

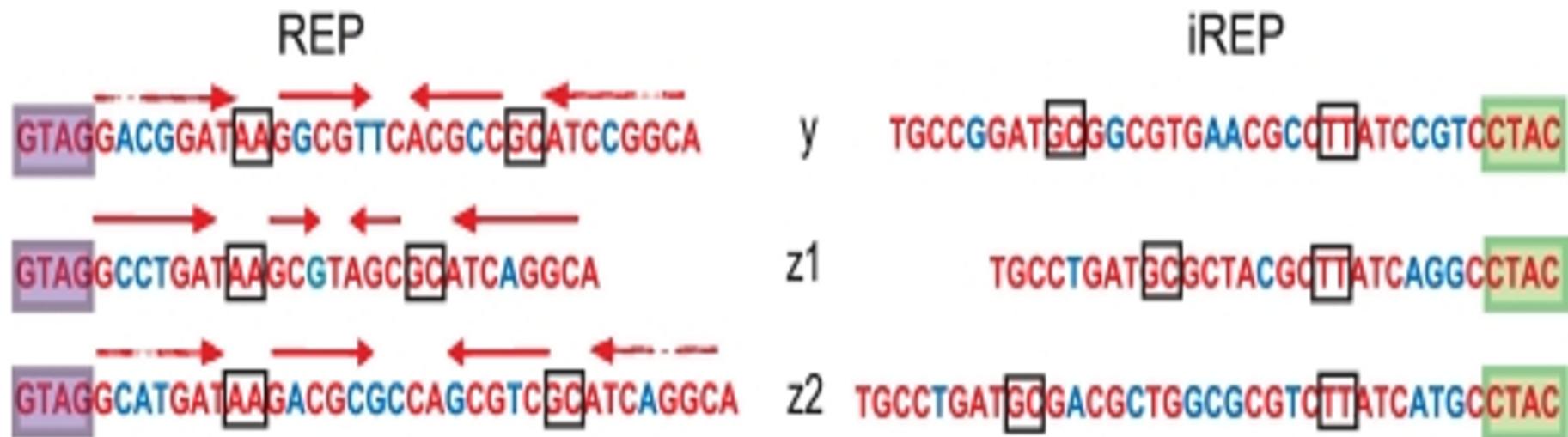
RNA-Seq study of
mRNA metabolism in operons:
role of BIME
in *Escherichia coli*.

Nicolas Jeanne
Juin 2015

Tuteurs : Yves Quentin
Roland Barriot
Rapporteur : Christine Gaspin

REP & BIME features

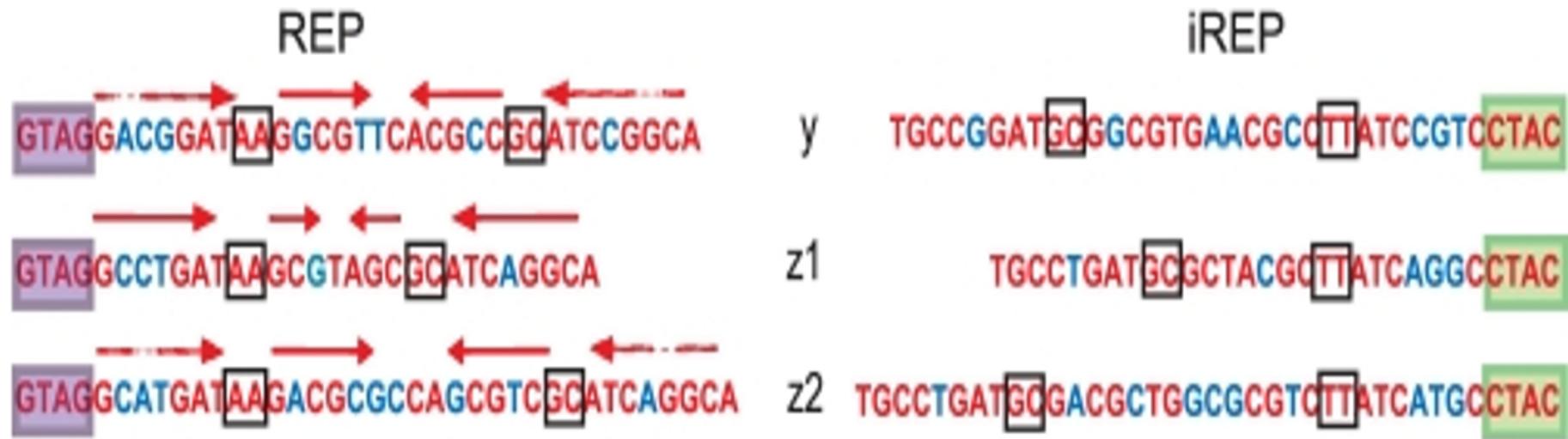
- Repeated Extragenic Palindrome.



Ton-Hoang et al. 2012

REP & BIME features

- Repeated Extragenic Palindrome.



Ton-Hoang et al. 2012

- BIME set-up (Bacterial Interspesed Mosaic Element)

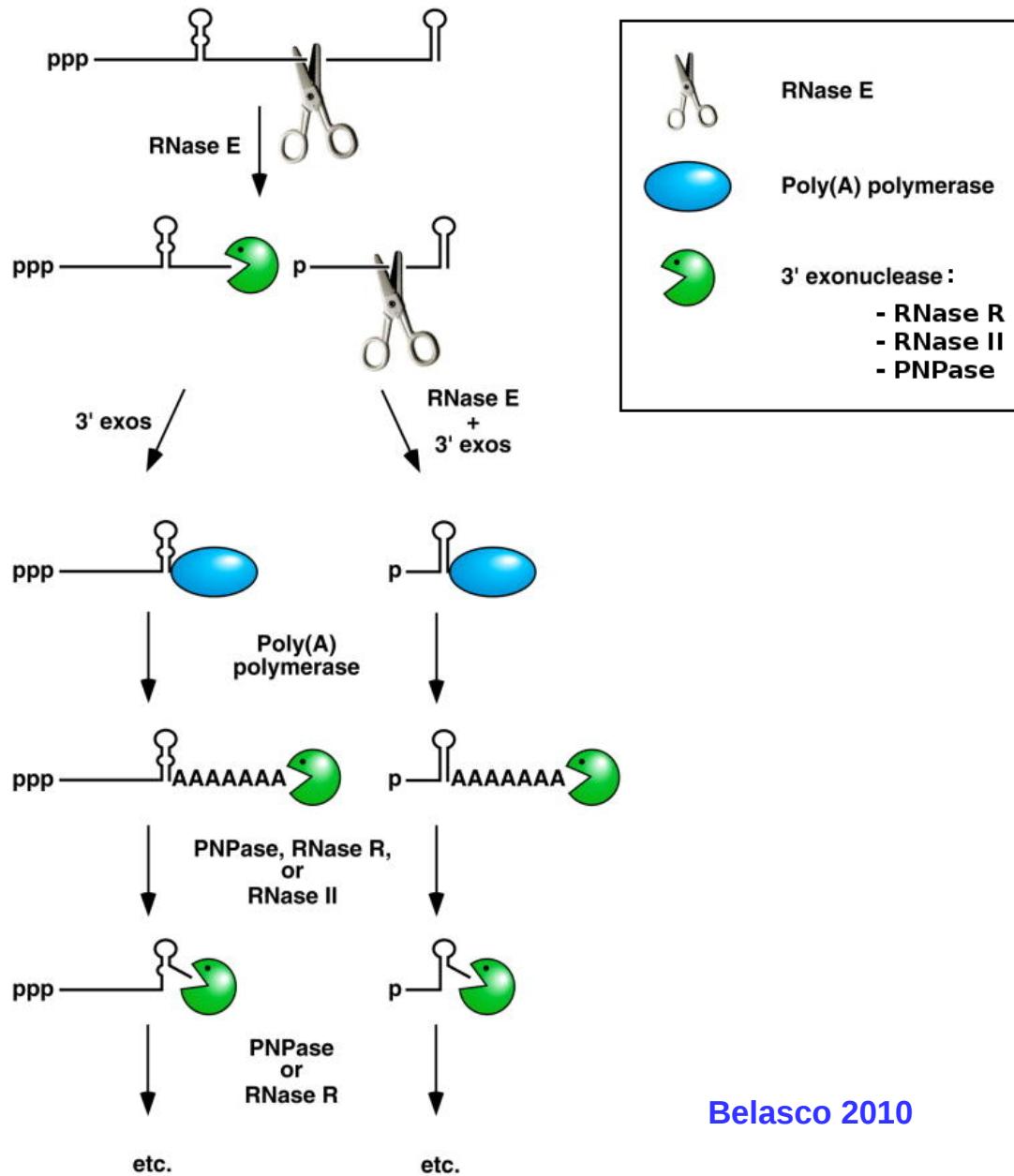
BIME properties?

- Structural roles:
 - ◆ Homologous recombination
 - ◆ Mobile DNA sequences insertion sites
- DNA level:
 - ◆ Proteic factors binding sites
 - ◆ Transcription initiation
 - ◆ Site specific recombination sites
- RNA level:
 - ◆ Traduction control
 - ◆ Transcription termination
 - ◆ 5' end transcript stabilisation

BIME properties?

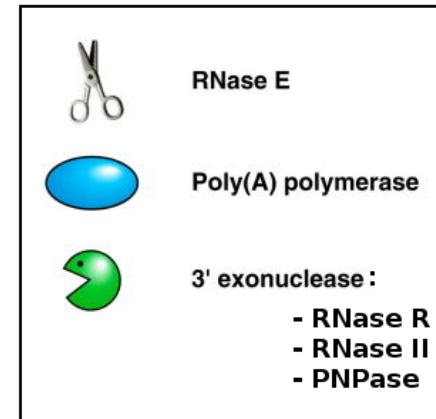
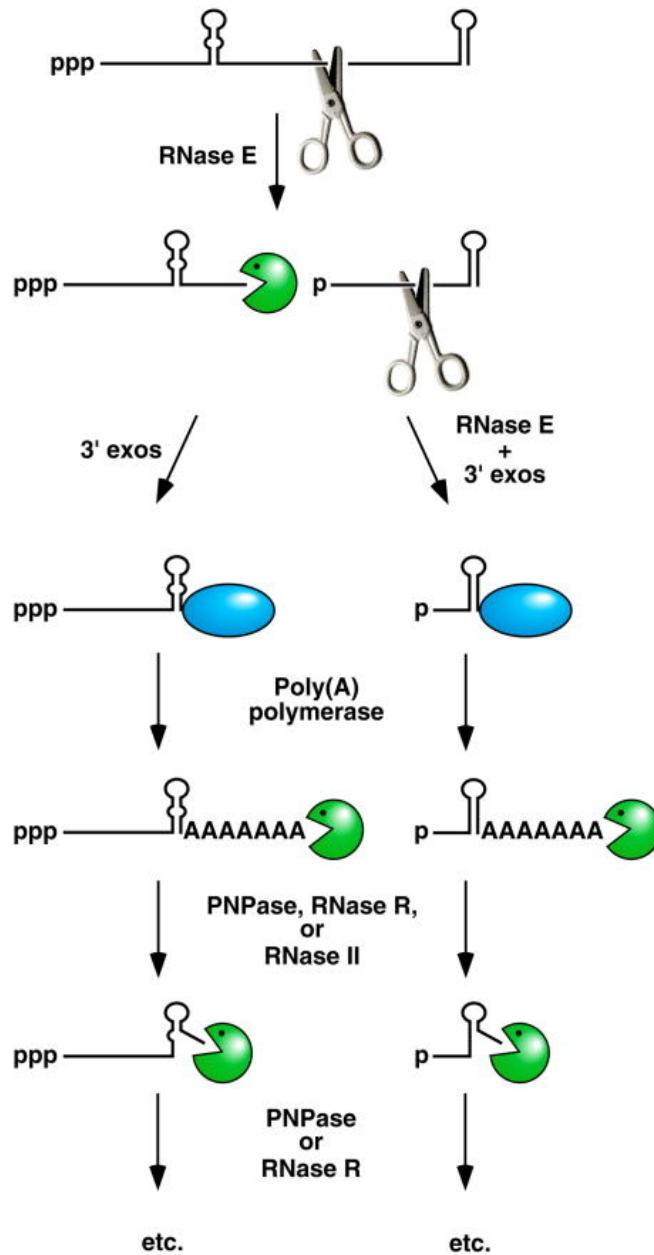
- Structural roles:
 - ◆ Homologous recombination
 - ◆ Mobile DNA sequences insertion sites
- DNA level:
 - ◆ Proteic factors binding sites
 - ◆ Transcription initiation
 - ◆ Site specific recombination sites
- RNA level:
 - ◆ Traduction control
 - ◆ **Transcription termination?**
 - ◆ **5' end transcript stabilisation?**

