Image-based high-throughput field phenotyping of crop roots

Corresponding author contact:

Alexander Bucksch
School of Biology
Georgia Institute of Technology
310 Ferst Drive
Atlanta, GA, 30332
Tel.: +1 (404) 993-2578
Email: Bucksch@gatech.edu

Research area: Breakthrough Technologies
Image-based high-throughput field phenotyping of crop roots

Summary

Development of automatic imaging methods is crucial for increased throughput and reproducibility of field-based phenotyping of plants. The automatic methods developed here, allow distinction of genotypes, including 13 newly accessible plant root traits, from images of field grown plants.

Alexander Bucksch*,$,1, 2, James Burridge*3, Larry M. York3, 4, Abhiram Das1, Eric Nord3, Joshua S. Weitz1, 5, Jonathan P. Lynch3

1 School of Biology, Georgia Institute of Technology, Atlanta, GA, USA
2 School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA, USA
3 Department of Plant Science, The Pennsylvania State University, University Park, PA, USA
4 Intercollege Graduate Degree Program in Ecology, The Pennsylvania State University, University Park, PA, USA
5 School of Physics, Georgia Institute of Technology, Atlanta, GA, USA

* Authors contributed equally
$ Corresponding Author
Footnotes
This research was supported by the NSF Plant Genome Research Program, NSF 0820624 (JPL and JSW), the Howard G Buffett Foundation (JPL), NSF/BREAD 4184–UM–NSF–5380 (JPL), and the Center for Data Analytics, Georgia Institute of Technology, Spatial Networks in Biology: Organizing and Analyzing the Structure of Distributed Biological Systems (AB and JSW). JSW holds a Career Award at the Interface of Science from the Burroughs Welcome Fund.
Abstract

Current plant phenotyping technologies to characterize agriculturally relevant traits have been primarily developed for use in laboratory and/or greenhouse conditions. In the case of root architectural traits, this limits phenotyping efforts, largely, to young plants grown in specialized containers and growth media. Hence, novel approaches are required to characterize mature root systems of older plants grown under actual soil conditions in the field. Imaging methods able to address the challenges associated with characterizing mature root systems are rare due, in part, to the greater complexity of mature root systems, including the larger size, overlap and diversity of root components. Our imaging solution combines a field imaging protocol and algorithmic approach to analyze mature root systems grown in the field. Via two case studies, we demonstrate how image analysis can be utilized to estimate localized root traits that reliably capture heritable architectural diversity as well as environmentally induced architectural variation of both monocot and dicot plants. In the first study we show that our algorithms and traits (including 13 novel traits inaccessible to manual estimation) can differentiate nine maize genotypes 8 weeks after planting. The second study focuses on a diversity panel of 188 cowpea genotypes to identify which traits are sufficient to differentiate genotypes even when comparing plants whose harvesting date differs up to 14 days. Overall, we find that automatically derived traits can increase both the speed and reproducibility of the trait-estimation pipeline under field conditions.
Introduction

Crop root systems represent an under-explored target for improvements as part of community efforts to ensure that global crop yields and productivity keep pace with population growth (Godfray et al., 2010; Gregory and George, 2011; Nelson et al., 2012). The challenge in improving crop root systems is that yield and productivity also depend on soil fertility, itself a major constraint to global food production (Lynch, 2007). Hence, desired improvements to crop root systems include enhanced water use efficiency and water acquisition given the increased likelihood of drought in future climates (IPCC, 2014). Over the long-term, the development of crop genotypes with improved root phenotypes requires advances in the characterization of root system architecture (RSA) and in the relationship between RSA and function.

The emerging discipline of plant phenomics aims to expand the scope, throughput and accuracy of plant trait estimates (Furbank, 2009). In the case of plant roots, structural traits may describe root system architecture as geometric or topological measures of the root shape at various scales, e.g. diameters and width of the whole root system or a single branch (Lynch, 1995; Den Herder et al., 2010). These traits can be used to predict yield under specific conditions such as drought or low fertility. Understanding the diversity and development of root architectural traits is crucial, because spatial and temporal root deployment affects plant fitness, especially water and nutrient acquisition (Rich and Watt, 2013). Thus, improving plant performance may benefit from improvements in the characterization of root architecture, including understanding how trait variation arises as a function of genotype and environmental conditions (Band et al., 2012; Shi et al., 2013).

Current efforts to understand the structure of crop root systems have already led to a number of imaging solutions (Lobet et al., 2013), able to extract root architecture traits under various conditions (Fiorani et al., 2012), including laboratory conditions (de Dorlodot et al., 2007) where plants are often grown in pots or glass containers (Zeng et al., 2008; Armengaud et al., 2009; Le Bot et al., 2010; Clark et al., 2011; Lobet et al., 2011; Naeem et al., 2011; Galkovskyi et al., 2012). In the case of pots, expensive MRI technologies represent one non-invasive approach to capture high resolution details of root architecture (Schulz et al., 2013), similar to the capabilities of micro X-Ray CT systems. X-Ray systems allow capturing of the root architecture at a fine scale in containers with a wide variety of soil types (Mairhofer et al., 2012; Mooney et al., 2012).
It has been shown that micro X-Ray CT paired with specifically designed algorithms has sufficient resolution to recover the root structure in many cases (Mairhofer et al., 2013). Nevertheless, at present, micro X-Ray CT systems are unable to image mature root systems because of technical restrictions in container size.

As an alternative, root systems can be imaged directly with a digital camera when grown in glass containers with transparent media such as gellan gum or transparent soil replacements (Downie et al., 2012). Such in situ imaging benefits from controlled lightning conditions during image acquisition, even more so when focusing on less complex root structures of younger plants that allow 3D reconstruction (Clark et al., 2011; Topp et al., 2013). Under such controlled conditions it is expected that imaging would enable the study of growth of roots over time (Spalding and Miller, 2013; Sozzani et al., 2014).

However, all of the above-listed solutions have been used, primarily, to assess root structures in the early seedling stage (French et al., 2009; Brooks et al., 2010; Sozzani et al., 2014) to approx. 10 days after germination (Clark et al., 2011), which makes it all but impossible to directly observe mature root systems. For example, primary and seminal roots make up the major portion of the seedling root system in maize during the first weeks after germination. Later in development, postembryonic shoot-borne roots become the major component of the maize root system (Hochholdinger, 2009), not yet accessible to laboratory phenotyping platforms. In addition to phenological limitations, current phenotyping approaches for root architecture require specialized growth conditions with aerial and soil environments that differ from field conditions; the effects of such differences on RSA are only sparsely reported in literature, e.g. (Hargreaves et al., 2009; Wojciechowski et al., 2009).

Indeed, high-throughput field phenotyping can be seen as the new frontier for crop improvement (Araus and Cairns, 2014) because imaging a mature root system under realistic field conditions poses new challenges and opportunities (Gregory et al., 2009; Zhu et al., 2011; Pieruschka and Poorter, 2012). Challenges are intrinsic to roots grown in the field because the in situ belowground imaging systems are, to date, unable to capture fine root systems (Daniels, 2005). As a consequence, initial attempts to characterize root systems in the field focused on the manual extraction of structural properties. Manual approaches analyzed the root system’s branching hierarchy in relation to root length and rooting depth (Fitter, 1991). In the late 1980’s, imaging
techniques were first used (Tatsumi et al., 1989) to estimate the space-filling behavior of roots; an estimation process that has recently been automated (Zhong et al., 2009). A weakness of such approaches is that exact space-filling properties, such as the fractal dimension are sensitive to the incompleteness of the excavated root network (Nielsen et al., 1997; Nielsen et al., 1999). In particular the box counting method was criticized for above ground branching networks of tree crowns (Da Silva et al., 2006). The same critiques apply to root systems, because fine secondary or tertiary roots can be lost, cut off or adhere to each other during the cleaning process, making it impossible to analyze the entire network.

As an alternative, the ‘shovelomics’ field protocol has been proposed to characterize the root architecture of maize under field conditions (Trachsel et al., 2011). In shovelomics, the researcher excavates the root at a radius of 20 cm around the hypocotyl and 20 cm below soil surface. This standardized process captures the majority of the root system biomass within the excavation area. After excavation the shoot is separated from the root 20 cm above the soil level and washed in water containing mild detergent to remove soil. The current procedure places the washed root on a phenotyping board consisting of a large protractor to measure dominant root angles with the soil level at depth intervals and marks to score length and density classes of lateral roots. A digital caliper is used to measure root stem diameters (Figure 1). Observed traits vary slightly from crop to crop but generally fit into the following categories by depth or root class: angle, number, density and diameter. In this way, field-based shovelomics allows the researcher to visually quantify the excavated structure of the root crown and compare genotypes via a common set of traits that do not depend on knowledge nor observation of the entire root system network. Of note, shovelomics is of particular use in developing countries, which have limited access to molecular breeding platforms (Delannay et al., 2012), and for which direct phenotypic selection is an attractive option.

Nonetheless, direct phenotypic selection for root traits of field grown root systems comes with a number of caveats and drawbacks. To date the quantification of mature root systems is highly dependent on the researcher, reducing the repeatability of measured quantities (Gil et al., 2007). Additionally, manual approaches impose limitations both on the number of accessible traits and the number of samples collected. For example, a typical shovelomics phenotyper can gather 10-12 traits from a sample in 2 minutes or ca. 200 samples in a typical workday (Trachsel et al., 2011). This means that evaluating a statistically more significant trial of 1000 samples may take
5 days or more. Such a time span introduces variation due to plant growth and phenology. On the other hand, shovelomics is adaptable to both monocot and dicot roots and therefore its procedure can be scaled over the huge variety of root morphologies in dicots and monocots. For these reasons, we contend that shovelomics represents a valued target for automated high-throughput phenotyping approach in the field.

We introduce an imaging approach for high throughput phenotyping of mature root systems under realistic soil conditions to address the limitations of manual data collection (Figure 2 I/II). Our new algorithms directly extract root architecture traits from excavated root images (Figure 2 III/IV). Rapid imaging of the roots and decoupled analysis of the collected data at a later time improves both the throughput and synchronization of experiments. In addition to traits that are developed from shovelomics, our algorithms extract previously inaccessible traits such as root tip diameter, spatial distribution and root tissue angle. Overall, our approach allows imaging directly at field sites with an easily reproducible imaging setup and involves no transportation cost or expensive hardware. We demonstrate the utility of our approach in the field-based analysis of maize and cowpea architecture.

RESULTS

Low-cost, reproducible, and field-based protocol for imaging plant roots

We extend the manual trait-estimation procedures described in shovelomics (Trachsel et al., 2011) in order to facilitate estimation of root traits based on digital images taken under field conditions. Similar to shovelomics, the root crown is excavated, washed in water, and then placed on a board to estimate root traits. Image based estimation of root traits requires improvements in the contrast between root and board object segmentation on the board and automated labeling. To address these challenges, we developed a new imaging board and imaging protocol. The new imaging board consists of a black background with diffuse reflectance properties on which a circle of known diameter is freely positioned which enables calculation of units (see Figure 2 I and Methods). Next, images of the root system were photographed using a digital camera mounted on a tripod (Figure 2 II & Methods). Finally, we developed an automated pipeline to derive measurements from the root crown and the excised root samples. In the first step, the objects on the imaging board are detected and segmented into individual image masks. Each individual mask represents the root crown, excised root sample,
experimental tag and scale marker (see Figure 2 III). As a precursor to trait estimation, the algorithm automatically detects the marker and tag using geometric priors; the geometry of the marker is then used to automatically calibrate the trait measurements. In the second step, a structural description of the root crown and the excised root is computed. The structural description, called Root-Tip-Paths (RTP), is a curve representation that samples the root architecture (Figure 2 IV). The sampling is necessary because the actual root networks of mature root systems occlude its interior and when the roots are excavated and washed smaller roots may bind together. Finally, we utilize the image masks and RTP to estimate architectural traits of dicot and sparse monocot root system based on angle, diameter, length, width and density measures (see Methods for complete description).

**The Digital Image Analysis Pipeline Enables Field-Based Estimation of Root Traits**
All trait measurements rely on the computation of root length, diameters, branching angles, density and spatial root distribution estimations. These measurements whose distributions are summarized in trait values characterize the root architecture. Technically, traits are distinguished based on whether they are computed from (i) the image mask; (ii) the “root-tip-path (RTP)” skeleton (see Figure 2 IV and Methods); (iii) the excised root component. The image mask (see Figure 2 III c) enables estimation of root width profile as well as root density. The RTP skeleton is a loop free sampling of the medial axis derived from the root shape visible in the image. The RTPs describe significant root parts by the number of paths passing through the same location in the image determined on the basis of detectable tips in the image.

As a consequence, the RTP skeleton allows the calculation of the length of significant root parts, the radius at each location as distance to one of the at least two closest pixels on the root mask boundary, branching frequencies and path length from the distance between branching points and angles between a root part emerging from the tissue of a parent root. Our segmentation of the original images enables separate analysis of an excised crown root(s) (see Figure 2 III d,e). The intermediate mask images used in our pipeline allows a user to visually identify possible miscalculations due to substandard imaging conditions. Tables 1 and 2 summarize all currently computed traits.
We determined the technical error of our pipeline with five persons that were untrained in the imaging protocol (see Methods for detailed description). We used a black block with a cylindrical groove to hold the maize root crown and increase the repeatability of the placement. Additionally we placed a bean root without the groove to demonstrate the concept of the root tissue angle (RTA) in comparison to the soil tissue angle (STA). For maize we obtained relative standard deviations below 10% except for avg. root density, diameter at 90% depth and the spatial root distribution in X. The majority of traits such as D- and DS-values showed relative standard deviations below 5.0%. For the freely placed bean we found, that the RTA measurements are more robust to placement errors than the STA measurements. The median RTA resulted in an average residual of 3.0° which corresponds to a relative standard deviation of 5.2%. In contrast the median STA showed an average residual of 9.4° corresponding to a relative standard deviation of 22.7%. The maximal width measurements resulted in a relative standard deviation less than 1% in maize, but 7.0% in bean depending on root placement. Overall our test showed in both cases the expected robustness of our traits to translational and rotational differences in the placement of the root. Nevertheless, we suggest the use of tools to guide the root placement to further improve the collected data quality. The full test results are available in the supplemental material (beanTest.xls and maizeTest.xls).

Validation of image based traits

We validated the diameter traits, which were derived with a caliper in the field on 1445 cowpea roots representing 188 genotypes. Figure 8 shows the correlation of taproot diameters at two depth levels derived as field measurements (Figure 8a), the direct comparison between manually and image-based derived diameters (Figure 8b) and image based measurements at two depth levels (Figure 8c). All three comparisons yield similar Pearson correlations of 0.49 ± 0.02 with p<0.001. The conceptual difference between field and image-based measurement is that manually measured diameters are taken at a certain depth below the soil line, not accounting for differences in plant age. Continuing growth induces errors emerging from simultaneous lateral thickening and elongation in depth. In contrast the image-based measurement is taken as a certain percentage of the maximal depth to compensate for differences in plant age, assuming a linear scaling of length and thickness within the given time frame. Additionally, we found strong correlations between image-based and field measured traits, which were not accessible in the image. For example the number of 3rd order roots scored in the field was correlated with a strong
Pearson $r=0.8$ to the number of detectable RTPs as shown in Figure 8(d). Because RTPs are derived from the number of tips visible in the image, the Pearson $r=0.8$ is explained by the exponentially growing amount of branches. Furthermore, the RTPs are correlated with the nodulation score for the cowpea data set ($r=0.59$), which again is explained by the exponentially growing amount of branches with increasing hierarchies allowing more nodulation.

In addition to the correlations found for cowpea, we found strong correlations for 9 available maize genotypes (Figure 4). Measurements of manually scored lateral root length and image-based estimated lateral root length of the brace root resulted in a Spearman $r=0.7$ ($p=0.0367$). Together with the caliper and image-based measurements of stem diameters (Spearman $r=0.83$, $p=0.0053$) the measurements validated our image based measurement pipeline (Figure S1 and S2). As expected, the crown/brace root angles were correlated with the D10 trait (Figure S3 and S4) and yielded a significant Spearman $r=-0.82/-0.69$ ($p=0.0072/0.0379$). The negative correlations are explained with the geometric constraint that the steeper the root angles to the soil line are the less width is accumulated at the D10 level. The surprising result in maize is the opposite correlation of the number of crown and brace roots at the D20-D60 values. For example the D20 value resulted in a negative Spearman $r=-0.73$ ($p=0.0379$) with the manually counted number of brace roots. In contrast the D20 value and the number of crown roots significantly correlated with a positive Spearman $r=0.88$ and a significance value of $p=0.0018$ (Figure S5 and S6). Under the assumption that crown roots maximize in number within the available volume enclosed by brace roots, then the correlation is positive because increasing width accumulation yields more crown roots at the given depth range. We conjecture that width accumulation of crown roots increases with an increasing number of crown roots and leads to shallower root angles because larger initial width accumulation is associated with shallower root angles for brace and crown roots (see above for D10 analysis). Additional analysis on the level of individual plants can be found in the Supplemental material in Figure S31-S36. The correlations to manual measured traits have known or hypothetical importance in soil resource acquisition (Lynch, 2011 and Lynch, 2013).

**Image based traits exhibit high differentiation potential**

We selected nine maize genotypes for diversity in lateral root density, nodal root angle and nodal root number (Figure 4). In the following we characterize the maize data set by comparing
relative phenotypic variation (RPV) as estimated using shovelomics and image based traits. The RPV of a trait for a given data set is a heuristic, defined as the ratio between the variance $V_d$ of the trait of all roots of the data set and the average trait variance per genotype $V_{avg}$.

$$RPV = \frac{V_d}{V_{avg}}$$

Traits are more likely to be useful in differentiating genotypes when their RPV is significantly greater than 1. An RPV of approximately 1 indicates that a given phenotypic trait is similarly distributed over given genotypes. In Figure 9 we show the RPV values of all traits for the maize crown roots. Manual measurements of the rootstock yielded only the stem diameter with an RPV above 2.5. In contrast, the absolute image based measurements such as root density; median/maximum width and spatial root distribution show RPV values above 2.5, suggesting differentiation between phenotypes in the given data set. Note that the stem diameter estimated from the images shows lower RPVs than the manually derived stem diameter. The difference in the RPV for the stem diameter is a consequence of the visual choice of a representative measurement location by the phenotyper. In contrast the relative measurements of D and DS values show often greater RPVs than the comparable crown root angle measured in the field.

The RPV of the excised root sample traits reveal differences in the measurement methodology. While a field phenotyper scores the average lateral length on a scale from 0 to 9 (RPV>3.5), the length of the RTPs is determined in cm with an RPV close to 1. This effect can be explained by a binning effect induced by the scoring scheme of the researcher, that does not reflect the variation in length (compare visually Figure 4). In contrast, we observed an RPV > 2 from images for the brace root (see SI Figure 11). Similarly to the RPV value we obtain the broad sense heritability $H^2$ from our trait data (Einspahr et al., 1963).

$$H^2 = \frac{V_{avg}}{V_{avg} + V_d}$$

The crown root traits resulted in $0.55 < H^2 \leq 0.81$ for traits measured manually and $0.56 < H^2 \leq 0.83$ from images (Figure S12). The largest $H^2$ was obtained for the lateral density (top) in case of the manual measurements and for the median width and D10 in case of the image
based measurement. For the brace root traits we observed $0.66 < H^2 \leq 0.85$ for traits measured manually and $0.56 < H^2 \leq 0.80$ from images (Figure S13). The most heritable traits were lateral length (bottom) for the manual measurements and avg. root density for the image based measurement. Note here that the heritability measure often results in a similar trait ranking than the RPV value, but does not reflect the strength of genotype differentiation.

We demonstrate the inter-genotype vs. intra-genotype variability for the crown root traits obtained with our imaging algorithms in Figure 6 and the manually derived traits in Figure 7. In both figures we highlight the genotypes OH43, 4226 and B103 for direct comparison. Here we derived the normalized mean values as the robust sample mean from the z-score values within the 3-Sigma-Interval (Maronna et al., 2006). The associated standard errors (of the mean, shown as error bars) in both Figures give insight into the precision of the population mean. Non-overlapping error bars visualize that the three example genotypes can be distinguished by the mean of a subset of traits. Pairwise comparison of genotypes per trait reveals that each genotype is distinguishable by at least eight image based traits and one manually measured trait (see bottom panel in Figure 6 and Figure 7). Most notably, we observed improved distinction for the distance to the first lateral for the image-based traits. While field practice assigns a researcher dependent distance score showing no variation between 8 of the chosen genotypes, the image-based measurement still distinguishes different distances. The comparison of root angle measurement in the field and D- and DS-values that describe the cumulative width over the depth shows that distinctions between genotypes can be made with both the absolute angle measurement in the field and the D- and DS-values derived from the images. The supplemental material contains the intra genotype variation and full resolution figures for brace and crown root images (Figure S7-S10).

**Differentiating genotypes in massive diversity panels**

We evaluated image-based traits on a cowpea diversity panel consisting of 188 genotypes (Figure 8). Such large diversity panels pose a trade-off between two additional problems:

1.) Long harvesting times introduce phenological effects requiring robustness to varying plant age in the trait definition. In our example the harvesting duration was 14 days.

2.) The large amount of genotypes requires high sensitivity of the traits to extract and differentiate the phenotypes present in the data set.
In our cowpea study, comparable differentiation of phenotypes with diameters measured in the field and in images was achieved. Overall, the RPV values of the cowpea study are lower (<2.4) than the RPV values of the maize study (up to 7 for the brace root and up to 4.5 for the crown root). Lower RPV values can be explained with the by far larger amount of genotypes (188 for cowpea versus 9 for maize), constituting a full diversity panel. Directly comparable measurement such as the stem diameter resulted in similar RPV values about 1.6 and the correlated traits number of RTP paths and 3rd order roots are similar strong around 1.9. The diameter measurements of tap root and comparable central path diameter show similar differentiation potential in their RPV values. Central path diameter at a fraction of the maximal depth is in the range between 1.3 and 1.4, which is in alignment with the RTP values of the taproot diameter at 5cm and 10 cm depth. The taproot diameter at 15cm depth is the trait with the strongest RPV value. However only about 30% of all measured roots reached a taproot length of 15 cm, which suggests that the RPV value is artificially high and the actual presence of a 15 cm deep taproot might constitute a novel trait. Similar, we observed stronger differentiation for all 2nd dominant angles (RPV>2.5) obtained with the image-based trait estimation. Again, a 2nd dominant angle is not obtained for all genotypes, which again suggests that the presence of such an angle might constitute a novel trait. The 2nd dominant root tissue angle showed the strongest differentiation potential with RPV=3.2. Mean and max. width differentiated with an RPV value around 1.5, similar to the related measures of basal and adventitious root angle, while D- and DS-values performed less strong (RPV < 1.4) compared to the maize study. The broad sense heritability (Einspahr et al., 1963) peaked at $H^2 = 0.76$ for the manual cowpea traits (nodulation score) and at $H^2 = 0.65$ for the image based traits (number of RTPs). Notably these are two strongly correlated traits as shown before in Figure 3d. The taproot diameter at 15 cm depth and the number of first order laterals at 15 cm depth below the soil line show a strong $H^2$. We considered these two heritability measures as unreliable, because only a few roots reached 15 cm depth or length. The heritability for all cowpea traits is shown in Figure S20.

