A low-cost greenhouse-based high-throughput phenotyping platform for genetic studies: a case study in maize under inoculation with plant growth-promoting bacteria

Rafael Massahiro Yassue[1], Giovanni Galli[1], Ronaldo Borsato Junior[1], Hao Cheng[2], Gota Morota[3,4,*], and Roberto Fritsche-Neto[1,5,*]

[1]Department of Genetics, Luiz de Queiroz College of Agriculture, University of São Paulo, São Paulo, Brazil

[2]Department of Animal Science, University of California, Davis, USA

[3]Department of Animal and Poultry Sciences, Virginia Polytechnic Institute and State University, Blacksburg, USA

[4]Center for Advanced Innovation in Agriculture, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 USA

[5]Quantitative Genetics and Biometrics Cluster, International Rice Research Institute, Los Baños, Philippines

Running title: A low-cost high-throughput maize phenotyping platform to study plant growth-promoting bacteria

Keywords: biostimulant, genomic correlation, genomic heritability, image-derived phenotyping, maize

Abbreviations: canopy coverage (CC); canopy volume (CV); genomic best linear unbiased prediction (GBLUP); ground control points (GCP); ground resolution of the orthomosaics (GRO); growing degree days (GDD); high-throughput phenotyping (HTP); high-throughput phenotyping plant height (PH_{HTP}); Luria-Bertani medium (LB); normalized difference vegetation index (NDVI); number of fully expanded leaves (NL); plant growth-promoting bacteria (PGPB); plant height (PH); real time kinematic (RTK); shoot dry mass (SDM); single-nucleotide polymorphism (SNP); unoccupied aerial vehicle (UAV).

Core ideas

- A low-cost greenhouse-based HTP platform was developed.

- Image-derived phenotypes presented moderate to high genomic heritabilities and correlations.

- Plant growth-promoting bacteria can improve plant resilience under nitrogen-limited conditions.

* Corresponding author

E-mail: morota@vt.edu (GM) and r.fritscheneto@irri.org (RFN)

ORCID: 0000-0002-7424-2227 (RMY), 0000-0003-2792-4695 (RBJ), 0000-0002-3400-7978 (GG), 0000-0001-5146-7231 (HC), 0000-0002-3567-6911 (GM), and 0000-0003-4310-0047 (RFN)

Email addresses: rafael.yassue@usp.br (RMY), giovannigalli@alumni.usp.br (GG), ronaldoborsatojr@usp.br (RBJ), qtlcheng@ucdavis.edu (HC), morota@vt.edu (GM), and r.fritscheneto@irri.org (RFN)

Abstract

Greenhouse-based high-throughput phenotyping (HTP) presents a useful approach for studying novel plant growth-promoting bacteria (PGPB). Despite the potential of this approach to leverage genetic variability for breeding new maize cultivars exhibiting highly stable symbiosis with PGPB, greenhouse-based HTP platforms are not yet widely used because they are highly expensive; hence, it is challenging to perform HTP studies under a limited budget. In this study, we built a low-cost greenhouse-based HTP platform to collect growth-related image-derived phenotypes. We assessed 360 inbred maize lines with or without PGPB inoculation under nitrogen-limited conditions. Plant height, canopy coverage, and canopy volume obtained from photogrammetry were evaluated five times during early maize development. A plant biomass index was constructed as a function of plant height and canopy coverage. Inoculation with PGPB promoted plant growth. Phenotypic correlations between the image-derived phenotypes and manual measurements were at least 0.6. The genomic heritability estimates of the image-derived phenotypes ranged from 0.23 to 0.54. Moderate-to-strong genomic correlations between the plant biomass index and shoot dry mass (0.24–0.47) and between HTP-based plant height and manually measured plant height (0.55–0.68) across the developmental stages showed the utility of our HTP platform. Collectively, our results demonstrate the usefulness of the low-cost HTP platform for large-scale genetic and management studies to capture plant growth.

3

Introduction

Recent studies have reported the benefit of using plant growth-promoting bacteria (PGPB) to increase yield and resilience against biotic and abiotic stresses (Arif et al., 2020; Kumar et al., 2020) through various molecular mechanisms, including nitrogen fixation and phytohormone production (Compant et al., 2010; Manoj et al., 2020). Importantly, Wintermans et al. (2016) and Vidotti et al. (2019a) found a differential response of genotypes under PGPB inoculation, suggesting that the response has a genetic basis. These findings opened frontiers for a plant breeding program to breed new cultivars having highly stable PGPB responses (Vidotti et al., 2019b). However, the difficulty of monitoring a large number of lines across phenological growth stages under different inoculation conditions constrains our ability to analyze the genetics of dynamic PGPB responses.

With the advancement in genotyping technologies, phenotyping is considered a new bottleneck in plant breeding (Furbank and Tester, 2011; Araus et al., 2018). Image-derived high-throughput phenotyping (HTP) presents a new avenue for automatic characterization of plants, owing to its capacity to generate difficult-to-measure phenotypes over time using advanced sensors and cameras (Mazis et al., 2020; Campbell et al., 2019). Greenhouse-based HTP platforms have been developed to evaluate a number of plant responses, such as morphological (Brichet et al., 2017), disease (Thomas et al., 2018), and physiological (Wang et al., 2018) under microbial inoculants (Chai et al., 2021) and biotic and abiotic stresses (Araus and Cairns, 2014; Campbell et al., 2018). Therefore, leveraging HTP to evaluate hundreds or thousands of genotypes non-destructively under different managements (Araus et al., 2018; Rouphael et al., 2018) is a promising approach to study PGPB. The choice of the HTP platform largely depends on the trade-off between the precision of phenotypes, the number of managements it can evaluate, and cost.

One major factor limiting the wide deployment of image-derived HTP in plant breeding

programs is the high cost of setting up an HTP platform, especially for small breeding programs or research institutions. In field trials, an unoccupied aerial vehicle (UAV) is the commonly used cost-effective HTP technology to collect high-throughput data (Xie and Yang, 2020). In greenhouses, conveyor (plant-to-sensor) and benchtop (sensor-to-plant) type systems are often used for automated HTP platforms (Li et al., 2021). The conveyor type automatically transports potted plants into an imaging room. In contrast, the benchtop type is equipped with a computer-controlled mechanical arm that can automatically locate the position of a plant for phenotyping. Although both conveyor and benchtop systems support diverse cameras, their installation costs are too expensive and may require modification of the existing greenhouse facilities. When there are budget constraints, researchers are motivated to build self-developed HTP platforms because large-scale greenhous-based HTP platforms are produced mainly by commercial companies (Czedik-Eysenberg et al., 2018), which are forbiddingly expensive.

Several efforts have been made to develop a novel low-cost custom greenhouse HTP platform (Zhou et al., 2018; Du et al., 2021). The most common approach is to use a sliding track or cable railing system to move the imaging system that consists of the camera in the x and y axes. The images are processed using image stitching or photogrammetry techniques to obtain 2-D or 3-D phenotypes. However, this type of HTP platform is yet to be widely adopted in genetic and management studies because the number of genotypes or managements that can be accommodated is limited. Therefore, the objective of this study was to build a low-cost non-commercial sensor-to-plant greenhouse-based HTP platform using a multispectral camera that has the capacity to accommodate hundreds of maize lines and develop an image-processing pipeline to obtain growth-related image-derived phenotypes. We assessed the utility of the image-derived phenotypes by evaluating 360 genotypes under different PGPB management in the early stages of maize development.

