Biological interpretation of RNA-seq bioinformatics: a number of pitfalls….

RNA-seq

=

big datasets, numerous tools with numerous parameters

=

many different interpretations from the same dataset

Biological interpretation of RNA-seq bioinformatics: a number of pitfalls….

« RNA-seq detects all RNAs »: Is my gene not expressed?

Biological interpretation of RNA-seq bioinformatics: a number of pitfalls….

« RNA-seq detects all RNAs »: Is my gene not expressed?

Starting material:

- Which tissue?

Laser microdissection vs whole plant

- Which extraction protocol?

Small RNAs, ImmunoPrecipitation

- Which library synthesis protocol?

polyA selection/ribosomal depletion/size selection

- Sequencing technology?

Illumina, PacBio, nanopore, IonTorrent…

- Sequencing depth?

Biological interpretation of RNA-seq bioinformatics: a number of pitfalls….

« RNA-seq detects all RNAs »: Is my gene not ~~expressed~~ detected?

Starting material:

- Which tissue?

Laser microdissection vs whole plant

- Which extraction protocol?

Small RNAs, ImmunoPrecipitation

- Which library synthesis protocol?

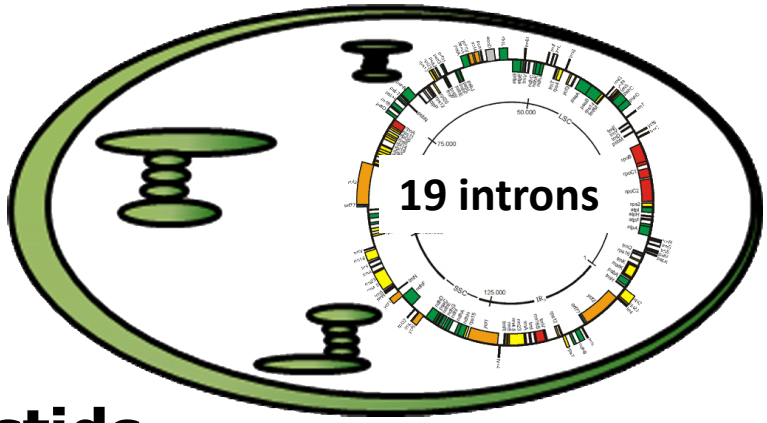
polyA selection/ribosomal depletion/size selection

- Sequencing technology?

Illumina, PacBio, nanopore, IonTorrent…

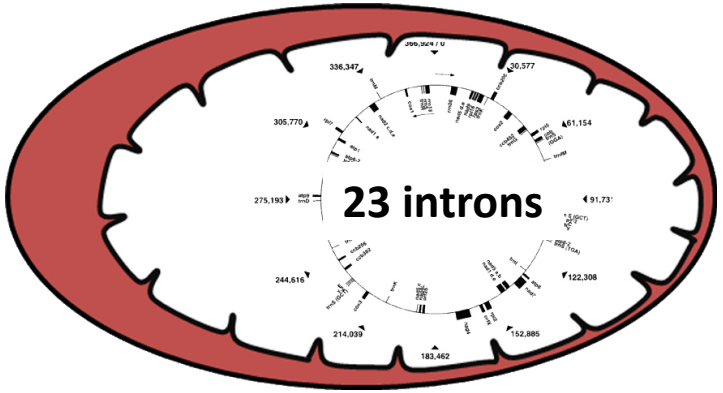
- Sequencing depth?

Mapping surprises: the impact of the tools and parameters on the biological interpretation
Splicing analysis in Ath organelles



Plastids

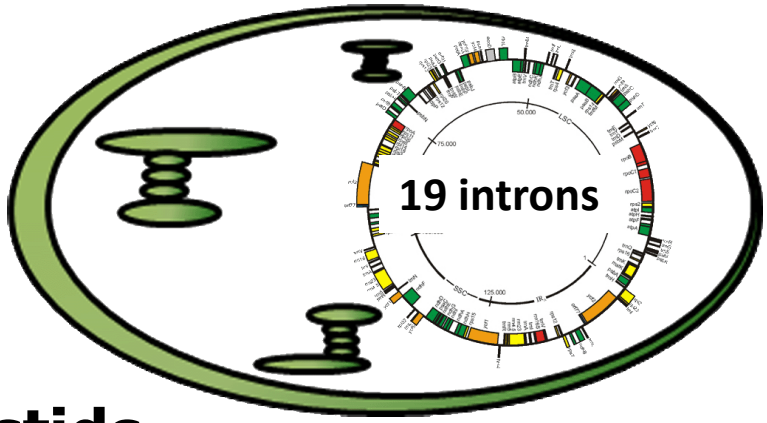
Organelles contain small genomes encoding key proteins and RNA necessary for their biology



Mitochondria

Mapping surprises: the impact of the tools and parameters on the biological interpretation

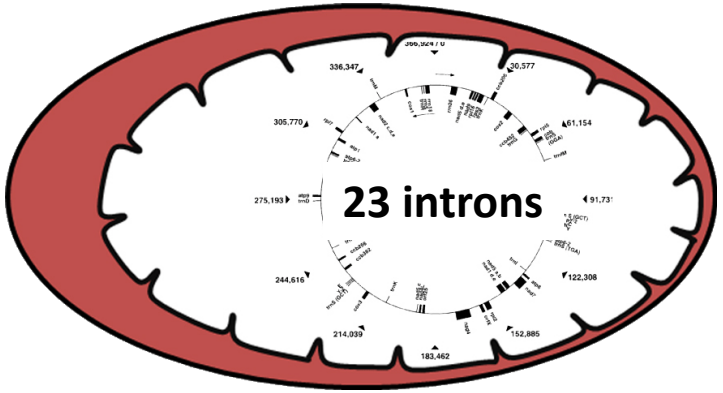
Splicing analysis in Ath organelles



Plastids

non polyA mRNAs

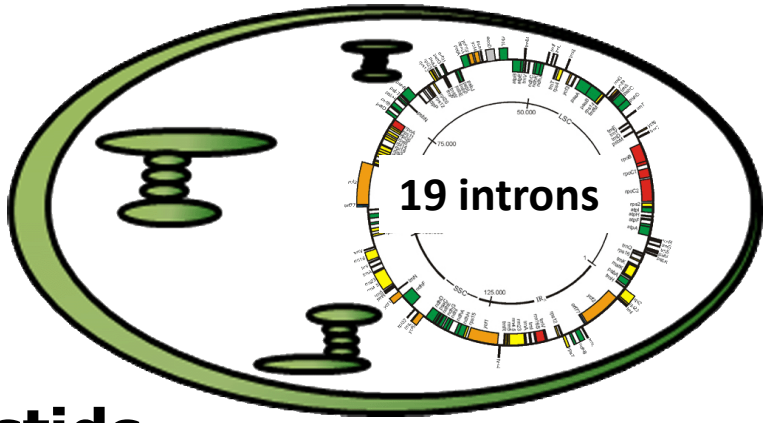
→ rRNA depletion



Mitochondria

Mapping surprises: the impact of the tools and parameters on the biological interpretation

Splicing analysis in Ath organelles



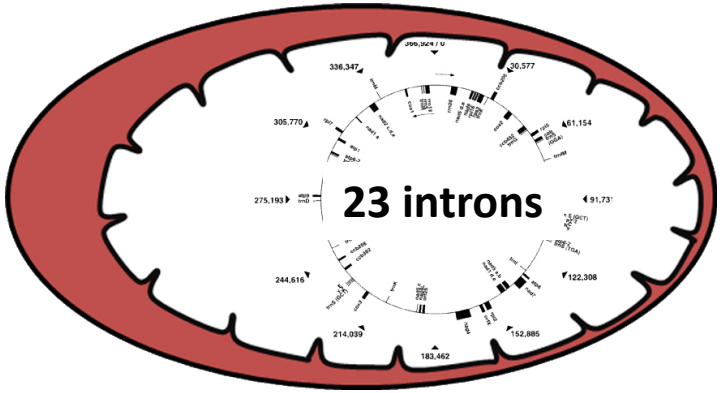
Plastids

non polyA mRNAs

→ rRNA depletion

Splicing analysis

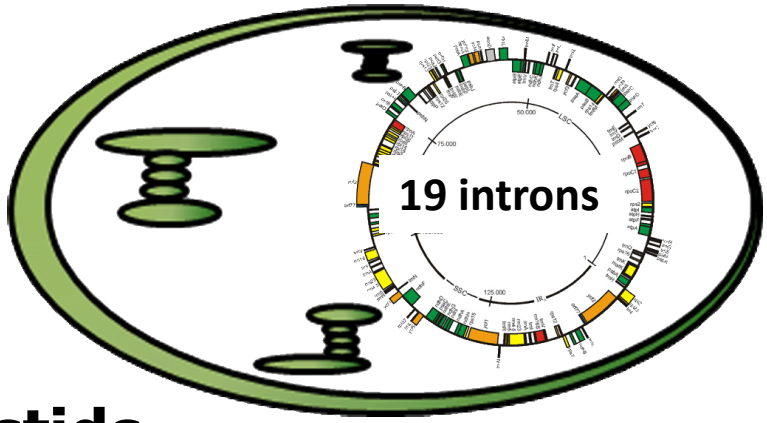
→ mapping on genome



Mitochondria

Mapping surprises: the impact of the tools and parameters on the biological interpretation

Splicing analysis in Ath organelles



Plastids

non polyA mRNAs

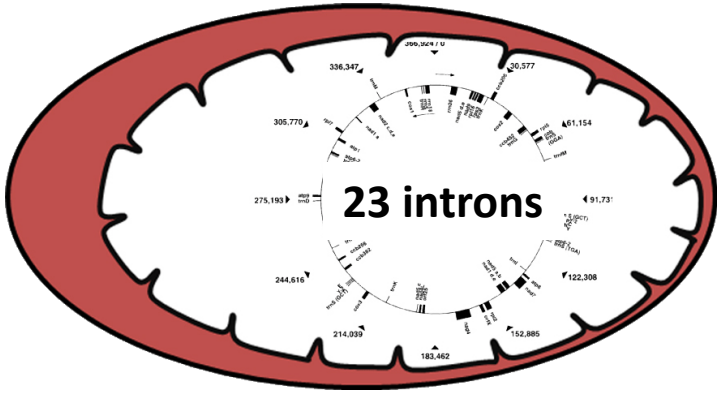
↳ rRNA depletion

Splicing analysis

↳ mapping on genome

Copy of mito genome in Chr2

↳ Mapping on all Chr



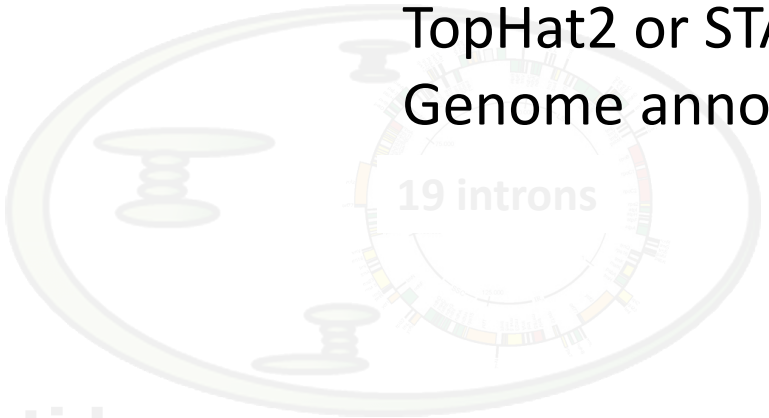
Mitochondria

Mapping surprises: the impact of the tools and parameters on the biological interpretation

Splicing analysis in Ath organelles

TopHat2 or STAR?

Genome annotation or not?

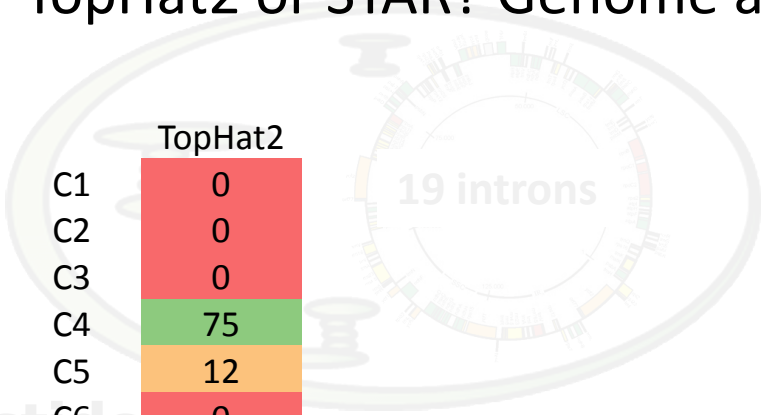


Mapping surprises: the impact of the tools and parameters on the biological interpretation

Splicing analysis in Ath organelles

TopHat2 or STAR? Genome annotation or not?

	TopHat2
C1	0
C2	0
C3	0
C4	75
C5	12
C6	0
C7	0
C8	0
C9	88
C10	26
C11	1
C12	62
C13	0
C14	0
C15	0
C16	0
C17	18
C18	1
C19	16



Mapping surprises: the impact of the tools and parameters on the biological interpretation
Splicing analysis in Ath organelles

TopHat2 or STAR? Genome annotation or not?

	TopHat2	TopHat2 + annot
C1	0	0
C2	0	11
C3	0	0
C4	75	75
C5	12	15
C6	0	19
C7	0	1
C8	0	2
C9	88	95
C10	26	27
C11	1	1
C12	62	62
C13	0	27
C14	0	6
C15	0	30
C16	0	75
C17	18	6
C18	1	1
C19	16	17

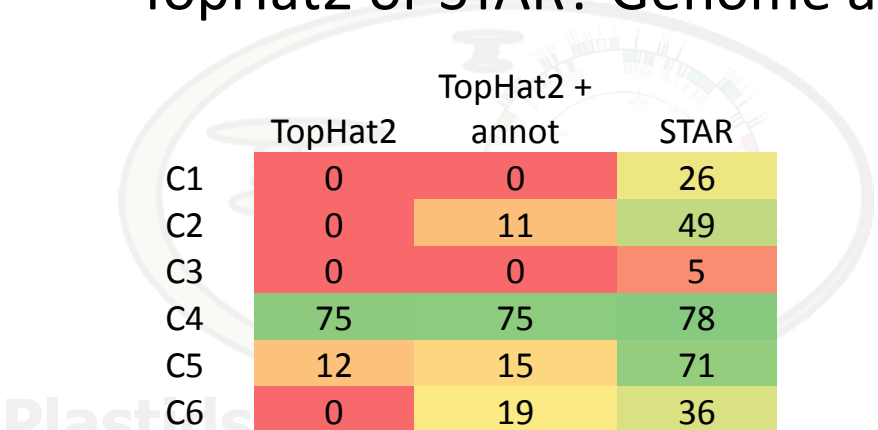
Nature paper!!!!
or biological non sense….

Mapping surprises: the impact of the tools and parameters on the biological interpretation

Splicing analysis in Ath organelles

TopHat2 or STAR? Genome annotation or not?

	TopHat2	TopHat2 + annot	STAR
C1	0	0	26
C2	0	11	49
C3	0	0	5
C4	75	75	78
C5	12	15	71
C6	0	19	36
C7	0	1	16
C8	0	2	61
C9	88	95	96
C10	26	27	85
C11	1	1	45
C12	62	62	77
C13	0	27	76
C14	0	6	88
C15	0	30	79
C16	0	75	85
C17	18	6	2
C18	1	1	4
C19	16	17	66



Mapping surprises: the impact of the tools and parameters on the biological interpretation

Splicing analysis in Ath organelles

TopHat2 or STAR? Genome annotation or not?

