

1      **Corresponding Author:**  
2            Ivan Baxter  
3            USDA-ARS  
4            Donald Danforth Plant Science Center  
5            (314) 587-1438  
6            [ivan.baxter@ars.usda.gov](mailto:ivan.baxter@ars.usda.gov)  
7

8      **Research Area:**

9            Breakthrough Technologies  
10          Ecophysiology and Sustainability  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

1   **A versatile phenotyping system and analytics platform reveals diverse temporal responses to water  
2 availability in *Setaria***

3

4   Noah Fahlgren\*, Maximilian Feldman\*, Malia A. Gehan\*, Melinda S. Wilson, Christine Shyu, Douglas  
5   W. Bryant, Steven T. Hill, Colton J. McEntee, Sankalpi N. Warnasooriya, Indrajit Kumar, Tracy Ficor,  
6   Stephanie Turnipseed, Kerrigan B. Gilbert, Thomas P. Brutnell, James C. Carrington, Todd C. Mockler,  
7   Ivan Baxter

8

9   Donald Danforth Plant Science Center, United States (NF, MF, MAG, MSW, CS, DWB, STH, CJM, SW,  
10   IK, TF, ST, KBG, TPB, JCC, and TCM); USDA-ARS, Donald Danforth Plant Science Center, United  
11   States (IB)

12

13   \*The following authors contributed equally to this work:

14           Noah Fahlgren

15           Maximilian Feldman

16           Malia A. Gehan

17

18   **One-sentence Summary:**

19   A high-throughput image-based phenotyping platform with controlled watering and new open-source trait  
20   extraction tools expose distinct responses to water availability in wild and domesticated *Setaria*.

21

22

1       \*Noah Fahlgren, Maximilian Feldman and Malia A. Gehan contributed equally to this work:

2       <sup>1</sup>This work was supported by the United States Department of Agriculture-National Institute of Food and  
3       Agriculture (MOW-2012-01361 to NF and 2014-67012-22269 to CS), National Science Foundation  
4       (MCB-1330562 to JCC, IIA-1355406 to TCM, and IOS-1202682 to MAG), United State Department of  
5       Energy (DOE-SC-008796 to TPB, TCM and IB and DE-SC0006627 to TCM), National Institutes of  
6       Health (AI043288 to JCC) and United States Department of Agriculture Agricultural Research Service  
7       Intramural Funds (IB)

8       <sup>2</sup>Corresponding author email address: [ivan.baxter@ars.usda.gov](mailto:ivan.baxter@ars.usda.gov)

9       The author responsible for the distribution of materials essential to the findings of this study in  
10      accordance with the policy described in the Instructions For Authors ([www.plantphysiol.org](http://www.plantphysiol.org)) is: Dr. Ivan  
11      Baxter (ivan.baxter@ars.usda.gov)

12      <sup>3</sup> NF, MF, MAG, MSW and IB designed and collected high-throughput phenotyping experiment. NF, MF,  
13      and MAG designed and wrote PlantCV image analysis tools. DWB, STH and CJM designed and wrote  
14      PhenoFront database access tool. NF, MF, MAG, CS, SW, IK, TF, ST, KBG, and IB designed and  
15      executed trait validation experiments. NF, MF, MAG, CS, and IB analyzed data. NF, MF, MAG, MSW,  
16      CS, and IB wrote the manuscript. TPB, JCC, TCM and IB, supervised and reviewed the manuscript.

17

1   **ABSTRACT**

2

3   Phenotyping has become the rate-limiting step in using large-scale genomic data to understand and  
4   improve agricultural crops. Here, the Bellwether Phenotyping Platform for controlled-environment plant  
5   growth and automated, multimodal phenotyping is described. The system has capacity for 1,140 plants,  
6   which pass daily through stations to record fluorescence images, near infrared images, and visible images.  
7   Plant Computer Vision (PlantCV) was developed as open-source software for quantitative image analysis  
8   that is a platform independent. In a four week experiment, wild *Setaria viridis*, and domesticated *Setaria*  
9   *italica* had fundamentally different temporal responses to water availability. While the lines produced  
10   similar levels of biomass under limited water conditions, *Setaria viridis* maintained the same water use  
11   efficiency under water replete conditions, while *Setaria italica* shifted to less efficient growth. Overall,  
12   the Bellwether Phenotyping Platform and PlantCV software detected significant effects of genotype, and  
13   environment on height, biomass, water-use efficiency, color, plant architecture, and near-infrared traits.  
14   All ~79,000 images acquired during the course of the experiment are publically available.

1   **INTRODUCTION**

2  
3   With growing and more affluent worldwide populations, agricultural crop yields will need to increase  
4   significantly by 2050 (Ray et al., 2013; Gerland et al., 2014). Sustainably increasing crop yields in the  
5   face of changing environments (IPCC, 2014), and doing so with a smaller environmental footprint, is one  
6   of the most pressing global challenges of the 21st century. With changing climate, more prevalent  
7   episodes of regional drought and damage due to precipitation extremes will limit agricultural productivity  
8   (IPCC, 2014). Salination of agricultural lands and low soil fertility are also increasingly important  
9   limitations to crop productivity in many regions.

10  
11   Advances in DNA sequencing technology has facilitated rapid advances in plant genomics, which has  
12   accelerated traditional breeding, molecular marker-assisted breeding, genome editing, and other  
13   approaches for crop improvement. Application of genomic technologies, however, is limited by the ability  
14   to reliably quantify plant traits (phenotypes). Traits can range in scale from gene expression to yield in the  
15   field. Variation in most traits is quantitative, with contributions from multiple genetic loci. Use of  
16   genetically defined populations, such as recombinant inbred lines (RILs), near-isogenic lines, or  
17   accession/association panels, can be used to identify the underlying genetic bases for phenotypes, but only  
18   when subtle changes in phenotype are accurately and consistently quantified (Benfey and Mitchell-Olds,  
19   2008). Consequently, plant phenotyping is widely regarded as the rate limiting step in crop improvement  
20   using genome-enabled approaches (Furbank and Tester, 2011; McCouch et al., 2013).

21  
22   Abiotic stresses can have variable or gradual effects that depend on plant developmental stage. Therefore,  
23   measuring phenotypes over time is necessary for in-depth understanding of plant stress responses  
24   (Richards and Thurling, 1978; Schoffl, 1998; Mahfoozi, 2001). Destructive methods of assessing plant  
25   phenotypes have been used in the majority of studies (Furbank and Tester, 2011), but such methods do  
26   not allow traits of discrete individuals to be compared over time. Accordingly, destructive sampling  
27   requires an exponential increase in sampling population size with each time-point added. Non-invasive  
28   and non-destructive techniques permit temporal examination of traits in individual plants, reducing the  
29   number of plants needed and permitting larger populations to be examined. The emerging field of plant  
30   phenomics, the high-throughput, non-destructive examination of plant growth and development in field or  
31   controlled-environment settings, provides new technologies that increase the accuracy and speed of data  
32   collection and analysis (Furbank and Tester, 2011). Current non-destructive phenomics technologies  
33   focus on a number of traits that, directly or indirectly, reflect chlorophyll content, carotenoid content,

1 photosynthesis, transpiration, plant water content, biomass, and height (White et al., 2012; Andrade-  
2 Sanchez et al., 2014).

3

4 Phenotyping in the field has all the advantages and disadvantages traditionally afforded field-based  
5 research. For example, field phenomics allows crop-sized plants to be examined in a natural  
6 environmental setting, but field-grown plants are subject to discrete seasonal growing periods and  
7 uncontrollable environmental conditions that increase experimental variability. Although a less natural  
8 setting, controlled-environment phenomics platforms (Hartmann et al., 2011; Klukas et al., 2014) allow 1)  
9 precise control of environmental variables and treatments, 2) experimental replication under reproducible  
10 conditions, and 3) better control of instrument placement and functionality. Controlled environment  
11 phenotyping also enables faster experimental turnover, which expedites the cyclical development of  
12 image processing algorithms and proxy measurement models that are applicable to both field and  
13 controlled environment phenomics.

14

15 The Bellwether Phenotyping Platform combines automated, controlled-environment plant growth with  
16 high-throughput, non-destructive imaging and is representative of the next wave of phenotyping platforms  
17 aimed at relieving bottlenecks of phenomics data collection. Following data collection, image processing  
18 and trait analysis are the next barrier to understanding the underlying biology in high-throughput  
19 phenotyping data. Standardized methods for processing image-based, high-throughput plant phenomics  
20 data lag behind high-throughput sequencing analysis tools in part because of the variety of commercial  
21 and non-commercial platforms used, the numerous research focus areas, and the variety of species and  
22 treatments. There is a growing plant phenomics community (<http://www.plant-phenotyping.org/>) and an  
23 excellent database of both commercial and open-source plant image processing software (Lobet et al.,  
24 2013). However, we chose to develop a new trait extraction software platform for daily high-throughput  
25 image analysis, Plant Computer Vision (PlantCV). PlantCV is written in a scripting language that has  
26 been shown to be accessible to biologists (Mangalam, 2002; Dudley and Butte, 2009), and with a  
27 community contribution schema that has been successful for other bioinformatics resources (Oliphant,  
28 2006; Hunter and Dale, 2007).

29

30 The utility of the system and software is demonstrated here using *Setaria* species and RILs grown under  
31 different water conditions. The C4 model monocot *Setaria viridis* (green millet) and the drought-tolerant  
32 domesticated crop *Setaria italica* (foxtail millet) (Zhang et al., 2007; Lata et al., 2010; Lata et al., 2011)  
33 are closely related to each other and to important crops with similar architecture, such as maize, sorghum,  
34 miscanthus and switchgrass (Brutnell et al., 2010; Li and Brutnell, 2011; Sage and Zhu, 2011). Given the

1 relatively limited understanding of monocot stress responses compared to *Arabidopsis* (Bray, 1997;  
2 Wilkins et al., 2010), and the importance of cereal crops as sources of food (Ray et al., 2013), *Setaria* has  
3 emerged as a useful model system to explore fundamental aspects of monocot biology (Nelson and  
4 Dengler, 1997; Kellogg, 1999; Doust, 2007a). Concerted data collection with the Bellwether Phenotyping  
5 Platform and analysis with PlantCV detected fundamentally different temporal responses to water  
6 availability between the wild and domesticated *Setaria* species. The wild *Setaria* line maintains water-use  
7 efficient growth while the domesticated line shifts to less water-use efficient growth in water sufficiency.  
8 This publicly available *Setaria* dataset of ~79,000 images represents the next generation of big data to  
9 query for phenotypes that impact important yield traits.

10

## 11 RESULTS AND DISCUSSION

12

### 13 Design of the Bellwether Phenotyping Platform

14

15 The custom design of the Bellwether Phenotyping Platform was key to the execution of the *Setaria* water-  
16 limitation experiment described below. The Bellwether Phenotyping Platform is comprised of a Conviron  
17 climate-controlled growth house integrated with a multi-camera digital imaging system (LemnaTec  
18 Scanalyzer<sup>3D-HT</sup>, Fig. 1A). This automated, high-throughput platform allows for repeated non-destructive  
19 image capture for multi-parametric analysis, and provides valuable information on the physiological  
20 changes of the plants over time. The integration of Conviron and LemnaTec technologies and other  
21 custom design decisions were motivated by the goal of tightly controlling the environment while  
22 analyzing large numbers of plants. In comparison to greenhouse environments where external factors such  
23 as cloud cover or extreme seasonal temperatures can greatly impact the environmental conditions and  
24 introduce unpredictable variables into an experiment, the control of the Bellwether Phenotyping Platform  
25 facilitates reproducible temporal phenotyping of plants.

26

27 A compact network of over 180 m of conveyor belts is used to store and transport 1,140 plants with a 4"  
28 pot size diameter, to and from watering and imaging stations inside the chamber. The conveyor belts are  
29 divided into four distinct modules (MOD1–4), which can be controlled independently or as a whole,  
30 providing flexibility in experimental design (Fig. 1A). A custom engineered automated door maintains the  
31 growth house environment while providing a route for plants to leave and enter the chamber (Fig. 1A).  
32 The pots are moved in Radio-frequency identification (RFID) chipped carriers, which are associated with  
33 plant barcodes that are used to organize collected image data and metadata, including water treatment  
34 instructions. Each MOD contains a precision dual-tank watering station that is equipped with a scale,

1 which can be configured to water to a fixed volume or to a target weight. When watered to an  
2 experimentally determined target weight, the real-time weight of each pot is used to calculate the amount  
3 of water required to reach the pre-determined target.

4

5 The decision to install four watering stations inside the growth house (Fig. 1A) allows for tight control of  
6 soil water status while maintaining imaging throughput. Due to the single-file movement of the carriers  
7 through the closed-loop conveyor system, a single watering station would have limited the number of  
8 water applications possible in a day and would have competed with imaging time. A single watering  
9 station would have also introduced a significant temporal divide between the first and last watered plants,  
10 adding variability into the treatment groups and confounding water usage analysis. Instead, four plants  
11 can be watered simultaneously, improving efficiency and drastically reducing the time required to water  
12 the entire population (less than two hours).

13 The digital imaging loop connected to the growth house consists of a dark adaptation tunnel and three  
14 distinct imaging chambers: visible light (VIS), photosystem II (PSII) fluorescence, and near-infrared  
15 (NIR) (Fig. 1A). The imaging stations are described in detail in the sections below. Visible light imaging  
16 is used to capture and quantify morphological parameters such as plant shape, color, size and biomass.  
17 PSII fluorescence is used for the analysis of photosynthetic efficiency, which can be used to estimate the  
18 response to plant treatment or stress, or to measure photosynthetic differences among genetically diverse  
19 lines (Maxwell and Johnson, 2000; Garg et al., 2002; Fernández-García et al., 2014; Talukder et al.,  
20 2014). Near-infrared imaging is used to estimate the tissue water content (Carter, 1991; Peñuelas and  
21 Filella, 1998; Seelig et al., 2008; Seelig et al., 2009). The throughput of the system depends on  
22 experimental design, the desired number of output images, the frequency of water treatment, and is  
23 limited by the transport rate of the conveyor. At full capacity (1,140 plants), using all imaging chambers  
24 to capture top-view and four side-view images (PSII top-view only), the entire population can be imaged  
25 in approximately 30 hours, typically yielding between 25,000 and 37,500 images per week.

26 ***Setaria* Water-Limitation Experiment**

27 With target-weight water treatments possible on the Bellwether Platform, the water available to the plant  
28 should be tightly controlled. To increase the precision of target-weight watering, pre-filled pots of soil  
29 were used. The use of pre-filled pots also helped to ensure that soil density was similar between pots. To  
30 test both the Bellwether Platform, and examine the response of *Setaria* to water-limited conditions, ten  
31 *Setaria* lines were grown under four different water regimes (full water capacity: 100% FC, 66% FC, 33%  
32 FC, and 0% FC) imposed 17 days after planting (DAP) and maintained for 17 days. Plants were watered

1 twice a day starting at ZT8 (Zeitgeber Time) and ZT23. For more experimental details see Materials and  
2 Methods. The lines used were *S. viridis* (accession A10), *S. italica* (accession B100) and eight RILs  
3 derived from a cross of *S. viridis* x *S. italica* (Devos et al., 1998; Wang et al., 1998; Bennetzen et al.,  
4 2012). An average of 15 plants per line per water treatment were loaded onto the system. Additional *S.*  
5 *viridis* and *S. italica* plants were also loaded for parallel destructive physiological tests.

6 The plants that received no water after 17 DAP died within 7 days of treatment and were removed from  
7 the system leaving 33% FC as the severest water deficit treatment. Comparisons between 33% and 100%  
8 FC were the focus of subsequent analysis. Figure 1B shows the volume of water added at each treatment  
9 time-point for *S. viridis* plants watered to 100% and 33% FC. The differences between the two water  
10 treatment set points can be seen clearly after the 33% pots complete their dry down from full water 17–19  
11 DAP. As the plants grow and transpire, more water is needed to return the pots to their set points. Water  
12 treatment outliers can be seen 18 DAP when the second water application was delayed and 24 DAP when  
13 the target-weight point was raised to account for increasing biomass of the plants (Fig. 1B). More  
14 watering treatments can be added if greater control over soil moisture content is desired.

### 15 ***Setaria* Image Acquisition**

16 During the experiment *Setaria* plants varied in height from 1.8 cm to 107.4 cm, so a scaled field of view  
17 (adjusted optical zoom-level) was used during image acquisition to maximize the image pixels (px)  
18 dedicated to plant material, and not background space. Consequently, plant traits measured in pixels from  
19 images taken at different zoom levels are not directly comparable because of the difference in the field of  
20 view. A reference object with known dimensions was imaged using the VIS side-view (SV) and top-view  
21 (TV) cameras at different zoom levels. To compare digital traits across zoom levels, scaling factors were  
22 calculated for both area and length traits using zoom-scaling functions (see Materials and Methods).

### 23 **PhenoFront Database Access Tool**

24 Image and water volume data from LemnaTec phenotyping systems are stored in a PostgreSQL database,  
25 but images themselves are stored separately in raw format. To provide distributed access to image and  
26 water treatment data an independent database interface tool (PhenoFront) was developed. It was necessary  
27 to develop a new database access tool rather than use existing open-source interface tools (Klukas et al.,  
28 2014) because the image storage structure of the LemnaTec version 4.0 database changed significantly  
29 from previous versions.

30 PhenoFront has a convenient web-based user interface to the phenotyping database and provides an  
31 experiment query-building tool that allows users to extract specific image and water treatment data.

1 Importantly, PhenoFront was built with diverse image types in mind (e.g. color vs. grayscale; 8-bit vs. 16-bit) and allows access to images without conversion to a single file format. Queries can be directly downloaded to the user's computer, or data can be downloaded to a remote server capable of high-throughput image processing using utilities such as Wget (<http://www.gnu.org/software/wget/wget.html>) or cURL (<http://curl.haxx.se/>).

6 Data sets extracted with PhenoFront are structured by data acquisition events (snapshot: weighing/watering, imaging) and a single metadata text file contains the experimental and system parameters associated with each event. Image metadata from the experiment presented here includes: experiment identification code; snapshot folder identification number; alphanumeric barcode encoded with species identification, line identification, water treatment group, and unique plant number; unique snapshot timestamp; weight and water volume information; individual image file names included in the snapshot directory. The structured data sets can be easily read by analysis software and are convenient for public data sharing.

14 Although strides have been made to create resources for image processing on the iPlant cyber infrastructure (Kvilekval et al., 2010), a dedicated centralized database for sharing published plant phenotyping image data has not been designated. Although a few plant phenomic datasets are publicly available (Fahlgren et al., 2015), this is the largest publically available dataset of above-ground plant tissue to date (79,200 images of 823 plants). Image data from this experiment is available via <http://plantcv.danforthcenter.org/pages/data.html> in Portable Network Graphics (PNG) and Joint Photographic Experts Group (JPEG) formats. JPEG format images with lossy compression consume approximately 11-fold less storage size, but comparisons of image datasets in JPEG and PNG format revealed significant differences in some analyzed traits (see Materials and Methods). Establishing a dedicated centralized database that is designed to deposit and access large plant phenomic datasets and standardized metadata is vital for aggregating and curating data for the community. The phenomics community will need to collectively consider formats, data and metadata structure, minimal information standards, ontologies, and other sharing issues. Implementing these phenotyping standards will expedite tool development and alleviate barriers to crop improvement.

## 28 **Phenotype Extraction Using Plant Computer Vision (PlantCV) Software**

30 High-throughput plant phenotyping has the potential to improve modeling of genotype by environment interactions and to expedite identification of germplasm that could increase yield and productivity of crop plants. With numerous commercial and custom-built image-based phenotyping platforms in existence, it

1 is unlikely that standard hardware will be used to capture image-based phenomics data in the near future.  
2 Lack of standard hardware thus requires each new platform to go through a significant initialization  
3 period, and a gambit of validation experiments, many of which are described below in this study. Despite  
4 different hardware platforms, image-analysis and trait-extraction is common ground between image-based  
5 phenotyping platforms. Therefore, open-source trait-extraction software with mechanism for community  
6 development will help to alleviate the phenotyping bottleneck on crop improvement.

7  
8 Available commercial and open-source plant image processing software range in their capabilities,  
9 flexibilities, and underlying programming languages and largely focus on analysis of single cells, leaves,  
10 roots, or plants with rosette architecture (<http://www.plant-image-analysis.org>; Lobet et al. 2013). There  
11 are analysis tools in the Plant Image Analysis database, such as LemnaGrid (<https://www.lemnatec.com>)  
12 (Munns and Tester, 2008; Berger et al., 2010; Golzarian et al., 2011; Honsdorf et al., 2014), the ImageJ  
13 (Abràmoff et al., 2004) plugin HTPheno (Hartmann et al., 2011), and the java-based open-source  
14 Integrated Analysis Platform (IAP) (Klukas et al., 2014) that are capable of analyzing larger model plants  
15 and crops with diverse architectures (Supplemental Table SI). We developed the open-source and open-  
16 development PlantCV image analysis platform to emphasize the following features: flexible user-defined  
17 analysis workflows; parallelizable image processing for fast throughput; and a scripting language  
18 implementation that lowers the barrier to community contributions that extend functionality. While some  
19 users may prefer graphical user interfaces for software, script-based programs are easier to develop, and  
20 the precise workflows are detailed directly in the scripts themselves, enabling reproducible research.

21  
22 Extraction of temporal plant trait data from images occurs in three steps: 1) isolation of plant material  
23 from background (Fig. 2A); 2) identification of features (traits) from isolated plants (Fig. 2B); and 3)  
24 analysis of traits across population, treatments, and time (Fig. 2C). PlantCV isolates plant material from  
25 background, quantifies plant traits, and populates an SQLite (<https://www.sqlite.org/>) database that is  
26 easily queried for further analysis across treatments, genotypes, and time. In debug mode, PlantCV  
27 creates annotated intermediate images for each step of image analysis pipelines (e.g. outlining plant  
28 perimeter), allowing users to verify that analysis steps are working as intended. Although *Setaria* images  
29 from the high-throughput Bellwether Phenotyping Platform are analyzed here, PlantCV has been used to  
30 analyze a variety of plant types (Fig. 2D–F), as well as images captured from non-commercial imaging  
31 stations (Fig. 2E and F). For the *Setaria* experiment here, traits such as height, biomass, and plant  
32 architecture, were manually measured and are highly correlated with traits computationally extracted by  
33 PlantCV (Figs. 3–5). Image datasets curated with manual measurements are available at

1    [http://plantcv.danforthcenter.org/pages/data-sets/2013/setaria\\_burnin2.html](http://plantcv.danforthcenter.org/pages/data-sets/2013/setaria_burnin2.html) and are a resource for  
2    generating further trait measurement algorithms.  
3  
4    PlantCV software is built upon the open-source libraries OpenCV (Bradski and Kaehler, 2008), NumPy  
5    (Oliphant, 2006), and Matplotlib (Hunter and Dale, 2007), and contains pipelines built with modular  
6    functions that are currently capable of automated analysis of color and grayscale images from VIS, PSII,  
7    and NIR cameras. PlantCV was written in the Python programming language with hopes that the greater  
8    phenomics community will utilize and extend its functionality. Although Python is not as efficient as  
9    compiled languages (e.g. C or Java) in terms of memory usage and speed of execution (Fourment and  
10   Gillings, 2008), it is currently more widely used by biologists, and is arguably easier to learn (Mangalam,  
11   2002; Dudley and Butte, 2009). On a single processor, PlantCV can process and extract data from 200–  
12   350 images per hour, depending on image size, and the modular script-based architecture allows for easy  
13   parallelization. At the Danforth Center, PlantCV is typically run on 10 CPUs. PlantCV is an open-source  
14   resource under a GNU General Public License (GPLv2) share-alike license that can be leveraged for the  
15   further development of more complex trait identification modules. PlantCV is also available via GitHub,  
16   providing a framework for community contribution that has been successful for other science and Python-  
17   based projects such as NumPy (~284 contributors), and Matplotlib (~294 contributors) (Oliphant, 2006;  
18   Hunter and Dale, 2007). Community contribution helps to maintain software permanence (Gentleman et  
19   al., 2004), which is often a problem for bioinformatics tools maintained by individual labs (Gilbert,  
20   2004). PhenoFront, PlantCV, PlantCV documentation, tutorials, parallelization scripts, downstream  
21   analysis scripts, and a contribution guide can be accessed at <http://plantcv.danforthcenter.org>.  
22

### 23    **VIS Image Processing and Traits**

24    The VIS imaging station in the Bellwether Phenotyping Platform captures visible light (400 – 700 nm)  
25    images with two high-resolution (2454 x 2056 pixel) charge-coupled device (CCD) cameras. One camera  
26    is mounted above the plant and the second camera is side-mounted for top- and side-view imaging,  
27    respectively. The plant carrier is positioned on a piston lifter with a turner that can rotate the plant 360  
28    degrees. During the *Setaria* experiment a single top-view and four side-view (0°, 90°, 180°, and 270°) VIS  
29    images were acquired per plant per time point. In total, *Setaria* plants were imaged 6,399 times from 11–  
30    33 DAP for a total of 31,968 VIS images. The set of images collected for a single plant in one imaging  
31    session is defined as a snapshot.  
32

33    In general, PlantCV VIS image processing pipelines automatically identify background material within  
34    the image, such as the carrier, pot, soil, and side paneling. To mitigate plant identification bias due to

1 genotype, developmental age, abiotic treatment, or other experimental parameters, plant material was  
2 isolated after several stages of background removal, rather than first thresholding for plant ‘greenness.’  
3 For a complete description of the VIS image processing steps, please refer to Supplemental Data and view  
4 the PlantCV online documentation.

