

Longitudinal modeling of omics data from HPV-induced carcinogenesis

Viktorian Miok

Contributors

Biostatistics department

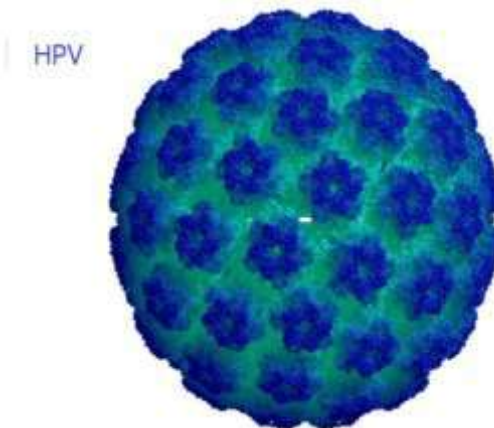
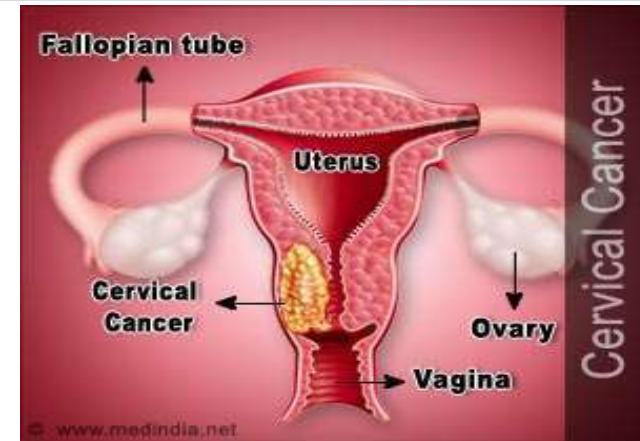
- Viktorian Miok
- Wessel van Wieringen
- Mark van de Wiel

Pathology department

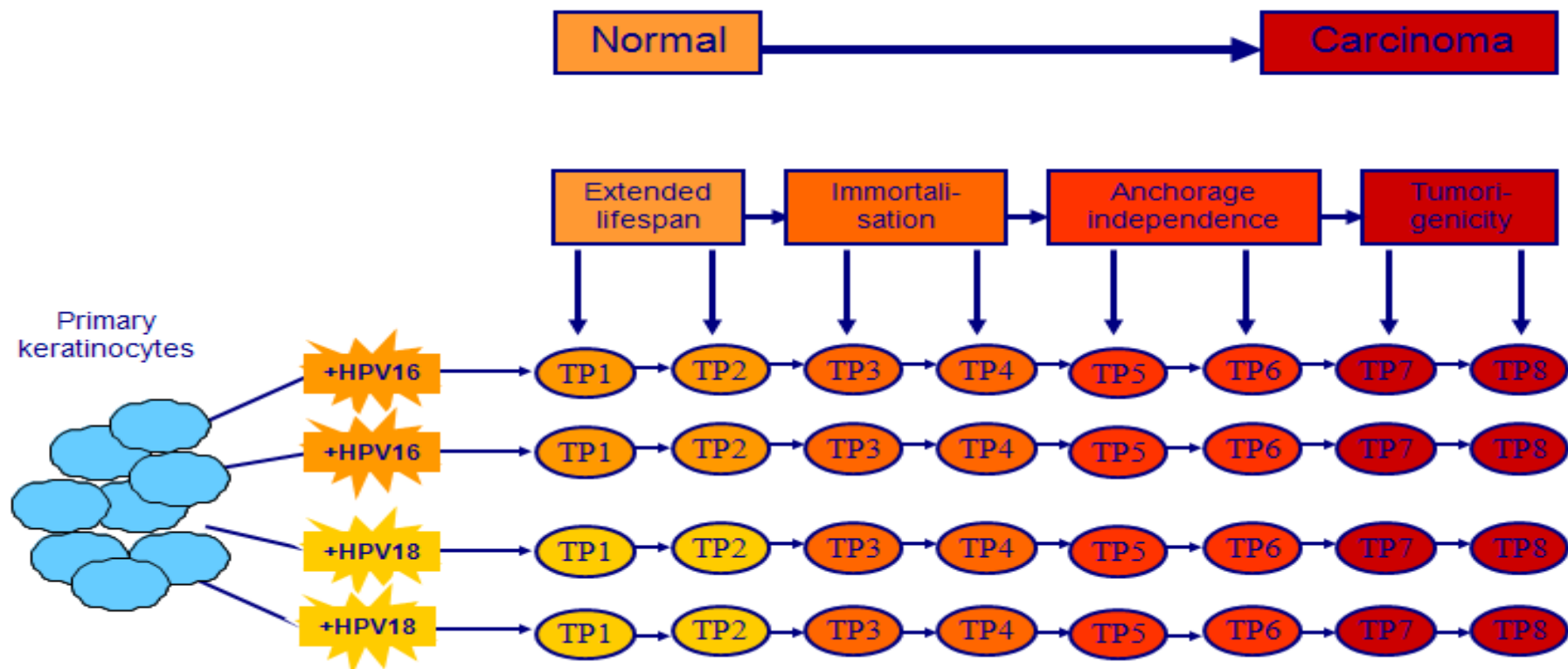
- Saskia Wilting
- Annelieke Jaspers
- Renske Steenbergen
- Peter Snijders