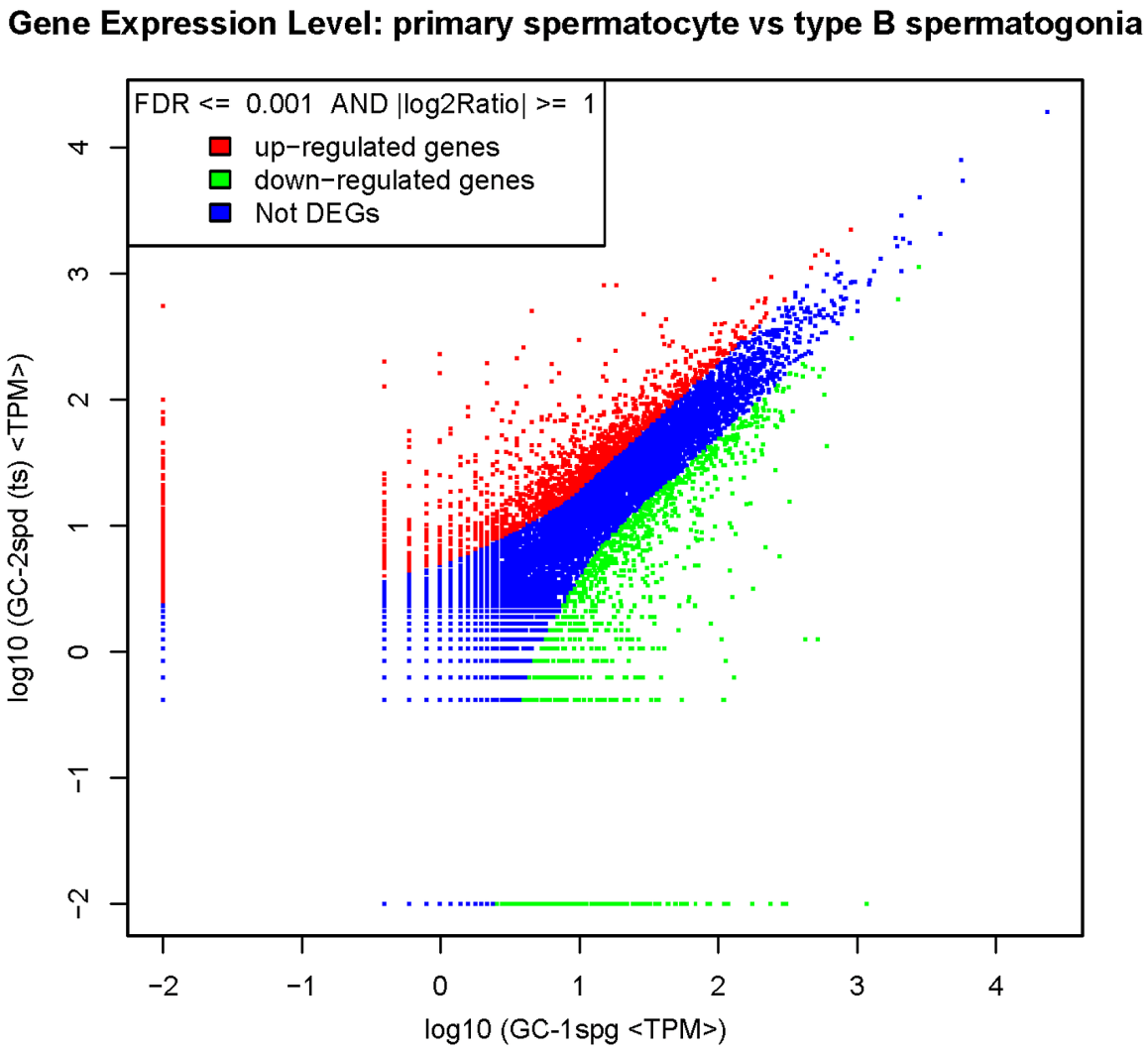
	TopHat2	TopHat2 + annot	STAR
C1	0	0	26
C2	0	11	49
C3	0	0	5
C4	75	75	78
C5	12	15	71
C6	0	19	36
C7	0	1	16
C8	0	2	61
C9	88	95	96
C10	26	27	85
C11	1	1	45
C12	62	62	77
C13	0	27	76
C14	0	6	88
C15	0	30	79
C16	0	75	85
C17	18	6	2
C18	1	1	4
C19	16	17	66

Use your biological expertise
or
borrow it from someone else!!!

Material and methods analysis

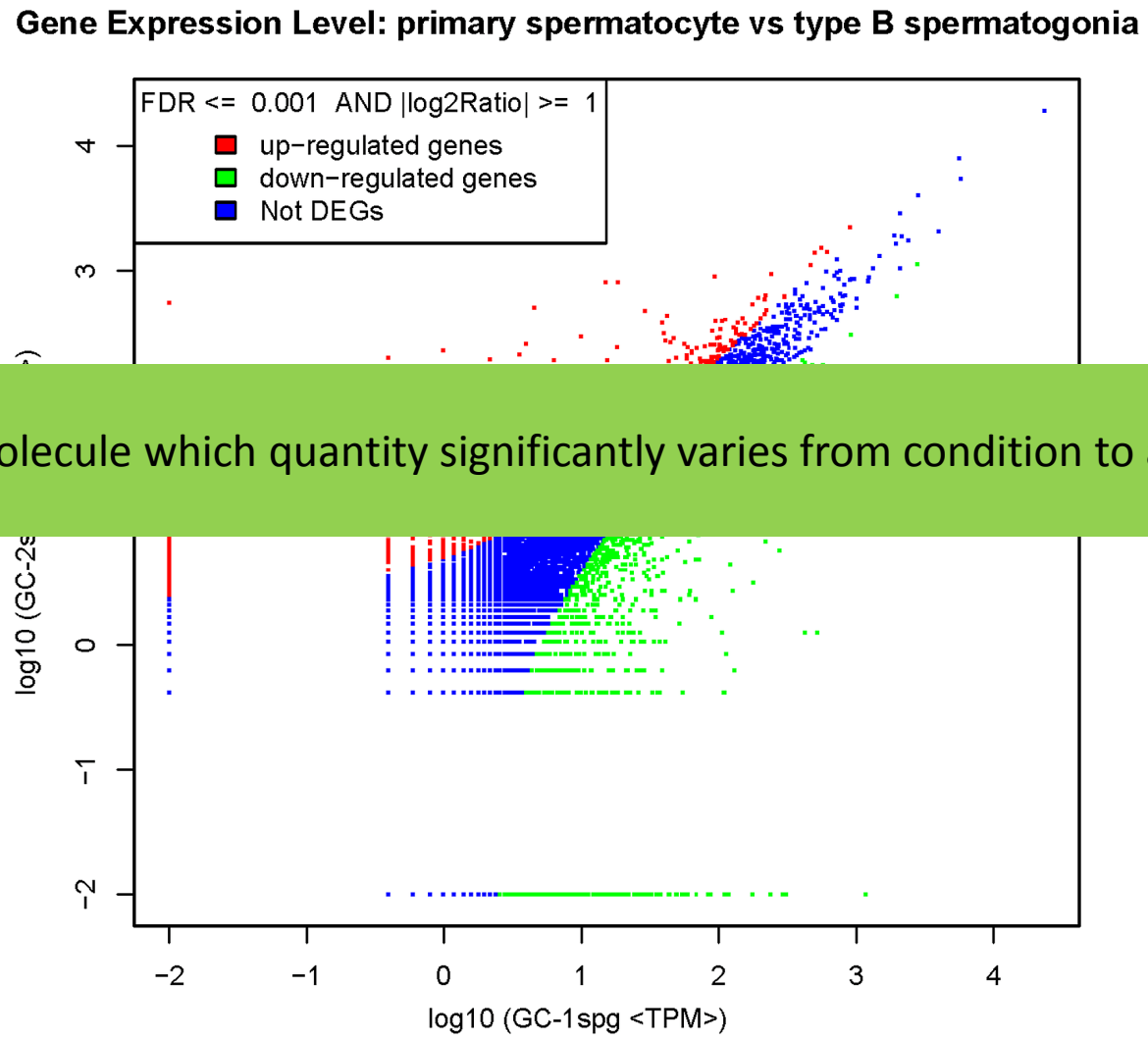
One biological interpretation of ….

“differentially expressed”



One biological interpretation of ….

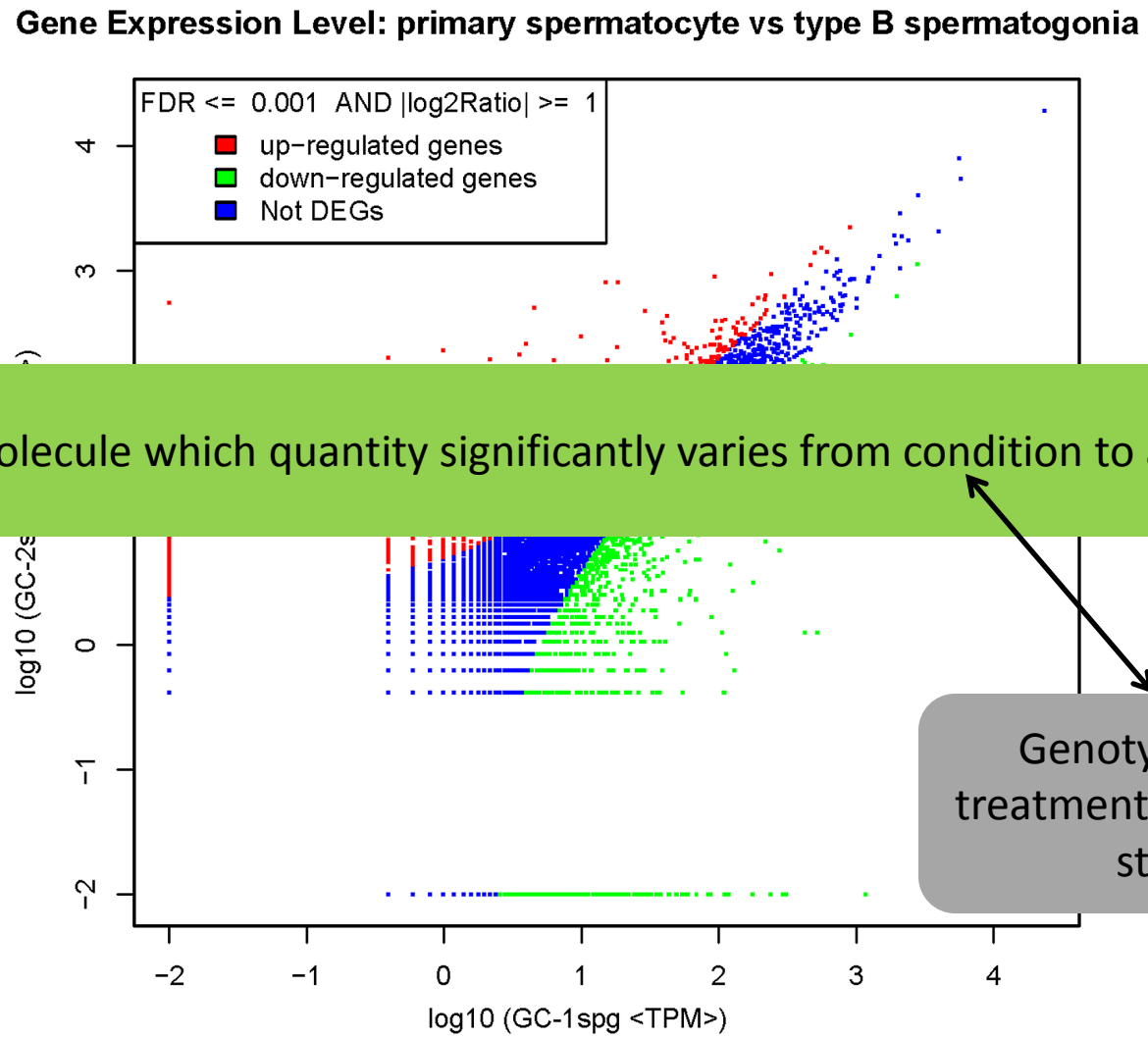
“differentially expressed”



RNA molecule which quantity significantly varies from condition to another

One biological interpretation of ….

“differentially expressed”

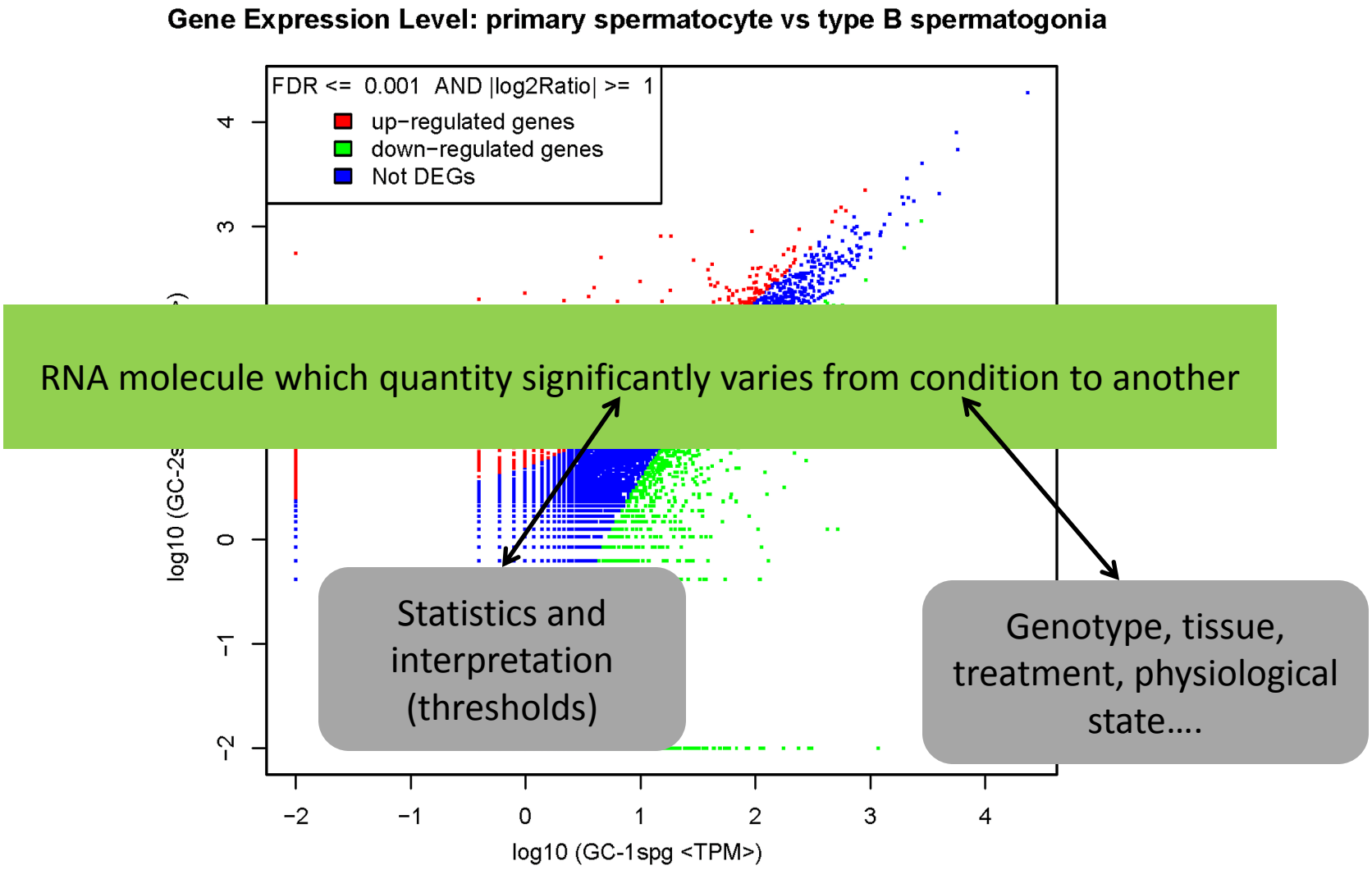


RNA molecule which quantity significantly varies from condition to another

Genotype, tissue, treatment, physiological state….

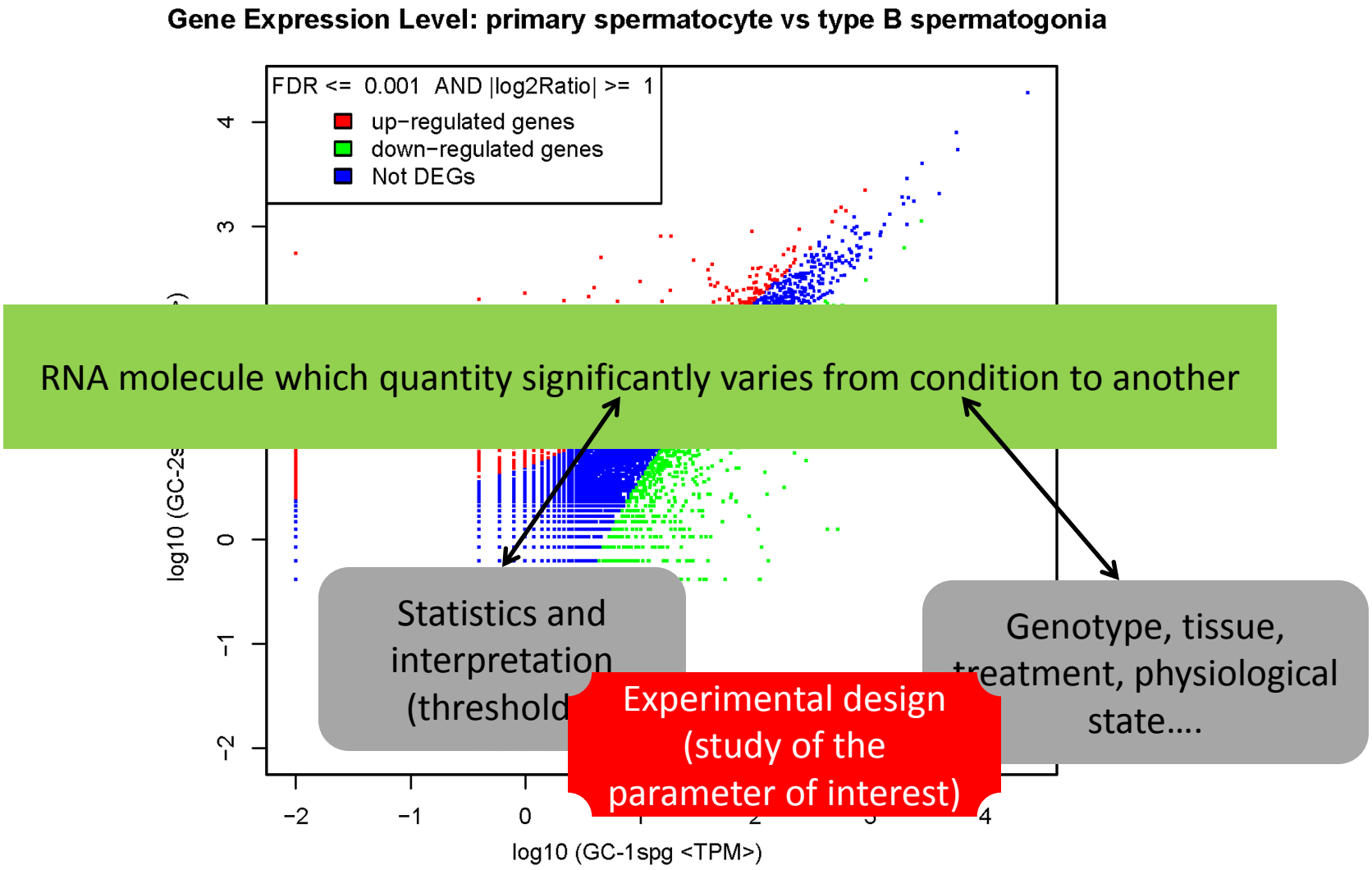
One biological interpretation of ….

