Identifying Divergent Selection in Structured Populations of *Daucus carota*

by

Joshua Eun-Song Craig

A thesis submitted in conformity with the requirements for the degree of Master of Science

Graduate Department of Ecology and Evolutionary Biology
University of Toronto

Identifying Divergent Selection in Structured Populations of *Daucus carota*

Joshua Craig

Master of Science

Ecology and Evolutionary Biology
University of Toronto

2022

Abstract

Determining causes of trait divergence between populations is a historically difficult task, as both selection and drift can instigate significant evolutionary change. Qst-Fst has been the most popular method for distinguishing between selection and drift but has numerous limitations making it unideal for most experimental designs. Recently, a new method analogous to Qst-Fst has been developed called Qpc. Qpc seeks to remedy certain limitations of Qst-Fst and provide a more accessible framework for testing trait divergence. In this paper, we apply Qpc to a comparatively small set of *Daucus carota* populations to test its applicability and efficacy, as well as identify drivers of evolution in *D. carota*. We affirm Qpc's ability to identify selection by identifying adaptively divergent traits in *D. carota* and provide comparisons of Qpc to previous methodology.

Author Contributions

Joshua Craig performed all the bioinformatic analyses, analyzed data, wrote the draft, and made all the figures. Shu Han (Julie) Gan, Joshua Craig, and John Stinchcombe designed the common garden experiment; Julie Gan carried out the common garden experiment and collected the phenotypic data, with advice and supervision from Joshua Craig and John Stinchcombe. John Stinchcombe made analysis suggestions and editorial comments on the draft.

Table of Contents

List of Tables

List of Figures

Chapter 1

Identifying Divergent Selection in Structured Populations of *Daucus carota*

1 Introduction

Evolution is predominantly driven by two forces: natural selection (Darwin 1859; Fisher 1930) and genetic drift (Kimura 1983). While both can instigate significant evolutionary change, their relative impacts on molecular and phenotypic evolution is contentious. The conflict between selection and drift is most apparent in geographic differentiation, as phenotypic divergence between geographically distinct populations can be explained by both selection and drift (Escobar-Páramo et al. 2005; Funk et al. 2016). This ambiguity is especially true for quantitative traits, which evolve through small shifts in allele frequencies at many loci and are highly susceptible to drift in early stages of population establishment (Höllinger et al. 2019). Considering the complexity of quantitative trait evolution, reliably distinguishing between the effects of selection and drift on population divergence is a longstanding goal in evolutionary biology.

Historically, finding evidence for neutral phenotypic evolution has been difficult (Lynch 1990). Traditional tests of neutral evolution require estimates of population size, divergence time, and heritability, all of which are difficult to reliably obtain (Lande 1977; Lynch 1990; Spitze 1993). Building on previous work by Wright (1949) and Lande (1992), Spitze (1993) developed a test of neutrality called Qst, which used the Gst (Nei 1973) of neutral allozyme variants as a null hypothesis to compare quantitative trait variance against. The nomenclature, Qst, was chosen to mirror its genetic equivalent Fst, which were shown by Lande (1992) to have equal means under neutrality. This was formalized by Whitlock (1999) into Qst-Fst, wherein Qst > Fst indicates more phenotypic variation than we would expect to occur under drift (divergent selection), Qst = Fst indicates variation equaling what we would expect to occur under drift (neutral evolution), and Qst < Fst indicates less variation than we would expect to occur under drift (stabilizing selection) (Merilä and Crnokrak 2001).

Qst-Fst analysis has since been employed extensively, and in numerous cases has found selection to be the causal force underlying population differentiation. A meta-analysis found that 70% of studies implementing Qst-Fst reported Qst to be greater than the average Fst (per study) [reviewed in (Leinonen et al. 2008)]. However, many studies use microsatellite markers to calculate Fst, which might underestimate neutral genetic variation due to their high mutation rates, biasing results in favour of selection (Edelaar and Björklund 2011; Edelaar et al. 2011). Furthermore, Qst-Fst comparisons have methodological drawbacks that limit its ability to make strong evolutionary inferences. These limitations are the result of various assumptions that natural populations might violate (e.g. traits are additive and autosomal, populations are in drift-migration equilibrium, no maternal effects, etc.) in addition to well known difficulties in estimating additive genetic variance (Whitlock 1999; Hendry 2002; O'Hara and Merilä 2005; Pujol et al. 2008). Because of these limitations, many studies have been restricted to traits in organisms with simplistic genetic architectures making generalizations tenuous (Merilä and Crnokrak 2001; O'Hara and Merilä 2005; Leinonen et al. 2008; Whitlock 2008; Whitlock and Guillaume 2009).

Obtaining accurate Qst values is difficult because it requires measures of additive genetic variance (VA) which typically require pedigree data or specific crossing designs to accurately infer (Pujol et al. 2008; Wilson et al. 2010). To circumvent this requirement, some studies involving natural populations or common garden populations would substitute Pst for Qst, which solely considers phenotypic variance (Brommer 2011). However, Pst is a sub-optimal substitution due to environmental differences in natural populations and non-additive genetic variances in common garden populations, making it an imperfect estimate of Qst (Pujol et al. 2008; Brommer 2011). Although imperfect, Qst remains the foremost strategy.

Overtime, Qst-Fst received various improvements, including a multivariate iteration that provided increased statistical power and accuracy (Martin et al. 2008). More recent iterations incorporate data from Genome-Wide Association Studies (GWAS) to produce improved Qst-Fst analogues. For example, Q_X, developed by Berg and Coop (2014), uses polygenic scores calculated from GWAS in lieu of Qst; which, while practical considering the prevalence of

genomic data, suffers from GWAS's susceptibility to population structure (Berg and Coop 2014).

A solution to the issues accompanying GWAS data and V_A estimation comes from Josephs et al. (2019) who published a new version of Qst-Fst analysis called Qpc. Their methods offer a new strategy for estimating V_A and controls for population structure, remedying drawbacks from earlier iterations of Qst/Fst methodologies (Josephs et al. 2019).

Qpc is an extension of Qst-Fst analysis optimized for genomic and common garden phenotype data. Josephs et al. (2019) use the principal components (PC) of a kinship matrix to test for selection and estimate V_A. This is possible because the set of PCs used for testing selection and the set of PCs used for estimating V_A are orthogonal, and measure between population variation and within population variation respectively. They identify leading PCs describing meaningful axes of relatedness, overlay a vector of trait values, and analyze the slope of the trait against the PC. Josephs et al. (2019) determined that under neutrality, the slope of a trait should be normally distributed, with a mean of 0 and a variance equaling the product of the PC's eigenvalue and additive genetic variance, such that:

$$q_m \sim N(0, \lambda_m V_A)$$

Where q_m is the slope of the trait against PC m, lambda is the eigenvalue of PC m, and V_A is the additive genetic variance estimated from downstream PCs. We consider λ_m because PCs that describe more variation in relatedness should describe more trait variation. Because selection would increase the variance exhibited by q_m, we would expect the variance of a trait under selection to significantly deviate from $\lambda_m V_A$. To test whether a trait's variance is greater than expected under drift, Josephs et al. (2019) use an F test, whereby rejection of the null hypothesis indicates adaptive differentiation, i.e.:

$$q^2_m / \lambda_m V_A \sim F(1, N)$$

Where the degrees of freedom are given by the number of PCs used for testing selection and the number of PCs used for estimating V_A (N).

We used Qpc to investigate the relative roles selection and drift have on quantitative trait differentiation in 33 lines of *Daucus* species. Additionally, we tested how selection and drift impact physiological responses to variable soil conditions (i.e., phenotypically plastic responses) by growing both a control and a fertilizer (treatment) cohort in a common garden experiment, as intense selection might alter crop performance across different environments (genotype by environment interactions) (Gage et al. 2017). The 33 lines of *D. carota* offer the opportunity to apply a Qst-Fst type analysis on complex traits in non-model organisms with the added benefit of being an economically and agriculturally important species. They also offer the opportunity to test the applicability and robustness of Qpc, since Qpc is a relatively new method and could become a crucial tool for future research in evolution.

In applying Qpc, we aimed to address the following research goals: 1) to determine which force more strongly contributes to the evolution of complex traits in *D. carota*: selection or drift, 2) to determine what traits are under selection in wild and domestic populations of *D. carota,* and 3) to test the applicability and effectiveness of Qpc on smaller datasets, as Qpc was initially tested on GWAS sized datasets. Although our study uses comparatively small sample sizes, we predict that Qpc will recover signals of adaptive differentiation between cultivated and wild populations and detect differentiation between wild and eastern lines.

2 Methods

2.1 Study System

Cultivated carrot (*Daucus carota* L. subsp. *sativus*) is an agriculturally and economically important crop. Their wild progenitor is the weedy Queen Anne's lace (*D. carota* subs. *carota*). Carrots are the richest source of vitamin A in the American diet and are one of the most produced vegetables globally (www.fao.org/faostat). In the United States, carrot crop occupies approximately 34000 hectares of land, while wild carrot occurs in temperate climes globally (Iorizzo et al. 2013). Wild carrot is widely distributed and has even been observed growing

among cultivated carrots (Wijnheijmer et al. 1989; Hauser and Bjørn 2001). Wild carrots and cultivated carrots are also sexually compatible, and there is evidence of reciprocal gene flow between them (Mandel et al. 2016).