Materials and Methods

Low-cost high-throughput phenotyping platform

A low-cost greenhouse HTP platform was built, wherein the camera was positioned in a way that it obtained images from directly above the plants. The system was built in a conventional greenhouse with dimensions of 3.5 × 11 × 6 m height, width, and length, respectively. A cooling wall and ventilation were used to maintain the desired temperature, and additional luminosity was supplied using LED lamps.

The image capture system was inspired by the UAV flight plans. It consists of two fixed parallel tracks (9 m) and one mobile perpendicular track (5 m). They were positioned 2.5 m above the ground. The two parallel tracks were fixed to the greenhouse roof, as well as two support tracks to ensure stability and alignment. The parallel tracks move the perpendicular track along the x-axis, whereas the perpendicular track moves the sensors along the y-axis. Each track contained an individual 96-watt electric motor. These electric motors were remotely controlled to achieve the desired overlap (Figure 1). The speed of the tracks was set at 0.16 m/s.

A medium-density fiberboard platform was designed to accommodate the multispectral camera, light sensor, and battery. The fiberboard platform was attached to the y-axis mobile track. Four ground control points (GCP) geo-referenced with real-time kinematic (RTK) were used to assemble the orthomosaics. Top-view image data collection was performed between 12:00 and 13:00 with an overlap of 80 % frontal and 70% lateral views. The multispectral camera used was a Parrot Sequoia (Parrot SA, Paris, France), including green (550 nm), red (660 nm), red-edge (735 nm), and near-infrared (790 nm).

6

Image processing and data extraction

Multispectral images were processed by assembling orthomosaics and the dense point cloud using Agisoft Metashape software (Agisoft LLC, St. Petersburg, Russia). The images were imported, aligned, and optimized using GCP. This was followed by the calculation of the dense point clouds and the stitching of orthomosaics.

The orthomosaics were analyzed using QGIS software (QGIS Development Team, 2021) to obtain a shapefile for each plot. The plots were manually identified, and a geometry point was assigned at the center of the plant. Then, a round positive buffer of 0.10 m was drawn for each plot. The shapefile of each plot was manually adjusted to reduce overlaps across plants. We applied image segmentation to the orthomosaics using the normalized difference vegetation index (NDVI) (Rouse et al., 1974) with a threshold of 0.35 to separate canopy vegetation from the background. The reflectance of each plot was calculated as the mean of each wavelength (green, red, red-edge, and near-infrared) using the R package FIELDimageR (Matias et al., 2020). The NDVI was calculated using the following formula: $NDVI = (NIR - RED)/ (NIR + RED)$, where NIR and RED are the reflectances at the near-infrared and red wavelengths, respectively. Canopy coverage (CC) was calculated from the sum of the pixels in the canopy vegetation and transformed to cm^2 based on the resolution of the orthomosaics ($mm\ pixel^{-1}$).

Dense cloud points were used to estimate plant height (PH_{HTP}) and canopy volume (CV). Each point from the dense cloud point was composed of GPS coordinates (latitude, longitude, and altitude in the universal transverse mercator). The dense cloud point data were processed using the R package lidR (Roussel and Auty, 2021). A round positive buffer of 0.01 m was generated at the center of each plant to obtain the corresponding points of each plot. PH_{HTP} was constructed from the difference between the 90 percentile of the top of the point cloud altitude and the pot altitude before plant germination (0 leaves) (Figure 2) (Galli et al., 2021). The image-derived plant biomass index, f(biomass), was derived

7

163 from the product of PH_{HTP} (cm) and CC (cm^2) (Li et al., 2020). For CV, the dense cloud

164 points were filtered by colors using the "Select Points by Color" function in the Agisoft

165 Metashape software to remove the background. Plants were then reconstructed from the

166 point cloud data, and the CV was estimated using the α-shape algorithm (Lafarge and

167 Pateiro-Lopez, 2020). The algorithm requires an α value that controls the tightness of the

168 3-D reconstruction of the points. The optimal value of α that yielded the greatest correlation

169 with manual measurements was 0.01 (Moreno et al., 2020).

Plant growth-promoting bacteria experiment

171 A tropical maize association panel containing 360 inbred lines was used to study the response

172 to PGPB. Of these, 179 inbred lines were from the Luiz de Queiroz College of Agriculture-

173 University of Sao Paulo (ESALQ-USP) and 181 were from the Instituto de Desenvolvimento

174 Rural do Paraná (IAPAR). More information about this panel is available on the Mendeley

175 platform (https://data.mendeley.com/datasets/5gvznd2b3n).

176 The inbred lines were evaluated under two managements: with (B+) and without (B-)

177 PGPB inoculation under nitrogen stress. The B+ management consisted of a synthetic pop-

178 ulation of four PGPB. *Bacillus thuringiensis* RZ2MS9, *Delftia* sp. RZ4MS18 (Batista et al.,

179 2018, 2021), *Pantoea agglomerans* 33.1 (Quecine et al., 2012), and *Azospirillum brasilense*

180 Ab-v5 (Hungria et al., 2010) were selected based on a preliminary experiment that showed

181 their ability to promote growth when co-inoculated. Each species was grown individually

182 in Luria-Bertani (LB) medium at 28°C with agitation at 150 rpm for 24 h. The synthetic

183 population was composed of an adjusted volume of each bacterial culture medium containing

184 approximately 10^8 colony-forming units/mL. The B- management consisted of an inoculum

185 with liquid LB only. Each plot containing three seeds was individually inoculated with 1

186 ml of the respective management, agitated, and sown afterwards. Each line was replicated

187 twice across time, and each replication was composed of an augmented block design with six

188 blocks and three common checks.

189 A total of 13,826 single-nucleotide polymorphisms (SNPs) were available for the maize

190 inbred lines using a genotyping-by-sequencing method following the two-enzymes (PstI and

191 MseI) protocol (Sim et al., 2012; Poland et al., 2012). DNA was extracted using the

192 cetyltrimethylammonium bromide method (Doyle and Doyle, 1987). SNP calling was per-

193 formed using the TASSEL 5.0 software (Bradbury et al., 2007) with B73 (B73-RefGen_v4)

194 as the reference genome. The SNP markers were filtered if the call rate was less than 90%,

195 non-biallelic, and the minor allele frequency was less than 5%. Missing marker codes were

196 imputed using the Beagle 5.0 software (Browning et al., 2018). Markers with pairwise linkage

197 disequilibrium higher than 0.99, were removed using the SNPRelate R package (Zheng et al.,

198 2012).

Manually measured and high-throughput phenotypes

200 The experiments were performed at ESALQ-USP in Brazil (22°42'39 "S; 47°38'09 "W, alti-

201 tude 540 m). The final evaluation was conducted when most plants had developed six fully

202 expanded leaves, approximately 33 days after sowing. The growth-related manually mea-

203 sured traits that were evaluated were plant height (PH) and shoot dry mass (SDM). Plant

204 height was measured from the soil to the last expanded leaf's ligule, and SDM was obtained

205 from the dry mass of the leaves and stalk.

206 The image-derived phenotypes were collected over time to capture plant growth, as pre-

207 viously described. For each replication, measurements were made at six time points defined

208 by the number of expanded leaves: 0 (before germination), 2, 3, 4, 5, and 6 (Hanway, 1966).

