Received: 1 February 2021 | Accepted: 27 July 2021

DOI: 10.1002/ppj2.20023

The Plant Phenome Journal OPEN ACCESS

REVIEW AND INTERPRETATION

# Images carried before the fire: The power, promise, and responsibility of latent phenotyping in plants

**Mitchell J. Feldmann**[1] | **Joseph L. Gage**[2] | **Sarah D. Turner-Hissong**[3] | **Jordan R. Ubbens**[4]

[1] Dep. of Plant Sciences, Univ. of California - Davis, Davis, CA 95616, USA

[2] Institute for Genomic Diversity, Cornell Univ., Ithaca, NY 14853, USA

[3] Bayer Crop Science, Chesterfield, MO 63017, USA

[4] Global Institute for Food Security, Univ. of Saskatchewan, Saskatoon, SK S7N 0W9, Canada

**Correspondence**
Mitchell J. Feldmann, Dep. of Plant Sciences, Univ. of California, Davis, CA 95616, USA.
Email: mjfeldmann@ucdavis.edu

Assigned to Associate Editor Seth Murray.

**Abstract**

Understanding the genetic basis of plant traits requires comprehensive and quantitative descriptions of the phenotypic variation that exists within populations. Cameras and other sensors have made high-throughput phenotyping possible, but image-based phenotyping procedures involve a step where a researcher selects the traits to be measured. This feature selection step is inherently prone to human biases. Recently, a set of phenotyping approaches, which are referred to collectively as latent phenotyping techniques, have arisen in the literature. Latent phenotyping techniques isolate a latent source of variance in the data, such as stress or genotype, and then quantify the effect of this latent source of variance using latent variables without defining any conventional traits. In this review, we discuss the differences between, and challenges of, both traditional and latent phenotyping.

## 1 | INTRODUCTION

Affordable cameras, data storage, and computational resources have drastically simplified the process of imaging plants in greenhouse or field settings (Chitwood & Topp, 2015; Fahlgren, Gehan, et al., 2015; Gehan & Kellogg, 2017; Granier et al., 2006; Horgan, 2001; Ke et al., 2013; Li et al., 2014; Manacorda & Asurmendi, 2018; Merieux et al., 2021; Omori & Iwata, 1998; Rymaszewski et al., 2018; van der Heijden et al., 2012; Vasseur et al., 2018; Vőfély et al., 2019; York & Lobet, 2017). Although image acquisition is relatively straightforward, methods for processing image data lag in usability and generalizability. When a research group develops an image-based phenotyping tool to study particular traits within a specific organism, adoption of that tool by other groups can be hindered by differences in image acquisition, population type, parameter settings (e.g. thresholds), tissue specificity, and curated metadata (Granier et al., 2006; Lobet, 2017; van der Heijden et al., 2012; Zhou et al., 2018). As a result, publicly available image-based phenotyping tools can prove to be ad-hoc solutions with limited utility beyond the group that developed them (Lobet, 2017). A recent study, in which different groups were given the same image dataset and asked to test the same set of hypotheses, observed sizable variation in the statistical results obtained by each group (Botvinik-Nezer et al., 2020; Pieruschka & Schurr, 2019; Zhou et al., 2018). In this review, we discuss the potential benefits and shortcomings of latent phenotyping, a family of machine learning techniques that presents opportunities to overcome issues related to selection bias, consistency, and broader applicability often encountered with image or sensor-based phenotyping. Further, we discuss how current dogmatic thought about complex traits may obscure the ability to capture a holistic view of true variability,

**Abbreviation:** CNNs, convolutional neural networks; IP, intellectual property; PCA, principal component analysis; PLSR, partial least squares regression; RGB, red, green, blue

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *The Plant Phenome Journal* published by Wiley Periodicals LLC on behalf of American Society of Agronomy and Crop Science Society of America

emphasizing the need for continued development of quantitative and mathematical approaches which will help to decipher genetic mechanisms underlying trait variation and to support crop improvement efforts. The ideas presented in this review share parallels with the allegory of Plato's cave: Prisoners trapped in a cave perceive shadows projected by flames onto the cave wall as reality, despite the truth that these shadows do not provide an accurate representation of the outside world. In the case of plant phenotyping, researchers view limited representations of the true plant phenotype via conventional metrics and/or latent space representations.

**Core Ideas**

- Latent phenotyping is an emerging field that captures fine scale variation in plants.
- Latent phenotypes can minimize the effect of human bias when describing plant form from high dimensional data.
- Latent phenotyping is useful for crop improvement and basic biology.

## 2 | THE CAVE, THE FIRE, AND THE SHADOWS IN PLANT PHENOTYPING

Quantitative plant phenotyping has historically been dominated by conventional traits that are intuitive to define, measure, and interpret: yield, time to flowering, plant height, dimensions of various organs, and many others (Atwell et al., 2010; Buckler et al., 2009; Dong et al., 2018; Fernandes et al., 2018; Liang et al., 2020; Li et al., 2014; Murray et al., 2009; Peiffer et al., 2013, 2014; Poland & Nelson, 2011; Rice et al., 2020; Valluru et al., 2019). These conventional traits capture only a limited slice of true variability yet represent the traits of primary interest to plant breeders because of their commercial and economic value. However, other traits that are also key selection targets, such as quality (e.g., taste, smell, texture), root and shoot system architectures (e.g., leaf and root angle, branching complexity, compactness), and organ form (e.g., size and shape), are more challenging to define with a single variable (Fan et al., 2021; Farneti et al., 2017; Klee & Tieman, 2013; Larsen et al., 2019; Li et al., 2020; Migicovsky et al., 2016; Migicovsky, Sawler, et al., 2017; Oh et al., 2021; Schwieterman et al., 2014; Tieman et al., 2017; Verma et al., 2017; Yano et al., 2019; Zhao et al., 2019). These traits, unlike yield, plant height, or time to flowering, are inherently multidimensional, resulting in a proliferation of potentially oversimplified definitions and interpretations.

For those whose life work involves looking at plants, immediate identification of unique and useful variation needs only a quick glance; indeed, the 'breeder's eye' is often spoken of as an undefinable instinct for selecting superior varieties (Bernardo, 2001, 2016, 2020; Dillmann & Guérin, 1998; Van Ginkel et al., 2008; Zhou et al., 2018). Though the plant phenotype is infinite (Chitwood & Topp, 2015) and can never be quantified in its entirety, images and sensor data that describe plants at high resolution provide an avenue to integrate multiple subtle signals and perspectives into a more comprehensive description of plant form and function (Das et al., 2015; Fahlgren, Feldman, et al., 2015; Fahlgren, Gehan, et al., 2015; Gehan et al., 2019; Gehan et al., 2017; Gehan & Kellogg, 2017; Herritt et al., 2020; Liu et al., 2020; Seethepalli et al., 2020; Tovar et al., 2018; Willis et al., 2017; York & Lobet, 2017).

Over the last 20 years, researchers have applied methods for comprehensively describing plant form with varying strategies and input requirements. These strategies include supervised morphometric methods and more abstract concepts that rely upon unsupervised methods, e.g., latent space phenotyping, persistent homology (Table 1). Many of the strategies that the community has drawn upon originated in disparate fields of study, from psychology, e.g., generalized procrustes analysis and structural equation modeling (Bookstein, 1997; Chitwood, 2020; Gower, 1975; Klingenberg, 2015; Klingenberg & McIntyre, 1998; Rohlf, 1999; Slice & Stitzel, 2004) to computer vision and machine learning, e.g., eigenshapes, convolutional neural networks (CNNs) (Horgan, 2001; Horgan et al., 2001; Sirovich & Kirby, 1987; Turk & Pentland, 1991; Wang et al., 2020). In fact, early work in latent phenotyping closely mirrored approaches used for human face recognition, but instead to capture internal color pattern variation of carrot roots (Horgan, 2001; Horgan et al., 2001).

More recently, Chitwood & Topp, 2015 discussed the idea of a plant 'cryptotype': a function of discrete trait values that can be used to discriminate between a priori defined groups of individuals. Practically, developing a cryptotype function depends on training a supervised model with pre-existing features and labels. Ubbens et al. (2020) developed the supervised method Latent Space Phenotyping, which uses CNNs to differentiate between treatments based on longitudinal image data. Images are first embedded in a low-dimensional latent space, and their paths through that space over time are used to assess temporal change in plant responses to treatment. In a similar but unsupervised fashion, Gage et al. (2019) used both CNNs and principal component analysis (PCA) on flattened lidar point clouds to produce unsupervised latent phenotypes from a single time point evaluation of field-grown maize hybrids. Further, several studies have demonstrated that PCA is a powerful, unsupervised tool for quantifying plant organ shape, as demonstrated in carrot (Horgan et al., 2001; Turner et al., 2018), radish (Iwata et al., 2000, 2004),

**TABLE 1** Abstract concepts in plant phenotyping and some examples of the conventional and latent traits often used to quantify them in existing literature

| Abstract concept | Conventional traits | Latent traits |
|---|---|---|
| Root/shoot architecture and form | Taproot length, overall root length, number of branches, root angles, biomass distribution, branching angle, leaf angle, compactness, width, height, area, aspect ratio, circularity, roundness, solidity | Principal component (PC) scores and autoencoder latent representation, biomass profile PC scores; persistent homology barcodes, elliptical Fourier descriptors, etc. |
| Abiotic stress | Biomass, height, greenness, NDVI, or other reflectance characteristics, leaf angle, flowering time | Latent space phenotyping |
| Biotic stress | Number of lesions, % area of chlorosis or necrotic tissue | NA |
| Chemotypic profiles | Concentration of specific compounds in the plant ionome, metabolome, herbivory defense compounds | PC scores |

*Note*. NA, nonapplicable; NDVI, normalized difference vegetation index. This is not an exhaustive list of concepts or measurable traits that may be described by conventional or latent approaches.

strawberry (Feldmann et al., 2020; Li et al., 2020; Zingaretti et al., 2021), apple (Lv et al., 2019; Migicovsky, Li, et al., 2017), grape (Chitwood et al., 2014; Yuan et al., 2016), maize (Miller et al., 2017; Warman et al., 2021), and rice (Iwata et al., 2015; Suzuki & Hirata, 2011). Other unsupervised methods include persistent homology (Li, An, et al., 2018; Li et al., 2017; Li, Frank, et al., 2018; Schlautman et al., 2020) and morphometric approaches (Chitwood & Otoni, 2017; Falk et al., 2020; Gupta et al., 2020; Iwata, 2011; Iwata et al., 2002, 2015; Klein et al., 2017; Manacorda & Asurmendi, 2018; Migicovsky, Li, et al., 2017; Sainin et al., 2016). Undoubtedly, the most common latent phenotype in plants is "shape", which is due, in part, to the access to inexpensive cameras and computers (Fahlgren, Gehan, et al., 2015; Tovar et al., 2018), the availability of user-friendly software (Bonhomme et al., 2014; Gehan et al., 2017; Iwata & Ukai, 2002), and the ability to visually interpret results (Chitwood et al., 2012; Horgan et al., 2001; Iwata et al., 2002; Langlade et al., 2005; Manacorda & Asurmendi, 2018; Turner et al., 2018; Vőfély et al., 2019; Xu & Bassel, 2020). Though methods such as the ones mentioned above are common in plant phenotyping, latent phenotyping approaches are only beginning to gain traction for the description of other traits in plants. The following sections provide more detail on conventional image- and sensor-based phenotyping methods, latent phenotyping techniques, and the technical considerations, benefits, and drawbacks of latent phenotyping methods compared to their conventional alternatives.

