the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2023) **116,** 38–57 doi: 10.1111/tpj.16357

Efficient sugar utilization and transition from oxidative to substrate-level phosphorylation in high starch storage roots of African cassava genotypes

Christian E. Lamm[1,*], Ismail Y. Rabbi[2], David Barbosa Medeiros[3], Laise Rosado-Souza[3], Benjamin Pommerrenig[4], Ismail Dahmani[3], David Rüscher[1], Jörg Hofmann[1], Anna M. van Doorn[2], Armin Schlereth[3], H. Ekkehard Neuhaus[4], Alisdair R. Fernie[3], Uwe Sonnewald[1] and Wolfgang Zierer[1]

[1]*Friedrich-Alexander-Universität Erlangen-Nürnberg, Division of Biochemistry, Erlangen, Germany,*
[2]*International Institute of Tropical Agriculture, Ibadan, Nigeria,*
[3]*Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm, Germany, and*
[4]*University of Kaiserslautern, Plant Physiology, Kaiserslautern, Germany*

Received 20 January 2023; revised 19 May 2023; accepted 14 June 2023; published online 17 June 2023.
*For correspondence (e-mail christian.lamm@fau.de).

SUMMARY

Cassava's storage roots represent one of the most important sources of nutritional carbohydrates worldwide. Particularly, smallholder farmers in sub-Saharan Africa depend on this crop plant, where resilient and yield-improved varieties are of vital importance to support steadily increasing populations. Aided by a growing understanding of the plant's metabolism and physiology, targeted improvement concepts already led to visible gains in recent years. To expand our knowledge and to contribute to these successes, we investigated storage roots of eight cassava genotypes with differential dry matter content from three successive field trials for their proteomic and metabolic profiles. At large, the metabolic focus in storage roots transitioned from cellular growth processes toward carbohydrate and nitrogen storage with increasing dry matter content. This is reflected in higher abundance of proteins related to nucleotide synthesis, protein turnover, and vacuolar energization in low starch genotypes, while proteins involved in sugar conversion and glycolysis were more prevalent in high dry matter genotypes. This shift in metabolic orientation was underlined by a clear transition from oxidative- to substrate-level phosphorylation in high dry matter genotypes. Our analyses highlight metabolic patterns that are consistently and quantitatively associated with high dry matter accumulation in cassava storage roots, providing fundamental understanding of cassava's metabolism as well as a data resource for targeted genetic improvement.

Keywords: ***Manihot esculenta*, source/sink, starch metabolism, sugar metabolism, energy metabolism, glycolysis, oxidative phosphorylation, TCA-cycle, proteomics, metabolomics.**

INTRODUCTION

The perennial shrub cassava (*Manihot esculenta*) with its edible, starch-rich storage roots exhibits many beneficial qualities, making it a valued food security crop in tropical and subtropical regions all over the world. Its robustness and tolerance to abiotic stresses, flexible harvest time, and the ability to propagate it by stem cuttings explain its relevance especially to smallholder farmers in sub-Saharan Africa (Howeler et al., 2013). However, compared to other production areas and despite increasing populations, overall cassava yield per hectare in Africa remained low throughout the past 60 years: While yield increased by over 170% from 1961 to 2020 in Asia, an increase of only approximately 50% could be realized in Africa (FAO, 2022). Although reasons for this tremendous difference are manifold and include structural and socio-economic obstacles such as limited access to improved varieties, irrigation, fertilizer, and modern agricultural practices and devices, it illustrates the enormous potential of cassava as a staple crop (Hershey et al., 2012; Kawano & Cock, 2005). In addition to structural improvements, this potential may be unlocked by conventional breeding or genetic engineering of the plant, and both approaches are being pursued with increasing efforts. While breeding programs will doubtlessly lead to the generation of cassava varieties with improved traits, this process suffers from several limitations, such as long timelines, the

plant's heterozygous nature or difficult seed production, to name only a few (Ceballos et al., 2004). When it comes to biotechnological improvements, a solid foundation has been provided by the development of efficient genetic transformation protocols (Bull et al., 2009; Taylor et al., 2012). Furthermore, well-described genetic elements such as promoters and contemporary genome editing methods enrich the cassava biotechnology toolkit (Bull et al., 2018; Gomez et al., 2019; Zierer et al., 2022). In line with this, work targeted to biofortification and virus resistance already showed very promising results (Beyene et al., 2018; Gomez et al., 2019; Li et al., 2015; Narayanan et al., 2021; Sayre et al., 2011). For other fields of research, our still very limited knowledge of cassava's physiology, biochemistry, and metabolism is hindering development of concepts leading to successful improvements: Although basic mechanisms of photoassimilate allocation and interconversion to starch and the storage root proteome during bulking or in light of postharvest deterioration were investigated previously (Mehdi et al., 2019; Owiti et al., 2011; Vanderschuren et al., 2014; Wang et al., 2016), many metabolic and/or physiological determinants of starch accumulation still remain unclear. To generate a deeper understanding of metabolic processes that promote or restrict starch accumulation in cassava storage roots and therefore influence yield directly, we analyzed a set of eight contrasting cassava genotypes. These were chosen from a larger breeding population for their similarities in root weight and low carotenoid content, but greatly differing dry matter content, a trait that is indicative of starch content. The selected genotypes were field-grown in three independent trials in Nigeria and investigated by metabolomics and proteomics experiments. This setup enabled us to pinpoint metabolic alterations in starch biosynthetic- and glycolytic pathways that are associated with higher starch levels in storage roots. Likewise, we demonstrate that increased focus on the pentose phosphate pathway, oxidative phosphorylation, vacuolar energization as well as enhanced protein turnover serve as hallmarks of low dry matter cassava genotypes. Our results reveal metabolic transitions associated with starch storage on one hand and cellular growth/proliferation on the other hand. This provides leverage points for future improvement of cassava as a staple crop, either by breeding efforts or by biotechnological approaches.

RESULTS

Selection of *Manihot esculenta* breeding lines for comparative proteome and metabolome analyses

A breeding population consisting of 52 cassava genotypes with differential traits was grown in repeated field trials over three growing periods of approximately 1 year in Ibadan or Ikenne, Nigeria. After harvesting in 2019, 2020, and 2021, agronomic data was recorded, and sample material was collected for subsequent molecular and biochemical analyses.

To pinpoint changes in the metabolic and proteomic composition of cassava storage roots that lead to an altered accumulation of starch, a subset of genotypes was chosen from the breeding population.

To avoid complications in data interpretation from trait–trait interactions, we focused on the dry matter content (DM%). This trait can be used as proxy for starch content in cassava storage roots, as both correlate strongly (Figure 3a). As high carotenoid content is known to negatively affect starch accumulation (Beyene et al., 2018), we selected genotypes with similar total storage root yield per plot (RTWT), but low carotenoid content based on *k*-means clustering. Hence, the genotypes were categorized as 'yellow' or 'white' depending on their root color, while we identified three *k*-means clusters for RTWT for all investigated genotypes (Figure 1a). The cluster referring to medium root fresh weights contained 11 genotypes that were characterized by white root color, but differential dry matter content ranging from 21 to 41 percent of the initial fresh weight (Figure 1b). From these 11 genotypes, a final subset was chosen for metabolome and proteome analyses: For the harvest year 2019, eight genotypes were selected, while for 2020 and 2021 only seven of those genotypes were available for analyses (Table 1).

To exclude the presence of unwanted trait–trait interactions, agronomic data was plotted against DM%, and linear regressions were calculated. Since we selected for similar RTWT deliberately, this trait did not correlate with dry matter content (Figure 1c). Furthermore, the selected genotypes showed no consistent correlation between dry matter content and storage root number per plot (Figure 1d). Using this data, average fresh weight per storage root was calculated, which again did not correlate with dry matter content (Figure 1e). Finally, no correlation could be found between shoot fresh weight per plot and DM% (Figure 1f). Hence, the selected genotypes display a wide range of dry

Figure 1. Selection of genotypes based on *k*-means clustering and dry matter content distribution. (a) Storage root fresh weight (FW) best linear unbiased estimators (BLUEs) of 51 cassava genotypes over three successive field trials. *k*-means clustering was used to categorize the genotypes according to storage root color (Y, yellow; W, white) and storage root FW BLUEs (low, medium, and high root weight [RTWT]). (b) Distribution of dry matter content BLUEs within 'white' and 'medium RTWT' clusters. Genotype accessions are indicated on the X-axis. Error bars indicate the standard error of BLUEs. (c–f) In either field trial, no consistent correlation was found between dry matter content (DM%) and storage root fresh weight per plot (RTWT, c), root number per plot (d), average fresh weight per root (e), and shoot fresh weight per plot (f). R^2-values of linear regression are depicted for each harvest year, asterisks indicate significance ($P \leq 0.05$).

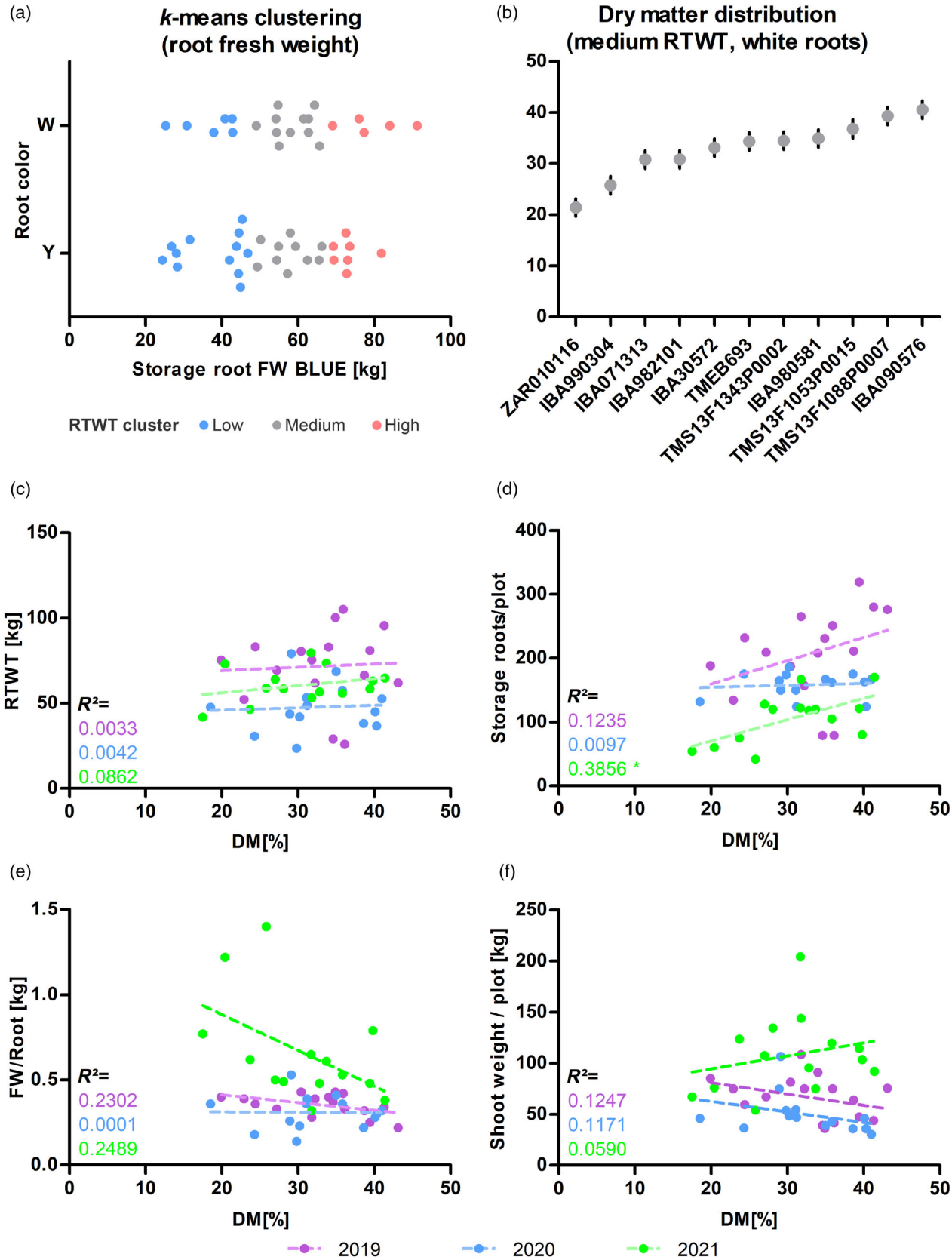

Table 1 Genotypes investigated in this study. The table indicates the respective storage root's dry matter content in percent for each plot per harvest years 2019, 2020, and 2021, as well as the overall average ± standard deviation

	Dry matter content 2019		*Dry matter content 2020*		*Dry matter content 2021*		
Genotype	Plot A	Plot B	Plot A	Plot B	Plot A	Plot B	Average dry matter content ± standard deviation
IITA-TMS-ZAR010116	24.4	19.9	24.3	18.5	23.7	17.5	21.4 ± 3.1
IITA-TMS-IBA990304	27.2	22.9	28.9	29.1	25.8	20.4	25.7 ± 3.5
IITA-TMS-IBA071313	30.4	34	30.2	31.1	31.8	27	30.8 ± 2.3
IITA-TMS-IBA982101	31.8	32.2	29.8	31.2	31.7	28.1	30.8 ± 1.6
TMS13F1343P0002	34.9	34.6	35.8	35	32.8	33.7	34.5 ± 1.1
IITA-TMS-IBA980581	35.9	36.1	Unavailable	Unavailable	Unavailable	Unavailable	36 ± 0.1
TMS13F1088P0007	38.7	43.1	38.6	40.1	39.4	35.8	39.3 ± 2.4
IITA-TMS-IBA090576	39.4	41.3	41	40.3	41.4	39.8	40.5 ± 0.8

matter contents, which is not associated with differences in the other traits.