“differentially expressed”



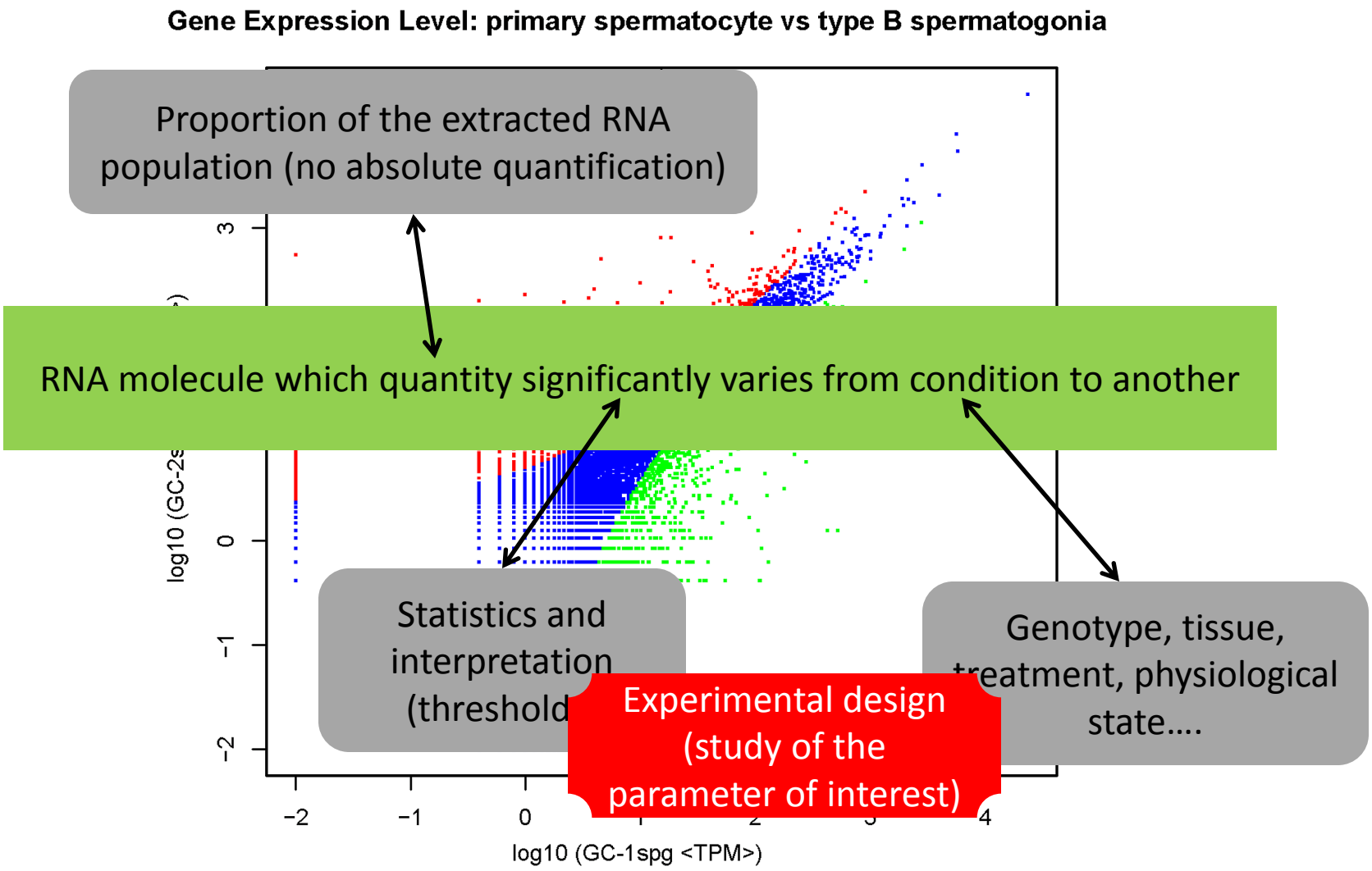
One biological interpretation of ….

“differentially expressed”



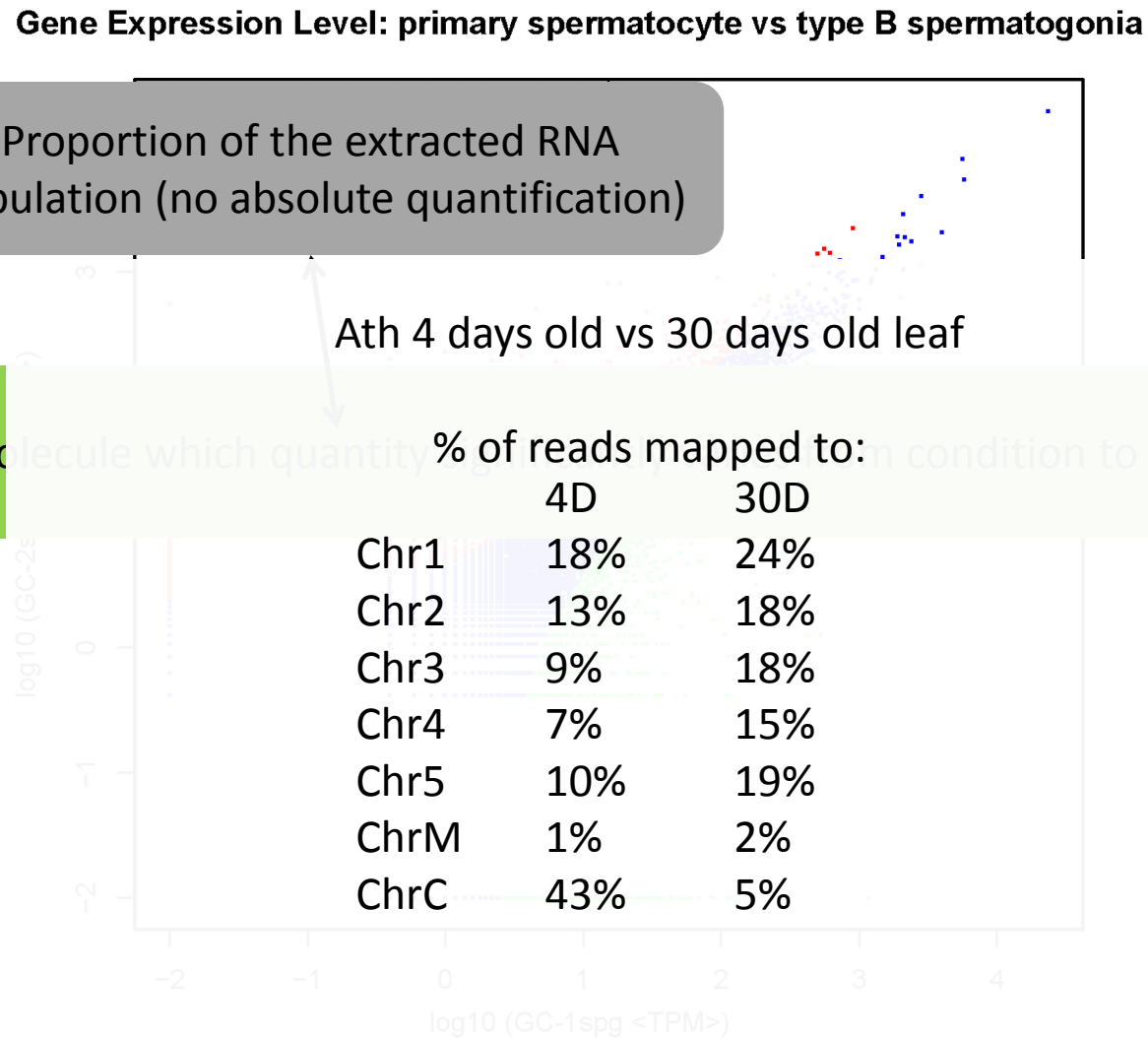
One biological interpretation of ….

“differentially expressed”



One biological interpretation of ….

“differentially expressed”



Proportion of the extracted RNA population (no absolute quantification)

Ath 4 days old vs 30 days old leaf

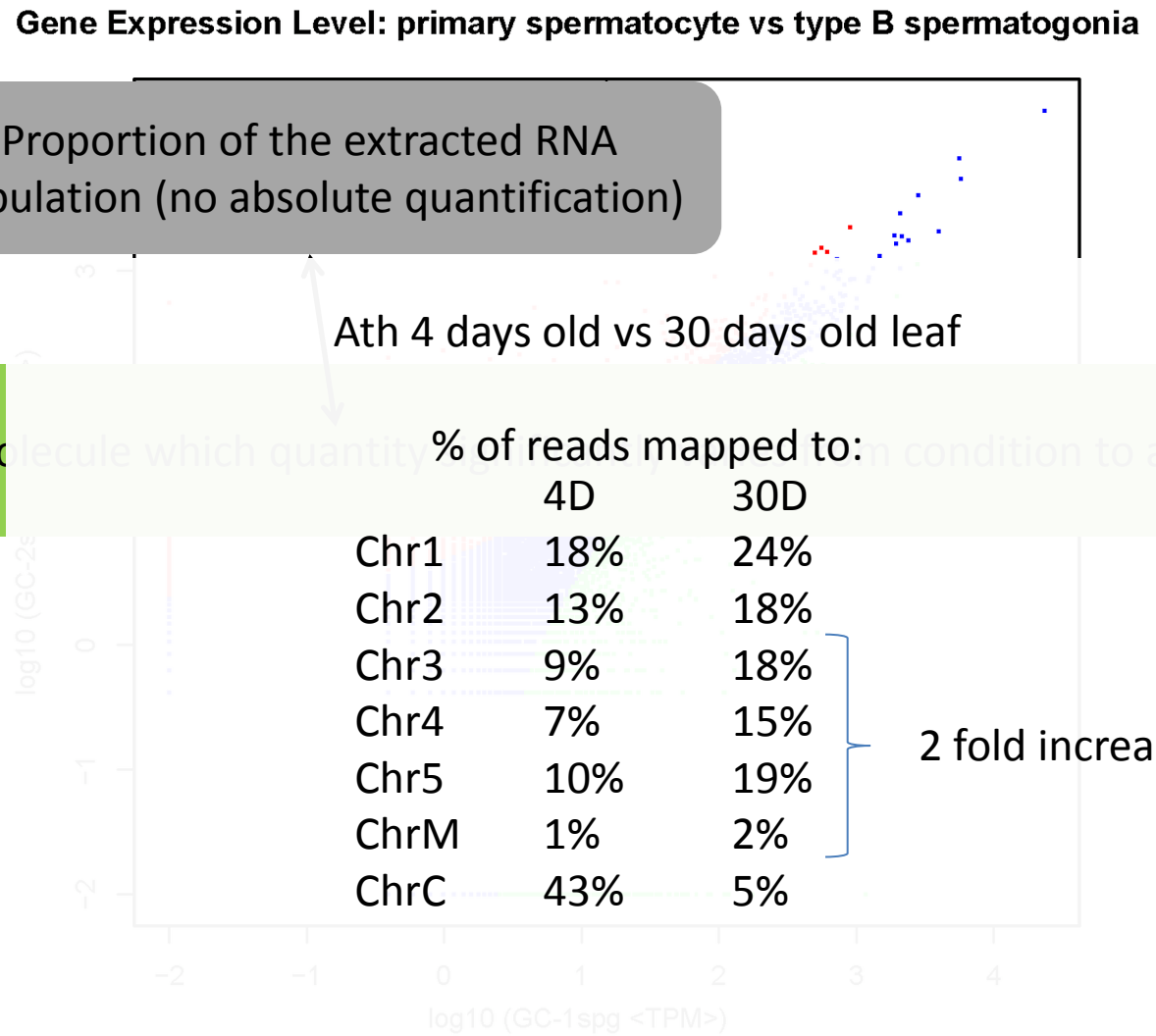
RNA molecule which quantity is significantly different from condition to another

% of reads mapped to:

	4D	30D
Chr1	18%	24%
Chr2	13%	18%
Chr3	9%	18%
Chr4	7%	15%
Chr5	10%	19%
ChrM	1%	2%
ChrC	43%	5%

One biological interpretation of ….

“differentially expressed”



Proportion of the extracted RNA population (no absolute quantification)

Ath 4 days old vs 30 days old leaf

RNA molecule which quantity is significantly different from condition to another

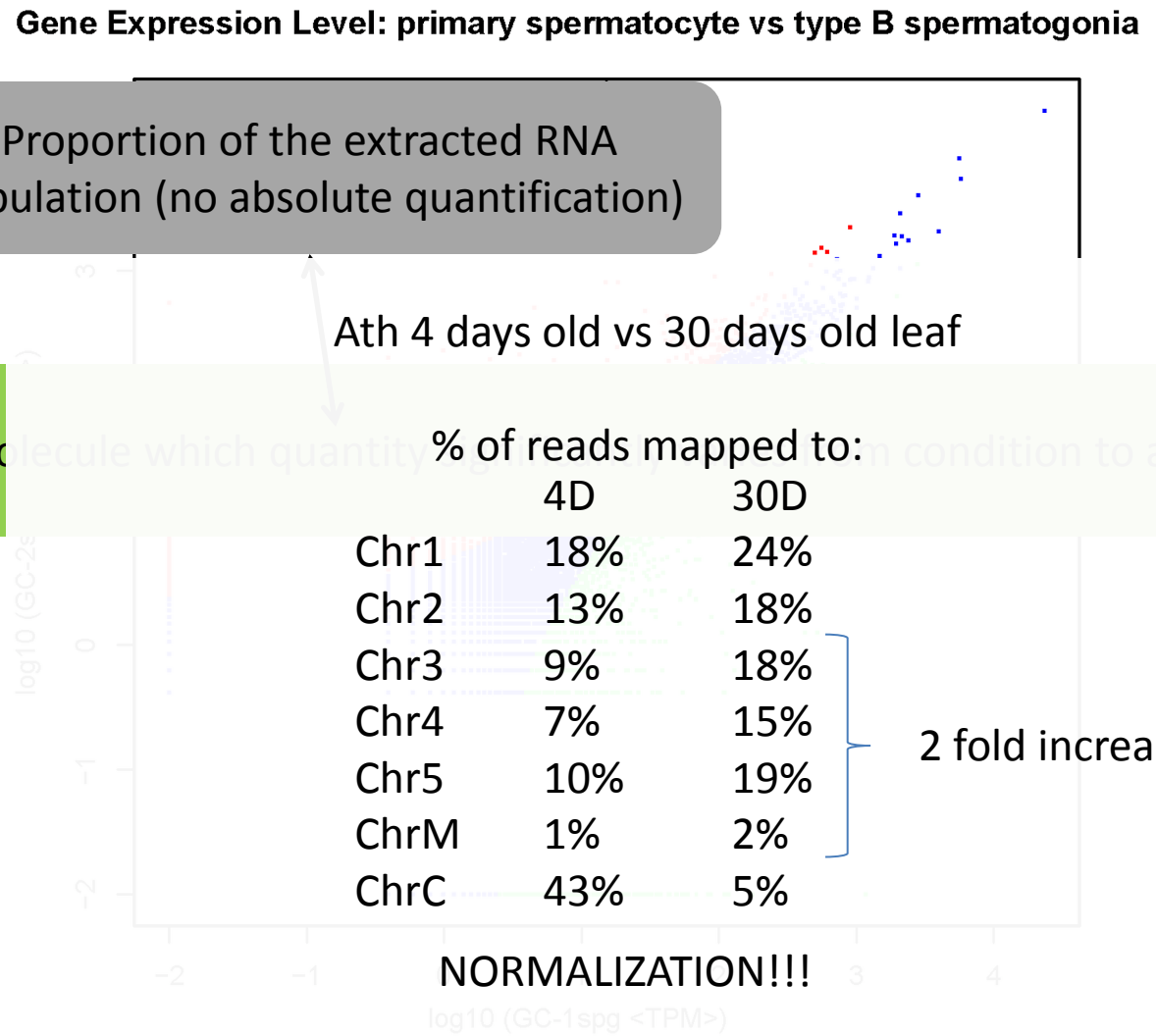
% of reads mapped to:

	4D	30D
Chr1	18%	24%
Chr2	13%	18%
Chr3	9%	18%
Chr4	7%	15%
Chr5	10%	19%
ChrM	1%	2%
ChrC	43%	5%

2 fold increase!

One biological interpretation of ….

“differentially expressed”



Proportion of the extracted RNA population (no absolute quantification)

Ath 4 days old vs 30 days old leaf

RNA molecule which quantity is significantly different from condition to another

% of reads mapped to:

	4D	30D
Chr1	18%	24%
Chr2	13%	18%
Chr3	9%	18%
Chr4	7%	15%
Chr5	10%	19%
ChrM	1%	2%
ChrC	43%	5%

2 fold increase!

NORMALIZATION!!!

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

How?

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

How?

Over-expressed RNA

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

How?

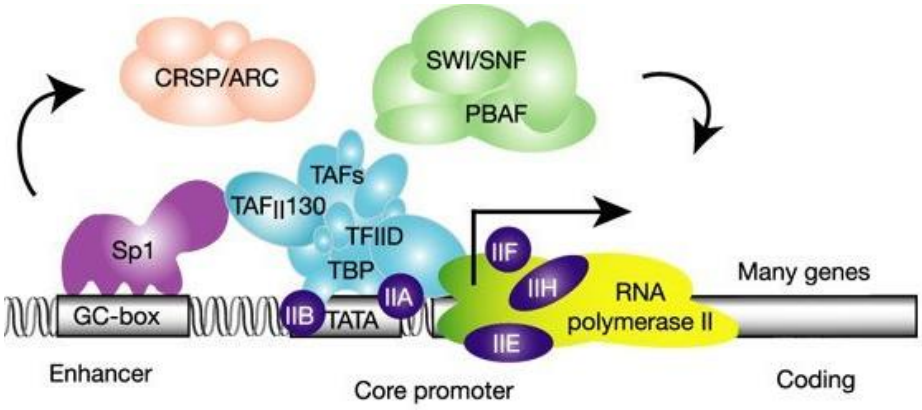
Over-expressed RNA  Transcription induction

One biological interpretation of ….
“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

How?