209 Since the genotypes presented expected inconsistencies in growth stages, it was determined

210 as the mode of the population at a given time. A time point before the germination step

211 was used to obtain the PH_{HTP}. Heat accumulation was calculated from the growing degree

212 days (GDD) based on the formula: $GDD = \sum_{i=1}^{m} T_i - T_{base}$, where T_i is the daily mean air

9

213   temperature and T_{base} is the base temperature of 10°C. Mean air temperature was calculated

214   using the following formula: $T_i = \frac{T_{max} - T_{min}}{2}$, where T_{max} and T_{min} are the maximum and

215   minimum temperatures, respectively, of day i (Gilmore and Rogers, 1958). The R package

216   pollen (Nowosad, 2019) was used to calculate GDD. Phenotypic correlations were estimated

217   using Pearson correlations between the image-derived phenotypes (PH$_{\text{HTP}}$, CC, f(biomass),

218   and CV) and manually measured phenotypes (PH and SDM).

Likelihood-ratio and Wald tests

219

The following model was used to test the effects of genotype, management (B+ and B-), and their interaction.

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_1\mathbf{r} + \mathbf{X}_2\mathbf{b} + \mathbf{X}_3\mathbf{m} + \mathbf{Z}_1\mathbf{g} + \mathbf{Z}_2\mathbf{gm} + \boldsymbol{\epsilon}$$

220   where $\mathbf{y}$ is the vector of phenotypes; $\mathbf{1}$ is the vector of ones; $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$ are the incidence

221   matrices for the fixed effects; $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are the incidence matrices for the random effects; μ

222   is the overall mean; $\mathbf{r}$, $\mathbf{b}$, and $\mathbf{m}$ are the fixed effects for replication, block within replication,

223   and management (B+ and B-), respectively; $\mathbf{g} \sim \text{N}(0, \mathbf{G}\sigma_g^2)$ is the vector of random effect

224   of genotype; $\mathbf{gm} \sim \text{N}(0, \mathbf{G}\otimes\mathbf{I}\sigma_{gm}^2)$ is the vector of random effects of the interaction between

225   genotype and management; and $\boldsymbol{\epsilon} \sim \text{N}(0, \mathbf{I}\sigma_\epsilon^2)$ is the random residual effect. Here $\mathbf{G}$ is the

226   additive genomic relationship matrix (VanRaden, 2008); $\mathbf{I}$ is the identity matrix; σ_g^2 is the

227   additive genomic variance; σ_{gm}^2 is the genotype-by-management interaction variance; and

228   σ_ϵ^2 is the residual variance. The significance of random and fixed effects was assessed using

229   the Wald and likelihood-ratio tests, respectively. The analysis was performed using the R

230   package ASReml-R (Butler et al., 2017).

10

Bayesian genomic best linear unbiased prediction

Univariate and bivariate Bayesian genomic best linear unbiased prediction (GBLUP) models were used to estimate genomic heritability and genomic correlation separately for B+ and B-. These Bayesian models were the same as those used for the Wald and likelihood-ratio tests, but the management ($\mathbf{m}$) and genotype-by-management interaction terms ($\mathbf{gm}$) were dropped. For the univariate model, a flat prior was assigned to $\mathbf{r}$ and $\mathbf{b}$. The variance components, σ_g^2 and σ_ϵ^2, were drawn from a scaled inverse χ^2 distribution. For the bivariate model, $\mathbf{y}$ is the vector of phenotypes of two responses; $\mathbf{g} \sim \mathrm{N}\left(0, \Sigma_g \otimes \mathbf{G}\right)$ is the vector of genotypes; $\boldsymbol{\epsilon} \sim \mathrm{N}\left(0, \Sigma_\epsilon \otimes \mathbf{I}\right)$ is the residual; $\otimes$ is the Kronecker product; and Σ_g and Σ_ϵ are the variance-covariance matrices for additive genomic and residual effects taking the forms of

$$\Sigma_g = \begin{bmatrix} \sigma_{g_1}^2 & \sigma_{g_{12}} \\ \sigma_{g_{21}} & \sigma_{g_2}^2 \end{bmatrix}, \ \ \Sigma_\epsilon = \begin{bmatrix} \sigma_{\epsilon_1}^2 & \sigma_{\epsilon_{12}} \\ \sigma_{\epsilon_{21}} & \sigma_{\epsilon_2}^2 \end{bmatrix},$$

where subscripts 1 and 2 refer to the first and second phenotypes. An inverse Wishart distribution was assigned to Σ_g and Σ_ϵ with degrees of freedom $\nu = 4$ and scale matrix S such that the prior means of Σ_g and Σ_ϵ equal half of the phenotypic variance. All the Bayesian GBLUP models were fitted using 60,000 Markov chain Monte Carlo samples, 10,000 burn-in, and a thinning rate of 60 implemented in JWAS software (Cheng et al., 2018a,b). Model convergence was assessed using trace plots of the posterior distributions of the variance components.

11

Heritability and genomic correlation

The variance components obtained from the univariate Bayesian GBLUP were used to esti-
mate genomic heritability using the following formula.

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{n_r}},$$

where n_r is the number of replications (2). The estimates of genomic correlation were ob-
tained from the estimated variance-covariance matrix in the bivariate Bayesian GBLUP
model.

12

Results

Image processing and data extraction

A total of 756 plots in each replication across time were evaluated during plant development. Each collection of images took approximately 10 min. The ground resolution of the orthomosaics was approximately 2.30 mm pix^{-1}, and the GCP error was approximately 4 cm (Table 1). Despite the difference between days after sowing, accumulated GDD were similar between replication 1 and replication 2. Additionally, the ground resolution of the orthomosaic values was consistent, and GCP errors were low across different numbers of leaves.

Plant growth-promoting bacteria experiment

Statistical test and phenotypic correlation

The management and genotype effects were statistically significant for all image-derived phenotypes across the different stages of maize development (Supplementary Tables S1-S4). This suggests that the presence of PGPB and genetic diversity significantly affect plant development and growth. However, the genotype-by-management interaction was not statistically significant. Similarly, the main effects of management and genotype were consistently significant, but the genotype-by-management interaction was absent for manually measured PH and SDM (Supplementary Table S5). Figure 3 shows the growth patterns of the image-derived phenotypes with (B+) or without (B-) PGPB inoculation. The B+ management produced higher mean values than the B- management for all image-derived phenotypes, suggesting that PGPB inoculation promotes plant growth as expected. Moderate phenotypic correlations were observed between the HTP and manually measured phenotypes (Table 2). Phenotypic correlations between PH$_{\mathrm{HTP}}$ and PH ranged from 0.23 to 0.64 (B+) and 0.36 to

0.57 (B-). Image-derived phenotypes CC, f(biomass), and CV were equally correlated with SDM. The later growth stages tended to show higher phenotypic correlations (4, 5, and 6 leaves). Overall, B+ and B- showed a similar pattern of phenotypic correlations.

Genomic heritability

Estimates of genomic heritability varied across image-derived phenotypes and stages of maize development (Table 3). Earlier developmental stages tended to show higher estimates of heritability. Among image-derived phenotypes, PH_{HTP} showed the highest estimates of genomic heritability ranging from 0.35 to 0.54 (B+) and 0.34 to 0.48 (B-). In contrast, CV showed the lowest genomic heritability estimates, particularly when the number of leaves was five. The genomic heritability estimates of manually measured PH were 0.61 (B+) and 0.57 (B-), while those of SDM were 0.30 (B+) and 0.28 (B-). Overall, the difference in the habitability estimates between B+ and B- was small.

