Statistical Analyses

For each array, the raw data comprise the logarithm of median feature pixel intensity at wavelengths 635 nm (red) and 532 nm (green). No background is subtracted. In the following description, log ratio refers to the differential expression between leaves and flowers. It is either log2(red/green) or log2(green/red) according to the experiment design.

Normalization:
An array-by-array normalization is performed to remove systematic biases. First, we exclude spots that are considered badly formed features. Then we perform a global intensity-dependent normalization using the loess procedure (see Yang et al., 2002) to correct the dye bias. Finally, for each block, the log-ratio median calculated over the values for the entire block is subtracted from each individual log-ratio value to correct print tip effects on each metablock.
Differential expression analysis:

To determine differentially expressed genes, we perform a paired t test on the log ratios. The number of observations per spot varies between two and four and is inadequate for calculating a gene-specific variance. For this reason, we assume that the variance of the log-ratios is the same for all genes, and we exclude 256 spots displaying extremes of variance (too small or too large). The raw P values are adjusted by the Bonferroni method, which controls the Family Wise Error Rate (FWER) (Chupeau et al., 1974).

For an example, see article: Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis