Royal Holloway
University of London

A transcriptomic analysis of photomorphogenesis in *Arabidopsis thaliana*

Hugh Shanahan,
Department of Computer Science,
Royal Holloway,
University of London

Chips, Computers and Crops
浙江大學
杭州
26th-27th September 2008

http://www.cs.rhul.ac.uk/home/hugh/Hangzhou

Outline

- Photomorphogenesis in *A. thaliana*
- Strategy of analysis :- Systems Biology (well, a qualitative picture) from one set of microarrays
- A time-line of photomorphogenesis
    - Regulation phase
    - Cell cycle phase
    - Growth and photosynthesis
- Lessons for other species

Photomorphogenesis in *Arabidopsis thaliana*

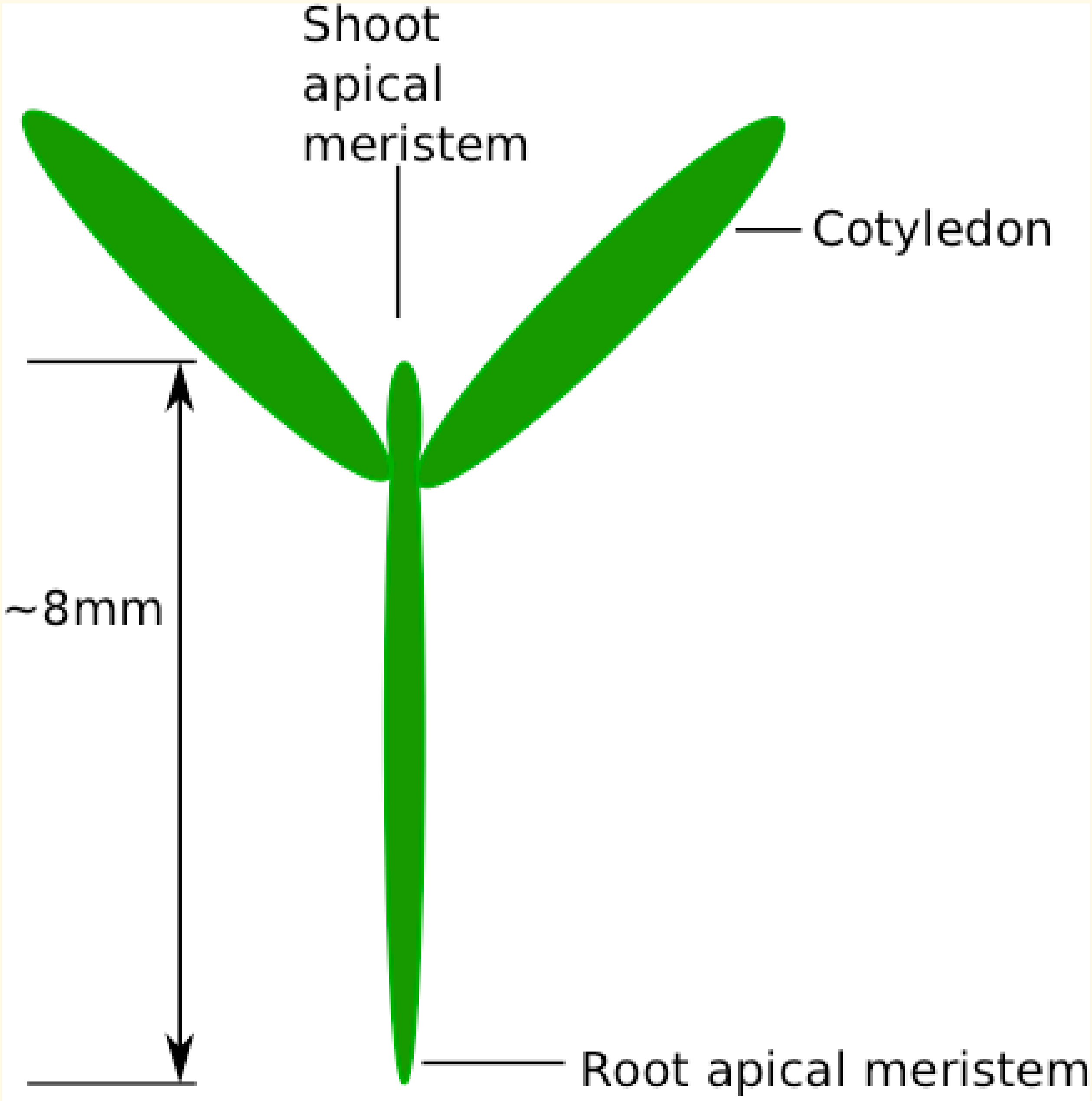
- Before exposure to light, seedling grows via skotomorphogenesis after germination - slow root growth, no growth in shoot apical meristem or cotyledon.
- When exposed to light, cotyledon grows through reproduction.
- Shoot apex meristem (stem cells) grow by differentiation.
- Shoot apex meristem source of true leaves.
- Little understood process.

Photomorphogenesis in *Arabidopsis thaliana*

- Before exposure to light, seedling grows via skotomorphogenesis after germination - slow root growth, no growth in shoot apical meristem or cotyledon.
- When exposed to light, cotyledon grows through reproduction.
- Shoot apex meristem (stem cells) grow by differentiation.
- Shoot apex meristem source of true leaves.
- Little understood process.

Photomorphogenesis in *Arabidopsis thaliana*

- Before exposure to light, seedling grows via skotomorphogenesis after germination - slow root growth, no growth in shoot apical meristem or cotyledon.
- When exposed to light, cotyledon grows through reproduction.
- Shoot apex meristem (stem cells) grow by differentiation.
- Shoot apex meristem source of true leaves.
- Little understood process.



The data

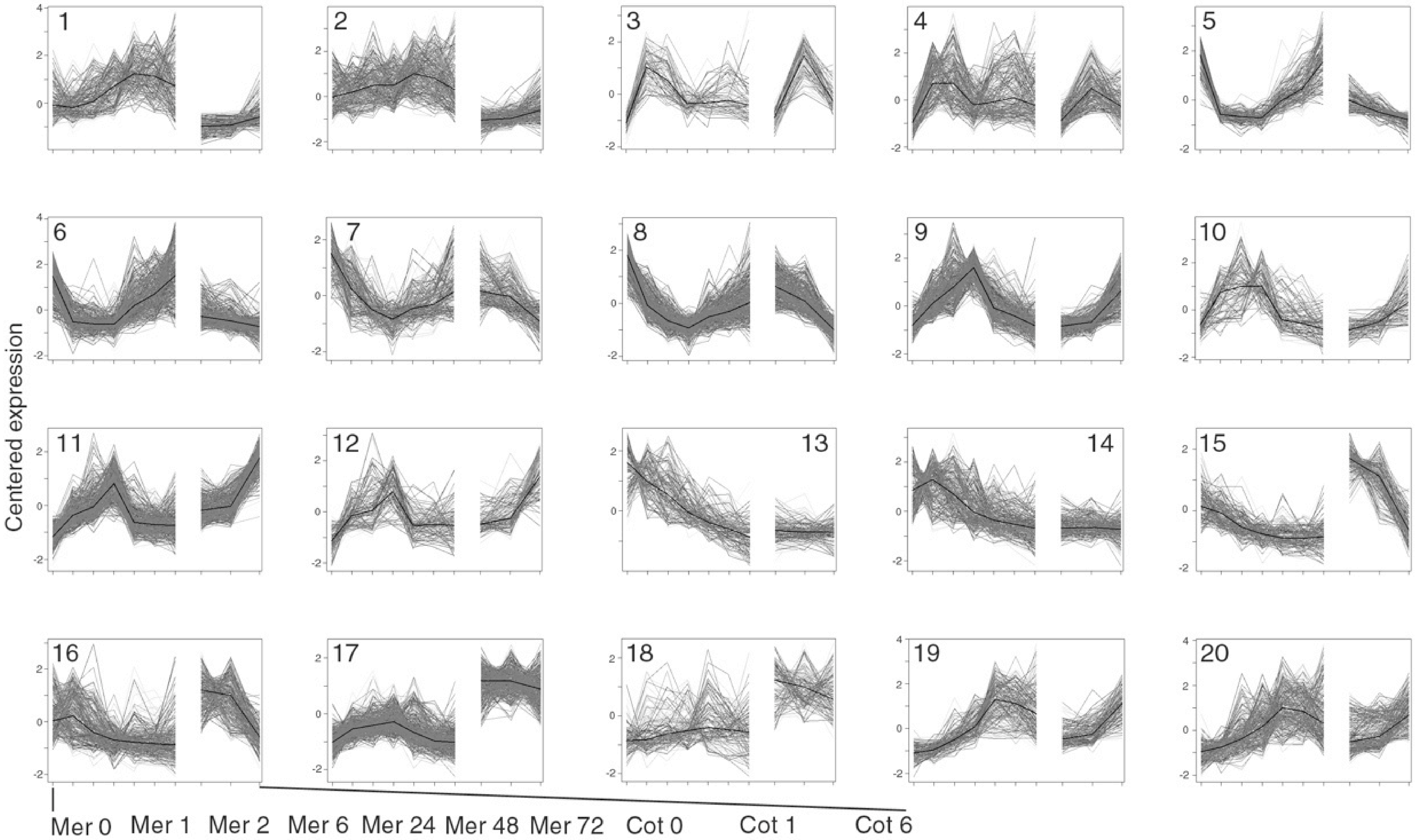
- RNA material was gathered from the shoot apical meristem and cotyledon of Arabidopsis seedlings at
    - 0 hour (in darkness)
    - 1 and 6 hours (Cot and Mer with replicates)
    - 2, 24, 48 and 72 hours (Mer only)
- Samples hybridised with Affymetrix ATH1 GeneChip array.
- No amplification of RNA material !

