

Experimental design

Solveig Mjelstad Olafsrud
Introduction to Microarray technology course
May 2011

Experimental design...

- The purpose of the experimental design is to **plan the experiment** in a way that makes sure it can **answer your biological question**
- The experimental design is documented in the **experimental plan**

What is an experimental plan?

- A **written document** explaining the purpose of the experiment, and **how** it should be performed
- Think of the experimental plan as answering the old journalists requirements of 5 Ws and 1 H: **who, what, when, where, why** and **how**



Biological aspects (1)

- What is your **hypothesis** or **question**?
 - ✓ Does the samples reflect this?
- What else is **known beforehand** on the topic?
 - ✓ Expression levels on known interesting genes available?
- What **samples** are available?
 - ✓ More samples at a later time?
 - ✓ All biopsies should be taken from the **same part** of the tissue!
 - ✓ Do you have enough RNA from each sample or is pooling of samples required?



Biological aspects (2)

- Choose the **right model system**
- **Comparing the right things**
 - ✓ Does the cell type express the genes of interest?
 - ✓ Make sure there are no confounding effects



Biological aspects (3)

- **Replicates**
 - ✓ 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?
- **Biological vs technical replicates**
 - ✓ 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?



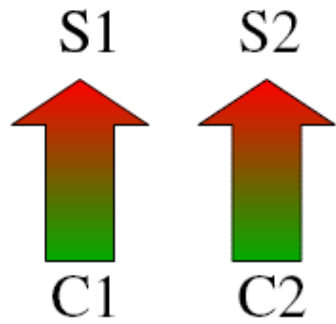
Technical aspects (1)

- Choice of **platform**
 - ✓ Closely linked to biological question
 - ✓ Should be decided upon before doing the detailed planning of sampling and extraction
 - ✓ One channel vs two channel?
 - One channel – one array – one sample
 - Two channel – one array – two samples
 - ✓ Natural pairing between samples to be compared?
 - ✓ Make sure key genes are on the array!

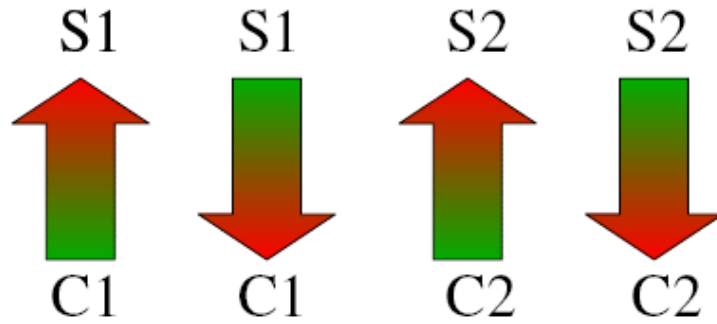
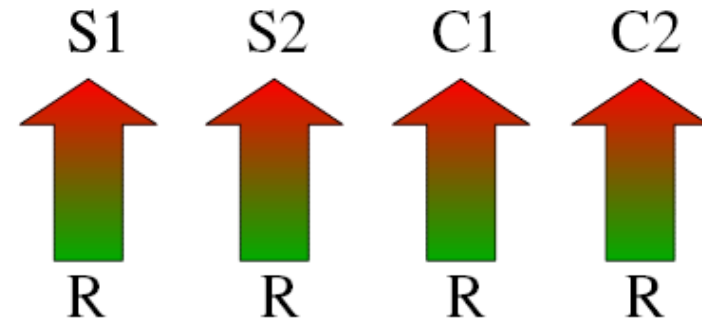
Technical aspects (2)

- Additional questions for **two channel platform**:
 - ✓ 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

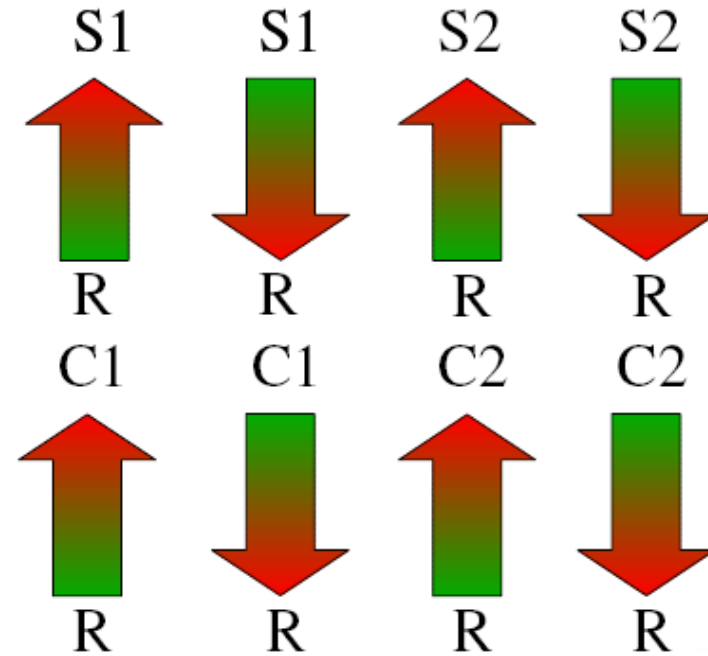
Direct



Indirect



Direct w/dye swap



Technical aspects (3)

