Unoccupied aerial systems discovered overlooked loci capturing the variation of entire growing period in maize

Alper Adak¹ | Seth C. Murray¹ | II Steven L. Anderson² | Sorin C. Popescu³ | Lonesome Malambo³ | M. Cinta Romay⁴ | Natalia de Leon⁵

¹ Dept. of Soil and Crop Sciences, Texas A&M Univ., College Station, TX 77843-2474, USA
² Dept. of Environmental Hort., Institute of Food and Agricultural Sciences, Mid-Florida Research and Education Center, University of Florida, Apopka, FL, USA
³ Dept. of Ecosystem Science and Management, Texas A&M Univ., College Station, TX 77843-2120, USA
⁴ Institute for Genomic Diversity, Cornell University, Ithaca, NY, USA
⁵ Department of Agronomy, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706, USA

Abstract

Traditional phenotyping methods, coupled with genetic mapping in segregating populations, have identified loci governing complex traits in many crops. Unoccupied aerial systems (UAS)-based phenotyping has helped to reveal a more novel and dynamic relationship between time-specific associated loci with complex traits previously unable to be evaluated. Over 1,500 maize (Zea mays L.) hybrid row plots containing 280 different replicated maize hybrids from the Genomes to Fields (G2F) project were evaluated agronomically and using UAS in 2017. Weekly UAS flights captured variation in plant heights during the growing season under three different management conditions each year: optimal planting with irrigation (G2FI), optimal dryland planting without irrigation (G2FD), and a stressed late planting (G2LA). Plant height of different flights were ranked based on importance for yield using a random forest (RF) algorithm. Plant heights captured by early flights in G2FI trials had higher importance (based on Gini scores) for predicting maize grain yield (GY) but also higher accuracies in genomic predictions which fluctuated for G2FD (−0.06~0.73), G2FI (0.33~0.76), and G2LA (0.26~0.78) trials. A genome-wide association analysis discovered 52 significant single nucleotide polymorphisms (SNPs), seven were found consistently in more than one flights or trial; 45 were flight or trial specific. Total cumulative marker effects for each chromosome’s contributions to plant height also changed depending on flight. Using UAS phenotyping, this study showed that many candidate genes putatively play a role in the regulation of plant architecture even in relatively early stages of maize growth and development.

Abbreviations: BLUP, best linear unbiased prediction; DAP, days after planting; DTA, days to anthesis; DTS, days to silking; GxE, genotype × environment; G2F, Genomes to Fields; G2FD, optimal dryland planting without irradiation; G2FI, optimal planting with irrigation; G2LA, stressed late planting; GAPIT, genome association and prediction integrated tool; GBS, genotyping-by-sequencing; GDD, growing degree days; GY, grain yield; HTP, high-throughput phenotyping; LD, linkage disequilibrium; LM, linear model; QTL, quantitative trait loci; RF, random forest; rGPA, genomic prediction accuracies; SNP, single nucleotide polymorphism; UAS, unoccupied aerial systems; UAS GEBVs, genetically estimated breeding value for each flight; UAS PEBVs, phenotypically estimated breeding value for each flight.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. The Plant Genome published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

Plant Genome. 2021;14:e20102. https://doi.org/10.1002/tpg2.20102
1 | INTRODUCTION

Quantitative variation of complex traits in maize (Zea mays L.) have been challenging to dissect since they show strong environmental interactions and are generally inconsistent between populations and different screening environments (Beavis et al., 1991; Koester et al., 1993; Peiffer et al., 2014; Sari-Gorla et al., 1999; Veldboom & Lee, 1996; Wang et al., 2006). Plant height, traditionally measured terminally at the end of the growing season with a ruler, is a prime example of a quantitative, complex trait; it is relatively easy to measure across many plots and it has high repeatability and heritability (Anderson et al., 2018, 2019, 2020; Mahan et al., 2018; Peiffer et al., 2014; Rood & Major, 1981; Veldboom & Lee, 1996). Genetic mapping and theory suggest an omnigenic model supported by the genetically polygenic inheritances observed and the variable contributions of pedigree as a source of variation, consistent with a large number of loci with minor effects governing these traits (Boyle et al., 2017; Mackay, 2001; Peiffer et al., 2014; Wallace et al., 2016; Wang et al., 2006). Recently, emerging high-throughput phenotyping (HTP) tools have demonstrated that complicated phenotypically dynamic growth patterns are occurring (Anderson et al., 2020; Miao et al., 2020; Pauli et al., 2016; Wang et al., 2019; Wu & Lin, 2006) and that the importance of various loci change temporally throughout plant growth and development (Furbank & Tester, 2011).

Using ruler-based terminal plant height, a limited number of genes, well-known to play a role in hormone synthesis and signaling pathways in maize, have been found. These larger-effect plant height genes are involved in processes such as regulating gibberellin signaling and biosynthesis (Bensen et al., 1995; Lawit et al., 2010; Winkler & Helentjaris, 1995), hindering of polar movement of auxin transport (Multan et al., 2003), and brassinosteroid synthesis (Hartwig et al., 2011; Makarevitch et al., 2012). Disruption of such hormone synthesis and signaling genes can cause significant reductions in plant heights. Although application to breeding have remained limited because of an antagonistic effect on yield (Bensen et al., 1995; Thornsberry et al., 2001; Winkler & Helentjaris, 1995), genetic variants at large-effect hormone loci may have been fixed or found to not segregate widely as a consequence of directional selection over time (Peiffer et al., 2014). As a phenotype, terminal plant height can also be a strong predictor of yield in some environments even in elite commercial hybrids (Farfan et al., 2013). It has been observed that early plant height measures (i.e., in seedling, jointing, and flowering growth stages; V0–VT) may provide novel insights into maize yield (Machado et al., 2002) and have the potential to predict yield at earlier time points (Anderson et al., 2019; Miao et al., 2020; Yin et al., 2011a,b; Zhang et al., 2017). It is likely that many quantitative trait loci (QTL) studies have been limited in explaining erratic plant height variation because they relied on end-of-season terminal height phenotyping data that did not monitor plant architecture and environmental stresses throughout plant growth (Su et al., 2019; Wang et al., 2010). Partitioning the genetic component from the total variation at each growth stage in a timely and repeated manner might further help to explain the hidden heritability issue in genetic dissection (Gibson, 2010).

Complex traits are orchestrated by the interplay of many genes, some constitutively expressed during all growth periods, while others are expressed at specific time periods (Feldman et al., 2017; Tessmer et al., 2013; Bac-Molenaar et al., 2015; Schmid et al., 2005; Sun & Wu, 2015; Li & Sillanpää, 2015). High-throughput phenotyping can measure the physical characteristics of plants in a temporal manner and has previously been used to discover temporal QTL under field conditions for rice (Oryza sativa L.) (Tanger et al., 2017; Yang et al., 2014), triticale (×Triticosecale spp.) (Würschum et al., 2014), cotton (Gossypium spp.) (Pauli et al., 2016), barley (Hordeum vulgare L.) (Neumann et al., 2017), and wheat (Triticum aestivum L.) (Lyra et al., 2020). In order to capture temporal gene–trait associations in maize, two studies have used automated greenhouse-based HTP platforms (Junker et al., 2015) under environmentally controlled conditions (Zhang et al., 2017; Muraya et al., 2017), while two others used unoccupied aerials system (UAS)-based HTP platforms under field conditions (Anderson et al., 2020; Wang et al., 2019) but only for inbred lines and not hybrids, which are what farmers grow.

