

Morphological and metabolic profiling of a tropical-adapted potato association panel subjected to water recovery treatment reveals new insights into plant vigor

David Toubiana^{1,*} , Rodrigo Cabrera¹, Elisa Salas², Chiara Maccera², Gabriel Franco dos Santos¹, Danny Cevallos², Hannele Lindqvist-Kreuze², Juan M. Lopez¹ and Helena Maruenda^{1,*} 

¹Departamento de Ciencias – Química, Centro de Espectroscopia de Resonancia Magnética Nuclear (CERMN), Pontificia Universidad Católica del Perú, Av. Universitaria 1801, Lima, Lima 32, Peru, and

²Genetics and Crop Improvement, International Potato Center, Av. La Molina 1895, Lima, Lima 12, Peru

Received 12 December 2019; revised 20 May 2020; accepted 12 June 2020; published online 24 June 2020.

*For correspondence (e-mails dtoubiana@pucp.edu.pe and hmaruen@pucp.edu.pe).

SUMMARY

Potato (*Solanum tuberosum* L.) is one of the world's most important crops, but it is facing major challenges due to climatic changes. To investigate the effects of intermittent drought on the natural variability of plant morphology and tuber metabolism in a novel potato association panel comprising 258 varieties we performed an augmented block design field study under normal irrigation and under water-deficit and recovery conditions in Ica, Peru. All potato genotypes were profiled for 45 morphological traits and 42 central metabolites via nuclear magnetic resonance. Statistical tests and norm of reaction analysis revealed that the observed variations were trait specific, that is, genotypic versus environmental. Principal component analysis showed a separation of samples as a result of conditional changes. To explore the relational ties between morphological traits and metabolites, correlation-based network analysis was employed, constructing one network for normal irrigation and one network for water-recovery samples. Community detection and difference network analysis highlighted the differences between the two networks, revealing a significant correlational link between fumarate and plant vigor. A genome-wide association study was performed for each metabolic trait. Eleven single nucleotide polymorphism (SNP) markers were associated with fumarate. Gene Ontology analysis of quantitative trait loci regions associated with fumarate revealed an enrichment of genes regulating metabolic processes. Three of the 11 SNPs were located within genes, coding for a protein of unknown function, a RING domain protein and a zinc finger protein ZAT2. Our findings have important implications for future potato breeding regimes, especially in countries suffering from climate change.

Keywords: *Solanum tuberosum* L., drought, potato morphology, potato tuber central metabolism, NMR, correlation based network analysis, zinc finger protein ZAT2, RING domain protein.

INTRODUCTION

Potato (*Solanum tuberosum* L.) is one of the world's most important crops, exceeded only by rice and wheat in terms of human consumption. In 2017 alone, approximately 388 million metric tons were produced globally (FAO database). Potato is mainly grown for the production of tubers for various forms of human consumption (boiled, fried, mashed, potato fries and potato chips) (Haverkort and Struik, 2015). Potato tubers are a preferred source of carbohydrates (starch), and are low in fat and abundant in vitamins B₁, B₃, B₆ and C. They are also a good source of potassium, magnesium, phosphorus and iron (International Potato Center, CIP, Peru and US Department of

Agriculture Database, USDA). Since the European discovery and conquest of the Americas the potato has spread throughout the world, and due to its nutritional value and its globally distributed cultivation with high yields it has become one of the essential species in terms of food security for a growing world population (Birch *et al.*, 2012).

The agronomic plasticity of the potato crop allows its cultivation under different climatic conditions. Nevertheless, the potato is a cool-season crop, producing its highest yields in regions with a moderate climate at an optimal growth temperature of approximately 20°C (Haverkort and Struik, 2015). Different climate simulation models predict a mean temperature increase of between 1.1 and 6.4°C up to

the end of the 21st century (Christensen *et al.*, 2007), with elongated and higher frequencies of extreme heat episodes (Ahuja *et al.*, 2010; Mittler and Blumwald, 2010; Mittler *et al.*, 2012; Li *et al.*, 2013). Simultaneously, heat waves will be accompanied by droughts or floods, drastically affecting growth conditions for crop plants (Atkinson and Urwin, 2012), eventually decreasing food security and increasing its vulnerability.

Although the simulation models prognosticate that the full impact of climate change is yet to be reached, consequences for global potato production can already be seen (Hijmans, 2003). Developing countries in particular suffer from lower yields than countries in Europe and North America. According to the Food and Agriculture Organization of the United Nations (<http://www.fao.org/>), developing countries achieve potato yields of 13–18 t ha⁻¹ versus 22–33 t ha⁻¹ in industrialized countries. To counteract the issue of drought and ensure high yields despite adverse climates, it is necessary to develop drought-resistant potato varieties. The effect of drought on the potato plant is a manifestation of the frequency and intensity of the water shortage and the phenology of the plant at the time of the stress (van Loon, 1981). To properly dissect drought tolerance, accurate and objective phenotyping is required. Most of the agronomic traits of potato are quantitatively inherited (Watanabe, 2015). One approach to identify the genomic regions responsible for the traits is quantitative trait locus (QTL) mapping using bi-parental crosses (Collard *et al.*, 2005). This approach, however, offers limited genetic resolution because of the restricted number of meiotic events and small amount of allelic diversity captured in one cross (Salvi and Tuberosa, 2005). An alternative approach is association mapping, which overcomes both constraints as it utilizes a large population of more distantly related individuals exploiting ancestral recombination events, and takes into account all alleles present in the population to identify significant marker–phenotype associations (Gupta *et al.*, 2014). Genomic regions associated with traits can be identified based on non-random associations of alleles at nearby loci. The quality of phenotypic data, population size and the degree of linkage disequilibrium (LD) present in the populations determine the success of association discovery. In recent years there have been growing efforts to exploit association panels, which are designed to capture a wide phenotypic variability (Yu and Buckler, 2006; Scossa *et al.*, 2016). This approach has successfully elucidated the genetic basis of metabolic natural variance, including that of carotenoids, glucosinolates and organic acids in different plant species (Riedelsheimer *et al.*, 2012; Gonzalez-Jorge *et al.*, 2013; Lipka *et al.*, 2013; Verslues *et al.*, 2014).

Here, we utilized a potato association panel comprising 258 diverse, tropically adapted tetraploid genotypes in an augmented block design field trial under normal irrigation (NI) and a water recovery (REC) treatment (see definitions

below). During and after the field trial morphological traits were measured. In addition, after the field experiment was finalized, the tubers were investigated for the phenotypic variation of metabolites. The main objectives of the current study were twofold: (i) to assess the effects of the REC treatment on plant performance and tuber metabolism, and (ii) to study the relationship between plant morphological traits and tuber metabolism, potentially identifying candidate metabolite markers for REC treatment.

RESULTS

The response of plants to water stress differs depending on the type and timing of the applied water deficiency. For instance, during terminal drought soil water content is progressively decreased and, once the water is depleted, the plant dies prematurely as a result of severe disruption of the photosynthetic apparatus, leading to altered carbohydrate metabolism and reduction of nutrient uptake (reviewed in Bodner *et al.*, 2015). Intermittent drought, on the other hand, is defined as a progressive decrease of soil water content followed by a period of re-watering, allowing the plant to recover and hence tolerate the stress (Bodner *et al.*, 2015); this treatment is also referred to as REC. The phenotypic response to drought in potato depends on the genotype as well as its developmental stage, the environment and their interaction with one another (Obidiegwu *et al.*, 2015). Normal irrigation treatment serves as a control for understanding the potential of the plant when water is not restricted.

To understand the effects of intermittent drought on the performance of the potato plant we applied a NI treatment and a REC treatment on a potato association panel in an augmented block design (for details see Experimental Procedures).

Apart from being an agronomically important crop, potato represents a suitable model for studying sink–source interactions in plants. Traditionally, source tissues are defined as tissues capable of exporting net energy and sink tissues as tissues importing energy from source tissues (Turgeon, 1989). Potato tubers are high-capacity sink tissues. Photosynthetic products are transported in the form of photoassimilates to the sink tissues (tuber), where starch granules are synthesized in the amyloplasts via the polymerization of glucose. Indeed, the interactions between the sink and source organs in potato are rather complex. They are regulated at different levels and can be easily perturbed or stimulated; for example, elevated temperatures (30°C) were shown to have a negative impact on starch synthesis (Geigenberger *et al.*, 1998). In the current study, we investigated the effect of REC conditions on source–sink interactions in potato by exploring 45 different plant morphological traits (Data S1 in the online Supporting Information) and 42 metabolites: 19 amino acids, six organic acids, six sugars, two N-compounds, two purines,

one sugar alcohol, one alkaloid, choline, two pyrimidines and two unknowns (for a detailed overview see Table S1 and Data S2).

Statistical and multivariate analyses

Descriptive statistics were estimated for all morphological traits (Figure S1) and metabolites (Figure S2) in three different modes, that is, for NI samples, for REC samples and for NI and REC samples combined. The 45 morphological traits analyzed for the current study were chosen as indicators for the response of potato plants induced by water-deficiency stress (van Loon, 1981; Cabello *et al.*, 2013; Obidiegwu *et al.*, 2015).

For the morphological traits a noticeable variance was detected for the total aerial part of fresh weight (TAFW) and harvest index of fresh weight (HI_FW) during the NI treatment (Figure S1a). During the REC treatment chlorophyll-associated traits as well as tuber-related traits showed high variances (Figures S1b). In the mixed mode the variance for almost all traits increased as expected due to the adaptation of the plants to the stress (Figure S1c). One-factorial analysis of variance (ANOVA) of the mixed model revealed significant differences for 33 out of the 46 traits, while 13 traits did not show significant differences, namely chlorophyll content first evaluation (CC_EV1), canopy reflectance using the normalized difference vegetation index, first evaluation (CR_NDVI_EV1), canopy reflectance using the normalized difference vegetation index, third evaluation (CR_NDVI_EV3), the slope of the canopy reflectance NDVI (CR_NDVI_Slp), the drone NDVI first evaluation (Drone_NDVI_EV1), harvest index fresh weight (HI_FW), dry matter content of aerial part (APDMC), plant height first evaluation (PLAHE_EV1), plant uniformity, plant vigor, stem diameter first evaluation (SD_EV1), stem diameter evaluation third evaluation (SD_EV3) and TAFW. The pronounced occurrence of traits corresponding to the first evaluation time point was indicative of low-level adjustment as the REC treatment had just been applied. For the metabolites the highest variances were detected for citrate, asparagine, sucrose, glutamate, γ -aminobutyric acid (GABA), malate and proline, hinting at the importance of genotype-specific fine tuning of nitrogen (N) to carbon (C) metabolism (Nunes-Nesi *et al.*, 2010). In particular, the occurrence of citrate, malate, glutamate and asparagine suggested an adjustment of tricarboxylic acid (TCA) cycle activity with respect to the GABA shunt (Fait *et al.*, 2008) and the GOGAT cycle. One-factorial ANOVA with the mixed model at a significance threshold of $P \leq 0.05$ revealed significant differences for 31 of the 42 metabolites. Metabolites that did not change significantly were the two purines (adenosine and guanosine), the amino acids histidine, lysine and threonine, the two alternative fructose structures (α -fructose and β -fructose), the pyrimidine cytidine and finally the N-compound allantoin.

Next, we used the morphological traits and metabolite profiles to run principal component analysis (PCA). In the PCA of the morphological traits a clear separation between NI and REC samples was visible on PC1, accounting for 21.3% of the cumulative variance observed (Figure 1b). The traits with the greatest loadings were: total tuber yield adjusted (TTYA), total tuber weight per plant (TTWPL), total tuber weight per plot (TTWP), total tuber yield not adjusted (TTYNA) and tuber fresh weight per plant (TFW). The predominant high loadings of traits associated with tuber morphology were indicative of significant variation of the potato tuber yield attributed to the conditional changes. The PCA of the metabolite profiles demonstrated a separation between NI and REC samples on PC2 (Figure 1a), accounting for 17% of the total observed variance. The associated PCA loadings revealed the greatest impact of citrate, choline, tyrosine, phenylalanine and tryptophan.

