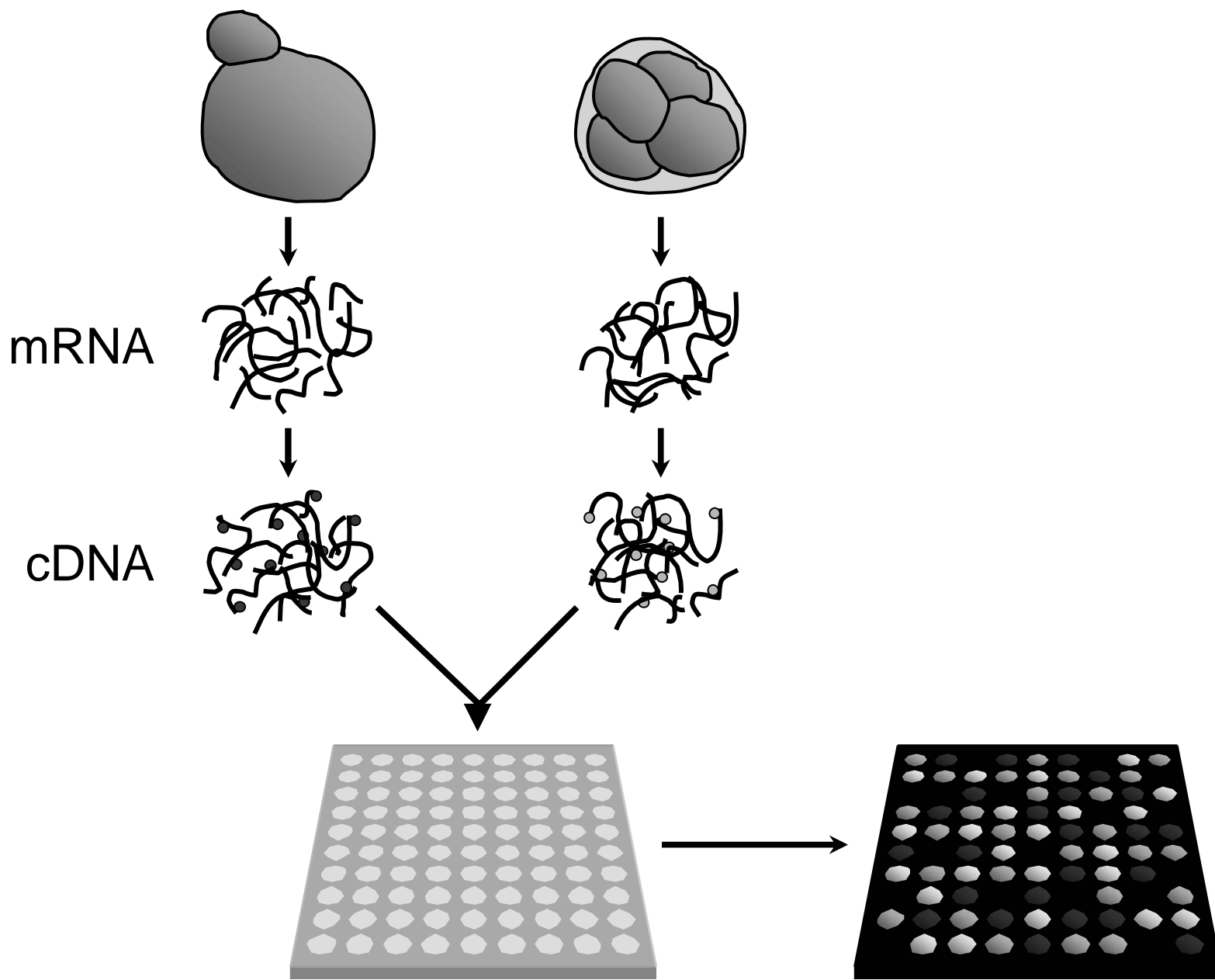


Experimental Design Considerations for Microarrays

Kathleen Kerr
Department of Biostatistics
University of Washington



mRNA

cDNA

DNA microarray

Re-created from Brown and Botstein,
Nature Genetics Supplement, 1999

The “design” of an experiment is

1. The samples selected for comparison
2. The specification of the units to which the samples will be applied
3. The rules by which the samples are allocated to the experimental units
4. The specifications of the measurements to be made