Over-expressed RNA → Transcription induction



One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

How?

Over-expressed RNA → Transcription induction

→ Increased stability:
miRNA, RNA binding proteins, exosome

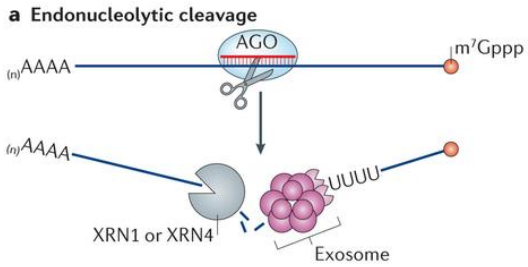
One biological interpretation of ….
“differentially expressed”

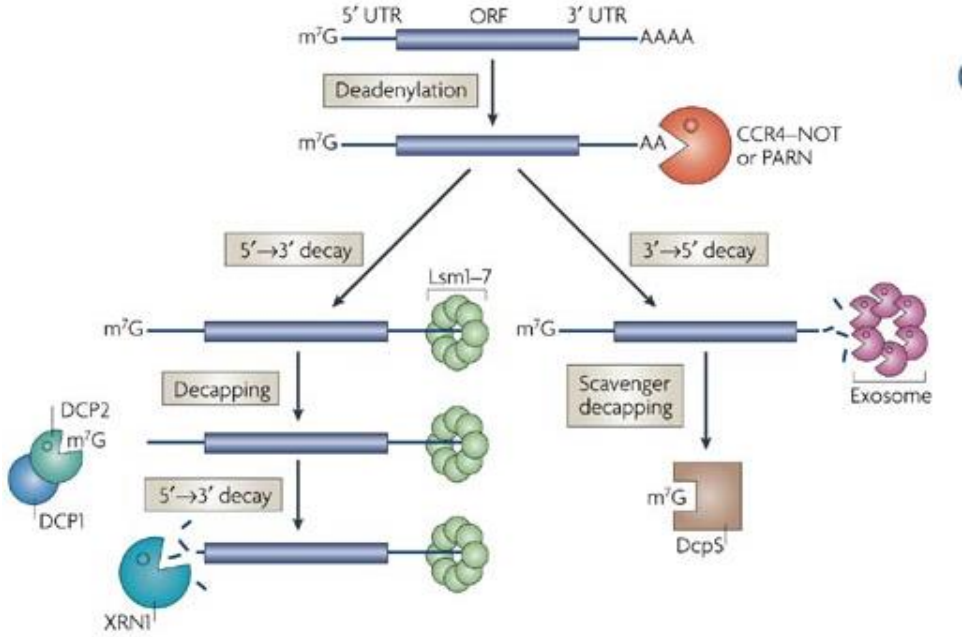
RNA molecule which quantity significantly varies from condition to another

How?

Over-expressed RNA → Transcription induction

→ Increased stability:
miRNA, RNA binding proteins, exosome





One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

How?

Over-expressed RNA → Transcription induction

→ Increased stability:
miRNA, RNA binding proteins, exosome

↓

Additional experiments are required to know “how?”

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context

Experimental design (study of the parameter of interest)

One biological interpretation of ….

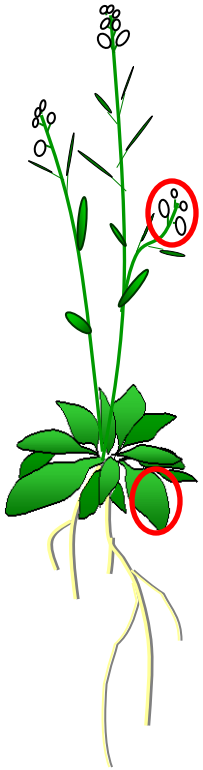
“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context

Experimental design (study of the parameter of interest)



Leaves
/
Floral buds

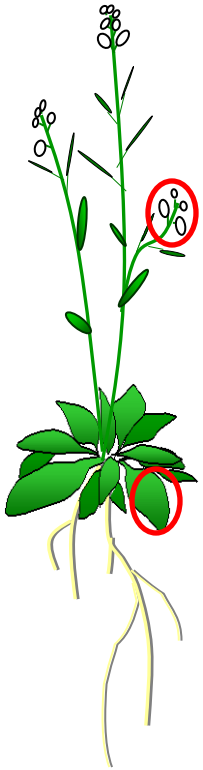
One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



Leaves
/
Floral buds

Experimental design
(study of the
parameter of interest)

30 fold over-expression of *At2g34430* in leaves vs. flower buds

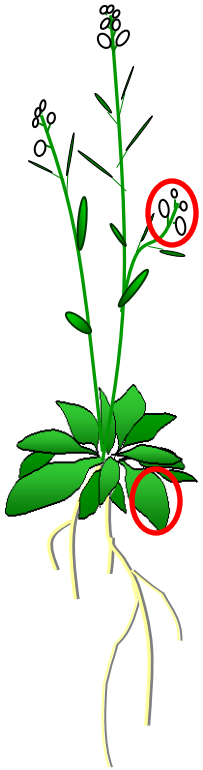
One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



Leaves
/
Floral buds

Experimental design
(study of the
parameter of interest)

30 fold over-expression of *At2g34430* in leaves vs. flower buds

At2g34430 = *LHCB1.4* :
“Light-harvesting chlorophyll-protein complex II subunit B1”

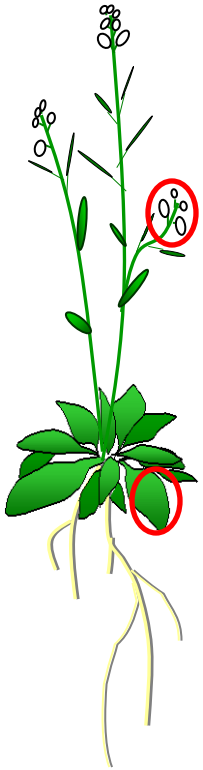
One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



Leaves
/
Floral buds

Experimental design
(study of the
parameter of interest)

30 fold over-expression of *At2g34430* in leaves vs. flower buds

At2g34430 = *LHCB1.4* :
“Light-harvesting chlorophyll-protein complex II subunit B1”

In leaves, there is more photosynthesis than in floral buds
so
More LHCB1.4 is required to capture light

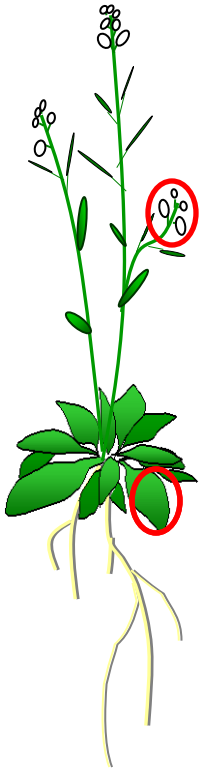
One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



Leaves
/
Floral buds

Experimental design
(study of the
parameter of interest)

30 fold over-expression of *At2g34430* in leaves vs. flower buds

At2g34430 = *LHCB1.4* :
“Light-harvesting chlorophyll-protein complex II subunit B1”

In leaves, there is more LHCB1.4 than in floral buds
so
There is more photosynthesis in leaves than in floral buds

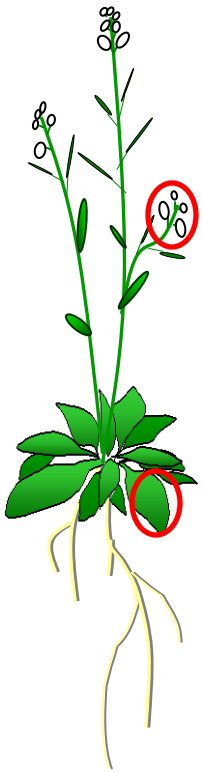
One biological interpretation of ….

“differentially expressed”

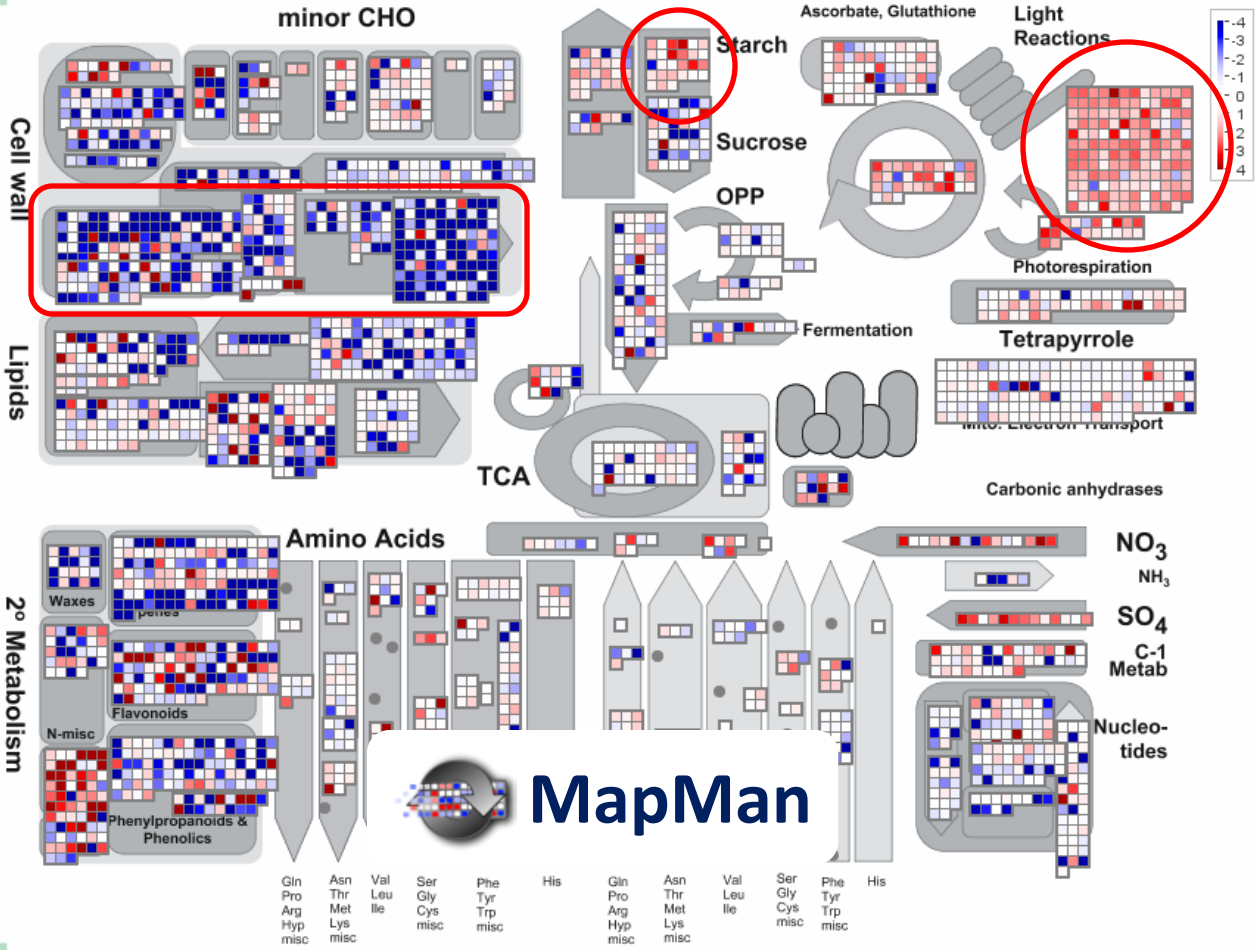
RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



Leaves
/
Floral buds



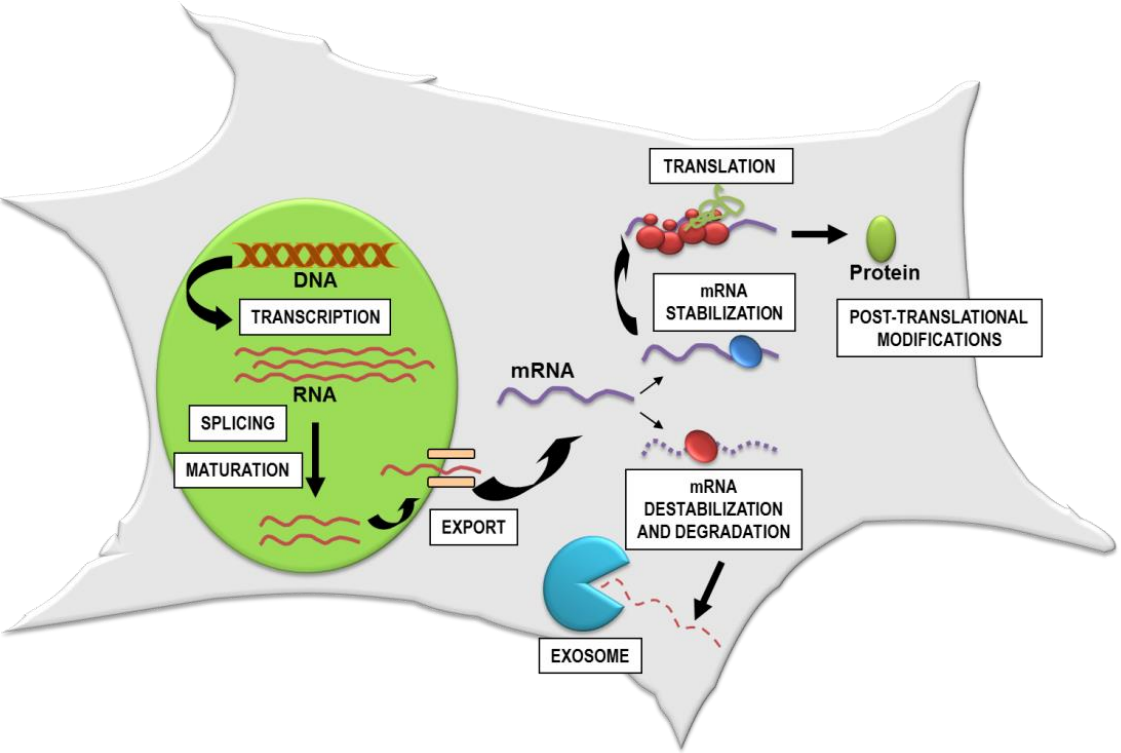
One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



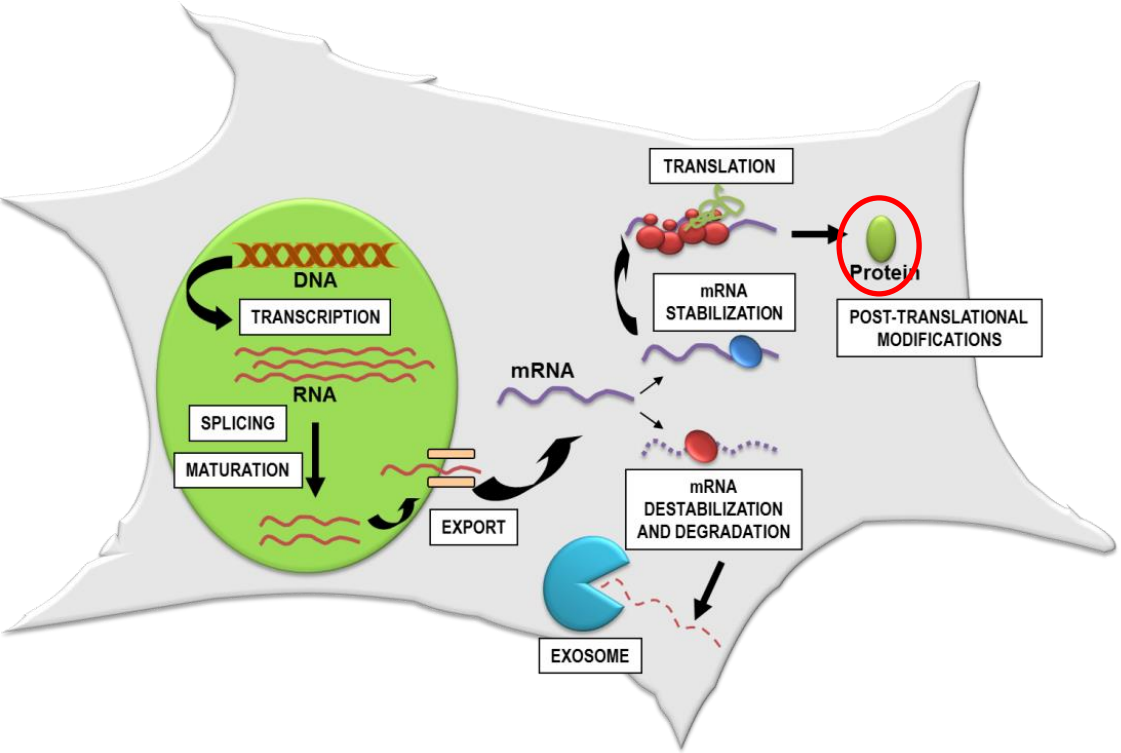
One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



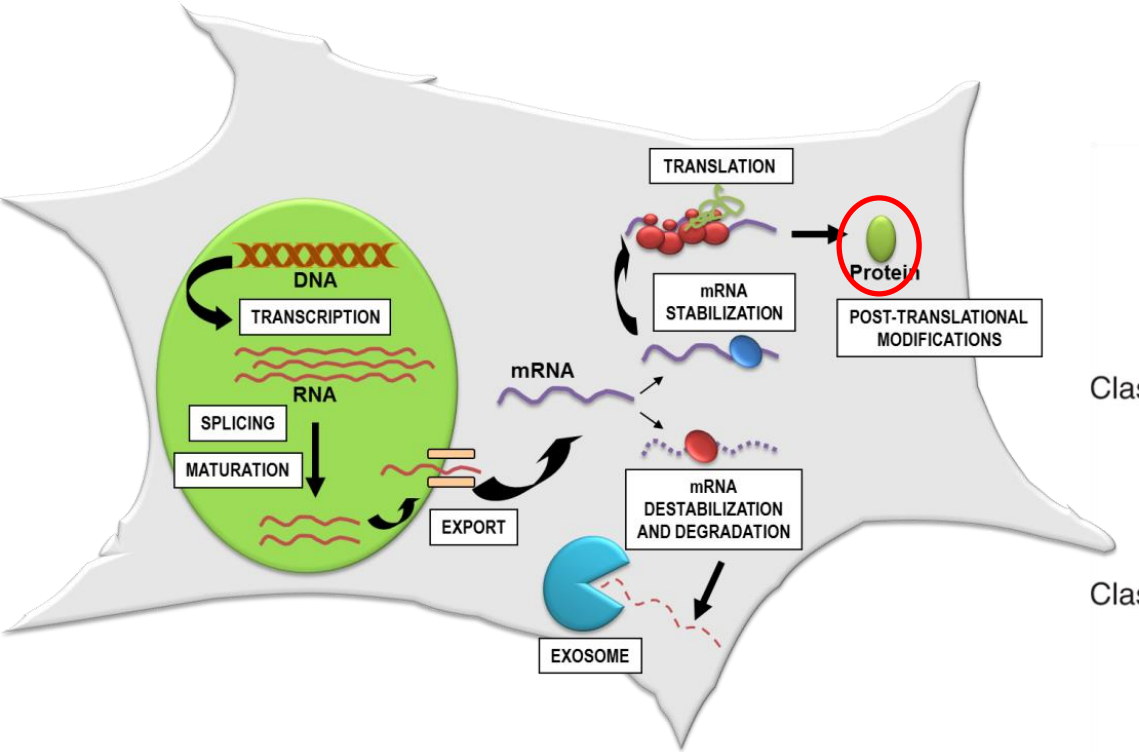
One biological interpretation of ….

