The genetic control of market class in carrot (*Daucus carota* subsp. *sativus*)

By
Scott Holston Brainard

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Plant Breeding and Plant Genetics)

At the
UNIVERSITY OF WISCONSIN-MADISON
2021

Date of final oral examination: February 15th, 2021

The dissertation is approved by the following members of the Final Oral Committee:

Irwin L. Goldman, Professor, Horticulture
Julie C. Dawson, Associate Professor, Horticulture
Patrick J. Krysan, Professor, Horticulture
Phillip W. Simon, Professor, Horticulture
Edgar P. Spalding, Professor, Botany

Dedicated to

Joel Brainard

Acknowledgements

This work would not have been possible without aide from all sides. First, I would like to thank the generous financial assistance of the USDA-NPGS and the Swedish Foundation Ekhagastiftelsen. Second, my advisors, Dr. Irwin Goldman and Dr. Julie Dawson, assisted me in every step of my research. In addition, the Goldman Lab as a whole has been a tremendously supportive environment, and I thank the lab-mates I've shared this space with over the last four years: Chris D'Angelo, Lynn Maher, Claire Luby, Solveig Hanson, Katharine Wigg, Andrey Alfaro, Emilee Gaulke, Adam D'Angelo, Chandler Meyer, and Liam Dixon. In addition, the undergraduates who have worked in our lab, in particular Larissa Kahan, Matt Mirkes, and Taylor Sefzik, who took innumerable photos of carrots with me, displayed a truly indefatigable dedication to this work.

I have also benefited greatly from the guidance of the three other members of my graduate committee: Dr. Phil Simon, Dr. Edgar Spalding, and Dr. Patrick Krysan. Dr. Spalding, and his lab members Julian Bustamante and Dr. Nathan Miller, were invaluable in teaching me how to perform image analysis, and the second chapter of this thesis would not exist without their support. Dr. Simon, together with Dr. Doug Senalik and Dr. Shelby Ellison, taught me a tremendous amount about carrot genetics, and similarly, the third chapter of this dissertation would not have been possible without them. Dr. Krysan has been a deeply generous guide through the world of molecular genetics. In addition, I thank Dr. Steve Goldstein, Dr. Dean Sanders, and the entire CHTC helpdesk, for patiently teaching me everything I know about

bioinformatics, and Christoph Reimers, for his friendship, his creativity, and his continuous companionship in debugging.

I am also fortunate to be hitched to a truly inspiration circle of friends, among them Andy Petzinger, Dermot Delude-Dix, Maurice Weeks, and Leah Rethy; their daily support over the last 4 years is unprecedented.

Finally, I could not have completed much of anything, let alone a doctoral degree, without the love and support of my brother Junior, my sister Katherine, my mother Jane Carol, and my father Joel.

Dissertation Abstract

Carrot (*Daucus carota* subsp. *sativus*) is a nutritionally significant vegetable crop. An important target of selection in carrot breeding programs is suite of morphological root traits which together define market class—i.e., the market into which a specific variety is intended to be sold (e.g., juicing, dicing, storage, fresh market, baby carrot production). The size and shape the taproot, which can range from long and tapered to short and blunt, have been used for at least several centuries to classify cultivars in this way according to human preference and production methods. Mechanization in the cultivation, harvesting and post-harvest handling of the crop has made these traits increasingly relevant for both farmers and breeders. However, these quantitative phenotypes have historically been challenging to objectively evaluate, and thus subjective visual assessment of market class remains the primary method by which selection for these traits is performed. This has hindered not only the establishment of metric-based standards for market classes, but also the investigation the genetic basis of such quantitative phenotypes. In order to dissect the genetic control of the shape features that define market class in carrot, a tool is required that quantifies the specific shape features used by humans in distinguishing between classes.

Advancements in digital image analysis have recently made possible this high-throughput quantification of size and shape attributes, and Chapter 2 of this dissertation describes the functioning and performance of a phenotyping pipeline which implements such methods. This is the first such platform to include a series of a preprocessing algorithms whereby RGB images are converted to binary masks, which are then standardized to remove curvature and residual

root hairs. Phenotyping is then performed, which includes the quantification of traits that could be measured by hand, such as length and width, as well as measurement of higher-dimensional traits, through the implementation of principal components analysis of the root contour and its curvature. Of particular importance is the idnetification of a previously undescribed phenotype – root fill – as the most significant source of variation across carrot germplasm. This platform's high-throughput performance and accuracy was validated in two experimental panels: a diverse, global collection of germplasm was used to assess its capacity to identify market classes through clustering analysis, and diallel mating design between inbred breeding lines of differing market classes was used to estimate the heritability of the key phenotypes that define market class.

Together with the recent development of a high-quality reference genome for carrot, it is now feasible to utilize modern methods of genetic analysis in the investigation of the genetic control of root morphology. To this end, in Chapter 3 of this dissertation, the digital phenotypes of the diversity panel described in Chapter 2 are combined with a set of dense molecular markers developed using high-throughput sequencing. The use of both genome wide association analysis and genomic predictions based on genomic-estimated breeding values is described. Novel QTL were identified for four of the traits underlying market class; of particular interest is an extremely well-defined peak of chromosome 2 for the novel, and previously uncharacterized "root fill" trait. This comparative analysis provides the first convincing evidence that the traits underlying market class are highly polygenic in nature, under the influence of many small effect quantitative trait loci (QTL), but that relatively large proportions of additive genetic variance for

many of the component phenotypes support high predictive ability of genomic-estimated breeding values. This study thereby represents a novel advance in our understanding of the genetic control of market class in carrot root. In addition, concrete guidelines are presented outlining the practical potential of using genomic predictions for quantitative traits in horticultural crops.

Table of Contents

List of Tables

List of Figures

List of Supplementary Tables

List of Supplementary Figures

Chapter 1:
Review of literature

Background

Carrot (*Daucus carota* subsp. *carota* L.) is a biennial, diploid species (2n = 2x = 18), which is cultivated as an annual crop for its swollen taproot. Carrot is an economically important crop, both in the U.S., and across the world. Globally, 1,128,695 ha are planted annually (FAO, 2020), and although only 2% of this production area is in the U.S., carrot has a domestic farm gate value of 863.5 million USD (United States Department of Agriculture, 2020). This is in no small part due to carrot's nutritional importance as a significant source of pro-vitamin A (Simon, 2000). This latter fact, in turn, is a reflection of the success of carrot breeding efforts, which has resulted in a 50% increase in carotene content over the last 45 years (Iorizzo et al., 2016).

Market class in carrot

Since at least the 17th century, specific varieties of carrot have been selected for on the basis of the morphology of this taproot (Banga, 1957; Fig. 1). A number of factors seemed to have influenced early selection for different combinations of root size and shape traits. Cultivation practices certainly played a role, with soil depth influencing farmers' preferences for growing carrots of a specific length. In addition, whether the crop was grown in the field or in pots in a greenhouse environment, and at what latitude it was grown, influenced the amount of top growth one would need or desire, and thus the over size of the root. Many of the most extreme phenotypes in root shape are indeed a result a selection for forcing varieties grown primarily in greenhouse environments for winter production, giving rise to very short, almost-spherical varieties such as Davanture, Grelot, and Parisienne. Whether the crop was intended to be sold fresh, versus being stored and potentially transported long distances could

determine the desired girth of the carrot, which would affect its durability (e.g., the Long Stump Winter class shown in Fig. 1C). Whether the crop was intended to be used a fodder crop was also highly relevant in terms of shaping interest in varieties that potentially sacrificed eating-quality for yield—e.g., the less-conical, broadly shouldered Berlikum and Flanders types. More nuanced characters such as aspect ratio, the shape of the tip, or the broadness of the shoulder were also selected for, and appear to have been driven by both aesthetic and culinary preferences, as well as assumptions regarding the linkage of these traits to other phenotypes, such as root quality, and plant vigor (e.g., the perceived higher eating-quality of the Danvers and St. Valery classes shown in Fig. 1A and Fig. 1B). Interestingly, some more recently developed market classes have been documented as descending explicitly from intra-market class crosses; e.g., Imperator is the product of a cross between Nantes and Chantenay.

With the mechanization of carrot production, harvest, and post-harvest handling, these particular combinations of carrot root size and shape attributes have become increasingly standardized into a set of modern market classes. These classes have historically been defined subjectively, and to date, no objective criteria exist which explicitly delineate either the complete set of traits which together constitute any given market class, nor standards for judging whether an individual genotype's value for any one of these components traits falls within the range acceptable for a given class. For example, USDA grading standards refer only to carrots being "well formed", and provide size specifications that define minimum and maximums for length and shoulder width. Instead, market classification has been subjective,

performed largely by breeders responding to both the needs of farmers and their production systems, and consumers and their specific tastes and preferences.

The result of such a unique situation is that while market class itself has become an increasingly important breeding target for carrot, the genetic architecture underlying it has been the object of relatively limited study. This is a consequence of fact that the use of many of the most common tools of modern genetic analysis require precise, objective, and ideally quantitative measures of phenotypes which vary along continuous scales. Instead of thematizing the specific question of the genetic control of market class *per se*, investigations into carrot roots have to date therefore typically focused on the genetic regulation of the basic processes controlling storage root formation in general, and not the more subtle variation in these molecular pathways that give rise to the multitude of sizes and shapes seen in cultivated germplasm.

This has been frequently approached by attempting to characterize loci that were likely under selection during the domestication of cultivated carrot from wild accessions, leading to the identification of several candidate genes involved in the development of the swollen taproot as a storage root organ (Iorizzo et al., 2013; Macko-Podgórni et al., 2017). While wild carrot (*Daucus carota* subsp. *carota*) often has a taproot structure, it is spindle-shaped and heavily branched, which does not appear to undergo any secondary growth. Domestication of wild carrot was associated with the elimination of this branching habit, in which lateral roots emerge from the xylem. In addition, a primary selection target has been for a heavily swollen,

succulent main taproot, with high eating quality, which has typically resulted in a greater ratio of phloem to xylem in the core of the root.

This research dovetails with independent lines of inquiry that have attempted to characterize the molecular pathways leading specifically to the development of supernumerary cambial tissues during secondary growth of carrot. Benjamin et al. (1997) provided an early physiological description of the derivation of the storage root in carrot from secondary growth of supernumerary cambial layers, arising from cylindrical vascular cambium in the hypocotyl and root. A general finding in many crops has been that transcription factors regulating shoot apical meristem development may also play a role in these secondary growth processes (Schrader et al., 2004); research into the underlying molecular physiology in carrot, however, is still at a very nascent stage, particularly relative to work in other crop species. For example, as reviewed by Goldman, 2020, research on sweet potato has made significant strides in characterizing the hormonal regulation of secondary root growth (Ku et al., 2008; Ravi et al., 2014). Cytokinin and auxin levels have been found to regulate gene families that control storage root division both through upregulation of cell division (Noh et al., 2013), as well as starch and sucrose metabolism pathways (Firon et al., 2013; Tao et al., 2012). In carrot, comparatively little specific characterization has occurred. An early study by Ebener et al. (1993) suggested that a gene associated with secondary root growth, *DcPRP1*, appeared to be upregulated by auxin, while Wang et al. (2015), identified a number of gibberellin-regulated genes that were associated with root elongation. More recently, Wang et al. (2019) identified a set of genes encoding expansin-like proteins, whose regulation is likely involved in taproot

formation in cultivated carrot. Expansins are a widely studied class of proteins that facilitate turgor pressure-mediated cell elongation through a weakening of the cell wall by loosening attachments between xyloglucan and microfibrils (Cosgrove, 2000). It is thus logical that their expression patterns would be associated with growth and expansion of the carrot taproot.

Both of the lines of research described above are undeniably valuable. The identification of loci that underwent selection during domestication, as well as molecular analysis of the regulation of secondary growth in carrots, deepen our understanding of the genetic control of carrot taproot development. However, neither is oriented towards addressing the particular question of the genetic bases of market class. Market class as such is specifically a set of categories that only apply to cultivated carrot genotypes, though not all cultivars fall unambiguously into a specific market class. Thus, the loci that distinguish cultivated carrots from wild accessions are likely in themselves not responsible for the genetic regulation of this intra-cultivated carrot diversity. Similarly, the common processes regulating taproot expansion during secondary growth are to variable extents shared across all market classes, and it is rather the fine-tuned regulation of these processes that likely is involved in determining whether a given genotype gives rise to a specific market class type.

A prerequisite to any genetic analysis of the set of traits which define market class, however, is a phenotyping methodology allowing for their precise quantification. The development and validation of such a platform was therefore one of the main aims of this research.

Digital image phenotyping

Precise measurement of morphological phenotypes is critical for any attempt at studying the genetic architecture of quantitative plant traits. Indeed, as the costs of genotyping even non-model plant species for thousands of polymorphic molecular markers continues to decrease rapidly, the time and labor required for accurate phenotyping is increasingly emerging as the limiting factor in genetic mapping and association studies. In this context, digital image-based phenotyping holds substantial promise in overcoming many of the practical limitations of traditional quantification of continuously variable traits (Fahlgren et al., 2015).

Though the term "digital image-based phenotyping" encompasses a broad array of specific technical workflows, in general terms, these techniques provide four advantages over traditional measurements, whether made by hand, or by eye. First, and perhaps most obviously, due to the decreasing cost of high-resolution digital cameras, it is now feasible to acquire extremely precise images. This facilitates the accurate detection and measurement of plant structures—easily to within fractions of millimeters (Kuijken et al., 2015). Second, by automating image acquisition, multiplexing multiple samples in a single image, and integrating technologies such as QR matrix barcoding into the image-acquisition pipeline, digital-phenotyping workflows often fall within the context of "high-throughput" technologies, which facilitate the measurement of greater numbers of samples than would be possible by hand (Furbank & Tester, 2011). Third, through the implementation of robust machine vision algorithms, the phenotyping of digital images is by definition standardized, removing bias that is often introduced by having multiple individuals scoring and measuring traits. Fourth, these

machine vision algorithms can easily quantify aspects of plant morphology that are not possible when measuring individual samples by hand, and may more accurately capture the genetic component of a given phenotype. For instance, as suggested by Howarth et al. (1992), and previously implemented in maize (Miller et al., 2017), tree leaves (Biot et al., 2016), carrots (Horgan et al., 2001; Turner, Senalik, et al., 2018) and radish (Iwata et al., 1998), dimensionality reduction techniques such as principal components analysis (PCA) can provide a powerful method for assessing complex, high-dimensional sources of variation in an image.

Diallel analysis

The phenotypes extracted from the analysis of digital images were utilized in this thesis in the context of three distinct approaches for analyzing the genetic control of quantitative traits. The first of these was the analysis of a diallel mating design. Initially developed in the context of the inbred-hybrid breeding system in corn, the conceptual framework for the diallel was proposed as a method for identifying high-performing hybrid crosses. This is the origin of the terms "general combining ability" (GCA) and "specific combining ability" (SCA) which refer respectively to the average performance of the hybrid progeny of a given inbred line, and the deviations from the GCA effects of two inbred parents in any particular hybrid progeny family (Sprague & Tatum, 1942). A diallel was thus originally proposed as a method for systematically determining GCA and SCA effects for a set of inbred lines by performing all pair-wise crosses between them, and evaluating the F_1 progeny families (Hayman, 1954a, 1954b).

The most common set of methods for analyzing the performance of these F_1 families is that proposed by of Griffing (1956), who described a set of linear models that allow for analyzing diallel experiments in which parental lines can modeled as either fixed or random effects (i.e., depending on whether the parental lines constitute a particular population within which predictions are to be made, versus a random sample from a larger set of possible parental material). These models allowed for the inclusion or exclusion of reciprocal crosses (the former allowing for the estimation of maternal and paternal effects), and are also robust to replicated experimental designs such as randomized complete block designs, as well as multi-environmental trials (Lin et al., 1977). These experimental designs are clearly highly resource intensive, however, since the number of F_1 families to be evaluated is proportional to the square of the number of parental lines. Even the evaluation of 10 parents can quickly become logistically untenable. As a result, Bayesian methods have been developed for imputing missing data from more sparse datasets (Lenarcic et al., 2012), and such methods have been validated as effective in carrot (Turner et al., 2018).

Utilization of genomic-scale data in carrot

Since the initial release of the first high-quality reference assembly for carrot (Iorizzo et al., 2016), researchers have begun to utilize data from next-generation sequencing platforms to explore the potential of both genome-wide association analysis (GWAS), and genomic prediction. These methods aim to deepen our understanding of the genetic control of key traits in carrot, as well as improve the efficiency of selecting for them in a breeding context. However, while utilizing similar phenotypic and genotypic data, and based on related statistical

theory, these two methodologies have several critical features that distinguish them, and therefore their utility in the context of understanding, and breeding for, quantitative traits. GWAS attempts to detect significant associations between molecular markers and phenotypic variation using a linear mixed model where the fixed effects are the markers themselves, and the random effect is a polygenic term that is used to control for what would otherwise be confounding population structure. These associations are then used either as the basis for marker-assisted selection, or future fine-mapping studies that attempt to characterize causal genes. Genomic selection also makes used of linear mixed models, in this case, the experimental design is modeled using fixed effects, while the random polygenic term is used to estimate genotypic effects directly. These estimates are then used to predict either the *per se* performance of an individual, or, more commonly, its breeding value (i.e., the "genomic-estimated breeding value"; GEBV).

Genome-wide association studies

With regards to GWAS, it is important to note that as an outcrossing species, carrot is both naturally highly heterozygous, and prone to inbreeding depression. The production of fully homozygous inbred lines has therefore been a major target of modern hybrid breeding efforts, but significant variation for tolerance of inbreeding remains (Simon et al., 2008). As a consequence, inbred lines are relatively rare in relation to open-pollinated varieties and outcrossing landraces, and recombinant inbred lines are currently still being developed. Because of the limited availability of the genetic resources that form the basis for more traditional linkage mapping approaches, GWAS offers a powerful method for leveraging

phenotypic diversity for root shape traits in an attempt to identify linked quantitative trait loci (QTL).

The lack of controlled, known pedigrees in GWAS present several potential complications. First, minor allele frequencies (MAF) can be much lower than in typical linkage mapping populations. This can limit power to detect QTL, due to the fact that the genetic variance attributable to a given QTL (assuming for simplification that it is bi-allelic) is equal to $4a^2pq$, where a is equal to the additive effect of the QTL, and p and q are the frequencies of the two alleles (Tabangin et al., 2009). As MAF decreases, so does this genetic variance term. This reduces power, because these parameters simultaneously determine the non-centrality of the *t*-distribution of the alternative hypothesis that there is a QTL at a given locus. In this case, the non-centrality parameter = $\frac{2a}{\sigma}\sqrt{npq}$, where n is the size of the population; thus, as genetic variance decreases, so too does the distance between the means of the alternative and null *t*-distributions. As a result of this statistical result, it is common practice to threshold MAF at a certain value, so as to only include SNPs for which power remains above some desired level, given a specified false-positive rate. This has the additional benefit of limiting the amount of variance in the allele frequency distributions of the panel, and thus improving the chances of having markers in high linkage disequilibrium with one another (a prerequisite for high linkage disequilibrium being that markers have similar allele frequency distributions). While this of course reduces marker density, next-generation sequencing has made it practically feasible to identify orders of magnitude more polymorphic markers than are necessary for association analysis.

Second, because of the variable degrees of relatedness between individuals in a diversity panel, identification of spurious associations between markers and a trait of interest can arise simply due to the fact that many markers may be in linkage with a causal marker within a closely related group of individuals (Sul et al., 2018). Clearly, a filtering procedure equivalent to what is employed with regards to allele frequency would not be ideal, due to the constraints this would place on population size. As a result, numerous methods have been proposed for adding terms to the linear models that are used in associated SNPs with the phenotype of interest. One approach is to include fixed effect terms, such as the principal component scores of PCA on the marker matrix (Price et al., 2006), or the grouping assignments of clustering algorithms such as STRUCTURE (Pritchard et al., 2000). An increasingly common approach, as mentioned above, is to use genetic markers themselves as covariates, by including the centered genotype matrix as a random effects term (Yu et al., 2006). All of these methods have been validated as controlling for the confounding effects of population structure, with minimal costs to statistical power. In addition, it is important to note that in carrot specifically, extremely limited degrees of population structure have been found in global collections that have been used in GWAS previously (Ellison et al., 2018). In cultivated accessions specifically, carrot germplasm represents a nearly unstructured population that is therefore well-suited to GWAS (Corak et al., 2019).

While GWAS has been used extensively in studying agronomically important traits, with numerous published studies that have focused specifically on root shape and root system

architecture (Alahmad et al., 2019; Beyer et al., 2019; Wang et al., 2019), only three genome-wide association studies have been reported in carrot to date. Two have focused on the genetic control of secondary metabolites: one on terpene synthase genes (Keilwagen et al., 2017), and a second on the genetic control of carotenoid accumulation (Ellison et al., 2018). The third study represent the only assessment to date of the genetic control of root shape in carrot utilizing a diversity panel: Macko-Podgórni et al. (2020) performed GWAS using 103 open-pollinated carrot accessions, and identified a single locus on chromosome 1 associated with crown width. As the authors describe, this study was limited by a relatively small number of accessions in the diversity panel, and a high degree of environmental variance due to a limited experimental design: genotypes were only grown in one in environment, in one year, without replication, and as a result, phenotyping was performed on only a single root per genotype. Nevertheless, this study demonstrates the potential utility of GWAS in the context of studying complex, quantitative traits such as carrot root shape, and thus motivated the more extensive analysis presented here.

Genomic prediction

While there has been a relatively limited use of GWAS in carrot, there been even more restricted evaluation of the utility of genomic prediction methods in carrot. Across a wide range of species, however, genomic selection has shown significant promise. For instance, in maize (Beyene et al., 2015; Combs & Bernardo, 2013b) and tomato (Yamamoto et al., 2017), genomic selection has been shown to outperform phenotypic selection when evaluated in terms of the rate of genetic gain, while the two methods were found to be equally efficient in wheat (Rutkoski et al., 2015). To date, only one study has considered the potential utility of

GEBVs in carrot breeding, and this analysis was restricted to two traits: flavor and canopy height (Corak et al., 2019). While this study did not compare the efficacy of genomic prediction relative to phenotypic selection, it did suggest that even for moderately heritable traits, GEBVs could attain non-negligible predictive abilities.

A key consideration in any evaluation of the potential of genomic prediction to aide in selecting for a specific set of traits in a particular crop is an evaluation of the predictive ability of genomic-estimated breeding values (GEBVs) under a range of parameters know to influence reliability. In particular, four variables have frequently been cited as having an impact on predictive ability. First, heritability of the trait in question is of course key (Ben Hassen et al., 2018; Combs & Bernardo, 2013a; Gorjanc et al., 2016); in this regard, increasing replication, improving experimental design, and reducing error in the methodology used to phenotype a given trait will all clearly improve genomic predictions.

Second, the influence of marker density has also been evaluated. Because markers are used in genomic selection not as fixed effects that can be associated with putative QTL, but rather as a covariance matrix for estimating genotypic effects, it has long been assumed that simply increasing marker density would have diminishing returns in terms of increasing prediction accuracy. Because genomic selection is typically utilized in applied contexts, it is critical from a cost-perspective to determine for any given trait and species, what marker density is sufficient for attaining desired predictive ability. A relatively robust finding in this regard is that even several thousand markers are often sufficient for reaching high prediction accuracies. (Asoro et

al., 2011; Erbe et al., 2013; Guo et al., 2018; Q. Wang et al., 2017; Wu et al., 2016; Zhang et al., 2015).

Third, the role of population size in determining predictive ability has been analyzed from a variety of perspectives. First, the effect of the total population size, independent of how it is apportioned between training and validation sets has been examined, generally finding that larger populations lead to higher accuracies (Ben Hassen et al., 2018; Heslot et al., 2015; Xu et al., 2018). In addition, the relative size of either the training population (Tayeh et al., 2015), or the relative size of the validation population (Zhang et al., 2017) has also been considered, finding in both cases that the more individuals in the training population, the higher the prediction accuracies. However, this effect can be diluted or in some cases even reversed if the additional training individuals are highly unrelated to the validation population (Clark et al., 2012; Lorenz & Smith, 2015). This has led to the consideration of a fourth and final parameter: the degree of genetic similarity between the training and validation populations. Because genomic predictions are based on covariance between individuals in the training population and individuals in the validation population, higher degrees of relatedness between these groups has long been hypothesized to lead to higher predictive ability. And indeed, in simulation studies which have considered this question, this appears to be born out whether predictive ability is compared across populations with known degrees of variable relatedness (such as full- vs. half-sib families), or between population groupings defined through clustering algorithms such STRUCTURE (Berro et al., 2019; Lozada et al., 2019; Riedelsheimer et al., 2013; Sverrisdóttir et al., 2018).

Conclusion

The three methods of quantitative genetic analysis described above – the diallel mating design, GWAS, and genomic prediction – each represent a distinct experimental approach to attempting to explain genetic control over traits of interest. Though they all represent potential approaches to utilizing phenotypic data extracted from digital image analysis, they rely upon different forms of statistical models as well as genotypic information. As a result, the conclusions one is able to draw from each of these experiments are also distinct. Diallel designs are clearly best suited to an inbred-hybrid breeding context, where hybrid prediction is most relevant, and understanding the relative contributions of additive and dominance variance is of utmost importance. GWAS also has practical implications for breeding, but is more often a preliminary step in the identification of QTL which then must be validated as being durable within a specific breeding background, and ideally fine-mapped to locate the specific causative genes influencing phenotypic variation. Genomic prediction is, like the diallel mating design, a method specifically oriented towards improving efficiency in a breeding program. However, instead of constructing a statistical model using a limited set of parents, and developing models specifically useful for predicting the performance of hybrid combinations, GEBVs aim to provide a more generalizable method for optimizing parent selection itself. By utilizing a diversity of methods in this thesis, not only will all of these sources of information be available to geneticists and breeders attempting to study and select for carrot market class, but

comparative analyses will also be made possible, allowing for a judgement as to which methods hold the most promise for subsequent research into these quantitative traits.

References

Alahmad, S., El Hassouni, K., Bassi, F. M., Dinglasan, E., Youssef, C., Quarry, G., Aksoy, A., Mazzucotelli, E., Juhász, A., Able, J. A., Christopher, J., Voss-Fels, K. P., & Hickey, L. T. (2019). A Major Root Architecture QTL Responding to Water Limitation in Durum Wheat. *Frontiers in Plant Science*, *10*, 436. https://doi.org/10.3389/fpls.2019.00436

Asoro, F., Newell, M., Beavis, W., Scott, P., & Jannink, J.-L. (2011). Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats. *Plant Gen.*, *4*, 132–144. https://doi.org/10.3835/plantgenome2011.02.0007

Banga, O. (1957). *Origin of the European cultivated carrot*. Instituut voor de Veredeling van Tuinbouwgewassen.

Ben Hassen, M., Cao, T. V, Bartholomé, J., Orasen, G., Colombi, C., Rakotomalala, J., Razafinimpiasa, L., Bertone, C., Biselli, C., Volante, A., Desiderio, F., Jacquin, L., Valè, G., & Ahmadi, N. (2018). Rice diversity panel provides accurate genomic predictions for complex traits in the progenies of biparental crosses involving members of the panel. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *131*(2), 417–435. https://doi.org/10.1007/s00122-017-3011-4

Benjamin, L. R., McGarry, A., & Gray, D. (1997). *The root vegetables: beet, carrot, parsnip and turnip*.

Berro, I., Lado, B., Nalin, R. S., Quincke, M., & Gutiérrez, L. (2019). Training Population Optimization for Genomic Selection. *The Plant Genome*, *12*(3), 190028. https://doi.org/https://doi.org/10.3835/plantgenome2019.04.0028

Beyene, Y., Semagn, K., Mugo, S., Tarekegne, A., Babu, R., Meisel, B., Sehabiague, P., Makumbi,

D., Magorokosho, C., Oikeh, S., Gakunga, J., Vargas, M., Olsen, M., Prasanna, B. M., Banziger, M., & Crossa, J. (2015). Genetic Gains in Grain Yield Through Genomic Selection in Eight Bi-parental Maize Populations under Drought Stress. *Crop Science*, *55*(1), 154–163. https://doi.org/https://doi.org/10.2135/cropsci2014.07.0460

Beyer, S., Daba, S., Tyagi, P., Bockelman, H., Brown-Guedira, G., Mohammadi, M., & IWGSC. (2019). Loci and candidate genes controlling root traits in wheat seedlings—a wheat root GWAS. *Functional & Integrative Genomics*, *19*(1), 91–107. https://doi.org/10.1007/s10142-018-0630-z

Biot, E., Cortizo, M., Burguet, J., Kiss, A., Oughou, M., Maugarny-Calès, A., Gonçalves, B., Adroher, B., Andrey, P., Boudaoud, A., & Laufs, P. (2016). Multiscale quantification of morphodynamics: MorphoLeaf software for 2D shape analysis. *Development*, *143*(18), 3417 LP – 3428. https://doi.org/10.1242/dev.134619

Clark, S. A., Hickey, J. M., Daetwyler, H. D., & van der Werf, J. H. J. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genetics, Selection, Evolution : GSE*, *44*(1), 4. https://doi.org/10.1186/1297-9686-44-4

Combs, E., & Bernardo, R. (2013a). Accuracy of Genomewide Selection for Different Traits with Constant Population Size, Heritability, and Number of Markers. *The Plant Genome*, *6*(1), plantgenome2012.11.0030. https://doi.org/https://doi.org/10.3835/plantgenome2012.11.0030

Combs, E., & Bernardo, R. (2013b). Genomewide Selection to Introgress Semidwarf Maize Germplasm into U.S. Corn Belt Inbreds. *Crop Science*, *53*(4), 1427–1436.

https://doi.org/https://doi.org/10.2135/cropsci2012.11.0666

Corak, K. E., Ellison, S. L., Simon, P. W., Spooner, D. M., & Dawson, J. C. (2019). Comparison of Representative and Custom Methods of Generating Core Subsets of a Carrot Germplasm Collection. *Crop Science*, *59*(3), 1107–1121. https://doi.org/https://doi.org/10.2135/cropsci2018.09.0602

Cosgrove, D. J. (2000). Loosening of plant cell walls by expansins. *Nature*, *407*(6802), 321–326. https://doi.org/10.1038/35030000

Ebener, W., Fowler, T. J., Suzuki, H., Shaver, J., & Tierney, M. L. (1993). Expression of DcPRP1 is linked to carrot storage root formation and is induced by wounding and auxin treatment. *Plant Physiology*, *101*(1), 259–265. https://doi.org/10.1104/pp.101.1.259

Ellison, S. L., Luby, C. H., Corak, K. E., Coe, K. M., Senalik, D., Iorizzo, M., Goldman, I. L., Simon, P. W., & Dawson, J. C. (2018). Carotenoid presence is associated with the Or gene in domesticated carrot. *Genetics*, *210*(4), 1497–1508. https://doi.org/10.1534/genetics.118.301299

Erbe, M., Gredler, B., Seefried, F. R., Bapst, B., & Simianer, H. (2013). A Function Accounting for Training Set Size and Marker Density to Model the Average Accuracy of Genomic Prediction. *PLOS ONE*, *8*(12), 1–11. https://doi.org/10.1371/journal.pone.0081046

Fahlgren, N., Feldman, M., Gehan, M. A., Wilson, M. S., Shyu, C., Bryant, D. W., Hill, S. T., McEntee, C. J., Warnasooriya, S. N., Kumar, I., Ficor, T., Turnipseed, S., Gilbert, K. B., Brutnell, T. P., Carrington, J. C., Mockler, T. C., & Baxter, I. (2015). A Versatile Phenotyping System and Analytics Platform Reveals Diverse Temporal Responses to Water Availability in Setaria. *Molecular Plant*, *8*(10), 1520–1535.

https://doi.org/https://doi.org/10.1016/j.molp.2015.06.005

FAO. (2020). *FAOSTAT*.

Firon, N., LaBonte, D., Villordon, A., Kfir, Y., Solis, J., Lapis, E., Perlman, T. S., Doron-Faigenboim, A., Hetzroni, A., Althan, L., & Adani Nadir, L. (2013). Transcriptional profiling of sweetpotato (Ipomoea batatas) roots indicates down-regulation of lignin biosynthesis and up-regulation of starch biosynthesis at an early stage of storage root formation. *BMC Genomics*, *14*, 460. https://doi.org/10.1186/1471-2164-14-460

Furbank, R. T., & Tester, M. (2011). Phenomics--technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, *16*(12), 635–644. https://doi.org/10.1016/j.tplants.2011.09.005

Goldman, I. L. (2020). The Root Vegetables: Beet, Carrot, Parsnip, and Turnip. In *The Physiology of Vegetable Crops* (pp. 399–420). CABI.

Gorjanc, G., Jenko, J., Hearne, S. J., & Hickey, J. M. (2016). Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. *BMC Genomics*, *17*(1), 30. https://doi.org/10.1186/s12864-015-2345-z

Griffing, B. (1956). Concept of General and Specific Combining Ability in Relation to Diallel Crossing Systems. *Australian Journal of Biological Sciences*, *9*(4), 463. https://doi.org/10.1071/bi9560463

Guo, X., Cericola, F., Fè, D., Pedersen, M. G., Lenk, I., Jensen, C. S., Jensen, J., & Janss, L. L. (2018). Genomic Prediction in Tetraploid Ryegrass Using Allele Frequencies Based on Genotyping by Sequencing. *Frontiers in Plant Science*, *9*, 1165. https://doi.org/10.3389/fpls.2018.01165

Hayman, B. I. (1954a). The Analysis of Variance of Diallel Tables. *Biometrics*, *10*(2), 235–244. https://doi.org/10.2307/3001877

Hayman, B. I. (1954b). The Theory and Analysis of Diallel Crosses. *Genetics*, *39*(6), 789–809. https://pubmed.ncbi.nlm.nih.gov/17247520

Heslot, N., Jannink, J.-L., & Sorrells, M. E. (2015). Perspectives for Genomic Selection Applications and Research in Plants. *Crop Science*, *55*(1), 1–12. https://doi.org/https://doi.org/10.2135/cropsci2014.03.0249

Horgan, G. W., Talbot, M., & Davey, J. C. (2001). Use of statistical image analysis to discriminate carrot cultivars. *Computers and Electronics in Agriculture*, *31*(2), 191–199. https://doi.org/10.1016/S0168-1699(00)00191-5

Howarth, M. S., Brandon, J. R., Searcy, S. W., & Kehtarnavaz, N. (1992). Estimation of tip shape for carrot classification by machine vision. *Journal of Agricultural Engineering Research*, *53*, 123–139. https://doi.org/https://doi.org/10.1016/0021-8634(92)80078-7

Iorizzo, M., Ellison, S., Senalik, D., Zeng, P., Satapoomin, P., Huang, J., Bowman, M., Iovene, M., Sanseverino, W., Cavagnaro, P., Yildiz, M., Macko-Podgórni, A., Moranska, E., Grzebelus, E., Grzebelus, D., Ashrafi, H., Zheng, Z., Cheng, S., Spooner, D., … Simon, P. (2016). A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nature Genetics*, *48*(6), 657–666. https://doi.org/10.1038/ng.3565

Iorizzo, M., Senalik, D. A., Ellison, S. L., Grzebelus, D., Cavagnaro, P. F., Allender, C., Brunet, J., Spooner, D. M., Van Deynze, A., & Simon, P. W. (2013). Genetic structure and domestication of carrot (Daucus carota subsp. sativus) (Apiaceae). *American Journal of*

Botany, *100*(5), 930–938. https://doi.org/10.3732/ajb.1300055

Iwata, H., Niikura, S., Matsuura, S., Takano, Y., & Ukai, Y. (1998). Evaluation of variation of root shape of Japanese radish (Raphanus sativus L.) based on image analysis using elliptic Fourier descriptors. *Euphytica*, *102*(2), 143–149. https://doi.org/10.1023/A:1018392531226

Keilwagen, J., Lehnert, H., Berner, T., Budahn, H., Nothnagel, T., Ulrich, D., & Dunemann, F. (2017). The Terpene Synthase Gene Family of Carrot (Daucus carota L.): Identification of QTLs and Candidate Genes Associated with Terpenoid Volatile Compounds. *Frontiers in Plant Science*, *8*, 1930. https://doi.org/10.3389/fpls.2017.01930

Ku, A. T., Huang, Y.-S., Wang, Y.-S., Ma, D., & Yeh, K.-W. (2008). IbMADS1 (Ipomoea batatas MADS-box 1 gene) is involved in tuberous root initiation in sweet potato (Ipomoea batatas). *Annals of Botany*, *102*(1), 57–67. https://doi.org/10.1093/aob/mcn067

Kuijken, R. C. P., van Eeuwijk, F. A., Marcelis, L. F. M., & Bouwmeester, H. J. (2015). Root phenotyping: from component trait in the lab to breeding. *Journal of Experimental Botany*, *66*(18), 5389–5401. https://doi.org/10.1093/jxb/erv239

Lenarcic, A. B., Svenson, K. L., Churchill, G. A., & Valdar, W. (2012). A General Bayesian Approach to Analyzing Diallel Crosses of Inbred Strains. *Genetics*, *190*(2), 413 LP – 435. https://doi.org/10.1534/genetics.111.132563

Lin, C. S., Binns, M. R., & Thompson, B. K. (1977). The use of regression methods to study genotype-environment interactions. *Heredity*, *38*(3), 309–319. https://doi.org/10.1038/hdy.1977.93

Lorenz, A. J., & Smith, K. P. (2015). Adding Genetically Distant Individuals to Training

Populations Reduces Genomic Prediction Accuracy in Barley. *Crop Science*, *55*(6), 2657–2667. https://doi.org/https://doi.org/10.2135/cropsci2014.12.0827

Lozada, D. N., Mason, R. E., Sarinelli, J. M., & Brown-Guedira, G. (2019). Accuracy of genomic selection for grain yield and agronomic traits in soft red winter wheat. *BMC Genetics*, *20*(1), 82. https://doi.org/10.1186/s12863-019-0785-1

Macko-Podgórni, A., Machaj, G., Stelmach, K., Senalik, D., Grzebelus, E., Iorizzo, M., Simon, P. W., & Grzebelus, D. (2017). Characterization of a genomic region under selection in cultivated carrot (Daucus carota subsp. sativus) reveals a candidate domestication gene. *Frontiers in Plant Science*, *8*, 12.

Macko-Podgórni, A., Stelmach, K., Kwolek, K., Machaj, G., Ellison, S., Senalik, D. A., Simon, P. W., & Grzebelus, D. (2020). Mining for Candidate Genes Controlling Secondary Growth of the Carrot Storage Root. *International Journal of Molecular Sciences*, *21*(12), 1–16. https://doi.org/10.3390/ijms21124263

Miller, N. D., Haase, N. J., Lee, J., Kaeppler, S. M., de Leon, N., & Spalding, E. P. (2017). A robust, high-throughput method for computing maize ear, cob, and kernel attributes automatically from images. *The Plant Journal*, *89*(1), 169–178. https://doi.org/10.1111/tpj.13320

Noh, S. A., Lee, H.-S., Kim, Y.-S., Paek, K.-H., Shin, J. S., & Bae, J. M. (2013). Down-regulation of the IbEXP1 gene enhanced storage root development in sweetpotato. *Journal of Experimental Botany*, *64*(1), 129–142. https://doi.org/10.1093/jxb/ers236

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association

studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959. http://www.genetics.org/content/155/2/945.abstract

Ravi, V., Chakrabarti, S. K., Makeshkumar, T., & Saravanan, R. (2014). Molecular Regulation of Storage Root Formation and Development in Sweet Potato. In *Horticultural Reviews: Volume 42* (pp. 157–208). https://doi.org/https://doi.org/10.1002/9781118916827.ch03

Riedelsheimer, C., Endelman, J. B., Stange, M., Sorrells, M. E., Jannink, J.-L., & Melchinger, A. E. (2013). Genomic predictability of interconnected biparental maize populations. *Genetics*, *194*(2), 493–503. https://doi.org/10.1534/genetics.113.150227

Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., & Sorrells, M. E. (2015). Genetic Gain from Phenotypic and Genomic Selection for Quantitative Resistance to Stem Rust of Wheat. *The Plant Genome*, *8*(2), plantgenome2014.10.0074. https://doi.org/https://doi.org/10.3835/plantgenome2014.10.0074

Schrader, J., Nilsson, J., Mellerowicz, E., Berglund, A., Nilsson, P., Hertzberg, M., & Sandberg, G. (2004). A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity. *The Plant Cell*, *16*(9), 2278–2292. https://doi.org/10.1105/tpc.104.024190

Simon, P. W. (2000). Domestication, Historical Development, and Modern Breeding of Carrot. In *Plant Breeding Reviews* (pp. 157–190). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470650172.ch5

Simon, P. W., Freeman, R. E., Vieira, J. V, Boiteux, L. S., & Briard, M. (2008). Carrot. *Vegetables*

II: Fabaceae, Liliaceae, Solanaceae, and Umbelliferae, 327–357. https://doi.org/10.1007/978-0-387-74110-9

Sprague, G. F., & Tatum, L. A. (1942). General vs. Specific Combining Ability in Single Crosses of Corn1. *Agronomy Journal*, *34*(10), 923–932. https://doi.org/https://doi.org/10.2134/agronj1942.00021962003400100008x

Sul, J. H., Martin, L. S., & Eskin, E. (2018). Population structure in genetic studies: Confounding factors and mixed models. *PLOS Genetics*, *14*(12), e1007309. https://doi.org/10.1371/journal.pgen.1007309

Sverrisdóttir, E., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., Janss, L., Bryan, G., & Nielsen, K. L. (2018). The Value of Expanding the Training Population to Improve Genomic Selection Models in Tetraploid Potato. *Frontiers in Plant Science*, *9*, 1118. https://doi.org/10.3389/fpls.2018.01118

Tabangin, M. E., Woo, J. G., & Martin, L. J. (2009). The effect of minor allele frequency on the likelihood of obtaining false positives. *BMC Proceedings*, *3 Suppl 7*(Suppl 7), S41. https://doi.org/10.1186/1753-6561-3-S7-S41

Tao, X., Gu, Y.-H., Wang, H.-Y., Zheng, W., Li, X., Zhao, C.-W., & Zhang, Y.-Z. (2012). Digital gene expression analysis based on integrated de novo transcriptome assembly of sweet potato [Ipomoea batatas (L.) Lam]. *PloS One*, *7*(4), e36234. https://doi.org/10.1371/journal.pone.0036234

Tayeh, N., Klein, A., Le Paslier, M.-C., Jacquin, F., Houtin, H., Rond, C., Chabert-Martinello, M., Magnin-Robert, J.-B., Marget, P., Aubert, G., & Burstin, J. (2015). Genomic Prediction in Pea: Effect of Marker Density and Training Population Size and Composition on Prediction

Accuracy. *Frontiers in Plant Science*, *6*, 941. https://doi.org/10.3389/fpls.2015.00941

Turner, S. D., Maurizio, P. L., Valdar, W., Yandell, B. S., & Simon, P. W. (2018). Dissecting the Genetic Architecture of Shoot Growth in Carrot (Daucus carota L.) Using a Diallel Mating Design. *G3: Genes|Genomes|Genetics*, *8*(2), 411 LP – 426. https://doi.org/10.1534/g3.117.300235

Turner, S. D., Senalik, D. A., Simon, P. W., Ellison, S. L., Miller, N. D., & Spalding, E. P. (2018). An Automated Image Analysis Pipeline Enables Genetic Studies of Shoot and Root Morphology in Carrot (Daucus carota L.). *Frontiers in Plant Science*, *9*(1703), 1–17. https://doi.org/10.3389/fpls.2018.01703

United States Department of Agriculture. (2020). *National Agricultural Statistics Service*. https://www.nass.usda.gov/Statistics_by_Subject/

Wang, G.-L., Jia, X.-L., Xu, Z.-S., Wang, F., & Xiong, A.-S. (2015). Sequencing, assembly, annotation, and gene expression: novel insights into the hormonal control of carrot root development revealed by a high-throughput transcriptome. *Molecular Genetics and Genomics : MGG*, *290*(4), 1379–1391. https://doi.org/10.1007/s00438-015-0999-5

Wang, H., Wei, J., Li, P., Wang, Y., Ge, Z., Qian, J., Fan, Y., Ni, J., Xu, Y., Yang, Z., & Xu, C. (2019). Integrating GWAS and Gene Expression Analysis Identifies Candidate Genes for Root Morphology Traits in Maize at the Seedling Stage. In *Genes* (Vol. 10, Issue 10). https://doi.org/10.3390/genes10100773

Wang, Q., Yu, Y., Yuan, J., Zhang, X., Huang, H., Li, F., & Xiang, J. (2017). Effects of marker density and population structure on the genomic prediction accuracy for growth trait in Pacific white shrimp Litopenaeus vannamei. *BMC Genetics*, *18*(1), 45.

https://doi.org/10.1186/s12863-017-0507-5

Wang, Y.-H., Que, F., Wang, G.-L., Hao, J.-N., Li, T., Xu, Z.-S., & Xiong, A.-S. (2019). iTRAQ-Based Quantitative Proteomics and Transcriptomics Provide Insights Into the Importance of Expansins During Root Development in Carrot. *Frontiers in Genetics*, *10*(March), 1–14. https://doi.org/10.3389/fgene.2019.00247

Wu, X.-L., Xu, J., Feng, G., Wiggans, G. R., Taylor, J. F., He, J., Qian, C., Qiu, J., Simpson, B., Walker, J., & Bauck, S. (2016). Optimal Design of Low-Density SNP Arrays for Genomic Prediction: Algorithm and Applications. *PLOS ONE*, *11*(9), 1–36. https://doi.org/10.1371/journal.pone.0161719

Xu, Y., Wang, X., Ding, X., Zheng, X., Yang, Z., Xu, C., & Hu, Z. (2018). Genomic selection of agronomic traits in hybrid rice using an NCII population. *Rice*, *11*(1), 32. https://doi.org/10.1186/s12284-018-0223-4

Yamamoto, E., Matsunaga, H., Onogi, A., Ohyama, A., Miyatake, K., Yamaguchi, H., Nunome, T., Iwata, H., & Fukuoka, H. (2017). Efficiency of genomic selection for breeding population design and phenotype prediction in tomato. *Heredity*, *118*(2), 202–209. https://doi.org/10.1038/hdy.2016.84

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, *38*(2), 203–208. https://doi.org/10.1038/ng1702

Zhang, A., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., San Vicente, F., Olsen, M., Prasanna, B. M., Crossa, J., Yu, H., & Zhang, X. (2017). Effect of Trait

Heritability, Training Population Size and Marker Density on Genomic Prediction Accuracy Estimation in 22 bi-parental Tropical Maize Populations. *Frontiers in Plant Science*, *8*, 1916. https://doi.org/10.3389/fpls.2017.01916

Zhang, X., Pérez-Rodríguez, P., Semagn, K., Beyene, Y., Babu, R., López-Cruz, M. A., San Vicente, F., Olsen, M., Buckler, E., Jannink, J.-L., Prasanna, B. M., & Crossa, J. (2015). Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. *Heredity*, *114*(3), 291–299. https://doi.org/10.1038/hdy.2014.99

Figures

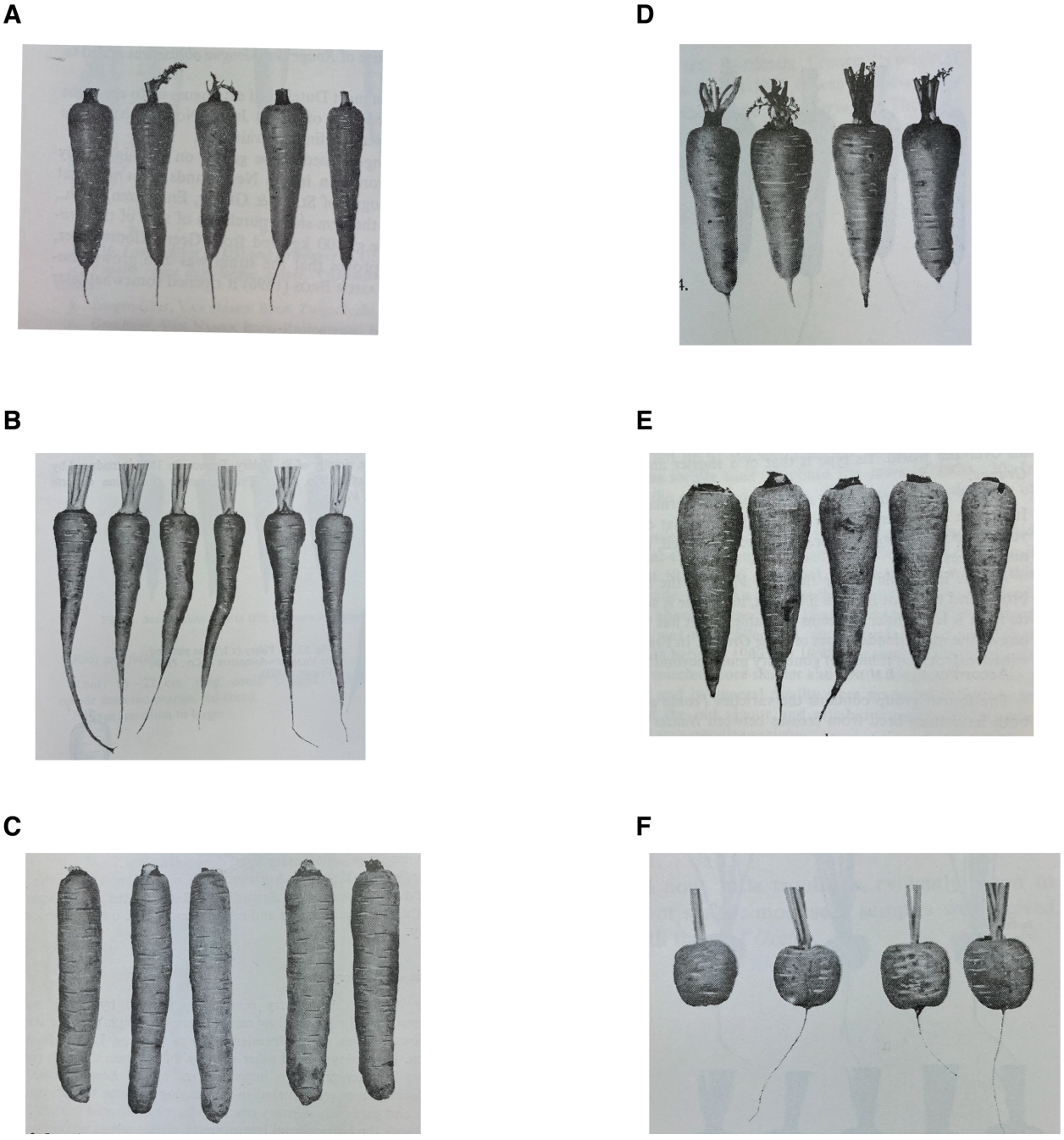


Figure 1. Early documentation of market classes in carrot. (A) Danvers (B) St. Valery (C) Long Stump Winter (D) Grosse Normande (E) Flanders (F) Grelot French Forcing. (Banga, 1957).

Chapter 2:
A high-throughput phenotyping platform for analyzing the genetic control of root shape attributes in carrot

Contributing Authors: Scott H. Brainard, Julian A. Bustamante, Julie C. Dawson, Edgar P. Spalding, Irwin L. Goldman

Author Contributions: SHB, JCD, and ILG designed the experiments. SHB performed the experiments, analyzed the data, and wrote the chapter. SHB, JAB and EPS designed the phenotyping platform. All authors have read and approved the chapter.

Abstract

Carrot (*Daucus carota* subsp. *sativus*) is an economically and nutritionally important vegetable crop. Carrot root shape, which ranges from long and tapered to short and blunt, has been used for at least several centuries to classify carrot cultivars. The subjectivity involved in determining market class hinders the establishment of metric-based standards and is ill-suited to dissecting the genetic basis of such quantitative phenotypes. Advances in digital image acquisition and analysis has enabled new methods for quantifying sizes of plant structures and shapes. In order to dissect the genetic control of the shape features that define market class in carrot, a tool is required that quantifies the specific shape features used by humans in distinguishing between classes. Here we report the construction and demonstration of such a platform. The pipeline includes a series of a preprocessing algorithms whereby color images are converted to binary masks, which are then standardized to remove curvature and residual root hairs. Root phenotyping is then performed, including both traits that are measurable by hand, such as length and width, as well as principal components analysis of the root contour and its curvature. This platform's high-throughput performance and accuracy was validated in two experimental panels: a diverse, global collection of germplasm was used to assess its capacity to identify market classes through clustering analysis, and diallel mating design between inbred breeding lines of differing market classes was used to estimate the heritability of the key phenotypes that define market class.

Introduction

Carrot (*Daucus carota* subsp. *sativus*) is an economically and nutritionally important vegetable crop. Over 40 million metric tons of carrots are grown annually across the globe (FAO, 2020). The 33,000 hectares of carrot cultivated annually in the United States generates a net farm gate value of over $730 million USD (United States Department of Agriculture, 2020) and is a significant source of pro-vitamin A in the human diet (Simon, 2000).

Carrot root shape, which ranges from long and tapered to short and blunt, has been used for at least several centuries to classify carrot cultivars (Banga, 1957, Simon et al., 2008). Culinary practices and horticultural traditions have led to the establishment of several modern market classes, which are based on use, size, and shape differences. For example, cultivars that produce large and bulky roots grown for a full season are typically used in canning, freezing, dehydrating, juicing, and other processing operations. These carrots are represented by market classes such as Danvers, Chantenay, and Berlicum, and are often cultivated at 500,000-1,000,000 plants per hectare. Fresh market types, represented by market classes such as Imperator, Kuroda, and Nantes, are grown at a significantly higher density (1,500,000 - 3,000,000 plants per hectare) due to their slimmer root profile.

The shape differences that are key to proper classification of roots range from obvious to very subtle. Ordinal scales, or adjectival descriptors such as circular, obovate, obtriangular, and narrow oblong are frequently used by breeders. The distinctions between market classes in some cases depends on subjective assessment of the curvature of the shoulder at the crown of the root, the variable filling of the tip, and specific combinations of these shape parameters and

root size. Furthermore, while the names of market classes would imply an assignment of cultivars to well-defined, and discrete categories, all of the traits which define market class are quantitative, and intermediate cultivars may possess characteristics of more than one class. The subjectivity involved in determining market class thus hinders the establishment of metric-based standards and are ill-suited to dissecting the genetic basis of such quantitative phenotypes.

Recent advances in measuring plant phenotypes by analyzing digital images has led to the development of methods for quantifying sizes and shapes of plant structures, including those that are not captured by simple angles, lengths, widths, and their ratios. Often, the aim of such algorithms is automated classification, e.g., distinguishing between crop and weed species (Le et al., 2020) or performing leaf segmentation (Kumar and Domnic, 2019; Victorino and Gómez, 2019). Similar pipelines have also been developed that seek to classify vegetables into cultivar types (Hameed et al., 2018; Visa et al., 2014), including in carrots (Horgan, 2001; Horgan et al., 2001; Howarth et al., 1992). However, the features of the contour utilized in such machine-learning applications are distinct from the specific shape features used by humans in distinguishing between market classes. In order to dissect the genetic control of the shape features that define a class, a tool that quantifies rather than classifies is required.

In this respect, contour analysis has also been used in many species, e.g., to quantify maize ears and kernels (Miller et al., 2017), tree leaves (Biot et al., 2016), and radish roots (Iwata et al., 1998). A previous study demonstrated that contour analysis of carrot roots could produce

shape metrics that served as genetically-mappable phenotypes (Turner et al., 2018). image analysis pipelines measure their target phenotypes but not other, even related, traits. For example, while the work of Turner et al. (2018) addressed a phenotyping challenge in carrots, their platform lacked the capacity to quantify variation in the root shoulder and tip, which are key determinants of market class in carrot. In addition, their pipeline was designed to study linkage mapping populations, and as such, did not need to include pre-processing standardization steps that are necessary when evaluating highly diverse collections of germplasm. In order to use digital phenotyping methods to investigate market-class related traits, therefore, an image analysis platform is needed that can quantify these particular shape features, and be able to operate with sufficient throughput to make a large-scale study of the genetic determinants of market class feasible.

Here we report the construction and demonstration of such a platform, and validate its performance in two experimental populations. First, a global collection of carrot accessions obtained from the USDA National Plant Germplasm System (NPGS) was used to test the performance of the phenotyping algorithms in phenotypically diverse context. This analysis confirmed that the phenotypes generated by this pipeline allow for precise quantification of the key features of market class in carrot. In addition, a diallel mating design was constructed to evaluate the genetic control of the carrot root phenotypes measured by this platform. A set of diallel crosses evaluates the progeny derived from pairwise matings between a defined set of inbred parental lines in an effort to estimate genetic variance components that are of significant import in both breeding and population genetics (Griffing, 1956). This experimental

design allowed for an estimation of the heritability of the various component traits which constitute market class – the first reported estimation of such parameters for these important morphological traits. As such, these analyses provided proof of concept datasets which confirmed the value of this phenotyping pipeline as a genetic research tool.