As a result of domestication, carrots exhibit multiple root colours and morphologies. While wild carrots are white, evidence suggests yellow and purple coloured carrots arose sometime 1,100 years ago, and orange carrots in the 1500s, representing eastern and western cultivars respectively (Banga 1957, 1963). The coloured roots are a product of intense artificial selection and indicate an accumulation of carotenoids which contribute to the carrot's nutritional value (Arscott and Tanumihardjo 2010).

Recently, a high-quality reference genome for *D. carota* was published, alongside the full genome sequences of 35 additional lines (Iorizzo et al. 2016). The paper provides a global sample of *Daucus* genomes and includes both wild and cultivated lines. This genomic dataset provides the necessary diversity to apply Qst-Fst analysis, and constitutes the genomic data required for Qpc.

2.2 Common Garden Design and Phenotyping

We acquired germplasm for each line studied in Iorizzo et al. (2016) from the Germplasm Resources Information Network (GRIN). These included a global sample of cultivated and wild lines, four USDA inbred lines, and several congeners which were used as an outgroup. In total, the paper included 35 lines. Of the 35 lines analyzed by Iorizzo et al. (2016), two were unavailable from GRIN, and therefore not considered in our analyses. Excluding the inbred and outgroup lines, STRUCTURE analysis found four distinct groupings: western wild, western cultivated, eastern wild, and eastern cultivated (Pritchard et al. 2000; Iorizzo et al. 2016). Our analyses focused on these four groups.

We conducted a common garden experiment in the greenhouses at the Earth Sciences Centre of the University of Toronto. We implemented a randomized block design across two cohorts: a control cohort (n = 480) and a fertilizer treatment cohort (n = 480) for a total sample size of n = 960. We planted 20 individuals from each line, with 10 assigned to the control cohort and 10

assigned to the fertilizer cohort. Germplasm was planted in pots with a total volume of 950mL (7.3cm x 7.3cm x 20cm) to allow for sufficient root growth. Greenhouse conditions were set at a 16/8 day/night cycle, with day temperatures between 23-26 C and night temperatures between 17-22 C. For the treatment cohort, we used fertilizer with a nutrient ratio of 1-9-3 NPK to help facilitate root growth. We applied the fertilizer one week after planting at a concentration of $2kg/10m^2$.

We phenotyped the plants after 15 weeks of growth. We recorded the number of dead leaves, live leaves, shoot height, shoot diameter, shoot weight, root length, root diameter, and root weight. For shoot and root weight, we dried the samples and used the dry weight to control for confounding variables associated with wet weight. Lastly, we calculated the shoot to root mass ratio (shoot mass / root mass) to include as an additional trait in our analyses.

To prepare the phenotype data for Qpc, we calculated the mean value of each phenotype per line for both the control and treatment cohorts. To study whether plastic responses to fertilizer differed geographically (west vs. east) or due to cultivation history (wild vs. cultivated) we calculated the difference between the treatment and control means (treatment - control).

2.3 Genomic Data and Qpc

Using the publicly available SNP data from Iorizzo et al. (Iorizzo et al. 2016), we took a random sample of 47,000 SNPs to construct the kinship matrix (K) required for Qpc. We estimated K by acquiring the VCF file from Iorizzo et al. (2016), downsampling the file size, and constructing a SNP matrix with the VCFTOOLS 012 flag, which encodes each biallelic polymorphism as either heterozygous (1) or homozygous (0 or 2), and where the rows are individuals, and the columns are loci (Danecek et al. 2011). We then converted the 012 matrix to allele frequencies and centered it, such that the resulting matrix represented the pairwise genetic correlations of each line (K).

We used the R package *quaint* to calculate Qpc as described in Josephs et al. (2019). We then visualized Qpc results using base R (R Team 2013). Occasionally, Qpc plots and their reported

Qpc p-values seemed discrepant, but this is due to the plots being a visual representation of Qpc only, and not a statistical test itself.

We calculated Qpc with several variations of K. We first used a matrix including all four groups:wild and cultivated, and Western and Eastern groups. We then subset this matrix into four smaller matrices so we could calculate Qpc using more focused comparisons, and better distinguish whether trait differentiation was due to geography of cultivation history (e.g. west vs. east or wild vs. cultivated). For example, by excluding all eastern lines from K, we could calculate Qpc between western wild and western cultivated lines only. Our final array of matrices included wild samples (allowing a western wild vs. eastern wild comparison), cultivated samples (a western vs. eastern cultivated comparison), western samples (a western wild vs. western cultivated comparison), and a eastern samples (an eastern wild vs. eastern cultivated comparison). We tested each combination of phenotype and matrix to look for differentiation.

To determine the number of principal components (PC) appropriate for testing selection, we used a scree plot to visually evaluate the number of leading axes explaining significant amounts of variation, as these are the PCs most likely to display between population variation. For our complete matrix this included the first four PCs. For our subset matrices this was the first PC only. For both our complete and subset matrices, we used the remaining PCs to estimate V_A, as subsequent PCs individually explained little variation and are the most likely to display within population variation.

3 Results

In total, 412 plants survived until harvesting. Most deaths occurred in the treatment cohort, including all individuals from two lines. We kept these lines in our control analysis, but excluded them from our treatment and difference (treatment - control) analysis.

We used PCA biplots to examine the genetic groupings of our populations. The first two PCs from our complete matrix describe distinct axes of relatedness: PC1 distinguishes western wild lines from the remaining lines, and PC2 distinguishes eastern wild from eastern cultivated lines

(Fig. 1). PC3 mainly distinguishes western wild from eastern cultivated lines, and PC4 eastern wild from western cultivated (Fig.1).

Applying Qpc to our complete dataset offers an overview of trait differentiation among *D. carota* of different geographies and cultivation histories. Crucially, there are two main signals Qpc elucidates: that there is adaptive differentiation between the western wild lines and the remaining lines, and between the eastern wild and eastern cultivated lines. The p-values for each Qpc test are presented in Table 1 and Table 2. Table 1 shows results for our complete dataset, and Table 2 shows results for our subset data.

In our complete control dataset, Qpc was significant for root diameter across PC1 ($p = 0.047$), PC2 ($p = 0.009$) and PC3 ($p = 0.031$) (Fig. 2). The Qpc plots suggested that differences in root diameter between western wild and the remaining samples drove differentiation along PC1; differences between eastern cultivated and the remaining samples appeared to reflect differentiation along PC2; differences along PC3 in root diameter appeared to be driven by eastern wild, and a handful of western wild samples, from the remainder. Additionally, root weight was significant across PC2 ($p = 0.034$; again, seemingly driven by differences between eastern cultivated samples), and dead leaf count was nearly significant across PC3 ($p = 0.082$) and PC4 ($p = 0.065$) (Fig. 3). In our complete treatment dataset, Qpc was nearly significant for root diameter across PC1 ($p = 0.056$) and PC3 ($p = 0.051$), and marginally significant for PC2 ($p = 0.098$) (Fig. 4). In each of these comparisons, differences between eastern cultivated and western wild samples seem to produce the nearly significant results. Lastly, live leaf count ($p = 0.030$) and total leaf count ($p = 0.054$) were significant and nearly significant across PC3, respectively (Fig. 5). There were no significant differences in our analyses of the plastic response to fertilizer (i.e., The (treatment - control) dataset).

Because our study used small sample sizes, one of our goals was to test whether Qpc could recapture known signals of selection. To do this, we applied Qpc to wild vs. cultivated matrices. These analyses were intended as a more rigorous test of the characterizations our complete matrix offered about which phenotypic differences appear to be driving significant Qpc results. However, it is important to note that because the samples (and SNPs) included in K change

between these analyses, as well as which PC axes are being used for hypothesis testing, these analyses are not formally comparable to the previous set using the entire dataset. Nonetheless, these subset matrices clarified signals found in our complete matrix.

In our control cultivated lines, root diameter ($p = 0.06$), shoot weight ($p = 0.08$), and root weight ($p = 0.08$) were nearly significant, suggesting that adaptive differentiation occurred between eastern and western cultivated lines (Fig. 6). In our control eastern samples, root diameter ($p = 0.005$) and root weight ($p = 0.048$) were significant while total leaf was nearly significant ($p = 0.08$) (Fig. 7), suggesting non-neutral differentiation between eastern cultivated and eastern wild samples. Furthermore, in our treatment eastern lines, root diameter was significant ($p = 0.01$) and height was nearly significant ($p = 0.053$) (Fig. 8), again suggesting non-neutral differences between cultivated and wild lines. However, Qpc was not significant for any trait in our western lines, meaning Qpc was unable to recapture the artificial selection between western wild and western cultivated lines. While this could be attributed to our small sample size, it is possible that gene flow between wild and cultivated lines decreased genetic and phenotypic differences, as hybridization is known to inadvertently occur between wild and cultivated *Daucus* (Wijnheijmer et al. 1989; Hauser and Bjørn 2001; Mandel et al. 2016). Lastly, Qpc did not detect any signals of selection between wild *D. carota* populations, nor were there any significant hits in our subset difference matrices.

4 Discussion

In this paper, we demonstrated the applicability and efficacy of Qpc on small datasets and tested for divergent selection in structured populations of *D. carota*. We first affirmed that Qpc can detect selection by applying Qpc to matrices comparing wild and cultivated populations. We then tested for divergent selection between populations from different environments and different cultivation histories. Lastly, we identified several advantages and disadvantages of Qpc compared to other iterations of Qst-Fst analysis and provide experimental design suggestions for future studies.

