

# Time series experiments

microarray.no



# Time series experiments

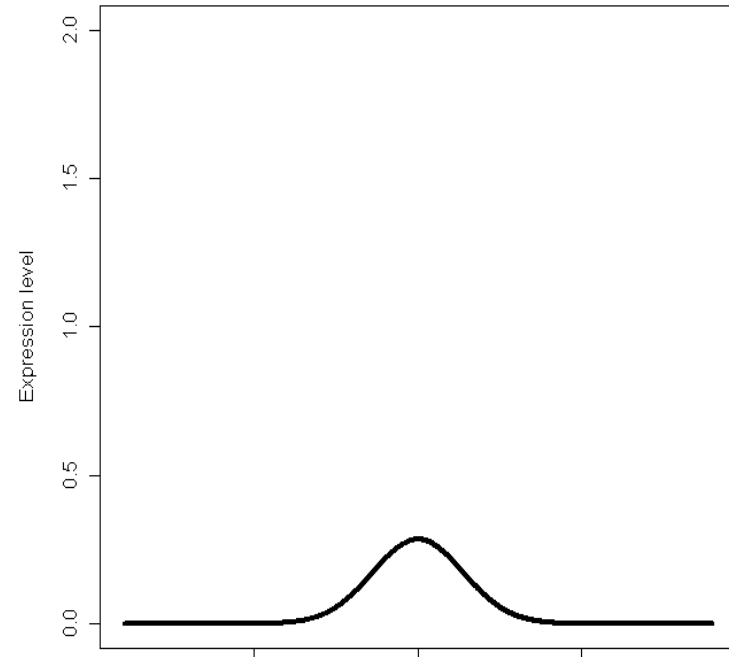
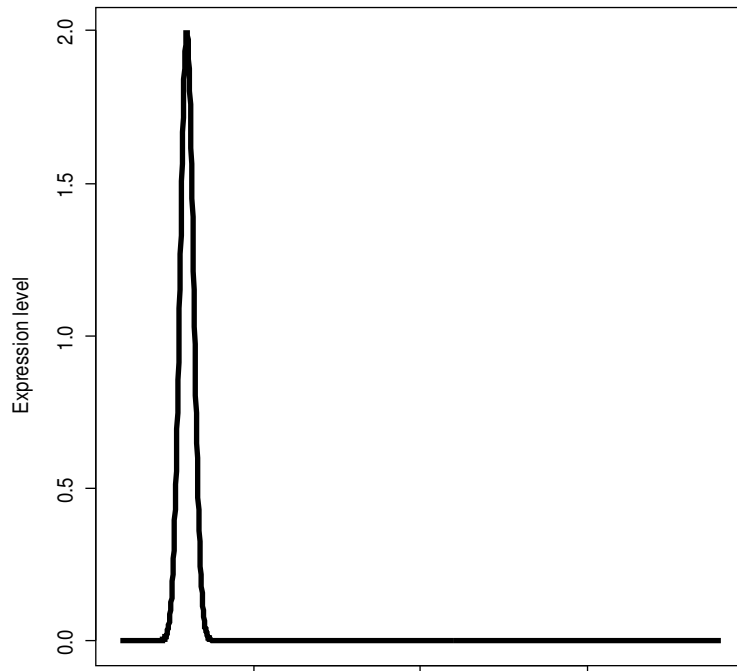
- Why is this a separate lecture:
  - The price of microarrays are decreasing – more time series experiments are coming
  - Often a more complex experimental design
  - Many time points makes this a large experiment
  - The analysis method is often specified by what you are looking for

# Time series experiments

- Why consider doing a time series experiment:
  - Biology is dynamic
    - The genes can change their expression in different ways during time
  - Observe cascades and secondary effects