Repeated data collection of maize plant height by UAS has already shown dynamic correlations with yield at different time points in a segregating population of hybrid lines (Anderson et al., 2019). Within the scope of the Genomes to Fields (G2F) project, UAS HTP data captured the temporal variations of growth patterns in diverse hybrids as well as other agronomic traits (i.e., flowering times and yield) under three different environmental management conditions (optimal planting with irrigation [G2FI], optimal planting...
without irrigation [G2FD], and late planting with irrigation [G2LA]) during 2017. In breeding programs, a predictive model that can quickly identify the highest yielding cultivars to advance before or without harvest yield data in larger nurseries would facilitate faster decisions, saving time and resources and possibly shortening the breeding cycle.

The objectives of this study were to (a) use temporal plant heights of the maize hybrids belonging to 22 flight dates (corresponding to 60 time points for three trials) as predictor variables in order to predict maize yield via a random forest (RF) algorithm; (b) dissect the underlying temporal QTL associated with the variation captured by each flight via a genome-wide association study; and (c) estimate and compare the temporal genomic predictions for plant height of each flight for each trial.

2 MATERIALS AND METHODS

2.1 Genetic materials, experimental conditions, and high-throughput phenotyping

The G2F project is an umbrella initiative involving collaborators from a variety of disciplines aiming to perform high-throughput genotyping and phenotyping to understand gene × gene (G×G) and genotype × environment (G×E) interactions in maize (https://www.genomes2fields.org). Under this project, a HTP platform via UAS was used to capture plant growth during various development stages over three management conditions per year in College Station, TX. In 2017, 280 hybrids were planted on 3 March under G2FI, while 230 hybrids of the same set were planted on 3 March under G2FD, and on 6 April for G2LA in College Station, Texas. To visualize the heat stress to which the G2LA has been exposed, cumulative growing degree days (GDD, °C) and photoperiod (h) were illustrated for the first 150-d period following optimal (G2LE and G2FD) and delayed (G2LA) planting times (Supplemental Figure S1). Growing degree days were calculated per day by subtracting the average of the daily maximum and minimum temperature from the base temperature (10 °C). If GDD was less than zero for any given day, it was set to zero. All hybrids in each trial were grown based on a randomized complete-block design with two replications. Days to anthesis (DTA) and silking (DTS) were recorded for each hybrid as the number of days from the time of planting until at least 50% of each hybrid showed anthesis and silking emergence in the two row plots. Agronomic field data and weather data is available for 2017 (https://doi.org/10.25739/w560-2114) (McFarland et al., 2020).

Preliminary studies of image processing and plant height extraction from HTP platform for the three trials of 2017 were reported previously by Anderson et al. (2019). First, point clouds were clipped based on the trial boundaries of G2FI, G2FD, and G2LA. Then noise points, which are located at the far bottom and top of the point clouds were illustrated by using the lateral view function in CloudCompare v2.10 (Girardeau-Montaut, 2016). These noise points can cause big fluctuations in plant height data extraction, so they were manually selected using segment tool function of CloudCompare v2.10 and removed from point cloud data. Executable functions of LASStools (Isenburg, 2015; rapidlasso, 2017) and FUSION/LDV McGaughey, 2016 software were transported in R. A custom batch script was developed with several steps to extract the plot-based plant height data. R/UASTools package was used to create the polygons (ESRI shapefile) for each two adjacent row plots in each trial using the unique plot IDs (Anderson & Murray et al., 2020); two adjacent row plots-based polygon construction was also illustrated in plots created from point cloud data in R was described as follows: (a) points of clipped point cloud data of each trial were first sorted using the function of lassort.exe of LASStools software; (b) the noisy points that were so close to the canopy structure of plants in row plots were removed using the function of lasnoise.exe of LAS-tools before constructing the digital surface model; (c) ground points were identified using the hierarchical robust interpolation algorithm (Kraus & Pfeifer, 1998) using the function of GridSurfaceCreate.exe of FUSION software; (d) maximum points from the ground points were determined using the function of lasthin.exe of LASStools software; and finally (e) a digital terrain model was generated based on the maximum points from the previous steps using the function of GridSurfaceCreate.exe of FUSION software. Next, canopy surface models generated from the digital terrain model (output of step [c]) were extracted from the digital surface model (the output of step [b]). Row plots in canopy surface models then were clipped based on the previously generated ESRI shapefile using the function of PolyClipData.exe of FUSION software. As a final step, plant heights were extracted for each clipped cloud points of each row plot as a percentile metrics using the function of CloudMetrics.exe of FUSION software. This study only used 99-percentile-based plant heights (Malambo et al., 2018). All flight dates were converted into days after planting (DAP) and specified DAP in the text, figures, and tables in this study.

The point cloud data (and all raw data) is available on Cyverse (Murray et al., 2019). In total, 21, 20, and 19 flights were used for G2FI, G2FD, and G2LA trials, respectively. For
plant height, obvious outliers (>3-m tall) were removed from the phenotypic data. Missing phenotypes (two hybrids at 36 and 39 DAP in G2LA) were imputed using missForest package in R.

2.2 Statistical models and RF application to HTP data

In order to calculate phenotypically estimated breeding values for each flight (UAS PEBVs), a full random model was fit using standard least squares (restricted maximum likelihood method) in JMP version 15 Pro (SAS Institute Inc.). Best linear unbiased predictions (BLUPs) of hybrids along with all random effects of variance components were estimated based on Equation 1:

\[ y_{grij} = \mu + \text{Genotype}_g + \text{Replation}_r + \text{Range}_i + \text{Row}_j + \epsilon_{grij} \]  

where \( y \) is the response vector of plant heights belonging to each flight time, \( \mu \) is grand mean, \( g \) is the vector of genetic effects of the hybrids \( [g ~ NID(0, \sigma^2_g)] \), \( r, i, \) and \( j \) are replication \( [r ~ NID(0, \sigma^2_r)] \), field range \( [i ~ NID(0, \sigma^2_i)] \) and field row \( [j ~ NID(0, \sigma^2_j)] \) effects, respectively, accounting for spatial effects in the randomized complete block design; \( \epsilon \) is the vector of error \( [\epsilon ~ NID(0, \sigma^2_\epsilon)] \).

Genotype, replication, range, and row are the incidence matrices of the variance components in the model.

Repeatability was estimated as follows (Equation 2):

\[ R = \frac{\sigma^2_g}{\sigma^2_g + \sigma^2_\epsilon \frac{\text{no. of reps}}{\text{no. of reps}}} \]

Pearson correlation coefficients were also calculated using UAS PEBVs, belonging to each flight of each trial separately to show the correlations among plant heights of different time points.

The UAS PEBVs belonging to each flight of each trial were used as predictors to predict continuous yield. Both linear model (LM) and random forest (RF) regressions were applied to each trial separately. The training and test populations were set as 70 and 30%, respectively. The Caret package was used to implement two regression models in R by setting the number of predictors sampled for splitting at each node (mtry) and number of trees grown (ntree) based on lowest RMSE. To illustrate the importance of the UAS PEBVs belonging to each flight, varImp function was used after RF model, where higher varImp scores indicate higher importance of the explanatory variable (flight dates) in prediction the yield. To show the relationships between each predictor (flight dates) and outcome (yield) in the RF model, partial dependence plots were generated by using the pdp package in R (Friedman, 2001).

To assess prediction skill of the LM and RF models, prediction function of Caret caret::predict() was used to predict yield of test data set using 1,000 bootstraps. Correlations between actual yield of test data and predicted yield of test data were calculated for LM (\( r_{\text{lm}} \)) and RF (\( r_{\text{rf}} \)) models at each bootstrap. The Wilcoxon signed rank test was then applied to each trial to compare correlation results of the LM and RF models.

