Supporting Information

Sulpice et al. 10.1073/pnas.0903478106

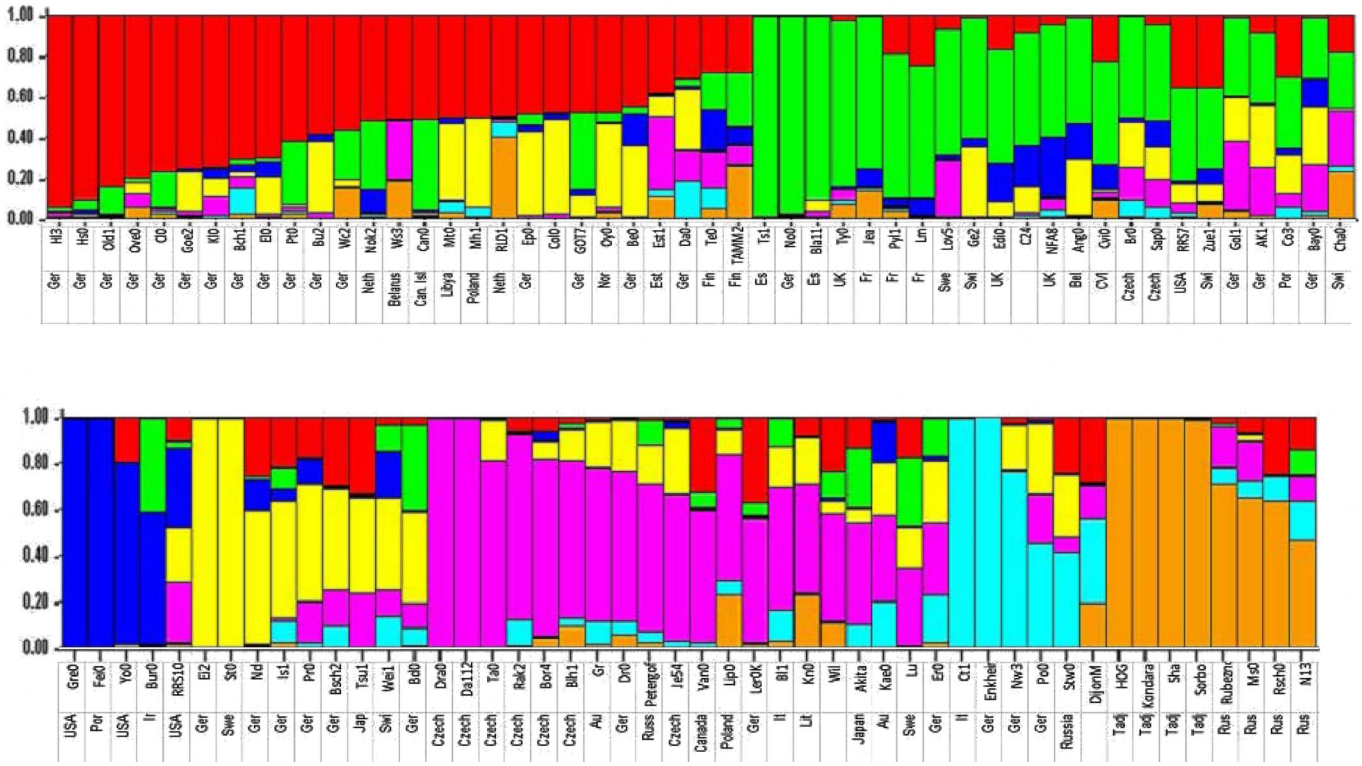


Fig. S1. Population structure of the population used in the study. The program Structure 2.1 (1) was used to determine population structure and assign accessions to subpopulations, using 419 markers that were randomly distributed and had <25% missing data. The optimal number of subpopulation was simulated by setting *K* (number of subpopulation) from 1 to 15. The length of burn-in period as well as MCMC reps after burn-in were set to 100,000 for each run, and each run was iterated 10 times. An accession was assigned to the subpopulation or group to which it showed the highest probability of membership. When *K* was varied from 1 to 15 the posterior probability [Ln P(D)] improved steadily for $K = 7–8$ and reached plateau for $K \geq 8$.

1. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. *Am J Hum Genet* 67:170–181.

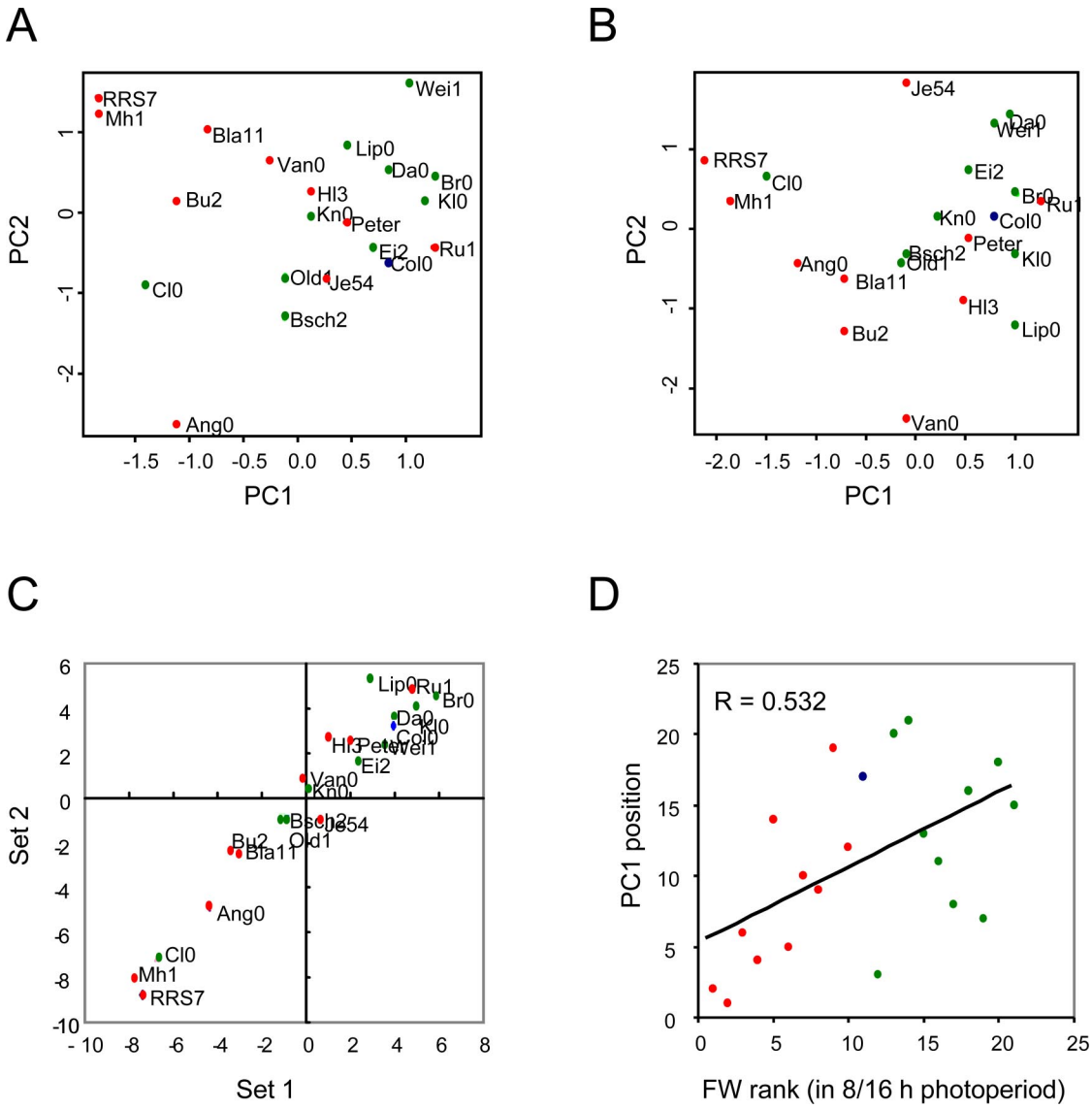
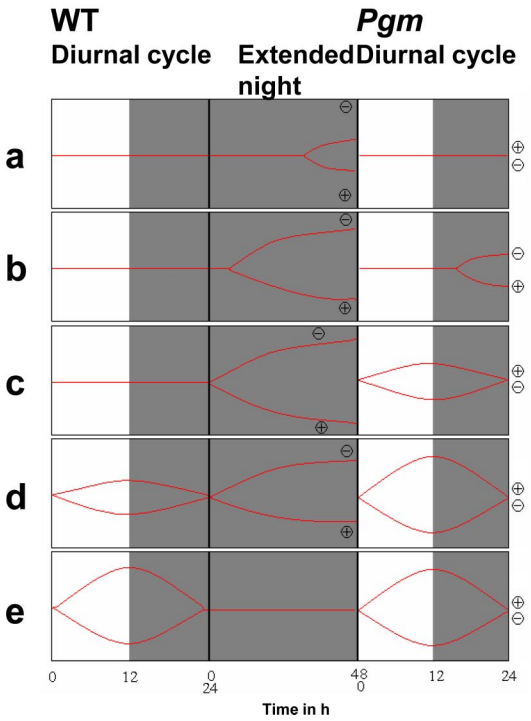


Fig. S2. Principle component analysis (PCA) of transcript levels at the end of the night for set 1 and set 2 of the C-responsive genes in 21 accessions. Set 1 contained 42 genes that show marked changes of transcripts at different times during the diurnal cycle or during the first hours of an extended night in wild-type plants. All these transcripts also respond to changes of the CO_2 concentration in wild-type Col0 and show accentuated responses in the starchless pgm mutant (1, 2). The other set included 52 genes whose transcripts change >log2 1.4 within 30 min of adding sucrose to carbon-starved seedlings (3). These presumably represent upstream components in the transcriptional response to sucrose. (*A*) PCA, using the 42 genes from for set 1. The PC1 and PC2 explained 41% and 11% of the variance, respectively. (*B*) PCA, using the 52 genes from for set 2. The PC1 and PC2 explained 34 and 14% of the variance, respectively. (*C*) Comparison of the separation of the 21 accessions in the PC1 of the PCA with genes of set 1 and set 2. (*D*) Relation between the loadings of the accessions in the PC1 and their rosette FW. Accessions were separated in 2 groups of slow (red circles) and fast (green circles) growing accessions from their behavior in an 8 h light/16 h dark regime. Col0 is depicted with a blue circle.

1. Blasing OE, et al. (2005) Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in Arabidopsis. *Plant Cell* 17:3257–3281.
2. Usadel B, et al. (2008) Global transcript levels respond to small changes of the carbon status during a progressive exhaustion of carbohydrates in Arabidopsis rosettes. *Plant Physiol* 146:1834–1861.
3. Osuna D, et al. (2007) Temporal responses of transcripts, enzyme activities and metabolites after adding sucrose to carbon-deprived Arabidopsis seedlings. *Plant J* 49:463–491.

A



B

Range of response	Genotype	Induced		Repressed	
Group a. Very low range	WT	$\lvert \max_{WT} - \min_{WT} \rvert < 0.5$			
	WT ExN	$0 - 6\text{ h} < 0.5$			
		$48\text{h} > 24\text{h} - 0.2 > 8\text{h} - 0.2$		$48\text{h} < 24\text{h} + 0.2 < 8\text{h} + 0.2$	
	pgm	$\lvert \max_{pgm} - \min_{pgm} \rvert < 0.5$			
Group b. Low range	WT	$\lvert \max_{WT} - \min_{WT} \rvert < 0.5$			
	WT ExN	min		max	
	pgm	$\lvert \max_{pgm} - \min_{pgm} \rvert < 1.0$			
Group c. Intermediate range	WT	$\lvert \max_{WT} - \min_{WT} \rvert < 0.5$			
	WT ExN	min		max	
	pgm	$1 < \lvert \max_{pgm} - \min_{pgm} \rvert < 0.5$			
Group d. High range	WT	max = ED	min = EN	max = EN	min = ED
	WT ExN	min		max	
	pgm	max = EN	min = ED	max = ED	min = EN
		$\lvert \max_{WT} - \min_{WT} \rvert < \lvert \max_{pgm} - \min_{pgm} \rvert$			
		$\lvert \max_{pgm} - \min_{pgm} \rvert > 2.0$			
Group e. Very high range	WT	$\lvert \max_{WT} - \min_{WT} \rvert ? \lvert \max_{pgm} - \min_{pgm} \rvert$			
	ExN	< 0.5			
	pgm	$\lvert \max_{pgm} - \min_{pgm} \rvert > 1.2$			

