PCP
PLANT & CELL PHYSIOLOGY

Effects of Parental Temperature and Nitrate on Seed Performance are Reflected by Partly Overlapping Genetic and Metabolic Pathways

Hanzi He[1], Leo A. J. Willems[1], Albert Batushansky[2], Aaron Fait[2], Johannes Hanson[3,4], Harm Nijveen[1,5], Henk W.M. Hilhorst[1] and Leónie Bentsink[1,4,*]

[1]Wageningen Seed Lab, Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, NL-6708 PB Wageningen, The Netherlands
[2]The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, 84990, Midreshet Ben-Gurion, Israel
[3]Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, SE-90187 Umeå, Sweden
[4]Department of Molecular Plant Physiology, Utrecht University, NL-3584 CH Utrecht, The Netherlands
[5]Laboratory of Bioinformatics, Wageningen University, Droevendaalsesteeg 1, NL-6708 PB Wageningen, The Netherlands
*Corresponding author: E-mail, leonie.bentsink@wur.nl.
(Received November 17, 2015; Accepted December 22, 2015)

Regular Paper

Seed performance is affected by the seed maturation environment, and previously we have shown that temperature, nitrate and light intensity were the most influential environmental factors affecting seed performance. Seeds developed in these environments were selected to assess the underlying metabolic pathways, using a combination of transcriptomics and metabolomics. These analyses revealed that the effects of the parental temperature and nitrate environments were reflected by partly overlapping genetic and metabolic networks, as indicated by similar changes in the expression levels of metabolites and transcripts. Nitrogen metabolism-related metabolites (asparagine, γ-aminobutyric acid and allantoin) were significantly decreased in both low temperature (15°C) and low nitrate (N0) maturation environments. Correspondingly, nitrogen metabolism genes (*ALLANTOINASE*, *NITRATE REDUCTASE 1*, *NITRITE REDUCTASE 1* and *NITRILASE 4*) were differentially regulated in the low temperature and nitrate maturation environments, as compared with control conditions. High light intensity during seed maturation increased galactinol content, and displayed a high correlation with seed longevity. Low light had a genotype-specific effect on cell surface-encoding genes in the *DELAY OF GERMINATION 6*-near isogenic line (NIL*DOG6*). Overall, the integration of phenotypes, metabolites and transcripts led to new insights into the regulation of seed performance.

Keywords: Light intensity • Metabolites • Nitrate • Seed maturation • Temperature • Transcriptome.

Abbreviations: CYP, cytochrome P450; DSDS50, days of seed dry storage required to reach 50% germination; *DOG*, *DELAY OF GERMINATION*; GABA, γ-aminobutyric acid; GAP, glycosylphosphatidylinositol (GPI)-anchored protein; G_{max}, maximum germination percentage; HL, high light; LL, low light; N0, low nitrate; NIL, near isogenic line; PCA, principal component analysis; RFO, raffinose family oligosaccharide; SL, standard light; t_{50}, the rate of germination; TCA, tricarboxylic acid.

Introduction

Seed maturation is an important phase of seed development during which embryo growth ceases, storage products accumulate and desiccation tolerance and seed dormancy are acquired. Parental environmental cues during seed maturation, such as temperature, light, nitrate and water, affect seed development and therefore the performance of the next generation (Donohue 2009). Temperature is one of the most important determinants of seed characteristics. Several studies have shown that the temperature during seed maturation determines the depth of primary dormancy. Low temperatures during seed maturation lead to deep primary dormancy, whereas warm temperatures lead to shallow dormancy in *Arabidopsis thaliana* (Schmuths et al. 2006, Kendall et al. 2011, He et al. 2014). Light is a pervasive environmental factor that affects development throughout the whole life cycle of the plant. Photoperiod and light intensity are two major influential factors of light signaling. The effect of photoperiod during seed maturation is species dependent. In Arabidopsis, Donohue et al. (2005) and Chen et al. (2014) did not find an effect of photoperiod, whereas Munir et al. (2001) and He et al. (2014) have shown that short days induced high dormancy levels. In other species, such as *Portulaca oleracea* (Gutterman 1974) and *Amaranthus retroflexus* (Kigel et al. 1979), germinability is promoted by short days, and the depth of dormancy increased with day length. Nitrate is an important nitrogen source for plants, but also a signal molecule that controls seed dormancy (Alboresi et al. 2005). Nitrate reduced ABA levels in dry seeds when provided to the mother plant during seed development (Matakiadis et al. 2009).

Information from genetic and transcriptome studies has increased our understanding of the maturation process and its cross-talk with environments. Transcriptome analysis has revealed that the ABA catabolic gene *CYP707A2* is responsive to exogenous nitrate levels and plays a central role in nitrate-mediated control of ABA levels during seed development and germination (Matakiadis et al. 2009). Kendall et al. (2011)

Plant Cell Physiol. 57(3): 473–487 (2016) doi:10.1093/pcp/pcv207, Advance Access publication on 6 January 2016, available online at www.pcp.oxfordjournals.org

identified genes associated with both dormancy [*DELAY OF GERMINATION* (*DOG1*)] and the hormone balance (*GA2ox6*, *NCED4* and *CYP707A2*) that may underlie the dormancy changes caused by low seed maturation temperatures. Global transcript analysis of *A. thaliana* (accession Cvi) seeds in a range of dormant and dry after-ripened states during dormancy cycling demonstrated that ABA signaling genes and ABA synthesis/catabolism key genes (*NCED* genes, *CYP707A2*, *GA2ox2* and *GA3ox1*) were differentially expressed across various dormant states (Cadman et al. 2006). These data support an ABA–gibberellin balance mechanism controlling dormancy cycling.

Metabolic analysis has revealed that seed maturation is associated with significant decreases in sugar, organic acid and amino acid levels, suggesting their efficient incorporation into storage reserves such as oil, seed storage protein and starch (Fait et al. 2006). One metabolite that occupies a central position in amino acid metabolism in plants is glutamate (Forde and Lea 2007). Glutamate is the precursor of many amino acids, such as arginine, proline and γ-aminobutyric acid (GABA). In addition, glutamine synthesized from glutamate, catalyzed by glutamine synthase (GS), is vital for nitrogen fixation. In stressful conditions, carbon (C) and nitrogen (N) metabolism must be tightly co-ordinated to sustain optimal growth and development (Zheng 2009). It has been shown that carbon shortage, due to the lack of or reduction in photosynthesis, could be compensated by catabolized amino acids that feed into the tricarboxylic acid (TCA) cycle (Galili et al. 2014). The regulation of the C/N balance can also be achieved by regulating the concentration of compounds with a low C/N ratio such as allantoin (C/N ratio of 1) (Miflin and Habash 2002). Fait et al. (2011) demonstrated that glutamate to GABA conversion during seed development has a profound effect on the C/N balance and storage reserve accumulation in the seed as well as on its germination performance.

Recently we have shown that the maturation of seeds in different environmental conditions resulted in variable seed performance (He et al. 2014). We showed that temperature, nitrate and light intensity were the most discriminative parental environments. In addition to that, individual genotypes responded differentially to the environmental conditions, which triggered us to investigate the underlying genetic and molecular pathways. Here, we describe the metabolic profiling of seeds matured in environments with low temperature, low nitrate, or low and high light intensities. To understand further the mechanistic details, a subset of these samples was also used for transcriptome analysis. Comparative analyses between both approaches showed that the effects of temperature and nitrate on seed performance are reflected by partly overlapping genetic and metabolic networks, while changed light intensity had more specific effects on transcripts and metabolites.

Results

Seeds of seven genotypes including L*er*, five *DOG* near isogenic lines (NILs; NIL*DOG1*, NIL*DOG2*, NIL*DOG3*, NIL*DOG6* and NIL*DOG22*) and the *dog1-1* mutant that had matured under various temperature, nitrate and light intensity regimes were previously analyzed for seed performance (He et al. 2014). These analyses revealed that low temperature (15°C), low nitrate (N0) and low (LL) and high light (HL) were the most influential seed maturation environments. Plant and seed phenotypes of L*er* are summarized in **Table 1**. Seeds grown in these environments were selected to study the underlying metabolic pathways, using a metabolomics approach.

Genetic effects on metabolite profiles

Primary metabolites of dry dormant seeds were analyzed using untargeted gas chromatography–time of flight–mass spectrometry (GC-TOF-MS). This resulted in 24,010 significant mass signals; after processing the raw data, we obtained 124 representative masses each consisting of reconstituted mass spectra. These 124 predominantly primary metabolites were identified by matching mass spectra and retention times to an in-house-constructed library and the NIST05 (National Institute of Standards and Technology; http://www.nist.gov/srd/) libraries. This led to the identification of 41 metabolites, which included amino acids, sugars, organic acids and some precursors and derivatives of central metabolism compounds (**Supplementary Table S1**). In addition to the primary metabolite profile, ABA content was measured for the same samples.

Principle component analysis (PCA) of the primary metabolite profiles showed that the maturation environmental effects were larger than the genotypic effects (**Fig. 1**). The strong genotypic effect on seed performance, especially seed dormancy, is biased since the genotypes have been selected for these differences. However, the fact that different factors explain the phenotypic and metabolic variance indicates that seed performance cannot only be explained by the metabolic profile. A possible reason is the large plasticity of cellular metabolism that buffers the effect of the environment (Almaas et al. 2005, Fu et al. 2009).

To assess the relationship between genotype and maturation environment, cluster analysis was performed for phenotype (16 plant and seed performance traits) and metabolite profiles separately. The results of cluster analysis based on phenotypes revealed a clear genetic effect of the genotypes (**Supplementary Table S2**). The *dog1-1* mutant samples clustered as one group in both the phenotypic and metabolic analyses (**Supplementary Table S2**) despite the various maturation environments. Since the lack of a functional *DOG1* gene over-rides these environmental effects, *dog1-1* was excluded from further metabolic analysis.

Correlations between seed performance traits and metabolites

With the intention of investigating the metabolic changes in relation to their altered seed performance, correlation plots were made among metabolite levels, ABA content and seed performance traits using L*er* and the five NILs for each maturation environment.

Table 1 Plant and seed phenotypes of the Landsberg *erecta* plants matured in low light intensity (LL), standard light intensity (SL), high light intensity (HL), low temperature (15°C), 20°C/nitrate 5 mM (20°C/N5) and low nitrate (N0)

Phenotypes	LL	Control	HL	15°C	Control	N0
Plant height (cm)	22.00	31.33	34.33	36.17	33.33	29.67
Silique per plant	89.67	135.00	210.00	197.00	162.33	129.67
Seeds per silique	47.00	63.33	65.33	61.00	72.67	64.33
Seed size (mm^2)	0.16	0.15	0.16	0.17	0.15	0.15
Seed weight (mg per 1,000 seeds)	20.99	21.40	23.22	20.75	22.19	20.39
DSDS50 (d)	31.67	12.00	10.67	36.33	6.67	20.67
Longevity (%)	0.11	0.23	0.48	0.01	0.22	0.10
G_{max} 22°C (%)	0.97	0.97	0.99	1.00	0.99	1.00
t_{50} 22°C (hours)	47.21	46.14	47.86	53.66	42.45	56.68
G_{max} 10°C (%)	0.97	1.00	1.00	0.95	0.98	0.98
t_{50} 10°C (h)	104.97	91.19	89.83	150.03	89.65	121.49
G_{max} 30°C (%)	0.61	0.43	0.56	0.22	0.99	0.98
G_{max} mannitol (%)	0.91	0.86	0.98	0.21	0.87	0.39
G_{max} salt (%)	0.39	0.30	0.16	0.05	0.18	0.01
G_{max} ABA (%)	0.05	0.03	0.15	0.31	0.36	0.26
Nitrate content (mg g^{-1} seeds)	–	–	–	–	0.71	0.07

Data of two independent experiments are shown; each experiment has one control condition (He et al. 2014).