“differentially expressed”

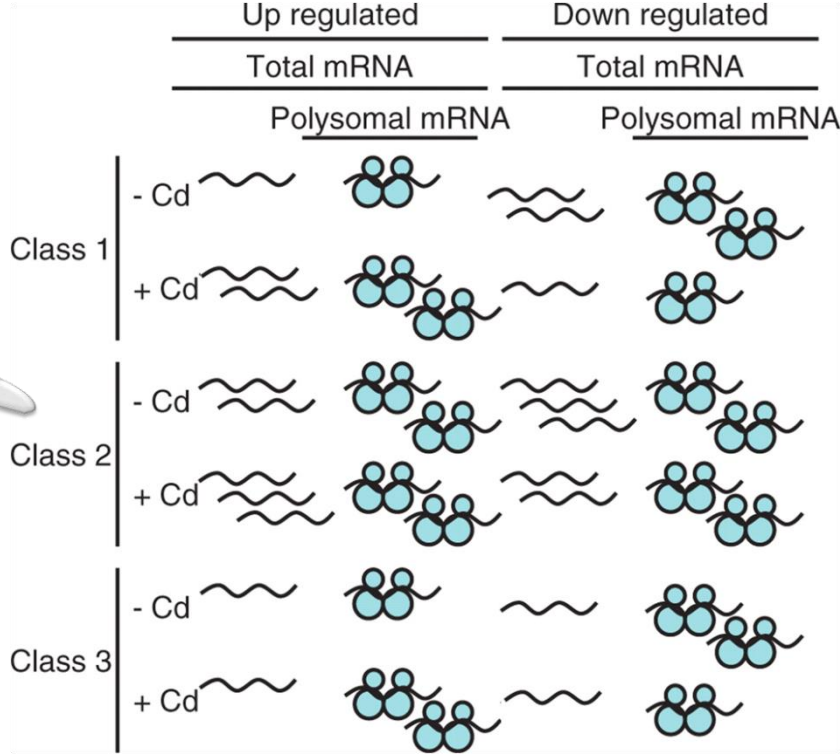
RNA molecule which quantity significantly varies from condition to another

What for?

Differentially expressed RNA = Response to changes of the cellular context



Cadmium treatment in Ath and translational regulation



One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

Biomarker?

One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

Biomarker?

Differentially expressed RNA = Specific (?) response to this change of the cellular context

One biological interpretation of ….

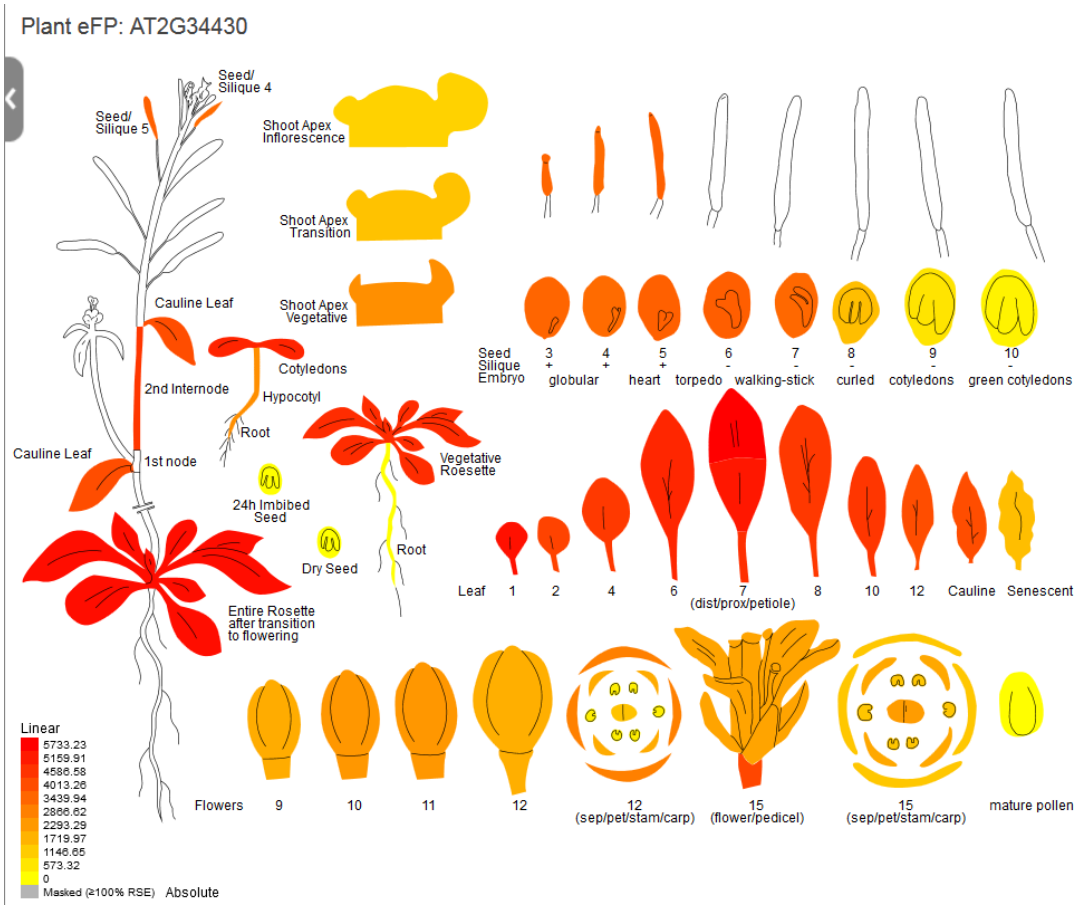
“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

Biomarker?

Differentially expressed RNA = Specific (?) response to this change of the cellular context

Expression profile



One biological interpretation of ….

“differentially expressed”

RNA molecule which quantity significantly varies from condition to another

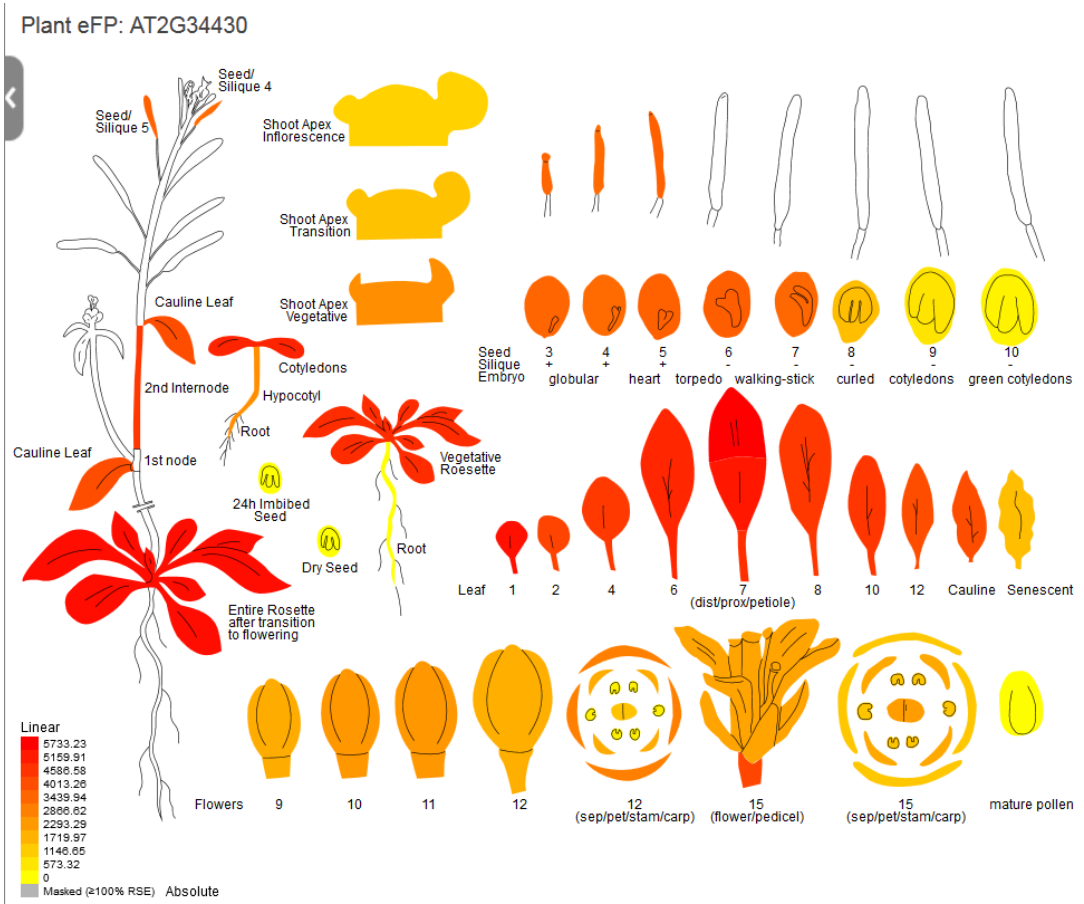
Biomarker?

Differentially expressed RNA = Specific (?) response to this change of the cellular context

Expression profile

A single RNA is never specific

Combination of transcripts
(multivariate analysis)



One biological interpretation of ….
“differentially expressed”

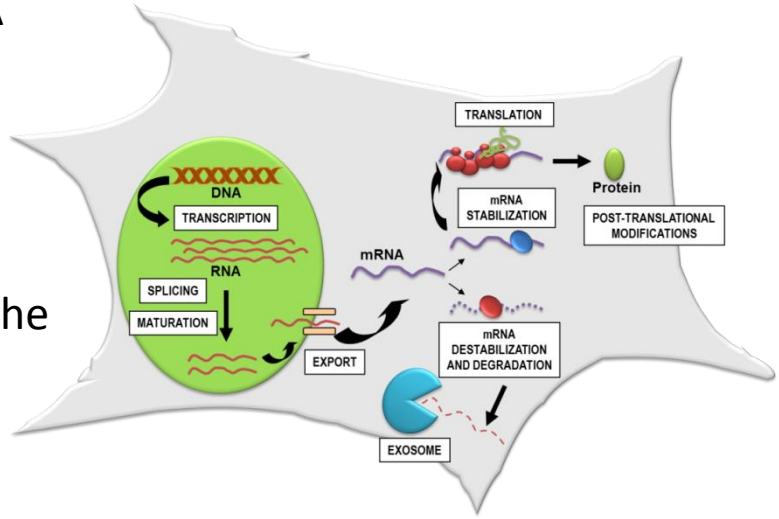
Differentially accumulated RNA

RNA involved in the response to the change of the cellular context

Candidate RNA

Direct or indirect involvement?
Cause or consequence?
How?

Additional experiments to be done…



One RNA vs several RNAs

One biological interpretation of ….

“differentially expressed”

Differentially accumulated RNA

One biological interpretation of ….
“differentially expressed”

Differentially accumulated RNA

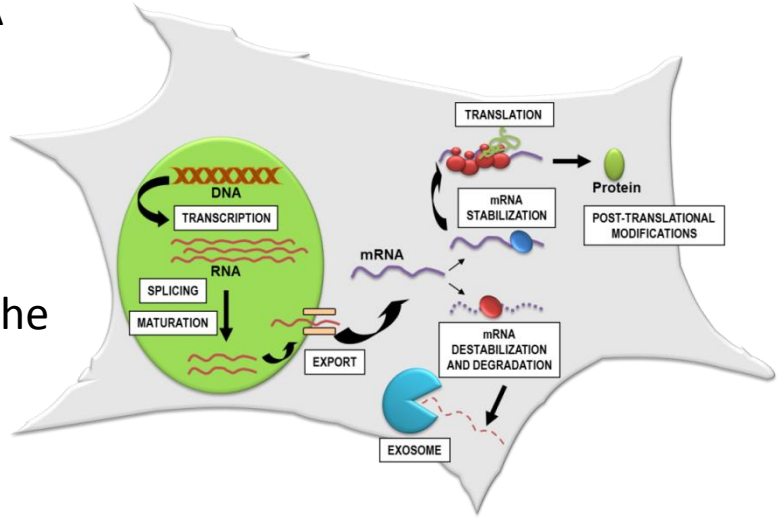
RNA involved in the response to the change of the cellular context

One biological interpretation of ….

“differentially expressed”

Differentially accumulated RNA

RNA involved in the response to the change of the cellular context

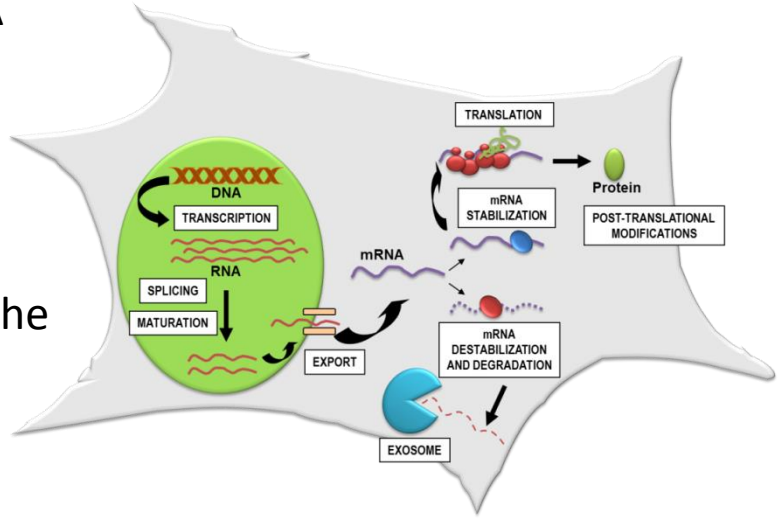


One biological interpretation of ….

“differentially expressed”

Differentially accumulated RNA

RNA involved in the response to the change of the cellular context

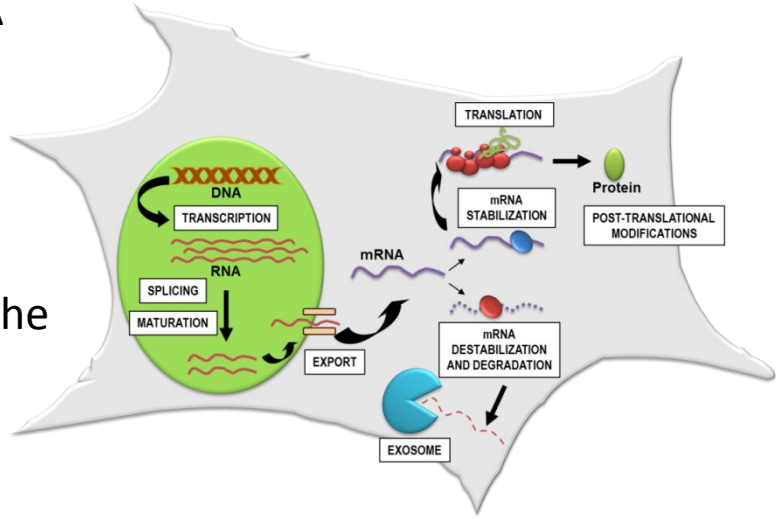


Direct or indirect involvement?
Cause or consequence?
How?

One biological interpretation of ….
“differentially expressed”

Differentially accumulated RNA

RNA involved in the response to the change of the cellular context



Direct or indirect involvement?
Cause or consequence?
How?