4.1 Divergent Selection in D. carota

In testing Qpc on our wild vs. cultivated matrices, we demonstrated that Qpc is capable of detecting selection in comparatively small sample sizes (e.g. compared to GWAS) with sufficiently strong selection. Overall, our results indicated that *D. carota* exhibits trait divergence between west vs. east lines and wild vs. cultivated lines. Root characteristics exhibited the clearest signs of differentiation, including root diameter and root weight. While this is possibly due to root size being a target of artificial selection, it is possible that root size is correlated with other targets of selection like carotenoid accumulation (Clotault et al. 2012).

Although we were able to recapture signals of artificial selection between the eastern wild and eastern cultivated lines, we were unable to do so for the western wild and western cultivated lines. This is likely attributable to our small sample size. However, studies have observed gene flow between wild and cultivated *D. carota*, with records of wild plants invading agricultural fields and inadvertently leading to hybridization (Wijnheijmer et al. 1989; Hauser and Bjørn 2001; Mandel et al. 2016). This admixture would homogenize differences between wild and cultivated populations making it difficult to detect signals of selection.
In addition to recapturing signals of selection between wild and cultivated lines, our results also identified signals of adaptive differentiation between cultivated lines of *D. carota*. This is primarily supported by our west vs. east cultivated matrix, wherein root diameter, root weight, and shoot weight were nearly significant. Although this suggests that divergent selection may have led to trait differentiation between west and east cultivated carrots, an alternative interpretation is that the observed difference in phenotype is a result of artificial selection on different pools of genetic variation. However, given the minimal phenotypic and genetic differences between western and eastern wild lines, the former interpretation seems more likely.

When looking for an effect of plasticity in our difference dataset, Qpc was not significant for any phenotype, suggesting that all populations responded similarly to fertilizer, and that neither artificial selection nor geography altered how fertilizer impacted phenotype. Our results here are similar to those of Julie Gan's undergraduate thesis, in that she detected evidence of a significant main effect of fertilizer on plant performance traits, but no evidence of G*E. Our results confirm hers, but use SNPs rather than accession designations to characterize genetic variation.

4.2 Advantages and Disadvantages of Qpc

We identified several advantages Qpc has over traditional Qst-Fst methods. First, Qpc is not reliant on external means of estimating V_A, remedying drawbacks in methods that require crosses or family trait measurements to estimate V_A (O'Hara and Merilä 2005). By extension, the data required for Qpc might be more easily obtainable than that required for Qst given the increasing accessibility of high throughput sequencing and SNP detection. Second, as we demonstrated, Qpc may be better suited to smaller data sets than standard Qst-Fst, the latter of which typically requires >20 populations (O'Hara and Merilä 2005). Most Qst-Fst studies focus on pairwise comparisons among a small number of structured populations which might not constitute sufficient scope to appropriately apply Qst-Fst. While there is a Qst-Fst method designed to improve statistical power for fewer populations (i.e. QstFstComp), the standard method is more common and intuitive (Ovaskainen et al. 2011; Li et al. 2019). In many cases Qpc might be more appropriate.

However, Qpc exhibits many of the same limitations as other Qst methods. Qpc relies on the same stringent assumption that traits are purely additive, as non-additive genetic variance may result in false positives. Additionally, any environmental variation decreases Qpcs power to detect divergent selection, as environmental variance contributes variance to later PCs, skewing measures of V_A (Josephs et al. 2019). This generally confines experimental designs to common garden experiments. As such, we attempted to minimize the role of environmental variation in phenotypic traits by using a replicated, blocked common garden design, and analyzing the means of accessions, which should better reflect that genotype's "true" phenotype.

Aspects of our own design could have been changed to improve Qpc's accuracy. Our foremost drawback is that our phenotype data and genotype data were not from the same individual, but rather from siblings. Determining what GRIN considers to be an "accession" of *D. carota* is challenging. After a few email exchanges, it appears that wild accessions are collections of samples from a given locality, with outcrossing within them, in part to maintain variation within samples from that locality. The samples we obtained from GRIN, however, were full siblings (from the same seed lot) of the samples sequenced by Iorizzo et al. (2016), from which we obtained our SNP data. Using different individuals for each data type decreases Qpc's accuracy,

so ensuring that phenotype and genotype data come from the same individual is ideal. However, given that our use of siblings would make it more difficult to detect divergent selection, we can say that the signals we did detect are robust. This reasoning is similarly true when considering our small sample sizes.

Uniquely, Qpc is unable to detect stabilizing selection. While Qst-Fst is able to detect stabilizing selection when Qst < Fst, Qpc solely considers the upper tail of the F distribution, meaning traits under stabilizing selection may be indistinguishable from drift when interpreted with Qpc.

5 Conclusion

When implementing Qpc, many of the same caveats carry over from Qst-Fst. Ideally, experiments should incorporate a common garden design, with phenotypic data and genotypic data being taken from the same individual. Minimizing environmental effects and adhering to the assumptions that traits are additive are also paramount. Given these similarities, Qpc might still be preferable if the study is interested in drawing explicit comparisons between a low number of populations, or pairs of populations. Additionally, Qpc might be ideal if trait data and SNP data are readily obtainable, forgoing the need for extraneous means of estimating V_A.

Qst-Fst is a popular method, and studies seeking to disentangle the effects of selection and drift are numerous. Although new, Qpc is a promising iteration of Qst-Fst, albeit relatively untested. Future studies might incorporate Qpc to further test its rigour, as it's intuitiveness and power make it a strong option when choosing between methods.

Our Qpc results were able to recapture artificial selection between wild and cultivated *Daucus*, as well as identify signals suggesting adaptive differentiation between western and eastern cultivars. Although there was no evidence to suggest that wild samples experienced differential selection per location, we identified adaptive divergence among cultivated samples that suggest non-neutral evolution after cultivation. While there are alternative explanations as to why traits among eastern and western cultivars might exhibit excess trait variance, further research into specific selection mechanisms and genomic responses to selection might elucidate them.

Tables

Table 1. Table of Qpc qm-values and p-values from the complete matrix. P-values < 0.1 are highlighted in yellow. P-values < 0.05 are highlighted in red. Qm values are the ratio of variances used for the F test. Following the qm value in brackets are the degrees of freedom for the numerator and denominator respectively.

Phenotype	control PC1	control PC2	control PC3	control PC4
Root Diameter	qm(1, 19)=4.47 P = 0.05	qm(1, 19)=8.31 P = 0.01	qm(1, 19)=5.39 P = 0.03	qm(1, 19)=0 P = 0.64
Germination Days	qm(1, 19)=0.25 P = 0.62	qm(1, 19)=0.26 P = 0.61	qm(1, 19)=0.16 P = 0.7	qm(1, 19)=0 P = 0.87
Dead Leaf	qm(1, 19)=1.22 P = 0.28	qm(1, 19)=0 P = 0.98	qm(1, 19)=3.35 P = 0.08	qm(1, 19)=0 P = 0.07
Live Leaf	qm(1, 19)=1.36 P = 0.26	qm(1, 19)=0.38 P = 0.55	qm(1, 19)=0.77 P = 0.39	qm(1, 19)=0 P = 0.62
Total Leaf	qm(1, 19)=0.22 P = 0.65	qm(1, 19)=0.28 P = 0.61	qm(1, 19)=2.47 P = 0.13	qm(1, 19)=0 P = 0.63
Height	qm(1, 19)=0.01 P = 0.92	qm(1, 19)=0.89 P = 0.36	qm(1, 19)=0.11 P = 0.75	qm(1, 19)=0 P = 0.84
Length	qm(1, 19)=1.95 P = 0.18	qm(1, 19)=0.06 P = 0.8	qm(1, 19)=2.73 P = 0.11	qm(1, 19)=0 P = 0.84
Shoot Weight	qm(1, 19)=0 P = 0.99	qm(1, 19)=2.31 P = 0.14	qm(1, 19)=0.8 P = 0.38	qm(1, 19)=0 P = 0.55
Root Weight	qm(1, 19)=0.75 P = 0.4	qm(1, 19)=5.21 P = 0.03	qm(1, 19)=0.93 P = 0.35	qm(1, 19)=0 P = 0.65

Phenotype	treatment PC1	treatment PC2	treatment PC3	treatment PC4
Root Diameter	qm(1, 17)=4.19 P = 0.06	qm(1, 17)=3.07 P = 0.1	qm(1, 17)=4.39 P = 0.05	qm(1, 17)=0.03 P = 0.87
Germination Days	qm(1, 17)=0.42 P = 0.53	qm(1, 17)=0 P = 0.96	qm(1, 17)=1.12 P = 0.3	qm(1, 17)=0.02 P = 0.88
Dead Leaf	qm(1, 17)=0.02 P = 0.88	qm(1, 17)=0 P = 1	qm(1, 17)=0.03 P = 0.86	qm(1, 17)=1.68 P = 0.21
Live Leaf	qm(1, 17)=0.54 P = 0.47	qm(1, 17)=0.34 P = 0.57	qm(1, 17)=5.57 P = 0.03	qm(1, 17)=0.79 P = 0.39
Total Leaf	qm(1, 17)=0.53 P = 0.48	qm(1, 17)=0.23 P = 0.64	qm(1, 17)=4.3 P = 0.05	qm(1, 17)=0.08 P = 0.78
Height	qm(1, 17)=0.46 P = 0.51	qm(1, 17)=0.05 P = 0.82	qm(1, 17)=0.53 P = 0.48	qm(1, 17)=0 P = 0.99
Length	qm(1, 17)=1.23 P = 0.28	qm(1, 17)=0.01 P = 0.93	qm(1, 17)=0.08 P = 0.78	qm(1, 17)=0.01 P = 0.91
Shoot Weight	qm(1, 17)=0.24 P = 0.63	qm(1, 17)=0.04 P = 0.84	qm(1, 17)=0.11 P = 0.75	qm(1, 17)=0.19 P = 0.67
Root Weight	qm(1, 17)=1.21 P = 0.29	qm(1, 17)=2.33 P = 0.15	qm(1, 17)=1.14 P = 0.3	qm(1, 17)=0.03 P = 0.85