Strategy

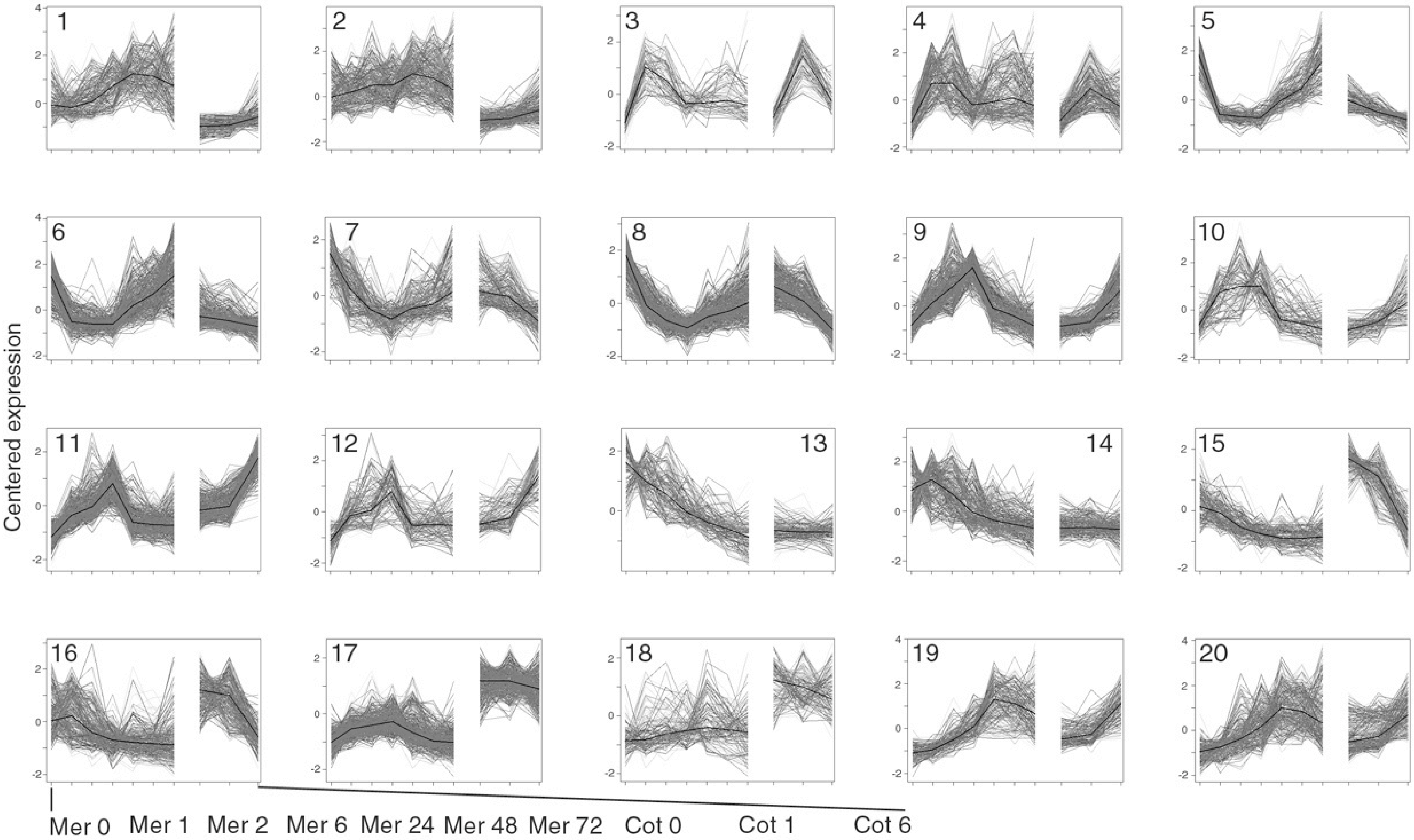
- Construct stringent test to determine genes which are clearly differentially expressed.
- Identify kinetic behaviour of different classes of differentially expressed genes (i.e. cluster genes according to their time behaviour).
- On the basis of the clusters, identify relevant functional groupings of genes and then examine how all the genes in that functional grouping behave (i.e. including those that are not differentially expressed according to our strict criteria) as a function of time.

Initial Results

- Selected 5,620 genes (out of 22,810).
  - (Very conservatively) 1/4 of the transcriptome is differentially expressed during photomorphogenesis.
- Majority selected through time variation (2/3 time, 1/3 tissue).
- Very small number selected using time-tissue variation (10).



K-means clusters (reordered)

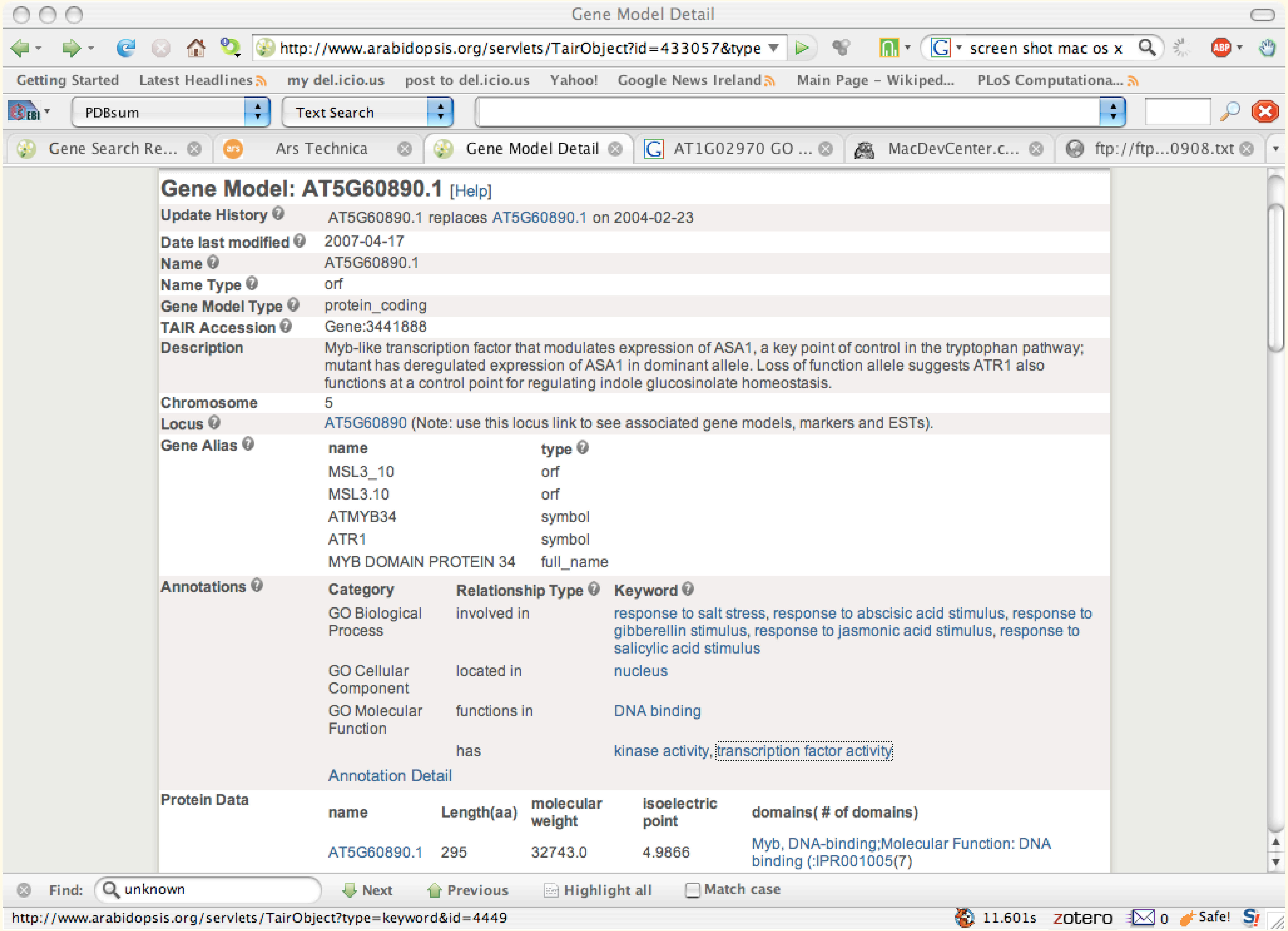


Yes, but what does it mean ?

K-means clusters (reordered)

Functional Classification :- not as straightforward as one thinks....

- Gene Ontology is useful but far too specific at its lowest nodes - False Discovery Rate calculation.
- Initially interested in general picture, what is the highest level annotation ?
- GO slim should cover more general cases, however annotations of genes can have multiple parents, e.g. a gene with kinase function and binds to DNA will sit in both classes.
- Ultimately, we developed our own general functional annotation using GO and MapMan.
- Lesson :- GO has a huge amount of information, but when looking at the big picture, you need to make the decisions !



Gene Model: AT5G60890.1 [Help]

Update History	AT5G60890.1 replaces AT5G60890.1 on 2004-02-23
Date last modified	2007-04-17
Name	AT5G60890.1
Name Type	orf
Gene Model Type	protein_coding
TAIR Accession	Gene:3441888
Description	Myb-like transcription factor that modulates expression of ASA1, a key point of control in the tryptophan pathway; mutant has deregulated expression of ASA1 in dominant allele. Loss of function allele suggests ATR1 also functions at a control point for regulating indole glucosinolate homeostasis.
Chromosome	5
Locus	AT5G60890 (Note: use this locus link to see associated gene models, markers and ESTs).

Gene Alias

name	type
MSL3_10	orf
MSL3.10	orf
ATMYB34	symbol
ATR1	symbol
MYB DOMAIN PROTEIN 34	full_name

Annotations

Category	Relationship Type	Keyword
GO Biological Process	involved in	response to salt stress, response to abscisic acid stimulus, response to gibberellin stimulus, response to jasmonic acid stimulus, response to salicylic acid stimulus
GO Cellular Component	located in	nucleus
GO Molecular Function	functions in	DNA binding
	has	kinase activity, transcription factor activity

Annotation Detail

Protein Data

name	Length(aa)	molecular weight	isoelectric point	domains(# of domains)
AT5G60890.1	295	32743.0	4.9866	Myb, DNA-binding;Molecular Function: DNA binding (:IPR001005(7)

Functional Classification :- not as straightforward as one thinks....