Additional experiments to be done…

One biological interpretation of ….
“differentially expressed”

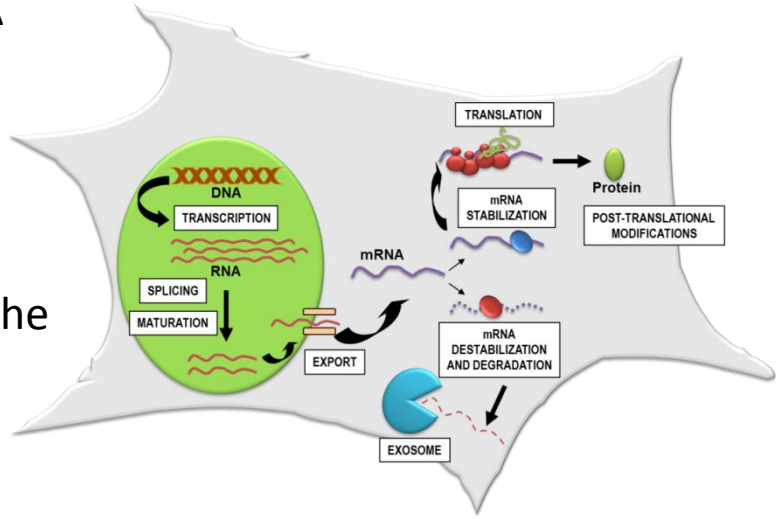
Differentially accumulated RNA

RNA involved in the response to the change of the cellular context

Candidate RNA

Direct or indirect involvement?
Cause or consequence?
How?

Additional experiments to be done…



One biological interpretation of ….
“differentially expressed”

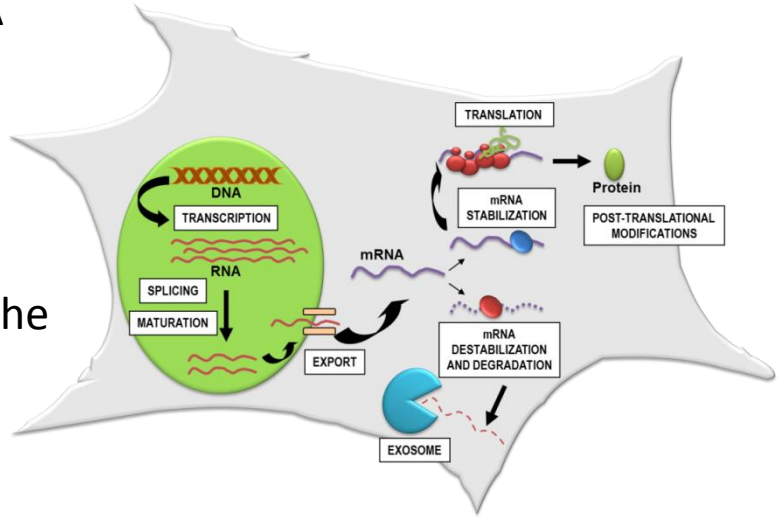
Differentially accumulated RNA

RNA involved in the response to the change of the cellular context

Candidate RNA

Direct or indirect involvement?
Cause or consequence?
How?

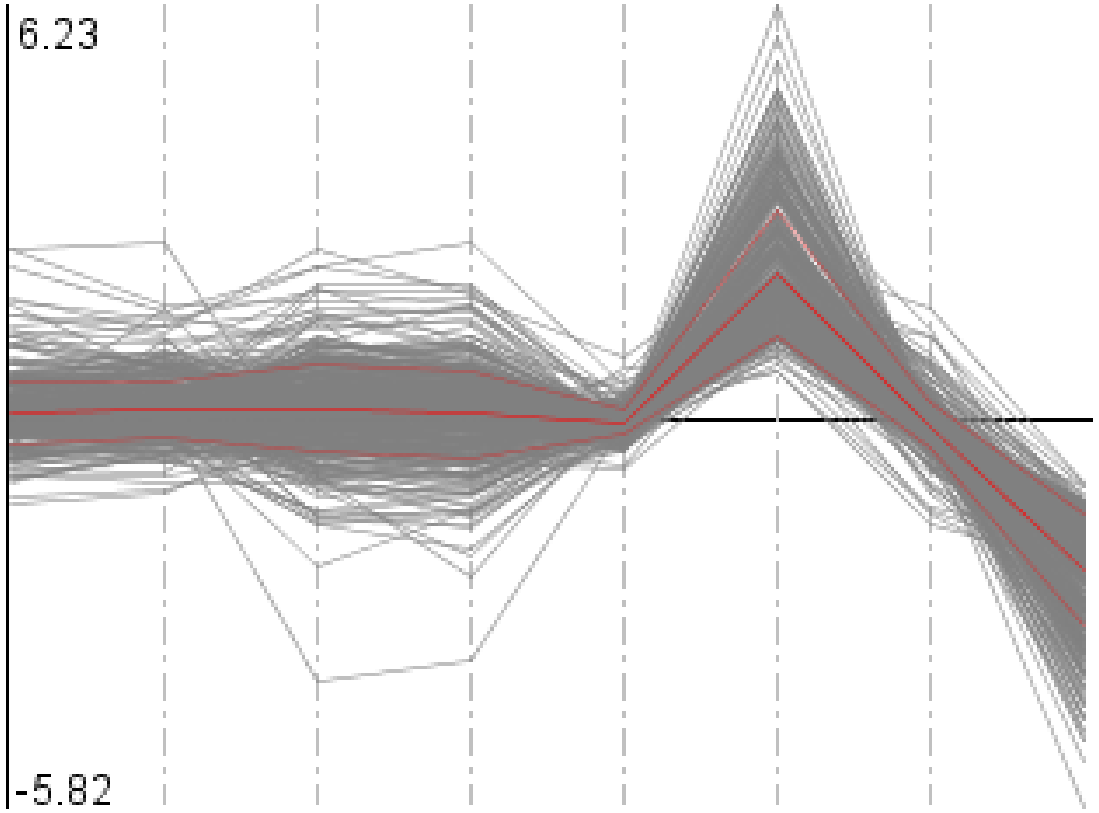
Additional experiments to be done…



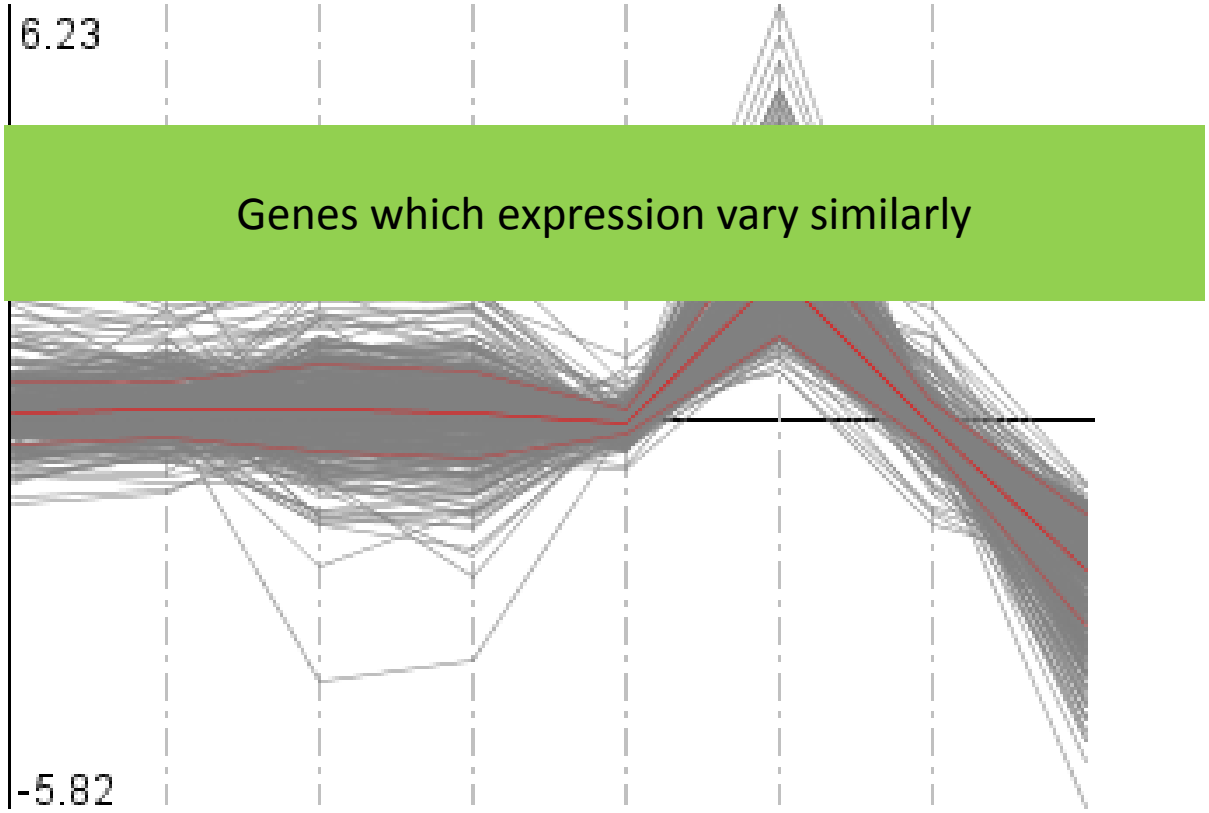
One RNA vs several RNAs

One biological interpretation of

"co-expressed"



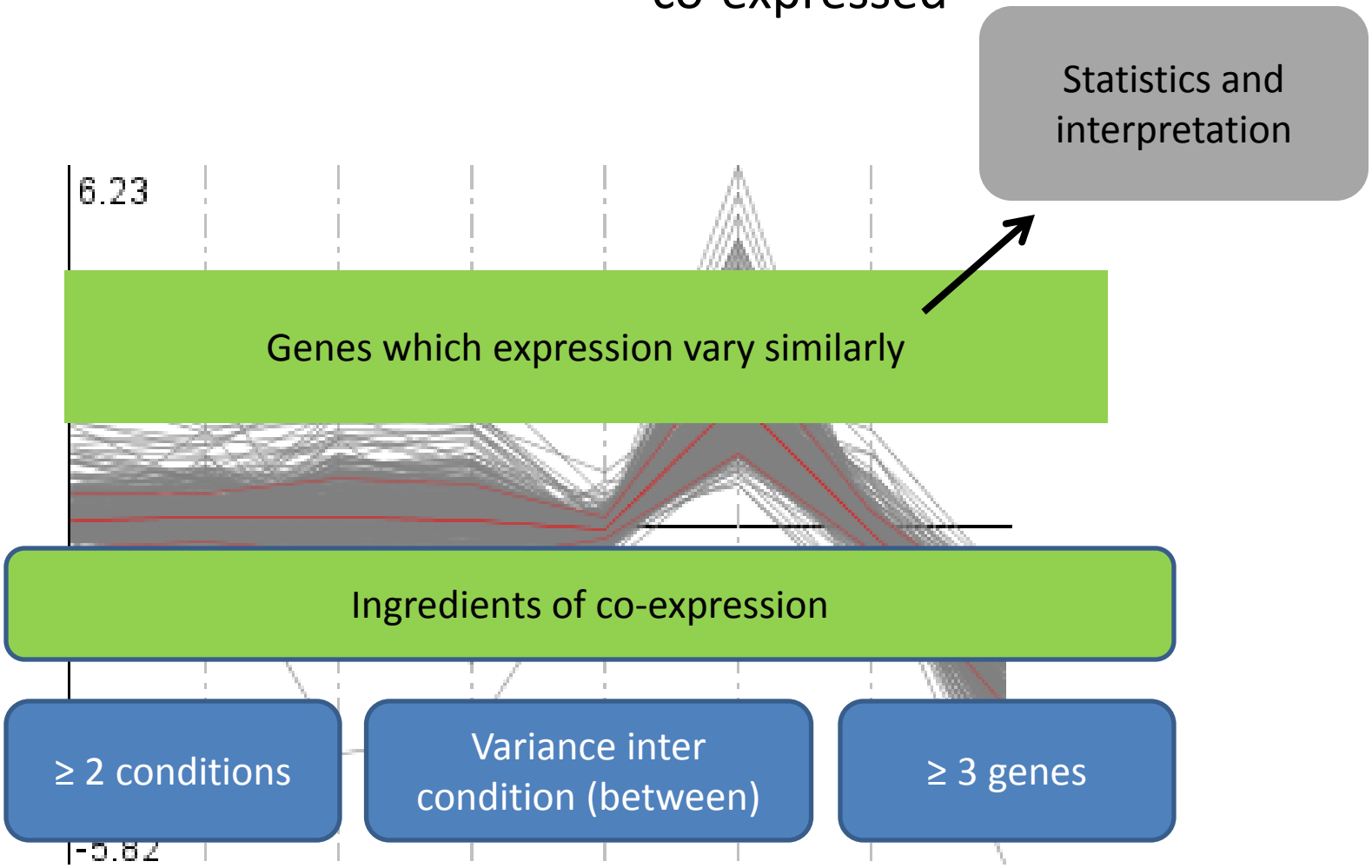
One biological interpretation of “co-expressed”



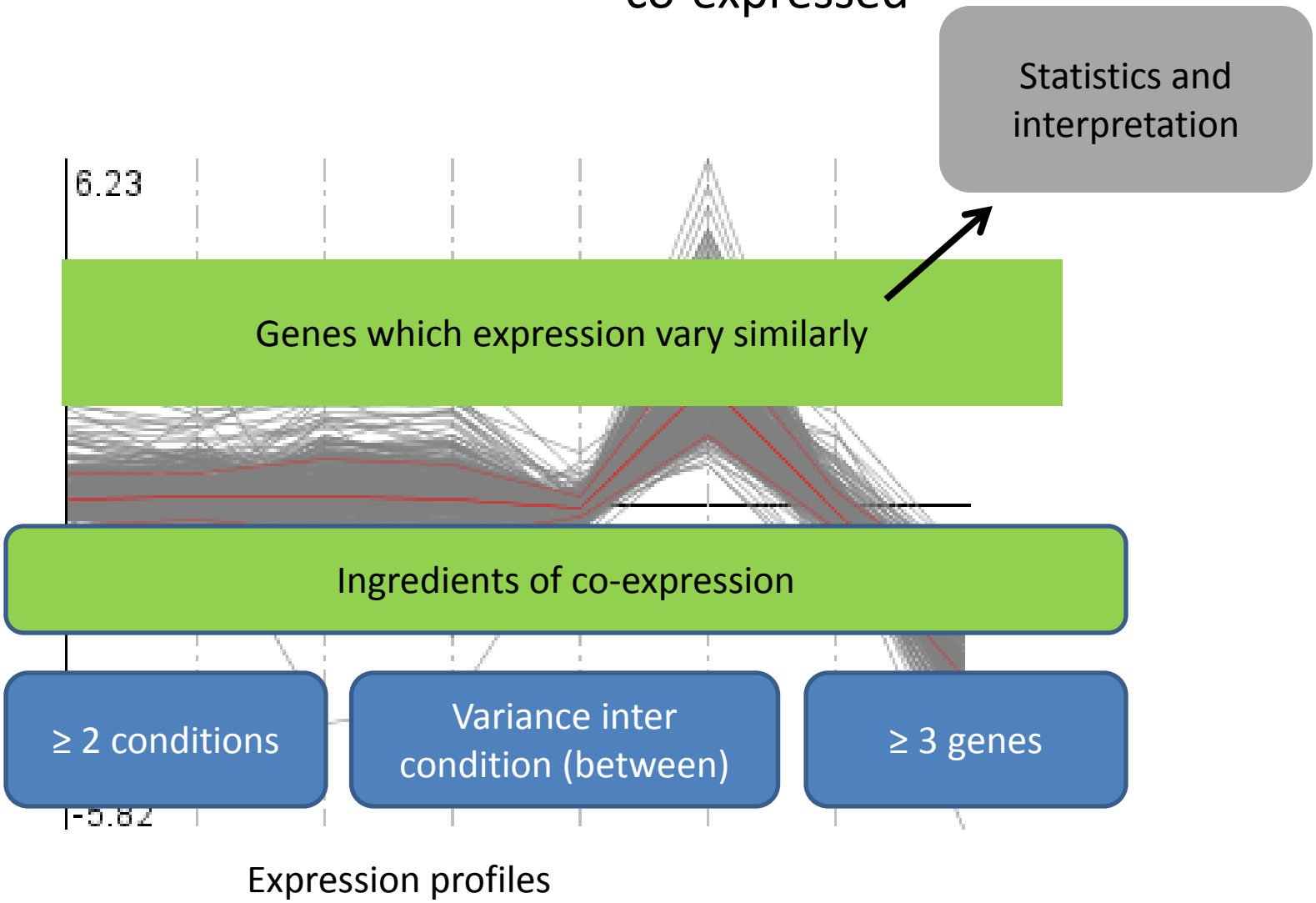
One biological interpretation of “co-expressed”