Phenotype	difference PC1	difference PC2	difference PC3	difference PC4
Root Diameter	qm(1, 17)=0.3 P = 0.59	qm(1, 17)=0.25 P = 0.62	qm(1, 17)=0.19 P = 0.66	qm(1, 17)=0.05 P = 0.83
Germination Days	qm(1, 17)=0.01 P = 0.92	qm(1, 17)=0.8 P = 0.38	qm(1, 17)=0.12 P = 0.74	qm(1, 17)=0.02 P = 0.89
Dead Leaf	qm(1, 17)=0.69 P = 0.42	qm(1, 17)=0 P = 0.96	qm(1, 17)=0.56 P = 0.47	qm(1, 17)=0.04 P = 0.84
Live Leaf	qm(1, 17)=0.08 P = 0.78	qm(1, 17)=0.03 P = 0.88	qm(1, 17)=1.59 P = 0.22	qm(1, 17)=0.05 P = 0.83
Total Leaf	qm(1, 17)=0.06 P = 0.8	qm(1, 17)=0.04 P = 0.84	qm(1, 17)=0.7 P = 0.41	qm(1, 17)=0.01 P = 0.92
Height	qm(1, 17)=0.47 P = 0.5	qm(1, 17)=0.01 P = 0.92	qm(1, 17)=0.34 P = 0.57	qm(1, 17)=0 P = 0.98
Length	qm(1, 17)=0.16 P = 0.69	qm(1, 17)=0.04 P = 0.85	qm(1, 17)=0.55 P = 0.47	qm(1, 17)=0 P = 0.95
Shoot Weight	qm(1, 17)=0.24 P = 0.63	qm(1, 17)=0.05 P = 0.83	qm(1, 17)=0 P = 0.95	qm(1, 17)=0.06 P = 0.8
Root Weight	qm(1, 17)=0.12 P = 0.73	qm(1, 17)=0.13 P = 0.72	qm(1, 17)=0.05 P = 0.83	qm(1, 17)=0.01 P = 0.93

Table 2. Table of Qpc qm-values and p-values from the subset matrices. P-values < 0.1 are highlighted in yellow. P-values < 0.05 are highlighted in red. Qm values are the ratio of variances used for the F test. Following the qm value in brackets are the degrees of freedom for the numerator and denominator respectively.

Phenotype	control wild	control cultivated	control east	control west
Root Diameter	qm(1, 9)=0.37 P = 0.56	qm(1, 11)=4.24 P = 0.07	qm(1, 10)=12.57 P = 0.01	qm(1, 10)=2.87 P = 0.12
Germination Days	qm(1, 9)=0.08 P = 0.78	qm(1, 11)=0.11 P = 0.75	qm(1, 10)=0.51 P = 0.49	qm(1, 10)=0.11 P = 0.74
Dead Leaf	qm(1, 9)=3.36 P = 0.1	qm(1, 11)=1.06 P = 0.33	qm(1, 10)=1.69 P = 0.22	qm(1, 10)=0.35 P = 0.57
Live Leaf	qm(1, 9)=0.28 P = 0.61	qm(1, 11)=1.31 P = 0.28	qm(1, 10)=1.34 P = 0.27	qm(1, 10)=0.59 P = 0.46
Total Leaf	qm(1, 9)=0.04 P = 0.84	qm(1, 11)=0.02 P = 0.9	qm(1, 10)=3.55 P = 0.09	qm(1, 10)=0.14 P = 0.71
Height	qm(1, 9)=0.03 P = 0.86	qm(1, 11)=2.73 P = 0.13	qm(1, 10)=2.21 P = 0.17	qm(1, 10)=0 P = 1
Length	qm(1, 9)=0.24 P = 0.64	qm(1, 11)=0.4 P = 0.54	qm(1, 10)=1.84 P = 0.21	qm(1, 10)=1.54 P = 0.25
Shoot Weight	qm(1, 9)=0.02 P = 0.9	qm(1, 11)=3.73 P = 0.08	qm(1, 10)=2.59 P = 0.14	qm(1, 10)=0 P = 0.98
Root Weight	qm(1, 9)=0.01 P = 0.93	qm(1, 11)=3.69 P = 0.08	qm(1, 10)=5.04 P = 0.05	qm(1, 10)=0.35 P = 0.57

Phenotype	treatment wild	treatment cultivated	treatment east	treatment west
Root Diameter	qm(1, 8)=0.87 P = 0.38	qm(1, 10)=0.56 P = 0.47	qm(1, 9)=8.68 P = 0.02	qm(1, 9)=1.95 P = 0.2
Germination Days	qm(1, 8)=0.28 P = 0.61	qm(1, 10)=0.04 P = 0.85	qm(1, 9)=0.01 P = 0.94	qm(1, 9)=0.41 P = 0.54
Dead Leaf	qm(1, 8)=0.01 P = 0.91	qm(1, 10)=0.59 P = 0.46	qm(1, 9)=1.47 P = 0.26	qm(1, 9)=0.08 P = 0.79
Live Leaf	qm(1, 8)=0.03 P = 0.88	qm(1, 10)=0.34 P = 0.57	qm(1, 9)=3.06 P = 0.11	qm(1, 9)=0.11 P = 0.74
Total Leaf	qm(1, 8)=0.08 P = 0.79	qm(1, 10)=0 P = 0.99	qm(1, 9)=2.67 P = 0.14	qm(1, 9)=0.36 P = 0.56
Height	qm(1, 8)=0.37 P = 0.56	qm(1, 10)=1.32 P = 0.28	qm(1, 9)=4.92 P = 0.05	qm(1, 9)=0.25 P = 0.63
Length	qm(1, 8)=0.61 P = 0.46	qm(1, 10)=0.23 P = 0.65	qm(1, 9)=0.15 P = 0.7	qm(1, 9)=1.05 P = 0.33
Shoot Weight	qm(1, 8)=0.08 P = 0.79	qm(1, 10)=0.08 P = 0.78	qm(1, 9)=0.02 P = 0.89	qm(1, 9)=0.11 P = 0.75
Root Weight	qm(1, 8)=0.11 P = 0.75	qm(1, 10)=1.13 P = 0.31	qm(1, 9)=2.32 P = 0.16	qm(1, 9)=0.44 P = 0.52

Phenotype	difference wild	difference cultivated	difference east	difference west
Root Diameter	qm(1, 8)=0.2 P = 0.67	qm(1, 10)=0.42 P = 0.53	qm(1, 9)=0.23 P = 0.64	qm(1, 9)=0.22 P = 0.65
Germination Days	qm(1, 8)=0 P = 0.95	qm(1, 10)=1.15 P = 0.31	qm(1, 9)=1.28 P = 0.29	qm(1, 9)=0 P = 0.98
Dead Leaf	qm(1, 8)=0.58 P = 0.47	qm(1, 10)=0.03 P = 0.87	qm(1, 9)=0 P = 0.95	qm(1, 9)=0.38 P = 0.55
Live Leaf	qm(1, 8)=0.34 P = 0.57	qm(1, 10)=0 P = 0.98	qm(1, 9)=2.54 P = 0.15	qm(1, 9)=0.04 P = 0.85
Total Leaf	qm(1, 8)=0.02 P = 0.9	qm(1, 10)=0.01 P = 0.92	qm(1, 9)=0.77 P = 0.4	qm(1, 9)=0.04 P = 0.84
Height	qm(1, 8)=0.26 P = 0.62	qm(1, 10)=0.11 P = 0.75	qm(1, 9)=0.05 P = 0.84	qm(1, 9)=0.22 P = 0.65
Length	qm(1, 8)=0.21 P = 0.66	qm(1, 10)=0.13 P = 0.73	qm(1, 9)=0.07 P = 0.8	qm(1, 9)=0.2 P = 0.66
Shoot Weight	qm(1, 8)=0.05 P = 0.82	qm(1, 10)=0.27 P = 0.61	qm(1, 9)=0.46 P = 0.51	qm(1, 9)=0.11 P = 0.75
Root Weight	qm(1, 8)=0.12 P = 0.74	qm(1, 10)=0.23 P = 0.64	qm(1, 9)=0.44 P = 0.52	qm(1, 9)=0.09 P = 0.77