5

6 **Plant height:** Plant height was defined as the maximum vertical extent of the plant from the top of the  
7 pot. For 173 randomly selected VIS side-view images, plant height (calibrated using the width of the  
8 plant carrier) was manually measured using ImageJ (Abràmoff et al., 2004). For the same image set, plant  
9 height was automatically estimated using PlantCV with height scaled to correct for the camera zoom level  
10 (see Material and Methods). Ordinary least squares regression analysis confirmed that PlantCV measured  
11 plant height was a good estimator of manually measured height and was robust across changes in camera  
12 zoom level (adjusted  $R^2 = 0.998$ ; Fig. 3A).

13

14 To analyze *Setaria* height more robustly at the population level, estimated plant height was averaged  
15 between all four VIS side-view images for each snapshot and the effect of genotype and low water  
16 availability on height was calculated. Intrinsic height differences under full water conditions were  
17 observed for the *Setaria* genotypes, particularly from 23–33 DAP. *S. viridis* was slightly taller than *S.*  
18 *italica* from 23–33 DAP with a mean height ranging from 50.3–101.2 cm compared to 43.0–98.6 cm,  
19 respectively (Fig. 3B). The height of RIL102 was within the range of the two parent lines but the other  
20 seven RILs were transgressive; RIL161 was shorter on average than the parents while the others were  
21 taller (Supplemental Fig. S1). Height for the tallest lines was underestimated because the camera field of  
22 view in this experiment was not zoomed out enough to completely capture the tallest plants. For example,  
23 truncated plant height due to camera settings is noticeable from 15–20 DAP after which the field of view  
24 was expanded and plants were fully imaged again (Fig. 3B and Supplemental Fig. S1). Although  
25 thorough surveillance can be used to avoid lapses in coverage due to sub-optimal camera settings, the  
26 large number of potentially diverse plants on high-throughput systems makes complete oversight  
27 challenging. A better approach will be to add real-time analysis that raises alarm to potential framing  
28 issues once preset threshold boundaries are reached. The parallelized rate of PlantCV analysis permits  
29 future experiments to implement real-time boundary detection.

30

31 In addition to intrinsic height phenotypes, the impact of limited water availability on plant height was also  
32 measured. *S. viridis* plants watered to 33% FC starting 17 DAP were significantly shorter than plants at  
33 100% FC from 20 DAP through the end of the experiment (95% confidence interval: 1.4–8.0 cm shorter;  
34 Fig. 3B). In contrast, the height of *S. italica* plants from both 100% and 33% FC groups was not

1 significantly different, suggesting that *S. italica* plant height is not tightly coupled to water availability  
2 (Fig. 3B). Of the ten lines analyzed, *S. viridis* has the largest difference in height when comparing the  
3 33% to 100% FC treatment groups (Supplemental Fig. S1). The difference in height for the RILs was  
4 either similar to *S. italica* or was intermediate between the two parents (Supplemental Fig. S1).

5

6 **Above ground biomass:** Throughout the experiment, 41 *S. viridis* and *S. italica* plants randomly selected  
7 from the full capacity water group were collected and the above ground fresh- and dry-weight biomass  
8 were recorded. The images acquired immediately prior to collection were analyzed with PlantCV to  
9 measure plant traits that could be used to model biomass. Linear modeling was done with three PlantCV  
10 measurements: Side-view area, the sum of the above ground plant pixel area from all four side-view VIS  
11 images; Top-view area, the plant pixel area from the single top-view VIS image; and plant height, as  
12 described above. The initial model included side-view area, top-view area and height, adjusted for zoom  
13 level, and all pairwise interaction terms. A stepwise model selection procedure was done using the  
14 Akaike's Information Criterion (Bozdogan, 1987), which resulted in the reduced model:

$$M_{fw} = 3.966 \times 10^{-5} A_{sv} - 0.3193$$

15 where  $M_{fw}$  is fresh-weight biomass and  $A_{sv}$  is side-view area (adjusted  $R^2 = 0.981$ ; Fig. 4A). Dry-weight  
16 biomass was also modeled efficiently with side-view plant pixel area (adjusted  $R^2 = 0.975$ ; Supplemental  
17 Fig. S2). It was hypothesized that residual variation in biomass could be explained by plants that were  
18 partially out of frame or by genotype. PlantCV records when plants extend to or past the border of each  
19 image, and the model including the out-of-frame status of each plant was accepted over the reduced  
20 model ( $p < 0.01$ , F-test) but only slightly improved the model (adjusted  $R^2 = 0.985$ ). In contrast, genetic  
21 background did not improve the model significantly, suggesting that biomass estimation for *S. viridis* and  
22 *S. italica* is robust to species differences, at least within the conditions tested.

23

24 **Growth rates and response to water availability:** Image-based biomass estimates were used to quantify  
25 the impact of limited water availability on *Setaria* fresh-weight biomass (Fig. 4 and Supplemental Fig.  
26 S3). Non-linear least squares regression was used to estimate the growth parameters (three-component  
27 logistic model) for each group (genotype by treatment) (Fig. 4B and C). Under water deficit *S. italica* and  
28 *S. viridis* grew at similar rates (maximum absolute growth rate at ~25 DAP was 1.6 and 1.7 g/day,  
29 respectively; Fig. 4B and C). In contrast, under 100% FC treatment the maximum absolute growth rate of  
30 *S. viridis* was 0.6 g/day more than *S. italica* (3 and 2.4 g/day, respectively) and was reached two days  
31 earlier (Fig. 4C). As a result, the impact of water-limitation on fresh-weight biomass was larger and  
32 observed more quickly in *S. viridis* compared to *S. italica*. *S. viridis* watered to 33% FC starting at 17  
33 DAP accumulated significantly less biomass relative to the control group from 19 DAP through the end

1 of the experiment (95% confidence interval: 0.8–2.7 g; Fig. 4B). *S. italica* watered to 33% FC  
2 accumulated significantly less biomass relative to the control group at 26 DAP through the end of the  
3 experiment (95% confidence interval: 1.5–6.4 g; Fig. 4B). Plant growth and response to the environment  
4 is a dynamic process and non-destructive phenotyping was key to understanding the differences between  
5 *S. viridis* and *S. italica*. For example, although *S. viridis* grows faster and earlier than *S. italica* under  
6 100% FC treatment, biomass at 33 DAP was similar, so endpoint biomass measurements would have  
7 missed the growth differences between *S. viridis* and *S. italica* (Fig. 4B and C).

8

9 **Integrated water-use efficiency:** The ability to measure and record the volume of water applied to  
10 individual plants (Fig. 1B) in combination with the estimated biomass of plants over time (Fig. 4B)  
11 provides a framework for quantifying plant intrinsic water-use efficiency (WUE). Here intrinsic WUE is  
12 operationally defined as the amount of biomass accumulated per ml of water applied (units: mg/ml). No  
13 significant difference in WUE was observed for *S. viridis* plants watered to 100% or 33% FC (Fig. 4D).  
14 Therefore, the amount of *S. viridis* accumulated biomass is proportional to the volume of water supplied.  
15 In contrast, *S. italica* plants had lower WUE at 100% FC relative to 33% FC, with significant differences  
16 observed from 22–27 DAP (Fig. 4D). This suggests that *S. italica* uses water less efficiently under water-  
17 sufficient conditions.

18

19 **Plant architecture:** The traits discussed above do not require knowledge of specific plant architecture, but  
20 architectural phenotypes such as tillering and leaf angle can have a significant impact on plant  
21 performance (Warnasooriya and Brutnell, 2014). Known plant architecture differences between the wild  
22 *S. viridis* and domesticated *S. italica* include the number of axillary branches and tiller number  
23 (Darmency et al., 1987; Doust and Kellogg, 2006; Doust, 2007b; Doust et al., 2009). Computational 3D  
24 reconstruction of plant architecture from 2D images would promote accurate estimates of tiller count, but  
25 would have a significant impact on imaging throughput because of the increase in images per plant, per  
26 time-point needed to enable 3D reconstruction (Phattaralerphong and Sinoquet, 2005; Paproki et al.,  
27 2012). Modeling is an alternative approach that can be used to predict the number of tillers given one or  
28 more morphological parameters that are more easily measured.

29

30 In a first attempt at generating a predictive model for *Setaria* architecture, tiller number, tiller angle (angle  
31 between the two outermost tillers), and leaf angle (angle between the leaf and tiller of the first leaves on  
32 the two outermost tillers) values were measured using ImageJ (Abràmoff et al., 2004) on 58 randomly  
33 selected images of plants 25 and 26 DAP. The results were compared to two morphological traits that are  
34 simple to calculate using PlantCV: height-width ratio (HW, height divided by extent-x; Fig. 5A) and

1 solidity (plant area divided by convex hull area). HW was significantly correlated ( $p < 0.01$ ) with all three  
2 manually measured traits, but the best model for HW, which accounted for 37% of the variance in an  
3 ANOVA analysis, included only tiller angle and leaf angle (see Material and Methods). Solidity was not  
4 significantly correlated with any of the manually measured traits. While HW ratio is not inherently a  
5 measure of a single plant architecture trait, it can serve as an easily obtainable proxy for further analysis.  
6 To generate a predictive model for tiller count, HW and several other traits were examined. Tillers were  
7 manually counted for 195 randomly selected images. HW, solidity, fresh-weight biomass, width (extent-  
8 x), and height were tested as potential explanatory variables for manual tiller counts. Ordinary least  
9 squares regression was used to generate a model for the number of tillers:

$$TC = 0.2289M_{fw} - 2.165HW + 5.164$$

10 where  $TC$  is the number of tillers, and  $M_{fw}$  is fresh-weight biomass and  $HW$  is height-width ratio. The  
11 tiller model explained 64.7% of the variation in tiller number (Fig. 5B). To further test the tiller model,  
12 tillers were counted for a second set of 646 randomly selected images. The accuracy of the tiller model to  
13 predict tiller number in the second set of 646 images was assessed, and the median difference between  
14 predicted and manual tiller number was  $\pm 1.29$  tillers (Fig. 5C). Although there are clearly other variables  
15 that contribute to the tiller model, the predicted tiller count model generated with spatial-independent  
16 morphological characteristics can serve as a proxy measurement of *Setaria* architecture.

17  
18 The tiller count model was used to predict tiller numbers in *S. viridis*, *S. italica* and the eight *S. viridis* x  
19 *S. italic* RILs under 100% FC and 33% FC water treatments. For all *Setaria* genotypes, tiller number  
20 decreased with reduced water treatment, but the intensity and speed of response varied (Fig. 5C and  
21 Supplemental Fig. S4). *S. viridis* predicted tiller count is greater than *S. italica* under both 100% FC and  
22 33% FC water conditions (Fig. 5D), which is consistent with known architectural differences between the  
23 two genotypes (Darmency et al., 1987; Doust and Kellogg, 2006; Doust, 2007b; Doust et al., 2009).  
24 Predicted tiller number appears to be an introgressive trait for the RILs analyzed (Supplemental Fig. S4)  
25 and can be genetically mapped in a larger *S. viridis* x *S. italica* RIL population.

26  
27 **Color analysis:** Color is often used to gauge plant health in response to biotic and abiotic treatment, or to  
28 identify lines that have possible defects in pigment development and thus changes in photosynthesis  
29 (Albrecht-Borth et al., 2013; Kremnev and Strand, 2014; Kunz et al., 2014; Satou et al., 2014). For each  
30 identified plant pixel, PlantCV records intensity values of color data for each of the three color channels  
31 of RGB (Red, Green, Blue), HSV (Hue, Saturation, Value), and LAB (Lightness, Green-Magenta, Blue-  
32 Yellow) color space, yielding a set of three histograms for each color space. The RGB data is presented  
33 here because images are captured in RGB color space. Results from interpolated HSV and LAB color

space are included in Supplemental Data, but interpretations of these results concur with RGB color space (Supplemental Figs. S5–S8). The VIS images obtained via PhenoFront are 8-bit color images, therefore each pixel in each channel can have a maximum intensity value of  $2^8$  or 256. PlantCV was used to normalize the *Setaria* color histograms by the number of plant pixels (plant size), and these values are averaged by the four side-view angles.

6

Color data was evaluated to determine if water treatment effects could be distinguished 25 to 26 DAP in *S. viridis*, where there is a significant difference in biomass under reduced water (Supplemental Figs. S5 and S6). Principal Component Analysis (PCA) visibly separated full water from reduced water treatment along PC2, which captures 16% of variation (Supplemental Figs. S5 and S6). Spearman correlation analysis found that PC2 was significantly negatively correlated ( $p < 0.01$ ; rho=-0.7056) with water treatment 25 to 26 DAP in *S. viridis*. Principal component regression found that a two-component model of PC1 and PC2 covers 62% of the variation seen in treatment 25 to 26 DAP (Supplemental Figs. S5 and S6). Therefore, *Setaria* color can be used to distinguish water treatments 25 to 26 DAP. PCA using all RGB color data as explanatory variables from *S. viridis* 17 and 18 DAP did not visibly separate full water from 33% FC (Supplemental Figs. S5 and S6) and a two-component regression model included only 0.45% of the variance explained by treatment. Consequently, *Setaria* color does not appear to be an earlier indicator of the effects of drought than biomass.

19

Color analysis was able to discriminate plant genotypes before treatment was applied, which was not the case for estimated plant biomass (Supplemental Fig. S3). PCA of RGB color data prior to water treatment (11 to 12 DAP) found that PC1 captures 41% of variation in color, and separates RIL161 (Fig. 6A) from the other genotypes. The RIL161 line was distinguishable from the other *Setaria* lines and appears to be pale in pigmentation (Fig. 6B). The seven other RIL lines (Fig. 6A) grouped along PC1 and PC2 with the parent lines, *S. viridis* (Fig. 6A, left) and the majority of *S. italica* (Fig. 6A, right). 17 out of 112 *S. italica* plants clustered with RIL161 plants (Fig. 6A, right). Investigation of *S. italica* plants that clustered with the RIL161 plants revealed that they are paler in color compared to *S. italica* that group with *S. viridis* and the other RILs (Fig. 6B). The color PC1 trait appears to be a transgressively segregating phenotype in RIL161 compared to the majority of parent phenotypes and may be genetically mapped in the larger *S. viridis* x *S. italica* RIL population. The color PC1 trait may also be associated with a photosynthetic trait, since RIL161 plants are not only pale but also shorter on average compared to other lines (Supplemental Fig. S1).

33

#### 34 **PSII Image Traits**

1 For more information on PSII image processing please refer to Supplemental Data. Comparison with  
2 plant architecture traits from the VIS system revealed that  $F_v/F_m$  is strongly correlated with plant height in  
3 a Spearman correlation analysis ( $\rho=-0.852$ ,  $p < 0.01$ ; Supplemental Fig. S9). To further test if plant  
4 height significantly influences  $F_v/F_m$  measurements, a 3D-printed staggered platform was designed and  
5 built to examine  $\sim 4 \text{ cm}^2$  of excised *Nicotiana benthamiana* leaf tissue at fixed heights (Supplemental Fig.  
6 S9B and C). *N. benthamiana* was used because leaves are large enough for tissue from a single leaf to be  
7 used at all four platform levels. PSII analysis of the 3D-printed platform confirmed that height  
8 significantly negatively correlates with  $F_v/F_m$  ( $r=-0.976$ ;  $p < 0.01$ ; Supplemental Fig. S9D). One  
9 explanation for this correlation is that  $F_{\min}$  may be disproportionately underestimated relative to  $F_{\max}$  in  
10 shorter plants resulting in an artificially inflated  $F_v/F_m$  value. Although PlantCV is still a useful tool for  
11 analyzing photosynthetic efficiency, alterations in the physical configuration of the PSII imaging station  
12 are necessary to prevent height differences from confounding photosynthetic efficiency measurements.

13

#### 14 **NIR Image Processing and Traits**

15 The NIR imaging station captures near-infrared light (900–1700 nm) using two cameras (320 x 256 pixel)  
16 mounted for top- and side-view imaging with a rotating lifter to allow for multiple side-view images. Four  
17 side-view (0°, 90°, 180°, and 270°) NIR images were captured per plant per time point. The NIR top-view  
18 camera was partially functioning during the experiment, so top-view images were not taken at all time  
19 points. Additionally, the background signal of soil in the top-view NIR images makes plant signal  
20 isolation difficult, so only side-view NIR images were used in the analysis. In total, *Setaria* plants were  
21 imaged 6,399 times from 11–33 DAP for a total of 25,596 NIR side-view images.

22

23 PlantCV isolation of plant material from other background components in NIR grayscale images is  
24 achieved through several image-processing techniques, including background estimation/removal and  
25 object of interest sharpening to improve the effectiveness of binary thresholding. Once the plant is  
26 identified, the grayscale NIR pixel-level intensity is summarized into a histogram containing 256 bins,  
27 where each bin contains the proportion of plant pixels exhibiting the corresponding grayscale intensity  
28 value. Shape attributes described above for VIS imaging are also recorded for each NIR image for quality  
29 control purposes. Comparison of shape measurements between imaging stations allows for the detection  
30 of outliers due to artifacts in the image processing pipelines. Similarity of NIR signal within snapshot sets  
31 and image classes (genotype by treatment by DAP) using Chi-squared and Earth Mover's Distances were  
32 also used to assess data quality. For more information on NIR image processing please refer to  
33 Supplemental Data.

34

1   **NIR signal analysis:** The effects of camera zoom level can be clearly observed as the camera imaging  
2 configuration changes from 3.1X to 1.4X optical zoom at 21 DAP (Fig. 7A). Comparison of 100% and  
3 33% FC water treatments as the mathematical difference of treatment-averaged histograms illustrates a  
4 shift in the distribution of NIR signal histogram toward increased NIR signal reflectance in drought  
5 treated plants (Fig. 7B).

6

7   PCA was used to summarize the changes observed in the NIR signal histograms. Due to the shifts in the  
8 NIR signal between zoom levels, PCA was done within each zoom level. While the first principal  
9 components (PC1s) at each zoom level are not directly comparable, the loadings indicate that they are  
10 driven by the same phenomenon: the two sides of the loading histograms with opposite signs indicate a  
11 shift from higher absorbance pixels, which have been associated with higher water content in previous  
12 studies, to lower absorbance pixels (Fig. 7C and D) (Carter, 1991; Peñuelas and Filella, 1998; Seelig et  
13 al., 2008; Seelig et al., 2009). PC1 captures a larger proportion of variation (33.5%) within NIR signals at  
14 widest zoom level (1.4X), corresponding to later growth stages and more extended drought treatment,  
15 than PC1s from earlier, narrower zoom levels (14.9% and 25.9% at 3.9X and 3.1X, respectively).

16   Significant differences of PC1 values ( $n=8$ ;  $p < 0.01$ ) compared between treatments (100% and 33% FC)  
17 were observed in both *S. viridis* and *S. italica* parental genotypes 23 DAP (Fig. 8). The NIR system  
18 measures wavelengths of light that are absorbed by water, but signal transmittance and reflectance can  
19 also change with differences in leaf thickness or tissue composition (Seelig et al., 2008; Seelig et al.,  
20 2009). Therefore, shifts in NIR signal due to differences in leaf water content cannot be distinguished  
21 from changes in leaf thickness or tissue composition in response to water deficit. More advanced imaging  
22 analysis, perhaps integrating VIS data, or additional imaging hardware might be utilized to distinguish  
23 between these two water-deficit responses.

24

## 25   **Integrated Analysis**

26

27   While each extracted trait can be treated individually, multimodal traits can be used to detect artifacts or  
28 gain greater insight to biological responses. For example, correlations between plant height and  $F_v/F_m$   
29 signal were used to identify the physical limitations of the PSII imaging station, and examining biomass  
30 with watering data expanded understanding of water use efficiency. Examining correlation between traits  
31 may identify simple biometric relationships inherent to a biological system (such as the positive  
32 correlation between plant biomass and height; Supplemental Fig. S10) or more intricate relationships that  
33 are dependent on treatment or temporal factors (relationship between NIR PC1 and VIS color PC1 and  
34 PC2; Supplemental Fig. S10). An important component of trait integration is an understanding of which

1 traits are precisely reporting on genotype or treatment effects. On four days that span the experimental  
2 period, variance due to genotype, treatment, and genotype by treatment across the traits discussed was  
3 examined using a simple linear model (Table I). Emphasizing the temporal dependence of the measured  
4 traits, there were notable shifts in which factor accounted for the largest components of variation at  
5 different time-points. In the six days from 19–25 DAP, as the cumulative water deficit became more  
6 severe, the treatment term accounted for a much larger portion of the biomass, tiller count, and NIR PC1  
7 variance, a trend that continues through 31 DAP. For WUE, treatment had the largest effect 19 DAP,  
8 reflecting the dynamic and differing response to water availability between *S. viridis* and *S. italica* (Fig.  
9 4). Temporal measurements allowed the identification of mechanistic differences in drought response,  
10 which would be indistinguishable with only end-point measurements.

11  
12 Genotype accounted for at least 17.9% of the variance at one or more time-point for all of the traits  
13 described in Table I, including PCA traits from color and NIR signal, suggesting that these traits will be  
14 tractable targets for temporal genetic and gene by environment analysis. Combined with the variance  
15 accounted for by treatment, it is clear that traits measured by the VIS and NIR systems can be used to  
16 investigate plant responses to a changing environment. An outstanding question is whether the distinct  
17 image types are contributing independent insights to an experiment. While analysis of eight RIL lines was  
18 sufficient to identify distinct responses in the two parents for multiple traits, it is not enough to determine  
19 if the traits are correlated or segregating independently from each other. Analysis of larger genetic  
20 populations is necessary to identify the underlying genetic loci for each trait and to determine if they  
21 segregate independently.

22  
23 It is important to note that the analysis routines presented here are just a starting point for the data that can  
24 be extracted from the images collected. The most significant finding from the analysis is that wild *S.*  
25 *viridis* and domesticated *S. italica* have different responses to water availability. This result comes from  
26 analysis of pixel-estimated biomass that does not require integration of plant architecture for accurate  
27 measure. However, integrating plant architecture, by including identification of leaf and stem pixels as  
28 well as segmenting the plant by developmental age, could extract further information from VIS and NIR  
29 images. Making the images and PlantCV software publically available is intended to encourage other  
30 researchers to join in on addressing these questions.

31  
32 **CONCLUSIONS**  
33

1 There are three essential components of the Bellwether Phenotyping Platform: tightly-controlled and  
2 recorded environmental variables, automated image capture, and scriptable analysis by PlantCV. With all  
3 of the components in place, experiments can be conducted on a scale not previously possible. For  
4 example, without the four automatic watering stations in the Bellwether Phenotyping Platform,  
5 maintaining the moisture level of 1,140 pots in a high-light, high-temperature environment would require  
6 considerable labor for the length of the experiment. The design decisions that led to the high capacity of  
7 the Bellwether Phenotyping Platform permit repeated measurements of hundreds of lines with replicates  
8 and multiple treatments. Analysis by PlantCV allows for flexible and timely image processing and  
9 analysis. Open-source PlantCV software is platform independent, allowing it to scale from controlled-  
10 environment growth chamber phenotyping to field phenotyping. PlantCV is currently used to process  
11 images from four diverse custom-built imaging platforms at the Donald Danforth Plant Science Center.  
12 The image dataset provided by this study can be utilized to extend PlantCV function by extracting further  
13 architectural traits that are also important to crops with similar structure, such as maize and sorghum.

14

15 In this study, four different water regimes were imposed with multiple replicates for each line/treatment  
16 combination. Although only ten *Setaria* lines were examined in the present study, this work showed that  
17 *S. viridis* and *S. italica* have fundamentally different responses to water availability. The wild accession *S.*  
18 *viridis*, adjusts its growth to utilize all of the available water, while the domesticated *S. italica* is less  
19 efficient at converting water to biomass under high-water availability. Although, the domesticated crop *S.*  
20 *italica* is known to be drought-tolerant from previous studies (Li and Brutnell, 2011), differences between  
21 *S. viridis* and *S. italica* response to water availability would not have been detected if only end-point  
22 biomass or WUE were examined, highlighting the importance of temporal examination of traits. The eight  
23 randomly selected RILs displayed variation between the parents suggesting, unsurprisingly, that the  
24 observed traits are controlled by multiple loci. Therefore, future phenotyping studies using the Bellwether  
25 Phenotyping Platform that examine a larger population of *S. viridis* x *S. italica* RILs will likely be  
26 successful in extracting multiple time-dependent drought-related QTL.