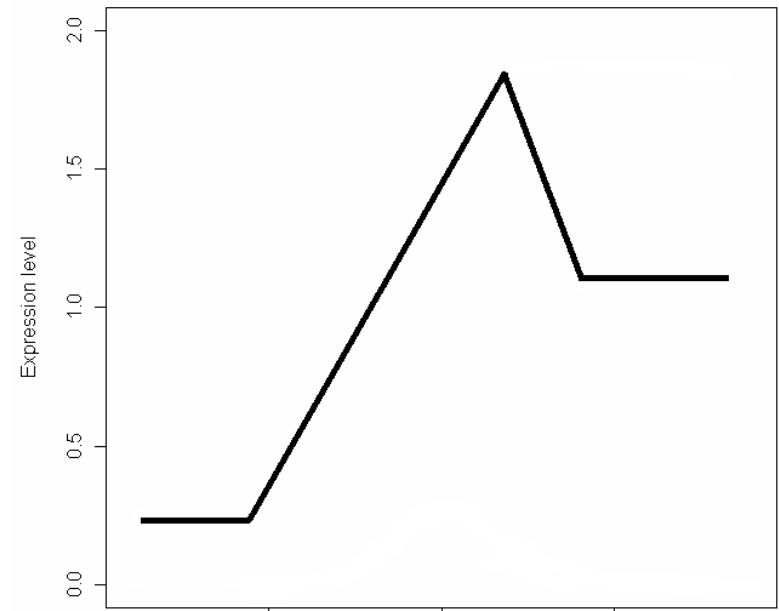
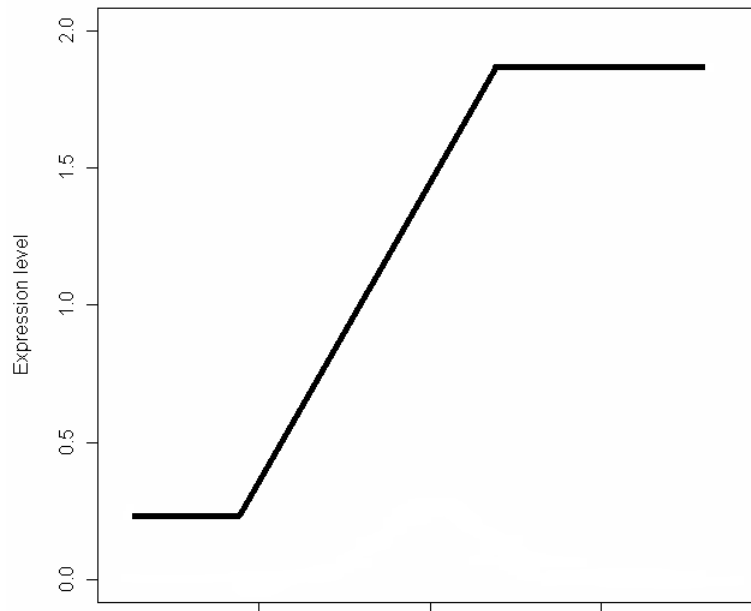
# Time series experiments

- Peaks



# Time series experiments

- Level change



# Experimental design (Ex.d)

- Often larger and complex
- What to consider in an experimental design of a time series study:
  - Time points
  - Distance between time points
  - Replicates
  - Control
  - Technology

# Ex.d.: Time points and the distance between them

- How to choose:
  - What is already known about the system/genes
    - Theoretical and experimental knowledge

Early /late response

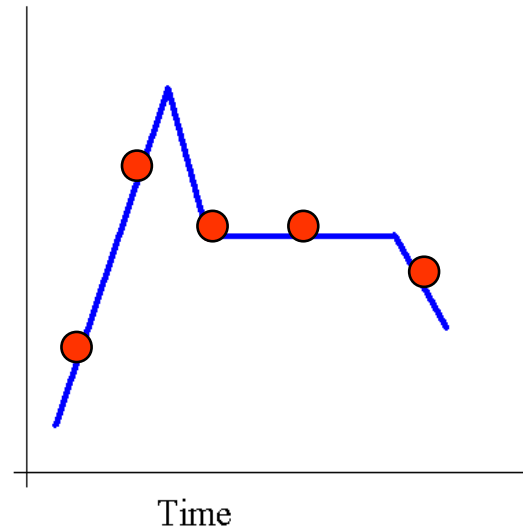
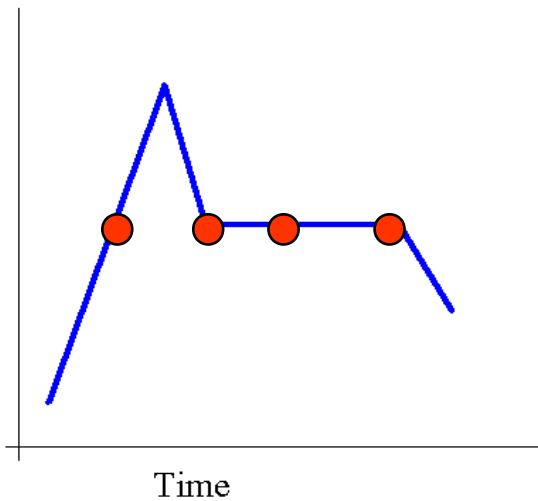
The direction of the response

