Exploration of high-throughput data for heat tolerance selection in Capsicum annuum

Nathan Fumia, Michael Kantar, Ya-ping Lin, Roland Schafleitner, Véronique Lefebvre, Ilan Paran, Andreas Börner, Maria José Diez, Jaime Prohens, Arnaud Bovy, et al.

To cite this version:

Nathan Fumia, Michael Kantar, Ya-ping Lin, Roland Schafleitner, Véronique Lefebvre, et al.. Exploration of high-throughput data for heat tolerance selection in Capsicum annuum. The Plant Phenome Journal, 2023, 6 (1), pp.e20071. 10.1002/ppj2.20071 . hal-04176395

Received: 15 November 2022 | Accepted: 13 April 2023

DOI: 10.1002/ppj2.20071

SPECIAL SECTION: PHENOTYPING TOOLS AND BREEDING FOR CLIMATE ADAPTATION AND EMISSION MIGRATION

The Plant Phenome Journal OPEN ACCESS

Exploration of high-throughput data for heat tolerance selection in *Capsicum annuum*

Nathan Fumia[1] | **Michael Kantar**[1] | **Ya-ping Lin**[2] | **Roland Schafleitner**[2] | **Véronique Lefebvre**[3] | **Ilan Paran**[4] | **Andreas Börner**[5] | **Maria José Diez**[6] | **Jaime Prohens**[6] | **Arnaud Bovy**[7] | **Filiz Boyaci**[8] | **Gancho Pasev**[9] | **Pasquale Tripodi**[10] | **Lorenzo Barchi**[11] | **Giovanni Giuliano**[12] | **Derek W. Barchenger**[2]

[1]Department of Tropical Plant and Soil Science, University of Hawaii at Manoa, Honolulu, Hawaii, USA

[2]World Vegetable Center, Tainan, Taiwan

[3]Unité de Génétique et Amélioration des Fruits et Légumes (GAFL), L'Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE), Montfavet, France

[4]Institute of Plant Sciences, Agricultural Research Organization, The Volcani Center, Bet Dagan, Israel

[5]Genome Diversity, Department Genebank, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Seeland, Germany

[6]Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València (UPV), Valencia, Spain

[7]Plant Breeding, Wageningen University & Research, Wageningen, The Netherlands

[8]Bati Akdeniz Agricultural Research Institute, Antalya, Turkey

[9]Maritsa Vegetable Crops Research Institute, Plovdiv, Bulgaria

[10]Research Centre for Vegetable and Ornamental Crops, Council for Agricultural Research and Economics (CREA), Pontecagnano Faiano, Italy

[11]Department of Agricultural, Forest and Food Sciences (DISAFA), Plant Genetics, University of Torino, Grugliasco, Italy

[12]Casaccia Research Centre, Italian National Agency for New Technologies, Energy, and Sustainable Economic Development, Rome, Italy

Correspondence
Derek W. Barchenger, World Vegetable Center, Shanhua, Tainan, Taiwan.
Email: derek.barchenger@worldveg.org

Assigned to Associate Editor Melanie Ooi.

Funding information
European Union's Horizon 2020 Research and Innovation Programme, Grant/Award Number: 677379(G2P-SOL); Rural Development Administration of Korea; UK aid from the UK government; The College of Tropical Agriculture and Human Resources, University of Hawaii; USDA Cooperative State Research, Education and Extension (CSREES), Grant/Award Number: HAW08039-H; Long-term strategic donors to the World Vegetable

Abstract

Recently, there has been a substantial increase in high-throughput technologies that generate highly complex large datasets for use in the sciences. Plant breeding and genetics have benefited from this data explosion where many public and private institutions now implement genomic and phenomic data to predict performance thus informing germplasm selection. However, the multitude of methodologies and data generates a situation of strategic uncertainty. We set out to compare different methods of genomic and phenomic selection in the *Capsicum* core collection, developed through the G2P-SOL project, producing a combination of unique and similar selected genotypes for heat tolerance. Combined, the methods tested identified a total of 33 genotypes that show tremendous promise for use as parents in heat tolerance breeding: with 13 of these being present in more than 1 selection method. Combining classical and multispectral phenotyping methods produced better selection results

---

Center (Taiwan, Germany, Thailand, Philippines, Korea and Japan); Hawaii Agriculture Research Center and the University of Hawaii Office of Sustainability

than either method alone. When each method was conducted without being informed by the other, similar results were obtained. Our weighted rank-sum selection index identified 10 entries across environments that show heat tolerance, 8 of which are also selected within heat environments. This suggests that different breeding programs can reach similar results despite having different logistical constraints. Our case study within pepper germplasm using phenomic and genomic data exhibits the potential to compensate for the dearth of germplasm knowledge with high-throughput data as well as the converse, to compensate for logistical or financial constraint to new technologies with breeder knowledge.

1 | INTRODUCTION

Global food systems are under increasing pressure with concentrating human populations, an increasingly stochastic climate, and future projections predicting declines in food production for major production regions (Godfray et al., 2010; Myers et al., 2017; Ray et al., 2019). Further, future geographic regions for many major staples are expected to shift, potentially leading to the abandonment of current production locales (Estoque et al., 2019; Pironon et al., 2019). Understanding the change in major abiotic stress factors and beginning the process of breeding for adaptation to these changes is essential to meeting future food security needs.

Plant breeding is based on efficiently selecting for traits that are desirable to humans (Bernardo, 2014). Selection indices, the combination of multiple traits into a single metric, have been a part of breeding for decades (Hazel, 1943; Henderson, 1950; Lush, 1935; Lush, 1948; Smith, 1936). Indices are often normalized and corrected for nongenic effects; occasionally having additional weights associated with trait importance or even economic relevance (e.g., Smith–Hazel Index). Recently, indices and genomic selection (GS) have been extended to high-throughput phenotypes through complex predictive modeling. For example, Rutkowski et al. (2016) used multivariate linear regression to incorporate secondary traits from aerial high throughput phenotyping (HTP) platforms and realized improved genomic prediction accuracies, and Montesinos-Lopez et al. (2021) found that when genotype-by-environment interaction is present, genomic best linear unbiased prediction (G-BLUP) possesses better predictive accuracy across multiple species than alternative deep-learning algorithms. These studies show that although more complex models have been examined, predictive models that leverage the power and simplicity of linear regression to understand secondary traits within complex genotype-by-environment interaction are comparable and easier to interpret (Montesinos-López et al., 2021; Rutkoski et al., 2016). The goal of a selection index is to generate simultaneous gains in multiple traits whilst accounting for trait covariation and interaction, ideally leading to rapid market introduction of novel varieties. For an index to be broadly applicable, it should be interpretable and easy to calculate.