27

## 28 MATERIALS AND METHODS

29

### 30 Plant Growth

31

32 ***Setaria* experiment growth details:** 4-inch diameter white/gray pots from Hummert were prefilled with  
33 ~473 ml of MetroMix360 soil, and 0.5 g of Osmocote Classic 14-14-14 fertilizer (Everris, US) was  
34 manually distributed to the top of each pot. *S. viridis* (A10), *S. italica* (B10) and 8 RILs (RIL020,

1 RIL070, RIL098, RIL102, RIL128, RIL133, RIL161, RIL187) from the *S. viridis* x *S. italica* population  
2 (Devos et al., 1998; Wang et al., 1998; Bennetzen et al., 2012) were planted, barcoded, then allowed to  
3 germinate for 9 days in a Conviron growth chamber before being loaded into the Conviron growth area of  
4 the Bellwether Phenotyping Platform. Barcoded information included genotype identification, water  
5 treatment group, and a unique pot identification number. During germination, plants were grown under a  
6 long day photoperiod (16h day/8h night) at 31°C day/21°C night. During the germination period the  
7 maximum light intensity of the Conviron chamber was used (230  $\mu\text{mol}/\text{m}^2/\text{s}$ ). At 9 DAP germinated  
8 plants were loaded into the Conviron growth chamber of the Bellwether Phenotyping Platform using a  
9 ‘random block’ design. In the phenotyping facility, plants were grown under a long day photoperiod (16h  
10 day/8h night) at 31°C day/21°C night at a light intensity of 500  $\mu\text{mol}/\text{m}^2/\text{s}$  using metal halide and high-  
11 pressure sodium lamps to deliver a broad spectrum of light for plant growth: 20.7% at 400-500 nm,  
12 33.3% at 500-600 nm, 26.8% at 600-700 nm, 3.3% at 700-800 nm and 15.8% at 800-900 nm (measured  
13 with Apogee Instruments Spectroradiometer PS-100).

14

15 **Water treatments:** The 4-inch pre-filled pots were dried down to remove excess moisture from the soil.  
16 After dry down, the average dry pot, soil, and Osmocote weight was 73 g. To determine soil volume  
17 water content, measured amounts of reverse osmosis water were added to the dry soil and allowed to  
18 absorb for 1 hour. Volume water content (VWC) was determined by a Decagon Devices GS3 soil  
19 moisture sensor and ProCheck Instantaneous data-logger. The full-capacity water treatment group was  
20 assigned a target weight of 625 g. Given a carrier weight of 335 g and soil/pot dry weight of 73 g, the full  
21 capacity treatment group corresponds to 217 ml of available water and a soil volume water content of  
22 ~48%. Initially, five treatment group target weights were defined as 100%, 75%, 50%, 25%, and 0% of  
23 the soil/pot wet weight, which corresponded to the target weights 625 g, 552.5 g, 480 g, 407.5 g, and 0 g.  
24 However, accounting for only relative water availability the treatments were 100% FC (217 ml; 48%  
25 VWC), 66.5% FC (144.5 ml; 31% VWC), 33% FC (72 ml; 14% VWC), and 0 ml for the remaining  
26 groups.

27

## 28 **Hardware**

29

30 **Conviron growth chamber specifications:** The 70  $\text{m}^2$  growth house combines a large growth area with  
31 the precise climate control. The growth house temperature can be controlled from 12°C – 35°C with lights  
32 off, to 15°C – 42°C with lights on while relative humidity is controlled by external chemical dryers. Metal  
33 halide and high-pressure sodium lamps deliver a broad spectrum of light for plant growth with an  
34 intensity of 0 – 800  $\mu\text{moles}/\text{m}^2/\text{s}$ .

1  
2     **VIS camera:** VIS imaging (400 – 700 nm) in the Bellwether Phenotyping Platform utilizes two 2/3"  
3     Progressive Scan CCD cameras with a frame-rate of 17 frame/s, and image resolution of 2454x2053  
4     pixels. The optics of the side and top-view VIS cameras feature a 3 motor lens, focal length of 12.5-  
5     75mm, iris speed of 4.0s, focus motor speed of 5.0s, zoom speed of 5.0s, and horizontal angle view of  
6     6.7° to 38.8°. For the *Setaria* experiment imaging configurations please refer to Supplemental Table SII  
7     for LemnaTec hardware control software settings.

8  
9     **PSII camera:** PSII imaging (680-900nm) in the Bellwether Phenotyping Platform uses one top-view  
10    CCD camera, with a frame-rate of 15 frames/s, image resolution of 1388x964 pixels, and field of view of  
11    120X170mm. The red light pulse (630nm) can vary in intensity from 1000 to 6000  $\mu\text{mol}/\text{m}^2/\text{s}$ . Multiple  
12    top-view images are taken 30-500  $\mu\text{s}$  from the initiation of the 1-2s light pulse. The null ( $F_0$ ), minimum  
13    ( $F_{\min}$ ) and maximum ( $F_{\max}$ ) fluorescence images are saved. For the *Setaria* experiment imaging  
14    configurations please refer to Supplemental Table SII for LemnaTec hardware control software settings.

15  
16    **NIR camera:** NIR imaging (900 to 1700 nm) features two NIR-300/F model cameras with InGaAs  
17    sensors. NIR camera image resolution is 320x256 pixels. For the *Setaria* experiment imaging  
18    configurations please refer to Supplemental Table SII for LemnaTec hardware control software settings.

19  
20 **PlantCV Image Processing and Data Capture**

21  
22    For a list of current measurements extracted by PlantCV please refer to Supplemental Table SIII. For  
23    more information on image processing pipelines please refer to Supplemental Methods and for an  
24    example pipeline workflow see Supplemental Fig. S11.

25  
26    **VIS images:** For *Setaria*, four levels of optical zoom were used to capture VIS side-view images (1.4X,  
27    1.6X, 3.1X and 3.9X zoom). The same image processing strategy was applied to all zoom levels, except  
28    that with 1.3X and 1.5X zoom configurations black background sidebars become visible in images, which  
29    requires extra steps for processing. For VIS top-view images six levels of zoom were used (1.0X, 1.3X,  
30    2.3X, 3.1X, 3.5X, and 3.9X zoom) and the image processing strategy was generally the same as side-view  
31    images except that a pre-made mask was applied to remove brass pieces of the lifter that were visible  
32    from the top-view. A Perl script parallelized the execution of PlantCV image analysis pipelines over a set  
33    of images and also wrote resulting shape and color data to an SQLite database. Zoom-correction of shape  
34    traits, calculation of multivariate traits (e.g. height-width ratio), and size normalization of plant color

1 histograms, was executed with secondary Python analysis scripts over the SQLite database (Supplemental  
2 Table SIV).

3

4 **PSII images:** For the *Setaria* experiment, PSII images were captured at two lifter positions (LemnaTec  
5 lifter positions 630 and 500). The same image processing strategy was applied to both lifter positions. For  
6 more information on PSII image processing please refer to Supplemental Methods. A Perl script  
7 parallelized the execution of PlantCV image analysis pipelines over a set of images and also wrote the  
8 resulting PSII signal data to an SQLite database. Size normalization of PSII histograms was executed  
9 with secondary Python analysis scripts over the SQLite database (Supplemental Table SIV).

10

11 **NIR images:** For the *Setaria* experiment, NIR images were captured at 3.9X, 3.1X and 1.4X zoom. The  
12 same general image processing strategy was applied to all zoom levels. For more information on the NIR  
13 image processing pipelines used please refer to the Supplemental Methods. A Perl script parallelized the  
14 execution of PlantCV image analysis pipelines over a set of images and also wrote the resulting NIR  
15 signal data to an SQLite database. Size normalization of NIR histograms was executed with secondary  
16 Python analysis scripts over the SQLite database (Supplemental Table SIV).

17

## 18 **Statistical Analysis and Modeling**

19

20 All statistical analyses were done in R (R Core Team, 2014). Additional packages used include: analogue  
21 (Simpson, 2007), car (Fox and Weisberg, 2011), emdist (Urbanek and Rubner, 2012), ggplot2 (Wickham,  
22 2009), grid (R Core Team, 2014), gridExtra (Auguie, 2012), lattice (Sarkar, 2008), lme4 (Bates et al.,  
23 2014), lubridate (Grolemund and Wickham, 2011), MASS (Venables and Ripley, 2002), mvtnorm (Genz  
24 and Bretz, 2009; Genz et al., 2014) nlme (Pinheiro et al., 2014) pls (Mevik et al., 2013), plyr (Wickham,  
25 2011), qvalue (Dabney and Storey), rgl (Adler et al., 2014), scales (Wickham, 2014), and VennDiagram  
26 (Chen, 2014). R scripts for all analyses are provided in Supplemental Table SIV.

27

28 **JPEG versus PNG format images:** The image datasets generated during an experiment have a relatively  
29 large file storage footprint. An experimental image set extracted from the database with PhenoFront in  
30 Portable Network Graphics (PNG) format with lossless compression is 200–300 gigabytes for a typical 4–  
31 5 week experiment. To reduce the image storage footprint with a minimal loss of information content,  
32 PhenoFront has the option to extract images in Joint Photographic Experts Group (JPEG) format with  
33 lossy compression, resulting in an approximately 11-fold reduction in storage size. Individual image  
34 compression ratios vary from ~7:1 to 20:1 depending on zoom level. The dataset presented here was

1 analyzed in both PNG and JPEG formats and a subset of plant phenotypes were compared to quantify  
2 discrepancies between the image formats. Pearson's correlation between JPEG and PNG traits extracted  
3 by PlantCV was calculated using the cor.test function in the R stats package. Plant shape-based traits  
4 (side-view area, top-view area and plant height from VIS images) were highly correlated ( $r > 0.997$ ),  
5 however the 2<sup>nd</sup> and 3<sup>rd</sup> PCs from RGB color analysis were less linear and less correlated ( $r = -0.893$  and -  
6 0.781, respectively; Supplemental Fig. S12). Therefore, JPEG images are sufficient for analysis when  
7 only shape and morphological features are needed (e.g. real-time image configuration feed-back), but  
8 images with lossless compression are preferable when signal/spectral data is analyzed.  
9

10 **Camera zoom correction:** To compare digital measurements across camera zoom levels, scaling factors  
11 for pixel area and length were calculated by imaging a 13.2 x 13.2 x 3.7 cm box at a series of zoom levels  
12 in the VIS imaging cabinet (Supplemental Fig. S13). For images at each zoom level, box pixel area and  
13 width was measured with ImageJ (Abràmoff et al., 2004). Pixel area was scaled relative to the 1X zoom  
14 level area and pixel length was scaled relative to actual box dimensions. For both relative area and  
15 pixels/cm scaling factors, regression analysis was done to estimate a function that would predict scaling  
16 factors for each zoom level. Non-linear least squares regression was done to fit an exponential model and  
17 ordinary least squares regression was done to fit a 2<sup>nd</sup> order polynomial model to both relative area and  
18 pixels/cm data sets. For each scaling factor, the model that minimized the Akaike's Information Criterion  
19 (Bozdogan, 1987) was chosen. For relative pixel area, the resulting model was an exponential function:

$$A_r = 0.8854e^{0.0007224z}$$

20 where  $A_r$  is relative area and  $z$  is the VIS camera zoom setting (Supplemental Fig. S13A). For pixel/cm,  
21 the resulting model was a 2nd-order polynomial function:

$$s = 14.27 + 0.002077z + 0.00000217z^2$$

22  $s$  is the px/cm scaling factor and  $z$  is the VIS camera zoom setting (Supplemental Fig. S13B).  
23

24 **Height modeling:** To determine the accuracy of PlantCV-estimated plant height, a validation set was  
25 created by manually measuring height for 173 randomly selected VIS side-view images from all zoom  
26 levels using ImageJ (Abràmoff et al., 2004) where height in cm was calibrated using the width of the  
27 plant carrier (12 cm wide). The manual height measurement was drawn as a straight vertical line from the  
28 top of the pot to the highest point of the plant and was an average of three independent measurements.  
29 Ordinary least squares regression was used to evaluate how well PlantCV estimated manually measured  
30 height (Fig. 3A; Supplemental Table SIV).

31

1   **Biomass modeling:** To estimate fresh- and dry-weight above ground biomass with PlantCV, a training set  
2   was created by harvesting the above ground tissue of 41 randomly selected *S. viridis* and *S. italica* plants  
3   from the 100% FC group after imaging throughout the experimental period. After collection, fresh-weight  
4   measurements were recorded and plant tissue was placed in polypropylene micro-perforated bags (PJP  
5   MarketPlace #361001) and dried for ~3 days at 60°C before dry weight was measured. To generate a  
6   model for predicting fresh- and dry-weight biomass, multiple linear regression analysis was done,  
7   followed by model selection, to determine the best predictor variables of biomass. The initial full model  
8   for fresh-weight biomass included terms for total VIS side-view pixel area, VIS top-view pixel area, plant  
9   height, and all interactions. A stepwise procedure was used to test the effect of dropping terms from the  
10   full model using the Akaike's Information Criterion (Bozdogan, 1987). The initial reduced model had  
11   coefficients that were not significant, so a second reduced model that had total VIS side-view area as the  
12   only explanatory variable was generated. Analysis of variance was done to compare the reduced models  
13   and indicated that the residual error for the models was not significantly different, so the smaller model  
14   was kept (see Results and Discussion; Fig. 4A). Total VIS side-view was used to predict dry weight  
15   without model selection (Supplemental Fig. S2):

$$M_{dw} = 4.634 \times 10^{-6} A_{sv} - 5.295 \times 10^{-2}$$

16   For both fresh- and dry-weight models, adding an indicator variable for plant out-of-frame status  
17   improved the models, but adding an indicator variable for genotype (*S. viridis* or *S. italica*) did not  
18   improve the models.

19  
20   **Growth-rate modeling:** *Setaria* plant growth patterns appeared to be asymptotic. Non-linear least squares  
21   regression analysis was used to model a three-component logistic growth function for each group  
22   (genotype by treatment) (Paine et al., 2012). Absolute growth rate over time (first derivative of the growth  
23   function) was computed for each group (Paine et al., 2012). The growth models were used to estimate the  
24   maximum absolute growth rate and the day it was reached for each group.

25  
26   **Plant architecture modeling:** For 58 randomly selected VIS side-view images (25 and 26 DAP), tillers  
27   were counted and ImageJ (Abràmoff et al., 2004) was used to manually measure tiller-angle and leaf  
28   angle. Tiller angle was defined by the angle between two outermost tillers. Leaf angle was measured as  
29   the average between first leaf angles of the two outermost tillers. All measurements were performed at  
30   least three times and the average values were used for further analysis. The averaged results were  
31   compared to height-width ratio (HW, height divided by extent-x) and solidity (plant area divided by  
32   convex hull area) by cor.test function in R. HW was significantly correlated ( $p < 0.01$ ) with all three

1 manually measured traits. A stepwise selection by Akaike's Information Criterion (Bozdogan, 1987) was  
2 then used to generate a model that best explains HW:

3 
$$HW = -0.017599A_{leaf} - 0.01298A_{tiller} + 2.696608$$

4 where HW is height-width ratio,  $A_{leaf}$  is leaf angle and  $A_{tiller}$  is tiller angle.

5

6 To generate the tiller count model, tillers were counted from 195 randomly selected VIS images. The  
7 number of tillers was counted at least three times and the average values were used for further analysis.  
8 Tiller number was correlated with biomass, height-width ratio, solidity, extent-x, and height. Height-  
9 width ratio was calculated based on zoom-corrected plant height and extent-x values generated by  
10 PlantCV. A stepwise selection by Akaike's Information Criterion (Bozdogan, 1987) was then used to  
11 generate a model that best explains tiller count. Tillers were counted for second set of 660 images and the  
12 predictive power of the tiller count model was assessed.

13

14 **Color analysis:** 17 to 18 DAP and 25 to 26 DAP, histograms of red, green, and blue color signal were  
15 extracted by PlantCV from *S. viridis* under 100% and 33% FC. Size-normalized data from each angle was  
16 averaged by PlantCV. Histogram values that were zero in all samples were removed, and the remaining  
17 data was used as explanatory variables in PCA. Principle component analysis was calculated using the  
18 prcomp function in the R software stats package. The principal component regression function pcr from  
19 the pls R package, regressed treatment group over the principal components of color data 17 to 18 DAP  
20 and 25 to 26 DAP.

21

22 Color data was also assessed to see if *Setaria* genotypes could be separated before water treatment was  
23 applied. Histograms of red, green, and blue color signal were extracted by PlantCV from all ten *Setaria*  
24 genotypes under 100% and 33% FC 11 to 12 DAP. Size-normalized data from each angle was averaged  
25 by PlantCV. Histogram values that were zero in all samples were removed, and the remaining data was  
26 used as explanatory variables in PCA.

27

28  **$F_v/F_m$  signal analysis:** An area-normalized histogram of  $F_v/F_m$  signal was calculated from PSII top-view  
29 images using PlantCV. The Spearman correlation between median  $F_v/F_m$  signal value and estimated plant  
30 height extracted by PlantCV was calculated with the cor.test function in the R stats package. To further  
31 test the correlation between  $F_v/F_m$  signal and height a 3D-printed 'plant' was made. 5 cm tall rectangular  
32 'stem' and flat 'leaf' platform pieces were designed using openSCAD (<http://www.openscad.org/>) and 3D  
33 printed using a Stratasys Fortus 250mc and ABSplus plastic. Four stem and leaf pieces were assembled  
34 using Oatey all-purpose cement, so that each platform was visible from a top view angle. *N. benthamiana*

1 leaf tissue was excised and segmented into four ~4 cm<sup>2</sup> samples (avoiding the leaf mid-vein). Elmer's  
2 multipurpose spray adhesive was used to make platform pieces tacky and leaf samples were adhered to  
3 each platform level. The bottom stem piece was positioned in a pot filled with blue non-fluorescent gravel  
4 so the first leaf piece was flush with the gravel (referred to as 0 cm; Supplemental Fig. S9C). 3D plant  
5 was PSII imaged and the process was repeated with a second *N. benthamiana* leaf. Resulting PSII images  
6 were analyzed with PlantCV with masks to segment each platform level. Pearson's correlation between  
7 leaf height and median F<sub>v</sub>/F<sub>m</sub> was calculated using the cor.test function in the R stats package. Designs for  
8 3D plant model pieces can be found at <https://www.thingiverse.com/thing:574287> and paths to analysis  
9 scripts can be found in Supplemental Table SIV.

10  
11 **NIR signal analysis:** Principle components were calculated using the prcomp function encoded by the R  
12 software package. PCA was performed within each zoom level for *Setaria* with full and 33% FC water  
13 treatments. The resulting PC scores for each of the four side view images were averaged. Significance  
14 testing between treatments over time was applied to a rolling two day time period because plants were  
15 imaged over two days. The number of replicates is not identical across genotypes and days. Therefore the  
16 maximum equal number of individuals was chosen randomly from the larger set and used as a  
17 representative sample for hypothesis testing. For each genotype on each experimental day, the  
18 distributions of phenotypic values between treatments (100% FC versus. 33% FC) were compared using  
19 the t.test function in R.

20  
21 **Integrated trait analysis:** The contribution of experimental factors to phenotypic variance was assessed  
22 using the lmer function encoded by the lme4 library in R (Bates et al., 2014). Variance partitioning was  
23 estimated by linear modeling of phenotypic values as a function of genotype, treatment (100% and 33%  
24 FC) and the interaction of genotype and treatment (specified as random effects) at four time-points (15-  
25 16, 19-20, 25-26, 31-32 DAP). Variance estimates associated with each factor were extracted from the  
26 resultant merMod model object using the VarCorr function and divided by total phenotypic variance to  
27 estimate proportion of variance explained by each experimental factor. Trait correlation was assessed at  
28 the same four time points and treatments using the cor function in R.

29

## 30 **SUPPLEMENTAL DATA**

31

32 The following materials are available in the online version of this article.

33

34 **Supplemental Methods.** Extended image-processing methods.

- 1   **Supplemental Figure S1.** Plant height estimated with PlantCV for ten *Setaria* genotypes.  
2  
3   **Supplemental Figure S2.** Above ground dry-weight biomass modeled using shoot and leaf pixel area  
4   from four side-view VIS images.  
5  
6   **Supplemental Figure S3.** Modeled fresh-weight biomass estimated with PlantCV for ten *Setaria*  
7   genotypes.  
8  
9   **Supplemental Figure S4.** Tiller count estimated with PlantCV for ten *Setaria* genotypes.  
10  
11   **Supplemental Figure S5.** Color can separate watering treatments 25 to 26 DAP, but not 17 to 18 DAP in  
12   RGB, HSV and LAB color space.  
13  
14   **Supplemental Figure S6.** Loadings of color data 25 to 26 DAP and 17 to 18 DAP.  
15  
16   **Supplemental Figure S7.** Before water treatment is applied (11 to 12 DAP) color can distinguish *Setaria*  
17   genotypes in RGB color (Fig. 6), which is also consistent in HSV and LAB color space.  
18  
19   **Supplemental Figure S8.** Loadings of color data 11 to 12 DAP.  
20  
21   **Supplemental Figure S9.** Height significantly correlates with median  $F_v/F_m$ .  
22  
23   **Supplemental Figure S10.** Matrices displaying Pearson's Correlation Coefficient calculated between  
24   traits at four roughly equidistant time points.  
25  
26   **Supplemental Figure S11.** A visual illustration of the pipeline used to threshold plant tissue from  
27   background within grayscale NIR images.  
28  
29   **Supplemental Figure S12.** JPEG images are sufficient for analysis when only shape and morphological  
30   features are needed.  
31  
32   **Supplemental Figure S13.** Zoom-correction scaling factors for pixel area and pixel length.  
33  
34   **Supplemental Table SI.** Comparison of image-analysis software.

1  
2   **Supplemental Table SII.** Camera configuration parameters.  
3  
4   **Supplemental Table SIII.** Description of measurements output by PlantCV.  
5  
6   **Supplemental Table SIV.** PlantCV image processing and analysis scripts.  
7  
8   **ACKNOWLEDGEMENTS**  
9  
10 We would like to thank the Bellwether Foundation for the generous donation that allowed for the  
11 construction of the Bellwether Phenotyping Platform, DDPSC Facilities for careful maintenance of the  
12 Bellwether Platform, Kevin Nagel for system training, Pu Huang, and Dustin Mayfield-Jones for manual  
13 measurement assistance.

1      **REFERENCES**  
2  
3

4      **Abràmoff MD, Magalhães PJ, Ram SJ** (2004) Image processing with ImageJ. *Biophotonics Int* **11**: 36–  
5      43.

6      **Adler D, Murdoch D** (2014) rgl: 3D visualization device system.  
7

8      **Albrecht-Borth V, Kauss D, Fan D, Hu Y, Collinge D, et al.** (2013) A Novel Proteinase, SNOWY  
9      COTYLEDON4, Is Required for Photosynthetic Acclimation to Higher Light Intensities in *Arabidopsis*.  
10     *Plant Physiol* **163**: 732–745.