Cervical cancer study

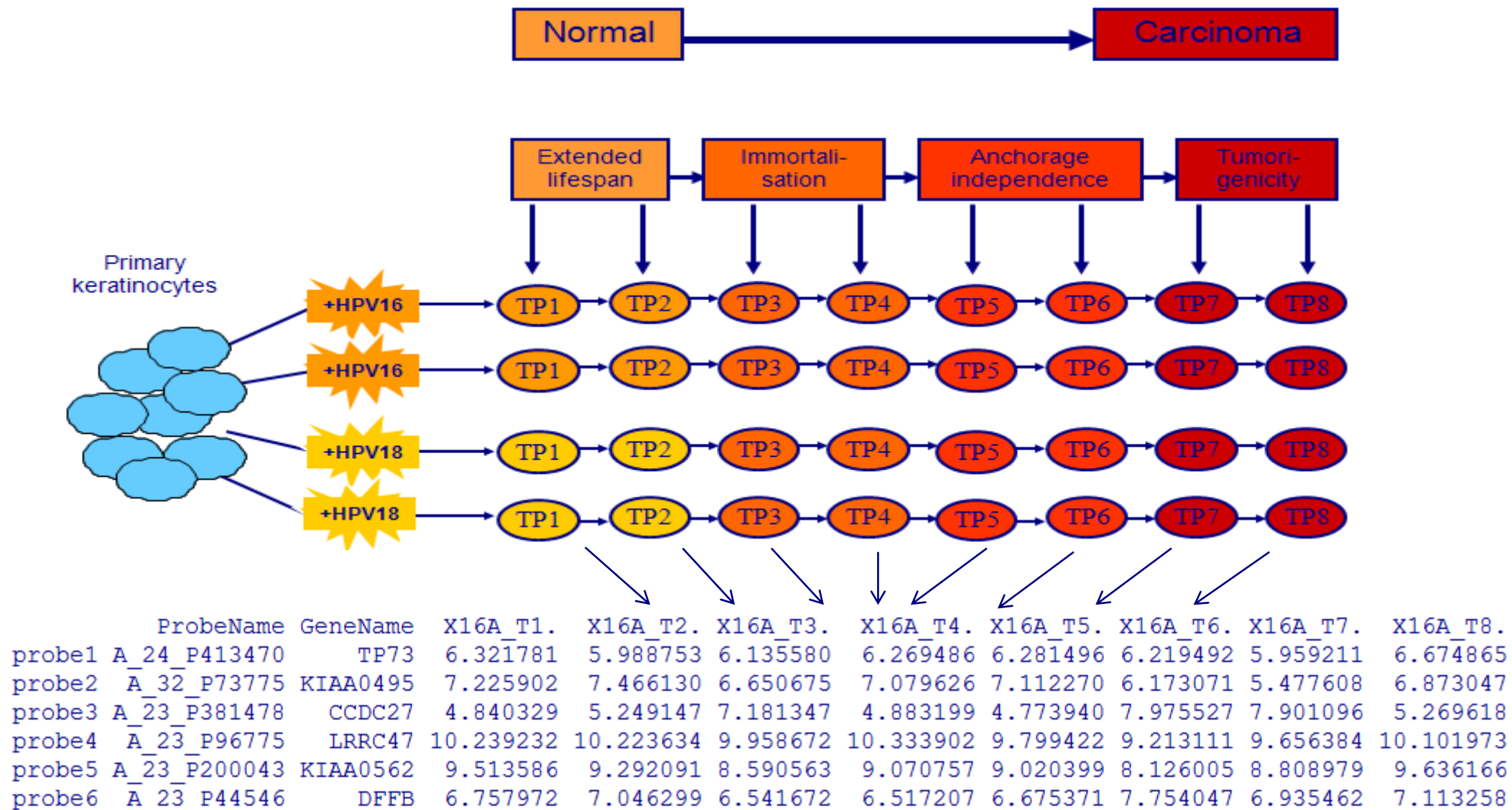
- Third most common cancer in women worldwide.
- Caused by HPV virus (70% cases HPV16 and HPV18) and followed by additional (epi)genetic abnormalities.
- Cell line model – in vitro model system of HPV-induced transformation.
- Integration – high-throughput multi level molecular data sets.
- Understand molecular mechanism driving cervical carcinogenesis



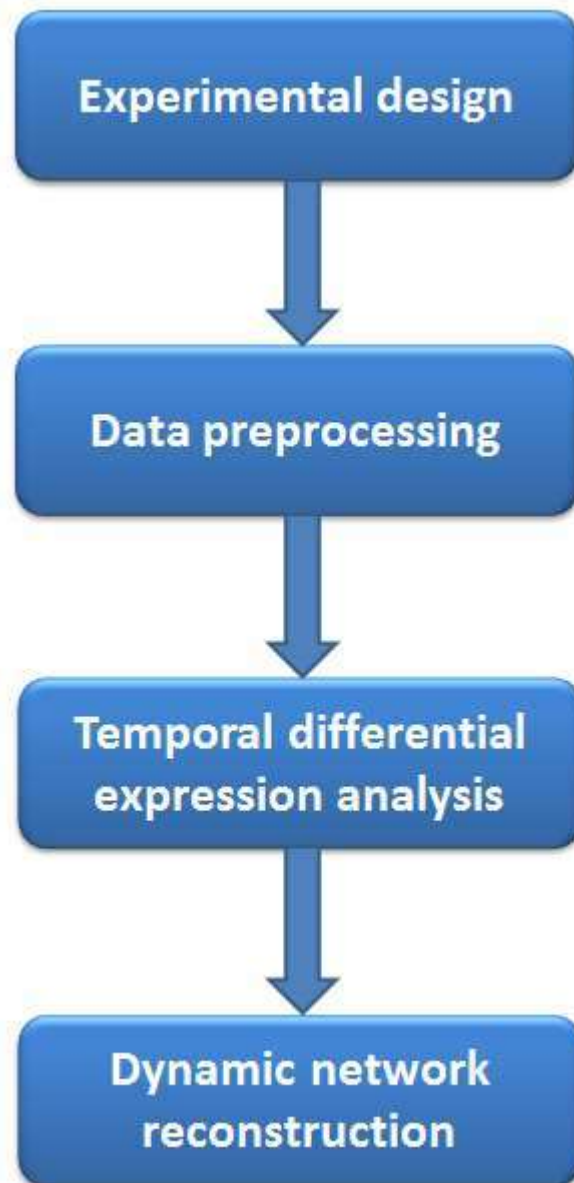
Time-course experiment



Time-course experiment



What we did?