- Avoid **systematic errors**
 - ✓ Technical noise added to experiment at each step
 - ✓ **Ideal: each step, one person, one protocol one day**
 - ✓ Important to identify these steps in order to be able to control them
 - Know your batches!
 - ✓ Sample groups should be **balanced** across the batches
 - ✓ **Randomize** the order of treatment within a batch
 - ✓ Use the **same batch** of arrays and reagents **within an experiment**
- Plan **biological replicates**
 - ✓ Statistical significance of findings
 - ✓ Choose results to validate

Physical aspects

- Limitations to the **equipment**
 - ✓ What is the limiting point?
- Get to know your **bathces**
 - ✓ How many samples are possible to do at the same time at the different steps?
- How many **people** are/should be involved?
 - ✓ Who does what?



Economical aspects

- What is the **budget**?
 - ✓ Experimental costs
 - ✓ Analysis costs
 - ✓ Time costs!
- **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?



General strategy

- For each step to control for systematic bias:
 - ✓ **Distribute** the biological groups systematically in a **balanced fashion**: for two groups of equal size, every second from each group
 - ✓ Divide it into roughly **equal size batches** limited by your capacity for the step
- Tip: In excel (or similar program) colour code the sample names by biological group and the column next to it by batch
- **Randomization and balancing** with respect to biology of interest: Possible to separate technical variation from biological variation



How to order

Bad

A1
A2
A3
A4
A5
B1
B2
B3
B4
B5

Better

A1
B1
A2
B2
A3
B3
A4
B4
A5
B5

Best

A1
B4
A3
B2
A2
B3
A5
B1
A4
B5

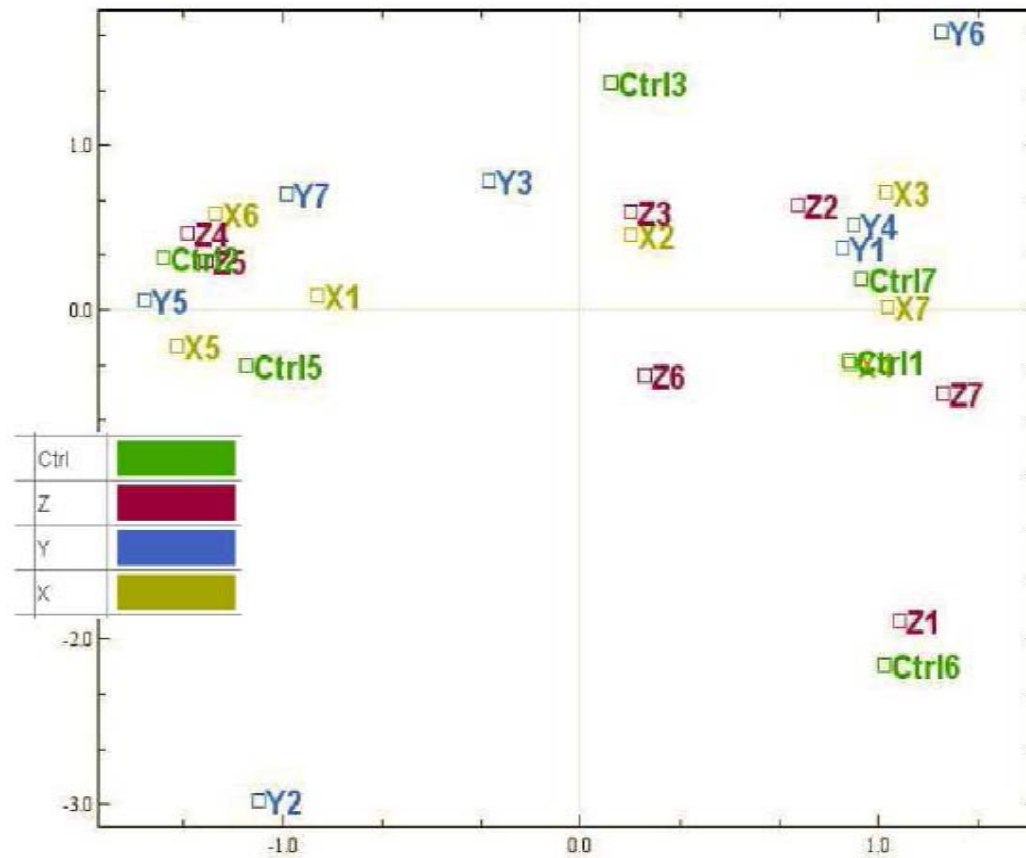
Experimental plan: an example

Biology
A1
A2
A3
A4
A5
A6
B1
B2
B3
B4
B5
B6
C1
C2
C3
C4
C5
C6

Biology	Sampling order
A1	1
B4	2
C2	3
A3	4
B6	5
C4	6
A5	7
B2	8
C6	9
A2	10
B3	11
C1	12
A4	13
B5	14
C3	15
A6	16
B1	17
C5	18