Will changes in some genes influence others

- What is your field of interest

## Ex.d.: Replicates

- How many replicates to use:
  - Many time points but few replicates



- Should be enough to do the analysis you want!

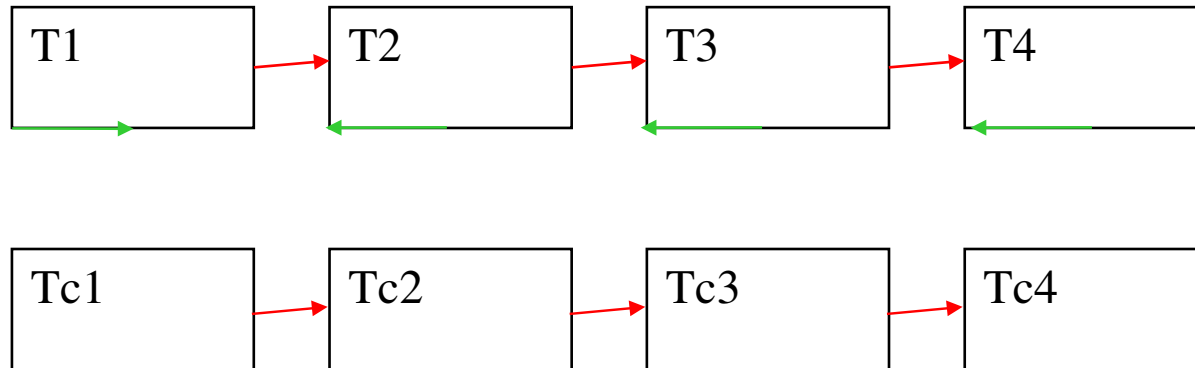
## Ex.d.: Control

- What to use as a control and how many controls?
  - One control point or one time series control
  - Differences between treatments
  - The other time points as a control

## Ex.d.: Control

- One control
  - Use one time point or a untreated sample as control
    - Ok if no other time effects are expected, such as growth, phase of expression
- A untreated time series as control
  - Same time points, technology, sample source and same handling except the treatment
  - Find genes that change due to treatment over time, filter out some of the effects due to handling and time

## Ex.d.: Control



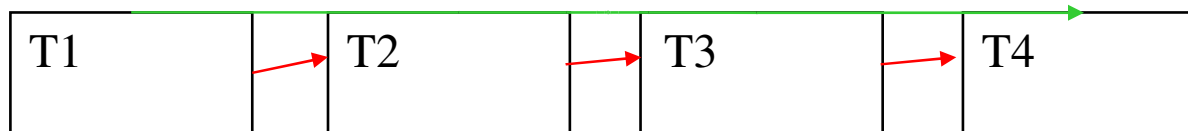
- **Vertically:** Differences due treatment, but you reduce the effect of the handling
- **Horizontally:** Differences due to treatment, time and some handling effect

## Ex.d.: Control

- Differences between treatments - compare two or more time series
- Have the same:
  - Time points
  - Conditions
  - Sample source
  - Technology
  - Handling

## Ex.d.: Control

- The other time points as a control:
  - Changes due to treatment, time but it does not exclude the effect of the handling
  - Less arrays used
  - See **short** time effects and **long** time effects'



- Ok if no other time effects are expected, such as growth, phase of expression

## Ex.d.: Technology

- What technology to choose
  - Homemade/commercial
  - One channel
  - Two channel: Reference design is recommended
- What comparisons are to be done
- Budget

# Experimental design

- Often larger and complex
- What to consider in an experimental design of a time series study:
  - Time points
  - Distance between time points
  - Replicates
  - Control
  - Technology

# Pre-processing

- The same steps as for non time series microarray experiments, but not always the same algorithm or the same use of the algorithms:
  - QC
  - Filtering
  - Normalization
  - .....