## 2.1 | Conventional phenotyping from images and sensors

Historically, conventional image based phenotyping provided an efficient strategy to measure key traits that are easy to define and interpret (Anderson et al., 2020; Demir et al., 2018; Li et al., 2014; Poland & Nelson, 2011; Tao et al., 2020; Wang et al., 2018, 2019; Zhou et al., 2020). These traits do not reflect the plant's full complexity, or even the complexity captured in an image. Like the shadows seen by prisoners in Plato's Cave, conventional phenotyping methods provide a limited representation of a plant's status. Methodologically, conventional image-based plant phenotyping aims to condense high-dimensional image data into tabular trait values which quantify some relevant characteristics of the individual (Das et al., 2015; Li et al., 2014; Seethepalli et al., 2020; Tabb & Medeiros, 2017; Walter et al., 2015; Zingaretti et al., 2021). For example, it is common to reduce an image of a plant into a single value specifying the number of vegetation pixels to estimate biomass, or to measure plant height as the number of contiguous pixels in the vertical axis. Often, a collection of these component traits is used to summarize some higher-level abstract concept about the individual and the population, such as the plant's level of stress, its general architecture, or other physiologically and agronomically relevant characteristics (York & Lobet, 2017; York, 2019). Table 1 presents some abstract physiological concepts and the corresponding image-based traits measured to proxy them.

The extraction of conventional trait values from images is a highly designed procedure, whether it is specified programmatically in the steps of an image processing pipeline (Das et al., 2015; Gehan et al., 2017; Seethepalli et al., 2020; Zingaretti et al., 2021), or by human annotation of training data for machine learning techniques (Ishikawa et al., 2018; Stewart et al., 2019; Visa et al., 2014; Zhou et al., 2018). In fact, these human-designed traits are based on the observations of researchers as to how different abstract concepts manifest in the visual appearance of the plants. This human-designed nature means that these measurements are

interpretable and concretely related to one, or more, of these observations.

User-defined traits are typically easy to interpret and have meaningful units associated with them (e.g., cm, g, mL, etc.), which is an understated advantage. Being able to interpret and communicate results is a major benefit of conventional phenotyping methodologies. Along these lines, standard units of measure are used across all fields of study and thus can be more effectively and clearly delivered to a diverse and non-expert audience. While the importance or value of a plant height difference of 10 cm may be field specific, the meaning of the measurement will be clear to anyone familiar with a centimeter. Thanks to their interpretability, the use of conventional image-based traits has seen a boom in publication and application over recent years.

However, conventional phenotyping methods are potentially non-unique, not comprehensive, and compromised by human biases. It takes multiple, and potentially many, descriptors to accurately describe canonical shapes (i.e., unit circle and square), let alone more complex structures (e.g., root and shoot systems). For instance, a circle and a square can have identical aspect ratios (aspect ratio $= \frac{\text{width}}{\text{length}}$), so other descriptors that capture information about the curvature of the shape, the number of corners, or the variation in interior angles are needed to discriminate between these two shapes. In general, as the complexity of a biological phenomenon increases, a greater number of conventional traits are needed to convey its meaning accurately and precisely. Users therefore need to make educated decisions regarding which conventional traits are required to capture the most salient and biologically relevant information. Finally, many conventional traits are scale invariant ratios (e.g., aspect ratio and circularity $= 4\pi\left(\frac{\text{area}}{\text{perimeter}}\right)^2$) and can result in multiple shapes returning the same value despite having visible differences in scale (Schindelin et al., 2012; Schneider et al., 2012). The user-driven decision of which traits to measure can introduce biases that, along with challenges in data collection, can lead to substantial ascertainment bias towards populations, germplasm, and variability. For example, in one strawberry population, fruit shape may be adequately described by the aspect ratio. However, aspect ratio may not be a discriminating descriptor in a second population. This could bias downstream analyses to think that the second population has no variation in shape, although this may be an artifact of the chosen perspective.

When many of these characteristics are collected, it can be tempting to apply data reduction techniques to summarize plant phenotypes based on the variability in that collection of traits. However, the use of PCA or other common dimensionality reduction techniques on a matrix of individually selected, univariate attributes require caution. One concern is that a dimension reduction technique may unintentionally amplify biases in the data, leading to a more biased analysis. For example, length, width, aspect ratio, area, circularity, roundness $= 4\left(\frac{\text{area}}{\pi\sqrt{\max[W,\ L]}}\right)$, solidity $= \frac{\text{area}}{\text{convex area}}$, etc. are extracted individually and are ratios of other variables. If PCA is conducted on these traits, the analysis is exposed to area, length, and width multiple times (Schindelin et al., 2012). Whereas a change in length or width necessarily causes a change in aspect ratio, and vice versa, traits resulting from multivariate analyses do not share this limitation. As opposed to ratios like circularity and solidity, the harmonics of elliptical Fourier analyses, which are used to describe the outlines of 2D shapes, are not functions of each other and each harmonic can be modified independently without changing any others (Bonhomme et al., 2014; Chitwood & Otoni, 2017; Iwata & Ukai, 2002; Migicovsky, Li, et al., 2017; Visa et al., 2014). There are also more intricate scenarios involving arrays of correlated traits, such as leaf angle and inflorescence branching angle (Rice et al., 2020). In these examples, dimension reduction enhances the biological signal of "branching angle of lateral organs" while not explicitly double- or triple-counting the same trait. Similarly, researchers have implicitly posited that latent traits also include the plant ionome (the profile of ions in the plant, Fikas et al., 2019) the defense metabolome (the profile of specialized metabolites that participate in herbivory and pathogen responses, Katz et al., 2021), and the composition of seed fatty acids (Carlson et al., 2019).

In general, conventional image and sensor-based phenotyping strategies are highly interpretable, can reduce time and cost relative to their hand-measured counterparts, and enable opportunities for greater reproducibility. Conventional image analyses tends to follow the common theme of measuring many single compounds and then using PCA to capture global patterns of correlated effects in genome wide association studies or quantitative trait locus mapping. However, given the quantity of conventional traits produced from images and sensors, there is a greater potential for implicit biases during data acquisition or dimensionality reduction. Due to challenges with multicollinearity, using many conventional traits to describe a plant's phenotype provides diminishing returns and greater likelihood of bias or reduced generalizability. The following section elaborates on how latent phenotyping methods address some of these issues.

## 2.2 | Latent phenotyping from images and sensors

Latent variable models explain patterns in observed data by modeling unobserved latent variables (Blei, 2014). The concept of latent variables (also called latent features or latent factors) is widespread in statistics, and statistical models incorporating latent variables have a rich set of applications in

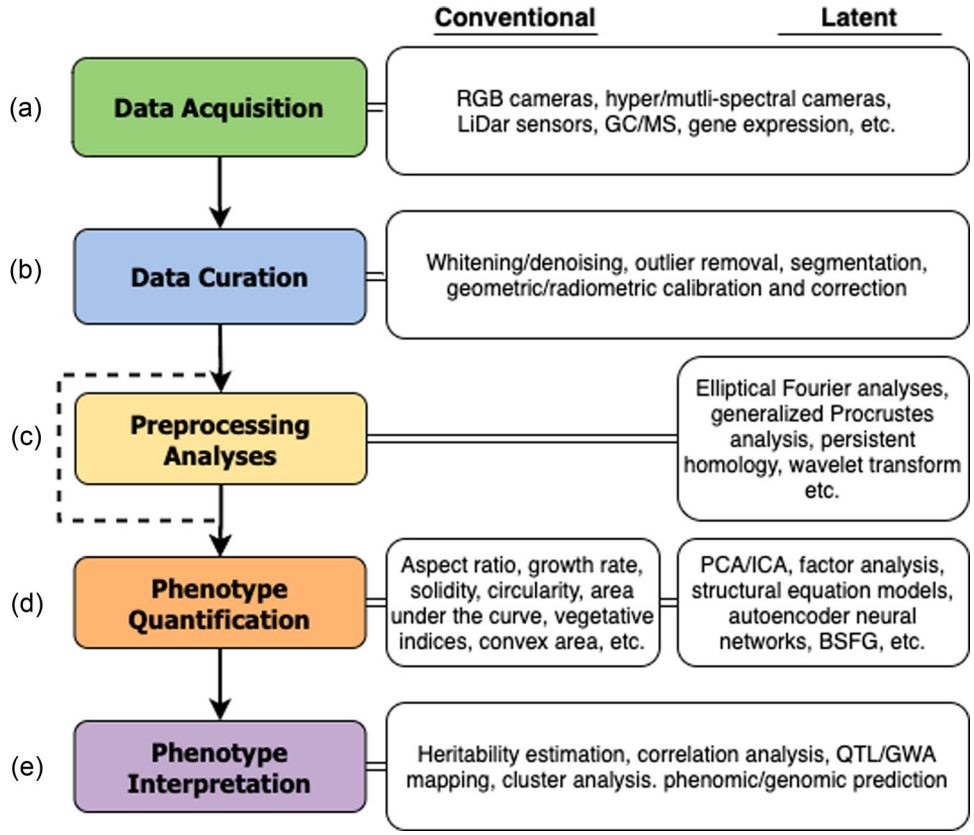


**FIGURE 1** A schematic of the latent phenotyping procedure. (a) The first step is to collect sensor data on a population of interest. (b) The second step is to process that data into a format that will be usable in downstream analyses and to make the target phenomenon more salient by reducing noise or highlighting specific features. (c) The third step is optional and depends on that study at hand. For traditional approaches, this step is skipped and for latent approaches this step further accentuates the phenomenon of interest within the data. (d) The fourth step is to quantify the traits, or latent traits, of interest from the cleaned data. (e) The final step is to try to either make biological sense of the quantified latent traits or to use them for a specific application, e.g., phenomic prediction. The preprocessing methodologies listed here are by no means exhaustive and represent disparate fields of research with many individual contributions that present their own assumptions and biases

the biological sciences (Grotzinger et al., 2019; Runcie & Mukherjee, 2013; Tardieu et al., 2017; Vőfély et al., 2019). As an example, partial least squares regression (PLSR) is commonly applied to high-dimensional biological datasets that contain a high degree of multicollinearity, such as -omics data (Edlich-Muth et al., 2016; Kleinbaum et al., 2013; Lane et al., 2020; Rincent et al., 2018; Runcie & Crawford, 2019; Runcie & Mukherjee, 2013). The PLSR approach and related models, such as partial least squares discriminant analysis, accomplish dimensionality reduction by projecting the raw data onto lower-dimensional latent variables, which retains the ability to predict the response variables. This property of reducing the input data to a set of lower-dimensional features is shared by many familiar and commonly applied models, which include PLSR, partial least squares discriminant analysis, PCA, independent component analysis, latent Dirichlet allocation, linear discriminant analysis, factor analysis, and others. A non-exhaustive list of methodological options is provided for these types of analyses in Figure 1. The choices of which method to employ will vary substantially depending on the data acquisition strategy (e.g., red, gree, blue [RGB] or lidar images), data cleaning and curation choices (e.g., background or outlier removal), pre-processing to highlight the phenomena of interest (e.g., elliptical Fourier analysis or wavelet transform), quantification of the latent space (e.g., PCA), and targeted interpretation of the latent space (e.g., correlation or genetic analyses). Such representations also appear in machine learning models, such as autoencoders (Figure 2), where a latent space is derived from the lower-dimensional activations of a layer in the center of a neural network. In Figure 2, image data of *Brassica napus* plants were analyzed with a convolutional autoencoder neural network with two latent dimensions. The goal of an autoencoder, broadly, is to learn a lower-dimensional representation, two variables in this case, that can be used to approximately reconstruct the original input data.