11     **Andrade-Sánchez P, Gore MA, Heun JT, Thorp KR, Carmo-Silva AE, et al.** (2014) Development  
12    and evaluation of a field-based high-throughput phenotyping platform. *Funct Plant Biol* **41**: 68-79.  
13

14     **Auguie B** (2012) gridExtra: functions in Grid graphics.  
15

16     **Bates D, Mächler M, Bolker B, Walker S** (2014) Fitting Linear Mixed-Effects Models using lme4.  
17

18     **Benfey PN, Mitchell-Olds T** (2008) From genotype to phenotype: systems biology meets natural  
19    variation. *Science* **320**: 495–7.  
20

21     **Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L,  
22    Vaughn JN, Grimwood J, et al** (2012) Reference genome sequence of the model plant *Setaria*. *Nat  
23    Biotech* **30**: 555–561.  
24

25     **Berger B, Parent B, Tester M** (2010) High-throughput shoot imaging to study drought responses. *J Exp  
26    Bot* **61**: 3519–28.  
27

28     **Bozdogan H** (1987) Model selection and Akaike's Information Criterion (AIC): The general theory and  
29    its analytical extensions. *Psychometrika* **52**: 345–370.  
30

31     **Bradski G** (2000) The opencv library. *Doctor Dobbs Journal* **25**(11): 120-126.  
32

33     **Bray EA** (1997) Plant responses to water deficit. *Trends Plant Sci* **2**: 48–54.  
34

35     **Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X-G, Kellogg E, Van Eck J**  
36    (2010) *Setaria viridis*: a model for C4 photosynthesis. *Plant Cell* **22**: 2537–44.  
37

38     **Carter GA** (1991) Primary and Secondary Effects of Water Content on the Spectral Reflectance of  
39    Leaves. *Am J Bot* **78**: 916.  
40

41     **Chen H** (2014) VennDiagram: Generate high-resolution Venn and Euler plots.  
42

43     **Dabney A, Storey JD** (2015) qvalue: Q-value estimation for false discovery rate control.  
44

45     **Darmency H, Ouin C, Pernes J** (1987) Breeding foxtail millet (*Setaria italica*) for quantitative traits  
46    after interspecific hybridization and polyploidization. *Genome* **29**: 453–456.  
47

48     **Devos KM, Wang ZM, Beales J, Sasaki T, Gale MD** (1998) Comparative genetic maps of foxtail millet  
49    (*Setaria italica*) and rice (*Oryza sativa*). *Theor Appl Genet* **96**: 63–68.  
50

51

- 1   **Doust AN** (2007a) Architectural evolution and its implications for domestication in grasses. *Ann Bot*  
2   **100**: 941–50.
- 3
- 4   **Doust AN** (2007b) Grass architecture: genetic and environmental control of branching. *Curr Opin Plant*  
5   **Biol** **10**: 21–5.
- 6
- 7   **Doust AN, Kellogg EA** (2006) Effect of genotype and environment on branching in weedy green millet  
8   (*Setaria viridis*) and domesticated foxtail millet (*Setaria italica*) (Poaceae). *Mol Ecol* **15**: 1335–1349.
- 9
- 10   **Doust AN, Kellogg EA, Devos KM, Bennetzen JL** (2009) Foxtail Millet: A Sequence-Driven Grass  
11   Model System. *Plant Physiol* **149**: 137–141.
- 12
- 13   **Dudley JT, Butte AJ** (2009) A quick guide for developing effective bioinformatics programming skills.  
14   *PLoS Comput Biol* **5**: e1000589.
- 15
- 16   **Fahlgren N, Gehan MA, Baxter I** (2015) Lights, camera, action: high-throughput plant phenotyping is  
17   ready for a close-up. *Curr Opin Plant Biol* **24**: 93–99.
- 18
- 19   **Fernández-García N, Olmos E, Bardisi E, García-De la Garma J, López-Berenguer C, Rubio-**  
20   **Asensio JS** (2014) Intrinsic water use efficiency controls the adaptation to high salinity in a semi-arid  
21   adapted plant, henna (*Lawsonia inermis L.*). *J Plant Physiol* **171**: 64–75.
- 22
- 23   **Fourment M, Gillings MR** (2008) A comparison of common programming languages used in  
24   bioinformatics. *BMC Bioinformatics* **9**: 82.
- 25
- 26   **Furbank RT, Tester M** (2011) Phenomics--technologies to relieve the phenotyping bottleneck. *Trends*  
27   *Plant Sci* **16**: 635–44.
- 28
- 29   **Garg AK, Kim J-K, Owens TG, Ranwala AP, Choi Y Do, Kochian L V, Wu RJ** (2002) Trehalose  
30   accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proc Natl Acad Sci*  
31   **99**: 15898–15903.
- 32
- 33   **Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y,**  
34   **Gentry J, et al** (2004) Bioconductor: open software development for computational biology and  
35   bioinformatics. *Genome Biol* **5**: R80.
- 36
- 37   **Genz A, Bretz F** (2009) Computation of multivariate normal and t probabilities. Springer Science &  
38   Business Media. Vol: 195.
- 39
- 40   **Genz A, Bretz F, Miwa T, Mi X, Leisch F, Scheipl F, Hothorn T** (2014) mvtnorm: Multivariate  
41   Normal and t Distributions.
- 42
- 43   **Gerland P, Raftery AE, Sevcikova H, Li N, Gu D, et al.** (2014) World population stabilization unlikely  
44   this century. *Science* **346**: 234–237.
- 45
- 46   **Gilbert D** (2004) Bioinformatics software resources. *Brief Bioinform* **5**: 300–304.
- 47
- 48   **Golzarian MR, Frick RA, Rajendran K, Berger B, Roy S, Tester M, Lun DS** (2011) Accurate  
49   inference of shoot biomass from high-throughput images of cereal plants. *Plant Methods* **7**: 2.
- 50

- 1   **Grolemund G, Wickham H** (2011) Dates and Times Made Easy with lubridate. *J Stat Softw* **40**: 1–25.
- 2
- 3   **Hartmann A, Czauderna T, Hoffmann R, Stein N, Schreiber F** (2011) HTPheno: an image analysis  
4   pipeline for high-throughput plant phenotyping. *BMC Bioinformatics* **12**: 148.
- 5
- 6   **Honsdorf N, March TJ, Berger B, Tester M, Pillen K** (2014) High-throughput phenotyping to detect  
7   drought tolerance QTL in wild barley introgression lines. *PLoS One* **9**: e97047.
- 8
- 9   **Hunter JD** (2007) Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **9**: 90–  
10   95.
- 11
- 12   **IPCC** (2014) Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral  
13   Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental  
14   Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University  
15   Press.
- 16
- 17   **Kellogg EA** (1999) C4 Plant Biology. *C4 Plant Biol.* 411–444.
- 18
- 19   **Klukas C, Chen D, Pape J-M** (2014) Integrated Analysis Platform: An Open-Source Information  
20   System for High-Throughput Plant Phenotyping. *Plant Physiol* **165**: 506–518.
- 21
- 22   **Kremnev D, Strand Å** (2014) Plastid encoded RNA polymerase activity and expression of  
23   photosynthesis genes required for embryo and seed development in *Arabidopsis*. *Front Plant Sci* **5**: 385.
- 24
- 25   **Kunz H-H, Gierth M, Herdean A, Satoh-Cruz M, Kramer DM, Spetea C, Schroeder JI** (2014)  
26   Plastidial transporters KEA1, -2, and -3 are essential for chloroplast osmoregulation, integrity, and pH  
27   regulation in *Arabidopsis*. *Proc Natl Acad Sci* **111**: 7480–7485.
- 28
- 29   **Kvilekval K, Fedorov D, Obara B, Singh A, Manjunath BS** (2010) Bisque: a platform for bioimage  
30   analysis and management. *Bioinforma* **26**: 544–552.
- 31
- 32   **Lata C, Bhatty S, Bahadur RP, Majee M, Prasad M** (2011) Association of an SNP in a novel DREB2-  
33   like gene SiDREB2 with stress tolerance in foxtail millet (*Setaria italica L.*). *J Exp Bot* **62**: 3387–401.
- 34
- 35   **Lata C, Sahu PP, Prasad M** (2010) Comparative transcriptome analysis of differentially expressed  
36   genes in foxtail millet (*Setaria italica L.*) during dehydration stress. *Biochem Biophys Res Commun* **393**:  
37   720–7.
- 38
- 39   **Li P, Brutnell TP** (2011) *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoideae  
40   grasses. *J Exp Bot* **62**: 3031–7.
- 41
- 42   **Lobet G, Draye X, Périlleux C** (2013) An online database for plant image analysis software tools. *Plant  
43   Methods* **9**: 38.
- 44
- 45   **Mahfoozi S** (2001) Developmental Regulation of Low-temperature Tolerance in Winter Wheat. *Ann Bot*  
46   **87**: 751–757.
- 47
- 48   **Mangalam H** (2002) The Bio\* toolkits -- a brief overview. *Brief Bioinform* **3**: 296–302.
- 49
- 50   **Maxwell K, Johnson GN** (2000) Chlorophyll fluorescence—a practical guide. *J Exp Bot* **51**: 659–668.

- 1 McCouch S, Baute GJ, Bradeen J, Bramel P, Bretting PK, Buckler E, Burke JM, Charest D,  
2 Cloutier S, Cole G, et al (2013) Agriculture: Feeding the future. *Nature* **499**: 23–4.
- 3
- 4 Mevik B-H, Wehrens R, Liland KH (2013) pls: Partial Least Squares and Principal Component  
5 regression.
- 6
- 7 Munns R, Tester M (2008) Mechanisms of salinity tolerance. *Annu Rev Plant Biol* **59**: 651–81.
- 8
- 9 Nelson T, Dengler N (1997) Leaf Vascular Pattern Formation. *Plant Cell* **9**: 1121–1135.
- 10
- 11 Oliphant TE (2007) Python for Scientific Computing. *Computing in Science & Engineering* **9**:10-20.
- 12
- 13 Paine CET, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, Turnbull LA (2012) How to fit  
14 nonlinear plant growth models and calculate growth rates: an update for ecologists. *Methods Ecol Evol* **3**:  
15 245–256.
- 16
- 17 Paproki A, Sirault X, Berry S, Furbank R, Fripp J (2012) A novel mesh processing based technique  
18 for 3D plant analysis. *BMC Plant Biol* **12**: 63.
- 19
- 20 Peñuelas J, Filella I (1998) Visible and near-infrared reflectance techniques for diagnosing plant  
21 physiological status. *Trends Plant Sci* **3**: 151–156
- 22
- 23 Phattaralerphong J, Sinoquet H (2005) A method for 3D reconstruction of tree crown volume from  
24 photographs: assessment with 3D-digitized plants. *Tree Physiol* **25**: 1229–1242.
- 25
- 26 Pinheiro J, Bates D, DebRoy S, Sarkar D, Team TRC (2014) nlme: Linear and Nonlinear Mixed  
27 Effects Models.
- 28
- 29 R Core Team (2014) R: A Language and Environment for Statistical Computing.
- 30
- 31 Ray DK, Mueller ND, West PC, Foley JA (2013) Yield Trends Are Insufficient to Double Global Crop  
32 Production by 2050. *PLoS One* **8**: e66428.
- 33
- 34 Richards R, Thurling N (1978) Variation between and within species of rapeseed (*Brassica campestris*  
35 and *B. napus*) in response to drought stress. I. Sensitivity at different stages of development. *Aust J Agric  
36 Res* **29**: 469.
- 37
- 38 Sage RF, Zhu X-G (2011) Exploiting the engine of C(4) photosynthesis. *J Exp Bot* **62**: 2989–3000.
- 39
- 40 Sarkar D (2008) Lattice: Multivariate Data Visualization with R.
- 41
- 42 Satou M, Enoki H, Oikawa A, Ohta D, Saito K, Hachiya T, Sakakibara H, Kusano M, Fukushima  
43 A, Saito K, et al (2014) Integrated analysis of transcriptome and metabolome of *Arabidopsis* albino or  
44 pale green mutants with disrupted nuclear-encoded chloroplast proteins. *Plant Mol Biol* **85**: 411–428.
- 45
- 46 Schöffl F (1998) Regulation of the Heat-Shock Response. *Plant Physiol* **117**: 1135–1141.
- 47
- 48 Seelig H -D., Hoehn A, Stodieck LS, Klaus DM, Adams III WW, Emery WJ (2008) The assessment  
49 of leaf water content using leaf reflectance ratios in the visible, near-, and short-wave-infrared. *Int J  
50 Remote Sens* **29**: 3701–3713.

- 1 Seelig H-D, Hoehn A, Stodieck LS, Klaus DM, Adams WW, Emery WJ (2009) Plant water  
2 parameters and the remote sensing R 1300/R 1450 leaf water index: controlled condition dynamics during  
3 the development of water deficit stress. *Irrig Sci* **27**: 357–365.
- 4
- 5 Simpson GL (2007) Analogue Methods in Palaeoecology: Using the analogue Package. *J Stat Softw* **22**:  
6 1–29.
- 7
- 8 Talukder S, Babar M, Vijayalakshmi K, Poland J, Prasad P, Bowden R, Fritz A (2014) Mapping  
9 QTL for the traits associated with heat tolerance in wheat (*Triticum aestivum L.*). *BMC Genet* **15**: 97.
- 10
- 11 Urbanek S, Rubner Y (2012) emdist: Earth Mover's Distance.
- 12
- 13 Venables WN, Ripley BD (2002) Modern Applied Statistics with S. Fourth. Springer, New York
- 14
- 15 Wang ZM, Devos KM, Liu CJ, Wang RQ, Gale MD (1998) Construction of RFLP-based maps of  
16 foxtail millet, *Setaria italica (L.) P. Beauv.* *Theor Appl Genet* **96**: 31–36.
- 17
- 18 Warnasooriya SN, Brutnell TP (2014) Enhancing the productivity of grasses under high-density  
19 planting by engineering light responses: from model systems to feedstocks. *J Exp Bot* **65**(11):2825–2834.
- 20
- 21 White JW, Andrade-Sanchez P, Gore MA, Bronson KF, Coffelt TA, Conley MM, Feldmann KA,  
22 French AN, Heun JT, Hunsaker DJ, et al (2012) Field-based phenomics for plant genetics research. *F*  
23 *Crop Res* **133**: 101–112.
- 24
- 25 Wickham H (2009) ggplot2: elegant graphics for data analysis.
- 26
- 27 Wickham H (2011) The Split-Apply-Combine Strategy for Data. *J Stat Softw* **40**: 1–29
- 28
- 29 Wickham H (2014) scales: Scale functions for graphics.
- 30
- 31 Wilkins O, Bräutigam K, Campbell MM (2010) Time of day shapes *Arabidopsis* drought  
32 transcriptomes. *Plant J* **63**: 715–27.
- 33
- 34 Zhang J, Liu T, Fu J, Zhu Y, Jia J, Zheng J, Zhao Y, Zhang Y, Wang G (2007) Construction and  
35 application of EST library from *Setaria italica* in response to dehydration stress. *Genomics* **90**: 121–31.

Variance Type	Biomass	Height	Tiller Count	WUE	RGB.PC1	RGB.PC2	NIR.PC1	DAP
Genotype	14.8	61.3	53.2	20.6	75.1	65.8	38.9	15
Treatment	0	0	0	0	0	0	0	
Genotype x Treatment	5.2	2.4	0	5	0	1.8	2.5	
Residual	80	36.3	46.8	74.5	24.9	32.4	58.6	
Genotype	17.9	22.8	36.5	20.2	62.6	39.4	29.5	19
Treatment	4.8	0.4	2.8	27.8	0	0.3	6.7	
Genotype x Treatment	7.2	5.3	5.2	4.4	0	1	5	
Residual	70.1	71.5	55.5	47.5	37.4	59.4	58.8	
Genotype	14.4	65.4	11.5	22.2	66.3	60.9	30.5	25
Treatment	55.4	1.8	66.9	8.2	13.9	0	56.7	
Genotype x Treatment	4.6	4	6.7	5.6	0.9	7.3	4.1	
Residual	25.5	28.8	14.9	64	18.9	31.8	8.7	
Genotype	2.7	42.8	9.8	7.7	62.5	8.3	8.3	31
Treatment	83.7	7.3	78.9	8	7.7	1	75.5	
Genotype x Treatment	2.1	8	1.9	0	5.6	13.9	3.8	
Residual	11.5	41.9	9.4	84.3	24.1	76.8	12.5	

1

2 **Table I.** Proportion of variance explained by experimental factors. Variance associated with factors  
3 genotype, experimental treatment, the interaction of genotype and treatment was assessed at four time  
4 points (15, 19, 25, 31 DAP). Unexplained variance is denoted as residual variance.

1   **FIGURE LEGENDS**

2

3   **Figure 1.** Bellwether Phenotyping Platform at the Donald Danforth Plant Science Center. A, Diagram of  
4   the Bellwether Phenotyping Platform. B, The volume of water added to *S. viridis* plants during morning  
5   (ZT23) and evening (ZT8) water applications.

6

7   **Figure 2.** PlantCV image analysis software developed at the Donald Danforth Plant Science Center. A,  
8   Automated identification of plant material from background. Plant is identified (blue outline, right) from  
9   the original image (left). B, Example of traits extracted by PlantCV. Shape characteristics such as convex  
10   hull (left), and height (middle) are extracted from visible images. Differences in color can be visualized  
11   by pseudocoloring by HSV, LAB or RGB color channels (right). *S. viridis* is pseudocolored based on the  
12   value channel of HSV color space. C, PlantCV analysis of traits over time under 100% FC and 33% FC  
13   water treatments. Traits such as biomass accumulation, growth-rate, and color can be tracked over time  
14   under experimental conditions, such as drought, from data collected by PlantCV. *S. viridis* is shown  
15   pseudocolored based on the value channel of HSV color space. D, Example of PlantCV extracted traits  
16   from plant species other than *Setaria*. *Brachypodium distachyon*, *Arabidopsis thaliana*, and *Camelina*  
17   *sativa* are pictured from left to right. Plants are pseudocolored based on the value channel of HSV. E,  
18   Example of PlantCV analysis of an image of cassava captured by a cell phone camera. F, Example of  
19   PlantCV analysis of an image of sweet potato captured by a digital camera.

20

21   **Figure 3.** Plant height estimated with PlantCV. A, Plant height estimated with PlantCV compared to  
22   manually measured height. Data shown are for 173 randomly selected VIS side-view images at four  
23   different zoom levels. The ordinary least squares linear model with standard error is plotted. B, Estimated  
24   plant height for *S. viridis* and *S. italica* plants from 11–33 DAP. Plants watered to 100% or 33% FC are  
25   shown. Local regression (LOESS) fitted curves with standard error are plotted for each genotype by water  
26   treatment group. The arrows indicate the day that the 33% FC water treatment started.

27

28   **Figure 4.** *S. viridis* and *S. italica* respond differently to water-limited conditions. A, Above ground fresh-  
29   weight biomass modeled using shoot and leaf pixel area from four side-view VIS images. Data shown are  
30   for 41 plants where destructive fresh-weight biomass was measured throughout the experiment. The  
31   ordinary least squares linear model with standard error is plotted. B–D, data shown are for *S. viridis* and  
32   *S. italica* imaged from 11–33 DAP. Plants watered to 100% or 33% FC are shown. Arrows indicate the  
33   day that the 33% FC water treatment started. B, Modeled fresh-weight biomass. Three-component logistic  
34   growth curves with 95% confidence intervals are plotted. C, Absolute growth rates over time with 95%

1 confidence intervals. D, Water-use efficiency, accumulated fresh-weight biomass divided by cumulative  
2 water added. LOESS fitted curves with standard error are plotted. C and D, see color key in B.

3

4 **Figure 5.** Tiller count modeled from spatial-independent morphological characteristics. A, Examples of  
5 height-width ratio variation from randomly selected images of plants 25–26 DAP. B, Added-variable  
6 plots of fresh-weight biomass and height-width ratio from 195 randomly selected images. C, Relationship  
7 between manually measured tiller count from 646 randomly selected images and model-predicted tiller  
8 count. D, Model-predicted tiller count for *S. viridis* and *S. italica* plants from 11–33 DAP. Plants watered  
9 to 100% or 33% FC are shown. Arrows indicate the day that the 33% FC water treatment started.

10

11 **Figure 6.** Before water treatment is applied (11 to 12 DAP) color can distinguish *Setaria* genotypes. A,  
12 PCA of RGB color for all *Setaria* genotypes before treatment is applied (11 to 12 DAP). PC1 and PC2 are  
13 plotted for *S. viridis* (orange), *S. italica* (green), RIL161 (purple) and 7 other *S. viridis* x *S. italica* RILs  
14 (gray). B, *S. italica* (yellow circle in A; yellow boxed image) found amongst R161 (blue circle in A; blue  
15 boxed image) are more similar in color than *S. italica* (red circle in A; red boxed image) that groups with  
16 other lines (gray). VIS images are pseudocolored based on the value channel in HSV color space.

17

18 **Figure 7.** Measurements derived from the NIR camera system. A, Average NIR signal distribution of  
19 well-watered plants throughout the experiment. Bins corresponding to signal intensity levels 0 – 255 are  
20 plotted along the x-axis, while experimental days are plotted along the y-axis. The proportion of pixels  
21 that fall into each bin are expressed as a grayscale shade (0% black, larger percentages are increasingly  
22 more white). Color bars along the y-axis illustrate zoom range (3.9X = light blue, 3.1X = blue, 1.4X =  
23 dark blue). B, Difference of average NIR signal distribution between plants receiving 100% and 33% FC.  
24 Signal bins shaded blue have an increased proportion of pixels in images of plants receiving 100% FC  
25 water; signal bins shaded red have a greater proportion of pixels in images receiving 33% FC water; white  
26 corresponds to no difference between images. C, A plot of PC1 loadings (3.9X = light blue, 3.1X = blue,  
27 1.4X = dark blue). D, A plot of the average signal difference between water treatment levels 100% and  
28 33% FC (3.9X = light blue, 3.1X = blue, 1.4X = dark blue).

29

30 **Figure 8.** NIR PC1 throughout time while zoom = 1.4X for *Setaria* genotypes *S. viridis* (A10) and *S.*  
31 *italica* (B100). Plants watered to 100% or 33% FC are shown. Local regression (LOESS) fitted curves  
32 with standard error are plotted for each genotype by water treatment group.

33

1   **Supplemental Figure S1.** Plant height estimated with PlantCV for ten *Setaria* genotypes. Estimated plant  
2   height for *S. viridis*, *S. italica*, and eight RIL plants from 11–33 DAP. Plants watered to 100% or 33% FC  
3   are shown. Local regression (LOESS) fitted curves with standard error are plotted for each genotype by  
4   water treatment group. The arrows indicate the day that the 33% FC water treatment started.  
5

6   **Supplemental Figure S2.** Above ground dry-weight biomass modeled using shoot and leaf pixel area  
7   from four side-view VIS images. Data shown are for 41 plants where destructive dry-weight biomass was  
8   measured throughout the experiment. The ordinary least squares linear model with standard error is  
9   plotted.  
10

11   **Supplemental Figure S3.** Modeled fresh-weight biomass estimated with PlantCV for ten *Setaria*  
12   genotypes. Plants watered to 100% or 33% FC are shown. Arrows indicate the day that the 33% FC water  
13   treatment started. Three-component logistic growth curves with 95% confidence intervals are plotted.  
14

15   **Supplemental Figure S4.** Tiller count estimated with PlantCV for ten *Setaria* genotypes. Estimated tiller  
16   count for *S. viridis*, *S. italica*, and eight RIL plants from 11–33 DAP. Plants watered to 100% or 33% FC  
17   are shown. Local regression (LOESS) fitted curves with standard error are plotted for each genotype by  
18   water treatment group. The arrows indicate the day that the 33% FC water treatment started.  
19

20   **Supplemental Figure S5.** Color can separate watering treatments 25 to 26 DAP, but not 17 to 18 DAP in  
21   RGB, HSV and LAB color space. A, PC1 and PC2 of *S. viridis* color PCA 25 to 26 DAP under 100% FC  
22   (blue) and 33% FC water treatment (red). RGB PC1 accounted for 65.06% of variation and PC2  
23   accounted for 15.98% of variation. B, Principal component regression showed that a two component  
24   model including PC2 separates a significant proportion of variation due to 100% FC (blue) and 33% FC  
25   (red) water treatments 25 to 26 DAP in RGB, HSV and LAB color space. C, PC1 and PC2 of *S. viridis*  
26   color PCA 17 to 18 DAP under 100% FC (blue) and 33% FC water (red) treatment. RGB PC1 accounted  
27   for 46.94% of variation and PC2 accounted for 17.77 % of variation. D, 17 to 18 DAP, PC1 of RGB,  
28   HSV, or LAB color did not separate 100%FC (blue) and 33% FC (red) water treatments.  
29

30   **Supplemental Figure S6.** Loadings of color data 25 to 26 DAP and 17 to 18 DAP. A, PC1 loadings of *S.*  
31   *viridis* color data PCA 25 to 26 DAP. B, PC1 loadings of *S. viridis* color data PCA 17 to 18 DAP. C, PC2  
32   loadings of *S. viridis* color data PCA 25 to 26 DAP. D, PC2 loadings of *S. viridis* color data PCA 17 to 18  
33   DAP.  
34

1 **Supplemental Figure S7.** Before water treatment is applied (11 to 12 DAP) color can distinguish *Setaria*  
2 genotypes in RGB color (Fig. 6), which is also consistent in HSV and LAB color space. A, PCA of HSV  
3 (left) and LAB (right) color for all *Setaria* genotypes before treatment is applied (11 to 12 DAP). PC1 and  
4 PC2 are plotted for *S. viridis* (orange), *S. italica* (gray), RIL161 (purple) and 7 other *S. viridis* x *S. italica*  
5 RILs (gray). B, PCA of HSV (left) and LAB (right) color for all *Setaria* genotypes before treatment is  
6 applied (11 to 12 DAP). PC1 and PC2 are plotted for *S. viridis* (gray), *S. italica* (green), RIL161 (purple)  
7 and 7 other *S. viridis* x *S. italica* RILs (gray).

8  
9 **Supplemental Figure S8.** Loadings of color data 11 to 12 DAP. A, PC1 loadings of *S. viridis* color data  
10 PCA 11 to 12 DAP. B, PC1 loadings of *S. viridis* color data PCA 11 to 12 DAP.