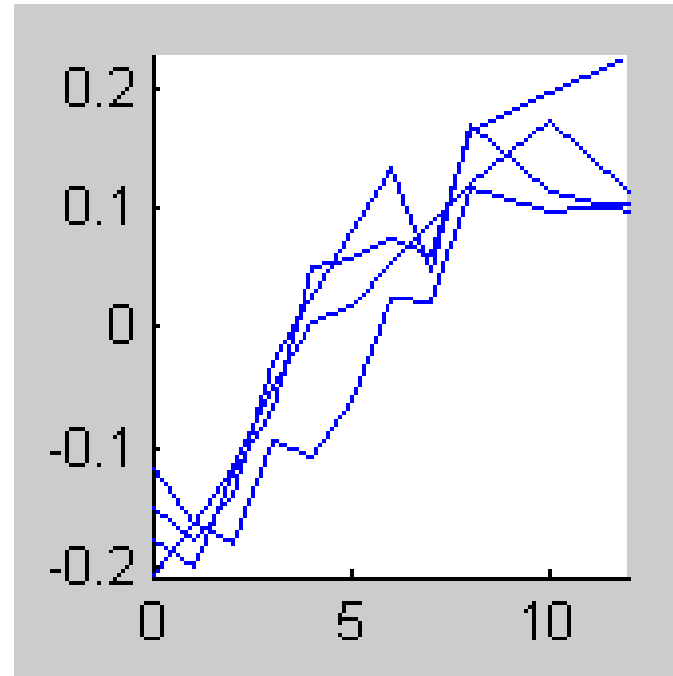
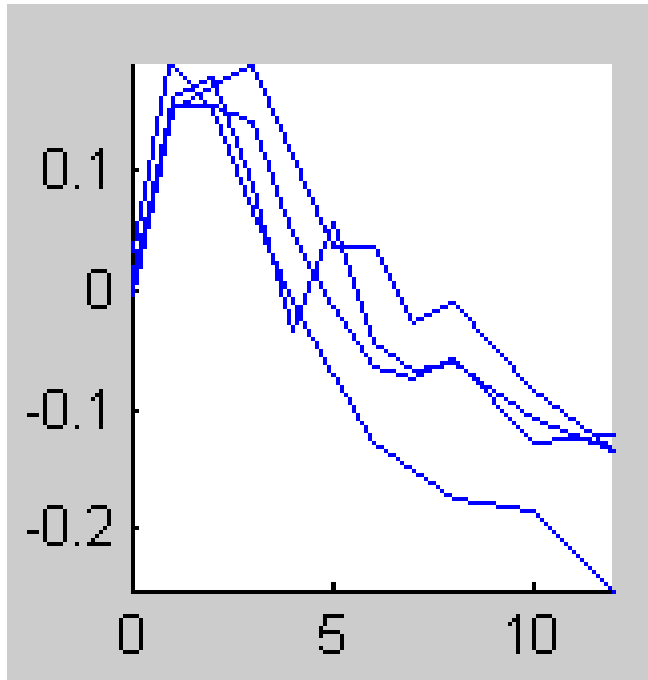
# Analysis of time series data

- Clustering
- T-test between time points
- PCA
- PLS
- Cyclic
- Profile
- .....

# Analysis: Clustering

- Filter the data set before clustering
- Cluster the genes
- Use correlation as a distance measure (as a starting point)
- There exist no right number of clusters – *a priori* knowledge is needed

# Analysis: Clustering



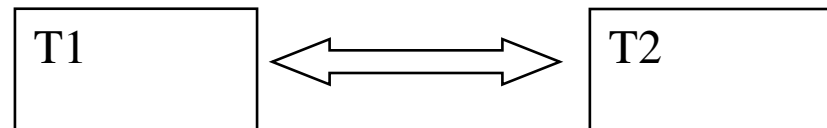
microarray.no

# Analysis: T-test

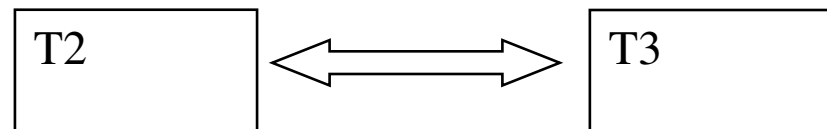
- One time point vs. another time point
- You have 3 time points: T1, T2 and T3
- What should you compare?