mRNA: 45K probes arrays
miRNA: 60K probes arrays
CN: 180K probes arrays

mRNA: 27637 genes
miRNA: 1187 genes
CN: 27637 genes

mRNA: 3642 genes
miRNA: 106 genes

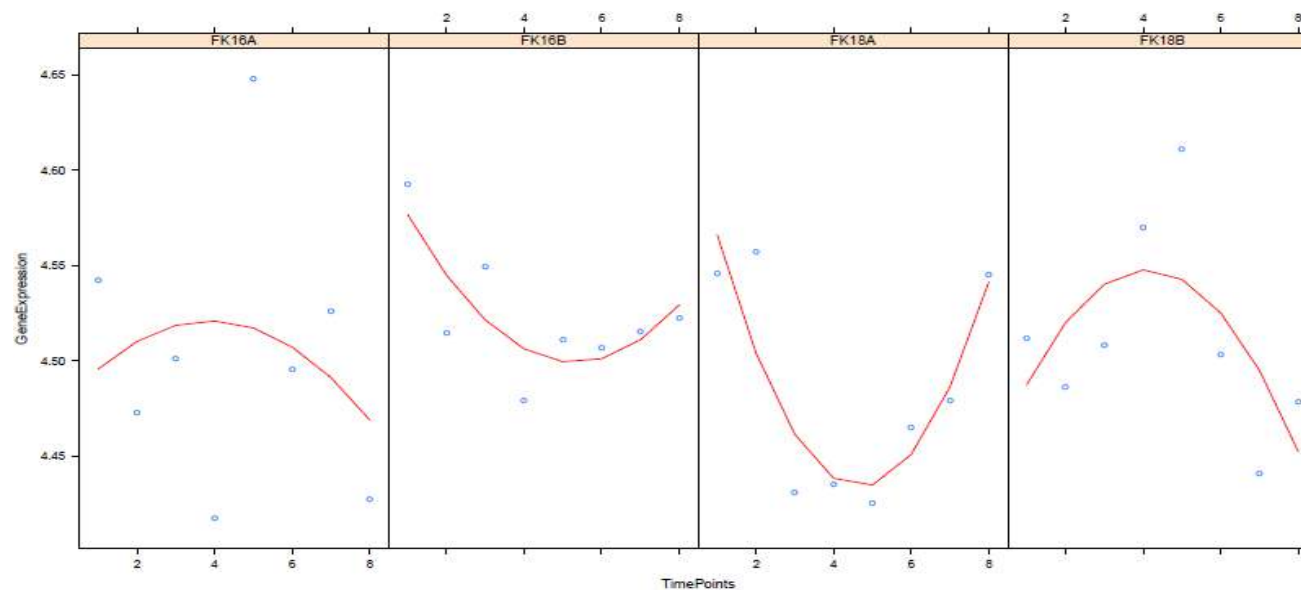
mRNA: 64 genes linked to p53 signaling pathway
miRNA: 106 genes which target mRNA