mRNA degradosome



Belasco 2010

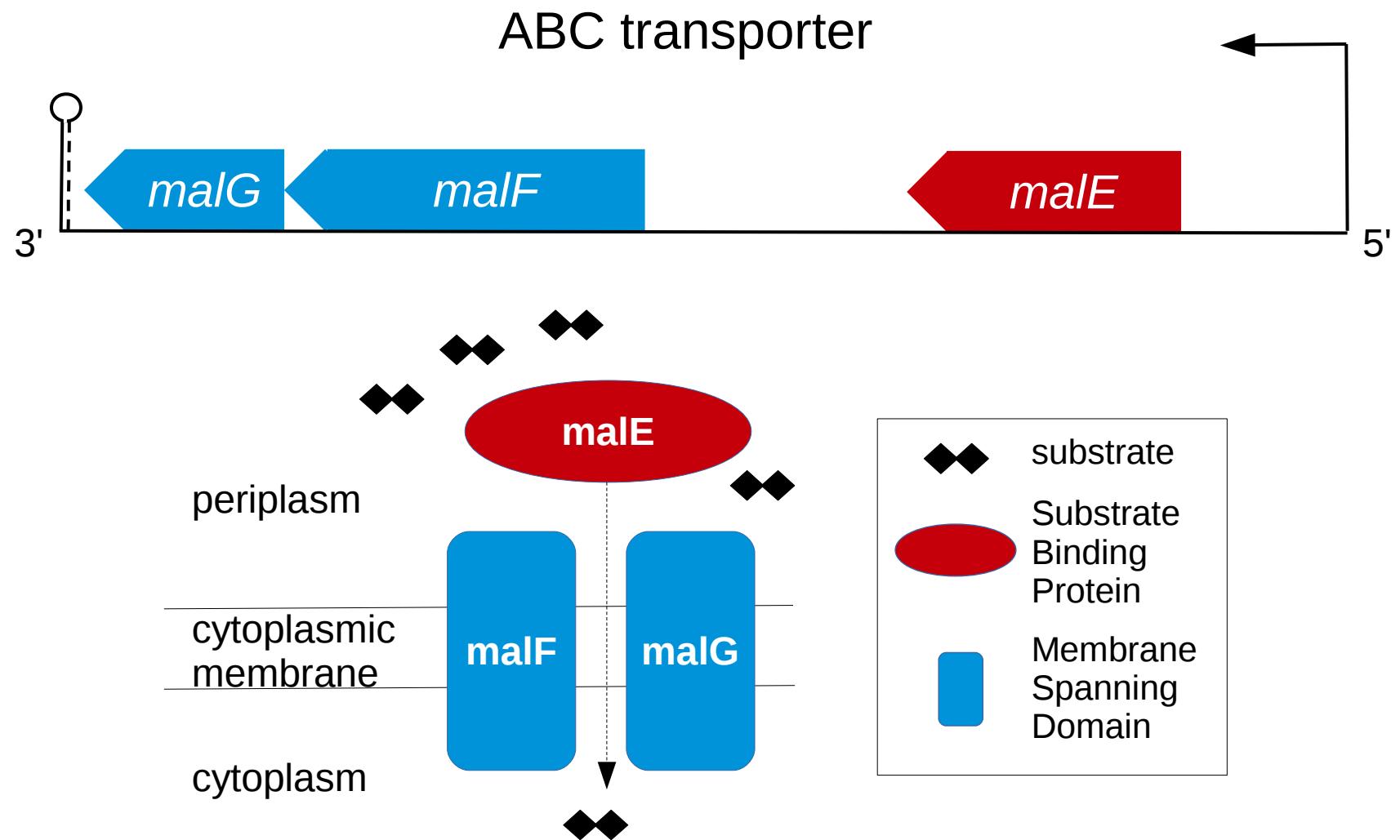
mRNA degradosome



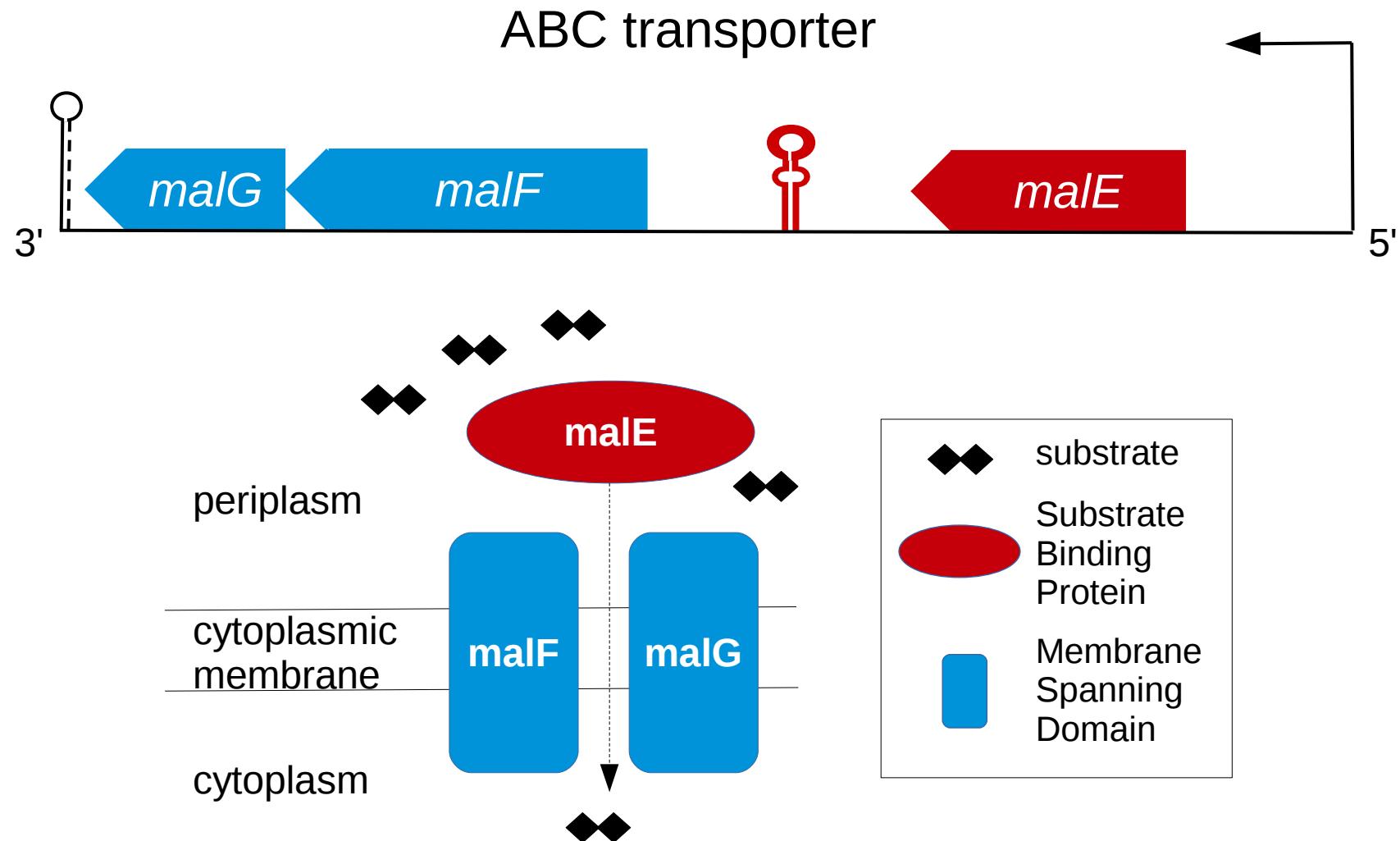
BIME secondary
structures?

