

# Chapter 11

## Got All the Answers! What Were the Questions? Avoiding the Risk of “Phenomics” Slipping into a Technology Spree



Vincent Vadez, Jana Kholova, Grégoire Hummel, and Uladzimir Zhokhavets

**Abstract** For many crops, the genomics revolution has given hope that breeding would become easier, faster, and more efficient. Relevant phenotyping is now the main bottleneck and new technologies provide opportunities for easier, faster, more sensitive, and more informative phenotyping. However, the phenotyping agenda must be driven by scientific questions rather than by a technological push, especially for complex constraints, such as drought. In this chapter, we provide a viewpoint on phenotyping and what it should take into account. Phenotyping is a full-fledge research effort, calling for a multidisciplinary effort between technology providers and several research disciplines, and which needs to address the issue of linking scales. Two phenotyping platforms are described; a lysimetric platform (LysiField) to assess the patterns of plant water use and relate these to grain yield, and an imaging platform (LeasyScan) to characterize crop canopy traits responsible for water savings. In both cases, the chapter discusses the thought process and the hypotheses around key traits for drought adaptation that were put in the development of these platforms. The chapter concludes with perspectives on the integration of high-throughput phenotyping (HTP) technology with breeding, starting with an analysis of the cost as a prerequisite to decide on its usage and adoption in breeding. It takes a few examples of current opportunities in the domain of imaging, trying to bring closer together what the technology can bring and what breeding pragmatically needs. In conclusion, while new technologies provide opportunities to make phenotyping easier, faster, better, cheaper, the risk of becoming the end that justifies the means can be avoided by driving the technology with research questions, made

---

V. Vadez (✉)

Institut de Recherche pour le Développement (IRD), UMR DIADE, University of Montpellier, 911 Av Agropolis BP65401, 34394 Montpellier, France  
e-mail: [vincent.vadez@ird.fr](mailto:vincent.vadez@ird.fr); [v.vadez@cgiar.org](mailto:v.vadez@cgiar.org)

J. Kholova

Crop Physiology Laboratory, ICRISAT, Greater Hyderabad, Patancheru 502324, Telangana, India  
URL: <http://www.icrisat.org>

G. Hummel · U. Zhokhavets

Phenospex, Jan Campertstraat 11/NL-6416 SG, Heerlen, The Netherlands  
URL: <https://www.phenospex.com>

© Springer Nature Switzerland AG 2021

J. Zhou et al. (eds.), *High-Throughput Crop Phenotyping*,  
Concepts and Strategies in Plant Sciences,  
[https://doi.org/10.1007/978-3-030-73734-4\\_11](https://doi.org/10.1007/978-3-030-73734-4_11)

223

possible through a cross-discipline approach between genetics, breeding, modeling, engineering, physiology, and statistics.

**Keywords** Multi-discipline trait · Drought · Breeding · Genetic gain · Transpiration

## 11.1 Introduction

Although we know more and more about the "genotype," obtaining information on the "phenotype" remains a challenge due to the complexity of biological systems and the requirement for physical measurement of plant traits that are difficult to perform quickly. Obtaining phenotypic information that is relevant, accurate, fast, repeatable, and manageable in large numbers has been, and will remain, a basic challenge of any breeding program. There is also a need, and may be an opportunity with new technologies, to move away from the very "integrated" phenotypes that have been the bulk of phenotyping so far (e.g., yield, biomass, height) towards the causal building blocks of these integrated phenotypes. In the past few years, a revolution in plant phenotyping is taking place and technological progress has made possible an increase in the throughput and precision of phenotyping. We argue that the current phenotyping revolution while creating fantastic opportunities to capture phenotypic traits quickly and non-destructively, is also running the risk of becoming driven by the technology itself, rather than being driven by research questions around the critical plant traits to phenotype. This chapter analyzes the opportunities and challenges facing this phenotyping revolution, presents two phenotyping platforms that break up phenotypes in smaller building blocks, and addresses the link to breeding applications and forthcoming opportunities.

The first section is a viewpoint on the principles that should be applied to plant phenotyping. At present, phenotyping is seen as a tool to generate data for the breeding community. Contrary to this view, we argue that phenotyping is a full-fledged scientific approach that requires a careful analysis of the traits to be measured and their relevance for the targeted constraint (especially for complex traits). This calls for a multidisciplinary approach if "phenomics" is to be relevant for crop improvement, a view that is shared by many others (e.g., Deery et al. 2014; White and Snow 2012; Araus and Cairns 2014). In this section, we discuss the fact that some phenotypes are "consequential" (for instance, staygreen), whereas others are "causal" (for instance, leaf developmental traits). This section then explores the challenges and opportunities of linking information at different levels of plant organization, i.e., from either specialized phenotyping platforms, targeting predominantly traits at a lower level of plant organization, up to the field for agronomic trait phenotyping. The issue of scale in phenotyping is addressed by bringing up crop simulation modelling as an integration tool to bridge these scales, advocating linkages between trait-based and field-based evaluations of genotypes (Chapuis et al. 2012).

The following section presents the LysiField and LeasyScan platforms. The LysiField platform is a lysimeter platform that measures plant water use over the entire cropping cycle, instead of measuring roots per se, with a throughput of about 500–600 cylinders weighted per day since it remains a manual operation. The LeasyScan platform is a three-dimensional (3D) laser scanning system that generate 3D images from which several canopy features are extracted, like the leaf area, and that includes a lysimeter component for an automatic pot weighing to assess plant transpiration. This section will detail the thought process and research questions that led to the development of LysiField, how these research questions have shaped how the platform stands (Vadez et al. 2014). It then goes on describing how the knowledge gained from LysiField generated new research questions that have led to the development of another phenotyping platform (LeasyScan) to measure traits at a lower level of biological organization (crop canopy traits) more quickly and precisely. This section presents the principles of the scanning operation plant transpiration in situ. Scans are obtained at a high rate (approximately 2500 plots scanned per hour) on several parameters per plant, while tray weights are polled every 15'. Data management then becomes a major challenge (Cobb et al. 2013) and a description is given of the data handling process. This section presents the web-based interface that is used to visualize the data, and the data management tools are used to query data from the database and initiate the data analysis. We also discuss critical planning aspects during the development of a phenotyping platform such as the need to test the technology prior to acquisition and the need for a close user-provider relationship during and after the development of platforms.