The absolute trait of the 2nd dominant angle distinguished genotypes with large angular differences as shown in the example in Figure 10. The large standard error for the first dominant angle possibly emerges from the difference in root age and is also observable in the angles of the field measurement (Figure S22). On the selected examples in Figure 8, D and DS values again
distinguish the genotypes by means of the standard error. Nevertheless, on the large diversity panel we found that there exists no single trait to distinguish all genotypes. In our example study we found that all 188 cowpea genotypes could be distinguished by at least one image-based trait, while five genotypes couldn’t be distinguished with the manual field measurement. Predominantly we found that D- and DS values and diameter measurements of the stem and along the central path are most suited to distinguish genotypes. The D- and DS values differentiate cowpea phenotypes better than corresponding angle measurements collected by the field researcher, because the traits are derived relative to the maximum depth of an individual root. Thus the difference in plant age is less influencing the actual measurement. Large versions of the figures, the corresponding data sets, the intra-genotype variation (Figure S18-S22) and the pair-wise comparison matrices can be found in the supplemental material (differentiationImaging.csv/ differentiationShovelomics.csv).

**Conclusion and future directions**

We have developed image processing algorithms in alignment with the needs of current field research practice to estimate root phenotypes in real soil conditions. In doing so, we have (i) developed a low-cost, reproducible protocol for image based root analysis; (ii) developed imaging processing methods to automatically classify multiple root components (and experimental tags) from a single image; (iii) developed a novel set of localized root system traits based on the novel principle of root-tip-paths; (iv) demonstrated that these traits can differentiate among genotypes of both cowpea and maize. In contrast to algorithms available in existing software packages like RootNav (Pound et al., 2013), SmartRoot (Lobet et al., 2011) or DART (Le Bot et al., 2010), which focus on the detailed measurement of an individual root system by a user, our algorithms allow the automatic extraction of many root traits in a high-throughput fashion. Naturally our approach comes at the cost of a greater dependency on image quality and a larger amount of root samples to achieve significant average trait values. Still, semi-automatic approaches are equally dependent on the visibility of root tips in images.

Overall we found that angular, diameter, width and length measurements derived from images show similar or in some cases superior RPV and differentiation in terms of the standard error of the mean compared to manual shovelomics quantification. In our example studies we showed that image analysis is an attractive alternative to extract crop root architecture traits also under
field conditions. Additionally, we identified a subset of traits (central path diameter/D- and DS-values) that have the potential to analyze large field studies with long harvesting durations. We hope that our algorithms enable large-scale plasticity studies on roots in the field because we see high throughput plant phenotyping under field conditions as enabled by our approach.

We developed RTPs as a representation of significant root parts, from which additional information can be computed. Nevertheless, our approach has shown that new conclusions can be drawn from the computed data, because of increased trait robustness and measurement resolution. Additionally, we provide 13 traits previously not accessible in the field such as tip diameter, symmetry measures or root tissue angle, which enriches the set of traits to be used by field researchers. The underlying assumptions for the RTPs allow us to objectively simulate the scoring and measurement of field researchers. Our traits include localized diameter, angle, length or width measurements at certain depth levels of major root components. In that sense our algorithms are complementary to approaches like GiA Roots (Galkovskyi et al., 2012) that predominantly focus on global root traits. Beside global traits Root Reader (Clark et al., 2011) offers mechanisms to measure local traits for individual roots. However, these mechanisms are designed for sparse root networks where root branches do not occlude each other. We note that neither was optimized for application to field images. Nonetheless, we provide a direct comparison of GiA Roots to our algorithms in the Supplemental material. GIA Roots did not distinguish 128 genotype combinations of the cowpea data set, while our algorithms could distinguish all cowpea genotypes. Recall here that the manually measured traits distinguished all but five genotype combinations. our image based trait measurement is a first step towards localized traits under field conditions in the sense that many of our traits are defined with respect to depth and length of the root.

Indeed, measurements of root tip diameters and diameters at 90% depth couldn’t be assessed in a Shovelomics setting, but are accessible in gel-based systems of younger roots. The technological possibility to extract root tip information is demonstrated on the example of 10 rice genotypes of a previously published study (Iyer-Pascuzzi et al., 2010) in Figure S23 of the Supplemental Information. These images were previously used for genotype differentiation based on global root traits with the algorithms available in GiaRoots (Galkovskyi et al., 2012). Beside obvious genotype distinction in many traits, the figure shows that mean tip diameters and the maximal diameter at 90% depth have good differentiation capabilities, when derived with our algorithms.
Indeed, many traits can be derived for both settings and may enable translational studies from the field to the lab.

The traits extracted with our imaging pipeline may also enhance input to data-driven simulation systems like *SimRoot* (Lynch et al., 1997). Such a combination would allow the reconstruction of the complete root system in the future even though the excavation process cuts off some of the root system. The challenge of such data-driven simulation is the relation between traits extracted from images and elementary phenes driving the simulation (Postma et al., 2013; York et al., 2013). Combined model based and data supported reconstructions have already been successfully used for tree crowns (Cote et al., 2009). Such a simulation would enable computational experiments of crop root responses to varying environmental conditions such as drought, N, P or K during root growth, even more extensive than known for cauliflower (Vansteenkiste et al., 2014).

In summary, our imaging approach and algorithms enable the novel possibility to phenotype mature root systems under real conditions from objective data. This approach has addressed an identified need for phenotyping protocols for mature root systems operated in a high-throughput fashion under real soil conditions to extract root architectural traits (Dhondt et al., 2013). We envision future extensions to leverage cloud and grid computing platforms as infrastructure (Moore et al., 2013) to calculate image based traits from thousands of images in parallel. In this way, field experiments may enhance their statistical power and significance to levels currently achieved in laboratory environments.

**Methods and Material**

**Plant and Growth Conditions**

We selected nine maize lines from a planting of a subset of the Wisconsin Diversity Panel (Hansey et al., 2011) grown for 8-12 weeks at the Ukulima Root Biology Center located in Alma, Limpopo Province, South Africa. Five lines (B73, H99, Mo17, OH43, W64A) were selected as being well studied lines that are also included in the EUROOT maize panel (http://www.euroot.eu/). Four replicates of each line were planted in an incomplete block design. A further four lines (4226, B103, Tr, W23) were selected with contrasting root system architecture; specifically to represent the range of nodal root angles, nodal root numbers and
lateral branching density as shown in Figure 8. These lines were replicated three times. For all these lines four root crowns per plot were excavated, and two were scored, resulting in a total of 44 imaged roots and 4-6 measurements per genotype. Images were taken separately for crown and brace roots, each containing an excised crown root or brace root respectively resulting in 88 images in total.

Cowpea roots were collected in January 2013 at the Ukulima Root Biology Center (URBC, Limpopo, South Africa) from a 188-line diversity panel grown for 5-6 weeks under non-limiting conditions. The supplemental material contains a list of all genotypes used in this study (CowpeaGenotypeMap.xlsx).

We cite for completeness the growth conditions of the rice data set given in (Iyer-Pascuzzi et al., 2010):

“Seeds were dehulled and sterilized with 10% peroxide for 10 min, followed by 70% ethanol for 1 min, and rinsing three times with sterile water. Sterilized seeds were planted at approximately 1 cm below the surface of the gel. Plants were grown in 2 L ungraduated borosilicate cylinders (Fisher) filled with 750 mL Yoshida's nutrient solution (Yoshida et al., 1976) and 0.3% Gelzan (the highest grade of gellan gum available; Sigma). Cylinders are approximately 520 mm high, with an o.d. of 82.5 mm. To prevent the agar from moving during imaging, six to eight glass pipette tips were bonded to the bottom of each cylinder with Sylgard 184 (Robert McKeown Co.). Plants were grown for 14 d at 12-h day/night, 28°C day, 25°C night. Images were taken 14 dap…. Twelve rice (*Oryza sativa*) genotypes were used: 93-11 (9311; *indica*), Caipo (Cai; *tropical japonica*), Basmati 217 (Ba217; Basmati), Jefferson (Jef; *tropical japonica*), Teqing (Teqing; *indica*), Lemont (Lemont; *tropical japonica*), Moroberekan (Moro; *tropical japonica*), Nipponbare (Nipp; *temperate japonica*), IR64 (IR64; *indica*), Carolina Gold (Gold; *japonica*), Bala (*indica*), and Azucena (*tropical japonica*). Genotypes were chosen based on parentage of available QTL mapping populations and to represent both of the two major subpopulations of rice (*indica* and *japonica*). Analyses were performed on six to 18 individuals of each genotype.”
**Low-cost, reproducible, and field-based protocol for imaging plant roots**

We extend the shovelomics trait-estimation procedures described in (Trachsel et al., 2011) with a new imaging board and imaging protocol. The new imaging board consists of a black background with diffuse reflectance properties on which a circle of known diameter is freely positioned (see Figure 2A). Optionally, a rectangular marker that encodes the experimental identification of the root crown in a sequence of letters and numbers can be placed on the board. Suitable boards were simple cutting boards, ordinary black cloth or plastic painted with blackboard coating. These boards are easily cleaned with a brush or damp cloth to remove soil and root fragments from previously imaged roots and are low-cost and suitable for fieldwork.