In machine learning, discovering this transformation from the input space to another, typically lower-dimensional form which is more appropriate for a learning task is known broadly as 'representation learning' (Bengio et al., 2013). In each case, the set of latent variables, or learned representation, is learned

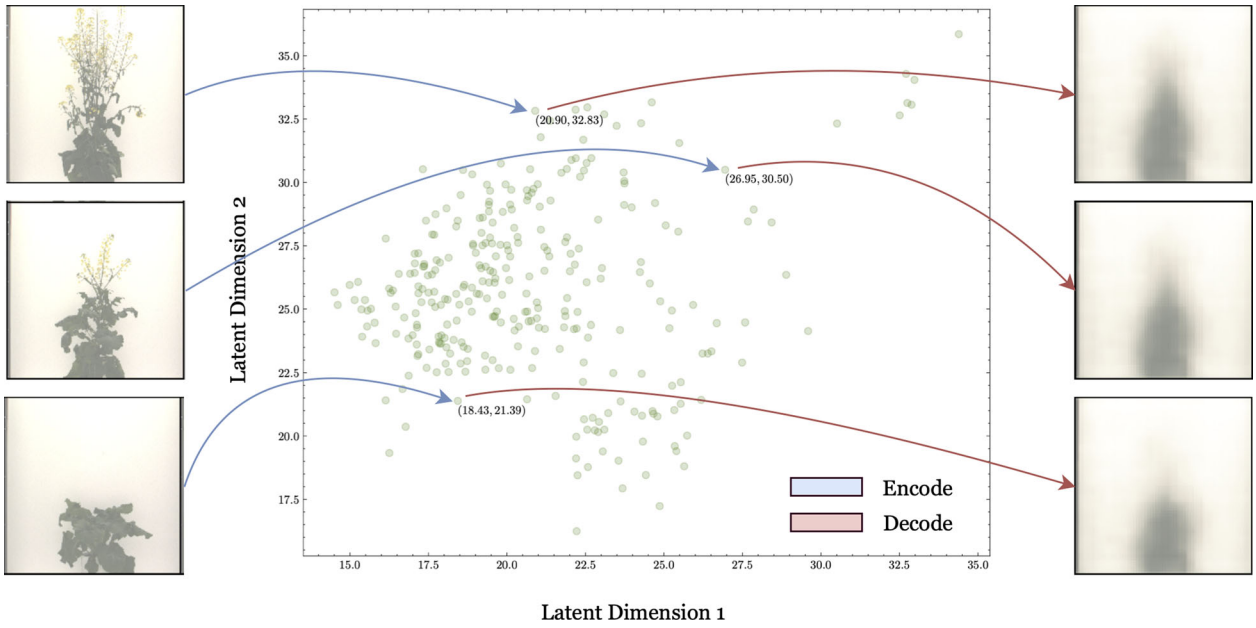


**FIGURE 2** A convolutional autoencoder trained on a dataset of mature *Brassica napus* plants. In this example, a high-dimensional image is encoded as a point in a two-dimensional latent space, and the decoder attempts to recreate the original image using only this information. The encoder needs to pack as much pertinent information as possible into these two dimensions to help the decoder produce accurate outputs

from the raw data, as opposed to being defined beforehand. Thus, each of these features do not directly represent any one input feature, but rather encode some linear or nonlinear combination of the many input features into a single dimension. Just as the latent variables in these models correspond to some learned abstract representation of the processes generating the input data, latent phenotypes are defined by some learned abstract representation of the latent source of variance present in the dataset. This variability is typically measured across the full input. For example, in the case of images, the input data will correspond to the values of each individual pixel of each image, producing an extremely high-dimensional dataset where each color channel, e.g., RGB for most standard digital cameras, in each pixel adds an additional input dimension. Compared to conventional traits, which are based on measurements made by researchers, latent traits are inferred. Consequently, this means that latent phenotypes are reflective of patterns present in the data, not measurements made by a human observer, and are, ideally, less susceptible to user-imposed biases and assumptions. While conventional phenotyping seeks to quantify a particular concept by summarizing it as a collection of one or more measurable traits, latent phenotyping instead seeks to indirectly quantify the latent source of variance in the high-dimensional raw data, whether that source is stress, genotype, or otherwise.

For the purposes of defining latent phenotyping, statistical methods which assume the presence of underlying latent factors that explain correlations in the observations, such as factor analysis, as well as factor models such as PCA, independent component analysis, etc., are included. For example, because environment (E) and management (M) can be controlled or, at least, accounted for using appropriate experimental designs and sampling strategies, genotype (G), interactions with genotype (i.e., GxE, GxM, GxExM), and irreducible error can be isolated as the sole sources of variance (Lane & Murray, 2021; Piepho et al., 2012; Schulz-Streeck et al., 2013). Thus, methods such as PCA which seek to summarize the major axes of variance are likely to find heritable components, which implies a latent effect from genotype. In cases where there are many sources of variance, a more computationally sophisticated latent phenotyping technique, e.g., latent space phenotyping (Ubbens et al., 2020), may be required to disentangle the sources of variance and uncover the genetic signal in more complex latent traits, e.g., whole plant response to abiotic stress.

As with conventional phenotyping approaches, the specific path a user takes to acquire latent phenotypes will depend on the structure and type of raw data, e.g., RGB, hyperspectral sensors, lidar, etc., and the underlying biology of the target phenomenon the user wishes to capture. Both the beauty and the danger of this framework is that there are nearly limitless approaches that a researcher could take (Backhaus et al., 2010; Barbier de Reuille et al., 2015; Biot et al., 2016; Botvinik-Nezer et al., 2020; Buckner et al., 2021; Chitwood et al., 2012; Clark et al., 2011; Das et al., 2015; de Reuille et al., 2014; Falk et al., 2020; Galkovskyi et al., 2012; Granier et al., 2006; Hartmann et al., 2011; Horgan et al., 2001; Iyer-Pascuzzi et al., 2010; Klingenberg, 2011; Li et al., 2020; Li, An, et al., 2018; Liu et al., 2020; Nowak et al., 2021; Ristova et al., 2013; Slice, 2007; Slice & Stitzel, 2004; van der Heijden et al., 2012; Wu

et al., 2016; Xu & Bassel, 2020; Zingaretti et al., 2021). The most common approaches to date are PCA and autoencoders, which both aim to compress data into a smaller number of axes based on a specific set of rules. For example, the use of PCA to describe internal carrot color variation (Horgan, 2001), strawberry fruit shape (Feldmann et al., 2020), and variation in carrot shape and biomass distributions (Turner et al., 2018); the use of PCA and autoencoders to describe plant architecture from heatmaps of maize plots (Gage et al., 2019); and the use of a latent embedding emitted by deep neural networks (a distinct multi-stage technique called 'latent space phenotyping') to capture the genetic variation attributable to drought stress in the genus *Setaria* and the visual drought response in *Brassica napus* (Ubbens et al., 2020). Importantly, these approaches can be deployed for both classification and prediction of target traits, as well as to capture biologically interesting aspects of plant growth and development.

Notable advantages of latent phenotyping include the ability to describe complex, multi-dimensional phenotypes or biological processes at the whole plant level and the reduction of direct human input in the measurement process. Most strikingly, latent traits are likely to be more independent of explicit human biases in the choice of what aspect of the data to measure, potentially allowing them to learn complex, nonlinear visual concepts instead of tying the measurement to a rigid, human-defined concept or pre-existing hypotheses (Baxter, 2020). This does not mean that latent phenotyping approaches are free from other types of user-imposed biases during model selection, implicit biases accrued during data collection, and over-generalized interpretations. Furthermore, compared to conventional phenotyping strategies, latent phenotyping techniques typically require less pre-processing, such as noise removal or segmentation, and can be applied to most types of input data. Although the adoption and active development of these techniques in plant phenotyping is relatively new, it seems likely that latent phenotyping will become increasingly important and widespread in the future due to these key advantages.

## 2.3 | Technical considerations of latent phenotyping

As with all attempts to model data, many techniques can be used to model latent phenotypes, and awareness of differences among these techniques is critical to understand how specific methodologies will influence results and the interpretation of those results. For example, PCA is only capable of modeling linear aspects of the input variables. As a result, nonlinear patterns in the data will be represented differently when using PCA versus other techniques that are capable of modeling nonlinear patterns, such as autoencoders. A hazard with nonlinear models is that the embeddings they provide exist on a nonlinear manifold. This means that they cannot be interpreted in subsequent steps, such as linear regression, as if they were on a linear scale, e.g. a point which embeds on a particular axis in latent space at the value 4 is not, in any meaningful sense, "twice as dissimilar" to an embedding at 8 as it is to a different embedding at 6. In fact, any comparison of Euclidean distance between points in these encoded variables as a measure of similarity is not reasonable because of the non-linear nature of the analysis. In contrast to other nonlinear models, variational autoencoders explicitly learn an approximation to a normal distribution, and therefore impose this distribution over the latent variables (Kingma & Welling, 2013). Forcing the latent variables to be normally distributed allows for tractable sampling from the posterior distribution, which is the original intent behind the variational autoencoder, but the utility of this arbitrarily distributed latent variable for a downstream phenotyping task is questionable. For these reasons and others, care must be taken when selecting a model, and the researcher should be informed about the model's assumptions and its capabilities prior to implementation.

Using learned latent traits instead of conventional traits comes with the substantial drawback of losing the inherent interpretability that comes with conventional measurements. For example, consider a conventional phenotyping result that height was taller by an average of 36 mm for a breeding line. This makes intuitive sense, as plant height is something experienced by an observer. In contrast, a study which makes use of latent phenotyping requires extra steps towards interpretation. The raw trait values, such as component scores on principal components, are unitless scalar values which only have meaning if mapped back through a particular function, such as a trained decoder network. Various studies have attempted to visualize these features using eigenshapes, the visualization of eigenvectors, and saliency mapping (Feldmann et al., 2020; Gage et al., 2019; Li, An, et al., 2018; Migicovsky, Li, et al., 2017; Turner et al., 2018; Ubbens et al., 2020; Vőfély et al., 2019). This begs the user to define their questions more rigorously before using latent phenotyping approaches, which may distort an otherwise simple question regarding plant height, yield, or any other easily defined, conventional trait.

Data dependency is a critical issue that is not only affected by the amount of observed data, but the saliency of the biological phenomenon and the reliability of data acquisition strategy. Latent phenotyping approaches such as PCA and autoencoders will map noise when only noise is present. A possible consequence of this is that the resultant principal component scores and encodings will have low repeatability or heritability, which may have more to do with data acquisition or model misuse than any biologically relevant phenomenon (Feldmann et al., 2020; Turner et al., 2018). Latent phenotyping approaches may be able to reveal whether data is of poor quality, but they are not able to "fix" overly noisy, unreliable data. However, if data is collected in a

careful, considered manner, these approaches can deliver powerful insights for both prediction and inference. Thus, the data acquisition and model selection strategies are just as relevant to latent phenotyping as they are to conventional phenotyping from images and sensors.