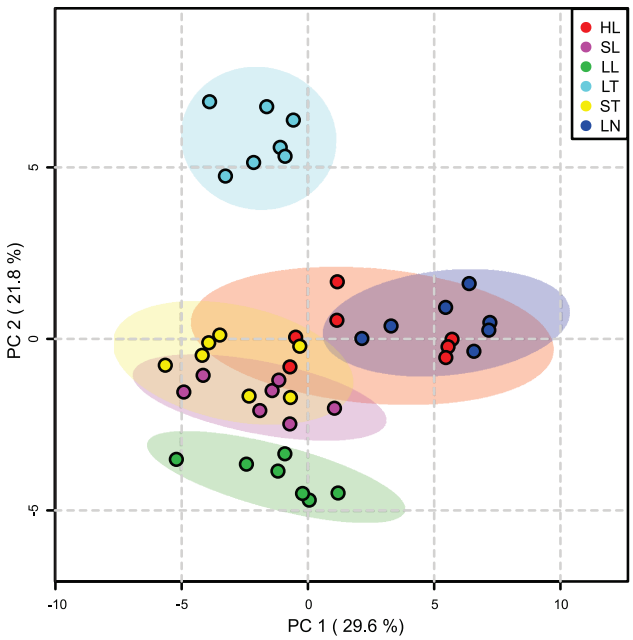


Fig. 1 Principle component analysis (PCA) of all the 124 detected metabolites. Three biological replicates of the seven genotypes are used in the analysis; only the averages of three biological replicates are presented. Different colors indicate different maturation environments as shown in the key. HL, high light intensity; SL, standard light intensity; LL, low light intensity; LT, low temperature (15°C); ST, standard temperature/standard nitrate (20°C/N5); and LN, low nitrate (N0).

Overall, the four correlation plots were different (**Fig. 2**), indicating that the different maturation environments affecting seed performance are reflected by different metabolic pathways. However, there were also some common effects. ABA levels correlated positively with dormancy levels (DSDS50: days of seed dry storage required to reach 50% germination) in all four environments, with correlation coefficients (r) of 0.67, 0.67, 0.45 and 0.34 in, respectively, 15°C, N0, HL and LL. These results confirm the importance of ABA in the acquisition of seed dormancy (Karssen et al. 1983).

The correlation patterns in 15°C and N0 have some similarities (**Fig. 2A, B**). DSDS50 showed similar correlations with all the metabolites in both environments (**Fig. 2A, B**). ABA content correlated negatively with almost all the phenotypes except DSDS50. Similarly, galactinol and myo-inositol contents both displayed negative correlations with almost all phenotypes except DSDS50; yet in N0 the negative correlations were higher than in 15°C. In LL and HL (**Fig. 2C, D**), similar correlations for DSDS50 with all metabolites were observed. Four TCA cycle intermediates and 2-hydroxyglutarate, which is derived from the TCA cycle intermediate 2-oxoglutarate, all had negative correlations with almost all measured seed performance traits in both environments. Allantoin and urea are two nitrogen-rich compounds. In low temperature (15°C) and low nitrate (N0) conditions, allantoin correlated negatively with DSDS50 ($r = -0.55$ and -0.34). However, in LL, allantoin and urea correlated positively with DSDS50 ($r = 0.60$ and 0.49) (**Fig. 2C**).

The correlation plots also showed that each environment had its unique effects (**Fig. 2**). Nitrate content correlated positively with most of the amino acids and nitrogen-rich compounds in N0 (**Fig. 2B**). Only NAcGlu, the intermediate between glutamate and the urea cycle, showed a negative correlation with nitrate content. Longevity showed a strong negative correlation with galactarate ($r = -0.71$) and gluconate ($r = -0.60$). Galactinol in altered light intensity, especially in HL, correlated highly positively with seed longevity ($r = 0.82$) (**Fig. 2D**).

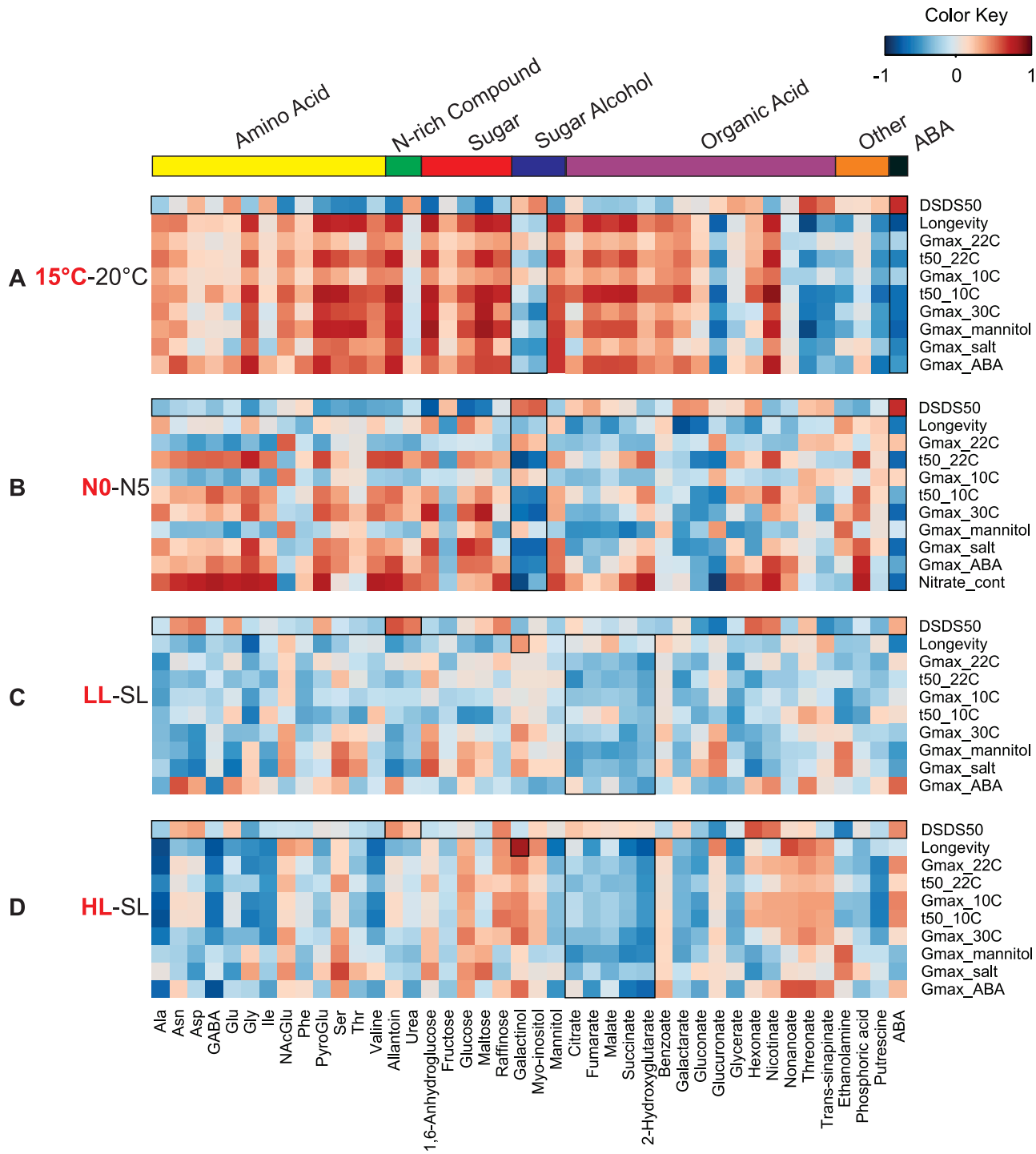


Fig. 2 Correlation analysis of metabolites and seed performance traits of L*er* and five NILs in four maturation environments. (A) Low temperature (15°C) and standard temperature (20°C, control); (B) low nitrate (N0) and standard nitrate (N5; control, N5 and 20°C are the same control); (C) low light intensity (LL) and standard light intensity (SL, control); and (D) high light intensity (HL) and standard light intensity (SL, control). The black rectangles indicate the correlations discussed in the text.

Overview of metabolite abundance change in metabolic pathways

Thirty-seven out of the 41 identified primary metabolites were mapped on metabolic pathways to view the metabolic changes in a more dynamic manner (**Fig. 3**). Low temperature significantly altered TCA cycle activity: all four TCA intermediates were decreased in 15°C; the level of fumarate and succinate was halved.

Low nitrate and low temperature both decreased the contents of nitrogen-rich amino acids, such as asparagine and nitrogen-related compounds, e.g. GABA. Also pyroglutamate and allantoin were decreased in both conditions, and glutamate,

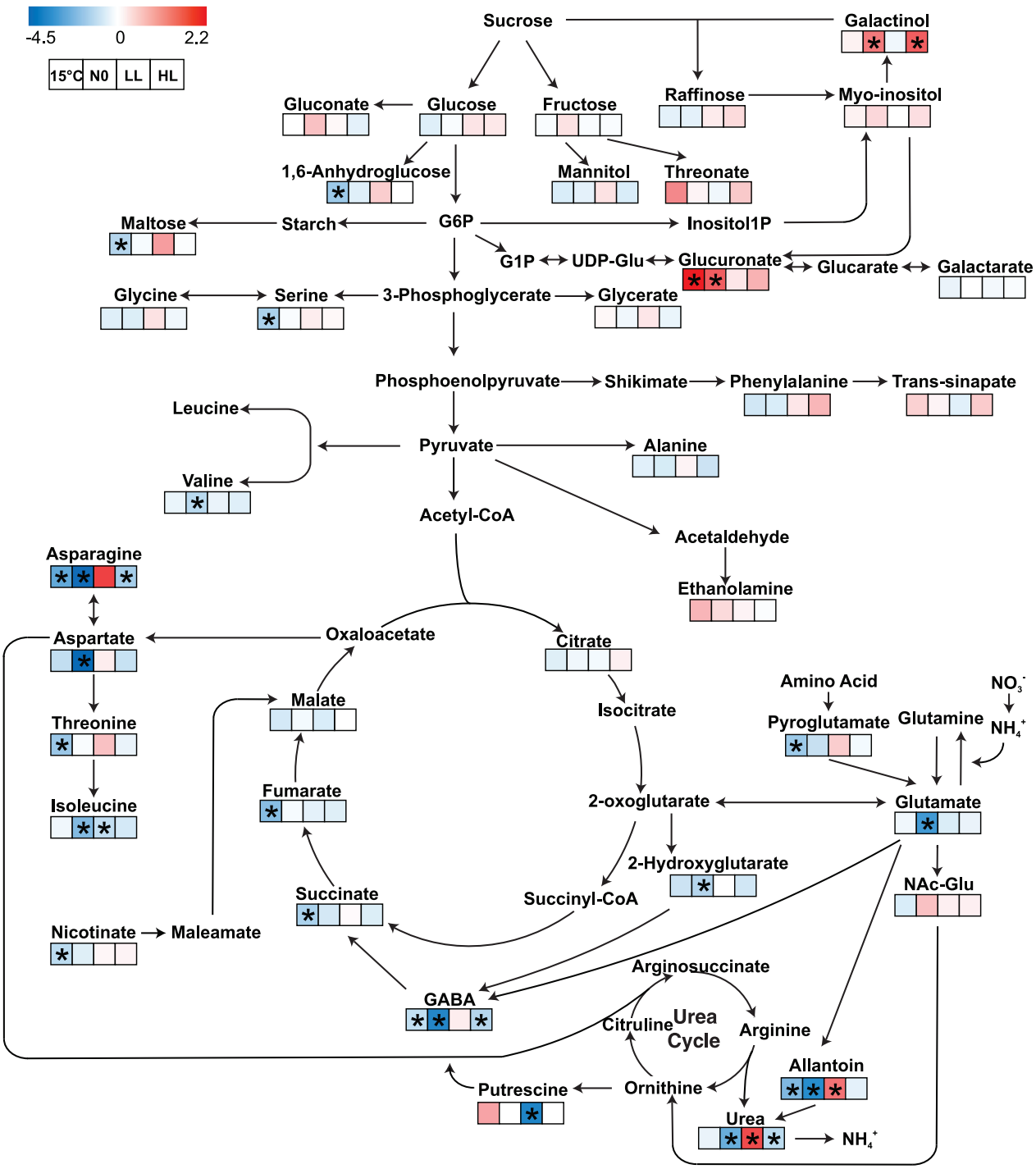


Fig. 3 Changes of metabolite levels in a framework of metabolic biosynthetic pathways. The values below each metabolite box indicate the fold change of *Ler* and five NILs. Asterisks indicate a >2-fold ($P < 0.05$) significant increase/decrease between changed environment and control. Environments are low temperature (15°C), low nitrate (N0), low light intensity (LL) and high light intensity (HL).

aspartate and urea only decreased in N0. In both conditions, glucuronate levels were significantly increased (>2-fold). Together with the correlation analyses (**Fig. 2A, B**), these results indicate that 15°C and N0 affect seed performance apparently via similar metabolic pathways.

Nitrogen-rich compounds, which were decreased in N0 and low temperature (**Fig. 3**), correlate with increased seed dormancy, i.e. a negative correlation of nitrogen-rich compounds with DSDS50 (**Fig. 2A, B**). However, in LL, allantoin and urea increased (**Fig. 3**) with increased seed dormancy, i.e. a positive correlation of allantoin and urea with DSDS50 (**Fig. 2C**). A possible explanation for this is that under the low light intensity plants produce fewer siliques and fewer seeds per silique (**Table 1**) (He et al. 2014). However, the amount of nutrients supplied to the plant was the same as in the control condition. Consequently, the amount of assimilated nutrients per seed

may be higher than under the control condition. Thus, the low light intensity condition could resemble the high nitrate condition to some extent, whereas the high light intensity condition, apparently, is similar to the low nitrate condition. This is supported by cluster analysis where N0 clusters with HL (**Supplementary Table S2**; metabolites cluster 3).

Metabolite networks show decreased correlations of metabolites in low temperature and high light intensity

To obtain more insight into the associations between metabolites and phenotypes, correlation-based networks were constructed. Under standard conditions, the network showed that the relationship between metabolites was strong and intertwined (**Fig. 4A**). In particular, amino acids and nitrogen-rich compounds were highly interconnected (**Fig. 4A**). The low maturation temperature led to a strong decrease in the number of nodes and edges (**Fig. 4B**; **Supplementary Table S3**). Moreover, correlations involving succinate, fumarate, citrate and malate did not pass the threshold set for the network at low maturation temperature, suggesting a significant change in TCA cycle activity that shifted from normal to stress metabolism (**Fig. 4B**). This confirms the earlier observation that all four TCA intermediates were decreased in 15°C (**Fig. 3**). Taken together, energy metabolism appears to be restricted in 15°C. Moreover, the appearance of the unique triad raffinose–myo-inositol–galactinol suggests the involvement of these three metabolites in the response to low temperature stress (**Fig. 4B**). They became highly correlated in low temperature ($r > 0.8$) (**Supplementary Table S4**). The temperature-induced changes also affected metabolism of N compounds, though to a weaker extent. Thus aspartate, one of the main players of the network under normal temperature, lost part of its edges, while glutamate kept most of it edges under both regimes (**Fig. 4A, B**). This reinforces the central role of glutamate in relation to other amino acids (for an extensive review, see Forde and Lea 2007).

The effect of light intensity was also investigated using the network approach. The metabolic cluster analysis (**Supplementary Table S2**; metabolic cluster 4) indicated that the standard light (SL) and LL maturation environmental effects are rather similar (**Table 1**). Thus, to improve the interpretation of the analysis, the network was constructed by

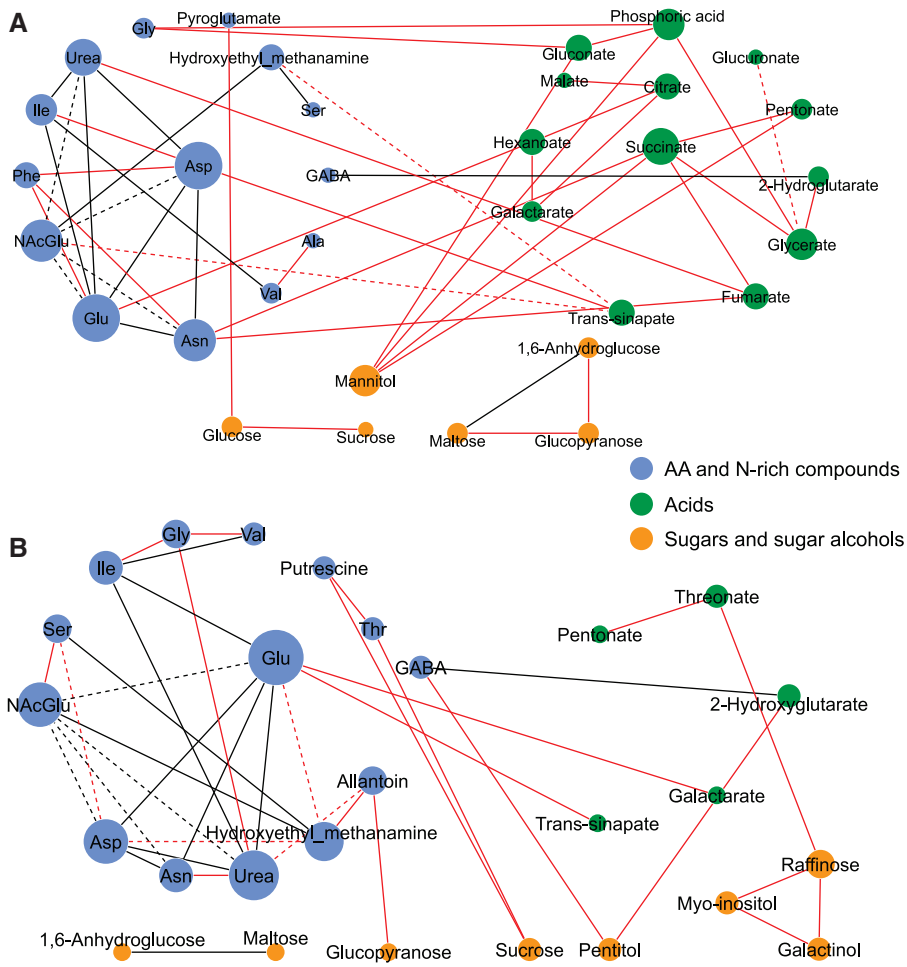


Fig. 4 Network of metabolites at standard temperature (20°C) (A) and low temperature (15°C) (B). Node color represents the compound class. Node size represents the nodal degree. Edge type represents positive (solid) and negative correlation (dash). Edge color represents uniqueness (red) or commonality (black) of the edges in both networks. Hydroxyethyl_methanamine, pentonate and glucopranose are predicted metabolites.

combining the SL and LL data in comparison with the HL network. The close relationship between amino acids under HL was absent from the SL/LL network (**Supplementary Fig. S1**). The links between sugars and TCA cycle intermediates (succinate–fumarate–malate) also dropped below those of HL. Due to the limited number of significant correlations, it was not possible to generate the N0 network, suggesting that metabolic correlations are significantly disrupted because of the low nitrate maturation environment.

The effect of the parental environment on the dry seed transcriptome

To investigate the effect of the environment on seed stored mRNAs, we analyzed genome-wide gene expression in freshly harvested dry seed of NIL*DOG1*, NIL*DOG3* and NIL*DOG6* grown in low temperature (15°C), N0 and LL. All three genotypes responded in a similar way to the environmental factors and all three environments affected seed dormancy in the same direction (increased DSDS50) (**Fig. 5A–C**). The number of genes differentially expressed in 15°C was the highest compared with N0 and LL, especially for the up-regulated genes (**Fig. 5D**). In 15°C and N0 conditions, the three genotypes responded in a similar way, with a significant number of genes that were commonly up-/down-regulated (**Fig. 5E, F, H**, 5I; **Supplementary Table S5**) so in this case the genotype-specific response to a certain environment is relatively small. Over-representation analysis was performed on significantly changed genes (>2-fold, adjusted *P*-value < 0.05) using Pageman analysis (Usadel et al. 2006) (**Fig. 5K**). Several functional categories, among which N metabolism, RNA-related and cell-related categories, were commonly down-regulated in 15°C and N0. Besides that, down-regulated genes of NIL*DOG1* and NIL*DOG6* in 15°C are enriched in (general) stress, abiotic stress and heat stress classes.