Genomic correlation

The genomic correlations between HTP and manually measured phenotypes showed a similar tendency to those of phenotypic correlations (Table 2). High genomic correlations were observed between PH_{HTP} and PH in the later stages of maize development for both B+ and B-. The image-derived f(biomass) showed the strongest genomic correlations with SDM, followed by CC. No differences in genomic correlations were observed between B+ and B-. Additionally, moderate-to-strong phenotypic and genomic correlations were observed across the developmental stages for each image-derived phenotype (Figure 4). As expected, the number of adjacent leaves showed higher correlations.

14

Discussion

A greenhouse HTP platform was developed to evaluate the influence of PGPB on plant growth using image-derived phenotypes. HTP platforms play an important role in plant breeding programs because they allow the evaluation of plant growth and development in a non-destructive, time-efficient, and less laborious manner. The image-capture system and processing were designed to be similar to those of UAVs. The roof structure of the greenhouse was used to attach the tracks to save costs and enable easy installation without restructuring the greenhouse itself. The total cost to develop our greenhouse-based HTP system was approximately US$5,000. Our expenses were higher than those of a recently developed HTP system for soybean (Zhou et al., 2018). However, the size of the HTP platform developed in this study is larger and can accommodate more genotypes. In terms of cost per m^2, our HTP platform is relatively cost-efficient because the cost associated with our HTP system was $75 per m^2, whereas that of Zhou et al. (2018) was $40 per m^2. The image overlap during the capture was controlled by the opening angle of the camera, speed of the track, and the y-axis distance, so that different cameras can be easily utilized by adjusting these factors. The coordinate system used for GCP was a universal transverse mercator obtained from RTK GPS, which may not always work indoors because of the greenhouse roof. An alternative option is to use a local coordinate system.

The image analysis pipeline consisted of aligning the images, obtaining dense cloud points, and mosaicking. The most laborious steps were to manually identify each plot and adjust its shapefile to avoid overlapping plots. Several approaches have been proposed to automate the plot identification step, such as the fieldShape function in the FIELDimageR R package (Matias et al., 2020) or a negative buffer area (Galli et al., 2020). However, these methods did not produce adequate results in our case, probably because of leaf overlapping (Ahmed et al., 2019). An alternative emerging approach is to implement semantic segmentation and

15

312 object detection based on deep learning (Xie et al., 2017; Zou et al., 2020).

313 The effects of genotype and management were significant and consistent between the
314 image-derived and manually measured phenotypes. This suggests that image-derived phe-
315 notypes can be used to assess the differences within genotypes or managements. Additionally,
316 the image-derived phenotypes were capable of capturing plant growth at different stages of
317 plant development. The HTP genomic heritability estimates tended to be lower than those
318 of manually measured phenotypes and decreased as the plants developed. This was likely
319 due to the difficulty in accurately phenotyping taller plants. The magnitude of the genomic
320 correlations and genomic heritabilities were similar between management groups B+ and
321 B-. This was expected because the genotype-by-management interaction term was not sig-
322 nificant. Our HTP platform was able to consistently capture genetic variability within each
323 management.

324 No significant interaction between genotype and management for both HTP and manu-
325 ally measured phenotypes may also indicate the absence of phenotypic plasticity for PGPB
326 responses in our population. Our findings agree with those of Vidotti et al. (2019a) and
327 Vidotti et al. (2019b), who did not find significant genotype and management interactions in
328 hybrid maize using different genotypes and PGPB from this study. This might be because
329 both managements were tested under nitrogen-limited conditions or the experiment only
330 covered the early developmental stages. For example, Guo et al. (2020) reported that low
331 nutrients in optimal irrigated growth conditions might contribute to the absence of genotype-
332 by-water availability interaction in wheat. On the other hand, the significant management
333 effect suggests that PGPB can promote plant growth. Nevertheless, further studies are
334 needed to vary nitrogen levels, assess PGPB responses at the later stages of development,
335 and validate our results in field trials.

336 Moderate-to-strong phenotypic and genomic correlations between PH_{HTP} and PH re-
337 vealed that image-derived PH_{HTP} can be a good predictor for manually measured PH. Simi-

16

338 larly, a moderate genomic correlation between f(biomass) and SDM suggests that f(biomass)
339 can be used as a secondary or correlated phenotype for SDM in genomic predictions (Rutkoski
340 et al., 2016). We also investigated the utility of spectral indices (e.g., NDVI) as a proxy for
341 SDM. However, the phenotypic correlation between NDVI and SDM was low. A potential
342 reason for this might be the difficulty in accurately calibrating images using a calibrated re-
343 flectance panel or a sunshine (light) sensor. The reduction of sunlight inside the greenhouse
344 due to the polyethylene roof may have limited the calibration accuracy. Unlike Li et al.
345 (2020), this was the main reason why we did not include NDVI to calculate f(biomass).

346 The architecture of maize plants makes image-derived phenotyping harder because stalks
347 and leaves grow beyond their pots and interfere with neighboring pots. This can be min-
348 imized by increasing the distance between the pots and distributing them equidistantly if
349 a larger greenhouse is available. Another limiting factor that may reduce the correlation
350 between PH_{HTP} and PH is related to plant morphology. For instance, during maize growth,
351 the leaf development stage directly impacts plant height projection. Alternatively, we can
352 measure PH_{HTP} at the leaf ligule of the last fully expanded leaf. However, locating the leaf
353 ligule in the HTP platform is a challenging task because PH_{HTP} is based on plant height
354 projection (Figure S1).

355 There are several greenhouse-based HTP platforms available that differ in terms of pre-
356 cision, resolution, and applications (Li et al., 2021). The advantage of our HTP platform
357 is its low cost compared to commercial platforms, while having the capacity to phenotype
358 many lines. Despite the fact that our image-derived phenotypes were slightly less correlated
359 with manually measured phenotypes than other related studies found in the literature, our
360 results confirm that image-derived phenotypes can provide valuable information for captur-
361 ing temporal PGPB responses in maize. Further research is warranted to evaluate the utility
362 of image-derived phenotypes to study PGPB responses in longitudinal genomic predictions
363 and genome-wide association studies (Campbell et al., 2019; Baba et al., 2020).

17

Conclusions

We developed a low-cost high-throughput phenotyping platform capable of capturing plant growth across developmental stages. This platform was used to study the symbiosis between PGPB and maize. We found a moderate-to-strong phenotypic and genomic correlation between the image-derived and manually measured phenotypes, where PGPB promoted growth in the population. The findings reported in this study will help small plant breeding programs or public research institutions to integrate phenomics, genetic, and management studies under a limited budget.

Acknowledgments

Author contributions

Rafael Massahiro Yassue: Conceptualization; Data curation, Formal analysis, Investigation; Methodology; Visualization; Writing-original draft; Writing-review & editing. Giovanni Galli: Investigation; Methodology; Writing-review & editing. Ronaldo Borsato Junior: Investigation; Writing-review & editing. Hao Cheng: Software, Writing-review & editing. Gota Morota: Conceptualization; Methodology; Supervision; Writing-original draft; Writing-review & editing. Roberto Fritsche-Neto: Conceptualization; Funding acquisition; Supervision; Writing-review & editing.