Standard operating procedures and common best practices for latent phenotyping approaches might be obscure to biological researchers, leading to naive applications of such tools. As an example, linear regression has a set of assumptions that make its use appropriate and interpretable given the assumptions are met. Similarly, PCA, PLSR, CNNs, and autoencoders have unique sets of assumptions that may render them useless if these underlying assumptions are violated. Latent phenotyping will not tell the user when a core assumption has been irreparably violated unless the user knows how to probe the data and the model with consideration of those assumptions. It is unlikely that every biological researcher with an interest in latent phenotyping will have the time or capacity to learn all the various assumptions and best practices themselves, thereby emphasizing the necessity to partner with experts in the application of these methods who can help to safeguard against nonsensical or biased results.

## 2.4 | Practical considerations for latent phenotyping

When embarking on a study, researchers consider the purpose of the study, the context of that purpose, and whether there are meaningful takeaways for general crop improvement (prediction) and/or basic biology (inference). This dichotomy in the philosophy of science, while often conflated, can be useful for discussing how latent phenotyping approaches may be valuable in practice.

If the aim is prediction, it is sufficient to identify latent traits that are correlated with and can be used to predict a trait or whole-plant phenotype of interest, such as yield, biomass, or fruit quality, without the necessity of an underlying causal relationship between predictors and response using PLSR, or a similar approach, and may rely on tens, hundreds, or even thousands of predictor traits. This process, termed "phenomic prediction," uses phenotypic and/or latent traits as independent variables—analogous to markers in genomic prediction—to predict the value of some dependent variable (Edlich-Muth et al., 2016; Fernandes et al., 2018; Krause et al., 2018; Lane et al., 2020; Momen et al., 2019; Rincent et al., 2018; Sandhu et al., 2021). Phenomic prediction has both the potential to improve genomic prediction by incorporating more reliable, correlated traits and to potentially reduce the need for or extent of genotyping (Fernandes et al., 2018; Lane & Murray, 2021; Lane et al., 2020; Rincent et al., 2018). These approaches, like other high-dimensional regression approaches, are subject to overfitting and multicollinearity. However, these can be addressed by using training-validation-test set optimization approaches (Berro et al., 2019), proper application of cross-validation (Runcie & Cheng, 2019), or model reduction, variable selection, and regularization strategies, e.g., Lasso, Ridge regression, and elastic nets. Further, as with genomic prediction, phenomic prediction runs the risk of conflating statistical significance with biological importance without appropriate contextual consideration and independent validation (Bernardo, 2001, 2016, 2020; Sanders, 2019; Shmueli, 2010).

On the other hand, when the goal is to understand an underlying mechanism or sources of variability for a target trait, methods that rely on statistical associations have limited utility unless paired with the necessary biological context and knowledge of the underlying assumptions for a chosen specific analysis (Bernardo, 2001, 2016, 2020). This makes the application of latent phenotyping more challenging for basic biological questions than for prediction. Despite these challenges, latent phenotyping approaches have been used to successfully identify genetic loci linked with tomato fruit and leaves (Chitwood et al., 2012; Li, Frank, et al., 2018; Wang et al., 2019), rice grains (Iwata et al., 2015), the maize, rice, and soybean ionomes (Chu et al., 2016; Fikas et al., 2019; Liu et al., 2021), the *Brassica* defensive metabolome (D'Oria et al., 2021; Katz et al., 2021), oat seed fatty acid concentrations (Carlson et al., 2019), inflorescence development in maize and sorghum (Leiboff & Hake, 2019; Rice et al., 2020), carrot shoot and roots (Turner et al., 2018), response to drought in *Setaria* (Ubbens et al., 2020), and strawberry fruit shape (Nagamatsu et al., 2021). Some of these successful examples have relied on prior information from preceding univariate analyses to validate the loci discovered using latent phenotyping and to aid their interpretation of those newly discovered loci (Fikas et al., 2019; Katz et al., 2021; Ubbens et al., 2020). Notably, latent phenotyping approaches have yet to be applied to biotic stress resistance in plants (Table 1). It could be that reasonably powered studies require unreasonable sample sizes because of inconsistent symptom development (e.g., chlorosis, wilting, spotting), or that subtle changes in plant status during disease onset and development are of lesser interest than the binary difference between being alive or dead, which is simpler to assess.

Regardless of the experimental goals, providing intuitive presentation and interpretation of results is critical to maximize the impact of empirical findings. This is especially true when measurements are acquired using methods that rely on abstract concepts, as is the case for latent phenotypes. Strategies to relate latent phenotypes to observable quantities are often as straightforward as a statistical comparison between latent phenotypes and human-understandable features, e.g., classification based on a categorical scale or regression against a numerical scale (Casanova et al., 2017; Chitwood, 2020; Clark et al., 2015; Ishikawa et al., 2018;

Kadir, 2015; Lane et al., 2020; Li, An, et al., 2018; Nascimento et al., 2021; Neto et al., 2006; Stewart et al., 2019; Zingaretti et al., 2021). However, it is also possible that latent phenotypes capture aspects governing a trait that are otherwise difficult to quantify, and therefore have complex, non-linear relationships with or cannot be directly related back to a familiar concept (Li et al., 2017; Migicovsky, Li, et al., 2017; Rice et al., 2020). When there is no clear observable counterpart for latent phenotypes, visual aids can fill the gap by providing human-interpretable representations of the data. Effective examples of this approach in practice include the presentation of PCA results alongside a morphospace of theoretical shapes based on eigenvectors (Bonhomme et al., 2014; Iwata, 2011; Iwata et al., 2002, 2015; Iwata & Ukai, 2002; Migicovsky, Li, et al., 2017; Miller et al., 2017) and the direct comparison of raw input images with reconstructions produced from an embedding (Figure 2).

Irrespective of whether measurements are easily observable or more abstract, perhaps the most critical set of questions a researcher can ask revolves around whether results meet certain standards of replicability and variability (Bernardo, 2020; Moehring et al., 2014). Similarly, as indicated, thought should also be given to the scope of diversity included in the study (e.g. diversity panels versus elite breeding populations) and the relevance to addressing the question(s) of interest, i.e. if there are useful or general takeaways when entries are more or less phenotypically diverse (Berro et al., 2019; Brandariz & Bernardo, 2019; Campbell et al., 2017).

A final consideration is if there are any large-scale uses for latent phenotyping results. From a logistical perspective, applied use will depend on whether a given method is scalable, cost-effective, and broadly applicable to diverse study systems or at least within a target system. By defining plant phenotypes through a more comprehensive set of traits, latent phenotyping approaches have the potential to change the way researchers practice breeding and deploy marketing strategies. For instance, these comprehensive traits could conceptually be used for variety identification and facilitating intellectual property (IP) applications and protections (e.g., plant patents, plant variety protections, and plant breeders rights) by including more quantitative descriptors, which have been shown to be predictive of both variety and species in multiple studies (Chitwood, 2020; Chitwood & Otoni, 2017; Ishikawa et al., 2018; Li, An, et al., 2018; Pereira et al., 2019). It is unlikely that latent or conventional traits will replace DNA forensics for the enforcement of IP laws, but they may aid plant specific IP applications by providing detailed and nuanced descriptions of plant properties. Regardless of downstream goals, latent phenotyping and other machine learning methods are undoubtedly an exciting area of research, but not a panacea in the pursuit of understanding phenotypic variation (Bernardo, 2016).

## 3 | CONCLUSIONS

In recent years, the growth of plant phenomics, the availability of powerful open-source software, and the competency to acquire complex datasets has elicited the need for a thorough discussion of the state of the art. As with all statistical techniques and applications, careful consideration of the hypothesis, experimental procedure, and data structure is required to deploy latent phenotyping approaches. All the challenges inherent to traditional methods still exist; the data must be collected systematically and in such a way to capture relevant information for the questions and hypotheses related to the specific study. While sensors are much less subject to explicit biases than are humans, implicit and explicit biases may still be reflected in the quantification of sensor data and exacerbated by the naive application of sophisticated tools. To support reproducibility and the adoption of best practices across labs with varying levels of experience, there is a need to report latent phenotyping analyses and protocols precisely and accurately in the literature, alongside relevant code repositories(Fanelli, 2018; Harris, 2018; Hutson, 2018; Miyakawa, 2020; Peng, 2015; Stoddart, 2016). Furthermore, it is important to note that not all research questions require a latent phenotyping approach, such as conventional disease traits, yield, and time-to-flowering, and that the naive application of these methods can lead to nonsensical results and biased interpretations. These methods rely on mathematical and statistical models that all have unique and diverse assumptions about the input, which may render certain methods irrelevant or inappropriate for a given data set. Successful application of latent phenotyping in practice will require communication among developers, end users, and stakeholders, to address important questions in crop improvement and basic biology True validation of these approaches in independent populations managed by independent groups with similar objectives is the final frontier for both conventional and latent phenotyping. While the promise of latent phenotyping is great, it is still critical to plan sampling procedures that match experimental goals to pay close attention to methodological assumptions, and to enable the repeatability of analyses and reproducibility of results.

## ACKNOWLEDGMENTS

The authors thank Daniel Chitwood and Brandon Hurr for discussing, reviewing, and improving the manuscript. The authors also thank the anonymous reviewers and Seth Murray for their comments, which greatly improved the quality of the manuscript. This material is based upon work supported by the NSF Postdoctoral Research Fellowship in Biology under Grant No. IOS-1906619. This research was undertaken thanks in part to funding from Canada First Research Excellence Fund.

## AUTHOR CONTRIBUTIONS

Mitchell J. Feldmann: Conceptualization; Project administration; Supervision; Visualization; Writing-original draft; Writing-review & editing. Joseph L. Gage: Conceptualization; Project administration; Supervision; Visualization; Writing-original draft; Writing-review & editing. Sarah D. Turner-Hissong: Conceptualization; Project administration; Supervision; Visualization; Writing-original draft; Writing-review & editing. Jordan R. Ubbens: Conceptualization; Project administration; Supervision; Visualization; Writing-original draft; Writing-review & editing

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

## ORCID

*Mitchell J. Feldmann* https://orcid.org/0000-0002-7257-3401
*Joseph L. Gage* https://orcid.org/0000-0001-5946-4414
*Sarah D. Turner-Hissong* https://orcid.org/0000-0002-0961-0767
*Jordan R. Ubbens* https://orcid.org/0000-0002-3162-5252

## REFERENCES

Anderson, S. L., II, Murray, S. C., Chen, Y., Malambo, L., Chang, A., Popescu, S., Cope, D., & Jung, J. (2020). Unoccupied aerial system enabled functional modeling of maize height reveals dynamic expression of loci. *Plant Direct*, *4*(5), e00223. https://doi.org/10.1002/pld3.223

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A. M., Hu, T. T., Jiang, R., Muliyati, N. W., Zhang, X., Amer, M. A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., …, & Nordborg, M. (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature*, *465*(7298), 627–631. https://doi.org/10.1038/nature08800

Backhaus, A., Kuwabara, A., Bauch, M., Monk, N., Sanguinetti, G., & Fleming, A. (2010). LEAFPROCESSOR: A new leaf phenotyping tool using contour bending energy and shape cluster analysis. *The New Phytologist*, *187*(1), 251–261. https://doi.org/10.1111/j.1469-8137.2010.03266.x