Temporal differential expression

$$GE = CL + CN + Time$$

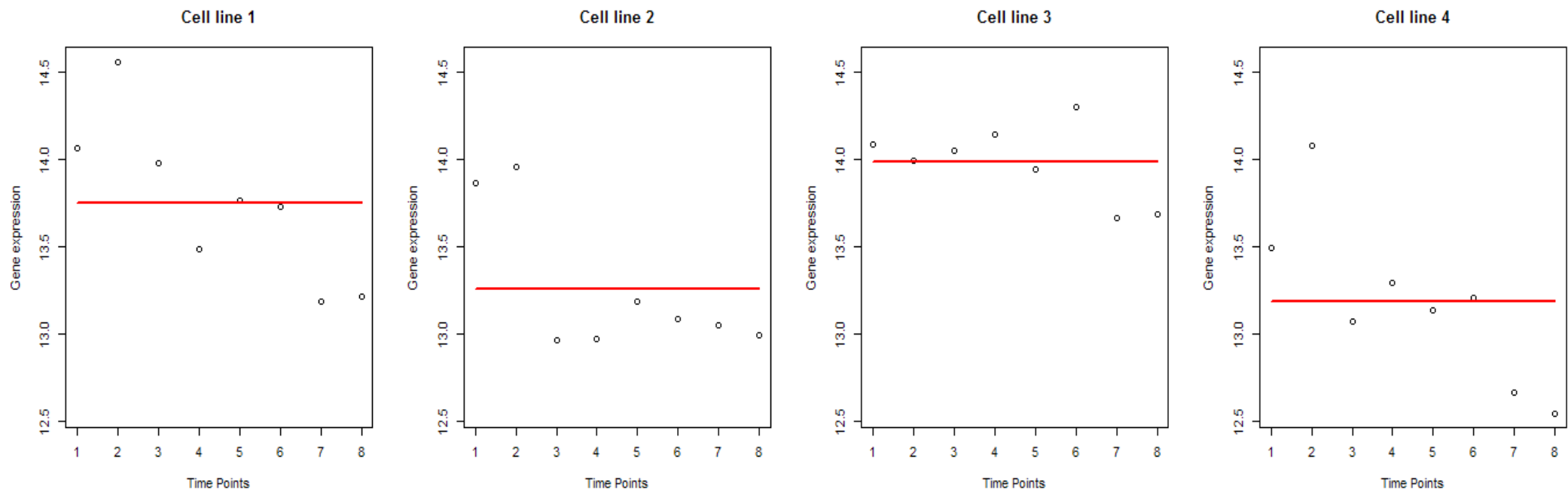
fixed (cell line, CN) effect

random (time) effect



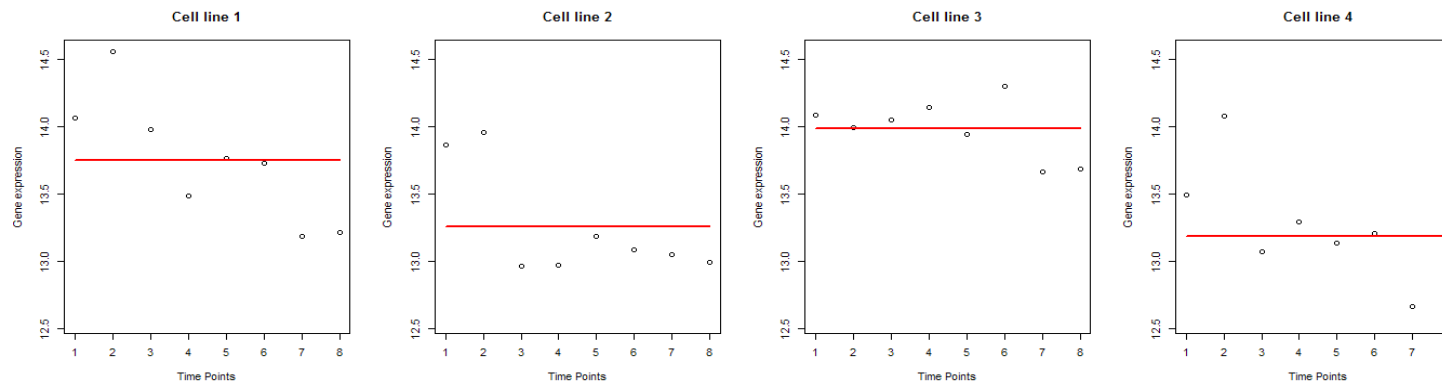
- Cell line effect

SP1



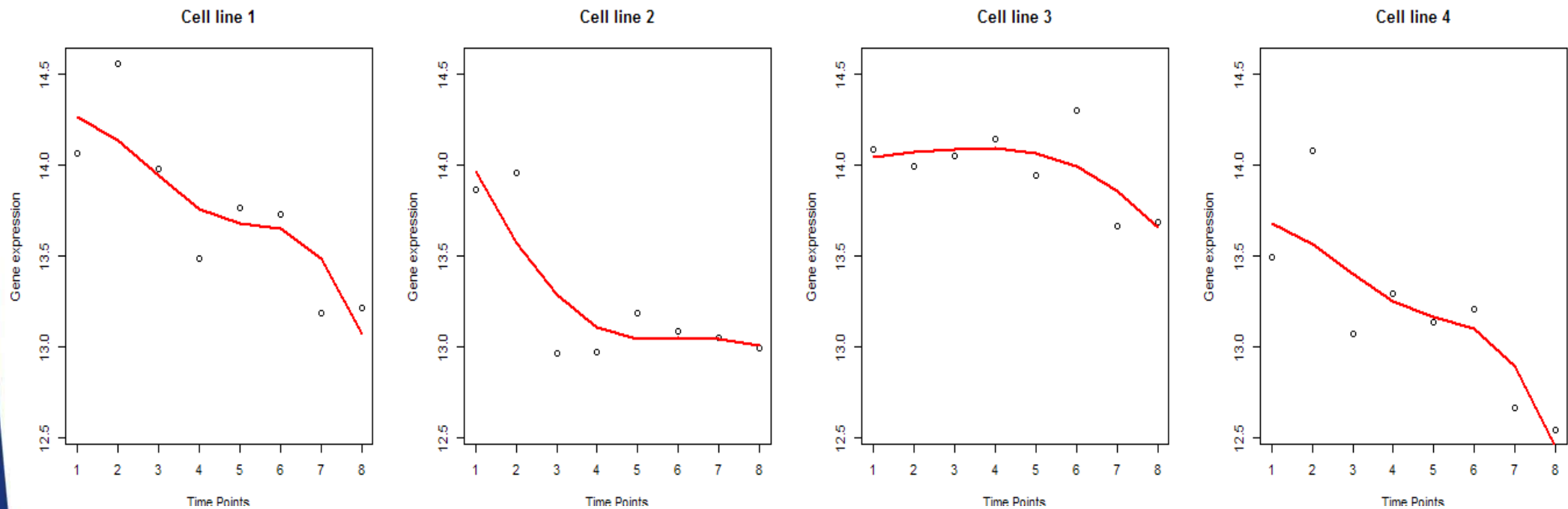
- Cell line effect

SP1



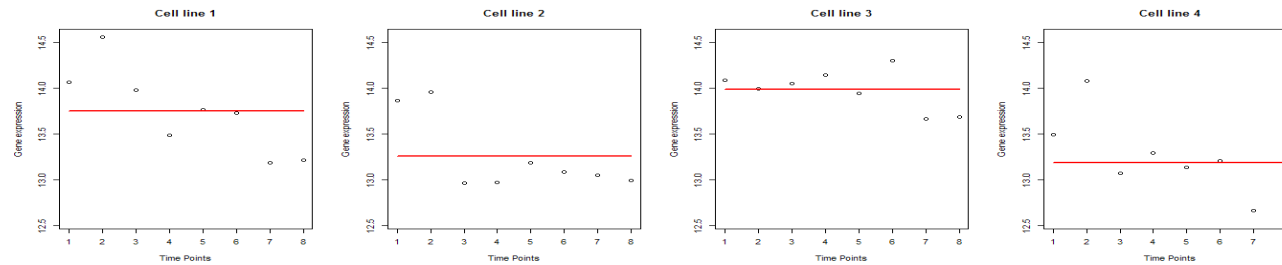
- Cell line and Time effect

SP1



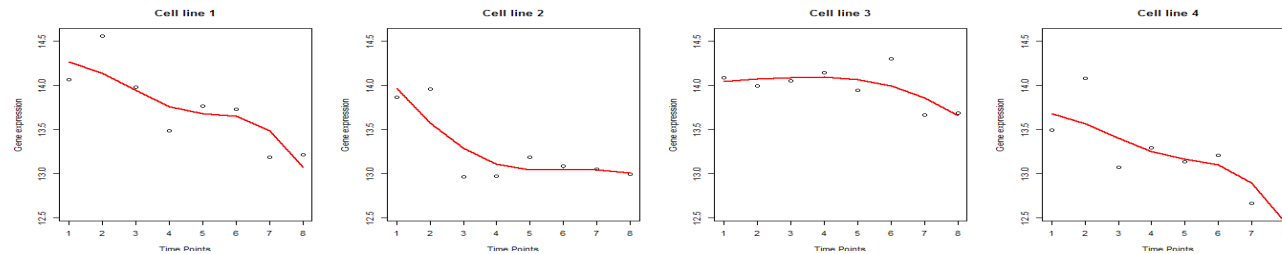
- Cell line effect

SP1



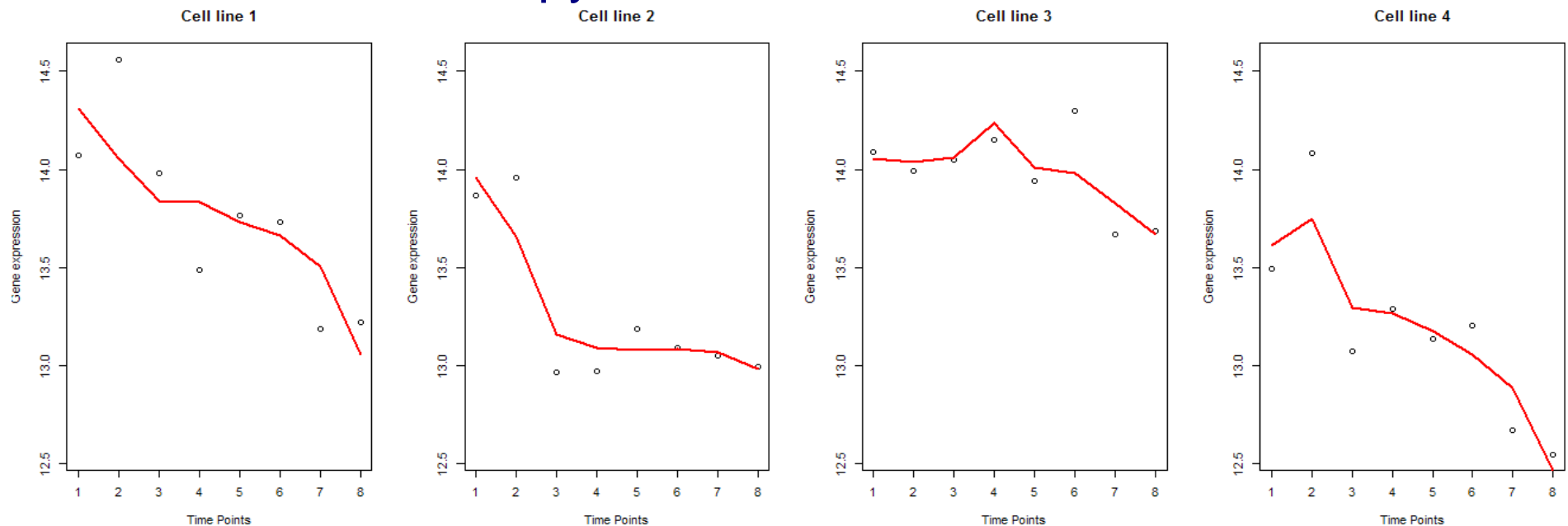
- Cell line and Time effect

SP1