2.3 SNP discovery and association mapping

More than 1,500 inbred lines with a total of 955,690 SNPs (GBS v2.7) were produced using Illumina Hi-seq 2000/2500 at the Institute for Genomic Diversity, Cornell University, Ithaca, NY, USA (Glaubitz et al., 2014). The imputed ZeaGBS v2.7 data is also available through CyVerse with AGPv4 physical coordinates (McFarland et al., 2020; www.panzea.org; Elshire et al., 2011). Genotyping-by-sequencing (GBS) data of 158, 118, and 118 hybrids from 2017 was synthesized from sequence information of the available parental inbreds for the G2FI, G2FD, and G2LA trials, respectively. Tassel 5 software was used to call polymorphic markers (Bradbury et al., 2007). First, markers were set as missing if the either or both inbreds of hybrids used in this study were heterozygous. Second, sequence information of each hybrid was generated using the Create_Hybrid_Genotype function in Tassel 5. Third, markers with missing data (>10%) and minor allele frequency of <1% were filtered out. Missing markers were imputed using SNP.impute=major code in GAPIT (Genome Association and Prediction Integrated Tool) package in R (Lipka et al., 2012). Finally, 153,252 markers remained.

Plant heights of hybrids, using the BLUPs in Equation 1, captured by each flight were associated with filtered GBS data in the three trials individually, implementing the fixed and random model circulating probability unification (Liu et al., 2016) with principal components in GAPIT (Lipka et al., 2012). In GAPIT, the Model.selection=TRUE code was used to determine the optimal number of principal components based on Bayesian information criterion. The first three principal components were used in genome-wide association study (Supplemental Figure S2). To calculate the explained variation by the SNPs with a higher score than conservative
Bonferroni correction \([-\log_{10}(p) > 6.5; 0.01/(\text{no. of markers})]\), a Random.model=TRUE code was used in GAPIT function. In addition, a false-positive discovery rate was set \([-\log_{10}(p) > 5]\) to determine the same SNPs (if any), which were discovered with Bonferroni correction in one flight, were important in any other different flight. MaizeGBD (http://www.maizegdb.org/) was used to look up SNPs positions for determining the candidate genes in those regions using AGPV4 physical coordinates of B73. The Gramene database (http://www.gramene.org) was used to determine the functions of candidate genes. Linkage disequilibrium (LD) was examined for colocalized SNPs that were discovered in more than one flight or trial using \(R^2(>0.8)\) in LD heatmap package in R (Shin et al., 2006).

### 2.4 Genomic prediction of flights

Genome-wide prediction was applied to each flight of each trial to calculate genetic estimated breeding values (UAS\(_{\text{GBEVs}}\)) using ridge regression best linear unbiased prediction package in R (Endelman, 2011). Training and test data sets were arranged at 70 and 30%, respectively. Equation 3 was used to estimate GEBVs:

\[
y = \mu + Z\beta + \epsilon
\]  

(3)

where \(y\) is the vector of observations as BLUPs of plant heights of flights, \(\mu\) is the overall mean, \(Z\) is the marker matrix, \(\beta\) is the marker effects matrix, and \(\epsilon\) is the residual effects vector with the assumptions of \(\beta \sim NID(0, \sigma_\beta^2)\) and \(\epsilon \sim NID(0, \sigma_\epsilon^2)\).

Prediction accuracies for genomic predictions (\(r_{\text{gpa}}\)) were obtained for each flight using correlations between \(\text{UAS}_{\text{PEBVs}}\) and \(\text{UAS}_{\text{GBEVs}}\) of test data set, which were calculated by using cross-validation with 500 iterations.

### 3 RESULTS AND DISCUSSION

#### 3.1 Variance components, repeatability, and breeding value estimations of hybrids using HTP data

Using plot-based plant height extraction from the 2017 G2F project in College Station, TX, BLUPs of plant heights belonging to maize hybrids for each flight date (UAS\(_{\text{PEBVs}}\)) were estimated using Equation 1. The \(\text{UAS}_{\text{PEBVs}}\) were categorized into two groups: low and high yielding hybrids using the average yield threshold of each trial. As expected from previous work, taller hybrids, especially at flights after flowering, were relatively higher yielding than hybrids that were shorter (Figure 1).

Late plantings (G2LA) being taller is consistent with our normal observations in our breeding program over the years. Our best reason for this result is that hybrids were exposed to longer day length with higher heat accumulation (as illustrated in Supplemental Figure S1) when planted late (e.g., G2LA in this research). We do not know the entire mechanism but speculate it has to do with a combination of photoperiod (our plantings in Texas have been the only G2F location where plants flower as the days are getting longer, the later planting flowers closer to the longest day of the year) and with temperature (more rapid accumulations of GDD). These environmental conditions caused by late planting may cause the maize hybrids to produce more photosynthetic activity resulting in taller plant height and biomass but less yield.

The \(\text{UAS}_{\text{PEBVs}}\) variation explained by pedigree fluctuated between 5 and 80% of total variation explained by the model. Repeatability ranged from 12 to 93% across flights and trials (Figure 2). The hybrids with continuous taller plant heights across flights resulted in higher yield (Figure 1) even though repeatability results were not consistently estimated for plant heights through the flights. In addition, correlation coefficients among the plant heights (\(\text{UAS}_{\text{PEBVs}}\)) belonging to later flights (after flowering times) had higher and more consistent correlation coefficients than the correlation coefficients of the plant heights (\(\text{UAS}_{\text{PEBVs}}\)) belonging to earlier flights (before flowering times) (Supplemental Figure S3).

It is likely that differences in repeatability for each flight were related to the image and stitching quality of each flight (Anderson et al., 2019; Malambo et al., 2018) as well as weed pressures especially in earliest flights (e.g., 35 DAS in G2FD). It is also likely that plant height at any point in time was impacted by the activity of many genes and interplays between genes in response to changing environmental conditions during various growth periods (Veldboom & Lee, 1996; Sibov et al., 2009; Anderson et al., 2020; Dijak et al., 1999; Han et al., 2018). Since UAS allowed temporal variation of plant height to be estimated here, this variation can be used in dissecting underlying genetic mechanisms such as discovering time-specific and colocalized genes in association mapping.

#### 3.2 Random forest algorithm for determining the importance ranks of flights

Plant heights captured by various flights from the three different trials were used as predictors to predict GY using the LM and RF algorithm. The RF model was consistently found to perform better than the LM model in predicting yield in each trial based on \(R^2\), RMSE, and mean absolute error (Figure 3). Prediction accuracies of the RF model (\(r_{\text{rf}}\)) were found to be greater than prediction accuracies of LM model (\(r_{\text{lm}}\)) in each trial (Figure 3), indicating that the RF model works slightly
better in predicting GY using temporal plant heights. Variable importance scores (varImp) were used to determine variable (flight dates) importance (Figure 3). Results of varImp scores showed that different flight dates were most important across different trials; earlier flights (e.g., 48 DAP in G2FI and 52 DAP in G2LA) and later flights (e.g., 82 and 101 DAP in G2FD and 78 and 98 DAP in G2LA) had the highest scores (Figure 3). The partial dependence plots are a functional illustration for evaluating the relationship between each explanatory variable (flight dates) and the response variable (yield); the relationships might be linear, parabolic, or something even more complex (Friedman, 2001). The partial dependence plot shows the changing predicted yield values based on the plant heights at a particular time point; illustrations of this change are critical to understand the relationships between plant heights at different flight dates and yield. Partial dependence plots of every flight date in each trial were illustrated (Figure 4). For instance, the most important variable (based on varImp scores) was found to be 101 DAP in G2FD, 48 DAP in G2FI, and 57 DAP in G2LA (Figure 3). Plant height between 2.00 and 2.05 m at 101 DAP in G2FD, plant height between 0.75 and 0.80 m at 48 DAP in G2FI, and plant height between 1.60 and 1.70 m at 57 DAP in G2LA were found to be desired plant heights that attached to maize hybrids with higher yield values (Figure 4). There were additional plant height thresholds belonging to each flight times that can be used as selection criteria for higher yielding hybrids, considering the relationships between plant height at particular time points and yield (Figure 4). The majority of plant heights belonging to the flight dates were found to have nearly linear relationships with predicted yield values except for first three flights in G2FD and first four flights in G2FI (Figure 4). However, it is noteworthy to mention that linear relationships between plant heights of majority flights in G2FI is more obvious than G2FD and G2LA, indicating that stress factors in G2FD and G2LA caused the trend of linear relationships between plant heights and higher GY to change. In addition, taller plant height in earlier flight dates (33, 36, 39, and 48 DAP in G2LA) were better indicators of higher yield in G2LA than G2FI and G2FD (Figure 4). Different contributions of plant heights belonging to different growth periods and environments in predicting GY suggests that discovering genes controlling the plant height variation at different growth periods and environments will provide supportive genetic information to manipulate yield in maize.