Barbier de Reuille, P., Routier-Kierzkowska, A. - L., Kierzkowski, D., Bassel, G. W., Schüpbach, T., Tauriello, G., Bajpai, N., Strauss, S., Weber, A., Kiss, A., Burian, A., Hofhuis, H., Sapala, A., Lipowczan, M., Heimlicher, M. B., Robinson, S., Bayer, E. M., Basler, K., Koumoutsakos, P., …, & Smith, R. S. (2015). MorphoGraphX: A platform for quantifying morphogenesis in 4D. *eLife*, *4*, 05864. https://doi.org/10.7554/eLife.05864

Baxter, I. (2020). We aren't good at picking candidate genes, and it's slowing us down. *Current Opinion in Plant Biology*, *54*, 57–60. https://doi.org/10.1016/j.pbi.2020.01.006

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Bernardo, R. (2001). What if we knew all the genes for a quantitative trait in hybrid crops? *Crop Science*, *41*(1), 1–4. https://doi.org/10.2135/cropsci2001.4111

Bernardo, R. (2016). Bandwagons I, too, have known. *Theoretical and Applied Genetics*, *129*(12), 2323–2332. https://doi.org/10.1007/s00122-016-2772-5

Bernardo, R. (2020). Reinventing quantitative genetics for plant breeding: Something old, something new, something borrowed, something BLUE. *Heredity*, *125*(6), 375–385. https://doi.org/10.1038/s41437-020-0312-1

Berro, I., Lado, B., Nalin, R. S., Quincke, M., & Gutiérrez, L. (2019). Training population optimization for genomic selection. *The Plant Genome*, *12*(3), 1–14. https://doi.org/10.3835/plantgenome2019.04.0028

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 2-D shape analysis. *Development*, *143*(18), 3417—3428. https://doi.org/10.1242/dev.134619

Blei, D. M. (2014). Build, compute, critique, repeat: Data analysis with latent variable models. *Annual Review of Statistics and Its Application*, *1*(1), pp. 203–232. https://doi.org/10.1146/annurev-statistics-022513-115657

Bonhomme, V., Picq, S., Gaucherel, C., & Claude, J. (2014). Momocs: Outline analysis using R. *Journal of Statistical Software*, *56*(13), https://doi.org/10.18637/jss.v056.i13

Bookstein, F. L. (1997). Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Medical Image Analysis*, *1*(3), 225–243. https://doi.org/10.1016/S1361-8415(97)85012-8

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., …, & Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88. https://doi.org/10.1038/s41586-020-2314-9

Brandariz, S. P., & Bernardo, R. (2019). Small ad hoc versus large general training populations for genomewide selection in maize biparental crosses. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *132*(2), 347–353. https://doi.org/10.1007/s00122-018-3222-3

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., Ersoz, E., Flint-Garcia, S., Garcia, A., Glaubitz, J. C., Goodman, M. M., Harjes, C., Guill, K., Kroon, D. E., Larsson, S., Lepak, N. K., Li, H., Mitchell, S. E., Pressoir, G., …, & McMullen, M. D. (2009). The genetic architecture of maize flowering time. *Science*, *325*(5941), 714–718. https://doi.org/10.1126/science.1174276

Buckner, E., Tong, H., Ottley, C., & Williams, C. (2021). High-throughput image segmentation and machine learning approaches in the plant sciences across multiple scales. *Emerging Topics in Life Sciences*, *5*(2), 239–248.

Campbell, M. T., Du, Q., Liu, K., Brien, C. J., Berger, B., Zhang, C., & Walia, H. (2017). A comprehensive image-based phenomic analysis reveals the complex genetic architecture of shoot growth dynamics in rice (*Oryza sativa*). *The Plant Genome*, *10*(2), 1–14. https://doi.org/10.3835/plantgenome2016.07.0064

Carlson, M. O., Montilla-Bascon, G., Hoekenga, O. A., Tinker, N. A., Poland, J., Baseggio, M., Sorrells, M. E., Jannink, J. - L., Gore, M. A.,

& Yeats, T. H. (2019). Multivariate genome-wide association analyses reveal the genetic basis of seed fatty acid composition in oat (*Avena sativa* L.). *G3: Genes, Genomes, Genetics*, *9*(9), 2963–2975. https://doi.org/10.1534/g3.119.400228

Casanova, D., da Silva, N. R., Cordeirod, D. F., Kolb, R. M., & Bruno, O. M. (2017). Shape analysis of plant leaves using complex networks for species identification. *Biological Shape Analysis*, *2017*, 57–70. https://doi.org/10.1142/9789813225701_0003

Chitwood, D. H. (2020). The shapes of wine and table grape leaves: An ampelometric study inspired by the methods of Pierre Galet. *bioRxiv*, 2020.05.08.085175. https://doi.org/10.1101/2020.05.08.085175

Chitwood, D. H., Headland, L. R., Kumar, R., Peng, J., Maloof, J. N., & Sinha, N. R. (2012). The developmental trajectory of leaflet morphology in wild tomato species. *Plant Physiology*, *158*(3), 1230–1240. https://doi.org/10.1104/pp.111.192518

Chitwood, D. H., & Otoni, W. C. (2017). Morphometric analysis of Passiflora leaves: The relationship between landmarks of the vasculature and elliptical Fourier descriptors of the blade. *GigaScience*, *6*(1), 1–13.

Chitwood, D. H., Ranjan, A., Martinez, C. C., Headland, L. R., Thiem, T., Kumar, R., Covington, M. F., Hatcher, T., Naylor, D. T., Zimmerman, S., Downs, N., Raymundo, N., Buckler, E. S., Maloof, J. N., Aradhya, M., Prins, B., Li, L., Myles, S., & Sinha, N. R. (2014). A modern ampelography: A genetic basis for leaf shape and venation patterning in grape. *Plant Physiology*, *164*(1), 259. https://doi.org/10.1104/pp.113.229708

Chitwood, D. H., & Topp, C. N. (2015). Revealing plant cryptotypes: Defining meaningful phenotypes among infinite traits. *Current Opinion in Plant Biology*, *24*, 54–60. https://doi.org/10.1016/j.pbi.2015.01.009

Chu, Q., Watanabe, T., Shinano, T., Nakamura, T., Oka, N., Osaki, M., & Sha, Z. (2016). The dynamic state of the ionome in roots, nodules, and shoots of soybean under different nitrogen status and at different growth stages. *Journal of Plant Nutrition and Soil Science*, *179*(4), 488–498. https://doi.org/10.1002/jpln.201600059

Clark, J. Y., Corney, D. P. A., Notley, S., & Wilkin, P. (2015). Image processing and artificial neural networks for automated plant species identification from leaf outlines. *Biological Shape Analysis*, 50–64. https://doi.org/10.1142/9789814704199_0003

Clark, R. T., MacCurdy, R. B., Jung, J. K., Shaff, J. E., McCouch, S. R., Aneshansley, D. J., & Kochian, L. V. (2011). Three-dimensional root phenotyping with a novel imaging and software platform. *Plant Physiology*, *156*(2), 455–465. https://doi.org/10.1104/pp.110.169102

Das, A., Schneider, H., Burridge, J., Ascanio, A. K. M., Wojciechowski, T., Topp, C. N., Lynch, J. P., Weitz, J. S., & Bucksch, A. (2015). Digital imaging of root traits (DIRT): A high-throughput computing and collaboration platform for field-based root phenomics. *Plant Methods*. 11(1). https://doi.org/10.1186/s13007-015-0093-3

Demir, N., Sönmez, N. K., Akar, T., & Ünal, S. (2018). Automated measurement of plant height of wheat genotypes using a DSM derived from UAV imagery. *Proceedings*, *2*(7), 350. https://doi.org/10.3390/ecrs-2-05163

de Reuille, P. B., Robinson, S., & Smith, R. S. (2014). Quantifying cell shape and gene expression in the shoot apical meristem using MorphoGraphX. *Methods in Molecular Biology*, *1080*, 121–134. https://doi.org/10.1007/978-1-62703-643-6_10

Dillmann, C., & Guérin, D. (1998). *Comparison between maize inbred lines: genetic distances in the expert's eye*. https://hal.archives-ouvertes.fr/hal-00885866/document

Donald, C. M. (1968). The breeding of crop ideotypes. *Euphytica/Netherlands Journal of Plant Breeding*, *17*(3), 385–403.

Dong, H., Zhao, H., Li, S., Han, Z., Hu, G., Liu, C., Yang, G., Wang, G., Xie, W., & Xing, Y. (2018). Genome-wide association studies reveal that members of bHLH subfamily 16 share a conserved function in regulating flag leaf angle in rice (*Oryza sativa*). *PLoS Genetics*, *14*(4), e1007323. https://doi.org/10.1371/journal.pgen.1007323

D'Oria, A., Courbet, G., Lornac, A., Pluchon, S., Arkoun, M., Maillard, A., Etienne, P., Diquélou, S., & Ourry, A. (2021). Specificity and plasticity of the functional ionome of *Brassica napus* and *Triticum aestivum* exposed to micronutrient or beneficial nutrient deprivation and predictive sensitivity of the ionomic signatures. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.641678

Edlich-Muth, C., Muraya, M. M., Altmann, T., & Selbig, J. (2016). Phenomic prediction of maize hybrids. *Bio Systems*, *146*, 102–109. https://doi.org/10.1016/j.biosystems.2016.05.008

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/10.1016/j.molp.2015.06.005

Fahlgren, N., Gehan, M. A., & Baxter, I. (2015). Lights, camera, action: High-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology*, *24*, 93–99. https://doi.org/10.1016/j.pbi.2015.02.006

Falk, K. G., Jubery, T. Z., Mirnezami, S. V., Parmley, K. A., Sarkar, S., Singh, A., Ganapathysubramanian, B., & Singh, A. K. (2020). Computer vision and machine learning enabled soybean root phenotyping pipeline. *Plant Methods*, *16*, 5. https://doi.org/10.1186/s13007-019-0550-5

Fanelli, D. (2018). Opinion: Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2628–2631. https://doi.org/10.1073/pnas.1708272114

Fan, Z., Hasing, T., Johnson, T. S., Garner, D. M., Barbey, C. R., Colquhoun, T. A., Sims, C. A., Resende, M. F. R., & Whitaker, V. M. (2021). Strawberry sweetness and consumer preference are enhanced by specific volatile compounds. *Horticulture Research*, *8*(1), 66. https://doi.org/10.1038/s41438-021-00502-5

Farneti, B., Di Guardo, M., Khomenko, I., Cappellin, L., Biasioli, F., Velasco, R., & Costa, F. (2017). Genome-wide association study unravels the genetic control of the apple volatilome and its interplay with fruit texture. *Journal of Experimental Botany*, *68*(7), 1467–1478. https://doi.org/10.1093/jxb/erx018

Feldmann, M. J., Hardigan, M. A., Famula, R. A., López, C. M., Tabb, A., Cole, G. S., & Knapp, S. J. (2020). Multi-dimensional machine learning approaches for fruit shape phenotyping in strawberry. *GigaScience*, *9*(5),. https://doi.org/10.1093/gigascience/giaa030

Fernandes, S. B., Dias, K. O. G., Ferreira, D. F., & Brown, P. J. (2018). Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *131*(3), 747–755. https://doi.org/10.1007/s00122-017-3033-y

Fikas, A. A., Dilkes, B. P., & Baxter, I. (2019). Multivariate analysis reveals environmental and genetic determinants of element covariation in the maize grain ionome. *Plant Direct*, *3*(5), e00139. https://doi.org/10.1002/pld3.139