Since the three genotypes showed a high number of overlapping genes (**Fig. 5E–I**), and in 15°C and N0 shared some functional categories (**Fig. 5K**), the core sets of commonly up- and down-regulated genes in the three genotypes were selected and compared between 15°C and N0. Eight of the up-regulated genes (total 14) in N0 are also up-regulated in seeds grown in 15°C, and half (17) of the genes down-regulated (total 34) in N0 are shared with genes down-regulated in seeds grown in 15°C (**Fig. 5L, M**). This overlap of up-/down-regulated genes, together with some common changes in metabolites (**Figs. 2**, **3**), suggests that the effects of low temperature (15°C) and low nitrate (N0) on seed performance are reflected by partly overlapping pathways. As the dormancy responses of the three NILs were similar in the three environments (**Fig. 5A–C**), the commonly responding genes are possibly responsible for the increase in seed dormancy. Among the commonly up-regulated genes is *GROWTH-REGULATING FACTOR 7* (*GRF7*) (**Supplementary Table S6**), which was reported to function as a transcriptional repressor of ABA-responsive genes (Kim et al. 2012) and to confer the ability to respond to karrikins (Nelson et al. 2010). Another up-regulated gene *ALLANTOINASE* (*ALN*) is known to degrade seed-stored allantoin to produce NH_3. This can explain the decrease in allantoin in both 15°C and N0 (**Fig. 3**). Probably, allantoin is degraded to compensate for N-deficient environments. *NITRATE REDUCTASE 1* (*NR1*) and *NITRITE REDUCTASE 1* (*NIR1*) are commonly down-regulated genes in 15°C and N0. *NITRATE TRANSPORTER 1.1* (*NRT1.1*) was up-regulated in 15°C [1.78-, 1.96- amd 1.56-fold up-regulated compared with control for NIL*DOG1*, NIL*DOG3* and NIL*DOG6*, respectively (adjusted *P*-value < 0.05)] and N0 (>2-fold, adjusted *P*-value < 0.05). Expression of *CYP707A2* in both 15°C and N0 of the three genotypes was down-regulated significantly (adjusted *P*-value < 0.05) by approximately 1.5-fold (1.44- to 1.68-fold). Furthermore, genes encoding a number of (other) enzymes were commonly up-/down-regulated in 15°C and N0 (**Supplementary Table S6**). Three of these enzymes are involved in carbohydrate metabolism: *BETA GLUCOSIDASE 28* (up-regulated), *UDP-GLUCOSYL TRANSFERASE 71B1* (down-regulated) and *AT5G41670* which is a 6-phosphogluconate dehydrogenase family protein (down-regulated). The lactate/malate dehydrogenase family protein (AT1G04410) was also down-regulated.

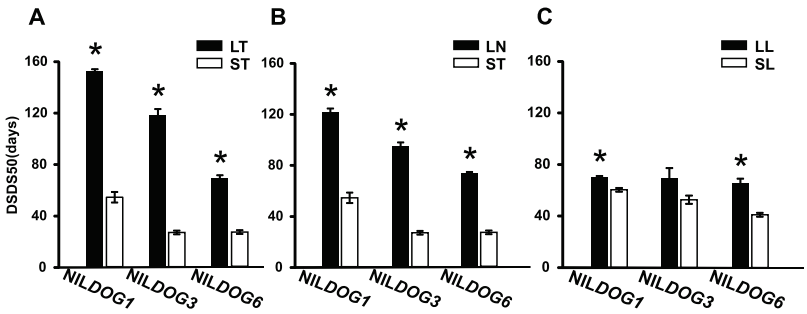
DOG1, *GA2ox6* and two NAC transcription factor genes (*ANAC3* and *ANAC019*), which are involved in ABA-mediated signaling pathways, were specifically up-regulated at 15°C. This suggests that the functions of these genes are specific in response to low temperature, which has been shown previously for *DOG1* and *GA2ox6* (Kendall et al. 2011). The uniquely down-regulated genes in N0 contain a MYB-like transcription factor (AT1G68670) and a PP2C family protein (AT5G26010).

There were very few genes differentially expressed for NIL*DOG1* and NIL*DOG3* in LL (**Fig. 5G, J**), while in NIL*DOG6* many genes were differentially regulated by light intensity. Over-representation analysis showed that, in particular, down-regulated genes of *DOG6* are cell wall-related genes (**Fig. 5K**). Among these, 11 encode cell wall proteins and eight encode membrane-anchored proteins (**Supplementary Table S7**). These eight membrane-anchored proteins were either predicted or confirmed to be glycosylphosphatidylinositol (GPI)-anchored proteins (GAPs; Borner et al. 2003), except for AT1G28290.

Integrated analysis of phenotype, metabolites and transcripts

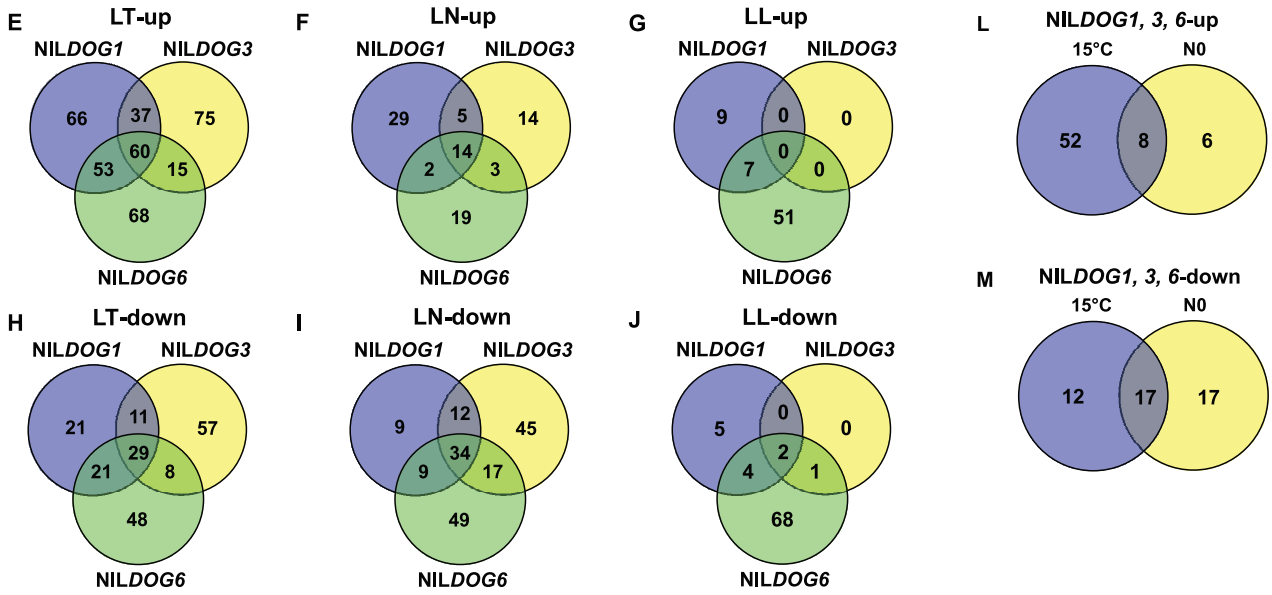
With the aim of discovering genes and metabolites that are possibly associated with seed performance traits, correlation analysis was performed with the gene expression levels [2log(intensities) > 5], all detected metabolite contents and seed performance traits in all environments (15°C, 20°C, N0, LL and SL) for NIL*DOG1*, NIL*DOG3* and NIL*DOG6* together. DSDS50, longevity, t_{50} 10°C, G_{max} (maximum germination percentage) mannitol and G_{max} salt were highly correlated with expression profiles of a number of genes and/or metabolites ($r > 0.75$, adjusted *P*-value < 0.05) (**Fig. 6**; **Supplementary Table S8**).

DSDS50 correlated ($r > 0.75$, adjusted *P*-value < 0.05) with 13 genes, but no significant correlation with metabolites was identified (**Fig. 6**; **Supplementary Table S8**). Among the genes was *MOTHER OF FT AND TFL1* (*MFT*), which was shown to



D Numbers of genes up/down regulated in low temperature (15°C), low nitrate (N0) and low light intensity (LL) for NIL*DOG1*, NIL*DOG3* and NIL*DOG6*

	NIL*DOG1*		NIL*DOG3*		NIL*DOG6*	
	up	down	up	down	up	down
15°C	216	82	187	105	196	106
N0	50	64	36	108	38	109
LL	16	11	0	3	58	75



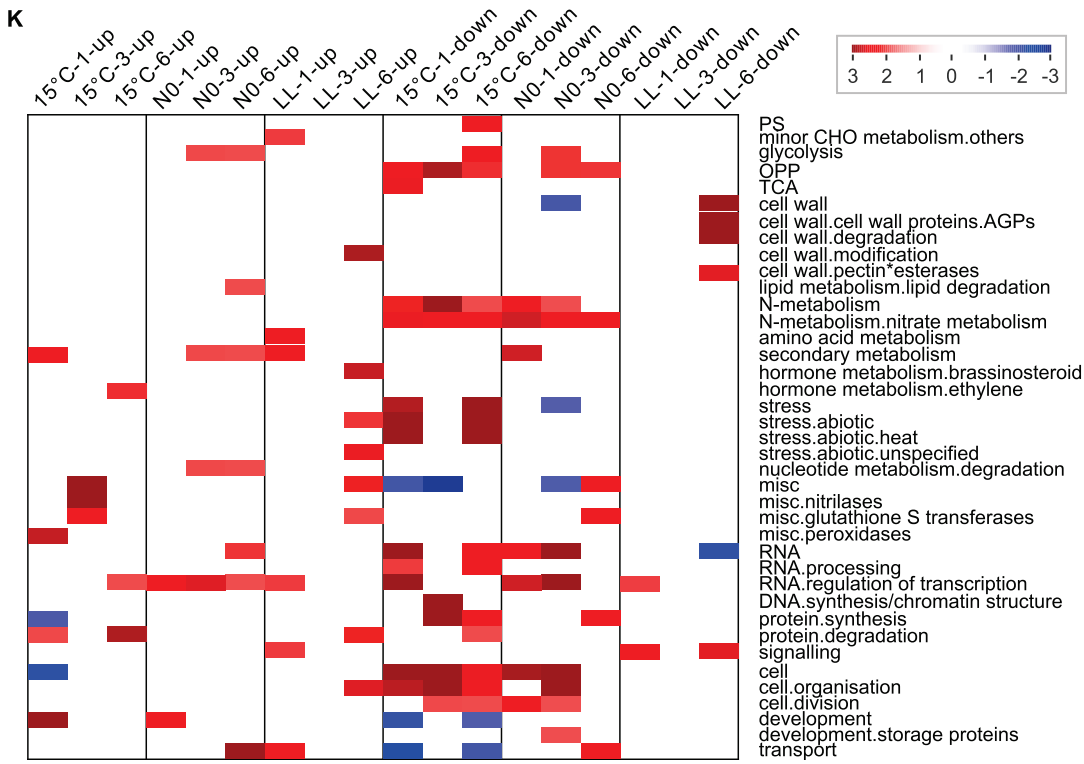


Fig. 5 Dormancy levels and transcriptome changes of NIL*DOG1*, NIL*DOG3* and NIL*DOG6* seeds grown in different seed maturation environments. Dormancy levels of seeds grown in low temperature (15°C) (A), low nitrate (N0) (B) and low light intensity (LL) (C) environments. (continued)

induce seed dormancy in both *A. thaliana* (Vaistij et al. 2013) and wheat (Nakamura et al. 2011) and could negatively regulate ABA signaling to promote germination in Arabidopsis (Xi et al. 2010). Another gene involved in ABA-mediated responses is *ABA INSENSITIVE RING PROTEIN 2* (*AIRP2*) (Cho et al. 2011) (**Supplementary Table S8**). *DOG1* correlated with DSDS50 ($r = 0.73$, adjusted P-value = 0.076).

G_{max} mannitol correlated strongly with 63 genes, eight metabolites (pyroglutamate, nicotinate, maltose and five unknowns) and two other seed performance traits (t_{50} 10°C and G_{max} salt) (**Supplementary Table S8**). *CYP707A2* and two genes encoding seed storage albumin proteins significantly correlated with G_{max} mannitol, indicating that the potential for germination in mannitol is possibly related to seed storage albumins and ABA catabolism. The highly correlated genes and metabolites have the potential to be involved in the regulation of seed performance and are, therefore, worthy of further investigation.