HTP has introduced new and more kinds of phenotypes for breeding (Araus & Cairns, 2014). These phenotypes include image data, growth data, abiotic stress data, and yield-related components accompanied by recent work exploring the relationship of these new data with traditional phenotypes to leverage their use in genetics and breeding (Montesinos-Lopez et al., 2017; Sun et al., 2017; Ziyomo & Bernardo, 2013). The dimensionality of HTP often makes interpretation difficult, although machine learning techniques reduce the complexity of the data and enable decision-making. There are many machine learning algorithms (e.g., classify *k*-means, support vector machines, and random forests), which have been used to help understand abiotic stress in plants (Ghosal et al., 2018; Singh et al., 2016, 2018).

Toward the goal of leveraging the positive aspects of the many different selection methods in pepper (*Capsicum annuum* L.), we apply and compare three techniques. First, we use an anomaly analysis through *k*-means clustering, where we explore individuals that show normal behavior under both heat (therefore anomalous) and normal conditions using all the phenotypic data. We then used this to train a random forest model that can classify heat-resistant types with only access to trials under optimal conditions and found that this model can detect germplasm near perfect accuracy. Following this phenomic-based analysis, we conduct genomic-based prediction to predict component traits and form a weighted rank-sum selection index (WRSSI) to identify genomic-based heat-tolerant lines. Lastly, we move with comparisons of selected lines from three selection methodologies: (1) breeder's intuition; (2) phenomics-based through anomaly analysis; and (3) genomics-based through predictive modeling and selection index.

2 | MATERIALS AND METHODS

2.1 | Germplasm and trial design

Three-hundred *Capsicum* entries were selected as a representative sample of the global pepper core collection. Only members of *C. annuum* were selected for this experiment to increase uniformity in germination rate and growth, essential for high-throughput automated phenotyping. These entries were chosen as they represent 84.1% of all *C. annuum* diversity (Tripodi et al., 2021). Trials were designed as a randomized complete block with three replications, each of them with four plants at a spacing of 0.45 by 0.45 m between plants with 0.7 m between plots and 1.5 m between rows. For the experiment, the entries were evaluated during three seasons. We classify two seasons as heat stress seasons and one as cool season, serving as the control. For the two heat stress seasons, sowing and transplanting were on March 20 and April 29, 2020 and March 4 and April 7, 2021, respectively. Sowing and transplanting dates for the control season were September 7 and October 19, 2020. The 2020 heat-stress season was characterized by long periods of stable, but high temperatures, and exposed the entries to moderate heat stress. The 2021 heat-stress season had a period of extreme heat that corresponded with flowering for most of the entries, exposing the entries to severe heat stress (Figure 1). All three experiments were conducted at the World Vegetable Center in Shanhua, Tainan, Taiwan (lat. 23.1˚N, long. 120.3˚E, elevation 12 m.a.s.l.). Phenotypic information was gathered through a combination of both manual and automated data collection. The manual data (classical phenotypes) included days to flowering, fruit maturity, yield components of fruit length, width, weight, and yield, as described by Barchenger et al. (2018, 2020) as well as pollen concentration and activity using impedance flow cytometry following the protocol of Lin et al. (2022). In addition, we manually collected leaf temperature, which was recorded between 12:45 and 13:45 p.m. using handheld infrared thermometers held ∼5 cm above the leaf. The automated data were collected using the PlantEye 3D-Spectral Scan F500 (Phenospex, Heerlen, the Netherlands). The 3D plant multispectral data in the near-infrared range (NIR: 720–750 nm) and at three color bands (RED: 620–645 nm, GREEN: 530–540 nm, and BLUE: 460–585 nm) were automatically collected using the PHENA analytics platform (Phenospex) and visualized and analyzed by HortControl 3.0 (Phenospex). In total, 75 phenotypic observations, including plant morphology, color, leaf temperature, pollen quality, and yield component data, were recorded.

2.2 | Selection methodologies

To compare multiple different types of selection, we created a flowchart to arrive at different sets of cultivars using different selection methodologies (Figure 2). These selection methodologies include a selection of lines based upon three methodologies with varying complexities: (1) breeder's intuition; (2) phenomics-based selection through anomaly analysis in *k*-means clustering; and (3) genomics-based through predictive modeling and selection index formation. Breeder's intuition leverages observation for informed selection. Phenomics-based selection utilizes classically and multispectral collected phenotypes to select lines as anomalies within specified *k*-means clusters. Genomics-based selection combines all data (e.g., classic and multispectral phenotypes, genomic marker data) to identify the most accurate model and utilize normalized best linear unbiased predictions (BLUPs) for each trait weighted by univariate model accuracies to form a WRSSI.

Core Ideas

- Combining classical phenotyping and multispectral phenotyping performed better than either method alone.
- Selection Indices combining genomic and phenomic data provided increased prediction accuracy in stress environments.
- There was overlap between selection methods, indicating progress can be made by using new technologies.

2.2.1 | Breeder's intuition selection

To set a baseline of selection, our on-site lead breeder makes selection based upon intuition. This is the art of breeding where the knowledge of germplasm and consistent field performance observations serve as the basis of selection. These selections are made based upon just that, field observations of line performance across a gradient of temperature through three separate field trials as well as summary statistics for different component traits.

2.2.2 | Phenomics-based selection

Our second selection scheme is based upon all phenotypic data and its subsets: classical and multispectral combined, classical alone, and multispectral alone. We leverage *k*-means clustering (Hartigan, 1975; Hartigan & Wong, 1979) to specify two clusters ($K = 2$) to represent groupings of (1) normal phenotypic values and (2) stress phenotypic values. These clusters are formed by the aggregation of all 75 phenotypic observations (N dimensions) from all three trials for every

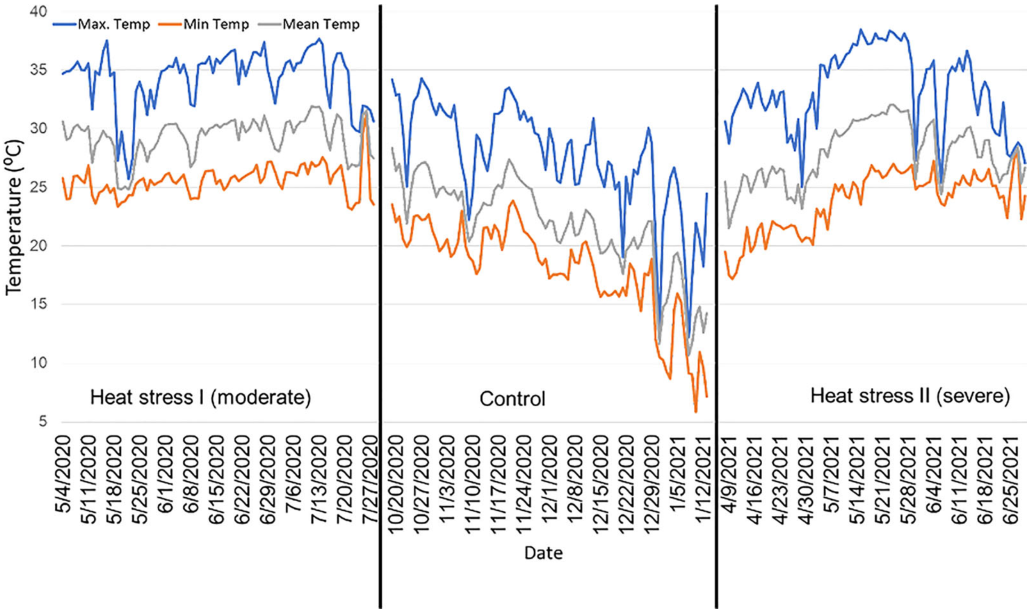

FIGURE 1 Temperature measured during the experiment over the course of three different field seasons, representing moderate heat stress, the control, and extreme heat stress.

genotype (*M* points). The algorithm, therefore, is iteratively searching for a *K*-partition with local optima by moving points between clusters to minimize the within-cluster sum of squares. Succinctly, the phenomics data is the input matrix of *M* genotypes (rows) and *N* phenotypes (columns) where every observation *i* is shifted between *K* clusters until within-cluster sums of squares are minimized. Optimal clusters designated from "fvis_nbclust()" in package "factoextra" at two clusters ($K = 2$) using methods (1) within-cluster sums of squares ("wss") and (2) average silhouette ("silhouette"); as well as trying to define expected phenomic values of normal versus temperature stressed. *k*-Means analysis performed with a maximum number of iterations at 1000 with a nstart, or random sets chosen, of 25.

Input data is from three trials: control, increased temperature stress, and severe temperature stress. All trials used the same 300 genotypes, each with 3 replications for a total of 9 phenotypic observations per trait per genotype. Selection of genotypes can be made upon this method by (1) identifying the cluster primarily formed by observations of genotype by control trial interaction, and (2) scanning observations of genotype by stress trials interaction that cluster with the genotype by control trial interaction observations. In effect, this is a phenomics selection tool to identify stress plastic genotypes, the phenotypes of which are affected less than the difference between the clusters. Moreover, these stress plastic genotypes have a phenomics evaluation under the within the sum of squares for the control cluster.

We perform cluster prediction using random forest machine learning and used predicted clusters of a given genotype by trial interaction to understand variance components and component trait contribution in cluster formation and assignment. Therefore, the random forest model followed the form of cluster assignment as the response, building predictive trees through all phenotypes observed at the genotype-by-environment level (nine observations per genotype: three replicates in each of the three environments).

To improve the understanding of phenomics applications in breeding, we applied the above techniques to subset phenotypic data as two groups: (1) classically collected phenotypes and (2) multispectral phenotypes (Table S1). This separation permits a comparison of the applicability of phenotyping techniques in abiotic stress breeding of peppers and aims to answer the question: Is more data really useful? Group 1 (classically collected phenotypes; 12 phenotypes) are primarily yield-related traits such as fruit size, width, weight, and pollen attributes, whereas Group 2 (technology collected; 63 phenotypes) is data collected from multispectral camera vision. Selections from these three groups (full data, classic phenotypes, multispectral phenotypes) were made by the identification of anomalies. These anomalies, as stated earlier, are those observations from genotype by stress environment interaction that cluster with the genotype by control environment cluster. Therefore, the anomalies are genotype performance in stress conditions that perform within the sum of squares difference of the control observations.

2.2.3 | Phenomics and genomics indices

Our next selection scheme is the implementation of GS, using phenomics and genomics data. Genotype processing is the

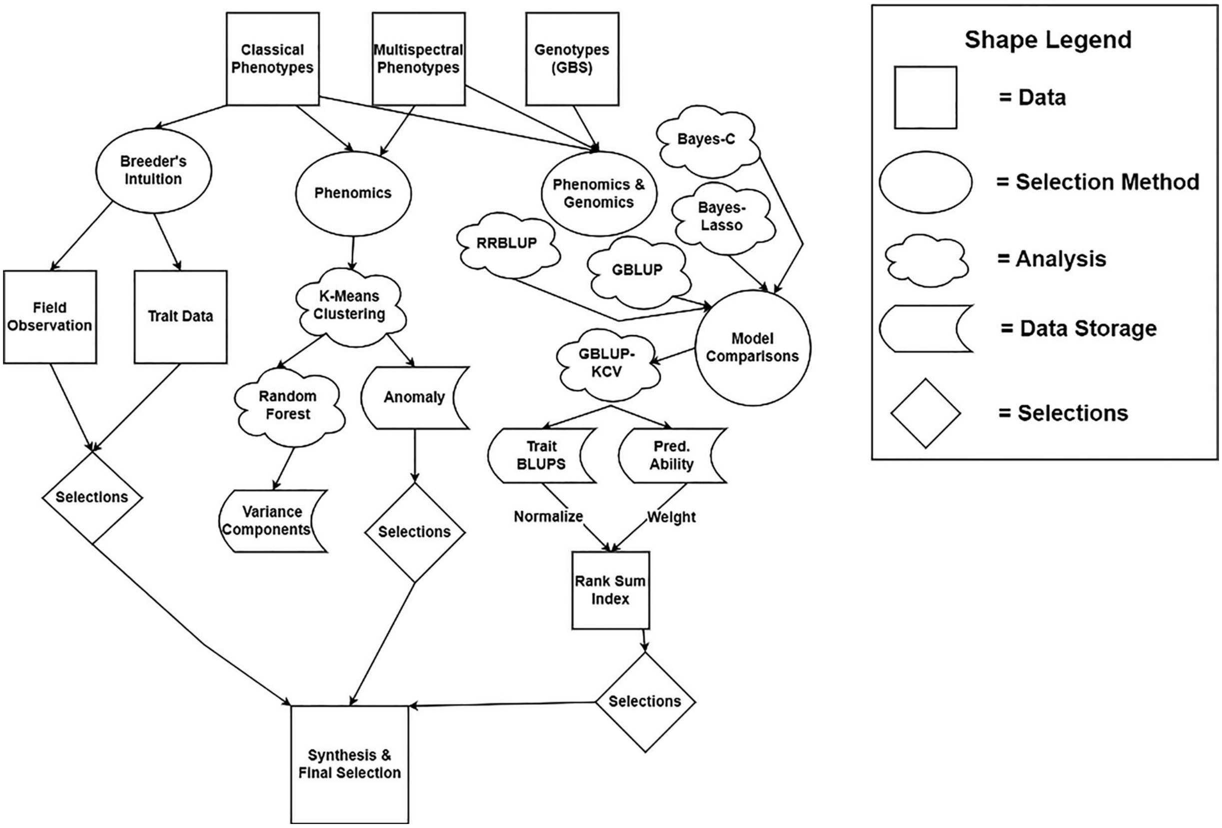

FIGURE 2 Flowchart showing the experimental methods used to make selections.

first step in this process. Marker information for 431 pepper genotypes from the World Vegetable Center were input as .vcf files and filtered for missing data (Tripodi et al., 2021). The genotype matrix was filtered for three criteria: removal of markers with greater than 50% missing data, removal of individuals with greater than 50% missing marker data, and removal of markers with extreme minor allele frequencies ($\text{MAF} < 0.05$). Filtering reduced the number of markers from 23,462 to 5840 markers, which were well distributed across the genome. Individuals with genotype information but no phenotype data (131 genotypes) were removed from the genotype matrix, whereas genotypes with phenotypic information (300 genotypes) are used in model training and identification. Any additional missing data in the genotype matrix was imputed using Markov chain implementation as provided by the R package "NAM" (Xavier et al., 2015).