Metabolic profiling reveals metabolites correlating with dry matter content

Individual metabolome analyses were carried out for each of the three field experiments. Storage root sample material collected from two plots per genotype was subjected to individual metabolite profiling by GC–MS for each of the three field experiments. In addition, levels of sucrose (Suc), glucose (Glc), and fructose (Frc) were determined by an NADP-coupled optical test, while starch content was determined by wet milling. In total, the contents of 53 metabolites were determined. To identify metabolic processes associated with the accumulation of dry matter, Pearson's correlation coefficients of metabolite levels and dry matter contents were calculated (Figure 2, Table S1). As expected, starch content was confirmed as the best positive correlator to dry matter content in all harvest years with an average correlation coefficient of 0.97, underlining the possibility to use dry matter content as proxy (Figures 2 and 3a). Apart from starch, the raffinose-family

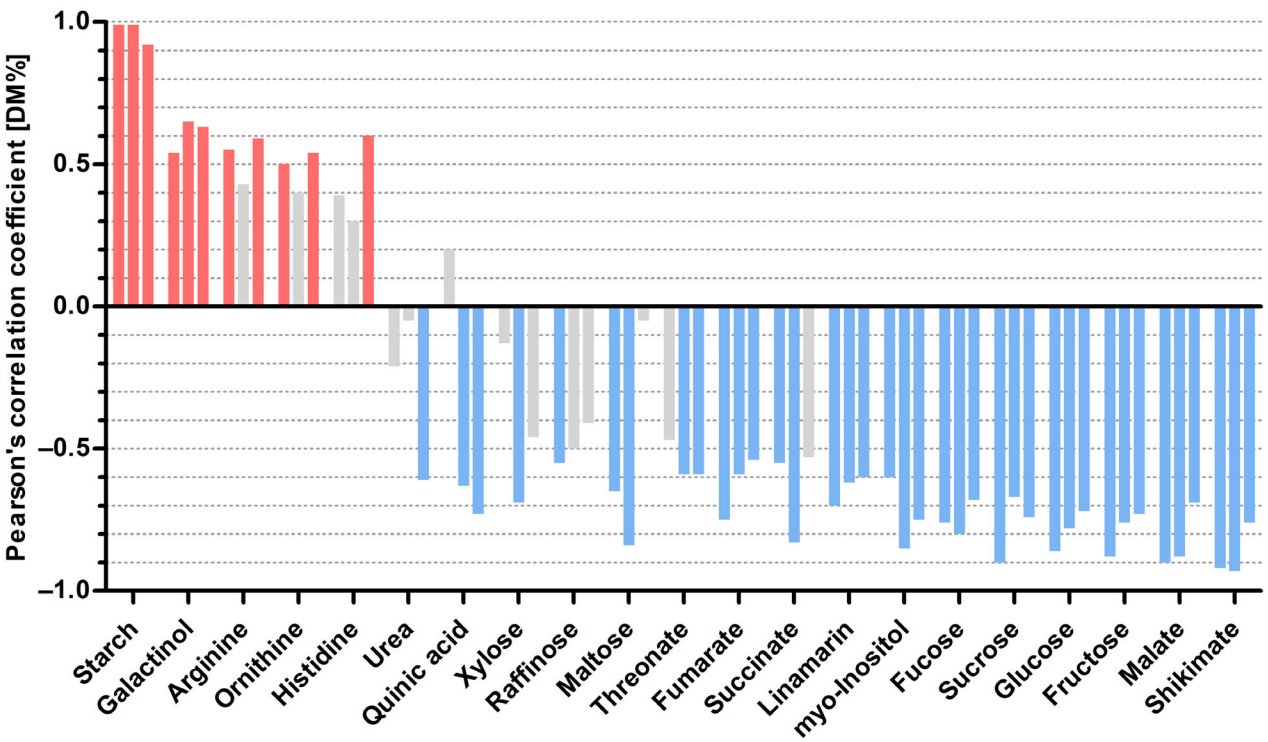

Figure 2. Metabolites show a significant correlation to dry matter content in at least one harvest year. Bars indicate Pearson's correlation coefficient for each harvest year per metabolite, and red and blue color indicates positive and negative correlation, respectively. Gray bars indicate a non-significant correlation.

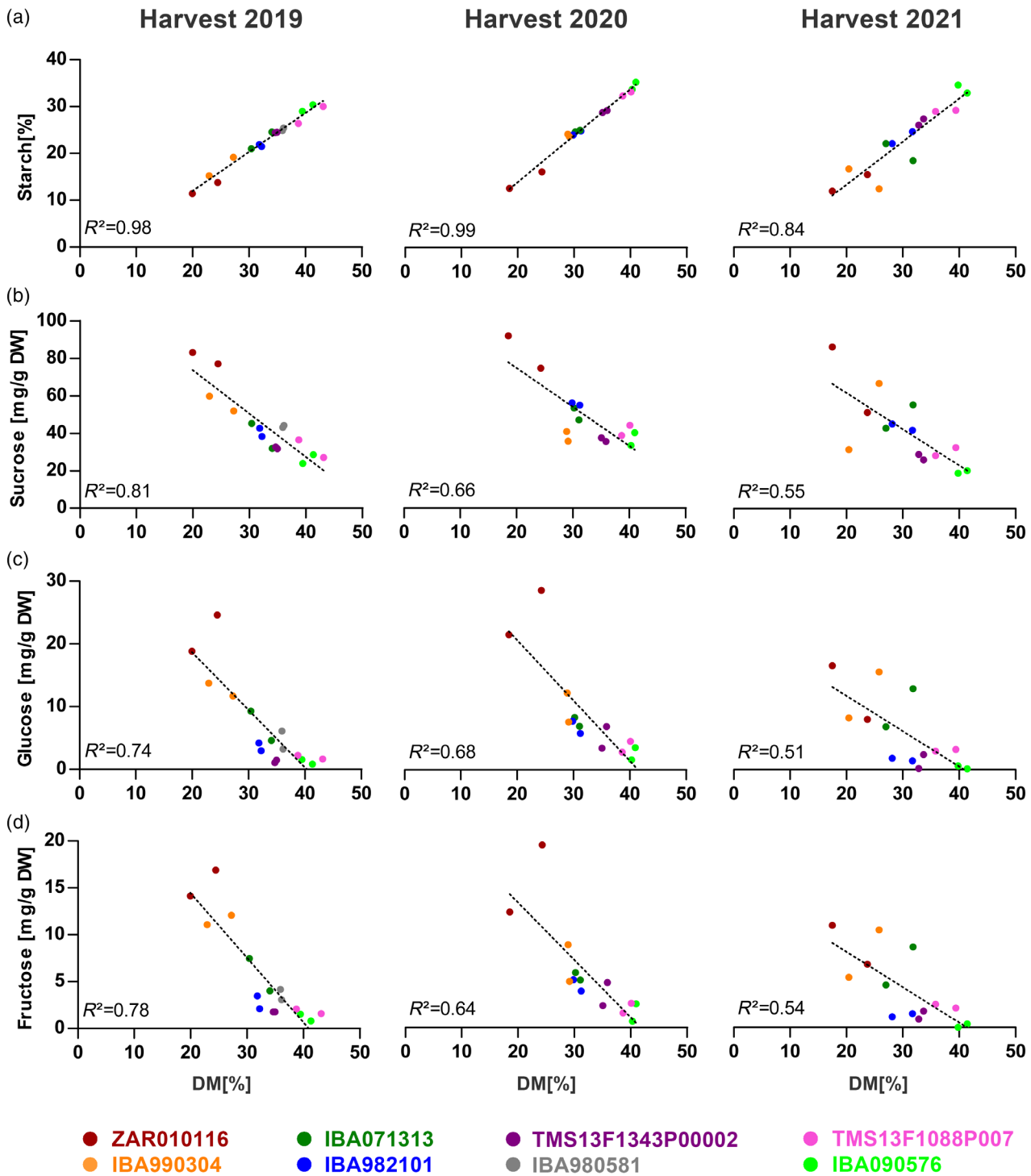

Figure 3. Correlation of dry matter content and starch/sugar levels. Dry matter contents were correlated with starch (a), sucrose (b), glucose (c), and fructose (d) for each harvest year. Colors indicate replicate plots of each genotype as defined below the graphs.

oligosaccharide (RFO) building block galactinol could be identified as only other metabolite with a positive correlation coefficient in all three harvest years. The amino acids arginine (Arg) and ornithine (Orn) were found to correlate positively with dry matter content in 2019s and 2021s samples, while this was not the case for the samples harvested in 2020. Furthermore, Histidine correlated positively with dry matter content in samples of 2021 only.

For metabolites correlating negatively with dry matter content, the monosaccharides Glc and Frc as well as fucose could be found in all 3 years. Together with the fact that the main transport sugar Suc was consistently found as negative correlator as well, this indicates that the genotypes that tend to accumulate lower dry matter levels do not suffer from starvation of carbon building blocks (Figures 2 and 3b–d).

Shikimate, an intermediate metabolite toward the synthesis of aromatic amino acids and other aromatic metabolites was found in all years to yield the strongest negative correlation. Likewise, the cyanogenic glucoside linamarin was consistently found among the negative correlators, as was *myo*-inositol. While the former serves as a defense compound against herbivores, the latter is, among other functions, associated with the biosynthesis of RFOs like raffinose, which was also found as negative correlator in samples from harvest year 2019.

In addition, the tricarboxylic acid (TCA) cycle intermediates fumarate (Fum) and malate (Mal) were consistently identified in all harvest years as compounds showing negative correlation coefficients with dry matter content. With succinate (Succ), another TCA cycle compound could be shown to correlate negatively with dry matter content in at least two of the field trials.

Furthermore, the ascorbate degradation product threonate (Green & Fry, 2005), the secondary metabolite quinic acid, and the starch degradation product maltose (Malt) could be found as compounds with negative correlation coefficients with dry matter content in two out of three harvest years.

Finally, the cell wall sugar xylose and the Arg degradation product urea were found to correlate negatively with dry matter content in a single harvest year.

Since the nitrogen-rich amino acids Arg, Orn, and His correlated positively in at least one of the harvest years and urea, potentially derived from Arg catabolism, was found among the negatively correlating metabolites in one experiment, an elemental analysis was conducted to investigate a possible relationship between total nitrogen content and dry matter content in the storage roots. However, no correlation could be found, indicating that accumulation of Arg, Orn, and His in high dry matter genotypes does not originate from an overall increase in root nitrogen content due to differences in nitrogen availability or uptake (Figure S1). Similarly, total protein content and dry matter content showed no correlation, either (Figure S1).

Comparative proteomics of eight cassava breeding genotypes from three individual field trials

To add an additional layer of information to the existing metabolome data and gain insight into the metabolism of the cassava genotypes, we conducted three individual comparative proteomics experiments, each referring to one harvest year.

After evaluating and filtering the data, 2958, 3428, and 3298 unique protein groups remained for relative quantitation for harvest years 2019, 2020, and 2021, respectively. Of these, 2124 protein groups could be identified reproducibly in each of the three experiments (Figure 4a).

Principle component analyses revealed proper clustering of both field plots of each genotype, while different genotypes were well-distributed across the plot. This proves biological reproducibility and technical soundness in all three harvest years and provides good indications that differences between the genotypes can be resolved by this approach (Figure 4b–d). Furthermore, genotypes tended to distribute along the axis of component 1 according to their dry matter content, indicating that dry matter content is in fact dependent on the protein composition of the respective genotype.

Correlation of protein abundance and dry matter content

To elucidate biochemical pathways or mechanisms that restrict or promote the accumulation of starch in cassava, the relative protein abundances were correlated with the respective dry matter content. To this end, Spearman's ranked correlation coefficient was calculated, and protein abundance was considered to correlate with dry matter content when the significance of the correlation was ≤ 0.05. This yielded 262, 246, and 198 positively correlating protein groups for the harvest years 2019, 2020, and 2021, respectively. Likewise, 703, 1264, and 508 protein groups were found as negative correlators for the three harvest years 2019, 2020, and 2021. To further increase confidence and exclude possible environmental effects of the field trials, only proteins that correlated with dry matter in at least two of the three field experiments were considered for the following evaluation, which resulted in 145 positively and 574 negatively correlating protein groups (Figure 5a,b and Supplemental Table S2 and S3).