- Gene Ontology is useful but far too specific at its lowest nodes - False Discovery Rate calculation.

- Initially interested in general picture, what is the highest level annotation ?

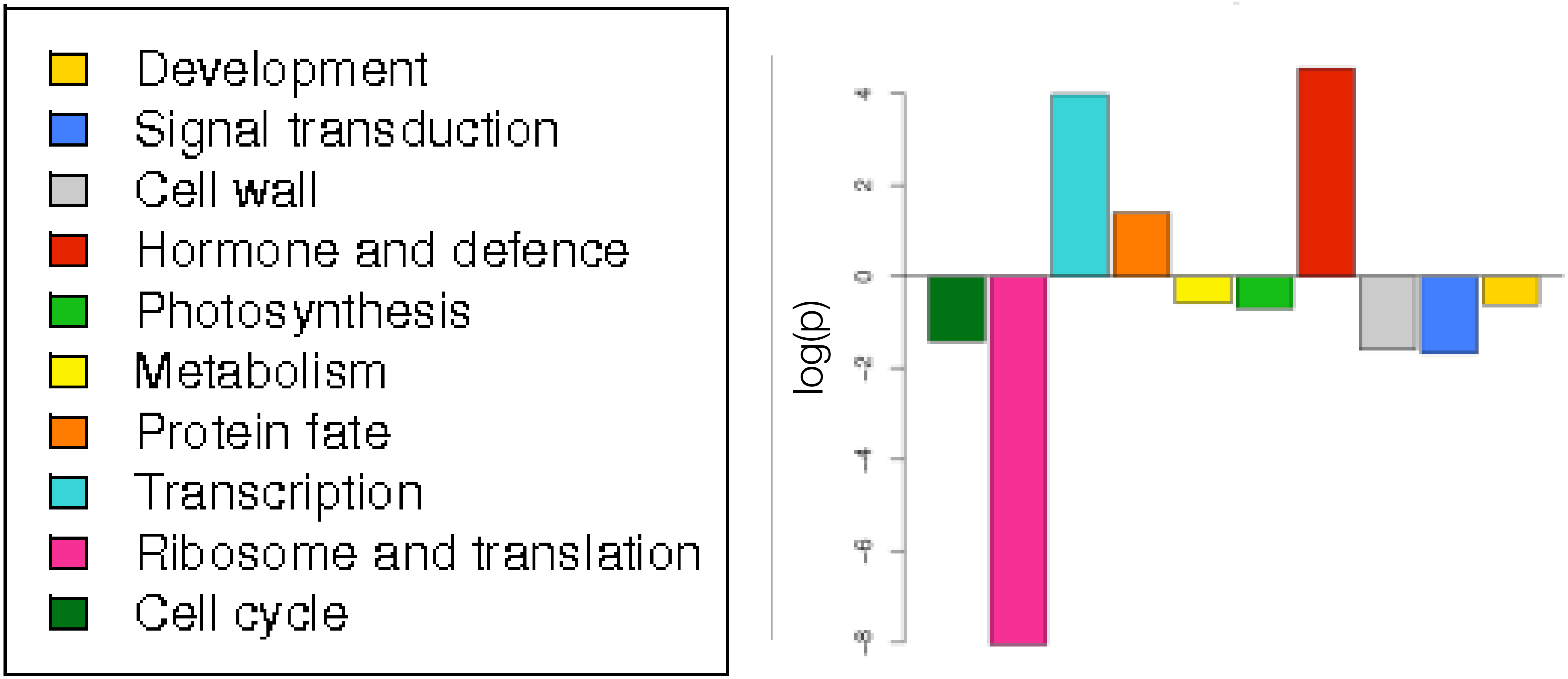
- GO slim should cover more general cases, however annotations of genes can have multiple parents, e.g. a gene with kinase function and binds to DNA will sit in both classes.

- Ultimately, we developed our own general functional annotation using GO and MapMan.

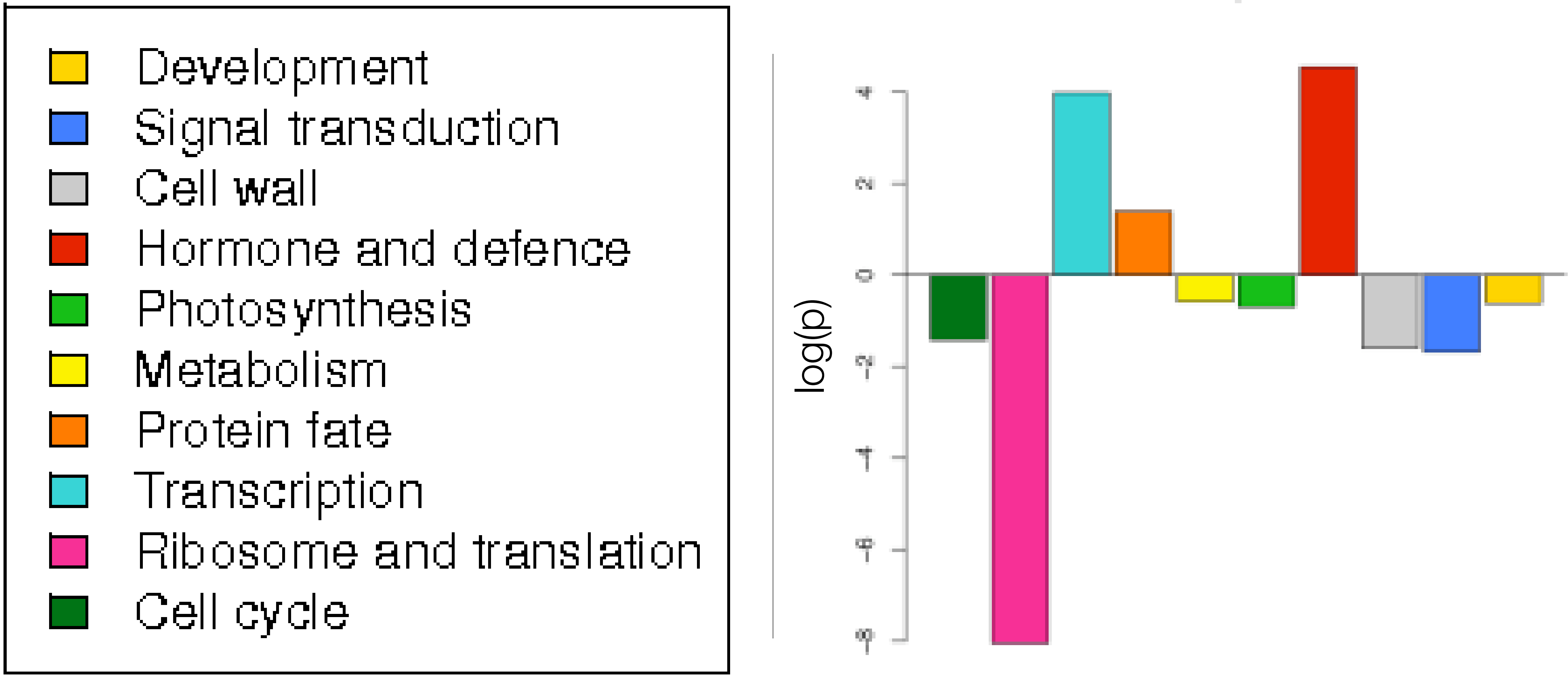
- Lesson :- GO has a huge amount of information, but when looking at the big picture, you need to make the decisions !

Example

Example


Development
Signal transduction
Cell wall
Hormone and defence
Photosynthesis
Metabolism
Protein fate
Transcription
Ribosome and translation
Cell cycle
log(p)
4
2
0
-2
-4
-6
-8