11  
12 **Supplemental Figure S9.** Height significantly correlates with median  $F_v/F_m$ . A, Median  $F_v/F_m$  is  
13 significantly negatively correlated ( $p < 0.01$ ; rho=-0.852) with estimated plant height. Average estimated  
14 height was extracted from four side-view VIS images by PlantCV from all ten *Setaria* genotypes over  
15 time under both 100% FC and 33% FC water treatments.  $F_v/F_m$  values for each plant pixel was calculated  
16 by PlantCV from  $F_0$ ,  $F_{min}$  and  $F_{max}$  PSII images. Median  $F_v/F_m$  is the median bin value in the  $F_v/F_m$  per  
17 pixel histogram for each plant. B, 3D-printed ‘plant’ components consist of two 5 cm ‘stem’ pieces and  
18 one flat ‘leaf’ platform piece. Designs for 3D-printed ‘plant’ components are available at  
19 <https://www.thingiverse.com/thing:574287>. C, Assembled 3D-printed plant with excised *N. benthamina*  
20 leaf tissue. *N. benthamina* leaf was positioned at 0 cm (flush with gravel in pot), 5 cm, 10 cm, and 15 cm.  
21 D, Median  $F_v/F_m$  is significantly negatively correlated ( $p < 0.01$ ; r=0.976) with height in the 3D-printed  
22 plant. 3D-printed plant with excised *N. benthamina* leaf tissue was imaged by the PSII imaging station  
23 two times (tissue from a second leaf during second imaging).  $F_v/F_m$  values for each plant pixel at each  
24 platform level was calculated by PlantCV from  $F_0$ ,  $F_{min}$  and  $F_{max}$  PSII images. Median  $F_v/F_m$  is the median  
25 bin value in the  $F_v/F_m$  per pixel histogram calculated for plant tissue at each platform level.

26  
27 **Supplemental Figure S10.** Matrices displaying Pearson’s Correlation Coefficient calculated between  
28 traits at four roughly equidistant time points. Darker red shading indicates a stronger positive correlation,  
29 whereas increased blue shading indicates stronger negative correlation.

30  
31 **Supplemental Figure S11.** A visual illustration of the pipeline used to threshold plant tissue from  
32 background within grayscale NIR images. Color boxes group together sets of related processes (Original  
33 images = dark blue; Original images processed with background subtraction = orange; Laplacian filtering

1 = dark green; Sobel filtering = yellow; Combined derivative filtering = light green; bitwise OR join =  
2 fuchsia; image masking and object identification = purple; final processed images = light blue).

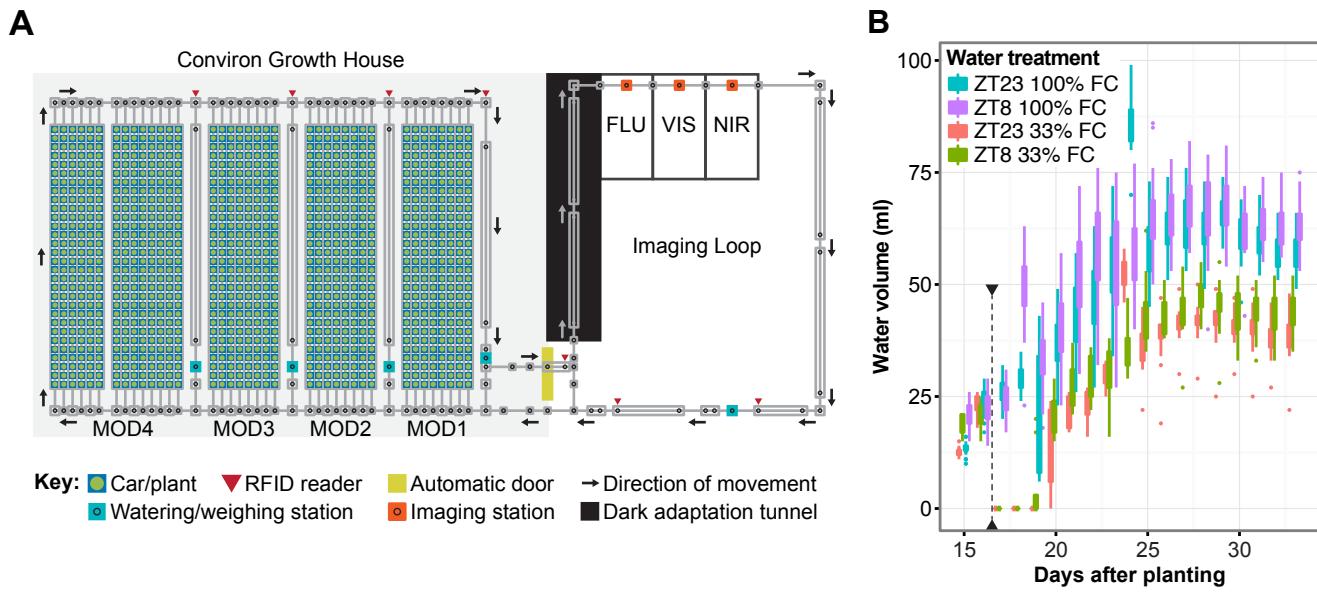
3

4 **Supplemental Figure S12.** JPEG images are sufficient for analysis when only shape and morphological  
5 features are needed. PNG format images are preferred when signal/spectral data is analyzed. A–F,  
6 Correlation analysis included all *Setaria* genotypes and the 100% and 33% FC water treatment groups. A,  
7 VIS side-view plant pixel area. B, VIS top-view plant pixel area. C, Estimated plant height. D, PC1 of  
8 RGB color PCA. E, PC2 of RGB color PCA. F, PC3 of RGB color PCA.

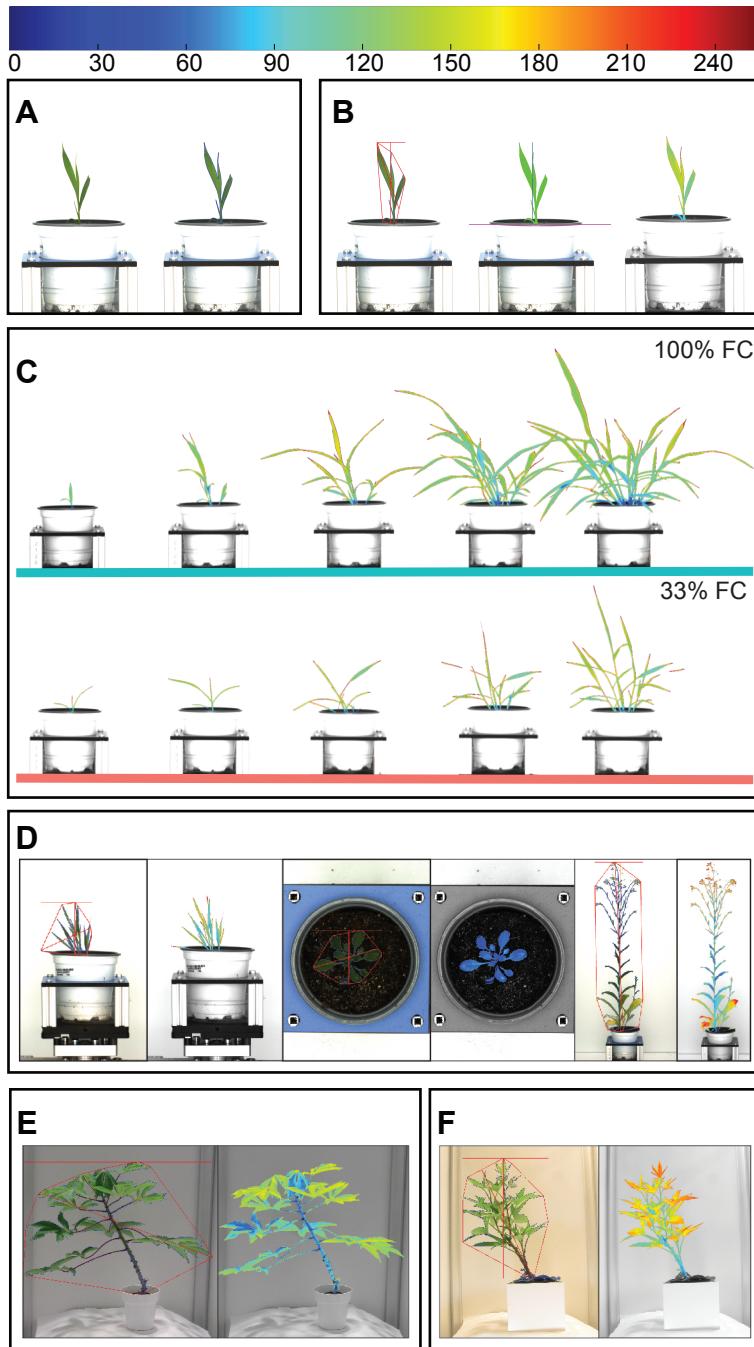
9

10 **Supplemental Figure S13.** Zoom-correction scaling factors for pixel area and pixel length. A, Zoom-  
11 correction scaling factor model for pixel area calculated from a reference object of known area. B, Zoom-  
12 correction scaling factor model for pixel length calculated from a reference object of known length.

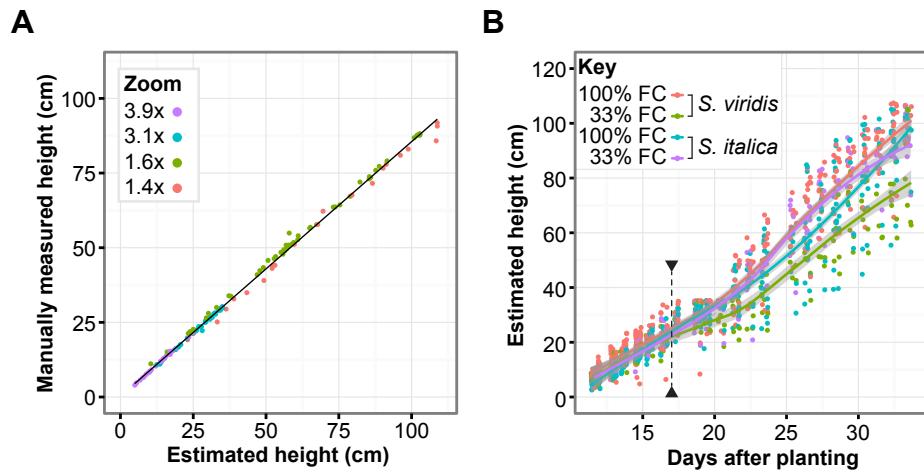
13



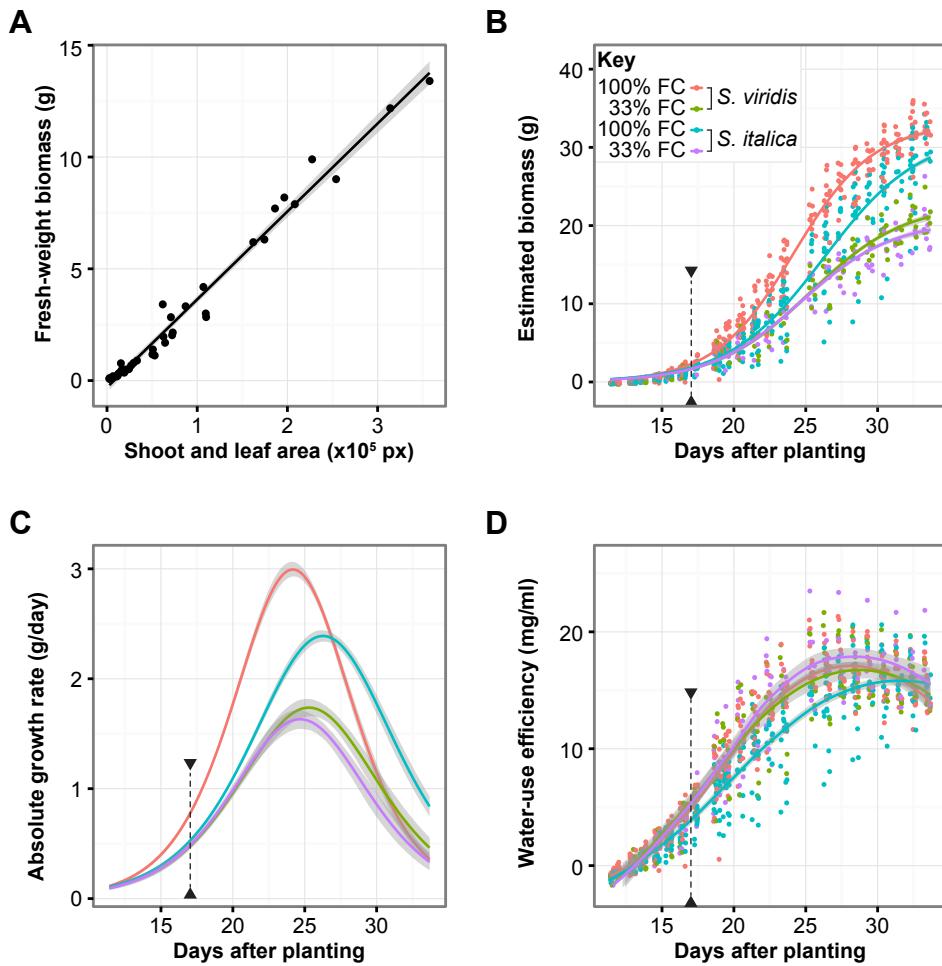
**Figure 1.** Bellwether Phenotyping Platform at the Donald Danforth Plant Science Center. A, Diagram of the Bellwether Phenotyping Platform. B, The volume of water added to *S. viridis* plants during morning (ZT23) and evening (ZT8) water applications.



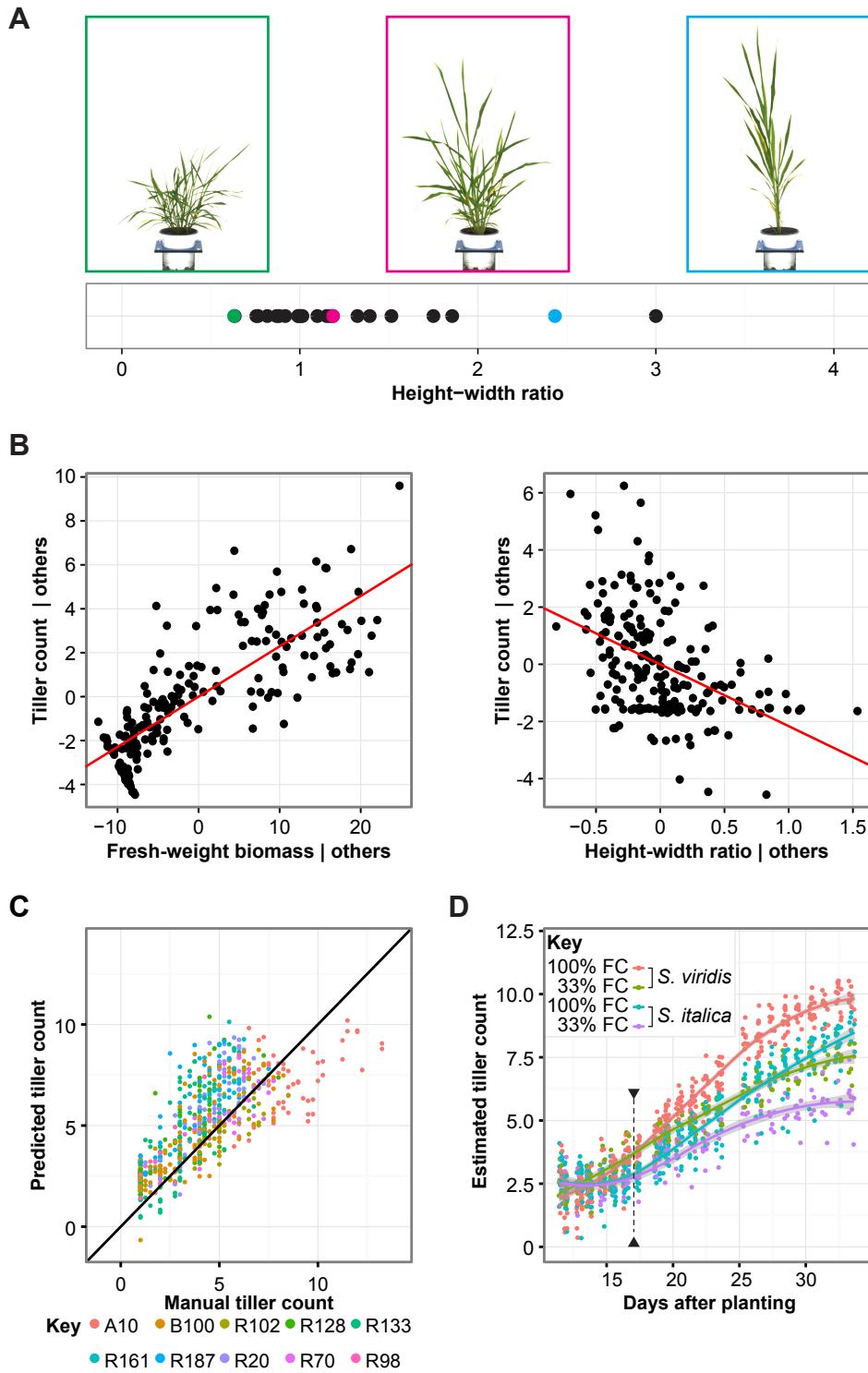
**Figure 2.** PlantCV image analysis software developed at the Donald Danforth Plant Science Center. A, Automated identification of plant material from background. Plant is identified (blue outline, right) from the original image (left). B, Example of traits extracted by PlantCV. Shape characteristics such as convex hull (left), and height (middle) are extracted from visible images. Differences in color can be visualized by pseudocoloring by HSV, LAB or RGB color channels (right). *S. viridis* is pseudocolored based on the value channel of HSV color space. C, PlantCV analysis of traits over time under 100% FC and 33% FC water treatments. Traits such as biomass accumulation, growth-rate, and color can be tracked over time under experimental conditions, such as drought, from data collected by PlantCV. *S. viridis* is shown pseudocolored based on the value channel of HSV color space. D, Example of PlantCV extracted traits from plant species other than *Setaria*. *Brachypodium distachyon*, *Arabidopsis thaliana*, and *Camelina sativa* are pictured from left to right. Plants are pseudocolored based on the value channel of HSV. E, Example of PlantCV analysis of an image of cassava captured by a cell phone camera. F, Example of PlantCV analysis of an image of sweet potato captured by a digital camera.



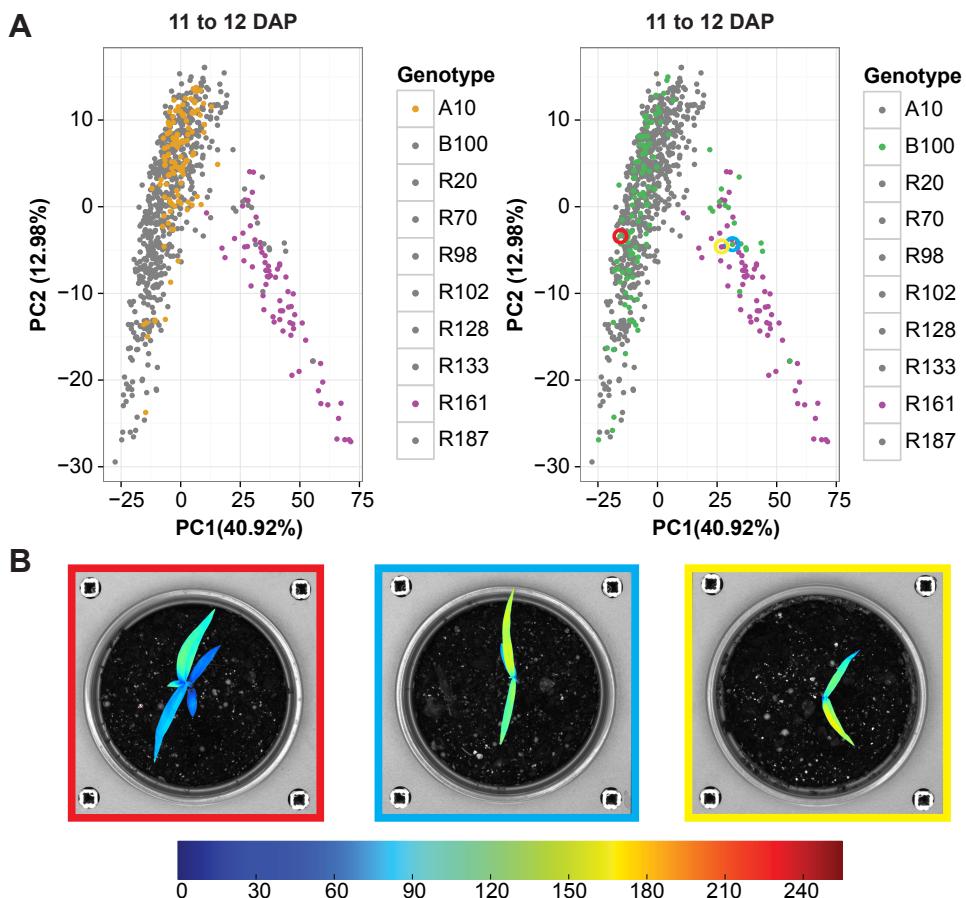
**Figure 3.** Plant height estimated with PlantCV. A, Plant height estimated with PlantCV compared to manually measured height. Data shown are for 173 randomly selected VIS side-view images at four different zoom levels. The ordinary least squares linear model with standard error is plotted. B, Estimated plant height for *S. viridis* and *S. italica* plants from 11–33 DAP. Plants watered to 100% or 33% FC are shown. Local regression (LOESS) fitted curves with standard error are plotted for each genotype by water treatment group. The arrows indicate the day that the 33% FC water treatment started.



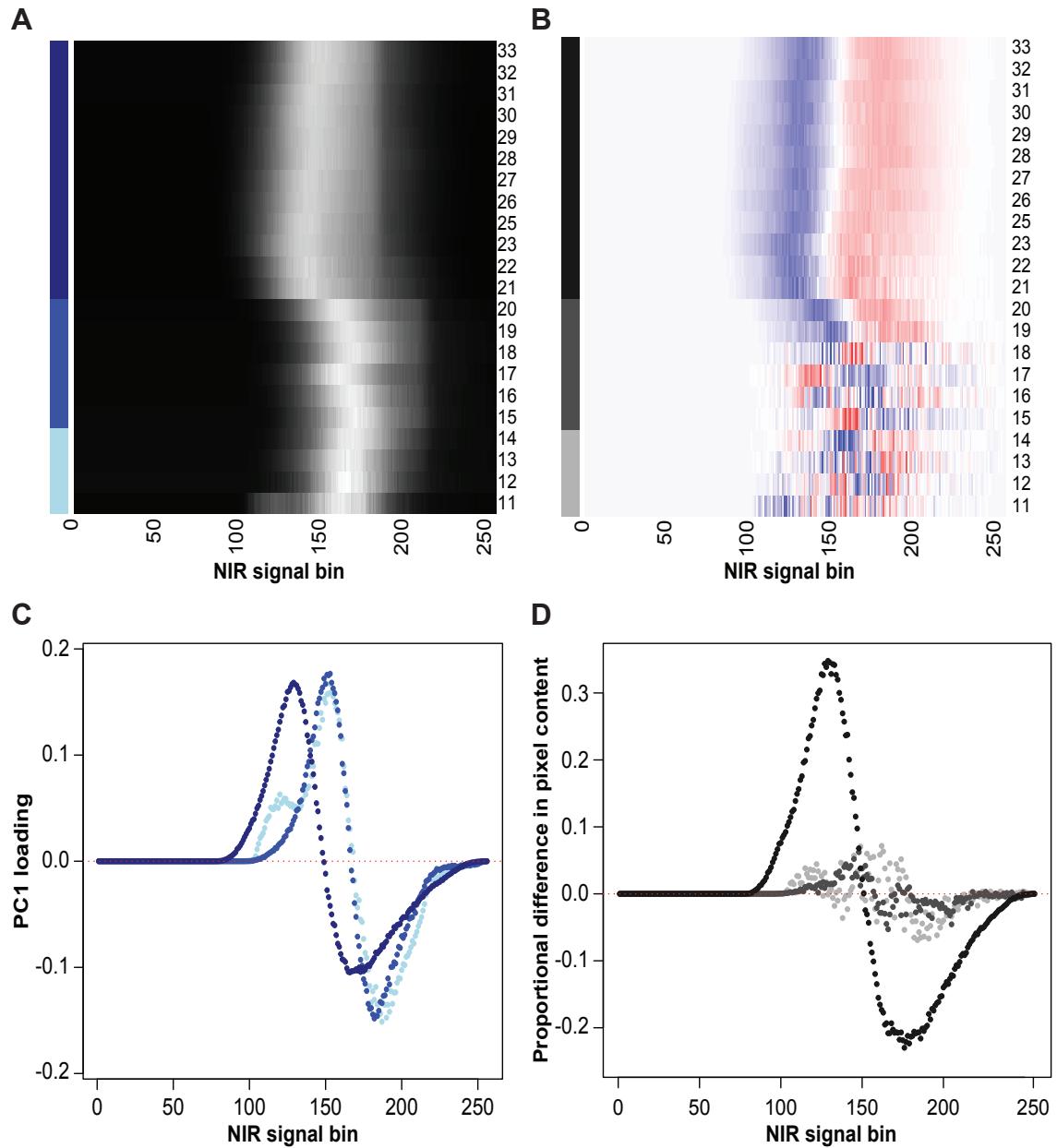
**Figure 4.** *S. viridis* and *S. italica* respond differently to water-limited conditions. A, Above ground fresh-weight biomass modeled using shoot and leaf pixel area from four side-view VIS images. Data shown are for 41 plants where destructive fresh-weight biomass was measured throughout the experiment. The ordinary least squares linear model with standard error is plotted. B–D, data shown are for *S. viridis* and *S. italica* imaged from 11–33 DAP. Plants watered to 100% or 33% FC are shown. Arrows indicate the day that the 33% FC water treatment started. B, Modeled fresh-weight biomass. Three-component logistic growth curves with 95% confidence intervals are plotted. C, Absolute growth rates over time with 95% confidence intervals. D, Water-use efficiency, accumulated fresh-weight biomass divided by cumulative water added. LOESS fitted curves with standard error are plotted. C and D, see color key in B.