# Analysis: T-test

- Early response:



- Late response:



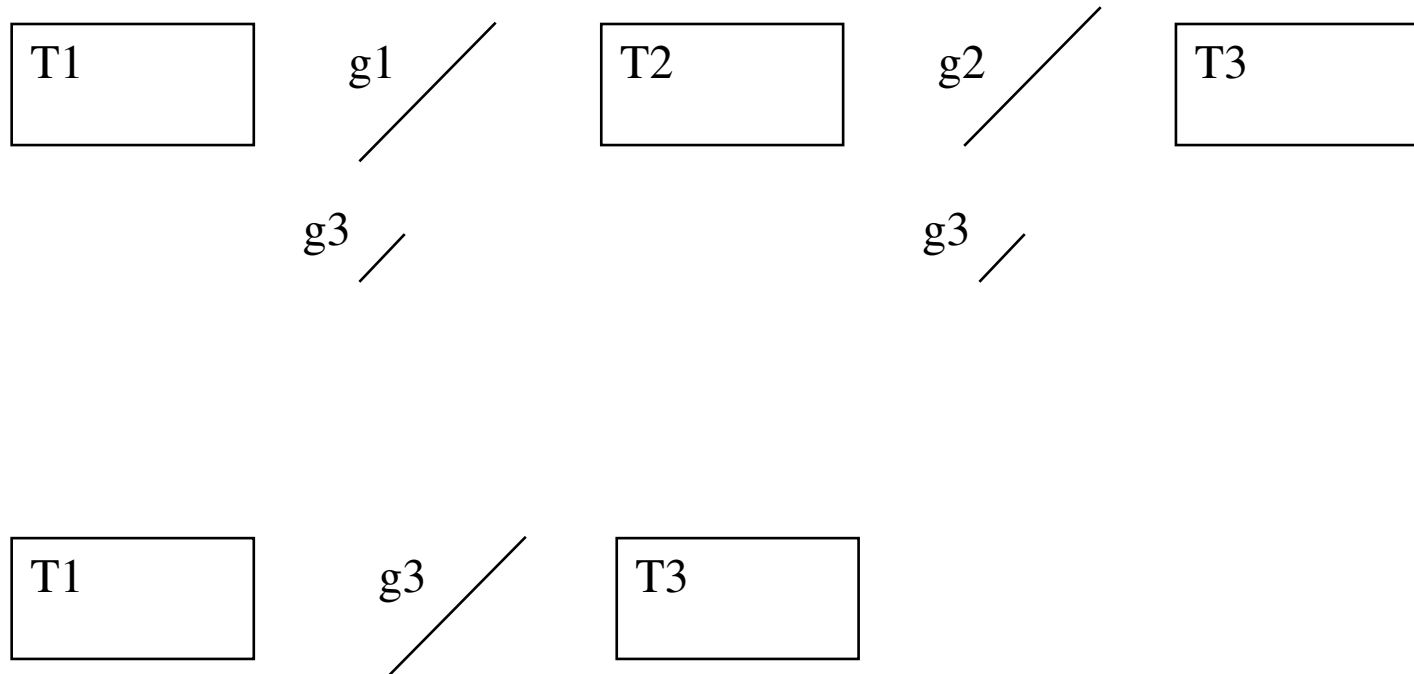
- (T1-T2)(T2-T3)