The protocol for imaging with this modified board is as follows. First, the root crown is placed on the board, oriented such that the stem is perpendicular to the horizontal edge of the board. A predefined number of excised roots, e.g. basal or nodal roots can be placed on the board. The placement of these elements is arbitrary so long as none of the imaged objects (i.e., circle, marker, root crown and excised roots) intersect another object. (see Figure 2). Including multiple root components on the same image enables downstream analysis of multiple root traits.

Next, images of the root system should be photographed using a digital camera mounted on a tripod (Figure 3) at about 30cm to 50cm above the imaging board. We used high-end consumer cameras in this study. The known dimensions of the circle correct for tilting of the camera normal to the imaging plane and enable automated correction of the trait estimations. As we describe below our methods provide mechanisms to deal with irregular ambient light and visual noise that may lead to varied image quality. Nonetheless, we recommend cleaning the imaging board regularly, while maintaining relatively constant diffuse lighting.

Finally, it is worth noting that our image based phenotyping pipeline is technologically constrained by the use of digital cameras, which rely on the reflection of light from surfaces and therefore cannot resolve large occlusions of the root system. As a consequence, our pipeline relies on the subjective impression of the field phenotyper to select a representative nodal root to detail the fine root structure. Measuring complete mature root architectures would require either automated cutting of the root into its individual organs or field ready 3D imaging systems, such as X-ray, neither of which is currently available. Given these limitations we assumed the maize roots to be symmetric to the shoot main axis if only imaged from one side. This assumption...
allows that the width-related traits (e.g. median width, D and DS values – defined in Table 1) and root density estimates are the parameters closest to the ones observed in the field.

**The Digital Image Analysis Pipeline**

**Identifying distinct components:** The original color image taken from the image board (a) is converted into a grey-scale image. The conversion uses the weighting from the PAL television system, where the grey value or chrominance $Y$ is derived as the weighted sum of $Y = 0.30R + 0.59G + 0.11B$, where $R, G, B$ are the red, green and blue components of each pixel respectively. The resulting grey image is segmented into a foreground and background by first estimating a global threshold with the Otsu method (Otsu, 1975). This thresholding method assumes a bimodal distribution of chrominance values, from which an initial global separation threshold is estimated. The initial threshold value is passed into an adaptive thresholding algorithm that locally refines the global threshold in a user-defined window around each pixel. Here, the scaled threshold is the sum of global threshold and the variance of a Gaussian distribution used to weight pixels within the thresholding window. If the weighted mean of the thresholding window is greater than the global threshold, than the pixel is classified as foreground. Otherwise the pixel is classified as background. Note that, such scaling influences the measurement slightly, because the root appears slightly thicker in the binary image than in the original. The resulting binary image is a segmentation of objects in the foreground and the background. The segmentation is shown in Figure 2 III b, where the foreground corresponds to white pixels and the background to black pixels. It is noteworthy that our thresholding reduces the number of user-controlled parameters to one if used with our field imaging protocol. The only user-controlled parameter is the scale parameter that controls weighing of neighboring pixels. All other parameters can be fixed to standard values in our setup.

To distinguish the objects on the board from noise in the foreground we first use connected component labeling (Samet, 1981) to identify clusters of neighboring foreground pixels called components. We assume that none of the 4 placed object classes (Figure 2 III) intersect and no additional objects are placed on the board. It is a prerequisite that noise pixels do not group into clusters that take up more area than the placed objects. Keeping the board clean from soil artifacts while images are taken can enforce this noise constraint. The root crown is defined as the component containing the largest number of pixels. Circle marker, excised root and the
experimental tag are distinguished using prior knowledge on their likely shape. The circle is detected as the component with a side length ratio of the bounding box closest to 1 and distinguished by a foreground percentage of 78%, corresponding to the fraction of area taken up by a circle inscribed into a square. Excised roots are detected as the next set of largest components. They are distinguished by the ratio of foreground to background pixels within the bounding box of the extracted object. We assume that the tag has a higher foreground to background ratio then the root and should be the 2\textsuperscript{nd} smallest component in terms of overall number of pixels. As a result we obtain an image mask with separated foreground and background for each object on the imaging board.

\textbf{Traits derived from the image mask:} The root crown image provides the input to compute the root width profile as illustrated in Figure 11, root width and depth, as well as root density. The root width profile is computed per row of the image raster. The root width per row is given as the distance between the first and the last foreground pixel in the row. From the width profile we calculate so-called D-values as the cumulative width at a certain fraction of the maximal depth. For example D-10 and D-90 represent the sum of all root widths at rows whose vertical location lies between 0-10\% and between 0-90\%, respectively, of the total number of rows in the root crown. Furthermore, each D-value has a corresponding DS-value that is the slope of the tangent at each D-value. Both D- and DS-values are novel traits in the literature and represent samples of a polynomial as shown in Figure 11. In practice the latter can be thought of a normalized root-top angle, as is the case for maize. The ratio between foreground pixels and background pixels in a row is the root density per row, which is averaged to yield the overall root density. The measurements computed from the image mask are similar to the manual measured shovelomics in that they quantify the global shape of the root system architecture in absolute values of the whole root system or at certain depth levels. Traits derived in relation to the maximal depth and/or width in case of D- and DS-values is inaccessible in the field. The latter are said to be relative traits while all other traits are absolute traits.

The individual masks of the root crown and the excised root are used to derive estimates of root width, height, density and angles. Traits are computed on the basis of the extracted root masks.
for the crown and the excised root with their corresponding root-tip-paths (RTP) as described in the next section.

**Computing Root-Tip-Paths**

The previously computed mask images of the root crown and excised roots serve as input to the RTP algorithm. First, we compute the medial axis (Figure 2 IVc) of the root object using morphological hit-and-miss transform (Serra, 1982) with a 3-by-3 cross mask. The calculation of the medial axis implies knowledge of the distance field (Figure 2 IVb). The distance field defines for each root-pixel the shortest Manhattan distance to a non-root pixel. The distance map can be used to define the medial axis. The medial axis is the collection of points within the root that are equidistant to at least 2 non-root pixels; in this way the medial axis approximates center-lines within the segmented root system (see SI Figure 22). Ideally, the medial axis contains all branching points and tips visible on the extracted root image. This would be the case for plant roots that have been “separated” and placed on a board. Instead for mature roots imaged in the field, occlusion of single roots leads to loops in the medial axis, which does not correspond to the branching structure of the root system. The RTP approach aims to estimate architectural traits between the computations of the medial axis that may include loops, as extracted from the segmented root image, and the loop-free architecture of plant roots. We define a single RTP as the shortest path along the medial axis from a given root tip to the root top, where the root top is the location where the root got separated from the shoot (see SI).

