# Supplementary Methods, Section 1

# PLAGE in detail [1]

1. The expression profile of each gene is bought to a common scale by ensuring a distribution of mean zero and variance one.
2. For each pathway a matrix *Y* is created only containing gene expression data from the genes included in the pathway. In the matrix *Y* each column represents a sample and each row represents a gene.
3. We perform the singular value decomposition of matrix *Y*. This involves writing *Y* in the form:

*Y =WDC*

*W* = Matrix of eigenvectors ordered by size of corresponding eigenvalue. Each column is an eigenvector.

*D* = Diagonal matrix of eigenvalues ordered from largest to smallest.

*C* = Matrix of weights. Each column is a vector of coefficients for one sample indicating the overall level of each pathway.

1. The pathway scores for one pathway in all samples are given by the first row of *C*. The row corresponding to the largest eigenvalue and hence explaining the most variation in the pathway across the samples.

# References

[1] Tomfohr J, Lu J, Kepler T (2005) Pathway level analysis of gene expression using singular value decomposition. *BMC Bioinformatics* **6**, 225.