SUPPLEMENTARY FIGURES

Supplementary Figure 1. Distribution Plot of Significant Spearman's Correlation Coefficients

Supplementary Figure 1 Legend.

Grey bars represent numbers of significant correlations ($p < 0.05$) in wild-type plants. White bars represent numbers of significant correlations ($p < 0.05$) in Col wild-type plants. The distribution of correlation coefficient is nearly symmetrical for positive and negative correlations.

Supplementary Figure 2. Impact of Individual Mutants on Pearson Product-Moment Correlation between Metabolic Phenotypes

Supplementary Figure 2 Legend

(A) to **(E)** Impact of *ats1-1* (yellow-green cross) and *fatb-ko* (orange squares) on the correlations among fatty acids and their ratios.

(F), **(G)** and **(J)** Impact of *5-fcl* (blue x) on the correlations among leaf Gly, seed Gly and seed Ser.

(H) Impact of *tha1-1* (purple squares) on the correlation between Cys and Thr in seeds.

(I) Impact of *pig1-1* (purple x) on the correlation between Gln and Ile in seeds.

Black lines represent linear regressions before indicated mutant data were excluded. Pearson correlation coefficient *r* and *p*-value before and after indicated mutant data were excluded are shown in black and red, respectively.

Supplementary Figure 3. Hierarchical Clustering of 148 Samples by 81 Variables without Standardization

Supplementary Figure 3 Legend:

Hierarchical clustering was performed using Ward's minimum variance methods with JMP 6.0. Data from different variables were NOT standardized. Traits with a positive *z*-score or numeric code are shown in red squares; traits with a negative *z*-score or numeric code are shown in blue squares. Plants are color-coded as in Figure 2. Seven clusters are identical to those in Figure 2: *5-fcl*, *ats1-1*, *dpe2-1*, *fatb-ko*, *sex1-1* and *sex4-5*, *tha1-1*, extreme Col and Ws wild-type plants (pink). Two clusters are very similar to those in Figure 2: *pig1-1* and *tt7-1*. The only changes in these two clusters are: (1) the Ws wild type clustered with *pig1-1* (purple) in Figure 2 regroups with Col wild-type plants; (2) the Col wild type in the *npq1-2* cluster (red) in Figure 2 regroups with *tt7-3*. Three clusters have substantial changes: the cluster of *arc10* and *arc12* in Figure 2 disappears; *arc10* regroups with Ws wild-type plants; *arc12* regroups with Col wild-type plants and *npq1-2*; and *npq1-2* mixes with Col wild-type plants. Chlpt, chloroplast; HL, high light; num, number; var, variation.

SUPPLEMENTARY TEXT

Classification of Mutants *via k*-means Clustering

k-means clustering was used to provide independent statistical evaluation of the entire data set. The number of *k*-means clusters was set to 12 according to the number of hierarchical clusters. Similar groupings were obtained using *k*-means clustering: 9 out of 12 hierarchical clusters reappeared in *k*-means clustering (Supplementary Table 8). However, *arc10* split from *arc12* and grouped with *sex1-1* and *sex4-5* while *arc12* grouped with the *lkr-sdh* mutant and Ws wild-type plants (Supplementary Table 8).

Factors Contributing to Spurious Phenotypic Correlations

Some correlations were caused by easily identified artifacts or biases in the dataset. These include mathematical artifacts, such as the positive correlation between fatty acid ratio (16:3+*trans*-16:1d3)/(18:0+18:2) and fatty acid 16:3 (Supplementary Table 6). Some correlations were mostly caused by phenotypic and metabolic differences between the two ecotypes, such as the association between presences of an inflorescence and other phenotypic or metabolic variables and the association between leaf color and other phenotypic or metabolic variables (Supplementary Table 6). Other examples of ecotypic correlations were not as obvious, such as the correlation between leaf Val and leaf Thr and the correlation between seed Trp and seed Ser (Supplementary Table 6). This is consistent with the fact that Ws wild-type plants had significantly higher mol% of Val and lower mol% of Thr in leaves and higher mol% of Trp and Ser in seeds than Col wild-type plants (Student's *t*-test, $p < 0.05$) (Tables 2-5).

Some correlations were caused by, or contributed to, by individual mutants, as illustrated with ten pairs of metabolic variables in Supplementary Figure 1. The absolute values of the correlation coefficient were larger than 0.70 when the full data set was used and they dropped to near or below 0.50 when selected mutants were excluded (Supplementary Figure 1). Fatty acid mutants *fatb-ko* and *ats1-1* strongly contributed to the correlations the following variables: (16:3+*trans*-16:1d3)/(18:0+18:2) and 16:0/(18:1d9+18:1d11); 16:0/(18:1d9+18:1d11) and 16:3/18:2; 16:0/(18:1d9+18:1d11) and 18:2; 18:1d9 and 18:2; 18:3 and *cis*-16:1 (Supplementary Figure 2 panels A, B, C, D,

E). The *5-fcl* mutant caused the correlations among leaf Gly, seed Gly, and seed Ser, as shown in Supplementary Figure 1 (panels F, G and J). The mutant causing the correlation between seed Gln and seed Ile was *pig1-1* (Supplementary Figure 2 panel I).

Correlation of Morphological Traits between Two Photoperiods

Two photoperiods were used in this study: the 16-h light/8-h dark photoperiod and the 12-h light/12-h dark photoperiod. To study morphological correlations between the two photoperiods, Spearman's ρ correlation analysis was performed. Among six pairs of variables, inflorescence, leaf color, leaf shape and mature leaf size under two photoperiods had strong and significant positive correlations ($\rho > 0.60$, $p < 0.001$). Rosette size under two photoperiods had a significant but less strong positive correlation between the two photoperiods ($\rho = 0.40$, $p < 0.001$). Leaf color variation and leaf number under two photoperiods did not have significant correlations.