Gage, J. L., Richards, E., Lepak, N., Kaczmar, N., Soman, C., Chowdhary, G., Gore, M. A., & Buckler, E. (2019). In-field whole-plant maize architecture characterized by subcanopy rovers and latent space phenotyping. *The Plant Phenome*, *2*, 1–11. https://doi.org/10.2135/tppj2019.07.0011

Galkovskyi, T., Mileyko, Y., Bucksch, A., Moore, B., Symonova, O., Price, C. A., Topp, C. N., Iyer-Pascuzzi, A. S., Zurek, P. R., Fang, S., Harer, J., Benfey, P. N., & Weitz, J. S. (2012). GiA Roots: Software for the high throughput analysis of plant root system architecture. *BMC Plant Biology*, *12*, 116. https://doi.org/10.1186/1471-2229-12-116

Gehan, M. A., Fahlgren, N., Abbasi, A., Berry, J. C., Callen, S. T., Chavez, L., Doust, A. N., Feldman, M. J., Gilbert, K. B., Hodge, J. G., Hoyer, J. S., Lin, A., Liu, S., Lizárraga, C., Lorence, A., Miller, M., Platon, E., Tessman, M., & Sax, T. (2017). PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ*, *5*, e4088. https://doi.org/10.7717/peerj.4088

Gehan, M. A., & Kellogg, E. A. (2017). High-throughput phenotyping. *American Journal of Botany*, *104*(4), 505–508. https://doi.org/10.3732/ajb.1700044

Gehan, M., Fahlgren, N., Jarolmasjed, S., & Schuhl, H. (2019). *PlantCV: Open-source image analysis software for models and crops*. Plant and Animal Genome XXVII Conference. https://pag.confex.com/pag/xxvii/meetingapp.cgi/Paper/34762

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, *40*(1), 33–51. https://doi.org/10.1007/BF02291478

Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S. J., Dauzat, M., Hamard, P., Thioux, J. - J., Rolland, G., Bouchier-Combaud, S., Lebaudy, A., Muller, B., Simonneau, T., & Tardieu, F. (2006). PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. *The New Phytologist*, *169*(3), 623–635. https://doi.org/10.1111/j.1469-8137.2005.01609.x

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., Ip, H. F., Marioni, R. E., McIntosh, A. M., Deary, I. J., Koellinger, P. D., Harden, K. P., Nivard, M. G., & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513–525. https://doi.org/10.1038/s41562-019-0566-x

Gupta, S., Rosenthal, D. M., Stinchcombe, J. R., & Baucom, R. S. (2020). The remarkable morphological diversity of leaf shape in sweet potato (*Ipomoea batatas*): The influence of genetics, environment, and G×E. *The New Phytologist*, *225*(5), 2183–2195. https://doi.org/10.1111/nph.16286

Harris, R. (2018). *Rigor Mortis: How Sloppy Science Creates Worthless Cures, Crushes Hope, and Wastes Billions*. Basic Books.

Hartmann, A., Czauderna, T., Hoffmann, R., Stein, N., & Schreiber, F. (2011). HTPheno: An image analysis pipeline for high-throughput plant phenotyping. *BMC Bioinformatics*, *12*, 148. https://doi.org/10.1186/1471-2105-12-148

Herritt, M. T., Pauli, D., Mockler, T. C., & Thompson, A. L. (2020). Chlorophyll fluorescence imaging captures photochemical efficiency of grain sorghum (*Sorghum bicolor*) in a field setting. *Plant Methods*, *16*, 109. https://doi.org/10.1186/s13007-020-00650-0

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/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/10.1016/S0168-1699(00)00191-5

Hutson, M. (2018). Artificial intelligence faces reproducibility crisis. *Science*, *359*(6377), 725–726. https://doi.org/10.1126/science.359.6377.725

Ishikawa, T., Hayashi, A., Nagamatsu, S., Kyutoku, Y., Dan, I., Wada, T., Oku, K., Saeki, Y., Uto, T., & Tanabata, T., & Others. (2018). Classification of strawberry fruit shape by machine learning. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, *42*(2), 463–470.

Iwata, H. (2011). Quantitative genetic analyses of crop organ shape based on principal components of elliptic Fourier descriptors. *Biological Shape Analysis*, 50–66. https://doi.org/10.1142/9789814355247_0005

Iwata, H., Ebana, K., Uga, Y., & Hayashi, T. (2015). Genome-wide association study of biological shape based on elliptic Fourier analysis: A case study in rice grain shape variation. *Biological Shape Analysis*, 106–118 https://doi.org/10.1142/9789814704199_0007

Iwata, H., Nesumi, H., Ninomiya, S., Takano, Y., & Ukai, Y. (2002). Diallel analysis of leaf shape variations of citrus varieties based on elliptic Fourier descriptors. *Breeding Science*, *52*(2), 89–94. https://doi.org/10.1270/jsbbs.52.89

Iwata, H., 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

Iwata, H., Niikura, S., Matsuura, S., Takano, Y., & Ukai, Y. (2004). Genetic control of root shape at different growth stages in radish (*Raphanus sativus* L.). *Breeding Science*, *54*(2), 117–124. https://doi.org/10.1270/jsbbs.54.117

Iwata, H., & Ukai, Y. (2002). SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *The Journal of Heredity*, *93*(5), 384–385. https://doi.org/10.1093/jhered/93.5.384

Iyer-Pascuzzi, A. S., Symonova, O., Mileyko, Y., Hao, Y., Belcher, H., Harer, J., Weitz, J. S., & Benfey, P. N. (2010). Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. *Plant Physiology*, *152*(3), 1148–1157. https://doi.org/10.1104/pp.109.150748

Kadir, A. (2015). Leaf identification using polar Fourier transform and linear Bayes normal classifier. *Biological Shape Analysis*, 40–49. https://doi.org/10.1142/9789814704199_0002

Katz, E., Li, J. - J., Jaegle, B., Ashkenazy, H., Abrahams, R. S., Bagaza, C., Holden, S., Pires, J. C., Angelovici, R., & Kliebenstein, D. J. (2021). Genetic variation, environment and demography intersect to shape Arabidopsis defense metabolite variation across Europe. *eLife*, *10*. https://doi.org/10.7554/eLife.67784

Ke, Q. H., Zhang, J. M., & Tian, Y. (2013). Arabidopsis growth analysis using image processing technology. *Advanced Materials Research*, *712–715*, 2359–2363. https://doi.org/10.4028/www.scientific.net/amr.712-715.2359

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. *arXiv*. http://arxiv.org/abs/1312.6114v10

Klee, H. J., & Tieman, D. M. (2013). Genetic challenges of flavor improvement in tomato. *Trends in Genetics*, *29*(4), 257–262. https://doi.org/10.1016/j.tig.2012.12.003

Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Rosenberg, E. S. (2013). *Applied regression analysis and other multivariable methods*. Cengage Learning.

Klein, L. L., Caito, M., Chapnick, C., Kitchen, C., O'Hanlon, R., Chitwood, D. H., & Miller, A. J. (2017). Digital morphometrics of two North American grapevines (*Vitis*: Vitaceae) quantifies leaf variation between species, within species, and among individuals. *Frontiers in Plant Science*, *8*, 373. https://doi.org/10.3389/fpls.2017.00373

Klingenberg, C. P. (2011). MorphoJ: An integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*(2), 353–357. https://doi.org/10.1111/j.1755-0998.2010.02924.x

Klingenberg, C. P. (2015). Analyzing fluctuating asymmetry with geometric morphometrics: Concepts, methods, and applications. *Symmetry*, *7*(2), 843–934. https://doi.org/10.3390/sym7020843

Klingenberg, C. P., & McIntyre, G. S. (1998). Geometric morphometrics of developmental instability: Analyzing patterns of fluctuating asymmetry with procrustes methods. *Evolution; International Journal of Organic Evolution*, *52*(5), 1363–1375. https://doi.org/10.1111/j.1558-5646.1998.tb02018.x

Krause, M. R., González-Pérez, L., Crossa, J., Pérez-Rodríguez, P., Montesinos-López, O., Singh, R. P., Dreisigacker, S., Poland, J., Rutkoski, J., Sorrells, M., Gore, M. A., & Mondal, S. (2018). Hyperspectral reflectance-derived relationship matrices for genomic prediction of grain yield in wheat. *bioRxiv*, https://doi.org/10.1101/389825

Lane, H. M., & Murray, S. C. (2021). High throughput can produce better decisions than high accuracy when phenotyping plant populations. *Crop Science*, 61. https://doi.org/10.1002/csc2.20514

Lane, H. M., Murray, S. C., Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Rooney, D. K., Barrero-Farfan, I. D., De La Fuente, G. N., & Morgan, C. L. S. (2020). Phenomic selection and prediction of maize grain yield from near-infrared reflectance spectroscopy of kernels. *The Plant Phenome Journal*, *3*(1). https://doi.org/10.1002/ppj2.20002

Langlade, N. B., Feng, X., Dransfield, T., Copsey, L., Hanna, A. I., Thébaud, C., Bangham, A., Hudson, A., & Coen, E. (2005). Evolution through genetically controlled allometry space. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(29), 10221–10226. https://doi.org/10.1073/pnas.0504210102

Larsen, B., Migicovsky, Z., Jeppesen, A. A., Gardner, K. M., Toldam-Andersen, T. B., Myles, S., Ørgaard, M., Petersen, M. A., & Pedersen, C. (2019). Genome-wide association studies in apple reveal loci for aroma volatiles, sugar composition, and harvest date. *The Plant Genome*, *12*. https://doi.org/10.3835/plantgenome2018.12.0104

Leiboff, S., & Hake, S. (2019). Reconstructing the transcriptional ontogeny of maize and sorghum supports an inverse hourglass model of inflorescence development. *Current Biology: CB*, *29*(20), 3410–3419.e3. https://doi.org/10.1016/j.cub.2019.08.044

Liang, Z., Qiu, Y., & Schnable, J. C. (2020). Genome–phenome wide association in maize and arabidopsis identifies a common molecular and evolutionary signature. *Molecular Plant*, *13*(6), 907–922. https://doi.org/10.1016/j.molp.2020.03.003

Li, B., Cockerton, H. M., Johnson, A. W., Karlström, A., Stavridou, E., Deakin, G., & Harrison, R. J. (2020). Defining strawberry shape uniformity using 3D imaging and genetic mapping. *Horticulture Research*, *7*, 115. https://doi.org/10.1038/s41438-020-0337-x

Li, L., Zhang, Q., & Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors*, *14*(11), 20078–20111. https://doi.org/10.3390/s141120078

Li, M., An, H., Angelovici, R., Bagaza, C., Batushansky, A., Clark, L., Coneva, V., Donoghue, M. J., Edwards, E., Fajardo, D., Fang, H., Frank, M. H., Gallaher, T., Gebken, S., Hill, T., Jansky, S., Kaur, B., Klahs, P. C., Klein, L. L., …, & Chitwood, D. H. (2018). Topological data analysis as a morphometric method: Using persistent homology to demarcate a leaf morphospace. *Frontiers in Plant Science*, *9*, 553. https://doi.org/10.3389/fpls.2018.00553

Li, M., Duncan, K., Topp, C. N., & Chitwood, D. H. (2017). Persistent homology and the branching topologies of plants. *American Journal of Botany*, *104*(3), 349–353. https://doi.org/10.3732/ajb.1700046

Li, M., Frank, M. H., Coneva, V., Mio, W., Chitwood, D. H., & Topp, C. N. (2018). The persistent homology mathematical framework provides enhanced genotype-to-phenotype associations for plant morphology. *Plant Physiology*, *177*(4), 1382–1395. https://doi.org/10.1104/pp.18.00104

