

Experimental Setup and Design

Kjell Petersen

Computational Biology Unit

microarray.no



Biological generality

Replicates

Batch

Biological generality

Replicates

Batch

microarray.no

Experimental Setup and Design

- Biological aspects
 - What is your hypothesis or question?
 - What else is known beforehand on the topic ?
- Technical aspects
 - Compare the right things to each other
 - Choice of platform
 - Avoid systematic errors
 - Ideal: each step, one person, one protocol, one day
 - Plan biological replicates
 - statistical significance of findings
 - choose results to validate
- What is the budget
 - Experimental and analysis costs

Biological vs technical replicates

- Generally we use biological replicates to answer biological questions and technical replicates to answer technical questions
- What is a biological replicate?
 - Are cell cultures that originates from the same biological source biological or technical replicates?
- What biological generality of your results are you aiming for?
- What is your long term plan ? Except for getting things published...

Comparing the right things

- Make sure there are no confounding effects
 - E.g. samples from patients vs control where all patients are old and all controls are young
- Biopsies should be taken from the same part of tissue
 - E.g. If one sample cuts through a major vein, differences seen could be due to different cell type compositions
- Choose the right model system
 - Do the cell types present express the genes of interest, e.g. not all treatments can be evaluated by blood samples, or in cell cultures

Choice of platform

- Choice of platform is closely linked to the biological question and has some implications on the rest of the design, so should be decided upon before doing the detailed planning of sampling and RNA extraction.
 - One channel vs two channel?
 - Natural pairing between samples to be compared?
 - Make sure key genes are on the array

Platform independent questions

- What is the difference you want to measure?
 - Do the samples reflect this ?
- Do we have to pool RNA to get enough for hybridisation?
- Do the arrays have spike probes and do we want to use them?

Avoid systematic errors

- Technical noise is added to the experiment every time you touch a sample at any step. It is important to control this, so that equal amounts of noise is added to all samples across the sample groups
- Common steps to control, include sampling, RNA extraction, labeling and hybridisation, but there may be others in your particular experiment
- Important to identify all steps that can add noise and that can be controlled.
- To control the steps you must also get to know your batches
 - Capacity per day
 - Persons involved
- The sample groups should generally balance across the batches
- Randomize the order of treatment within a batch

How to order

Bad

A1
A2
A3
A4
A5
B1
B2
B3
B4

Better

A1
B1
A2
B2
A3
B3
A4
B4
A5

Best

A1
B4
A3
B2
A2
B3
A5
B1
A4

microarray.no

General strategy

- For each step to control for systematic bias:
 - Distribute the biologic groups systematically in a balanced fashion: for two groups of equal size, every second from each group
 - Within each biological group, reshuffle the order of these samples
 - Divide it into roughly equal same size batches limited by your capacity for the step
- Tip: Colour code the sample names by biological group and the column next to it by batch

Experiment plan example

Biology	Sampling order	Extraction order
A1	A1	A4
A2	B4	B3
A3	A3	A5
A4	B2	B4
A5	A2	A1
B1	B3	B1
B2	A5	A3
B3	B1	B5
B4	A4	A2

microarray.no

Exp plan with batches

Biology	Sampling order	Extraction order
A1	A1	A4
A2	B4	B3
A3	A3	A5
A4	B2	B4
A5	A2	A1
B1	B3	B1
B2	A5	A3
B3	B1	B5
B4	A4	A2

How many replicates?

- How many do you have ready now?
 - Will you get more later?
- How large differences are you looking for?
- What is the expected expression difference of targeted biology in these samples ?
- Will “no change” be a desired significant result ?

Budget

- Pilot necessary?
 - Plan to balance batch effect if pilot runs should be used in final study
- Prioritize some groups/contrasts, and run more biological replicates ?
- Time is money also in academia, remember analysis costs.