Statistics and interpretation

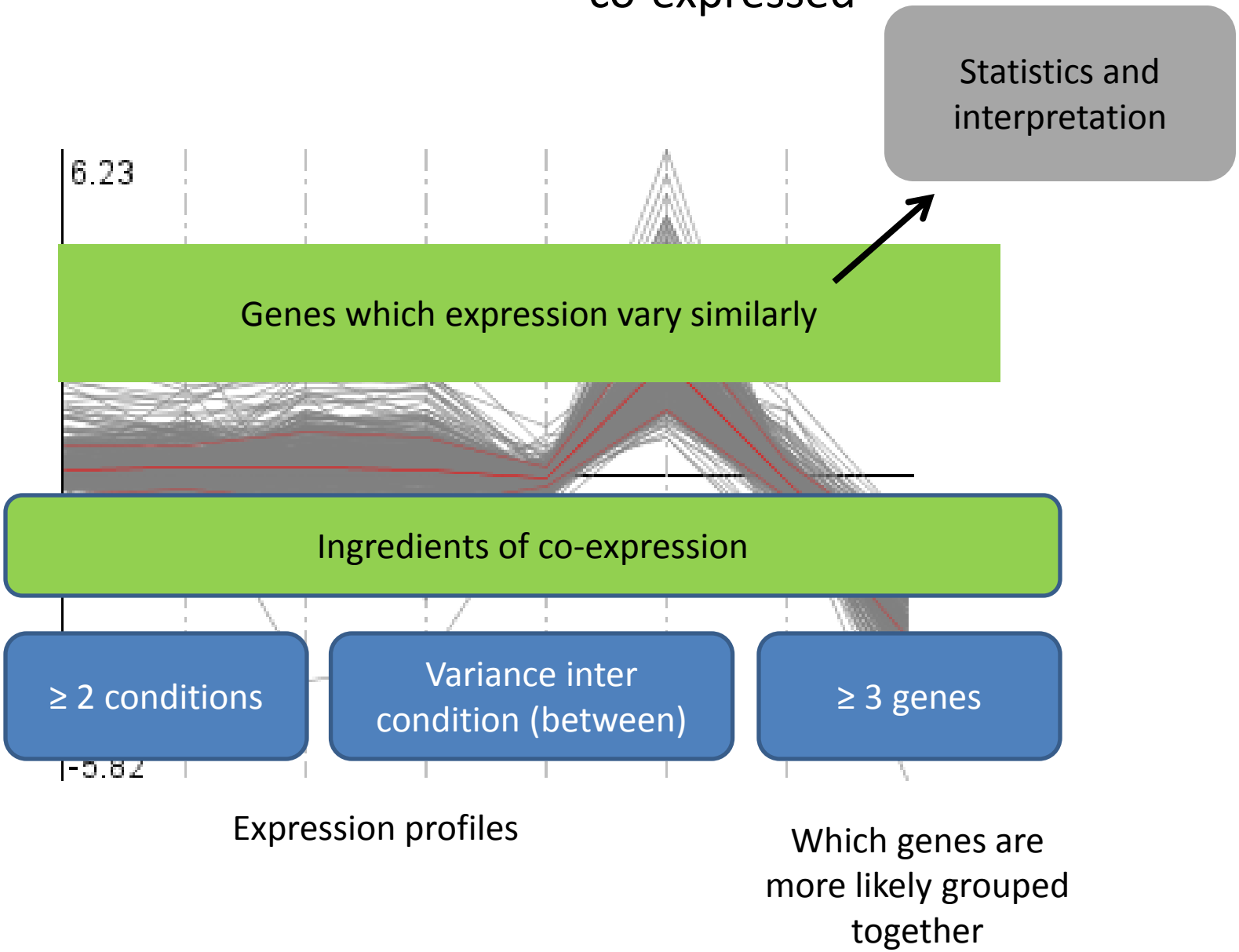
One biological interpretation of “co-expressed”



One biological interpretation of "co-expressed"



One biological interpretation of “co-expressed”



One biological interpretation of “co-expressed”

Genes which expression vary similarly

What for?

Co-expressed genes= genes involved in the same responses to the same changes of the cellular context?

= genes involved in the same biological processes?

One biological interpretation of "co-expressed"

Genes which expression vary similarly

What for?

>Co-expressed genes→ functional partners?

One biological interpretation of "co-expressed"

Genes which expression vary similarly

What for?

>Co-expressed genes→ functional partners?

Every protein works with partners

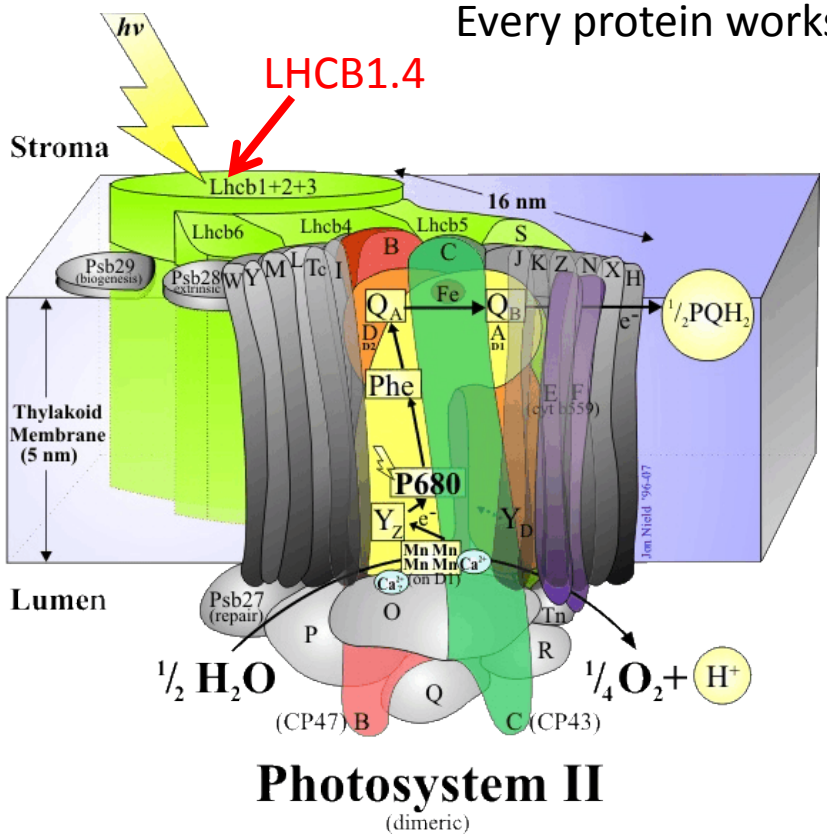
One biological interpretation of “co-expressed”

Genes which expression vary similarly

What for?

>Co-expressed genes→ functional partners?

Every protein works with partners



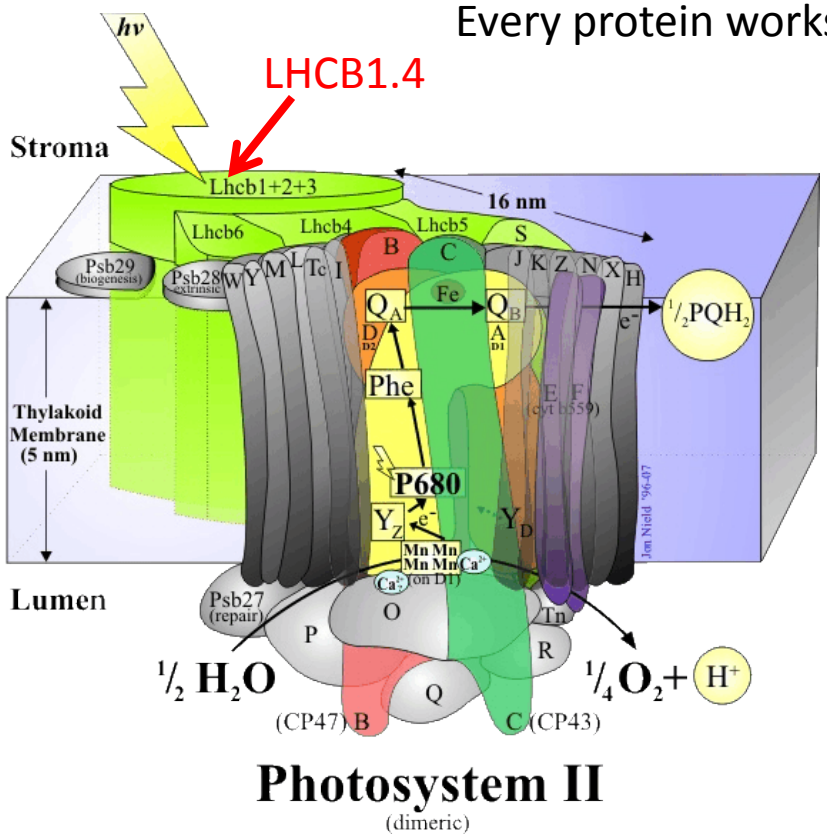
One biological interpretation of “co-expressed”

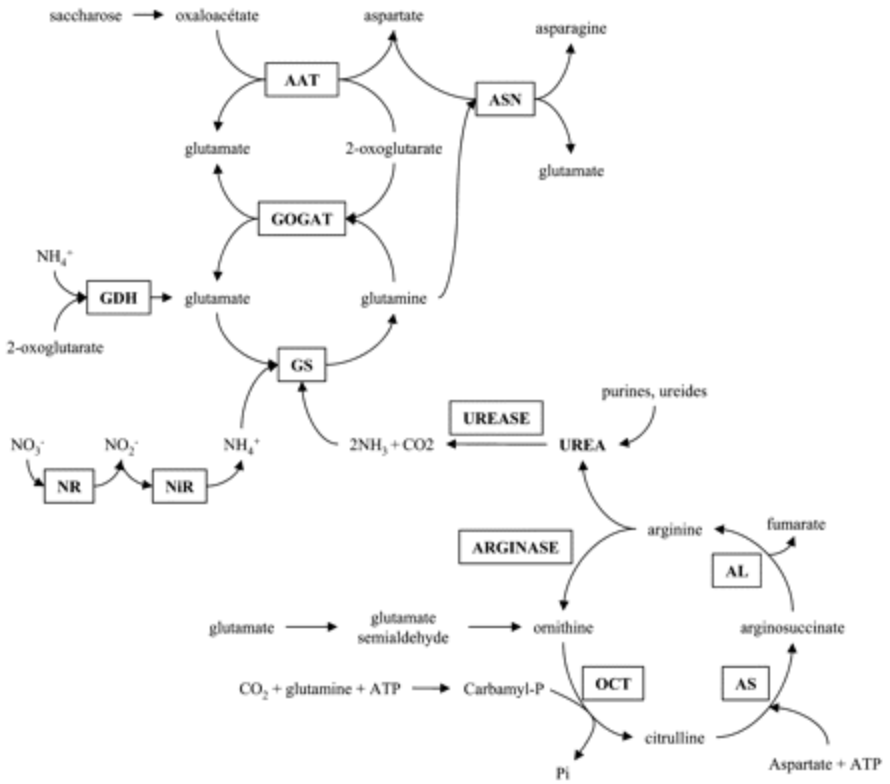
Genes which expression vary similarly

What for?

>Co-expressed genes→ functional partners?

Every protein works with partners





Nitrate assimilation pathway

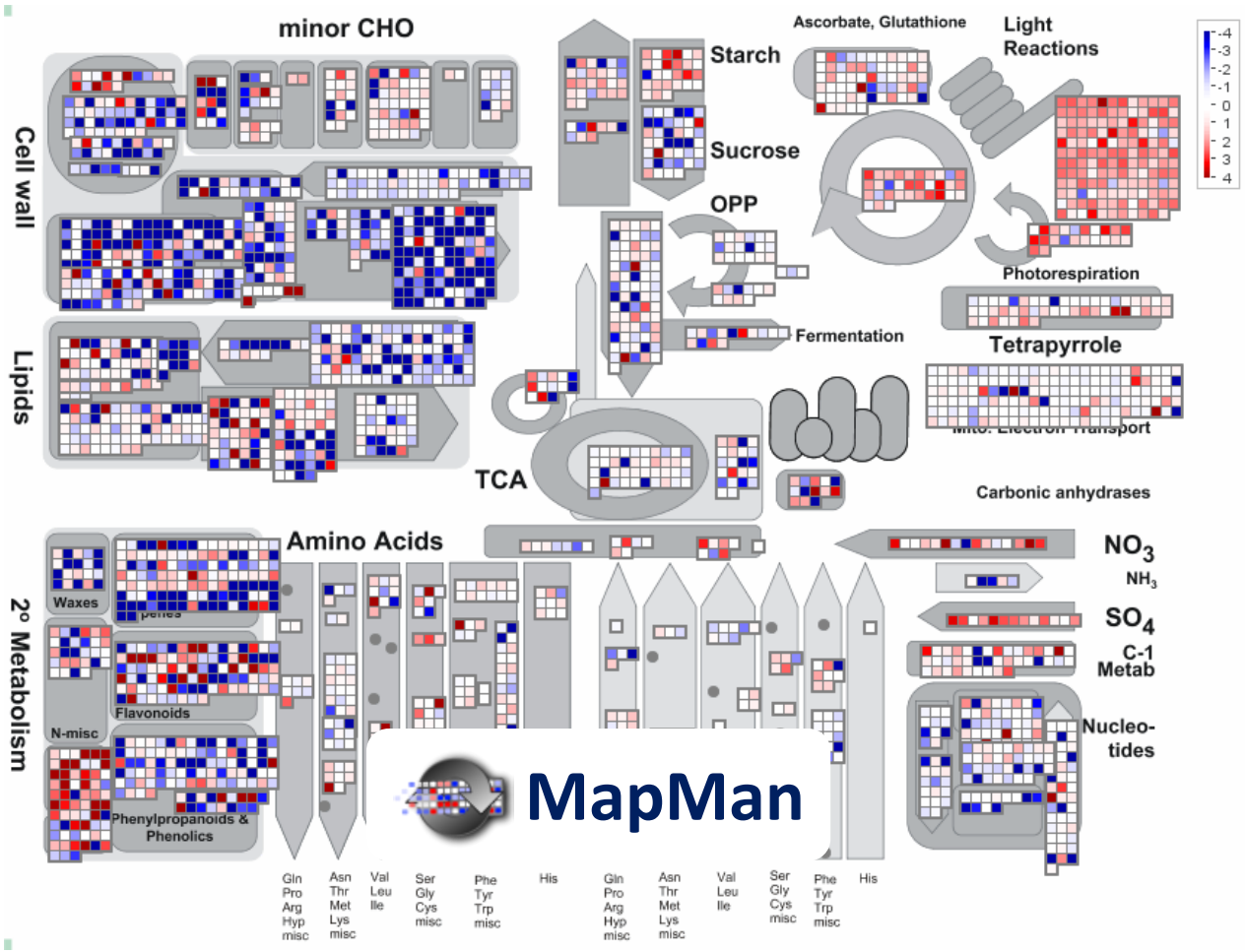
One biological interpretation of

“co-expressed”

Genes which expression vary similarly

What for?

>Co-expressed genes→ functional partners?

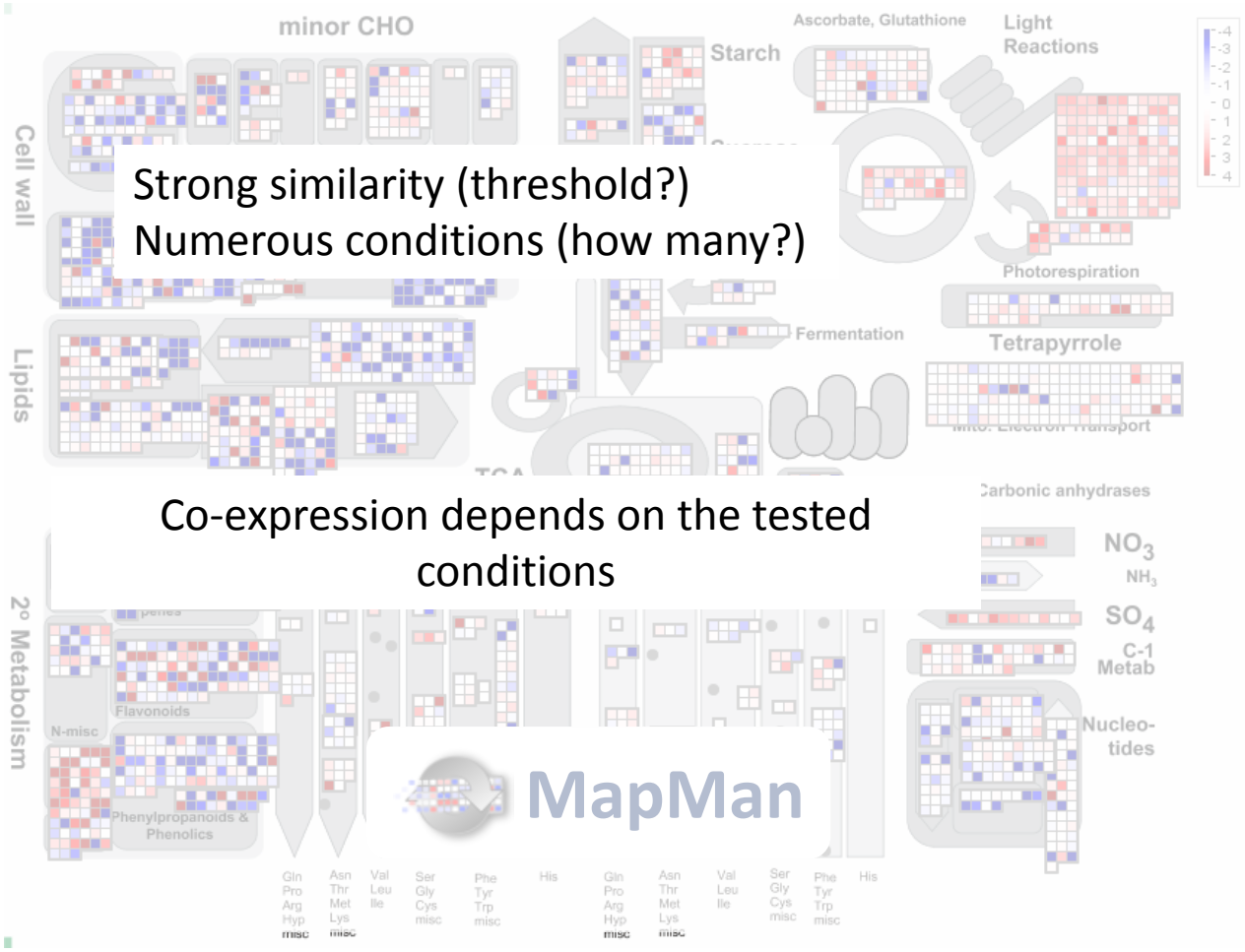


One biological interpretation of “co-expressed”

Genes which expression vary similarly

What for?