To understand the genotypic versus environmental contribution to the observed phenotypic variation, we calculated the norm of reaction for all traits (Figure 2). In a norm of reaction plot, the slope is indicative of the genotypic or environmental contribution. A low slope suggests a highly heritable trait regardless of the environment, whereas a high slope indicates a large environmental contribution to the phenotype. For the morphological traits a large slope was observed for traits related to chlorophyll content, particularly during the second and third evaluation periods, as well as for traits associated with the tuber (Figure 2a,b). A large genotypic contribution was visible for traits measured during the first evaluation period (Figure 2a,d). This finding was expected, as this was the evaluation period shortly after initiation of the REC treatment. Metabolic profiling of the tubers was performed after the field experiment was terminated, so that at this evaluation period the tubers will have experienced the full impact of the REC treatment. Proline demonstrated the steepest slope for metabolites, indicative of an elevated environmental contribution to the differences observed for the NI and REC treatments (Figure 2e,f). Sugars and other amino acids, such as threonine, histidine and lysine, revealed a greater genotypic contribution (Figures 2e,h).

Correlation analysis emphasizes functional groups of amino acids and morphological traits

To identify patterns of coordinated behavior, a pairwise-correlation analysis between metabolites and morphological traits using Pearson correlation was carried out (Figures 3, S3 and S4). The resulting correlation coefficients were ordered according to hierarchical clustering. Strong correlations were identified between 16 of the 19 measured amino acids in the mixed model (Figure 3) Only aspartate, glutamate and proline did not demonstrate strong correlation coefficients to other amino acids. Proline, however, revealed mild negative correlations to the aromatic amino

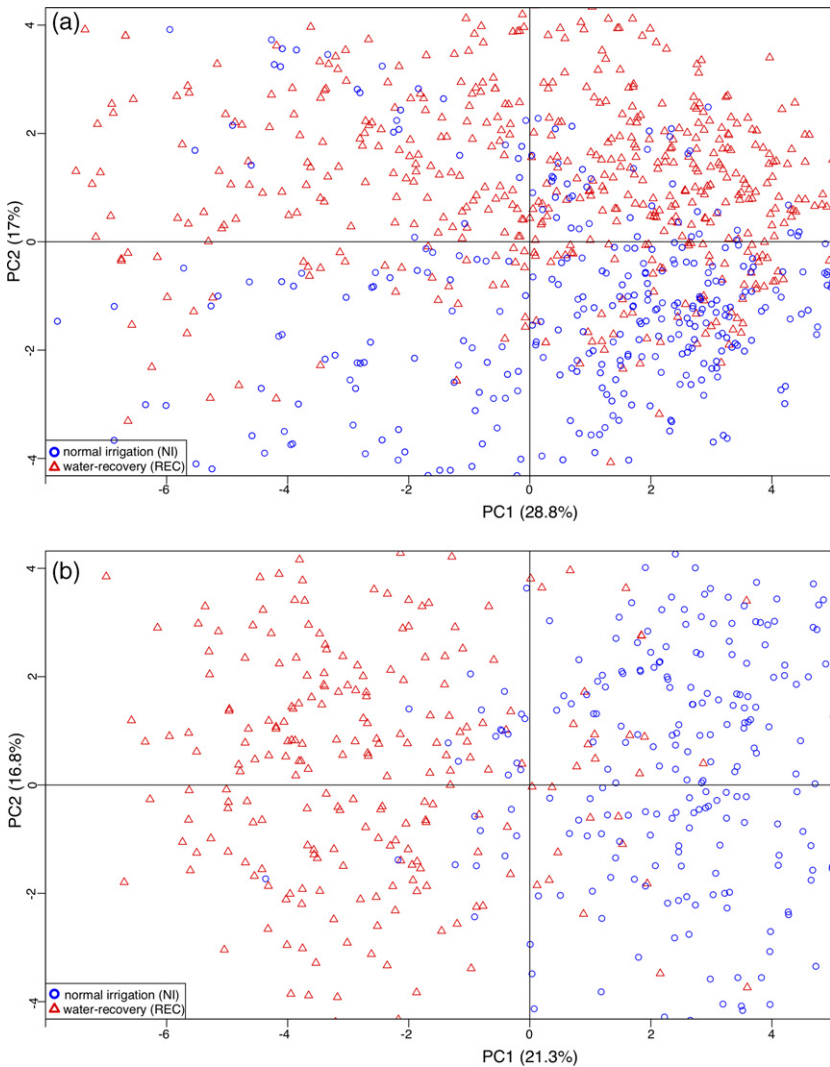


Figure 1. Principal component analysis. Principal component analysis of metabolite profiles (a) and the morphological traits (b) measured on the potato association panel. The field studies, normal irrigation (NI) and water recovery (REC) were carried out in Ica, Peru in 2017 in an augmented block design. Blue circles correspond to NI samples and red triangles to REC samples.

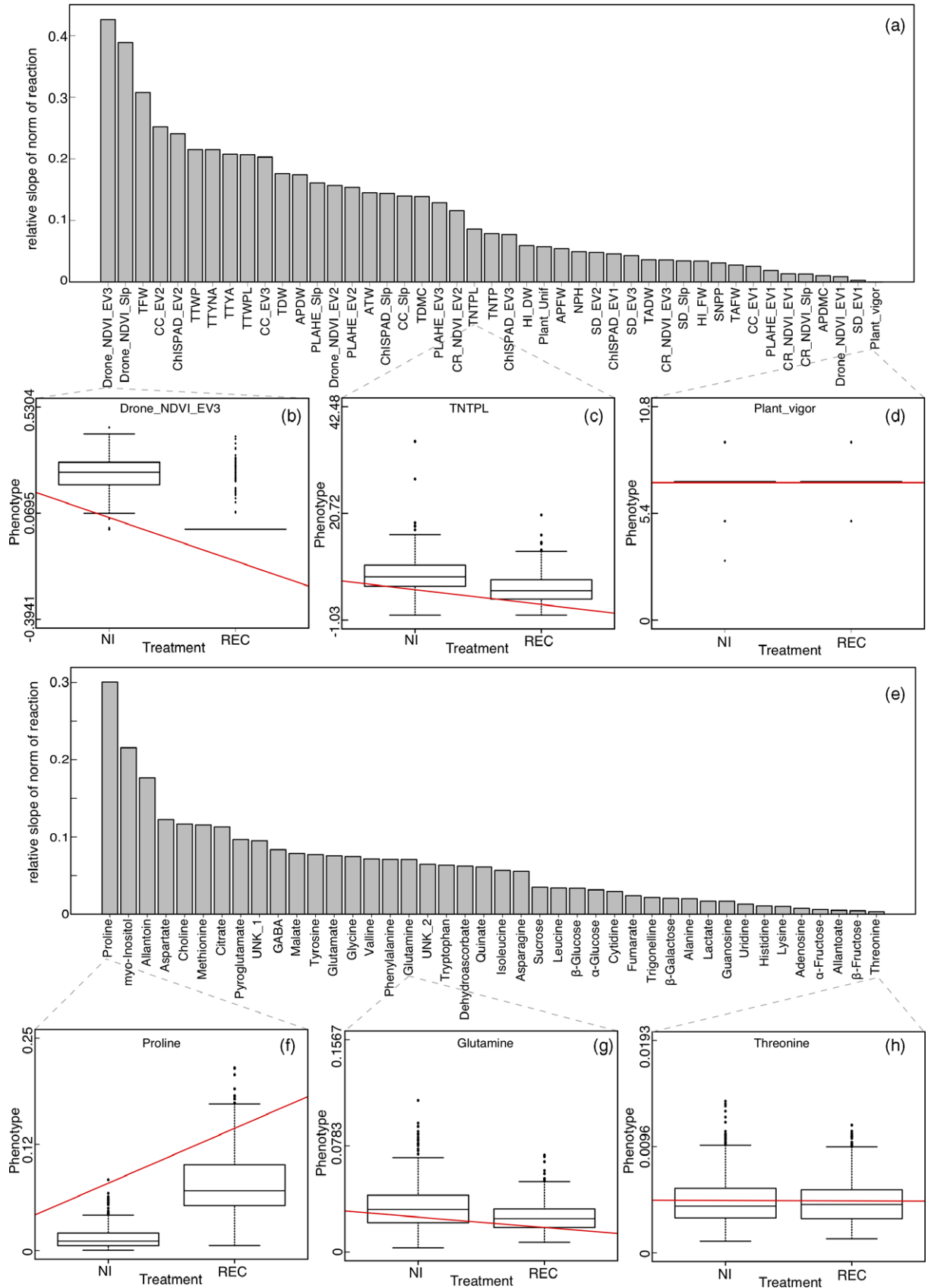
acids phenylalanine, tryptophan and tyrosine. The organic acids citrate, fumarate, choline and allantoate and unknown compound 1 (UNK1) demonstrated similar negative correlations to the same group of amino acids. In the

correlation analysis using the samples corresponding to the single conditions, the clustering of amino acids was less pronounced. For example, the NI correlation analysis revealed two groups of amino acids that correlated

Figure 2. Norm of reaction.

(a) Bar graph of the absolute slope values computed for all morphological traits. (b)–(d) Box plots of selected morphological traits for normal irrigation (NI) and water recovery (REC) treatments; red lines represent slope. The different traits were chosen to demonstrate different genotypic versus environment contributions to the observed variety. (e) Bar graph of the absolute slope value computed for all metabolites. (f)–(h) Box plots of selected metabolites for NI and REC treatments; red lines represent slope. The different metabolites were chosen to demonstrate different genotypic versus environment contributions on the observed variety.

The NI and REC samples of the potato association panel were grown in an augmented block design in Ica, Peru in 2017. APDMC, dry matter content of aerial part; APDW, dry weight of aerial part; APFW, fresh weight of aerial part; ATW, average tuber weight; CC_EV1/2/3, canopy cover evaluation 1/2/3; CC_Slp, canopy cover slope; ChisPAD_EV1/2/3, chlorophyll content index evaluation 1/2/3; ChisPAD_Slp, chlorophyll content index slope; CR_NDVI_EV1/2/3, canopy reflectance evaluation 1/2/3; CR_NDVI_Slp, canopy reflectance slope; DW, dry weight; Drone_NDVI_EV1/2/3, drone normalized difference vegetation index evaluation 1/2/3; Drone_NDVI_Slp, drone normalized difference vegetation index slope; HI_DW, harvest index dry weight; HI_FW, harvest index fresh weight; NPH, number of harvested plants; PLAHE_EV1/2/3, plant height evaluation 1/2/3; PLAHE_Slp, plant height slope; Plant_unif, plant uniformity; Plant_vigor, plant vigor; SD_EV1/2/3, stem diameter evaluation 1/2/3; SD_Slp, stem diameter slope; SNPP, stem number per plant; TADW, total aerial part dry weight; TAFW, total aerial part fresh weight; TDMC, tuber dry matter content per plot; TDW, tuber dry weight per plant; TFW, tuber fresh weight per plant; TNTP, total number of tubers per plot; TNTPL, total number of tubers per plant; TTWP, total tuber weight per plot; TTWPL, total tuber weight per plant; TTYA, total tuber yield adjusted; TTYNA, total tuber yield not adjusted.