In general, the slightly convex shape distribution in varImp scores of flight dates in G2FD and G2FI trial (Figure 3) showed that 82–112 DAP in G2FI and 73–88 and 95–118 DAP in G2FD contributed to higher GY than other flight dates. The same flight dates varied in terms of varImp scores...
The Plant Genome

FIGURE 2  Stacked bar graph showed the explained percentage variations by each component calculated by Eq. 1 with repeatability and total variance (as number) for each trial. X axis is the flight as days after planting (DAP). Left y axis was scaled as percentage to show the repeatability values (white diamond) and percent variations by components. Right y axis was scaled as number to show the total variation explained as number (black round). DTA, DTS, and GY are the abbreviations of days to anthesis, silking (days), and plot-based grain yield (t ha$^{-1}$), respectively. Total variance of DTA, DTS, and GY in number are 5.67, 5.80, and 3.48 for G2FD; 4.22, 3.99, and 2.24 for G2FI; and 4.59, 4.32, and 1.49 for G2LA. G2FI, G2FD, and G2LA trials represent the optimal planting time with irrigation, optimal planting time without irrigation, and delayed planting time with irrigation respectively.

3.3 SNP–flight associations and functions of candidate genes

Over the three trials, the 99 percentile plant heights over 60 UAS flights, as well as GY, were genetically mapped using 153,252 SNP markers. Principal component analysis suggested that there were effectively three populations (Supplemental Figure S2). After accounting for population structure across all three trials, a total of 52 SNPs with $p$ values above the Bonferroni correction were detected and annotated; seven were discovered to be common in more than one flight date, while 45 of these were flight-time specific (Figure 5; Supplemental Table S1).

Of the seven SNPs that showed association with the trait in multiple flights and trials, three were located on chromosomes 1 and 10 and showed association in two flights, four were located on chromosomes 1 and 6 showed association across three flights (Supplemental Table S1). Additionally, two SNPs on chromosomes 1 showed association in two trials (Supplemental Table S1). Of the SNPs that showed association in two trials, none showed association in the same flight in different trials (Supplemental Table S1). In other words, the same SNPs were discovered in an earlier flight of the delayed planting trial (67 DAP in G2LA) but later flights in optimal planting (73 and 95 DAP in G2FD) (Supplemental Table S1); these were not the same calendar date flights. Discovering the common SNPs for plant heights belonging to different flights in more than one trial suggested that the same genetic cause of phenotypic variation in plant height can occur in different times under different management conditions such as delayed and optimal planting (Supplemental Table S1).
FIGURE 3 Model evaluations and variable importance scores. (a) The $R^2$ of linear (LM) and random forest (RF) model, higher is better. (b) Root mean square error (RMSE) of each model in each trial, lower is better. (c) Mean absolute error (MAE) of each model in each trial, lower is better. (d) Correlation results between predicted yield and actual yield of test data obtained by 1,000 bootstrap belonging to each model in each trial. Wilcoxon sign rank test results showed the comparison of correlations belonging to both analysis models in each trial. (e) Variable importance scores (varImp) show the variable importance scores of the predictors (flight dates) where higher varImp score indicates more importance variable in prediction the yield. G2FI, G2FD and G2LA trial represent the optimal planting time with irrigation, optimal planting time without irrigation, and delayed planting time with irrigation, respectively.
FIGURE 4  The partial dependence plots for each predictor (flight dates as days after planting; DAP) in each trial. Y axis of each plot shows the predicted continuous yield (t ha$^{-1}$), while x axes are the predictor values that are plant heights of each flight dates. Red lines show the relationships between each predictor and predicted yield values. G2FI, G2FD, and G2LA trial represent the optimal planting time with irrigation, optimal planting time without irrigation, and delayed planting time with irrigation, respectively.
**FIGURE 5** Combined Manhattan plots for plant heights of each flight using best linear unbiased predictions (BLUPs) of the mixed effects spatial model. The heat map at the bottom of each Manhattan plot shows the single nucleotide polymorphism (SNP) density (within 1 Mb window size) through the chromosomes. Scale of this heat map was given on the right side of Manhattan plot. Star shapes were used for colocalized SNPs detected in more than one trial. Triangle shapes were used for colocalized SNPs detected in more than one flight within any trial. Round shapes were used for unique SNPs. Each unique color within each trial represents the association between SNPs and plant heights of each flight; color charts are given inside the rectangles at the top of each Manhattan plots. G2FI, G2FD, and G2LA trial represent the optimal planting time with irrigation, optimal planting time without irrigation, and delayed planting time with irrigation, respectively.
FIGURE 6 Linkage disequilibrium (LD) blocks of seven colocalized single nucleotide polymorphisms (SNPs) located in chromosomes 1, 6, and 10 associated with plant height. Black lines indicated the pairwise LD regions of the SNPs with 0.8 or higher \( R^2 \) within the physical lengths. Black stars showed the locations (base pair; bp) of the above SNPs.

This results in the discovery of the same loci at different time points of growth belonging to different management conditions.

The ranges in LD observed in the regions surrounding the seven colocalized SNPs cannot exclude linked SNPs as causal (Figure 6). In chromosome 1, SNPs (76,153,074 and 76,153,056 bp) were found to be significant in flights of 73 and 95 DAP in G2FD and 67 DAP in G2LA, both were high stress trials. While these two SNPs discovered in G2FD 73 and 95 DAP flights with the highest logarithm of the odds scores (>6.55), these SNPs were also nearly significant later in 98, 112, 118, and 132 DAP flights (logarithm of the odds scores >5) in G2FD. The LD regions surrounding these two SNPs extends to 131 kbp (Chr1:76.08–76.21) (Figure 5) and contains four candidate genes in the LD interval. One of the candidate genes in LD interval is \( \text{GRMZM2G005630} \), located 26 kbp away from these SNPs. \( \text{GRMZM2G005630} \) encodes the EID1-like F-box protein 2 that regulates ABA-dependent signaling that manages seed germination, root growth, and transition to flowering time as well as the accumulation of anthocyanin under drought conditions in \( \text{Arabidopsis} \) (Koops et al., 2011). F-box genes are one of the most diverse and largest gene families in higher plants; maize has \( \sim 359 \) F-box genes (Zhang et al., 2019; Jia et al., 2013). The F-box protein family regulates plant growth and development as well as biotic and abiotic stresses in maize (Jia et al., 2013).

Two additional SNP associations were found close to each other on chromosome 1 (91,993,193 and 92,056,014 bp) and were significant at 88, 91, and 101 DAP in G2FI trials. The LD region is \( \sim 66 \) kbp (Chr:91.99–92.06) (Figure 6) and contains three candidate genes. One of the candidate genes in LD is \( \text{GRMZM5G806839} \) that controls the AP2 transcription.
factor (also known as ereb44) in maize. GRMZM5G806839 mediates response to multiple abiotic and biotic stresses (Kizis et al., 2001) such as southern corn leaf blight response (Bipolaris maydis) in maize (Bian et al., 2014).