Materials and Methods

The digital image-based phenotyping methodology developed in this study followed a three-stage workflow: image acquisition, image pre-processing, and image analysis.

Image acquisition

Images were captured with a DSLR camera with a 24mm fixed-length lens, mounted above a template containing two 22.5 cm x 75 cm black-bordered rectangles. This facilitated imaging two roots simultaneously; further partitioning would permit multiplexing more than two roots in a single image, depending on root size and desired resolution. Six fluorescent Interfit (Atlanta, GA) F5 lights provided overhead illumination in order to maximize contrast between root and background, and eliminate shadow. Each rectangle was divided into an upper and lower portion by blue, 1.25 cm in Gaffer's tape. The upper portion contained a 100 mm scale bar, and a QR matrix barcode which encoded identifying information pertaining to the specific carrot root being photographed. The lower portion contained the corresponding root, placed on either a black felt or white vinyl background depending on the exterior pigmentation of the carrot. The top of the root was aligned to be parallel with the blue tape in order to precisely divide root and shoot growth.

The DSRL camera was connected via a USB cable to a computer running SmartShooter (Tether Tools, Phoenix, AZ) a tethered shooting application which allows for high-resolution live previewing of DSLR cameras' view-frames. This facilitated accurate positioning of carrots relative to the blue line. Upon image acquisition, SmartShooter wrote a raw and a lossless JPG image to disc at a user-specified "source" location.

A Python application then handled initial image processing and file management. This application ran inside of a custom Python v3.7.7 virtual environment in order to easily utilize a suite of open-source image-processing libraries. First, the `watchdog` library was used to detect each new JPG as it was created, and the `lensfunpy` wrapper for the C++ library `lensfun` was used to remove distortion due to the curvature of the lens (Fig. 1A). Next, each black-bordered box within the image was identified using the Python bindings for the `OpenCV` library, and the QR code within the upper portion of each box was scanned using the bindings for the `zbar` library (Fig. 1B). As a preliminary form of quality control, the `click` package was used to display the attribute-value pairs encoded by the QR code; when the user accepted these as accurate, the corresponding image was subsequently displayed within a browser window using Node.js, with a transparent overlay of the region detected as corresponding to the carrot root. This overlay was also generated using the `OpenCV` library: in brief, RGB images were converted to grayscale, a bilateral filter was applied to smooth the image while preserving edges, and a binary threshold was applied to generate the so-called "binary mask"—i.e., a black and white image in which white pixels designate the presence of carrot, and black pixels designate

background (Fig. 1C). If binary masks were visually judged to correctly identify the root, they were then saved as high-resolution PNG files to a pre-specified "destination" directory path defined according to information encoded in each QR code (e.g., "Location/Year/Genotype"). All identifying information specific to each root was included in the PNG file name, in addition to a scaling parameter corresponding to the detected pixel length of the 100 mm scale bar.

Image pre-processing

Following image acquisition, pre-processing steps were performed to standardize image, specifically by removing curvature and residual, unexpanded root tips. The straightening procedure was performed in MATLAB R2019b (MATLAB, 2019), and consisted of two stages. First, the midline of the carrot root was estimated, and second, widths were determined by "slicing" the binary mask normal to the tangent vectors along this midline. To estimate the midline, the minimum distance between all points within the carrot root and contour of the carrot root was determined using the Euclidean distance-transform described by Maurer et al. (2003). Next, the tip of the carrot was found via an iterative algorithm that identifies points of maximum curvature within increasingly narrow segments of the root contour. Finally, following the procedure described by Miller et al., (2007), the midline was traced: starting at the tip, an ordered series of midline coordinates was sequentially found by identifying the local maximum in the smoothed distance-transform, stepping in the direction of this local maximum, and then repeating the procedure. This "walk" along the maximum of the smoothed distance transform surface stably traces the midline, under the condition that all midlines must end by passing through the center of the shoulder of the carrot (Fig. 1D).

In the second stage, the width of the carrot along its length was calculated. Regardless of any degree of local skew or curvature in the root, the "true" width of the carrot at any point follows the vector normal to the tangent of the midline at that point. Thus, the binary mask was sampled along this vector, starting at the midline and moving outward in both directions until an intersection with the contour is reached (Fig. 1E). This method of sampling prevented the inclusion of multiple segments of the carrot root in a single slice through its width, in cases where the tip of the carrot may curl back upon itself. A straightened version of the carrot was then obtained by arranging these widths into a single array (or "width vector"), all centered on their midpoints.

After straightening, a secondary pre-processing step was taken to remove any trailing, unexpanded portion of the taproot which extends past the tip of the carrot (Fig. 1F-G). Any geometric definition of this cut point (e.g., on the basis of the derivative of width along the length of the root) is hindered due to extensive variation in tip shape and tip fill.
A random forest classifier was therefore trained (using the `sklearn` Python library) with a subset of the images that were photographed both with their residual taproot attached, and with the taproot removed. This model was subsequently used to detect the appropriate point at which to "de-tip" the straightened versions of the binary masks.

Image analysis

Following these pre-processing steps, images were analyzed for several different types of phenotypes. First, phenotypes that could be measured by hand, such as length, maximum width, and width at different points along the length of the carrot, were calculated for each image. These were determined by simply measuring the length of different slices through the straightened, de-tipped binary mask, and converting pixels to mm using the appropriate scaling parameter.

Second, traits pertaining to individual roots that would be extremely difficult to measure by hand, such as tip angle, convex hull area of the shoulder, and curvature values of the shoulder and tip were also calculated. Tip angle was defined here as the interior angle formed by the line segments connecting the tip of the carrot to contour points located 10% up the length of the carrot toward its top, while shoulder hull area was the area encompassed by background pixels in the rectangular region bounding the top 10% of the carrot. Curvature values were computed at each point along the root profile in both the shoulder and tip regions (the first and last 50 contour points of the root, respectively) as described by Driscoll et al. (2012). In principle, since curvature is inversely proportional to the radius of such a circle, curvature values at any particular boundary point can be estimated by inscribing a circle within the root using splines fitted to segments of the contour surrounding this boundary point, for which derivatives can be calculated. Curvature was then computed as:

$$K = \left\{\left(\frac{dx}{dt}\frac{d^2y}{dt^2} - \frac{dy}{dt}\frac{d^2x}{dt^2}\right)\left(\left[\frac{dx}{dt}\right]^2 + \left[\frac{dy}{dt}\right]^2\right)\right\}^{-\frac{3}{2}}$$

These vectors of curvature values in the shoulder and tip region can then be summed, to generate a metric of total curvature, or decomposed into principal components using PCA to identify sources of variation in curvature values across a wide sampling of carrots.

Finally, in order to identify and quantify size-independent variation in the shape of the entire carrot root, PCA was also used on the raw contour data of the entire width profile. In this case, the relevant covariance matrix was constructed using straightened, de-tipped masks that were first normalized such that all roots possessed equal lengths and maximum widths. Each root was represented by 1000 standardized widths sampled evenly along the carrot's length, and each width along a carrot's length was divided by its maximum width, such that each carrot had a maximum width of 1.

Validation of accuracy and reliability

To validate the accuracy of the pre-processing pipeline and phenotyping algorithms, 100 roots (10 each drawn from 10 carrot genotypes representing divergent market class types) were measured by hand prior to being photographed. Length measurements were made from the top of the shoulder to the point at which the unexpanded, residual tip of the carrot began using a tape measure, while maximum width was measured using calipers. Secondly, to determine the extent of the variance in phenotypes extracted from different 2D projections of a given root, 100 roots were photographed three times, with each root being rotated 45° following the acquisition of the first and second photograph.

Visualization of the phenotypic correlates of principal components analysis

In order to evaluate the performance of the principal components analyses, a diverse collection of carrot germplasm was grown at the Desert Research and Extension Center (Holtville, CA, U.S.) in 2019. This collection was obtained from the USDA National Plant Germplasm System, and consisted of a total of 683 cultivated accessions, comprising breeding lines, open-pollinated cultivars, and land races. These accessions represent a substantial amount of the global diversity in cultivated carrot, and thus provided an excellent basis for visualizing the variation captured by the first principal components of the normalized width profile, as well as the curvature in the tip and shoulder regions. Two replications of this collection were grown in 1-m plots, with 5-15 roots sampled randomly from each plot at harvest; in total, 8687 roots were imaged in this analysis. Phenotype values for each genotype were then estimated with a fixed effects model which modeled PC scores as a linear function of genotype and replicate.

Experimental design of half-diallel

To demonstrate the utility of the phenotyping methodologies described here in the genetic analysis of root morphology, a half-diallel mating design was used to determine the heritability of digitally-phenotyped traits. Seven inbred carrot lines and one open-pollinated variety were used as parents in this diallel. Two inbred lines (B2566 and L1408) were developed by the USDA-ARS Vegetable Crops Research Unit, which breeds primarily for the fresh market carrot industries (Simon et al., 1987); five inbreds (W279, W289, W287, W278 and W280) were produced by the University of Wisconsin-Madison carrot breeding program, which breeds primarily for processing (canning and juicing) industries (Goldman, 1996); the open-source

variety OSSI-Ball is a Parisienne-type for specialty markets (Luby and Goldman, 2016). Together, these eight accessions represent a diversity of the market classes, and consequently the root shapes, that appear in public sector carrot breeding in the U.S. (Supp. Fig. 3). An additional benefit of using these inbred lines, is that both sterile and fertile versions, relying on a cytoplasmic-genic male sterility system (Eisa and Wallace, 1969; Peterson and Simon, 1986), were available, greatly aiding in the logistics of performing the requisite crosses. In the case of OSSI-Ball, only fertile roots were available, and thus this genotype was used as a male parent in each cross. A representative cross between an inbred line used for Imperator production (L1408) and a specialty market type (OSSI-Ball) illustrates the typical manner in which root size and shape phenotypes combined in crosses between divergent market class types (Supp. Fig. 4).

L1408 and B2566 were grown at Miller Farms in Hancock, WI (44°08’N, 89°32'W), and all other genotypes at Jack’s Pride Farm in Randolph, WI (43°37’N, 89°00’W) in the summers of 2017 and 2018. Harvested roots were vernalized at 4°C for 12 weeks with shoot growth removed before being planted in pots (15.2 cm tall x 13.8 cm in diameter, filled with a blend of one-third Pro-Mix High Porosity (Premier Tech, Quakertown, PA) and two-thirds field soil). These vernalized roots were grown under at 20°C and 16 h daylength at the Walnut Street Greenhouses in Madison, WI. Following flowering, pairs of roots—one sterile, one fertile—were selected for crossing: immediately following the appearance of anthers, the umbels of each pair were enclosed in a cloth bags, to which blue bottle fly pupae were added at weekly intervals to ensure high rates of pollination. This process was carried out during both the winter of 2017-

2018 and 2018-2019 in order to obtain progeny from each of the 28 pairwise cross combinations.

Following seed-set, umbels were separated, and seed harvested by hand from the sterile parent. F_1 seed from each cross was grown in two replicated 3m plots at Jack's Pride Farm in the summer of 2019. Roots were thinned to a density of 1 seedling per 5 cm 21 days after planting (DAP), and roots from each plot were harvested from the middle of each row 107 DAP. The logistical complexity of obtaining pairwise crosses between a large number of inbred parents with differing flowering times limited the number of parents to the eight used in this study. Nevertheless, such an experimental design (of eight parents with two replicates) falls within the range recommended by Pederson (1971) on the basis of simulation studies, and a large sample size within each replicate of F_1 progeny (15-20 roots) was obtained to minimize bias as much as possible in the estimation of variances.

Estimation of heritabilities

For each of the two replicates of every F_1 family, 15-20 roots were randomly selected for digital phenotyping. Because neither reciprocal crosses nor parental lines were included along with the F_1 progeny, the Method IV, Model I diallel analysis was utilized, as described by Griffing (1956): $y_{ijk} = \mu + g_i + g_j + s_{ij} + \varepsilon_{ijk}$. μ is the population mean, g_i is the so-called "general combining ability (GCA) effect of the i^{th} parent, g_j is the GCA of the j^{th} parent, s_{ij} is the "specific combining ability" (SCA) effect of the ij^{th} cross (where $s_{ij} = s_{ji}$), and ε_{ijk} is the residual error of

the *ijk*th observation. All of these terms, including the effect of replicate, are considered as fixed effects.

Variance components for GCA, SCA, and error terms were calculated on the basis of expected mean squares: $EMS_{GCA} = V_e + V_{SCA} + (p-2)V_{GCA}$; $EMS_{SCA} = V_e + V_{SCA}$; $EMS(\varepsilon) = V_e$. p is equal to the number of parental lines, in this case 8, and V_e is equal to the error calculated on an entry-mean basis ($\frac{\sigma_\varepsilon^2}{r}$), with r, the number of replicates, equal to 2. Additive (V_A) and dominance (V_D) components were subsequently estimated as described by Pederson, 1971: $V_A = \frac{4}{1+F} V_{GCA}$ and $V_D = \frac{4}{(1+F)^2} V_{SCA}$.

Broad- (H^2) and narrow-sense (h^2) heritabilities were finally calculated as: $H^2 = \frac{V_G}{V_G + \frac{V_\varepsilon}{r}}$ and $h^2 = \frac{V_A}{V_G + \frac{V_\varepsilon}{r}}$ with $r = 2$ replications; following Falconer (1996), total genotypic variance, V_G, was defined as $V_A + V_D$. Baker's ratio was calculated as: $\frac{2MS_{GCA}}{2MS_{GCA} + MS_{SCA}}$; as such, it varies between 0 (in the case of all variance being attributed to SCA) to 1 (in the case of all variance being attributed to GCA) (Baker, 1978).

Data statement

Python code for the image acquisition platform and scripts for producing binary masks are available at https://github.com/shbrainard/carrot-phenotyping. MATLAB algorithms for straightening binary masks and performing PCA on contours or curvature values are available at: https://github.com/jbustamante35/carrotsweeper. Phenotypic data and example R code (R

Core Team, 2020) pertaining to both the market class-based clustering analysis and the diallel analysis are available via the Harvard Dataverse (Brainard, 2020).

Results

Figure 1 depicts the digital image-based phenotyping workflow. An example raw image, in this case containing two carrots and their respective matrix-barcoded identifying information, is shown in Figure 1A. The multiplexed main image is divided into single-sample sub-images. The regions of interest are the machine-readable QR code data, and the root below the line of blue tape (Fig. 1B). The former is used to name, save, and track image files and phenotypes, while the latter is converted to a binary mask (Fig. 1C). Figure 1D-G illustrate image pre-processing steps: first a midline that terminates at the tip is generated (Fig. 1D); next, 1000 axial width measurements are made along vectors normal to this midline (Fig. 1E); a straightened root profile is constructed using this midline and resulting widths (Fig. 1F); finally, the image is cropped at an apical position determined by a trained machine learning algorithm (Fig. 1G). Technical aspects of the algorithms used in each step are provided in the Experimental Procedures section.

Accuracy of image-derived phenotypes

Scatter plots of the root phenotypes obtained algorithmically from digital images versus hand measurements confirms that the two methods provide highly similar results (Fig. 2). For both length and maximum width, strong correlations were apparent across the range of phenotypes

measured. Root mean squared error (RMSE; the square root of the squared residuals) of the linear model $y = x$ was 8.30 mm for length and 2.00 mm for width.

In addition to parity between human- and computer-based phenotypes, the variance associated with the arbitrary positioning of a given carrot root beneath the camera was also evaluated. Some variance due to placement should be expected, particularly in estimated widths, because this method analyzes a two-dimensional projection of a three-dimensional root, and no carrot is a perfect cone. Throughput of the pipeline would be reduced if more than one perspective of each root needed to be analyzed. Pairwise comparisons were made between phenotypes extracted from three photos of individual, diverse roots that were variably rotated along their long axis. These comparisons demonstrated that deviation from symmetry about the long axis of the carrot was responsible for minimal variance in maximum width measurements and affected length even less (Supp. Fig. 1). The variation between pairs of images was very similar to the variation between human and algorithmic measurements presented in Fig. 2. These results indicate that a single two-dimensional projection of a three-dimensional root is sufficient to obtain reliable estimates of key morphological phenotypes.

We measured the diverse USDA-NPGS carrot collection to determine how variation due to digital phenotyping error compared with variation between genotypes, which is typically the most relevant criteria from both a genetic and plant breeding perspective. Least significant differences (LSDs) between accessions were calculated for both length and maximum width, with MS_{Error} being estimated by ANOVA utilizing a linear model in which accession was

included as a fixed effect. The LSD values of 35.8 mm for length, and 7.5 mm for maximum width (with $\alpha = 0.05$), indicate that the measurement error associated with this digital phenotyping platform is many times smaller than the detection threshold for distinguishing distinct carrot genotypes from each other.

Principal components analysis of contours

An automated tool for measuring root length and maximum width would advance research into the genetic control of root development, but understanding the genetic variation that underlies variation between market classes requires additional information regarding shape in particular. A method was therefore developed to produce for each root image a set of contour points derived from width measurements made at 1000 points along the main axis. Normalizing each contour data set with respect to the maximum in both the *x* and *y* (width and length) dimensions makes the shape information they contain comparable across diverse cultivated accessions drawn from the USDA-NPGS carrot collection. A total of 8687 images representing 683 accessions were collected and analyzed in this manner to produce a contour data set for each root. Principal components analysis (PCA) of these sets of contour points showed that the first principal component (PC1) explained 84.9% of the total variance in the contours, with PC2 and PC3 explaining only 9.6% and 2.5% of the variance, respectively. The eigenvectors of a given PCA were used to generate simulated root profiles based on artificial PC scores. Figure 3 shows that decreasing PC1 while holding all other PC scores equal to their mean reduces or delays the rate of taper toward the tip, i.e. the root maintains width further along its length

compared to roots with a higher PC1. This trait will therefore be referred to as “root fill” in subsequent analyses.

Principal components analysis of shoulder and tip curvatures

The PCA of whole-root contours presented above did not require pre-determining what aspect of shape contributes most significantly to variation across this population. This is in contrast to deliberately measuring shapes of interest, such as the distribution of curvature in the shoulder and tip regions that humans subjectively consider when distinguishing cultivars from differing market classes. The curvature quantified in this case is the instantaneous rate of change of angle of a vector normal to the contour as the vector moves along the contour. Put another way, curvature at each contour point is the reciprocal of the radius of the circle that is tangent to the contour at that point. After fitting smoothed splines to the top 50 (root shoulder) and bottom 50 (root tip) contour points, curvature at each point was calculated to construct the respective covariance matrices. PC1 of shoulder-region curvature explained 87.3% of the total variation. PC1 of tip-region curvature explained 84.2% of the total variation. Representative examples of roots with PC1 values in the 1st and 99th percentiles for shoulder curvature are shown in Figure 4B,C. The Imperator type (Fig. 4B) has almost no curvature in the shoulder, while the broadly shouldered Parisienne type (Fig. 4C) has much more substantial curvature. Representative examples of roots with PC1 values in the 1st and 99th percentiles for tip curvature are shown in Figure 4E,F. The extremely blunt-tipped Nantes type in Fig. 4E can be contrasted with the highly acuminate Danvers type in Fig. 4F.

Fig. 4A and 4D show correlograms of these curvature-derived metrics, with alternative methods of measuring shoulder and tip shape: PC scores derived from curvature values ("Curvature PCA"), PC scores derive from normalized contours ("Contour PCA"), the sum of curvature values, and either shoulder hull area or tip angle. As would be expected, PC scores derived from curvature values are strongly correlated with the sum of curvatures in the shoulders (r = 0.931) and tips (r = 0.954). Less expected was the poor correlation between curvature PC scores and contour PC scores derived from the same region (r = -0.351 for the shoulder region and r = 0.537 for the tip region). This suggests that the process of fitting smoothed splines to the contours gives rise to a meaningful difference between the phenotypes that are measured by quantifying "variation in the curvature" and "variation in the contour" of the shoulder and tip regions.

In general, it is also clear that for any particular pair of phenotypes, the correlation is stronger in the tip region, compared to the shoulders. In particular, while tip angle appears to be a moderately accurate surrogate measure for variation in tip curvature, hull area has a relatively weak correlation with shoulder curvature. This is understandable, since broadly shouldered carrots and carrots completely lacking shoulder curvature should have large hull areas, whereas tip angle varies monotonically with the curvature of the tip.

Cluster analysis of representative carrot market class types

To test the discriminatory power of the phenotypes quantified by this new pipeline, a clustering analysis was performed using length, maximum width, root fill, curvatures of the tip and

shoulder, and aspect ratio (i.e., length divided by maximum width). Thirty five roots were drawn from each of five economically important and phenotypically diverse market classes. The market classes were: Chantenay, a short, bulky processing type; Imperator, a long and slender type used in baby carrot production; Danvers, a medium length, pointed type typical for fresh-market sale; Nantes, a medium length blunt type often used as a storage root; and Ball, a very short, rounded type often sold to specialty markets (Supp. Fig. 2). In addition to being representative of the most economically important market classes, this set of carrots was selected because they represent diverse combinations of component phenotypes. For instance, while root fill is independent of size (being calculated from length- and width-normalized contours), in a particular collection of carrots it may by predictive of length, due to non-random associations between shape and size. While not every combination of phenotypes can be sampled in only five market classes, the classes selected here were therefore chosen in order to minimize such correlations. The results of this clustering analysis are shown in Fig. 5B. In this bi-plot visualization, PCA was used to generate linear combinations of the six traits, and the first two PCs were used as the *x*- and *y*- axes to visually illustrate both the degree of clustering within classes, and the phenotypic "distance" between market classes. Clear clustering is observed for all five market classes, indicating that this phenotyping pipeline effectively characterizes the key phenotypic components of market classes. This can be compared to a clustering analysis which used only those two traits most readily measured by hand, length and maximum width (Fig. 5A). While the most phenotypically divergent roots are still distinguishable (e.g., Imperator and Ball), the exclusion of shape descriptors markedly increases the overlap between market classes that are similar in their overall dimension (e.g.,

Nantes and Danvers, which differ primarily in their degree of tip fill. This confirms that the phenotyping platform measures shape parameters that parallel the morphological differences between market classes.

Diallel mating design

All pair-wise crosses were made between eight parental lines to construct a half-diallel population. 760 F1 roots representing this population were measured with the pipeline described above. Mean length was 18.12 cm, with a standard deviation of 4.46 cm; mean maximum width was 4.72 cm with a standard deviation of 0.80 cm; and mean L/W ratio was 4.01 with a standard deviation of 1.34. For all traits, MS_{GCA} was larger than MS_{SCA}, although the degree to which this was the case varied substantially from phenotype to phenotype, from over 15x greater in the case of length-width ratio, to only 1.43x times greater in the case of root biomass (Table 1). This finding is captured well in Baker's ratio: in general, values were found to be close to unity, suggesting meaningful degrees of additive gene action for all traits considered.

Broad-sense heritability values were ≥ 0.94 for all traits save biomass and tip curvature. This somewhat surprising finding indicates a high degree of genetic influence over phenotypes that are exposed to a high degree of environmental influence, due to the roots' direct contact with the inherently heterogeneous soil profile. In this regard, two factors should be kept in mind: first, the soils in which this trial was grown is a Houghton muck, a deep saprist histosol with more than 50% organic matter. Given the aggressive tillage prior to seed bed preparation, this

leads to a highly uniform soil profile with minimal compaction. Together with conventional weed control and fertilizer application, this produces one of them most uniform environments for growing root crops in Wisconsin. Thus, it is relatively unsurprising that environmental variation was minimized, and heritability maximized, in such a production system. Secondly, the fact that seven of the eight parents in this diallel were inbred lines likely contributed to phenotypic uniformity within full-sib families due to the genetic uniformity within each particular hybrid combination.

Narrow sense heritabilities displayed a wide range of values, from 0.14 for biomass, to 0.84 for L/W ratio. Length, maximum width, and shoulder curvature all exhibited middling values (0.66, 0.72, 0.76, respectively). In general, these values conform to the experience of modern breeding for both the fresh market and processing industries, which have successfully produced highly typified long and slender carrots for the Imperator market, and much shorter, broader, heavily tapered carrots for the processing industries.

It is interesting to consider why the biomass and aspect ratio – both traits that would intuitively be understood as primarily functions of the overall dimensions of the root (length and maximum width) – possess such markedly different narrow-sense heritabilities. While aspect ratio is primarily the function of two highly heritable traits (length and width), biomass is a function not only of length and width, but also root fill. In this particular population, the correlation between maximum width and how far down the carrot that width is maintained

(e.g., PC1) is very weak: only 0.12. That is, whether a carrot has a very wide shoulder or narrow shoulder is not highly correlated with widths elsewhere along the carrot.

Discussion

Market class in carrot

It is challenging to define market class in horticultural crops like carrot, which is a composite of several closely related morphological traits, with more precision than shown in the sketches in Supp. Fig. 2. Although it is possible to list a suite of parameters that are clearly involved in defining a given root class, quantifying them in relation to each other, integrating them into a single metric to determine which classes are most similar to each other, or evaluating how much variation exists within a given class: all of these tasks are impossible without a robust set of digital phenotyping tools, such as those described here. This is perhaps best exemplified in the first principal component of the straightened, length- and width-normalized contours – referred to here as "root fill". This trait gives breeders and researchers a method for visualizing and quantifying a source of shape variation that would otherwise be confounded with root size. Finally, this platform is not designed around a machine learning algorithm for classifying carrots into predefined – and therefore static – market classes. As such, it can be adapted to any range of root shapes. This is particularly useful for a character like market class, which is determined by current agricultural practices and culinary preferences. While these classes are therefore malleable, and will certainly change over time, their component phenotypes will still be quantifiable by way of the pipeline described here.

In order to answer the question as to whether these phenotypes together accurately describe the market classes which predominate today, we performed a PCA-based clustering analysis that clearly identified clusters corresponding to five major market classes. In addition, this analysis demonstrated the degree to which significant phenotypic diversity exists within market classes; even though individuals from the same market class cluster together, at the border regions of each cluster there is limited overlap between classes. This is unsurprising, since all of the traits composing market class are quantitative in nature, and furthermore since most market classes are descended recently from a relatively narrow genetic base (Banga, 1957).

Trait heritabilities in a diallel mating design

To demonstrate its potential utility in terms of genetic analysis that is particularly relevant from a breeding perspective, this study presents the first diallel experiment in carrot composed of parental inbred lines drawn from the predominant U.S. market classes, that simultaneously evaluates both size and size-independent shape parameters. On a trait-by-trait basis, this aims to provide an understanding of the predominant form of gene action influencing these traits (additive versus dominance variance). Taken together, these results will guide breeding decisions in programs that are utilizing these inbreds to generate hybrid progeny.

From a practical perspective, the relatively high narrow-sense heritabilities reported here reflect breeders' success in efficiently selecting for these traits. From a genetic perspective, however, this finding does not necessarily indicate simple control, in the sense these traits are necessarily primarily controlled by only a few, large-effect quantitative trait loci (QTL). Large

additive genetic variance components could also be associated with highly polygenic traits that simply lack dominance and epistatic variance (Huang and Mackay, 2016). Furthermore, the heritabilities reported here only reflect phenotypes measured in a particular – albeit common and important – environment in Wisconsin. As such, they should be understood to represent a "ceiling" for heritabilities: they are quite high, and thus further research studying these traits across multiple years in multiple environments should therefore be encouraged. When this is carried out, estimates of environmental variance will likely rise, and thus heritabilities should be expected to fall. The degree to which multi-environment trials might lead to lower estimates of heritability is not clear, given the fact that the estimates reported here are roughly in accord with values previously estimated for root length (Brar and Sukhija, 1981; Prasad and Prasad, 1980). Additionally, it is important to stress that heritability is a parameter of the population under consideration, and not purely a function of the trait they describe. Estimates of heritability will vary depending on the particular cultivars or population studied, as has been found in carrot with respect to heritabilities for nematode resistance in carrot (Huang et al., 1986; Vieira et al., 2003) as well carotenoid components (Fernandes Santos and Simon, 2006). For this reason, we selected as parents a diverse set of inbred lines, with the aim that heritabilities would likely reflect the germplasm used in current breeding programs. Nevertheless, the precise heritabilities realized in any population not derived from these particular genotypes will necessarily deviate to some degree from those calculated here.