Conflict of interest

None declared.

References

Ahmed, I., Eramian, M., Ovsyannikov, I., van der Kamp, W., Nielsen, K., Duddu, H. S., Rumali, A., Shirtliffe, S., and Bett, K. (2019). Automatic detection and segmentation of lentil crop breeding plots from multi-spectral images captured by UAV-mounted camera. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE.

Araus, J. L. and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*, 19(1):52–61.

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., and Cairns, J. E. (2018). Translating high-throughput phenotyping into genetic gain. *Trends in Plant Science*, 23(5):451–466.

Arif, I., Batool, M., and Schenk, P. M. (2020). Plant microbiome engineering: Expected benefits for improved crop growth and resilience. *Trends in Biotechnology*, 38(12):1385–1396.

Baba, T., Momen, M., Campbell, M. T., Walia, H., and Morota, G. (2020). Multi-trait random regression models increase genomic prediction accuracy for a temporal physiological trait derived from high-throughput phenotyping. *PloS One*, 15(2):e0228118.

Batista, B. D., Dourado, M. N., Figueredo, E. F., Hortencio, R. O., Marques, J. P. R., Piotto, F. A., Bonatelli, M. L., Settles, M. L., Azevedo, J. L., and Quecine, M. C. (2021). The auxin-producing bacillus thuringiensis RZ2ms9 promotes the growth and modifies the root architecture of tomato (solanum lycopersicum cv. micro-tom). *Archives of Microbiology*.

Batista, B. D., Lacava, P. T., Ferrari, A., Teixeira-Silva, N. S., Bonatelli, M. L., Tsui, S., Mondin, M., Kitajima, E. W., Pereira, J. O., Azevedo, J. L., and Quecine, M. C. (2018).

Screening of tropically derived, multi-trait plant growth- promoting rhizobacteria and evaluation of corn and soybean colonization ability. *Microbiological Research*, 206:33–42.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23(19):2633–2635.

Brichet, N., Fournier, C., Turc, O., Strauss, O., Artzet, S., Pradal, C., Welcker, C., Tardieu, F., and Cabrera-Bosquet, L. (2017). A robot-assisted imaging pipeline for tracking the growths of maize ear and silks in a high-throughput phenotyping platform. *Plant Methods*, 13(1).

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics*, 103(3):338–348.

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., and Thompson, R. (2017). *ASReml-R Reference Manual Version 4*.

Campbell, M., Momen, M., Walia, H., and Morota, G. (2019). Leveraging breeding values obtained from random regression models for genetic inference of longitudinal traits. *The Plant Genome*, 12(2):180075.

Campbell, Z. C., Acosta-Gamboa, L. M., Nepal, N., and Lorence, A. (2018). Engineering plants for tomorrow: how high-throughput phenotyping is contributing to the development of better crops. *Phytochemistry Reviews*, 17(6):1329–1343.

Chai, Y. N., Ge, Y., Stoerger, V., and Schachtman, D. P. (2021). High-resolution phenotyping of sorghum genotypic and phenotypic responses to low nitrogen and synthetic microbial communities. *Plant, Cell & Environment*, 44(5):1611–1626.

21

432 Cheng, H., Fernando, R., and Garrick, D. (2018a). JWAS: Julia implementation of whole-
433 genome analysis software. In *Proceedings of the world congress on genetics applied to*
434 *livestock production.*

435 Cheng, H., Kizilkaya, K., Zeng, J., Garrick, D., and Fernando, R. (2018b). Genomic pre-
436 diction from multiple-trait bayesian regression methods using mixture priors. *Genetics*,
437 209(1):89–103.

438 Compant, S., Clément, C., and Sessitsch, A. (2010). Plant growth-promoting bacteria in
439 the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and
440 prospects for utilization. *Soil Biology and Biochemistry*, 42(5):669–678.

441 Czedik-Eysenberg, A., Seitner, S., Güldener, U., Koemeda, S., Jez, J., Colombini, M., and
442 Djamei, A. (2018). The 'phenobox', a flexible, automated, open-source plant phenotyping
443 solution. *New Phytologist*, 219(2):808–823.

444 Doyle, J. and Doyle, J. (1987). A rapid dna isolation procedure for small quantities of fresh
445 leaf tissue. *PHYTOCHEMICAL BULLETIN*, 17(RESEARCH).

446 Du, J., Fan, J., Wang, C., Lu, X., Zhang, Y., Wen, W., Liao, S., Yang, X., Guo, X., and Zhao,
447 C. (2021). Greenhouse-based vegetable high-throughput phenotyping platform and trait
448 evaluation for large-scale lettuces. *Computers and Electronics in Agriculture*, 186:106193.

449 Furbank, R. T. and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping
450 bottleneck. *Trends in Plant Science*, 16(12):635–644.

451 Galli, G., Horne, D. W., Collins, S. D., Jung, J., Chang, A., Fritsche-Neto, R., and Rooney,
452 W. L. (2020). Optimization of UAS-based high-throughput phenotyping to estimate plant
453 health and grain yield in sorghum. *The Plant Phenome Journal*, 3(1).

22

Galli, G., Sabadin, F., Costa-Neto, G. M. F., and Fritsche-Neto, R. (2021). A novel way to validate UAS-based high-throughput phenotyping protocols using in silico experiments for plant breeding purposes. *Theoretical and Applied Genetics*, 134(2):715–730.

Gilmore, E. C. and Rogers, J. S. (1958). Heat units as a method of measuring maturity in corn 1. *Agronomy Journal*, 50(10):611–615.

Guo, X., Svane, S. F., Füchtbauer, W. S., Andersen, J. R., Jensen, J., and Thorup-Kristensen, K. (2020). Genomic prediction of yield and root development in wheat under changing water availability. *Plant Methods*, 16(1).

Hanway, J. J. (1966). *How a corn plant develops.* Iowa State University.

Hungria, M., Campo, R. J., Souza, E. M., and Pedrosa, F. O. (2010). Inoculation with selected strains of azospirillum brasilense and a. lipoferum improves yields of maize and wheat in brazil. *Plant and Soil*, 331(1-2):413–425.

Kumar, A., Singh, S., Gaurav, A. K., Srivastava, S., and Verma, J. P. (2020). Plant growth-promoting bacteria: Biological tools for the mitigation of salinity stress in plants. *Frontiers in Microbiology*, 11.

Lafarge, T. and Pateiro-Lopez, B. (2020). *alphashape3d: Implementation of the 3D Alpha-Shape for the Reconstruction of 3D Sets from a Point Cloud.* R package version 1.3.1.

Li, D., Quan, C., Song, Z., Li, X., Yu, G., Li, C., and Muhammad, A. (2021). High-throughput plant phenotyping platform (HT3P) as a novel tool for estimating agronomic traits from the lab to the field. *Frontiers in Bioengineering and Biotechnology*, 8:1533.

Li, F., Piasecki, C., Millwood, R. J., Wolfe, B., Mazarei, M., and Stewart Jr, C. N. (2020). High-throughput switchgrass phenotyping and biomass modeling by uav. *Frontiers in Plant Science*, 11:1532.