Another SNP association on chromosome 1 (204,145,649 bp) was found at 98 and 105 DAP in G2FI. The surrounding LD region of this SNP is ~170 kbp (Chr1:203.98–204.15) and contains seven candidate genes in the interval, of which none have been previously identified. One of the candidate genes, GRMZM2G016210, contains this SNP in its genic region and includes the umc1122 simple sequence repeat marker. The flanking region between umc1035 and umc1122 markers was previously found to control plant height at different growth stages with different effect sizes in maize (Yan et al., 2003). Qph1, which is a rare SNP in Bractice2 gene, is one of the well-known QTL controlling plant height that has also been mapped near umc1122 (Xing et al., 2015). This QTL is congruent with the idea of a major QTL often described based on studies largely conducted in optimal environmental conditions of the U.S. Midwest and Europe would only be detected in Texas under the most favorably managed trial.

Two chromosome 1 SNP associations (286,738,454 and 286,738,763 bp) were found at 82, 91, and 101 DAP in G2FI. The LD block around the SNPs is ~66 kbp (Chr1:286.73–286.79) (Figure 6) and it contains three candidate genes. GRMZM2G028286 is one of the candidate genes in the LD region and it is reported to control the cellulose biosynthesis pathway (Kianifariz, 2017). In a previous study, GRMZM2G028286 was found to be downregulated when the NUT1 mutant gene is active; the NUT1 mutant gene causes an erratic tassel phenotype (e.g. tassel browning and sterility) especially under drought stress because of a defect in water transport (Dong et al., 2020) as well as reduced plant height in maize (Dong et al., 2020). Another candidate gene in this LD region is GRMZM2G028151 (ereb184-AP2-EREBP transcription factor) that is an AP2-like ethylene responsive transcription factor and it was reported to be upregulated under heat stress (Casaretto et al., 2016).

The chromosome 1 SNP association (287,796,640 bp) was found at 67 and 105 DAP in G2LA and G2FD trials. The LD region surrounding this SNP is ~253 kbp (Chr1:287.66–287.92) and contains 11 candidate genes. One of the candidate genes within the LD interval is GRMZM2G122139, and this SNP is also found in its genic region. This candidate gene encodes Cytosolic purine 5-nucleotidase. This enzyme was purified from maize microsomes (Carter & Tipton, 1985) and wheat seedling leaves and potato (Solanum tuberosum L.) (Polya, 1974, 1975), suggesting a relationship with a cyclic nucleotide regulatory system in higher plants. Another candidate gene, GRMZM2G072806, encodes the ubiquinone oxidoreductase enzyme; its activity plays a role on NAD(H) biosynthesis by transporting a hydride ion (H\(^+\)). This gene could disrupt regular enzymatic activities causing oxidative stress (Moller, 2001).

A SNP association on chromosome 6 (108,551,216 bp) was found at 88, 105, and 112 DAP in the G2FI trial. The LD region around this SNP is ~180 kbp (Chr6:108.40–108.58) (Figure 6) and contains nine genes. One candidate gene in this interval is Zm00001d036994. This candidate gene has been reported to encode trichrome birefringence-like proteins that were localized in Golgi and transfers the acetyl groups to xylan resulting in the elongation of the xylan (Xiong et al., 2013). Two trichome birefringence-like-related recessive mutants in rice were found to reduce the xylan monoacetylation in cell wall and cause the reduced plant height and susceptibility to leaf blight (Gao et al., 2017).

Chromosome 10 SNP associations (88,841,430 and 88,845,809 bp) were found at 35 DAP in the G2FD trial. The LD region around these SNPs is ~6 kbp (Chr10: 88.84–88.85) (Figure 6) and contains two candidate genes. One of the candidate genes in this interval, GRMZM2G089484 (MAP kinase), was reported to be a possible candidate associated with northern corn leaf blight resistance caused by Exserohilum turcicum (Ding et al., 2015). Another candidate gene, GRMZM2G422090 (GTP-binding protein), is related to the control of cell elongation resulting from the stimulation of the auxin hormone (Terryn et al., 1993).

Five SNPs that are in the genic region of GRMZM2G010356, GRMZM2G455869, GRMZM2G051050, GRMZM2G058522, and GRMZM2G028014 were discovered for GY in G2FD and in G2FI (Figure 7). GRMZM2G010356, also known as tps17, involves the terpene synthase biosynthesis that contributes the defense mechanisms against parasitoids and pathogens in maize (Block et al., 2019; Liang et al., 2018). GRMZM2G455869, also known as MYB37, regulates the elongation of bundle sheath cells in maize (Chang et al., 2012). GRMZM2G051050, also known as Ypt/Rab domain of the gyp1p superfamily protein, was found to be related with yield under high infestation of Mediterranean corn borer (Sesamia nonagrioides Lefebvere) in maize (Jiménez-Galindo et al., 2018). GRMZM2G058522, also known as superoxide dismutase, creates an initial defense response against reactive oxygen species that stemmed from exposure of numerous environmental stresses such as drought, intense UV lights, air pollutants, and chilling temperatures (Alscher, Erturk & Heath, 2002). GRMZM2G028014 is one of member in F-box protein family in maize that is responsible for regulation of protein degradation, signal perception and transduction pathways inside and often outside of cells regions in maize (Jia, Wu, Li, Huang & Zheng, 2013).

### 3.4 Contributions of UAS to association studies

High-throughput phenotyping by UAS captured phenotypic variation arising from the response of diverse hybrids to...
FIGURE 7  Single nucleotide polymorphisms (SNPs) associated with hybrid grain yield (t ha\(^{-1}\)) for G2FI, G2FD, and G2LA trials. G2FI, G2FD, and G2LA trial represent the optimal planting time with irrigation, optimal planting time without irrigation (dryland), and delayed planting time with irrigation respectively. The heat map at the bottom of the Manhattan plot shows the SNP density (within 1 Mb window size) through the chromosomes. Scale of this heat map was given on the right side of Manhattan plot.

FIGURE 8  Temporal cumulative marker effects of each chromosome (negative and positive) for each flight of each trial. The cumulative effects of whole-genome-wide markers (y axis, mm) were dependent upon the flights (x axis, DAP). G2FI, G2FD, and G2LA trial represent the optimal planting time with irrigation, optimal planting time without irrigation, and delayed planting time with irrigation, respectively.

Various abiotic stresses across plant development stages. When this variation was combined with genomic information in an association mapping study, SNPs were detected and linked candidate genes were identified for plant heights of different flights and GY involved in coping with many environmental stress factors (Figures 5 and 7; Supplemental Table S1). The resulting candidate gene annotations for these SNPs showed that segregating loci controlling numerous biological
functions orchestrate differences between hybrids temporally throughout plant development, yet there were substantial differences under the two abiotic stresses. Some of these loci conditioned variation across multiple time periods, while others appeared at a specific time point and under specific growing conditions. Plant height is a result of many environmental stress factors during the development periods, causing temporal modifications in plant height depending on different growth stages. The temporal variations in plant height can therefore be evaluated as a phenotypic consequence of the response of maize hybrids to environmental stresses during growth. When temporal phenotypic data collected by UAS are combined with genomic (e.g. GBS) data, phenotypic plasticity of plant height occurred within growth stages and underlying genes can be scrutinized with higher accuracy and resolution.