Once genotype was filtered and imputed, we moved to test different GS models for the predictive ability of our many component traits. Trait BLUPs were used as the response in univariate GS mixed models following linear parametric (ridge-regression BLUP [RR-BLUP] and G-BLUP) and nonlinear parametric (Bayes-LASSO and Bayes-C) methods of genomic prediction to account for their different assumptions (Jannink et al., 2010; Voss-Fels et al., 2019; Meuwissen et al., 2001; Bernardo, 2020; Habier et al., 2007). Nonlinear parametric methods of Bayes-LASSO and Bayes-C were used to try to increase the prediction accuracy of the models (Park & Casella, 2008; de los Campos et al., 2009; Heslot et al., 2012). These models were fit using R packages "rrBLUP" (Endelman, 2011) and "BGLR" (Perez & de los Campos, 2014). This cumbersome process was undertaken to find the model within each group (linear parametric and non-linear parametric) with the highest mean predictive ability (average across all 75 traits and 3 environments).

The predictive abilities of each model by every trait within each environment were recorded (Tables S2–S4). The average model accuracy (e.g., an average of G-BLUP performance for each trait across environments) was used to compare against the other model used in each model type (e.g., G-BLUP or RR-BLUP in linear parametric). However, the accuracies of these models are likely overestimated because of model overfitting, an expected outcome, but useful in this model selection process. To account for the inevitable overfitting in our full data models, we introduced techniques of leave-one-out cross-validation (LOOCV) to the best model in linear parametric (G-BLUP) and nonlinear parametric (Bayes-C) methods to obtain more practical univariate BLUPs within each environment and inform confidence in the predictive ability of a given trait. Considering the relative similarity of predictive ability between G-BLUP-LOOCV and Bayes-C-LOOCV, we move forward with G-BLUP-KCV (*K*-fold cross-validation) predictions for selection index formation given its computational efficiency.

We employed the technique of rank-sum index to form a model accuracy–adjusted (i.e., weighted) rank-sum selection index. Predicted phenotypic values (BLUP) from the most accurate univariate models, G-BLUP-KCV, were used to rank genotypes within each trait for BLUP of phenotypic performance. Each trait BLUP was normalized (0,1) to adjust values to the same order of magnitude. Moreover, because predictive abilities of each univariate model varied for each trait, we adjust the weighting of the ranking by $1 - r_j$ (model accuracy r for trait j), entrusting more weight to those traits with high predictive ability and deflating the importance of those with low predictive ability. This method was iterated for every trait within each environment (control, increased temperature, severe temperature). Selections from each group were made via truncation selection.

2.3 | Genotype-by-environment prediction

The univariate model accuracies through cross-validation show substantial decrease in the mean predictive ability as compared to full data models. We then applied these methods of KCV for multivariate GS by environment interactions. We applied a genotype-by-environment model to make the best prediction for selection considering changing environmental stress by the specification of variance–covariance structure in genotype by environment, using compound symmetry. In short, compound symmetry is a variance–covariance structure specified through (1) a genetic correlation structure with equal genotypic variance and equal correlation among pairs of environments as well as (2) residual variance assumed the same across environments (Covarrubias-Pazaran, 2016). These predictions were performed primarily to identify component traits with large genotype-by-environment variation and used for internal purposes of phenotyping priority, similar to the random forest models to understand trait contribution to cluster assignment.

3 | RESULTS AND DISCUSSION

3.1 | Phenomics-based selection

Different clusters emerged for the control (normal temperature—cluster 2) and the treatment (cluster 1—temperature stressed). Some genotypes from the temperature stress condition trials grouped with control (normal condition phenotypes); this implies that these entries perform in the increased temperature conditions like that of most entries in the control condition, which form the bulk of the cluster (Figure 3a). Individuals that were stressed but grouped with controls did not show clear phenotypic symptoms of heat stress (Table 1). This clustering was repeated using only the classic phenotypes again showing clear clustering that differentiated the control (normal temperature—cluster 2) and the treatments (cluster 1—temperature stressed). Again, there was a small subset of entries that grouped with the control despite being in the stress treatment (Table 1; Figure 3b). Clustering was then performed with only multispectral phenotypes, again individual genotypes under temperature stress clustered with the control (Table 1; Figure 3c). Two key insights from this were that combining classical and multispectral phenotypes provided more clear differentiation of clusters than either type of phenotyping alone, indicating that more data provides better information. When used independently of one another, classical phenotyping and multispectral phenotypes return the same best genotypes, indicating that breeders can remain within different program constraints to reach similar results.

Random forest regression was then applied to each genotype to assess if it was possible to accurately predict the group (e.g., control or stress cluster) assignment based solely on using the phenotypic data from the experimental trials and to understand component trait contributions and to identify the most important chromosomes in heat response (Figures S1 and S2). This additional information allowed for the exploration of common mechanisms from the same genomic regions; we found that there were several signals of stress tolerance across the entire genome, with different genotypes having beneficial alleles from different genomic regions (Figures S1 and S2). Again, having more phenotypes provided an increased benefit, despite correlations between traits (Figure S3). In our case, this came in the form of an increased correlation between model accuracy in predicting cluster assignment and a number of phenotypes used in the model ($R^2 = 0.426$). Variables with the highest contribution include leaf area, biomass, and mean normalized difference vegetation index in the full and multispectral models with yield and days to anthesis in the classical phenotype-based model.

3.2 | Genomic selection

In addition to being phenotyped, every entry was also genotyped (see Tripodi et al., 2021). Both genetic and phenotypic information was used to calculate BLUPs, which then form the training and validation sets used to compare multiple techniques of assessing breeding value. We included genotype-by-environment interactions to select genotypes with the highest multivariate performance for classically collected phenotypes, multispectral collected phenotypes, and all phenotypes combined.