- Cell line, Time and Copy Number effect

SP1

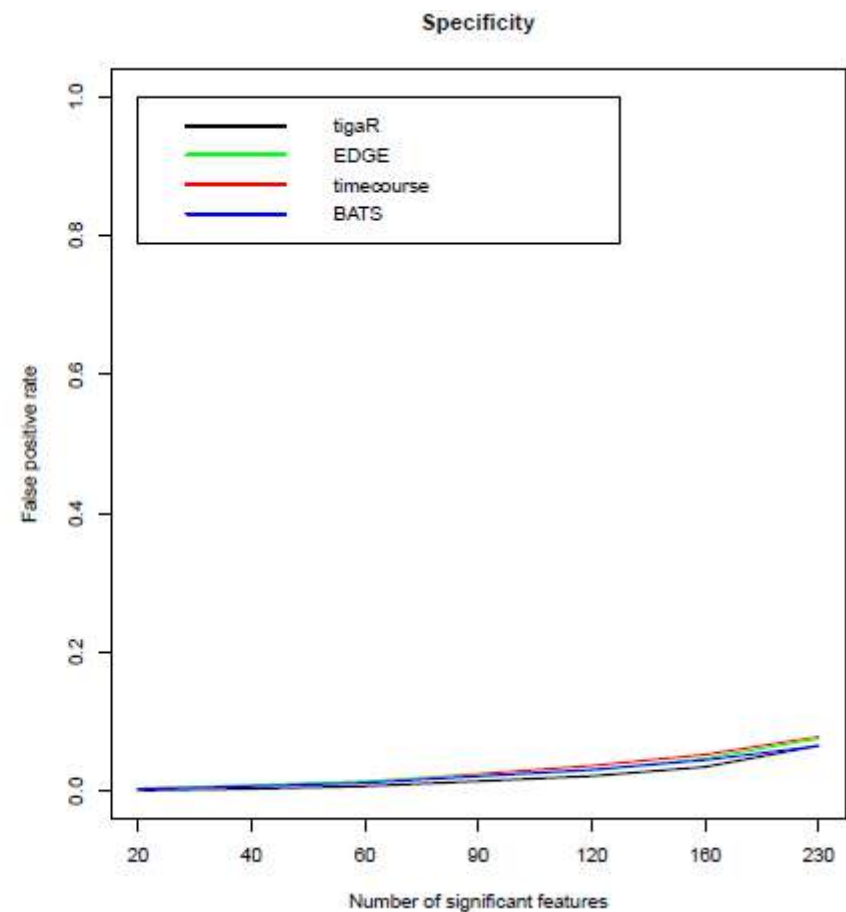
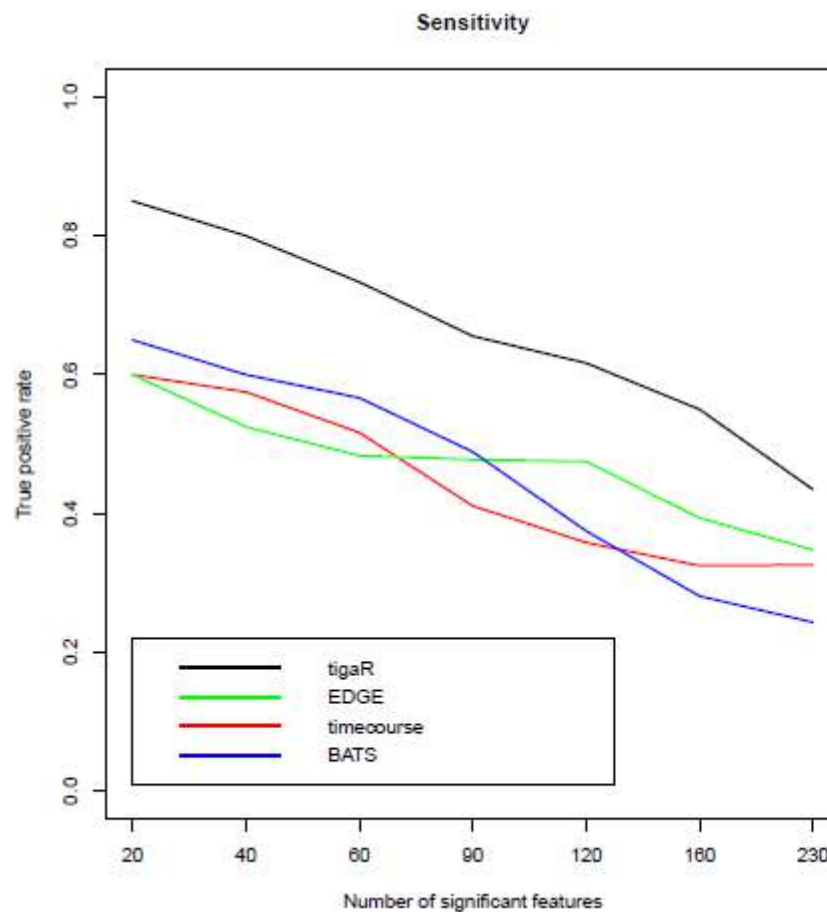


Comparison

- Comparison of following methods:
 - **timecourse** – Tai and Speed, Annals of Statistics, 2006.
 - **EDGE** – Storey et al., PNAS, 2005.
 - **BATS** – Angelini et al., BMC Bioinformatics, 2008.
 - **tigaR** – Miok et al., BMC Bioinformatics, 2014.
- Method is applied on two data sets
 - Data from our experiment (only mRNA data)
 - Data from Storey et al., PNAS, 2005.