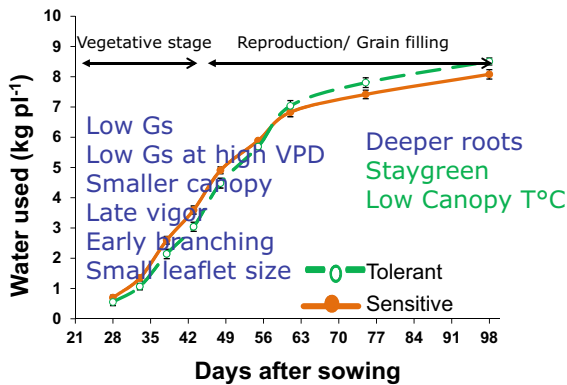
The last section addresses phenotyping costs, and how this becomes a critical factor in the adoption of modern and technology-intensive methods in the breeding process. This section also presents a few examples using imaging technology to mirror what the technology can provide and what the immediate needs of breeding programs are.

## **11.2 Phenotyping: Basic Principles**

### ***11.2.1 Phenotyping is a Research Approach***

Understand the basic biological and physiological processes behind phenotypes is critical, especially for complex phenotypes. For instance, canopy temperature measurements can be used as a proxy phenotype for the transpiration capacity of genotypes. However, done at a late stage in a crop exposed to drought, this transpiration capacity could reflect: (i) the capacity to extract water from deep soil layer thanks to deeper rooting, or (ii) the fact that there is water remaining in the soil profile. In turn, the latter could be the consequence of a slower water use at earlier stages and have different causes, including (ii-a) a smaller leaf canopy size; and/or (ii-b) a lower canopy conductance under certain conditions. A smaller canopy size

could come from reduced tillering or branching, a lower leaf number, a smaller leaf or leaflet size. This example illustrates how a phenotype can be explained by a cascade of possible other phenotypes reflecting biological processes underneath. Similarly, the expression of a staygreen phenotype is actually the consequence of several phenotypes having contributed to differences in plant water use much earlier (Borrell et al. 2014). The concept of “phenes,” i.e., the phenotypic equivalent of genes (Lynch and Brown 2012), representing building blocks of more complex phenotypes, with a cause/consequence order between phenes measured at different levels of plant organization, reflects the difficulty and complexity of choosing the “right” phene (Fig. 11.1). As such, several “causal” phenes could influence a “consequential” trait (for instance, tillering and leaf size on staygreen (Borrell et al. 2014)), or how phene-to-phene interactions could influence a “consequential” trait (e.g., leaf size and leaf thickness effect on transpiration rate (Kholova et al. 2012), in a similar manner than pleiotropy and epistasis in genetics. We think there is no alternative to carefully ordering the cause-consequence structure of phenes to make phenotyping relevant and useful to genetics and breeding. Therefore, phenotyping is not only about generating trait data using well-set protocols, it is truly a scientific approach that involves the deciphering of complex biological cause/consequence relationships in a phenotype.



**Fig. 11.1** Profile of water extraction under terminal drought conditions in a set of terminal drought tolerant ( $n = 12$  lines) and sensitive ( $n = 8$  lines) chickpea genotypes having a narrow range of flowering time (re-drawn from Vadez et al. 2014—J Exp. Bot. <https://doi.org/10.1093/jxb/eru040>). Tolerant genotypes extracted less water during vegetative stage and more during reproduction and grain filling. The figure lists possible “causal” canopy phenotypes (in blue), and “consequential” phenotypes (in green), possibly explaining the differences in the patterns of water use

### ***11.2.2 Research-Driven, not Technology-Driven Development of Phenotyping Platforms***

In the past decade, the capacity to image plants has progressed dramatically. This includes simple digital measurement with RGB (Red-Green-Blue) cameras, more specific measurement of temperature with Infrared (IR)-cameras, more complex multi-spectral sensors including the simpler versions to assess vegetative index (e.g., Normalized Difference Vegetation Index (NDVI)), and highly complex hyperspectral or fluorescence measurements. The revolution in phenotyping tools offers both a terrific opportunity and also a major challenge. The opportunity is to extract information from genetic material on “phenotypes” that are non-visible to the human eye, visible but too complex to be measured from simple observations, or new phenotypes that were not considered before. However, the risk is of losing perspective on the target phenotype in favor of the many phenotypes that can be acquired and may be unrelated to the target phenotype.

### ***11.2.3 People-Skills, Cross-Discipline Interactions***

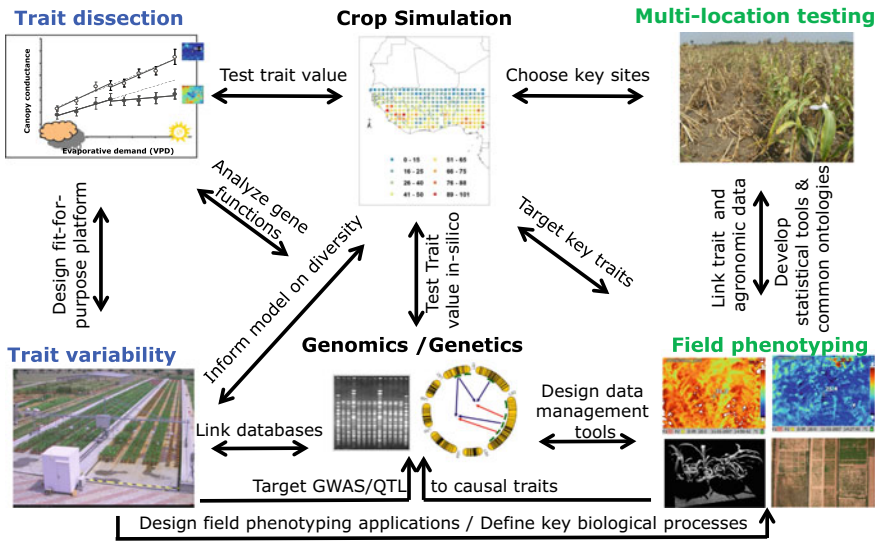
Phenotyping is also a cross-road where technology providers, physiologists, pathologists, geneticists, breeders, data analysts, statisticians, and others, come to interact. It calls for a multidisciplinary effort. For instance, scientists working in the area of stomata patchiness (stomata at the leaf level are regulated by patches that are interconnected) have provided evidence of a close relationship between these processes and those in computation (Peak et al. 2004), and this was only possible because biologists interacted with computer scientists. Linking phenomics information to genomics is a first step as we learn more about “phenes.” However, there is much to be done to make this link workable and useful, first in terms of data format, suitable databases, meta-data informing trials, ontologies, and statistical tools to analyze complex data (Cobb et al. 2013) or multi-trait analysis (Brown et al. 2014; Korol et al. 2001). Linking these dimensions goes beyond finding a technical fix to connect these spheres of information: It is about co-designing the linkages across disciplines and their technical features so that the linkage can be truly functional, leading to new and relevant knowledge. For instance, designing marker-trait analysis that takes into account environmental conditions as a covariate, or discussing population size beforehand to avoid logistical constraints of phenotyping populations larger than, say, 500 individuals, or defining what precision is needed in the measurements. Last but not least, a very close and iterative interaction between technology suppliers and biologists is needed to ensure the phenotyping platforms/sensors truly address the phenotyping needs.

### ***11.2.4 An Issue of Scales: Combining Platform- and Field-Based Phenotyping***

Earlier, we discussed the importance of deciphering phenotypes at different levels of biological organization, and then of structuring phenotypes into cause-consequence relationships. This is where specialized platforms have a role to play, in assessing the variation for critical causal traits and harnessing the genetics of these building blocks. Because such platforms cannot be developed in all breeding programs, there is a need to have a connection between platform-based trait phenotyping that the breeder can access in a simple and high-throughput manner and field-based phenotyping. The connection between trait and field phenotyping can also be established when traits can be measured in the field itself by imagery sensors, and then linked to agronomic assessment, e.g., grain or stover yield, in a network of testing locations. A few recent papers describe a number of such applications for trait phenotyping in the field (Araus and Cairns 2014; White et al. 2012; Deery et al. 2014), including tractor-based supporting devices and airborne devices. For instance, genotypic differences in the response of leaf expansion to vapor pressure deficit (VPD) (e.g., Welcker et al. 2011) could be proxied by NDVI measurements in the field. In this case, NDVI assessment in the field would integrate over time the cumulative effects of a physiological process that can be measured in a specialized platform. These applications also need to monitor the degree of causality/consequence of the different phenotypes that are measured (Fig. 11.1). The next step as we move up in the degree of plant integration is to establish links with agronomic assessments, where it can be tested in which environment a given trait, measured in a specialized platform, would have an effect on yield. Taking again the example of the sensitivity of leaf expansion to high VPD, it was shown that this trait was correlated to the drought sensitivity index measured in the field (Chapuis et al. 2012). In short, there is a great prospect for using the information from specialized platforms to inform and enrich field-based phenotyping application and use these traits in selection (see Fig. 11.2).

### ***11.2.5 Linking Phenotyping to Crops Simulation modeling***

Once traits benefitting crops under certain water stress patterns have been identified, testing their effects of traits via experimental means is restricted to a few traits at a time and a few environmental and climatic scenarios. In addition, the complexity in the resulting phenotypes originates from the interaction among traits and from their interactions with the environment (Buckler et al. 2009; Schuster 2011). This is where crop models can serve to “integrate” complex behavioral/developmental processes of plants that are all related through water need/use (Fig. 11.2). Models that are suitable for this must contain algorithms that reflect observable and quantifiable biological observations (Sinclair and Seligman 2000; Hammer et al. 2010). This is only then that models can be sensitive to changes in the conditions and can accurately predict



**Fig. 11.2** Schematic linkage relationships between disciplinary effort toward crop improvement. The left part of the schema deals with trait dissection variability (blue), and how specific platforms are designed to phenotype for critical traits at a large scale. The top left graph displays genetic variation for a water-saving trait (Kholova et al. 2010). The bottom right picture presents the LeasyScan platform (Vadez et al. 2015). The central part of the schema represents the interface with crop simulation and genetic analysis (black). The top central map represents the yield increase arising from the modification of a genetic trait, displayed in form of model output in  $1^\circ$  latitude  $\times$   $1^\circ$  longitude averaged across 50 years of weather information (Vadez et al. 2017). The right part of the schema are the field applications of phenotyping (green), where consequential phenotypes are measured along with agronomic traits. The top right picture represents the expression of a staygreen phenotype in sorghum lines introgressed with a staygreen QTL. The bottom right pictures represent different possible applications to field phenotyping. The different arrows represent the main linkage relationships between disciplinary domains and indicate the type of actions needed to make the links functional

effects. There is now convincing evidence that crop models are relevant to guide breeding targets (Kholová et al. 2014; Vadez et al. 2012; Reynolds et al. 2018). Using a mechanistic crop model, Soltani and colleagues (1999) showed that an early decline in leaf expansion and transpiration upon soil drying in chickpea led to about 5% yield increases under water stress conditions. Therefore, these traits had a limited interest where they were tested and did not justify an investment in breeding. In another study with chickpea, a rapid root growth rate decreased yield by an average of 5%, whereas an increase in the depth of root water extraction by 20 cm increased yield by an average of 10% (Vadez et al. 2012). This example shows the efficacy of a model for comparing genetic options, before deciding what to possibly invest in. In the last example in sorghum, the capacity to restrict transpiration under high VPD was simulated and showed yield advantage in all situations where it was tested, yet with higher effect in zones facing severe water stress (Kholová et al. 2014). The modeling approach is powerful because it is now possible to simulate the effects of

certain Quantitative Trait Loci (QTLs) on yield, based on the percentage effect of a given QTL on particular traits (Chapman et al. 2003; Welcker et al. 2007; Chenu et al. 2009; Cooper et al. 2009). We believe that investment in HTP platforms could be guided by prior crop simulation of the value of the trait that is targeted in these platforms.

## 11.3 Phenotyping Platforms

### 11.3.1 Lysimetric System to Assess Plant Water Use

Roots are intuitively basic for crops and especially for the adaptation of a crop to water deficit because nutrients and water are absorbed through them. However, they are difficult to work with (Vadez 2014). For water stress research, the root capacity to extract water was the basis of the idea to develop a lysimeter-based system (Vadez et al. 2008, 2014), in which consecutive weighings of lysimeters provide data on plant water extraction to support transpiration at different times (Fig. 11.3). Because the goal was to measure water use in crops grown under field conditions (something that is difficult to do precisely in the field) and with a high throughput, certain basic principles had to be followed. The platform was set outdoors and the tubes were designed and placed so that soil volume and surface area were similar to field population densities. Therefore, two types of tubes were developed to cater for different crops: Small lysimeters (1.2 m length and 20 cm diameters) were designed for crops sown at approximately 20 plant  $m^{-2}$  like chickpeas (Zaman-Allah et al. 2011b), whereas the large lysimeters (2.0 m length and 25 cm diameters) were designed for crops sown at approximately 10 plant  $m^{-2}$  (Vadez et al., 2011, 2013a). Lysimeters were also treated as micro-plots and kept undisturbed from one crop to the next, following a field-like rotation, alternating either experimental or fallow crops.

**Fig. 11.3** Overview of the lysimetric platform at ICRIASAT (LysiField), showing the large tubes (25 cm diameter, 2.0 m length), which are set in trenches and allow a planting density of about 10 plant  $m^{-2}$ . A pigeon pea crop is seen on the left trench and a sorghum crop in the central trench





The lysimetric platform was originally designed to screen genotypes for the capacity to extract water from the soil profile, instead of measuring roots. Genetic variation for total plant water extraction was found in all species that were tested (for a review see Vadez et al. 2014). However, the range of variation (30% among a subset of the sorghum reference collection) was not related to yield differences under stress conditions. The relationship between water extraction during the grain filling period and grain yield under stress conditions was much more critical (see Vadez et al. 2014 for a review), e.g., in pearl millet (Vadez et al. 2013a) chickpea (Zaman-Allah et al. 2011b) or peanut (Ratnakumar et al. 2009). For at least three crops, the availability of water during the grain filling period was not related to a higher capacity to extract water, but to earlier water-saving under non-stress conditions. For instance, tolerant chickpea genotypes had a smaller leaf canopy at the vegetative stage (Zaman-Allah et al. 2011a). Tolerant peanut genotypes also developed a smaller leaf canopy (Ratnakumar and Vadez 2011) and tolerant pearl millet had both a lower canopy conductance and the capacity to further reduce the conductance under high VPD conditions (Kholova et al. 2010). There have been similar findings in other crops such as cowpea (Belko et al. 2012) and sorghum (Borrell et al. 2014). In short, even if the cylinders were weighted only about once a week, the lysimetric system provided sufficient precision to pinpoint small but critical differences in the patterns of plant water use. From then on, the focus shifted towards traits that explain these small water use differences and influence the rate at which a crop uses the soil profile's moisture, including (i) canopy size and dynamic of canopy development; (ii) canopy conductance; and (iii) canopy conductance under high VPD (see a review in Vadez et al. 2013b). In other words, specific patterns of plant water use were a "consequential phenotype," and further attention shifted to the "causal phenotypes," which required a different type of measurement (see Fig. 11.1 for an example in chickpea).

### ***11.3.2 The LeasyScan Platform: 3D Scanning Plus Transpiration Assessment***

#### **11.3.2.1 Description**

LeasyScan's principle is to have a continuous and simultaneous monitoring of plant water use and leaf canopy development. In brief, the platform is using a set of scanners (PlantEye, Phenospex, Heerlen, Netherlands) which are moved above the plants using a carrier device and generate 3D point clouds of the crop canopy, from which the leaf area and several other plant parameters are extracted after a segmentation process of the 3D data cloud (Fig. 11.4). Validation of scanned leaf area versus observation has been successfully done before acquiring the equipment and has been re-validated later on while working on higher planting densities (Figs. 4 and 5 in Vadez et al. 2015). Leaf canopy development traits that influence plant water use are a combination of



**Fig. 11.4** Overview of the LeasyScan platform at ICRISAT. A groundnut crop is seen on side strips. The central strip shows the installation of load cells to allow the continuous weighing of the pots. Eight scanners (small white boxes) can be seen, attached to an irrigation boom that travels over the crop, on top of a center and two side walls. The central metal box has a key role to ensure a steady platform movement

(i) vigor, i.e., how quick the canopy develops; and (ii) size, i.e., how large a canopy develops (Fig. 1 of Vadez et al. 2013b).

The PlantEye sensor projects a very thin laser line in the near infrared (NIR) region of the light spectrum (940 nm) on plants and captures the reflected light with an integrated complementary metal oxide-semiconductor (CMOS)-camera. Since most of the light is reflected from plants, the device can operate day and night. All artifacts from sunlight or background noise are automatically removed with inter-graded optical- and algorithm-based sunlight filters. During the scanning process, the scanner linearly moves over the plants and generates 50 height profiles/s, those are then automatically merged into a 3D point cloud with a resolution of around  $0.8 \times 0.8 \times 0.2$  mm into the xyz-direction, respectively. The measurements are triggered and stopped via mechanical barcodes (metal plates 20 mm  $\times$  50 mm) positioned on the platform. PlantEye computes a diverse set of plant parameters on the flight by meshing neighboring points with a nearest neighbor search. From this triangle mesh a subsequent surface triangulation algorithm computes 3D leaf area (which is the area of the leaf independently of its position and orientation in the 3D space and relative to the sensor), plant height, leaf angle distribution within a second.

At the LeasyScan platform, the scanners are pre-set to image an area of 65 cm width and a length of either 40 or 60 cm. The volume in which the 3D image is generated is then a cuboid of 65  $\times$  40  $\times$  100 cm or 65  $\times$  60  $\times$  100 cm. Each scanning unit is referred to as a “sector.” Every 12 consecutive sectors constitute a “field.” Sector-wise binning of data point clouds is performed using a system of barcodes every 5 m (12 times 40 cm + 20 cm gap or 8 times 60 cm + 20 cm gap) to re-set the scanner position in height and length. As in the lysimetric facility, our choice was to remain as close as possible to the field conditions where plants are cultivated in each sector at a density similar to the field (for instance 24–32 plant  $m^{-2}$

chickpea or 16 plant  $\text{m}^{-2}$  for pearl millet or sorghum). The scanners are mounted on top of an irrigation boom, which is electronically controlled to be fully automated and speed-controlled. At a movement speed of  $3 \text{ m min}^{-1}$ , eight scanners are capable of scanning 4800 sectors (the name of an experimental unit) in slightly  $<2 \text{ h}$ .

These parameters can be visualized through a web-based software interface (HortControl<sup>R</sup>), which allows the selection of sectors and performs basic grouping functions to assess how the experiment is progressing. In addition, the platform is equipped with a set of 12 environmental sensors (Campbell Scientific, Logan, Utah, USA) that continuously monitor relative humidity (RH%) and temperature ( $^{\circ}\text{C}$ ), integrating values every 30 min, one light sensor, one wind sensor. Each scanner is wirelessly connected to a local area network (LAN) through which the analyzed data are downloaded onto a server, along with the 3D images. 3D images can be reused at any time; for example, to re-calculate new parameters based on a new algorithm for additional plant traits or for better-optimized scanning software. Therefore, the scanning images become a repository of plant measurements that can be reused at a later date. An important factor to decide on the scanning system was to understand the signal-noise ratio for our targeted phenotype (leaf area), and then check not only the resolution of the sensor itself but also the noise of the environment, e.g., wind, diurnal rhythm of leaves, rain, reflection.

### 11.3.2.2 Integration of Canopy Growth with Plant Transpiration

A basic necessity in the development of the LeasyScan platform was to combine the measurements of leaf development parameters (which can be encapsulated in “volumetric growth”) with a continuous assessment of plant transpiration (or “massic growth” considering transpiration as a proxy for photosynthesis), to obtain a continuous measurement of the canopy conductance and shift from earlier destructive measurements (Kholova et al. 2012). In earlier studies, a low canopy conductance under high VPD was closely related to terminal drought stress adaptation in several crops (see a review in Vadez et al. 2014), but this phenotype depended on time-consuming leaf area measurements, especially in a crop like chickpea (Zaman-Allah et al. 2011b). One part of that phenotype, the leaf area, is described above. The other part of that phenotype, plant transpiration, is typically measured manually by gravimetrically determining transpiration (e.g., Zaman-Allah et al. 2011b). Using scales (also called load cells) then allowed to have a continuous weighing of the pots, avoiding time-consuming weighing of pots. Notably, in the development of this platform, we also sought the possibility to study intra- or inter-specific variations in crop water loss during the night (following recent results in wheat (Schoppach et al. 2014)), the interaction between water use and the 3D architecture of the crop canopy, possible relationships between leaf movements during the day (especially in legumes or, for example, in *Arabidopsis* (Dornbusch et al. 2012)), patterns of plant water use during the day, and of course the interplay between volumetric (leaf area dynamics) and massic (transpiration) growth.

The scales (PSX Rugged Scale 50, Phenospex, Heerlen, Netherlands) that were initially used had a capacity of 50 kg, with 0.02% accuracy. The accuracy of these temperature-corrected scales ( $-10^{\circ}\text{C}$  and  $+40^{\circ}\text{C}$  range) was tested under artificial rapid increase in temperature ( $14^{\circ}/\text{h}$ , i.e., much above our experimental conditions) and showed that the error remained within the stipulated 0.02% error range. The scales provided a reading with a 0.02% precision every second and these were integrated over one hour, giving readings with a precision of 0.1 g. An initial prototype of scale was developed where the frequency of measurements was limited to one every hour. After validation that key phenotypes like the capacity to restrict transpiration under high VPD could be measured (see Fig. 11.9 in Vadez et al. 2015), 1500 load cells were installed, each with a 150 kg capacity (Fig. 11.4). This increased capacity now allows to grow plants on large trays ( $60\text{ cm} \times 40\text{ cm} \times 30\text{ cm}$ , length-width-height) containing about 90 kg of soil and allowing to grow several plants in conditions that mimics the field. It also allows to irrigate at less frequent intervals, yet maintaining plants away from water stress.

### 11.3.2.3 Data Generation, Storage, and Visualization in HortControl

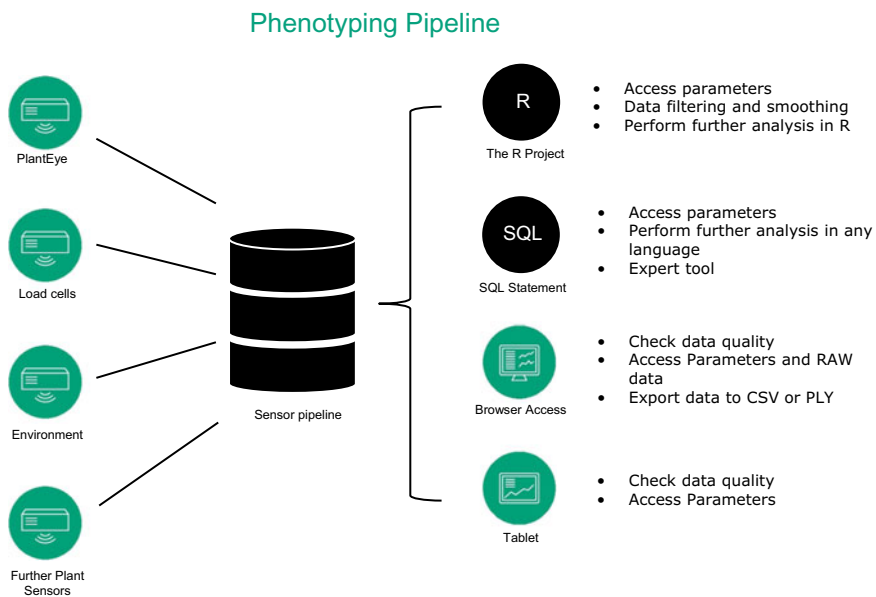
Scanning takes place every 2 h so that about 50,000 scans are captured every day and about 10 traits are calculated for each scan. This is in addition to the environmental data that are measured every minute. With regards to load cells, these are polled by a micro-processor every half-second and these data are integrated and loaded on the database every 15 min so that each load cell delivers 96 data points every day. All data gathered from PlantEye sensors, scales, and associated climate sensors are stored in a central PostgreSQL database. The data can be accessed and visualized with the web-based HortControl software that allows to follow up the progress in the different parameters that are measured. The three types of data collected in the platform, i.e., scan, weight, climate, are not collected at the same frequency. Therefore, the data are downloaded independently. R-scripts have then been developed to aggregate data at a time scale suited for all three types of data (Kar et al. 2020a). HortControl is used as a data visualization tool to monitor the experiment, for instance, to ensure scan data are properly computed, or possibly to detect load cell errors. In particular, the 3D image of any sector at any time during the experiment can be called for quality control, which is particularly useful to pinpoint possible outliers (for example, in case of sector to sector overlapping or other unexpected disturbance). It also allows the simultaneous plotting of the environmental conditions to the parameter evolution, for instance, to qualitatively estimate reasonable wind thresholds in each species.

### 11.3.2.4 Database Access, Processing, and Analysis

One major challenge of this platform, and of any high-throughput platform, is the analysis of the data. This issue was discussed at length in a recent review (Cobb et al. 2013). At the same time, well-documented datasets represent a potential treasure

trove to investigate plant growth processes on a large scale (for example, in meta-analysis (Poorter et al. 2010)). In that regard, we focused on linking measurement data with the most critical environmental parameters affecting plant growth (i.e., temperature, relative humidity, and light). It is also critical to have detailed meta-data accompanying the datasets if they are to be re-used in the future.