High-throughput phenotyping

This report presents a high-throughput method for phenotyping carrot roots using 2D digital photos. This approach was validated as accurate (compared against human measurements, Fig. 2) and robust to the arbitrary rotation of the 3D root, through comparison with multiple images of the same root (Supp. Fig. 1). Interestingly, while variation about the line $y = x$ in Fig. 2 appears randomly distributed for the case of maximum width, there is a slight upward bias in algorithm-measured lengths for particularly long carrots. This is potentially a consequence of the determination of the point at which the carrot root ends and the unexpanded, residual tip begins. This unexpanded portion of the taproots was removed from images where it existed (as shown in Fig. 1F,G), since it is not relevant to defining market class. Anatomically, it can be distinguished from the carrot root proper as lacking the secondary cambial layers leading to root thickening (Goldman, 2020). As a result, practically, it is often intentionally or unintentionally broken off during harvest, or post-harvest handling. Because it does not break off all roots, however, it would represent a significant source of residual variation if not controlled for. A difference between human judgment and machine learning methods in determining where this residual tap root begins could have led to a slight systematic bias.

In addition, binary masks are straightened before length measurements are calculated. Carrot roots may possess varying degrees of curvature depending on their orientation in the soil profile and their angle and direction of growth. In addition, orientation of the root parallel to the blue tape within the rectangular template in some cases introduced varying degrees of skewness. Therefore, straightening is essential to minimizing experimental error. However, while the algorithms for straightening the carrot are robust to any arbitrary curvature of a

carrot root, they are of course different from the method by which a human measures carrot length, and therefore likely contribute to some of the variation observed in Fig. 2A. Nevertheless, these results confirm both the accuracy of the digital phenotyping approach, as well as its general accordance with human assessments of root size. As such, this tool has substantial value in any attempt at assessing phenotypic diversity for, or dissecting the genetic control of, morphological traits, as these traits are often impossible or challenging to measure by hand.

Beyond its utility in simply describing phenotypic variability, the automated potential of such a phenotyping approach means it has substantial promise within plant breeding programs. In plant breeding, a significant bottleneck remains the resources required to screen large populations, whether these be time, money, or labor. From this perspective, the pre-processing and image analysis stages of this pipeline are already explicitly automated. Furthermore, the acquisition algorithms described here are robust to virtually any object; only the specific RGB thresholding value used to distinguish plant tissue from background requires adjustment to allow for the accurate production of a binary mask. With regards to the pre-processing and analysis algorithms, the only requirement to their broad application is that the object be non-branching, and within horticultural and agronomic crops this would allow for a wide array of fruit, roots and tubers.

In the development of this phenotyping approach, three main classes of phenotypes were measured. First, concrete phenotypes that could easily be measured by hand were calculated

(length, maximum width, aspect ratio). Second, phenotypes that could not easily be reliably calculated by hand, but are often used as subjectively-judged determinants of market class, were measured (tip fill and shoulder curvatures). Finally, PCA of the entire size-normalized contour (root fill), as well as PCA of curvature values in the shoulders and tip, was performed in order to more objectively quantity sources of variation in carrot shape. This PCA-based analysis of root shape mirrors closely previous studies performed in Japanese radish (*Raphanus sativa*, L.), which used elliptical Fourier descriptors to quantify contour variation of normalized binary masks (Iwata et al., 2000; Iwata et al., 1998). There were some important differences between the two analyses, however: the panel of radish roots possessed much less diversity than the collection presented here, and the images were not straightened prior to contour analysis. Nevertheless, as in this analysis, the first principal component captured the vast majority of the total variance in radish shape (73.9%, compared to 84.9% found in this study). Interestingly, however, the variation did not describe root fill, but rather the thickening in the central portion of the root, which likely reflects the different source of variance produced through selection for carrot- vs radish-specific market class traits.

In this regard, it is important to note that botanically, carrots – like many storage roots – are formed through a swelling of the taproot, driven by the production of supernumerary cambia, i.e., secondary growth characterized by the production and expansion of additional xylem and phloem tissues (Goldman, 2020). As such, the phenotyping methods described here hold significant potential not only in the context of plant breeding, but in understanding plant tissues that have been modified through domestication. Precisely studying variation across carrot

market class therefore represents a unique opportunity to deepen our understanding of the genetic bases of secondary root development in general. Future studies based around linkage mapping populations or association panels could utilize this phenotyping pipeline to identify QTL associated with these underlying physiological processes.

References

Baker, R. J. (1978). Issues in diallel analysis. *Crop Science*, *18*(4), 533–536.

Banga, O. (1957). Origin of the European cultivated carrot. *Euphytica*, *6*(1), 54–63. https://doi.org/10.1007/BF00179518

Biot, E., Cortizo, M., Burguet, J., Kiss, A., Oughou, M., Maugarny-Calès, A., Gonçalves, B., Adroher, B., Andrey, P., Boudaoud, A., & Laufs, P. (2016). Multiscale quantification of morphodynamics: MorphoLeaf software for 2D shape analysis. *Development*, *143*(18), 3417 LP – 3428. https://doi.org/10.1242/dev.134619

Brainard, S. (2020). *Market class clustering and diallel analysis* (V1 ed.). Harvard Dataverse. https://doi.org/doi:10.7910/DVN/LOVXZA

Brar, J. S., & Sukhija, B. S. (1981). Studies on genetic parameters in carrot (Daucus carota L.). *Journal of Research Punjab Agricultural University*, *18*, 287–291.

Driscoll, M. K., McCann, C., Kopace, R., Homan, T., Fourkas, J. T., Parent, C., & Losert, W. (2012). Cell shape dynamics: From waves to migration. *PLoS Computational Biology*, *8*(3), e1002392. https://doi.org/10.1371/journal.pcbi.1002392

Eisa, H. M., & Wallace, D. H. (1969). Morphological and anatomical aspects of petaloidy in the carrot (Daucus carota L.). *J Amer Soc Hort Sci*, *94*, 545–548.

Falconer, D. S. (Douglas S. (1996). *Introduction to quantitative genetics*. Fourth edition. Essex, England : Longman, 1996.

FAO. (2020). *FAOSTAT*. http://www.fao.org/faostat/en/

Fernandes Santos, C. A., & Simon, P. W. (2006). Heritabilities and Minimum Gene Number Estimates of Carrot Carotenoids. *Euphytica*, *151*(1), 79–86.

https://doi.org/10.1007/s10681-006-9130-7

Goldman, I. L. (1996). A list of germplasm releases from the University of Wisconsin carrot breeding program, 1964--1994. *HortScience*, *31*(5), 882–883.

Goldman, I. L. (2020). The Root Vegetables: Beet, Carrot, Parsnip, and Turnip. In *The Physiology of Vegetable Crops* (pp. 399–420). CABI.

Griffing, B. (1956). Concept of General and Specific Combining Ability in Relation to Diallel Crossing Systems. *Australian Journal of Biological Sciences*, *9*(4), 463. https://doi.org/10.1071/bi9560463

Hameed, K., Chai, D., & Rassau, A. (2018). A comprehensive review of fruit and vegetable classification techniques. *Image and Vision Computing*, *80*(0262–8856), 24–44. https://doi.org/10.1016/j.imavis.2018.09.016

Horgan, G. W. (2001). The statistical analysis of plant part appearance — a review. *Computers and Electronics in Agriculture*, *31*(2), 169–190. https://doi.org/https://doi.org/10.1016/S0168-1699(00)00190-3

Horgan, G. W., Talbot, M., & Davey, J. C. (2001). Use of statistical image analysis to discriminate carrot cultivars. *Computers and Electronics in Agriculture*, *31*(2), 191–199. https://doi.org/https://doi.org/10.1016/S0168-1699(00)00191-5

Howarth, M. S., Brandon, J. R., Searcy, S. W., & Kehtarnavaz, N. (1992). Estimation of tip shape for carrot classification by machine vision. *Journal of Agricultural Engineering Research*, *53*, 123–139. https://doi.org/https://doi.org/10.1016/0021-8634(92)80078-7

Huang, S. P., Vecchia, P. T., & Ferreira, P. E. (1986). Varietal Response and Estimates of Heritability of Resistance to Meloidogyne javanica in Carrots. *Journal of Nematology*, *18*(4),

496–501. https://pubmed.ncbi.nlm.nih.gov/19294218

Huang, W., & Mackay, T. F. C. (2016). The Genetic Architecture of Quantitative Traits Cannot Be Inferred from Variance Component Analysis. *PLoS Genetics*, *12*(11), e1006421–e1006421. https://doi.org/10.1371/journal.pgen.1006421

Iwata, H., Niikura, S., Matsuura, S., Takano, Y., & Ukai, Y. (1998). Evaluation of variation of root shape of Japanese radish (Raphanus sativus L.) based on image analysis using elliptic Fourier descriptors. *Euphytica*, *102*(2), 143–149. https://doi.org/10.1023/A:1018392531226

Iwata, Hiroshi, Niikura, S., Matsuura, S., Takano, Y., & Ukai, Y. (2000). Diallel Analysis of Root Shape of Japanese Radish (Raphanus sativus L.) Based on Elliptic Fourier Descriptors. *Breeding Science*, *50*(2), 73–80. https://doi.org/10.1270/jsbbs.50.73

Kumar, J. P., & Domnic, S. (2019). Image based leaf segmentation and counting in rosette plants. *Information Processing in Agriculture*, *6*(2), 233–246. https://doi.org/https://doi.org/10.1016/j.inpa.2018.09.005

Le, V. N. T., Ahderom, S., Apopei, B., & Alameh, K. (2020). A novel method for detecting morphologically similar crops and weeds based on the combination of contour masks and filtered Local Binary Pattern operators. *GigaScience*, *9*(3), 1–16. https://doi.org/10.1093/gigascience/giaa017

Luby, C. H., & Goldman, I. L. (2016). Release of Eight Open Source Carrot (Daucus carota var. sativa) Composite Populations Developed under Organic Conditions. *HortScience*, *51*(4), 448–450.

MATLAB. (2019). *Version 9.6.0.1072779 (R2019a)*. The MathWorks Inc.

Maurer, C. R., Qi, R., & Raghavan, V. (2003). A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(2), 265–270. https://doi.org/10.1109/TPAMI.2003.1177156

Miller, N. D., Haase, N. J., Lee, J., Kaeppler, S. M., de Leon, N., & Spalding, E. P. (2017). A robust, high-throughput method for computing maize ear, cob, and kernel attributes automatically from images. *The Plant Journal*, *89*(1), 169–178. https://doi.org/10.1111/tpj.13320

Miller, N. D., Parks, B. M., & Spalding, E. P. (2007). Computer-vision analysis of seedling responses to light and gravity. *The Plant Journal*, *52*(2), 374–381. https://doi.org/10.1111/j.1365-313X.2007.03237.x

Pederson, D. G. (1971). The estimation of heritability and degree of dominance from a diallel cross. *Heredity*, *27*(2), 247–264. https://doi.org/10.1038/hdy.1971.88

Peterson, C. E., & Simon, P. W. (1986). Carrot Breeding. In M. J. Bassett (Ed.), *Breeding Vegetable Crops*. AVI Publishing Company, Inc.

Prasad, A., & Prasad, L. (1980). Genotypic and phenotypic variability in a collection of carrot varieties. *Progressive Horticulture*, *11*, 21–25.

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Simon, P. W. (2000). Domestication, Historical Development, and Modern Breeding of Carrot. In *Plant Breeding Reviews* (pp. 157–190). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470650172.ch5

Simon, P. W., Freeman, R. E., Vieira, J. V, Boiteux, L. S., & Briard, M. (2008). Carrot. In *Vegetables II: Fabaceae, Liliaceae, Solanaceae, and Umbelliferae* (pp. 327–357). Springer-Verlag New York. https://doi.org/10.1007/978-0-387-74110-9

Simon, P. W., Peterson, C. E., Bassett, M. J., Strandberg, J. O., White, J. M., & Rubatzky, V. E. (1987). B2566 carrot inbred. *HortScience*, *22*(2), 327.

Turner, S. D., Senalik, D. A., Simon, P. W., Ellison, S. L., Miller, N. D., & Spalding, E. P. (2018). An Automated Image Analysis Pipeline Enables Genetic Studies of Shoot and Root Morphology in Carrot (Daucus carota L.). *Frontiers in Plant Science*, *9*(1703), 1–17. https://doi.org/10.3389/fpls.2018.01703

United States Department of Agriculture. (2020). *National Agricultural Statistics Service*. https://www.nass.usda.gov/Statistics_by_Subject/

Victorino, J., & Gómez, F. (2019). Contour analysis for interpretable leaf shape category discovery. *Plant Methods*, *15*(1), 1–12. https://doi.org/10.1186/s13007-019-0497-6

Vieira, J. V, Charchar, J. M., Aragão, F. A. S., & Boiteux, L. S. (2003). Heritability and gain from selection for field resistance against multiple root-knot nematode species (Meloidogyne incognita race 1 and M. javanica) in carrot. *Euphytica*, *130*(1), 11–16. https://doi.org/10.1023/A:1022372410941

Visa, S., Cao, C., Gardener, B. M., & van der Knaap, E. (2014). Modeling of tomato fruits into nine shape categories using elliptic fourier shape modeling and Bayesian classification of contour morphometric data. *Euphytica*, *200*(3), 429–439. https://doi.org/10.1007/s10681-014-1179-0

Tables

Table 1. Parameters estimated from ANOVA of the half-diallel using Griffing's Method IV, Model 1. Variance components and heritabilities are reported for the primary size and shape traits which define market class.

	Length (mm)	**Width (mm)**	**L/W Ratio**	**Biomass (m^2)**	**Root fill (PC score)**	**Tip curvature (PC score)**	**Shoulder curvature (PC score)**
$\mathbf{MS_{GCA}}^{\dagger}$	5200.03	175.14	5.84	2.79	1.77	1.18e-04	5.16e-05
$\mathbf{V_A}$	1445.25	50.36	1.82	0.28	0.37	1.57e-05	1.54e-05
$\mathbf{MS_{SCA}}^{\dagger}$	864.29	24.05	0.38	1.94	0.67	7.28e-05	5.41e-06
$\mathbf{V_D}$	651.05	15.94	0.33	1.43	0.55	1.83e-05	4.30e-06
$\mathbf{V_\varepsilon}$	213.24	8.11	0.06	0.51	0.12	5.45e-05	1.07e-06
$\boldsymbol{H^2}$	0.95	0.94	0.98	0.87	0.94	0.55	0.97
$\boldsymbol{h^2}$	0.66	0.72	0.84	0.14	0.38	0.26	0.76
$\boldsymbol{BR}^{\ddagger}$	0.92	0.94	0.97	0.74	0.84	0.77	0.95

†Mean sum of squares from the ANOVA

‡Baker's Ratio

Figures

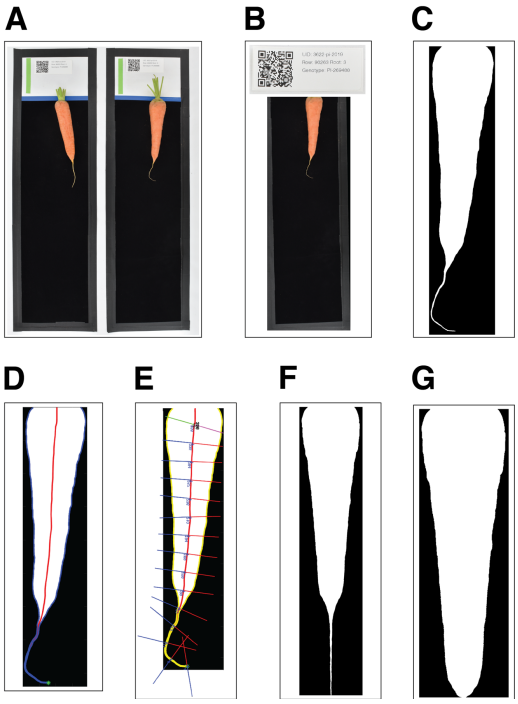


Figure 1. Workflow outlining the pre-processing of digital images of carrot roots. **(A):** Each black-bordered box within the overall image was identified; **(B)** QR codes within the upper portion of each box were scanned and the encoded text displayed as a form of quality control; **(C)** Carrot pixels were distinguished from background pixels to generate binary masks; **(D)** The midline of the carrot root was estimated by tracing a path from the carrot tip to the center of the shoulder, following the maximum of the smoothed Euclidean distance transform; **(E)** Width measurements were made by sampling the binary mask normal to vectors tangent to the midline; **(F-G)** A random forest classifier was trained to detect the point at which to "de-tip" any residual, unexpanded tap root.

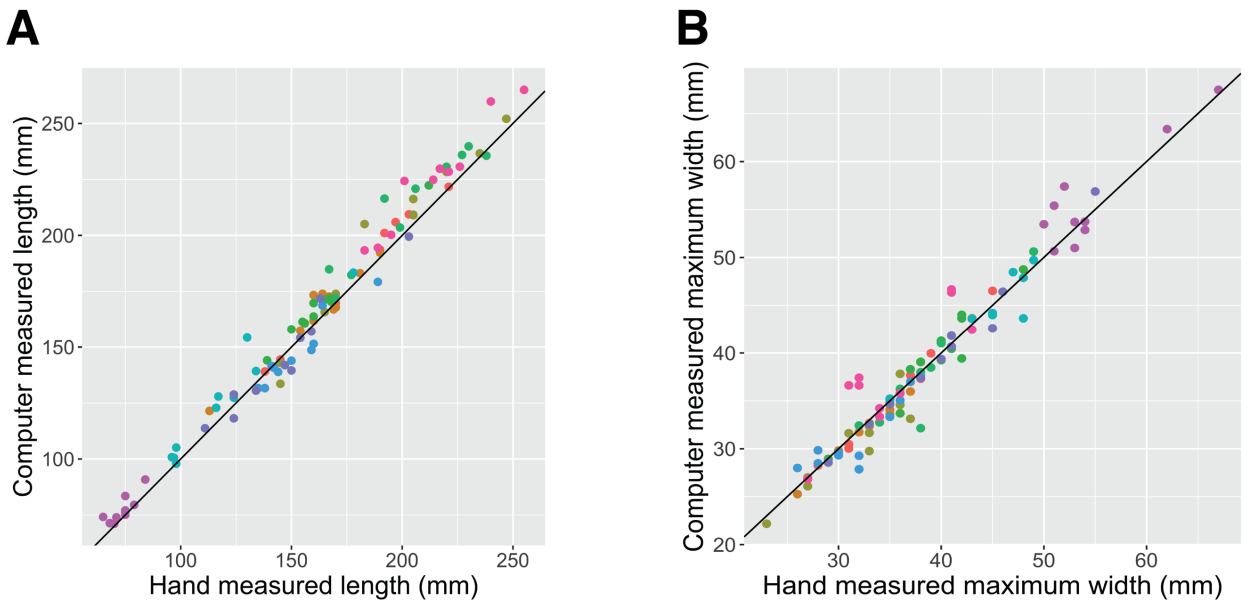


Figure 2. Hand- (*x*-axis) versus digital-image based (*y*-axis) measurements of length (**A**) and maximum width (**B**) of 100 carrot roots. Each point represents a unique carrot; colors indicate one of 10 carrot accessions representing a variety of market classes. RMSE of the linear model $y = x$ was 8.30 mm for length and 2.00 mm for maximum width.

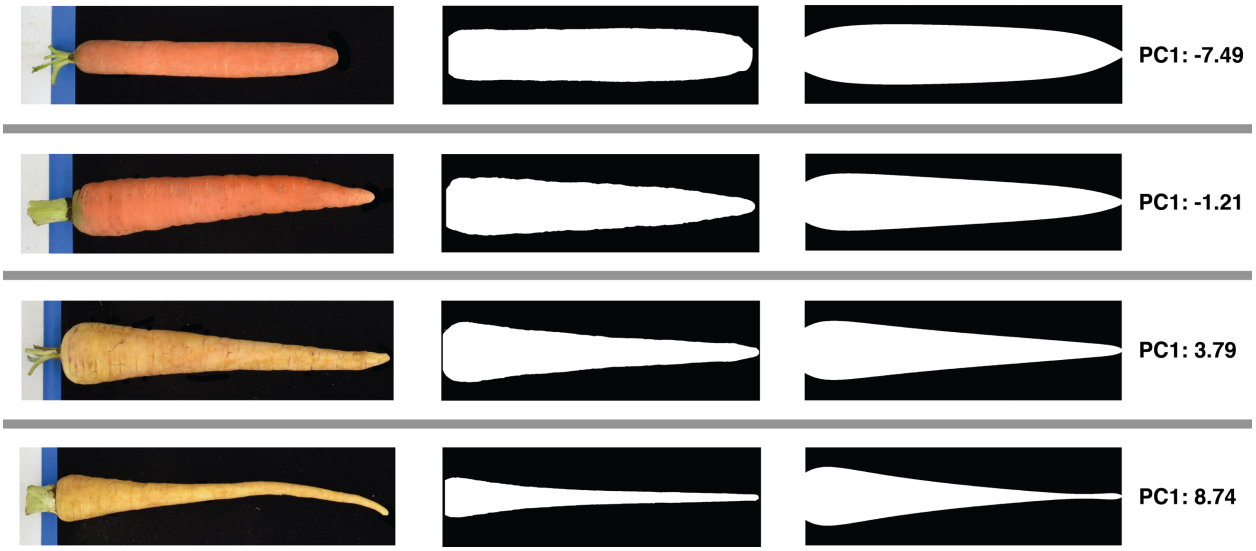


Figure 3. Quantification of size-independent variation in carrot root shape using PCA of length- and width- normalized contours. All carrots were standardized to a maximum width of 1 and a length of 1000, and contours were decomposed into five principal components. Rows correspond to four representative roots sampled from each quartile of the range of scores along the first principal component, and illustrate the form of phenotypic variance captured by this first component. From left to right: raw color photos of roots taken during image acquisition; straightened binary masks of the corresponding root; simulated root profiles generated by taking the product of the first PC score pertaining to this root (far right) and the mean of all other PC scores with the transpose of the eigenvectors generated during eigendecomposition. These simulated profiles demonstrate that variation along the first principal component reflect the degree of "root fill", or extent to which a carrot preserves its maximum width down its length.

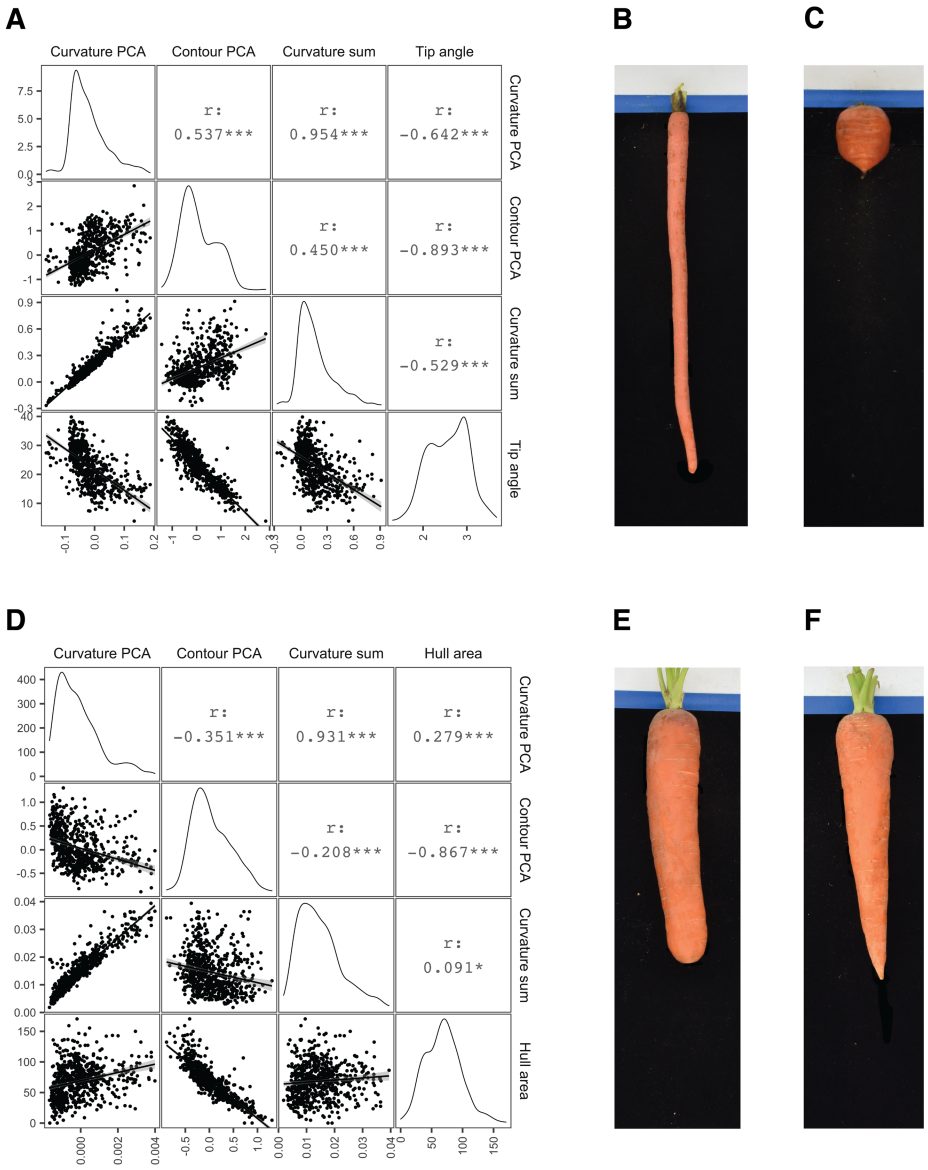


Figure 4. Correlograms comparing several measures of shoulder shape (**A**) and tip shape (**D**): PC1 scores derived from curvature values ("Curvature PCA"), PC1 scores derive from normalized contours ("Contour PCA"), the sum of curvature values, and either shoulder hull area (**A**) or tip angle (**D**). For Pearson correlation coefficients (*r*) shown above the diagonal, ***: *p*-value < 0.001; *: *p*-value < 0.1. Images on right illustrate representative roots drawn from the extremes of the first principal component scores corresponding to shoulder curvature values (**B** & **C**) and tip curvature values (**E** & **F**).

