the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2020) doi: 10.1111/tpj.14865

PERSPECTIVE

The Cassava Source–Sink project: opportunities and challenges for crop improvement by metabolic engineering

Uwe Sonnewald[1], Alisdair R. Fernie[2], Wilhelm Gruissem[3,4], Pascal Schläpfer[3], Ravi B. Anjanappa[3], Shu-Heng Chang[4], Frank Ludewig[1,†], Uwe Rascher[5], Onno Muller[5], Anna M. van Doorn[5,‡], Ismail Y. Rabbi[6] and Wolfgang Zierer[1,*]

[1]*Department of Biology, Division of Biochemistry, Friedrich-Alexander-University Erlangen-Nuremberg, Staudtstrasse 5, Erlangen 91058, Germany,*

[2]*Max-Planck-Institute of Molecular Plant Physiology, Am Mühlenberg 1, Potsdam 14476, Germany,*

[3]*Department of Biology, Plant Biotechnology, ETH Zurich, Universitaetstrasse 2, Zurich 8092, Switzerland,*

[4]*Advanced Plant Biotechnology Center, Institute of Biotechnology, National Chung Hsing University, Xingda Road, South District, Taichung City 402, Taiwan,*

[5]*Forschungszentrum Jülich GmbH, Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Leo-Brandt-Str, Jülich 52425, Germany, and*

[6]*International Institue for Tropical Agriculture, Oyo Road, Ibadan, Oyo State, 200001, Nigeria*

Received 14 April 2020; revised 22 May 2020; accepted 2 June 2020.

*For correspondence (e-mail wolfgang.zierer@fau.de).

[†]Present address: KWS Saat SE, Grimsehlstraße 31, Einbeck, 37574, Germany

[‡]Present address: International Institue for Tropical Agriculture, Oyo Road, Ibadan, Oyo State, 200001, Nigeria

SUMMARY

Cassava (*Manihot esculenta* Crantz) is one of the important staple foods in Sub-Saharan Africa. It produces starchy storage roots that provide food and income for several hundred million people, mainly in tropical agriculture zones. Increasing cassava storage root and starch yield is one of the major breeding targets with respect to securing the future food supply for the growing population of Sub-Saharan Africa. The Cassava Source–Sink (CASS) project aims to increase cassava storage root and starch yield by strategically integrating approaches from different disciplines. We present our perspective and progress on cassava as an applied research organism and provide insight into the CASS strategy, which can serve as a blueprint for the improvement of other root and tuber crops. Extensive profiling of different field-grown cassava genotypes generates information for leaf, phloem, and root metabolic and physiological processes that are relevant for biotechnological improvements. A multi-national pipeline for genetic engineering of cassava plants covers all steps from gene discovery, cloning, transformation, molecular and biochemical characterization, confined field trials, and phenotyping of the seasonal dynamics of shoot traits under field conditions. Together, the CASS project generates comprehensive data to facilitate conventional breeding strategies for high-yielding cassava genotypes. It also builds the foundation for genome-scale metabolic modelling aiming to predict targets and bottlenecks in metabolic pathways. This information is used to engineer cassava genotypes with improved source–sink relations and increased yield potential.

Keywords: source, sink, Cassava, *Manihot esculenta*, biotechnology, yield.

INTRODUCTION

The world population is expected to increase from a currently projected 7.8–9.7 billion people in 2050 and may reach a peak at 11 billion by the end of this century (United Nations, D.o.E.a.S.A., Population Division, 2019). This population growth will mainly take place in Sub-Saharan Africa (SSA). Together with the increasing per capita calorific consumption, the need for improving crop yield is more acute than ever. In the past, agricultural productivity significantly benefited from technological and scientific progress that facilitated agronomic and breeding improvements.

The breeding of semi-dwarf rice and wheat genotypes during the Green Revolution significantly reduced the problem of lodging and thereby increased grain yield. Together with the use of fertilizers, pesticides and irrigation systems, this led to a massive increase in agricultural output in the USA, Europe, Asia, and Latin America. A similar development was largely missed in SSA, where smallholder farmers produce approximately 80% of the local agricultural output. Most of them cannot afford larger investments in improved crop varieties, pesticides or fertilizers to increase yield. In addition, until recently, important African crop plants such as cassava and cowpea did not benefit from technological developments. Consequently, *de facto* food production per capita in many African countries declined. Among the tropical root and tuber crops, cassava is the most important root crop, reaching an annual world production of 290 million tons in 2017. Cassava contributes 2.6% of the global caloric intake from all sources and is the fourth most important staple food in the world after maize, rice, and wheat. With almost 59.5 million tons in 2017, Nigeria is the world's largest cassava producer (FAOSTAT, 2019). However, in terms of yield (tons ha^{-1}), Nigeria produces less than 80% of the world average and approximately three-fold less than Laos, which is the highest producer per hectare (FAOSTAT, 2017). Indeed, although the cassava area harvested in Nigeria has almost doubled since 2007, storage root yield per hectare declined by more than 20% in 2017 (FAOSTAT, 2019; Otekunrin and Sawicka, 2019). Nigeria is not unique in this respect because this problem exisits in all of SSA. Major threats for cassava yield are pathogens causing major diseases such as bacterial leaf blight, cassava mosaic disease (CMD) or cassava brown steak disease (McCallum *et al.*, 2017), and physiological postharvest deterioration of storage roots (Zainuddin *et al.*, 2018), as well as low or no input of fertilizer and chemicals for pest and weed control (FAO, 2013).

Cassava (*Manihot esculenta* Crantz) is a woody perennial plant and belongs to the family of *Euphorbiaceae*. Propagation and planting of cassava occur via stem cuttings. Cassava grows well in poor soils and is adapted to low rain falls. The flexible harvesting time of cassava makes it a popular and important food security crop. Despite its importance for SSA and as a global food source, cassava has received little scientific interest until recently. As a result, our current understanding of the most basic biology of cassava root development and yield determinants is very limited.

In other plant species, efficient assimilate allocation between photosynthetically active tissues (source tissues, mainly leaves) and storage tissues (sink tissues) such as seeds, roots or tubers is a major yield determinant. Based on molecular and physiological studies, many factors have been identified that impact source-to-sink interaction and assimilate allocation. This facilitated the design of transgenic plants with improved trait characteristics and yield under experimental conditions. Assuming no significant biotic and abiotic stress, storage root growth, root-starch content, and post-harvest characteristics determine the final yield of cassava storage roots and dry matter (starch). Developing roots usually compete with other sink tissues for photoassimilates. The import capacity of photoassimilates and the conversion of these metabolites into storage molecules (mainly starch in case of cassava) determine the competitiveness, or sink strength, of the organ. Strategies to facilitate storage root growth and starch accumulation are most promising for increasing cassava yield and securing a sufficient food supply in SSA and other regions in the world. To enable these strategies in cassava, it is important to understand metabolic processes in source and sink tissues, to analyze the genomic diversity for storage root yield and starch content in different cassava genotypes, to test transgenic strategies for increasing sink strength that have already been successful in other crop plants, and to understand the seasonal dynamics and gene $\times$ environment interactions under field conditions.