KEGG pathway enrichment analysis

Next, the proteome data were annotated with KEGG Orthology terms, and the 145 positively- or the 574 negatively correlating proteins were used for Benjamini–Hochberg corrected Fisher's exact tests (FDR ≤ 0.05). A significant enrichment of the KEGG pathway glycolysis/gluconeogenesis (3.1-fold enriched) could be detected for positive correlators. For negative correlators, the KEGG pathways photosynthesis (6.4-fold), phagosome (2.3-fold), oxidative phosphorylation (OxPhos) (2.3-fold), and amino sugar and nucleotide sugar metabolism (0.2-fold) were significantly enriched/depleted (Figure 5c and Table S4). While the enrichment of pathways like glycolysis, OxPhos, and amino sugar metabolism is indeed related to the storage root carbon metabolism, pathways like phagosome

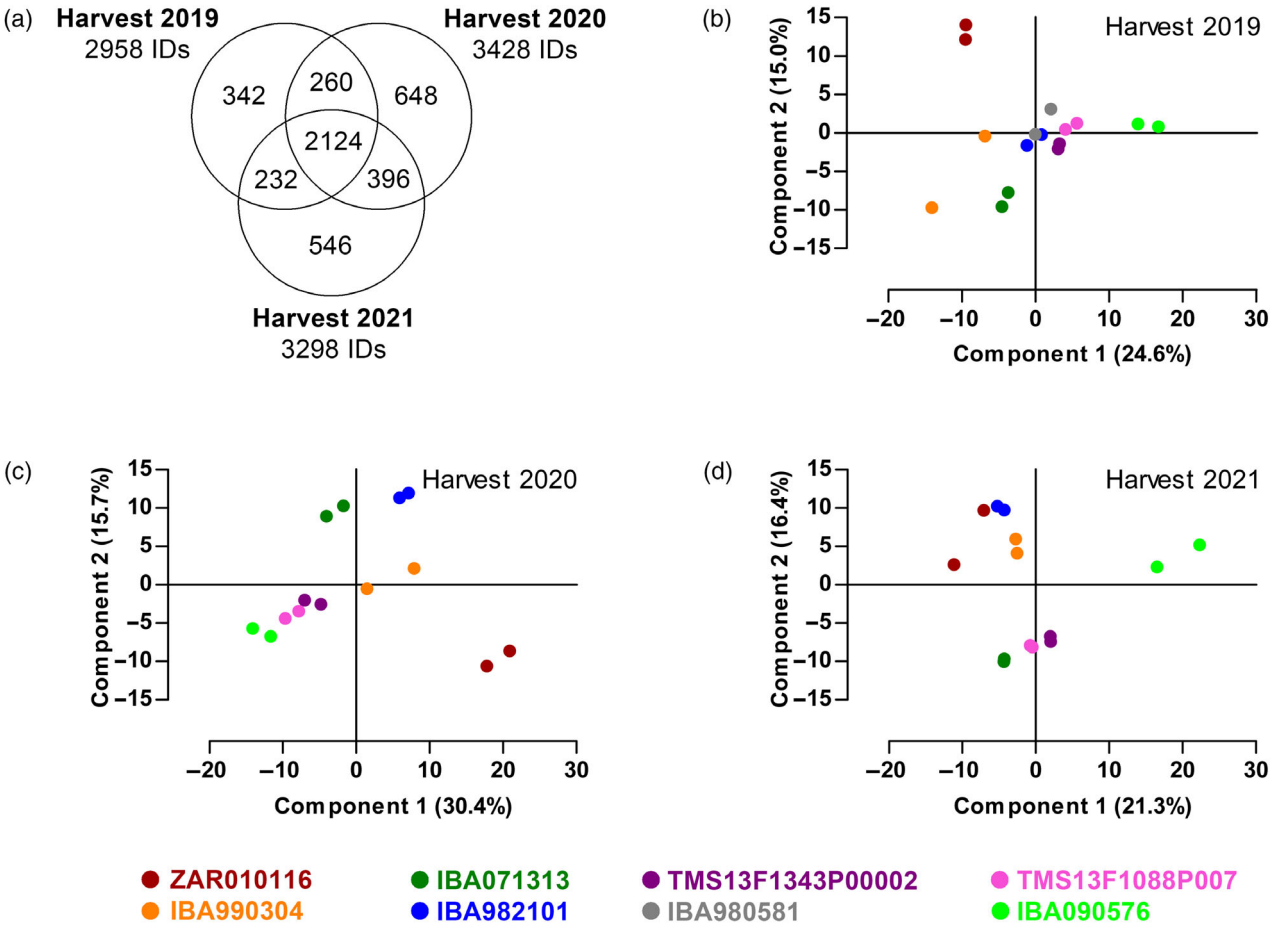

Figure 4. General results of comparative proteomics. The Venn diagram shows number of total protein groups that were identified in each experiment and indicates reproducibly detected protein group numbers (a). Principal component analyses of each of the three experiments for harvest year 2019 (b), 2020 (c), and 2021 (d). Colors indicate replicate plots of each genotype as defined below the graphs.

and photosynthesis appear surprising. However, the proteins that are included in these enriched categories seem to be wrongly annotated: The pathway photosynthesis contains four proteins, three of which are F-type ATPase subunits with predicted subcellular localization in mitochondria, making them more likely components of OxPhos. The fourth protein contained in this KEGG pathway refers to a root ferredoxin: NADP(H) oxidoreductase implicated in nitrite reduction (Hachiya et al., 2016). Within the KEGG pathway phagosome, a total of ten V-type ATPase subunits could be found, as well as one tubulin and five small GTPases that serve as molecular switches for various functions, including vesicular trafficking (Nielsen, 2020).

STRING database analyses

With the aim of structuring the data, a STRING database analysis was performed (Szklarczyk et al., 2021). The STRING database contains information on known protein–protein interactions, as well as indirect functional associations of proteins taken from literature and other publicly available sources. An input protein list is screened for such associations and a graphical network view is generated, in which functional connections are drawn as strings between nodes representing individual proteins. Hence, using the amino acid sequences of the 145 positively or 574 negatively correlating cassava proteins as input, STRING identified the *Arabidopsis thaliana* protein with the highest sequence similarity, before searching the database for known associations. The model organism's database was chosen over a dedicated cassava database to profit from the superior annotation details. When analyzing the proteins that correlated positively with dry matter content, three major clusters of functionally associated proteins could be identified (Figure 6a). One of these clusters predominantly contained small heat-shock proteins/chaperones, while the other two were characterized by ribosomal proteins and enzymes of the glycolytic pathway, corroborating the results of the previous enrichment analysis.

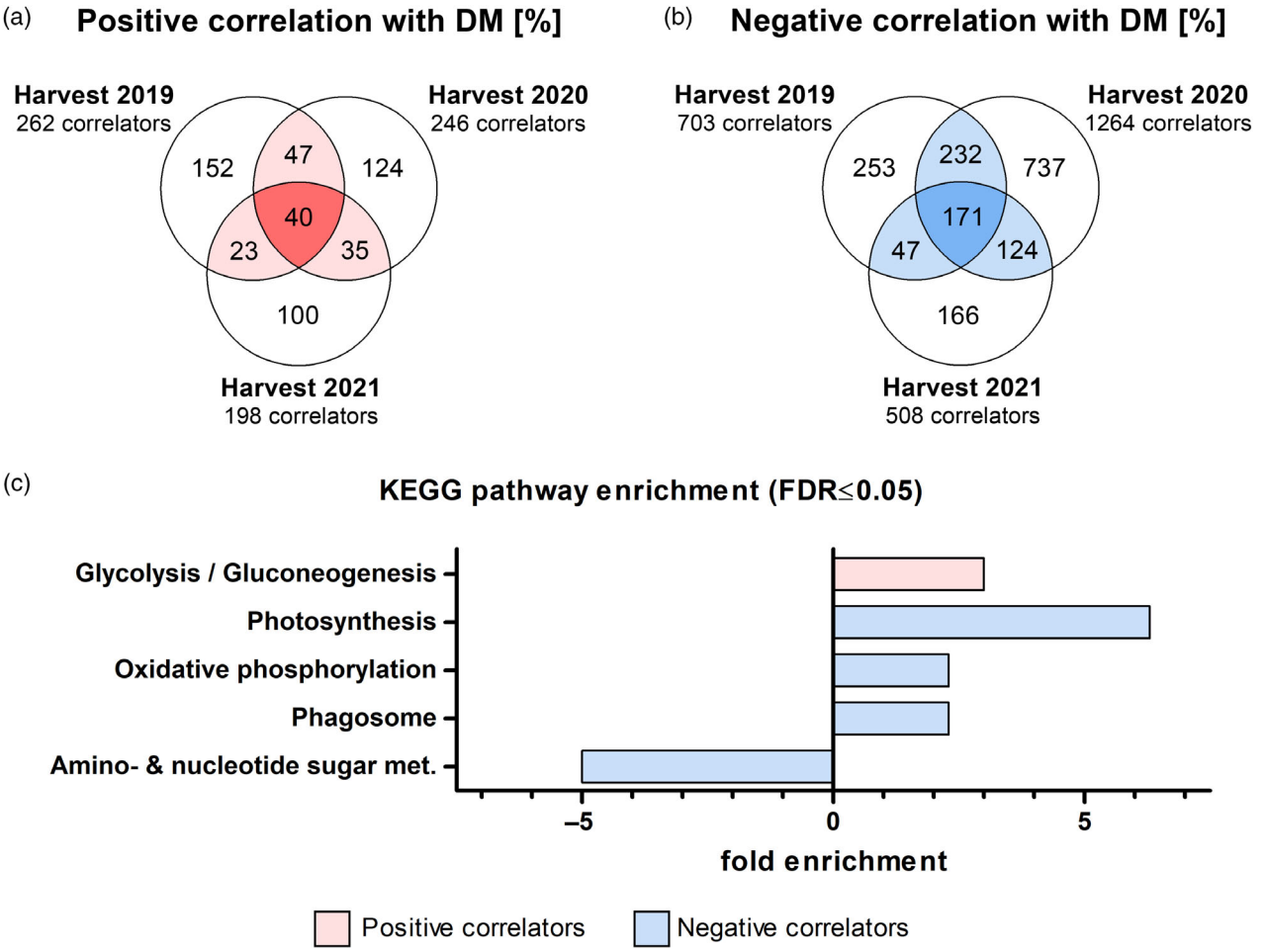

Figure 5. Correlation of protein abundance with dry matter content. The Venn diagrams show the number of positively (a) or negatively (b) correlating protein groups of each proteomics experiment ($P \leq 0.05$). Correlations in at least 2 years were deemed necessary for inclusion in the following analyses. (c) KEGG pathway enrichment analysis of positive and negative correlators. Fisher's exact test was used to determine significantly enriched pathways (FDR≤0.05). Red bars indicate pathways enriched among positively correlating proteins, while blue bars indicate pathways enriched among negative correlators.

When it comes to the negatively correlating proteins, the larger number of input sequences led to a more complex depiction (Figure 5b). Nevertheless, four main complexes could be seen. The largest one referred to proteins associated with RNA processing, ribosomes, or translation initiation, while one minor cluster was characterized by proteins involved in proteasomal degradation, implying increased protein turnover. Another cluster included several V-type ATPase subunits that are involved in vacuolar membrane energization. The last clear-cut cluster featured proteins related to the TCA cycle, mitochondrial electron transport chain, and OxPhos.

Taken together, the STRING database analysis confirmed the findings of the enrichment analysis, indicating that the glycolytic pathway on the one hand, and processes related to the TCA cycle and OxPhos on the other hand may serve as pivoting point for dry matter accumulation. The analysis furthermore turned up other noteworthy processes, like the energization of the vacuolar membrane that seemed to be enhanced in low dry matter genotypes.

Enzymes involved in starch biosynthesis correlate to dry matter content

Since the selected genotypes differed greatly in the accumulation of starch/dry matter content, a closer look at the enzymes involved in sucrose-to-starch conversion might reveal metabolic adaptations promoting increased starch levels.