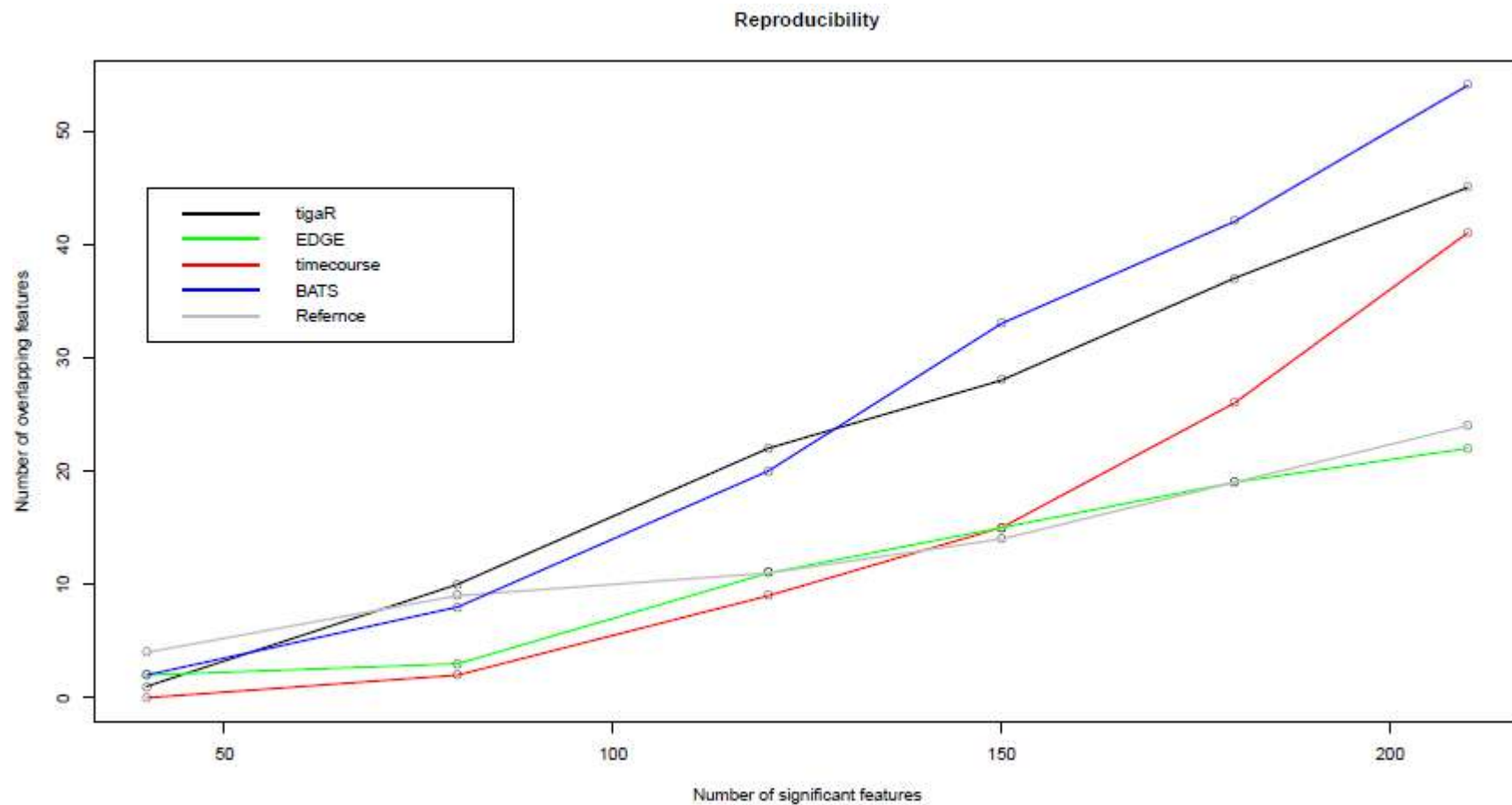
Sensitivity and specificity

- Truth – overlap of significant genes among methods.

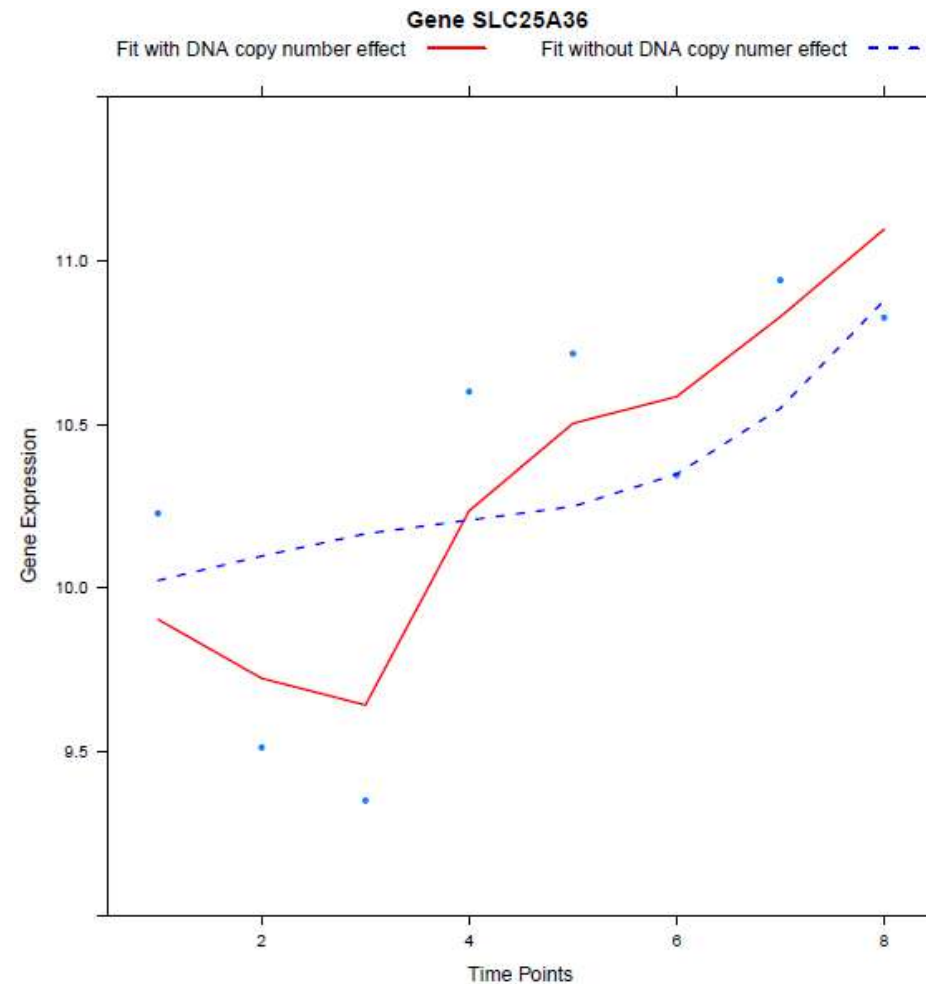


Reproducibility

- Equally divided data set in two groups.