Discussion

We have analyzed the metabolite profiles of seeds matured in four different environments (15°C, N0, LL and HL) and determined their relationship with their corresponding seed performance traits. A subset of genotypes and environments were selected for transcriptome analysis to investigate the effect of the environment on seed stored mRNAs. A correlation-based integrative analysis was performed on phenotypes, metabolites and transcripts to uncover genes and metabolites associated with phenotypes.

The effects of parental temperature and nitrate on seed performance are reflected by partly overlapping genetic and metabolic pathways

We have identified a strong overlapping effect of seeds matured in 15°C and N0 based on the correlation of metabolites with phenotypes (**Fig. 2A, B**), fold change of metabolites (**Fig. 3**), over-representation analysis of the transcriptome (**Fig. 5K**) and overlapping up-/down-regulated genes (**Fig. 5L, M**). This could be either a direct effect of the treatments or an indirect effect of low temperature and low nitrate on plant development and maturation (delayed and prolonged maturation period; He et al. 2014). It is possible that the uptake, mobilization and re-location of nutrients in the plants were altered, or that the C/N balance was modified. This indirect response could cause changes that partially resemble nitrate-deficient environments. For example, some amino acids and N-related metabolites were decreased in both 15°C and N0 (**Fig. 3**). The functional category 'N metabolism' was over-represented in both environments (**Fig. 5K**). In addition, three genes that are involved in nitrogen metabolism, *NITRILASE 4*, *NIR1* and *NR1*, were all affected. *NITRILASE 4* was up-regulated in 15°C and N0, while *NIR1* and *NR1* were both down-regulated in 15°C and N0. This is in agreement with Kendall et al. (2011), who also showed down-regulation of the nitrate metabolism gene *NITRATE REDUCTASE 2* (*NR2*) when seeds matured in 10°C. Furthermore, *NITRATE TRANSPORTER 1.1* (*NRT1.1*) was also up-regulated in 15°C and N0, although to a lesser extent in 15°C. This suggests that low temperature also affected nitrate transport and assimilation, and therefore influenced accumulation of amino acids and N-rich compounds in dry seeds.

NRT1.1 displays an unusual dual-affinity transport activity (Liu et al. 1999) depending on phosphorylation of the Thr101 residue catalyzed by the CIPK23 kinase (Gojon et al. 2011). It was shown that in addition to nitrate transport activity, *NRT1.1* contributed to other important physiological functions, including auxin transport (Krouk et al. 2010), carbohydrate accumulation (Schulze et al. 1994), cytosol acidification (Meraviglia et al. 1996) and modified root or shoot development (Guo et al. 2002). Besides acting as a signaling sensor to release seed dormancy (Alboresi et al. 2005), *NRT1.1* has also been suggested to control several other responses by its signaling effect, such as increased stomatal opening (Guo et al. 2003). Thus it was hypothesized by Gojon et al. (2011) that *NRT1.1* could mediate several different signaling pathways, through different sensing mechanisms. Possibly *NRT1.1* senses temperature changes in the environment and thus may cause similar effects to nitrate-deficient environments. Moreover, *NRT1.1* is regulated by auxin in both shoots and roots (Guo et al. 2002). Another nitrate transporter NRT1.2 is able to transport ABA, in addition to nitrate (Kanno et al. 2012). Since we showed that ABA content increased with seed dormancy (**Fig. 2**), *NRT1.1* could also be regulated by ABA levels in seeds. Clearly, the molecular mechanisms involved require further investigation.

Carbon/nitrogen balance in low temperature and low nitrate environments

Glucuronate accumulated to very high levels in 15°C and N0 environments (**Fig. 3**). Possibly, the oxidation of myo-inositol results in the formation of glucuronate present in cell wall polysaccharides, and this pathway is important in controlling the carbohydrate flux (Valluru and Van den Ende 2011). This indicates that the C/N balance is dramatically changed in both

Fig. 5 Continued

Averages of three replicates are presented. Error bars show the SE. Asterisks indicate significant differences between treatment and control of each genotype ($P < 0.05$). (D) Numbers of genes up-/down-regulated in low temperature (15°C), low nitrate (N0) and low light intensity (LL) for NIL*DOG1*, NIL*DOG3* and NIL*DOG6*. The threshold is set at a fold change of two and $P < 0.05$. (E–G) Venn diagrams comparing up-regulated genes in freshly harvested dry seeds of NIL*DOG1*, NIL*DOG3* and NIL*DOG6* grown in 15°C, N0 and LL, respectively. (H–J) Venn diagrams comparing down-regulated genes in freshly harvested dry seeds of NIL*DOG1*, NIL*DOG3* and NIL*DOG6* grown in 15°C, N0 and LL, respectively. (K) Over-representation analyses of differentially expressed genes of NIL*DOG1*, NIL*DOG3* and NIL*DOG6* seeds grown in low temperature (15°C), low nitrate (N0) and low light intensity (LL). Red color indicates over-representation, and blue color indicates under-representation. The color key is shown in the figure. (L, M) Venn diagrams of all up-regulated genes (A) and all down-regulated genes (B) in freshly harvested dry seeds of NIL*DOG1*, NIL*DOG3* and NIL*DOG6* matured in 15°C and N0.

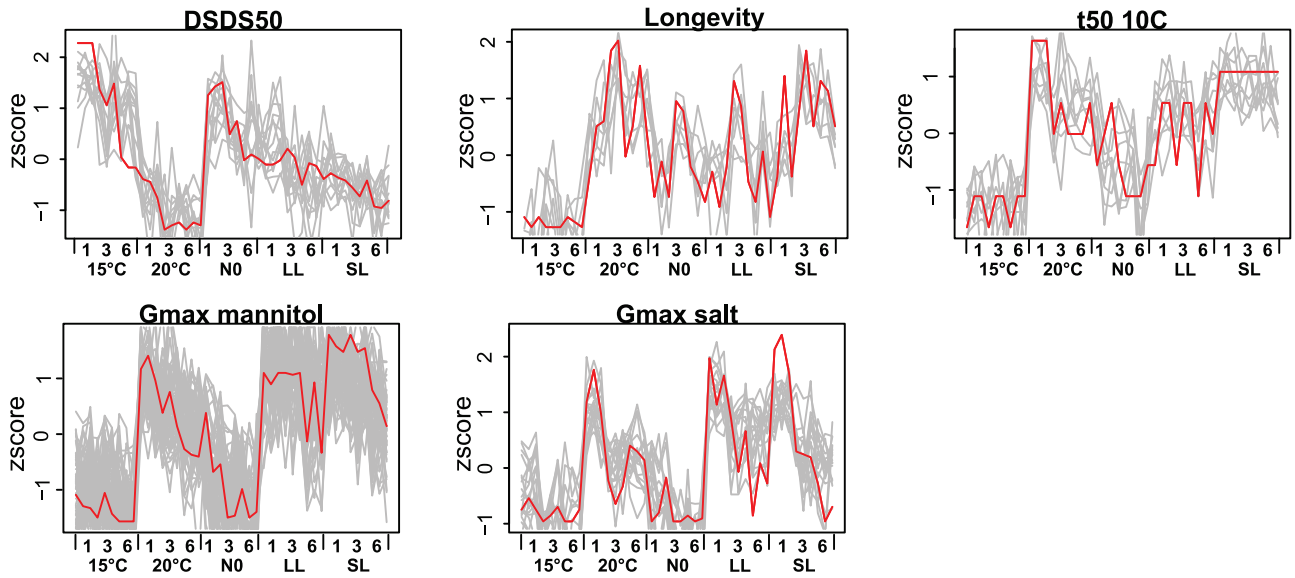


Fig. 6 Correlations among seed performance traits, all the expressed transcripts [2log(intensities) > 5] and all detected metabolites for five environments (15°C, 20°C, N0, LL and SL) using NIL*DOG1*, NIL*DOG3* and NIL*DOG6*. Correlation coefficients were calculated using Spearman's rank correlation ($r > 0.75$). The *P*-value was adjusted by Benjamini–Hochberg correction (adjusted *P* value < 0.05). The seed performance traits that have correlated transcripts/metabolites/seed performance traits are shown. Red lines indicate the *Z*-score transformed seed performance trait value. Gray lines indicate values of correlated transcripts/metabolites/seed performance traits. A list of these traits is presented in **Supplementary Table S8**.

15°C and N0. Further evidence of a changed C/N balance is the significant decrease of GABA in both environments (**Fig. 3**). It was shown by Fait et al. 2011 is a four-carbon non-proteinogenic amino acid that is associated with primary C and N metabolism. It is metabolized via a pathway known as the GABA shunt which by-passes two steps of the TCA cycle by using glutamate as a source (reviewed by Michaeli and Fromm 2015). Therefore, via transamination reactions, GABA could possibly be used to mitigate the reduction in N and consequently contribute to the C/N balance (Fait et al. 2008).

Specific effect of low light intensity on NIL*DOG6*

We have shown that low light intensity had a specific effect on the NIL*DOG6* transcriptome (**Fig. 5G, J**), and that, in particular, the expression of genes encoding cell wall and membrane proteins was affected (**Supplementary Table S6**). GAPs were the main category of membrane-anchored proteins. A number of studies have highlighted the importance of GAPs at the plant cell surface. Smyth (2004) has reviewed the biological role of arabinogalactan proteins and concluded that their amphiphilic molecular nature makes them prime candidates for linking the cell wall, the plasma membrane and the cytoplasm. It is likely that *DOG6* controls dormancy and germination by a currently unknown effect on the cell surface, and this molecular pathway is regulated by light intensity.

Importance of galactinol, myo-inositol and raffinose in changing environments

It was observed that the raffinose pathway metabolites raffinose, galactinol and myo-inositol were more strongly correlated in low temperature (**Fig. 4**; **Supplementary Table S4**), and that galactinol and myo-inositol correlated negatively with a number of seed performance traits in low temperature (15°C) and low nitrate (N0) (**Fig. 2A, B**). The importance of raffinose pathway metabolites in cold acclimation was demonstrated by Cook et al. (2004). Our results suggest a co-ordinated function of these metabolites in low temperature conditions (**Fig. 4**; **Supplementary Table S4**).

In addition, galactinol showed a very high correlation (0.82) with seed longevity in high light intensity (HL) (**Fig. 2D**). These observations strongly indicate the importance of these metabolites in changing environments. Raffinose family oligosaccharides (RFOs) were discovered several decades ago to accumulate in the late stages of soybean (*Glycine max*) seed maturation and desiccation, indicating that they may play a role as protectants (Castillo et al. 1990). The mechanisms of protection remains to be elucidated, although two have been suggested: (i) raffinose can contribute with sucrose and stachyose to the increase of viscosity of the inner parts of the seed, thus maintaining a low but safe hydrated environment (glassy matrix) (Bernal-Lugo and Leopold 1995, Hoekstra et al. 2001); or (ii) RFOs can stabilize liposomes in the membrane to protect cellular integrity during desiccation (Hincha et al. 2003). More recently, raffinose was suggested as a storage metabolite in Arabidopsis seeds, accumulating during desiccation and readily consumed upon imbibition (Baud et al. 2002, Fait et al. 2006). RFOs are important in various stress tolerance defense mechanisms (reviewed by ElSayed et al. 2014). These authors showed that RFOs can act as antioxidants to counteract the accumulation of reactive oxygen species (ROS) under stress conditions. The seed aging process is accompanied by an increase in oxidative stress, such as DNA damage accumulation (Waterworth et al. 2010). It can

therefore be hypothesized that galactinol plays a role in oxidative stress protection during seed aging in high light intensity. This is in agreement with a recent report on the correlation between galactinol and seed longevity (de Souza Vidigal 2015).