Based on our previous work, we have now a basic understanding of cassava source and sink metabolism. Cassava performs C3 photosynthesis (Arrivault *et al.*, 2019) and leaf photoassimilates are loaded apoplasmically into the phloem (Mehdi *et al.*, 2019). Nodal-derived fibrous roots, emerging on planted cassava stem pieces, develop into storage roots by secondary growth. Storage roots are characterized by a well-organized vascular cambium between phloem and xylem. Longitudinal, cambium-derived vascular ray cells bridge the two cell types ensuring exchange of water, nutrients, and carbohydrates (Figure 1). Unloading of the photoasimilates in storage roots follows a symplasmic route to the storage parenchyma cells and is facilitated by vascular rays (Mehdi *et al.*, 2019). Alternating ray initial cells and fusiform initial cells in the vascular cambium give rise to the vascular ray cells and xylem/phloem cells, respectively. By contrast to the ray initial cells, which are connected to the root symplast, the fusiform initial cells are nutritionally supported by apoplastic transport. Cassava genes encoding starch biosynthesis enzymes, which are similar to those important for starch biosynthesis in potato tubers, are highly expressed during storage starch synthesis (Yang *et al.*, 2011; Wang *et al.*, 2016). Hence, sink–source concepts developed for potato are likely applicable to cassava as well. Over the last 30 years, the impressive number of transgenic plants that has been generated has greatly enhanced our fundamental understanding of source to sink interactions. In parallel, quantitative genetic approaches have been initiated to address the same question. However, with a few notable exceptions, most engineering approaches have failed to establish a convincing applied biotechnology perspective. Three main reasons can

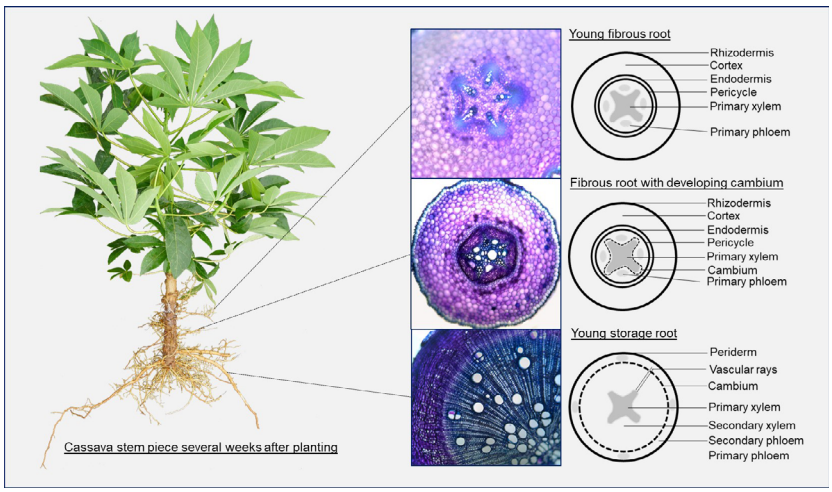

Figure 1. Illustration of cassava fibrous and storage roots. Cassava is typically planted from stem sticks. Nodal-derived fibrous roots are formed within 2–4 days after planting. Fibrous roots display a primary vascular anatomy characterized by a central vascular cylinder containing star-shaped primary xylem alternating with primary phloem. Approximately 25–30 days after planting, the first signs of secondary root growth can be found. Nodal-derived fibrous roots develop a vascular cambium and start to enlarge by building new xylem and phloem cells. The central vascular cylinder is eventually broken by the expanding phloem/xylem and a typical secondary root growth anatomy can be observed. Slightly enlarged storage roots characterized by a well-organized vascular cambium between phloem and xylem can routinely be observed approximately 30–40 days after planting. The periderm can also be observed at this stage. Longitudinal, cambium-derived vascular rays bridge the phloem and xylem cells ensuring continued exchange of water, nutrients and carbohydrates. Storage roots now continuously enlarge, forming predominantly xylem parenchyma cells in which starch and other molecules can be stored.

explain these failures: first, a missing comprehensive understanding of the inherent complexity of source–sink relationships; second, the reliance on single genes in first generation interventions to alter these relationships; and, third, the complexity of gene expression under the dynamically changing environmental conditions, such as the seasonal change between wet and dry periods in many production regions. Today, current knowledge combined with increasing sophistication in both mathematical modelling and multi-gene targeting strategies now allows us to deploy novel strategies for improving source to sink relations in cassava to increase cassava root yield. Combinatorial approaches to tackle multiple metabolic processes simultaneously and the integration of developmental and metabolic processes into novel concepts of source–sink relations have recently been discussed (Sonnewald and Fernie, 2018; Fernie *et al.*, 2020).

Following the concept of simultaneously improving source-, allocation-, and sink-processes, the Cassava Source-Sink (CASS) project focuses on the concerted engineering of metabolic and physiological processes in source, transport, and sink tissues by designing multigene constructs and testing their performance for storage root and starch yield increases in different cassava genotypes to support smallholder farmers in SSA (Figure 2). To achieve our goal, we combine systems biological studies of diverse African cassava genotypes and biotechnological approaches to boost metabolic reactions and facilitate developmental processes targeted at increased storage root and starch production. We use advances in next generation sequencing and MS-based profiling of metabolites, transcripts, and proteins, and also combine computational metabolic modelling with expert knowledge, to predict limiting steps in cassava metabolism and physiology. Empowered by rapid gene synthesis, we design multigene constructs and introduce them into cassava genotypes by *Agrobacterium tumefaciens*-mediated transformation. Following their molecular characterization, we are testing these genotypes in confined field trials for their above ground growth rates and biomass allocation using non-invasive phenotyping approaches. Based on the molecular and biochemical data together with the phenotypic data from the field, we then review and refine our strategic engineering concept. This iterative process greatly advances cassava biotechnology and generates novel genotypes for targeted cassava breeding based on validated molecular and biochemical processes.

LEARNING FROM GENETIC DIVERSITY OF CASSAVA

To date, the CASS project has pursued three parallel approaches to identify target genes and biochemical processes that are relevant for cassava storage root yield and dry matter composition. Two of these generated information about the inherent variation in natural cassava populations that had not been previously evaluated, whereas the third approach produced predictions of metabolic bottlenecks and target genes from genome-scale metabolic models (Figure 2).