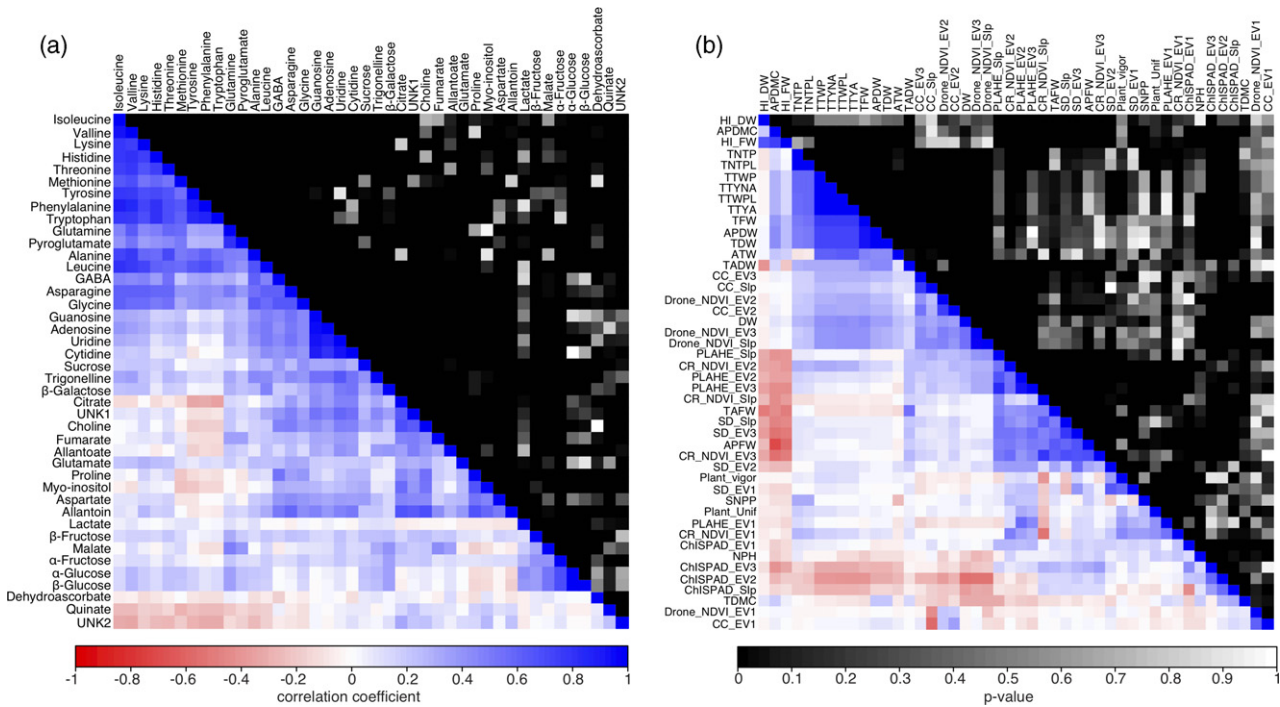


Figure 3. Heatmap of correlation analysis.

Heatmap representation of the correlation analysis of the 42 quantified metabolites (a) and the 46 morphological traits (b) across all normal irrigation and water recovery samples of the potato association panel grown in an augmented block design in Ica, Peru in 2017. The lower triangle illustrates the correlation coefficients, where a red rectangle represents a negative correlation and a blue rectangle a positive correlation. The upper triangle (shaded area) represents the corresponding *P*-values. Variables on the x and y axes are ordered as determined by hierarchical clustering.

APDMC, dry matter content of aerial part; APDW, dry weight of aerial part; APFW, fresh weight of aerial part; ATW, average tuber weight; CC_EV1/2/3, canopy cover evaluation 1/2/3; CC_Slp, canopy cover slope; ChisPAD_EV1/2/3, chlorophyll content index evaluation 1/2/3; ChisPAD_Slp, chlorophyll content index slope; CR_NDVI_EV1/2/3, canopy reflectance evaluation 1/2/3; CR_NDVI_Slp, canopy reflectance slope; DW, dry weight; Drone_NDVI_EV1/2/3, drone normalized difference vegetation index evaluation 1/2/3; Drone_NDVI_Slp, drone normalized difference vegetation index slope; HI_DW, harvest index dry weight; HI_FW, harvest index fresh weight; NPH, number of harvested plants; PLAHE_EV1/2/3, plant height evaluation 1/2/3; PLAHE_Slp, plant height slope; Plant_unif, plant uniformity; Plant_vigor, plant vigor; SD_EV1/2/3, stem diameter evaluation 1/2/3; SD_Slp, stem diameter slope; SNPP, stem number per plant; TADW, total aerial part dry weight; TAFW, total aerial part fresh weight; TDMC, tuber dry matter content per plot; TDW, tuber dry weight per plant; TFW, tuber fresh weight per plant; TNTP, total number of tubers per plot; TNTPL, total number of tubers per plant; TTWP, total tuber weight per plot; TTWPL, total tuber weight per plant; TTYA, total tuber yield adjusted; TTYNA, total tuber yield not adjusted.

strongly to each other; one group containing alanine, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, pyroglutamate, tryptophan, tyrosine and valine, and the other asparagine, GABA, glutamate and

glycine (Figure S3a). In the REC correlation analysis, four amino acid groups were identified pairing glutamine with pyroglutamate, asparagine with glycine and GABA, aspartate with proline, and the remaining amino acids

Figure 4. Correlation-based metabolite networks.

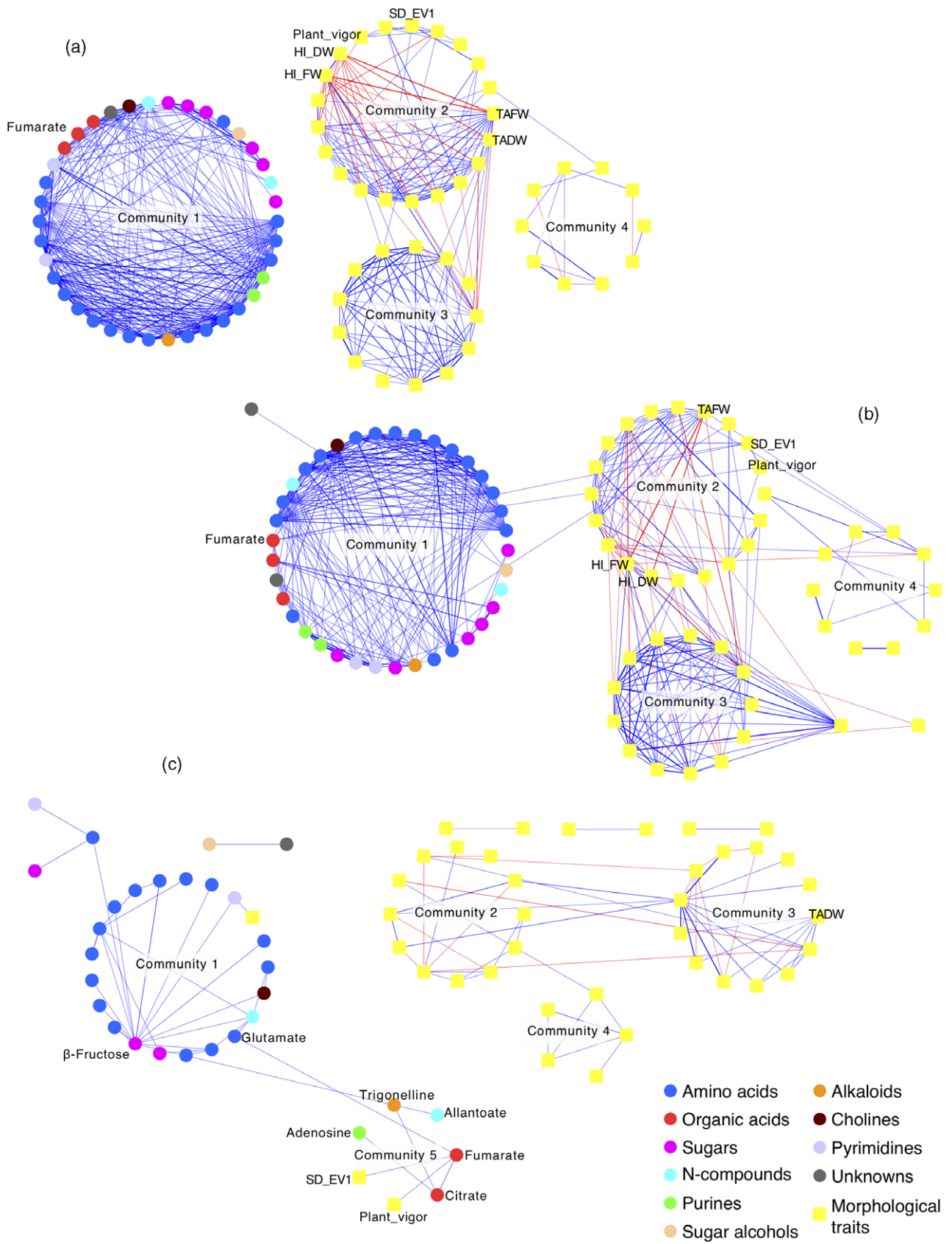
Network visualization of 42 metabolites and 46 morphological traits correlated to each other as measured in a potato association panel grown in an augmented block design in Ica, Peru in 2017. Metabolites are represented as elliptical nodes color coded by the compound class and morphological traits are represented as yellow, rectangular nodes. Edges between nodes represent significant correlation coefficients, where blue edges illustrate positive correlations and red edges illustrate negative correlations. Nodes are clustered into communities according to the walktrap community detecting algorithm.

(a) Network corresponding to normally irrigated (NI) samples.

(b) Network corresponding to water recovery (REC) samples.

(c) Difference network, where REC \subset NI.

APDMC, dry matter content of aerial part; APDW, dry weight of aerial part; APFW, fresh weight of aerial part; ATW, average tuber weight; CC_EV1/2/3, canopy cover evaluation 1/2/3; CC_Slp, canopy cover slope; ChisPAD_EV1/2/3, chlorophyll content index evaluation 1/2/3; ChisPAD_Slp, chlorophyll content index slope; CR_NDVI_EV1/2/3, canopy reflectance evaluation 1/2/3; CR_NDVI_Slp, canopy reflectance slope; DW, dry weight; Drone_NDVI_EV1/2/3, drone normalized difference vegetation index evaluation 1/2/3; Drone_NDVI_Slp, drone normalized difference vegetation index slope; HI_DW, harvest index dry weight; HI_FW, harvest index fresh weight; NPH, number of harvested plants; PLAHE_EV1/2/3, plant height evaluation 1/2/3; PLAHE_Slp, plant height slope; Plant_unif, plant uniformity; Plant_vigor, plant vigor; SD_EV1/2/3, stem diameter evaluation 1/2/3; SD_Slp, stem diameter slope; SNPP, stem number per plant; TADW, total aerial part dry weight; TAFW, total aerial part fresh weight; TDMC, tuber dry matter content per plot; TDW, tuber dry weight per plant; TFW, tuber fresh weight per plant; TNTP, total number of tubers per plot; TNTPL, total number of tubers per plant; TTWP, total tuber weight per plot; TTWPL, total tuber weight per plant; TTYA, total tuber yield adjusted; TTYNA, total tuber yield not adjusted.



(Figure S3b). Interestingly, isoleucine, methionine, phenylalanine, tryptophan, tyrosine and valine demonstrated strong negative correlations to proline and aspartate.

Three functional groups were identified for the correlation analyses of the morphological traits. Strong positive correlations were observed for traits associated with tuber yield in all three correlation analyses (Figures 3b and S4). Another group was centered on traits associated with canopy reflectance, plant height and stem diameter. These traits showed strong negative correlations to the harvest index traits. The third group contained traits second evaluation of chlorophyll content (CHISPAD_EV2), third evaluation of chlorophyll content (CHISPAD_EV3) and slope of chlorophyll content. While these traits were positively correlated to each other they were actually negatively correlated to tuber yield traits and canopy cover.

Network analysis reveals a correlation between fumarate and plant vigor

Relationships between heterogeneous plant traits may be captured by correlation-based networks (CN). Here, we employed a CN approach based on the combined datasets metabolite to metabolite, metabolite to morphological traits and morphological traits to morphological traits; associations were visualized by a graph of nodes and edges. We constructed two initial networks corresponding to NI and REC samples as described previously in Toubiana *et al.* (2013), where nodes represent compounds and the edges between them the significant correlation coefficients (r). Thresholds for both networks were set to $r \geq 0.4$ and a q -value ≤ 0.05 to remove spurious correlations. Consequently, the NI sample network was composed of 88 nodes and 194 edges (Figure 4a), while the REC sample network was composed of 90 nodes and 272 edges (Figure 4b). The greater number of edges in the water deficit stress network supported the notion of readiness for rewiring metabolic routes at the initiation of the abiotic stress, as also observed in grapevine leaves (Hochberg *et al.*, 2013) and tomato seeds (Rosental *et al.*, 2016).

Next, the nodes in the network were grouped into communities using the walktrap community algorithm (Pons and Latapy, 2005), identifying densely connected subgraphs (i.e. communities) based on short random walks (Figure 4). In both networks, the presence of four salient communities was detected, of which community 1 corresponded to the metabolites and communities 2–4 to the

morphological traits. This view clearly emphasized the separation of metabolites and morphological traits in the NI network, while highlighting relationships between a distinct group of metabolites and morphological traits in the REC network. Furthermore, the networks lucidly accentuated the positive nature of relationships between metabolites, suggestive of their genetically tightly coordinated regulation. Negative correlations, on the other hand, were detected for morphological traits only. In particular, the harvest index revealed significant negative correlations to other morphological traits. These results are in keeping with earlier findings showing negative correlations of harvest index to other morphological traits (Schauer *et al.*, 2006), suggesting a negative tradeoff between aerial parts of the plant and tuber development.

To highlight differences between the networks, a difference network was constructed (Figure 4c) containing the difference set of the REC network over the NI network (REC \setminus NI). The network illustrated the 20 edges that were present in the REC network but not in the NI network. Under this reduced network topography one additional community was suggested, containing the first evaluation of stem diameter (SD_EV1) and plant vigor linked to the following metabolites: adenosine, allantoate, citrate, fumarate and trigonelline. However, of all the compounds present in community 5 only fumarate revealed (positive) correlations to the morphological traits.

Gene Ontology enrichment analysis associates QTLs with metabolic processes

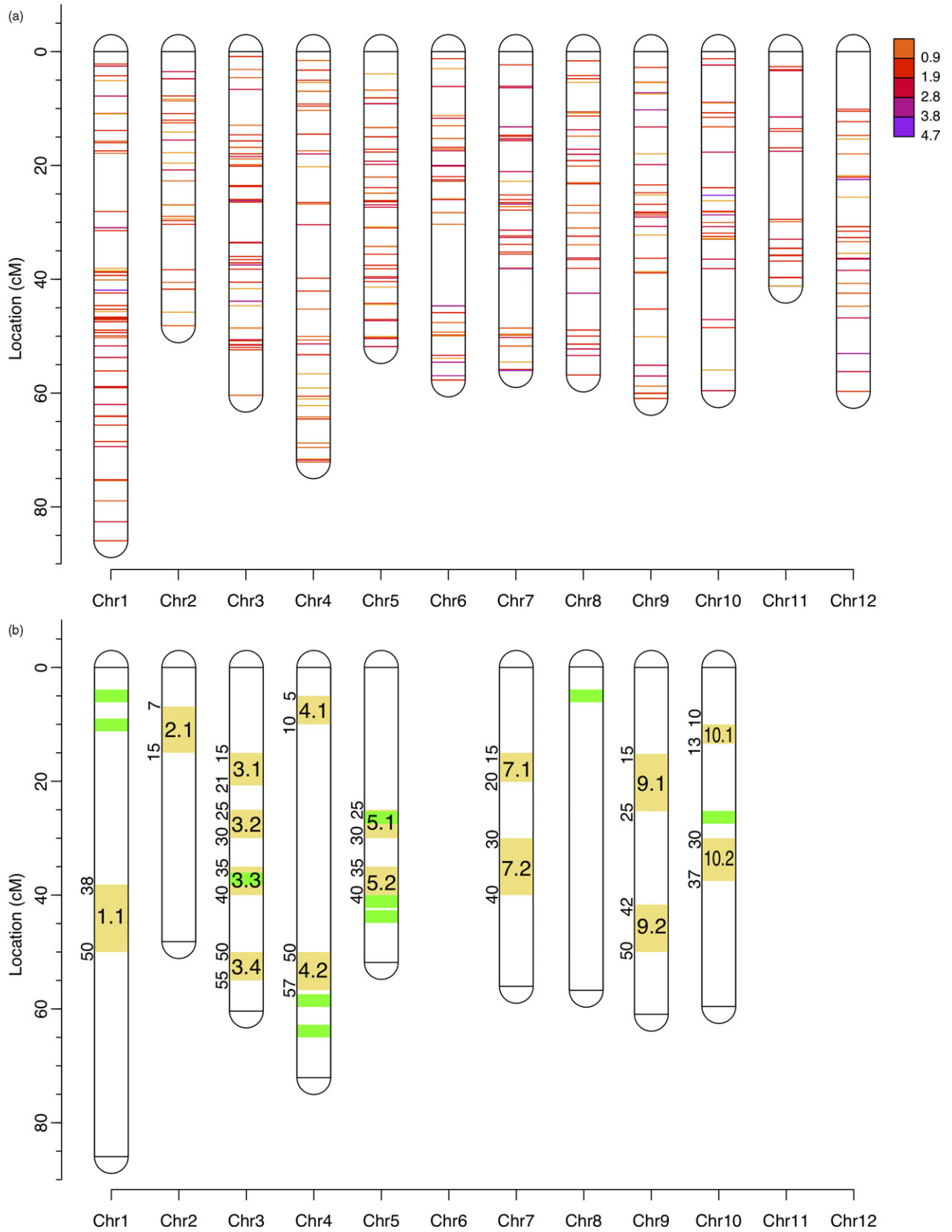
We explored the genetic basis of natural variation for all metabolites by a genome-wide association study (GWAS) using a marker set consisting of 29 408 bi-allelic single nucleotide polymorphisms (SNPs). For each metabolite a GWAS was performed individually. A false discovery rate was employed (Benjamini and Hochberg, 1995) to adjust for multiple hypothesis testing. For all metabolites, a total of 3532 significant SNPs (Data S3) were identified, corresponding to 2667 individual SNPs, and revealing QTL hotspots on chromosomes 1, 3–5, 7, 9 and 10 (Figure S5). Decreasing the P -value threshold to a stringent level of ≤ 0.005 reduced the number of highly significant markers to 844.

Based on the reduced number of SNPs we inspected hotspot regions for metabolite QTLs (Figure 5, Table 1) using the SPUD db genome browser for potato (Hirsch

Figure 5. Metabolite quantitative trait locus (QTL) map.

(a) Hotspot QTL map of metabolite QTL (at a P -value threshold ≤ 0.005) on the potato genome: color intensities correspond to P -values. High-density QTL regions (hotspots) show a higher density of colored bands.

(b) Gene Ontology (GO) term analysis of metabolite QTL hotspots: yellow highlighted regions within chromosomes correspond to high-density QTL regions for which GO term enrichment analysis was performed. Values next to chromosomes correspond to hotspot limits in cM. Hotspot regions are illustrated chronologically for each chromosome. The most significant GO terms are listed in Table 1. Green highlighted regions represent 2 Mb windows around single nucleotide polymorphisms associated with fumarate.



et al., 2014). For each QTL hotspot region we defined chromosome segments containing all corresponding SNPs. Subsequently, all genes embedded within each segment were identified. To provide a biological perspective for the identified genes, we performed Gene Ontology (GO) term enrichment analysis (Table 1). The GO initiative was developed to provide a system in which sets of genes can be classified hierarchically in a graph-like structure (Harris et al., 2008). The GO term analysis revealed that, in particular, segments on chromosomes 1, 2, 5 and 7 were enriched with genes associated with metabolic processes. Segments on chromosomes 3 and 10 contained genes integral to membranes. Segments on chromosomes 3 and 4 showed monooxygenase and auxin response activities, respectively.

To understand the genetic relationship identified via network analysis, we specifically investigated SNPs associated with fumarate. A total of 11 markers were identified (Figures 5b and 6a) on chromosomes 1, 3, 4, 5, 8 and 10. Using the SPUD db genome browser, we investigated the genomic regions of all 11 SNPs. Eight SNPs were located in non-coding regions (Figure 6b) with the remaining three SNPs being located within genes, namely: (i) the SNP at position 5 096 024 on chromosome 1 was located within a gene coding for a protein containing a RING domain, (ii) the SNP at position 59 109 325 on chromosome 4 was located within a gene coding for a zinc finger protein (ZAT2) and (iii) the SNP at position 41 382 562 on chromosome 5 was located within a gene coding for a protein of

unknown function. All three SNPs were located within intron regions (Figure 6b). Furthermore, based on the short-distance LD threshold (see Experimental Procedures for details), we inspected 2 Mb windows around the 11 SNPs (Figure 5b). A total of 1190 genes were identified in those chromosomal regions (Data S4). Also, here a GO term enrichment analysis was performed, revealing that the identified genes were mainly associated with small molecule metabolic processes ($P = 0.00091$) and carboxylic acid metabolic processes ($P = 0.033$).

Phylogenetic analysis provides eco-physiological insights

In an effort to explore the eco-physiological relationships of potato grown under REC treatment, we employed phylogenetic analysis of the RING domain and the zinc finger ZAT2 protein identified via GWAS. The protein sequences of both proteins were blasted against the NCBI database and the 49 sequences with the highest percentage sequence identity were used for further analysis. After multiple sequence alignment, phylogenetic trees were generated (Figure 7). The protein sequence of the RING domain protein sequence revealed high similarity (>99%) to homologs in two tomato species (*Solanum lycopersicum* and *Solanum pennellii*) and to two homologs in a pepper species (*Capsicum annuum*). Their close relationship was also visible in the phylogenetic tree as they were clustered together (Figure 7a). The phylogenetic analysis of the zinc finger protein showed significant clustering of five homologs in tobacco plants (*Nicotinia tabacum*, *Nicotinia*

Table 1 Gene Ontology (GO) term analysis of quantitative trait locus hotspots in Figure 5(b)

Hotspot	Segment	No. of genes	No. of genes with unknown function	Most significant GO terms
1.1	38–50 cM	371	188	tRNA processing, tRNA metabolic process, ncRNA processing, ncRNA metabolic processing
2.1	7–15 cM	285	130	Regulation of primary metabolic process, regulation of metabolic process, small molecule metabolic process
3.1	15–21 cM	206	103	Developmental process, single-organism developmental process
3.2	25–30 cM	160	105	Integral component of membrane, intrinsic component of membrane
3.3	35–40 cM	332	104	Peptidyl-lysine methylation, histone methylation, regulation of molecular function
3.4	50–55 cM	759	170	Monooxygenase activity, enzyme inhibitor activity, endopeptidase regulator activity
4.1	5–10 cM	542	126	Cell redox homeostasis, cellular homeostasis, homeostatic process
4.2	50–57 cM	484	160	Response to auxin, response to hormone, response to endogenous stimulus
5.1	25–30 cM	158	102	Primary metabolic process, cellular macromolecular metabolic process, cellular process
5.2	35–40 cM	148	95	Metabolic process, binding
7.1	15–20 cM	179	106	Primary metabolic process, cellular process, cellular metabolic process
7.2	30–40 cM	364	180	RNA biosynthetic process, phosphorylation, cellular biosynthetic process, nitrogen compound metabolic process
9.1	15–25 cM	288	174	Protein amino acid phosphorylation, phosphorylation
9.2	30–35 cM	204	107	Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process
9.3	42–50 cM	156	500	Protein binding
10.1	10–13 cM	90	62	NA
10.1	30–37 cM	309	163	Intrinsic to membrane, integral to membrane, membrane part

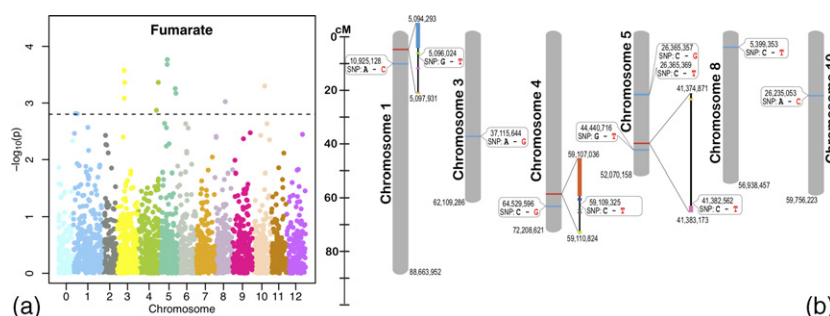


Figure 6. Single nucleotide polymorphisms (SNPs) corresponding to fumarate.