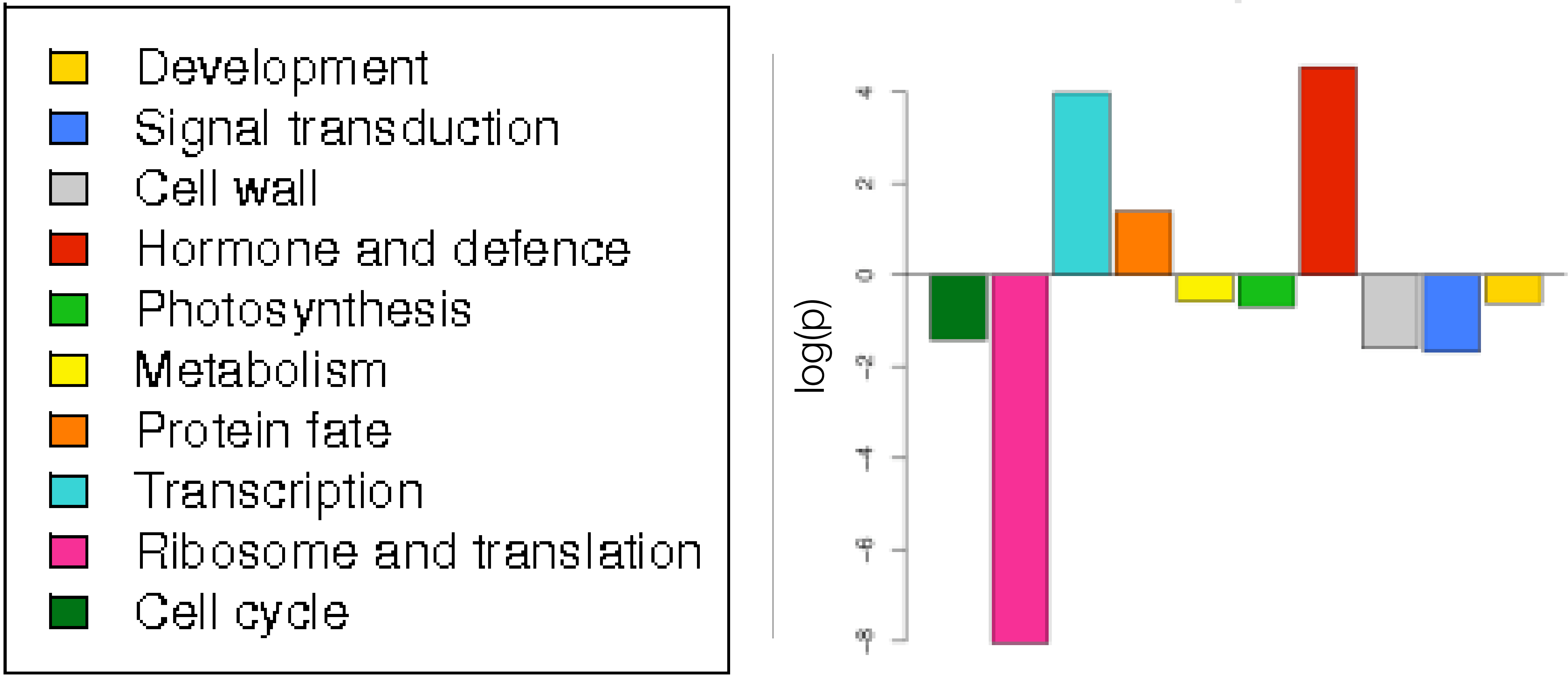
Example



Over-representation

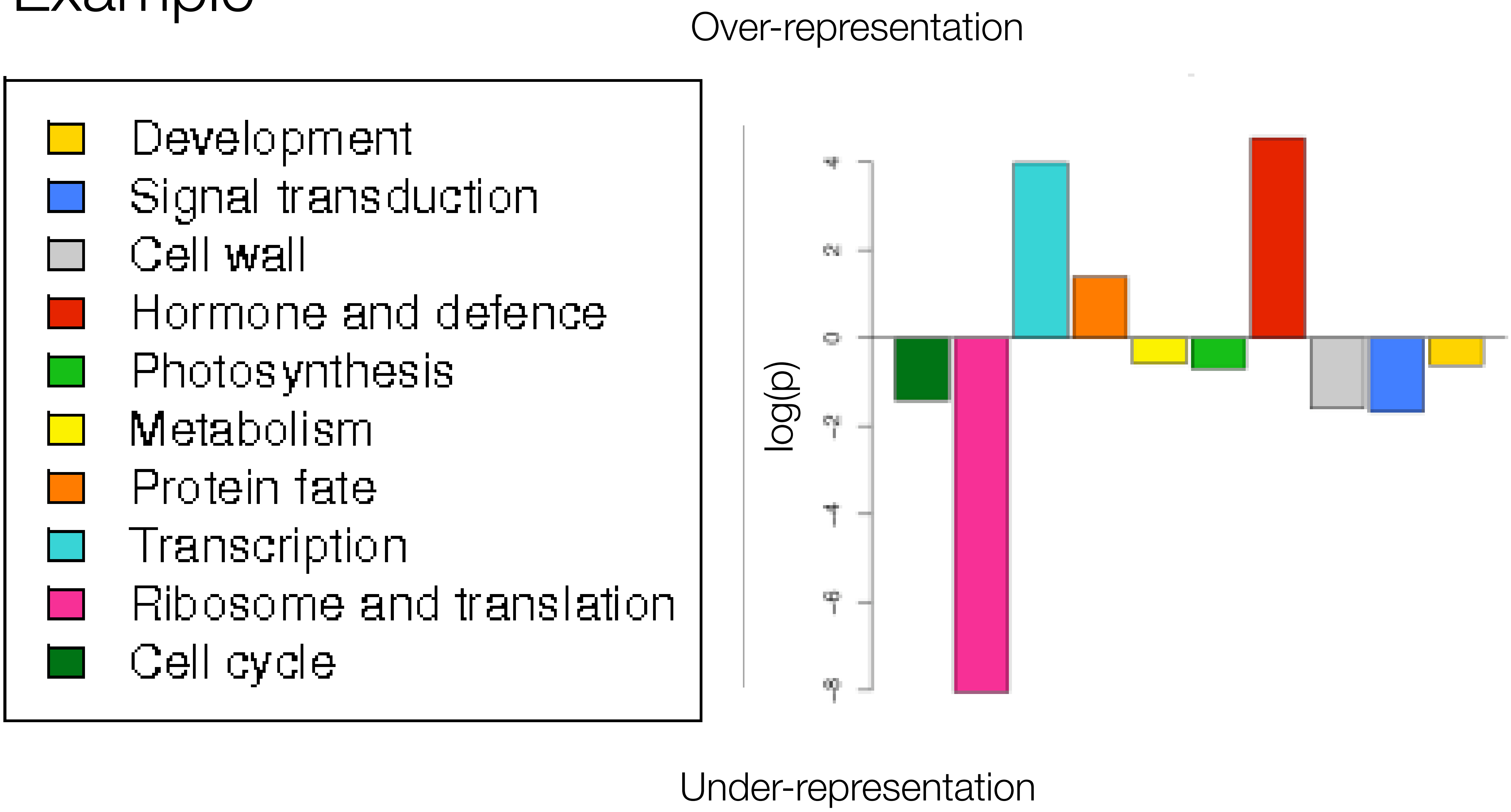
Example

Over-representation



Under-representation

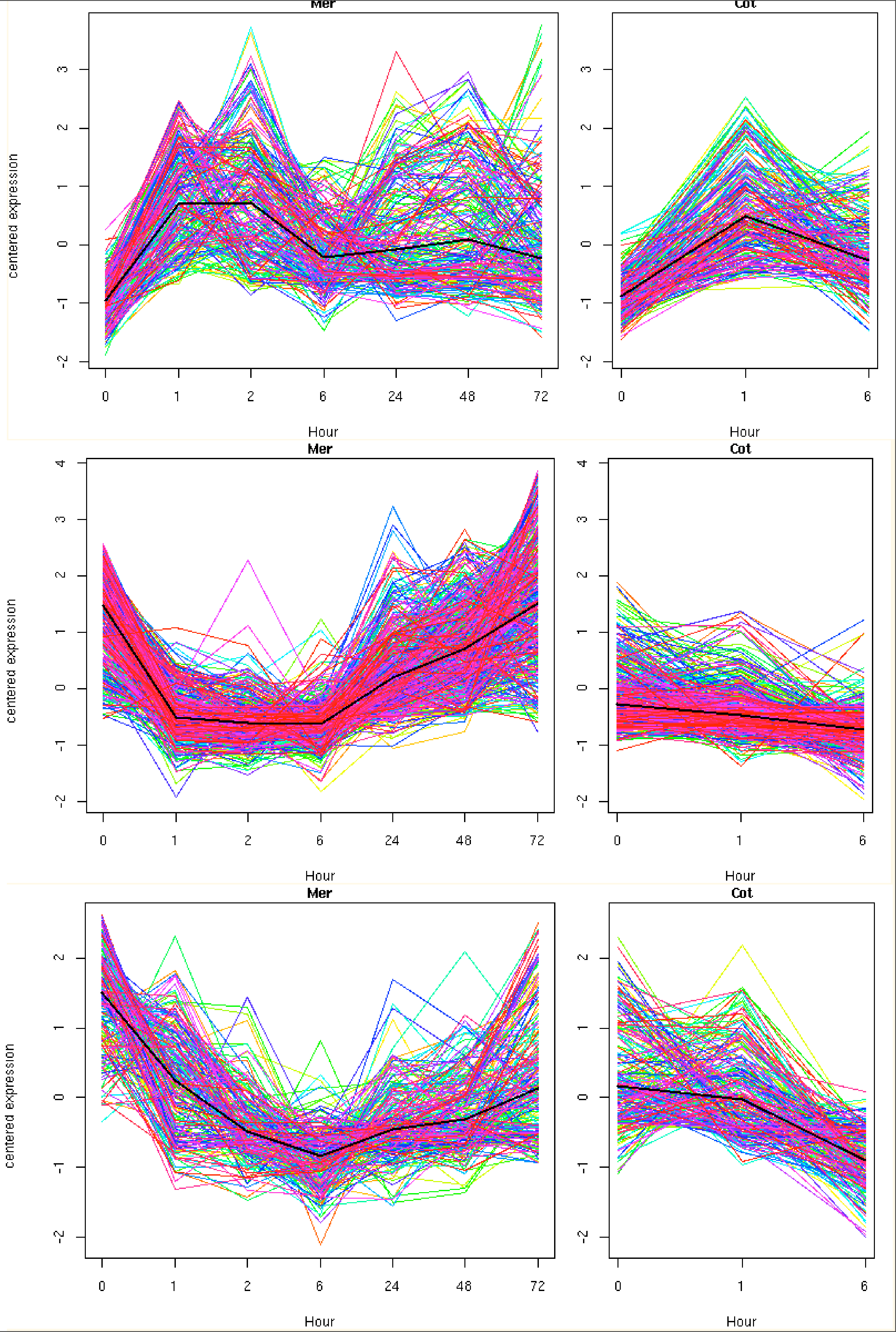
Example



p-value over(under) abundance computed using hypergeometric distribution

First Phase (0-1 hours)