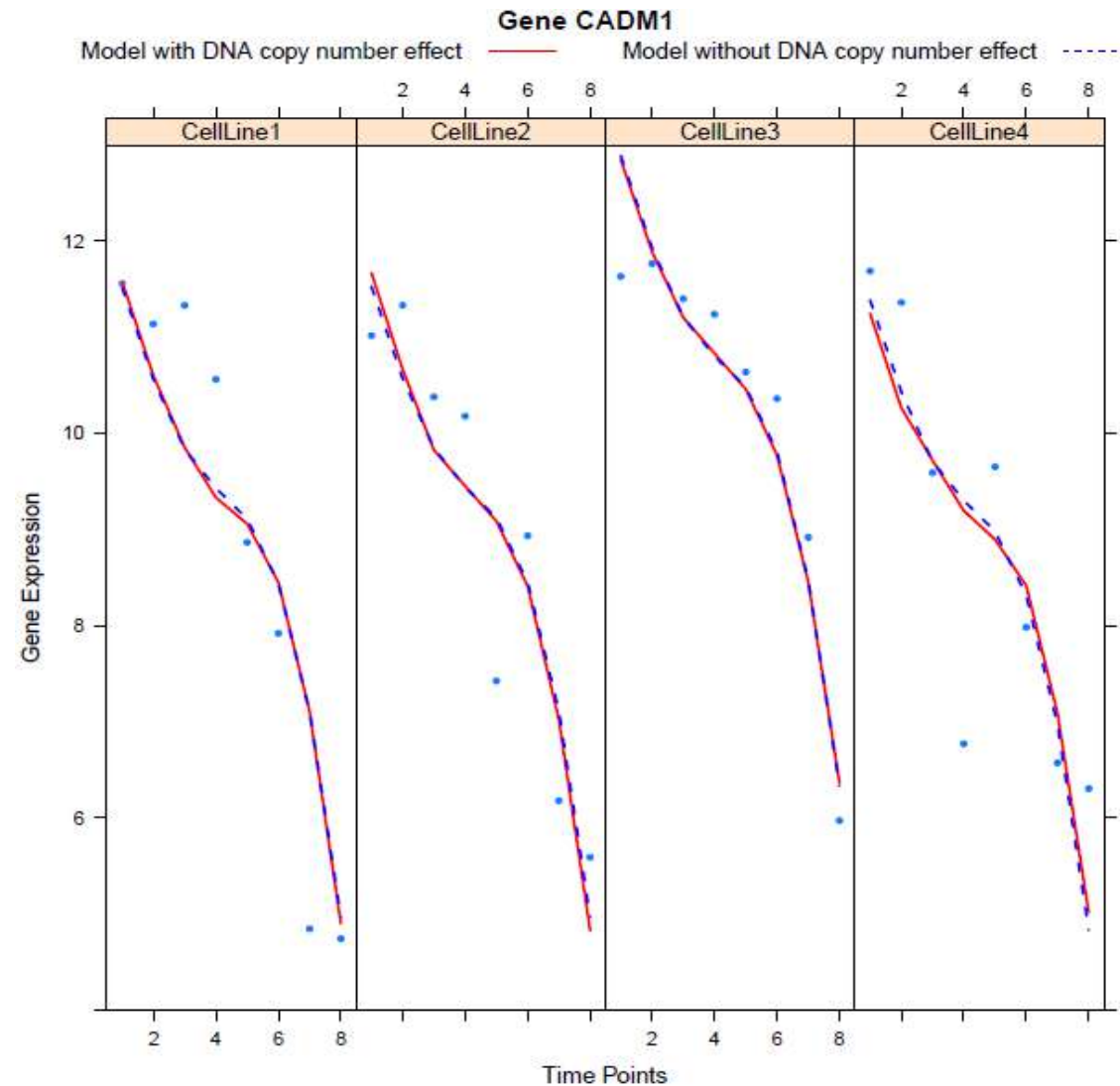


SLC25A36 – gene with CN effect



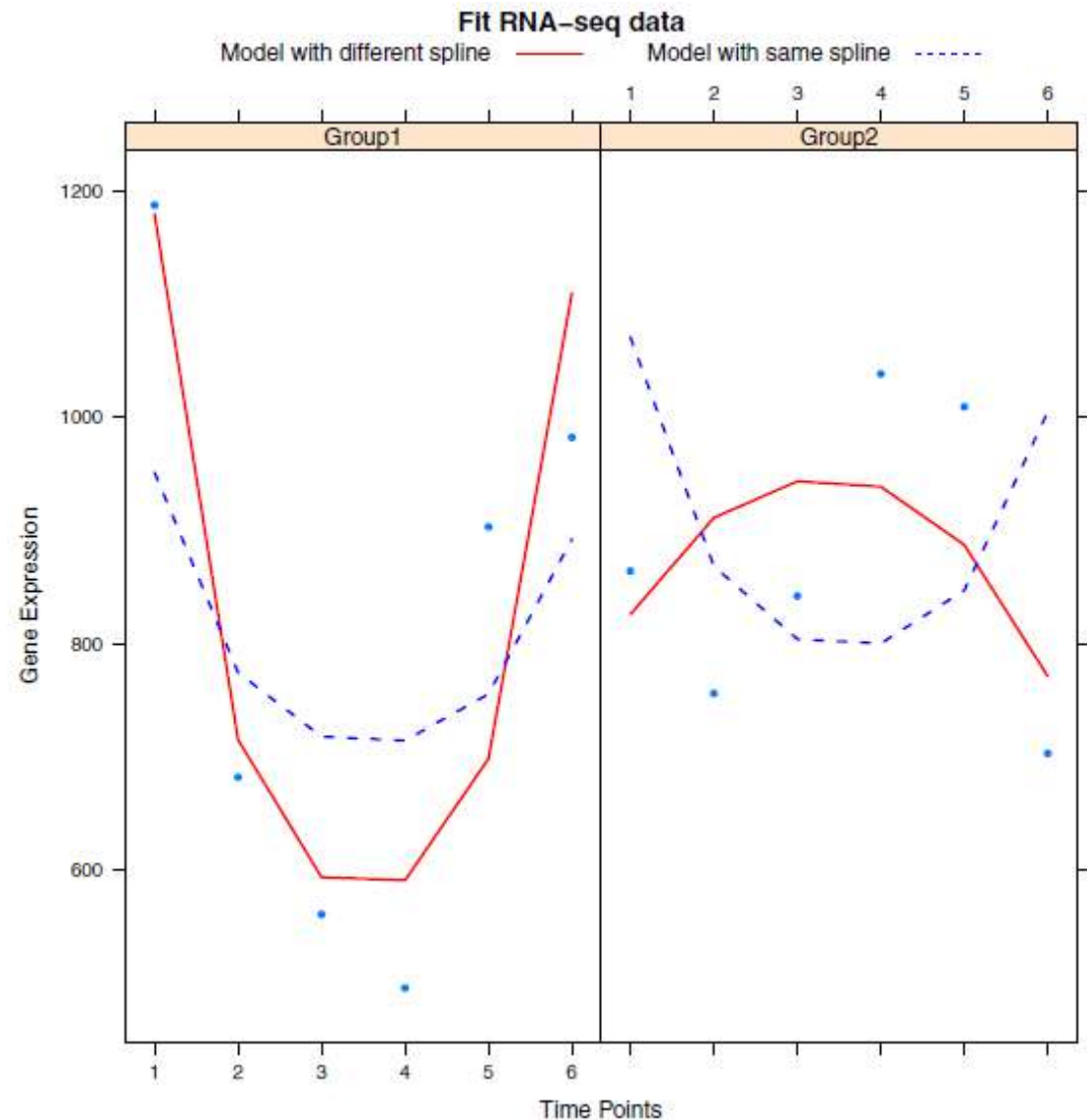
Wilting et al., Genes, Chromosomes and Cancer, 2008.

CADM1- gene without CN effect

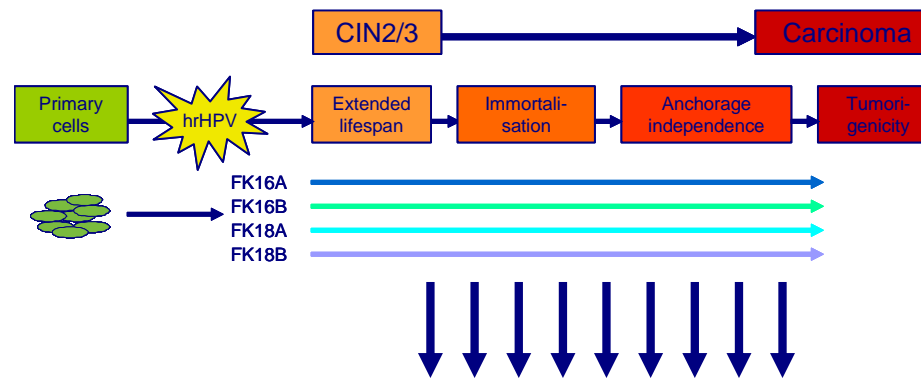


RNA-seq data

- Changing link function method can deal with count data.
- Two group time-course RNA-seq data.

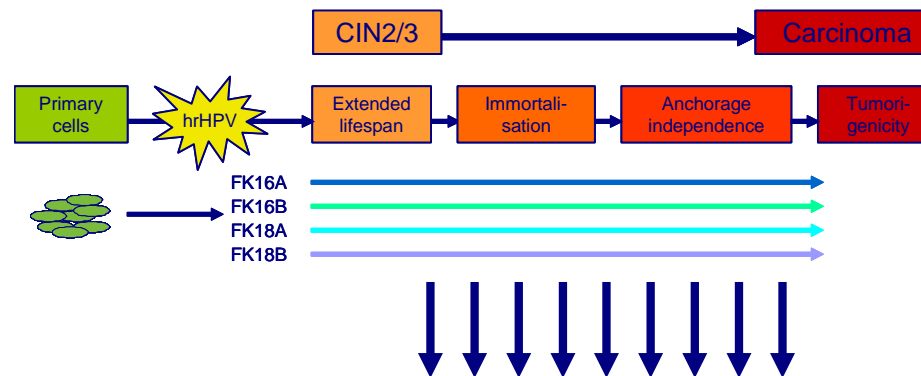


tigaR analysis



Measured expression of 1187 miRNAs and 27637 mRNAs

tigaR analysis