Figures

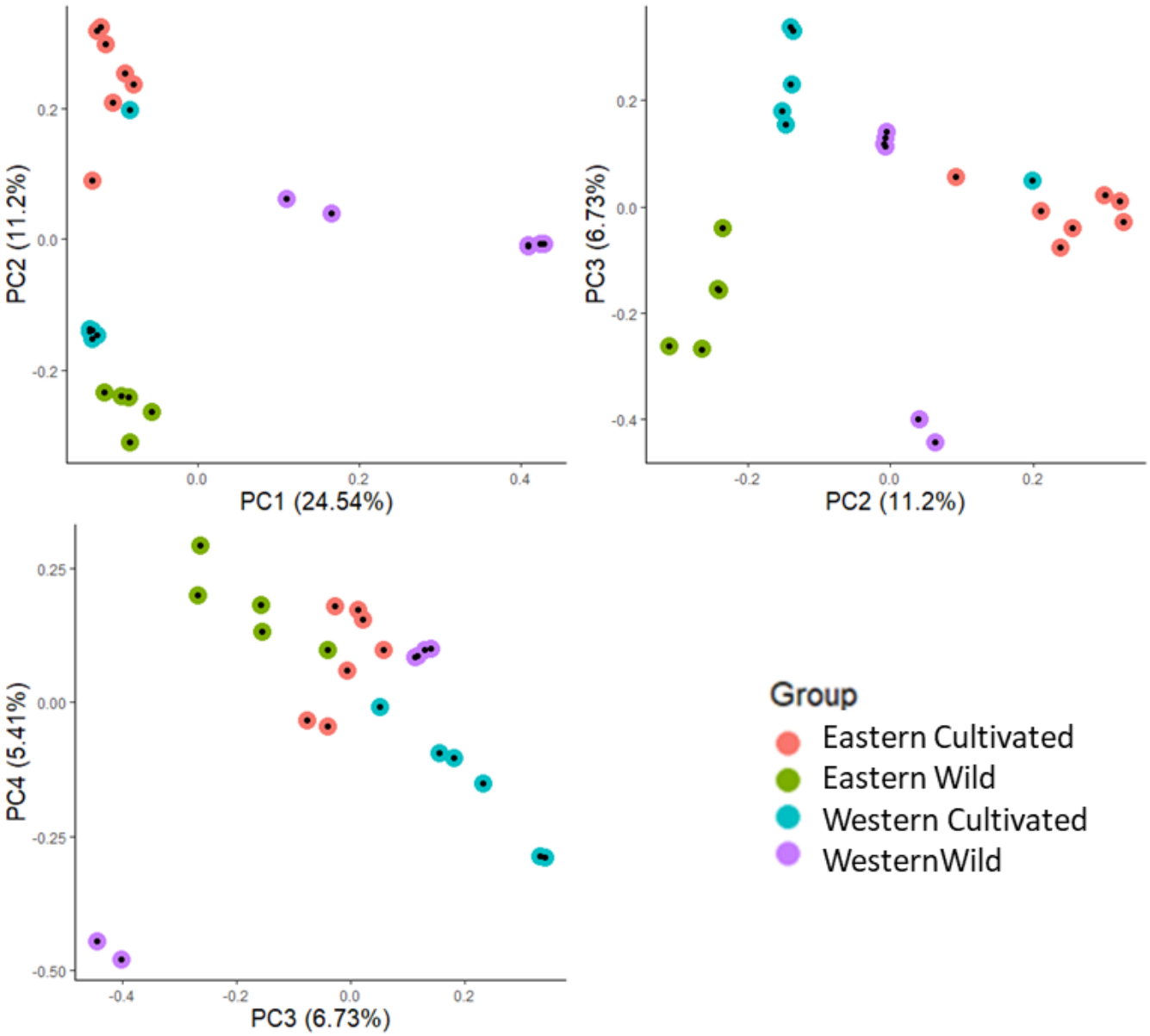

Figure 1. PCA biplots of all four *Daucus* populations used in our analyses. PCs 1-4 were used for testing selection. Subsequent PCs were used for estimating additive genetic variation.

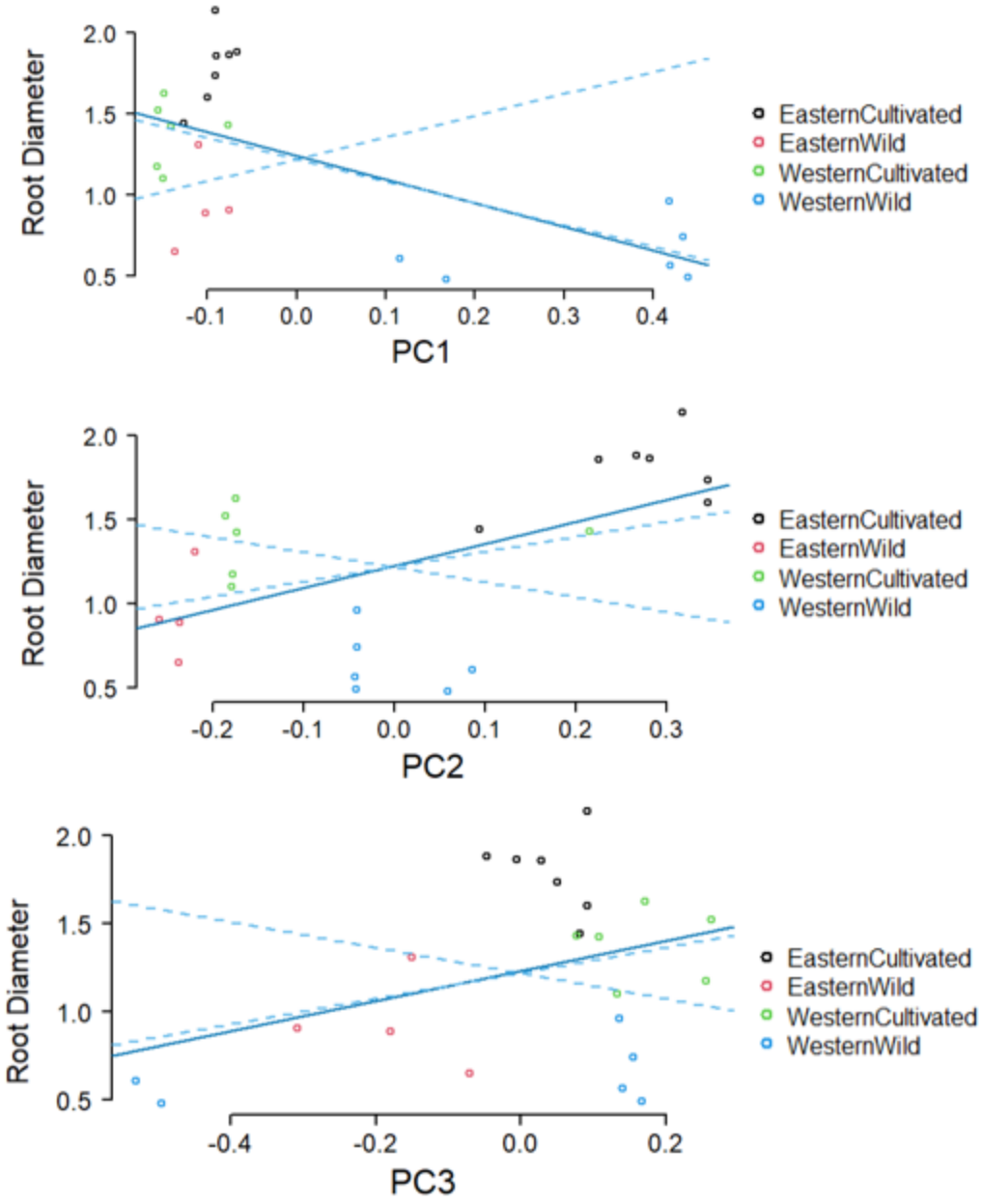

Figure 2. Qpc for root diameter in the complete control group. Root diameter was significant across PC1, PC2, and PC3. The dotted lines indicate a 95% confidence interval around the neutral expectation of the trait. The dotted lines are purely for visualization purposes and are unrelated to the QPC statistical test.