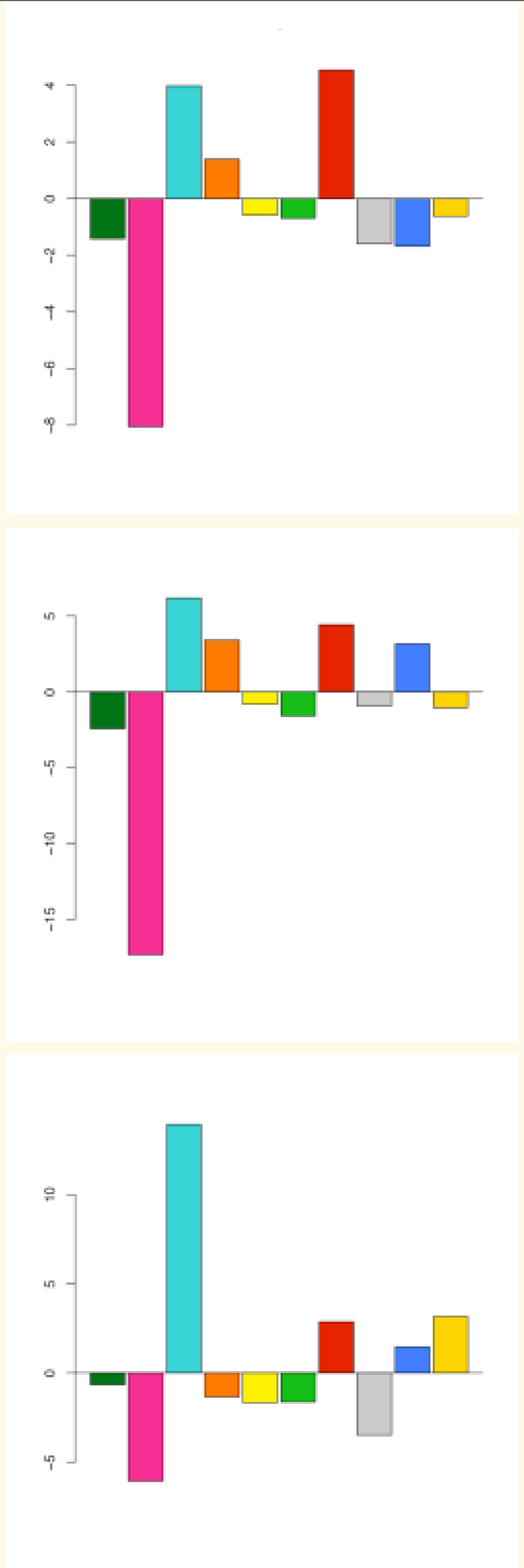
- Transcription Factors (down reg.)
- Ubiquitination, (down reg.)
- Hormones, kinases (up reg.)

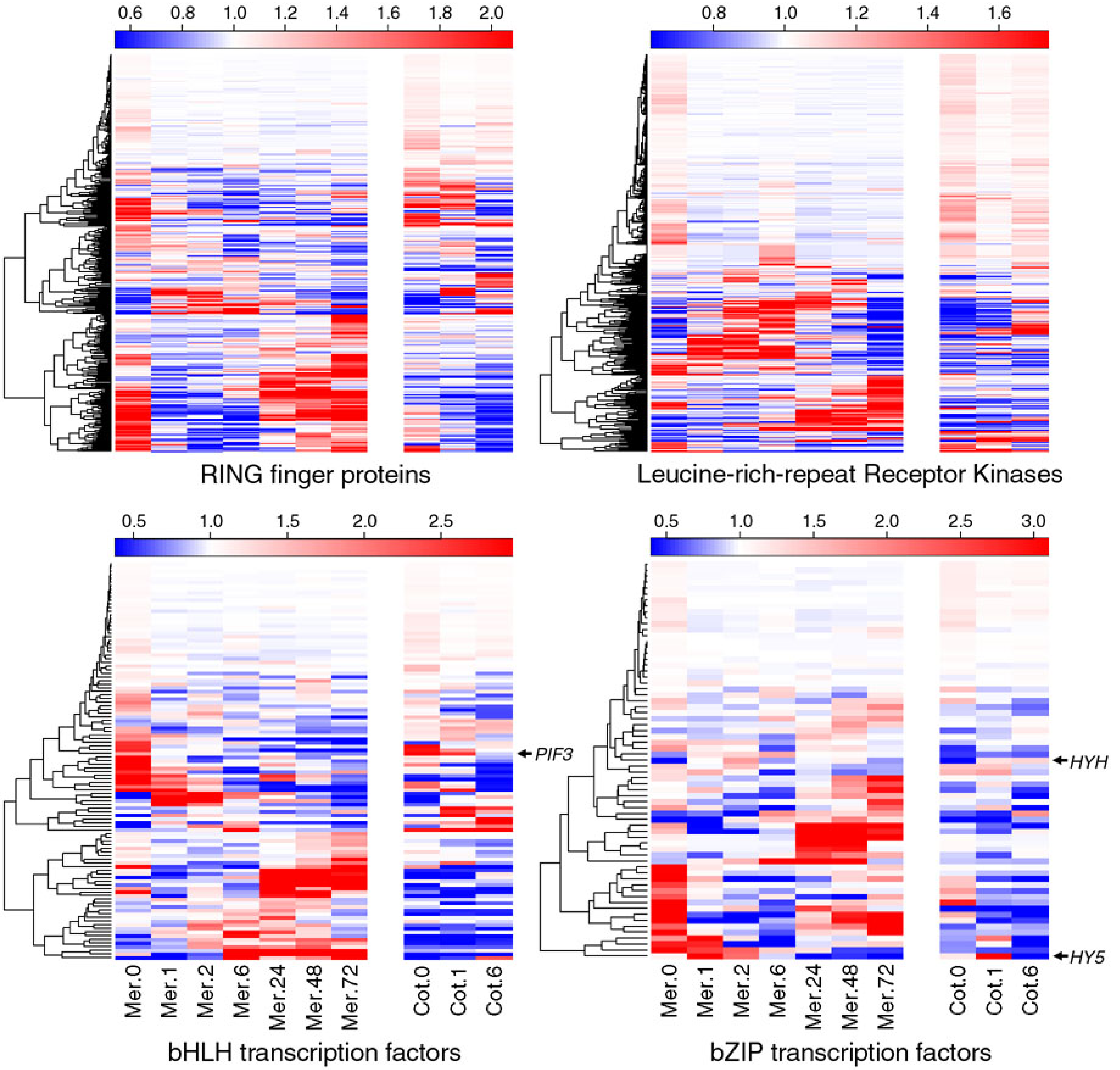


First Phase (0-1 hours)

- Transcription Factors (down reg.)
- Ubiquitination, (down reg.)
- Hormones, kinases (up reg.)

🟨 Development
🟦 Signal transduction
⬜ Cell wall
🟥 Hormone and defence
🟩 Photosynthesis
🟨 Metabolism
🟧 Protein fate
🟦 Transcription
🟪 Ribosome and translation
🟩 Cell cycle





RING finger proteins

Leucine-rich-repeat Receptor Kinases

bHLH transcription factors

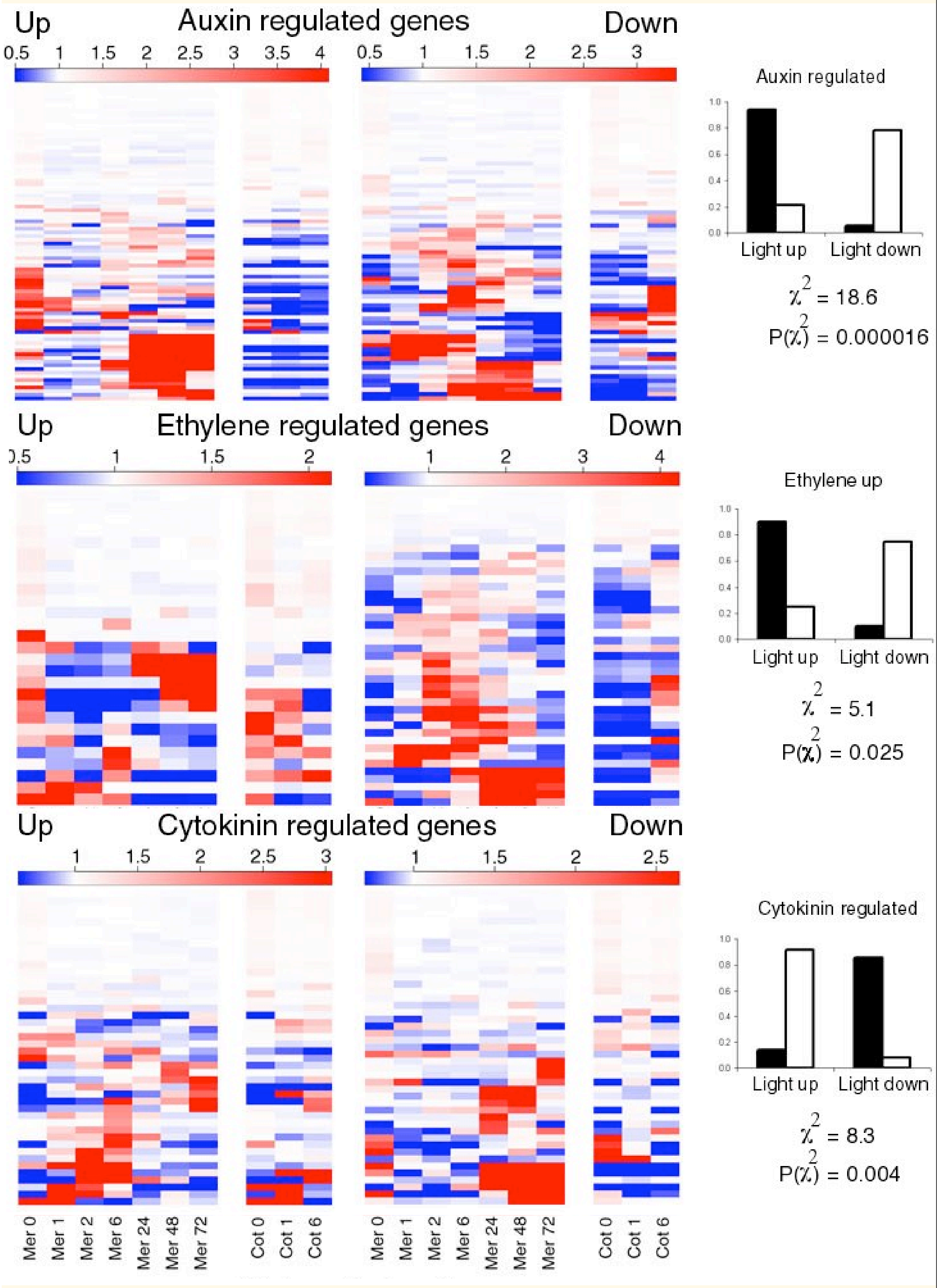
bZIP transcription factors

Regulator (?) classes

Including non-selected genes

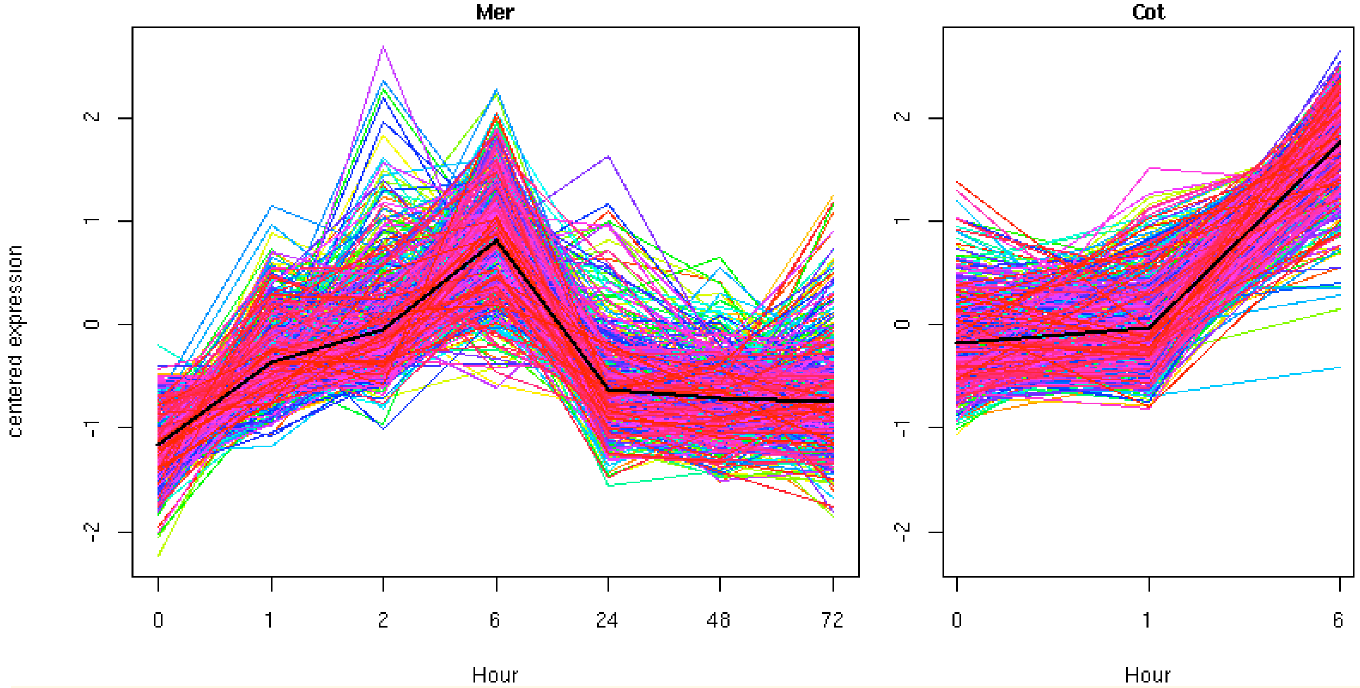
Hormones

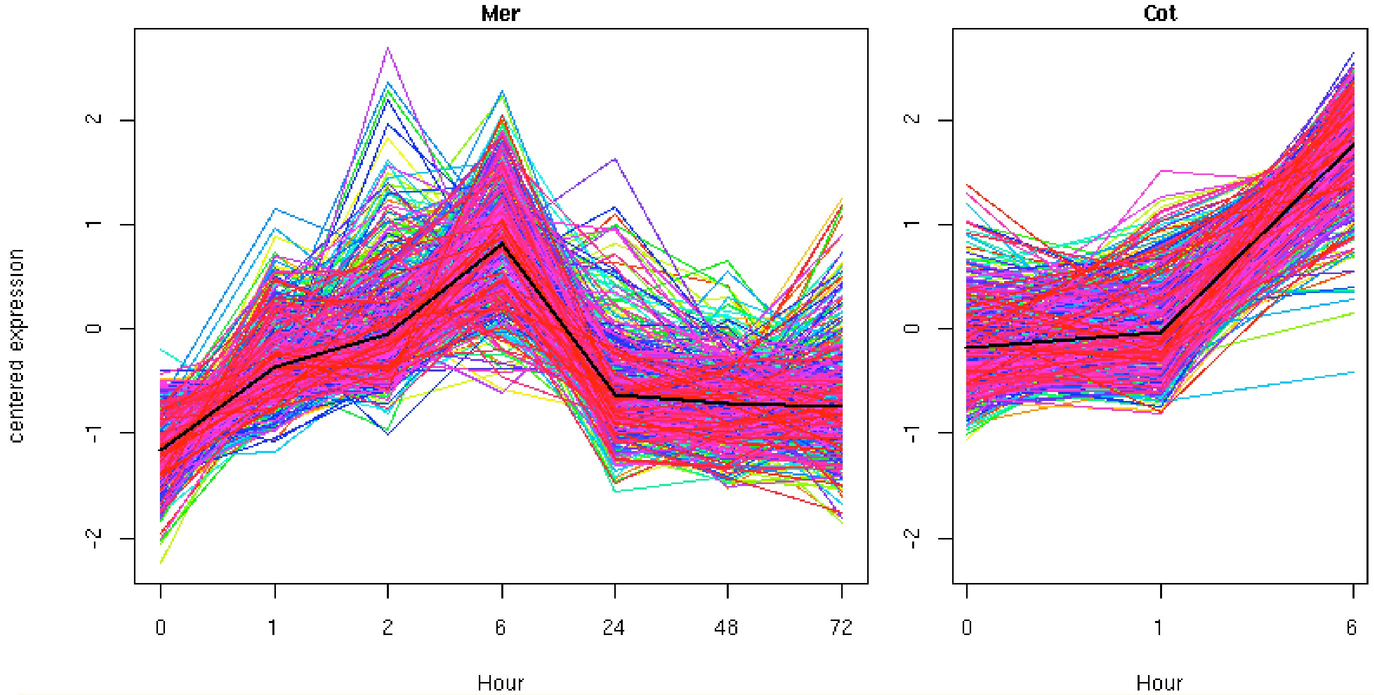
- Auxin, Ethylene promote growth (elongation).
- Cytokinin repress elongation.
- Cannot track hormone concs.
- Follow activity of genes regulated by hormones (strictly differentially expressed).
- Auxin, Ethylene “repressed”.
- Cytokinin “promoted”.



Second Phase (1-24 hours)