The “design” of an experiment is

1. The samples selected for comparison
2. The specification of the units to which the samples will be applied
3. The rules by which the samples are allocated to the experimental units
4. The specifications of the measurements to be made

Replication

First level of replication:

Genes spotted multiple times per array

Second level of replication:

Multiple arrays to study the same samples
contrast with:

Replication in the classical sense:

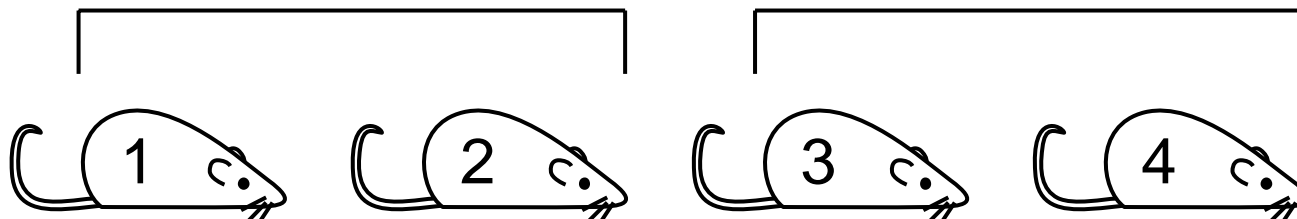
Random sampling of individuals from the populations of interest or randomly assigning individuals to treatment groups. Without this kind of replication, inference is limited to the particular RNAs in the study.

Treatments:

A

B

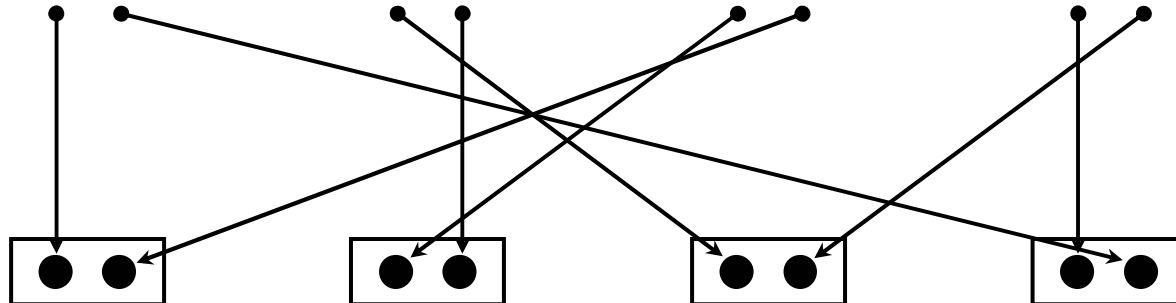
Replicates:



Dyes:

R G R G R G R G

Arrays:



Replication

Measurement
Error

First level of replication:

Genes spotted multiple times per array

Second level of replication:

Multiple arrays to study the same samples

Biological
Variability

Replication in the classical sense:

Random sampling of individuals from the populations of interest or randomly

assigning individuals to treatment groups.

Without this kind of replication, inference is limited to the particular RNAs in the study.

Replication

Measurement
Error

First level of replication:

Genes spotted multiple times per array

Second level of replication:

Multiple arrays to study the same samples

replication: repeated measures? subsampling?

Biological
Variability

Replication in the classical sense:

Random sampling of individuals from the

populations of interest or randomly

assigning individuals to treatment groups.

Without this kind of replication, inference is

limited to the particular RNAs in the study.

Replication

Pool RNAs?

Pooling RNAs reduces biological variability. It also eliminates the ability to see that variability.