477 Manoj, S. R., Karthik, C., Kadirvelu, K., Arulselvi, P. I., Shanmugasundaram, T., Bruno,
478 B., and Rajkumar, M. (2020). Understanding the molecular mechanisms for the enhanced
479 phytoremediation of heavy metals through plant growth promoting rhizobacteria: A re-
480 view. *Journal of Environmental Management*, 254:109779.

481 Matias, F. I., Caraza-Harter, M. V., and Endelman, J. B. (2020). FIELDimageR: An R
482 package to analyze orthomosaic images from agricultural field trials. *The Plant Phenome*
483 *Journal*, 3(1).

484 Mazis, A., Choudhury, S. D., Morgan, P. B., Stoerger, V., Hiller, J., Ge, Y., and Awada, T.
485 (2020). Application of high-throughput plant phenotyping for assessing biophysical traits
486 and drought response in two oak species under controlled environment. *Forest Ecology*
487 *and Management*, 465:118101.

488 Moreno, H., Rueda-Ayala, V., Ribeiro, A., Bengochea-Guevara, J., Lopez, J., Peteinatos,
489 G., Valero, C., and Andújar, D. (2020). Evaluation of vineyard cropping systems using
490 on-board RGB-depth perception. *Sensors*, 20(23):6912.

491 Nowosad, J. (2019). *pollen: Analysis of Aerobiological Data*. R package version 0.71.

492 Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012). Development of
493 high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-
494 sequencing approach. *PLoS ONE*, 7(2):e32253.

495 QGIS Development Team (2021). *QGIS Geographic Information System*. QGIS Association.

496 Quecine, M. C., Araújo, W. L., Rossetto, P. B., Ferreira, A., Tsui, S., Lacava, P. T., Mondin,
497 M., Azevedo, J. L., and Pizzirani-Kleiner, A. A. (2012). Sugarcane growth promotion by
498 the endophytic bacterium pantoea agglomerans 33.1. *Applied and Environmental Micro-*
499 *biology*, 78(21):7511–7518.

500   Rouphael, Y., Spíchal, L., Panzarová, K., Casa, R., and Colla, G. (2018). High-throughput
501   plant phenotyping for developing novel biostimulants: From lab to field or from field to
502   lab? *Frontiers in Plant Science*, 9.

503   Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., et al. (1974). Monitoring vegetation
504   systems in the great plains with erts. *NASA special publication*, 351(1974):309.

505   Roussel, J.-R. and Auty, D. (2021). *Airborne LiDAR Data Manipulation and Visualization*
506   *for Forestry Applications*. R package version 3.1.4.

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

511   Sim, S.-C., Durstewitz, G., Plieske, J., Wieseke, R., Ganal, M. W., Deynze, A. V., Hamilton,
512   J. P., Buell, C. R., Causse, M., Wijeratne, S., and Francis, D. M. (2012). Development
513   of a large SNP genotyping array and generation of high-density genetic maps in tomato.
514   *PLoS ONE*, 7(7):e40563.

515   Thomas, S., Behmann, J., Steier, A., Kraska, T., Muller, O., Rascher, U., and Mahlein,
516   A.-K. (2018). Quantitative assessment of disease severity and rating of barley cultivars
517   based on hyperspectral imaging in a non-invasive, automated phenotyping platform. *Plant*
518   *Methods*, 14(1).

519   VanRaden, P. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy*
520   *Science*, 91(11):4414–4423.

521   Vidotti, M. S., Lyra, D. H., Morosini, J. S., Granato, Í. S. C., Quecine, M. C., de Azevedo,
522   J. L., and Fritsche-Neto, R. (2019a). Additive and heterozygous (dis)advantage GWAS

25

models reveal candidate genes involved in the genotypic variation of maize hybrids to Azospirillum brasilense. *PLOS ONE*, 14(9):e0222788.

Vidotti, M. S., Matias, F. I., Alves, F. C., Pérez-Rodríguez, P., Beltran, G. A., Burgueño, J., Crossa, J., and Fritsche-Neto, R. (2019b). Maize responsiveness to Azospirillum brasilense: Insights into genetic control, heterosis and genomic prediction. *PLOS ONE*, 14(6):e0217571.

Wang, H., Qian, X., Zhang, L., Xu, S., Li, H., Xia, X., Dai, L., Xu, L., Yu, J., and Liu, X. (2018). A method of high throughput monitoring crop physiology using chlorophyll fluorescence and multispectral imaging. *Frontiers in Plant Science*, 9.

Wintermans, P. C. A., Bakker, P. A. H. M., and Pieterse, C. M. J. (2016). Natural genetic variation in Arabidopsis for responsiveness to plant growth-promoting rhizobacteria. *Plant Molecular Biology*, 90(6):623–634.

Xie, C., Wang, J., Zhang, Z., Zhou, Y., Xie, L., and Yuille, A. (2017). Adversarial examples for semantic segmentation and object detection. In *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE.

Xie, C. and Yang, C. (2020). A review on plant high-throughput phenotyping traits using uav-based sensors. *Computers and Electronics in Agriculture*, 178:105731.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28(24):3326–3328.

Zhou, J., Chen, H., Zhou, J., Fu, X., Ye, H., and Nguyen, H. T. (2018). Development of an automated phenotyping platform for quantifying soybean dynamic responses to salinity stress in greenhouse environment. *Computers and Electronics in Agriculture*, 151:319–330.

546   Zou, H., Lu, H., Li, Y., Liu, L., and Cao, Z. (2020). Maize tassels detection: a benchmark

547   of the state of the art. *Plant Methods*, 16(1).

Tables

Table 1: Replication (Rep), number of fully expanded leaves (NL), days after sowing (DAS), ground resolution of orthomosaic (GRO), ground control points (GCP) error, and accumulated growing degree days (GDD) across five evaluations during maize growth development.

Rep	NL	DAS	GRO (mm pixel^{-1})	GCP error (m)	GDD (°C)^a
1	2	11	2.32	0.04	169.6
1	3	15	2.29	0.06	229.3
1	4	18	2.27	0.05	268.5
1	5	22	2.25	0.05	320.4
1	6	27	2.25	0.04	394.8
2	2	14	2.34	0.03	172.1
2	3	21	2.35	0.03	250.1
2	4	27	2.31	0.03	310.6
2	5	30	2.31	0.03	338.8
2	6	37	2.29	0.03	410.3

^{a} The base temperature used for GDD estimation was 10°C

Table 2: Phenotypic (r_p) and genomic (r_g) correlations between high-throughput phenotyping and manually measured phenotypes across maize development with (B+) or without (B-) plant growth-promoting bacteria inoculation. PH_{HTP}: image-derived plant height; PH: manually measured plant height; CC: canopy coverage; SDM: shoot dry mass; f(biomass): plant biomass index; CV: canopy volume; and NL: number of fully expanded leaves.