Integration of phenotypes, metabolites and transcripts

As 'omics' technologies develop, an integrated view spanning the multiple levels of the cellular control hierarchy will be most informative and can help in proposing novel hypotheses and directions for future research. In recent years, a number of studies on higher plants have begun to combine extensive transcript and metabolite data sets to understand the underlying regulatory processes. Networks are largely used on the basis of co-response of different transcripts, metabolites and proteins (Toubiana et al. 2013). By using the 'guilt by association' principle, genes in a metabolic pathway that are co-ordinately regulated can be investigated, and therefore novel gene annotations can be achieved (Usadel et al. 2005, Bassel et al. 2011). For example, by correlating metabolites and the most differentially changed transcripts, Osorio et al. (2012) found correlations of functional gene categories with metabolite compound classes, and suggested the importance of some organic acids in pepper fruit development. The integrated analysis of metabolome and transcriptome data sets in citrus showed correlations of certain sugars/organic acids and the corresponding enzymes in the catalytic reactions, and denoted that the reactions of sugars are more prevalent in the flesh, while the reactions of organic acids are more active in the rind and have greater influence on the metabolic network (Ding et al. 2015). In Arabidopsis seeds, Angelovici et al. (2009) addressed the response of seed metabolism and transcription to developmentally inducible lysine metabolism. A large set of protein synthesis genes including ribosomal-related genes, translation initiation and elongation factor genes, and genes belonging to networks of amino acids and sugar metabolism were regulated. These authors concluded that the inducible lysine metabolism was primarily associated with altered expression of genes belonging to networks of amino acids and sugar metabolism.

The networks that we produced increased our understanding of the metabolite changes of seeds developed under low temperature (**Fig. 4**). For instance, the correlation network identified the co-ordinated regulation of the RFO cluster. In addition, some amino acids, e.g. alanine and GABA, were only loosely linked to other amino acid clusters in the network (**Fig. 4**), suggesting a stand-alone regulation of their level in response to the stress. However, correlation-based analysis faces some limitations. First, regulation of seed performance at the transcript, metabolite and protein levels shows temporal and spatial dependence. Secondly, data are often collected at a steady state (dry seeds in our case), while metabolic pathway flux can change substantially without perturbation of the final concentration at steady state, and therefore data collected at a single time point could not always reflect the real situation (Wang et al. 2015). Finally, since correlation alone provides no proof of causality, the putative causal genes need further validation such as demonstrated by Hannah et al. (2010).

Nevertheless, our correlation-based integrative analysis was utilized to identify possible genes and/or metabolites associated with seed performance traits. Interestingly, *MFT* correlated highly with DSDS50. *MFT* is a homolog of *FLOWERING LOCUS T* (*FT*) and *TERMINAL FLOWER 1* (*TFL1*) (Kobayashi et al. 1999), and belongs to the phosphatidylethanolamine-binding protein (PEBP) family. This protein family has been identified in numerous tissues in a wide variety of organisms, including bacteria, yeast, nematodes, mammals and plants (Schoentgen and Jolles 1995). The various functions described for members of this family include acceleration of flowering (Yoo et al. 2004), the control of the morphological switch between shoot growth and flower structures (Che et al. 2006) and the regulation of several signaling pathways such as the mitogen-activated protein (MAP) kinase pathway (Banfield and Brady 2000). Vaistij et al. (2013) concluded that *MFT* promoted dormancy during seed development by acting downstream of or parallel to the ABA and gibberellin response pathways, as in *mft-2* elevated ABA coincided with an increase in the expression of *REPRESSOR-OF-GA* (*RGA*). A wheat homolog of MFT was shown to be up-regulated in dormant seeds grown at low temperature (13°C) (Nakamura et al. 2011). This agrees with our results showing that, in *A. thaliana*, *MFT* expression was up-regulated (approximately 1.5-fold, data not shown) when seeds had matured at a low temperature (15°C), and its expression is highly correlated with DSDS50 levels (**Supplementary Table S8**). This suggests a role for *MFT* in low temperature signaling. Further characterization of the underlying molecular mechanism will be of great interest.

In conclusion, the integrative analysis of phenotypes, metabolites and transcripts on seeds matured in different environmental conditions was shown to be well suited for revealing underlying networks and generating new hypotheses. Understanding how environments regulate seed performance is of great importance in the agricultural area as field-grown crops often experience unpredictable environments. Hence, knowledge of how environments affect seed quality at the genetic and metabolite levels could provide information for practical applications such as post-harvest treatment of seeds and, perhaps more importantly, for prediction of seed performance.

Materials and Methods

Plant materials

Ler, NIL*DOG1*-Cvi, NIL*DOG2*-Cvi, NIL*DOG3*-Cvi, NIL*DOG6*-Kas-2, NIL*DOG22*-An-1 and the *dog1-1* mutant are described in He et al. (2014)

Growth conditions

Plant growth conditions before flowering are as described in He et al. (2014). In brief, three replicates per genotype were grown in controlled conditions [20°C/18°C (day/night) under a 16 h photoperiod and 70% relative humidity] using Rockwool blocks (4 × 4 cm) watered with a standard solution as describe by He et al. (2014). Upon the start of flowering, half of the plants were transferred to 15°C or 0 mM nitrate (N0), low light intensity (LL) or high light intensity (HL), while all the other conditions stayed the same, and the other half were maintained at 20°C, 5 mM nitrate and SL (control condition). Plants of each condition were grown in three blocks, with five plants per block.

Seed phenotyping

Germination was scored using the Germinator package (Joosen et al. 2010). To measure seed dormancy level (DSDS50: days of seed dry storage required to reach 50% germination), germination tests were performed weekly until all seed batches germinated >90%. Testing of germination under stress conditions was performed on fully after-ripened seeds. An accelerated aging test was carried out to measure seed longevity. The detailed methods of measuring DSDS50 and longevity are illustrated in He et al. (2014). The germination stress conditions were 10°C, 30°C, −0.8 MPa mannitol (Sigma-Aldrich), 125 mM NaCl (Sigma-Aldrich) and 0.2 μM ABA (Duchefa Biochemie).

G_{max} at the end of the germination assay and the rate of germination (t_{50}) values were extracted from the germination assay using the Germinator package (Joosen et al. 2010). For the germination at 30°C, in mannitol, NaCl and ABA, G_{max} was used as there were suficient variations. For germination at 10 and 22°C (as control), G_{max} of most genotypes reached 100%; therefore, to better distinguish the small differences between genotypes, t_{50} was used for those two conditions.

Metabolite extraction and derivatization methods

The metabolite extraction was performed on dry mature seeds of L*er*, NIL*DOG1* and *dog1-1* based on a previously described method (Roessner et al. 2000) with some modifications. For each genotype, metabolite extractions were performed on three biological replicates. For each sample, 5 mg of seeds pre-cooled in liquid nitrogen were homogenized in 2 ml tubes with two metal balls (2.5 mm) using a micro dismembrator (Mo Bio Laboratory). A 233 μl aliquot of methanol/chloroform (4 : 3) was added, together with 50 μl of standard (0.13 mg ml^{-1} ribitol) and mixed thoroughly. After 10 min of sonicationm 66 μl of MQ water was added to the mixture followed by vortexing and centrifugation (5 min, 15,000 r.p.m.). The methanol phase was collected in a glass vial. Then 166 μl methanol/chloroform (1 : 1) was added to the remaining organic phase and kept on ice for 10 min. A 66 μl aliquot of MQ water was added followed by vortexing and centrifugation (5 min, 15,000 r.p.m.). Again the methanol phase was collected and mixed with the previously collected phase and 60 μl was dried overnight using a speedvac (room temperature, Savant SPD121). Dried samples were derivatized online as described by Lisec et al. (2006) using a Combi PAL autosampler (CTC Analytics). The derivatized samples were analyzed by a gas chromatography–mass spectrometry–time-of-flight (GC-TOF-MS) system consisting of an Optic 3 high-performance injector (ATAS) and an Agilent 6890 gas chromatograph (Agilent Technologies) coupled to a Pegasus III time-of-flight mass spectrometer (Leco Instruments). A 2 μl aliquot of each sample was introduced into the injector. The details of the GC-TOF-MS method are described by Carreno-Quintero et al. (2012) with some minor modifications. The detector voltage was set at 1,650 V.

GC-MS data processing methods

Raw data were processed using the chromaTOF software 2.0 (Leco instruments) and further processed using the Metalign software (Lommen 2009), to extract and align the mass signals. A signal to noise ratio of 2 was used. The output was further processed by the Metalign Output Transformer (METOT; Plant Research International, Wageningen) and mass signals that were present in <3 samples were discarded. Centrotypes were created using the MSclust program (Tikunov et al. 2012) to reduce signal redundancy. The mass spectra of these centrotypes were used for the identification by matching to an in-house-constructed library and the NIST05 (National Institute of Standards and Technology; http://www.nist.gov/srd/mslist.htm) libraries. This identification is based on spectra similarity and comparison of retention indices calculated by using a third-order polynomial function (Strehmel et al. 2008). Forty-one metabolites were identified by matching to the in-house-constructed library and the NIST05 library. β-D-Methylfructofuranoside (4TMS), glucopyranose (5TMS), hydroxyethyl-methanamine (2TMS), pentitol (5TMS), pentonic acid (5TMS) and unknown metabolites matched with the NIST05 library but were not confirmed using standards.

ABA extraction and detection method

The ABA extraction and detection method is as described in He et al. (2014). In brief, 10 mg of dry seeds were ground and extracted with methanol/water/acetic acid (80 : 19 : 1) and [$^2H_6$]ABA as internal standard. The tubes were vortexed and sonicated for 10 min. Samples were centrifuged for 10 min and the liquid phase was transferred to a glass vial. The samples were re-extracted with methanol/water/acetic acid (80 : 19 : 1). Both fractions were combined in the vial and dried in a speedvac. The residue was dissolved in 100 μl of methanol/acetic acid 99 : 1 (v/v) and 900 μl of 1% acetic acid in water. The samples were loaded on HLB columns (Oasis®, Waters, 30 mg 1 ml) which were previously equilibrated. The columns/samples were washed with 1 ml of methanol/water/acetic acid (10 : 89 : 1). After washing, 1 ml of methanol/water/acetic acid (80 : 19 : 1) was added to elute the samples. The samples were dried in a vacuum concentrator and re-suspended in 100 μl of water/acetonitrile/formic acid (94.9 : 5 : 0.1). ABA analysis was performed with a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospray ionization source and coupled to an Acquity UPLC system (Waters).