Fig. S3. Selection of the genes for sets 1 and 2. Selection criteria for genes in set 1 were based on the changes in transcript level in Col0 (WT) and pgm mutant in the normal diurnal cycle and in the extended night in WT. (*A*) Idealized representation of the behavior of the 5 classes of genes that were included in gene set 1. Five groups of sugar-responsive genes were designed very low range (a), low range (b), intermediate range (c), high range (d), and very high range (e) in respect to their response to changes in the C status of the rosettes. Genes that are induced under high sugar status show a curve shape that is labeled with (+), whereas genes that are repressed under high sugar conditions show a curve that is labeled with (−). A time axis is indicated beneath the figure, and periods of illumination and darkness are depicted as white and gray backgrounds, respectively. (*B*) Mathematical criteria used to fulfill the conditions described in *A*. All numerical values refer to a log2 scale used for display and comparison of ATH1 array data. Maximal (max) and minimal (min) correspond to the highest and lowest log2 values per gene, respectively. To generate the second and nonoverlapping set, the genes in set 1 were removed and a second set was then identified based on the response 30 min after adding sucrose to C-starved seedlings (1). The transcripts that increased or decreased most strongly in response to sucrose after 30 min were selected as set 2. They are presumably representatives of an early transcriptional response to sucrose. All genes selected for set 1 and 2 are listed in Table S3. ED, end of day; EN, end of night; ExN extended night.

1. Thompson AR, Vierstra RD (2005) Autophagic recycling: Lessons from yeast help define the process in plants. *Curr Opin Plant Biol* 8:165–173.