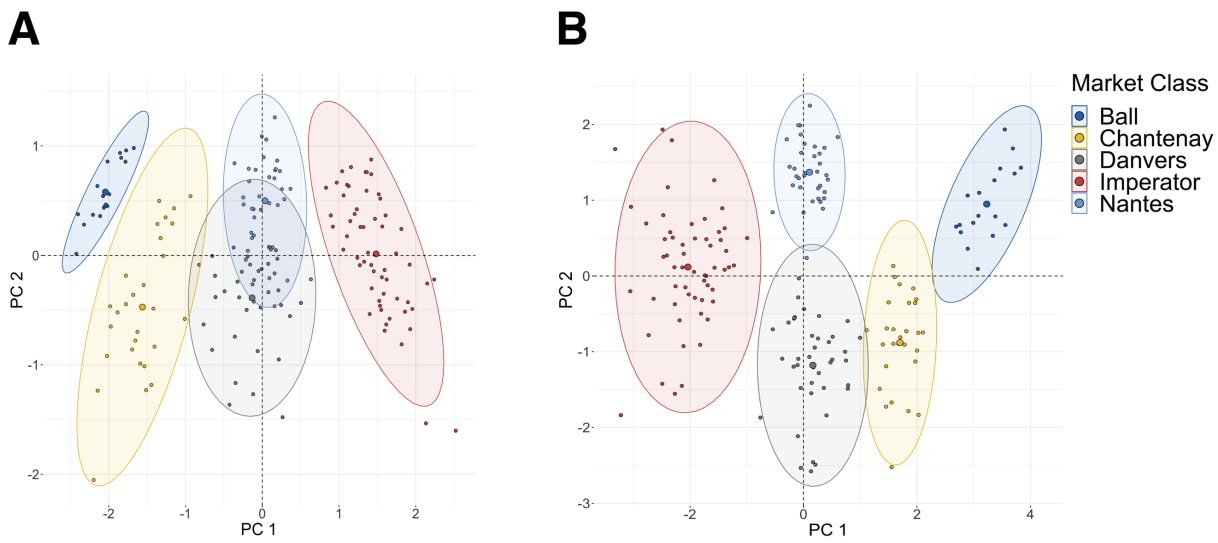
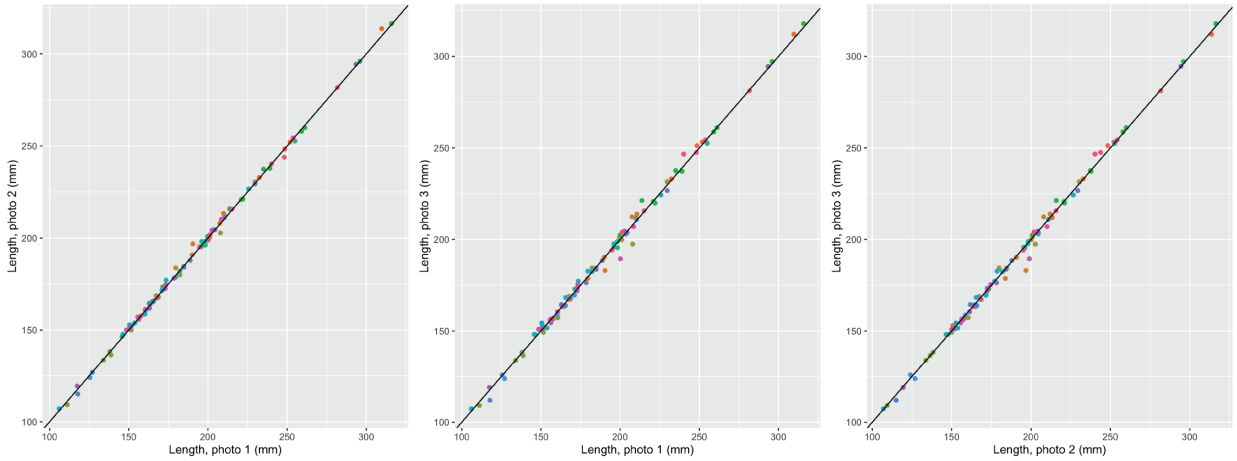


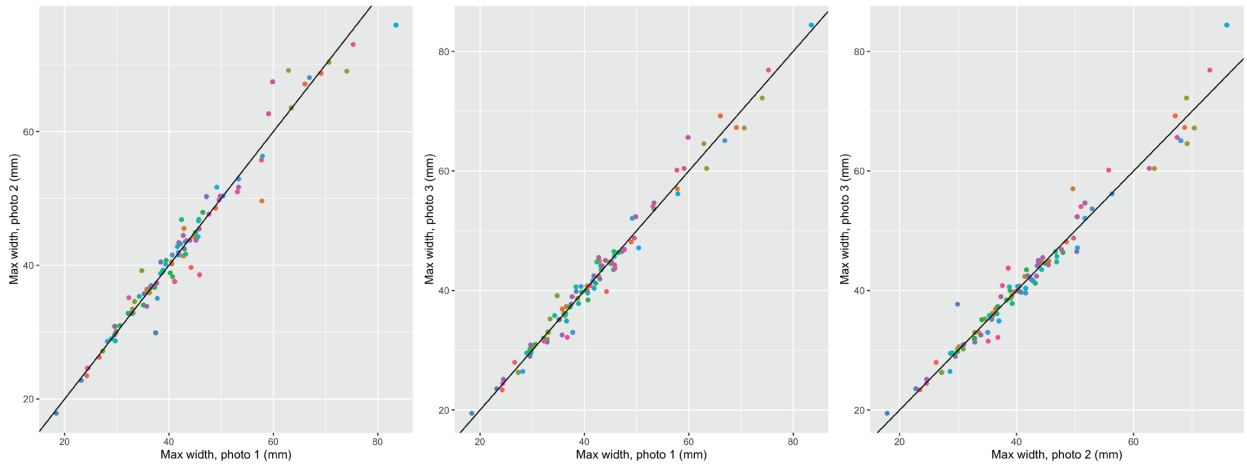
Figure 5. (A) PCA-based clustering of 175 roots sampled from five major market classes on the basis of only length and maximum width; **(B)** PCA-based clustering of these same roots using PCA of curvature values in the tip and shoulder, root length, maximum width, aspect ratio, and root fill.

Supplementary Figures

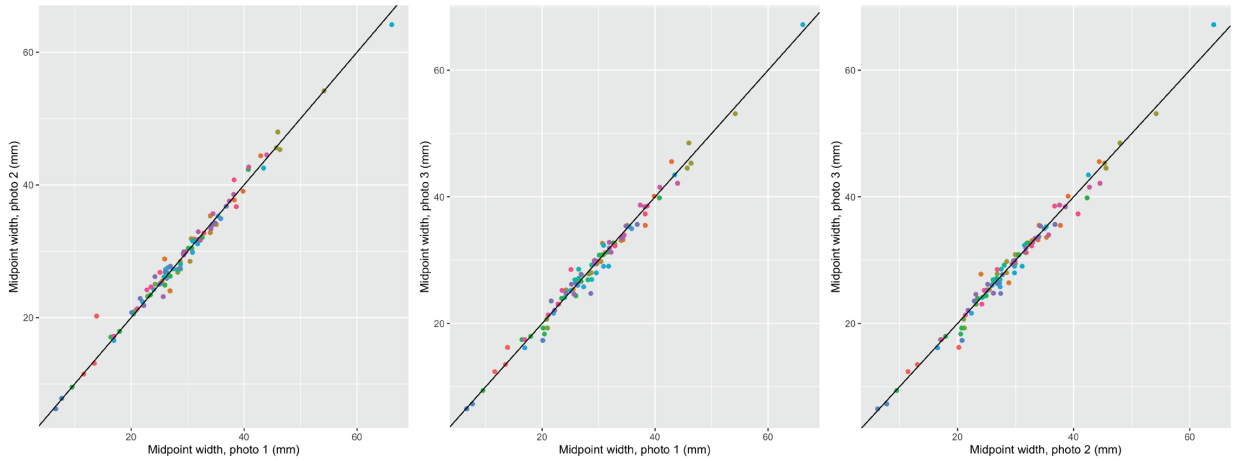
A



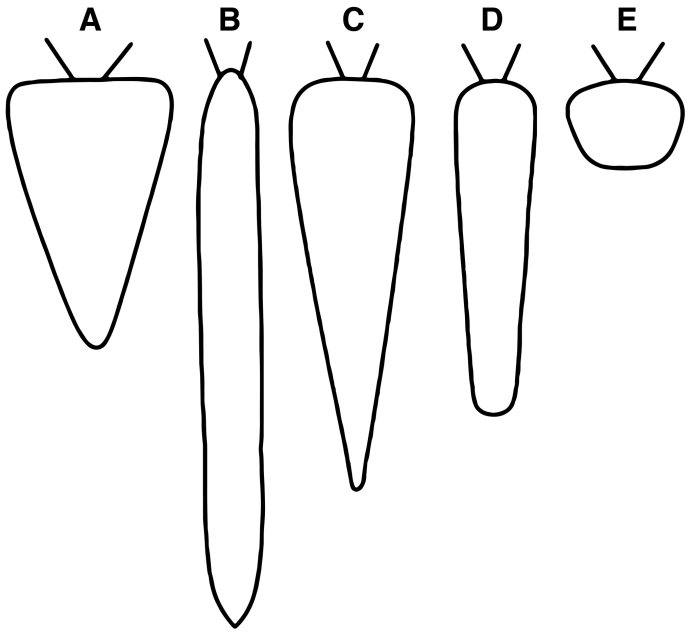
B



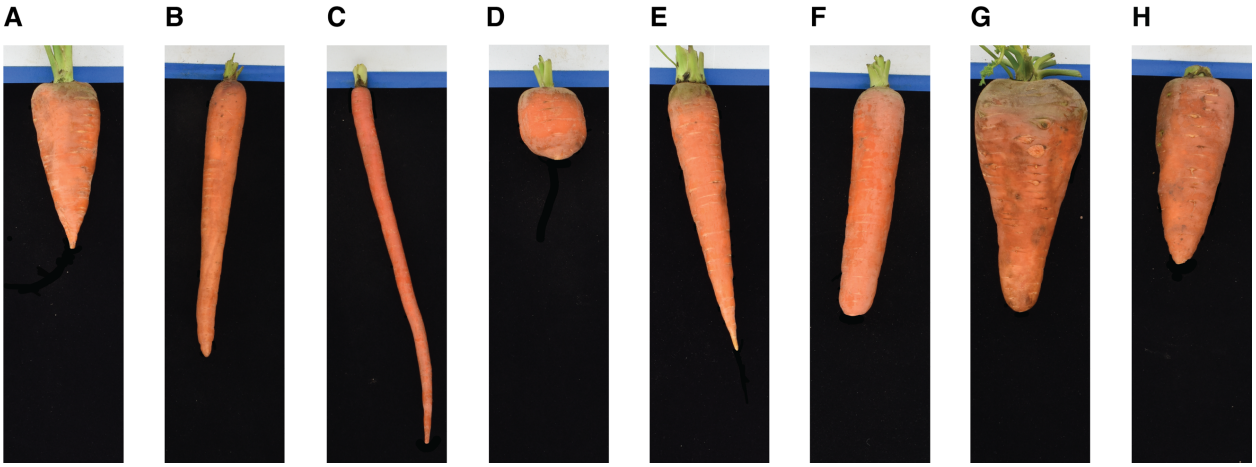
C



Supplementary Figure 1. Pairwise comparisons of phenotypes **(A)** length; **(B)** maximum width; **(C)** midpoint width, extracted from multiple photos of individual roots. Each point represents a single carrot that was photographed three times from three different angles (as in Fig. 3, color corresponds to distinct genotypes). Left-most panels compare photo 2 against photo 1, middle panels photo 3 versus photo 1, and right-most panels photo 3 versus photo 2.



Supplementary Figure 2. Illustration of stereotypical contours of roots from five major carrot market classes (adapted from Simon et al., (2008)) used in the clustering analysis: **(A)** Chantenay; **(B)** Imperator; **(C)** Danvers; **(D)** Nantes; **(E)** Parsienne/Ball.



Supplementary Figure 3. Carrot genotypes used in the half-diallel mating scheme. **(A)** W279; **(B)** B2566; **(C)** L1408; **(D)** OSSI-Ball; **(E)** W289; **(F)** W287; **(G)** W278; **(H)** W280. A, G & H exemplify the Chantenay-type processing carrot; E & F typify the Danvers, and Nantes market classes, and typically sold as fresh-market or storage carrots; D is a specialty Parisienne (or Ball) type; and C is an Imperator type used in the production of baby carrots.

Supplementary Figure 4. Representative roots from one of the pair-wise crosses included in the diallel mating scheme: **(A)** L1408; **(B)** L1408xOSSI-Ball; **(C)** OSSI-Ball.

Chapter 3:
Comparing genome-wide association analysis and genomic-estimated breeding values as methods for explaining heritable variation in the size and shape of carrot roots

Authors: Scott H. Brainard, Shelby L. Ellison, Douglas A. Senalik, Philipp W. Simon, Julie C. Dawson, Irwin L. Goldman

Authors' contributions: SHB, JCD, and ILG designed the experiments. SLE, DAS, and PWS grew the diversity panel. SHB performed the phenotyping; SHB, SLE and DAS performed the genotyping. SHB analyzed the data and wrote the chapter, with ILG and JCD. All authors have read and approved the chapter.

Abstract

The size and shape of carrot roots are not only major components affecting yield, but are the primary determinants of market class. These quantitative phenotypes have historically been challenging to objectively evaluate, and thus subjective visual assessment of market class remains the primary method by which selection for these traits is performed. However, advancements in digital image analysis have recently made possible the high-throughput quantification of size and shape attributes. Together with a high-quality reference genome, it is now feasible to utilize modern methods of genetic analysis in investigating the genetic control of root morphology. To this end, this study uses both genome wide association analysis and genomic predictions based on genomic-estimated breeding values. This analysis suggests that the components of market class are highly polygenic traits, likely under the influence of many small effect quantitative trait loci (QTL). In addition, relatively large proportions of additive genetic variance for many of the component phenotypes support high predictive ability of genomic-estimated breeding values. This study represents a novel advance in our understanding of the genetic control of market class in carrot root, as well as the possibilities of using genomic predictions for quantitative traits in horticultural crops.

Introduction

Carrot (*Daucus carota* subsp. *sativus*) is a widely cultivated vegetable crop of both significant economic importance—globally, annual carrot production exceeds 40 million metric tons (FAO, 2020)—and nutritional value, representing a significant source of pro-vitamin A in the human diet (Simon, 2000). Carrot roots are sold into many different markets as a fresh product, a

storage root, and a processing crop. In this regard, the size and shape of the edible, swollen taproot are key traits: in addition to influencing yield, root size and shape are the principle determinants of market class in carrot (Banga, 1957; Simon et al., 2008), affecting both harvestability, post-harvest handling, and marketability.

While extensive diversity for root size and shape exists within cultivated carrot germplasm, these quantitative traits have historically been challenging to objectively evaluate. To this day, distinguishing among market classes relies on a subjective visual assessment of the curvature of the carrot root shoulder and tip, its aspect ratio, as well as its length and width. No method for quantification of standard size and shape categories is currently recognized. In this context, quantitative digital image analysis holds significant potential in not only automating phenotyping tasks, but enabling the precise measurement of the determinative components of root shape.

An image analysis pipeline designed specifically for this task was recently developed and validated in order to provide a high-throughput method for accurately evaluating both size and shape parameters in a diverse collection of carrot germplasm (see Chapter 2). This pipeline allows for the precise characterization of the morphological phenotypes which in turn could allow for the establishment of criteria for distinguishing market classes from one another. In particular, it was shown that principal components analysis (PCA)-based methods of quantifying shoulder and tip curvature, as well as size-independent variation in the full root contour, improve discrimination between market classes, relative to what is possible using only

measurements of root length, width, and aspect ratio. Together with the recent construction of a high-quality, chromosome-scale reference genome for carrot (Iorizzo et al., 2016), this platform, and the quantitative phenotypic data it provides, makes it possible for the first time to utilize both genome-wide association analysis (GWAS) and genomic-estimated breeding values (GEBVs) to analyze the genetic control of root shape in carrot. These methods have both become widely utilized in the study of plant genetics, both in terms of their ability to improve the efficiency of plant breeding, as well as providing a starting point for molecular characterization of the genetic control of key agronomic traits.

This study utilized both of these methods, thus allowing for a comparison of the efficacy of GWAS – which attempts to identify QTL through their linkage with genetic markers – and GEBVs – which are based on a direct estimation of additive genotypic effects. GWAS has become a widely used tool in quantitative genetic analysis, however even in cases where marker density is high and a heterogenous diversity panel is utilized, it is often underpowered to detect a majority of small effect QTL in the case of highly polygenic traits (Brachi et al., 2011). In contrast, GEBVs calculated using an infinitesimal model of gene action do not use a significance threshold for inclusion of a marker in a predictive model. Since being initially developed in the context of animal breeding (Meuwissen et al., 2001), the development of efficient methods for calculating a marker-derived relationship matrix (VanRaden, 2008) has led to the extensive use of GEBVs in agricultural breeding programs. There are many factors that can limit the accuracy of such GEBVs, ranging from population structure and population size, to trait heritability. To date, these two methods have not been compared in carrot, or with respect to morphological

root traits in general. This study represents a novel advance in terms of our understanding of the genetic control of market class in carrot root, with implications both for further research and breeding for these traits. In addition, the methodological comparisons between GWAS and genomic prediction provide support to the growing body of literature illustrating the utility of GEBVs for highly polygenic, quantitative traits, particularly in unstructured, outcrossing plant species.

Materials and methods

Plant materials

749 accessions (also referred to throughout this paper as "genotypes") were utilized in this study, composed mainly of Plant Introductions in the USDA NPGS collection of *D. carota* germplasm held in Ames, IA, as well as breeding lines from both the University of Wisconsin and USDA-ARS carrot breeding programs in Madison, WI. This panel included the majority of the collection described in the population genetic analysis by Ellison et al. (2018), and as such represented a diverse, global collection sampled from all major identified population sub-groups, including the large subdivisions of Western domesticated, Western wild accessions, and Eastern samples, as well as carrots originating in Tunisia and Portugal. A description of the geographic origin of each of the samples included in this analysis is included in Supplementary File S1.

In 2016 and 2018, the collection was grown at the Hancock Agricultural Research Station in Hancock, WI (44°08'N, 89°32'W); plots were planted on May 16th and May 24th, and harvested

on August 29th and 30th, respectively. In 2018-2019, the collection was grown at the University of California Desert Research and Extension Center in Holtville, CA (32°48'N, 115°26'W). In this environment, planting occurred on October 10th, and roots were harvested on February 25th. In both locations, genotypes were grown in 1 m long rows; in Wisconsin, one replicate per genotype was planted, while in California, two replicates of all genotypes were planted in a randomized complete block design (RCBD). One to fifteen roots were harvested at random per replicate. In Wisconsin, higher amounts of disease pressure led to fewer mature, undamaged roots being harvested per plot, on average. Following harvest, the tops of the carrots were removed, and they were stored at 4°C until phenotyping.

Phenotypic Evaluation

Roots were digitally phenotyped as described in Chapter 2. In brief, after being cleaned, roots were QR coded and placed against either a white vinyl or black felt backdrop depending on root pigmentation. Images were acquired using a Nikon 5600 DSLR camera tethered to a Unix computer. Python bindings for the OpenCV library were used to create binary masks of the roots by thresholding the hue-saturation-value color space. Custom MATLAB scripts were subsequently used to correct for any residual curvature in each root, and a random forest classifier was used to remove any unexpanded portion of the tap root.

Following acquisition and pre-processing, phenotypes were extracted from the straightened, de-tipped binary masks. Root length, maximum width, total biomass, and aspect ratio were calculated for each root individually. In addition, in order to quantify size-independent

parameters of contour shape, principal components analysis (PCA) was performed on the root contour following a normalization procedure whereby each carrot was standardized to have a maximum width of 1, and a length of 1000. The scores along the first principal component capture the degree of root fill, or how far down the length of the carrot the maximum width of the carrot is maintained. In addition, curvature values were computed at each point along the root contour in both the shoulder and tip regions as described by Driscoll et al. (2012), with PCA of these values providing a metric corresponding to shoulder broadness and tip fill, respectively. Together, this suite of root traits has been found to accurately classify roots based on visual assignment of carrot market class (see Chapter 2, Figure 5).

Finally, prior to association analyses and construction of genomic prediction models, the diversity panel was restricted to those accessions that exhibited little to no branching of the taproot. Both the nature of the root-straightening algorithm – which depends upon finding tip of the carrot – and many of the phenotypes themselves (length, tip curvature), implicitly require that the root be a single unbranched taproot. Small root hairs were removed through smoothing operations, but just as forked or split roots were discarded from the analysis, cultivars with highly branched fibrous root systems were also excluded on the basis of being inappropriate to an analysis of market class traits. Compounded with the failure of some roots to produce new leaf tissue following vernalization, this reduced the total size of the diversity panel used in subsequent analyses to 661 unique cultivated accessions of *D. carota*.

Estimation of genotype means

In all subsequent association analyses and genomic predictions, a so-called "two-stage" approach was adopted whereby each genotype was first represented by a single phenotypic value, estimated using a mixed effects linear model. Genotype was modeled as a fixed effect, and each of the four unique combinations of location and replicate were combined into a single fixed "environment" effect with four levels. Because of unequal subsampling within environments, an additional random effect term was included to model genotype x environment interactions. The resulting model took the form of an RCBD model with subsampling:

$$Y_{ijl} = \mu + G_i + E_j + GE_{ij} + \varepsilon_{ijl}$$

Here, G_i represents the *i*th genotype effect, E_j the *j*th environment effect, GE_{ij} the genotype x environment interaction (with $GE_{ij} \sim N(0, \sigma^2_{GE})$) and ε_{ijl} the residual variance (i.e., variance among subsamples, with $\varepsilon_{ijl} \sim N(0, \sigma^2_{\varepsilon})$). Models were fit for each trait independently using the `lme4` package in R v4.0.3 (R Core Team, 2020), and genotype means were extracted using the package `emmeans`.

DNA extraction, genotyping, marker development

Following 6 weeks of vernalization at 4°C, one root per accession was transferred to a greenhouse environment and planted in conical tubes containing Pro-Mix High Porosity potting mix (Premier Tech, Quakertown, PA). Roots were maintained at 20°C with a 16 hr photoperiod. Following emergence of new leaf tissue, leaf samples were obtained, and stored at -80°C until lyophilization. ~0.1 g of freeze-dried tissue was then macerated, and genomic DNA was

extracted using Macherey-Nagel NucleoSpin 96-well kits. DNA quantification (using Quantus PicoGreen dsDNA kits), library preparation, and sequencing was performed at the University of Wisconsin-Madison Biotechnology Center. In brief, restriction enzyme-digestion was performed with *ApeKI*, following which Illumina adapters and sample-specific barcodes were annealed. Samples were pooled into groups of 96 per flowcell lane, and sequenced on an Illumina NovaSeq 6000, generating on average 4 million, 150-bp paired-end reads.

Raw, multiplexed .fastq files corresponding to forward and reverse reads of each lane were checked for quality, and demultiplexed using a custom a Java application (http://github.com/shbrainard/gbstools/), which facilitates including both forward and reverse reads to TASSEL 5. SNPs were then called using the GBSv2 pipeline of TASSEL 5, as described by Bradbury et al. (2007) with v3 of the *D. carota* genome (Iorizzo et al., 2016) used as a reference. Missing data was imputed with `Beagle` v 5.1 (Browning et al., 2018), using default parameters, 20 iterations, and 300 phase states. Filtering, performed using `bcftools` v 1.11, was used to remove markers with minor allele frequency less than 0.05, markers with mean depth less than 10 reads, or greater than 500 reads, and any non-biallelic markers. This filtering resulted in a total of 146,821 SNPs that were used as the basis for subsequent analyses. Genome-wide linkage disequilibrium was also calculated using `bcftools`, as the square of the sample Pearson correlation between marker genotypes (r^2). Filtering on the basis of LD was performed using the `prune` plugin.

Calculation of the realized-relationship matrix

In both performing association analysis and genomic prediction, SNPs were used to estimate a realized relationship matrix, $\mathbf{A}_m$, calculated on the basis of the imputed marker data:

$$\mathbf{A}_m = \frac{\mathbf{Z}\mathbf{Z}^{\mathrm{T}}}{\sum_{k=1}^{m} 2p_k q_k}$$

with $\mathbf{Z}$ representing a matrix of centered genotypes (661 accessions x 146,821 markers). The scaling factor insures diagonal elements are equal to 1 + f, where *f* is equal to the intra-individual gametic correlation Kang et al. (2008). p_k and q_k indicate minor and major allele frequency for the k^{th} marker. Shrinkage estimation was also applied in the case of estimating breeding values, using default settings of the `A.mat` function of `rrBLUP` v 4.6.1, as described by Endelman & Jannink (2012).

Linkage disequilibrium decay and population structure

LD was assessed in two ways. First, correlation coefficients between each SNP and its 100 nearest neighboring markers were calculated, and recorded along with the physical genetic distance between each pair. Distances were then binned, and the decay in LD regressed against genetic distance using the decay function $LD(x) \sim y_f + (y_0 - y_f)e^{-e^{(log\alpha)x}}$, with initial estimates for y_f, $y_0$and $log\alpha$ estimated in R using self-starting regression function `SSasymp`. Second, genome-wide LD was visualized as a Manhattan plot by calculating the mean LD of each SNP again with its 100 nearest neighbors, having first thinned the marker dataset to only 1 SNP per kilobase, to avoid distortions due exclusively to uneven marker distribution across the genome.

Population structure was assessed by performing PCA on the centered marker matrix, and plotting the first two PCs against each other in a biplot. Scree plots of variance attributed to each component were also used to visually determine the correct number of PCs to include as fixed effects in the GWAS model.

Genome-wide association analysis

GWAS was performed using the `GWASpoly` package (Rosyara et al., 2016), which implements the mixed model described by Yu et al., (2006). This tool utilizes the so-called *Q + K* method (Zhang et al., 2010), whereby population structure and relatedness between individuals is controlled for through both fixed effects, as well as a random polygenic term calculated using all markers. This resulted in the model:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{S}\boldsymbol{\tau} + \mathbf{Z}\boldsymbol{u} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{y}$ is a vector of phenotypes, calculated as the estimated genotype values from the linear model described above. $\boldsymbol{\tau}$ is a vector of SNP effects. $\boldsymbol{u}$ is a vector of random polygenic effects, with a variance equal to $\sigma_G^2\mathbf{K}$, where σ_G^2 is the genetic variance, and **K** is the realized relationship matrix defined above, but without scaling by *p* and *k*. Because variance components were estimated for each marker independently, this model is equivalent to that proposed by Kang et al. (2008). $\boldsymbol{\varepsilon}$ is a vector of residual effects following a $N(0, \mathbf{I}\sigma_\varepsilon^2)$ distribution, and $\boldsymbol{\beta}$ is a vector of fixed population structure effects. **X**, **S**, and **Z** represent the respective incidence matrices. In this study, the first principal component of the marker matrix was used as a fixed effect, as proposed by Price et al., (2006) as an alternative to the groupings provided by a clustering algorithm such as STRUCTURE (Pritchard et al., 2000).

LOD thresholds were determined on a trait-by-trait basis by running 1000 simulated analyses using random permutations of each phenotype vector; this permutation testing was performed using the compute resources and assistance of the UW-Madison Center For High Throughput Computing (CHTC) in the Department of Computer Sciences. LOD thresholds were empirically calculated to control the family-wise error rate (FWER) at a level of α=0.05.

Genomic-estimated breeding values

In addition to GWAS, marker data were used to calculate GEBVs, using best linear unbiased predicators (BLUPs) (Henderson, 1963). The marker matrix used for estimation of kinship was thinned significantly. Using `bcftools`, as above, markers were thinned to a minimum density of 1 marker per 1 kilobase (kb), with no missing data, resulting in 12,370 SNPs. The $\mathbf{A_m}$ matrix was then calculated as above. BLUPs of the additive genotypic effects were then calculated using the `rr`BLUP package (Endelman, 2011). The prediction error variance (PEV) of the BLUPs calculated as the inverse of the $\mathbf{C_{22}}$ component of the coefficient matrix, scaled by the diagonal elements of the covariance structure defined by the realized relationship matrix (i.e., the variance of the given BLUP) (Henderson, 1973). Cross-validation of these predictions was also performed in order to assess predictive ability, by calculating the correlation of predicted values, and an estimate of the true genotypic value. In this study, by default, 10% of the phenotypic data was randomly masked in the calculation of BLUPs (the validation population; VP), and correlation coefficients between the BLUPs for these genotypes, and their true

phenotypic values was calculated, using the remaining 90% of the panel as a training population (TP). This was repeated 100 times, and average correlations reported as predictive ability.

Analysis of parameters affecting predictive ability

SNP density, degree of relatedness between TP and VP, and TP size were all evaluated in terms of their effects on predictive ability using the cross-validation approach described above. In each case, the same self-starting regression function `SSasymp` used above to model LD decay was fit to the resulting data. In addition, for any specific cross-validation analysis, all parameters not being varied were held constant at levels determined to not limit predictive ability.

For SNP density, VCF files were filtered according to a sequence of thinning parameters to generate progressively more sparse marker distributions. Markers were thinned to one SNP every 0.1 kb (resulting in 18,093 SNPs), 2 kb (resulting in 11,269 SNPs), 5 kb (resulting in 9,535 SNPs), 10 kb (resulting in 7,882 SNPs), 100 kb (resulting in 2,392 SNPs), 250 kb (resulting in 1,191 SNPs), 500 kb (resulting in 660 SNPs), 750 kb (resulting in 467 SNPs), 1 megabase (Mb) (resulting in 362 SNPs), 2 Mb (resulting in 195 SNPs), 3 Mb (resulting in 136 SNPs), 4 Mb (resulting in 103 SNPs), and 5 Mb (resulting in 86 SNPs). A separate VCF file was also generated containing only markers on chromosome 3 (resulting in 5,621 SNPs); this highly biased marker set provided an extreme case with which to evaluate the effects of distorting the genome-wide distribution of a relatively large number of markers.

To evaluate the effect of similarity between the TP and VP, a k-means clustering algorithm was applied to the PCA of the marker matrix containing 9,535 SNPs. The most distant sub-grouping of 60 accessions was adopted as the most unrelated VP, and was then progressively diluted by swapping individuals at random with random draws from the larger population, to thereby simulate a gradient of relatedness between the TP and VP. Dilution amounts were set to 1, 3, 5, 7, 10, 13, 15, 17, 24, 27, 31, 35, 45, and 50 individuals, with 50 replications performed at each dilution level. "Distance" between TP and VP was then calculated on the basis of a similarity matrix, defined as the inverse of the distance matrix constructed from the first 100 PCs of the PCA. Each individual in a given VP was then compared against the *n* most similar individuals in the TP, where *n* was allowed to vary between 1 and 660, depending on the distance metric being analyzed. These distances were then averaged across all individuals in the TP. The same 100-fold cross-validation approach as outlined above was repeated here for every TP/VP combination, and the resultant average predictive ability was then regressed onto the average similarity index. Finally, the same exponential function defined above for characterizing LD decay was then fit.

To evaluate the effect of varying TP size, two distinct approaches were taken. First, absolute TP size was varied, while holding the relative size of the VP constant at 10%. This was performed by sampling subsets of the full panel of sizes ranging from 10 to 660, repeating this sampling process 50 times at each population size, and performing the same 100-fold cross validation approach as with previous analyses. Separately, relative TP size was also varied, by holding the absolute size of the VP constant at 60 individuals, and varying the total TP size from 75 to 660.

As with the previous analysis, at each level of relative TP size, the sampling procedure was iterated 50 times, and for each iteration the 100-fold cross validation approach described above was performed. Finally, relative VP size was varied, by holding the total population size constant at 661 individuals, and sampling VPs ranging from 10 to 650, and as above, iterating each VP size 50 times, and performing 100-fold cross-validation at each iteration.

Data availability

Filtered SNP markers, as well as binary masks of the images used for phenotyping are available via the Harvard Dataverse: https://dataverse.harvard.edu/dataverse/usda-npgs-carrot-collection. Phenotyping methods described in Chapter 2 rely on custom Python and MATLAB pipelines available at https://github.com/shbrainard/carrot-phenotyping (for image acquisition and production of binary masks) and https://github.com/jbustamante35/carrotsweeper (for straightening binary masks and performing PCA on contours or curvature values).

Results

Phenotypic variation present in the diversity panel

Representative roots drawn from four common carrot market classes are shown in Fig. 1, illustrating the degree of phenotypic differentiation between classes, as well the particular way in which the components of root size and shape are assembled to define specific classes. For instance, Imperator-type roots (Fig. 1b) combine narrow maximum width with long root length, while the Chantenay type (Fig. 1d) combines large maximum width with low degrees of root fill;

as described above, this latter trait reflects the first PC score of a PCA analysis of the size-normalized root contour. While all of the carrots shown in Fig. 1 clearly exhibit distinct aspect ratios, Fig. 1a-c all have high degrees of root fill, while only Fig. 1d exhibits a rapid tapering along its length. This highlights the particular value of image analysis procedures such as PCA, which allow for the de-coupling of size and shape parameters, and extraction and quantification of high-dimensional phenotypes.

Linkage disequilibrium and population structure

The extent and rate of decay of linkage disequilibrium (LD) across the genome is an important determinant of the potential resolution of association analysis, and a decisive factor in determining marker density when performing GWAS (Alqudah et al., 2020; Otyama et al., 2019). A slow decrease in LD toward equilibrium, as the distance between pairs of markers increases, implies both that a relatively fewer number of markers is necessary to effectively capture the extent of historical recombination in the diversity panel, but also that large stretches of extended haplotype blocks will likely lead to larger intervals for QTL, with numerous non-causal SNPs found to be significantly associated with given traits (Myles et al., 2009).

Results of short-range LD decay are shown in Fig. 2a. Intercepts with r^2 values of 0.2 (blue line) and 0.1 (green line) are indicated at 796 and 19.7 kb, respectively, illustrating that within only several kb, there is a rapid approach to linkage equilibrium across the panel. As a result, the curated marker set (with filters as described above), was judged to be more than adequately

dense, with an average distance between markers of only 3711 bp. Short-range genome-wide LD patterns were visualized as a Manhattan plot shown in Fig. 2b. While some peaks in LD are seen, particularly on chromosome 2, average LD is relatively minimal, with a mean of only 0.038, and only 3.6% of all windows exceeding the threshold of $r^2 > 0.1$. This demonstrates both a consistent and relatively limited degree of LD across the genome. While genome-wide estimates of LD are never perfect using unphased genotypes, the dense marker dataset available in this study appeared well-suited to association analysis.

In addition to LD, which determines an upper bound on QTL resolution, and as such, informs appropriate marker density, population structure is another determinative characteristic of any association panel. The presence of uncontrolled population structure and admixture can lead to spurious inflation of *p*-values, even in the absence of severe linkage disequilibrium (Ewens & Spielman, 1995; Pritchard & Rosenberg, 1999). Multiple methods have been proposed for controlling for such population structure by adding both fixed (Price et al., 2006; Pritchard et al., 2000) and random (Yu et al., 2006) covariates to the model. As described above, a combination of both fixed and random covariates, the so-called K+Q model, was utilized in this study, whereby the markers themselves are used to estimate a random polygenic relationship matrix, and principal component analysis of the marker matrix is used to calculated fixed effects.

A PCA bi-plot was used to assess the degree of population structure, and the results mirror the minimal degree of structure observed by Ellison et al. (2018) (Fig. 3a). Aside from one main

cluster observed around high values along PC1 (which was previously identified as corresponding to Western domesticated carrot accessions), little clustering was detected. And indeed, the variance captured by the first component was only ~10% of total variance in the marker matrix, with all subsequent components explaining roughly 1% of total variance (Fig. 3b). Consequently, one PC was judged to be sufficient to be included as a fixed effect in the association analysis.

Genome-wide association analysis

Manhattan plots illustrating the results of GWAS for 4 root shape traits that compose market class are shown in Fig. 4. Three of these traits pertain specifically to the dimensions of the carrot root: length, maximum width (which occurs in the shoulder region of carrot roots), and their quotient, aspect ratio. These traits define the size of the root, and are in principle measurable by hand. In addition, however, significant associations were also found for root fill, which corresponds to the first PC score obtained by performing PCA on the length- and width-normalized root profile. This trait represents the most significant source of variation in contour of the full root—specifically, the extent to which a carrot maintains its maximum width down its length. Root fill is therefore explicitly a “shape trait”, insofar as it is calculated from the contours that have been standardized for their size, and as such, is not measurable by hand, though it reflects a key aspect of market class.

The most significant SNPs corresponding to each peak in a trait’s respective Manhattan plot are listed in Table 1, and box-and-whisker plots of the effect sizes of the most significant SNP per

trait (shown here as a function of allele dosage) are shown in Fig. 5. Across all traits, relatively limited numbers of peaks were detected, with significant SNPs being located in a single peak for length, root fill and maximum width, 4 peaks for L/W. Interestingly, while all of these peaks exceed a LOD threshold set to control the FWER at 0.05, their effect sizes are relatively small (Table 1, Fig. 5). Linear models that regress a given trait's phenotype onto these markers as fixed effects exhibited an average adjusted R^2 of only 0.03, reflecting the limited explanatory power of these SNPs. A notable exception was aspect ratio, for which 5 clear peaks were detected: 1 on chromosome 1, 2 on chromosome 3, one on chromosome 4, and one on chromosome 6. In this case, the R^2 of the complete model is 0.18 – significantly higher than all other traits, and approaching a level that could be practically useful in a marker-assisted selection context, if the regions were fine mapped for the development of markers near the causal loci. As will be shown below, however, this value is significantly lower than the predictive ability of GEBVs. In addition, the remaining two market class-related traits – shoulder and tip curvature – which reflect more subtle variation in the contour of these regions, were not significantly associated with any SNPs.

Effects of marker density

While the general result that increasing marker density increases predictive ability has been well-documented, the precise nature of the relationship will vary depending on the population and traits under consideration. Given that this diversity panel was genotyped at high density for the purpose of GWAS, it was therefore feasible to produce marker sets generated through progressively more stringent filtering criteria, and thereby determine the effect of SNP density

on predictive ability through cross-validation for each of the digitally phenotyped root traits evaluated in this study. The results of these analyses are shown in Fig. 6, and are well described by an exponential function: at low marker densities, any increase in density is met with relatively rapid increases in predictive ability. As marker density is increased, however, these increases asymptomatically approach a maximum predictive ability, which in the case of these roots traits is attained at roughly 2,500 SNPs. This can be contrasted with GWAS, which, as described above, necessitates a relatively dense array of markers across the genome in order to increase the likelihood that some subset of these will be in high LD with QTL. If the marker dataset used in the GWAS analyses shown above, is thinned to only 1 marker every 100kb (i.e., 2,392 markers), all significant associations between SNPs and QTL shown in the above Manhattan plots (Fig. 4) are no longer detected (Supp. Fig. 1).

Similar results were obtained by specifically utilizing markers on a single chromosome, and comparing predictive ability against a random distribution of an equivalent number of SNPs. As shown in Supp. Table. 1 utilizing exclusively markers on chromosome 3 only reduces predictive ability by an average of 12% across all traits. Some reduction in accuracy is to be expected, due to some degree of linkage between markers when all are located on a single chromosome. This finding is in line with those of Daetwyler et al. (2012), and highlights the fact that markers' effect on predictive ability of GEBVs is primarily a function of their ability to accurately model covariance between individuals, and not their linkage with QTL.

Effect of relatedness between the TP and VP

It has been well-documented that increasing relatedness between the TP and VP leads to increases in predictive ability (Edwards et al., 2019; Olatoye et al., 2020). This is the logical consequence of increasing the similarity in the causal variants segregating in the two populations, and thus the degree to which genetic covariance between the two populations can accurately be used to make predictions. The results of regressing predictive ability onto three different measures of relatedness (i.e., three different values of *n*) are shown in Fig. 7, and confirm this general result. For all traits, an exponential regression similar to that observed in the case of varying SNP density is again observed. Prediction accuracies are substantially reduced from their maximum when similarity between the TP and VP is minimized, and as similarity increases, prediction ability increased exponentially, approaching a maximum that itself is the average predictive ability reported above. This convergence to the average can be understood as a consequence of the fact that on average, a random sampling procedure will select a VP that is extremely similar to the TP for this diversity panel, due to the overarching lack of population structure.

In addition to this general trend, it is also clear that across all traits, both extremely high and extremely low values of *n* – i.e., the number of individuals in the TP that each member of the VP was compared against – give non-optimal results. In the case of the former, due to oversampling, the possible range of similarities between the TP and VP is significantly compressed at low values. Because average prediction accuracies are high, this has the additional result of inflating average prediction accuracies, given a specific level of similarity. At the other extreme, there is clearly substantial noise about the exponential regression for the

case of n = 1, a consequence of this metric of similarity providing an inaccurate representation of the overall similarity between the TP and VP due to undersampling; i.e., more than one individual must be similar to each individual in the VP in order to make accurate predictions. While these results do not provide a basis for determining a specific value of n that should be used in for any arbitrary TP/VP combination, they do justify intermediate values of n (e.g., in this study, 40) that are both highly precise in that they fit the exponential regression function extremely well, while also being highly informative, in that they allow for a discrimination between degrees of relatedness for a wide range of TP/VP combinations.

Effect of population size on predictive ability

The last parameter evaluated in terms of its effect on predictive ability was the size of the TP. In this case as well, studies have consistently shown that increasing the size of the TP will, on average, increase predictive ability. However, there are three distinct ways in which the effect of TP size can be modified, as described above: absolute population size can be varied, relative TP size can be varied, or relative VP size can be varied.

In any specific context, one may also wish to consider the interaction between all of these effects, but for the purposes of providing more generalizable conclusions, they were performed discretely in this analysis. Thus, first, the effect of varying absolute population size, with VP size held constant at 10%, was evaluated. For all traits, predictive ability reaches its asymptote at roughly 330 individuals (Fig. 8a). Similarly, when the VP size is held constant at 60 individuals and relative TP size is varied, all traits appear to follow a similar dynamic, with predictive ability

again attaining is maximum at roughly 330 (Fig. 8b). Finally, the effect of varying relative VP size, with the TP held constant at the total population size of 661 individuals (less the size of the VP) was analyzed. Again, all traits follow the relationship seen in the previous two cross-validations, and across all traits, predictive ability reaches its asymptote when the TP is less than roughly 50% of the total population, or 330 individuals (Fig. 8c).

Discussion

This study identified a novel set of QTL for four of the most relevant morphological components of root market class in carrot. For the most part these represented a relatively limited set of small-effect QTL. This is surprising, given the relatively high heritabilities observed for these phenotypes, both as estimated here using genomic data, and as previously described using a diallel mating design (see Chapter 2). This would suggest that the effect size of the identified QTL is being underestimated, that there are additional unidentified QTL, or both. In any of these cases, however, the observed results would therefore run counter to the Beavis effect (Beavis and Paterson, 1998): given that the traits appear highly heritable, the variance attributed to the small number of QTL found in these GWAS analyses should have been proportionally over-estimated. This is itself unlikely, however, due to the low percent variance explained by the QTL in each case.

There are multiple plausible explanations for these two apparently contradictory results. First, as described above, carrot is a highly heterozygous outcrossing species, and indeed, most of the

accessions included in this diversity panel are not inbred lines, but sourced from landraces, open-pollinated varieties, and populations. This produced a strikingly rapid decay in linkage disequilibrium within this diversity panel, with an average r^2 of 0.2 between pairs of markers reached with a distance of only 796 bp. Despite the dense marker distribution used in this study, therefore, it is likely that this rapid decay of LD led to an underestimation, instead of overestimation, both of the effect size of the QTL that were identified in this study, as well as of markers that did not exceed the LOD threshold. Further complicating this analysis is the fact that selection for root shape morphology has likely occurred in numerous genetic backgrounds, with different allelic combinations producing similar root shapes within, e.g., Western European, Eastern European and North American accessions, and the USDA-NPGS collection used in this study included accessions from all of these geographic regions. In this regard, linkage analysis could again provide a fruitful subsequent line of analysis, by addressing the under-estimation of effect sizes due to differences in frequencies between marker alleles and QTL alleles.

Tip fill (e.g., the blunt-tipped Nantes type in Fig. 1c vs. the pointed Chantenay root in Fig. 1d) and shoulder broadness (e.g., the highly curved Parisienne-type in Fig. 1a vs. the straight-shoulders of the Imperator in Fig. 1b) represented more subtle aspects of root shape variation, since they are restricted to specific regions of the root contour. No SNPs were found to be significantly associated with phenotypic variation for these two traits, despite their evident importance in distinguishing between market classes. For the components of market class in carrot for which QTL were identified, the significant SNPs explained a small percentage of the

total variation for these traits. This is not an uncommon result when the traits under consideration are quantitative and highly polygenic, as would appear to be the case here. Indeed, it is consistent with the only other published report of a GWAS that included carrot root traits (Macko-Podgórni et al., 2020), which detected a QTL on chromosome 1 which accounted for roughly 10% of the phenotypic variation. The panel utilized by Macko-Podgórni et al. (2020) differed significantly from the material used in this study, representing 103 accessions from the Warwick Crop Centre in Wellesbourne, UK. It is therefore not surprising that the QTL in that context was not detected in our analysis, but nevertheless, the low effect size is consistent with that reported here. Similarly, a previous study that attempted to use linkage mapping to detect QTL for root traits detected no QTL for width or aspect ratio, and the 3 QTL associated with length each explained less than 10% of the phenotypic variation (Turner et al., 2018). These results are not directly comparable with those presented here, since phenotypic effect sizes in a linkage mapping context are estimated as a function of the LOD; furthermore, these populations represented F_2 families descended from crosses of elite inbred lines with wild accessions, and as such likely reflect domestication loci. Nevertheless, the low effect size of the QTL is again consistent.

The predictive ability of GEBVs were evaluated using the same diversity panel and marker set used in the GWAS analysis, so as to make accurate comparisons between the two approaches. The key difference between GEBVs and GWAS can be summarized as the reliance of the latter on detecting significant associations between markers and QTL, while the former simply attempts to directly estimate additive genotypic effects using markers as the basis of a

covariance matrix for modeling relatedness. Predictive ability for many of the traits phenotyped using image analysis were found to be quite high, and therefore cross-validation studies were performed to gain a more precise understanding of how robust these predictive ability would be given various marker densities, population sizes, and degrees of relationship between training and validation sets.

Regarding marker density, the asymptote of the exponential relationship between SNP density and predictive ability was attained a relatively low marker density. Compared to the 146,821 markers used in the GWAS analysis, maximal predictive ability was attained at only several thousand markers. This is consistent with the ranges presented in numerous previous studies (Erbe et al., 2013; Wang et al., 2017; Wu et al., 2016; Zhang et al., 2015), and highlights the different role that markers play in an analysis based around testing significant associations (particularly in a species with rapid LD decay) and estimating genomic relatedness.

In the case of GWAS, a large diversity panel ideally encompasses a large amount of historical recombination. This fast erosion of linkage disequilibrium implies that in order to detect QTL it is essential to have very dense molecular markers across the genome. While increasing the number of markers used in GWAS increases the computational burden of the analysis, and potentially decreases power due to the increased number of comparisons being made, the attendant increase in LD between markers and putative QTL has been found to outweigh these costs, even when hundreds of thousands of markers are used. This situation can clearly be contrasted with genomic prediction, where the goal is not to identify associations between

markers and QTL, but rather to estimate additive genotypic effects on the basis of a marker-derived covariance matrix. Due to this, far fewer markers are required to attain maximum predictive ability of GEBVs, relative to the markers required to detect QTL in a GWAS context.

Furthermore, the results presented here demonstrate that from a practical perspective, the genotyping costs associated with implementing genomic selection are at least in principle less than that associated with GWAS, though in practice this would depend on a high-quality genotyping platform that generated only thousands, instead of hundreds of thousands of markers. In addition, it is important to note that the vast majority of the accessions present in this diversity panel represent landraces, open-pollinated cultivars, or populations; only a small minority are inbred lines. As a result, these accessions certainly contain variable degrees of genetic heterogeneity. Loci that were called as heterozygous in the marker dataset utilized here are most likely segregating in these accession, and utilizing the genotype of a single root will necessarily mask this intra-accession variation. Techniques such as PoolSeq, which utilize bulked DNA from multiple individuals for GBS sequencing, have been found to be potentially useful in such situations (Anand et al., 2016; Bélanger et al., 2016). By sequencing at a high depth, it is possible to estimate continuous measures of allele frequency in the linear models used to test for associations between SNPs and a given phenotype (instead of the categorical allele dosages used here).

With respect to population structure, it has long been understood that because predictions are based on a covariance matrix relating phenotyped individuals to non-phenotyped individuals,

that higher degrees of relatedness between the TP and VP will lead to higher predictive ability. However, this phenomenon has typically been investigated qualitatively, by either comparing prediction accuracies across populations with known degrees of variable relatedness (such as full- vs. half-sib families), or between population groupings defined through clustering algorithms such STRUCTURE (Lozada et al., 2019; Riedelsheimer et al., 2013; Sverrisdóttir et al., 2018). While it would be theoretically feasible to apply a clustering algorithm to the panel presented in this study, such an approach is of limited utility in the case of a loosely structured diversity panel such as that used in this study. Because this population is not composed of well-defined, discrete subpopulations, the relevant question is one which relates a quantitative measure of the degree of relatedness between the TP and VP to predictive ability, not a qualitative judgment as to whether the TP and VP are or are not members of the same sub-grouping. Furthermore, in cases where relatedness has been measured between the TP and VP, the appropriate metric is typically assumed to be a comparison of the means of the two groups; i.e., the value of *n*, as designated in this study, is set equal to the size of the TP, and each individual on the VP is compared against each individual in the TP (Berro et al., 2019).

The cross-validations performed in this study therefore represent an advance in terms of the precision with which conclusions regarding the effect of relatedness on predictive ability can be made. Not only is increasing similarity between the TP and VP associated with increases in predictive ability, but this relationship also follows the same exponential function found in the case of marker density. In addition, intermediate values of *n* were found to give the most precise and informative measure of similarity in terms of defining this exponential relationship.

Finally, regarding conclusions one can draw about this particular diversity panel, it is clear that one of the factors contributing to high prediction accuracies on average is that the mean level of relatedness between a randomly selected TP and VP is extremely high (measured with *n* = 40, average similarity is 23.1, with a standard deviation of 0.56). This is already in the range of relatedness that defines the asymptotic portion of the exponential function, and thus it is reasonable to assume that these GEBVs would be robustly accurate to any arbitrary construction of TP and VP, given a breeding population similar in structure to the diversity panel analyzed here. In this regard it is important to emphasize that while this diversity panel is relatively unstructured, and contains a large amount of genetic variation, breeding populations often have higher levels of overall relatedness among individuals, and may have clear family structures. They will almost certainly contain less diversity than a global germplasm collection. These results will not necessarily transfer, therefore, to any particular breeding context.

It is interesting to note that not all traits exhibited an exponential relationship between predictive ability and relatedness equally well. Certain traits, such as root fill and tip curvature, clearly exhibited this exponential relationship (Fig. 7 and Supp Fig. 2). Others, such as aspect ratio, length, and maximum width displayed a more linear relationship, while biomass appeared to follow an entirely linear trend. This variation is well correlated with the maximum predictive ability attained for each of the traits, and therefore would suggest that the asymptotic portion of the relationship is only evident when high predictive ability is attainable given a wide range of analyzed similarity levels.

Finally, the effects of TP size were considered in this study by explicitly considering two distinct cases. First, the consequence of varying absolute population size was evaluated. Second, the effect of changing the relative size of either the TP or the VP was considered. These effects typically are confounded with each other in studies that have considered how to optimize the size of TPs: Xu et al. (2018) considered only the effect of varying the absolute size of the TP, by maintaining the VP at 20% of the total population size; Tayeh et al. (2015) only considered the effect of varying the relative size of the TP, by holding the absolute VP size constant, and varying the size of the overall population; Zhang et al. (2017) varied both the absolute size of the TP, and varied the relative size of the VP at each of these levels, but then averaged across all of the relative size variations, reporting only the effect of changes in absolute TP size.

Here again, in all three cases, an exponential relationship was found between predictive ability and either the total size of the population, or the relative size of the VP or TP. While it is unsurprising that increasing the total number of individuals in the panel would increase predictive ability, it is interesting to note that in the case of varying relative TP size, (either by increasing the size of the TP while holding the VP constant, or conversely by decreasing the size of the VP while holding the TP constant), the key determinant is simply the number of individuals in the TP, scaled to the number of individuals one is attempting to predict. Regarding the point at which the asymptotic maximal predictive ability is attained, at roughly or 50% of the total diversity panel, it is interesting to note that this is consistent with other reports of the minimum TP size needed to attain maximal predictive ability in squash (Hernandez et al., 2020), and carrot (Corak et al., 2019).

While the results of these cross-validations are therefore very similar – i.e., increasing the number of individuals in the TP leads to an increase in predictive ability – they clearly differ in their precise interpretation, and most importantly, describe very different practical questions from a resource allocation perspective. Analyses which vary the relative size of the TP are more relevant for situations in which cost limitations constrain either phenotyping or genotyping capacity. For instance, if genotyping costs are most limiting, it may be more critical to know what the minimum total population size is at which one can attain either the maximum predictive ability for a given trait, or a predetermined minimum predictive ability; this would correspond to analyses which vary absolute population size. This can be contrasted with a scenario in which for a given trait phenotyping costs are most restrictive, and it is therefore most pressing to consider what the minimum TP size would be, given the need to predict a maximum number of non-phenotyped individuals; this in turn would correspond to analyses which vary relative TP size. Finally, if the total pool of germplasm available is itself limited, it may be most relevant to consider simply what the largest percentage of this population is that can be allocated to the VP without sacrificing predictive ability; this, logically, corresponds to scenarios which vary the relative size of the VP.

It is important to note the general pattern observed for all of the cross-validations reported here – that of an exponential relationship between predictive ability and the variable under consideration. Importantly, the asymptotic maximum predictive ability is reached at low values of SNP density, population similarity, or population size, relative to the marker density, diversity

panel size, and level of population structure in the carrot collection presented in this study. From a practical perspective, it is feasible to attain non-limiting levels of nearly all the determinative factors influencing predictive ability.

Practical considerations in comparing GWAS and GEBVs

Numerous studies have attempted to compare the effectiveness of GWAS and genomic selection strategies in the context of specific crops' breeding programs. In general terms, the conclusions are extremely consistent. GWAS offers a method for screening populations for preliminary gene discovery, and for guiding the development of mapping populations that could validate markers, which in turn could be utilized within a marker-assisted selection breeding scheme. Genomic selection offers the potential for more immediate use of marker information by estimating additive genotypic effects based on relatedness (Minamikawa et al., 2018; Srivastava et al., 2020; Tsai et al., 2020). Numerous reviews, however, conclude that GWAS is poorly suited to the detection of minor effect QTL that underlie quantitative traits (Caballero et al., 2015; Robinson et al., 2014). Even when QTL have been previously detected through interval mapping approaches, marker-assisted selection based on multiple linear regression using QTL-linked markers has been observed to have lower prediction accuracy than genome-wide prediction models (Hadasch et al., 2016).

It is relevant to note the practical implications for the mode of genomic selection that would be enabled on the basis of the predictions made in this study. Following the nomenclature of Falconer, 1996, genotypic value (G) is typically decomposed as: $G = A + D + I$, where A

designates additive or breeding value, D indicates the dominance deviation, and I the interaction or epistatic deviation. What has been estimated in this study are explicitly the additive components of genotypic value, and as such, the portion of a given individual's value that is transmissible to the next generation. While this is potentially also a reasonable surrogate for estimating genotypic value as a whole, given the large additive variance attributed to the traits considered here, this would need to be validated in future studies. The GEBVs calculated here should therefore not be assumed to be a sufficient metric by which to judge *per se* performance, but rather, a basis upon which to select individuals for their breeding value. In addition, while these GEBVs could also be utilized to make predictions as to which inbred lines to select in the production of hybrids, this particular case of estimating progeny value would more appropriately be performed utilizing a genomic prediction model that additionally estimates dominance deviations (Alvarez-Castro and Carlborg, 2007).

As a result, the immediate practical utility of the predictions made in this study would likely be most evident in population improvement efforts. In particular, the most frequent use of diversity panels such as the one utilized in this study is the identification of novel traits that currently do not exist in elite germplasm. Through the introgression of such a trait into breeding lines, market class attributes would likely be impacted; the GEBVs reported here could therefore significantly accelerate the pace at which a particular desired market class is recovered, following such wide crosses. Despite their promise, the actual gains from selection one can expect to attain clearly will vary from trait to trait. While some phenotypes presented here, such as root fill, length, and maximum width, show surprisingly high prediction accuracies,

others, such as biomass and tip curvature, are markedly harder to predict. This is unsurprising, since biomass is clearly a composite trait, much like yield. Subtle variations in the tip will, for their part, likely always be subject to greater environmental variation, and thus be challenging to select for. Nevertheless, despite the lower prediction accuracies for these traits, GEBVs still offer a method for utilizing genomic-scale data to aide in improving the efficiency of selection, whereas GWAS was unable to detect any significant associations for these traits. Given the high-throughput nature of the phenotyping platform used to collect training data, and the relatively limited amount of genotyping required to calculate GEBV, this study provides compelling evidence supporting the inclusion of genomic selection in breeding programs for carrot market class.

References

Alqudah, A. M., Sallam, A., Stephen Baenziger, P., & Börner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. *Journal of Advanced Research*, *22*, 119–135. https://doi.org/https://doi.org/10.1016/j.jare.2019.10.013

Alvarez-Castro, J. M., & Carlborg, O. (2007). A unified model for functional and statistical epistasis and its application in quantitative trait Loci analysis. *Genetics*, *176*(2), 1151–1167. https://doi.org/10.1534/genetics.106.067348

Anand, S., Mangano, E., Barizzone, N., Bordoni, R., Sorosina, M., Clarelli, F., Corrado, L., Martinelli Boneschi, F., D'Alfonso, S., & De Bellis, G. (2016). Next Generation Sequencing of Pooled Samples: Guideline for Variants' Filtering. *Scientific Reports*, *6*(1), 33735. https://doi.org/10.1038/srep33735

Banga, O. (1957). *Origin of the European cultivated carrot*. Instituut voor de Veredeling van Tuinbouwgewassen.

Beavis, W. D., & Paterson, A.H. (1998). QTL analyses: power, precision, and accuracy. *Molecular dissection of complex traits*: 145-162.

Bélanger, S., Esteves, P., Clermont, I., Jean, M., & Belzile, F. (2016). Genotyping-by-Sequencing on Pooled Samples and its Use in Measuring Segregation Bias during the Course of Androgenesis in Barley. *The Plant Genome*, *9*(1). https://doi.org/10.3835/plantgenome2014.10.0073

Berro, I., Lado, B., Nalin, R. S., Quincke, M., & Gutiérrez, L. (2019). Training Population Optimization for Genomic Selection. *The Plant Genome*, *12*(3), 190028.

https://doi.org/https://doi.org/10.3835/plantgenome2019.04.0028

Brachi, B., Morris, G. P., & Borevitz, J. O. (2011). Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology*, *12*(10), 232. https://doi.org/10.1186/gb-2011-12-10-232

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A One-Penny Imputed Genome from Next-Generation Reference Panels. *American Journal of Human Genetics*, *103*(3), 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Caballero, A., Tenesa, A., & Keightley, P. D. (2015). The Nature of Genetic Variation for Complex Traits Revealed by GWAS and Regional Heritability Mapping Analyses. *Genetics*, *201*(4), 1601–1613. https://doi.org/10.1534/genetics.115.177220

Corak, K. E., Ellison, S. L., Simon, P. W., Spooner, D. M., & Dawson, J. C. (2019). Comparison of Representative and Custom Methods of Generating Core Subsets of a Carrot Germplasm Collection. *Crop Science*, *59*(3), 1107–1121. https://doi.org/https://doi.org/10.2135/cropsci2018.09.0602

Daetwyler, H. D., Kemper, K. E., van der Werf, J. H. J., & Hayes, B. J. (2012). Components of the accuracy of genomic prediction in a multi-breed sheep population1. *Journal of Animal Science*, *90*(10), 3375–3384. https://doi.org/10.2527/jas.2011-4557

Driscoll, M. K., McCann, C., Kopace, R., Homan, T., Fourkas, J. T., Parent, C., & Losert, W. (2012). Cell shape dynamics: From waves to migration. *PLoS Computational Biology*, *8*(3),

e1002392. https://doi.org/10.1371/journal.pcbi.1002392

Edwards, S. M., Buntjer, J. B., Jackson, R., Bentley, A. R., Lage, J., Byrne, E., Burt, C., Jack, P., Berry, S., Flatman, E., Poupard, B., Smith, S., Hayes, C., Gaynor, R. C., Gorjanc, G., Howell, P., Ober, E., Mackay, I. J., & Hickey, J. M. (2019). The effects of training population design on genomic prediction accuracy in wheat. *Theoretical and Applied Genetics*, *132*(7), 1943–1952. https://doi.org/10.1007/s00122-019-03327-y

Ellison, S. L., Luby, C. H., Corak, K. E., Coe, K. M., Senalik, D., Iorizzo, M., Goldman, I. L., Simon, P. W., & Dawson, J. C. (2018). Carotenoid presence is associated with the Or gene in domesticated carrot. *Genetics*, *210*(4), 1497–1508. https://doi.org/10.1534/genetics.118.301299

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome*, *4*(3), 250–255. https://doi.org/10.3835/plantgenome2011.08.0024

Endelman, J. B., & Jannink, J.-L. (2012). Shrinkage estimation of the realized relationship matrix. *G3: Genes, Genomes, Genetics*, *2*(11), 1405–1413. https://doi.org/10.1534/g3.112.004259

Erbe, M., Gredler, B., Seefried, F. R., Bapst, B., & Simianer, H. (2013). A Function Accounting for Training Set Size and Marker Density to Model the Average Accuracy of Genomic Prediction. *PLOS ONE*, *8*(12), 1–11. https://doi.org/10.1371/journal.pone.0081046

Ewens, W. J., & Spielman, R. S. (1995). The transmission/disequilibrium test: history, subdivision, and admixture. *American Journal of Human Genetics*, *57*(2), 455–464.

Falconer, D. S. (Douglas S. (1996). *Introduction to quantitative genetics*. Fourth edition. Essex, England : Longman, 1996.

FAO. (2020). *FAOSTAT*. http://www.fao.org/faostat/en/

Hadasch, S., Simko, I., Hayes, R. J., Ogutu, J. O., & Piepho, H.-P. (2016). Comparing the Predictive Abilities of Phenotypic and Marker-Assisted Selection Methods in a Biparental Lettuce Population. *The Plant Genome*, *9*(1), plantgenome2015.03.0014. https://doi.org/https://doi.org/10.3835/plantgenome2015.03.0014

Henderson, C R. (1973). Sire Evaluation and Genetic Trends. *Journal of Animal Science*, *1973*(Symposium), 10–41. https://doi.org/10.1093/ansci/1973.Symposium.10

Henderson, Charles R. (1963). Selection index and expected genetic advance. *Statisitical Genetics and Plant Breeding*, *NAS NRCP Publication 982*.

Hernandez, C. O., Wyatt, L. E., & Mazourek, M. R. (2020). Genomic Prediction and Selection for Fruit Traits in Winter Squash. *G3: Genes, Genomes, Genetics*, *10*(10), 3601–3610. https://doi.org/10.1534/g3.120.401215

Iorizzo, M., Ellison, S., Senalik, D., Zeng, P., Satapoomin, P., Huang, J., Bowman, M., Iovene, M., Sanseverino, W., Cavagnaro, P., Yildiz, M., Macko-Podgórni, A., Moranska, E., Grzebelus, E., Grzebelus, D., Ashrafi, H., Zheng, Z., Cheng, S., Spooner, D., … Simon, P. (2016). A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nature Genetics*, *48*(6), 657–666. https://doi.org/10.1038/ng.3565

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., & Eskin, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics*, *178*(3), 1709–1723. https://doi.org/10.1534/genetics.107.080101

Lozada, D. N., Mason, R. E., Sarinelli, J. M., & Brown-Guedira, G. (2019). Accuracy of genomic

selection for grain yield and agronomic traits in soft red winter wheat. *BMC Genetics*, *20*(1), 82. https://doi.org/10.1186/s12863-019-0785-1

Macko-Podgórni, A., Stelmach, K., Kwolek, K., Machaj, G., Ellison, S., Senalik, D. A., Simon, P. W., & Grzebelus, D. (2020). Mining for Candidate Genes Controlling Secondary Growth of the Carrot Storage Root. *International Journal of Molecular Sciences*, *21*(12), 1–16. https://doi.org/10.3390/ijms21124263

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*, *157*(4), 1819–1829. http://www.genetics.org/content/157/4/1819.abstract

Minamikawa, M. F., Takada, N., Terakami, S., Saito, T., Onogi, A., Kajiya-Kanegae, H., Hayashi, T., Yamamoto, T., & Iwata, H. (2018). Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (Pyrus pyrifolia Nakai). *Scientific Reports*, *8*(1), 11994. https://doi.org/10.1038/s41598-018-30154-w

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., & Buckler, E. S. (2009). Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. *The Plant Cell*, *21*(8), 2194–2202. https://doi.org/10.1105/tpc.109.068437

Olatoye, M. O., Clark, L. V, Labonte, N. R., Dong, H., Dwiyanti, M. S., Anzoua, K. G., Brummer, J. E., Ghimire, B. K., Dzyubenko, E., Dzyubenko, N., Bagmet, L., Sabitov, A., Chebukin, P., Głowacka, K., Heo, K., Jin, X., Nagano, H., Peng, J., Yu, C. Y., … Lipka, A. E. (2020). Training Population Optimization for Genomic Selection in Miscanthus. *G3: Genes, Genomes, Genetics*, *10*(7), 2465–2476. https://doi.org/10.1534/g3.120.401402

Otyama, P. I., Wilkey, A., Kulkarni, R., Assefa, T., Chu, Y., Clevenger, J., O'Connor, D. J., Wright,

G. C., Dezern, S. W., MacDonald, G. E., Anglin, N. L., Cannon, E. K. S., Ozias-Akins, P., & Cannon, S. B. (2019). Evaluation of linkage disequilibrium, population structure, and genetic diversity in the U.S. peanut mini core collection. *BMC Genomics*, *20*(1), 481. https://doi.org/10.1186/s12864-019-5824-9

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Pritchard, J K, & Rosenberg, N. A. (1999). Use of unlinked genetic markers to detect population stratification in association studies. *American Journal of Human Genetics*, *65*(1), 220–228. https://doi.org/10.1086/302449

Pritchard, Jonathan K, Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959. http://www.genetics.org/content/155/2/945.abstract

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Riedelsheimer, C., Endelman, J. B., Stange, M., Sorrells, M. E., Jannink, J.-L., & Melchinger, A. E. (2013). Genomic predictability of interconnected biparental maize populations. *Genetics*, *194*(2), 493–503. https://doi.org/10.1534/genetics.113.150227

Robinson, M. R., Wray, N. R., & Visscher, P. M. (2014). Explaining additional genetic variation in complex traits. *Trends in Genetics*, *30*(4), 124–132. https://doi.org/10.1016/j.tig.2014.02.003

Rosyara, U. R., De Jong, W. S., Douches, D. S., & Endelman, J. B. (2016). Software for Genome-

Wide Association Studies in Autopolyploids and Its Application to Potato. *The Plant Genome*, *9*(2). https://doi.org/10.3835/plantgenome2015.08.0073

Simon, P W. (2000). Domestication, Historical Development, and Modern Breeding of Carrot. In *Plant Breeding Reviews* (pp. 157–190). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470650172.ch5

Simon, Philipp W, Freeman, R. E., Vieira, J. V, Boiteux, L. S., Briard, M., Nothnagel, T., Michalik, B., & Kwon, Y.-S. (2008). *Carrot - Vegetables II: Fabaceae, Liliaceae, Solanaceae, and Umbelliferae* (J. Prohens & F. Nuez (eds.); pp. 327–357). Springer New York. https://doi.org/10.1007/978-0-387-74110-9_8

Srivastava, R. K., Singh, R. B., Pujarula, V. L., Bollam, S., Pusuluri, M., Chellapilla, T. S., Yadav, R. S., & Gupta, R. (2020). Genome-Wide Association Studies and Genomic Selection in Pearl Millet: Advances and Prospects. In *Frontiers in Genetics* (Vol. 10, p. 1389). https://www.frontiersin.org/article/10.3389/fgene.2019.01389

Sverrisdóttir, E., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., Janss, L., Bryan, G., & Nielsen, K. L. (2018). The Value of Expanding the Training Population to Improve Genomic Selection Models in Tetraploid Potato. *Frontiers in Plant Science*, *9*, 1118. https://doi.org/10.3389/fpls.2018.01118

Tayeh, N., Klein, A., Le Paslier, M.-C., Jacquin, F., Houtin, H., Rond, C., Chabert-Martinello, M., Magnin-Robert, J.-B., Marget, P., Aubert, G., & Burstin, J. (2015). Genomic Prediction in Pea: Effect of Marker Density and Training Population Size and Composition on Prediction Accuracy. *Frontiers in Plant Science*, *6*, 941. https://doi.org/10.3389/fpls.2015.00941

Tsai, H.-Y., Janss, L. L., Andersen, J. R., Orabi, J., Jensen, J. D., Jahoor, A., & Jensen, J. (2020).

Genomic prediction and GWAS of yield, quality and disease-related traits in spring barley and winter wheat. *Scientific Reports*, *10*(1), 3347. https://doi.org/10.1038/s41598-020-60203-2

Turner, S. D., Senalik, D. A., Simon, P. W., Ellison, S. L., Miller, N. D., & Spalding, E. P. (2018). An Automated Image Analysis Pipeline Enables Genetic Studies of Shoot and Root Morphology in Carrot (Daucus carota L.). *Frontiers in Plant Science*, *9*(1703), 1–17. https://doi.org/10.3389/fpls.2018.01703

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Wang, Q., Yu, Y., Yuan, J., Zhang, X., Huang, H., Li, F., & Xiang, J. (2017). Effects of marker density and population structure on the genomic prediction accuracy for growth trait in Pacific white shrimp Litopenaeus vannamei. *BMC Genetics*, *18*(1), 45. https://doi.org/10.1186/s12863-017-0507-5

Wu, X.-L., Xu, J., Feng, G., Wiggans, G. R., Taylor, J. F., He, J., Qian, C., Qiu, J., Simpson, B., Walker, J., & Bauck, S. (2016). Optimal Design of Low-Density SNP Arrays for Genomic Prediction: Algorithm and Applications. *PLOS ONE*, *11*(9), 1–36. https://doi.org/10.1371/journal.pone.0161719

Xu, Y., Wang, X., Ding, X., Zheng, X., Yang, Z., Xu, C., & Hu, Z. (2018). Genomic selection of agronomic traits in hybrid rice using an NCII population. *Rice*, *11*(1), 32. https://doi.org/10.1186/s12284-018-0223-4

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A unified mixed-

model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, *38*(2), 203–208. https://doi.org/10.1038/ng1702

Zhang, A., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., San Vicente, F., Olsen, M., Prasanna, B. M., Crossa, J., Yu, H., & Zhang, X. (2017). Effect of Trait Heritability, Training Population Size and Marker Density on Genomic Prediction Accuracy Estimation in 22 bi-parental Tropical Maize Populations. *Frontiers in Plant Science*, *8*, 1916. https://doi.org/10.3389/fpls.2017.01916

Zhang, X., Pérez-Rodríguez, P., Semagn, K., Beyene, Y., Babu, R., López-Cruz, M. A., San Vicente, F., Olsen, M., Buckler, E., Jannink, J.-L., Prasanna, B. M., & Crossa, J. (2015). Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. *Heredity*, *114*(3), 291–299. https://doi.org/10.1038/hdy.2014.99

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M., & Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics*, *42*(4), 355–360. https://doi.org/10.1038/ng.546

Tables

Table 1. Significant SNPs associated with root shape traits. Associations were found for four traits (root fill, maximum width, length, and aspect ratio) using a diversity panel of 661 carrot accessions genotyped for 146,821 SNPs, and phenotyped using the methods of described in Chapter 2. Chromosome, position, LOD score, and additive effect are listed for the most significant SNP in each peak exceeding the permutation test-derived LOD threshold.

Trait	**Chromosome**	**Position (bp)**	**Score**	**Effect**
Root fill	2	47341762	8.09	-0.912
Max Width	3	58042921	6.309098	-1.71
	7	6622754	6.82	-2.35
Length	2	42684849	5.89	-17.2
	5	34380903	6.51	-13
L/W Ratio	1	52762815	6.11	0.655
	3	7295883	7.06	0.661
	3	56902806	6.39	0.703
	4	33224217	7.9	0.543
	6	37411769	6.37	1.02

Table 2. Prediction error variance (PEV), and predictive ability of random 100-fold cross validation (Pred. ability – avg), and when relatedness between TP and VP is minimized (Pred. ability – min) for seven carrot root traits, phenotyped using roots grown across two locations and three growing seasons.

Trait	**PEV**	**Pred. ability - avg**	**Pred. ability - min**
Biomass	0.71	0.67 (± 0.11)	0.29
Root fill	0.85	0.86 (± 0.03)	0.17
Max width	0.78	0.72 (± 0.06)	0.30
Length	0.80	0.77 (± 0.06)	0.29
L/W ratio	0.92	0.82 (± 0.05)	0.37
Tip curvature	0.41	0.25 (± 0.04)	0.01
Shoulder curvature	0.35	0.63 (± 0.05)	0.10

Figures

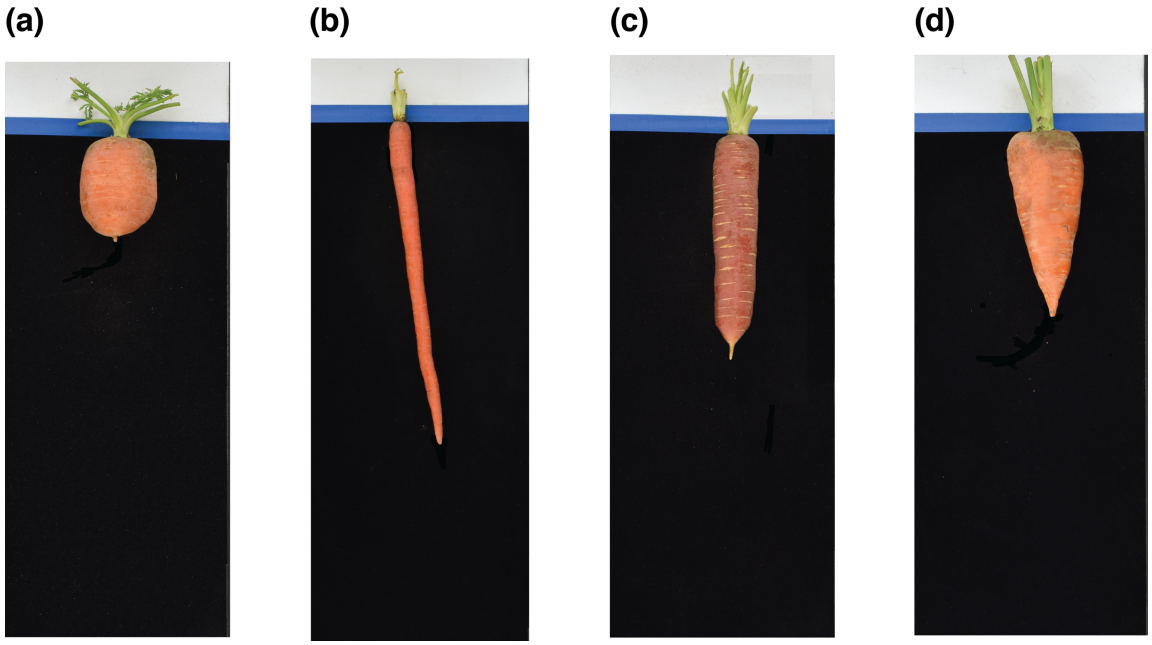


Fig. 1 Representative roots from four distinct market classes, exemplifying variation in the four traits for which GWAS detected significant association with markers. **(a)** and **(b)** illustrate variation for length and maximum width. **(a)** A Parisienne-type carrot; **(b)** An Imperator-type. **(c)** and **(d)** illustrate variation for aspect ratio and root fill. **(c)** A Nantes-type; **(d)** A Chantenay-type

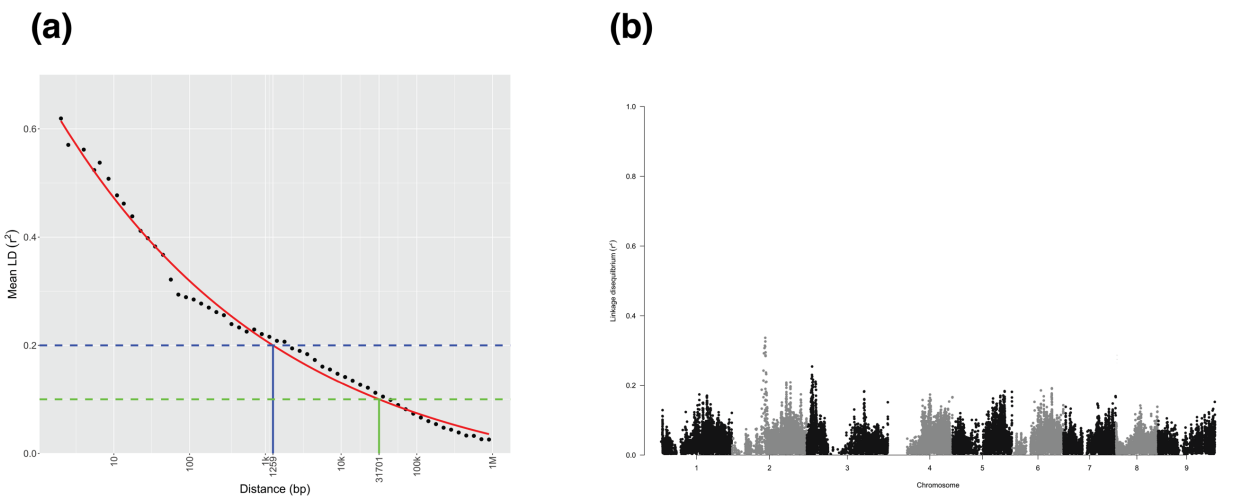


Fig. 2 Genome-wide linkage disequilibrium in the diversity panel of 661 carrot accessions, using 29,456 SNPs represented in terms of a decay function (a), and Manhattan plot (b). **(a)** Average LD is plotted against the genetic distance between pairs of markers (black dots; log scale), and a self-starting asymptomatic decay function is fit to the data (red). Intercepts with r^2 values of 0.2 (blue line) and 0.1 (green line) are indicated as 796 and 19.7k bp, respectively. **(b)** LD calculated on a sliding-window basis (the mean of a given SNP and its 100 nearest neighbors) is represented as a Manhattan plot

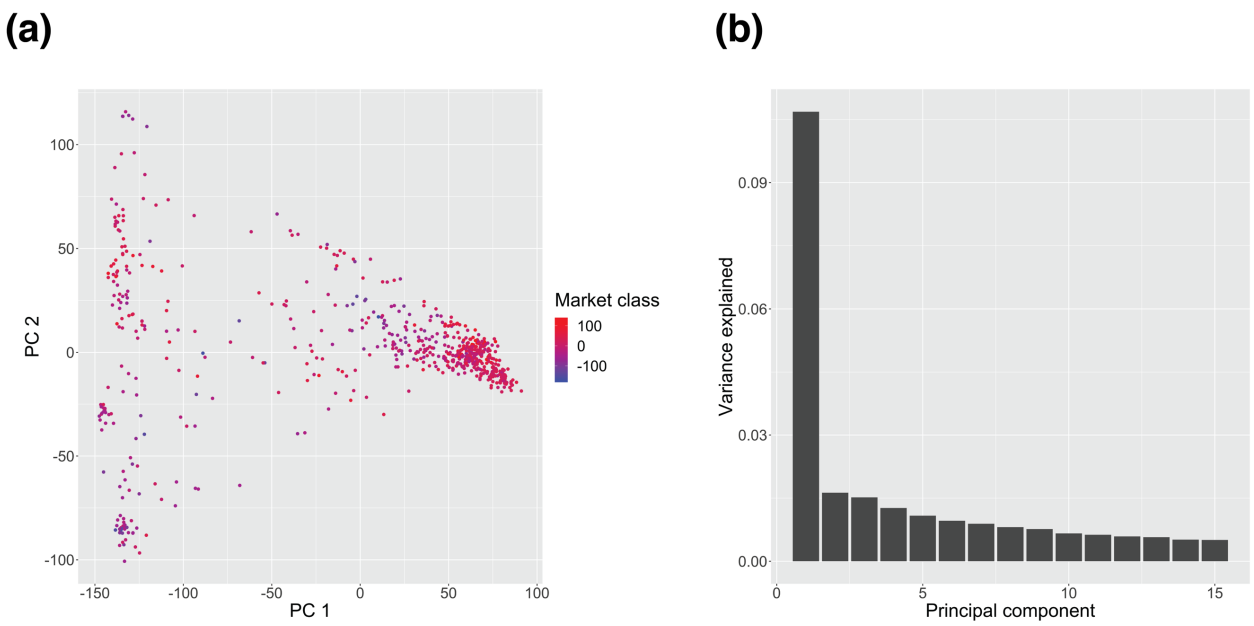


Fig. 3 PCA-based visualization of population structure in the carrot diversity panel. **(a)** PCA bi-plot representing all accessions in the diversity panel according to their scores along first (*x*-axis) and second (*y*-axis) principal component. Points are colored according to a quantitative measure of their market class following the methods of Brainard, et al. (2021). **(b)** Scree plot of the variance explained by the first 15 principal components in marked-based PCA

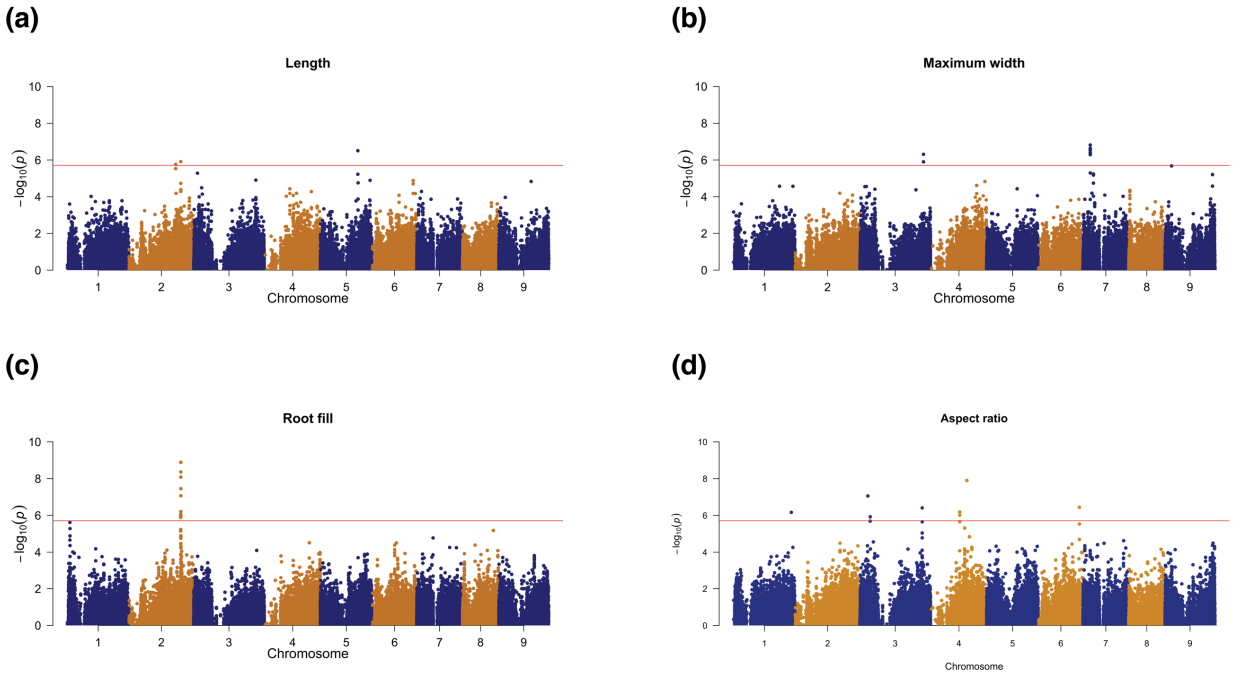


Fig. 4 Manhattan plots for GWAS results using a diversity panel of 661 carrot accessions and 146,821 SNPs. Significant associations were found for 4 of the digital phenotypes extracted through the image analysis pipeline. **(a)** Root length; **(b)** Maximum width; **(c)** Root fill (the score of the first PC of the length- and width-normalized root contour; **(d)** Aspect ratio (length / maximum width)

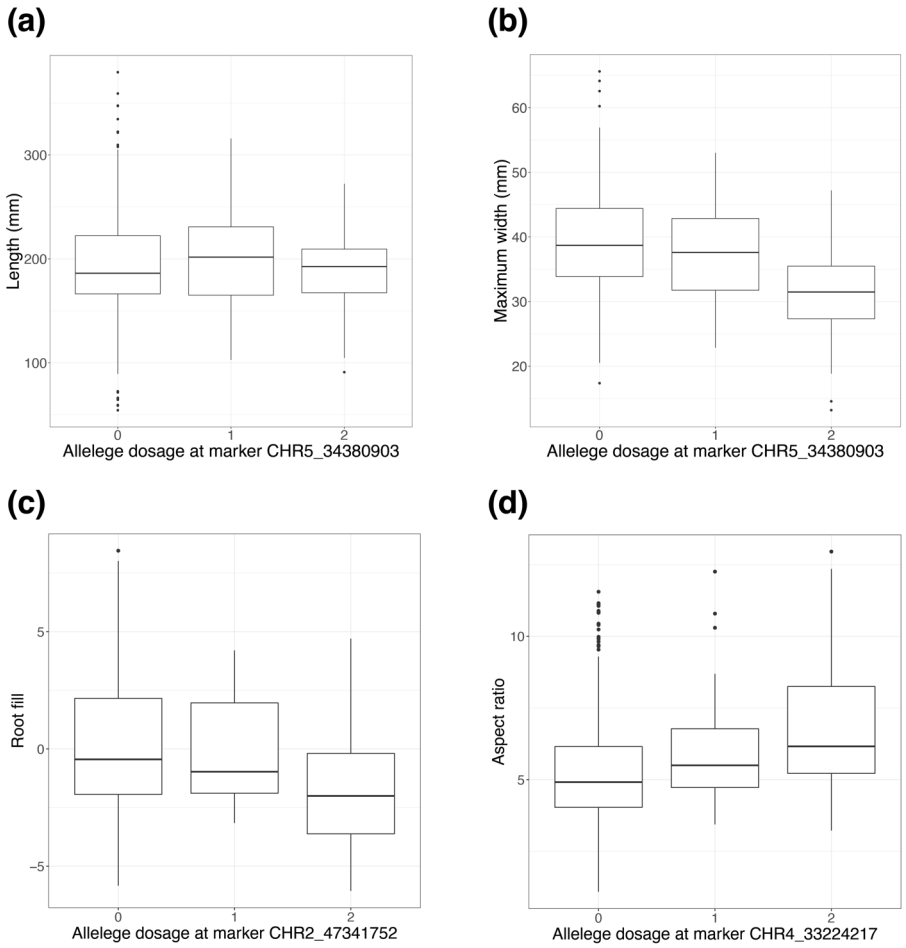


Fig. 5 Effect sizes of the most significant SNPs for each of the traits shown in Figure 3. **(a)** Root length; **(b)** Maximum width; **(c)** Root fill; **(d)** Aspect ratio