Indeed, proteins related to starch biosynthesis correlated with the dry matter content of the genotypes: High dry matter genotypes were characterized by an increased abundance of all enzymes involved in the sucrose synthase (SUS)-mediated cytosolic part of the pathway toward glucose-6-phosphate (G6P) except UDP-glucose-pyrophosphorylase (UGPase) and cytosolic fructokinase (FRK), for which no correlation could be found (Figure 7a

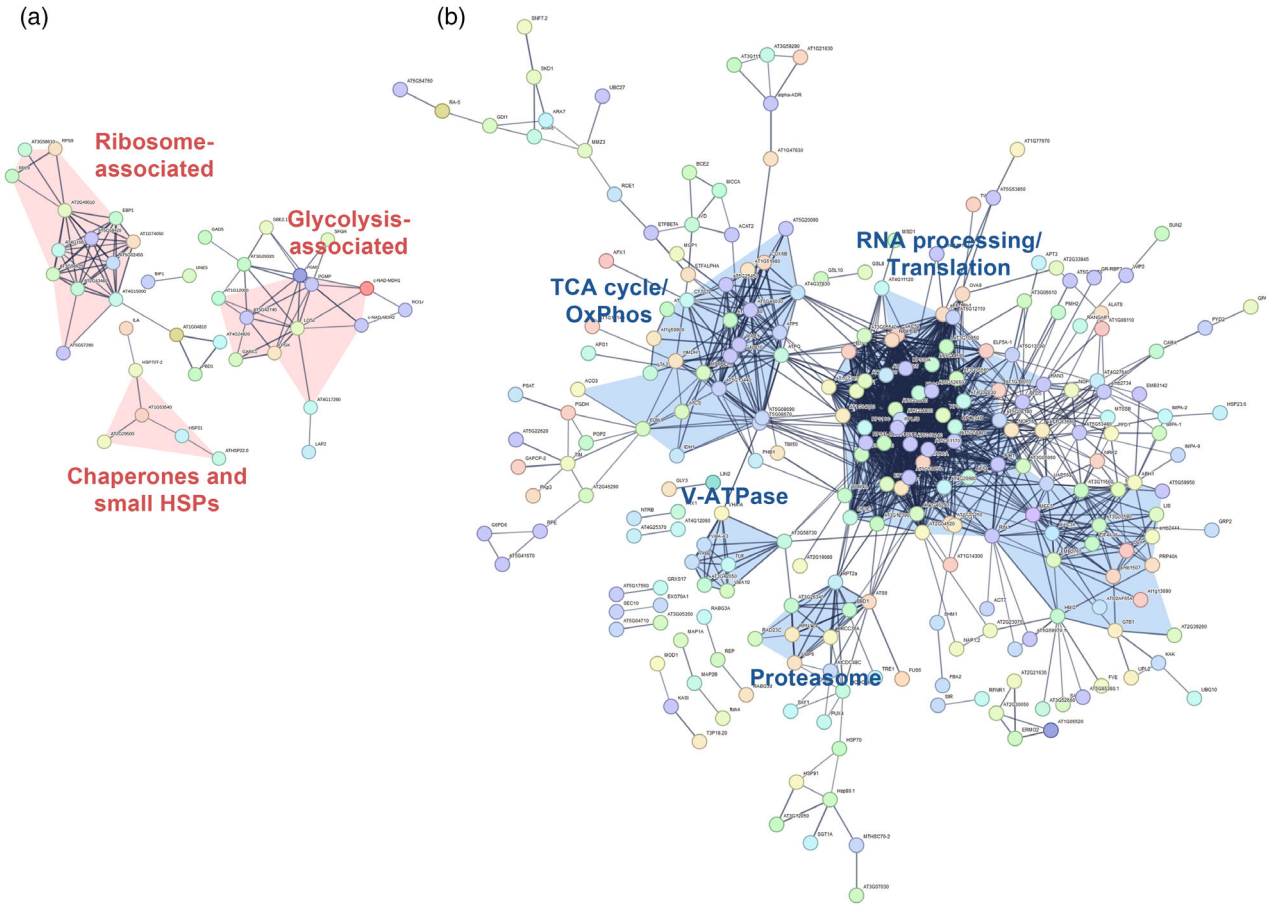

Figure 6. STRING database analyses. 145 proteins correlating positively (a) or 574 proteins correlating negatively (b) with dry matter content were queried against an *A. thaliana* database, functional associations were drawn from available experimental or co-expression information with high confidence. String thickness indicates confidence of association. Major clusters were manually highlighted by red (a) or blue (b) background and their respective general function.

and Table 2). In contrast to this, a cytosolic invertase (cINV) was found among the negative correlators. For the reactions occurring in the amyloplast, a positive correlation between the abundance of a plastid-localized phosphoglucomutase (pPGM) and dry matter content could be found, while the abundance of ADP-glucose pyrophosphorylase (AGPase) was found to be unaltered. Likewise, the starch-producing enzymes (soluble starch synthase SSS, granule-bound starch synthase GBSS, and branching enzyme BE) did not show correlations with dry matter content.

Interestingly, the nucleotide transporter NTT, which imports ATP necessary for ADP-Glucose (ADP-Glc) formation into the amyloplast, was also found among the positive correlators.

Taken together, these results indicate a shift from cINV- to SUS-mediated sucrose cleavage and enhanced funneling of sugars toward amyloplasts in high dry matter genotypes. At the same time, differential starch accumulation in the chosen genotypes is not driven by altered AGPase abundance. However, this does not rule out possible changes in AGPase enzyme activity, which is known to be regulated extensively (Figueroa et al., 2022).

Glycolysis and pentose phosphate pathway show contrasting enzyme abundance

Since the enrichment- and the STRING analysis both highlighted glycolysis and due to its importance for energy- and carbon-metabolism, the pathway was, together with the closely connected pentose phosphate pathway (PPP), investigated in further detail. In the canonical model, the starting metabolite of glycolysis, Glc, needs to be converted to fructose-6-phosphate (F6P) via G6P. However, as the main route of Suc breakdown in cassava storage roots is via SUS, F6P or G6P may enter the glycolytic pathway directly.

As shown in Figure 7b and Table 3, we found five glycolytic enzymes correlating positively with dry matter content. Apart from glucose-6-phosphate isomerase (GPI), which is also part of the sucrose-to-starch conversion pathway described above, pyrophosphate-dependent

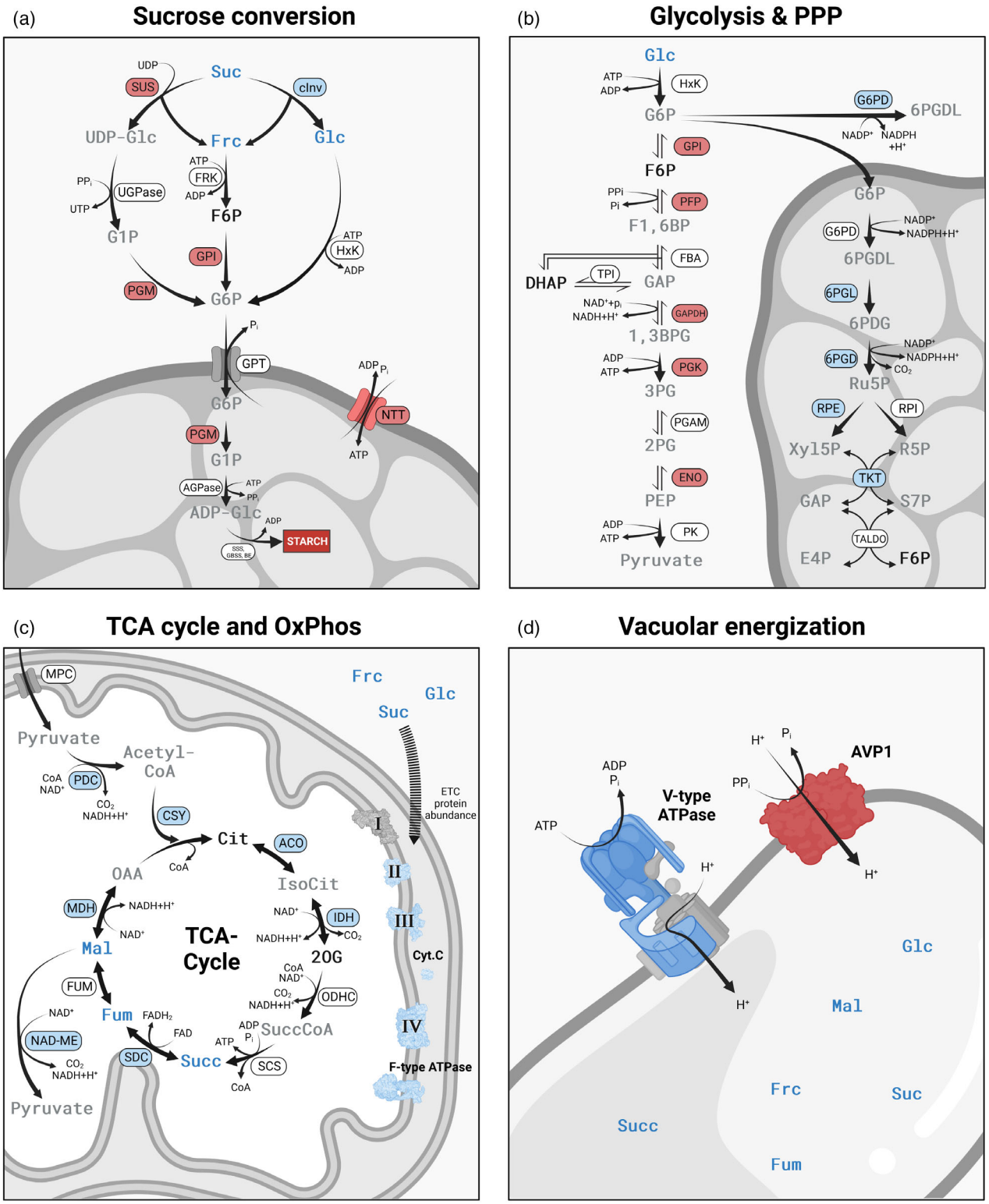

Figure 7. Overview over most perturbed pathways. The pathways of sucrose conversion (a), glycolysis and PPP (b), TCA cycle and OxPhos (c), and vacuolar energization (d) exhibited excessive alterations, as described in the main text body. Proteins and metabolites are color-coded according to their correlation to dry matter content, red: positive correlation, blue: negative correlation, black: no correlation, gray: not identified/measured.

Table 2 Enzymes of the sucrose converting pathway identified as correlators. The isoform was inferred from the Arabidopsis protein with the highest sequence similarity. Correlation coefficients from each of the three harvest years are indicated on the right, n.c.: no correlation

Enzyme	Isoform	Locus identifier	Spearman's correlation coefficients 2019/2020/2021
Sucrose synthase	SUS3	Manes.01G221900	n.c./0.57/0.60
Phosphoglucomutase	PGM1 (plastidial)	Manes.14G031100	0.64/0.69/0.57
	PGM2 (cytosolic)	Manes.15G131800	0.69/0.70/n.c.
Invertase	cINV1	Manes.18G103000	−0.76/−0.67/−0.60
Glucose-6-phosphate isomerase	SIS	Manes.05G160800	0.89/0.94/0.83
		Manes.18G027800	0.70/0.67/0.60
Nucleotide transporter	NTT1	Manes.11G049300	0.65/n.c./0.67

Table 3 Glycolytic- and PPP-enzymes identified as correlators. The isoform was inferred from the Arabidopsis protein with the highest sequence similarity. Correlation coefficients from each of the three harvest years are indicated on the right, n.c.: no correlation; n.i.: not identified

Enzyme	Isoform	Locus identifier	Spearman's correlation coefficients 2019/2020/2021
Pyrophosphate-dependent phosphofructokinase	PFP1β	Manes.16G002600	0.54/0.67/n.c.
Glyceraldehyde-3-phosphate dehydrogenase subunit C	GAPC	Manes.06G116400	0.73/0.82/n.c.
		Manes.14G054600	0.68/0.59/0.69
Phosphoglyceratekinase	PGK3	Manes.06G176500	0.64/0.71/n.c.
		Manes.14G008900	0.52/0.65/n.c.
Enolase	ENO2	Manes.14G033200	0.84/0.57/0.60
Glucose-6-phosphate dehydrogenase	G6PD6 (cytosolic)	Manes.15G156500	−0.52/−0.68/n.c.
6-Phosphogluconolactonase	PGL3 (plastidial)	Manes.05G193300	−0.61/n.c./−0.78
6-Phophogluconate dehydrogenase	PGD3 (plastidial)	Manes.02G013767	−0.71/−0.9/−0.73
Ribulose-5-phosphate-3-epimerase	RPE (plastidial)	Manes.06G090100	−0.73/−0.94/n.i.
Transketolase	TKL2 (plastidial)	Manes.05G047000	−0.66/−0.70/−0.56

phosphofructokinase (PFP), two isoforms of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglyceratekinase (PGK), and enolase (ENO) were found to correlate positively in at least two harvest years. Hence, cassava genotypes with higher dry matter accumulation strengthen the glycolytic pathway by increased enzyme abundance.

In contrast to this, enzymes related to the pentose phosphate pathway that provides reducing power and building blocks for nucleotide- and amino acid biosynthesis, were found to be generally less abundant in high dry matter genotypes. In detail, a cytosolic glucose-6-phosphate dehydrogenase (G6PD) as well as plastidial isoforms of 6-phosphogluconolactonase (6PGL), 6-phosphogluconate dehydrogenase (6PGD), ribulose-5-phosphate-3-epimerase (RPE), and transketolase (TKT) could be identified among the proteins correlating negatively with dry matter content (Figure 7b and Table 3).

TCA cycle enzymes and electron transport chain components correlate negatively to dry matter content

As the TCA cycle intermediates Succ, Mal, and Fum correlated with dry matter content and the enrichment analysis revealed OxPhos as a KEGG pathway associated with low dry matter genotypes, we investigated these mitochondrial metabolic pathways in further detail. Pyruvate generated during glycolysis can enter the mitochondrial matrix through mitochondrial pyruvate carrier (MPC). Here it is oxidatively decarboxylated to acetyl-CoA by pyruvate dehydrogenase complex (PDC). Within our proteome data, we found three subunits of this protein complex correlating negatively with dry matter content in at least two harvest years (Figure 7c and Table 4). Next, acetyl-CoA may enter the TCA cycle as an acetyl donor to regenerate oxaloacetate (OAA) to citrate (Cit). The enzyme catalyzing this reaction, citrate synthase (CSY), was also found to correlate negatively with dry matter content, as were the enzymes aconitase (ACO) and isocitrate dehydrogenase (IDH) that oxidize Cit to 2-oxoglutarate (2OG). Likewise, a subunit of the succinate dehydrogenase complex (SDC), as well as a malate dehydrogenase (MDH) were found among the negative correlators. Furthermore, the abundance of mitochondrial NAD-dependent malic enzyme (NAD-ME) correlated negatively with dry matter content. This enzyme drives the oxidative decarboxylation of Mal to pyruvate, thereby rerouting the TCA cycle.