The computation of the shortest paths forming the RTP skeleton leverages Dijkstra’s shortest path algorithm (Goldberg and Harrelson, 2005) from each tip to the start of the root top along the medial axis. Each pixel belonging to the medial axis is associated with a diameter $d_x$ derived from the previously computed distance field. In the following, we weigh the shortest path calculation with $w_{x1,x2}$ for each step from one medial axis pixel $p_1$ to an adjacent pixel $p_2$ as the difference between the maximal stem diameter $d_{max}$ and the average diameter of $p_1$ and $p_2$.

$$w_{x1,x2} = d_{max} - \frac{d_{x1} + d_{x2}}{2}$$

All RTPs of a root assemble a skeleton graph represented as a set of image pixels, where pixels with 2 neighbors are elongated root parts; pixels with 3 or more neighbors are a branching point where multiple root parts intersect. Pixels with only one neighbor correspond to a root tip or the
start of the root stem. The RTP skeleton is weighted by the number of paths as a measure of root thickness at all locations of the skeleton. Underlying the definition of RTPs is the assumption that RTPs to all visible tips in the image approximate the actual root shape well.

The RTP skeleton describes significant root parts by the number of paths passing through the same location in the image. Because RTPs are determined on the basis of detectable tips in the image, a significant root part is either roots with many child roots or larger amounts of adhered roots whose tips are visible in the image. The pseudocode shown in Box 1 demonstrates a simple procedure to compute RTPs from a medial axis input.

A central path with largest average diameter is assumed to be the taproot in dicots. Along the central path branches are counted and measured. For monocots such as maize a ‘central’ root itself is not accessible via the image and the central path only contains information about the start of the root crown, reflected in many paths emerging within a small portion of the central path. Hence, only the fine structure of dense monocot roots such as maize is analyzed with RTPs on an excised root sample. The fine structure of dense monocot roots such as maize is analyzed with RTPs on an excised root sample.

In our example in Figure 2 IId, all paths emerge from the location where the root stem was separated from the shoot. At this location the number of RTPs equals the number of tips detectable in the image. At every branching point of the RTP skeleton a certain number of paths split off into another branch, such that the number of paths present after the split is equal to the number of paths in the respective sub-tree of the branch (see SI Figure 28).

**Traits derived from the RTP skeleton:** The RTP skeleton is a loop free sampling of the medial axis derived from the root shape visible in the image. As a consequence the RTP skeleton allows the calculation of the length of significant root parts, the radius at each location as distance to one of the at least two closest pixels on the root mask boundary, branching frequencies and path length from the distance between branching points and angles between a root emerging from the tissue of a parent root. Computing the root length from the point of emergence to the tip requires care. The length from a branching point to a root tip includes the radius at the branching point, because the skeleton is centered within the root. The tip of the skeleton is also a center with respect to the root tissue. Therefore, the radius at the branching point has to be subtracted from the length and the tip radius has to be added. We used a threshold of three pixels to eliminate
spurious branches (Figure 4c,d) resulting from boundary disturbances of the foreground mask. Note that this filtering is applied after the detection of the central path. Angles are calculated as the angle between the tangent to the parent root at location of the branching point and the line determined by the least squares fit through the emerging path. We calculate two different angles: The soil tissue angle and the root tissue angle. The soil tissue angle is the angle formed by the fit line and the $x$-axis. Nevertheless, the soil tissue angle is dependent on the placement of the root on the imaging board as shown in Figure 12, if angular symmetry of root with the stem cannot be assumed.

This dependency on the root placement motivated the development of the root tissue angle. A tangent is estimated as the least-square fit through a definable window size around a branching point on the central path. This angle calculation takes the local curvature of the nodal root into account and is independent from the root placement on the board, because it is locally related to the root.

The spatial root distribution is represented as a vector. The spatial root distribution is calculated as the difference vector between the centers of the bounding box of the RTP skeleton and the mean coordinate of the RTP skeleton excluding central path. In the given coordinate system of the image the $x$-component of the vector represents the symmetry of the root to the $y$-axis, which is expected to be useful especially in dicots. The $y$-component reflects where the majority of roots are located. Diameter measurements of the central path in predefined intervals relative to the maximal length of the central path can be retrieved on dicots and on excised roots. The distance between branching points along an excised root enables the computation of branching frequency, lateral length and diameters from the RTPs.

**Imaging specifications**

The maize study was performed with a Canon Power Shot A1200 with a focal length of 6mm and a maximal aperture of 3.3. The cowpea case study was imaged with a Nikon D70s using a focal length of 18mm and a maximal aperture of 3.6. In both case studies only optical zoom was used to adjust the field of view in all studies, such that an interpolation effect because of software-enhanced resolution could be avoided.
Images for the maize diversity panel were taken at a resolution of 4000x3000 pixels. Images for which the scale maker was not extractable were removed from the data set and the comparison. Except for the brace root images of B73, 3-6 images of sufficient quality were obtained. Overall 85 images were used. The scale marker varied in 210.47 pixels in the x direction and 217.22 pixels in the y direction with standard deviations of 10.31 and 9.80 respectively. The threshold scale factor for deriving the image mask was 20.0 and we used an 80x80 window for the adaptive thresholding.

The images for the cowpea study were taken at a resolution of 3008x2000 pixels. From the original 1500 images taken we used 1444 that had sufficient quality, such that 6-8 samples per genotype were obtained. The scale marker width is 238.45 pixels in the x direction and 237.79 pixels in the y direction with standard deviations of 5.04 and 2.44 respectively. The threshold scale factor for deriving the image mask was 5.0 and we used an 80x80 window for the adaptive thresholding.

For completeness we cite the imaging setup given in (Iyer-Pascuzzi et al., 2010) for the rice data set:

“Plants were imaged using a PhotoCapture360 turntable and software from Ortery Technologies connected to a Canon PowerShot G7 digital camera and Dell Latitude 620 laptop computer. Cylinders were lit from the sides with fluorescent light bulbs, and from below with a UV light box (VWR); lighting conditions were similar for all experiments. Images from 20 angles per plant per day were acquired.” The threshold scale factor for deriving the image mask was 1.0 and we used a 80x80 window for the adaptive thresholding.”

The images to determine the technical error of the imaging protocol were taken with an Olympus SZ-12 camera having a focal length of 4 mm and a maximal aperture of 3.17 on a Sony VCT-R 640 tripod with ground level adjustment. The images were taken at an original resolution of 4288x3216 pixels. We cut out the image portion in which the root is visible. The background was painted with Blick Artists Black Gesso paint (1 US FL QT). Five researchers took ten
images of a bean and a maize root crown. All persons were exposed to the imaging procedure for the first time. All 50 images per root were correctly computed.
Acknowledgements

We thank Scott Stelpflug for assisting the shovelomics at the URBC in South Africa and pointing us to the book of Walter Fehr. We acknowledge Shawn Kaeppler, Andreas Hund, and Michael Goodisman for valuable advice and two anonymous reviewers for helpful comments on the manuscript. We thank Tsi’tso Mokoena for taking maize images at the URBC, Rodney Managa for initial validation measurements on maize root images, Austin Kirt and Andy Evenson for obtaining preliminary test images and Troy Hilley and Wesley Emeneker for support of the computational pipeline at Georgia Tech.

Online access to the data and a beta version of the source code

We provide access to the data used in this paper via an online application at www.dirt.biology.gatech.edu. This online tool is a system specific solution that allows other researchers to access and use the data. Source code for the algorithms is available as a free download on the website. The online platform is implemented in python 2.7 (Van Rossum and Drake, 2009) using the libraries numpy (Oliphant, 2006) and scipy (Jones et al., 2001) to derive the statistics, mahotas (Coelho, 2013) for image processing and the graph library graph-tools (http://graph-tool.skewed.de). The web application is implemented in Drupal (Graf, 2006). The source code of our algorithms is available at https://github.com/abucksch.
**Contributions**

AB designed the research, developed and implemented imaging algorithms and traits, wrote the manuscript; performed analysis and developed analysis methods; JB designed the research, collected image and shovelomics data, developed traits and co-wrote the manuscript; EN collected image and shovelomics data, analyzed data, selected the maize genotypes and contributed to writing the manuscript; LMY co-designed traits, collected image data and contributed to writing the manuscript; AD customized the algorithms for the Georgia Tech Grid computing environment, executed the computations, developed the public database and analyzed data; JSW analyzed data and co-wrote the manuscript. JPL contributed to the research design and contributed to writing the manuscript.
References


Figure 1: (a) A classic shovelomics scoring board to score the angle of maize roots with the soil tissue and (b) an example to score rooting depth and angle in common bean.