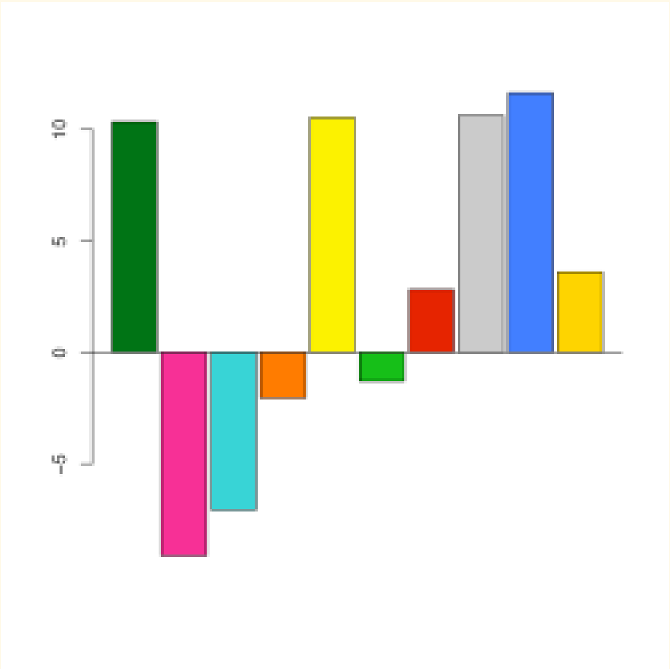
- Ribosomal activity
- cell cycle

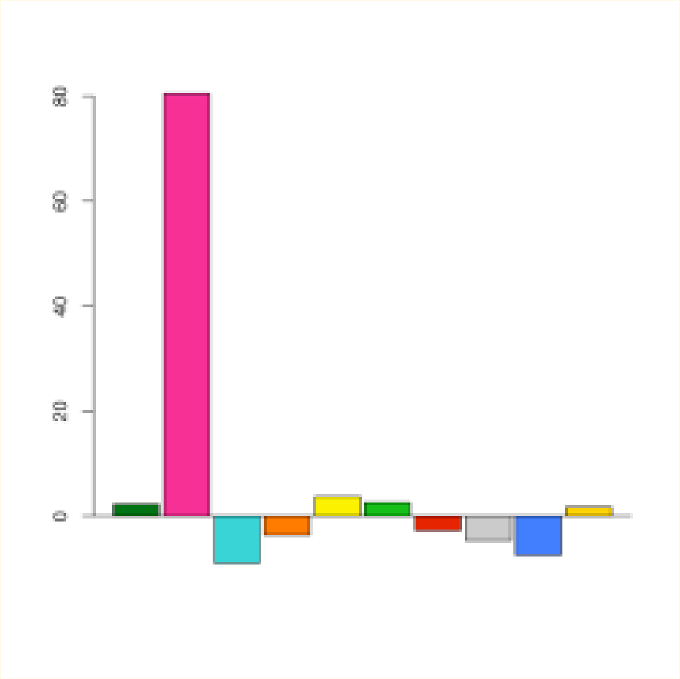




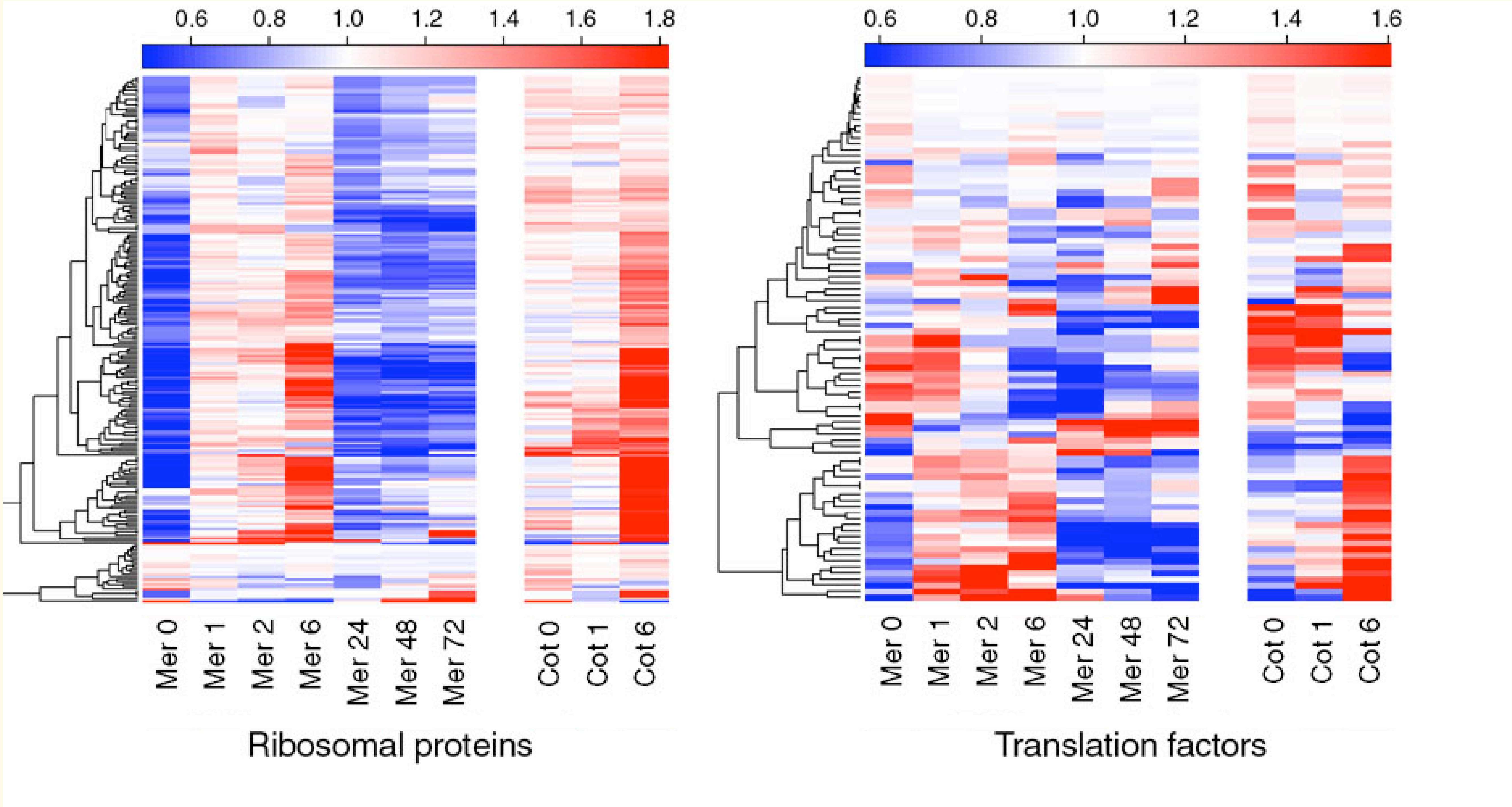
Second Phase (1-24 hours)

- Ribosomal activity
- cell cycle





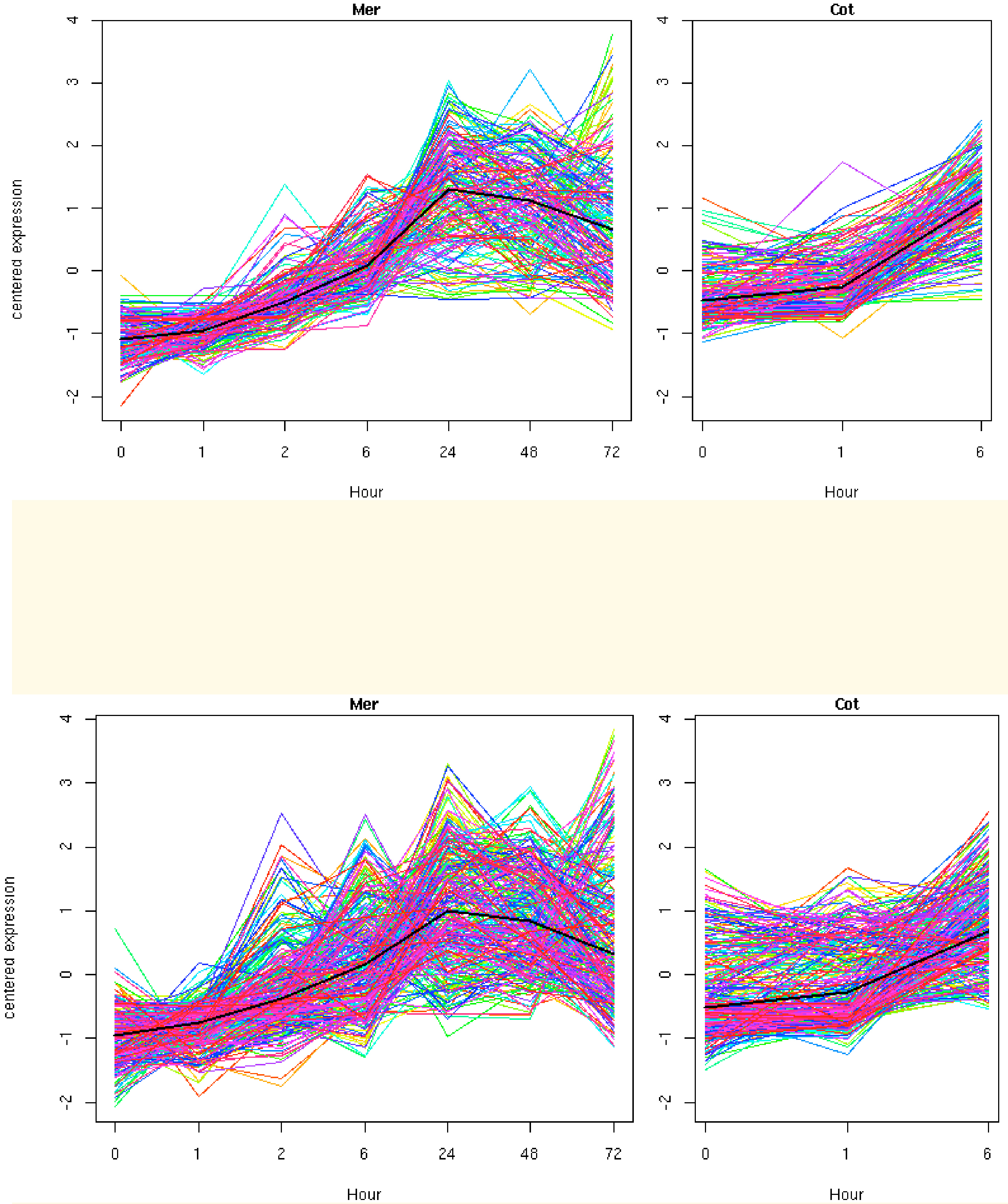
- 🟨 Development
- 🟦 Signal transduction
- ⬜ Cell wall
- 🟥 Hormone and defence
- 🟩 Photosynthesis
- 🟨 Metabolism
- 🟧 Protein fate
- 🟦 Transcription
- 🟪 Ribosome and translation
- 🟩 Cell cycle



Ribosome working overtime

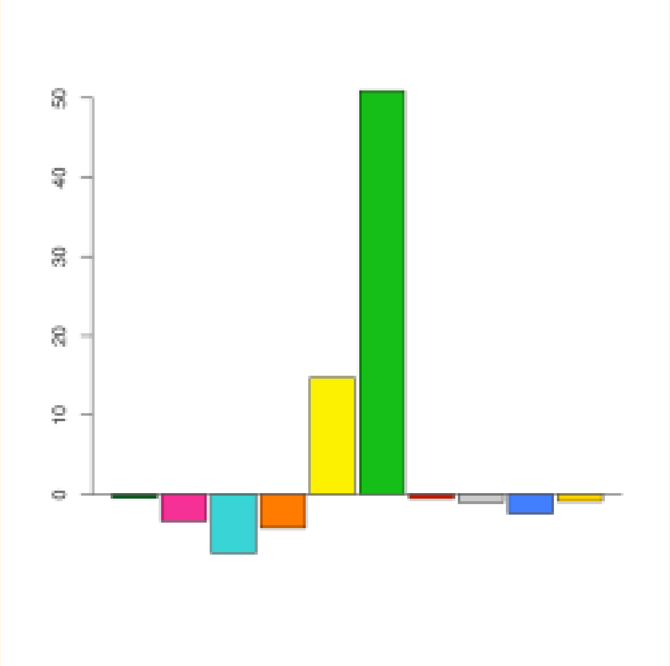
Third Phase (24- hours)