(a) Manhattan plot corresponding to a genome-wide association study (GWAS) of fumarate as generated by the software GWASPoly. The 12 chromosomes of potato are represented by different colors. Undetermined SNPs are represented by chromosome 0. The dashed line represents the significance cut-off using a false discovery rate.

(b) Illustrating the SNPs that were associated with fumarate and plant vigor on chromosomes 1, 3, 4, 5, 8 and 10. The SNPs located within non-coding regions of chromosomes are depicted in blue. Genes that contained SNPs are illustrated in red and are magnified (chromosomes 1, 4 and 5). Colored, bulky areas of magnified genes represent exons, black areas represent introns. Speech bubbles indicate the exact position of SNPs showing the reference nucleotide (black letter) and the substitution nucleotide (red letter). Numerical values represent base pair positions.

sylvestris, *Nicotinia attenuata* and *Nicotinia tomentosiformis*) and to five homologs in pepper plants (*C. annuum*, *Capsicum baccatum* and *Capsicum chinense*; Figure 7b).

DISCUSSION

Potato is the world's third most important food crop, essential for sustaining the demands of an increasing human population (Birch *et al.*, 2012). However, potato is a cool-season crop, with optimum yields being obtained at mild temperatures (Haverkort and Struik, 2015). Increased temperatures perturb starch synthesis (Geigenberger *et al.*, 1998) in the potato tuber, resulting in decreased yields as evidenced in sub-Saharan countries compared with countries with a mild climate (Birch *et al.*, 2012). Current climate models predict a mean temperature increase of 1.1–6.4°C (Christensen *et al.*, 2007), which, in turn, will pose problems for maintaining steady potato yields in the future. Solutions to counteract this trend are necessary.

In the current study, the profiling of metabolites and morphological traits across a tropical climate-adapted potato association panel was implemented and subjected to network analysis and GWAS. Descriptive statistics and the analysis of the norm of reaction (Figures 2, S1 and S2) revealed that the observed variance was trait specific, i.e. for some traits the observed difference was genotypic while for others it was environmental, while again for others the observed phenotypic range was due to both genotypic and environment contributions. The PCA of the morphological traits revealed a clear separation between treatments on PC1 (Figure 1b), accounting for 21.3% of the observed variance. A similar picture arose for the PCA of the metabolites (Figure 1a), where a separation of treatments was visible on PC2. However, here also a significant overlap between NI and REC samples was visible. The less pronounced effect of treatments on tuber metabolism may be attributed to its nature as a sink organ. Metabolism in

the tuber is inherently associated with starch synthesis (Nazarian-Firouzabadi and Visser, 2017). In the phototrophic leaf tissues starch is synthesized during the day and degraded at night to supply the energy needed for biological processes. In the heterotrophic tuber, on the other hand, starch is synthesized for long-term storage. During tuber development the accumulated starch is used to maintain the reduced energy needs of the dormant tuber and, once dormancy is broken, it fuels the outgrowth of new shoots. The need for the tuber to initiate the next life cycle may be the ecological factor dictating the development of an energetically robust but flexible sink organ. Therefore, the tuber appears to be more drought resistant. We believe that further investigations are necessary to explore the underlying mechanisms that allow tuber metabolism to be less affected by water stress.

Strong correlations were identified between 16 of the 19 measured amino acids in the mixed model (Figure 3) – an observation that has been reported in various plants and their respective tissues, under normal and stress conditions (Toubiana *et al.*, 2012; Hochberg *et al.*, 2013; Toubiana *et al.*, 2016; Maruenda *et al.*, 2018). Correlation-based network analysis (CNA) emphasized the separation of metabolites and morphological traits. The REC network showed a higher network connectivity than the NI network, supporting the notion of an increased readiness to rewire the metabolic network in response to abiotic stress. Similar observations have also been made in grapevine leaves (Hochberg *et al.*, 2013) and tomato seeds (Rosental *et al.*, 2016). Community detection in the NI and REC networks identified four communities, with community 1 corresponding to metabolites and communities 2–4 to morphological traits (Figure 4). Nevertheless, while the metabolite community in the NI network was completely separated (Figure 4a) from the morphological trait communities, some connections between metabolites and morphological

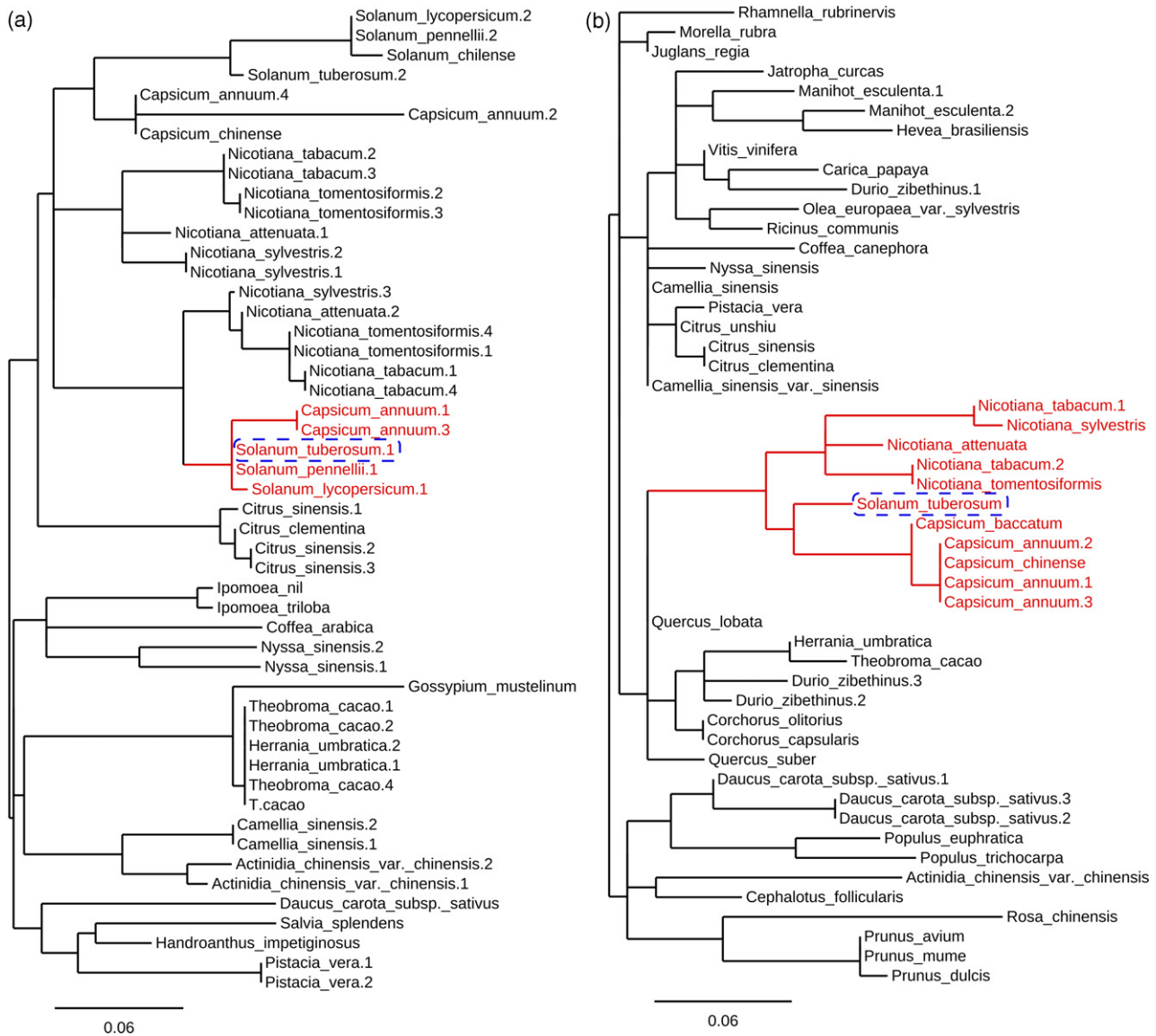


Figure 7. Phylogenetic trees.

(a) Phylogenetic tree of the RING domain gene and the 49 genes with the greatest sequence identity. The cluster containing the reference sequence in *Solanum tuberosum* is highlighted in red. The reference gene is highlighted within the blue dashed box. Branch annotations are species names.

(b) Phylogenetic tree of zinc finger gene and 49 genes with the greatest sequence identity. The cluster containing the reference sequence in *S. tuberosum* is highlighted in red. The reference gene is highlighted within the blue dashed box. Branch annotations are species names.

traits could be found in the REC network (Figure 4b). To highlight the differences between the NI and REC networks, a difference network was constructed subtracting the connections of the NI network from the REC network (Figure 4c). This network constellation generated a new community, where the connections between the traits plant vigor and the first evaluation of the stem diameter and fumarate was stressed. Fumarate is an intermediate of the TCA cycle but also occurs in its oxidized form as the outcome of complex II of the mitochondrial electron transport chain. The TCA cycle, the mitochondrial electron transport chain and glycolysis are inherently related as part

of the respiratory pathway complex (Fernie *et al.*, 2004). As a heterotrophic tissue, the tuber depends on the transport of sucrose via the phloem from active leaves. In the tuber, sucrose is used to produce glucose-6-phosphate, which is then transported into the amyloplasts for subsequent starch synthesis. Glucose-6-phosphate is an intermediate of glycolysis. The connection between fumarate in the tuber and the morphological traits plant vigor and stem diameter may reflect the complex biochemical mechanisms involved in tuber development with respect to plant performance. Fumarate and other organic acids often accumulate in sink organs to maintain intracellular ionic

balance and nutrient uptake to sustain drought stress (Guo *et al.*, 2018). Fumarate and other organic acids often act as compatible solutes involved in osmotic adjustment and membrane and protein protection from reactive oxygen species (ROS). In wheat seedlings (Guo *et al.*, 2018) and soybean (Silvente *et al.*, 2012) fumarate was suggested as a metabolic marker for drought response.

Finally, we used the metabolic trait profiles in a GWAS to identify genomic regions (loci) associated with metabolism and potentially identify candidate genes for drought stress regulation. We note that the GWAS was performed for one season only, therefore its results should be considered with caution. The study was carried out using different gene action models as provided by the software GWASpoly (Rosyara *et al.*, 2016). Using a false discovery rate to account for multiple hypothesis testing (q -value of ≤ 0.05) we recorded for all metabolites a total of 3532 significant SNPs corresponding to 2667 individual SNPs. Investigating the distribution of loci associated with the identified markers revealed a high dispersion across the potato genome. Central metabolism is characterized by a complex multilevel regulation, futile cycles and enzyme promiscuity, and redundant pathways in multiple cellular compartments (Collakova *et al.*, 2012) – all of which contribute to metabolic stability (Rontein *et al.*, 2002) and which probably affect the ability to measure trait heritability. Furthermore, population structure and cryptic relatedness of many plant populations targeted for GWAS often lead to weak associations between a trait and its prospective locus (Astle and Balding, 2009). It is therefore not surprising that this and other studies were either unsuccessful or just identified indistinct links between metabolite content and its underlying genetics.