During the first phase of the CASS project, and similar to the Bill & Melinda Gates Foundation-supported project

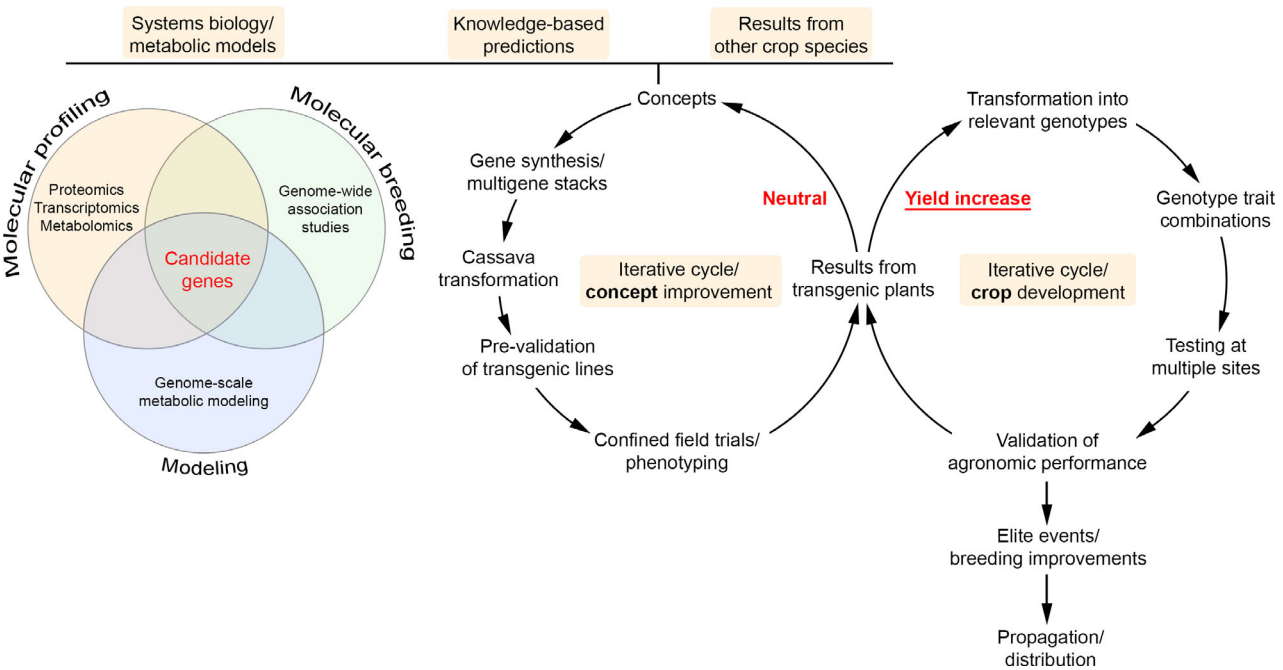

Figure 2. Schematic representation of the Cassava Source–Sink (CASS) strategy to improve cassava root and starch yield. Cassava genotypes are profiled for their metabolome, transcriptome and proteome. Data are compared with genome-wide association studies and metabolic predictions derived from genome-scale metabolic models. The genes and processes identified in all three approaches are candidate targets for genetic or biotechnological improvement. Together with knowledge-based predictions and results from other crop plants, the gene targets are tested in iterative cycles of cassava transformation, field testing, and agronomic performance evaluation. Top performing lines will enter iterative cycles of product development for high-yielding, resistant, and farmer-preferred genotypes.

African Cassava Whitefly (Perez-Fons *et al.*, 2019), considerable optimization of methods for metabolite profiling was required. This resulted in standard operating procedures across projects for metabolite and starch quality assessments in different cassava tissues (Rosado-Souza *et al.*, 2019) and for enzyme activity measurements (Arrivault *et al.*, 2019). In addition, six African cultivars of cassava were grown either in the greenhouse (at the University of Erlangen, Germany) or in the field (at the International Institute for Tropical Agriculture, Ibadan, Nigeria) to analyze source leaves, sink leaves, stems, and storage roots during storage root bulking for their metabolite, ion, and enzyme content (Obata *et al.*, 2020) and to establish a procedure for unmanned aerial vehicle (UAV)-based quantification of the three-dimensional canopy structure and its dynamical changes throughout the season (Van Doorn *et al.*, 2020). Although the number of genotypes in this study was small, a number of important conclusions could be drawn from this combined metabolomics and ionomics study. First, the high ratio of ribulose-1,5-bisphosphate carboxylase to phospho*enol*pyruvate carboxylase activities is consistent with a C3-type photosynthesis. The generally high Calvin–Benson cycle enzyme activities can also explain the high rates of photosynthesis in cassava. Second, despite the high rates of photosynthesis, cassava appears to be source-limited because root yield correlated with leaf metabolite profiles rather than metabolite profiles of stems or storage roots (De Souza and Long, 2018; Obata *et al.*, 2020). Third, compared to greenhouse-grown plants, field-grown plants produced more and larger storage roots, which was associated with higher ADP-glucose pyrophosphorylase (AGPase) activity and lower sucrose levels in storage roots. This is consistent with earlier studies reporting that overexpression of AGPase in cassava increases root yield, possibly via feedforward loops that enhance sink capacity in the high light, low nitrogen environment in the field (Ihemere *et al.*, 2006). Collectively, the results revealed that the carbon assimilation rate (De Souza *et al.*, 2019), the potassium battery that energizes phloem transport (Gajdanowicz *et al.*, 2011), root starch synthesis, leaf trehalose, and chlorogenic acid accumulation (Obata *et al.*, 2020) are potential targets for the genetic improvement of storage root yield. In parallel to this genotype-centered study, we completed comprehensive transcriptome, proteome, and metabolome analyses of cassava root development. Stems of the high-yielding cassava genotype TME 419 were planted and root growth was monitored. Two types of roots could be phenotypically distinguished 3–4 weeks after planting: the fibrous roots and possible early storage roots that have a darker brown color, likely reflecting newly formed periderm. This protective tissue layer at the root surface forms during secondary root growth and often appears darker because of

high levels of lignin, suberin, and other phenolic compounds (Campilho *et al.*, 2020). Fibrous roots and developing storage roots were analyzed by RNA-sequencing and metabolite profiling. Although transcriptome and metabolome data did not distinguish between root types early during development, this changed over time. Metabolomics and ionomics revealed that nitrate levels strongly decreased, whereas phosphate, starch, and amino acid levels strongly increased, which is characteristic for a shift from a nutrient uptake organ to a storage organ. The increase in storage root carbohydrate levels was reflected by changes in the expression of genes for cell wall enzymes, as well as sugar- and starch-related enzymes. Interestingly, the mode of sucrose utilization shifted from invertase (in fibrous roots) to sucrose synthase (in storage roots). These results suggested a shift in sugar transport from an apoplasmic mode in fibrous roots to a symplasmic mode in storage roots. This was subsequently confirmed in cassava plants expressing phloem-specific green fluorescent protein and by detailed histological analysis using apoplasmic and symplasmic tracer molecules (Mehdi *et al.*, 2019). In addition, many transcription factors were differentially expressed between fibrous roots and during storage root development. Further studies will allow the regulatory networks controlling the initiation and development of storage roots to be deciphered.

Together, these two studies highlight the value of a 'guilt-by-association' approach in identifying targets for metabolic engineering and in expanding the list from those that have shown positive effects in other crop plants to include new targets specific for cassava.