Belasco 2010

malEFG operon

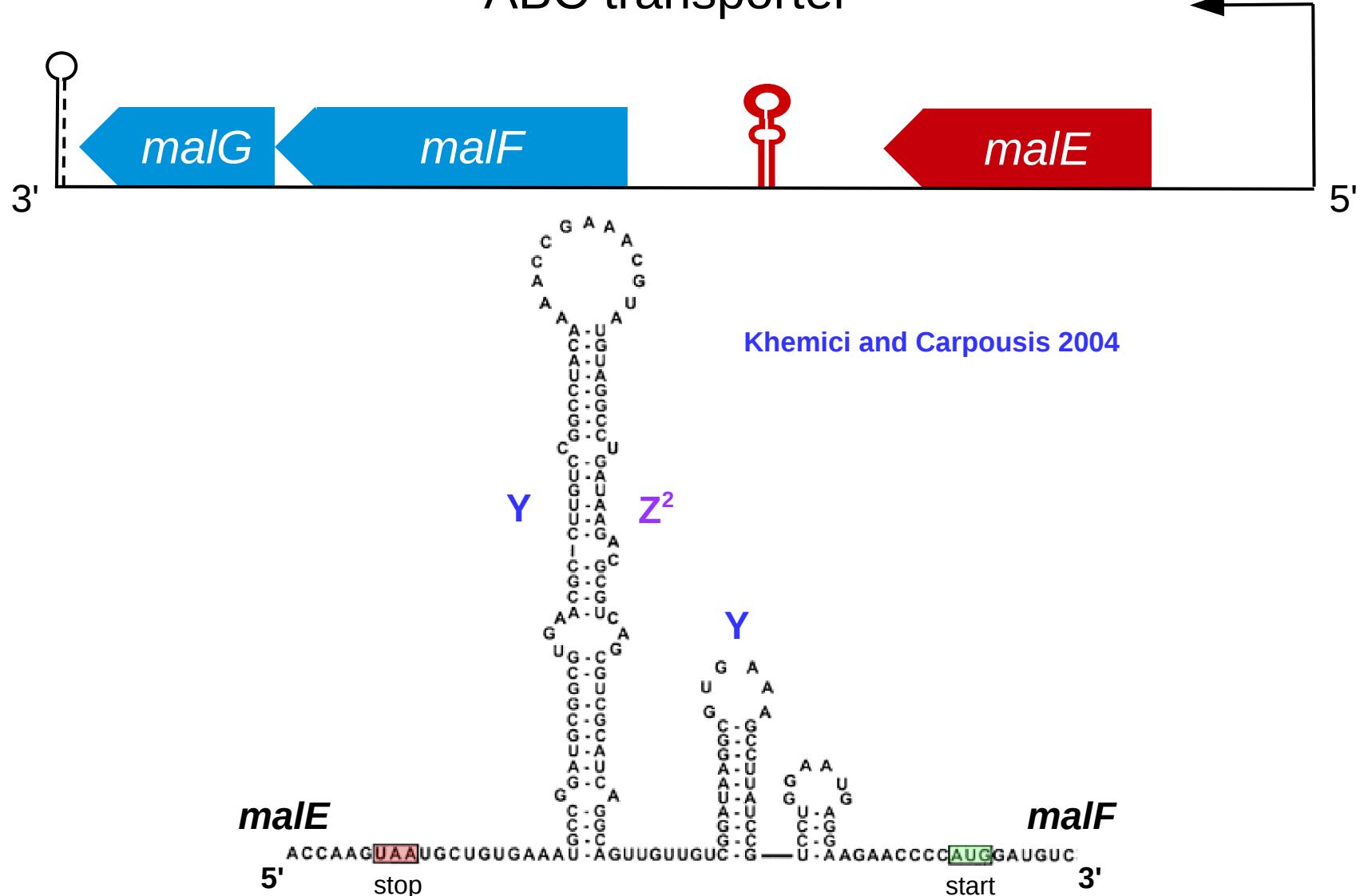


malEFG operon

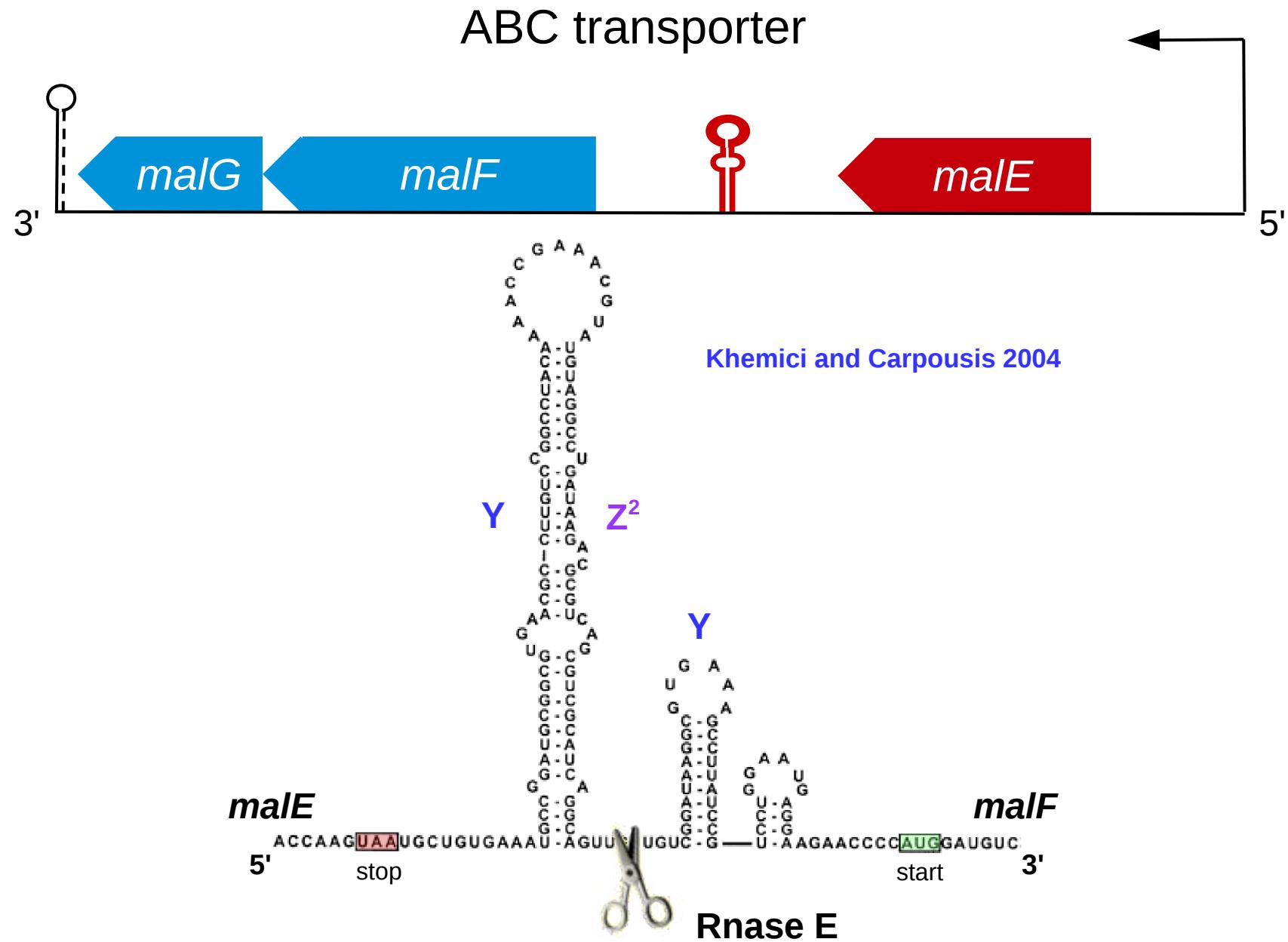


malEFG operon

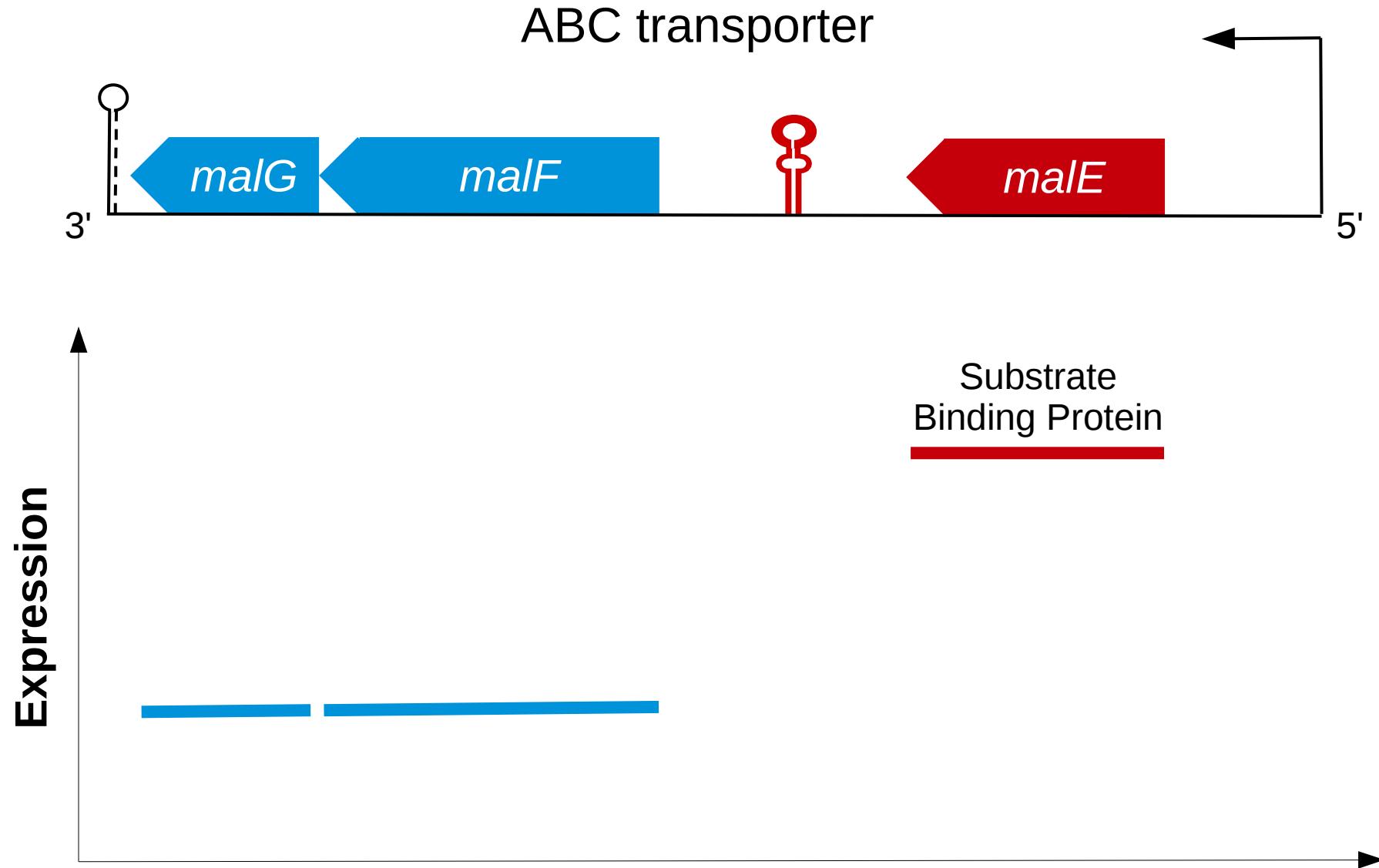
ABC transporter



malEFG operon



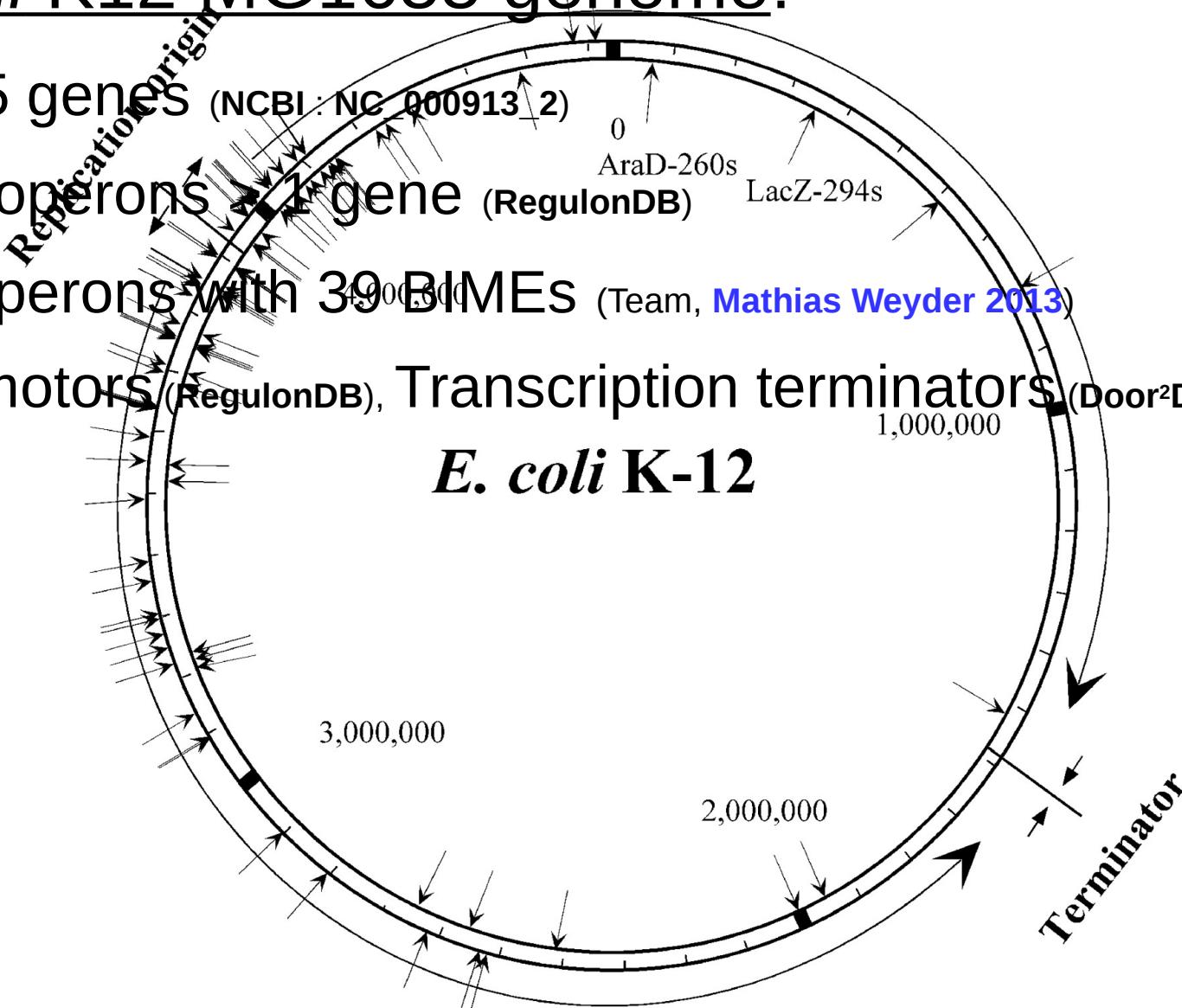
malEFG operon



BIME effect study in whole genome

- *E. coli* K12 MG1655 genome:

- 4605 genes (NCBI : NC_000913.2)
- 848 operons > 1 gene (RegulonDB)
- 36 operons with 39 BIMES (Team, Mathias Weyder 2013)
- Promotors (RegulonDB), Transcription terminators (Door²DB)

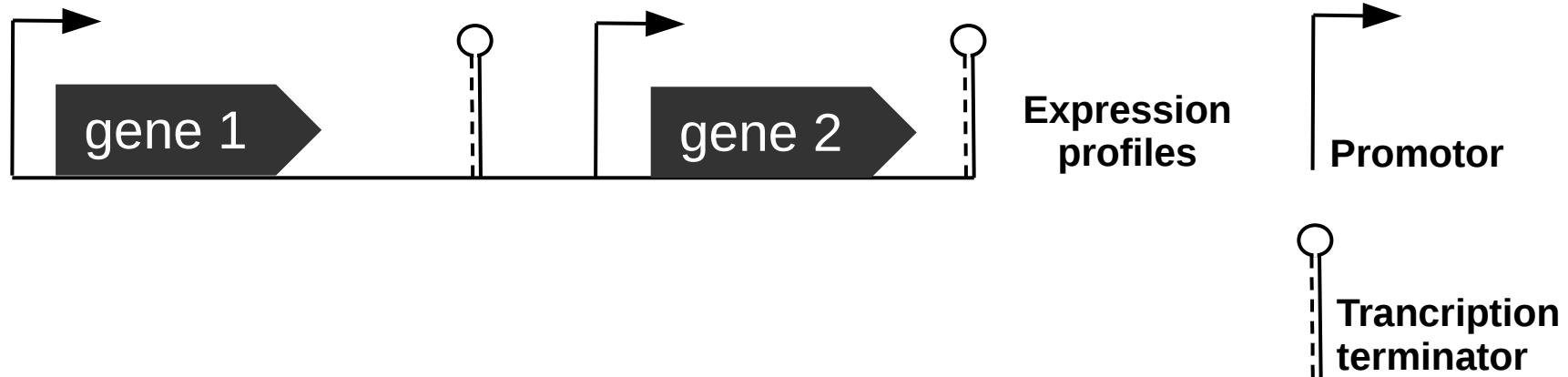


BIME effect study in whole genome

- *E. coli* K12 MG1655 genome:
 - ◆ 4605 genes (NCBI : NC_000913.2)
 - ◆ 848 operons > 1 gene (RegulonDB)
 - ◆ 36 operons with 39 BIMES (Team, Mathias Weyder 2013)
 - ◆ Promotors (RegulonDB), Transcription terminators (Door²DB)
- RNA-Seq:
 - ◆ Whole transcriptome study
 - ◆ Snapshot of transcription & degradation
 - ◆ 3 data sets from 2 experiments (GEO database)
 - ◆ 8, 3 and 3 replicates

Possible expression patterns

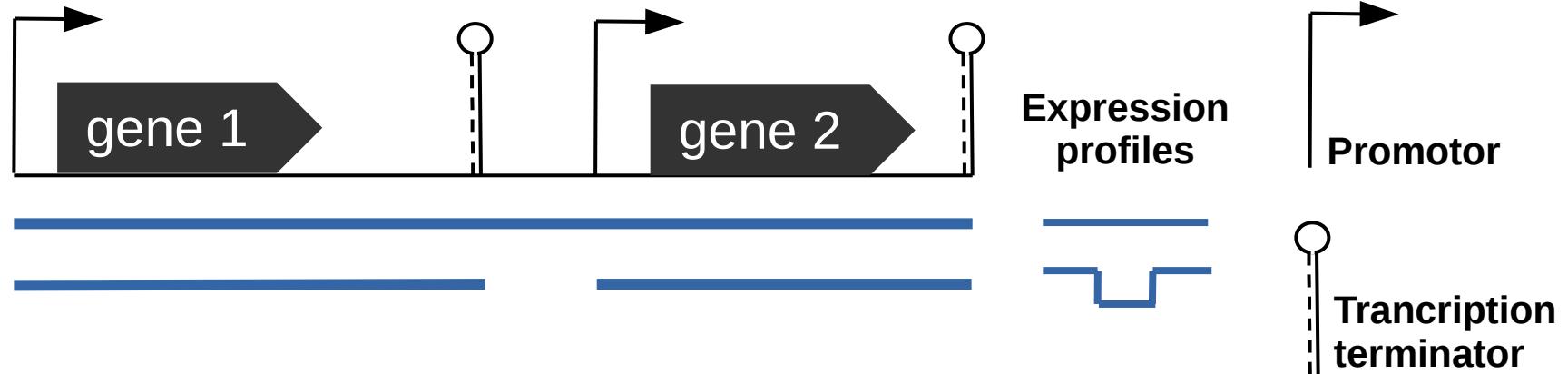
Without
BIME



Possible expression patterns

Without
BIME

Transcription
&
Degradation

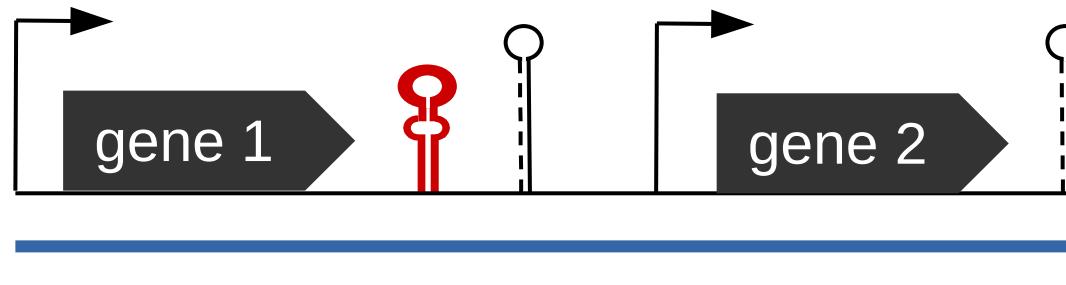


Possible expression patterns

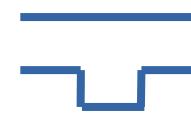
With
BIME

Transcription
&
Degradation

Transcription



Expression
profiles



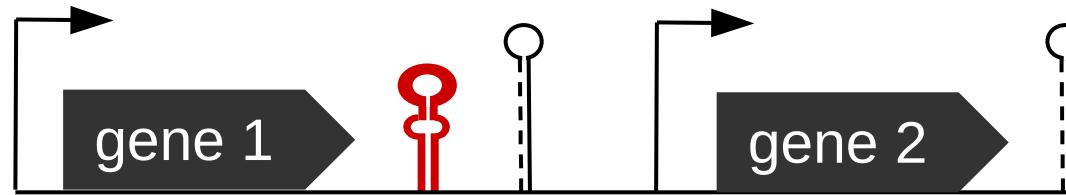
Possible expression patterns

With
BIME

Transcription
&
Degradation

Transcription

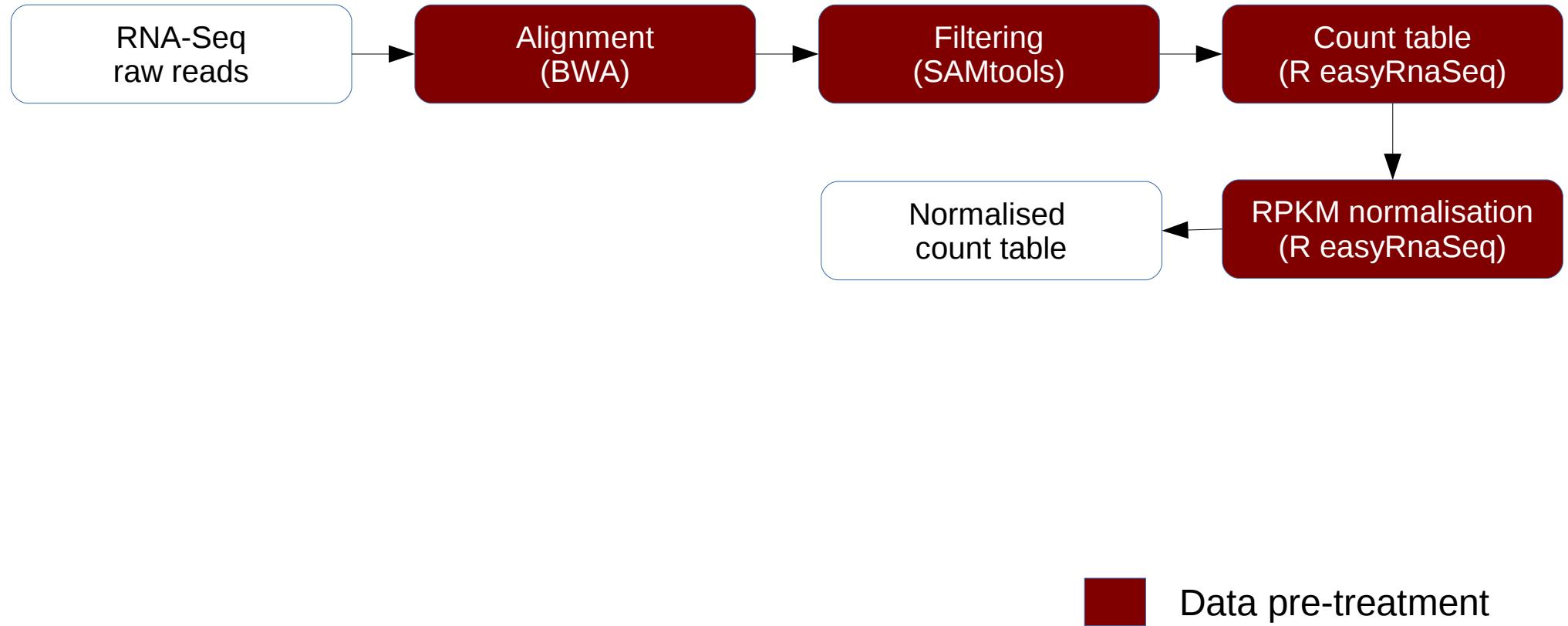
Degradation



Expression profiles



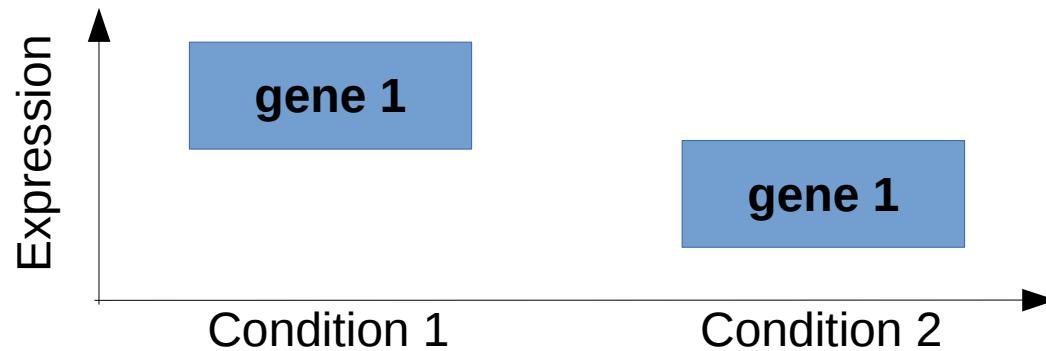
Analysis pipeline



Data pre-treatment

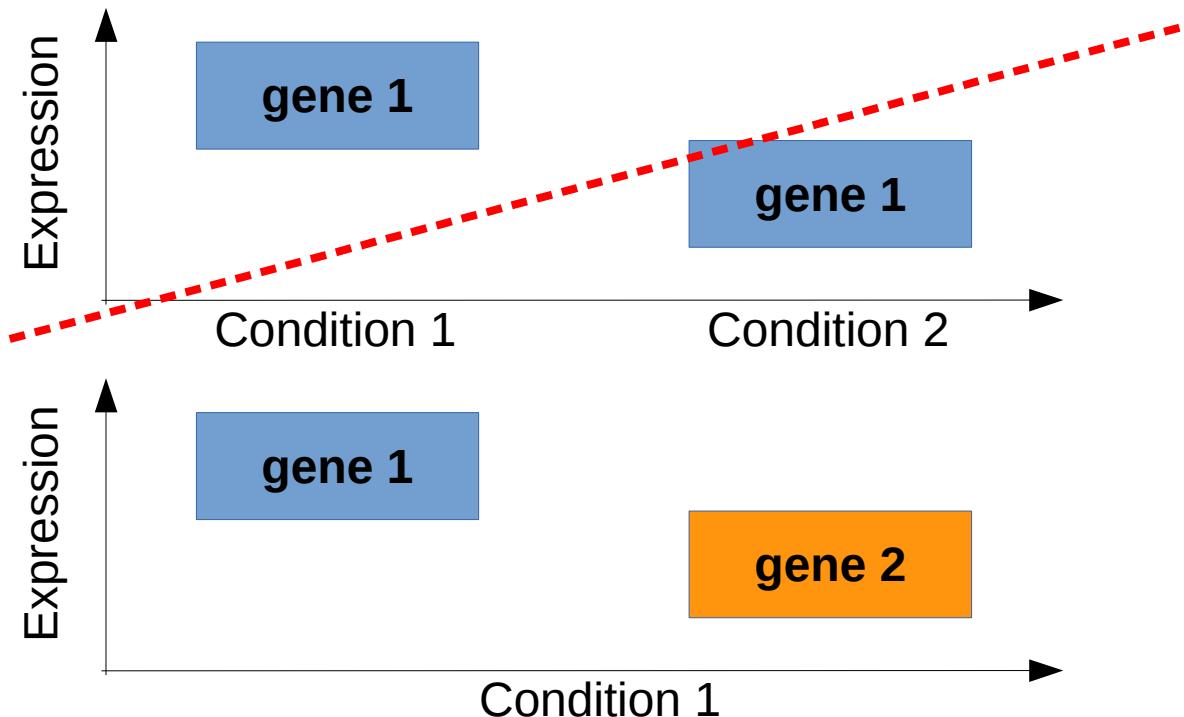
RNA-Seq normalisation

- Keep in mind:



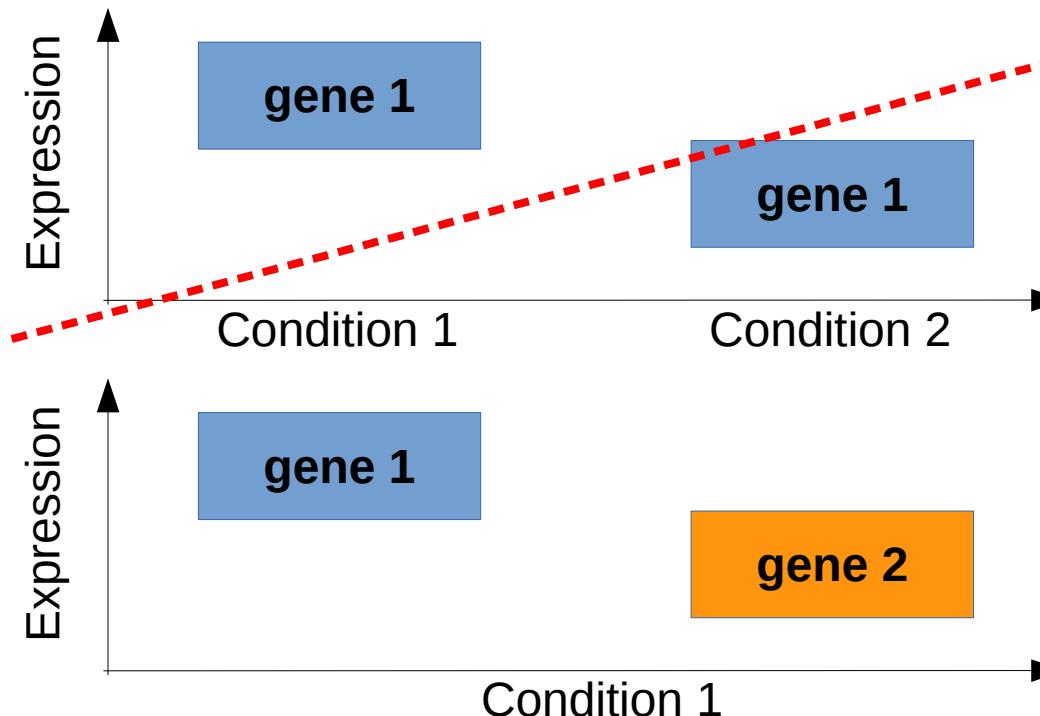
RNA-Seq normalisation

- Keep in mind:



RNA-Seq normalisation

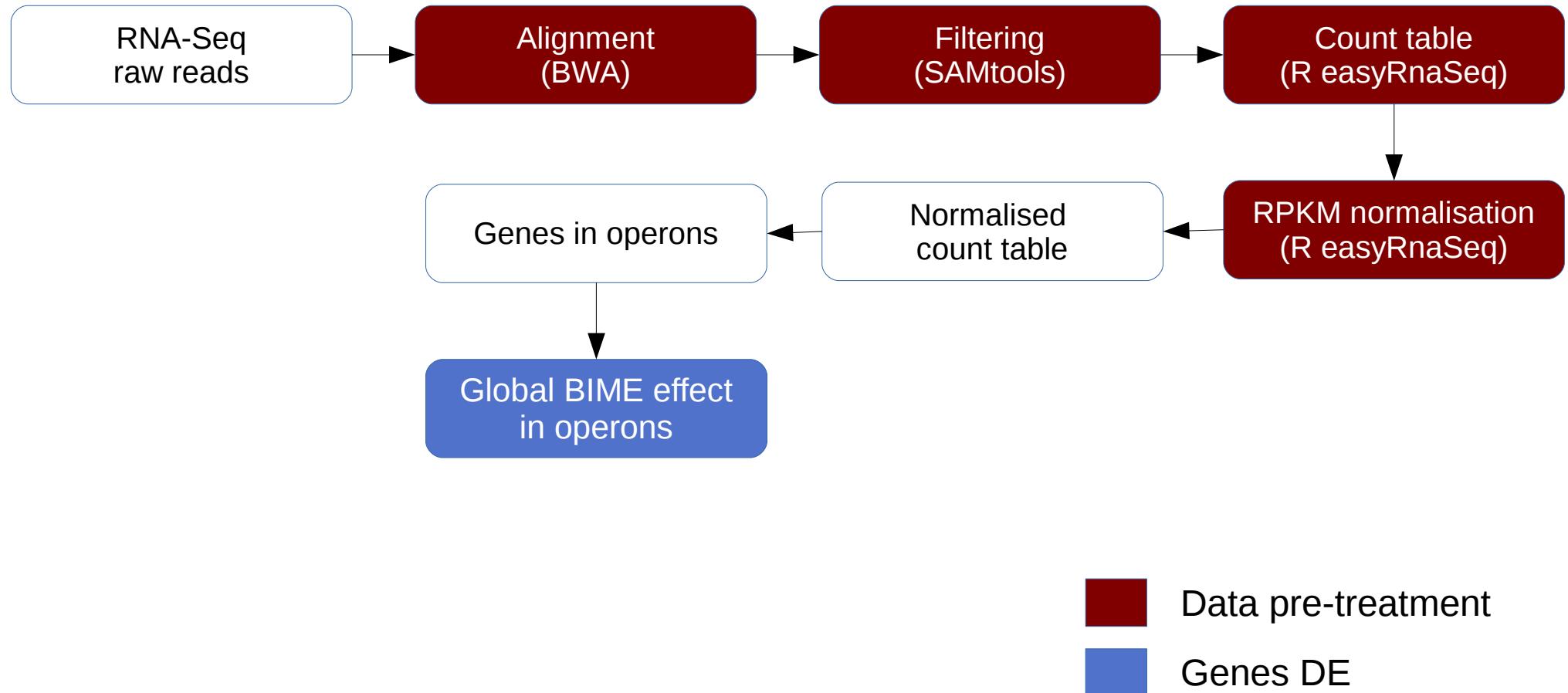
- Keep in mind:



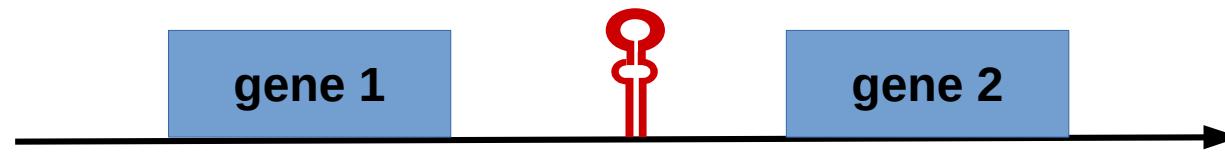
- RPKM normalisation:

- ◆ Transcript size
- ◆ Library size

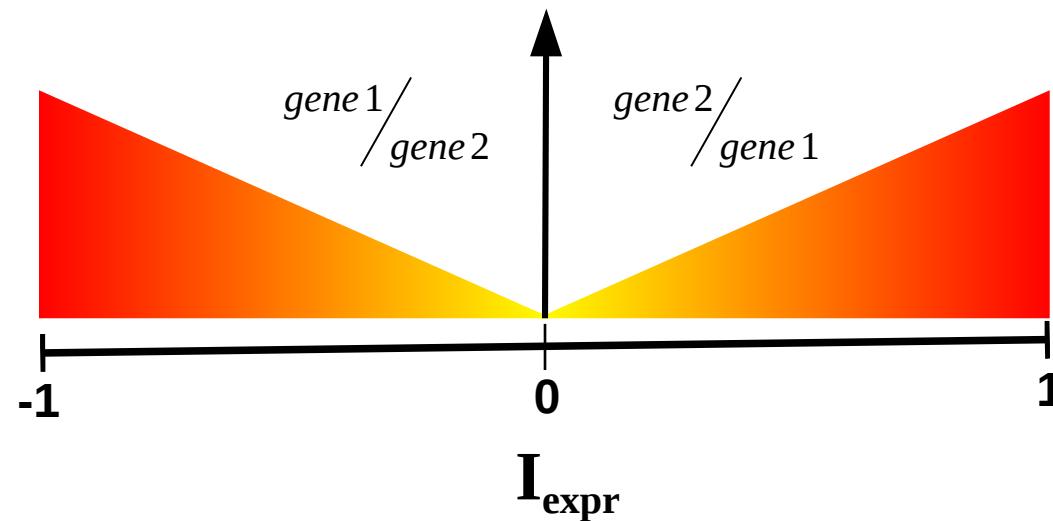
Analysis pipeline



Global BIME effect in operons

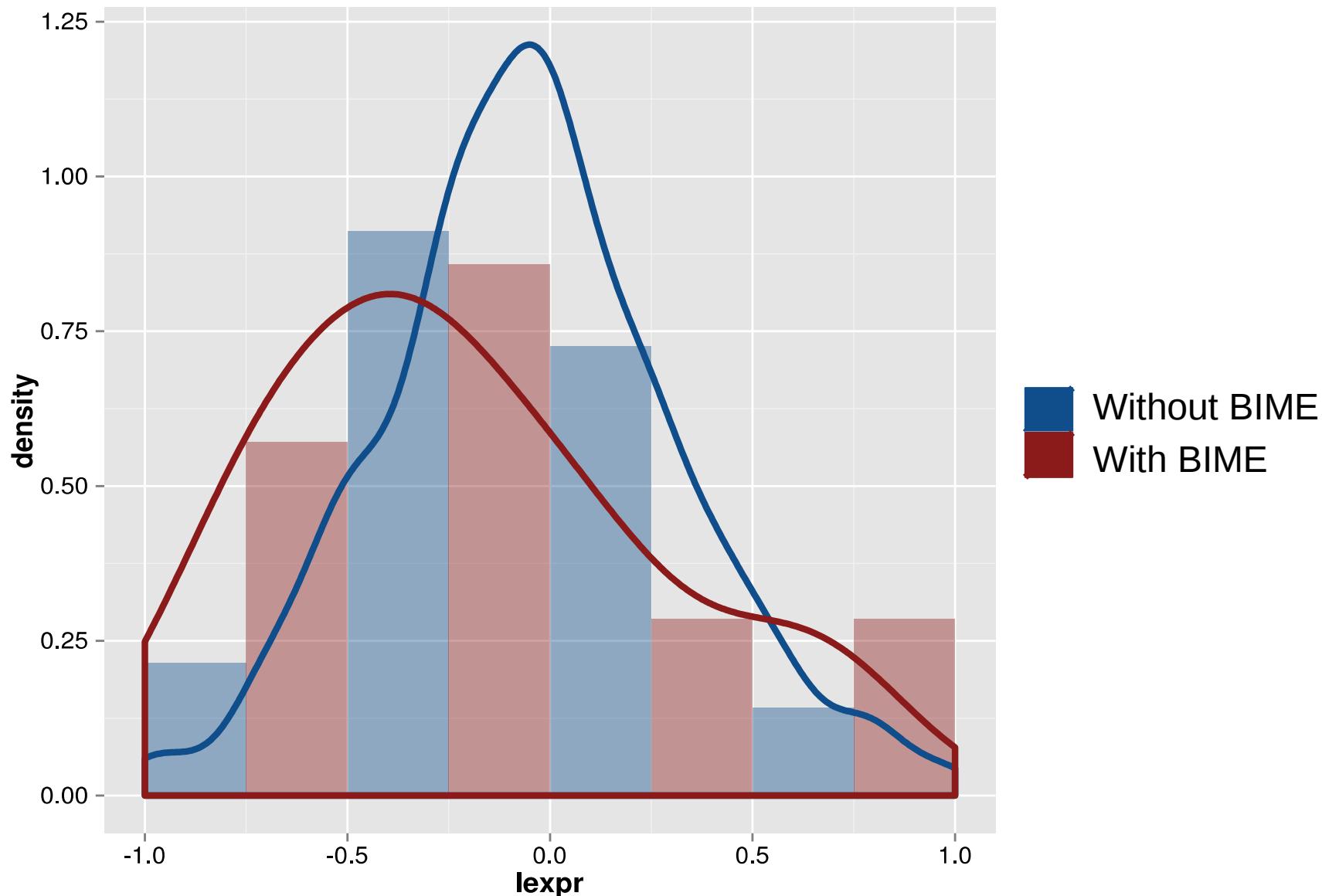


$$I_{expr} = \frac{RPKM_{gene2} - RPKM_{gene1}}{RPKM_{gene2} + RPKM_{gene1}}$$

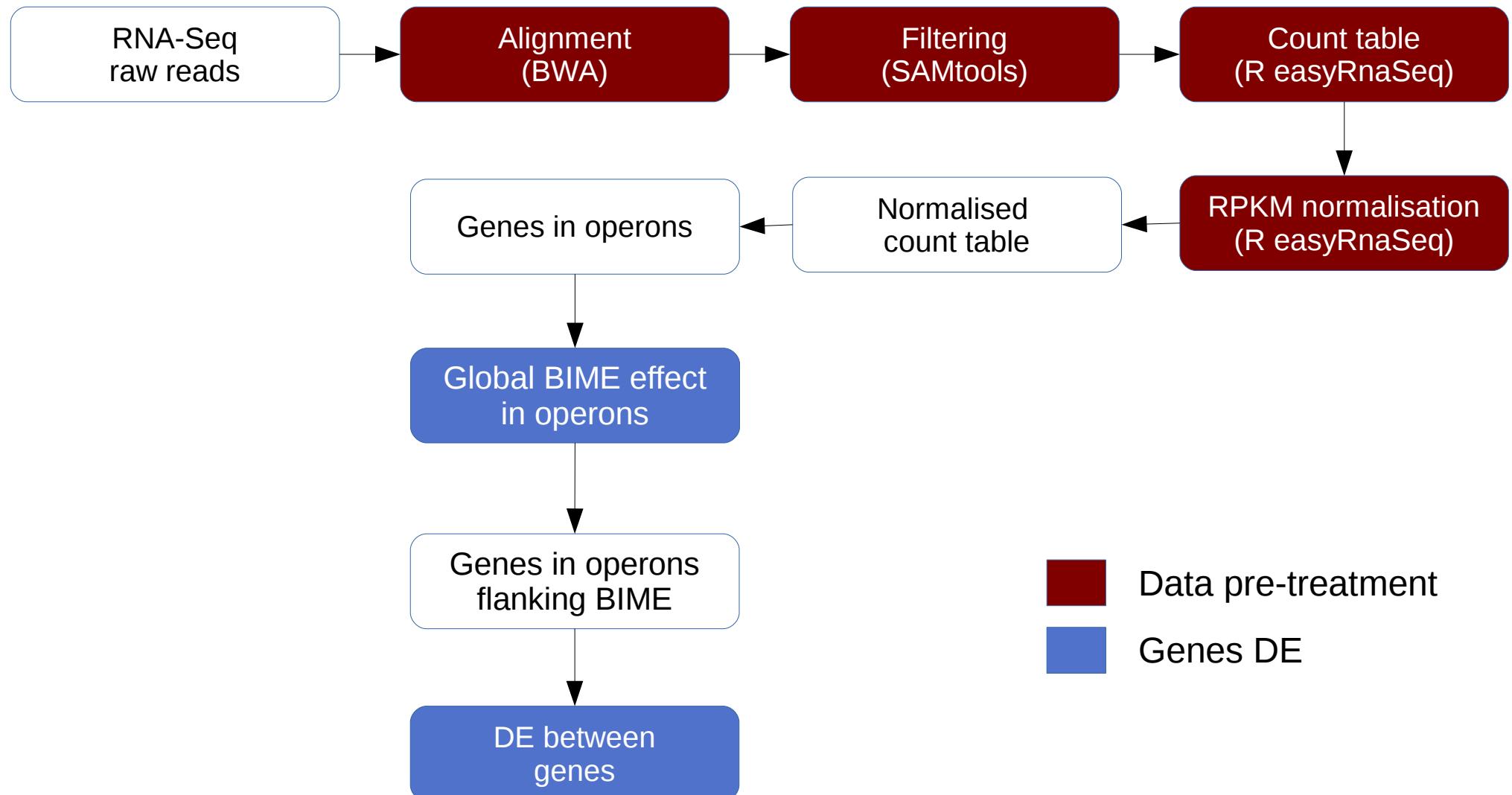


Global BIME effect in operons

Distributions of I_{expr} without promoters and terminators



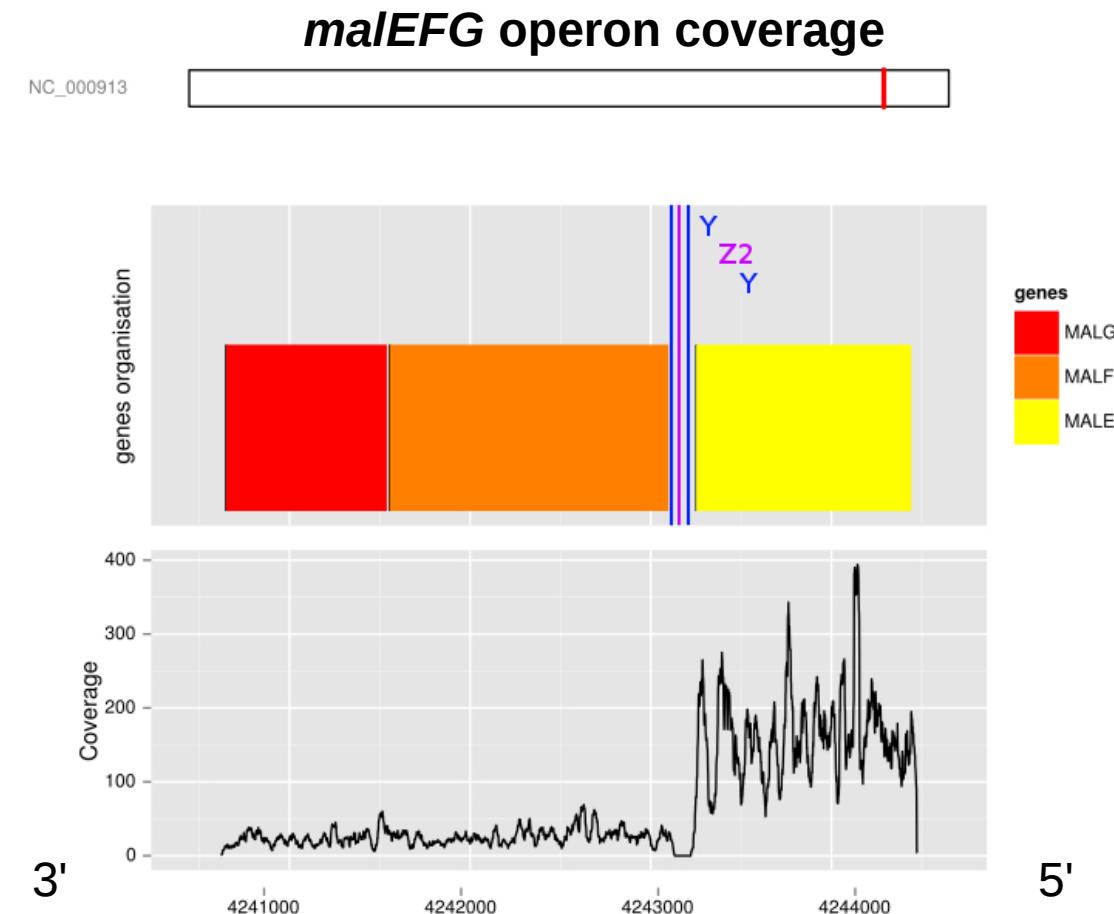
Analysis pipeline



Data pre-treatment
 Genes DE

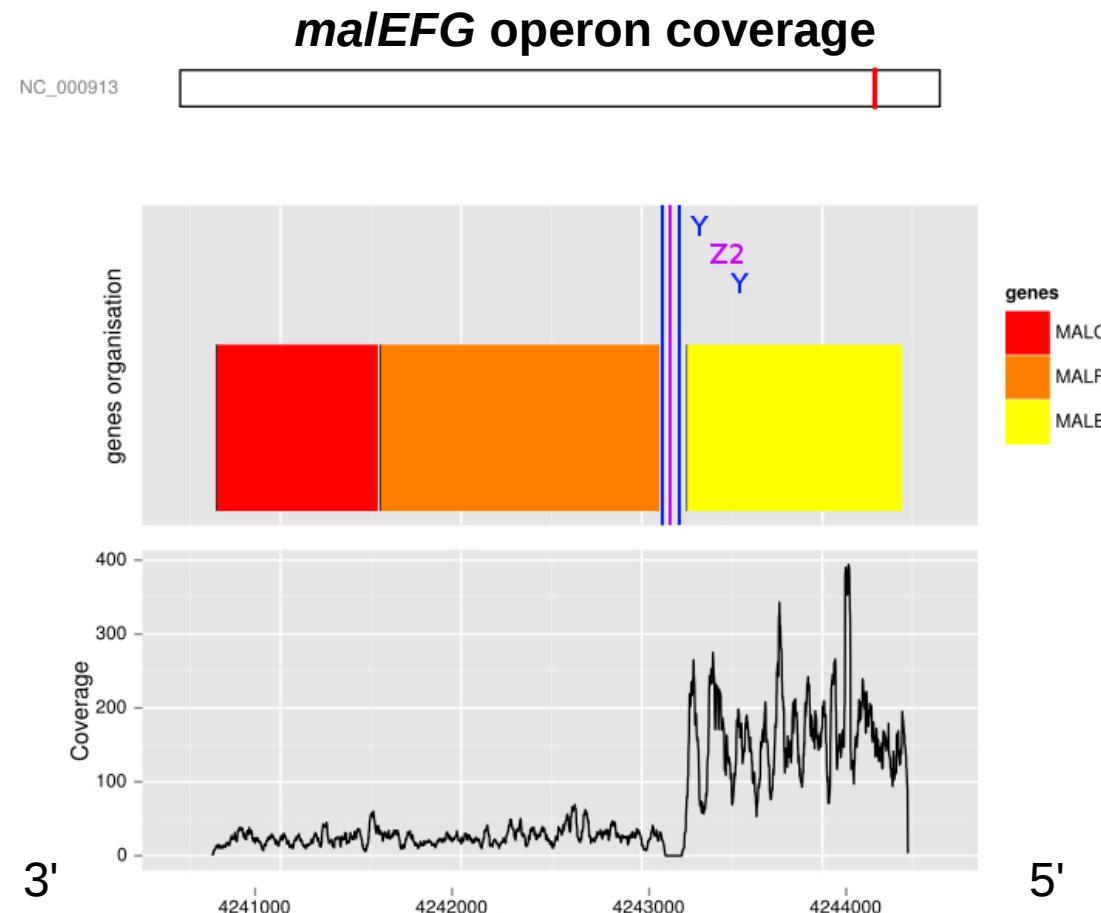
Genes flanking BIME Differential Expression?

- Wilcoxon rank test for DE (p-value threshold : 0.01)



Genes flanking BIME Differential Expression?

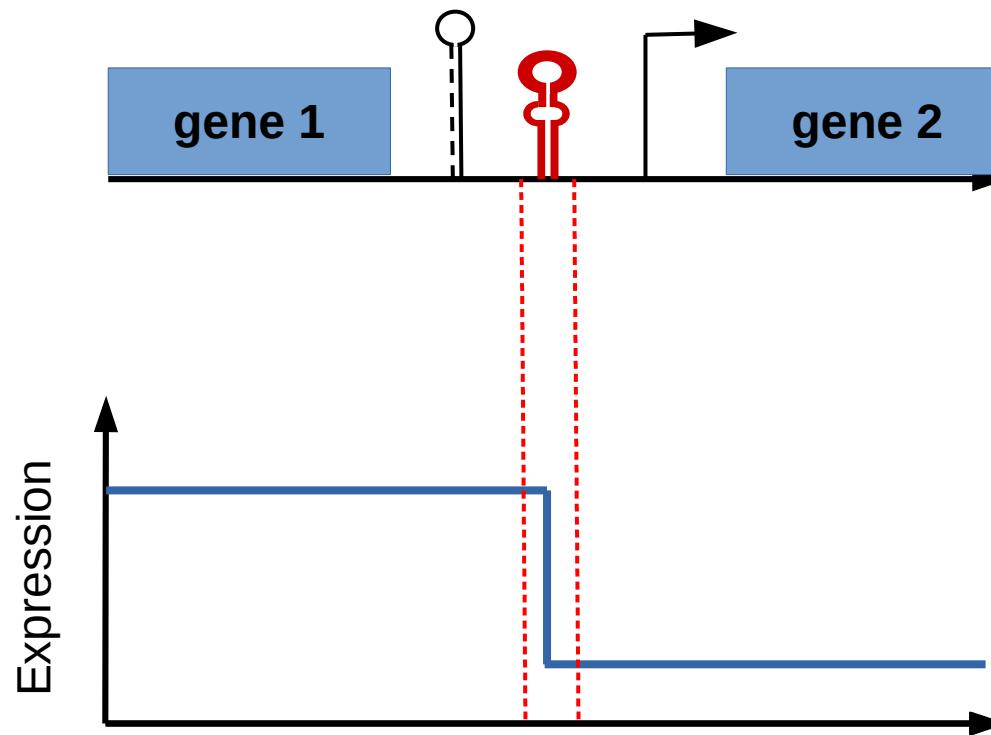
- Wilcoxon rank test for DE (p-value threshold : 0.01)



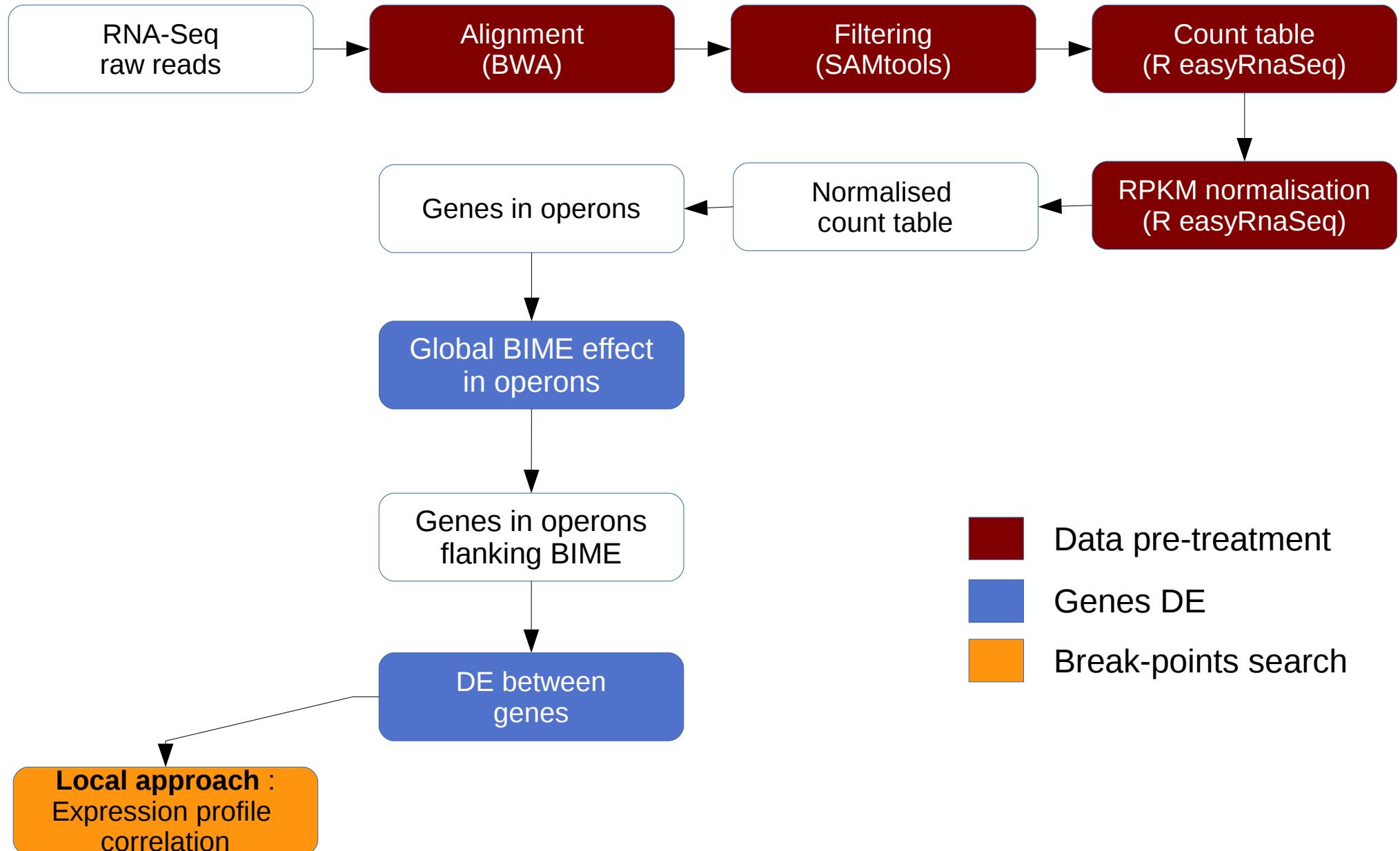
- 25 DE on 39 pairs of gene flanking a BIME (3 data sets)

Genes difference of expression linked to the BIME?

- Coverage break-points on BIME?