When increasing the stringency level for the GWAS results (P -value ≤ 0.005), the number of markers reduced significantly to 844 (Figure 5). The pronounced decrease of markers associated with metabolites at an increased stringency level hinted at tight gene-to-trait linkage at specific QTLs. The GO term analysis of metabolite QTL hotspots pinpointed chromosome segments enriched with genes associated with the regulation of metabolic processes (Figure 5, Table 1). In addition, segments on chromosomes 3 and 4 were enriched with genes with monooxygenase and auxin response activities. Overexpression of spinach monooxygenases in rice and tobacco increased their tolerance to salt, temperature and drought stress, respectively (Shirasawa *et al.*, 2006; Duan *et al.*, 2017). A recent review discussed the role of auxin in modifying root system architecture upon salt stress and water deficit (Korver *et al.*, 2018).

Intersecting the results of the CNA and the GWAS, our analysis highlighted the relationship of plant vigor to fumarate under REC conditions. We then investigated the 11 markers identified for fumarate, and based on the

estimated short-distance LD also inspected 2 Mb windows around the SNPs. A total of 1190 genes (Data S4) were identified in fumarate-associated chromosomal regions. The GO analysis of these genes showed enrichment for metabolic and carboxylic processes, which can be attributed to fumarate being an intermediate in the TCA cycle, the mitochondrial electron transport chain and its aforementioned functions associated with stress response. Eight of the 11 SNPs were located in non-coding regions and three SNPs were identified in introns of three genes (Figure 7): the gene on chromosome 1 coded for protein containing the RING finger domain, the gene on chromosome 4 coded for the zinc finger protein ZAT2, while the function of the gene identified on chromosome 5 is still undetermined.

Proteins containing RING finger motifs are known to play a significant role in the regulation of protein degradation via the ubiquitin–proteasome system (Joazeiro and Weissman, 2000). In maize it was shown that overexpression of the *ZmXerico1* and *ZmXerico2* RING finger domain-containing genes increases water deficit tolerance by controlling ABA homeostasis (Brugiére *et al.*, 2017). Also, in rice, the overexpression of the *OsRDGP1* RING domain-containing gene increased the tolerance to water deficit (Bae *et al.*, 2011). Zinc finger proteins have previously been associated with drought stress. For example, in Arabidopsis it was shown that the constitutive expression of *Zat10* enhanced the expression of ROS response transcripts (Mittler *et al.*, 2006) and in soybean that the C2H2-type zinc finger protein *GmZFP3* negatively regulated drought stress (Zhang *et al.*, 2016). Zinc finger proteins have also been associated with salt stress; for example, in Arabidopsis, Zang *et al.* (2016) demonstrated that overexpression of *GmZFP3* regulated sodium and potassium homeostasis, ROS scavenging and osmotic potential. The fact that the SNPs were identified in introns and not in exon regions may be due to two reasons: (i) SNPs are often only representatives of traits and it is likely that other close SNPs in high LD are causal for the observed difference, and (ii) SNPs cause differences in gene expression levels rather than modifications in protein formation (Tak and Farnham, 2015). To put the results into an eco-physiological perspective, we ran phylogenetic analyses for the RING domain and the zinc finger gene (Figure 7). For the RING domain protein a close relationship was found to tomato (*Lycopersicon*) and pepper (*Capsicum*) species – both members of the Solanaceae family. It was shown before that *S. pennellii* – one of the tomato species identified here – has developed mechanisms for drought tolerance (Egea *et al.*, 2018). Moreover, analysis of its genome has identified 389 stress-related genes (Bolger *et al.*, 2014). For the zinc finger domain a close relationship to tobacco (*Nicotiana*) and pepper species was observed. Also the genus *Nicotiana* belongs to the Solanaceae family. Particularly in *C. annuum*, abiotic stresses such as heat, drought,

cold and salinity were shown to induce the transcription of *CaRZFP1*, which contains an open reading frame coding for a RING zinc finger protein (Zeba *et al.*, 2006). Due to the close relationship of potato to other species in the Solanaceae, their well-studied physiological mechanisms to drought response and the known functions of the RING domain and the zinc finger genes, it may be hypothesized that these genes exert a similar function in potato.

CONCLUSIONS

In the current study we assessed the effects of water-recovery treatment on the plant performance of a potato association panel and on the central metabolism of the tuber. By applying advanced correlation-based network analysis, we studied the relational ties between morphological traits and tuber metabolism in order to identify potential metabolic markers for drought resistance. A difference network highlighted the differential connections of the REC versus the NI treatment. Specifically, the connection between plant vigor and fumarate was highlighted. Based on known metabolic pathways and their intertwined relationships, our findings suggest that fumarate may be used as a metabolic marker for drought stress in the potato tuber, especially with respect to plant vigor. Genes identified in QTLs associated with fumarate support our suggestion. Our findings may have important implications for future potato breeding regimes, especially in countries already suffering from the effects of the changing climate.

EXPERIMENTAL PROCEDURES

Population and field trial

The potato association panel used in this study comprised 258 genotypes representing six main CIP breeding populations as well as a group of varieties with different origins (Data S5). Population 'A' in Data S5 was developed during the period 1980–1990 emphasizing late blight resistance, derived from global breeding programs based on species the *Solanum demissum*, native Andean cultivars of *S. tuberosum*, the wild species *Solanum acaule* and *Solanum bulbocastanum*. Population 'B1' is a *S. tuberosum* group andigena derivative. Population 'B3' genotypes are population 'A' derivatives with emphasis on increasing frequencies and levels of quantitative resistance to late blight. Population 'LTVR' was developed with an emphasis on viral disease resistance (PVY, PVX and PLRV), short crop duration and adaptation to warm environments. The 'B3-LTVR' population incorporates hybrid genotypes originating from crosses between populations 'B3' and 'LTVR'. The 'LB-HT' population combines bred varieties for heat tolerance from Europe and North America, varieties for blight resistance from the 'B3' population and varieties from the 'LTVR' population. The group of varieties is composed of a group of potato varieties or key breeding lines: Desiree and Tomasa Condemayta.

All 258 genotypes of the potato association panel were grown in a field trial in Ica, Peru. The field trial was performed with triplicates in an augmented block design, with (i) control conditions (NI) and (ii) water-deficit conditions with recovery (REC), as described in the handbook for evaluation and data management of potato clones (Cabello *et al.*, 2017). In brief, a drought period

was initiated for the REC treatment when approximately 5% of plants had flower buds, that is, approximately 40–45 days after planting (DAP). After that, the moisture level was maintained at -4 bars = -0.4 MPa of the field capacity level until the end of the field trial (below field capacity but not above wilting point = -15 bars or -1.5 MPa), as opposed to -0.033 MPa in the NI treatment.

Not all genotypes survived the field trials. A total of 1195 samples were collected, of which 583 corresponded to NI samples and 612 to REC samples. The NI samples were represented by 194 different genotypes (out of 258), while REC samples comprised 204 genotypes (out of 258).

Linkage disequilibrium

The LD decay estimate was performed essentially as described in Lindqvist-Kreuzer *et al.* (2020). In brief, the tetraploid marker set was used to estimate LD decay. Short- and long-distance LD decay were estimated based on the intersection of the significance threshold of $r^2 = 0.1$ and the fitted spline at the 90th percentile. The short-distance threshold was computed as 2 Mb, while the long-distance one was estimated at 5.5 Mb. For the short-distance LD decay, an $r^2_{1/2\max, 90}$ value of 0.55 Mb was calculated, similar to what has been reported in European potato varieties (Vos *et al.*, 2017). The LD decay was estimated as moderate.

Morphological traits

The complex phenotypic response of potato plants to REC conditions is the manifestation of the interactive effects of the plant's genotypic potential, the developmental stage and the environment. The 45 morphological traits evaluated in the current study were chosen as they have been shown to be affected by REC conditions (van Loon, 1981; Cabello *et al.*, 2013; Obidiegwu *et al.*, 2015). All plant morphological traits of all genotypes were measured at different time points throughout the field trial according to the procedures and protocols described in Cabello *et al.* (2017). The traits were measured in bulk for each plot. Their respective measurements were processed as described in Cabello *et al.* (2017) for further analysis. In brief, some traits were measured during three evaluation periods, namely 41–45 DAP, 63–67 DAP and 84–88 DAP; other traits were measured only once. To assess changes of the traits measured during the different evaluation periods a trend was estimated and expressed as the slope of the trend line. The traits corresponded to two sections of the plant, that is, the above-ground aerial part and the below-ground part of the plant. In particular, traits above ground were associated with canopy cover, plant height, plant uniformity, plant vigor, stem diameter, harvest index (HI), several measurements associated with chlorophyll content and others. Below-ground measurements were specifically associated with tuber morphology, such as tuber number, tuber weight (fresh and dry), tuber dry matter and tuber yield per plant and plot. A list of all traits and their respective definitions can be found in Data S6 and at <http://www.cropontology.org/>.

Nuclear magnetic resonance metabolite profiling

Nuclear magnetic resonance metabolite profiling of the tubers was performed at the end of the field trials.

Sample preparation. Four to eight tubers, mechanically peeled and sliced, were immediately frozen in liquid nitrogen and stored in an ultra-freezer at -74°C until lyophilization. The dry material was ground to a fine powder using an electric grinder and stored at -74°C until further use. The extracts were prepared following a procedure reported earlier (Maruenda *et al.*, 2018), with minor

modifications. One hundred and sixty milligrams of ground tuber was extracted with 1.6 ml of deionized water under sonication for 45 min at 10°C. The suspension was centrifuged at 23 000 *g* for 20 min at 10°C and the supernatant evaporated to dryness in a refrigerated centrifugal vacuum concentrator for 16 h at 10°C. The solid obtained (20–25 mg) was resuspended in 100 μ l of sodium oxalate buffer pH 4 (0.9 ml) and the solution was evaporated to dryness for 16 h at 10°C. The NMR sample was prepared by re-dissolving the recovered solid in 0.9 ml of deuterium oxide containing 3 mM 3-(trimethylsilyl)-propionic-*d*₄ (TSP), centrifuged at 23 000 *g* for 5 min at 10°C, and filtered through a 13 mm polytetrafluoroethylene 0.45 μ m syringe filter. Three independent extractions of each potato sample were performed.

NMR analyses. All NMR experiments were recorded at 25°C on a Bruker Avance III HD 500 MHz spectrometer equipped with a TCI Cryoprobe, auto-sampler and automatic tuning and matching (<https://www.bruker.com/>). All ¹H NMR spectra were acquired with a 30° flip angle, 5 Hz field strength water pre-saturation, 384 scans, 64 k data points, a spectral width of 10 kHz and a total relaxation time of 5.28 s. The flip angle was calibrated for each individual sample using 360° pulse optimization. Spectra were Fourier transformed with 0.3 Hz exponential apodization, and phase and baseline corrected using Topspin 3.5 pl5 software.

Spectral assignment was performed by standard procedures using ¹H-¹H DQF-COSY, ¹H-¹H TOCSY, ¹H *J*-resolved, ¹H-¹³C HSQC, ¹H-¹³C HSQC-TOCSY and ¹H-¹³C HMBC experiments. In the cases of dehydroascorbic and choline, their presence was confirmed by standard addition.