Liu, H., Long, S. - X., Pinson, S. R. M., Tang, Z., Guerinot, M. L., Salt, D. E., Zhao, F. - J., & Huang, X. - Y. (2021). Univariate and multivariate QTL analyses reveal covariance among mineral elements in the rice ionome. *Frontiers in Genetics*, *12*, 638555. https://doi.org/10.3389/fgene.2021.638555

Liu, S., Barrow, C. S., Hanlon, M., Lynch, J. P., & Bucksch, A. (2020). DIRT/3D: 3D root phenotyping for field grown maize (*Zea mays*). *bioRxiv*. https://doi.org/10.1101/2020.06.30.180059

Lobet, G. (2017). Image analysis in plant sciences: Publish then perish. *Trends in Plant Science*, *22*(7), 559–566. https://doi.org/10.1016/j.tplants.2017.05.002

Lv, J., Wang, Y., Ni, H., Wang, Q., Rong, H., Ma, Z., Yang, B., & Xu, L. (2019). Method for discriminating of the shape of overlapped apple fruit images. *Biosystems Engineering*, *186*, 118–129. https://doi.org/10.1016/j.biosystemseng.2019.07.003

Manacorda, C. A., & Asurmendi, S. (2018). Arabidopsis phenotyping through geometric morphometrics. *GigaScience*, *7*(7), https://doi.org/10.1093/gigascience/giy073

Merieux, N., Cordier, P., Wagner, M.-H., Ducournau, S., Aligon, S., Job, D., Grappin, P., & Grappin, E. (2021). ScreenSeed as a novel high throughput seed germination phenotyping method. *Scientific Reports*, *11*(1), 1404. https://doi.org/10.1038/s41598-020-79115-2

Migicovsky, Z., Gardner, K. M., Money, D., Sawler, J., Bloom, J. S., Moffett, P., Chao, C. T., Schwaninger, H., Fazio, G., Zhong, G. - Y., & Myles, S. (2016). Genome to phenome mapping in apple using historical data. *The Plant Genome*, *9*(2),. https://doi.org/10.3835/plantgenome2015.11.0113

Migicovsky, Z., Li, M., Chitwood, D. H., & Myles, S. (2017). Morphometrics reveals complex and heritable apple leaf shapes. *Frontiers in Plant Science*, *8*, 2185. https://doi.org/10.3389/fpls.2017.02185

Migicovsky, Z., Sawler, J., Gardner, K. M., Aradhya, M. K., Prins, B. H., Schwaninger, H. R., Bustamante, C. D., Buckler, E. S., Zhong, G. - Y., Brown, P. J., & Myles, S. (2017). Patterns of genomic and phenomic diversity in wine and table grapes. *Horticulture Research*, *4*, 17035. https://doi.org/10.1038/hortres.2017.35

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: For Cell and Molecular Biology*, *89*(1), 169–178. https://doi.org/10.1111/tpj.13320

Miyakawa, T. (2020). No raw data, no science: Another possible source of the reproducibility crisis. *Molecular Brain*, *13*(1), 24. https://doi.org/10.1186/s13041-020-0552-2

Moehring, J., Williams, E. R., & Piepho, H.-P. (2014). Efficiency of augmented p-rep designs in multi-environmental trials. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *127*(5), 1049–1060. https://doi.org/10.1007/s00122-014-2278-y

Momen, M., Campbell, M. T., Walia, H., & Morota, G. (2019). Predicting longitudinal traits derived from high-throughput phenomics in contrasting environments using genomic legendre polynomials and B-splines. *G3: Genes, Genomes, Genetics*, *9*(10), 3369–3380. https://doi.org/10.1534/g3.119.400346

Murray, S. C., Rooney, W. L., Hamblin, M. T., Mitchell, S. E., & Kresovich, S. (2009). Sweet sorghum genetic diversity and association mapping for brix and height. *The Plant Genome*, *2*(1). https://doi.org/10.3835/plantgenome2008.10.0011

Nagamatsu, S., Tsubone, M., Wada, T., Oku, K., Mori, M., Hirata, C., Hayashi, A., Tanabata, T., Isobe, S., Takata, K., & Shimomura, K. (2021). Strawberry fruit shape: Quantification by image analysis and QTL detection by genome-wide association analysis. *Breeding Science*, *71*(2), 167–175. https://doi.org/10.1270/jsbbs.19106

Nascimento, M. G. P., Mayo, S. J., & Andrade, I. M. (2021). Distinguishing the Brazilian mangrove species *Avicennia germinans* and *A. schaueriana* (Acanthaceae) by elliptic Fourier analysis of leaf shape. *Feddes Repertorium*, *132*(2), 77–107. https://doi.org/10.1002/fedr.202000025

Neto, J. C., Meyer, G. E., Jones, D. D., & Samal, A. K. (2006). Plant species identification using Elliptic Fourier leaf shape analysis. *Computers and Electronics in Agriculture*, *50*(2), 121–134. https://doi.org/10.1016/j.compag.2005.09.004

Nowak, J., Eng, R. C., Matz, T., Waack, M., Persson, S., Sampathkumar, A., & Nikoloski, Z. (2021). A network-based framework for shape analysis enables accurate characterization of leaf epidermal cells. *Nature Communications*, *12*(1), 458. https://doi.org/10.1038/s41467-020-20730-y

Oh, Y., Barbey, C. R., Chandra, S., Bai, J., Fan, Z., Plotto, A., Pillet, J., Folta, K. M., Whitaker, V. M., & Lee, S. (2021). Genomic characterization of the fruity aroma gene, *FaFAD1*, reveals a gene dosage effect on γ-decalactone production in strawberry (*Fragaria* × *ananassa*). *Frontiers in Plant Science*, *12*, 639345. https://doi.org/10.3389/fpls.2021.639345

Omori, H., & Iwata, H. (1998). Development of image analysis software, BioImage ver.1, for Windows to measure leaf area. *Agricultural Information Research*, *7*(2), 71–80. https://doi.org/10.3173/air.7.71

Peiffer, J. A., Flint-Garcia, S. A., De Leon, N., McMullen, M. D., Kaeppler, S. M., & Buckler, E. S. (2013). The genetic architecture of maize stalk strength. *PLOS ONE*, *8*(6), e67066. https://doi.org/10.1371/journal.pone.0067066

Peiffer, J. A., Romay, M. C., Gore, M. A., Flint-Garcia, S. A., Zhang, Z., Millard, M. J., Gardner, C. A. C., McMullen, M. D., Holland, J. B., Bradbury, P. J., & Buckler, E. S. (2014). The genetic architecture of maize height. *Genetics*, *196*(4), 1337–1356. https://doi.org/10.1534/genetics.113.159152

Peng, R. (2015). The reproducibility crisis in science: A statistical counterattack. *Significance*, *12*(3), 30–32. https://doi.org/10.1111/j.1740-9713.2015.00827.x

Pereira, C. S., Morais, R., & Reis, M. J. C. S. (2019). Deep learning techniques for grape plant species identification in natural images. *Sensors*, *19*(22). https://doi.org/10.3390/s19224850

Piepho, H. - P., Möhring, J., Schulz-Streeck, T., & Ogutu, J. O. (2012). A stage-wise approach for the analysis of multi-environment trials. *Biometrical Journal. Biometrische Zeitschrift*, *54*(6), 844–860. https://doi.org/10.1002/bimj.201100219

Pieruschka, R., & Schurr, U. (2019). Plant phenotyping: Past, present, and future. *Plant Phenomics*, *2019*. https://doi.org/10.34133/2019/7507131

Poland, J. A., & Nelson, R. J. (2011). In the eye of the beholder: The effect of rater variability and different rating scales on QTL mapping. *Phytopathology*, *101*(2), 290–298. https://doi.org/10.1094/PHYTO-03-10-0087

Rice, B. R., Fernandes, S. B., & Lipka, A. E. (2020). Multi-trait genome-wide association studies reveal loci associated with maize inflorescence and leaf architecture. *Plant and Cell Physiology*, *61*(8), 1427–1437. https://doi.org/10.1093/pcp/pcaa039

Rincent, R., Charpentier, J. - P., Faivre-Rampant, P., Paux, E., Le Gouis, J., Bastien, C., & Segura, V. (2018). Phenomic selection is a low-cost and high-throughput method based on indirect predictions: Proof of concept on wheat and poplar. *G3: Genes, Genomes, Genetics*, *8*(12), 3961–3972. https://doi.org/10.1534/g3.118.200760

Ristova, D., Rosas, U., Krouk, G., Ruffel, S., Birnbaum, K. D., & Coruzzi, G. M. (2013). RootScape: A landmark-based system for rapid screening of root architecture in Arabidopsis. *Plant Physiology*, *161*(3), 1086–1096. https://doi.org/10.1104/pp.112.210872

Rohlf, F. J. (1999). Shape statistics: Procrustes superimpositions and tangent spaces. *Journal of Classification*, *16*(2), 197–223. https://doi.org/10.1007/s003579900054

Runcie, D., & Cheng, H. (2019). Pitfalls and remedies for cross validation with multi-trait genomic prediction methods. *G3: Genes, Genomes, Genetics*, *9*(11), 3727–3741. https://doi.org/10.1101/595397 bioRxiv

Runcie, D. E., & Crawford, L. (2019). Fast and flexible linear mixed models for genome-wide genetics. *PLoS Genetics*, *15*(2), e1007978. https://doi.org/10.1371/journal.pgen.1007978

Runcie, D. E., & Mukherjee, S. (2013). Dissecting high-dimensional phenotypes with bayesian sparse factor analysis of genetic covariance matrices. *Genetics*, *194*(3), 753–767. https://doi.org/10.1534/genetics.113.151217

Rymaszewski, W., Dauzat, M., Bédiée, A., Rolland, G., Luchaire, N., Granier, C., Hennig, J., & Vile, D. (2018). Measurement of *Arabidopsis thaliana* plant traits using the PHENOPSIS phenotyping platform. *BIO-PROTOCOL*, *8*(4). https://doi.org/10.21769/bioprotoc.2739

Sainin, M. S., Ahmad, F., & Alfred, R. (2016). *Comparison of half and full-leaf shape feature extraction for leaf classification*. *AIP Conference Proceedings 1761*, 020092. https://doi.org/10.1063/1.4960932

Sanders, N. (2019). A balanced perspective on prediction and inference for data science in industry. *Harvard Data Science Review*, *1*(1). https://hdsr.mitpress.mit.edu/pub/a7gxkn0a/release/2

Sandhu, K. S., Mihalyov, P. D., Lewien, M. J., Pumphrey, M. O., & Carter, A. H. (2021). Combining genomic and phenomic information for predicting grain protein content and grain yield in spring wheat. *Frontiers in Plant Science*, *12*, 613300. https://doi.org/10.3389/fpls.2021.613300

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. - Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Schlautman, B., Diaz-Garcia, L., & Barriball, S. (2020). Morphometric approaches to promote the use of exotic germplasm for improved food security and resilience to climate change: A kura clover example. *Plant Science: An International Journal of Experimental Plant Biology*, *290*, 110319. https://doi.org/10.1016/j.plantsci.2019.110319

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Schulz-Streeck, T., Ogutu, J. O., & Piepho, H. - P. (2013). Comparisons of single-stage and two-stage approaches to genomic selection. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *126*(1), 69–82. https://doi.org/10.1007/s00122-012-1960-1