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

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

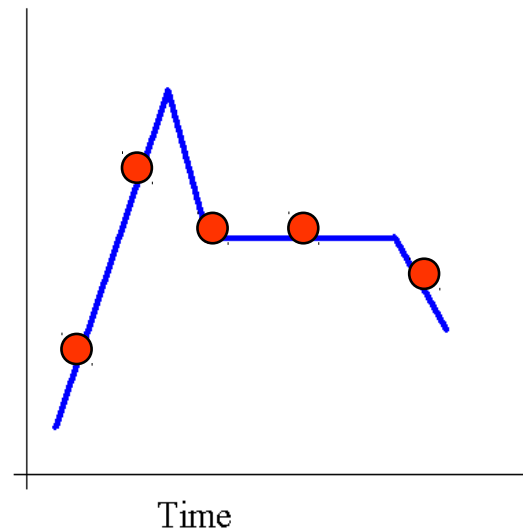
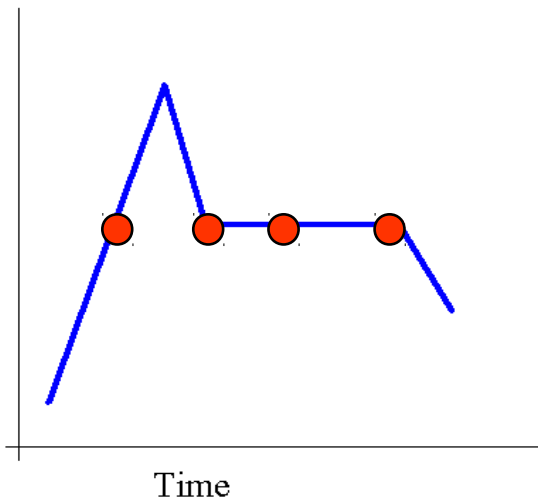
The direction of the response

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!

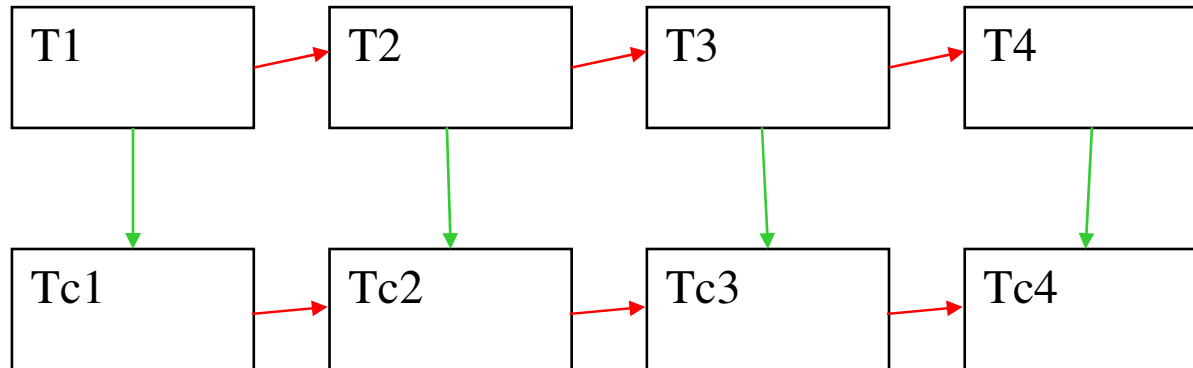
Ex.d.: Control

- 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



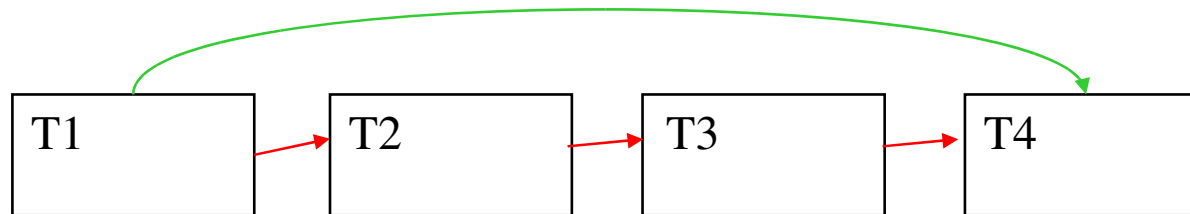
- **Vertically:** Differences due treatment, but you reduce the effect of the handling
- **Horizontally:** Differences due to treatment, time and some handling effect

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'