Figure 2: (I) Imaging board on the example of a maize root, the experiment tag is used to capture an experiment number, the scale marker allows the correction of camera tilting and transforming image coordinates into metric units. (II) Camera mounted on a tripod placed on top of the imaging board coated with blackboard paint. Note that images were taken with protection against direct sunlight not shown in the image. (III) Example of the segmentation of the original image into a binary image and then into a series of image masks that serve as input to estimate traits for monocot and dicot roots. The sample is that of a maize root, 40 days after planting at the Ukulima Root Biology Center located in Alma, Limpopo Province, South Africa. (IV) The imaging pipeline for dicot roots and sparse monocot roots: (a) Original image on the imaging board, (b) Derived distance map where the lighter grey level represents a larger diameter of the imaged object, (c) Medial axis includes loops. (d) Loop RTP, sample of the root branching structure. Colors are randomly assigned to each path. The sample is that of a cowpea root, approx. 30 days after planting at the Ukulima Root Biology Center located in Alma, Limpopo Province, South Africa.

Figure 3: Trait correlations in cowpea. (a) Results from field measurements of taproot diameters at 5 cm and 10 cm depth show correlations close to r=0.5 for Spearman and Pearson coefficients. Note not all plants reached 10 cm depth. (b) An example of confirming results with r=0.5 for Spearman and Pearson coefficients was achieved in correlation between the field measured 5 cm depth level and the diameters at 75% of the central path length. (c) The central path diameters extracted from images are shown with similar correlation coefficients than observed with manual field measurements. (d) A surprising correlation; the number of 3rd order basal roots in cowpea is highly correlated with the overall number of RTP’s. Both correlation measures together (Spearman r and Pearson r) suggest even linear dependency. The ANOVA analysis confirms that both measurements are taken from different quantities for both correlations. All diameter measurements are in mm and data points represent averages per genotype.

Figure 4: Examples of maize genotype selection from the Wisconsin Diversity panel. The figure shows increasing root system width and shovelomics angle from left to right. The top row shows the complete root crown, the bottom row shows the root crown with brace roots removed to reveal the crown roots. All six images show a selected representative nodal root at the right of the root crown.

Figure 5: RPV analysis of the crown root measurements. Traits are more likely to be useful in differentiating genotypes when their RPV is significantly greater than 1 (blue line). Trait definitions are found in Table 1 and 2.
Figure 6: Top panel: Image based phenotype differentiation: Normalized mean trait values of traits derived from crown root images. The inter-genotype variation for the crown roots of the 9 examined maize genotypes is shown. The points represent average normalized values. The connection between points allows the reader to visually identify the three genotypes shown in Figure 4. The error bars indicate the standard error of the mean. (STA= soil tissue angle, RTA= root tissue angle, TD = tip diameter, DD90max = maximum diameter at 90% - 100% depth, RTP = Root Tip Path). Note that relative traits only differentiate at certain depth levels. Bottom panel: The number of traits that distinguish a pair of genotypes in the maize study. For each combination at least eight distinguishing traits were found.

Figure 7: Top panel: Phenotype differentiation with shovelomics: normalized mean trait values show the inter-genotype variation for the crown roots of the 9 examined maize genotypes. The points represent average normalized values. The connection between points allows the reader to visually identify the three genotypes shown in Figure 4. All density and distance measures are researcher scores. Bottom panel: The number of traits that distinguish a pair of genotypes in the maize study. For each combination at least one distinguishing trait was found.

Figure 8: Six examples from the cowpea diversity panel. Different root architectures in mature cowpea with the genotype denoted in each image are illustrated.

Figure 9: RPVs of shovelomics and image based traits for the cowpea diversity panel. Traits are more likely to be useful in differentiating genotypes when their RPV is significantly greater than 1 (blue line). Note that the manual traits counting 1st order laterals do not include counts in the basal region of the root. All field scores ranged from 0 (low) - 9 (high).

Figure 10: Overall phenotype differentiation of the cowpea diversity panel. Black crosses show the trait values of the whole dataset, the lines are three selected examples with errorbars denoting the standard error of the mean. The 3 example genotypes distinguish by at least one D- and DS-value and by the second dominant angle. The D-values demonstrate, that UCR 779 differentiates from Petite’n’Green at higher depth levels and from Early Scarlet a deeper depth levels. UCR 779 differentiates additionally in central path diameters from Petite and Green and Early Scarlet. The connection between points allows the reader to visually identify the three of the six genotypes shown in Figure 8. (STA= soil tissue angle, RTA= root tissue angle, TD = tip diameter, DD90max = maximum diameter at 90% - 100% depth, RTP = Root Tip Path, dia. = diameter)

Figure 11: Image masks overlaid with the width profile (red) and the corresponding cumulative width and slope function from which the D-values and DS-values are calculated. (top) Cowpea root. (bottom) Maize root. Note that the width profile does have to match the outline of the root, because the outline is not always symmetric to a vertical and straight line. It is also visible in the cumulative curve that the D-values are robust to single roots sticking out of the crown.

Figure 12: Difference between the root tissue angle (green) and the soil tissue angle (red): The root tissue angle does not change if the root is rotated on the board. In contrast the soil tissue angle on the rotated root on the right is larger than on the left.
Box 1: Principal of the RTP algorithm as pseudocode

**Input:** $MA = \text{[size of original image]}$ //Medial axis as 2D array with weights derived from the distance field,

$RootTop = \text{(row, column)}$ //a pair of values denoting the location of the RootTop on MA

**Output:** $RTPImage$

$RTP = \text{[[(row, column)]]}$ //array of arrays of locations (pair of values) on MA

$RTPImage = \text{[size of MA]}$ //2D array initialized with zeros

# Computing Root-Tip-Paths

For each tip in $MA$ do:

- **compute** shortest path to $RootTop$
- **append** shortest path to $RTP$

# Building a loop free hierarchy

For each path in $RTP$:

For each $(row, column)$ in path:

- **Increment** $RTPImage[(row, column)]$ by 1

**return** $RTPImage$
Table 1: Overview of traits derived for the root stock

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
<th>Closest trait analogue in shovelomics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Root Density:</td>
<td>Ratio between foreground and background pixels of the extracted root.</td>
<td>Scoring of density</td>
</tr>
<tr>
<td>Soil tissue angle X</td>
<td>Angle between a root tip path at x% length and the x-axis calculated as arctan(m), where m is the slope of the fit line.</td>
<td>shovelomics angle at certain depth for dicots</td>
</tr>
<tr>
<td>Dominant soil tissue angle 1 and 2</td>
<td>Average of the 1st and 2nd significant peak in the histogram of calculated soil tissue angles binned in 10 degree steps.</td>
<td>shovelomics angle for dicots</td>
</tr>
<tr>
<td>Min/Median/Max. soil tissue angle:</td>
<td>Minimal, median and maximal of all calculated soil tissue angles.</td>
<td>N/A</td>
</tr>
<tr>
<td>Dominant root tissue angle 1 and 2</td>
<td>Average of the 1st and 2nd significant peak in the histogram of calculated root tissue angles binned in 10 degree steps.</td>
<td>N/A</td>
</tr>
<tr>
<td>Min/Median/Max. root tissue angle:</td>
<td>Minimal, median and maximal of all calculated root tissue angles.</td>
<td>N/A</td>
</tr>
<tr>
<td>Median/max Width:</td>
<td>Median/maximum of the calculated width in the width height diagram.</td>
<td>Only max width or selected representative width</td>
</tr>
<tr>
<td>D 10-90</td>
<td>Accumulated width over the depth at x% of the central path length. The change in width accumulation denotes a change of the root top angle.</td>
<td>N/A, but closely related to the root-top angle for maize.</td>
</tr>
<tr>
<td>DS 10-90</td>
<td>Slope of the graph of central path length vs. accumulated width at x percent of the accumulated width.</td>
<td>N/A</td>
</tr>
<tr>
<td>Spatial root distribution</td>
<td>Displacement of the center of mass between the bounding box of the RTP skeleton and the RTP skeleton excluding the central path.</td>
<td>N/A</td>
</tr>
<tr>
<td>(as separate X and Y components)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stem diameter</td>
<td>Diameter at the beginning of the root top. We use the first branching point of the filtered medial axis as a characteristic diameter.</td>
<td>Caliper measurement</td>
</tr>
<tr>
<td>Central path diameter X</td>
<td>The diameter at x% length of the central path.</td>
<td>Diameter at a defined depth below the soil line.</td>
</tr>
<tr>
<td>Max Diameter at 90% depth</td>
<td>Maximal tip diameter in the last 10% of the image.</td>
<td>Measured for dicots</td>
</tr>
<tr>
<td>Tip diameter distribution</td>
<td>Exponent of the exponential function fitted through the tip diameter histogram.</td>
<td>N/A</td>
</tr>
<tr>
<td>Median/mean tip diameter</td>
<td>Median and mean tip diameters of all tips.</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Table 2: Overview of traits computed for the excised root sample