>Co-expressed genes→ functional partners?



Strong similarity (threshold?)
Numerous conditions (how many?)

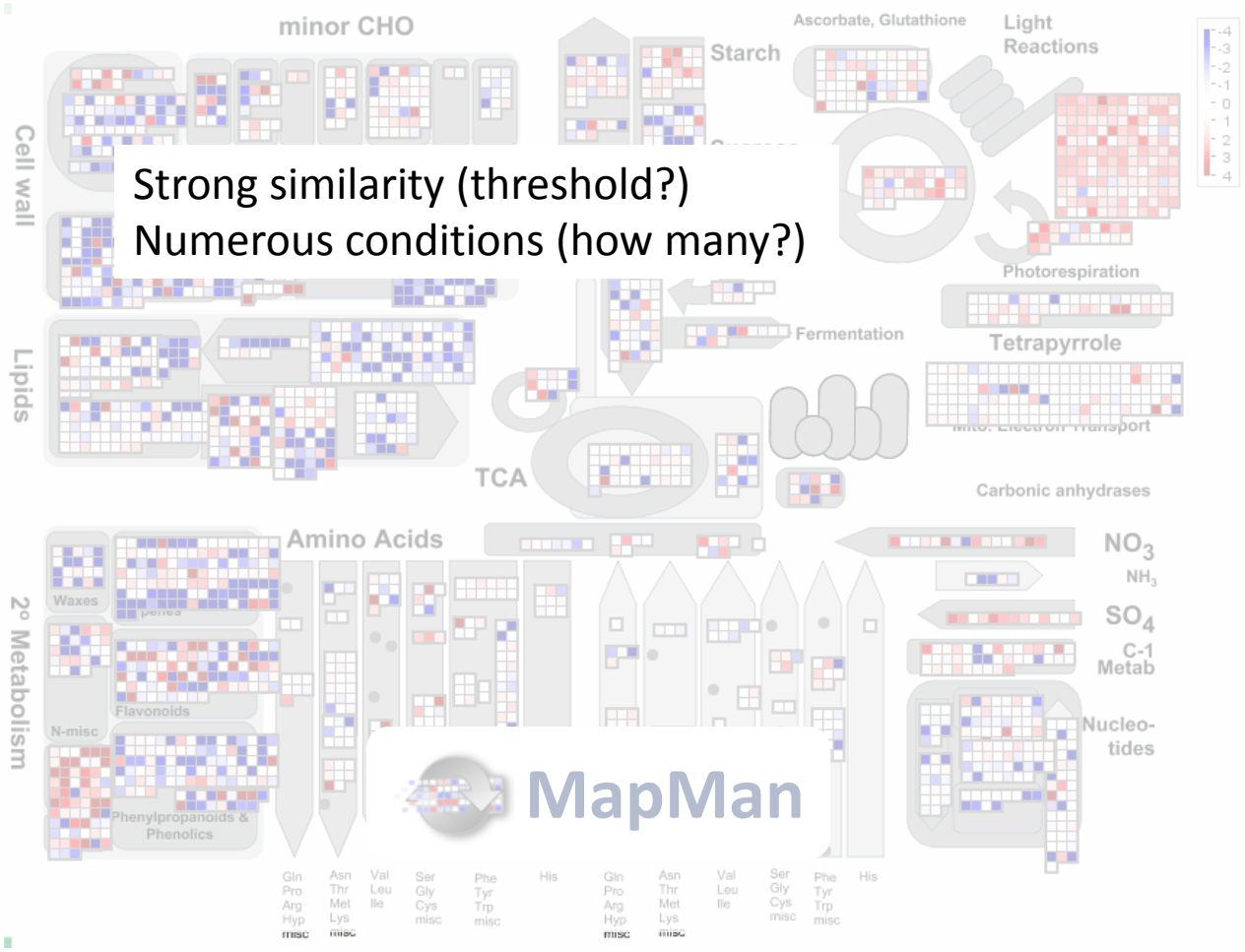
Co-expression depends on the tested conditions

One biological interpretation of “co-expressed”

Genes which expression vary similarly

What for?

>Co-expressed genes→ functional partners?



Strong similarity (threshold?)

Numerous conditions (how many?)

One biological interpretation of

“co-expressed”

Genes which expression vary similarly

How?

Co-expressed genes= under the same regulations ?

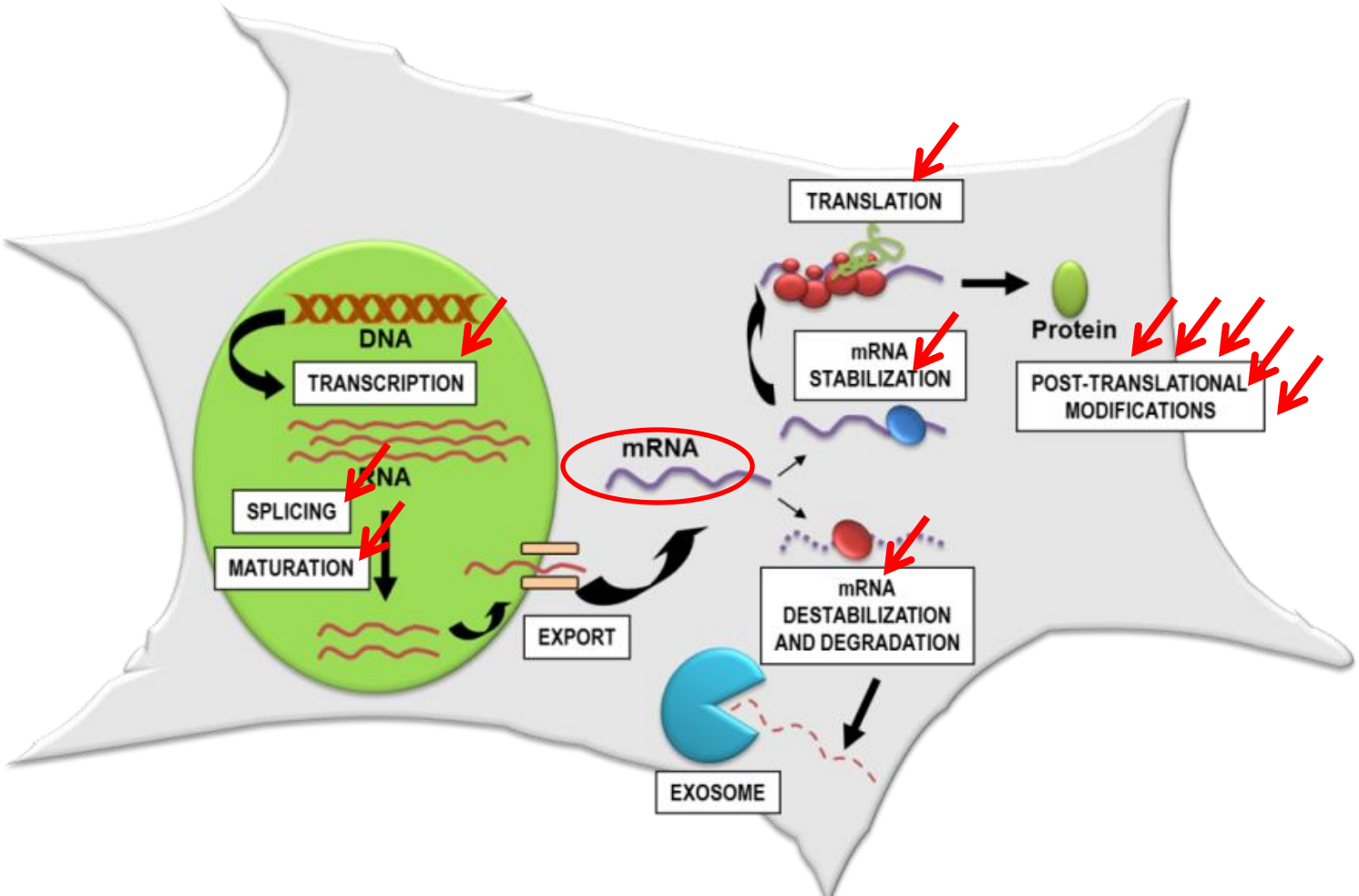
One biological interpretation of

"co-expressed"

Genes which expression vary similarly

How?

Co-expressed genes= ~~under the same regulations~~ ?

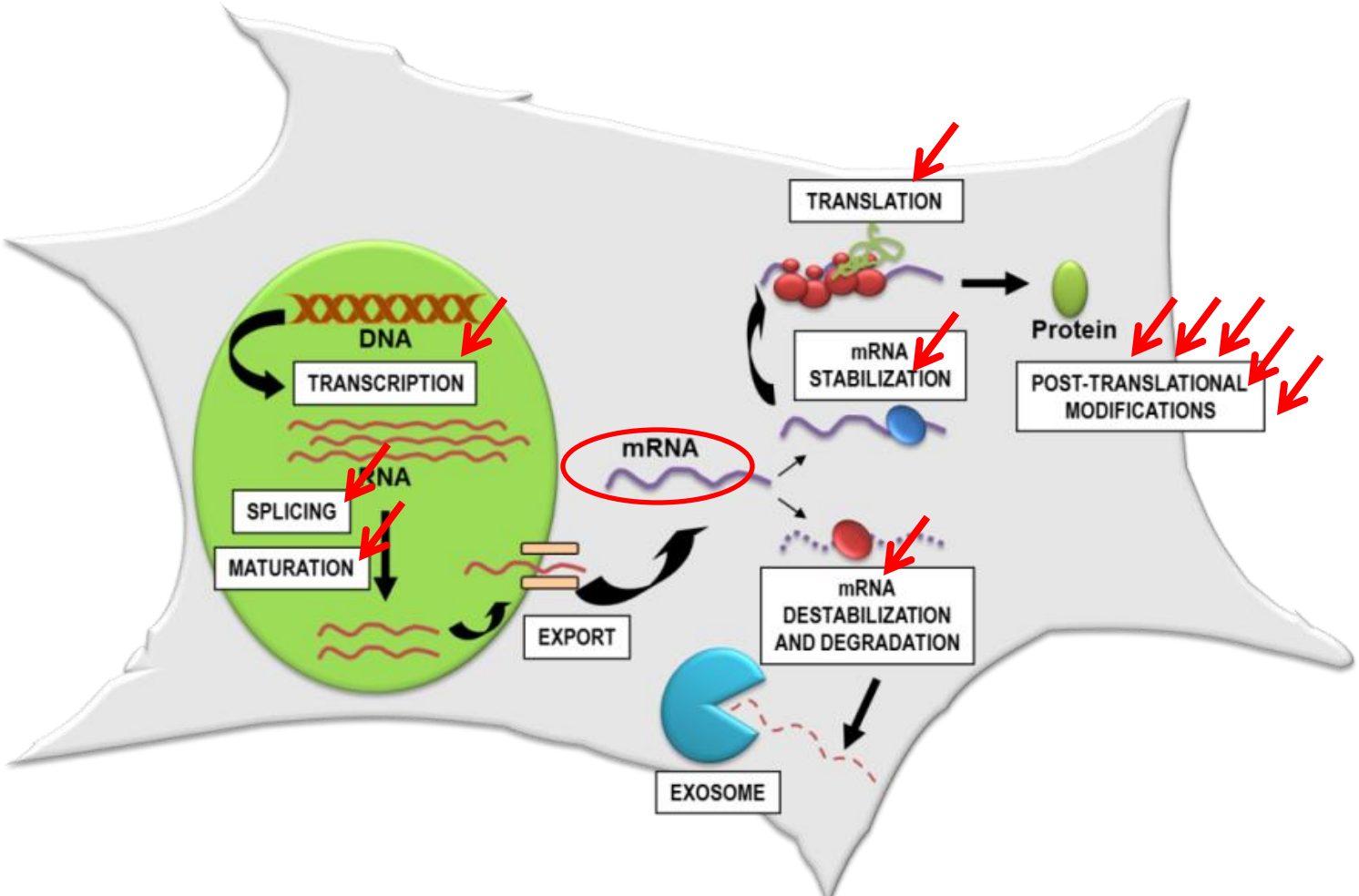


One biological interpretation of “co-expressed”

Genes which expression vary similarly

How?

Co-expressed genes= under the same regulations ?



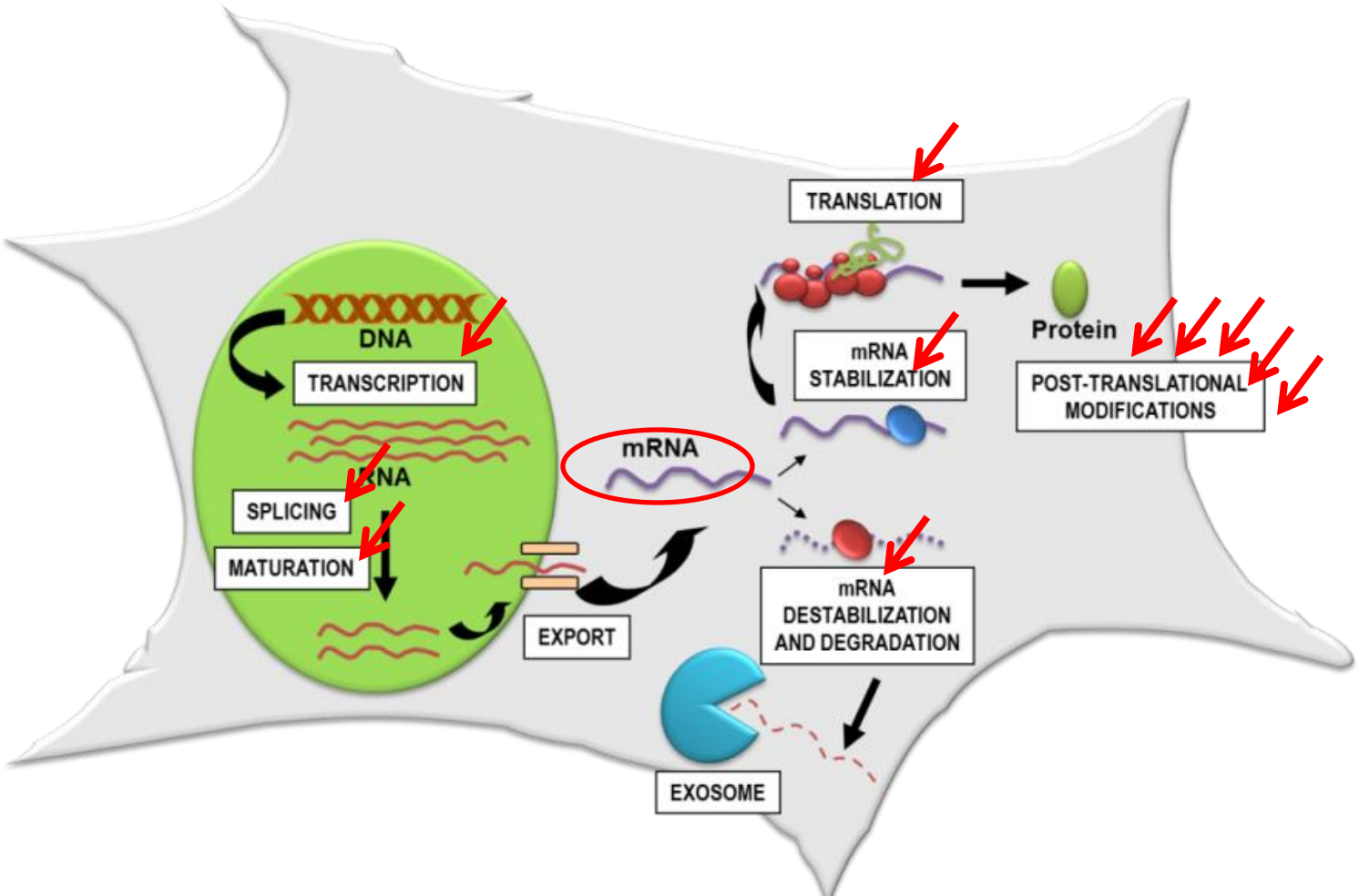
One biological interpretation of

"co-expressed"

Genes which expression vary similarly

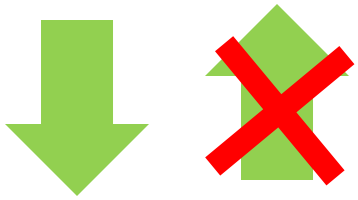
How?

Co-expressed genes= ~~under the same regulations~~ ?



One biological interpretation of “co-expressed”

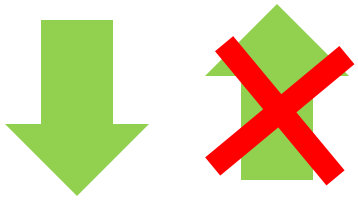
Co-expressed genes



Genes potentially involved in the same biological processes given the studied context

One biological interpretation of “co-expressed”

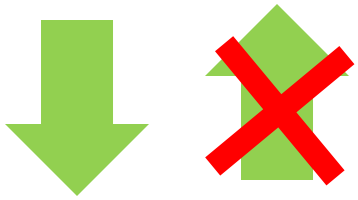
Co-expressed genes



Genes potentially involved in the same biological processes given the studied context

One biological interpretation of “co-expressed”

Co-expressed genes

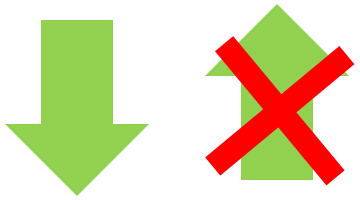


Genes potentially involved in the same biological processes given the studied context

Putative functional annotation

One biological interpretation of “co-expressed”

Co-expressed genes



Genes potentially involved in the same biological processes given the studied context

Putative functional annotation

Additional experiments to be done…