Genome-wide association studies (GWAS) are becoming increasingly important in crop breeding because they connect genotypes with phenotypes (Fernie and Gutierrez-Marcos, 2019). In cassava, GWAS of 672 cassava clones and 72 000 single nucleotide polymorphism loci used the yellow color intensity (yellowness) of storage roots to indirectly assess variation in carotenoid and dry matter content (Rabbi *et al.*, 2017). This study built on earlier work using bi-parental crosses to detect quantitative trait loci for these traits. Rabbi *et al.* (2017) identified two major loci for yellowness and one for dry matter that co-localized with one of the yellowness loci, although changes in the traits were negatively correlated. The identified genomic region identified was relatively large as a result of the low rate of recombination in the genomic region, although the genes encoding UDP-glucose pyrophosphorylase and sucrose synthase mapped to this region. This suggests a role for these enzymes in determining storage root yield, which would be consistent with genetic studies in sweet potato that link sucrose synthase to tuber yield (Gemenet *et al.*, 2020) and the fact that overexpressing the enzyme in potato increased tuber yield (Baroja-Fernandez *et al.*, 2009). Today, the availability of chromosome scale cassava reference genomes (Prochnik *et al.*, 2012; Kuon *et al.*, 2019) and application of next generation sequencing methods (Tecle *et al.*, 2014; Rabbi *et al.*, 2014a; Rabbi *et al.*, 2014b; Bredeson *et al.*, 2016) makes GWAS a more comprehensive tool for assessing the genetic and phenotypic variation of the cassava gene pool (Zhao *et al.*, 2011). Extending GWAS approaches to a range of additional metabolic and physiological traits to reveal their interactions would therefore be a highly useful strategy with respect to the identification of additional targets for the genetic improvement of cassava.

USING GENOME-SCALE METABOLIC MODELING TO REVEAL TARGET GENES

Several genetic loci for target traits have been identified in cassava using the approaches described above and potentially many more will be uncovered by expanding the combination of -omics and GWAS methods. Typically, complex traits are multigenic and genetic dissection can help prioritize some genes over others, although this analysis is time consuming. For example, previous work on complex traits such as starch content in storage organs has mostly focused on enzymes or metabolite transporters involved in starch metabolism. Considering that storage root starch content based on dry matter may already be as high as 85% in modern cassava varieties (Beyene *et al.*, 2018), further increases in overall dry yield may be difficult to achieve by focusing on starch metabolism alone. It would therefore be useful to identify potential feedback mechanisms in cassava and other crops that could trigger storage organ growth when starch content increases. Computational approaches have been used successfully for modelling complex traits such as biomass accumulation (Schwender *et al.*, 2004; Kromdijk *et al.*, 2016). Thus, it would be pragmatic to exploit genome-scale metabolic (GSM) modeling to rapidly simulate genetic scenarios without the time-consuming analysis of multiple generations of plants (Sonnewald and Fernie, 2018).

Plant GSM models that simulate growth have an advantage over dynamic models in that they consider the entire metabolism and thus can generate a larger scale prediction of candidate target genes for genetic engineering strategies. Although they are less precise in simulating metabolism than dynamic models and less accurate in simulating yield than phenomenological simulations trained on plant growth data, they nevertheless can provide mechanistic explanations at the level of metabolism. This is especially useful for cassava storage root development and starch production, which occurs over 8–10 months. For example, to produce 10% more storage root starch, a daily efficiency gain of one per mille in starch biosynthesis may be sufficient to reach this yield increase. A comparative transcriptome or proteome analysis between cultivars that have high and low storage root starch content may not be able

to explain such a small daily efficiency gain at the molecular level. GSM models would reveal small, but recurring differences in metabolism, even in biochemical pathways outside of the central metabolism. Thus, GSM models can guide genetic engineering strategies by identifying target genes and preferable combinations of genes based on changes in metabolism.

We used cassava genome information (Bredeson *et al.*, 2016; Kuon *et al.*, 2019) to assemble the necessary metabolic toolset for building cassava GSM models that simulate metabolism in source (leaf) and sink (storage root) tissues. These models help us to identify likely metabolic bottlenecks that can be engineered to increase leaf productivity, to facilitate metabolite transport, and to increase storage root growth and starch production in storage roots. They also help to highlight previously unconsidered enzyme co-factor concentrations and side product inhibitions. Different strategic outcomes can be clearly stated as problems that are implemented as mathematical equations. Metabolic aspects such as carbon fixation in source leaves, co-factor use in enzymatic reactions or respiration-induced carbon loss all contribute to storage root growth and starch accumulation. They all can be modelled independently, although the alteration of one aspect of metabolism often affects other aspects. Only their integrated analysis in GSM models produces predictions for a portfolio of useful modifications. These are implemented via genetic engineering strategies using various DNA constructs for the coordinated expression of genes with established functions and GSM-model predicted genes. Additional data are needed, however, to further improve the predictive power and precision of the GSM model-guided metabolic engineering approach. This is generally the case for GSM models of most organisms and particularly for crop plants. For example, detailed information on organ growth as a function of time, activities of key enzymes, and validation of predicted enzyme functions will be valuable instructions for enabling GSM models to become widely used breeding and genetic engineering tools.

CONSIDERATIONS PRIOR TO ENGINEERING CASSAVA PLANTS FOR HIGHER YIELDS

The CASS research project is based on the premise that simultaneous acceleration of leaf photosynthesis, phloem transport, storage root development, and starch biosynthesis will result in higher cassava yields. At the start, the project was confronted with limited knowledge about basic physiological processes, a lack of biochemical data for central metabolic pathways, and a limited toolbox for transgene expression. As a result, we initially built comprehensive biochemical and physiological databases and applied genetic engineering strategies that had already been successfully used in other crop species. Currently, few attempts have been made to simultaneously engineer source, transport, and sink processes. Therefore, we concentrated on studies in which individual metabolic pathways had been altered in other crop plants and were shown to have beneficial effects on biomass production. Based on previous experience, we focused on specific source and sink metabolic pathways and selected genes that likely have important functions in these pathways for expression in cassava. To facilitate multigene cloning strategies, we designed and synthesized genes that were codon-optimized for cassava. Specific promoters are required to enable the cell-, tissue-, and organ-specific expression of the selected genes. Because reports on promoter specificities are very limited for cassava, we decided to build a CASS-promoter tool box of leaf-, phloem-, and storage root-specific promoters. Candidate promoter sequences were either selected from characterized promoters in other plant species (preferably root and tuber crops) or predicted from cassava transcriptome and proteome data. To test their specificity, reporter genes were fused to selected promoter regions and expressed in transgenic cassava plants grown under greenhouse or field conditions.

Finding strong and highly specific promoters for storage roots has been challenging. Typical storage sink-specific promoters often originate from genes encoding enzymes involved in the synthesis of storage compounds (i.e. fatty acids, carbohydrates or proteins). Promoters of genes for starch biosynthesis enzymes, such as the promoter for the granule-bound starch synthase gene, show storage sink-specific activity in many plant species. However, their activity is also tightly regulated by the sugar status of the plant/organ and therefore dependent on developmental and environmental factors (Koch, 1996). Storage protein-specific gene promoters (e.g. the promoter of the potato patatin gene) have the most specific activity in cassava storage roots. Currently, we are focusing on the identification and testing of additional storage root promoters, as well as phloem- and cambium-specific promoters. In the case of phloem-specific expression, the applicability of the Arabidopsis *SUCROSE TRANSPORTER 2* (*AtSUC2*) promoter could be confirmed (Mehdi *et al.*, 2019).