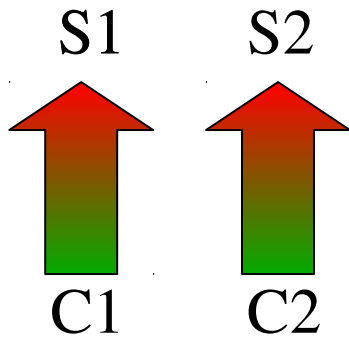
- Ok if no other time effects are expected, such as growth, phase of expression

Supplementary slides

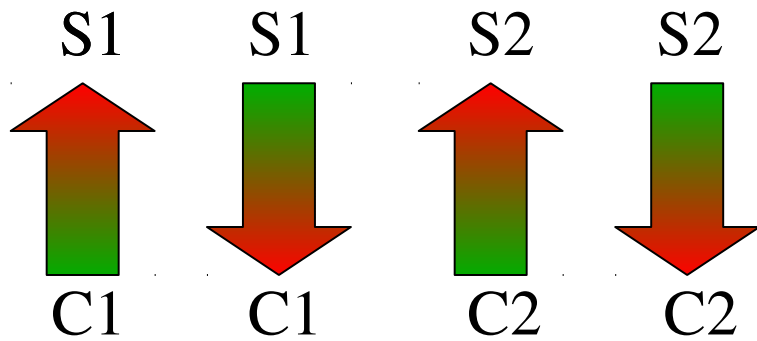
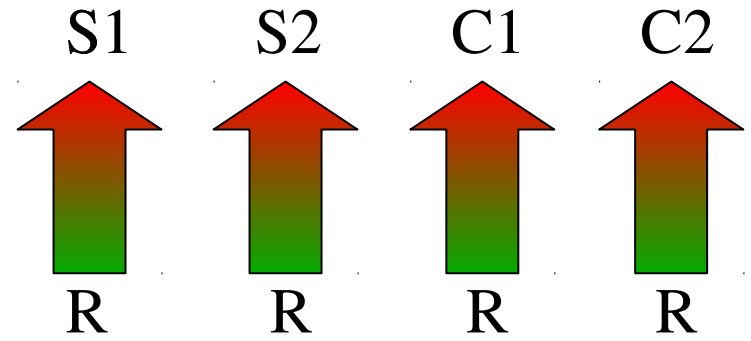
Additional questions for two channel platforms

- Are the samples naturally paired ?
 - Before and after treatment
- Direct comparison vs indirect ?
- Dye swaps ?
- Common reference design
 - Representable, independent RNA
 - Dye swaps strongly recommended
 - More noisy
 - More flexible

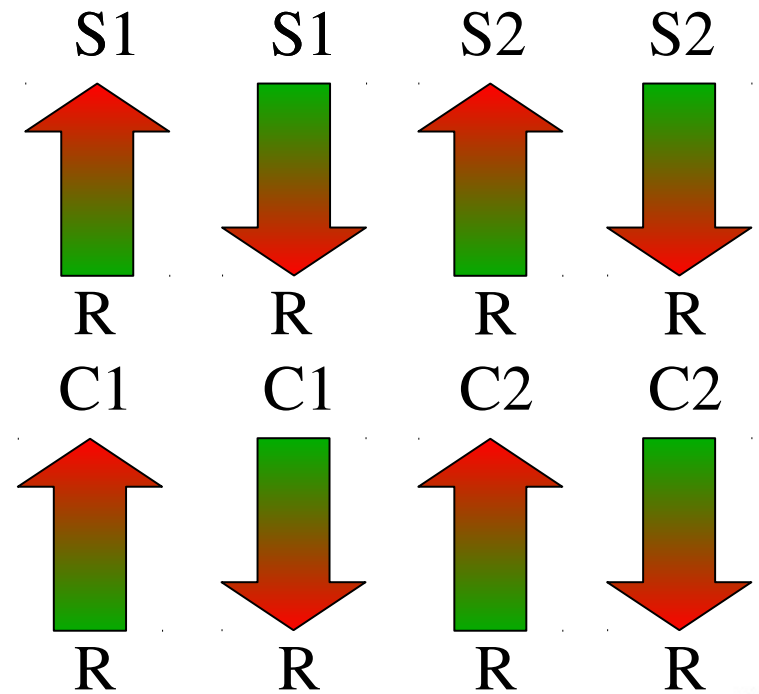
Direct



Indirect



Direct w/dye swap



microarray.no