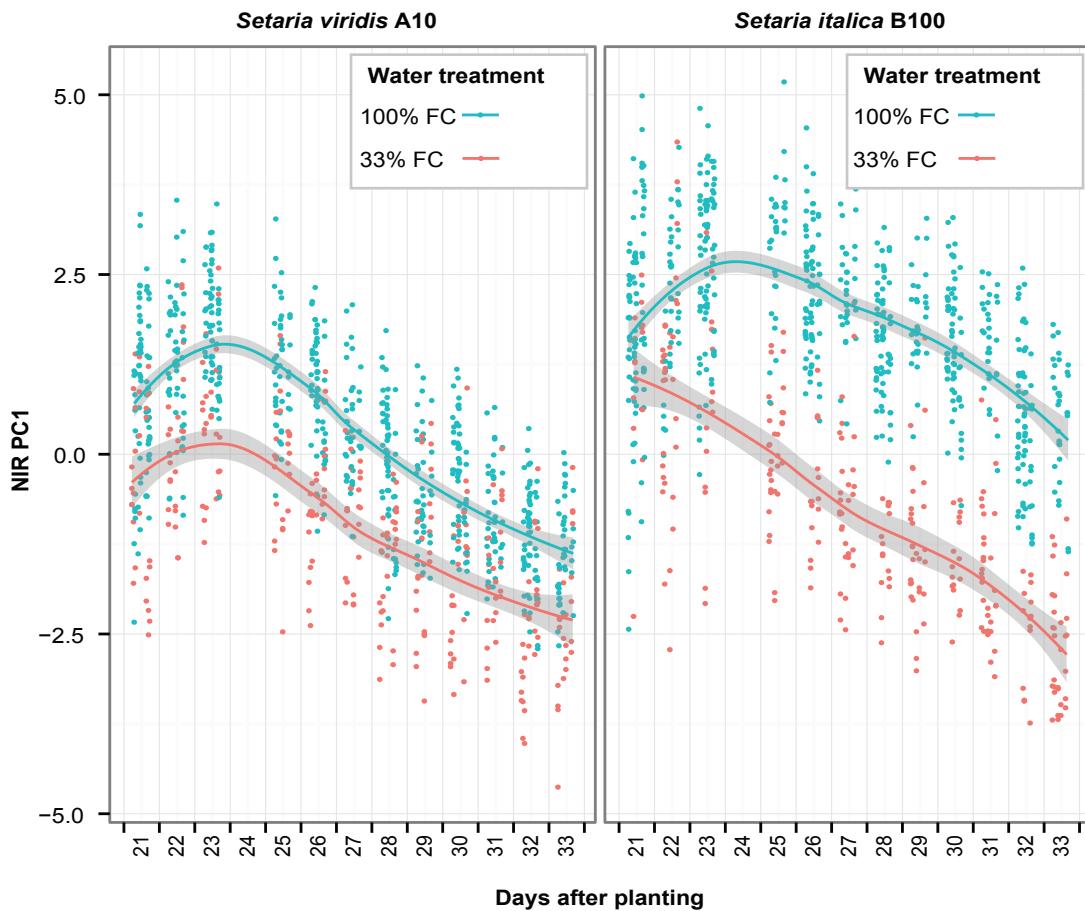
**Figure 5.** Tiller count modeled from spatial-independent morphological characteristics. A, Examples of height-width ratio variation from randomly selected images of plants 25–26 DAP. B, Added-variable plots of fresh-weight biomass and height-width ratio from 195 randomly selected images. C, Relationship between manually measured tiller count from 646 randomly selected images and model-predicted tiller count. D, Model-predicted tiller count for *S. viridis* and *S. italica* plants from 11–33 DAP. Plants watered to 100% or 33% FC are shown. Arrows indicate the day that the 33% FC water treatment started.



**Figure 6.** Before water treatment is applied (11 to 12 DAP) color can distinguish *Setaria* genotypes. A, PCA of RGB color for all *Setaria* genotypes before treatment is applied (11 to 12 DAP). PC1 and PC2 are plotted for *S. viridis* (orange), *S. italica* (green), RIL161 (purple) and 7 other *S. viridis* x *S. italica* RILs (gray). B, *S. italica* (yellow circle in A; yellow boxed image) found amongst R161 (blue circle in A; blue boxed image) are more similar in color than *S. italica* (red circle in A; red boxed image) that groups with other lines (gray). VIS images are pseudocolored based on the value channel in HSV color space.



**Figure 7.** Measurements derived from the NIR camera system. A, Average NIR signal distribution of well-watered plants throughout the experiment. Bins corresponding to signal intensity levels 0 – 255 are plotted along the x-axis, while experimental days are plotted along the y-axis. The proportion of pixels that fall into each bin are expressed as a grayscale shade (0% black, larger percentages are increasingly more white). Color bars along the y-axis illustrate zoom range (3.9X = light blue, 3.1X = blue, 1.4X = dark blue). B, Difference of average NIR signal distribution between plants receiving 100% and 33% FC. Signal bins shaded blue have an increased proportion of pixels in images of plants receiving 100% FC water; signal bins shaded red have a greater proportion of pixels in images receiving 33% FC water; white corresponds to no difference between images. C, A plot of PC1 loadings (3.9X = light blue, 3.1X = blue, 1.4X = dark blue). D, A plot of the average signal difference between water treatment levels 100% and 33% FC (3.9X = light blue, 3.1X = blue, 1.4X = dark blue).



**Figure 8.** NIR PC1 throughout time while zoom = 1.4X for *Setaria* genotypes *S. viridis* (A10) and *S. italica* (B100). Plants watered to 100% or 33% FC are shown. Local regression (LOESS) fitted curves with standard error are plotted for each genotype by water treatment group.

1 **SUPPLEMENTAL DATA**

2

3 **SUPPLEMENTAL METHODS**

4

5 **Image Processing**

6

7 The following describe the image processing pipelines and outlier flagging algorithms used to process  
8 this *Setaria* experiment. Paths to image processing and trait extraction pipelines can be found in Table 4.

9 For further details on individual functions please refer to the documentation at

10 <http://plantcv.danforthcenter.org/>.

11

12 **VIS image processing:** For 1.4X and 1.6X side-view zoom configurations, the RGB image first was  
13 converted to HSV color-space and the saturation channel was isolated. To remove most of the pot from  
14 the image, a threshold was applied to the isolated saturation channel, which converted the image to  
15 binary, with white pixels as objects. A median blur was then used to remove some of the noise from the  
16 binary image.

17

18 The original RGB image was then converted to LAB color-space and the blue-yellow channel was  
19 isolated. To remove most of the blue pot carrier from the image, a threshold was applied to the isolated  
20 blue-yellow channel, which converted the image to binary, with white pixels as objects. The two binary  
21 images from the saturation and blue-yellow channels were then joined together so that only pixels that  
22 were white in both images were isolated. The joined binary image was then used to mask the original  
23 RGB image so the pot and pot carrier were removed. This new masked image contained plant material as  
24 well as residual background from the black sidebars in the background. By masking a portion of the non-  
25 plant material from the original image a less stringent threshold could be applied to the masked image to  
26 better isolate plant material.

27

28 The masked RGB image was converted to LAB color-space and both the blue-magenta, and green-yellow  
29 channels were isolated and thresholded to produce two binary images that extract pixels containing plant  
30 material from background. The two binary images were combined so object pixels from either image  
31 were kept and small areas of noise were filled. The resulting binary image was used to mask the  
32 previously masked image that had removed the pot and pot carrier.

33

1 In the resulting masked image, areas that contain black bars were isolated with a region-of-interest tool,  
2 and plant material overlapping the black bars was separated with a series of thresholding fill, and masking  
3 steps. After thresholding and masking, a binary image with the white object pixels overlapping the  
4 sidebars was saved. The region of interest tool was used to define an area that encompasses everything  
5 except the black background bars in the image, and object pixels were detected and combined with the  
6 object pixels in the sidebars. The final binary object mask was used for object detection and the detected  
7 objects were grouped together for extraction of shape characteristics. The plant color analysis function  
8 also used this final mask to isolate plant pixels for color analysis. VIS shape and color traits currently  
9 extracted by PlantCV are described in more detail in Table 3 and the documentation of PlantCV software  
10 (<http://plantcv.danforthcenter.org/>). For analysis of images at 3.1X and 3.9X magnification, background  
11 sidebars were no longer visible in images so sidebar steps were not included in plant identification  
12 pipelines.

13

14 **PSII image processing:** First a pre-made mask for the pot carrier screws that auto-fluoresce was  
15 converted to a binary image and applied to an 8-bit color version of the 16-bit grayscale  $F_{max}$  image. The  
16 8-bit color  $F_{max}$  image was used to create a mask to isolate plant pixels, which was later applied to the 16-  
17 bit grayscale images. A binary threshold was applied to the 8-bit  $F_{max}$  image with masked screws and a  
18 median blur and fill step was used to reduce some of the noise. Objects (white pixels) were identified  
19 within the resulting image. The region of interest tool was then used to define a large area that  
20 encompasses the plant (centered on the pot carrier position that is consistent in all images). Objects that  
21 were within or partially within the region of interest were kept. A binary image with the kept objects was  
22 saved as an image mask. The kept objects were also grouped together for shape analysis. Shape data was  
23 used as a quality control of plant detection in image processing. The mask of the kept objects was then  
24 applied to the  $F_0$ ,  $F_{min}$  and  $F_{max}$  images and  $F_v/F_m$  measurements were calculated over a matrix of  
25 identified plant pixels.

26 The  $F_0$  image was captured before the fluorescence inducing light-pulse and was therefore used as a null  
27 standard in each image set. If the maximum signal in the null image surpasses a threshold, the image set  
28 was flagged so the user may further examine if there were problems during image capture. In this data set  
29 no  $F_0$  images were flagged as outliers.  $F_v/F_m$ , was calculated for each plant pixel in  $F_{min}$  and  $F_{max}$  images  
30 ( $F_{max}-F_{min}/F_{max}$ ), and a histogram of  $F_v/F_m$  values was built using 1000 bins (values from 0 to 1).

31 **NIR image processing:** NIR images derived from pots without plants captured throughout the experiment  
32 were averaged within each zoom level and were used to calculate background signal. These zoom specific  
33 averages of background signal intensity were then subtracted from the image matrix of interest (plant-

1 containing). Binary thresholding was then performed on the background-subtracted matrix, isolating plant  
2 material from image background components.

3  
4 Further refinement in plant capture was achieved using an image sharpening approach similar to the  
5 detection methods developed to enhance visualization of skeletal systems in x-ray images (Gonzalez and  
6 Woods, 2007). Laplacian filtering (second derivative) was applied to the original grayscale image to  
7 detect the boundary between plant material and background components. Edges between plant and  
8 background were subsequently sharpened by subtracting the Laplacian-filtered image from the original.  
9 Other, more ambiguous boundaries and texture changes within the image were then detected using Sobel  
10 filtering (first derivative) applied across both the x and y-axis of the original image. The resulting Sobel  
11 filtered (both x and y-axis) images were then combined through addition and subsequently smoothed  
12 using a median filter to remove noise artifacts generated during application of the derivative filter. The  
13 resulting smoothed, Sobel filtered image was then subtracted from the Laplacian filtered image; further  
14 sharpening the contrast between plant material and background components. Image sharpening enhanced  
15 the ability to isolate plant material from background using binary thresholding. After thresholding, the  
16 resultant binary images then underwent four parallel erosion steps (each step removed any focal pixel  
17 with less than two adjacent non-zero pixels) to remove remaining background noise.

18 This binary image was combined with the binary image obtained from the background-subtraction routine  
19 described above using a logical ‘OR’ join statement. Finally, problematic background components such  
20 as the cart/pot, and image chamber paneling were excluded from the image through application of image  
21 masks. Grayscale NIR pixel-level intensity of plant material was summarized into a histogram containing  
22 256 bins, where each bin contains the proportion of plant pixels exhibiting the corresponding grayscale  
23 intensity value.

24 To isolate pots with *Setaria* plants that did not germinate, values of side-view area and extent y were  
25 calculated at each zoom level with pots known to be empty. NIR images with side-view area and extent-y  
26 values equal to or less than the maximum value observed in empty-pot NIR images were removed from  
27 the dataset. In total 1083 images were filtered out of the *Setaria* dataset using NIR side-view area and  
28 extent-y criterion (691 at 3.9X zoom; 36 at 3.1X zoom; 356 at 1.4X zoom).

29 To identify potential NIR image plant detection problems, plant area ( $r = 0.992$ ), extent y ( $r = 0.971$ ) and  
30 perimeter ( $r = 0.962$ ) were correlated between VIS and NIR images. Images flagged for growing outside  
31 of the field of view in VIS images, but not outside the field of view in NIR images, were omitted from the  
32 shape outlier detection routine. NIR images where measurement residual error was greater than two

1 standard deviations from the predicted model values were flagged as potential outliers. Chi-square  
2 distance and Earth Mover's Distance metrics were used to compare histograms of NIR signal between  
3 image angles, and sample groups (genotype x treatment x date). Chi-square and Earth Mover's Distances  
4 were calculated in R using the distance function encoded by the analogue package (Simpson, 2007) and  
5 the emd2d function in the emdist library (Urbanek and Rubner, 2012), respectively. Image histogram  
6 distance values that were greater than two standard deviations from the mean were flagged as potential  
7 outliers. Manual examination of pseudo-colored (colored by plant area identified) NIR images flagged as  
8 outliers found that plant capture was adequate in these images so no additional images were removed  
9 from the dataset.

10 **SUPPLEMENTAL REFERENCES**

- 11 **Gonzalez, R.C., Woods, R.E.** (2008) *Digital Image Processing*, 3rd edition, Pearson Prentice Hall,  
12 Upper Saddle River.  
13  
14 **Simpson GL** (2007) Analogue Methods in Palaeoecology: Using the analogue Package. J Stat Softw **22**:  
15 1–29.  
16  
17 **Urbanek S, Rubner Y** (2012) emdist: Earth Mover's Distance.  
18  
19

<b>Feature</b>	<b>PlantCV</b>	<b>IAP</b>	<b>HTPheno</b>	<b>EBIImage</b>	<b>LemnaTec***</b>
<b>Image Types*</b>	Color, Grayscale, PSII	Color, Grayscale	Color	Color, Grayscale	Color, Grayscale, PSII
<b>Community Contribution</b>	Via GitHub	Contact Authors	Contact Authors	Contact Authors	--
<b>Database Type</b>	SQLite	MongoDB	--	--	--
<b>Extendibility</b>	Python Modules**	IAP Plugin	ImageJ Plugins	R Packages	--
<b>Image Analysis Functions</b>	48	51	12	32	111
<b>Interface</b>	Command Line	GUI or Command Line	ImageJ/GUI	R (GUI or Command Line)	GUI
<b>Language</b>	Python	Java	Java	R	--
<b>License</b>	GPLv2	GPLv2	GPLv2	LGPL	Commercial
<b>Scalability/ Parallelization</b>	Multi-thread	Grid	--	--	Grid

1      **Supplemental Table SI.** Comparison of image-analysis software. \*Color (e.g. VIS cameras), Grayscale  
 2      (e.g. NIR, IR, FLU, IR, or thermal-infrared cameras), PSII (saturating-pulse FLU camera system).  
 3      \*\*Examples of Python libraries to extend PlantCV function include OpenCV (Python), SciPY, NumPy,  
 4      and MatPlotLib. \*\*\*LemnaLauncher v4.x.  
 5

1      **A VIS Image Configurations**

2

<b>Days After Planting</b>	<b>Camera</b>	<b>Exposure</b>	<b>Gain</b>	<b>Iris</b>	<b>Focus</b>	<b>Zoom Setting</b>	<b>Optical Zoom</b>	<b>Lifter</b>
11 to 14	side-view	110	0	1500	1900	3500	3.9X	2
15 to 16	side-view	110	0	1500	1900	2500	3.1X	2
17 to 20	side-view	110	0	1500	1900	2500	3.1X	2
21	side-view	110	0	1500	1900	500	1.4X	1
22 to 31	side-view	110	0	1500	1900	700	1.6X	1
32 to 33	side-view	110	0	1500	1900	500	1.4X	1
11 to 14	top-view	110	0	1500	800	3500	3.9X	2
15 to 16	top-view	110	0	1500	800	3000	3.5X	2
17 to 20	top-view	110	0	1500	800	2500	3.1X	2
21	top-view	110	0	1500	800	1500	2.3X	2
22 to 31	top-view	110	0	1500	800	300	1.3X	2
32 to 33	top-view	110	0	1500	800	1	1X	2

3

4      **B PSII Image Configurations**

5

<b>Days After Planting</b>	<b>Camera</b>	<b>Gain</b>	<b>Focus</b>	<b>Pulse Power</b>	<b>Shutter</b>	<b>Lifter</b>
11 to 31	top-view	100	4800	1200	500	630
32 to 33	top-view	100	4800	1200	500	500

6

7      **C NIR Image Configurations**

8

<b>Days After Planting</b>	<b>Camera</b>	<b>Exposure</b>	<b>Gain</b>	<b>Iris</b>	<b>Focus</b>	<b>Zoom Setting</b>	<b>Optical Zoom</b>	<b>Lifter</b>
11 to 14	side-view	90	0	2000	1200	3500	3.9X	2
14 to 20	side-view	90	0	2000	1200	2500	3.1X	2
21 to 33	side-view	90	0	2000	1200	500	1.4X	1

9

10     **Supplemental Table SII.** Camera configuration parameters. A, LemnaTec control software VIS image  
 11 configuration settings for *Setaria* experiment. B, LemnaTec control software PSII image configuration  
 12 settings for *Setaria* experiment. C, LemnaTec control software NIR image configuration settings for  
 13 *Setaria* experiment.

14

1 **PlantCV Output Measurements**

2

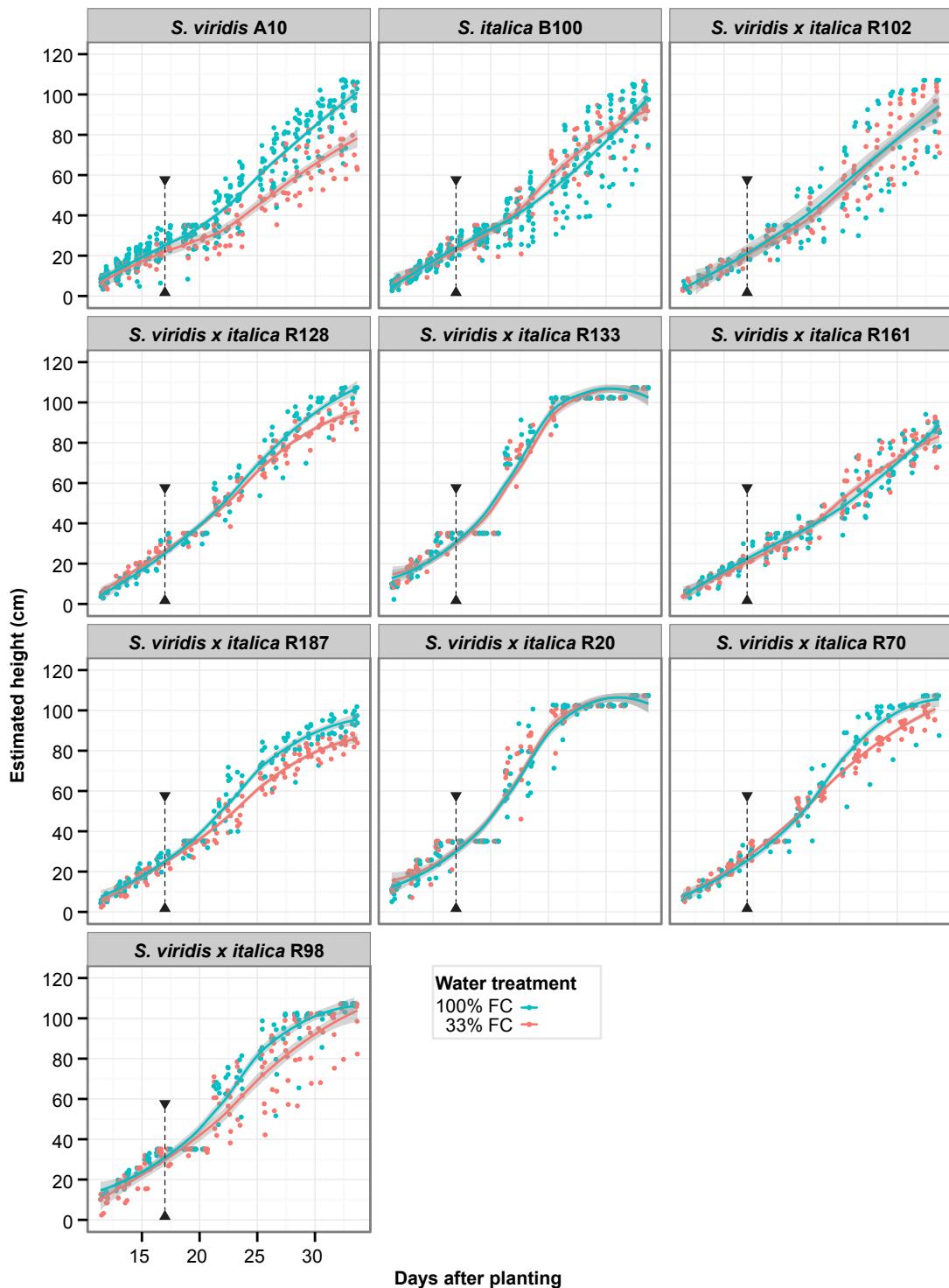
Camera	Measurement	Function Name	General Function Description
VIS	Area Above Bound	analyze_bound	object area above a preset boundary line
VIS	Area Below Bound	analyze_bound	object area below a preset boundary line
VIS	Centroid Height	analyze_object	y position of object centroid to preset boundary line
VIS	Centroid X Position	analyze_object	x position of object centroid
VIS	Centroid Y Position	analyze_object	y position of object centroid
VIS	Color	analyze_color	splits image into RGB, HSV, and LAB color space and outputs histogram of each color channel
VIS	Exent X	analyze_object	total span of object pixels in along x axis
VIS	Exent Y	analyze_object	total span of object pixels along y axis
VIS	Height Above Bound	analyze_bound	total span of object pixels along y axis above a preset boundary line
VIS	Height Above Bound	analyze_bound	total span of object pixels along y axis below a preset boundary line
VIS	Height-Width Ratio	analyze_object	total span of object pixels along y axis above a preset boundary line, divided by total span of object pixels in along x axis
VIS	Long Axis	analyze_object	total span of pixels from convex hull point furthest from the centroid, through the centroid to the intersecting point on convex hull
VIS	Perimeter	analyze_object	total length of pixels around object convex hull
VIS	Side View Area	analyze_object	total number of pixels in object
VIS	Solidity	analyze_object	total number of pixels in object divided by convex hull area
VIS	Top View Area	analyze_object	total number of pixels in object
PSII	Maximum FLU Signal	fluor_fvfm	maximum bin value of Fv/Fm histogram
PSII	Fv/Fm Signal	fluor_fvfm	histogram of Fv/Fm (fmax minus fmin, divided by fmax) values for each pixel
PSII	Median Fv/Fm Signal	fluor_fvfm	median bin value of Fv/Fm histogram
NIR	NIR Signal	analyze_NIR_intensity	histogram values of NIR signal

3

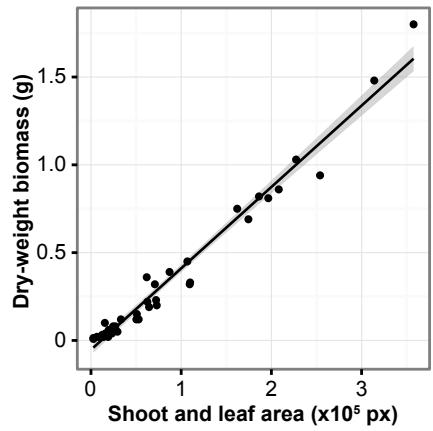
4 **Supplemental Table SIII.** Description of measurements output by PlantCV. For more information on  
5 each function please refer to PlantCV documentation (<https://plantcv.danforthcenter.org>).

