (%) from actual berry diameter
60
30
0
−30
−60
cMavruN
cMerloN
cAlvarB
cPinNoN
cPemGeR
cRoussB
wEscal16
wRivel1
cSauviB
cCabSaN
cBarbeN
cMourvN
cFesAlB
cMusPGN
cChardB
cDebinB
wCalme11
cMusPGRs
wEscal18
cCarigN
wCalme10
wEscal17
wCamSa4
wEscal14Bis
cPSL13
cKetsc27
wPSLH
wEscal20
wChala7
cKarPaRs
wPSL13
wCoBab2
cHadarB
cSyrahN
wEscal13
wLambrN
cGaidoB
cKypreN
cPSL5
cMesFrB
cSilvaN
cFerTaR
cGrenaN
cPalmA
cChevkN
cChaBlB
cHunisN
cAinBoB
cKraTzN