NL	PH_{HTP}:PH		CC:SDM		f(biomass):SDM		CV:SDM	
	r_p	r_g	r_p	r_g	r_p	r_g	r_p	r_g
				B+				
2	0.29	0.55	0.35	0.20	0.35	0.24	0.42	0.14
3	0.23	0.59	0.45	0.30	0.39	0.39	0.62	0.29
4	0.61	0.64	0.51	0.35	0.62	0.47	0.38	0.36
5	0.64	0.67	0.47	0.30	0.64	0.43	0.62	0.32
6	0.54	0.66	0.53	0.32	0.60	0.42	0.65	0.31
				B-				
2	0.38	0.60	0.35	0.35	0.38	0.41	0.29	0.21
3	0.36	0.67	0.51	0.41	0.48	0.46	0.49	0.29
4	0.53	0.68	0.59	0.47	0.59	0.43	0.40	0.36
5	0.57	0.63	0.62	0.45	0.66	0.47	0.50	0.32
6	0.56	0.67	0.63	0.32	0.63	0.44	0.61	0.31

Table 3: Genomic heritability estimates of image-derived phenotypes across maize development with (B+) or without (B-) plant growth-promoting bacteria inoculation. PH_{HTP}: image-derived plant height; PH: manually measured plant height; CC: canopy coverage; SDM: shoot dry mass; f(biomass): plant biomass index; CV: canopy volume; and NL: number of fully expanded leaves.

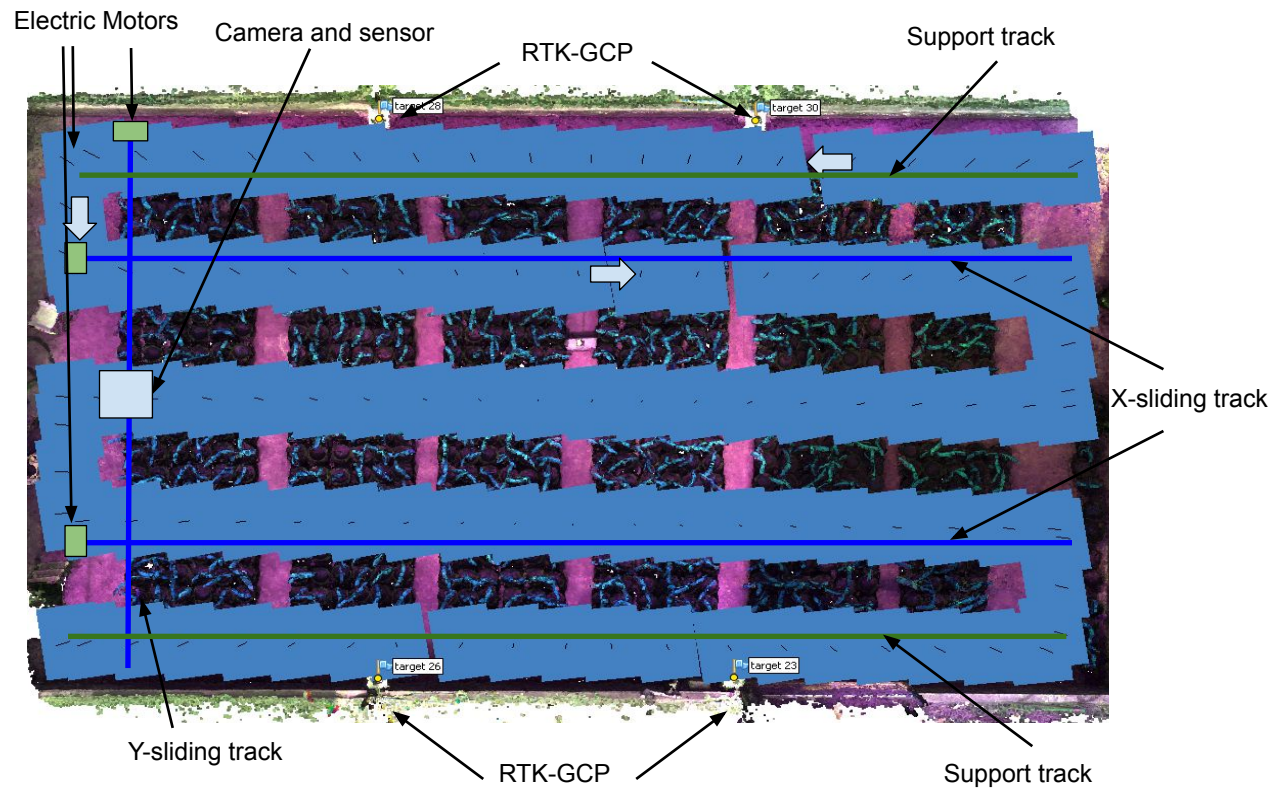| NL	PH_{HTP}		CC		f(biomass)		CV	
	B+	B-	B+	B-	B+	B-	B+	B-
2	0.54	0.48	0.46	0.43	0.46	0.45	0.31	0.29
3	0.36	0.44	0.33	0.37	0.32	0.40	0.23	0.33
4	0.43	0.36	0.35	0.36	0.36	0.34	0.23	0.22
5	0.41	0.44	0.23	0.22	0.25	0.24	0.21	0.21
6	0.35	0.34	0.23	0.23	0.24	0.25	0.27	0.25