SIMULTANEOUS ENGINEERING OF MULTIPLE METABOLIC PROCESSES

CASS biotechnology strategy aims at increasing cassava storage root and starch yield by simultaneously increasing source (photosynthesis, sucrose biosynthesis, phloem loading) and sink (sucrose-to-starch conversion) metabolism (Figure 3). To enable the simultaneous genetic manipulation of several target genes, a transformation construct pipeline based on the Golden Gate vector system (Engler *et al.*, 2008; Engler *et al.*, 2009; Engler *et al.*, 2014) was adapted for cassava. This required the synthesis of a

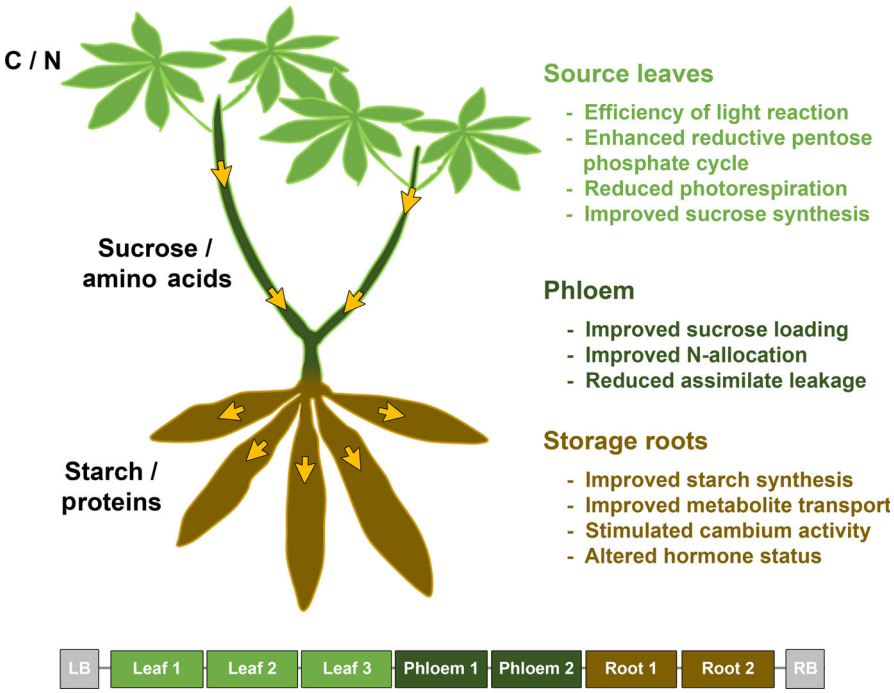

Figure 3. Schematic representation of target processes in leaves, phloem, and storage roots to facilitate assimilate allocation for increased accumulation of storage root compounds (starch, proteins). Leaf 1–3, gene stacks for manipulation of leaf metabolism; Phloem 1–2, gene stacks for manipulation of phloem metabolism; Root 1–2, gene stacks for the manipulation of root processes.

cassava-optimized transformation vector and the selection of suitable promoters for root- and leaf-specific expression of the transgenes.

A typical CASS construct combines genes involved in photosynthesis, transport, and sink metabolism; for example, the *Escherichia coli GLYCOLATE DEHYDROGENASE*, the *Arabidopsis TONOPLAST SUGAR TRANSPORTER* (*AtTST*), the *Pisum sativum GLUCOSE-6-PHOSPHATE/PHOSPHATE TRANSLOCATOR 2* (*PsGPT2*), and the *Arabidopsis NUCLEOTIDE TRANSLOCATOR 1* (*AtNTT1*). The *E. coli* glycerol dehydrogenase polyprotein (EcGlyDH), when expressed in chloroplasts, functions as a photorespiratory bypass by converting 2-phosphoglycolate to glycerate and CO_2 directly in the chloroplast, thereby reducing photorespiratory energy loss and increasing yield (Kebeish *et al.*, 2007; Nölke *et al.*, 2014; Dalal *et al.*, 2015; South *et al.*, 2019). In potato, for example, overexpression of *EcGlyDH* resulted in increased photosynthesis, higher sugar and starch levels, and an increase in tuber yield under greenhouse conditions (Nölke *et al.*, 2014). Therefore, overexpression of *EcGlyDH* in the leaves of a tropical C3 plant such as cassava should result in reduced photorespiratory energy loss and provide increased amounts of carbohydrates, mostly in the form of sucrose, for export, and storage. We combined *EcGlyDH* with *AtTST* to increase photosynthesis and sugar export from the leaf. By shifting glucose from the cytoplasm to the vacuole, AtTST alters the subcellular sugar partitioning. Because metabolic feedback regulation is a potential limitation of photosynthesis (Paul and Foyer, 2001), the expression of AtTST can alter cellular sugar signaling (Wingenter *et al.*, 2010). Together, these two adjustments to photosynthetic metabolism should increase cassava photosynthetic productivity to facilitate the production of fixed carbon and sucrose. However, we expect increased source capacity without high sink demand to result in a buildup of assimilates in leaves and feedback inhibition of photosynthesis. To avoid such buildup, we combined the leaf-specific expression of *EcGlyDH* and *AtTST* with root-specific overexpression of *PsGPT2* and *AtNTT1*. PsGPT2 and AtNTT1 transport glucose 6-phosphate and ATP, respectively, into the amyloplast to facilitate starch synthesis. Both hexose-phosphates and ATP are major building blocks of starch and simultaneous overexpression of both transporters greatly enhances starch accumulation in potato tubers (Zhang *et al.*, 2008; Jonik *et al.*, 2012). The above genes and their functions illustrate the use of selected targets in the CASS project. To date, we have designed more than 25 different combinations of genes for source–sink metabolic engineering and transformed the multi-stack gene constructs into cassava. The project is now continuously generating new combinations of source–sink multi-stack gene constructs and transgenic plants for testing their agronomic performance in the greenhouse and the field.

VALIDATION OF METABOLIC CONCEPTS

CASS has established a multi-national pipeline for cassava source–sink metabolic engineering that covers all of the steps from construct design to transformation, greenhouse testing, and confined field trials (CFT) for agronomic performance evaluations in Nigeria and Taiwan (Figure 4a). After successful construct design and assembly, multigene constructs are routinely transformed into the cassava genotypes 60444 and TME 7, a farmer-preferred variety, by Agrobacterium-mediated gene transfer (Bull *et al.*, 2009). Cassava friable embryonic calli are generated, transformed, and placed on selection medium. After several selection and regeneration steps, transgenic cassava shoots emerge that are propagated into plantlets (Figure 4b,c).