Processed data (e.g., leaf area) can be downloaded from a function in HortControl. These data are queried from the database via an R-command library interface working at the back-end (R, version 4.2.1, the R foundation) (Fig. 11.5). Among the essential features of the library is a process for smoothing data and for filtering the data to reject outliers. For instance, wind affects the quality of the 3D images. Data obtained when the wind is too high to have useful information (from shaky images) should be filtered out. The scanning data are tagged to the timing of each scan so that the time stamp can be linked to the environmental data provided. A data pre-processing and analysis pipeline has now been recently developed (Kar et al. 2020b) for scanning data, which allows to apply filters on raw data towards outlier detection, to input missing data, to choose for an optimal time window for genotype discrimination, and for spatially adjusting data. A similar pipeline has been developing to extract features from the transpiration profiles coming from the load cells and that characterize the transpiration response to high VPD conditions (Kar



**Fig. 11.5** The phenotyping processes flow; the information generated by the PlantEye<sup>R</sup> scanners along with the information from the environmental sensors are connected to one another through time stamps and stored in the data repository system. Data can be called from the repository through various means: using R software, web-based interface browser (Hortcontrol<sup>R</sup>), and is also compatible with other analytical tools (SQL)

et al. 2020a). We are also planning to develop an alternative time stamp, right from the "R" interface, calculated from the temperature conditions and based on either thermal time or equivalent time at 20° (Parent et al. 2010). This feature would allow us to compare growth traces across experiments and analyze environmental effects on canopy development, independent of temperature effects. In this way, the analysis will increasingly become an exercise of statistical treatment of massive data series.

## 11.4 Cost of HTP Methods and Their Linkage to Breeding Efforts

### 11.4.1 Cost of HT Phenotyping

While new technologies offer precision and throughput, technology cost is an important decision factor. High-throughput phenotyping could directly contribute to breeding efforts but the choice made by breeders to adopt or not a given HTP approach will often/always be driven by a cost consideration. The "breeder's equation" is as follow:

$$\Delta G_{\text{year}} = i * r * \sigma / L$$

where  $\Delta G_{\text{year}}$  is the genetic gain per year,  $i$  is the selection intensity,  $r$  is the selection accuracy (or the heritability),  $\sigma$  is the genetic variance for the desired trait, and  $L$  is the length of one generation. The cost of achieving a given  $\Delta G_{\text{year}}$  per unit of phenotyping cost would then be

$$\Delta G_{\text{year}*\$} = i * r * \sigma / L * C$$

where  $C$  is the phenotyping cost per progeny. Breeding being a numbers' game, optimizing this ratio can be done in several ways where HT phenotyping has a role to play:

- The coefficient "r" proxies the accuracy of the trait. Breeders would always ask if a given trait is more accurate than those they measure already such as yield. Let's assume a trait is a good predictor of an increased yield, it could be given priority over yield assessment if its heritability was higher than yield heritability, provided its cost is not prohibitive. One could assume that the decision would depend on the ratio  $r/C$ . A heritability doubled by a HTP method would afford an increase in cost per progeny of a similar magnitude.
- The coefficient "i" here becomes important because it concerns the throughput at which phenotyping efforts are made. Let's assume here again a trait that is a good predictor of an increased yield. Its advantage could be in the fact that thousands

of progeny lines could be tested, instead of only a smaller number of lines that can be tested for yield. Therefore, any HTP method that could cater for a large number would allow to dramatically increase the selection accuracy. Here also the cost factor determines the choice within the boundaries of the  $i/C$  ratio.

- Finally, the coefficient " $\sigma$ " could be concerned in cases where there is large genetic variance for a given trait. A genetic variance for a given trait that is larger than the genetic variance for yield would potentially favor this trait over yield, provided that there is also a close association between this trait and yield under relevant scenarios.

### ***11.4.2 Integration of HTP Methods into Breeding***

This effort is a mix of being pragmatic while seeking the advantage of new technologies. Technology developer often propose high-end solutions while breeders want simple tools, easy to use, and cheap. So, efforts are needed to connect these two domains. Here, a few examples of existing cases showcase a possible fit between the "offer" from the HTP standpoint and the "demand" from the breeding side.

*Drone/remote sensing imaging—HTP methods at the service of yield trial quality control*—The use of drone imaging to acquire plant features that would be otherwise difficult or simply impossible to acquire has grown exponentially (Potgieter et al. 2016). Except for few programs the use of drone in routine breeding still remains at a research phase, although opportunities exist that would bring a lot of benefit. The first among these would be the use of drone imaging to support the quality control of plot measurements. Indeed, breeding networks in the National Agriculture Research Systems (NARS) could benefit from imaging technology to quickly assess the quality of testing fields. For instance, by measuring plant counts and ensuring these are in accordance with targeted density, or by measuring NDVI around canopy closure to ensure homogeneity in the plots. This information could be used to remove heterogeneous plots or parts of the field in the analysis and it would increase the accuracy of the evaluations. Quality standards during data acquisition will be needed to ensure the quality of drone images. For breeding programs to have an easy access to drone technology, data processing and analysis pipeline will also be needed, allowing breeding programs to easily load their images and receive data with a rapid turnover time to be part of the selection decisions. Then only more sophisticated measurements can be taken from the research stage to the scale of a breeding program. Additional such traits could be yield estimates (Guo et al. 2018), or indices that reflect on the crop development, functioning and efficiency with indices reflecting light interception, radiation use efficiency.

### 11.4.3 *Quality Analysis*

NIRS spectroscopy is being routinely used in the assessment of quality in grain and in stover. NIRS measurement currently take place in the lab using benchtop NIRS equipment in most cases. NIRS probes can also be mounted on combine harvesters, as is done in the private sector for major crops like maize. There is also an opportunity to insert NIRS probes in smaller harvesting equipment like the Harvest Master (Juniper System Inc, Logan, UT, USA). Different portable NIRS now exist and start being tested for a direct evaluation of quality in the field (Blummel, pers. Comm.). Raman spectroscopy is also appearing as a new opportunity technology, complementary to NIRS in the domain of quality analysis (Altangerel et al. 2017). X-ray fluorescence (XRF) equipment are used to measure mineral content of grain such as Fe or Zn, which are important but deficient component of the diet of poor rural populations of Africa and Asia.