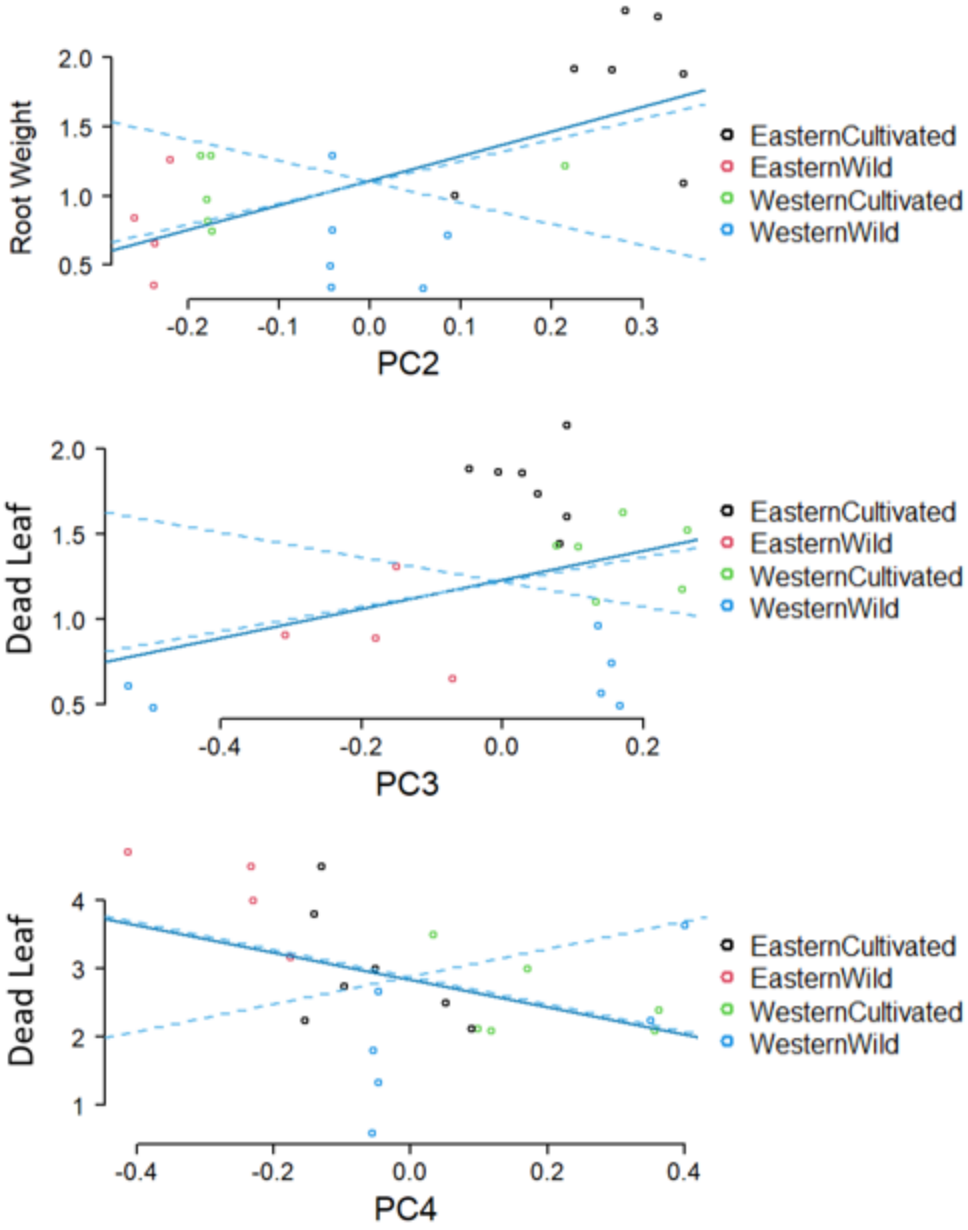

Figure 3. Qpc for root weight and dead leaf count in the complete control group. Root weight was significant across PC2, and dead leaf was nearly significant across PC3 and PC4. The dotted lines indicate a 95% confidence interval around the neutral expectation of the trait. The dotted lines are purely for visualization purposes and are unrelated to the QPC statistical test

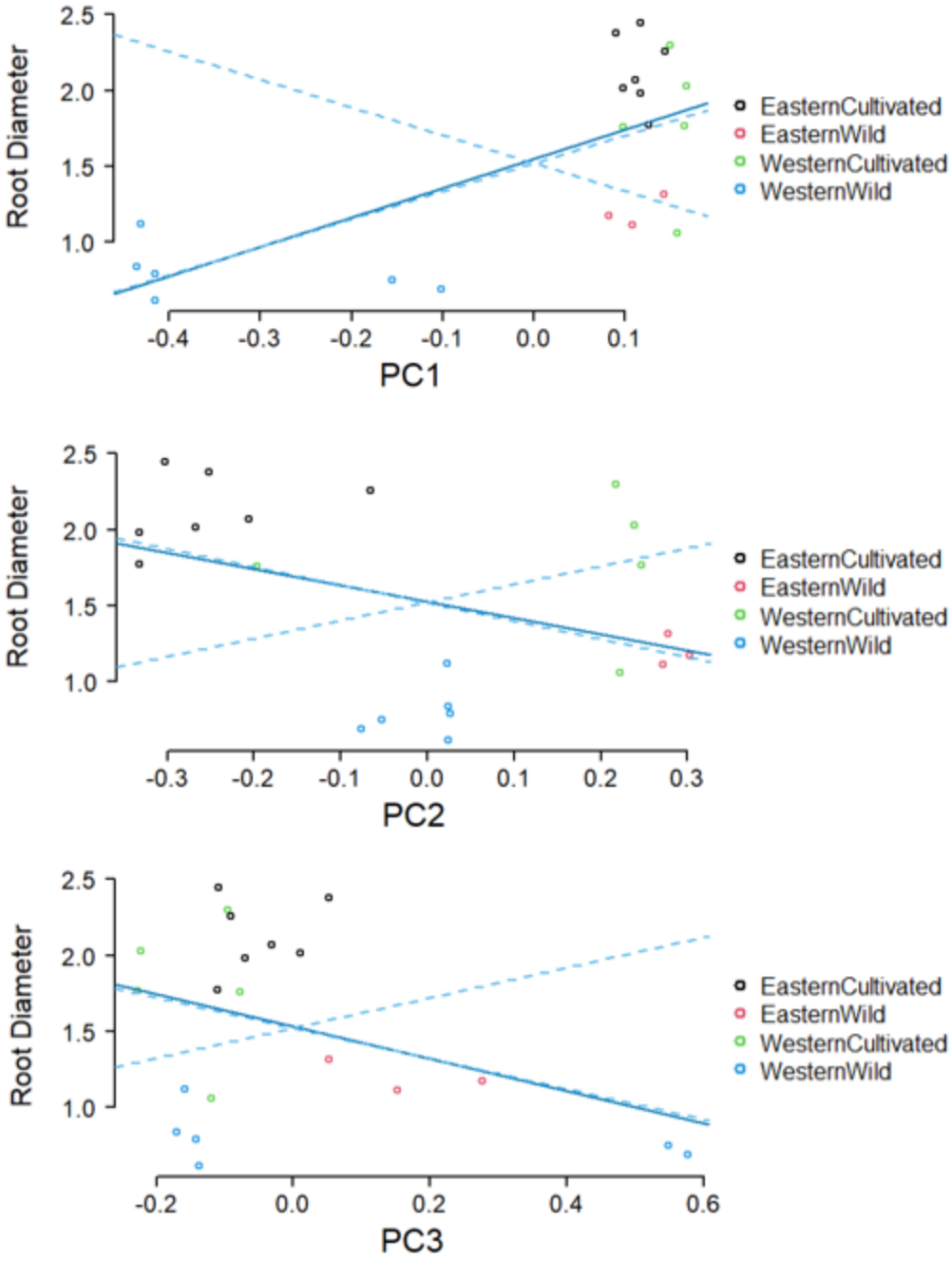

Figure 4. Qpc for root diameter in the complete treatment cohort. Root diameter was nearly significant across PC1, PC2, and PC3. The dotted lines indicate a 95% confidence interval around the neutral expectation of the trait. The dotted lines are purely for visualization purposes and are unrelated to the QPC statistical test

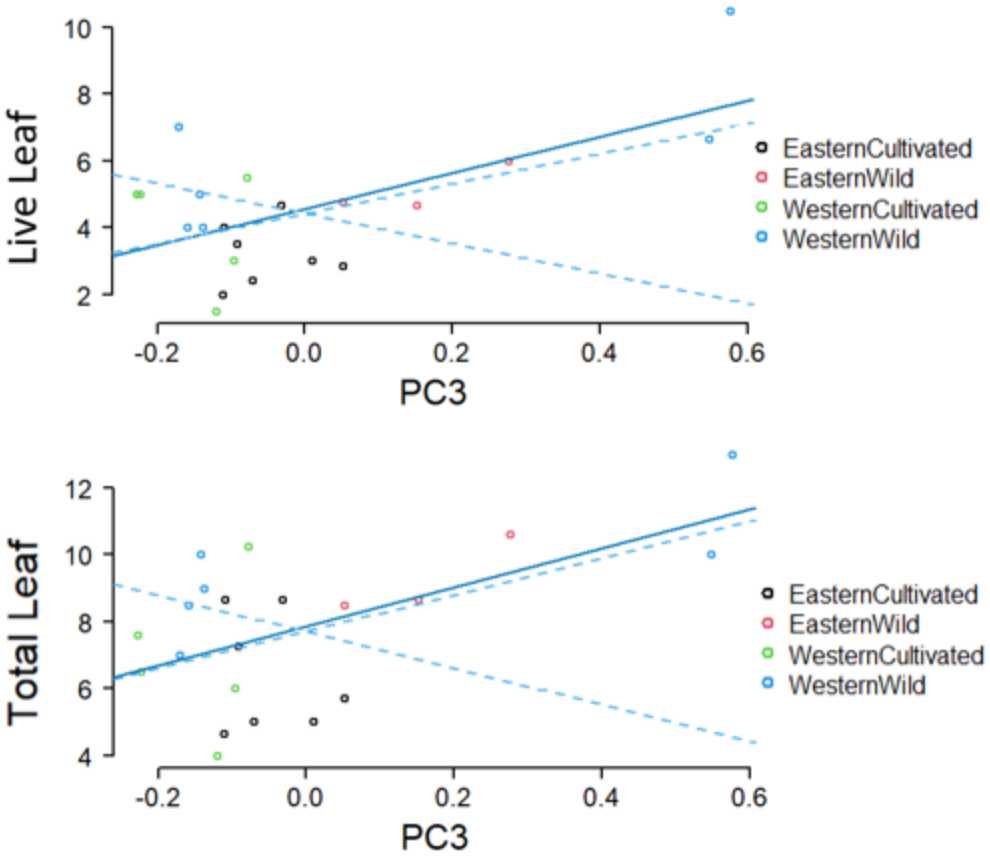

Figure 5. Qpc for live leaf and total leaf count in the complete treatment cohort. Qpc was significant for live leaf count, and nearly significant for total leaf count across PC3. The dotted lines indicate a 95% confidence interval around the neutral expectation of the trait. The dotted lines are purely for visualization purposes and are unrelated to the QPC statistical test

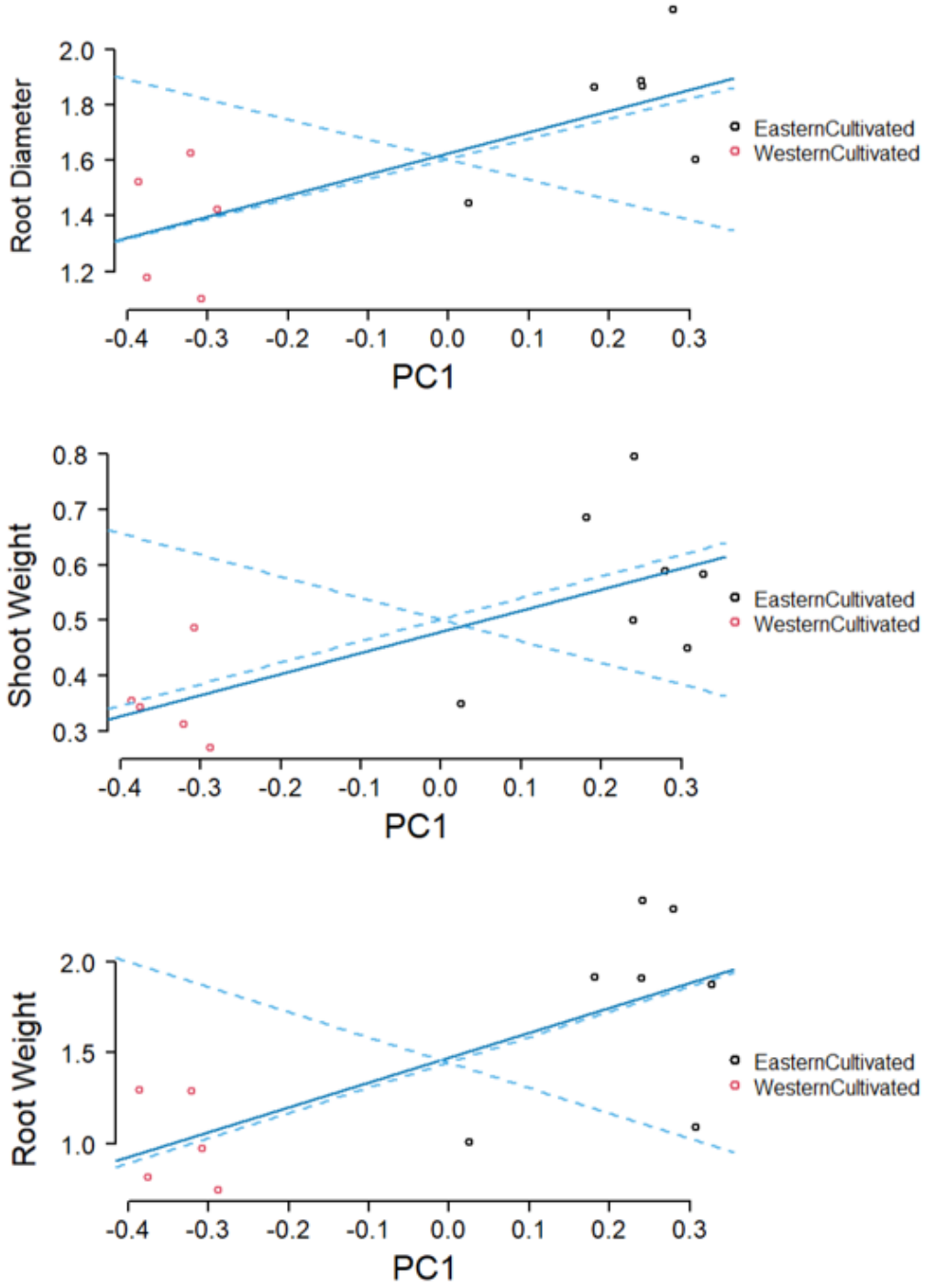

Figure 6. Qpc for cultivated control lines only. Qpc was nearly significant for root diameter, shoot weight, and root weight among cultivated control lines. The dotted lines indicate a 95% confidence interval around the neutral expectation of the trait. The dotted lines are purely for visualization purposes and are unrelated to the QPC statistical test

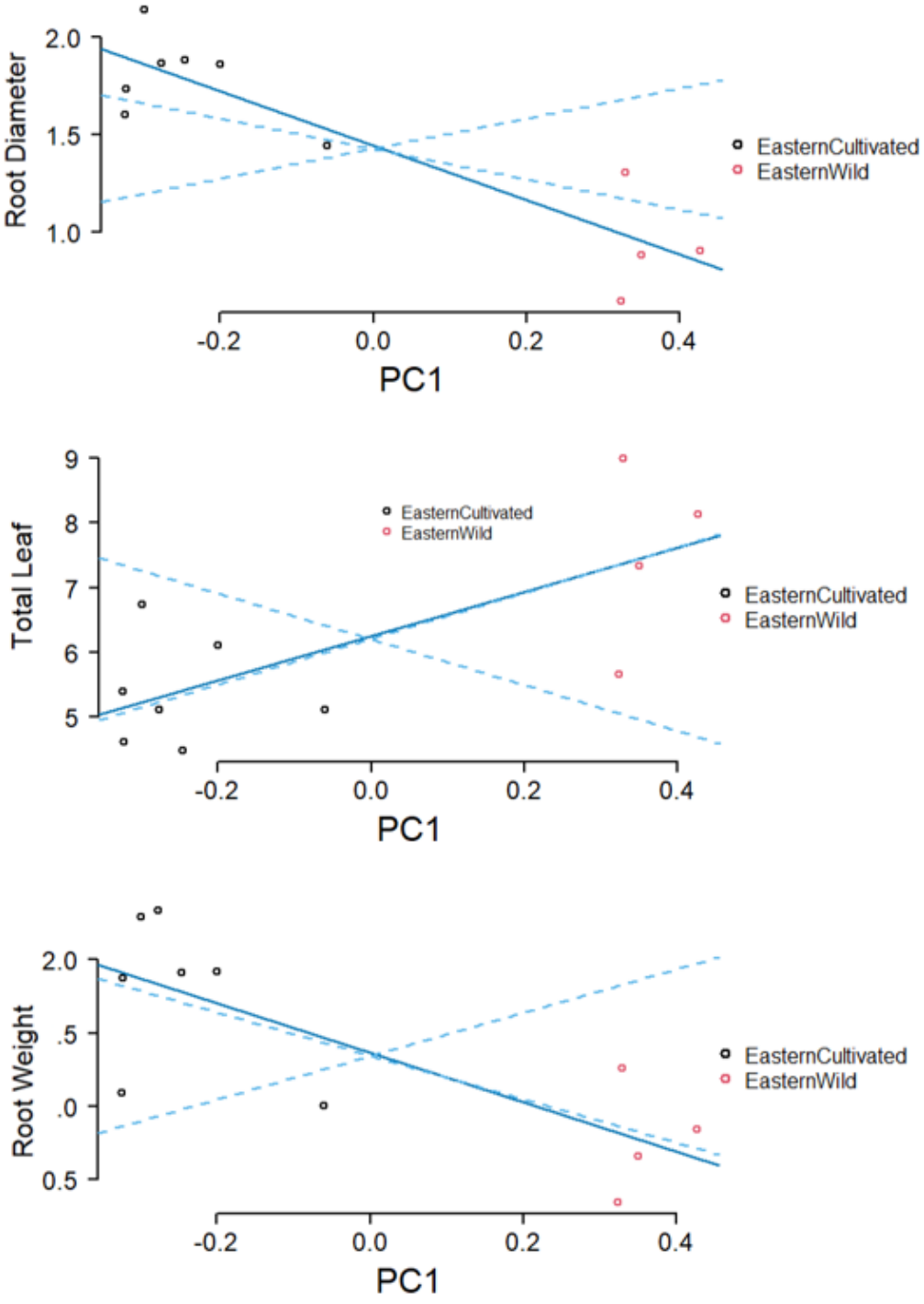

Figure 7. Qpc for eastern control lines only. Qpc was significant for root diameter and root weight, and nearly significant for total leaf count among eastern control lines. The dotted lines indicate a 95% confidence interval around the neutral expectation of the trait. The dotted lines are purely for visualization purposes and are unrelated to the QPC statistical test