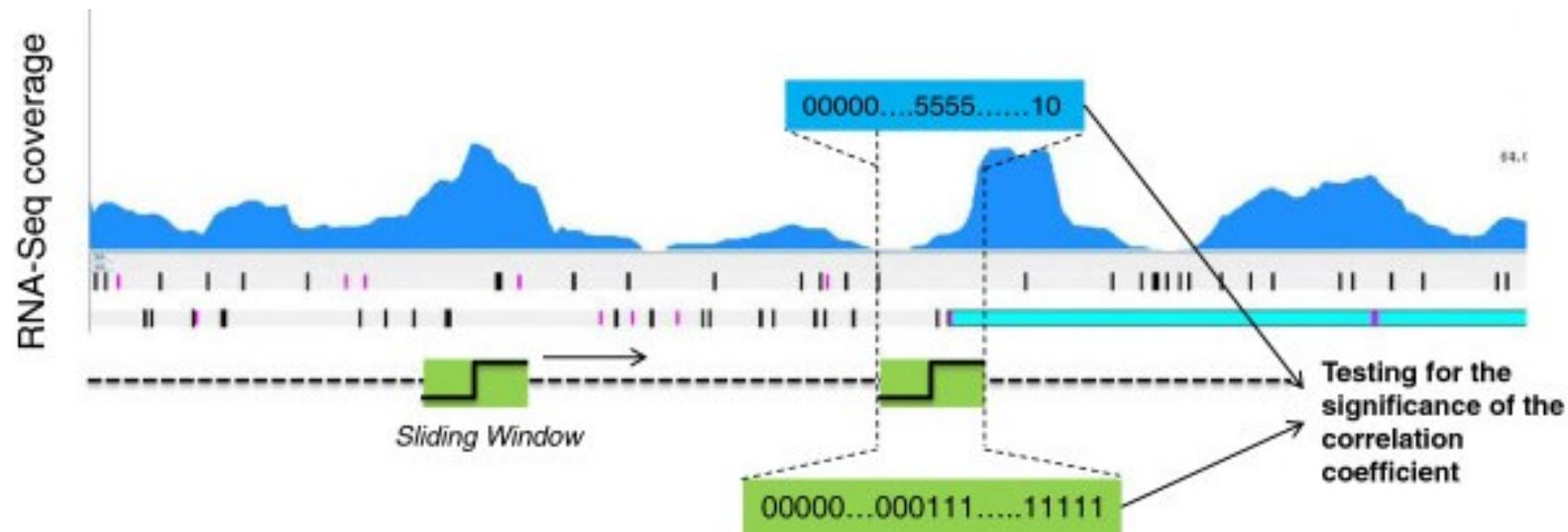


Analysis pipeline



Adaptation of operon detection method (Local approach)

Expression profile correlation

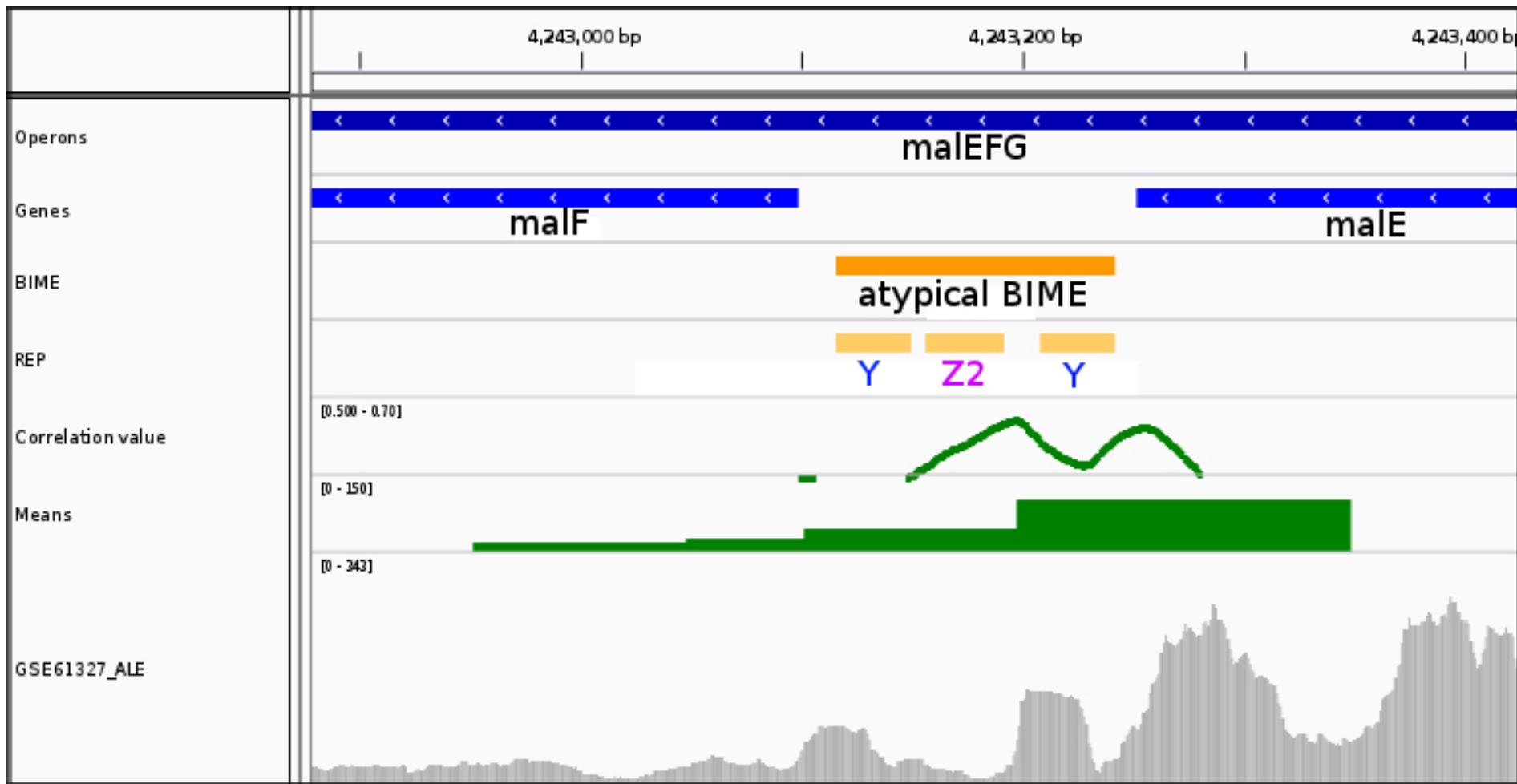


Fortino et al. 2014

- Window size: 300 bp
- 2 fold change of expression
- Correlation > 0.5 & p-value threshold < 10^{-7}

Adaptation of operon detection method (Local approach)

Expression profile correlation on *malE-malF* BIME



Local approach results

- 18 break-points on BIMEs for 25 DE

Terminator in intergenic region	Overlapping a BIME	Not overlapping a BIME	Absence
Coverage break-points on BIME	4/18	3/18	13/18

- Weakness:
 - RNA-Seq coverage not uniform
 - False positive detection
- Other approach?



Local approach results

- 18 break-points on BIMEs for 25 DE

Terminator in intergenic region	Overlapping a BIME	Not overlapping a BIME	Absence
Coverage break-points on BIME	4/18	3/18	13/18

Local approach results

- 18 break-points on BIMEs for 25 DE

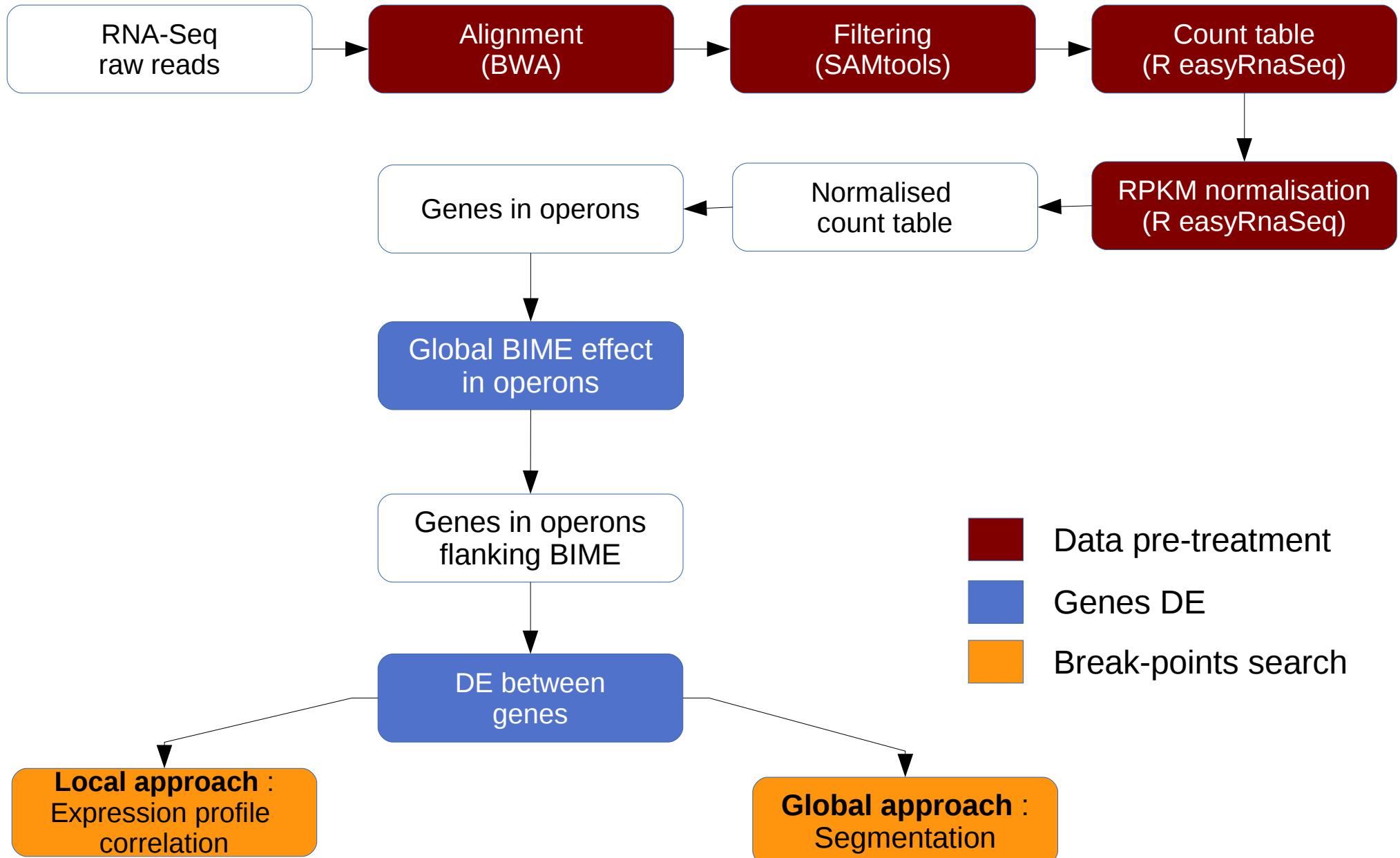
Terminator in intergenic region	Overlapping a BIME	Not overlapping a BIME	Absence
Coverage break-points on BIME	4/18	3/18	13/18

- Weakness:



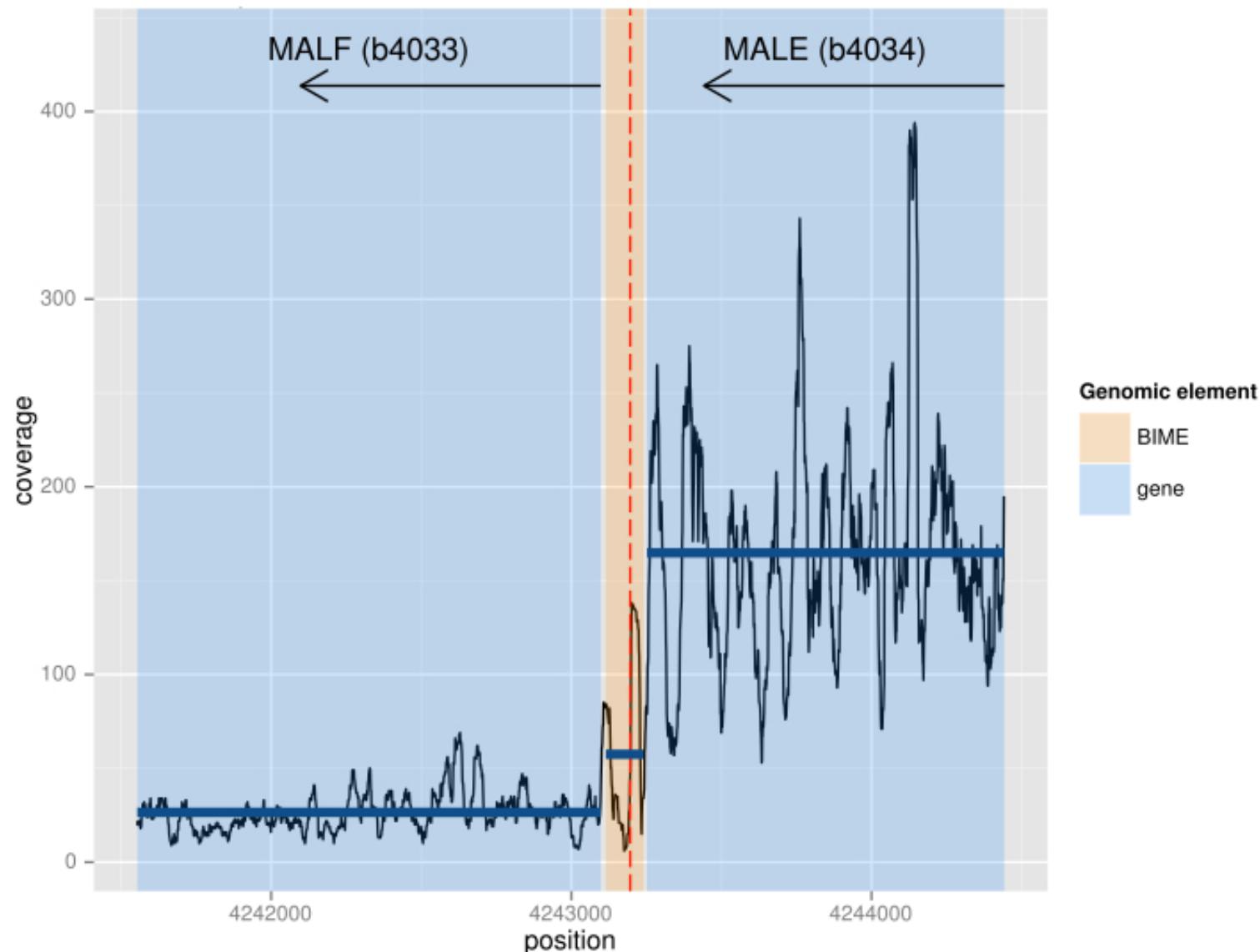
- RNA-Seq coverage not uniform
- False positive detection

Analysis pipeline



Segmentation (Global approach)

Segmentation on *malE-malF* BIME



Global approach results

- 17 break-points on BIMEs for 25 DE

Terminator in intergenic region	Overlapping a BIME	Not overlapping a BIME	Absence
Coverage break-points on BIME	3/17	3/17	12/17

- Weakness:
 - Sensitive to number of replicates
 - Fewer replicates → less detection

Question

Does BIME are involved in the mRNA metabolism
in operons for *E. coli* genome :

- as transcription terminator ?
- as stabiliser of 5' end of transcript ?