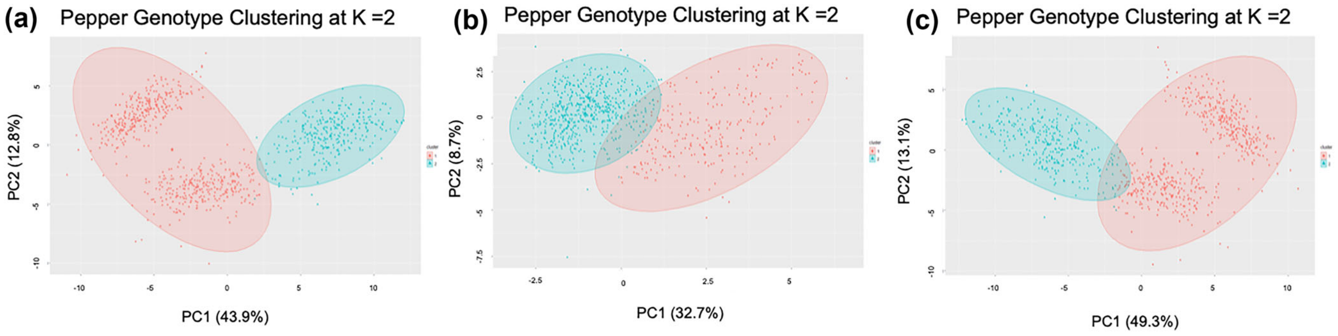

FIGURE 3 *k*-Means clustering of pepper accessions under control and heat stress environments on the principal components of (a) both classic and multispectral phenotyping; (b) classic phenotyping; (c) multispectral phenotyping. In each case, it is possible to see clustering based on the control and stress phenotypes, with some overlap between the two. Red indicates the treated environment, and blue indicates the control environment. The overlap suggests that some genotypes perform stably in stress and control environments.

TABLE 1 Individual genotypes which are clustered with controls when exposed to stress.

k-Means clustering with all phenotypes	*k*-Means clustering with classical phenotypes	*k*-Means clustering with multispectral phenotypes
GPC063760, GPC078540, GPC084560, GPC093310, GPC097130, GPC107030, GPC107960, GPC112760, GPC112830, GPC113230, GPC115410, GPC116220, GPC116260, GPC121020, GPC124980, GPC128790	GPC003240, GPC003310, GPC003370, GPC020570, GPC028750	GPC003240, GPC003310, GPC003370, GPC020570, GPC028750

Note: Interestingly, entries from classical phenotyping and multispectral phenotyping resulted in the same entries performing the same.

3.2.1 | Model comparisons

BLUPs were calculated for each trait separately using mixed model regression. There were strong correlations among the BLUP-predicted values between different groups of traits (Figure S3). To predict univariate trait performance per individual within each environment (control, increased temperature, severe temperature), phenotypic information was separated by environment and then informed with genotype information using RR-BLUP (rrBLUP), G-BLUP (rrBLUP), Bayes-LASSO (BGLR), and Bayes-C (BGLR) (Tables S2–S4).

In RR-BLUP, across environments, the mean univariate model accuracy was 0.899 in the control, 0.903 in increased temperature, and 0.887 in severe temperature (Tables S2–S4). The minimum accuracies for univariate traits were for the multispectral imaging traits leaf inclination in control (0.707), leaf inclination in increased temperature (0.588), and plant senescence reflectance index mean in severe temperature (0.723). The opposite is true for maximum accuracies found within each environment, typically being those classically collected, such as fruit length in control (0.989), fruit length in moderate stress (0.994), and fruit number under severe stress (0.998). We observed a clear relationship between environments (Figure 3).

G-BLUP identified no differences in environment-specific univariate trait prediction accuracies. However, the introduction of priors with Bayesian inference following double-exponential (Bayes-LASSO) and two-component mixtures prior with a point of mass at zero and a Gaussian slab (Bayes-C) show a slight shift in prediction accuracy. Bayes-LASSO had a mean univariate model prediction accuracy of 0.894 in control, 0.901 in increased temperature, and 0.878 in severe temperature. Bayes-C finds a mean univariate model prediction accuracy of 0.894 in control, 0.902 in increased temperature, and 0.890 in severe temperature.

3.2.2 | Finding the optimal model

To account for potential overfitting, cross-validation analysis was employed using LOOCV on the linear model (G-BLUP) and non-linear model (Bayes-C) with the best average predictive abilities. G-BLUP-LOOCV had a mean univariate model prediction accuracy of 0.487 in control, 0.486 in increased temperature, and 0.429 in severe temperature (Tables S2–S4).

In the control, univariate prediction accuracies were generally lower for multispectral traits than classical traits (Table S2). Under increased temperature, multispectral traits still had low accuracy but there were slight increases for classical traits (Table S3). Further increases in temperature further decreased the predictive accuracy of multispectral traits but did not impact classical traits (Table S4). Bayes-C-LOOCV finds a mean univariate model prediction accuracy (average correlation between experimental BLUPS and Bayes-C-LOOCV) of 0.466 in control, 0.470 in increased temperature, and 0.453 in severe temperature. Again, using this model we see lower prediction accuracies for multispectral traits than classical traits. The shift to increased temperature maintains the low predictive ability of multispectral traits with classically collected traits maintaining high accuracies (e.g., fruit width—0.846 and yield—0.516). Further increases in temperature decrease the predictive accuracy of multispectral traits with classically collected traits maintaining high accuracies (e.g., fruit width at 0.75, yield at 0.53). Unexpectedly, pollen activity accuracy increases with temperature; this may be due to having a better understanding of when pollen will not be viable.

The performance of LOOCV methods highlights the overfitting of univariate models by almost 50% ($r = 0.892$ in Bayes-C control vs. $r = 0.466$ in Bayes-C-LOOCV control). However, the computational demand of this method was significant, and thus, we move forward with KCV ($k = 10$) for index formation inputs. G-BLUP-KCV resulted in a mean univariate model predictive ability (average correlation between experimental BLUPS and G-BLUP-KCV) of 0.473 in control, 0.477 in increased temperature, and 0.439 in severe temperature. The similar accuracies of G-BLUP-KCV and Bayes-C-LOOCV, along with improved computational efficiency, highlight the potential use of G-BLUP-KCV for the univariate prediction of the testing population of genotypes ($n = 130$).

3.3 | Selection indices

We then formed a WRSSI, which is the summation of genotypic ranking by trait within each environment (control, temperature stress, severe temperature stress) and then weighted by model predictive ability, to compare overall genotype performance in all environments as well as performance in stress environments. We performed truncation selection based on these indices and found some disparities between selected genotypes among indices and intuitive methods. Genotype-by-environment modeling illustrated that the component traits show phenotypic variation (Figure 4a) and those show little phenotypic variation (Figure 4b) through changing environmental conditions. These same phenotypically variable traits exhibit increased correlation with the WRSSI as environments increase in heat (Figure 4c), whereas those narrowly variables show a reduction (Figure 4d).

Different entries performed best when coinciding with all environments or stress environments (Table 2; Figure 5). A large portion of the best performing lines under the control were also the best performing lines under moderate heat, but fewer lines remained the best performers when exposed to severe heat (Figure 5). There is a consistent group of low performing lines across all three treatments, which will not do well in any growing environment. The intermediate lines have complex patterning of rank changes, highlighting the complexity of heat tolerance. We observe a trend for many intermediate-ranked entries under moderate heat to become the best performing entries under severe heat, which may indicate that the best performers under severe heat would be moderate performers anywhere else. This observation matches previous work in pepper breeding where lines with more consistent performance were typically moderate and those with extreme performance excel in a single environment (Barchenger et al., 2020). Filtration on the mean stress environments caused the selection of individual genotypes with phenotypically plastic trait performances (i.e., being able to produce well in both control and stress environments), introducing different selections from the total mean (GPC020720, GPC033360) as well as shifting the ranking of genotypes for performance in temperature stress. The selection indices were able to find clear evidence that some traits were more useful than others when trying to identify the best individual genotypes under stress (Figure 3; Figure S4).