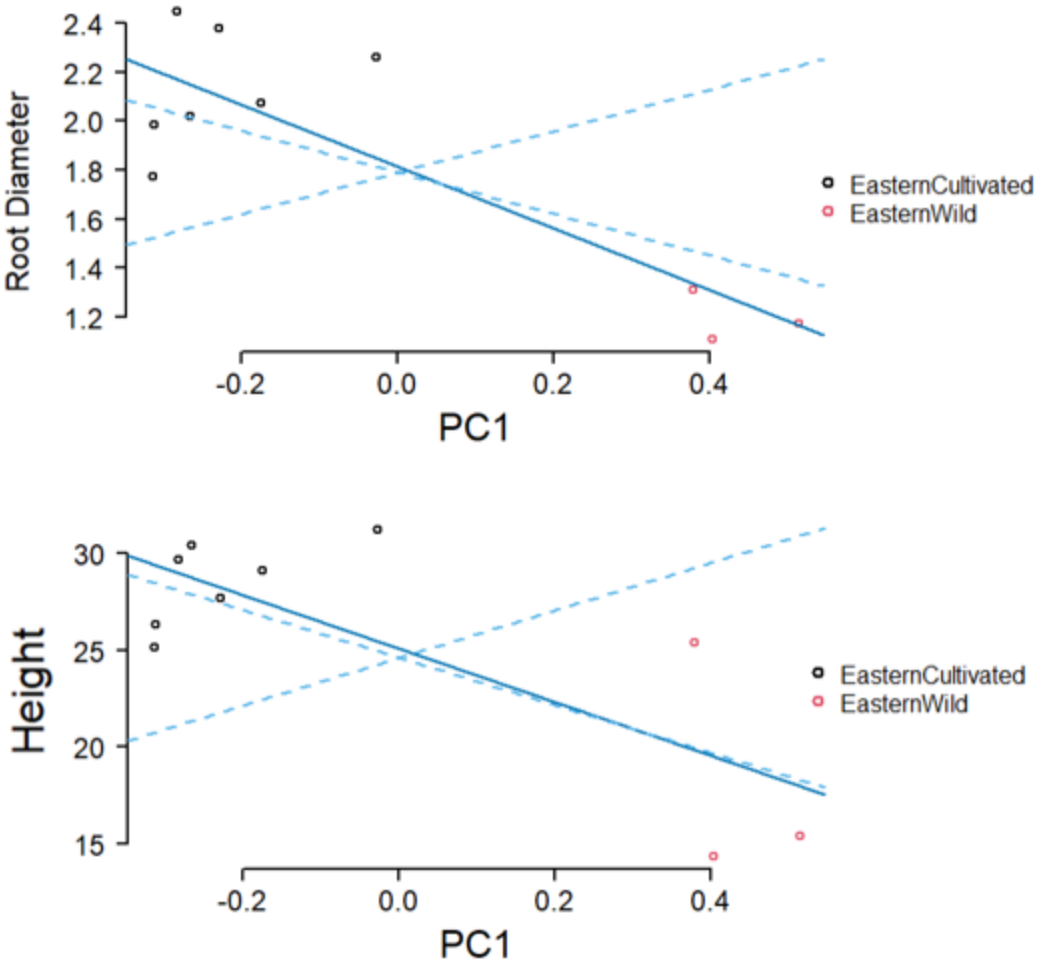

Figure 8. Qpc for eastern treatment lines only. Qpc was significant for root diameter and nearly significant for height among eastern treatment lines. The dotted lines indicate a 95% confidence interval around the neutral expectation of the trait. The dotted lines are purely for visualization purposes and are unrelated to the QPC statistical test

References

Arscott, S. A., and S. A. Tanumihardjo. 2010. Carrots of many colors provide basic nutrition and bioavailable phytochemicals acting as a functional food. Compr. Rev. Food Sci. Food Saf. 9:223–239. Wiley.

Banga, O. 1963. Origin and distribution of the western cultivated carrot. Genetica Agraria 17:357–370.

Banga, O. 1957. Origin of the European cultivated carrot. Euphytica 6:54–63. Springer Science and Business Media LLC.

Berg, J. J., and G. Coop. 2014. A population genetic signal of polygenic adaptation. PLoS Genet. 10:e1004412.

Brommer, J. E. 2011. Whither Pst? The approximation of Qst by Pst in evolutionary and conservation biology. J. Evol. Biol. 24:1160–1168.

Clotault, J., D. Peltier, V. Soufflet-Freslon, M. Briard, and E. Geoffriau. 2012. Differential selection on carotenoid biosynthesis genes as a function of gene position in the metabolic pathway: a study on the carrot and dicots. PLoS One 7:e38724.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Darwin, C. 1859. On the Origin of Species.

Edelaar, P., and M. Björklund. 2011. If F(ST) does not measure neutral genetic differentiation, then comparing it with Q(ST) is misleading. Or is it? Mol. Ecol. 20:1805–1812.

Edelaar, P., P. Burraco, and I. Gomez-Mestre. 2011. Comparisons between Q(ST) and F(ST) -- how wrong have we been? Mol. Ecol. 20:4830–4839.

Escobar-Páramo, P., S. Ghosh, and J. DiRuggiero. 2005. Evidence for genetic drift in the diversification of a geographically isolated population of the hyperthermophilic archaeon Pyrococcus. Mol. Biol. Evol. 22:2297–2303.

Fisher, R. 1930. The Genetical Theory of Natural Selection.

Funk, W. C., R. E. Lovich, P. A. Hohenlohe, C. A. Hofman, S. A. Morrison, T. S. Sillett, C. K.

Ghalambor, J. E. Maldonado, T. C. Rick, M. D. Day, N. R. Polato, S. W. Fitzpatrick, T. J. Coonan, K. R. Crooks, A. Dillon, D. K. Garcelon, J. L. King, C. L. Boser, N. Gould, and W. F. Andelt. 2016. Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). Mol. Ecol. 25:2176–2194.

Gage, J. L., D. Jarquin, C. Romay, A. Lorenz, E. S. Buckler, S. Kaeppler, N. Alkhalifah, M. Bohn, D. A. Campbell, J. Edwards, D. Ertl, S. Flint-Garcia, J. Gardiner, B. Good, C. N. Hirsch, J. Holland, D. C. Hooker, J. Knoll, J. Kolkman, G. Kruger, N. Lauter, C. J. Lawrence-Dill, E. Lee, J. Lynch, S. C. Murray, R. Nelson, J. Petzoldt, T. Rocheford, J. Schnable, P. S. Schnable, B. Scully, M. Smith, N. M. Springer, S. Srinivasan, R. Walton, T. Weldekidan, R. J. Wisser, W. Xu, J. Yu, and N. de Leon. 2017. The effect of artificial selection on phenotypic plasticity in maize. Nat. Commun. 8:1348.

Hauser, T. P., and G. K. Bjørn. 2001. Genet. Resour. Crop Evol. 48:499–506. Springer Nature.

Hendry, A. P. 2002. QST > = ≠ < FST?

Höllinger, I., P. S. Pennings, and J. Hermisson. 2019. Polygenic adaptation: From sweeps to subtle frequency shifts. PLoS Genet. 15:e1008035.

Iorizzo, M., S. Ellison, D. Senalik, P. Zeng, P. Satapoomin, J. Huang, M. Bowman, M. Iovene, W. Sanseverino, P. Cavagnaro, M. Yildiz, A. Macko-Podgórni, E. Moranska, E. Grzebelus, D. Grzebelus, H. Ashrafi, Z. Zheng, S. Cheng, D. Spooner, A. Van Deynze, and P. Simon. 2016. A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. Nat. Genet. 48:657–666.

Iorizzo, M., D. A. Senalik, S. L. Ellison, D. Grzebelus, P. F. Cavagnaro, C. Allender, J. Brunet, D. M. Spooner, A. Van Deynze, and P. W. Simon. 2013. Genetic structure and domestication of carrot (Daucus carota subsp. sativus) (Apiaceae). Am. J. Bot. 100:930–938.

Josephs, E. B., J. J. Berg, J. Ross-Ibarra, and G. Coop. 2019. Detecting Adaptive Differentiation in Structured Populations with Genomic Data and Common Gardens. Genetics 211:989–1004.

Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press.

Lande, R. 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. Evolution 46:381–389.

Lande, R. 1977. Statistical tests for natural selection on quantitative characters. Evolution 31:442–444.

Leinonen, T., R. B. O'Hara, J. M. Cano, and J. Merilä. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. J. Evol. Biol. 21:1–17.

Li, Z., A. Löytynoja, A. Fraimout, and J. Merilä. 2019. Effects of marker type and filtering criteria on Q ST-F ST comparisons. R Soc Open Sci 6:190666.

Lynch, M. 1990. The Rate of Morphological Evolution in Mammals from the Standpoint of the Neutral Expectation. Am. Nat. 136:727–741. The University of Chicago Press.

Mandel, J. R., A. J. Ramsey, M. Iorizzo, and P. W. Simon. 2016. Patterns of Gene Flow between Crop and Wild Carrot, *Daucus carota* (Apiaceae) in the United States. PLoS One 11:e0161971.

Martin, G., E. Chapuis, and J. Goudet. 2008. Multivariate QST-FST comparisons: a neutrality test for the evolution of the g matrix in structured populations. Genetics 180:2135–2149.

Merilä, J., and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. J. Evol. Biol. 14:892–903. Wiley.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U. S. A. 70:3321–3323.

O'Hara, R. B., and J. Merilä. 2005. Bias and precision in QST estimates: problems and some solutions. Genetics 171:1331–1339.

Ovaskainen, O., M. Karhunen, C. Zheng, J. M. C. Arias, and J. Merilä. 2011. A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. Genetics 189:621–632.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Pujol, B., A. J. Wilson, R. I. C. Ross, and J. R. Pannell. 2008. Are Q(ST)-F(ST) comparisons for natural populations meaningful? Mol. Ecol. 17:4782–4785.

Spitze, K. 1993. Population structure in Daphnia obtusa: quantitative genetic and allozymic variation. Genetics 135:367–374.

Team, R. C., and Others. 2013. R: A language and environment for statistical computing. Vienna, Austria.

Whitlock, M. C. 2008. Evolutionary inference from QST. Mol. Ecol. 17:1885–1896.

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

Whitlock, M. C., and F. Guillaume. 2009. Testing for spatially divergent selection: comparing QST to FST. Genetics 183:1055–1063.

Wijnheijmer, E. H. M., W. A. Brandenburg, and S. J. Ter Borg. 1989. Interactions between wild and cultivated carrots (*Daucus carota* L.) in the Netherlands. Euphytica 40:147–154. Springer Science and Business Media LLC.

Wilson, A. J., D. Réale, M. N. Clements, M. M. Morrissey, E. Postma, C. A. Walling, L. E. B. Kruuk, and D. H. Nussey. 2010. An ecologist's guide to the animal model. J. Anim. Ecol. 79:13–26.

Wright, S. 1949. The genetical structure of populations. Ann. Eugen. 15:323–354. Wiley.