3.5 Cumulative SNP effects

Genomic prediction and selection approaches rely on contributions of genome-wide genetic markers. However, the phenotypes used in these studies typically have been low dimensional and low throughput obtained from limited time intervals (one or a few times) or small sample sizes. Therefore, different contributions of each locus at different time points for the same phenotype of interest will remain unobserved, with the exception of those effects that make substantial enough contribution to be detected during the terminal measurements. As we were able to only obtain genotypes of 158 hybrids of the over 270 hybrids studies, this limited these studies potential detection ability. In this study, cumulative SNP marker effects of each chromosome were found to differ depending on the flight date (Figure 8). However, like repeatability (Figure 2) the cumulative marker effects likely varied over time because of both unexplained error and different sources of phenotypic variation at different flights. For instance, there is a peak in cumulative marker effects at 64 DAP in the G2LA trial (Figure 8). The 64 DAP also had the highest repeatability value in G2LA trial (Figure 2). This is a particular example where repeatability and total cumulative marker effects were highest in G2LA trial, which may help explain that variation by pedigree at 64 DAP in G2LA trial has stemmed mostly from marker effects contributed by many regions across the genome (Figure 8). Furthermore, the different results we observed between the three management conditions demonstrated that G×E interplay, illustrated here by the interaction of hybrids and trials, had a cryptic influence on phenotypic variation (Gage et al., 2017). Plasticity is occurring even within an environment temporally as demonstrated by fluctuating marker effects, repeatability values, and explained variation by pedigree. The flights revealed this temporal within environmental variation, which was combined with G×E interaction to better scrutinize phenotypic plasticity. This study was only carried out within 1 yr, but additional...
years would likely further increase the amount of GxE interaction effects found.

### 3.6 Temporal genomic predictions

Genomic prediction accuracies for plant heights belonging to each flight in each trial were illustrated in Figure 9. Genomic prediction accuracies for DTA, DTS, and GY were also calculated for each trial (Figure 9).

Our findings showed $r_{gpa}$ for plant heights captured by flights fluctuated within G2FD (~0.06–0.73), G2FI (0.33–0.76), and G2LA (0.26–0.78) trials (Figure 9). Genomic prediction accuracies for 40–60 DAP flights (pre-flowering) were surprisingly higher than in later flights for each trial (Figure 9). Dynamic patterns of $r_{gpa}$ tended to decrease toward later flights, especially in the G2FI optimal trial. In other words, flights before initiation of flowering times had higher $r_{gpa}$ as opposed to later flights in each trial (Figure 9). In the findings of this study, the time-dependent alteration of genomic prediction results (Figure 9) and the genome-wide marker effects (Figure 8) revealed that the same markers (loci) may have diverse effects on plant height depending on different growth periods in maize.

### 4 CONCLUSION

The estimations of breeding values for genetic materials using HTP combined in association with high-throughput genotyping data (e.g., GBS) has only begun to be explored. This study was among the first evaluating a high number of hybrid maize plots with UAS in a genome-wide association study to discover loci (and where LD permitted, candidate genes) and the prediction ability for the phenotype of interest. Within the scope of this study, early growth stages were found to have undiscovered phenotypic variation controlled by unique loci not found at later stages; this is important to increase our understanding of complex traits in crops. Major conclusions of this study were that (a) earlier plant height data had more importance for GY under optimal planting and irrigation, whereas later plant height data had more importance in terms of yield under stressed conditions; (b) genomic prediction accuracies ($r_{gpa}$) varied for plant heights belonging to different growth stages; and (c) a large number of plant height data captured by UAS can be used as predictors to predict yielding hybrids via RF algorithm.

### ACKNOWLEDGMENTS

Financial support for this research has been provided by research funds as follows: USDA–NIFA–AFRI Award nos. 2017-67013-26185, 2020-68013-32371, and 2021-67013-33915, USDA–NIFA Hatch funds, Texas A&M AgriLife Research, the Texas Corn Producers Board, the Iowa Corn Promotion Board, and the Eugene Butler Endowed Chair in Biotechnology. We would like to acknowledge Jacob Pekar, David Rooney, and Stephen Labar for their agronomic and technical support; graduate students and undergraduate and high school employees of the Texas A&M Quantitative Genetics and Maize Breeding Program for their hard work and effort maintaining fields and collecting phenotypic data.

### CONFLICT OF INTEREST STATEMENT

Authors declared that there is no conflict of interest.

### AUTHOR CONTRIBUTIONS

A.A.: Original draft, formal analysis, data curation, bioinformatic analysis, methodology, data extraction from orthomosaics, conceptualization, investigation, field observation, review and editing (lead); S.C.M.: conceptualization, funding acquisition, methodology, project administration, supervision, resources, review and editing (lead and supporting); S.L.A.: formal analysis, data curation, data extraction from orthomosaics, conceptualization, investigation, field observation, review and editing (supporting); S.C.M.: conceptualization, funding acquisition, resources, review and editing (supporting), software, supervision; L. M.: conceptualization, data curation, formal analysis, investigation, methodology, review and editing (supporting); M.C.R.: bioinformatic, conceptualization, review and editing (supporting); N.D.L.: Genome to field project (G2F) coordination, bioinformatic, conceptualization, investigation, project administration, review and editing (supporting).

### ORCID

Alper Adak https://orcid.org/0000-0002-2737-8041

### REFERENCES

Alscher, R. G. (2002). Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. *Journal of Experimental Botany, 53*, 1331–1341. https://doi.org/10.1093/jexbot/53.372.1331

Anderson, S. L., Mahan, A. L., Murray, S. C., & Klein, P. E. (2018). Four parent maize (FPM) population: Effects of mating designs on linkage disequilibrium and mapping quantitative traits. *The Plant Genome, 11*, 170102. https://doi.org/10.3835/plantgenome2017.11.0102

Anderson, S. L., & Murray, S. C. (2020). R/USAtools:plotshpcreate: Create multi-polygon shapefiles for extraction of research plot scale agriculture remote sensing data. *Frontiers in Plant Science, 11*, 511768. https://doi.org/10.3389/fpls.2020.511768

Anderson, S. L., Murray, S. C., Chen, Y., Malambo, L., Chang, A., Popescu, S., Cope, D., & Jung, J. (2020). Unoccupied aerial system enabled functional modeling of maize height reveals dynamic expression of loci. *Plant Direct, 4*, e00223. https://doi.org/10.1002/pld.2223

Anderson, S. L., Murray, S. C., Malambo, L., Ratcliff, C., Popescu, S., Cope, D., Chang, A., Jung, J., & Thomasson, J. A. (2019). Prediction of maize grain yield before maturity using improved temporal height estimates of unmanned aerial systems. *The Plant Phenome Journal, 2*, 1–15. https://doi.org/10.2135/tppj2019.02.0004
Bac-Molenaar, J. A., Vreugdenhil, D., Granier, C., & Keurentjes, J. J. B. (2015). Genome-wide association mapping of growth dynamics detects time-specific and general quantitative trait loci. *Journal of Experimental Botany*, 66, 5567–5580. https://doi.org/10.1093/jxb/erv176

Barrero Farfan, I. D., Murray, S. C., Labar, S., & Pietsch, D. (2013). A multi-environment trial analysis shows slight grain yield improvement in Texas commercial maize. *Field Crops Research*, 149, 167–176. https://doi.org/10.1016/j.fcr.2013.04.017

Beavis, W. D., Grant, D., Albertsen, M., & Fincher, R. (1991). Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theoretical and Applied Genetics*, 83, 141–145. https://doi.org/10.1007/BF00226242

Bensen, R. J., Johal, G. S., Crane, V. C., Tossberg, J. T., Schnable, P. S., Meeley, R. B., & Briggs, S. P. (1995). Cloning and characterization of the maize An1 gene. *The Plant Cell*, 7, 75–84. https://doi.org/10.1105/tpc.7.1.75

Bian, Y., Yang, Q., Balint-Kurti, P. J., Wisser, R. J., & Holland, J. B. (2014). Limits on the reproducibility of marker associations with southern leaf blight resistance in the maize nested association mapping population. *BMC Genomics*, 15, 1068. https://doi.org/10.1186/1471-2164-15-1068