3.4 | Breeders' intuition and recommendations

By exploring conventional and multispectral phenotypes, individually and together it was possible to begin to understand what can be gained from different types of information provided when selecting. It was also possible to compare these to the breeders' intuition, which is based on a combination of data and insight from observing the plants. We found that classic phenotypes performed as well as multispectral phenotypes when explored individually, but when combined they performed better than either set of phenotypes by themselves. This was true when genomic information was included resulting in the ability to identify promising entries increasing again. Although not unexpected, the ability to include an increasing amount of data and combine methods led to better selection outcomes for complex traits.

When exploring genomic and phenomic selection in the core collection of pepper, each method produced a clear list of entries that showed promise (Table 2). We have identified a total of 33 genotypes that show high promise for being used as parents in heat tolerance breeding for pepper, 13 of these

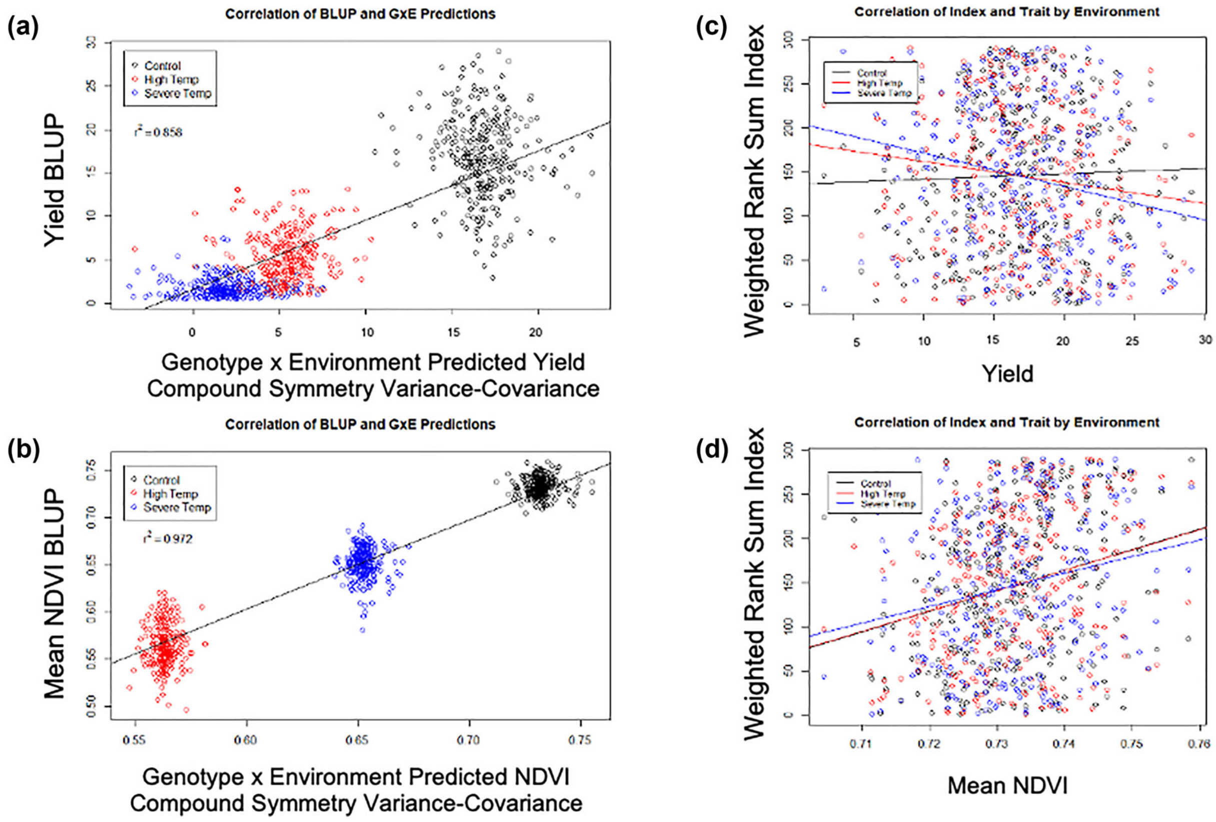

FIGURE 4 Relationships of key parameters of pepper accessions between best linear unbiased prediction (BLUP) in the different environments: (a) relationship between Yield BLUP and GxE Predicted Yield; (b) relationship between mean normalize difference vegetation index (NDVI) BLUP and GxE Predicted NDVI; (c) correlation between the weighted rank-sum index and yield where a low rank sum is best; (d) correlation between mean NDVI and the weighted rank-sum index.

TABLE 2 Individual genotypes that performed best using the weighted rank-sum selection index for each environment.

Stress environments	Across environments	Stress environments (with untested genotypes)	Across environments (with untested genotypes)
GPC014400, GPC014630, GPC023040, GPC022320, GPC000300, GPC020470, GPC040670, GPC036000, GPC022310, GPC071220	GPC020470, GPC040670, GPC014630, GPC014400, GPC022310, GPC023040, GPC071220, GPC008400, GPC000300, GPC001490	GPC018160, GPC028460, GPC057060, GPC035430, GPC049320, GPC028850, GPC042150, GPC037070, GPC123700, GPC003410	GPC028460, GPC018160, GPC042150, GPC057060, GPC049320, GPC028850, GPC035430, GPC003410, GPC008380, GPC007080

genotypes being present in more than one selection method. Although there is tremendous utility in using genomic and phenotypic technology (e.g., increased prediction accuracy), there was a large amount of overlap when using only classically measured phenotypes. For example, the WRSSI has a correlation of 0.30 ($p < 0.001$) with the cluster assignment using phenomics-based *k*-means analysis. This modest accuracy of predicted assignment in the relatively simple analysis technique of *k*-means with the complex GS WRSSI supports the idea that breeding programs can work within their constraints to achieve similar results in heat tolerance breeding. As breeding programs continue to invest in new technologies, there is clear evidence that knowing the germplasm breeders are working with can compensate for not having access to all the newest technologies, but the converse is also true, where having access to all the latest technologies can help a new breeder overcome a lack of understanding of their new germplasm.