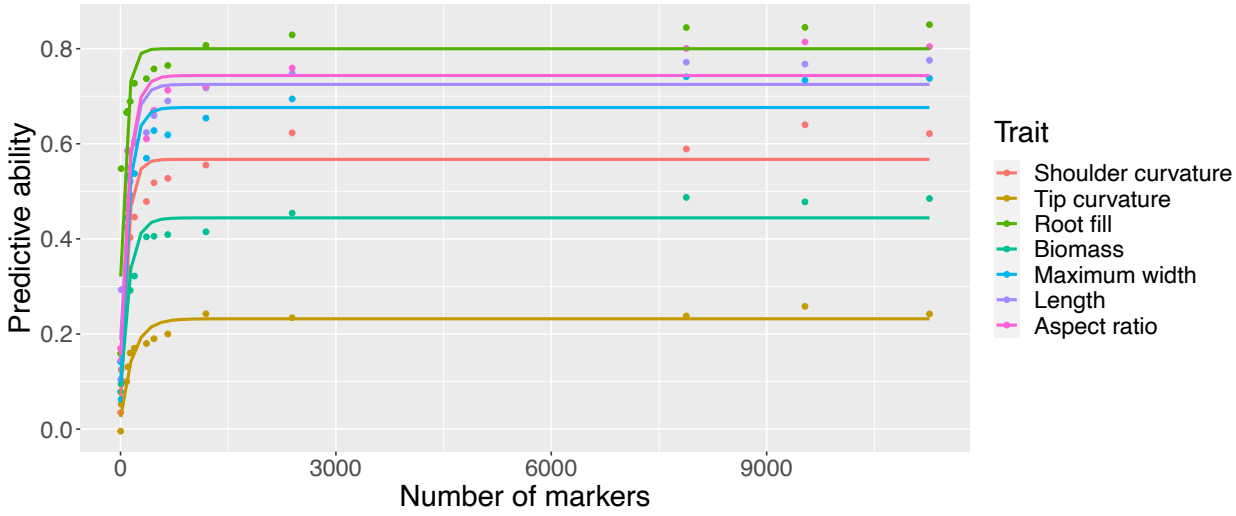


Fig. 6 Effect of marker density on predictive ability for six root size and shape traits, as well as total biomass, using the full carrot diversity panel. Curves follow an asymptotic exponential function, reaching their maximum at approximately 2500 markers

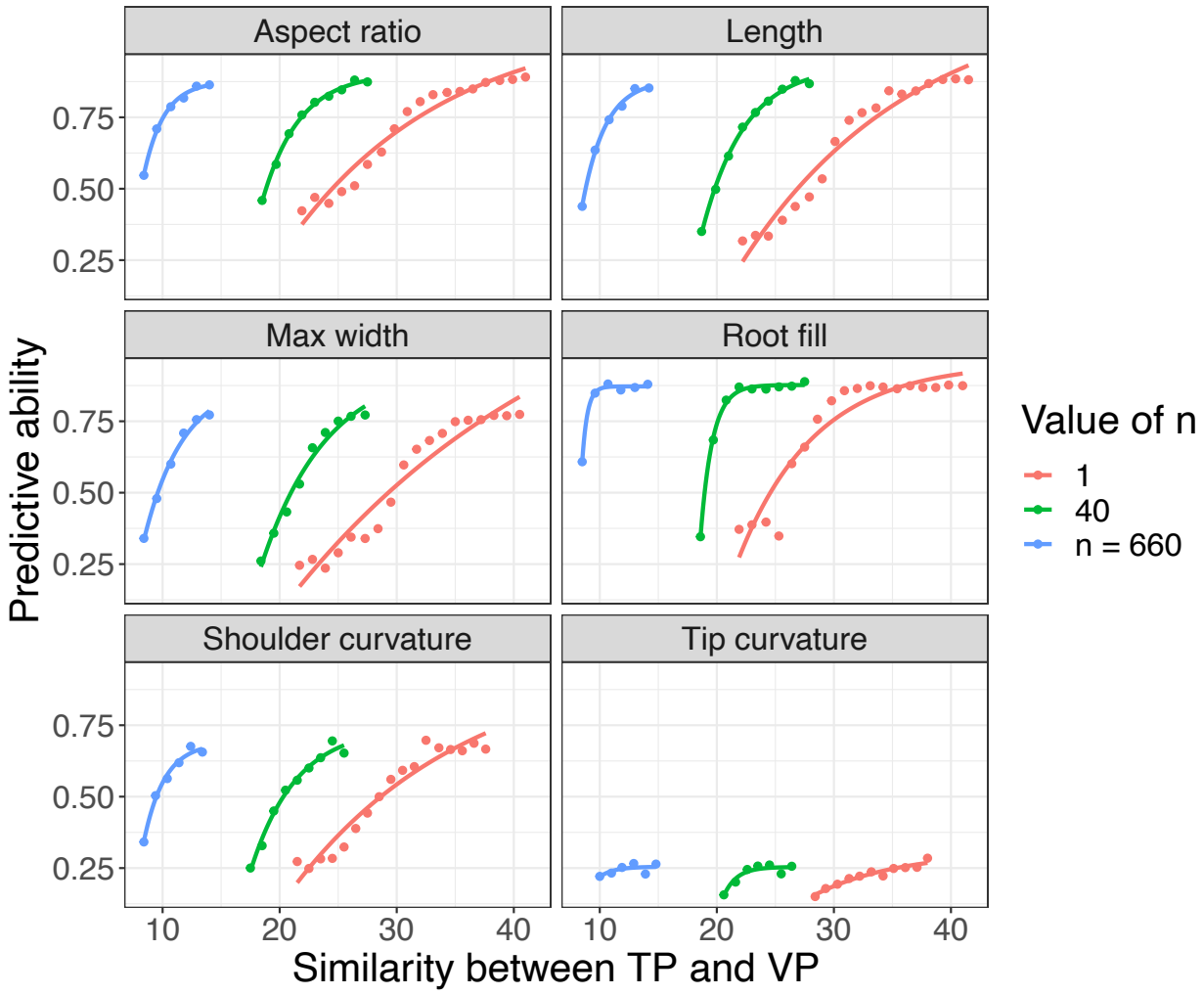


Fig. 7 Effect of relatedness between training population (TP) and validation population (VP) on prediction accuracy for the six key size and shape traits which constitute market class, using the full carrot diversity panel. Similarity is defined as the inverse of the Euclidean distance matrix, and *n* indicates the number of individuals in the TP against which each member of the VP is compared against

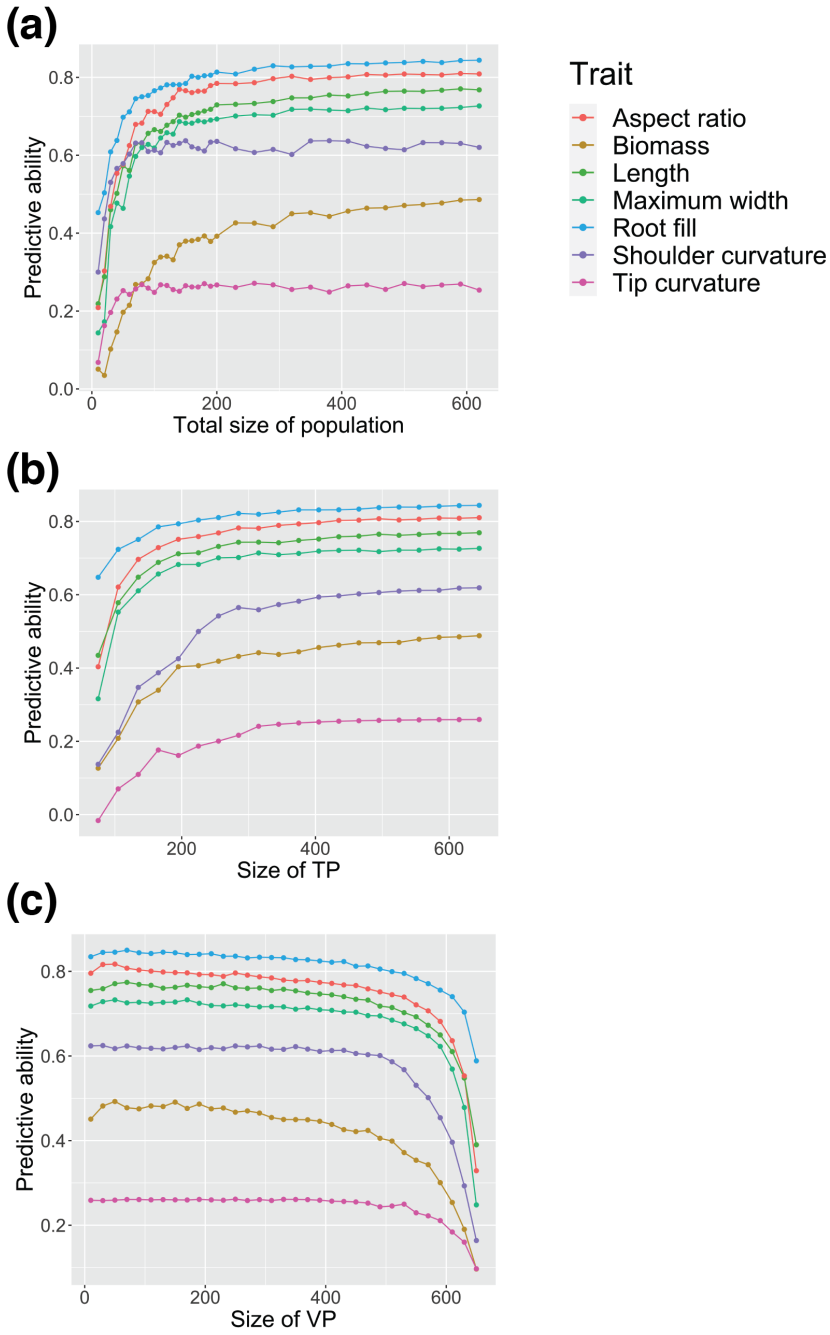


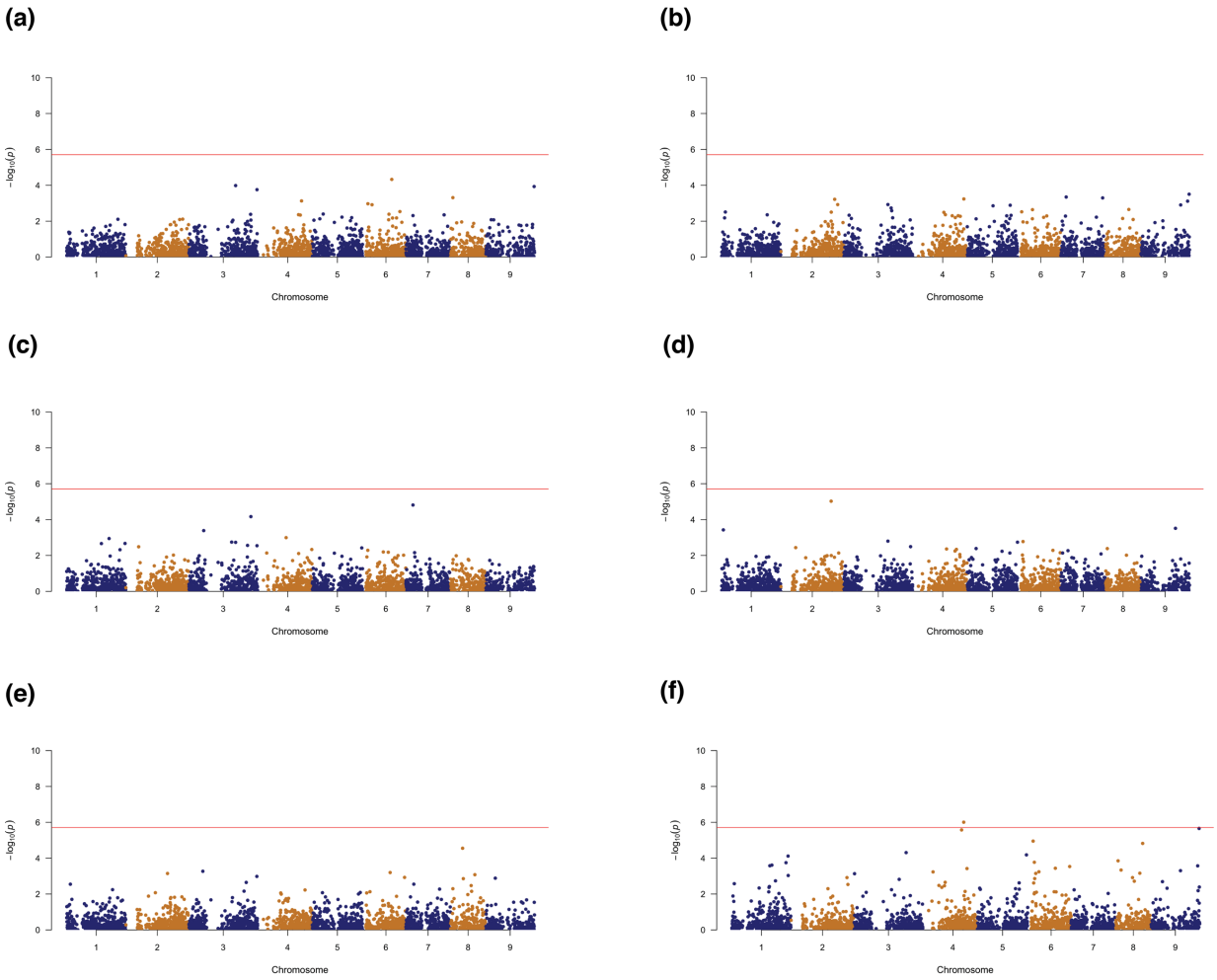
Fig. 8 Effect of population size on prediction accuracy for the six traits underlying carrot market class, as well as total biomass. **(a)** Absolute population size is varied, holding the VP at 10% of the total population; **(b)** The relative size of the TP is varied by holding the VP constant at 60 individuals; **(c)** Relative VP size is varied by keeping total population size constant at 661 individuals

Supplementary Tables

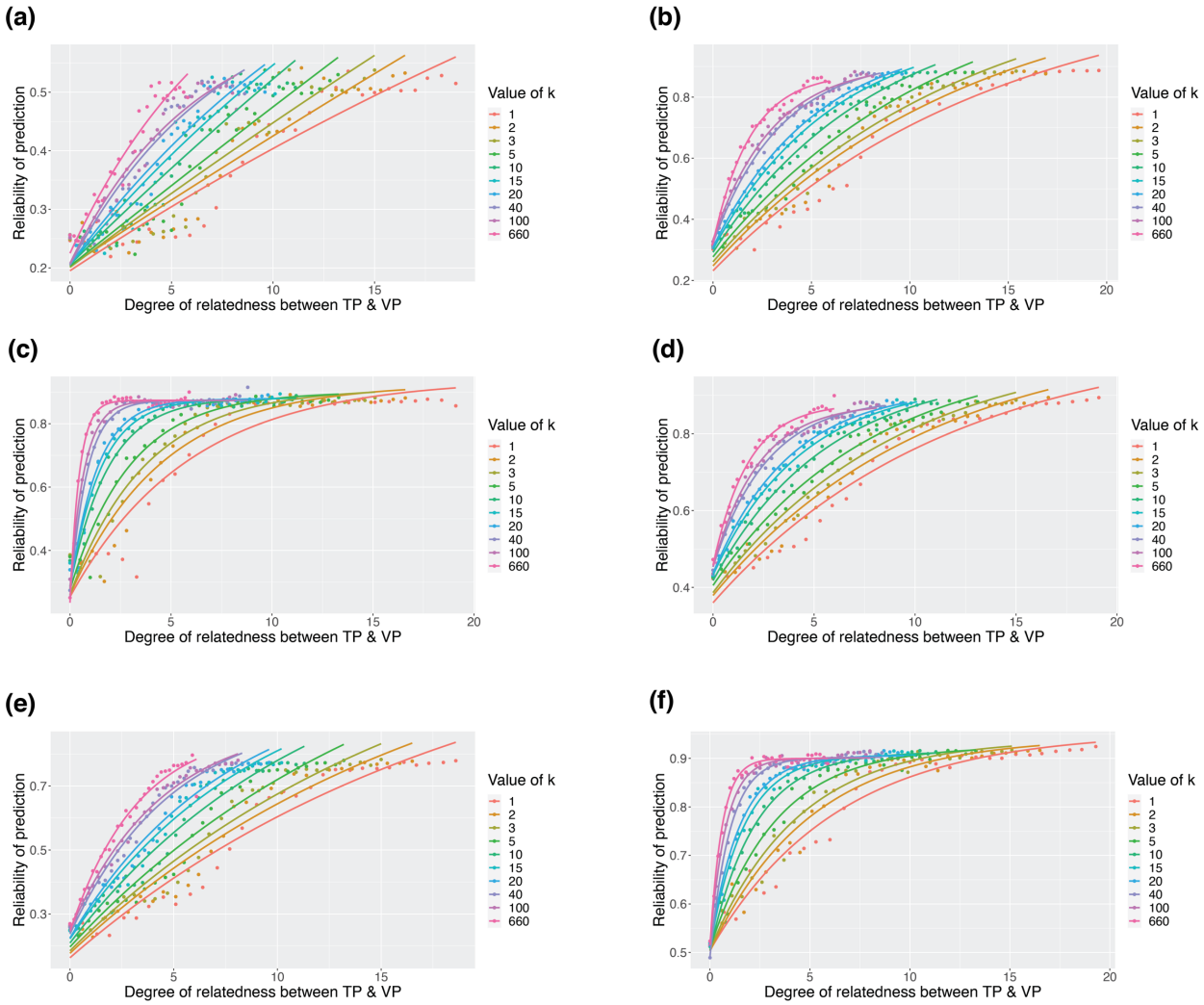
Supplementary Table 1. GEBVs for the seven carrot root traits listen in Table 2. The entire diversity panel of 661 carrot accessions was utilized for phenotyping, but only SNPs located on chromosome 3 were used in estimating genetic variance.

Trait	h^2	**Pred. accur. (avg)**	**Pred. accur. (min)**
Shoulder curvature	0.62	0.53	0.14
Tip curvature	0.04	0.09	0.07
Root fill	0.62	0.77	0.07
Biomass	0.36	0.38	0.10
Max width	0.61	0.65	0.08
Length	0.65	0.68	0.17
L/W ratio	0.80	0.72	0.15

Supplementary Figures



Supplementary Figure 1. Manhattan plots for GWAS results using a diversity panel of 661 carrot accessions and only SNPs located on chromosome 3. No significant associations were found for 4 of root traits for which associations were previously detected (root length **(a)**; maximum width **(b)**; root fill **(c)**; and aspect ratio **(d)**), nor for tip curvature **(e)**, or shoulder curvature **(f)**.



Supplementary Figure 2. Prediction accuracy versus population structure at a variety of values of *k*, illustrating the variable goodness of fit of the asymptotic exponential function for 6 traits: **(a)** biomass; **(b)** length; **(c)** root fill; **(d)** aspect ratio; **(e)** maximum width; **(f)** tip curvature. The entire diversity panel of 661 carrot accessions was used for phenotyping.

Appendix

For complete listing and description of genotypes utilized in the diversity panel, please refer to Supplementary_File_S3_1.csv

Chapter 4:
Future directions

This thesis is oriented toward answering the question: what is the nature of the genetic control of market class in carrot? While the various components of these experiments built upon recent advances in the fields of quantitative image analysis, bioinformatics, and statistical genomics, this particular research question had, to date, not been itself thematized in this way. The result has been an advance in the methods necessary to study the complex phenotypes underlying root shape, our understanding of the genetic architecture of market class traits, and novel strategies for improving the efficiency of breeding for particular market classes. At the same time, this research also represents a starting point for numerous subsequent lines of inquiry that will hopefully build upon the experiments presented here.

First, with regards to the digital phenotyping methods presented in Chapter 2, there are several clear directions in which the imaging platform could be improved to increase both phenotyping precision, as well as its high-throughput capacity. In terms of the former, although the 2D imaging methods presented here were found to be suitable for quantifying traits that can be measured in two-dimensions (by validating that phenotypic measurements did not vary as the root was rotated about its long axis), there are of course aspects of root morphology that only exist in three dimensions, which are also important to market class and carrot quality. The degree of concavity in the shoulder, at the point where the petioles attach to the root crown, for instance, is not an aspect of root shape that can be easily studied using the imaging pipeline developed here. A 3D imaging platform would thus expand the range of phenotypes accessible to image analysis.

An additional outstanding challenge is a phenotyping methodology that is able to track root development throughout a growing season. Understanding the temporal dynamics that underlie the formation of different market classes would be particularly useful in attempting to utilize gene expression data (e.g., in the form of RNAseq) to validate certain candidate genes as playing a causative role in driving the physiological processes leading to specific root sizes or shapes at maturity. However, at the moment, it remains a technical hurdle yet to be cleared to either develop a production system amenable to existing methods of digital image acquisition, or a digital imaging platform amenable to the field production of carrots. Regarding the former, it has been observed that carrots grown from seed under hydroponic conditions, artificial growth media, or even in field soils in pots exhibit abnormal growth, with atypical root swelling, excessive branching, and stunted lengths. Thus, developing controlled conditions under which repeated destructive harvests could be performed, or automated imaging at regular intervals could be carried out, remains technically challenging. Alternatively, regarding the latter, there have been advances in remote sensing technologies to model root system development below-ground, throughout the growing season. Implementation of such a system for carrot production would be extremely helpful, and remove the need to replicate field conditions in a greenhouse environment.

Secondly, from a high-throughput perspective, while the imaging system developed in this study offers significant advances over traditional manual phenotyping and record-keeping, it still requires monitoring and involvement by human researchers. Automation of the image-acquisition process remains a key bottleneck that would dramatically increase throughput, and

fully realize the potential of this digital phenotyping methodology. This would facilitate the dramatic expansion of the scale of experimental designs and replications that could be analyzed in a research context, as well as make the implementation of these methods feasible in the practical setting of a commercial breeding program.

There are several additional avenues that could productively be pursued with regards to the genetic analyses presented in this thesis as well. First, the diallel experiment described in Chapter 2 represents a convincing proof-of-concept that such mating designs can provide useful estimates of the heritabilities of the phenotypic components of market class, and that these heritabilities are, in many cases, relatively high for root size and shape traits. In this study, a predominant production environment for Wisconsin growers was selected – the muck soils of southern Wisconsin – but while common, this environment clearly represents a unique combination of abiotic and biotic conditions. To better refine these estimates, gain an understanding of potential genotype-by-environment effects for these traits, and make more broadly applicable conclusions regarding SCA effects for particular hybrid combinations, it would be worthwhile to extend this diallel mating scheme to multiple environments.

While the results of the GWAS analyses were relatively limited in this study, they could be potentially improved through several methods. First, with regards to phenotyping, while more than 10,000 roots were utilized in phenotyping the diversity panel, the degree of replication in the experimental design used was not extensive, with only two replications in the California environment, and one replication in both of the years that roots were grown in the Wisconsin

environment. To some extent this was due to logistical constraints on growing such a large collection containing over 700 accessions, and thus it could be profitable to grow a smaller subset of this collection with additional replication, to try to minimize environmental error. In this regard the methods of estimating variable degrees of relatedness between individuals using a dissimilarity matrix constructed from the PCA of GBS-derived markers, as implemented to evaluate GEBV predictive ability in Chapter 3, could be particularly useful. By deliberately sampling a subset that contains individuals that are maximally unrelated to every other individual in the subset, high proportions of total genetic diversity could be obtained while significantly lowering both genotyping and phenotyping costs, and likely not losing potentially informative loci. It is also possible that if a greenhouse environment were developed, as described above, allowing for normal root shape development under more controlled conditions, that this would also facilitate a minimization of environmental variance.

Second, with respect to genotyping, only a minority of the accessions included in this diversity panel constituted inbred lines. The vast majority are either landraces, open-pollinated cultivars, or populations. As a result, representing these accessions with the genotype of a single root necessarily obscures the heterozygosity present in these genotypes. Techniques such as PoolSeq, which utilize bulked DNA from multiple individuals for GBS sequencing, and thus continuous measures of allele *frequency* instead of categorical allele *dosages* in the ANOVAs used to detect associations between SNPs and a given phenotype, have been shown to increase power in such situations.

Neither of these steps are guaranteed to improve the results of the GWAS to the point where the majority of the phenotypic variances observed between market classes is captured in detected quantitative trait loci (QTL). In the case of highly polygenic traits – which the morphological components of market class appear to be – power to detect small-effect QTL will likely remain low, despite the relatively high heritabilities associated with these root shape phenotypes. It was a recognition of this limitation that led to the consideration of the comparative value of GEBVs in utilizing genomic-scale data in selection for these traits. With regard to this approach, both of the above considerations for how to improve power in a GWAS context would likely apply to further improving the accuracy of GEBVs: both having a greater amount of phenotypic data with which to characterize the training population, as well as a PoolSeq derived marker matrix with which to construct the additive relationship matrix, and thus estimate genotypic variance components, would be advantageous in making genomic predictions. In addition, as discussed in Chapter 3, beyond simply estimating GEBVs, variance components could also be estimated to construct covariance matrices describing dominance deviations, thus allowing prediction to be made for accessions' complete genotypic value, and not just their breeding values. This could expand the utility of such genomic selection methods to hybrid prediction, as well as selection among progeny families.

Despite the limitations of the current study, and the small effects of the QTL detected, the significant associations between SNPs and trait variation should not be disregarded, particularly given that for some traits, such as aspect ratio, these QTL did explain non-negligible degrees of phenotypic variation. As a result, it would be appropriate to begin evaluation of these QTL

using fine mapping techniques, in order to investigate candidate genes. Using the annotations provided in the current version of the carrot reference genome, it should be possible to search in the regions around these significant SNPs for putative genes. Fine mapping of these regions could then facilitate the identification of SNPs in coding regions or regulatory sequences, which could provide direct evidence of the role of such genes in the molecular pathways that produce the phenotypic variation seen in this study.

Finally, as another alternative to GWAS not explored in this study, linkage mapping offers a method that aims, as with GWAS, to detect associations between molecular markers and phenotypic variation. However, by designing populations using specific controlled crosses, it is possible to both estimate recombination frequencies between markers, and ensure high minor allele frequencies (MAF) across all loci. A limitation of GWAS diversity panels is that, due to the lack of controlled crosses between inbred parents, the range of MAFs can extend to much lower values than in a classic BC or F_2 mapping population. This can limit power to detect QTL, due to the fact that the genetic variance attributable to a given bi-allelic QTL is proportional to the minor allele frequency.

In addition, using recombination frequencies, genotypes across a grid of pseudomarkers laid out at a specified density can be used in linkage analysis, potentially improving resolution in identifying the genomic region associated with a marker, as opposed to simply finding a single marker that is highly associated with the trait of interest (as in GWAS). For all of these reasons, crosses between inbred lines from Prof. Goldman and Prof. Simon's breeding programs were

made in 2017, and the resulting F_1 populations have been advanced through single-seed descent for the last 3 years. The resulting segregating F_3 families that this work produced are currently being actively utilized as a basis for such linkage mapping.

In conclusion, this study represents an initial attempt at the rigorous phenotyping of market class traits in carrot, and the deployment of quantitative genetic analysis tools in an attempt to improve our understanding of their control. As such, while the results presented in this thesis are compelling in their own right, there are numerous avenues by which these methods can be refined, and additional approaches that could improve upon each aspect of this work. Hopefully the promising results of these experiments act as an impetus to many additional projects, which will continue to deepen our understanding of carrot market class, and aide the deployment of next-generation sequencing technologies and high-throughput phenotyping in the breeding for these important traits.