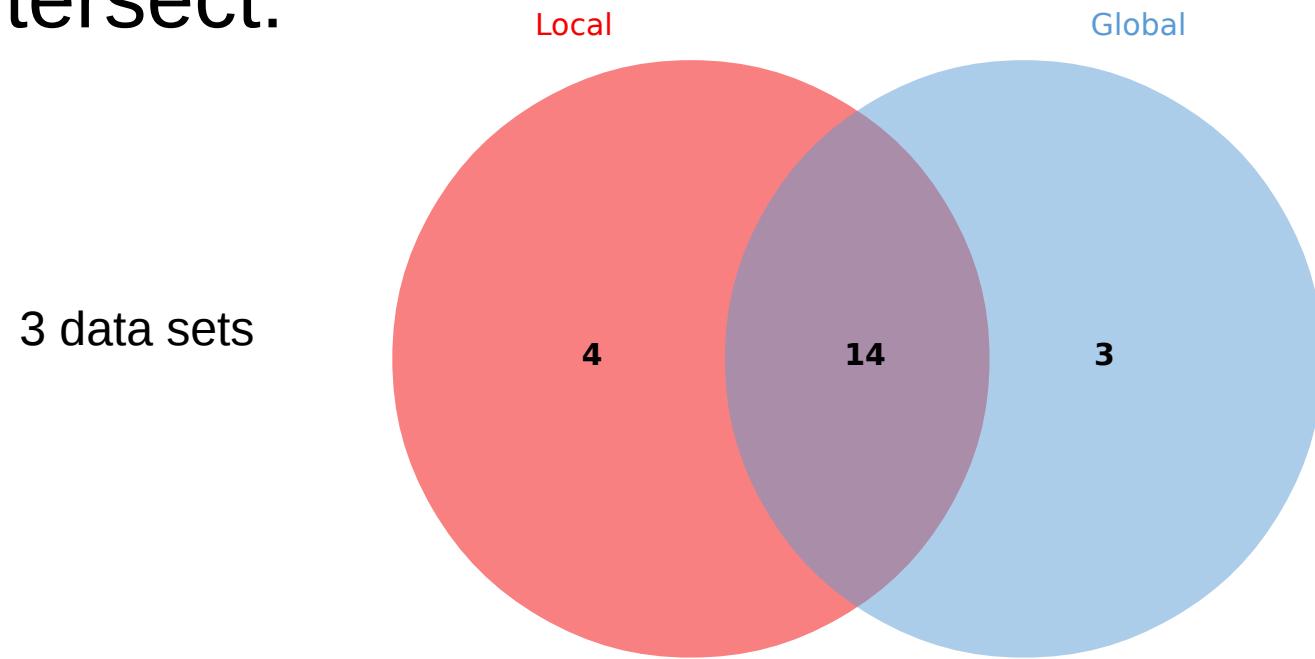
Break-point search results synthesis

- Methods intersect:



Break-point search results synthesis

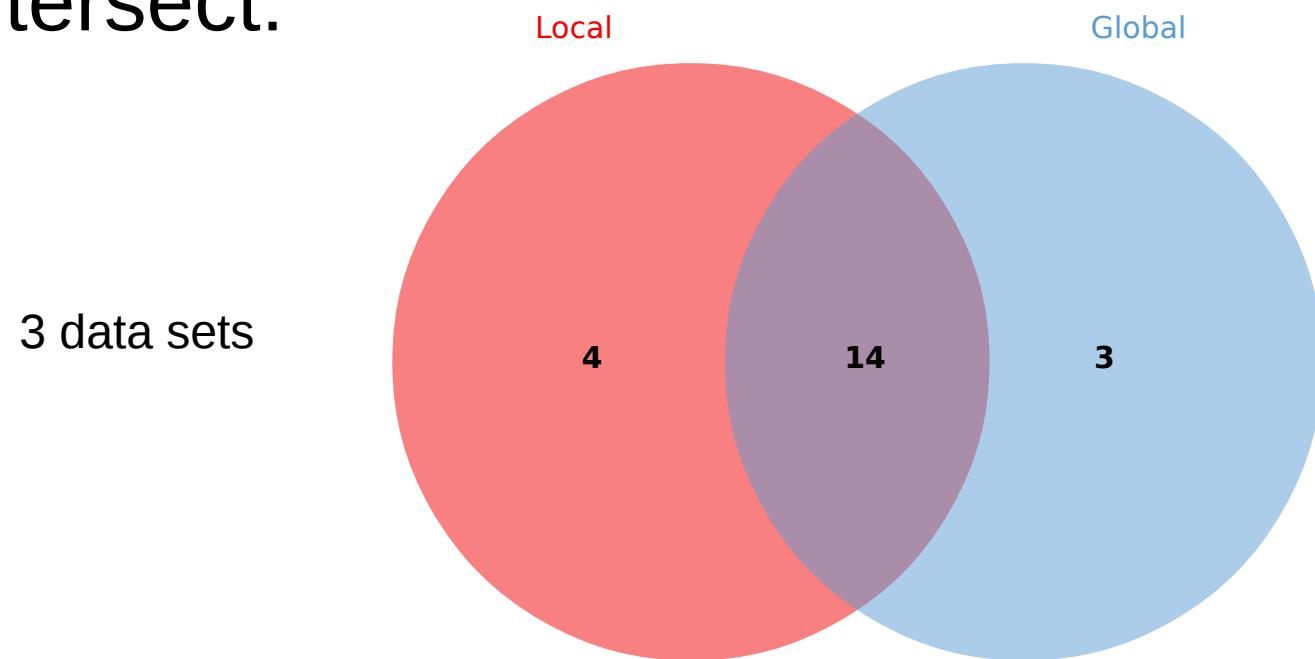
- Methods intersect:



Terminator in intergenic region	Overlapping a BIME	Not overlapping a BIME	Absence
Coverage break-points on BIME	4/21	4/21	15/21

Break-point search results synthesis

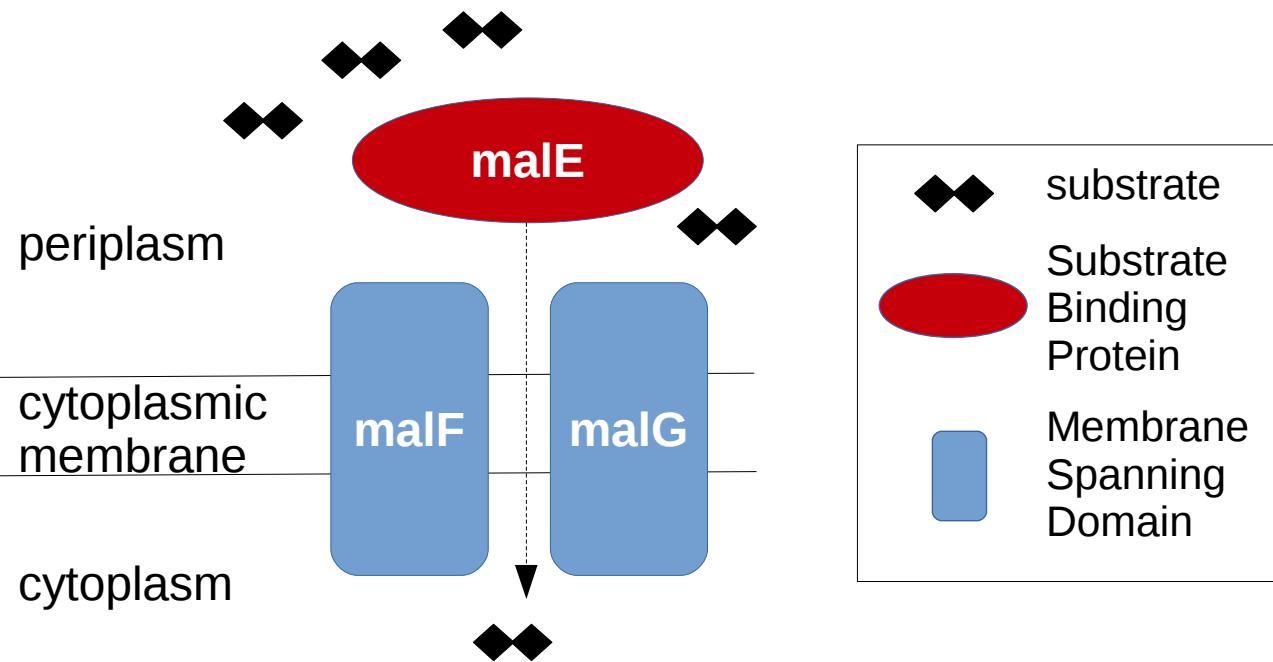
- Methods intersect:



Terminator in intergenic region	Overlapping a BIME	Not overlapping a BIME	Absence
Coverage break-points on BIME	4/21	4/21	15/21

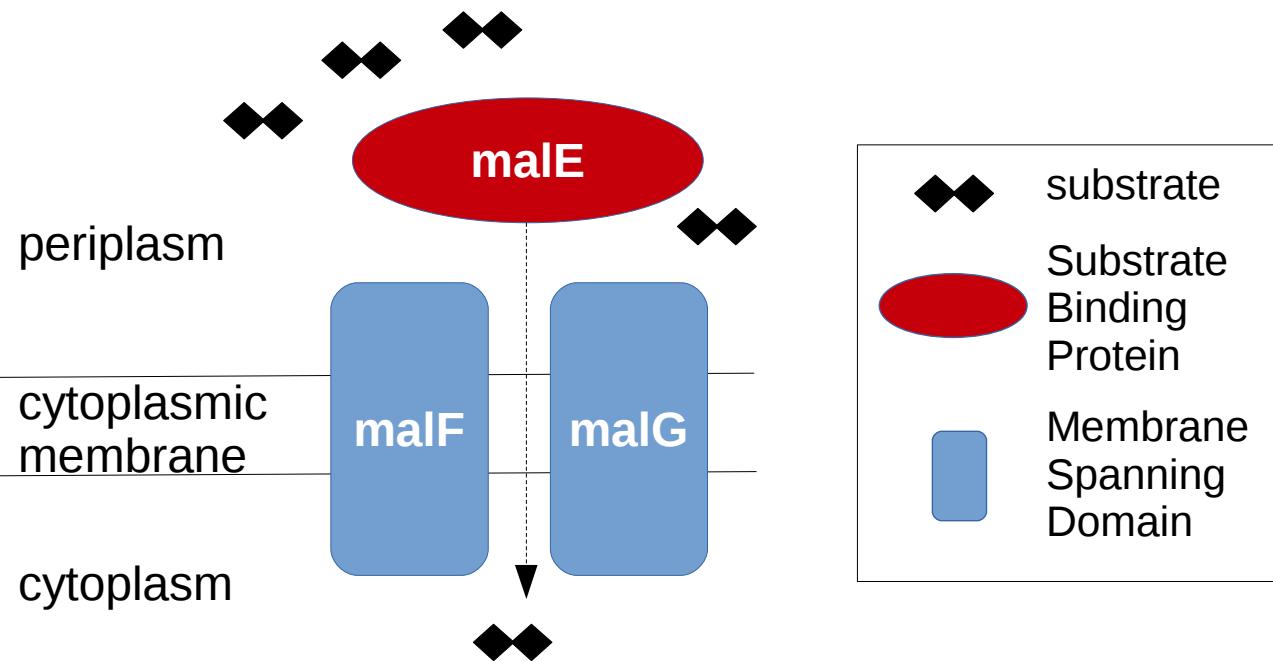
Biological synthesis

- *malEFG* model validated
- 7 ABC transporter with 8 BIMEs on 21 detections



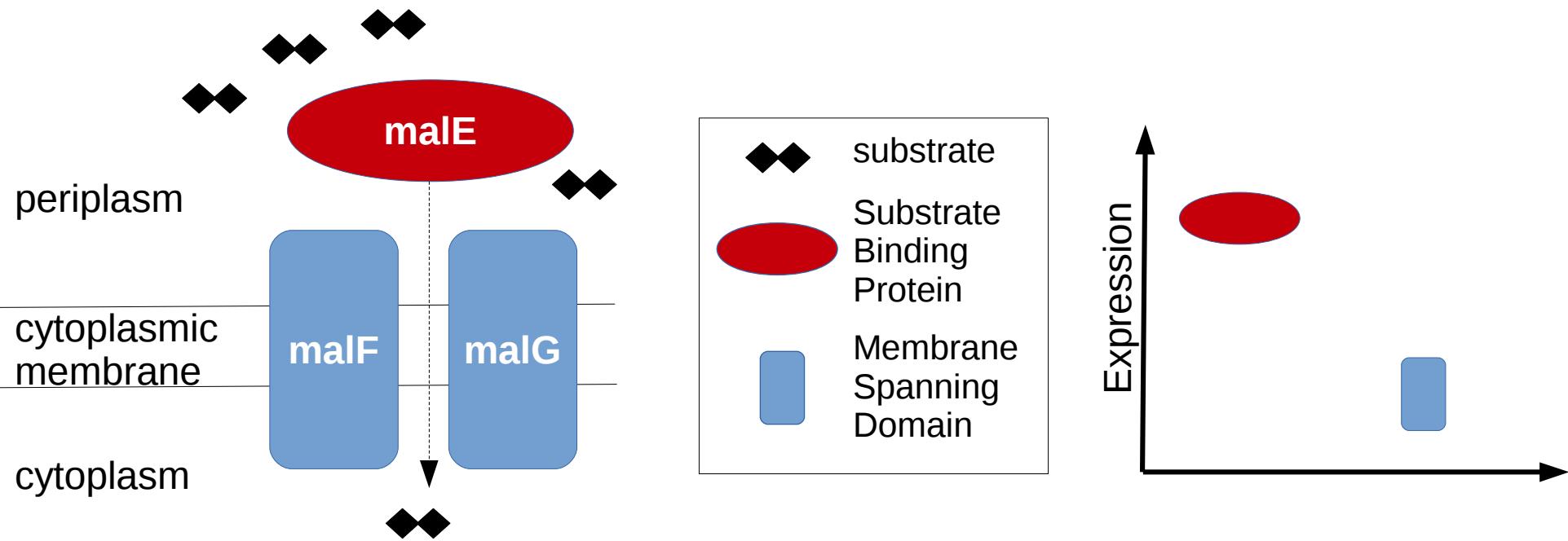
Biological synthesis

- *malEFG* model validated
- 7 ABC transporter with 8 BIMEs on 21 detections
- 6 without terminators



Biological synthesis

- *malEFG* model validated
- 7 ABC transporter with 8 BIMEs on 21 detections
- 6 without terminators → Substrate Binding Proteins



Conclusions

- RNA-Seq data fine for those studies
- BIME effect on upstream gene
- Segmentation method suited (Number of replicates)
- Easy to use with other data sets
- Annotation quality?
- Transcription terminator and/or 5' end transcript stabiliser?

Perspectives

- RNA-Seq experiments designs
 - ◆ Test transcription termination function:
WT & alteration of degradosome genes
 - ◆ Test transcript 5' end stabiliser function:
Transcription blocking & time sampling for monitoring of the transcripts degradation
- Benchmark with segmentation
 - ◆ reduce the number of replicates (same experiment)
- Other species studies

Acknowledgments

LMGM: Genomics of integrated systems

Gwennaele Fichant

Yves Quentin

Roland Barriot

Petra Langendijk-Genevaux

Mathias Weyder

Arnaud Frèche

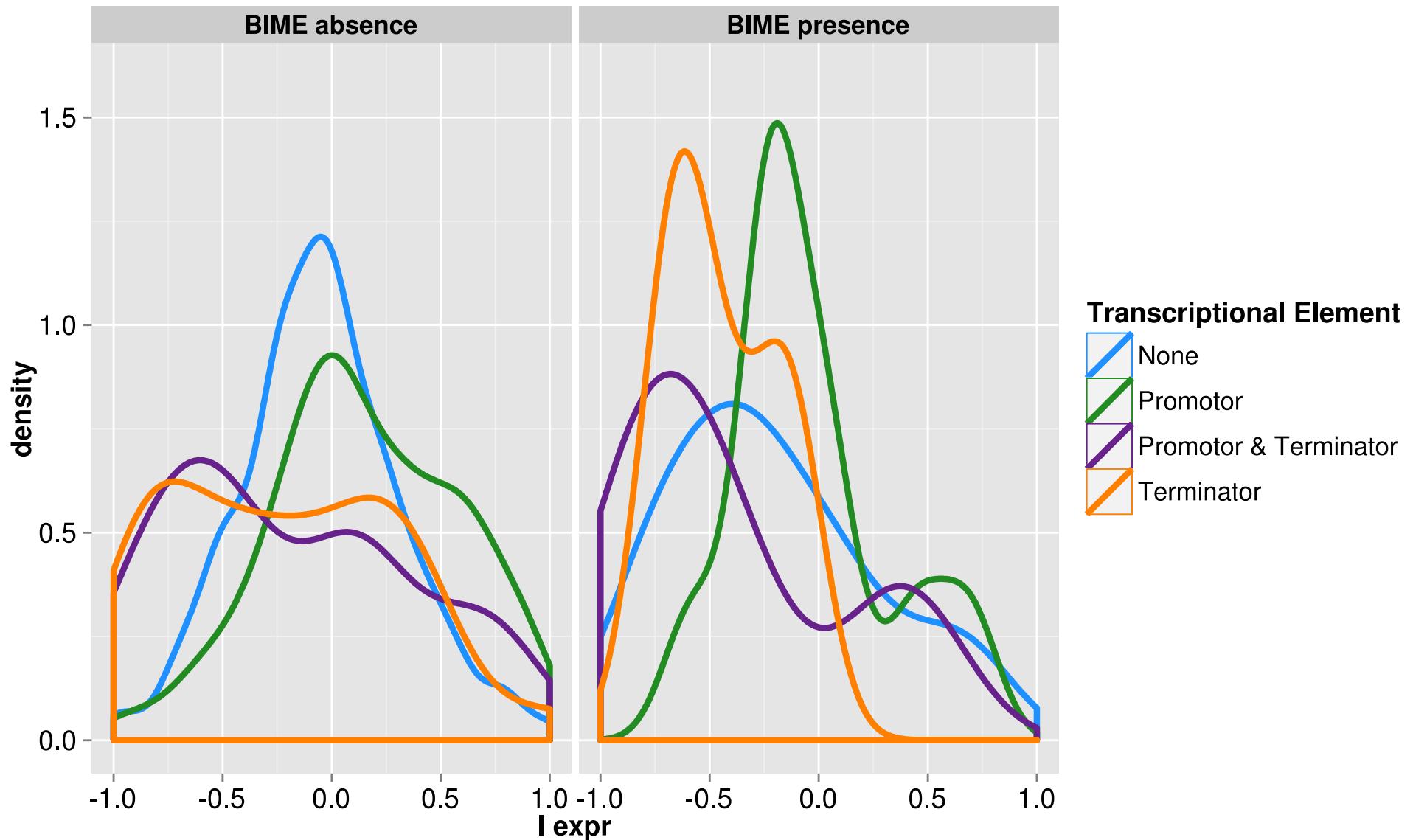
INRA: Applied Mathematics and Informatics