For many breeding programs of the public system in crops others than the main commercial commodities like maize, wheat, or rice, breeding for quality to respond to a market or consumer demand, or to address a nutritional issue, imply a major shift in what is being evaluated. While the technologies above are available, they are still largely disconnected from the breeding process pipeline. That is, agronomic traits are measured at harvest while quality traits are measured after harvest, often too late to be taken into consideration in breeding selection decisions. Therefore, efforts here are needed to streamline the assessment of quality with the usual traits, allowing the combination of the probes above in the breeding process, here also accounting for time and cost of including these additional sensors and of making additional measurements. It requires the re-designing of harvesting pipelines, the possible re-development of harvesting tools including quality probes.

## 11.5 Conclusion

While new technologies provide opportunities to make phenotyping easier, faster, less expensive, and more informative, they also run the risk of becoming the end that justifies the means. We can avoid this by driving the technology with research questions, made possible through a cross-discipline approach between genetics, breeding, modeling, engineering, physiology, pathology, data management, and statistics. Combination of trait-based phenotyping targeting “building blocks” of critical phenotypes (phenes) to field-based phenotyping for capturing these traits or their consequences holds great promise to generate relevant phenotyping information to breeding programs and match up the load of genomic data available for finding genes behind the phenes. Last but not least, the cost of these HTP technologies has to be taken into consideration if these are to be used in breeding pipelines.



**Acknowledgments** The authors are thankful for the funding from ICRISAT for the capital investment in the LeasyScan facility and to the Kirkhouse Trust fund for contributing to the acquisition of additional scales.

## References

- Altangerel N, Walker JW, González PM, Bailey DW, Estell RE, O'Scullly M. M (2017) Comparison of near infrared reflectance spectroscopy and raman spectroscopy for predicting botanical composition of cattle diets. *Rangeland Ecol Manag* 70:781–786
- Araus JL, Cairns JE (2014) Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci* 19:52–61
- Belko N, Zaman-Allah M, Cisse N, Diop NN, Zombre G, Ehlers JD, Vadez V (2012) Lower soil moisture threshold for transpiration decline under water deficit correlates with lower canopy conductance and higher transpiration efficiency in drought-tolerant cowpea. *Funct Plant Biol* 39:306–322
- Borrell AK, van Oosterom EJ, Mullet JE, George-Jaeggli B, Jordan DR, Klein PE, Hammer GL (2014) Stay-green alleles individually enhance grain yield in sorghum under drought by modifying canopy development and water uptake patterns. *New Phytol* 203:817–830
- Brown TB, Cheng R, Sirault XRR, Rungrat T, Murray KD, Trtilek M, Furbank RT, Badger M, Pogson BJ, Borevitz JO (2014) TraitCapture: genomic and environment modelling of plant phenomic data. *Curr Opin Plant Biol* 18:73–79
- Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Oropeza Rosas M, Rocheford TR, Romay MC, Romero S, Salvo S, Sanchez Villeda H, da Silva HS, Sun Q, Tian F, Upadhyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MD (2009) The genetic architecture of Maize flowering time. *Science* 325(5941):714–718. <https://doi.org/10.1126/science.1174276>
- Chapman S, Cooper M, Podlich D, Hammer G (2003) Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agron J* 95:99–113
- Chapuis R, Delluc C, Debeuf R, Tardieu F, Welcker C (2012) Resiliences to water deficit in a phenotyping platform and in the field: how related are they in maize? *Eur J Agron* 42:59–67
- Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C, Hammer GL (2009) Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: A “Gene-to-Phenotype” modeling approach. *Genetics* 183:1507–1523
- Cobb JN, DeClerck G, Greenberg A, Clark R, McCouch S (2013) Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theor Appl Genet* 126:867–887
- Cooper M, van Eeuwijk FA, Hammer GL, Podlich DW, Messina C (2009) Modeling QTL for complex traits: detection and context for plant breeding. *Curr Opin Plant Biol* 12:231–240
- Deery D, Jimenez-Berni J, Jones H, Sirault X, Furbank R (2014) Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* 5:349–379
- Dornbusch T, Lorrain S, Kuznetsov D, Fortier A, Liechti R, Xenarios I, Fankhauser C (2012) Measuring the diurnal pattern of leaf hyponasty and growth in Arabidopsis—a novel phenotyping approach using laser scanning. *Funct Plant Biol* 39:860–869
- Guo W, Zheng B, Potgieter AB, Diot J, Watanabe K, Noshita K, Jordan DR, Wang X, Watson J, Ninomiya S, Chapman SC (2018) Aerial imagery analysis—quantifying appearance and number of sorghum heads for applications in breeding and agronomy. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2018.01544>