PCA and correlation analysis

PCA and correlation analysis were performed using the online web tool MetaboAnalyst 3.0, http://www.metaboanalyst.ca/ (Xia et al. 2015). For the PCA, all the 124 detected metabolites in all the samples (seven genotypes, three biological replicates) were used. The correlation analysis was performed between the 41 identified primary metabolites, ABA levels and seed performance. Data of changed maturation environments were compared with their own controls (i.e. 15°C vs. 20°C, N0 vs. N5, LL vs. SL, and HL vs. SL).

Cluster analysis

The phenotypic data used for cluster analysis were 16 plant and seed performance traits on seven genotypes using the averages of three biological replicates. Forty-one metabolites which were identified by matching to the in-house-constructed library and the NIST05 library and five metabolites matched with the NIST05 library but not confirmed using standards [β-D-methylfructofuranoside (4TMS), glucopyranose (5TMS), hydroxyethyl-methanamine (2TMS), pentitol (5TMS) and pentonic acid (5TMS)] as well as the saturated metabolite sucrose were used for metabolic network construction. Thus, in total, the metabolite data used for cluster analysis were 47 metabolites on seven genotypes using the averages of three biological replicates. Cluster analysis was performed with the R statistical programming environment (R Core Team 2013), using the 'cluster' package. Samples were clustered based on a Partitioning Around Medoids (pam) algorithm. The distance matrix was calculated by using the formula $1 - r^2$, where r is the Spearman's rank correlation coefficient. The number of clusters was identified based on rule of thumb by the formula $k = \sqrt{(n/2)}$, where k is the number of clusters and n is the total number of elements used (genotypes). Standard temperature (20°C) is the control of both 15°C and N0, and standard light intensity (SL) is the control of both LL and HL.

Network construction

The same 47 metabolites as used for the cluster analysis were used for metabolic network construction. Metabolic networks were built for L*er*, NIL*DOG1*, NIL*DOG2*, NIL*DOG3*, NIL*DOG6* and NIL*DOG22* in different maturation environments. The SL/LL combined network used the average of metabolite contents of each metabolite in both environments. Spearman's rank correlation was used with the following threshold: $r > |0.49|$ and $P < 0.001$. The correlation matrix was calculated and converted to a network format using the 'igraph' package from the R statistical programming environment (R Core Team 2013). Networks were analyzed using Cytoscape software, version 3.1.0 (http://chianti.ucsd.edu/cytoscape-3.1.0/).

Microarray experiment

RNA isolation. RNA was isolated from NIL*DOG1*, NIL*DOG3* and NIL*DOG6* seeds, which were matured in low temperature (15°C), N0 and LL regimes.

Total RNA was extracted according to the hot borate protocol described by Maia et al. (2011). In brief, 3–3.5 mg of seeds for each treatment were homogenized and mixed with 800 μl of extraction buffer containing dithiothreitol (DTT) and PVP40 which had been heated to 80°C. Proteinase K was added and incubated for 15 min at 42°C. After adding 2 M KCl, the samples were incubated on ice for 30 min and centrifuged. Ice-cold 8 M LiCl was added to the supernatant and the tubes were incubated overnight on ice. After centrifugation, the pellets were washed with ice-cold 2 M LiCl and centrifuged for 10 min. The

pellets were re-suspended in 80 µl of water. The samples were phenol chloroform extracted, DNase treated and further purified. RNA quality and concentration were assessed by agarose gel electrophoresis and Nanodrop ND1000 spectrophotometry.

Microarray analysis. The quality control, RNA labeling, hybridization and data extraction were performed at ServiceXS B.V. Labeled single-stranded cDNA (ss-cDNA) was synthesized using the Affymetrix NuGEN Ovation PicoSL WTA v2 kit and Biotin Module using 50 ng of total RNA as template. The fragmented ss-cDNA was utilized for the hybridization on the Affymetrix ARAGene 1.1ST Array plate. The Affymetrix HWS Kit was used for the hybridization, washing and staining of the plate. Scanning of the array plates was performed using the Affymetrix GeneTitan scanner. All procedures were performed according to the instructions of the manufacturers (nugen.com and affymetrix.com). The resulting data were analyzed using the R statistical programming environment (R Core Team 2013) and the Bioconductor packages (Gentleman et al. 2004). The data were normalized using the RMA algorithm (Irizarry et al. 2003) utilizing the TAIRG v17 cdf file (http://brainarray.mbni.med.umich.edu). Significant differential expression changes were computed using the limma package (Smyth 2004) and *P*-values were adjusted for multiple testing with the Benjamini and Hochberg method to control for false positives (Benjamini and Hochberg 1995). Significant (adjusted *P* value < 0.05) average expression changes >2-fold were considered in this study to be differential and were analyzed further. Over-representation analysis was performed using Pageman (Carrari et al. 2006). Microarray quality and reproducibility data are presented in **Supplementary Fig. S2**. The microarray data described in this article have been deposited in the National Center for Biotechnology Information's Gene Expression Omnibus (accession No. GSE75837).

Integrated analysis of phenotype, metabolome and transcriptome

Correlation analysis was performed on expressed transcripts [2log(intensities)>5 for at least one sample], 124 detected metabolites and 16 plant and seed performance traits for NIL*DOG1*, NIL*DOG3* and NIL*DOG6* together, with R software, version 3.0.2. Correlation was determined using Spearman's rank correlation coefficient (*r*). The 'rcorr' function from the Hmisc package was used for calculating the *P*-value; *P*-values were adjusted (adjusted *P*-value) for multiple testing with the Benjamini and Hochberg method to control for false positives (Benjamini and Hochberg 1995). An adjusted *P*-value < 0.05 is regarded as significant. *Z*-score transformation was applied in figure plotting in **Fig. 6**.

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the Dutch Technology Foundation (STW) which is the applied science division of the Netherlands Organization for Scientific Research [STW; grant No. STW11314 to H.H]; the Albert Katz International School [PhD fellowship to A.B.]; the Israeli Science Foundation [grant No. 1471/13 for the work in the laboratory of A.F.]; Bio4Energy, a Strategic Research Environment appointed by the Swedish government [support for J.H.].

Acknowledgements

We thank Professor Fred van Eeuwijk (Wageningen University) for critical reading of the manuscript and fruitful input.

Disclosures

The authors have no conflicts of interest to declare.

References

Alboresi, A., Gestin, C., Leydecker, M.T., Bedu, M., Meyer, C. and Truong, H.N. (2005) Nitrate, a signal relieving seed dormancy in Arabidopsis. *Plant Cell Environ.* 28: 500–512.

Almaas, E., Oltvai, Z.N. and Barabasi, A.L. (2005) The activity reaction core and plasticity of metabolic networks. *PLoS Comput. Biol.* 1: 557–563.

Angelovici, R., Fait, A., Zhu, X.H., Szymanski, J., Feldmesser, E., Fernie, A.R., et al. (2009) Deciphering transcriptional and metabolic networks associated with lysine metabolism during Arabidopsis seed development. *Plant Physiol.* 151: 2058–2072.

Banfield, M.J. and Brady, R.L. (2000) The structure of *Antirrhinum centroradialis* protein (CEN) suggests a role as a kinase regulator. *J. Mol. Biol.* 297: 1159–1170.

Bassel, G.W., Lan, H., Glaab, E., Gibbs, D.J., Gerjets, T., Krasnogor, N., et al. (2011) Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proc. Natl. Acad. Sci. USA* 108: 9709–9714.

Baud, S., Boutin, J.P., Miquel, M., Lepiniec, L. and Rochat, C. (2002) An integrated overview of seed development in Arabidopsis thaliana ecotype WS. *Plant Physiol. Biochem.* 40: 151–160.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B: Methodol.* 57: 289–300.

Bernal-Lugo, I. and Leopold, A. (1995) Seed stability during storage: raffinose content and seed glassy state. *Seed Sci. Res.* 5: 75–80.

Borner, G.H.H., Lilley, K.S., Stevens, T.J. and Dupree, P. (2003) Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. *Plant Physiol.* 132: 568–577.

Cadman, C.S., Toorop, P.E., Hilhorst, H.W. and Finch-Savage, W.E. (2006) Gene expression profiles of Arabidopsis Cvi seeds during dormancy cycling indicate a common underlying dormancy control mechanism. *Plant J.* 46: 805–822.

Carrari, F., Baxter, C., Usadel, B., Urbanczyk-Wochniak, E., Zanor, M.I., Nunes-Nesi, A., et al. (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol.* 142: 1380–1396.

Carreno-Quintero, N., Acharjee, A., Maliepaard, C., Bachem, C.W., Mumm, R., Bouwmeester, H., et al. (2012) Untargeted metabolic quantitative trait loci analyses reveal a relationship between primary metabolism and potato tuber quality. *Plant Physiol.* 158: 1306–1318.

Castillo, E.M., Delumen, B.O., Reyes, P.S. and Delumen, H.Z. (1990) Raffinose synthase and galactinol synthase in developing seeds and leaves of legumes. *J. Agric. Food Chem.* 38: 351–355.

Che, P., Lall, S., Nettleton, D. and Howell, S.H. (2006) Gene expression programs during shoot, root, and callus development in Arabidopsis tissue culture. *Plant Physiol.* 141: 620–637.

Chen, M., MacGregor, D.R., Dave, A., Florance, H., Moore, K., Paszkiewicz, K., et al. (2014) Maternal temperature history activates *Flowering Locus T* in fruits to control progeny dormancy according to time of year. *Proc. Natl. Acad. Sci. USA* 111: 18787–18792.

Cho, S.K., Ryu, M.Y., Seo, D.H., Kang, B.G. and Kim, W.T. (2011) The Arabidopsis RING E3 ubiquitin ligase AtAIRP2 plays combinatory roles with AtAIRP1 in abscisic acid-mediated drought stress responses. *Plant Physiol.* 157: 2240–2257.

Cook, D., Fowler, S., Fiehn, O. and Thomashow, M.F. (2004) A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. *Proc. Natl. Acad. Sci. USA* 101: 15243–15248.

de Souza Vidigal, D. (2015) Adaptation and Acclimation of Seed Performance. PhD thesis, Wageningen University, Wageningen, The Netherlands.

Ding, Y., Chang, J., Ma, Q., Chen, L., Liu, S., Jin, S., et al. (2015) Network analysis of postharvest senescence process in Citrus fruits revealed by transcriptomic and metabolomic profiling. *Plant Physiol.* 168: 357–376.

Donohue, K. (2009) Completing the cycle: maternal effects as the missing link in plant life histories. *Philos. Trans. R. Soc. B: Biol. Sci.* 364: 1059–1074.

Donohue, K., Dorn, L., Griffith, C., Kim, E., Aguilera, A., Polisetty, C.R., et al. (2005) Environmental and genetic influences on the germination of *Arabidopsis thaliana* in the field. *Evolution* 59: 740–757.

ElSayed, A.I., Rafudeen, M.S. and Golldack, D. (2014) Physiological aspects of raffinose family oligosaccharides in plants: protection against abiotic stress. *Plant Biol.* 16: 1–8.

Fait, A., Angelovici, R., Less, H., Ohad, I., Urbanczyk-Wochniak, E., Fernie, A.R., et al. (2006) Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. *Plant Physiol.* 142: 839–854.

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.