Considering the size of the T-DNAs transferred into the cassava genome, we have encountered no problems with the transformation of 35-kb multigene constructs. Most of the single event transgenic lines have full-length T-DNA inserts and express the transgenes. Transgenic lines are also being maintained in tissue culture as a backup and for further use (Figure 4c). Pre-screening of plant growth and phenotype, as well as transgene expression in leaves and roots, can be carried out in the greenhouse (Figure 4d–f). However, cassava plants growing in pots in the greenhouse hardly form storage roots and the source–sink balance of the plants does not resemble the natural biomass distribution observed in the field. Therefore, testing of physiological, biochemical, and yield parameters needs to be performed under appropriate field conditions. CASS currently uses an approximately 2500 m^2 field site at the National Chung-Hsing University in Taiwan, which has a subtropical climate appropriate for cassava growth from February to November. After obtaining the necessary permissions for the cassava import and CFT, the tissue-culture plantlets arrive at the greenhouse of the Experimental Station (Figure 4g). The tissue-culture plantlets are then transferred to soil and hardened in the greenhouse for 8 weeks. Prior to planting, every individual plant is barcoded and labeled (Figure 4h). Cassava plants are planted on ridges in the field approximately 1.20 m apart. In 2019, healthy plant growth was observed over the entire growth season from March to November (Figure 4i). The CFT includes leaf level measurements of photosynthetic performance and UAV-based measurements at regular intervals using a structure from motion approach to quantify the three-dimensional growth dynamics during the growing season (Figure 4j). UAV measurements are used to reliably determine plant growth rates, plant height, and canopy volume, giving insight into the seasonal dynamics of early vigor and plant establishment, biomass gain during main growth period, and other variety specific properties such as leaf shedding and regrowth that may occur to the roots during starch allocation or as a response to unfavourable environmental conditions. Because shoot and root fresh weight showed a strong correlation ($R = 0.9$) at the final harvest in 2019, UAV-based phenotyping helps to identify well-performing lines prior to the final harvest, which is very useful for planning sampling and further experiments. Additionally, this approach opens the possibility of linking organ specific growth rates and integrating seasonal trait expression to gain a better understanding of the final biomass and root yield at time of harvest.

Wild-type 60 444 plants grew up to 3.5 m in height and produced up to 10 kg in root fresh weight during the 8 months in the field, which is approximately 2–4 months shorter than the typical 60 444 growth season (Figure 4k). CASS concluded the first CFT in 2019 with 89 independent events from different multigene constructs and the 2020 CFT with new multigene constructs is currently underway. In addition to Taiwan, we are also testing transgenic plants at the International Institute for Tropical Agriculture (IITA) in Nigeria. Currently no CMD-resistant cassava cultivar can be routinely transformed with high efficiency, although efforts are underway in CASS to establish high-throughput transformation pipelines with CMD-resistant cultivars. Although TME 7 has a dominant CMD tolerance, this is lost in tissue culture (Beyene *et al.*, 2018). To test our transgenic plants in Nigeria, we therefore constructed screenhouses that protect plants from the whitefly vector that transmits the CMD virus (Figure 4l). Cassava engineered for reduced starch breakdown grew healthily in the screenhouse during a first transgenic trial in 2018. New screenhouse trials are currently underway at IITA with source–sink engineered events (Figure 4m–o). Chosen events were selected for their superior performance in the 2019 CFT at National Chung Hsing University. Together, our cassava source–sink metabolic engineering pipeline is very efficient and productive for feeding the annual CFTs in Africa and Asia with transgenic plants expressing novel gene combinations in multi-stack constructs.

CONCLUDING REMARKS

CASS (https://cass-research.org/team) has built an international research consortium focused on the engineering of cassava source-to-sink relations to increase storage root and starch yield. The network covers all steps from molecular cloning to field testing. Systems biology approaches are used to gain new and detailed insights into cassava biology, which is a prerequisite for knowledge-driven engineering and construction of genome-scale metabolic models at the whole plant level. Biotechnology interventions to remove metabolic bottlenecks have involved multi-stack constructs for modifying the expression of a portfolio of target genes that are tested in transgenic plants in the field. As genome editing methods become routine in cassava, this will open up other avenues of research; for