Block, A. K., Vaughan, M. M., Schmelz, E. A., & Christensen, S. A. (2019). Biosynthesis and function of terpenoid defense compounds in maize (*Zea mays*). *Planta*, 249, 21–30. https://doi.org/10.1007/s00425-018-2999-2

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell*, 169, 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

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

Carter, S. G., & Tipton, C. L. (1985). Purification and characterization of a 5′-nucleotidase from *Zea mays* microsomes. *Phytochemistry*, 25, 33–37. https://doi.org/10.1016/S0031-9422(00)94496-0

Casarett, J. A., El-Keraeamy, A., Zeng, B., Stiegelmeyer, S. M., Chen, X., Bi, Y. M., & Rothstein, S. J. (2016). Expression of OsMYB55 in maize activates stress-responsive genes and enhances heat and drought tolerance. *BMC Genomics*, 17, 312. https://doi.org/10.1186/s12864-016-2659-5

Chang, Y. M., Liu, W. Y., Shih, A. C. C., Shen, M. N., Lu, C. H., Lu, M. Y. J., Yang, H. W., Wang, T. Y., Chen, S. C. C., Chen, S. M., Li, W. H., & Ku, M. S. B. (2012). Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. *Plant Physiology*, 160, 165–177. https://doi.org/10.1104/pp.112.203810

Djak, M., Modarres, A. M., Hamilton, R. I., Dwyer, L. M., Stewart, D. W., Mather, D. E., & Smith, D. L. (1999). Leafy reduced-stature maize hybrids for short-season environments. *Crop Science*, 39, 1106–1110. https://doi.org/10.2135/cropsci1999.0011183X0039000040025x

Ding, J., Ali, F., Chen, G., Li, H., Mahuku, G., Yang, N., Narro, L., Magorokosho, C., Makumbi, D., & Yan, J. (2015). Genome-wide association mapping reveals novel sources of resistance to northern corn leaf blight in maize. *BMC Plant Biology*, 15, 206. https://doi.org/10.1186/s12870-015-0589-z

Dong, Z., Xu, Z., Xu, L., Galli, M., Gallavotti, A., Dooner, H. K., & Chuck, G. (2020). Necrotic upper tips mimic heat and drought stress and encodes a protoxylem-specific transcription factor in maize. *Proceedings of the National Academy of Sciences*, 117, 20908–20919. https://doi.org/10.1073/pnas.2005014117

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, 117, e19379. https://doi.org/10.1371/journal.pone.0019379

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, 4, 250–255. https://doi.org/10.3835/planigenome2011.08.0024

Feldman, M. J., Paul, R. E., Banan, D., Barrett, J. F., Sebastian, J., Yee, M. C., Jiang, H., Lipka, A. E., Bruttell, T. P., Dinneny, J. R., Leakey, A. D. B., & Baxter, I. (2017). Time dependent genetic analysis links field and controlled environment phenotypes in the model C4 grass Setaria. *PLoS Genetics*, 13, e1006841. https://doi.org/10.1371/journal.pgen.1006841

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29, 1189–1232. https://doi.org/10.1214/aos/1013203451

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

Gage, J. L., Jarquín, D., Romay, C., Lorenz, A., Buckler, E. S., Kaeppler, S., Alkalifah, N., Bohn, M., Campbell, D. A., Edwards, J., Ertl, D., Flint-Garcia, S., Gardiner, J., Good, B., Hirsch, C. N., Holland, J., Hooker, D. C., Knoll, J., Kolkman, J., … de Leon, N. (2017). The effect of artificial selection on phenotypic plasticity in maize. *Nature Communications*, 8, 1348. https://doi.org/10.1038/s41467-017-01450-2

Gao, Y., He, C., Zhang, D., Liu, X., Xu, Z., Tian, Y., Liu, X. H., Zang, S., Pauly, M., Zhou, Y., & Zhang, B. (2017). Two trichome birefringence-like proteins mediate xylan acetylation, which is essential for leaf blight resistance in rice. *Plant Physiology*, 173, 470–481. https://doi.org/10.1104/pp.16.01618

Gibson, G. (2010). Hints of hidden heritability in GWAS. *Nature Genetics*, 42, 558–560. https://doi.org/10.1038/ng0710-558

Girardeau-Montaut, D. (2016). CloudCompare. Version 2.8. https://www.danielgm.net/cc/

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*, 9, e90346. https://doi.org/10.1371/journal.pone.0090346

Haghighattalab, A., Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., & Poland, J. (2017). Application of geographically weighted regression to improve grain yield prediction from unmanned aerial system imagery. *Crop Science*, 57, 2478–2489. https://doi.org/10.2135/cropscireview.2016.12.1016

Han, L., Yang, G., Yang, H., Xu, B., Li, Z., & Yang, X. (2018). Clustering field-based maize phenotyping of plant-height growth and canopy spectral dynamics using a UAV remote-sensing approach. *Frontiers in Plant Science*, 9, 1638. https://doi.org/10.3389/fpls.2018.01638

Hartwig, T., Chuck, G. S., Fujisoka, S., Klempien, A., Weizbauer, R., Potluri, D. P. V., Choe, S., Johal, G. S., & Schulz, B. (2011). Brassinosteroid control of sex determination in maize. *Proceedings of the National Academy of Sciences*, 108, 19814–19819. https://doi.org/10.1073/pnas.1108359108

Isenburg, M. (2015). LAStools: Efficient tools for LiDAR processing. *Version 130506*. Department of Comuter Science, University of North Carolina at Chapel Hill.
Neumann, K., Zhao, Y., Chu, J., Keilwagen, J., Reif, J. C., Kilian, B., & Graner, A. (2017). Genetic architecture and temporal patterns of biomass accumulation in spring barley revealed by image analysis. *BMC Plant Biology, 17*, 137. https://doi.org/10.1186/s12870-017-1085-4

Pauli, D., Andrade-Sanchez, P., Carmo-Silva, A. E., Gazave, E., French, A. N., Heun, J., Hunsaker, D. J., Lipka, A. E., Setter, T. L., Strand, R. J., Thorp, K. R., Wang, S., White, J. W., & Gore, M. A. (2016). Field-based high-throughput plant phenotyping reveals the temporal patterns of quantitative trait loci associated with stress-responsive traits in cotton. *G3: Genes, Genomes, Genetics, 6*, 865–879. https://doi.org/10.1534/g3.115.023515

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

Polya, G. M. (1974). Regulation of a plant 5′(3′)-ribonucleotide phosphohydrolase by cyclic nucleotides and pyrimidine, purine, and cytocinokin ribosides. *Proceedings of the National Academy of Sciences, 71*, 1299–1303. https://doi.org/10.1073/pnas.71.4.1299

Polya, G. M. (1975). Purification and characterization of a cyclic nucleotide-regulated 5′-nucleotidase from potato. *Biochimica et Biophysica Acta (BBA)-Enzymology, 384*(2), 443–457. https://doi.org/10.1016/0005-2744(75)90045-5

rapidlasso. (2017). LASTools: Efficient LiDAR processing software. Version 170628. rapidlasso GmbH. http://rapidlasso.com/LASTools

Rood, S. B., & Major, D. J. (1981). Diallel analysis of leaf number, rapid growth and yield in soybean. *Crop Science, 21*(6), 867–873. https://doi.org/10.2135/cropsci1981.0011883X002100060017x

Sari-Gorla, M., Krajewski, P., Di Fonzo, N., Villa, M., & Frova, C. (1999). Genetic analysis of drought tolerance in maize by molecular markers. II. Plant height and flowering. *Theoretical and Applied Genetics, 99*, 289–295. https://doi.org/10.1007/s001220051234

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., & Lohmann, J. U. (2005). A gene expression map of Arabidopsis thaliana development. *Nature Genetics, 37*(5), 501–506. https://doi.org/10.1038/ng1543

Shin, J. H., Blay, S., McNeney, B., & Graham, J. (2006). LDheatmap: An R function for graphical display of pairwise linkage disequilibrium between single nucleotide polymorphisms. *Journal of Statistical Software, 16*(3), 1–10.