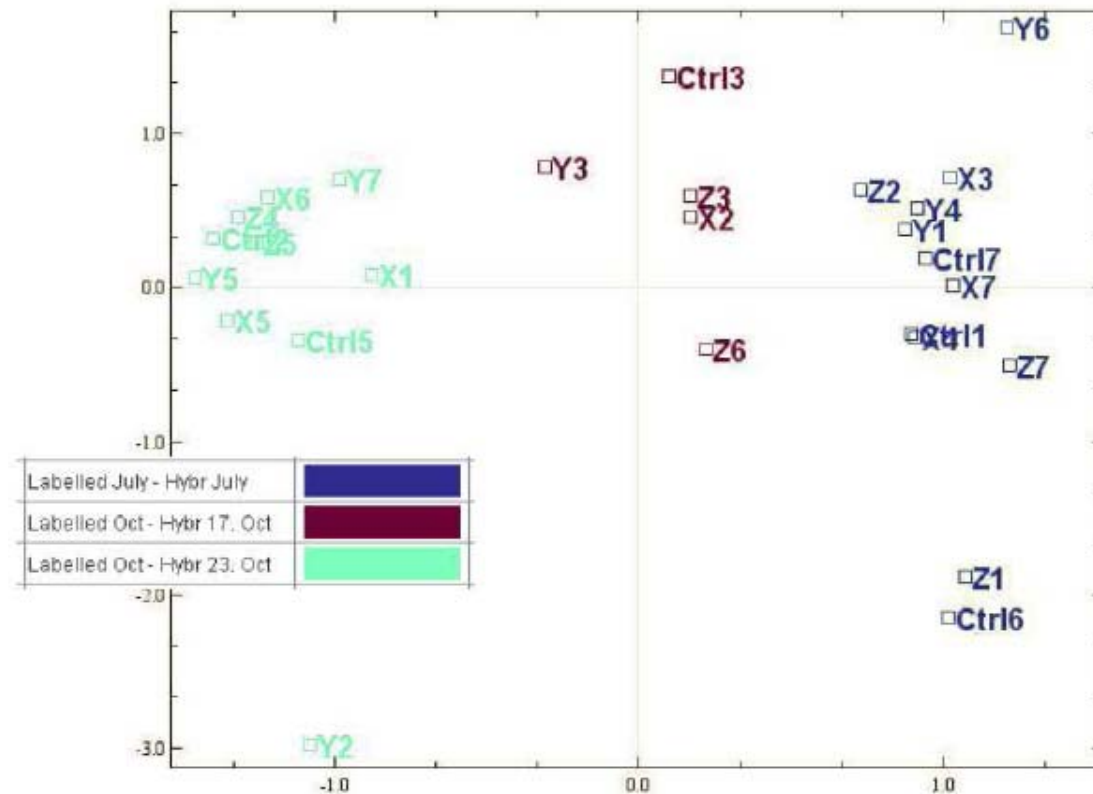
Biology	Sampling order	Extraction order
A2	10	1
B6	5	2
C1	12	3
A5	7	4
B5	14	5
C6	9	6
A6	16	7
B4	2	8
C5	18	9
A3	4	10
C3	15	11
B2	8	12
A4	13	13
C4	6	14
B1	17	15
A1	1	16
B3	11	17
C2	3	18

Batch effect: an example (1)



Samples labeled according to **biology**

Batch effect: an example (2)



Samples labeled according to **labelling date**

Summary

- Plan ahead!
- Randomization and balancing
- Write it down in an Experimental plan
- Follow the experimental plan!

Practical exercise 1

- You have 16 animals, 8 treated (T1-8) and 8 controls (C1-8)
- Your sampling capacity is 5 animals per day
- In what order will you do the sampling?

Practical exercise 2

- 16 samples, 8 treated(T1-T8), 8 controls (C1-C8)
- Can extract 12 samples at the same time because of centrifuge capacity
- How would you organize the extraction?

Practical exercise 3

- The same samples will be hybridized to the arrays, one sample per array
- You have 20 arrays available, 12 from batch A, 8 from batch B
- How will you assign the samples to arrays?