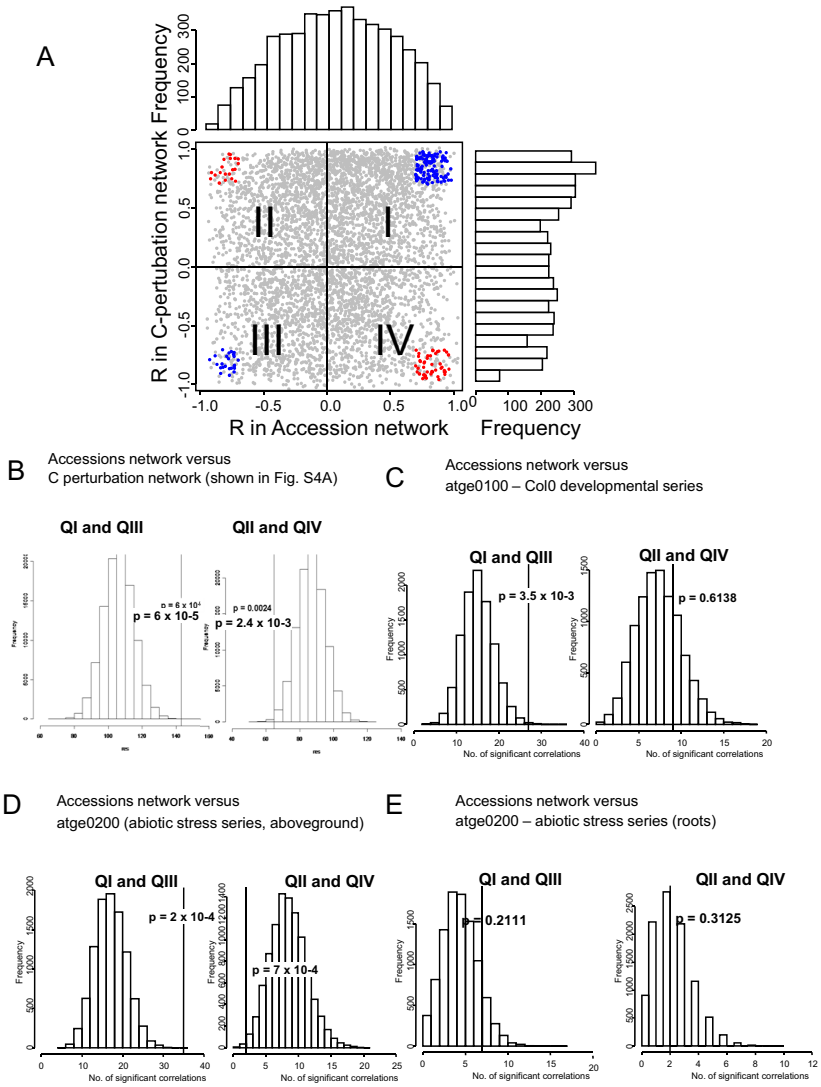


Fig. S4. Comparison of gene correlation networks generated by measuring transcript levels at the end of the night in 21 accessions, with networks obtained in Col0 containing different levels of endogenous sugars (C-perturbation network), and 3 publicly available datasets in the database csbdb (1). (*A*) Scatter plot of the R values of all pair-wise gene–gene correlations obtained by comparing the accession gene network and the C-perturbation network. The accession gene network is computed from the data shown in Fig. 2 and Table S5. The C-perturbation network is calculated from published ATH1 experiments in which 5-week-old rosettes of Col0 were subjected to 23 treatments that alter endogenous C, harvest at different times during the diurnal cycle in wild-type Col0 and the starchless pgm mutant, and a short extension of the night and illumination in the presence of different CO_2 concentrations in Col0 (2, 3). Changes of C make a major contribution to the global changes of transcript levels in this data set (2, 3). R values were calculated for all pair-wise correlations of the 94 genes in each correlation network using R (4). Gene pair-wise correlations that were significant in both networks at $P < 0.01$ and $R > 0.7$ are depicted in blue and red circles when they have the same (Quadrant I and III) or opposite (Quadrant II and IV) signs, respectively. (*B–E*) Evaluation of the significance of the enrichment of pair-wise correlations in the accession network (Fig. 2) compared with correlation networks obtained from in house and publically available expression array datasets for Col0. The datasets used to generate a C-perturbation network are listed in the legend of Table S3. R and *P* values were calculated for each gene pair in each network (*A*) and conserved (QI and QIII) and reversed (QII and QIV) significant gene pair-wise correlations identified using a threshold of $P < 0.01$ and $R > 0.7$. The number of significant correlations was compared with the randomly created population of correlations generated by shuffling 100,000 times the data in the correlation matrix of the C-responsive genes in the 21 accessions to calculate the *P* values for the enrichment in these quadrants. A similar approach was taken to compare the accession network with for 3 further data series, publicly available on the csbdb database (5). (*B*) Accession network versus the carbon perturbation network. (*C*) Accession network versus a developmental series for wild-type Col0. (*D*) Accession network versus an abiotic stress series for Col0 aboveground organs. (*E*) Accession network versus an abiotic stress series for Col0 roots.

1. Steinhauser D, Usadel B, Luedemann A, Thimm O, Kopka J (2004) CSB.DB: A comprehensive systems-biology database. *Bioinformatics* 20:3647–3651.
2. Blasing OE, et al. (2005) Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in Arabidopsis. *Plant Cell* 17:3257–3281.
3. Usadel B, et al. (2008) Global transcript levels respond to small changes of the carbon status during a progressive exhaustion of carbohydrates in Arabidopsis rosettes. *Plant Physiol* 146:1834–1861.
4. R Development Core Team (2008) R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna).
5. Steinhauser D, Usadel B, Luedemann A, Thimm O, Kopka J (2004) CSB.DB: A comprehensive systems-biology database. *Bioinformatics* 20:3647–3651.

Other Supporting Information Files

Table S1

Table S2

Table S3

Table S4

Table S5

Table S6