Metabolite quantification

All potato tuber spectra were submitted to a high-performance cluster (HPC) called LEGION (<http://legion.pucp.edu.pe/>) using a custom-made script. The script used the software package Bayesian Automated Metabolite Analyzer for NMR spectra (BATMAN) (Hao *et al.*, 2014), which was developed to quantify metabolites in NMR spectra. BATMAN uses a Markov chain Monte Carlo (MCMC) method to calculate the numerical approximation of the multi-dimensional integrals embedded in the NMR spectra. The MCMC method deconvolutes the spectra to identify and quantify individual metabolites. The BATMAN MCMC is divided into two phases. One is the burn-in phase, which allows the algorithm to find a region of high posterior probability. Once the posterior region has been identified the second phase, a posterior burning phase, is initiated, during which the actual approximations of metabolites is performed. We made adequate modifications to the burn-in and posterior-burn phases to identify and quantify 42 metabolites in the spectra of the potato tuber samples.

For the metabolites α -glucose, β -glucose, β -galactose, α -fructose, β -fructose, adenosine, alanine, asparagine, choline, fumaric acid, GABA, glutamine, guanosine, histidine, isoleucine, leucine, lysine, malic acid, methionine, myo-inositol, phenylalanine, pyroglutamic acid, proline, quinic acid, sucrose, threonine, tryptophan, tyrosine and valine, the BATMAN templates were generated from spectra of pure compounds acquired under the same experimental conditions as the potato tuber samples (buffer, pH and temperature). For the metabolites allantoic acid, allantoin, aspartic acid, citric acid, cytidine, dehydroascorbic acid glutamic acid, glycine, lactic acid, trigonelline, uridine, UNK1, UNK2 and reference TSP, which showed different signal patterns than in the pure compound spectra, due to either complex multiplicities, second-order *J* couplings, relaxation effects and/or

exchange processes, virtual spectra were generated for each compound using signal deconvolution in the tuber spectra. These virtual spectra were then used as templates for BATMAN quantification. The deconvolution was performed using the global spectra deconvolution (GSD) algorithm from MestreNova 11.0 software (Willcott, 2009).

A post-performance BATMAN script was developed to organize the HPC outcome into an easy to process format. BATMAN and in-house scripts were run under R (R Development Core Team, 2009). The obtained relative concentrations determined by BATMAN were normalized against TSP (3 mM) to retrieve absolute concentrations. The results were expressed in millimoles of metabolite per gram of dry potato powder (Data S2). The 1195 metabolite profiles contained 4.2% missing data, which were imputed employing PCA (Stacklies *et al.*, 2007). For subsequent analyses all metabolite data were log transformed.

Statistics

Descriptive statistics were calculated for both datasets (metabolite and morphological), followed by ANOVA. Subsequently, PCA, the norm of reaction and correlation analysis were performed. For the norm of reaction, linear regression models were generated for each trait, computing the intercept and slope variables.

Network analysis

Generation of the network was based on correlation analysis of all metabolites and morphological traits. The Pearson correlation was chosen to estimate correlation coefficients. To construct networks, first the corresponding *P*-value threshold of the correlation coefficient at a *q*-value of 0.05 was determined. Second, the adequate correlation coefficient threshold was chosen via stability tests of four different network properties, that is, average node degree, clustering coefficient, network density and diameter across a range of *P*-values. Next, the networks were clustered into communities by employing the walktrap community algorithm (Pons and Latapy, 2005). To detect differences between networks a difference network was generated according to set theory, REC $\not\subset$ NI.

All computations for network visualizations were generated in R (R Development Core Team, 2009). The software Cytoscape (Shannon *et al.*, 2003), version 3.0.0, was used for network visualization. Network properties and communities were computed using the R package igraph (Gabor and Tamas, 2006).

Genotyping and GWAS

The association panel was genotyped using genotyping by sequencing (GBS), and the SNPs were called and filtered as described in Lindqvist-Kreuzer *et al.* (2020).

After filtering, the final SNP dataset consisted of 29 408 bi-allelic SNPs. The population structure matrix *Q* and the kinship matrix *K* were combined for subsequent GWAS in GWASpoly (Rosyara *et al.*, 2016). The software allows testing for different models of gene action using polyploid or diploid data. With the normalized data for metabolites, a GWAS with 29 408 SNPs was performed for each metabolite. All different models (general, additive, dominant) available for diploids were used to identify QTLs. A false discovery rate at a *q*-value of ≤ 0.05 was applied to account for multiple hypothesis testing and identify significant markers.

Genomic position and GO term analysis

Quantitative trait loci identified via GWAS for metabolites were selected based on a stringent *P*-value of ≤ 0.005 . Association

mapping was employed to identify hotspots of QTLs within the potato genome. Hotspot regions were defined as chromosomal segments with a higher gene density. Using the SPUD db genome browser for potato (<http://solanaceae.plantbiology.msu.edu/cgi-bin/gbrowse/potato/>) (Hirsch *et al.*, 2014) we identified all genes within hotspot regions and performed GO term enrichment analysis. For the GO term enrichment analysis we used Agrigo (<http://bioinfo.cau.edu.cn/agriGO/>) (Du *et al.*, 2010; Tian *et al.*, 2017). We performed a singular enrichment analysis against the Gramene release 50 database.

Phylogenetic analysis

Protein sequences of genes PGSC0003DMP400039353 (ring finger domain) and PGSC0003DMP400001517 (zinc finger protein ZAT2) were blasted (blastp: protein-protein-BLAST) against the National Center for Biotechnology Information database (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). Protein sequences for the 49 genes with the greatest percentage identity for each gene were downloaded and subjected to phylogenetic analysis, which was performed at the Methods et Algorithmes pour la Bioinformatique LIRMM pipeline (http://www.phylogeny.fr/simple_phylogeny.cgi) (Dereeper *et al.*, 2008; Dereeper *et al.*, 2010). For multiple sequence alignment the pipeline implements MUSCLE alignment (Edgar, 2004), applies Gblocks to eliminate poorly aligned positions (Castresana, 2000) and uses an approximate likelihood ratio test to evaluate branches in the tree (Anisimova and Gascuel, 2009).

ACKNOWLEDGEMENTS

DT and GFdS are indebted to Cienciactiva-CONCYTEC and the Instituto Interamericano de Cooperación para la Agricultura (IICA) for supporting their stay at CERMN-PUCP. Our thanks to Dr Meredith Bonierbale (CIP) for her brief, but determining, support at the beginning of this project. This study was supported by the Programa Atracción de Investigadores Cienciactiva-CONCYTEC (008-2017-FONDECYT) and by Dirección de Gestión de la Investigación – PUCP (FONCAI-0023-2018). The research at CIP was undertaken as part of the CGIAR Research Program on Roots, Tubers and Bananas (RTB) and funding support for this work was provided by GIZ within the project Accelerating the Development of Early-Maturing-Agile Potato for Food Security through a Trait Observation and Discovery Network.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

DT and HM conceived and designed the project. HM and JL designed and implemented the NMR analyses. RC performed the metabolite identification and quantitation. JL acquired the NMR data and supervised the metabolite identification and quantification. DT performed the statistical, network analysis, GWAS and wrote the manuscript. GdS and RC prepared the samples. ESM, CM and DC designed and implemented the field trial and curated the phenotypic data. HLK coordinated genotyping using GBS and performed the filtering steps to obtain the final dataset

used in GWAS. All authors reviewed, edited and approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials.

SUPPORTING INFORMATION

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

Table S1. NMR assignments of metabolites identified in aqueous extracts of *Solanum tuberosum* L.

Figure S1. Descriptive statistics of morphological traits.

Figure S2. Descriptive statistics of metabolites.

Figure S3. Heatmap of metabolites correlation analysis.

Figure S4. Heatmap of morphological traits correlation analysis.

Figure S5. Quantitative trait locus linkage maps.

Data S1. Morphological trait profiles.

Data S2. Metabolite profiles.

Data S3. Metabolite single nucleotide polymorphism marker set.

Data S4. Fumarate single nucleotide polymorphism-associated genes.

Data S5. Potato association panel.

Data S6. Morphological trait definitions.

REFERENCES

- Ahuja, I., de Vos, R.C.H., Bones, A.M. and Hall, R.D. (2010) Plant molecular stress responses face climate change. *Trends Plant Sci.* **15**, 664–674.
- Anisimova, M. and Gascuel, O. (2009) Approximate likelihood ratio test for branches: a fast, accurate and powerful alternative. *Syst. Biol.* **55**, 539–552.
- Astle, W. and Balding, D.J. (2009) Population structure and cryptic relatedness in genetic association studies. *Stat. Sci.* **24**, 451–471.
- Atkinson, N.J. and Urwin, P.E. (2012) The interaction of plant biotic and abiotic stresses: from genes to the field. *J. Exp. Bot.* **63**, 3523–3543.
- Bae, H., Kim, S.K., Cho, S.K., Kang, B.G. and Kim, W.T. (2011) Overexpression of OsRDCP1, a rice RING domain-containing E3 ubiquitin ligase, increased tolerance to drought stress in rice (*Oryza sativa* L.). *Plant Sci.* **180**, 775–782.
- Bolger, A., Scossa, F., Bolger, M.E. *et al.* (2014) The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nat. Genet.* **46**, 1034–1038.
- Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B Met.* **57**, 289–300.
- Birch, P.R.J., Bryan, G., Fenton, B., Gilroy, E.M., Hein, I., Jones, J.T., Prashar, A., Taylor, M.A., Torrance, L. and Toth, I.K. (2012) Crops that feed the world 8: potato: are the trends of increased global production sustainable? *Food Secur.* **4**, 477–508.
- Bodner, G., Nakhforoosh, A. and Kaul, H.S. (2015) Management of crop water under drought: a review. *Agron. Sustain. Dev.* **35**, 401–425.
- Brugiare, N., Zhang, W.J., Xu, Q.Z. *et al.* (2017) Overexpression of RING domain E3 ligase ZmXerico1 confers drought tolerance through regulation of ABA homeostasis. *Plant Physiol.* **175**, 1350–1369.
- Cabello, R., Monneveux, P., De Mendiburu, F. and Bonierbale, M. (2013) Comparison of yield based drought tolerance indices in improved varieties, genetic stocks and landraces of potato (*Solanum tuberosum* L.). *Euphytica*, **193**, 147–156.
- Cabello, R., Cevallos, D., De Souza, J., Gutierrez, R., Farfan, E., Lozano, F., Eyzaguirre, R., Khan, A., Schafleitner, R. and Bonierbale, M. (2017) *Procedures for Standard Evaluation and Data Management of Advanced Potato Clones. Practical Guide to Assessing Potato Clones for Drought Tolerance under Field Conditions*. Lima: International Potato Center.

- Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552.
- Christensen, J.H., Hewitson, B., Busuioic, A. *et al.* (2007) Regional climate predictions. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Solomon, S., Quin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. and Miller, H.L., eds). New York: Cambridge University Press, pp. 847–940.
- Collakova, E., Yen, J.Y. and Senger, R.S. (2012) Are we ready for genome-scale modeling in plants? *Plant Sci.* **191**, 53–70.
- Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B. and Pang, E.C.K. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, **142**, 169–196.
- Dereeper, A., Guignon, V., Blanc, G. *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, 464–469.
- Dereeper, A., Audic, S., Claverie, J.M. and Blanc, G. (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol. Biol.* **12**, 8.
- Du, Z., Zhou, X., Ling, Y., Zhang, Z.H. and Su, Z. (2010) agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* **38**, W64–W70.
- Duan, F., Ding, J., Lee, D., Lu, X., Feng, Y. and Song, W. (2017) Overexpression of SoCYP85A1, a spinach cytochrome p450 gene in transgenic tobacco enhances root development and drought stress tolerance. *Front. Plant Sci.* **8**(1909). <http://dx.doi.org/10.3389/fpls.2017.01909>
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797.
- Egea, I., Albaladejo, I., Meco, V., Morales, B., Sevilla, A., Bolarin, M.C. and Flores, F.B. (2018) The drought-tolerant *Solanum pennellii* regulates leaf water loss and induces genes involved in amino acid and ethylene/jasmonate metabolism under dehydration. *Sci. Rep.* **8**(1). <http://dx.doi.org/10.1038/s41598-018-21187-2>.
- Fait, A., Fromm, H., Walter, D., Galili, G. and Fernie, A.R. (2008) Highway or byway: the metabolic role of the GABA shunt in plants. *Trends Plant Sci.* **13**, 14–19.
- Fernie, A.R., Carrari, F. and Sweetlove, L.J. (2004) Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Curr. Opin. Plant Biol.* **7**, 254–261.
- Gabor, C. and Tamas, N. (2006) The igraph software package for complex network research. *InterJournal*, **22**, 456.
- Geigenberger, P., Geiger, M. and Stitt, M. (1998) High-temperature perturbation of starch synthesis is attributable to inhibition of ADP-glucose pyrophosphorylase by decreased levels of glycerate-3-phosphate in growing potato tubers. *Plant Physiol.* **117**, 1307–1316.
- Gonzalez-Jorge, S., Ha, S.-H., Magallanes-Lundback, M. *et al.* (2013) Carotenoid cleavage dioxygenase4 is a negative regulator of beta-carotene content in Arabidopsis seeds. *Plant Cell*, **25**, 4812–4826.
- Guo, R., Shi, L.X., Jiao, Y., Li, M.X., Zhong, X.L., Gu, F.X., Liu, Q., Xia, X. and Li, H.R. (2018) Metabolic responses to drought stress in the tissues of drought-tolerant and drought-sensitive wheat genotype seedlings. *Aob Plants*, **10**(2). <http://dx.doi.org/10.1093/aobpla/ply016>.
- Gupta, P.K., Kulwal, P.L. and Jaiswal, V. (2014) Association mapping in crop plants: opportunities and challenges. *Adv. Genet.* **85**, 109–147.
- Hao, J., Liebecke, M., Astle, W., De Iorio, M., Bundy, J.G. and Ebbels, T.M.D. (2014) Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN. *Nat. Protoc.* **9**, 1416–1427.
- Harris, M.A., Deegan, J.L., Lomax, J. *et al.* (2008) The gene ontology project in 2008. *Nucleic Acids Res.* **36**, D440–D444.
- Haverkort, A.J. and Struik, P.C. (2015) Yield levels of potato crops: recent achievements and future prospects. *Field Crop. Res.* **182**, 76–85.
- Hijmans, R.J. (2003) The effect of climate change on global potato production. *Am. J. Potato Res.* **80**, 271–279.
- Hirsch, C.D., Hamilton, J.P., Childs, K.L., Cepela, J., Crisovan, E., Vaillancourt, B., Hirsch, C.N., Habermann, M., Neal, B. and Buell, C.R. (2014) Spud DB: a resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. *The Plant Genome*, **7**(1), plantgenome2013.12.0042. <http://dx.doi.org/10.3835/plantgenome2013.12.0042>.
- Hochberg, U., Degu, A., Toubiana, D., Gendler, T., Nikoloski, Z., Rachmilevitch, S. and Fait, A. (2013) Metabolite profiling and network analysis reveal coordinated changes in grapevine water stress response. *BMC Plant Biol.* **13**(1), 184. <http://dx.doi.org/10.1186/1471-2229-13-184>.
- Joazeiro, C.A.P. and Weissman, A.M. (2000) RING finger proteins: mediators of ubiquitin ligase activity. *Cell*, **102**, 549–552.
- Korver, R.A., Koevoets, I.T. and Testerink, C. (2018) Out of shape during stress: a key role for auxin. *Trends Plant Sci.* **23**, 783–793.
- Li, J., Lin, X., Chen, A., Peterson, T., Ma, K., Bertzky, M., Ciaia, P., Kapos, V., Peng, C. and Poulter, B. (2013) Global priority conservation areas in the face of 21st century climate change. *PLoS One*, **8**(1), e54839. <http://dx.doi.org/10.1371/journal.pone.0054839>.
- Lindqvist-Kreuz, H., de Boeck, B., Unger, P. *et al.* (2020). Global multi-environment resistance QTL for foliar late blight resistance in tetraploid potato with tropical adaptation. *bioRxiv*. <https://doi.org/10.1101/2020.02.16.950618>
- Lipka, A.E., Gore, M.A., Magallanes-Lundback, M. *et al.* (2013) Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. *G3: Genes - Genomes - Genetics*, **3**, 1287–1299.
- Maruenda, H., Cabrera, R., Canari-Chumpitaz, C., Lopez, J.M. and Toubiana, D. (2018) NMR-based metabolic study of fruits of *Physalis peruviana* L. grown in eight different Peruvian ecosystems. *Food Chem.* **262**, 94–101.
- Mittler, R. and Blumwald, E. (2010) Genetic engineering for modern agriculture: challenges and perspectives. *Annu. Rev. Plant Biol.* **61**, 443–462.
- Mittler, R., Finka, A. and Goloubinoff, P. (2012) How do plants feel the heat? *Trends Biochem. Sci.* **37**, 118–125.
- Mittler, R., Kim, Y., Song, L.H., Coutu, J., Coutu, A., Ciftci-Yilmaz, S., Lee, H., Stevenson, B. and Zhu, J.K. (2006) Gain- and loss-of-function mutations in Zat10 enhance the tolerance of plants to abiotic stress. *FEBS Lett.* **580**, 6537–6542.
- Nazarian-Firouzabadi, F. and Visser, R.G.F. (2017) Potato starch synthases: functions and relationships. *Biochem. Biophys. Rep.* **10**, 7–16.
- Nunes-Nesi, A., Fernie, A.R. and Stitt, M. (2010) Metabolic and signaling aspects underpinning the regulation of plant carbon nitrogen interactions. *Mol. Plant*, **3**, 973–996.
- Obidiegwu, J.E., Bryan, G.J., Jones, H.G. and Prashar, A. (2015) Coping with drought: stress and adaptive responses in potato and perspectives for improvement. *Front Plant Sci.* **6**, 542.
- Pons, P. and Latapy, M. (2005) Computing communities in large networks using random walks. *Lect. Notes Comput. Sci.* **3733**, 284–293.
- Riedelshheimer, C., Lisek, J., Czedik-Eysenberg, A., Sulpice, R., Flis, A., Grieder, C., Altmann, T., Stitt, M., Willmitzer, L. and Melchinger, A.E. (2012) Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl Acad. Sci. USA*, **109**, 8872–8877.
- Rontein, D., Dieuaide-Noubhani, M., Dufourc, E.J., Raymond, P. and Rolin, D. (2002) The metabolic architecture of plant cells - Stability of central metabolism and flexibility of anabolic pathways during the growth cycle of tomato cells. *J. Biol. Chem.* **277**, 43948–43960.
- Rosental, L., Perelman, A., Nevo, N., Toubiana, D., Samani, T., Batushansky, A., Sikron, N., Saranga, Y. and Fait, A. (2016) Environmental and genetic effects on tomato seed metabolic balance and its association with germination vigor. *BMC Genomics*, **17**(1). <http://dx.doi.org/10.1186/s12864-016-3376-9>.
- Rosyara, U.R., De Jong, W.S., Douches, D.S. and Endelman, J.B. (2016) Software for genome-wide association studies in autopolyploids and its application to potato. *The Plant Genome*, **9**(2), plantgenome2015.08.0073. <http://dx.doi.org/10.3835/plantgenome2015.08.0073>.
- Salvi, S. and Tuberosa, R. (2005) To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.* **10**, 297–304.
- Schauer, N., Semel, Y., Roessner, U. *et al.* (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* **24**, 447–454.
- Scossa, F., Brotman, Y., de Abreu e Lima, F., Willmitzer, L., Nikoloski, Z., Tohge, T. and Fernie, A.R. (2016) Genomics-based strategies for the use of natural variation in the improvement of crop metabolism. *Plant Sci.* **242**, 47–64.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504.
- Shirasawa, K., Takabe, T. and Kishitani, S. (2006) Accumulation of glycinebetaine in rice plants that overexpress choline monoxygenase from

- spinach and evaluation of their tolerance to abiotic stress. *Ann. Bot.* **98**, 565–571.
- Silvente, S., Sobolev, A.P. and Lara, M.** (2012) Metabolite adjustments in drought tolerant and sensitive soybean genotypes in response to water stress. *PLoS ONE*, **7**, e38554.
- Stacklies, W., Redestig, H., Scholz, M., Walther, D. and Selbig, J.** (2007) pcaMethods – a bioconductor package providing PCA methods for incomplete data. *Bioinformatics*, **23**, 1164–1167.
- Tak, Y.G. and Farnham, P.J.** (2015) Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin*, **8**(1). <http://dx.doi.org/10.1186/s13072-015-0050-4>.
- R Development Core Team** (2009) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Tian, T., Liu, Y., Yan, H.Y., You, Q., Yi, X., Du, Z., Xu, W.Y. and Su, Z.** (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45**, W122–W129.
- Toubiana, D., Fernie, A.R., Nikoloski, Z. and Fait, A.** (2013) Network analysis: tackling complex data to study plant metabolism. *Trends Biotechnol.* **31**, 29–36.
- Toubiana, D., Semel, Y., Tohge, T., Beleggia, R., Cattivelli, L., Rosental, L., Nikoloski, Z., Zamir, D., Fernie, A.R. and Fait, A.** (2012) Metabolic profiling of a mapping population exposes new insights in the regulation of seed metabolism and seed, fruit, and plant relations. *Plos Genet.* **8**(3), e1002612. <http://dx.doi.org/10.1371/journal.pgen.1002612>.
- Toubiana, D., Xue, W.T., Zhang, N.Y. et al.** (2016) Correlation-based network analysis of metabolite and enzyme profiles reveals a role of citrate biosynthesis in modulating N and C metabolism in *Zea mays*. *Front. Plant Sci.* **7**. <http://dx.doi.org/10.3389/fpls.2016.01022>.
- Turgeon, R.** (1989) The sink-source transition in leaves. *Annu. Rev. Plant Phys.* **40**, 119–138.
- van Loon, C.D.** (1981) The effect of water stress on potato growth, development, and yield. *Am. Potato J.* **58**, 51–69.
- Verslues, P.E., Lasky, J.R., Juenger, T.E., Liu, T.-W. and Kumar, M.N.** (2014) Genome-wide association mapping combined with reverse genetics identifies new effectors of low water potential-induced proline accumulation in arabidopsis. *Plant Physiol.* **164**, 144–159.
- Vos, P.G., Paulo, M.J., Voorrips, R.E., Visser, R.G.F., van Eck, H.J. and van Eeuwijk, F.A.** (2017) Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. *Theor. Appl. Genet.* **130**, 123–135.
- Watanabe, K.** (2015) Potato genetics, genomics, and applications. *Breed. Sci.* **65**, 53–68.
- Willcott, M.R.** (2009) MestRe Nova. *J. Am. Chem. Soc.* **131**, 13180.
- Yu, J.M. and Buckler, E.S.** (2006) Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* **17**, 155–160.
- Zang, D.D., Li, H.Y., Xu, H.Y., Zhang, W.H., Zhang, Y.M., Shi, X.X. and Wang, Y.C.** (2016) An Arabidopsis zinc finger protein increases abiotic stress tolerance by regulating sodium and potassium homeostasis, reactive oxygen species scavenging and osmotic potential. *Front. Plant Sci.* **7**. <http://dx.doi.org/10.3389/fpls.2016.01272>.
- Zeba, N., Ashrafuzzaman, M. and Hong, C.H.** (2006) Molecular characterization of the *Capsicum annuum* RING Zinc Finger Protein (CaRZFP1) gene induced by abiotic stresses. *J. Plant Biol.* **49**, 484–490.
- Zhang, D.Y., Tong, J.F., Xu, Z.L. et al.** (2016) Soybean C2H2-type zinc finger protein GmZFP3 with conserved QALGGH motif negatively regulates drought responses in transgenic Arabidopsis. *Front. Plant Sci.* **7**. <http://dx.doi.org/10.3389/fpls.2016.00325>.