Measured expression of 1187 miRNAs and 27637 mRNAs

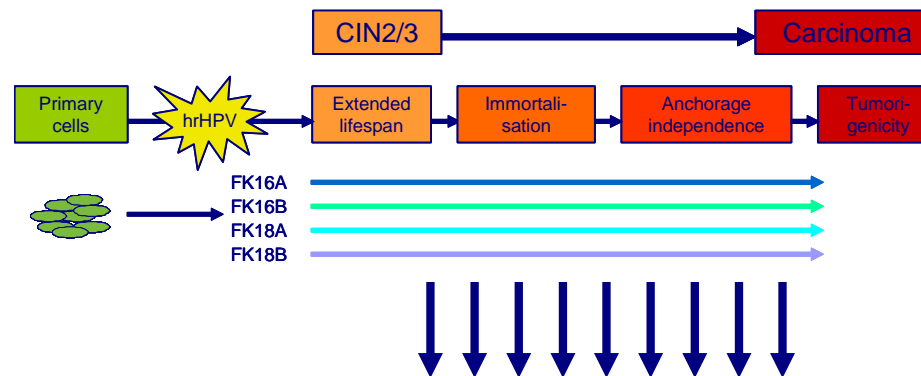


106 significant miRNAs and 3642 significant mRNAs

(concordant change in expression in at least 3 cell lines)



tigaR analysis



Measured expression of 1187 miRNAs and 27637 mRNAs



106 significant miRNAs and 3642 significant mRNAs

(concordant change in expression in at least 3 cell lines)



36 miRNAs and 1233 mRNAs linked with CN

(~34% of altered expression in both cases)



**Thank you for your
attention!**