Christine Gaspin



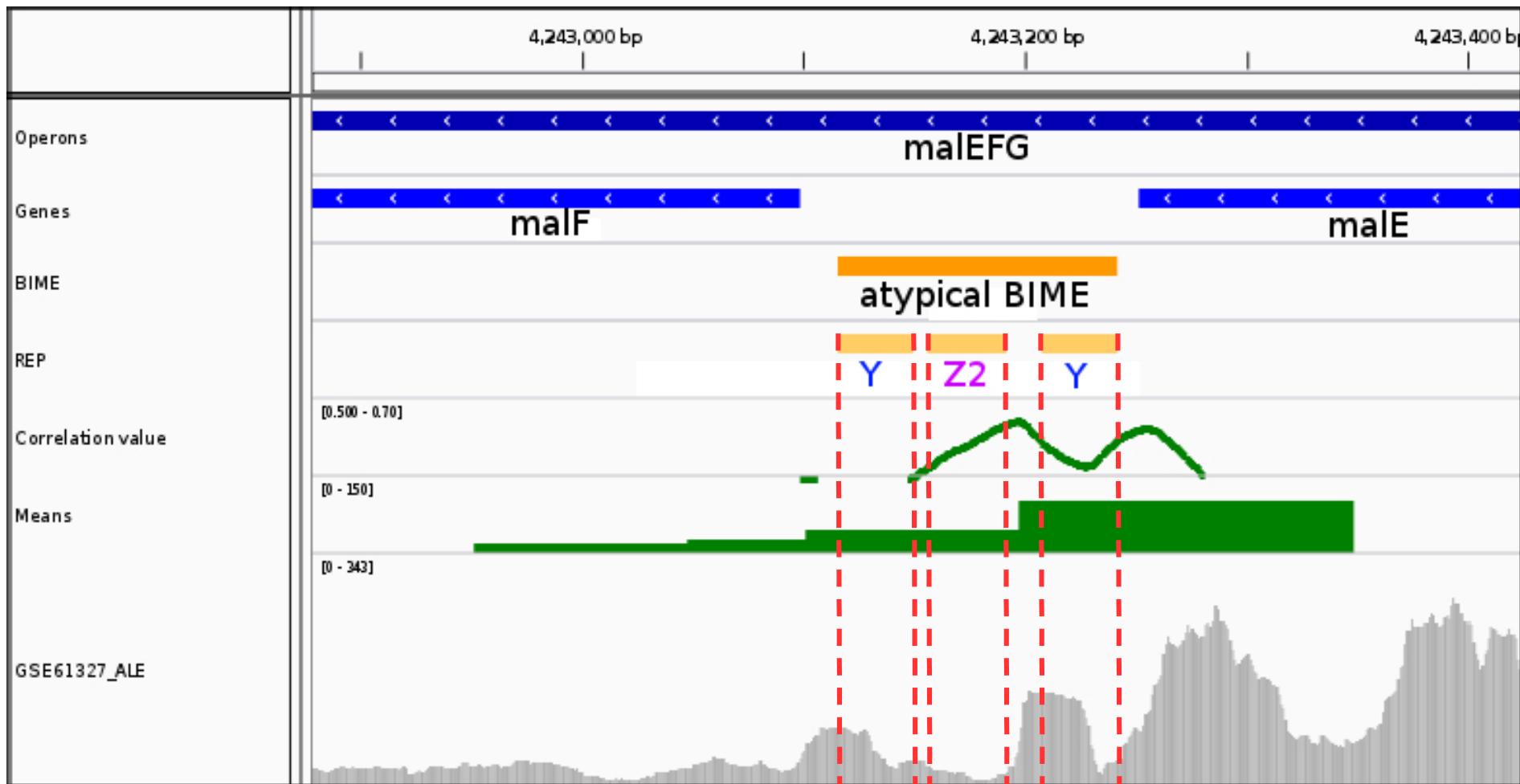
Global BIME effect in operons

Distributions of I_{expr} depending on Transcriptional Elements



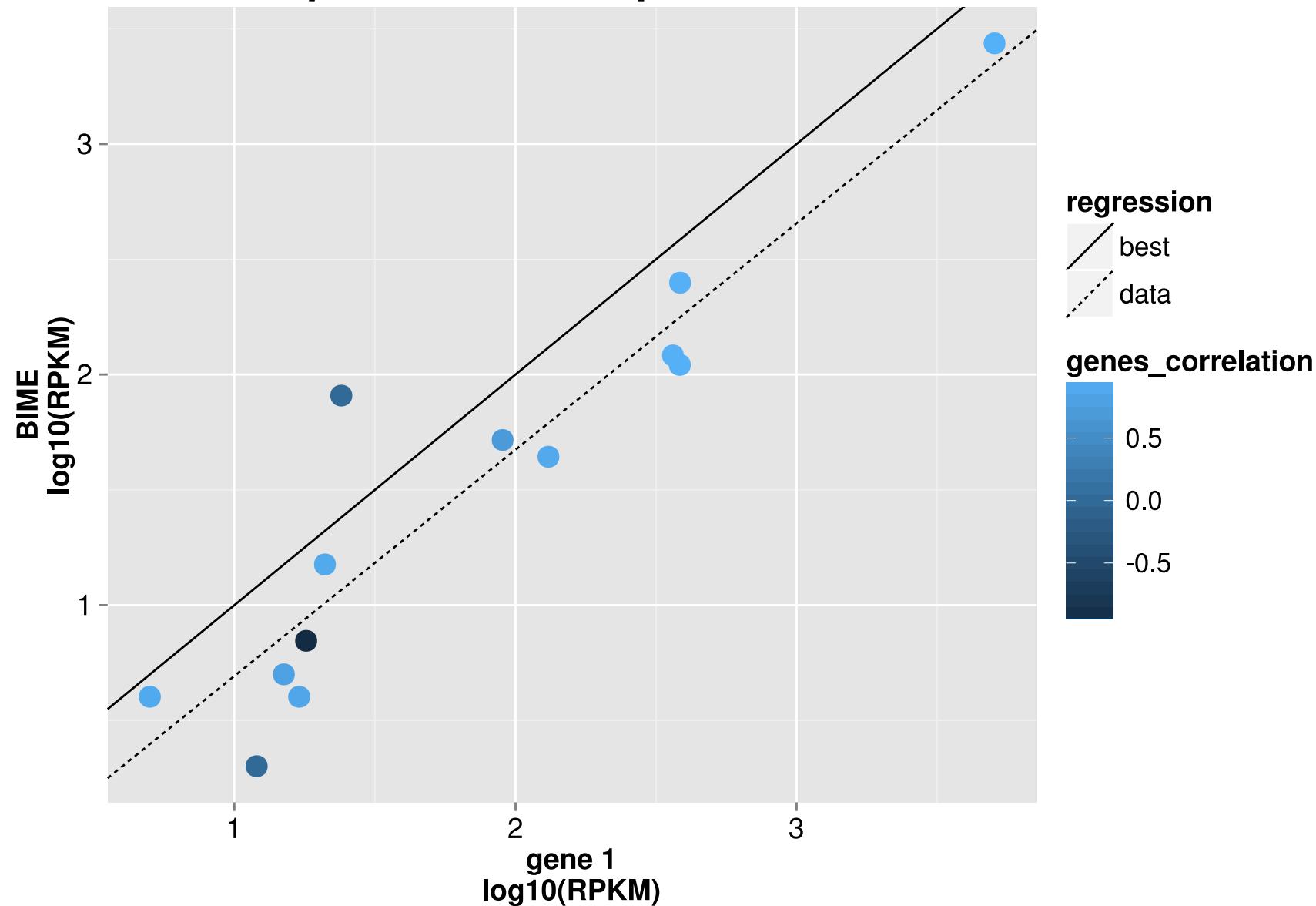
Adaptation of operon detection method (Local approach)

Expression profile correlation on *malE-malF* BIME



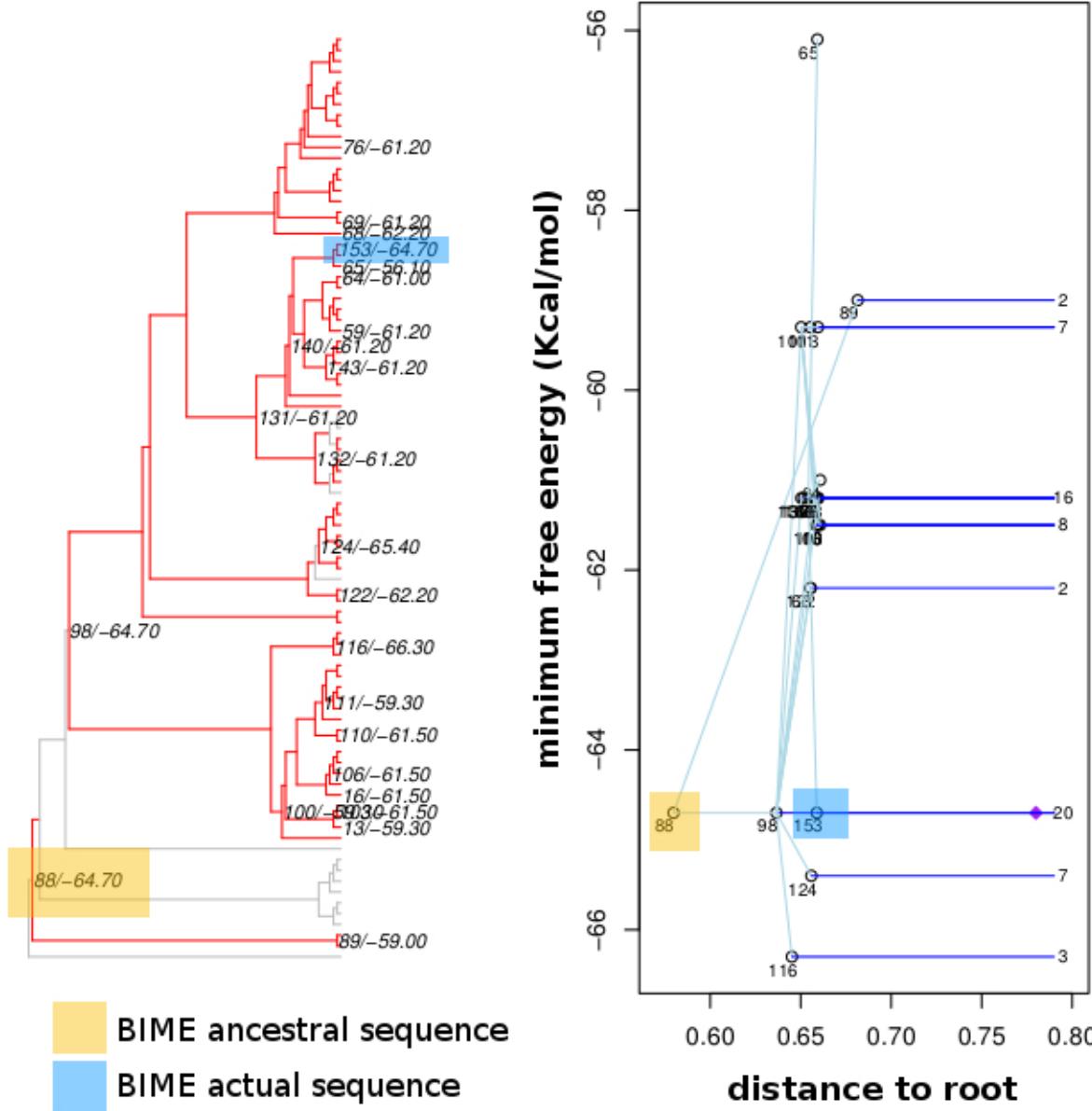
BIME sequencing bias

No DE in operons, no promotors or terminators



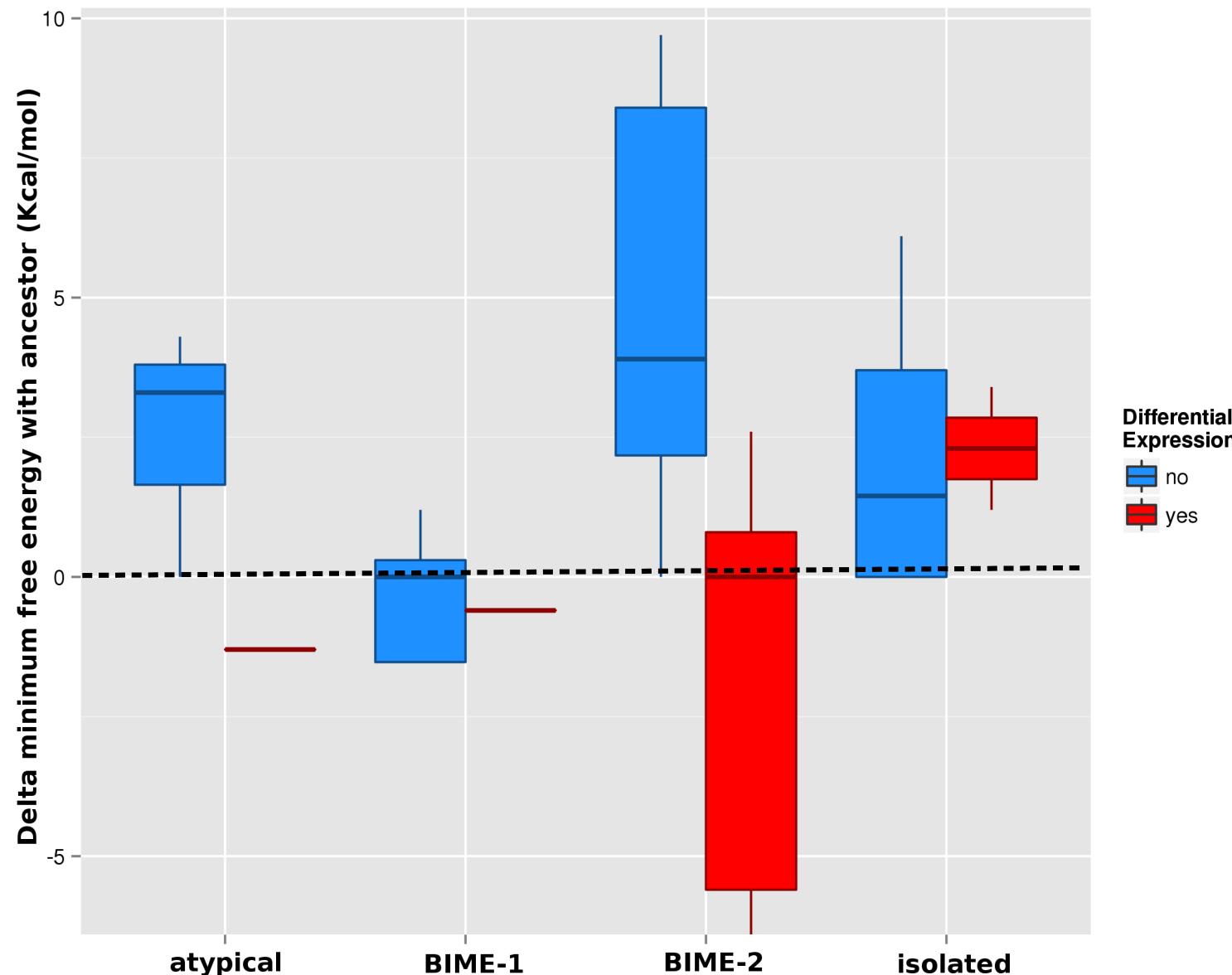
Ancestral sequences and secondary structures

malE-malF BIME 2nd structure energy evolution



Ancestral sequences and secondary structures

Delta of energy between ancestor & actual BIME



Segmentation efficiency & replicates