When selecting for heat tolerance based on visual observation (not data collection, but intuition), the plant breeder selects for adaptability. Heat tolerance is a complex trait, which involves complex component traits, each of which has interactions with each other and the environment and are almost certainly quantitatively controlled. Visual selection considers many traits at once but also ignores many traits that cannot be easily observed, such as pollen concentration

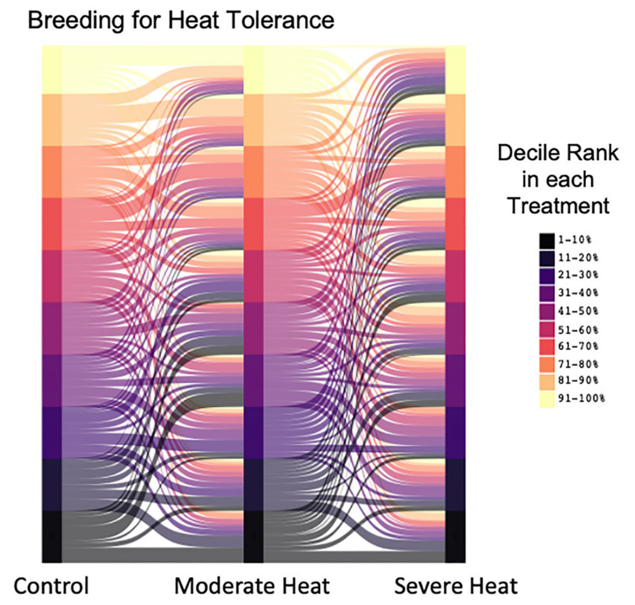

FIGURE 5 Rank change of weighted rank-sum selection index with increasing temperature stress for pepper entries. Majority of germplasm changes are between deciles of ranks in this study. We see those in the best (bottom) ranks (1%–10%) being relatively stable across treatments; however, in the other deciles, there was a large amount of rank change. The higher proportion of good ranks shifting to bad ranks in transition from moderate to severe stress than from control to moderate stress indicates that much of the germplasm has tolerance to some heat but not to severe heat.

and many multispectral traits. Among the intuitively selected entries, based purely on visual observation (Table S5), six were also identified as clustering similarly between stress and control environments (Table 1) or performed best using the WRSSI for the stress environment (Table 2). This overlap of selection based on the breeders' eye, classically or manually collected data and multispectral data supports the use of modern technologies to identify heat-tolerant entries. However, the question also arises: Is the use of advanced technology to phenotype for heat stress tolerance beneficial enough to warrant the investment? Given that an experienced plant breeder can generally select the best entries, it may not be worthwhile to utilize this type of technology for selection, depending upon breeding and research program constraints. However, the use of multispectral and HTP has the benefit to allow for genome-wide association studies for individual component traits that contribute to heat stress tolerance. The use of genome-wide association studies allows us to develop molecular markers to select for unseen traits, limiting the required investment in more manual and human observational selections. The combination of all techniques most certainly improves selection and could result in larger gains, but consideration must be given to the complexity of heat tolerance breeding where genetic variation in your population is the foundation of maintaining a sustainable breeding program under increasingly stochastic climate.

4 | CONCLUSIONS

We explore a combination of genomic-based selection, phenomic-based selection, and breeder intuition. Although these different methods are often placed in opposition to each other, we see clear evidence of complementarity. Using all the data provided better overall information, whereas using individual datasets provided comparable results. This highlights that despite constraints, breeding programs can leverage many different methods to make progress in sustainable cultivar development from stochastic climate.

AUTHOR CONTRIBUTIONS

Nathan Fumia: Conceptualization, Formal analysis, Writing – original draft. **Michael Kantar**: Conceptualization; Visualization; Writing – review & editing. **Ya-Ping Lin**: Formal analysis; Writing – review & editing. **Roland Schafleitner**: Conceptualization; Resources; Writing – review & editing. **Véronique Lefebvre**: Writing – review & editing. **ilan paran**: Writing – review & editing. **Andreas Börner**: Writing – review & editing. **Maria José Diez**: Writing – review & editing. **Jaime Prohens**: Writing – review & editing. **Arnaud Bovy**: Writing – review & editing. **Filiz Boyaci**: Writing – review & editing. **Gancho Pasev**: Writing – review & editing. **Pasquale Tripodi**: Writing – review & editing. **Lorenzo Barchi**: Writing – review & editing. **Giovanni Giuliano**: Writing – review & editing. **Derek Barchenger**: Conceptualization; Data curation; Formal analysis; Funding acquisition; Writing – review & editing.

ACKNOWLEDGMENTS

Funding for this research was provided by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 677379 (G2P-SOL project: Linking genetic resources, genomes and phenotypes of Solanaceous crops), the Rural Development Administration of Korea and long-term strategic donors to the World Vegetable Center: Taiwan, UK aid from the UK Government, United States Agency for International Development (USAID), Australian Centre for International Agricultural Research (ACIAR), Germany, Thailand, Philippines, Korea, and Japan. We would like to thank the Hawaii Agriculture Research Center and the University of Hawaii Office of Sustainability for their support of Nathan Fumia through the Sustainable Agriculture Fellowship. We would like to recognize the hosts and participants of the Summer Institute in Plant Breeding (2022) at the University of Minnesota and in particular, Dr. Aaron Lorenz for his section "Data Bootcamp for Genomic

Prediction in Plant Breeding.” Additional funding provided by The College of Tropical Agriculture and Human Resources, University of Hawaii at Manoa; USDA Cooperative State Research, Education and Extension (CSREES), Grant/Award Number: HAW08039-H.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this article.

DATA AVAILABILITY STATEMENT

All raw data are available at http://www.g2p-sol.eu and https://worldveg.tind.io/record/75778?ln=en. Analysis scripts and outputs are available at https://github.com/Nfumia/Capsicum_selection.

ORCID

Ya-ping Lin https://orcid.org/0000-0002-9575-2007
Jaime Prohens https://orcid.org/0000-0003-1181-9065
Derek W. Barchenger https://orcid.org/0000-0002-5948-3372

REFERENCES

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, *19*, 52–61. https://doi.org/10.1016/j.tplants.2013.09.008

Barchenger, D. W., Clark, R. A., Gniffke, P. A., Ledesma, D. R., Lin, S.-W., Hanson, P., & Kumar, S. (2018). Stability of yield and yield components of pepper (*Capsicum annuum*), and evaluation of publicly available predictive meteorological data in East and Southeast Asia. *HortScience*, *53*(12), 1776–1783. https://doi.org/10.21273/HORTSCI13581-18

Barchenger, D. W., Thandar, K., Myint, T. G. i., Hung, T. N., Hung, N. Q., Lin, S.-W., Wang, Y.-W., & Lin, T.-H. (2020). Yield and yield component performance of chile pepper in Myanmar and Vietnam. *HortTechnology*, *30*(3), 463–467. https://doi.org/10.21273/HORTTECH04580-20

Bernardo, R. (2014). *Essentials of plant breeding*. Stemma Press.