Fait, A., Nesi, A.N., Angelovici, R., Lehmann, M., Pham, P.A., Song, L.H., et al. (2011) Targeted enhancement of glutamate-to-γ-aminobutyrate conversion in Arabidopsis seeds affects carbon–nitrogen balance and storage reserves in a development-dependent manner. *Plant Physiol.* 157: 1026–1042.

Forde, B.G. and Lea, P.J. (2007) Glutamate in plants: metabolism, regulation, and signalling. *J. Exp. Bot.* 58: 2339–2358.

Fu, J., Keurentjes, J.J.B., Bouwmeester, H., America, T., Verstappen, F.W.A., Ward, J.L., et al. (2009) System-wide molecular evidence for phenotypic buffering in Arabidopsis. *Nat. Genet.* 41: 166–167.

Galili, G., Avin-Wittenberg, T., Angelovici, R. and Fernie, A.R. (2014) The role of photosynthesis and amino acid metabolism in the energy status during seed development. *Front. Plant Sci* 5: 447.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.

Gojon, A., Krouk, G., Perrine-Walker, F. and Laugier, E. (2011) Nitrate transceptor(s) in plants. *J. Exp. Bot.* 62: 2299–2308.

Guo, F.O., Young, J. and Crawford, N.M. (2003) The nitrate transporter AtNRT1.1 (CHL1) functions in stomatal opening and contributes to drought susceptibility in Arabidopsis. *Plant Cell* 15: 107–117.

Guo, F.Q., Wang, R.C. and Crawford, N.M. (2002) The Arabidopsis dual-affinity nitrate transporter gene *AtNRT1.1* (*CHL1*) is regulated by auxin in both shoots and roots. *J. Exp. Bot.* 53: 835–844.

Gutterman, Y. (1974) The influence of the photoperiodic regime and red–far red light treatments of *Portulaca oleracea* L. plants on the germinability of their seeds. *Oecologia* 17: 27–38.

Hannah, M.A., Caldana, C., Steinhauser, D., Balbo, I., Fernie, A.R. and Willmitzer, L. (2010) Combined transcript and metabolite profiling of Arabidopsis grown under widely variant growth conditions facilitates the identification of novel metabolite-mediated regulation of gene expression. *Plant Physiol.* 152: 2120–2129.

He, H., de Souza Vidigal, D., Snoek, L.B., Schnabel, S., Nijveen, H., Hilhorst, H., et al. (2014) Interaction between parental environment and genotype affects plant and seed performance in Arabidopsis. *J. Exp. Bot.* 65: 6603–6615.

Hincha, D.K., Zuther, E. and Heyer, A.G. (2003) The preservation of liposomes by raffinose family oligosaccharides during drying is mediated by effects on fusion and lipid phase transitions. *Biochim. Biophys. Acta* 1612: 172–177.

Hoekstra, F.A., Golovina, E.A. and Buitink, J. (2001) Mechanisms of plant desiccation tolerance. *Trends Plant Sci.* 6: 431–438.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249–264.

Joosen, R.V.L., Kodde, J., Willems, L.A.J., Ligterink, W., van der Plas, L.H.W. and Hilhorst, H.W.M. (2010) GERMINATOR: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *Plant J.* 62: 148–159.

Kanno, Y., Hanada, A., Chiba, Y., Ichikawa, T., Nakazawa, M., Matsui, M., et al. (2012) Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. *Proc. Natl. Acad. Sci. USA* 109: 9653–9658.

Karssen, C.M., Brinkhorstvanderswan, D.L.C., Breekland, A.E. and Koornneef, M. (1983) Induction of dormancy during seed development by endogenous abscisic acid: studies on abscisic acid deficient genotypes of *Arabidopsis thaliana* (L) Heynh. *Planta* 157: 158–165.

Kendall, S.L., Hellwege, A., Marriot, P., Whalley, C., Graham, I.A. and Penfield, S. (2011) Induction of dormancy in Arabidopsis summer annuals requires parallel regulation of *DOG1* and hormone metabolism by low temperature and CBF transcription factors. *Plant Cell* 23: 2568–2580.

Kigel, J., Gibly, A. and Negbi, M. (1979) Seed germination in *Amaranthus retroflexus* L. as affected by the photoperiod and age during flower induction of the parent plants. *J. Exp. Bot.* 30: 997–1002.

Kim, J.S., Mizoi, J., Kidokoro, S., Maruyama, K., Nakajima, J., Nakashima, K., et al. (2012) Arabidopsis GROWTH-REGULATING FACTOR7 functions as a transcriptional repressor of abscisic acid- and osmotic stress-responsive genes, including *DREB2A*. *Plant Cell* 24: 3393–3405.

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M. and Araki, T. (1999) A pair of related genes with antagonistic roles in mediating flowering signals. *Science* 286: 1960–1962.

Krouk, G., Lacombe, B., Bielach, A., Perrine-Walker, F., Malinska, K., Mounier, E., et al. (2010) Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. *Dev. Cell* 18: 927–937.

Lisec, J., Schauer, N., Kopka, J., Willmitzer, L. and Fernie, A.R. (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* 1: 387–396.

Liu, K.H., Huang, C.Y. and Tsay, Y.F. (1999) CHL1 is a dual-affinity nitrate transporter of Arabidopsis involved in multiple phases of nitrate uptake. *Plant Cell* 11: 865–874.

Lommen, A. (2009) MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Anal. Chem.* 81: 3079–3086.

Maia, J., Dekkers, B.J.W., Provart, N.J., Ligterink, W. and Hilhorst, H.W.M. (2011) The re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds and its associated transcriptome. *PLoS One* 6: e29123.

Matakiadis, T., Alboresi, A., Jikumaru, Y., Tatematsu, K., Pichon, O., Renou, J.P., et al. (2009) The Arabidopsis abscisic acid catabolic gene *CYP707A2* plays a key role in nitrate control of seed dormancy. *Plant Physiol.* 149: 949–960.

Meraviglia, G., Romani, G. and Beffagna, N. (1996) The *chl1* Arabidopsis mutant impaired in the nitrate-inducible NO_3^- transporter has an acidic intracellular pH in the absence of nitrate. *J. Plant Physiol.* 149: 307–310.

Michaeli, S. and Fromm, H. (2015) Closing the loop on the GABA shunt in plants: are GABA metabolism and signaling entwined?. *Front Plant Sci.* 6: 419.

Miflin, B.J. and Habash, D.Z. (2002) The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. *J. Exp. Bot.* 53: 979–987.

Munir, J., Dorn, L.A., Donohue, K. and Schmitt, J. (2001) The effect of maternal photoperiod on seasonal dormancy in *Arabidopsis thaliana* (Brassicaceae). *Amer. J. Bot.* 88: 1240–1249.

Nakamura, S., Abe, F., Kawahigashi, H., Nakazono, K., Tagiri, A., Matsumoto, T., et al. (2011) A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. *Plant Cell* 23: 3215–3229.

Nelson, D.C., Flematti, G.R., Riseborough, J.-A., Ghisalberti, E.L., Dixon, K.W. and Smith, S.M. (2010) Karrikins enhance light responses during germination and seedling development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 107: 7095–7100.

Osorio, S., Alba, R., Nikoloski, Z., Kochevenko, A., Fernie, A.R. and Giovannoni, J.J. (2012) Integrative comparative analyses of transcript and metabolite profiles from pepper and tomato ripening and development stages uncovers species-specific patterns of network regulatory behavior. *Plant Physiol.* 159: 1713–1729.

R Core Team (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Roessner, U., Wagner, C., Kopka, J., Trethewey, R.N. and Willmitzer, L. (2000) Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography–mass spectrometry. *Plant J* 23: 131–142.

Schmuths, H., Bachmann, K., Weber, W.E., Horres, R. and Hoffmann, M.H. (2006) Effects of preconditioning and temperature during germination of 73 natural accessions of *Arabidopsis thaliana*. *Ann. Bot.* 97: 623–634.

Schoentgen, F. and Jolles, P. (1995) From structure to function—possible biological roles of a new widespread protein family binding hydrophobic ligands and displaying a nucleotide-binding site. *FEBS Lett.* 369: 22–26.

Schulze, W., Schulze, E.D., Stadler, J., Heilmeier, H., Stitt, M. and Mooney, H.A. (1994) Growth and reproduction of *Arabidopsis thaliana* in relation to storage of starch and nitrate in the wild type and in starch-deficient and nitrate-uptake-deficient mutants. *Plant Cell Environ.* 17: 795–809.

Smyth, G.K. (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3: Article3.

Strehmel, N., Hummel, J., Erban, A., Strassburg, K. and Kopka, J. (2008) Retention index thresholds for compound matching in GC–MS metabolite profiling. *J. Chromatogr. B* 871: 182–190.

Tikunov, Y.M., Laptenok, S., Hall, R.D., Bovy, A. and de Vos, R.C.H. (2012) MSClust: a tool for unsupervised mass spectra extraction of chromatography–mass spectrometry ion-wise aligned data. *Metabolomics* 8: 714–718.

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.

Usadel, B., Kuschinsky, A.M., Steinhauser, D. and Pauly, M. (2005) Transcriptional co-response analysis as a tool to identify new components of the wall biosynthetic machinery. *Plant Biosyst.* 139: 69–73.

Usadel, B., Nagel, A., Steinhauser, D., Gibon, Y., Blasing, O., Redestig, H., et al. (2006) PageMan: an interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. *BMC Bioinformatics* 7: 535.

Vaistij, F.E., Gan, Y., Penfield, S., Gilday, A.D., Dave, A., He, Z., et al. (2013) Differential control of seed primary dormancy in Arabidopsis ecotypes by the transcription factor SPATULA. *Proc. Natl. Acad. Sci. USA* 110: 10866–10871.

Valluru, R. and Van den Ende, W. (2011) Myo-inositol and beyond—emerging networks under stress. *Plant Sci.* 181: 387–400.

Wang, H., Paulo, J., Kruijer, W., Boer, M., Jansen, H., Tikunov, Y., et al. (2015) Genotype–phenotype modeling considering intermediate level of biological variation: a case study involving sensory traits, metabolites and QTLs in ripe tomatoes. *Mol. Biosyst.* 11: 3101–3110.

Waterworth, W.M., Masnavi, G., Bhardwaj, R.M., Jiang, Q., Bray, C.M. and West, C.E. (2010) A plant DNA ligase is an important determinant of seed longevity. *Plant J.* 63: 848–860.

Xi, W.Y., Liu, C., Hou, X.L. and Yu, H. (2010) *MOTHER OF FT AND TFL1* regulates seed germination through a negative feedback loop modulating ABA signaling in Arabidopsis. *Plant Cell* 22: 1733–1748.

Xia, J., Sinelnikov, I.V., Han, B. and Wishart, D.S. (2015) MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Res.* 43: W251–W257.

Yoo, S.Y., Kardailsky, I., Lee, J.S., Weigel, D. and Ahn, J.H. (2004) Acceleration of flowering by overexpression of *MFT* (*MOTHER OF FT AND TFL1*). *Mol. Cells* 17: 95–101.

Zheng, Z.-L. (2009) Carbon and nitrogen nutrient balance signaling in plants. *Plant Signal. Behav.* 4: 584–591.