Schwieterman, M. L., Colquhoun, T. A., Jaworski, E. A., Bartoshuk, L. M., Gilbert, J. L., Tieman, D. M., Odabasi, A. Z., Moskowitz, H. R., Folta, K. M., Klee, H. J., Sims, C. A., Whitaker, V. M., & Clark, D. G. (2014). Strawberry flavor: Diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PLOS ONE*, *9*(2), e88446. https://doi.org/10.1371/journal.pone.0088446

Seethepalli, A., Guo, H., Liu, X., Griffiths, M., Almtarfi, H., Li, Z., Liu, S., Zare, A., Fritschi, F. B., Blancaflor, E. B., Ma, X. - F., & York, L. M. (2020). RhizoVision crown: An integrated hardware and software platform for root crown phenotyping. *Plant Phenomics*, *2020*, 3074916. https://doi.org/10.34133/2020/3074916

Shmueli, G. (2010). To explain or to predict? *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *25*(3), 289–310. https://doi.org/10.1214/10-STS330

Sirovich, L., & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America. A, Optics and Image Science*, *4*(3), 519–524. https://doi.org/10.1364/JOSAA.4.000519

Slice, D. E. (2007). Geometric morphometrics. *Annual Review of Anthropology*, *36*(1), 261–281. https://doi.org/10.1146/annurev.anthro.34.081804.120613

Slice, D. E., & Stitzel, J. (2004). Landmark-based geometric morphometrics and the study of allometry. *SAE Technical Paper Series*, https://doi.org/10.4271/2004-01-2181

Stewart, E. L., Wiesner-Hanks, T., Kaczmar, N., DeChant, C., Wu, H., Lipson, H., Nelson, R. J., & Gore, M. A. (2019). Quantitative phenotyping of northern leaf blight in UAV images using deep learning. *Remote Sensing*, *11*(19), 2209. https://doi.org/10.3390/rs11192209

Stoddart, C. (2016). Is there a reproducibility crisis in science? In *Nature*, https://doi.org/10.1038/d41586-019-00067-3

Suzuki, K., & Hirata, Y. (2011). Analysis of rice plant type using P-type Fourier descriptors and the search for related chromosomal regions. *Biological Shape Analysis*, https://doi.org/10.1142/9789814355247_0003

Tabb, A., & Medeiros, H. (2017). A robotic vision system to measure tree traits. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 6005–6012.

Tao, H., Feng, H., Xu, L., Miao, M., Yang, G., Yang, X., & Fan, L. (2020). Estimation of the yield and plant height of winter wheat using UAV-based hyperspectral images. *Sensors*, *20*(4), 1231. https://doi.org/10.3390/s20041231

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., & Bennett, M. (2017). Plant phenomics, from sensors to knowledge. *Current Biology*, *27*(15), R770–R783. https://doi.org/10.1016/j.cub.2017.05.055

Tieman, D., Zhu, G., Resende, M. F. R. Jr, Lin, T., Nguyen, C., Bies, D., Rambla, J., L., Beltran, K. S. O., Taylor, M., Zhang, B., Ikeda, H., Liu, Z., Fisher, J., Zemach, I., Monforte, A., Zamir, D., Granell, A., Kirst, M., Huang, S., & Klee, H. (2017). A chemical genetic roadmap to improved tomato flavor. *Science*, *355*(6323), 391–394. https://doi.org/10.1126/science.aal1556

Tovar, J. C., Hoyer, J. S., Lin, A., Tielking, A., Callen, S. T., Castillo Elizabeth, S., Miller, M., Tessman, M., Fahlgren, N., Carrington, J. C., Nusinow, D. A., & Gehan, M. A. (2018). Raspberry Pi-powered imaging for plant phenotyping. *Applications in Plant Sciences*, *6*(3), e1031. https://doi.org/10.1002/aps3.1031

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86. https://doi.org/10.1162/jocn.1991.3.1.71

Turner, S. D., Ellison, S. L., Senalik, D. A., Simon, P. W., Spalding, E. P., & Miller, N. D. (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. https://doi.org/10.3389/fpls.2018.01703

Ubbens, J., Cieslak, M., Prusinkiewicz, P., Parkin, I., Ebersbach, J., & Stavness, I. (2020). Latent space phenotyping: Automatic image-based phenotyping for treatment studies. *Plant Phenomics*, *2020*, 5801869. https://doi.org/10.34133/2020/5801869

Valluru, R., Gazave, E. E., Fernandes, S. B., Ferguson, J. N., Lozano, R., Hirannaiah, P., Zuo, T., Brown, P. J., Leakey, A. D. B., Gore, M. A., Buckler, E. S., & Bandillo, N. (2019). Deleterious mutation burden and its association with complex traits in sorghum (*Sorghum bicolor*). *Genetics*, *211*(3), 1075–1087. https://doi.org/10.1534/genetics.118.301742

van der Heijden, G., Song, Y., Horgan, G., Polder, G., Dieleman, A., Bink, M., Palloix, A., van Eeuwijk, F., & Glasbey, C. (2012). SPICY: Towards automated phenotyping of large pepper plants in the greenhouse. *Functional Plant Biology: FPB*, *39*(11), 870–877. https://doi.org/10.1071/FP12019

Van Ginkel, M., Reynolds, M. P., Trethowan, R., & Hernandez, E. (2008). Complementing the breeders eye with canopy temperature measurements. *International Symposium on Wheat Yield Potential*, 134.

Vasseur, F., Bresson, J., Wang, G., Schwab, R., & Weigel, D. (2018). Image-based methods for phenotyping growth dynamics and fitness components in *Arabidopsis thaliana*. *Plant Methods*, *14*(1). https://doi.org/10.1186/s13007-018-0331-6

Verma, S., Zurn, J. D., Salinas, N., Mathey, M. M., Denoyes, B., Hancock, J. F., Finn, C. E., Bassil, N. V., & Whitaker, V. M. (2017). Clarifying sub-genomic positions of QTLs for flowering habit and fruit quality in U.S. strawberry (*Fragaria* × *ananassa*) breeding populations using pedigree-based QTL analysis. *Horticulture Research*, *4*, 17062. https://doi.org/10.1038/hortres.2017.62

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/Netherlands Journal of Plant Breeding*, *200*(3), 429–439.

Vőfély, R. V., Gallagher, J., Pisano, G. D., Bartlett, M., & Braybrook, S. A. (2019). Of puzzles and pavements: A quantitative exploration of leaf epidermal cell shape. *The New Phytologist*, *221*(1), 540–552. https://doi.org/10.1111/nph.15461

Walter, A., Liebisch, F., & Hund, A. (2015). Plant phenotyping: From bean weighing to image analysis. *Plant Methods*, *11*, 14. https://doi.org/10.1186/s13007-015-0056-8

Wang, R., Qiu, Y., Zhou, Y., Liang, Z., & Schnable, J. C. (2020). A high-throughput phenotyping pipeline for image processing and functional

growth curve analysis. *Plant Phenomics*, *2020*, 1–8. https://doi.org/10.34133/2020/7481687

Wang, X., Singh, D., Marla, S., Morris, G., & Poland, J. (2018). Field-based high-throughput phenotyping of plant height in sorghum using different sensing technologies. *Plant Methods*, *14*(1). https://doi.org/10.1186/s13007-018-0324-5

Wang, X., Xuan, H., Evers, B., Shrestha, S., Pless, R., & Poland, J. (2019). High-throughput phenotyping with deep learning gives insight into the genetic architecture of flowering time in wheat. *GigaScience*, *8*(11). https://doi.org/10.1093/gigascience/giz120

Warman, C., Sullivan, C. M., Preece, J., Buchanan, M. E., Vejlupkova, Z., Jaiswal, P., & Fowler, J. E. (2021). A cost-effective maize ear phenotyping platform enables rapid categorization and quantification of kernels. *The Plant Journal: For Cell and Molecular Biology*, *106*(2), 566–579. https://doi.org/10.1111/tpj.15166

Willis, C., Lambert, M., McHenry, K., & Kirkpatrick, C. (2017). Container-based analysis environments for low-barrier access to research data. *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact*, 1–4.

Wu, T. - C., Belteton, S. A., Pack, J., Szymanski, D. B., & Umulis, D. M. (2016). LobeFinder: A convex hull-based method for quantitative boundary analyses of lobed plant cells. *Plant Physiology*, *171*(4), 2331–2342. https://doi.org/10.1104/pp.15.00972

Xu, H., & Bassel, G. W. (2020). Linking genes to shape in plants using morphometrics. *Annual Review of Genetics*, *54*, 417–437. https://doi.org/10.1146/annurev-genet-022620-094553

Yano, K., Morinaka, Y., Wang, F., Huang, P., Takehara, S., Hirai, T., Ito, A., Koketsu, E., Kawamura, M., Kotake, K., Yoshida, S., Endo, M., Tamiya, G., Kitano, H., Ueguchi-Tanaka, M., Hirano, K., & Matsuoka, M. (2019). GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(42), 21262–21267. https://doi.org/10.1073/pnas.1904964116

York, L., & Lobet, G. (2017). Phenomics of root system architecture: Measuring and analyzing root phenes. *The Plant Cell*, *29*. https://pdfs.semanticscholar.org/7160/4eb4c254ff8c141fd65b7c0854e2d35b9feb.pdf

York, L. M. (2019). Functional phenomics: An emerging field integrating high-throughput phenotyping, physiology, and bioinformatics. *Journal of Experimental Botany*, *70*(2), 379–386. https://doi.org/10.1093/jxb/ery379

Yuan, L. - M., Cai, J. - R., Sun, L., & Ye, C. (2016). A preliminary discrimination of cluster disqualified shape for table grape by mono-camera multi-perspective simultaneously imaging approach. *Food Analytical Methods*, *9*(3), 758–767. https://doi.org/10.1007/s12161-015-0250-3

Zhao, J., Sauvage, C., Zhao, J., Bitton, F., Bauchet, G., Liu, D., Huang, S., Tieman, D. M., Klee, H. J., & Causse, M. (2019). Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. *Nature Communications*, *10*(1), 1534. https://doi.org/10.1038/s41467-019-09462-w

Zhou, L., Gu, X., Cheng, S., Yang, G., Shu, M., & Sun, Q. (2020). Analysis of plant height changes of lodged maize using UAV-LiDAR data. *Agriculture*, *10*(5), 146. https://doi.org/10.3390/agriculture10050146

Zhou, N., Siegel, Z. D., Zarecor, S., Lee, N., Campbell, D. A., Andorf, C. M., Nettleton, D., Lawrence-Dill, C. J., Ganapathysubramanian, B., Kelly, J. W., & Friedberg, I. (2018). Crowdsourcing image analysis for plant phenomics to generate ground truth data for machine learning. *PLoS Computational Biology*, *14*(7), e1006337. https://doi.org/10.1371/journal.pcbi.1006337

Zingaretti, L. M., Monfort, A., & Pérez-Enciso, M. (2021). Automatic fruit morphology phenome and genetic analysis: An application in the octoploid strawberry. *Plant Phenomics (Washington, D.C.)*, *2021*, 9812910.

**How to cite this article:** Feldmann, M, Gage, J, Turner-Hissong, S, & Ubbens, J. Images carried before the fire: The power, promise, and responsibility of latent phenotyping in plants. *Plant Phenome J*, 2021;*4:*e20023. https://doi.org/10.1002/ppj2.20023