- In common
- Not in common



# Analysis: T-test

- The smaller changes: T1 vs. T3

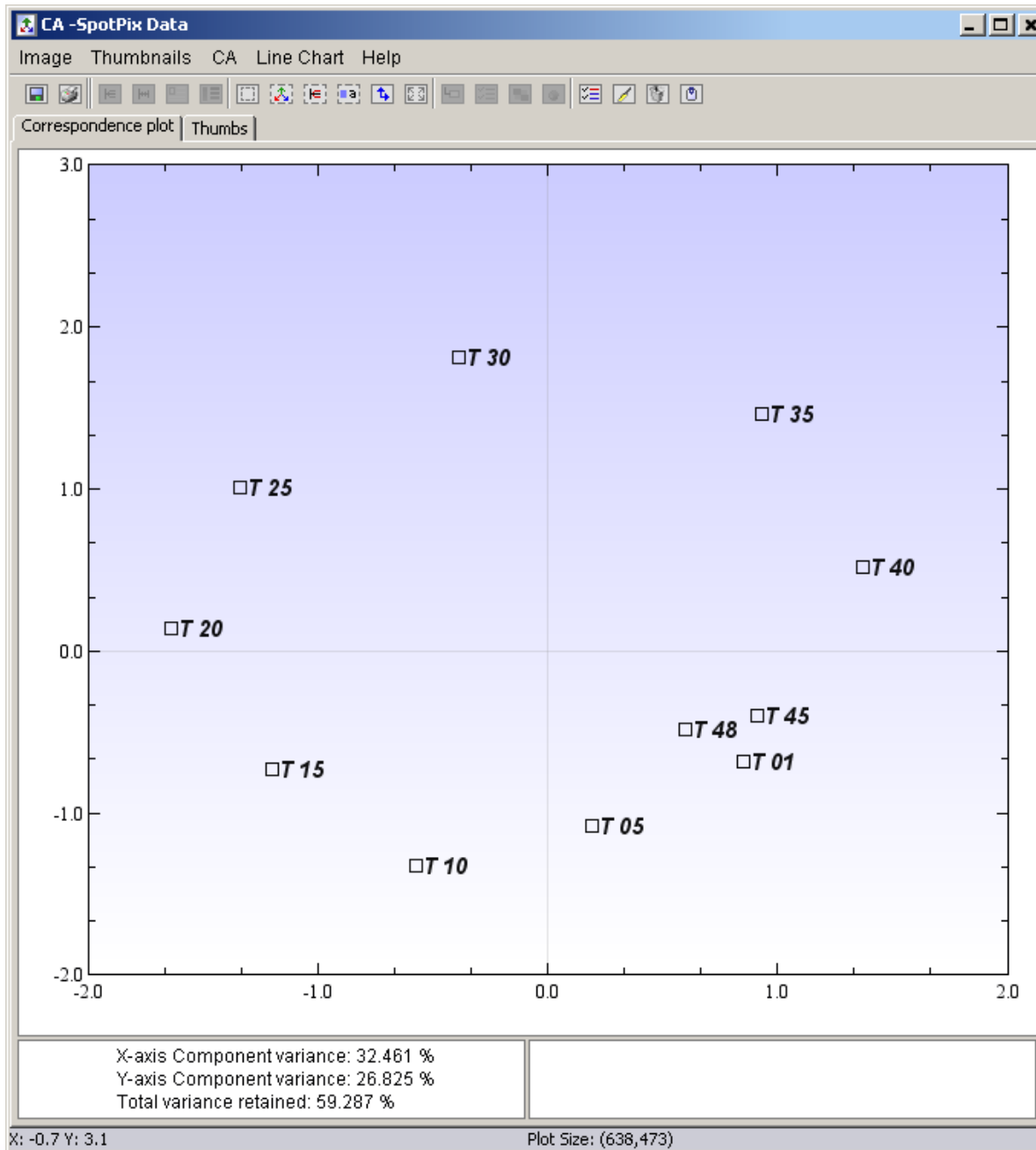


# Analysis: T-test

- If you have a large time series, it can be time consuming to do this
- Hard to select the ones to compare
- Solution:
  - Compare all consecutive pairs of time points and then make a profile based on this
  - Other types of analysis: PCA, PLS, ...

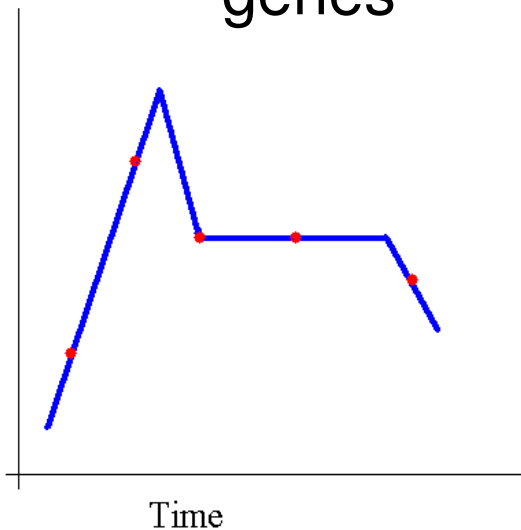
# PCA

- Projecting the genes or samples into (a low dimensionality) space that displays most of the variance in the (high dimensionality) data
  - i.e. PC1 and PC2