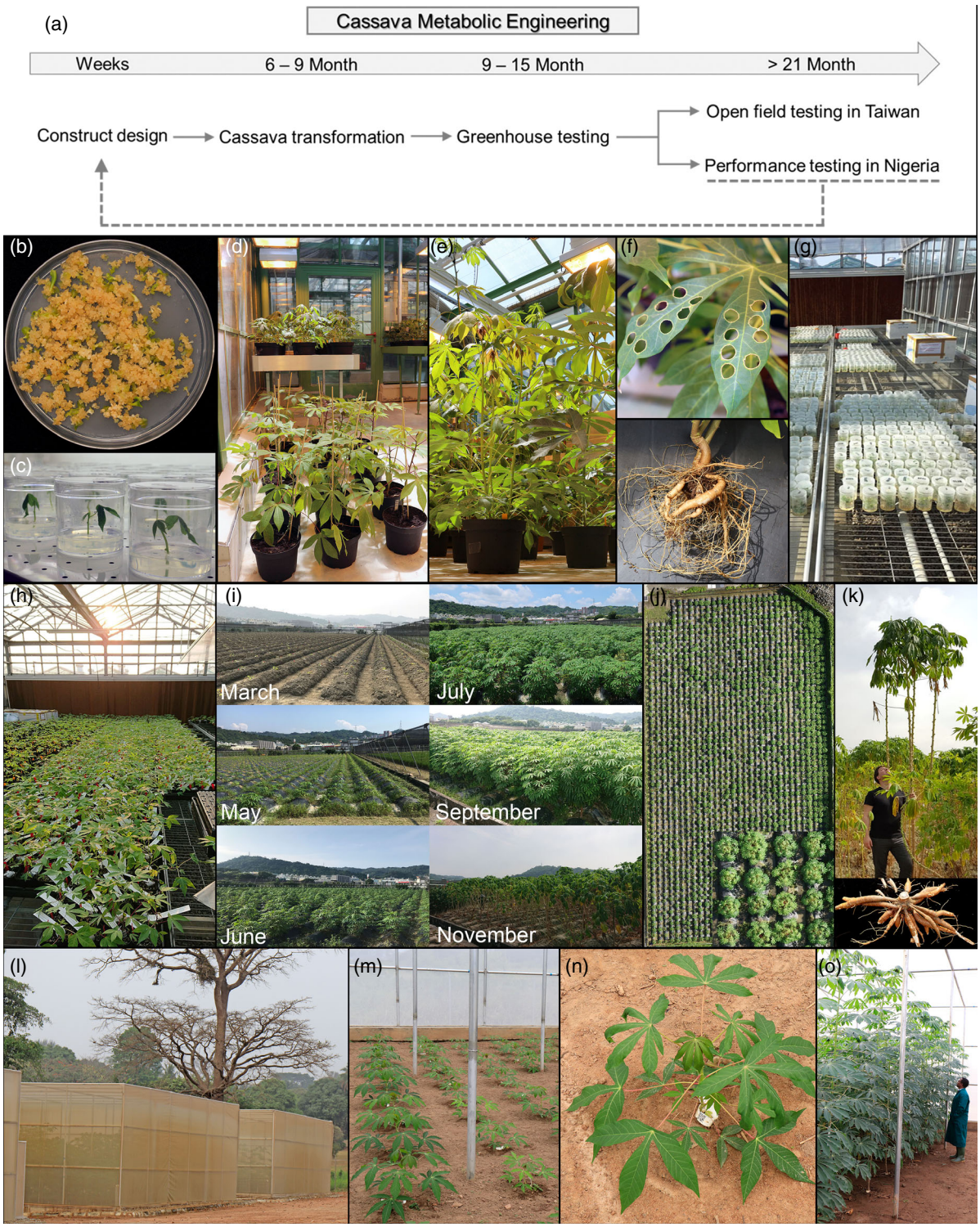

Figure 4. Overview on cassava source–sink metabolic engineering. (a) Timeline and process overview of cassava source-sink metabolic engineering. (b) Emerging cassava shoots after transformation of friable embryonic calli. (c) Cassava plantlets in tissue culture. (d,e) Potted cassava plants growing in a greenhouse in Germany. (f) Leaf and root sampling of greenhouse-grown plants. (g,h) Cassava plantlets in tissue culture jars are transferred into soil in pots equipped with unique labels and grown for a confined field trial. (i) Confined field trial at the National Chung Hsing University (NCHU) Experimental Station in Taiwan from March until November 2019. (j) Top-view of the confined field trial at NCHU Experimental Station via unmanned aerial vehicle-based areal phenotyping. A magnified image of individual plants is shown at the bottom right. (k) Example of shoot and root growth from transgenic plants (60444 background). (l) Screenhouses on a confined field trial site at the International Institute for Tropical Agriculture. (m) Plant growth in screenhouses shortly after planting in October 2019. (n) Single cassava plant shortly after planting. (o) Plant growth in screenhouses 6 months after planting.

The Plant Journal, (2020), doi: 10.1111/tpj.14865

example, *de novo* domestication (Fernie and Yan, 2019). The wild relative of cassava, *Manihot glaziobii*, has already been proposed as a target species for such an approach because it produces larger roots (Zsogon *et al*., 2017). However, even with currently available biotechnological tools and non-invasive field phenotyping, we are optimistic that CASS will make considerable progress in increasing cassava storage root and starch yield.

CASS clearly has a long way to go with respect to bringing locally-adapted, yield-improved and resilient cassava cultivars to the market that can contribute to agronomic transformation in Africa. However, we note that interest in using cassava is increasing, both as an industrial crop and staple food resource. We are convinced that this will encourage more laboratories to participate in collaborative cassava research. As the African saying goes: 'If you want to go fast, go alone. If you want to go far, go together'.

ACKNOWLEDGEMENTS

The Cassava source–sink project (CASS) is supported by grants from the Bill and Melinda Gates Foundation (OPP1113365 and OPP1207956). We thank all CASS members for their various contributions to the success of the project.

AUTHOR CONTRIBUTIONS

All authors contributed to writing the manuscript and designing the figures.

CONFLICT OF INTEREST

The authors declare that they have no competing interests. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Arrivault, S., Alexandre Moraes, T., Obata, T. *et al*. (2019) Metabolite profiles reveal interspecific variation in operation of the Calvin-Benson cycle in both C4 and C3 plants. *J. Exp. Bot.* **70**, 1843–1858.

Baroja-Fernandez, E., Munoz, F.J., Montero, M. *et al*. (2009) Enhancing sucrose synthase activity in transgenic potato (*Solanum tuberosum L.*) tubers results in increased levels of starch, ADPglucose and UDPglucose and total yield. *Plant Cell Physiol.* **50**, 1651–1662.

Beyene, G., Solomon, F.R., Chauhan, R.D. *et al*. (2018) Provitamin A biofortification of cassava enhances shelf life but reduces dry matter content of storage roots due to altered carbon partitioning into starch. *Plant Biotechnol. J.* **16**, 1186–1200.

Bredeson, J.V., Lyons, J.B., Prochnik, S.E. *et al*. (2016) Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat. Biotechnol.* **34**, 562–570.

Bull, S.E., Owiti, J.A., Niklaus, M., Beeching, J.R., Gruissem, W. and Vanderschuren, H. (2009) Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. *Nat. Protoc.* **4**, 1845–1854.

Campilho, A., Nieminen, K. and Ragni, L. (2020) The development of the periderm: the final frontier between a plant and its environment. *Curr. Opin. Plant Biol.* **53**, 10–14.

Dalal, J., Lopez, H., Vasani, N.B. *et al*. (2015) A photorespiratory bypass increases plant growth and seed yield in biofuel crop *Camelina sativa*. *Biotechnol. Biofuels*, **8**, 175.

De Souza, A.P. and Long, S.P. (2018) Toward improving photosynthesis in cassava: characterizing photosynthetic limitations in four current African cultivars. *Food Energy Secur.* **7**, e00130.

De Souza, A.P., Wang, Y., Orr, D.J., Carmo-Silva, E. and Long, S.P. (2019) Photosynthesis across African cassava germplasm is limited by Rubisco and mesophyll conductance at steady state, but by stomatal conductance in fluctuating light. *New Phytol.* **225**, 2498–2512.

Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S. (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One*, **4**, e5553.

Engler, C., Kandzia, R. and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, **3**, e3647.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D., Patron, N.J. and Marillonnet, S. (2014) A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* **3**, 839–843.

FAO (2013) Save and grow: cassava. A guide to sustainable production intensification. *Food and Agriculture Organization of the United Nations*, 1–129.

FAOSTAT (2017) *Food and Agriculture Organization of the United Nations*. http://www.fao.org/faostat/en/#data/QC

FAOSTAT (2019) *Food and Agriculture Organization of the United Nations*. http://www.fao.org/faostat/en/#data/QC

Fernie, A.R., Bachem, C.W.B., Helariutta, Y. *et al*. (2020) Synchronization of developmental, molecular and metabolic aspects of source-sink interactions. *Nat. Plants*, **6**, 55–66.

Fernie, A.R. and Gutierrez-Marcos, J. (2019) From genome to phenome: genome-wide association studies and other approaches that bridge the genotype to phenotype gap. *Plant J.* **97**, 5–7.

Fernie, A.R. and Yan, J. (2019) De novo domestication: an alternative route toward new crops for the future. *Mol. Plant*, **12**, 615–631.

Gajdanowicz, P., Michard, E., Sandmann, M. *et al*. (2011) Potassium (K^+) gradients serve as a mobile energy source in plant vascular tissues. *Proc. Natl Acad. Sci. USA*, **108**, 864–869.

Gemenet, D.C., da Silva Pereira, G., De Boeck, B. *et al*. (2020) Quantitative trait loci and differential gene expression analyses reveal the genetic basis for negatively associated beta-carotene and starch content in hexaploid sweetpotato [*Ipomoea batatas (L.) Lam.*]. *Theor. Appl. Genet.* **133**, 23–36.

Ihemere, U., Arias-Garzon, D., Lawrence, S. and Sayre, R. (2006) Genetic modification of cassava for enhanced starch production. *Plant Biotechnol. J.* **4**, 453–465.

Jonik, C., Sonnewald, U., Hajirezaei, M.R., Flugge, U.I. and Ludewig, F. (2012) Simultaneous boosting of source and sink capacities doubles tuber starch yield of potato plants. *Plant Biotechnol. J.* **10**, 1088–1098.

Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.J., Rosenkranz, R., Stabler, N., Schonfeld, B., Kreuzaler, F. and Peterhansel, C. (2007) Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. *Nat. Biotechnol.* **25**, 593–599.

Koch, K.E. (1996) Carbohydrate-modulated gene expression in plants. *Annu. Rev. Plant Physiol.* **47**, 509–540.

Kromdijk, J., Glowacka, K., Leonelli, L., Gabilly, S.T., Iwai, M., Niyogi, K.K. and Long, S.P. (2016) Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science*, **354**, 857–861.

Kuon, J.E., Qi, W., Schlapfer, P. *et al*. (2019) Haplotype-resolved genomes of geminivirus-resistant and geminivirus-susceptible African cassava cultivars. *BMC Biol.* **17**, 75.

McCallum, E.J., Anjanappa, R.B. and Gruissem, W. (2017) Tackling agriculturally relevant diseases in the staple crop cassava (*Manihot esculenta*). *Curr. Opin. Plant Biol.* **38**, 50–58.

Mehdi, R., Lamm, C.E., Ravi Bodampalli, A. *et al*. (2019) Symplasmic phloem unloading and radial post-phloem transport via vascular rays in tuberous roots of *Manihot esculenta*. *J. Exp. Bot.* **70**(20), 5559–5573.

Nölke, G., Houdelet, M., Kreuzaler, F., Peterhansel, C. and Schillberg, S. (2014) The expression of a recombinant glycolate dehydrogenase polyprotein in potato (*Solanum tuberosum*) plastids strongly enhances photosynthesis and tuber yield. *Plant Biotechnol. J.* **12**, 734–742.

Obata, T., Klemens, P.A.W., Rosado-Souza, L. *et al*. (2020) Metabolic profiles of six African cultivars of cassava (*Manihot esculenta Crantz*) highlight bottlenecks of root yield. *Plant J.* https://doi.org/10.1111/tpj.14693.

Otekunrin, O. and Sawicka, B. (2019) Cassava, a 21st century staple crop: how can Nigeria harness its enormous trade potentials? *Acta Sci. Agric.* **3**, 194–202.

Paul, M.J. and Foyer, C.H. (2001) Sink regulation of photosynthesis. *J. Exp. Bot.* **52**, 1383–1400.

Perez-Fons, L., Bohorquez-Chaux, A., Irigoyen, M.L., Garceau, D.C., Morreel, K., Boerjan, W., Walling, L.L., Becerra Lopez-Lavalle, L.A. and Fraser, P.D. (2019) A metabolomics characterisation of natural variation in the resistance of cassava to whitefly. *BMC Plant Biol.* **19**, 518.

Prochnik, S., Marri, P.R., Desany, B. *et al.* (2012) The cassava genome: current progress, future directions. *Trop. Plant Biol.* **5**, 88–94.

Rabbi, I., Hamblin, M., Gedil, M., Kulakow, P., Ferguson, M., Ikpan, A.S., Ly, D. and Jannink, J.-L. (2014a) Genetic mapping using genotyping-by-sequencing in the clonally propagated cassava. *Crop Sci.* **54**, 1384–1396.

Rabbi, I.Y., Hamblin, M.T., Kumar, P.L., Gedil, M.A., Ikpan, A.S., Jannink, J.L. and Kulakow, P.A. (2014b) High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Res.* **186**, 87–96.

Rabbi, I.Y., Udoh, L.I., Wolfe, M., Parkes, E.Y., Gedil, M.A., Dixon, A., Ramu, P., Jannink, J.L. and Kulakow, P. (2017) Genome-wide association mapping of correlated traits in cassava: dry matter and total carotenoid content. *Plant Genome*, **10**(3). https://doi.org/10.3835/plantgenome2016.09.0094

Rosado-Souza, L., David, L.C., Drapal, M. *et al.* (2019) Cassava metabolomics and starch quality. *Curr. Protoc. Plant Biol.* **4**, e20102.

Schwender, J., Goffman, F., Ohlrogge, J.B. and Shachar-Hill, Y. (2004) Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature*, **432**, 779–782.

Sonnewald, U. and Fernie, A.R. (2018) Next-generation strategies for understanding and influencing source-sink relations in crop plants. *Curr. Opin. Plant Biol.* **43**, 63–70.

South, P.F., Cavanagh, A.P., Liu, H.W. and Ort, D.R. (2019) Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science*, **363**(6422), eaat9077.

Tecle, I.Y., Edwards, J.D., Menda, N., Egesi, C., Rabbi, I.Y., Kulakow, P., Kawuki, R., Jannink, J.L. and Mueller, L.A. (2014) solGS: a web-based tool for genomic selection. *BMC Bioinformatics*, **15**, 398.

United Nations, D.o.E.a.S.A., Population Division (2019) World Population Prospects 2019. *Volume II: Demographic Profiles* (ST/ESA/SER.A/427).

Van Doorn, A.M., Muller, O., Gisel, A., Kulakow, P., Rabbi, I., Siegmann, B., Wilke, N. and Rascher, U. (2020) Quantifiying canopy structure of four cassava varieties by structure from motion. *Agriculture* (submitted).

Wang, X., Chang, L., Tong, Z. *et al.* (2016) Proteomics profiling reveals carbohydrate metabolic enzymes and 14-3-3 proteins play important roles in starch accumulation during cassava root tuberization. *Sci. Rep.* **6**, 19643.

Wingenter, K., Schulz, A., Wormit, A., Wic, S., Trentmann, O., Hoermiller, I.I., Heyer, A.G., Marten, I., Hedrich, R. and Neuhaus, H.E. (2010) Increased activity of the vacuolar monosaccharide transporter TMT1 alters cellular sugar partitioning, sugar signaling, and seed yield in Arabidopsis. *Plant Physiol.* **154**, 665–677.

Yang, J., An, D. and Zhang, P. (2011) Expression profiling of cassava storage roots reveals an active process of glycolysis/gluconeogenesis. *J. Integr. Plant Biol.* **53**, 193–211.

Zainuddin, I.M., Fathoni, A., Sudarmonowati, E., Beeching, J.R., Gruissem, W. and Vanderschuren, H. (2018) Cassava post-harvest physiological deterioration: from triggers to symptoms. *Postharvest. Biol. Technol.* **142**, 115–123.

Zhang, L., Hausler, R.E., Greiten, C., Hajirezaei, M.R., Haferkamp, I., Neuhaus, H.E., Flugge, U.I. and Ludewig, F. (2008) Overriding the co-limiting import of carbon and energy into tuber amyloplasts increases the starch content and yield of transgenic potato plants. *Plant Biotechnol. J.* **6**, 453–464.

Zhao, K., Tung, C.W., Eizenga, G.C. *et al.* (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2**, 467.

Zsogon, A., Cermak, T., Voytas, D. and Peres, L.E. (2017) Genome editing as a tool to achieve the crop ideotype and de novo domestication of wild relatives: case study in tomato. *Plant Sci.* **256**, 120–130.