Figures

Figure 1: Summary of the image acquisition using a low-cost high-throughput phenotyping platform for greenhouse experiments. The blues lines indicate the y and x sliding tracks. The small arrows show the direction of the camera path. Each blue square represents a photo taken by the multispectral camera.
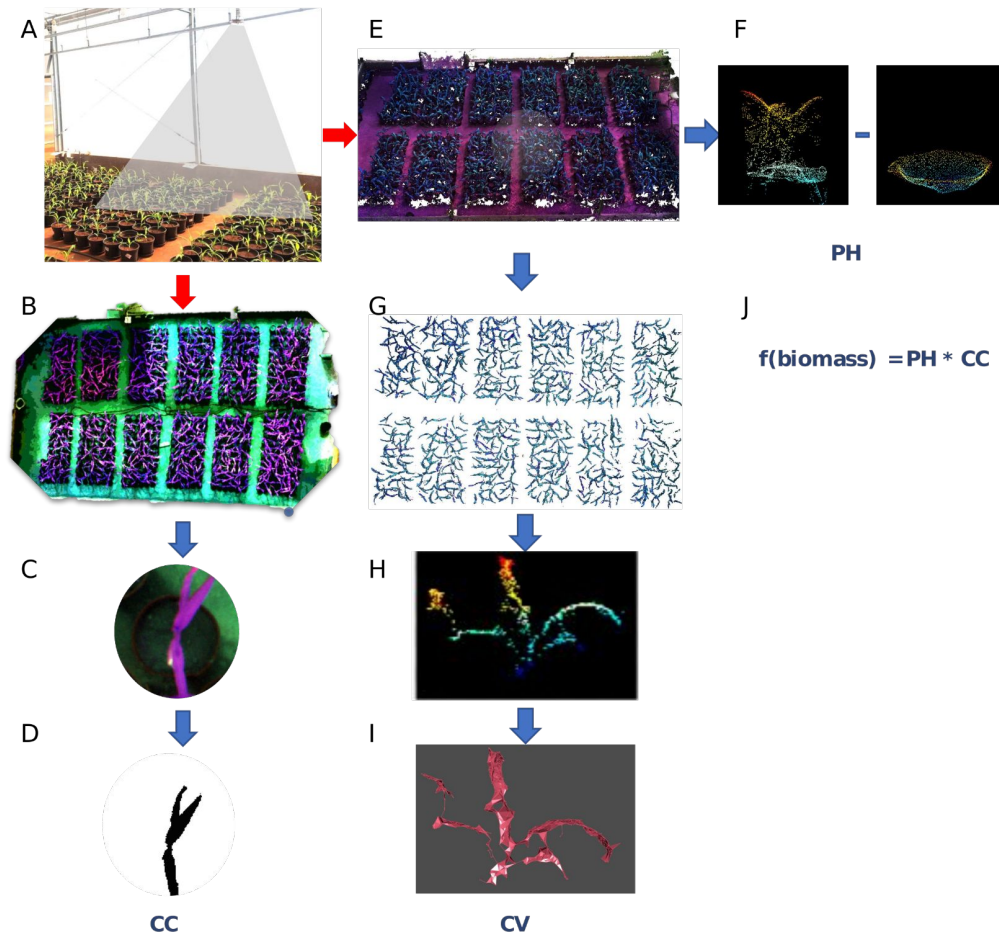
Figure 2: Summary of the multispectral image processing. (A) image acquisition; (B) mosaicking; (C) plot clip; (D) canopy coverage (CC), (E) dense cloud point; (F) plant height (PH) was calculated from the difference between the dense cloud point with the plants and the dense cloud point with only the pot; (G) dense cloud point after applying the filter to remove the background; (H) dense cloud point for each plot; (I) 3-D reconstruction of the dense cloud point to obtain canopy volume (CV); and (J) f(biomass) (plant biomass index) was obtained from PH and CC.
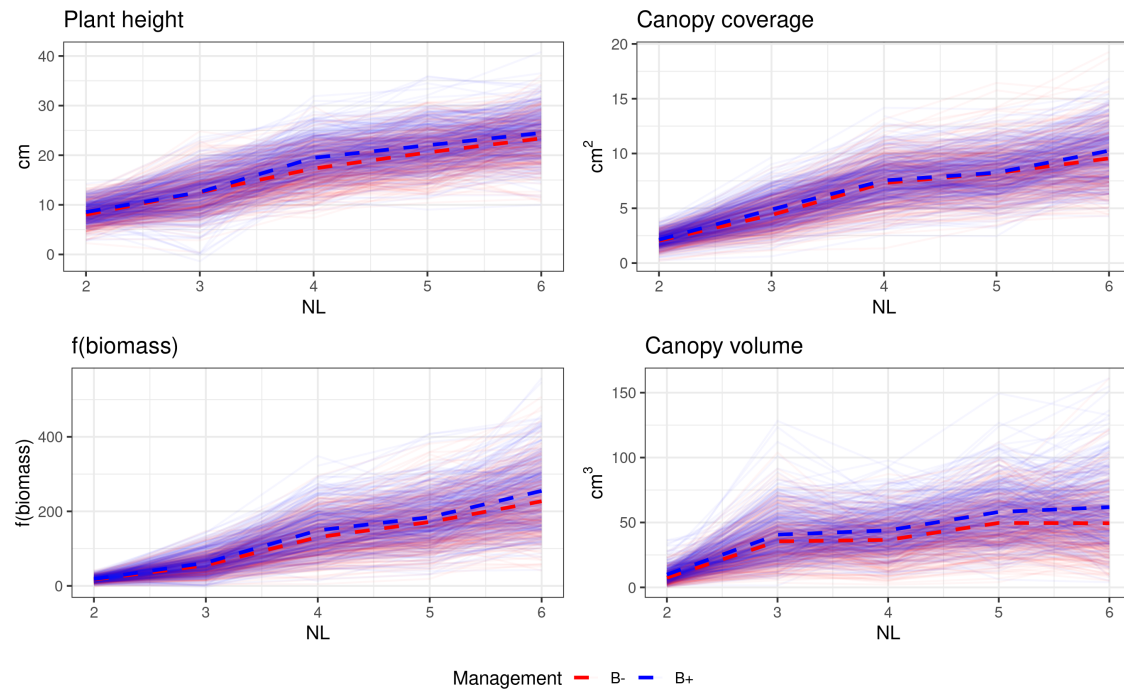
Figure 3: Growth patterns of genotypes across maize development with (B+) or without (B-) plant growth-promoting bacteria inoculation. The blue and red dashed lines represent the means of B+ and B- managements, respectively. Each thin colored line represents the mean of a genotype.
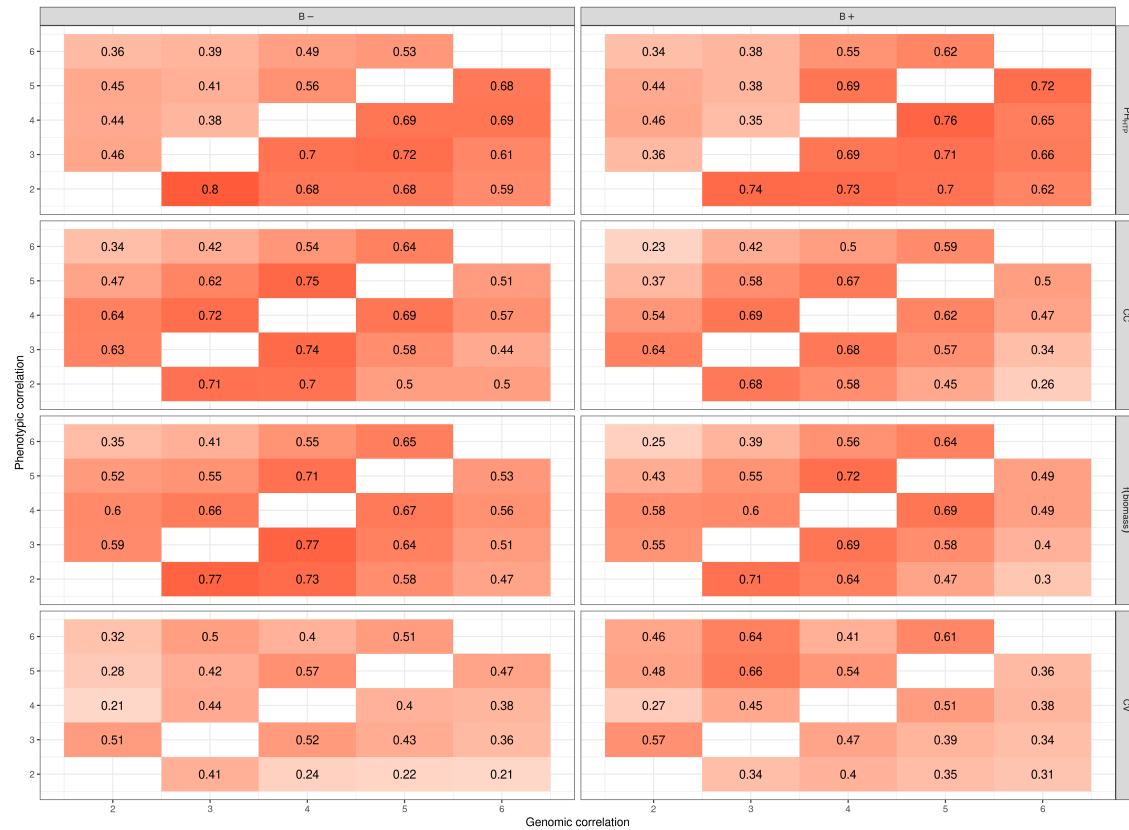
Figure 4: Graphical display of phenotypic and genomic correlations for image-derived phenotypes across maize development (number of leaves varied from 2 to 6). The upper and lower diagonal elements show phenotypic and genomic correlations, respectively. PH_{HTP}: plant height from high-throughput phenotyping; CC: canopy coverage, f(biomass): plant biomass index; and CV: canopy volume.