# Partial Least Square (PLS)

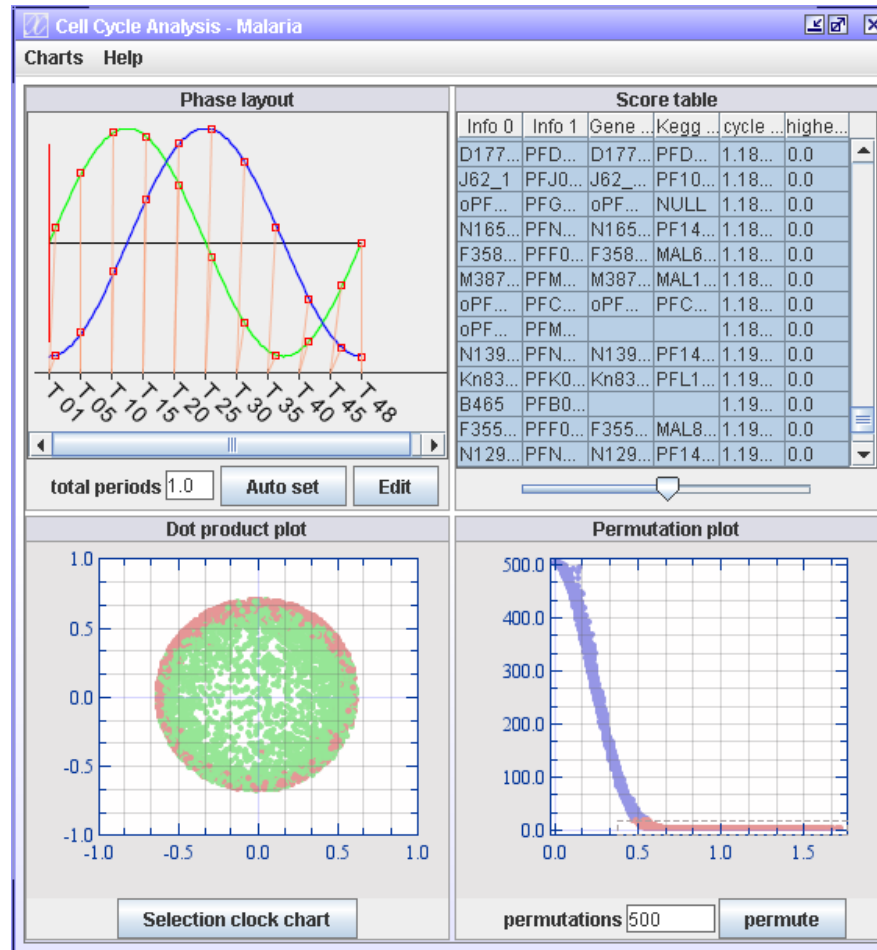
- Using external information to guide the projection
  - Instead of projecting the genes into a space that shows greatest variance, the genes are projected into a space that will help identify genes with particular characteristics, e.g. as below or cyclic genes



## Example: Cycling genes

- Expectation of cyclic genes
  - One top and one bottom within one cycle
  - Time between top and bottom is half a cycle
- Sine and cosine are known shapes that have these characteristics
- Use sine and cosine to search for genes that are correlated with these shapes
- The combination of the two allows us to identify all phase shifts of these shapes

# J-Express: Cellcycle



microarray.no

# GSEA with continuous profiles

- The genes are ranked based on their correlation to an interesting (continuous) profile
- After ranking the genes, looking for genesets that are overrepresented towards the top of the ranked list is the same as for differential expression

# GSEA with continuous profiles

- Use a gene profile from the dataset as the search profile
- Create a profile
  - Ascending or decending profile
  - Peak profile

# Summary

- Time series experiments are now affordable and within reach for more groups.
- Time series experiments are more complex
- Time series experiments demand more planning

# Acknowledgement

- Presentation made in collaboration with Ingrid Østensen (former NMC Oslo).