- Hammer GL, van Oosterom E, McLean G, Chapman SC, Broad I, Harland P, Muchow RC (2010) Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. *J Exp Bot* 61:2002–2185
- Kar S, Tanaka R, Korbu LB, Kholova J, Iwata H, Durbha SS, Adinarayana J, Vadez V (2020a) Automated discretization of ‘transpiration restriction to increasing VPD’ features from outdoors high-throughput phenotyping data. *Plant Method* 16:140. <https://doi.org/10.1186/s13007-020-00680-8>
- Kar S, Garin V, Kholová J, Vadez V, Durbha SS, Tanaka R, Iwata H, Urban MO, Adinarayana J (2020b) SpaTemHTP: a data analysis pipeline for efficient processing and utilization of temporal high-throughput phenotyping data. *Front Plant Sci* (In press)
- Kholova J, Hash CT, Kumar PL, Yadav RS, Kocova M, Vadez V (2010) Terminal drought-tolerant pearl millet *Pennisetum glaucum* (L.) R. Br. have high leaf ABA and limit transpiration at high vapour pressure deficit. *J Exp Bot* 61:1431–1440
- Kholova J, Nepolean T, Hash CT, Supriya A, Rajaram V, Senthilvel S, Kakkera A, Yadav R, Vadez V (2012) Water saving traits co-map with a major terminal drought tolerance quantitative trait locus in pearl millet *Pennisetum glaucum* (L.) R. Br *Molecular Breeding* 30:1337–1353
- Kholová J, Tharanya M, Kaliamoorthy S, Malayee S, Baddam R, Hammer GL, McLean G, Deshpande S, Hash CT, Craufurd PQ, Vadez V (2014) Modelling the effect of plant water use traits on yield and stay-green expression in sorghum. *Funct Plant Biol* 41(10–11):1019–1034
- Korol AB, Ronin YI, Itskovich AM, Peng JH, Nevo E (2001) Enhanced efficiency of quantitative trait loci mapping analysis based on multivariate complexes of quantitative traits. *Genetics* 157:789–1803
- Lynch JP, Brown KM (2012) New roots for agriculture: exploiting the root phenome. *Philos Trans Royal Soc B-Biol Sci* 367:1598–1604
- Parent B, Turc O, Gibon Y, Stitt M, Tardieu F (2010) Modelling temperature-compensated physiological rates, based on the co-ordination of responses to temperature of developmental processes. *J Exp Bot* 61:2057–2069
- Peak D, West JD, Messinger SM, Mott KA (2004) Evidence for complex, collective dynamics and emergent, distributed computation in plants. *PNAS* 101:918–922. <https://www.pnas.org/cgi/doi/10.1073/pnas.0307811100>
- Poorter H, Niinemets U, Walter A, Fiorani F, Schurr U (2010) A method to construct dose-response curves for a wide range of environmental factors and plant traits by means of a meta-analysis of phenotypic data. *J Exp Bot* 61:2043–2055
- Potgieter AB, Lobell DB, Hammer GL, Jordan DR, Davis P, Brider J (2016) Yield trends under varying environmental conditions for sorghum and wheat across Australia. *Agric For Meteorol* 228:276–285. <https://doi.org/10.1016/J.AGRFORMET.2016.07.004>
- Ratnakumar P, Vadez V (2011) Groundnut (*Arachis hypogaea*) genotypes tolerant to intermittent drought maintain a high harvest index and have small leaf canopy under stress. *Funct Plant Biol* 38:1016–1023
- Ratnakumar P, Vadez V, Nigam SN, Krishnamurthy L (2009) Assessment of transpiration efficiency in peanut (*Arachis hypogaea* L.) under drought using a lysimetric system. *Plant Biol* 11:124–130
- Reynolds M, Kropff M, Crossa J, Koo J, Kruseman G, Molero Milan A, Rutkoski J, Schulthess U, Singh B, Sonder K, Tonnang H, Vadez V (2018) Role of modelling in international crop research: overview and some case studies. *Agronomy* 8:291. <https://doi.org/10.3390/agronomy8120291>
- Schoppach R, Claverie E, Sadok W (2014) Genotype-dependent influence of night-time vapour pressure deficit on night-time transpiration and daytime gas exchange in wheat. *Funct Plant Biol* 41:963–971
- Schuster I (2011) Marker-assisted selection for quantitative traits. *CBAB* 11:50–55
- Sinclair TR, Seligman N (2000) Criteria for publishing papers on crop modeling. *Field Crops Research* 68:165–172
- Soltani A, Ghassemi-Golezani K, Khooei FR, Moghaddam M (1999) A simple model for chickpea growth and yield. *Field Crops Res* 62:213–224

- Vadez V (2014) Root hydraulics: the forgotten side of roots in drought adaptation. *Field Crops Res* 165:15–24
- Vadez V, Soltani A, Sinclair TR (2012) Modelling possible benefits of root related traits to enhance terminal drought adaptation of chickpea. *Field Crops Res* 137:108–115
- Vadez V, Kholova J, Yadav RS, Hash CT (2013a) Small temporal differences in water uptake among varieties of pearl millet (*Pennisetum glaucum* (L.) R. Br.) are critical for grain yield under terminal drought. *Plant Soil* 371:447–462
- Vadez V, Kholova J, Zaman-Allah M, Belko N (2013b) Water: the most important ‘molecular’ component of water stress tolerance research. *Funct Plant Biol* 40:1310–1322
- Vadez V, Kholova J, Medina S, Kakkera A, Anderberg H (2014) Transpiration efficiency: new insights into an old story. *J Experimental Botany*, eru040
- Vadez V, Rao S, Kholova J, Krishnamurthy L, Kashiwagi J, Ratnakumar P, Sharma K, Bhatnagar-Mathur P, Basu P (2008) Root research for drought tolerance in legumes: quo vadis. *J Food Legumes* 21:77–85
- Vadez V, Krishnamurthy L, Hash CT, Upadhyaya HD, Borrell AK (2011) Yield, transpiration efficiency, and water-use variations and their interrelationships in the sorghum reference collection. *Crop Pasture Sci* 62:645–655
- Vadez V, Kholova J, Hummel G, Zhokhavets U, Gupta SK, Hash CT (2015) LeasyScan: a novel concept combining 3D imaging and lysimetry for highthroughput phenotyping of traits controlling plant water budget. *J Exp Bot* 66(18):5581–5593. <https://doi.org/10.1093/jxb/erv251>
- Vadez V, Halilou O, Hissene HM, Sibiry-Traore P, Sinclair TR, Soltani A (2017) Mapping water stress incidence and intensity, optimal plant populations, and cultivar duration for african groundnut productivity enhancement. *Front Plant Sci* 8:432. <https://doi.org/10.3389/fpls.2017.00432>
- Welcker C, Bousuge B, Bencivenni C, Ribaut JM, Tardieu F (2007) Are source and sink strengths genetically linked in maize plants subjected to water deficit? A QTL study of the responses of leaf growth and of anthesis silking interval to water deficit. *J Exp Bot* 58:339–349
- Welcker C, Sadok W, Dignat G, Renault M, Salvi S, Charcosset A, Tardieu F (2011) A common genetic determinism for sensitivities to soil water deficit and evaporative demand: meta-analysis of quantitative trait loci and introgression lines of maize. *Plant Physiol* 157:718–729
- White TA, Snow VO (2012) A modelling analysis to identify plant traits for enhanced water-use efficiency of pasture. *Crop Pasture Sci* 63:63–76
- White JW, Andrade-Sanchez P, Gore MA, Bronson KF, Coffelt TA, Conley MM, Feldmann KA, French AN, Heun JT, Hunsaker DJ, Jenks MA, Kimball BA, Roth RL, Strand RJ, Thorp KR, Wall GW, Wang GY (2012) Field-based phenomics for plant genetics research. *Field Crops Res* 133:101–112
- Zaman-Allah M, Jenkinson DM, Vadez V (2011a) Chickpea genotypes contrasting for seed yield under terminal drought stress in the field differ for traits related to the control of water use. *Funct Plant Biol* 38:270–281
- Zaman-Allah M, Jenkinson DM, Vadez V (2011b) A conservative pattern of water use, rather than deep or profuse rooting, is critical for the terminal drought tolerance of chickpea. *J Exp Bot* 62:4239–4252