Table 4 Enzymes related to the TCA cycle and OxPhos identified as correlators. The isoform or subunit was inferred from the Arabidopsis protein with the highest sequence similarity. Correlation coefficients from each of the three harvest years are indicated on the right, n.c.: no correlation; n.i.: not identified

Enzyme (complex)	Isoform (subunit)	Locus identifier	Spearman's correlation coefficients 2019/2020/2021
Pyruvate dehydrogenase complex	PDHE1-A	Manes.01G194500	−0.81/−0.82/n.c.
	PDHE1-B	Manes.09G185400	−0.59/−0.54/n.c.
	PDC-E2 1	Manes.10G054100	−0.72/−0.64/−0.60
Citrate synthase	CSY4	Manes.10G152000	−0.76/−0.79/−0.60
Aconitase	ACO3	Manes.13G070400	n.c./−0.77/−0.54
Isocitrate dehydrogenase	IDH-1	Manes.01G153600	n.c./−0.78/−0.69
Succinate dehydrogenase complex (Complex II)	SDH7B	Manes.14G069501	−0.71/−0.54/n.i.
Malate dehydrogenase	mMDH1	Manes.03G103200	−0.59/−0.74/n.c.
NAD-malic enzyme	NAD-ME2	Manes.01G260100	−0.55/−0.54/n.c.
Cytochrome c	CYTC-2	Manes.10G115100	n.c./−0.93/−0.61
Cytochrome bc1 complex (Complex III)	RISP2	Manes.11G076300	n.c./−0.64/−0.59
	CYC1	Manes.15G157600	−0.71/−0.68/n.c.
	QCR7-2	Manes.16G119000	−0.76/−0.85/n.c.
Cytochrome c oxidase (Complex IV)	COX5B-2	Manes.11G052200	−0.85/−0.92/−0.85
		Manes.04G117400	−0.65/−0.85/−0.60
	COX6A	Manes.13G106700	n.c./−0.77/−0.79
	COX6B	Manes.18G064621	−0.55/−0.55/n.c.
F-type ATPase	F1 subunit beta	Manes.06G062900	−0.75/−0.61/−0.53
		Manes.14G110100	−0.50/−0.54/n.c.
	F1 subunit delta	Manes.17G020400	−0.61/−0.87/−0.75
	F0 subunit O	Manes.10G143400	−0.72/−0.66/n.c.
	F0 subunit d	Manes.07G102500	−0.55/−0.56/n.c.

During OxPhos, cofactors reduced within TCA cycle reactions transfer electrons to an electron transport chain located in the inner mitochondrial membrane to reduce O_2 to H_2O. The energy released here is used to generate a proton gradient across the membrane, allowing an F-type ATPase to use the proton motif force to drive ATP synthesis and provide the cell with energy. We found the respective KEGG pathway enriched among the proteins correlating negatively with dry matter content. In fact, we were able to identify components of the electron transport chain's complexes II to IV, as well as multiple subunits of the F-type ATPase, revealing a generally lower abundance of OxPhos components in high dry matter genotypes (Figure 7c and Table 4).

Altogether, this reveals major adaptations in the configuration of the TCA cycle and OxPhos between low- and high-dry matter genotypes and underlines the observed accumulation of Succ, Fum, and Mal in low-dry matter genotypes.

V-type ATPase subunits show increased abundance in low dry matter genotypes

Tonoplast-located vacuolar-type H^+ ATPases (V-ATPases) hydrolyze ATP to generate a proton gradient enabling transport and ultimately vacuolar storage of various metabolites. The prominent emergence of subunits of such V-ATPases in the enrichment- and STRING analyses prompted us to have a closer look at vacuolar energization. In total, the proteome analysis revealed ten V-ATPase subunits that correlated in at least two of the three experiments. These ten proteins correlated strictly negatively with dry matter content (Figure 7d and Table 5). While eight of them are part of the peripheral V_1-subcomplex responsible for ATP hydrolysis and cytosolic interaction with other cellular components, two proteins refer to the membrane-integral V_0-subunit VHA-a3 (Li et al., 2022). Noteworthy, we could identify the inorganic pyrophosphatase AVP1 correlating positively with dry matter content in two of three harvest years. Like V-type ATPases, this pyrophosphatase is embedded in the vacuolar membrane and translocates protons into the vacuole. However, AVP1 hydrolyzes pyrophosphate (PP_i) rather than ATP (Gaxiola et al., 2007). With the concomitant downregulation of OxPhos, this can be interpreted as an energy-preserving adaptation in high dry matter genotypes.

These findings indicate an increase in tonoplast energization by V-type ATPases in cassava genotypes with lower dry matter accumulation. This in turn indicates an increased vacuolar storage capability. Considering the strong negative correlation of several metabolites including Glc, Frc, and Suc, it is tempting to speculate about these sugars being retained in the vacuole.

Table 5 V-ATPase subunits and AVP1 identified as correlators. The isoform or subunit was inferred from the Arabidopsis protein with the highest sequence similarity. Correlation coefficients from each of the three harvest years are indicated on the right, n.c.: no correlation

Enzyme (complex)	Isoform (subunit)	Locus identifier	Spearman's correlation coefficients 2019/2020 /2021
V-type H + -transporting ATPase	Subunit A	Manes.14G008200	−0.91/−0.78/−0.67
		Manes.06G177500	−0.87/−0.85/−0.67
	Subunit B2	Manes.04G071200	−0.93/−0.95/−0.74
	Subunit D	Manes.07G088000	−0.78/n.c./−0.77
	Subunit E1	Manes.01G148200	−0.70/−0.79/−0.90
	Subunit G	Manes.10G045500	−0.65/−0.90/−0.76
		Manes.07G101100	−0.68/−0.88/−0.79
	Subunit H	Manes.09G128600	−0.84/−0.88/−0.72
	Subunit a3	Manes.04G085500	−0.64/−0.63/−0.71
		Manes.11G087300	−0.58/−0.56/n.c.
Inorganic pyrophosphatase	AVP1	Manes.06G045100	n.c./0.77/0.68

DISCUSSION

Correlation analyses of proteomic, metabolomic, and agronomic data reveal metabolic alterations in dry matter accumulation in cassava

Generating high-yielding and resilient crop plants, either by breeding or biotechnological approaches, represents one of the main goals of modern plant science. Since model organisms like *Arabidopsis thaliana* are unsuitable to represent certain attributes of crop plants, like storage roots, it is necessary to gain direct insights into the respective crop plant's metabolism to achieve such goals. In this study, we used a population of field-grown cassava genotypes with distinct starch accumulation properties to elucidate differences in proteomic and metabolomic configurations of storage roots that either restrict or promote starch yield. Using this approach, we analyzed and integrated data from three individual field experiments and were able to identify biochemical pathways, cellular functions, and metabolites that were reproducibly correlating with dry matter content, a trait that can be used as a proxy for starch content.

We identified four major pathways that featured broadly contrasting behavior in low and high dry matter genotypes. These are (i) the sucrose conversion pathway, (ii) glycolysis and PPP, (iii) the TCA cycle and OxPhos, and (iv) vacuolar energization. In the following, we outline scenarios explaining the contrasting accumulation of starch in storage roots of the investigated cassava genotypes.

Sucrose-to-starch interconversion as limiting factor

The source-sink model is commonly used to describe limitations in photoassimilate production and/or utilization in crop plants. Undoubtedly, bottlenecks in either source or sink can contribute to restricted yield, thus both provide targets for biotechnological improvement (Sonnewald & Fernie, 2018). Interestingly, one major finding of our study revealed the accumulation of the hexoses Glc and Frc in low dry matter genotypes, while shoot biomass was unaltered. Additionally, the long-distance transport of sugar Suc was found to correlate negatively with dry matter. Hence, a limiting effect of photoassimilate provision on starch synthesis in the selected (low dry matter) genotypes appears unlikely, as carbon building blocks seemed readily available. Thus, a limitation in utilization of these sugar molecules seemed reasonable. This hypothesis was corroborated by the abundance of the enzymes converting Suc to G6P for amyloplast import: While abundance of SUS is increased in high dry matter genotypes, cINV is increased in low dry matter genotypes. In potatoes, the mode of Suc cleavage determines whether carbon building blocks are used for energy metabolism (cINV) or starch storage (SUS) (Ferreira & Sonnewald, 2012; Hajirezaei et al., 2000). Indeed, ectopic overexpression of SUS by the strong viral 35S promoter was shown to enhance yield of potato tubers (Baroja-Fernández et al., 2009). Therefore, a similar substrate channeling mechanism may decide the fate of assimilate utilization in cassava.

Moreover, GPI as well as cytosolic and plastidic PGM were more abundant in high dry matter genotypes, as was the nucleotide transporter NTT that provides the amyloplast with ATP. Geigenberger et al. (2004) analyzed publicly available data of Suc-converting enzymes to calculate flux control coefficients, that is, coefficients indicating the impact of changed enzyme activity on starch accumulation in potatoes. The authors found that especially NTT largely affected starch biosynthesis, while AGPase and both cytosolic and plastidic PGM exhibited more moderate effects. In line with this, NTT was included in biotech approaches to successfully enhance sink strength in potatoes (Jonik et al., 2012; Zhang et al., 2008). Considering this, the efficient channeling of Suc toward amyloplasts could become a limiting factor for starch synthesis in cassava, and the identified proteins bear the potential to serve as targets for

biotechnological alteration to increase carbon flux into amyloplasts.

At this point, it is important to note that we solely focused on steady-state measurements of cassava storage root material harvested 12 months after planting. This approach inherently omits the metabolism of all other parts of the plant, including source and photoassimilate transport tissue, which serve as vital parts of the carbon metabolism and could markedly influence processes in sink tissues (and vice versa). However, at least for certain cassava cultivars, grafting studies provided evidence for major sink limitations already decades ago (Cours, 1951; Hunt et al., 1977), which is in agreement with our results discussed above. Although our experimental setup did not allow us to resolve metabolite fluxes and subcellular compartmentalization, the data strongly suggest a beneficial effect of enhanced sucrose interconversion. Since a certain source colimitation still cannot be excluded, biotechnological improvement of cassava should include a concomitant improvement of source and sink, as was suggested before (Ludewig & Sonnewald, 2016; Sonnewald & Fernie, 2018).

Future work dedicated to metabolite and enzyme dynamics, as well as distribution throughout tissues and within cells will surely lead to an even better understanding of sugar interconversion processes in storage roots.

Substrate-level phosphorylation can provide ATP in dense tissues

While high dry matter genotypes were characterized by enhanced expression of sucrolytic enzymes, they simultaneously exhibited an increased abundance of glycolytic enzymes. Since oxidation of Glc to pyruvate enables substrate-level phosphorylation, that is, ATP synthesis, this alteration can be seen as a means to provide energy, for example, for enhanced starch biosynthesis.

Another explanation for increased glycolysis can be given by the physiology of the storage root: Especially in dense tissues such as potato tubers or bulking cassava roots, oxygen concentrations decrease drastically toward the center of the tissue, leading to a downregulation of respiratory pathways (Geigenberger et al., 2000). Under conditions where oxygen becomes a limiting factor for OxPhos, a consequential shortage in ATP provision can be averted by substrate-level phosphorylation via higher glycolytic activity (van Dongen & Licausi, 2015). In line with this, we previously showed ATP levels decreasing toward the center of the storage root (Mehdi et al., 2019).

A general downregulation of the pentose phosphate pathway in high dry matter genotypes

While various glycolytic enzymes were found to be more abundant in high dry matter genotypes, the closely associated PPP was found to be downregulated at multiple positions. Since the PPP is a major source of building blocks for biosynthesis of nucleotides and amino acids as well as reducing power in the form of NADPH (Kruger & von Schaewen, 2003), the observed downregulation can be seen as an adaptation of storage root metabolism away from cellular proliferation toward a more storage-oriented mode with decreased necessity for anabolic pathways. Furthermore, it is interesting to note that the plastidial part of the PPP, as well as amyloplast starch synthesis, use G6P as starting substrate. In fact, we could identify plastidial PGM, which would convert G6P to glucose-1-phosphate (G1P) for starch synthesis, as positively correlating enzyme, as was discussed earlier. Hence, substrate channeling toward starch synthesis may take place at the level of G6P in amyloplasts of high dry matter cassava storage roots.

TCA cycle flux mode supports genotype-dependent metabolic programs

Regarding dry matter-correlating metabolites and proteins, the TCA cycle was among the most perturbed pathways in our analysis. In fact, we found consistently negative correlations between Succ, Fum, and Mal. Likewise, multiple enzymes showed higher abundance in low dry matter genotypes. Importantly, not all metabolites/enzymes of the cycle were among the correlators. This is relevant as it is a misconception that the pathway only sustains a cyclic flux. Instead, Sweetlove et al. (2010) proposed alternative, non-cyclic flux modes that are based on modeling approaches and published experimental data underlining the plasticity of the TCA cycle. In an environment without ATP-limitation, one of these models suggests a disruption of flux between 2-OG and Fum to enable sustained cell proliferation by the provision of carbon building blocks for nitrogen assimilation and amino acid biosynthesis. Our data resembles these models, as low dry matter genotypes were characterized by an increased abundance of enzymes involved in the reaction steps from Succ to 2-OG, while reactions from 2-OG to Succ remained unaltered.