Sibov, S. T., De Souza Jr, C. L., Garcia, A. A. F., Silva, A. R., Garcia, A. F., Mangolin, C. A., Benchimol, L. L., & De Souza, A. P. (2003). Molecular mapping in tropical maize (*Zea mays L.*) using microsatellite markers. 2. Quantitative trait loci (QTL) for grain yield, plant height, ear height and grain moisture. *Hereditas, 139*, 107–115. https://doi.org/10.1111/j.1600-5223.2003.01667.x

Sun, J., Poland, J. A., Mondal, S., Crossa, J., Juliana, P., Singh, R. P., Rutkoski, J. E., Jannink, J. L., Crespo-Herrera, L., Velu, G., Huerta-Espino, J., & Sorrells, M. E. (2019). High-throughput phenotyping platforms enhance genomic selection for wheat grain yield across populations and cycles in early stage. *Theoretical and Applied Genetics, 132*, 1705–1720. https://doi.org/10.1007/s00122-019-03309-0

Su, Y., Wu, F., Ao, Z., Jin, S., Qin, F., Liu, B., Pang, S., Liu, L., & Guo, Q. (2019). Evaluating maize phenotype dynamics under drought stress using terrestrial lidar. *Plant Methods, 15*, 11. https://doi.org/10.1186/s13007-019-0396-x

Sun, L., & Wu, R. (2015). Mapping complex traits as a dynamic system. *Physics of Life Reviews, 13*, 155–185. https://doi.org/10.1016/j.plrev.2015.02.007

Tanger, P., Klasssen, S., Mojica, J. P., Lovell, J. T., Movers, B. T., Baraoidan, M., Naredo, M. E. B., Mcnally, K. L., Poland, J. B., Bush, D. R., Leung, H., Leach, J. E., & Mckay, J. K. (2017). Field-based high throughput phenotyping rapidly identifies genomic regions controlling yield components in rice. *Scientific Reports, 7*, 42839. https://doi.org/10.1038/srep42839

Terryn, N., Van Montagu, M., & Inzé, D. (1993). GTP-binding proteins in plants. *Plant Molecular Biology, 22*, 143–152. https://doi.org/10.1007/BF00039092

Tessmer, O. L., Jiao, Y., Cruz, J. A., Kramer, D. M., & Chen, J. (2013). Functional approach to high-throughput plant growth analysis. *BMC Systems Biology, 7*, S17. https://doi.org/10.1186/1752-0509-7-S6-S17

Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D., & Buckler, E. S. (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nature Genetics, 28*, 286–289. https://doi.org/10.1038/ng1035

Veldboom, L. R., & Lee, M. (1996). Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: II. Plant height and flowering. *Crop Science, 36*, 1320–1327. https://doi.org/10.2135/cropsci1996.0011183X0036000050041x

Wallace, J. G., Zhang, X., Beyene, Y., Semagn, K., Olsen, M., Prasanna, B. M., & Buckler, E. S. (2016). Genome-wide association for plant height and flowering time across 15 tropical maize populations under managed drought stress and well-watered conditions in Sub-Saharan Africa. *Crop Science, 56*, 2365–2378. https://doi.org/10.2135/cropsci2015.10.0632

Wang, Y., Yao, J., Zhang, Z., & Zheng, Y. (2006). The comparative analysis based on maize integrated QTL meta map and meta-analysis of plant height QTLs. *Chinese Science Bulletin, 51*, 2219–2230. https://doi.org/10.1143/0011883X00210006017x

Wang, Z., Wu, X., Ren, Q., Chang, X., Li, R., & Jing, R. (2010). QTL mapping for developmental behavior of plant height in wheat (*Triticum aestivum L.*). *Euphytica, 174*, 447–458. https://doi.org/10.1007/s10681-010-0166-3

Wang, X., Zhang, R., Song, W., Han, L., Liu, X., Sun, X., Luo, M., Chen, K., Zhang, Y., Yang, H., Yang, G., Zhao, Y., & Zhao, J. (2019). Dynamic plant height QTL revealed in maize through remote sensing phenotyping using a high-throughput unmanned aerial vehicle (UAV). *Scientific Reports, 9*, 3458. https://doi.org/10.1038/s41598-019-39448-z

Winkler, R. G., & Helentjaris, T. (1995). The maize Dwarf3 gene encodes a cytochrome P450-mediated early step in Gibberellin biosynthesis. *The Plant Cell, 7*, 1307–1317. https://doi.org/10.1105/tpc.7.8.1307

Wu, R., & Lin, M. (2006). Functional mapping—How to map and study the genetic architecture of dynamic complex traits. *Nature Reviews Genetics, 7*, 229–237. https://doi.org/10.1038/nrg1804

Würschum, T., Liu, W., Busmeleyner, L., Tucker, M. R., Reif, J. C., Weisssmann, E. A., Hahn, V., Ruckelshausen, A., & Maurer, H. (2014). Mapping dynamic QTL for plant height in triticale. *BMC Genetics, 15*, 59. https://doi.org/10.1186/1471-2156-15-59

Xing, A., Gao, Y., Ye, L., Zhang, W., Cai, L., Ching, A., Llica, V., Johnson, B., Liu, L., Yang, X., Kang, D., Yan, J., & Li, J. (2015). A rare SNP mutation in Brachytic2 moderately reduces plant height and
increases yield potential in maize. *Journal of Experimental Botany*, 66, 3791–3802. https://doi.org/10.1093/jxb/erv182

Xiong, G., Cheng, K., & Pauly, M. (2013). Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the *Arabidopsis* mutant tbl29. *Molecular Plant*, 6, P1373–P1375. https://doi.org/10.1093/mp/sts014

Yan, J. (2003). Dynamic analysis of QTL for plant height at different developmental stages in maize (*Zea mays* L.). *Chinese Science Bulletin*, 48, 2601–2607. https://doi.org/10.1360/03wc0044

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., Fang, W., Feng, H., Xie, W., Lian, X., Wang, G., Luo, Q., Zhang, Q., Liu, Q., & Xiong, L. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications*, 5, 5087. https://doi.org/10.1038/ncomms6087

Yin, X., Jaja, N., Mcclure, M. A., & M. Hayes, R. (2011a). Comparison of models in assessing relationship of corn yield with plant height measured during early-to mid-season. *Journal of Agricultural Science*, 3, 14–24. https://doi.org/10.5539/jas.v3n3p14

Yin, X., Mcclure, M. A., Jaja, N., Tyler, D. D., & Hayes, R. M. (2011b). In-season prediction of corn yield using plant height under major production systems. *Agronomy Journal*, 103, 923–929. https://doi.org/10.2134/agronj2010.0450

Zhang, X., Gonzalez-Carranza, Z. H., Zhang, S., Miao, Y., Liu, C., & Roberts, J. A. (2019). F-Box proteins in plants. *Annual Plant Reviews*, 2, 307–328. https://doi.org/10.1002/9781119312994.apr0701

Zhang, X., Huang, C., Wu, D., Qiao, F., Li, W., Duan, L., Wang, K., Xiao, Y., Chen, G., Liu, Q., Xiong, L., Yang, W., & Yan, J. (2017). High-throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth. *Plant Physiology*, 173, 1554–1564. https://doi.org/10.1104/pp.16.01516

**SUPPORTING INFORMATION**

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

---

**How to cite this article:** Adak A, Murray SC, Anderson SL, Popescu SC, Malambo L, Romay MC, & de Leon N. Unoccupied aerial systems discovered overlooked loci capturing the variation of entire growing period in maize. *Plant Genome*. 2021;14:e20102. https://doi.org/10.1002/tpg2.20102