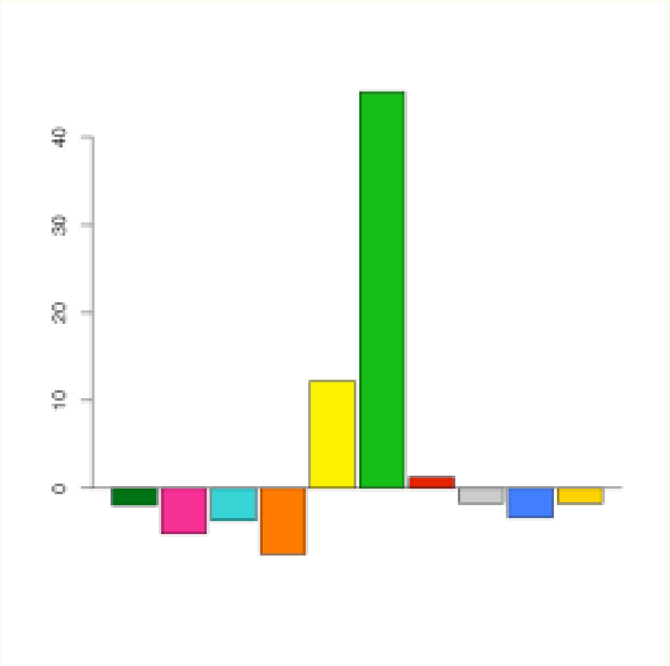
- Photosynthesis
- Cell wall loosening



Third Phase (24- hours)

- Photosynthesis
- Cell wall loosening

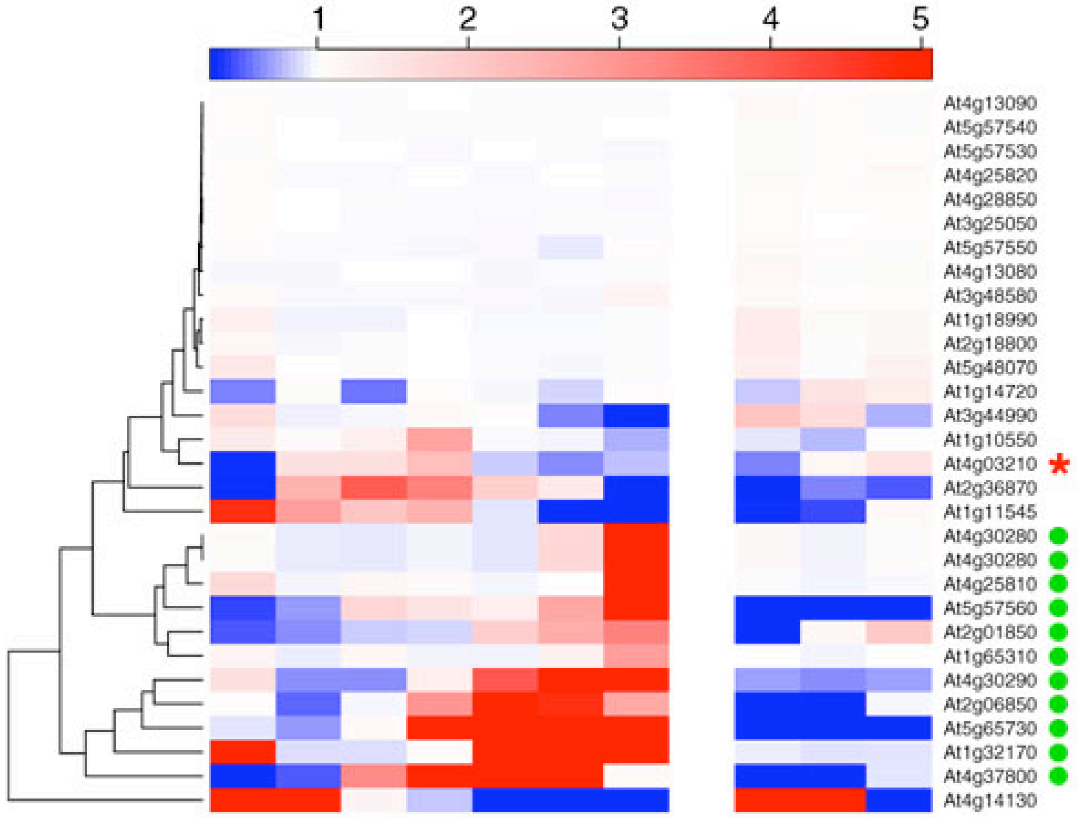




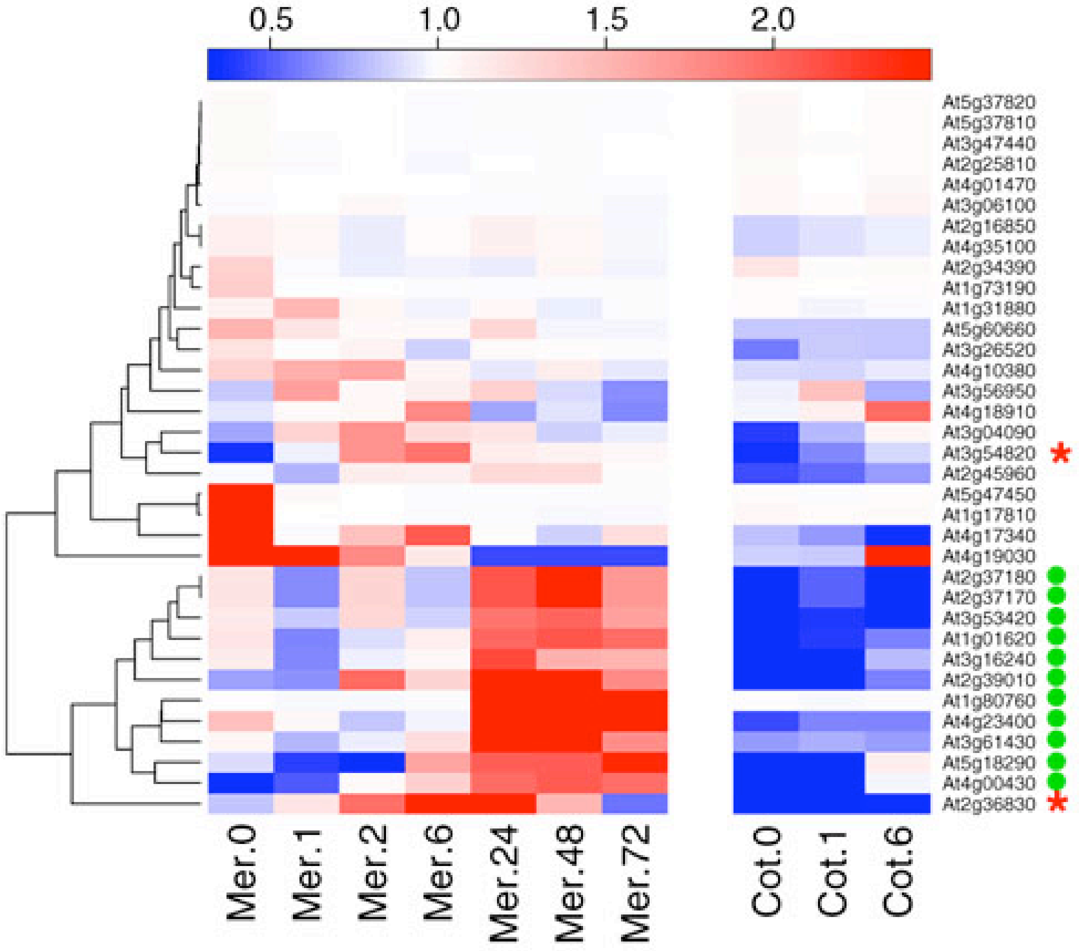
🟨 Development
🟦 Signal transduction
⬜ Cell wall
🟥 Hormone and defence
🟩 Photosynthesis
🟨 Metabolism
🟧 Protein fate
🟦 Transcription
🟪 Ribosome and translation
🟩 Cell cycle

Cell Wall

- Plant cells have rigid cell walls.
- Expansion implies that cell walls must become less rigid.
- Complicated process between cell wall modification and internal turgor pressure.
- Nonetheless, see late expression in genes controlling this behaviour.



Xyloglucan endotransglycosylase



Aquaporin

Summary

- After constructing a set of genes which are strictly differentially expressed we find:
- An early burst (0-1 hours after exposure to light) of genes in the meristem that are regulatory in nature and are in general down rather than up-regulated.
- Around 6 hours after exposure to light evidence for cell division and repression of growth.
- At later times, parts of meristem are already starting to behave like leaves and we see growth through expansion rather than division (up-regulation of relevant hormone-related genes, down-regulation of ribosomal genes).
- Promoter searches bear this out :- Abscissic Acid Response Element Motif very common at early times, ribosomal motif at later times.

Lessons learnt

- Selecting the right tissue may be very tedious but well worth the effort.
- Time series data is absolutely crucial in understanding developmental processes.
- In the absence of large co-expression sets for network building
    - Focus on functional classes of genes (GSEA or ANCOVA).
    - Gene Ontology may not be appropriate.
    - Shared known upstream motifs can provide additional evidence.

Acknowledgements

Acknowledgements

School of Biological Sciences,
Royal Holloway

- Enrique López-Juez
- Edyta Dillon
- Safina Khan
- Zoltan Magyar
- Laszlo Bögre

John Innes Centre
Norwich

- Saul Hazeldine

Department of Molecular Genetics,
Ghent University

- Gerrit Beemster

Institute of Biotechnology,
University of Cambridge

- James A. H. Murray

Acknowledgements

School of Biological Sciences,
Royal Holloway

- Enrique López-Juez
- Edyta Dillon
- Safina Khan
- Zoltan Magyar
- Laszlo Bögre

John Innes Centre
Norwich

- Saul Hazeldine

Department of Molecular Genetics,
Ghent University

- Gerrit Beemster

Institute of Biotechnology,
University of Cambridge

- James A. H. Murray

Presentation: http://www.cs.rhul.ac.uk/home/hugh/Hangzhou