In line with this, the negative correlation of NAD-ME and dry matter content underlines a rerouting of the TCA cycle in low dry matter genotypes: A recent study in Arabidopsis found distinct mitochondrial pyruvate pools, either imported from the cytosol, or generated by NAD-ME. While imported pyruvate fuels the TCA cycle, NAD-ME-derived pyruvate is exported from the mitochondria and used for cellular metabolisms like fatty acid or amino acid synthesis (Le et al., 2022).

In fact, an impact on nitrogen metabolism across cassava genotypes is also reflected in the positive correlation of the nitrogen-rich amino acids Arg and Orn with dry matter content, as well as the negative correlation of urea. This is corroborated by the consistent negative correlation of arginase ARGAH1 (Locus identifier Manes.01G267300), which hydrolyzes Arg in the mitochondrial matrix, yielding urea and Orn, which in turn can be interconverted into

glutamate (Witte, 2011). Despite the positive correlation of the nitrogen-rich amino acids Arg and Orn with dry matter content, no increase in total nitrogen levels could be observed in high dry matter genotypes. Hence, rather than suffering from nitrogen starvation, low dry matter genotypes seem to invest nitrogen in molecules other than Arg or Orn. In fact, in the STRING database analysis of the proteins correlating negatively with dry matter content, one major cluster related to translation and protein degradation, indicating increased protein biosynthesis and turnover. Moreover, among the proteins that correlate negatively with dry matter content, we found several enzymes of the PPP, namely G6PD, 6PGL, 6PGD, RPE, and TKT. One of the oxidative PPP's main functions is the provision of carbon skeletons for nucleotides, which represent another form of nitrogen-rich molecules.

Taken together, our data indicate distinct alterations in TCA cycle components between contrasting cassava genotypes that reflect different metabolic programs in terms of synthesis of carbon building blocks and nitrogen assimilation.

Transition from oxidative phosphorylation to substrate-level phosphorylation in high dry matter genotypes

After glycolysis and the TCA cycle, OxPhos is the third and last step of aerobic respiration in plants. The process relies on four multiprotein complexes, mainly embedded in the inner mitochondrial membrane. These act as electron transport chains to reduce O_2 to H_2O while generating a proton gradient, which is used to drive an F-type ATPase to generate large amounts of ATP (Meyer et al., 2019). As was the case for TCA cycle enzymes, our data revealed an extensive downregulation of several components belonging to complexes II–IV, as well as multiple F-type ATPase subunits in high dry matter genotypes, arguing for decreased OxPhos. In concert with the observed upregulation of glycolytic enzymes, this can be interpreted as a gradual transition from OxPhos to glycolytic substrate-level phosphorylation that coincides with increasing starch storage capability. As discussed before, reasons for this might be found in an improved adaptation of high dry matter genotypes toward limited oxygen availability, which is a likely state in bulky tissues such as cassava storage roots (van Dongen & Licausi, 2015). Conversely, low dry matter genotypes are characterized by higher levels of sucrose and hexoses. Previous studies showed that expression of -nuclear-encoded subunits of all OxPhos complexes decreases upon sucrose starvation, consequentially limiting mitochondrial protein complex formation and leading to decreased respiration. This state could be reverted by replenishing sucrose (Giegé et al., 2005; Gonzalez et al., 2007). Such a signaling effect of sugars on OxPhos could also play a role in the analyzed cassava genotypes.

Tonoplast energization as a driver for vacuolar storage and growth

V-ATPases serve as ATP-hydrolyzing proton pumps and are localized in the tonoplast and endomembrane compartments of the secretory pathway such as the Trans-Golgi network (Dettmer et al., 2006; Li et al., 2022). The multisubunit protein complexes consist of two subcomplexes and a total of 13 subunits (Li et al., 2022). Of these, we could identify 10 proteins referring to seven subunits with a negative correlation coefficient of protein abundance and dry matter content. Thereof, two proteins referred to the VHA-a3 subunit, which is the membrane-integral subunit targeting the complex to the tonoplast, linking the identified proteins to the vacuole rather than other possible compartments (Krebs et al., 2010). Since the proton motif force generated by V-ATPases is essential for proton-coupled metabolite import into the vacuole, an increase in abundance of manifold V-ATPase subunits in low dry matter genotypes could facilitate increased solute transport into the vacuole. As a matter of fact, the accumulation of various sugars and metabolites can also be interpreted in this direction. Due to their reducing nature, especially the hexoses Glc and Frc, which showed a drastic negative correlation with dry matter content, are not likely to accumulate in high concentrations in the cytoplasm and are rather stored in the vacuole. An increase in solutes in the vacuole would furthermore increase the osmotic potential of the organelle, leading to water influx and expansion of the vacuole. It is interesting to note that vacuolar expansion is tightly connected to cell expansion (Kaiser & Scheuring, 2020; Krüger & Schumacher, 2018). Hence, it is possible that cassava plants with lower dry matter contents are focusing on cellular metabolism supporting proliferation, while genotypes featuring high dry matter contents are characterized by the mere storage of carbohydrates in amyloplasts. Furthermore, the STRING-database analysis of proteins correlating negatively with dry matter content revealed one major cluster related to RNA processing, ribosomes, and translation initiation, corroborating the idea of increased cellular metabolism to support growth processes.

Our data revealed the tonoplast inorganic pyrophosphatase AVP1 as positive correlator. While the protein acting as homodimer also contributes to vacuolar energization like V-ATPases, it does so by hydrolyzing PP_i instead of ATP (Gaxiola et al., 2007; Lin et al., 2012). It has been shown that overexpressing Arabidopsis AVP1 in cotton leads to increased drought- and salinity resistance in greenhouse experiments as well as enhanced yield in field experiments (Pasapula et al., 2011), and similar results could be obtained for other (crop) plants including tobacco, maize, and rice (Gao, 2006; Li et al., 2008; Zhao et al., 2006). Since our data show that high dry matter

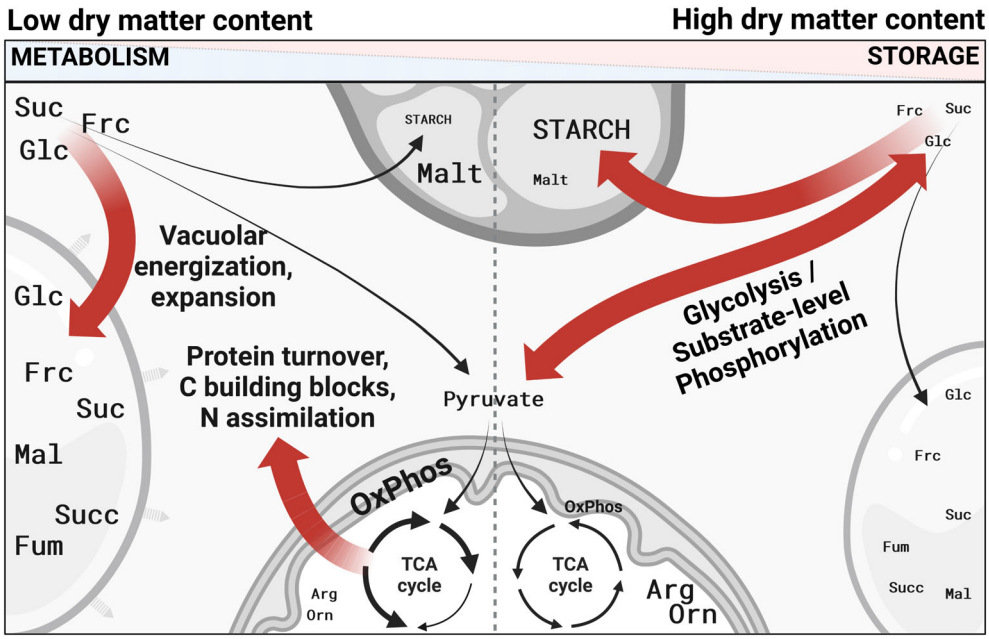

Figure 8. Summary of relevant pathways in low (left) and high (right) dry matter genotypes. Pathways that are generally strengthened are highlighted with red arrows, while pathways that are generally extenuated are shown by thinner black arrows. Accumulating metabolites are depicted in larger font. While low dry matter genotypes invest in a more active metabolism, this is gradually altered in favor of processes supporting starch storage in high dry matter genotypes.

genotypes are characterized by decreased OxPhos but rely on less efficient substrate-level phosphorylation for energy supply, it appears reasonable to decrease V-ATPase expression and instead utilize PP_i for vacuolar energization, as this would preserve ATP. Additionally, the increased stress resilience attributed to higher AVP1 abundance could induce a beneficial secondary effect under field conditions. In this scenario, the storage root's vacuoles might serve as water reservoirs to support the plant under unfavorable conditions, assigning an important, yet so far neglected, secondary role to cassava's storage roots.

CONCLUSION

The concomitant analysis of protein and metabolite composition of cassava genotypes with distinct dry matter contents revealed clear-cut alterations in several metabolic pathways. As the most prominent adaptations, a switch from OxPhos to substrate-level phosphorylation and an improved channeling of Suc toward starch synthesis was observed in high dry matter genotypes. In addition, genotypes associated with lower dry matter contents appeared to be more active in terms of growth and cellular metabolism, as could be seen by increases in mechanisms of vacuolar storage and expansion, the PPP, protein turnover, and alterations in the TCA cycle supporting generation of biosynthetic intermediates and N-assimilation (Figure 8). Furthermore, the observed alterations can also be interpreted as an adaptation toward limited oxygen availability in bulking storage roots. Under these conditions, adaptations of the energy metabolism toward substrate-level phosphorylation instead of OxPhos could be beneficial and allow for more efficient starch accumulation. Likewise, increased utilization of PP_i-dependent enzymes like PFP or AVP1 could aid high dry matter genotypes in terms of ATP preservation.

Importantly, differences between genotypes are clearly not confined to these examples. Furthermore, the highlighted processes are non-exclusive for genotypes accumulating high or low dry matter levels. Instead, these need to be understood as successive metabolic alterations that occur to a continuous extent throughout the different investigated cassava genotypes and lead to the observed high range of dry matter accumulation. Although the identification of key regulators that drive such metabolic shifts in the first place is still outstanding, the metabolic pathways and processes discussed above, as well as individual proteins, certainly offer targets for future efforts in breeding and/or biotechnological alterations to support the development of yield improved varieties of this important staple crop.

EXPERIMENTAL PROCEDURES

Field experiments and plant material

A core collection of 52 *Manihot esculenta* genotypes (CASS PYT52 collection) representing the genetic and phenotypic diversity available in the breeding germplasm was planted in three consecutive field seasons, namely 2018–2019 in Ibadan, Nigeria, and 2019–2020 as well as 2020–2021 in Ikenne, Nigeria. Each trial was established as Randomized Complete Block Design (RCBD) in two replicates per genotype. The experimental plot consisted of 36 plants,

established with an inter-row spacing of 1 m and intra-row spacing of 0.8 m. Plants were allowed to grow in field under natural conditions for 12 months, with weeding done when needed. Plant material from each plot was harvested from the four inner rows and pooled for agronomic data collection, including root weight and dry matter content, which was estimated as the percentage of dry weight relative to a 100 g of starting fresh weight. Storage root starch content was determined by wet-milling. From the selected genotypes, eight randomly selected storage roots were peeled and starch-storing xylem parenchyma samples were pooled before being freeze-dried on-site (Harvest 2019 and 2020) or flash-frozen in liquid nitrogen before being shipped to Erlangen, Germany (Harvest 2021). After crude grinding in a domestic blender, the plant material was lyophilized (only harvest year 2021), followed by homogenization in a bead mill.

k-Means clustering

k-means clustering was performed on the best linear unbiased predictors (BLUEs) of storage root yellowness and root fresh weight. BLUEs were calculated from a linear mixed effect model using the lme4 R package (Bates et al., 2015). Each trait was predicted using the genotype as fixed effect with year of harvest as random effect (intercept only) and an overall intercept of 0. The replication within each year was omitted as random effect as it led to a borderline singular fit of the model. The yellowness of the root was determined by the b value of a CIELAB colorimeter. For each clustering, the number of clusters (*k*) was determined using a within-cluster-sum-of-squares (WSS) elbow plot. *k* was set to two or three for yellowness and root fresh weight, respectively. For reproducibility's sake, a random seed of 1 was set for each clustering.

Metabolome analyses

For detailed metabolic analysis, metabolites were extracted from lyophilized and pulverized plant material of 16 (Harvest year 2019) or 14 (Harvest years 2020 and 2021) individual samples and subjected to gas chromatography–mass spectrometry (GC–MS) analysis as previously described by Rosado-Souza et al. (2019) and Obata et al. (2020). Data was reported according to recently updated reporting standards for metabolomics (Alseekh et al., 2021; Tables S1).

Soluble sugar measurements

Extraction and quantification of soluble sugars and starch were performed as described in Keller et al. (2021). 20 to 50 mg of freeze-dried and ground storage root material was extracted twice with 500 μl of 80% Ethanol at 80°C for 30 min. The ethanolic supernatants of the extracts were combined and evaporated in a vacufuge concentrator. Resulting pellets were dissolved in ddH_2O. Solved sugars were quantified with an NADP-coupled enzymatic test (Stitt et al., 1989) using a microplate reader. Remaining pellets from the extracts were used for starch quantification as described by Keller et al. (2021).

Elemental analysis

Nitrogen content of triplicate samples from harvest year 2021 was determined by CNS analyses performed by Raiffeisen-Laborservice, Ormont, Germany.

Total protein measurement

Total protein content of cassava storage root samples was determined in triplicate samples derived from the 2021 harvest. 20 mg of freeze-dried material were used for protein extraction with 200 μl of 50 mM Tris–HCl, 1 mM EDTA, and 0.01% Triton X-100. After centrifugation, the supernatant was subjected to chloroform/methanol precipitation (Wessel & Flügge, 1984), pellets were resuspended in H_2O and protein concentration was determined by a microtiter plate Bradford assay (Bio-Rad, Hercules, CA USA) according to the manufacturer's advice.

Comparative proteomics and data analysis

Protein was extracted from 20 mg of lyophilized and pulverized storage root material of each cassava genotype by the addition of 200 μl of 5% sodium dodecyl sulfate (SDS) in 50 mM triethylammonium bicarbonate buffer (TEAB). Protein extraction was supported by grinding with a micro-pestle attached to a laboratory stirrer followed by incubation at 60°C for 5 min. After centrifugation, the supernatant containing extracted proteins was transferred to a new reaction tube. Reduction, alkylation, and tryptic digest were performed on commercial S-Trap columns (Protifi, Huntington, NY USA) according to the manufacturer's instructions. After elution, peptides were dried in a vacuum concentrator and resolved in 100 μl 50 mM TEAB. Peptide concentration was determined using the Pierce Quantitative Fluorometric Peptide Assay (Thermo Fisher Scientific, Waltham, MA USA). 25 μg peptides of each sample were used for isobaric labeling using TMTpro labeling reagents (Thermo Fisher Scientific) according to the manufacturer's instruction. After quenching the reactions, all 16 individual samples were pooled. 100 μg of combined peptides were dried in a vacuum concentrator before being fractionated to eight fractions using the Pierce High pH Reversed-Phase Peptide Fractionation Kit (Thermo Fisher Scientific). Peptides were dried again before approx. 2 μg of each fraction were resolved in 10% formic acid and subjected to LC–MS on an Orbitrap Fusion Tribrid mass spectrometer connected to an UltiMate3000 nano-UHPLC system (Thermo Fisher Scientific) according to the method reported by Thompson et al. (2019). RAW data files were analyzed using PEAKS Studio 8.5 (Bioinformatics Solutions Inc., Waterloo, Canada) and the cassava proteome release 8.1 (acquired from https://phytozome-next.jgi.doe.gov/). For dynamic modifications, oxidation was selected, while carbamidomethylation of cysteines and TMT tags were selected as static modifications. Parent mass tolerance was set to 20 ppm, fragment mass tolerance was set to 0.5 Da, and quantitation mass tolerance was set to 0.003 Da. Relative protein abundance was calculated by the TMT reporter ion intensities acquired in MS^3 and an FDR of 1%. Further data analysis was done using Perseus v.1.6.15.0 (Tyanova et al., 2016). Briefly, proteins were arranged in protein groups and intensities were log2-transformed. All proteins containing invalid values were filtered out and intensities were normalized by subtraction of each protein's median intensity. Spearman's rank correlation was used to correlate protein intensities with dry matter content. For KEGG orthology annotation, cassava amino acid sequences were locally blasted (blast+ v.2.11.0, e-value $\leq 10^{-3}$) against the *A. thaliana* proteome (Cheng et al., 2017), and KEGG orthology IDs were assigned using the kofam-scan tool (v.1.3.0). These were used for Fisher's exact test to calculate enrichments of correlating proteins against all detected proteins.

STRING database analysis

For STRING database analyses (Szklarczyk et al., 2021), amino acid sequences of correlating proteins were queried against an *Arabidopsis thaliana* database to profit from the more extensive references of the model plant. As interaction sources, 'Experiments' and 'Co-expression' were chosen, while the minimum required

interaction score was set to high confidence. Furthermore, disconnected nodes were hidden in the final graphs. Clusters were highlighted manually.

ACKNOWLEDGEMENTS

Figures 7 and 8 were created with BioRender.com. We thank Otilia Ciobotea and Michaela Reiser for excellent technical assistance. We thank the Bill and Melinda Gates Foundation for funding this research through the grant INV-008053 'Metabolic Engineering of Carbon Pathways to Enhance Yield of Root and Tuber Crops' provided to Prof. Dr. Uwe Sonnewald. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

WZ, US, CEL, ARF, and HEN designed the research. IYR and AMD designed and executed the field trials, AS processed the sample material. ARF, DBM, LRS, and ID conceived and performed the metabolic analyses. CEL, with the help of JH, conceived and performed the proteomics experiments. HEN and BP conceived and performed the sugar measurements. DR performed the *k*-means cluster analysis. CEL and WZ wrote the manuscript.

DATA AVAILABILITY STATEMENT

Code used for *k*-means clustering is available on GitHub (https://github.com/DavidRuescher95/PYT52_PROTEOMICS.git).

The mass spectrometry proteomics data and documentation have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2022) partner repository with the dataset identifier PXD038813 and https://doi.org/10.6019/PXD038813.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Correlation of dry matter content with nitrogen- and dry matter content. Dry matter contents were correlated with total nitrogen levels as determined by elemental analysis (a) and with total protein contents (b).

Table S1. Correlation of dry matter content and metabolite levels in harvest years 2019 and 2020. Correlations were calculated as Pearson's correlation coefficient R, the p-value of the respective correlation is indicated on the right. Correlations with $P \leq 0.05$ were considered significant (black font), n.i., below quality standard, not included in analysis.

Table S2. Proteins correlating positively with dry matter content in at least two-third harvest years. Cassava locus identifiers and descriptions are shown next to the Arabidopsis locus identifier and name referring to the protein with the highest sequence similarity. Spearman's ranked correlation coefficient was calculated together with the respective p-value for each harvest year. n.i., not identified.

Table S3. Proteins correlating negatively with dry matter content in at least two-third harvest years. Cassava locus identifiers and descriptions are shown next to the Arabidopsis locus identifier and name referring to the protein with the highest sequence similarity. Spearman's ranked correlation coefficient was calculated together with the respective p-value for each harvest year. n.i., not identified.

Table S4. KEGG pathway enrichment analysis. Only significantly enriched pathways for positively (upper panel) and negatively correlating proteins (lower panel) are indicated together with the enrichment factor. Numbers on the right indicate the total background size (all identified proteins) and the number of proteins with the respective annotation among these, as well as the number of correlating proteins (foreground size) and the number of proteins with the respective annotation within this subset.

Table S5. Metabolite reporting checklist

Table S6. Metabolite annotation and documentation for GC–MS data of harvest year 2019.

Table S7. Metabolite annotation and documentation for GC–MS data of harvest year 2020.

Table S8. Metabolite annotation and documentation for GC–MS data of harvest year 2021.

REFERENCES

Alseekh, S., Aharoni, A., Brotman, Y., Contrepois, K., D'Auria, J., Ewald, J. *et al.* (2021) Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. *Nature Methods*, **18**, 747–756. Available from: http://www.nature.com/articles/s41592-021-01197-1

Baroja-Fernández, E., Muñoz, F.J., Montero, M., Etxeberria, E., Sesma, M.T., Ovecka, 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 Physiology*, **50**, 1651–1662. Available from: https://doi.org/10.1093/pcp/pcp108

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48. Available from: http://www.jstatsoft.org/v67/i01/

Beyene, G., Solomon, F.R., Chauhan, R.D., Gaitán-Solis, E., Narayanan, N., Gehan, J. *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 Biotechnology Journal*, **16**, 1186–1200. Available from: https://doi.org/10.1111/pbi.12862

Bull, S.E., Owiti, J.A., Niklaus, M., Beeching, J.R., Gruissem, W. & Vanderschuren, H. (2009) Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. *Nature Protocols*, **4**, 1845–1854. Available from: http://www.nature.com/articles/nprot.2009.208

Bull, S.E., Seung, D., Chanez, C., Mehta, D., Kuon, J.E., Truernit, E. *et al.* (2018) Accelerated ex situ breeding of GBSS- and PTST1-edited cassava for modified starch. *Science Advances*, **4**, eaat6086. Available from: https://doi.org/10.1126/sciadv.aat6086

Ceballos, H., Iglesias, C.A., Pérez, J.C. & Dixon, A.G.O. (2004) Cassava breeding: opportunities and challenges. *Plant Molecular Biology*, **56**, 503–516. Available from: https://doi.org/10.1007/s11103-004-5010-5

Cheng, C., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S. & Town, C.D. (2017) Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *The Plant Journal*, **89**, 789–804. Available from: https://doi.org/10.1111/tpj.13415

Cours, G. (1951) *Le manioc à Madagascar*. Paris: L'Université de Paris.

Dettmer, J., Hong-Hermesdorf, A., Stierhof, Y.-D. & Schumacher, K. (2006) Vacuolar H + -ATPase activity is required for endocytic and secretory trafficking in *Arabidopsis*. *Plant Cell*, **18**, 715–730. Available from: https://academic.oup.com/plcell/article/18/3/715/6114819

FAO. (2022) Crops and livestock products. License: CC BY-NC-SA 3.0 IGO. Extracted from: https://www.fao.org/faostat/en/#data/QCL. *Crops and livestock products. License: CC BY-NC-SA 3.0 IGO.* Extracted from: https://www.fao.org/faostat/en/#data/QCL. *Date of Access: 02.09.2022.*

Ferreira, S.J. & Sonnewald, U. (2012) The mode of sucrose degradation in potato tubers determines the fate of assimilate utilization. *Front. Plant Science*, **3**, 23. Available from: https://doi.org/10.3389/fpls.2012.00023

Figueroa, C.M., **Asencion Diez, M.D.**, **Ballicora, M.A.** & **Iglesias, A.A.** (2022) Structure, function, and evolution of plant ADP-glucose pyrophosphorylase. *Plant Molecular Biology*, **108**, 307–323. Available from: https://doi.org/10.1007/s11103-021-01235-8

Gao, F. (2006) Cloning of an H + -PPase gene from Thellungiella halophila and its heterologous expression to improve tobacco salt tolerance. *Journal of Experimental Botany*, **57**, 3259–3270. Available from: https://doi.org/10.1093/jxb/erl090

Gaxiola, R.A., **Palmgren, M.G.** & **Schumacher, K.** (2007) Plant proton pumps. *FEBS Letters*, **581**, 2204–2214. Available from: https://doi.org/10.1016/j.febslet.2007.03.050

Geigenberger, P., **Fernie, A.R.**, **Gibon, Y.**, **Christ, M.** & **Stitt, M.** (2000) Metabolic activity decreases as an adaptive response to low internal oxygen in growing potato tubers. *Biological Chemistry*, **381**, 723–740. Available from: https://doi.org/10.1515/BC.2000.093/html

Geigenberger, P., **Stitt, M.** & **Fernie, A.R.** (2004) Metabolic control analysis and regulation of the conversion of sucrose to starch in growing potato tubers. *Plant, Cell & Environment*, **27**, 655–673. Available from: https://doi.org/10.1111/j.1365-3040.2004.01183.x

Giegé, P., **Sweetlove, L.J.**, **Cognat, V.** & **Leaver, C.J.** (2005) Coordination of nuclear and mitochondrial genome expression during mitochondrial biogenesis in Arabidopsis. *Plant Cell*, **17**, 1497–1512. Available from: https://academic.oup.com/plcell/article/17/5/1497/6114545

Gomez, M.A., **Lin, Z.D.**, **Moll, T.**, **Chauhan, R.D.**, **Hayden, L.**, **Renninger, K.** ***et al.*** (2019) Simultaneous CRISPR/Cas9-mediated editing of cassava *eIF4E* isoforms *nCBP-1* and *nCBP-2* reduces cassava brown streak disease symptom severity and incidence. *Plant Biotechnology Journa*, **17**, 421–434. Available from: https://doi.org/10.1111/pbi.12987

Gonzalez, D.H., **Welchen, E.**, **Attallah, C.V.**, **Comelli, R.N.** & **Mufarrege, E.F.** (2007) Transcriptional coordination of the biogenesis of the oxidative phosphorylation machinery in plants. *The Plant Journal*, **51**, 105–116. Available from: https://doi.org/10.1111/j.1365-313X.2007.03121.x

Green, M.A. & **Fry, S.C.** (2005) Vitamin C degradation in plant cells via enzymatic hydrolysis of 4-O-oxalyl-l-threonate. *Nature*, **433**, 83–87. Available from: http://www.nature.com/articles/nature03172

Hachiya, T., **Ueda, N.**, **Kitagawa, M.**, **Hanke, G.**, **Suzuki, A.**, **Hase, T.** ***et al.*** (2016) Arabidopsis root-type ferredoxin:NADP(H) oxidoreductase 2 is involved in detoxification of nitrite in roots. *Plant Cell Physiology*, **57**, 2440–2450. Available from: https://doi.org/10.1093/pcp/pcw158

Hajirezaei, M., **Takahata, Y.**, **Trethewey, R.N.**, **Willmitzer, L.** & **Sonnewald, U.** (2000) Impact of elevated cytosolic and apoplastic invertase activity on carbon metabolism during potato tuber development. *Journal of Experimental Botany*, **51**, 439–445. Available from: https://doi.org/10.1093/jexbot/51.suppl_1.439

Hershey, C., **Álvarez, E.**, **Maung Aye, T.** ***et al.*** (2012) Eco-Efficient Interventions to Support Cassava's Multiple Roles in Improving the Lives of Smallholders. In: Hershey, C. & Neate, P. (Eds.) *Eco-Efficiency: From Vision to Reality*. Cali, Colombia: Centro Internacional de Agricultura Tropical International (CIAT), pp. 135–160. Available from: www.rtb.cgiar.org/resources/proposal-documents/final-proposal-annexes/

Howeler, R., **Lutaladio, N.** & **Graeme, T.** (2013) *Save and grow: Cassava*. Rome: Food and Agriculture Organization of the United Nations.

Hunt, L.A., **Wholey, D.W.** & **Cock, J.H.** (1977) Growth physiology of cassava. *Field Crop Abstracts*, **30**, 77–91.

Jonik, C., **Sonnewald, U.**, **Hajirezaei, M.-R.**, **Flügge, U.-I.** & **Ludewig, F.** (2012) Simultaneous boosting of source and sink capacities doubles tuber starch yield of potato plants. *Plant Biotechnology Journal*, **10**, 1088–1098. Available from: https://doi.org/10.1111/j.1467-7652.2012.00736.x

Kaiser, S. & **Scheuring, D.** (2020) To Lead or to follow: contribution of the plant vacuole to cell growth. *Frontiers in Plant Science*, **11**, 553. Available from: https://doi.org/10.3389/fpls.2020.00553

Kawano, K. & **Cock, J.H.** (2005) Breeding cassava for underprivileged. *Journal of Crop Improvement*, **14**, 197–219. Available from: https://doi.org/10.1300/J411v14n01_09

Keller, I., **Müdsam, C.**, **Rodrigues, C.M.**, **Kischka, D.**, **Zierer, W.**, **Sonnewald, U.** ***et al.*** (2021) Cold-triggered induction of ROS- and Raffinose metabolism in freezing-sensitive taproot tissue of sugar beet. *Frontiers in Plant Science*, **12**, 1886. Available from: https://doi.org/10.3389/fpls.2021.715767

Krebs, M., **Beyhl, D.**, **Görlich, E.**, **Al-Rasheid, K.A.S.**, **Marten, I.**, **Stierhof, Y.-D.** ***et al.*** (2010) Arabidopsis V-ATPase activity at the tonoplast is required for efficient nutrient storage but not for sodium accumulation. *Proceedings of the National Academy of Sciences*, **107**, 3251–3256. Available from: https://doi.org/10.1073/pnas.0913035107

Krüger, F. & **Schumacher, K.** (2018) Pumping up the volume – vacuole biogenesis in *Arabidopsis thaliana*. *Seminars in Cell & Developmental Biology*, **80**, 106–112. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1084952116304001

Kruger, N.J. & **von Schaewen, A.** (2003) The oxidative pentose phosphate pathway: structure and organisation. *Current Opinion in Plant Biology*, **6**, 236–246. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1369526603000396

Le, X.H., **Lee, C.P.**, **Monachello, D.** & **Millar, A.H.** (2022) Metabolic evidence for distinct pyruvate pools inside plant mitochondria. *Nature Plants*, **8**, 694–705. Available from: https://www.nature.com/articles/s41477-022-01165-3

Li, B., **Wei, A.**, **Song, C.**, **Li, N.** & **Zhang, J.** (2008) Heterologous expression of the TsVP gene improves the drought resistance of maize. *Plant Biotechnology Journa*, **6**, 146–159. Available from: https://doi.org/10.1111/j.1467-7652.2007.00301.x

Li, K.-T., **Moulin, M.**, **Mangel, N.**, **Albersen, M.**, **Verhoeven-Duif, N.M.**, **Ma, Q.** ***et al.*** (2015) Increased bioavailable vitamin B6 in field-grown transgenic cassava for dietary sufficiency. *Nature Biotechnology*, **33**, 1029–1032. Available from: http://www.nature.com/articles/nbt.3318

Li, Y., **Zeng, H.**, **Xu, F.**, **Yan, F.** & **Xu, W.** (2022) H^+-ATPases in plant growth and stress responses. *Annual Review of Plant Biology*, **73**, 495–521. Available from: https://doi.org/10.1146/annurev-arplant-102820-114551

Lin, S.-M., **Tsai, J.-Y.**, **Hsiao, C.-D.**, **Huang, Y.T.**, **Chiu, C.L.**, **Liu, M.H.** ***et al.*** (2012) Crystal structure of a membrane-embedded H + -translocating pyrophosphatase. *Nature*, **484**, 399–403. Available from: http://www.nature.com/articles/nature10963

Ludewig, F. & **Sonnewald, U.** (2016) Demand for food as driver for plant sink development. *Journal of Plant Physiology*, **203**, 110–115. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0176161716300943

Mehdi, R., **Lamm, C.E.**, **Bodampalli Anjanappa, R.**, **Müdsam, C.**, **Saeed, M.**, **Klima, J.** ***et al.*** (2019) Symplasmic phloem unloading and radial post-phloem transport via vascular rays in tuberous roots of Manihot esculenta J. Lunn, ed. *Journal of Experimental Botany*, **70**, 5559–5573. Available from: https://academic.oup.com/jxb/article/70/20/5559/5522314

Meyer, E.H., **Welchen, E.** & **Carrie, C.** (2019) Assembly of the complexes of the oxidative phosphorylation system in land plant mitochondria. *Annual Review of Plant Biology*, **70**, 23–50. Available from: https://doi.org/10.1146/annurev-arplant-050718-100412

Narayanan, N., **Beyene, G.**, **Chauhan, R.D.**, **Grusak, M.A.** & **Taylor, N.J.** (2021) Stacking disease resistance and mineral biofortification in cassava varieties to enhance yields and consumer health. *Plant Biotechnology Journal*, **19**, 844–854. Available from: https://doi.org/10.1111/pbi.13511

Nielsen, E. (2020) The small GTPase superfamily in plants: a conserved regulatory module with novel functions. *Annual Review of Plant Biology*, **71**, 247–272. Available from: https://doi.org/10.1146/annurev-arplant-112619-025827

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. *The Plant Journal*, **102**, 1202–1219. Available from: https://doi.org/10.1111/tpj.14693

Owiti, J., **Grossmann, J.**, **Gehrig, P.**, **Dessimoz, C.**, **Laloi, C.**, **Hansen, M.B.** ***et al.*** (2011) iTRAQ-based analysis of changes in the cassava root proteome reveals pathways associated with post-harvest physiological deterioration. *The Plant Journal*, **67**, 145–156. Available from: https://doi.org/10.1111/j.1365-313X.2011.04582.x

Pasapula, V., **Shen, G.**, **Kuppu, S.**, **Paez-Valencia, J.**, **Mendoza, M.**, **Hou, P.** ***et al.*** (2011) Expression of an Arabidopsis vacuolar H + -pyrophosphatase gene (AVP1) in cotton improves drought- and salt tolerance and increases fibre yield in the field conditions. *Plant Biotechnology Journal*, **9**, 88–99. Available from: https://doi.org/10.1111/j.1467-7652.2010.00535.x

Perez-Riverol, Y., **Bai, J.**, **Bandla, C.**, **García-Seisdedos, D.**, **Hewapathirana, S.**, **Kamatchinathan, S.** ***et al.*** (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Research*, **50**, D543–D552. Available from: https://academic.oup.com/nar/article/50/D1/D543/6415112

Rosado-Souza, L., **David, L.C.**, **Drapal, M.**, **Fraser, P.D.**, **Hofmann, J.**, **Klemens, P.A.W.** ***et al.*** (2019) Cassava metabolomics and starch quality.

Current Protocols in Plant Biology, **4**, e20102. Available from: https://doi.org/10.1002/cppb.20102

Sayre, R., **Beeching, J.R.**, **Cahoon, E.B.**, **Egesi, C.**, **Fauquet, C.**, **Fellman, J.** ***et al.*** (2011) The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. *Annual Review of Plant Biology*, **62**, 251–272. Available from: https://doi.org/10.1146/annurev-arplant-042110-103751

Sonnewald, U. & **Fernie, A.R.** (2018) Next-generation strategies for understanding and influencing source–sink relations in crop plants. *Current Opinion in Plant Biology*, **43**, 63–70. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1369526617301504

Stitt, M., **Lilley, R.M.**, **Gerhardt, R.** & **Heldt, H.W.** (1989) [32] Metabolite levels in specific cells and subcellular compartments of plant leaves. In: *Methods in enzymology*. San Diego: Academic Press, pp. 518–552. Available from: https://linkinghub.elsevier.com/retrieve/pii/0076687989740350

Sweetlove, L.J., **Beard, K.F.M.**, **Nunes-Nesi, A.**, **Fernie, A.R.** & **Ratcliffe, R.G.** (2010) Not just a circle: flux modes in the plant TCA cycle. *Trends in Plant Science*, **15**, 462–470. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1360138510000993

Szklarczyk, D., **Gable, A.L.**, **Nastou, K.C.**, **Lyon, D.**, **Kirsch, R.**, **Pyysalo, S.** ***et al.*** (2021) The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, **49**, D605–D612. Available from: https://academic.oup.com/nar/article/49/D1/D605/6006194

Taylor, N., **Gaitán-Solís, E.**, **Moll, T.**, **Trauterman, B.**, **Jones, T.**, **Pranjal, A.** ***et al.*** (2012) A high-throughput platform for the production and analysis of transgenic cassava (Manihot esculenta) plants. *Tropical Plant Biology*, **5**, 127–139. Available from: https://doi.org/10.1007/s12042-012-9099-4

Thompson, A., **Wölmer, N.**, **Koncarevic, S.**, **Selzer, S.**, **Böhm, G.**, **Legner, H.** ***et al.*** (2019) TMTpro: design, synthesis, and initial evaluation of a Proline-based isobaric 16-Plex tandem mass tag reagent set. *Analytical Chemistry*, **91**, 15941–15950. Available from: https://doi.org/10.1021/acs.analchem.9b04474

Tyanova, S., **Temu, T.**, **Sinitcyn, P.**, **Carlson, A.**, **Hein, M.Y.**, **Geiger, T.** ***et al.*** (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, **13**, 731–740. Available from: http://www.nature.com/articles/nmeth.3901

van Dongen, J.T. & **Licausi, F.** (2015) Oxygen Sensing and Signaling. *Annual Review of Plant Biology*, **66**, 345–367. Available from: https://doi.org/10.1146/annurev-arplant-043014-114813

Vanderschuren, H., **Nyaboga, E.**, **Poon, J.S.**, **Baerenfaller, K.**, **Grossmann, J.**, **Hirsch-Hoffmann, M.** ***et al.*** (2014) Large-scale proteomics of the cassava storage root and identification of a target gene to reduce postharvest deterioration. *Plant Cell*, **26**, 1913–1924. Available from: https://academic.oup.com/plcell/article/26/5/1913-1924/6099918

Wang, X., **Chang, L.**, **Tong, Z.**, **Wang, D.**, **Yin, Q.**, **Wang, D.** ***et al.*** (2016) Proteomics profiling reveals carbohydrate metabolic enzymes and 14–3-3 proteins play important roles for starch accumulation during cassava root Tuberization. *Scientific Reports*, **6**, 1–15. Available from: https://doi.org/10.1038/srep19643

Wessel, D. & **Flügge, U.I.** (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Analytical Biochemistry*, **138**, 141–143. Available from: https://linkinghub.elsevier.com/retrieve/pii/0003269784907826

Witte, C.-P. (2011) Urea metabolism in plants. *Plant Science*, **180**, 431–438. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168945210003195

Zhang, L., **Häusler, R.E.**, **Greiten, C.**, **Hajirezaei, M.-R.**, **Haferkamp, I.**, **Neuhaus, H.E.** ***et al.*** (2008) Overriding the co-limiting import of carbon and energy into tuber amyloplasts increases the starch content and yield of transgenic potato plants. *Plant Biotechnology Journal*, **6**, 453–464. Available from: https://doi.org/10.1111/j.1467-7652.2008.00332.x

Zhao, F.-Y., **Zhang, X.-J.**, **Li, P.-H.**, **Zhao, Y.-X.** & **Zhang, H.** (2006) Co-expression of the Suaeda salsa SsNHX1 and Arabidopsis AVP1 confer greater salt tolerance to transgenic rice than the single SsNHX1. *Molecular Breeding*, **17**, 341–353. Available from: https://doi.org/10.1007/s11032-006-9005-6

Zierer, W., **Anjanappa, R.B.**, **Lamm, C.E.**, **Chang, S.-H.**, **Gruissem, W.** & **Sonnewald, U.** (2022) A promoter toolbox for tissue-specific expression supporting translational research in cassava (Manihot esculenta). *Frontiers in Plant Science*, **13**, 1042379. Available from: https://doi.org/10.3389/fpls.2022.1042379