**Excised root traits**

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
<th>Closest trait analogue in shovelomics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodal root length:</td>
<td>Length of the central path along the excised root.</td>
<td>Scored from representative sample</td>
</tr>
<tr>
<td>Avg. nodal root diameter</td>
<td>Avg. root diameter along the central path of the excised root.</td>
<td>N/A</td>
</tr>
<tr>
<td>Distance to first lateral</td>
<td>Path length from the top pixel of the excised root to the first detected emerging lateral of the filtered medial axis.</td>
<td>Distance to first lateral as score</td>
</tr>
<tr>
<td>Lateral branching frequency</td>
<td>Estimate of the lateral branching frequency calculated as: (number of Laterals) / (nodal root path length).</td>
<td>Density scores</td>
</tr>
<tr>
<td>Avg. lateral length:</td>
<td>Average length of the detected lateral roots emerging from the central path along the excised roots</td>
<td>Length score</td>
</tr>
<tr>
<td>Lateral min/mean/max root tissue angle</td>
<td>Min/mean/Max angle of detected laterals emerging from the central path along the excised root.</td>
<td>N/A</td>
</tr>
<tr>
<td>Median/mean lateral diameter</td>
<td>Median/mean of all lateral tip diameters</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Figure 1: (a) A classic shovelomics scoring board to score the angle of maize roots with the soil tissue and (b) an example to score rooting depth and angle in common bean.
Figure 2: (I) Imaging board on the example of a maize root, the experiment tag is used to capture an experiment number, the scale marker allows the correction of camera tilting and transforming image coordinates into metric units. (II) Camera mounted on a tripod placed on top of the imaging board coated with blackboard paint. Note that images were taken with protection against direct sunlight not shown in the image. (III) Example of the segmentation of the original image into a binary image and then into a series of image masks that serve as input to estimate traits for monocot and dicot roots. The sample is that of a maize root, 40 days after planting at the Ukulima Root Biology Center located in Alma, Limpopo Province, South Africa. (IV) The imaging pipeline for dicot roots and sparse monocot roots: (a) Original image on the imaging board, (b) Derived distance map where the lighter grey level represents a larger diameter of the imaged object, (c) Medial axis includes loops. (d) Loop RTP, sample of the root branching structure. Colors are randomly assigned to each path. The sample is that of a cowpea root, approx. 30 days after planting at the Ukulima Root Biology Center located in Alma, Limpopo Province, South Africa.
Figure 3: Trait correlations. (a) Results from field measurements of taproot diameters at 5 cm and 10 cm depth show correlations close to $r=0.5$ for Spearman and Pearson coefficients. Note not all plants reached 10 cm depth. (b) An example of confirming results with $r=0.5$ for Spearman and Pearson coefficients was achieved in correlation between the field measured 5 cm depth level and the diameters at 75% of the central path length. (c) The central path diameters extracted from images are shown with similar correlation coefficients than observed with manual field measurements. (d) A surprising correlation; the number of 3rd order basal roots in cowpea is highly correlated with the overall number of RTP’s. Both correlation measures together (Spearman $r$ and Pearson $r$) suggest even linear dependency. The ANOVA analysis confirms that both measurements are taken from different quantities for both correlations. All diameter measurements are in mm and data points represent averages per genotype.
Figure 4: Examples of maize genotype selection from the Wisconsin Diversity panel. The figure shows increasing root system width and shovelomics angle from left to right. The top row shows the complete root crown, the bottom row shows the root crown with brace roots removed to reveal the crown roots. All six images show a selected representative nodal root at the right of the root crown.
Figure 5: RPV analysis of the crown root measurements. Traits are more likely to be useful in differentiating genotypes when their RPV is significantly greater than 1 (blue line). Trait definitions are found in Table 1 and 2.
Figure 6: Top panel: Image based phenotype differentiation: Normalized mean trait values of traits derived from crown root images. The inter-genotype variation for the crown roots of the 9 examined maize genotypes are shown. The points represent average normalized values. The connection between points supports the reader visually to identify the 3 genotypes shown in Error! Reference source not found.. The error bars indicate the standard error of the mean. (STA= soil tissue angle, RTA= root tissue angle, TD = tip diameter, DD90max = maximum diameter at 90% - 100% depth, RTP = Root Tip Path). Note that relative traits only differentiate at certain depth levels. Bottom panel: The number of traits that distinguish a pair of genotypes in the maize study. For each combination at least eight distinguishing traits were found.
Figure 7: Top panel: Phenotype differentiation with shovelomics: normalized mean trait values show the inter-genotype variation for the crown roots of the 9 examined maize genotypes. The points represent average normalized values. The connection between points supports the reader visually to identify the 3 genotypes shown in Error! Reference source not found. All density and distance measures are researcher scores. Bottom panel: The number of traits that distinguish a pair of genotypes in the maize study. For each combination at least one distinguishing trait was found.
Figure 8: Six examples from the cowpea diversity panel. Different root architectures in mature cowpea with the genotype denoted in each image are illustrated.
Figure 9: RPVs of shovelomics and image based traits for the cowpea diversity panel. Traits are more likely to be useful in differentiating genotypes when their RPV is significantly greater than 1 (blue line). Note that the manual traits counting 1st order laterals do not include counts in the basal region of the root. All field scores ranged from 0 (low) - 9 (high).
Figure 10: Overall phenotype differentiation of the cowpea diversity panel. Black crosses show the trait values of the whole dataset, the lines are three selected examples with errorbars denoting the standard error of the mean. The 3 example genotypes distinguish by at least one D- and DS-value and by the second dominant angle. The D-values demonstrate, that UCR 779 differentiates from Petite’n’Green at higher depth levels and from Early Scarlet a deeper depth levels. UCR 779 differentiates additionally in central path diameters from Petite and Green and Early Scarlet. The connection between points supports the reader visually to identify the 3 genotypes shown in Figure 8. (STA= soil tissue angle, RTA= root tissue angle, TD = tip diameter, DD90max = maximum diameter at 90% - 100% depth, RTP = Root Tip Path, dia. = diameter)
Figure 11: Image masks overlaid with the width profile (red) and the corresponding cumulative width from which D-values are calculated. The function of DS-values is calculated as tangent slopes per D-value. Here the estimation of tangents was performed in a 100 pixel neighborhood of the corresponding D-value. (top) Cowpea root. (bottom) Maize root. Note that the width profile does not have to match the outline of the root, because the outline is not always symmetric to a vertical and straight line. It is also visible in the cumulative curve that the D-values are robust to single roots sticking out of the crown.
Figure 12: Difference between the root tissue angle (red) and the soil tissue angle (green): The root tissue angle does not change if the root is rotated on the board. In contrast the soil tissue angle on the rotated root on the right is larger than on the left.