Camera	Description	Script Type	Notes	Path
PSII	Lifter 630 FLU TV	Image Processing and Output	N/A	/plantcv/scripts/image_analysis/flu_tv/flu_z630_L1.py
PSII	Lifter 500 FLU TV	Image Processing and Output	N/A	/plantcv/scripts/image_analysis/flu_tv/flu_z500_L1.py
NIR	3.9X Zoom NIR SV	Image Processing and Output	N/A	/plantcv/script/image_analysis/nir_sv/nir_sv_z3500_L1.py
NIR	3.1X Zoom NIR SV	Image Processing and Output	N/A	/plantcv/script/image_analysis/nir_sv/nir_sv_z2500_L1.py
NIR	1.4X Zoom NIR SV	Image Processing and Output	N/A	/plantcv/script/image_analysis/nir_sv/nir_sv_z500_L1.py
VIS	1.4X Zoom VIS SV	Image Processing and Output	N/A	/plantcv/scripts/image_analysis/vis_sv/vis_sv_z500_L1.py
VIS	1.6X Zoom VIS SV	Image Processing and Output	N/A	/plantcv/scripts/image_analysis/vis_sv/vis_sv_z700_L1.py
VIS	3.1X Zoom VIS SV	Image Processing and Output	N/A	/plantcv/scripts/image_analysis/vis_sv/vis_sv_z2500_L1.py
VIS	3.9X Zoom VIS SV	Image Processing and Output	N/A	/plantcv/scripts/image_analysis/vis_sv/vis_sv_z3500_L1.py
VIS	1.0X Zoom VIS TV	Image Processing and Output	N/A	/plantcv/script/image_analysis/vis_tv/vis_tv_z1_L1.py
VIS	1.3X Zoom VIS TV	Image Processing and Output	N/A	/plantcv/script/image_analysis/vis_tv/vis_tv_z300_L1.py
VIS	2.3X Zoom VIS TV	Image Processing and Output	N/A	/plantcv/script/image_analysis/vis_tv/vis_tv_z1500_L1.py
VIS	3.1X Zoom VIS TV	Image Processing and Output	N/A	/plantcv/script/image_analysis/vis_tv/vis_tv_z2500_L1.py
VIS	3.5X Zoom VIS TV	Image Processing and Output	N/A	/plantcv/script/image_analysis/vis_tv/vis_tv_z3000_L1.py
VIS	3.9X Zoom VIS TV	Image Processing and Output	N/A	/plantcv/script/image_analysis/vis_tv/vis_tv_z3500_L1.py
ALL	Parallelization	Paralellization	N/A	/plantcv/scripts/image_analysis/image_analysis.pl
PSII	Extract FLU signal	Data Extraction and DB Query	N/A	/plantcv/scripts/db/flu_fvfm_exporter.py
NIR	Extract NIR Signal	Data Extraction and DB Query	N/A	/plantcv/scripts/db/nir_signal_exporter.py
VIS, NIR	Extract Shape Data	Data Extraction and DB Query	N/A	/plantcv/scripts/db/vis_measurement_exporter.py
VIS	Extract Color Data	Data Extraction and DB Query	N/A	/plantcv/scripts/db/vis_color_exporter.py
None	Watering	Statistical Analysis Script	Fig. 1	/plantcv/scripts/manuscripts/manuscript_1/vis_analysis.R
ALL	Vis Shape and Zoom	Statistical Analysis Script	Fig. 3,4,5,12	/plantcv/scripts/manuscripts/manuscript_1/vis_analysis.R
VIS	Color	Statistical Analysis Script	Fig. 6	/plantcv/scripts/manuscripts/manuscript_1/color_early_late_21115.R
VIS	Color	Statistical Analysis Script	Fig. 7	/plantcv/scripts/manuscripts/manuscript_1/color_r161_a10_b100_21115.R
PSII	PSII	Statistical Analysis Script	Fig. 8	/plantcv/scripts/manuscripts/manuscript_1/PSII_setaria_3315.R
NIR	Analyze NIR Signal	Statistical Analysis Script	Fig. 9, 10	/plantcv/scripts/manuscripts/manuscript_1/NIR_image_analysis_burnin2.R
NIR	Outlier Identification	Statistical Analysis Script	N/A	/plantcv/scripts/manuscripts/manuscript_1/NIR_image_QCQA_burnin2.R
VIS	JPEG vs. PNG	Statistical Analysis Script	Fig. 11	/plantcv/scripts/manuscripts/manuscript_1/color_pngvsjpg_11-18-14.R
ALL	Variance	Statistical Analysis Script	Table 1	/plantcv/scripts/manuscripts/manuscript_1/variance_partitioning_script_burnin2.R

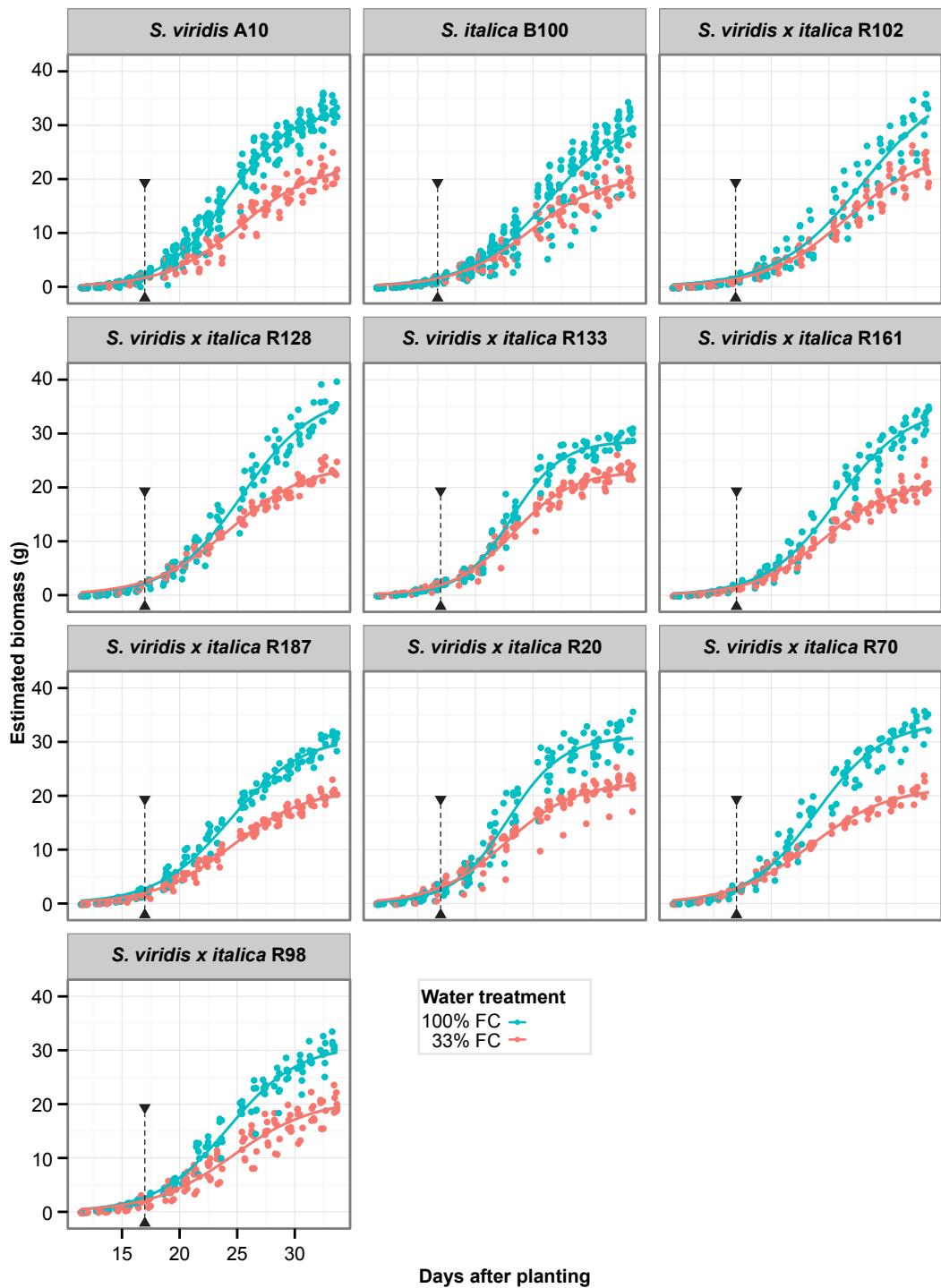
1   **Supplemental Table SIV.** PlantCV image processing and analysis scripts. For more information on image processing procedures and analysis  
2   methods please refer scripts and documentation at <https://plantcv.danforthcenter.org>.



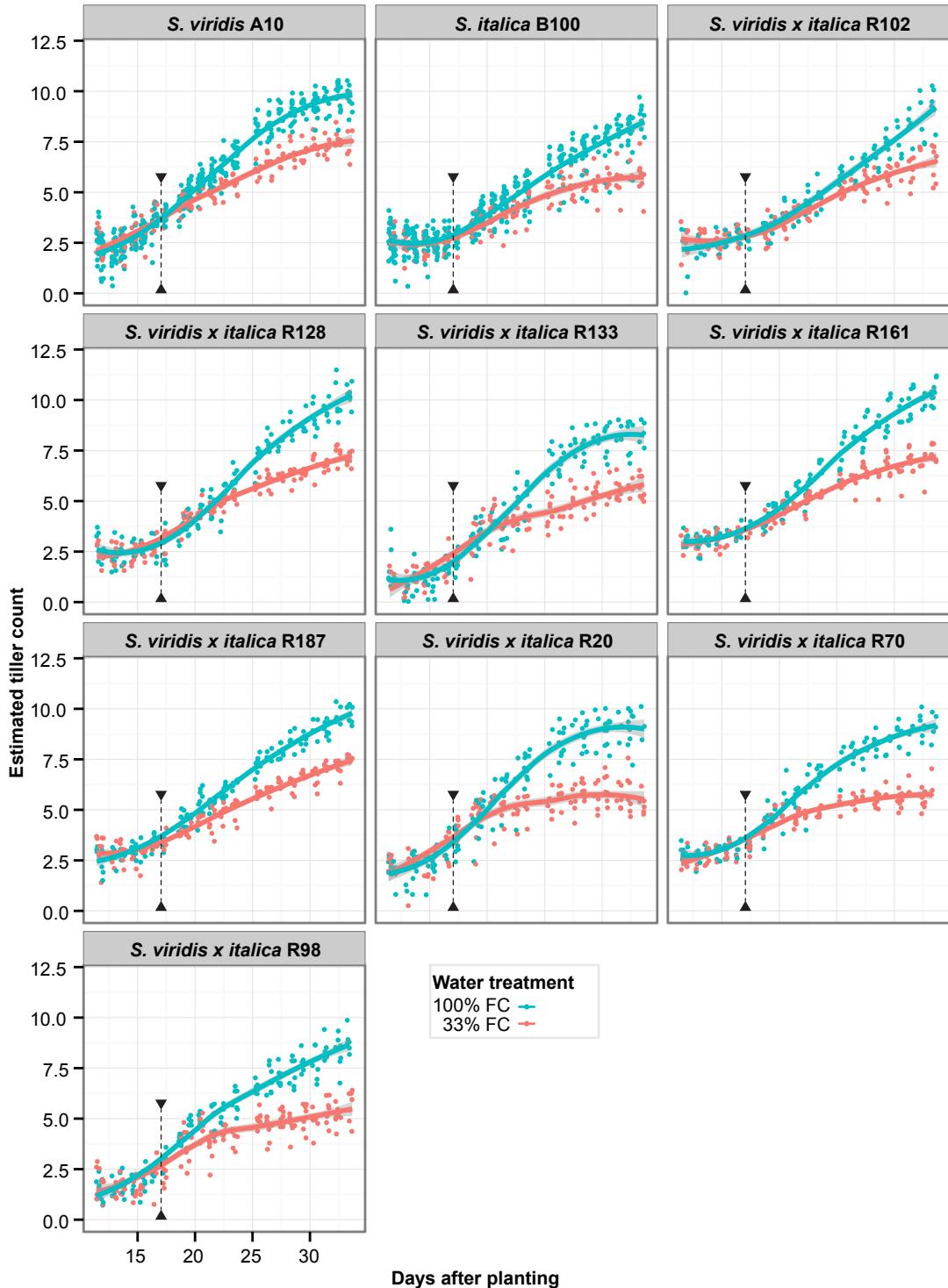
**Supplemental Figure S1.** Plant height estimated with PlantCV for ten *Setaria* genotypes. Estimated plant height for *S. viridis*, *S. italica*, and eight RIL plants from 11–33 DAP. Plants watered to 100% or 33% FC are shown. Local regression (LOESS) fitted curves with standard error are plotted for each genotype by water treatment group. The arrows indicate the day that the 33% FC water treatment started.



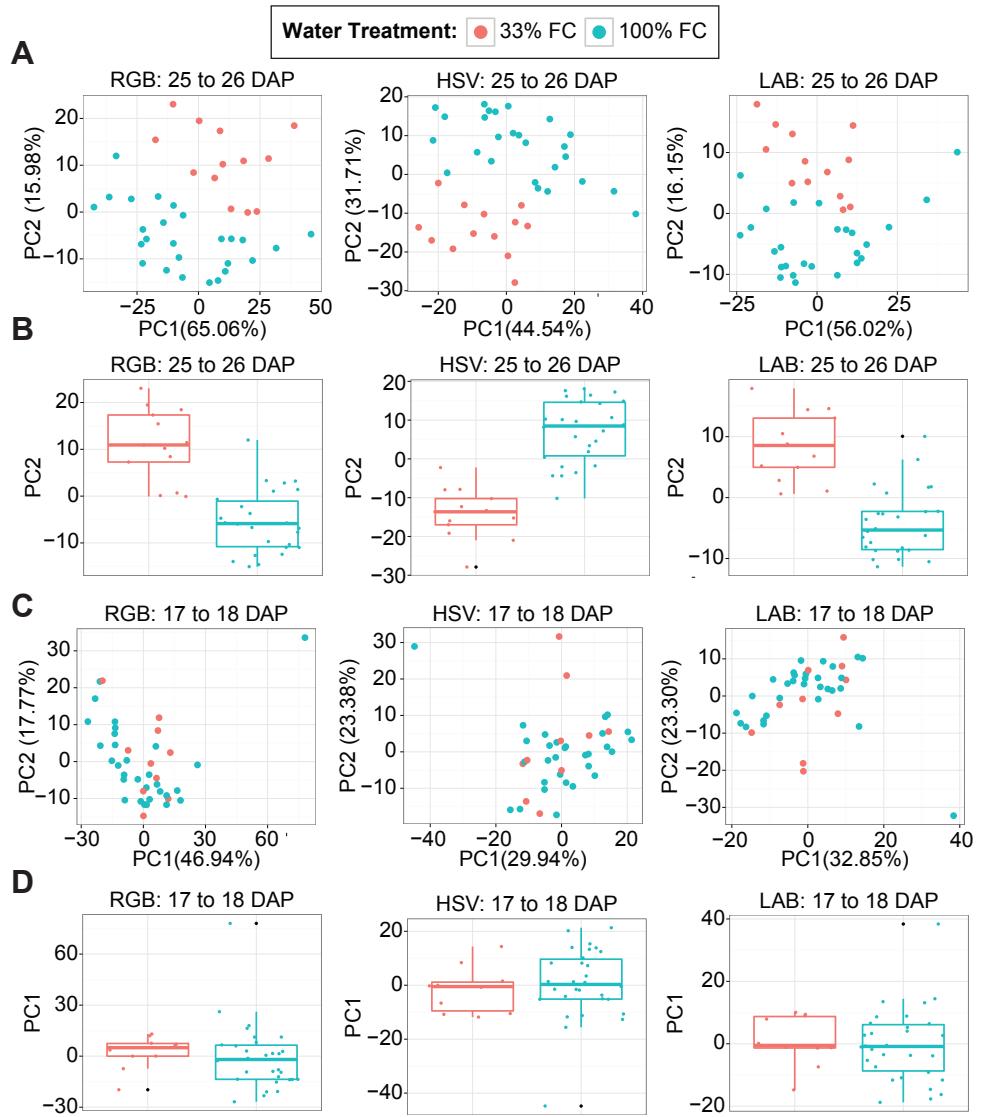
**Supplemental Figure S2.** Above ground dry-weight biomass modeled using shoot and leaf pixel area from four side-view VIS images. Data shown are for 41 plants where destructive dry-weight biomass was measured throughout the experiment. The ordinary least squares linear model with standard error is plotted.



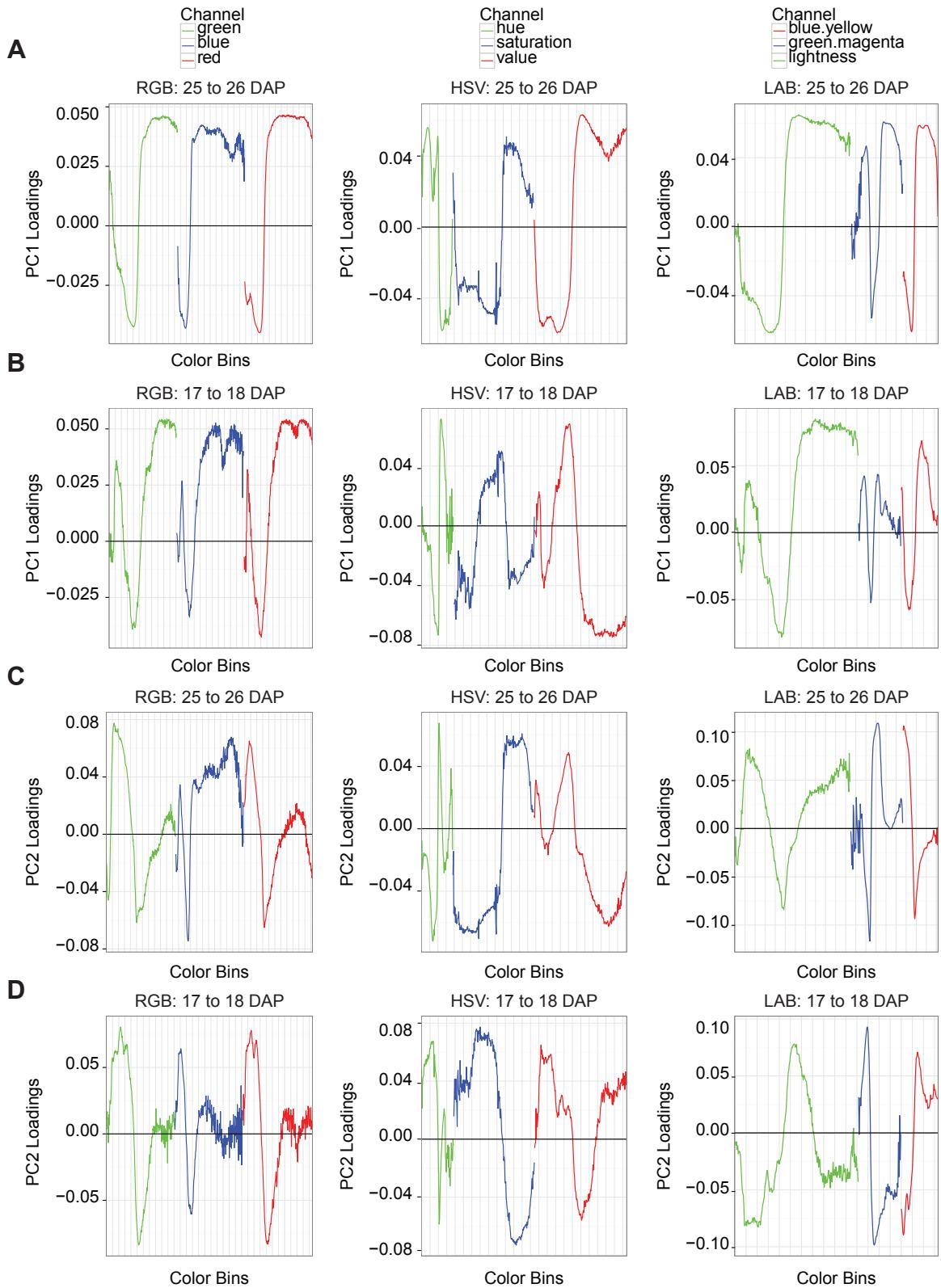
**Supplemental Figure S3.** Modeled fresh-weight biomass estimated with PlantCV for ten *Setaria* genotypes. Plants watered to 100% or 33% FC are shown. Arrows indicate the day that the 33% FC water treatment started. Three-component logistic growth curves with 95% confidence intervals are plotted.



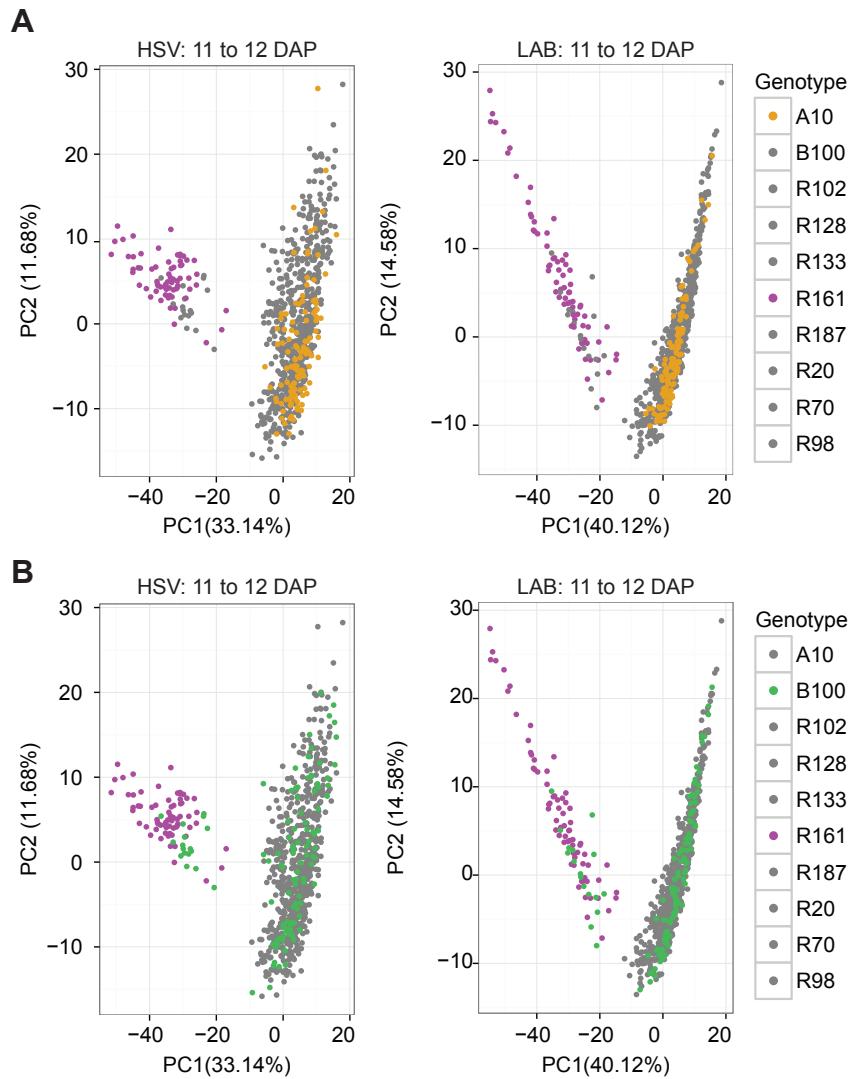
**Supplemental Figure S4.** Tiller count estimated with PlantCV for ten *Setaria* genotypes. Estimated tiller count for *S. viridis*, *S. italica*, and eight RIL plants from 11–33 DAP. Plants watered to 100% or 33% FC are shown. Local regression (LOESS) fitted curves with standard error are plotted for each genotype by water treatment group. The arrows indicate the day that the 33% FC water treatment started.