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

Bernardo, R. (2021). Predictive breeding in maize during the last 90 years. *Crop Science*, *61*, 2872–2881. https://doi.org/10.1002/csc2.20529

Covarrubias-Pazaran, G. (2016). Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS One*, *11*(6), e0156744. https://doi.org/10.1371/journal.pone.0156744

De Los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K., & Cotes, J. M. (2009). Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics*, *182*(1), 375–385. https://doi.org/10.1534/genetics.109.101501

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, *4*(3), 250–255. https://doi.org/10.3835/plantgenome2011.08.0024

Estoque, R. C., Gomi, K., Togawa, T., Ooba, M., Hijioka, Y., Akiyama, C. M., Nakamura, S., Yoshioka, A., & Kuroda, K. (2019). Scenario-based land abandonment projections: Method, application and implications. *Science of the Total Environment*, *692*, 903–916. https://doi.org/10.1016/j.scitotenv.2019.07.204

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 4613–4618. https://doi.org/10.1073/pnas.1716999115

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., & Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people. *Science*, *327*, 812–818. https://doi.org/10.1126/science.1185383

Habier, D., Fernando, R. L., & Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, *177*(4), 2389–2397. https://doi.org/10.1534/genetics.107.081190

Hartigan, J. A. (1975). Clustering algorithms. Wiley series in probability and mathematical statistics. In *Applied probability and statistics*. Wiley.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *28*(1), 100–108.

Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics*, *28*, 476–490. https://doi.org/10.1093/genetics/28.6.476

Henderson, C. R. (1950). Estimation of genetic parameters. *Annals of Mathematical Statistics*, *21*, 309–310.

Heslot, N., Yang, H.-P., Sorrells, M. E., & Jannink, J.-L. (2012). Genomic selection in plant breeding: A comparison of models. *Crop Science*, *52*(1), 146–160. https://doi.org/10.2135/cropsci2011.06.0297

Jannink, J.-L., Lorenz, A. J., & Iwata, H. (2010). Genomic selection in plant breeding: From theory to practice. *Briefings in Functional Genomics*, *9*(2), 166–177. https://doi.org/10.1093/bfgp/elq001

Lin, S.-W., Lin, T.-H., Yee, C. K. M., Chen, J., Wang, Y.-W., Nalla, M. K., & Barchenger, D. W. (2022). Impedance flow cytometry for selection of pollen traits under high temperature stress in pepper. *HortScience*, *57*(2), 181–190. https://doi.org/10.21273/HORTSCI16258-21

Lush, J. L. (1935). Progeny test and individual performance as indicators of an animal’s breeding value. *Journal of Dairy Science*, *18*, 1–19. https://doi.org/10.3168/jds.S0022-0302(35)93109-5

Lush, J. L. (1948). *The genetics of populations*. Iowa State College.

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Montesinos-López, O. A., Montesinos-López, A., Crossa, J., De Los Campos, G., Alvarado, G., Suchismita, M., Rutkoski, J., González-Pérez, L., & Burgueño, J. (2017). Predicting grain yield using canopy multispectral reflectance in wheat breeding data. *Plant Methods*, *13*, 4. https://doi.org/10.1186/s13007-016-0154-2

Montesinos-López, O. A., Montesinos-López, A., Mosqueda-González, B. A., Bentley, A. R., Lillemo, M., Varshney, R. K., & Crossa, J. (2021). A new deep learning calibration method enhances genome-based prediction of continuous crop traits. *Frontiers in Genetics*, *12*, 798840.

Myers, S. S., Smith, M. R., Guth, S., Golden, C. D., Vaitla, B., Mueller, N. D., Dangour, A. D., & Huybers, P. (2017). Climate change and global food systems: Potential impacts on food security and undernutrition. *Annual Review of Public Health*, *38*, 259–277. https://doi.org/10.1146/annurev-publhealth-031816-044356

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, *103*(482), 681–686. https://doi.org/10.1198/016214508000000337

Pérez, P., & De Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, *198*(2), 483–495. https://doi.org/10.1534/genetics.114.164442

Pironon, S., Etherington, T. R., Borrell, J. S., Kühn, N., Macias-Fauria, M., Ondo, I., Tovar, C., Wilkin, P., & Willis, K. J. (2019). Potential adaptive strategies for 29 sub-Saharan crops under future climate change. *Nature Climate Change*, *9*, 758–763. https://doi.org/10.1038/s41558-019-0585-7

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., & Chatterjee, S. (2019). Climate change has likely already affected global food production. *PLoS One*, *14*, 217148. https://doi.org/10.1371/journal.pone.0217148

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L. G., Crossa, J., Reynolds, M., & Singh, R. (2016). Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3: Genes, Genomes, Genetics*, *6*(9), 2799–2808.

Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science*, *21*, 110–124. https://doi.org/10.1016/j.tplants.2015.10.015

Singh, A. K., Ganapathysubramanian, B., Sarkar, S., & Singh, A. (2018). Deep learning for plant stress phenotyping: trends and future perspectives. *Trends in Plant Science*, *23*, 883–898. https://doi.org/10.1016/j.tplants.2018.07.004

Smith, H. F. (1936). A discriminant function for plant selection. *Annals of Eugenics*, *7*, 240–250. https://doi.org/10.1111/j.1469-1809.1936.tb02143.x

Sun, J., Rutkoski, J. E., Poland, J. A., Crossa, J., Jannink, J.-L., & Sorrells, M. E. (2017). Multitrait, random regression, or simple repeatability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield. *Plant Genome*, *10*(2). https://doi.org/10.3835/plantgenome2016.11.0111

Tripodi, P., Rabanus-Wallace, M. T., Barchi, L., Kale, S., Esposito, S., Acquadro, A., Schafleitner, R., Van Zonneveld, M., Prohens, J., Diez, M. J., Börner, A., Salinier, J., Caromel, B., Bovy, A., Boyaci, F., Pasev, G., Brandt, R., Himmelbach, A., Portis, E., … Stein, N. (2021). Global range expansion history of pepper (*Capsicum* spp.) revealed by over 10,000 genebank accessions. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(34), e2104315118. https://doi.org/10.1073/pnas.2104315118

Voss-Fels, K. P., Cooper, M., & Hayes, B. J. (2019). Accelerating crop genetic gains with genomic selection. *Theoretical and Applied Genetics*, *132*(3), 669–686. https://doi.org/10.1007/s00122-018-3270-8

Xavier, A., Xu, S., Muir, W. M., & Rainey, K. M. (2015). NAM: Association studies in multiple populations. *Bioinformatics*, *31*(23), 3862–3864. https://doi.org/10.1093/bioinformatics/btv448

Ziyomo, C., & Bernardo, R. (2013). Drought tolerance in maize: Indirect selection through secondary traits versus genomewide selection. *Crop Science*, *53*, 1269–1275. https://doi.org/10.2135/cropsci2012.11.0651

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Fumia, N., Kantar, M., Lin, Y.-P., Schafleitner, R., Lefebvre, V., Paran, I., Börner, A., Diez, M. J., Prohens, J., Bovy, A., Boyaci, F., Pasev, G., Tripodi, P., Barchi, L., Giuliano, G., & Barchenger, D. W. (2023). Exploration of high-throughput data for heat tolerance selection in *Capsicum annuum*. *The Plant Phenome Journal*, *6*, e20071. https://doi.org/10.1002/ppj2.20071