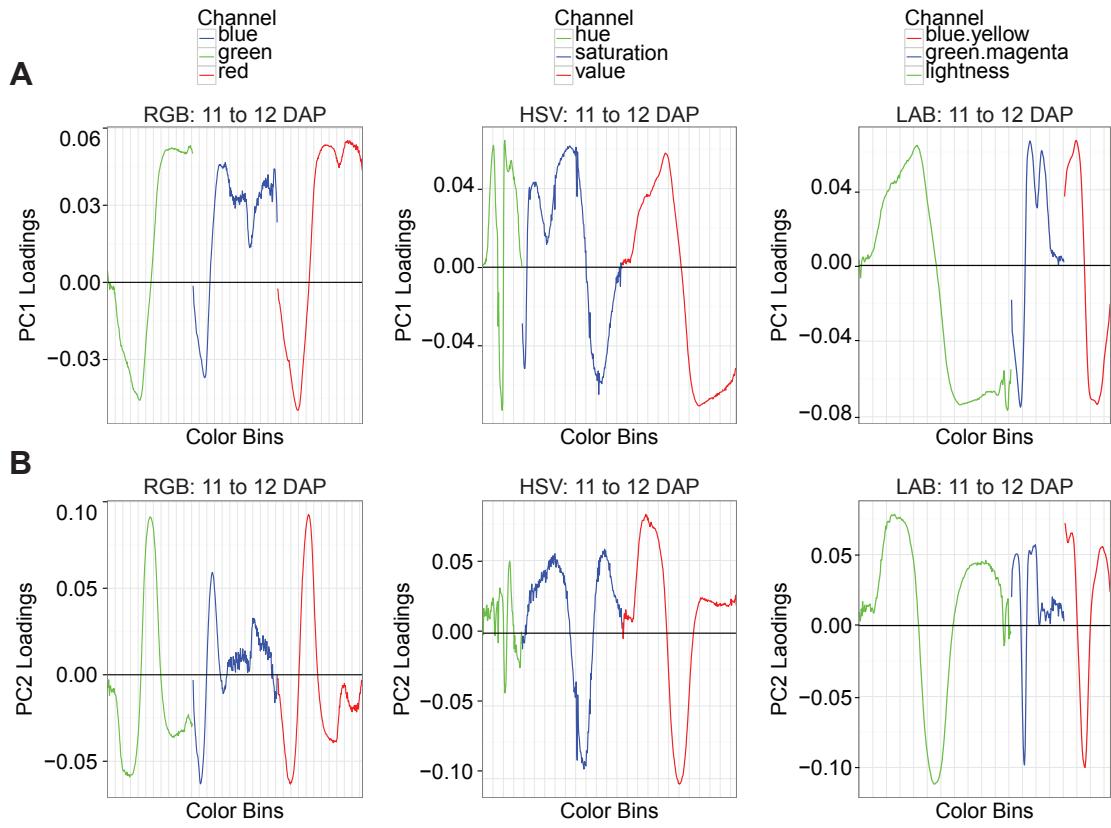
**Supplemental Figure S5.** Color can separate watering treatments 25 to 26 DAP, but not 17 to 18 DAP in RGB, HSV and LAB color space. A, PC1 and PC2 of *S. viridis* color PCA 25 to 26 DAP under 100% FC (blue) and 33% FC water treatment (red). RGB PC1 accounted for 65.06% of variation and PC2 accounted for 15.98% of variation. B, Principal component regression showed that a two component model including PC2 separates a significant proportion of variation due to 100% FC (blue) and 33% FC (red) water treatments 25 to 26 DAP in RGB, HSV and LAB color space. C, PC1 and PC2 of *S. viridis* color PCA 17 to 18 DAP under 100% FC (blue) and 33% FC water (red) treatment. RGB PC1 accounted for 46.94% of variation and PC2 accounted for 17.77 % of variation. D, 17 to 18 DAP, PC1 of RGB, HSV, or LAB color did not separate 100%FC (blue) and 33% FC (red) water treatments.



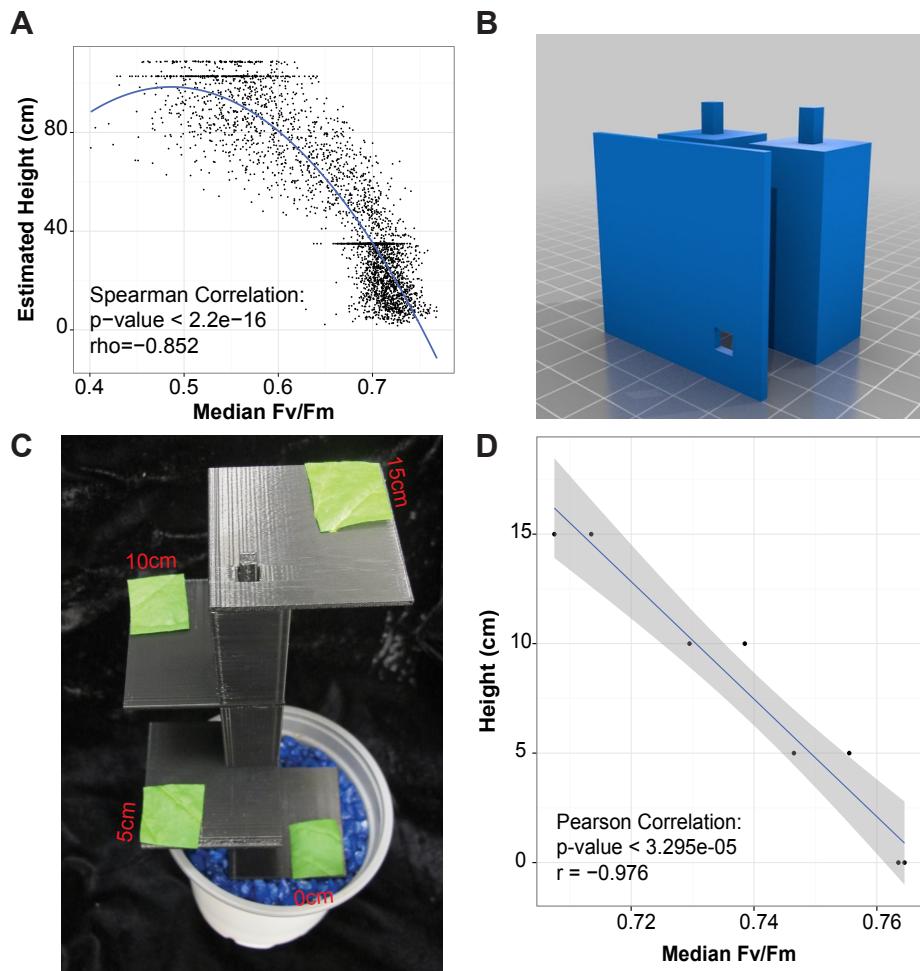
**Supplemental Figure S6.** Loadings of color data 25 to 26 DAP and 17 to 18 DAP. A, PC1 loadings of *S. viridis* color data PCA 25 to 26 DAP. B, PC1 loadings of *S. viridis* color data PCA 17 to 18 DAP. C, PC2 loadings of *S. viridis* color data PCA 25 to 26 DAP. D, PC2 loadings of *S. viridis* color data PCA 17 to 18 DAP.



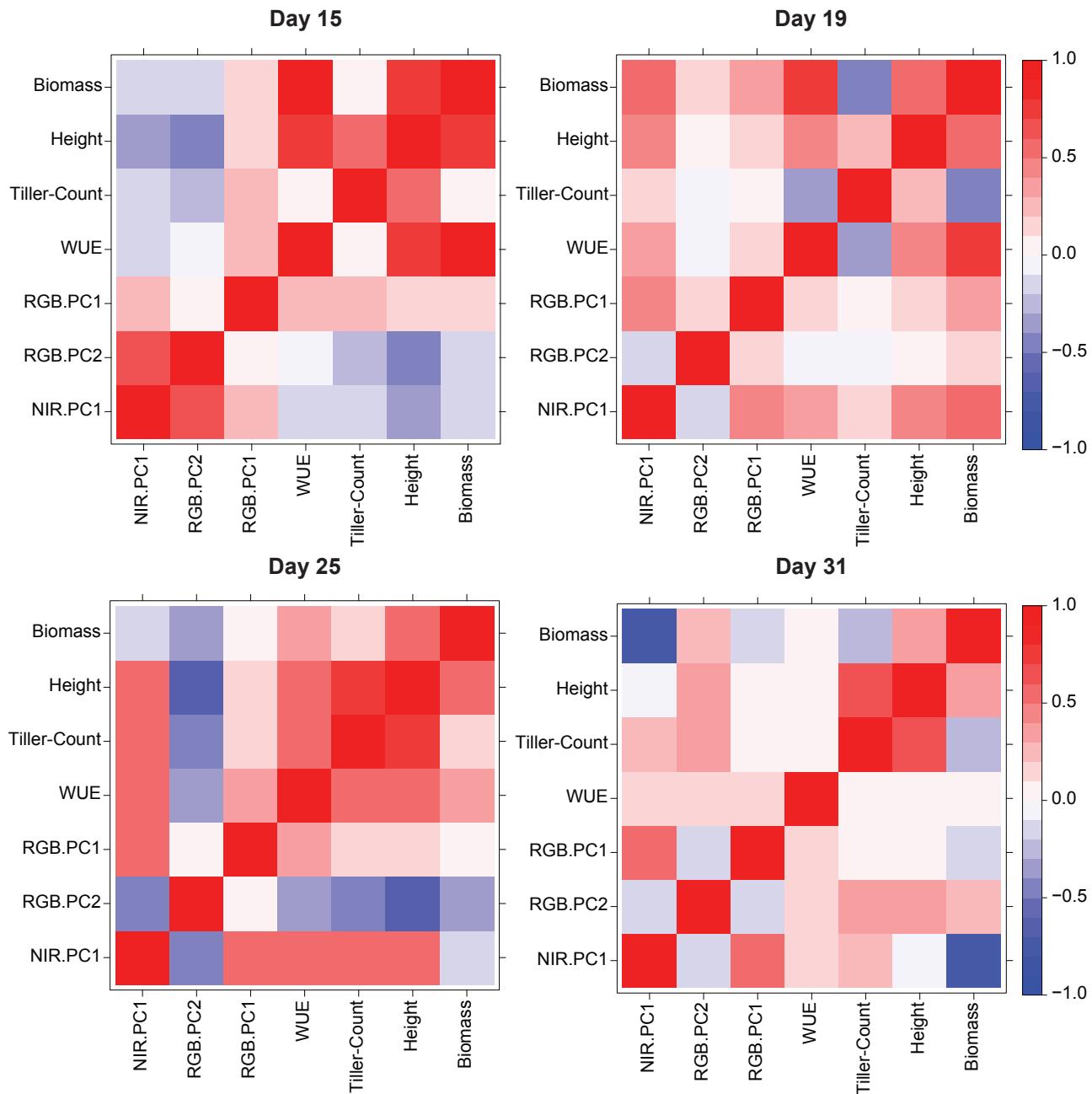
**Supplemental Figure S7.** Before water treatment is applied (11 to 12 DAP) color can distinguish *Setaria* genotypes in RGB color (Fig. 6), which is also consistent in HSV and LAB color space. A, PCA of HSV (left) and LAB (right) color for all *Setaria* genotypes before treatment is applied (11 to 12 DAP). PC1 and PC2 are plotted for *S. viridis* (orange), *S. italica* (gray), RIL161 (purple) and 7 other *S. viridis* x *S. italica* RILs (gray). B, PCA of HSV (left) and LAB (right) color for all *Setaria* genotypes before treatment is applied (11 to 12 DAP). PC1 and PC2 are plotted for *S. viridis* (gray), *S. italica* (green), RIL161 (purple) and 7 other *S. viridis* x *S. italica* RILs (gray).



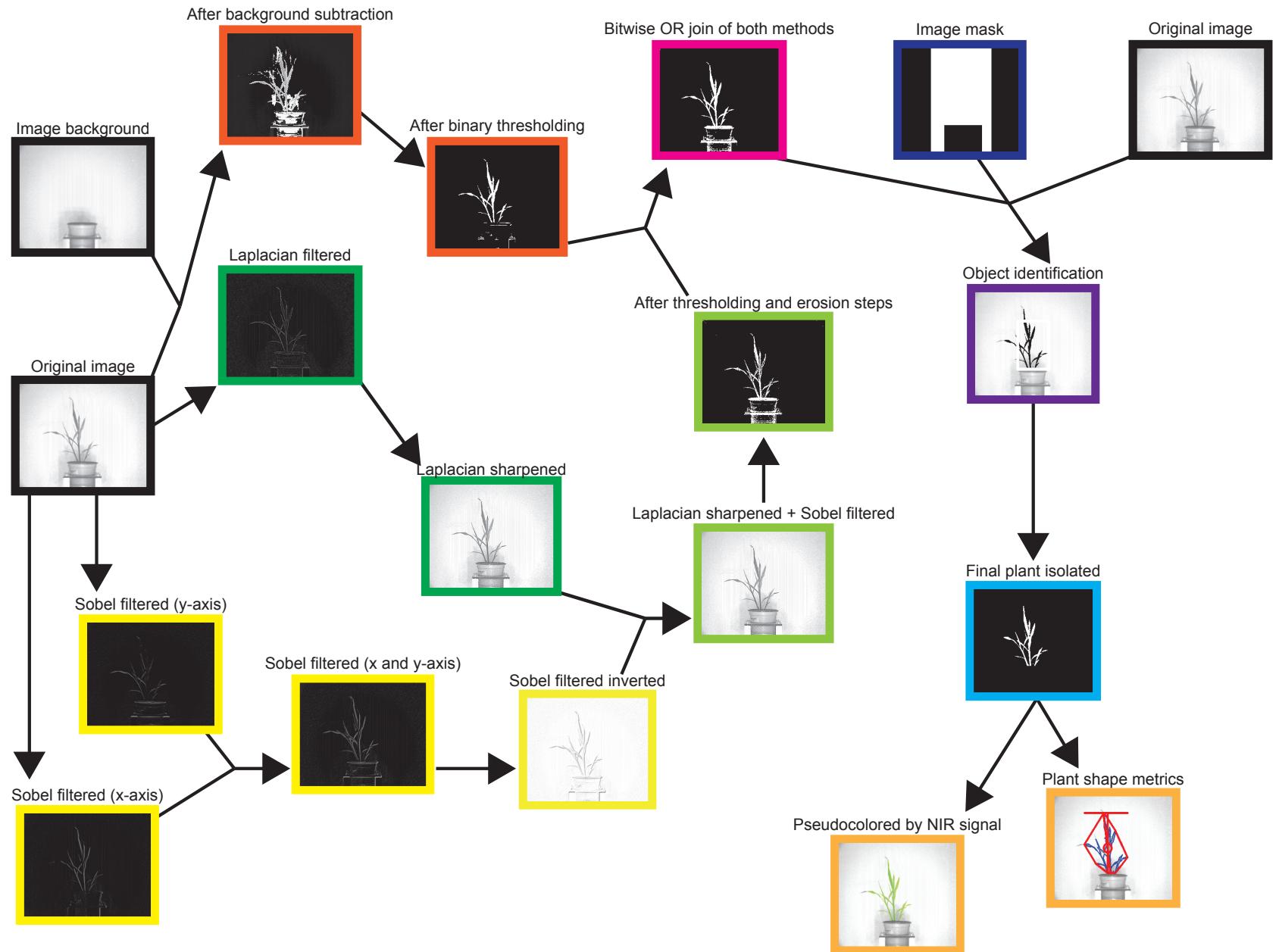
**Supplemental Figure S8.** Loadings of color data 11 to 12 DAP. A, PC1 loadings of *S. viridis* color data PCA 11 to 12 DAP. B, PC2 loadings of *S. viridis* color data PCA 11 to 12 DAP.



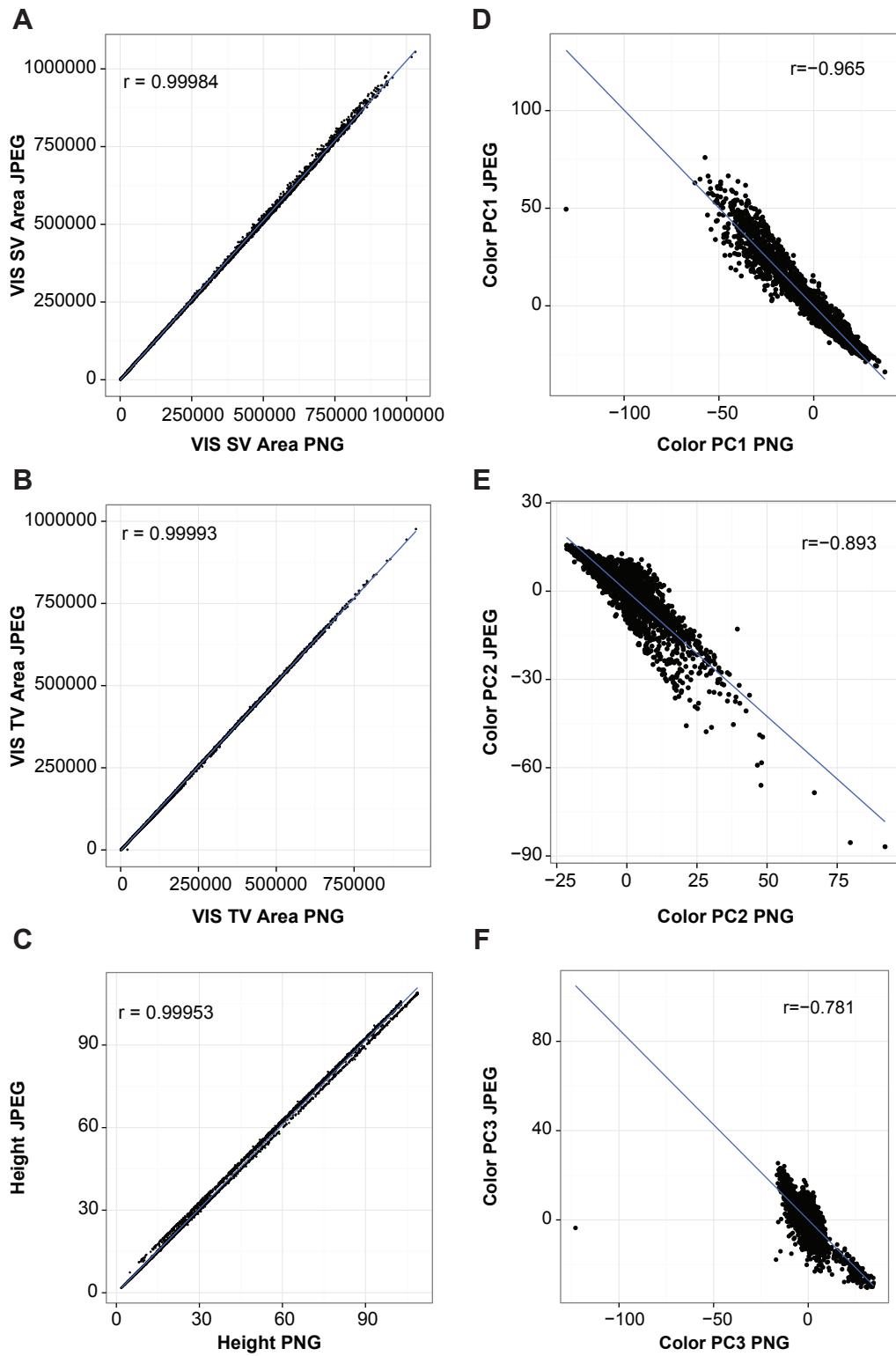
**Supplemental Figure S9.** Height significantly correlates with median Fv/Fm. A, Median Fv/Fm is significantly negatively correlated ( $p < 0.01$ ;  $\rho = -0.852$ ) with estimated plant height. Average estimated height was extracted from four side-view VIS images by PlantCV from all ten *Setaria* genotypes over time under both 100% FC and 33% FC water treatments. Fv/Fm values for each plant pixel was calculated by PlantCV from F0, Fmin and Fmax PSII images. Median Fv/Fm is the median bin value in the Fv/Fm per pixel histogram for each plant. B, 3D-printed ‘plant’ components consist of two 5 cm ‘stem’ pieces and one flat ‘leaf’ platform piece. Designs for 3D-printed ‘plant’ components are available at <https://www.thingiverse.com/thing:574287>. C, Assembled 3D-printed plant with excised *N. benthamina* leaf tissue. *N. benthamina* leaf was positioned at 0 cm (flush with gravel in pot), 5 cm, 10 cm, and 15 cm. D, Median Fv/Fm is significantly negatively correlated ( $p < 0.01$ ;  $r = 0.976$ ) with height in the 3D-printed plant. 3D-printed plant with excised *N. benthamina* leaf tissue was imaged by the PSII imaging station two times (tissue from a second leaf during second imaging). Fv/Fm values for each plant pixel at each platform level was calculated by PlantCV from F0, Fmin and Fmax PSII images. Median Fv/Fm is the median bin value in the Fv/Fm per pixel histogram calculated for plant tissue at each platform level.



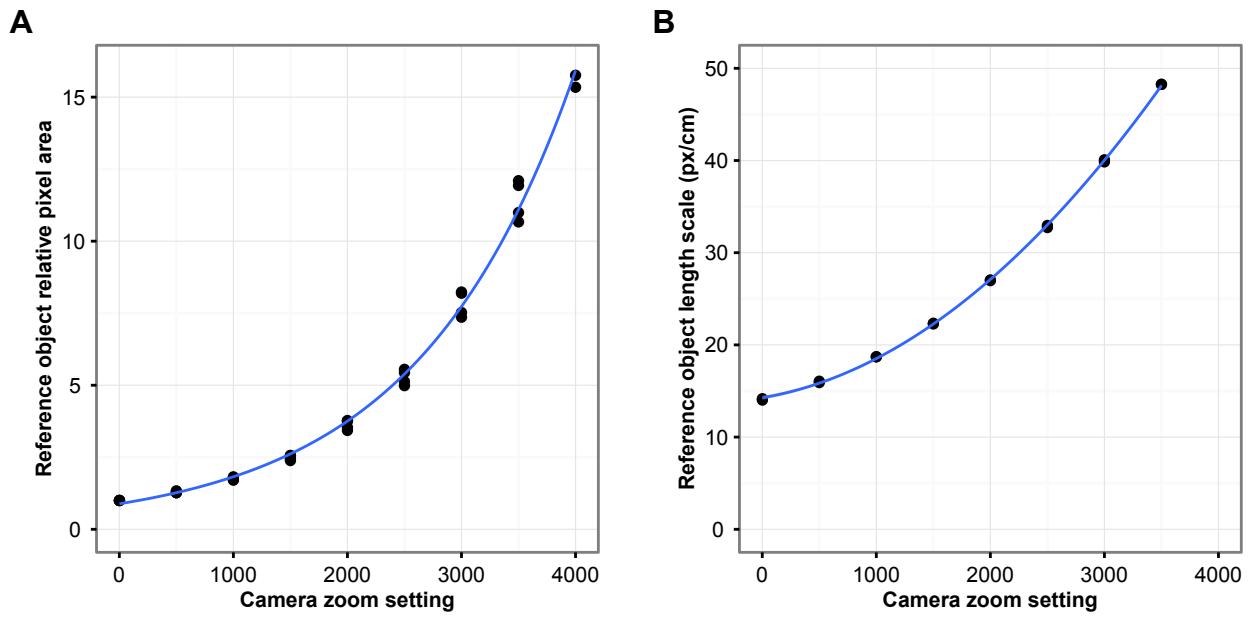
**Supplemental Figure S10.** Matrices displaying Pearson's Correlation Coefficient calculated between traits at four roughly equidistant time points. Darker red shading indicates a stronger positive correlation, whereas increased blue shading indicates stronger negative correlation.



**Supplemental Figure S11.** A visual illustration of the pipeline used to threshold plant tissue from background within grayscale NIR images. Color boxes group together sets of related processes (Original images = dark blue; Original images processed with background subtraction = orange; Laplacian filtering = dark green; Sobel filtering = yellow; Combined derivative filtering = light green; bitwise OR join = fuchsia; image masking and object identification = purple; final processed images = light blue).



**Supplemental Figure S12.** JPEG images are sufficient for analysis when only shape and morphological features are needed. PNG format images are preferred when signal/spectral data is analyzed. A–F, Correlation analysis included all *Setaria* genotypes and the 100% and 33% FC water treatment groups. A, VIS side-view plant pixel area. B, VIS top-view plant pixel area. C, Estimated plant height. D, PC1 of RGB color PCA. E, PC2 of RGB color PCA. F, PC3 of RGB color PCA.



**Supplemental Figure S13.** Zoom-correction scaling factors for pixel area and pixel length. A, Zoom-correction scaling factor model for pixel area calculated from a reference object of known area. B, Zoom-correction scaling factor model for pixel length calculated from a reference object of known length.