Replication

Pool RNAs?

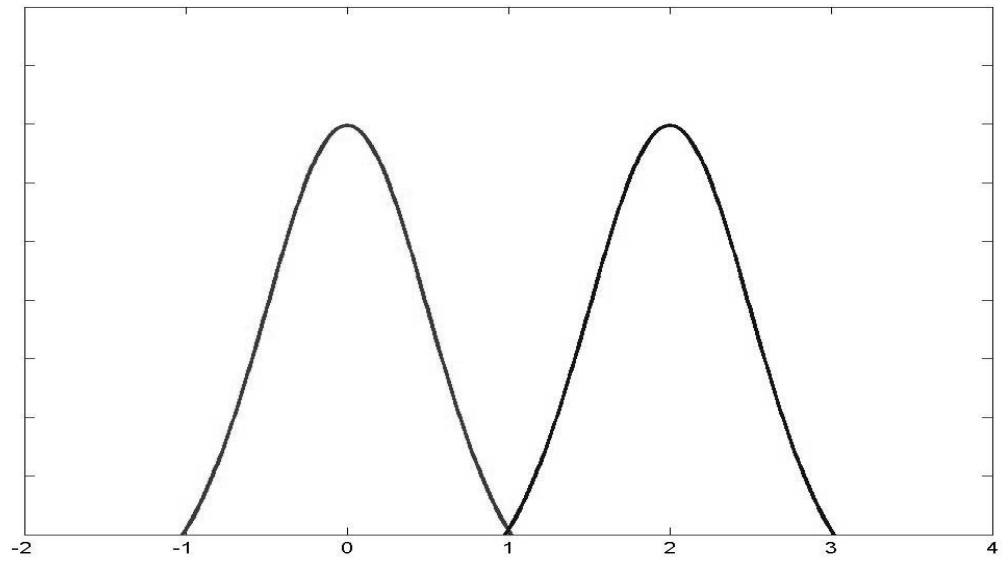
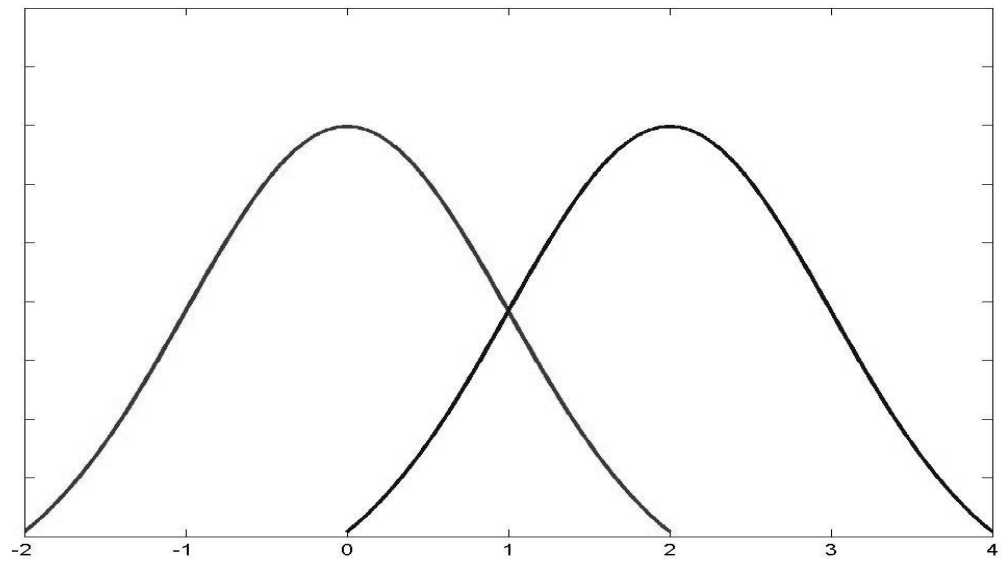
Pooling RNAs reduces biological variability. It also eliminates the ability to see that variability.

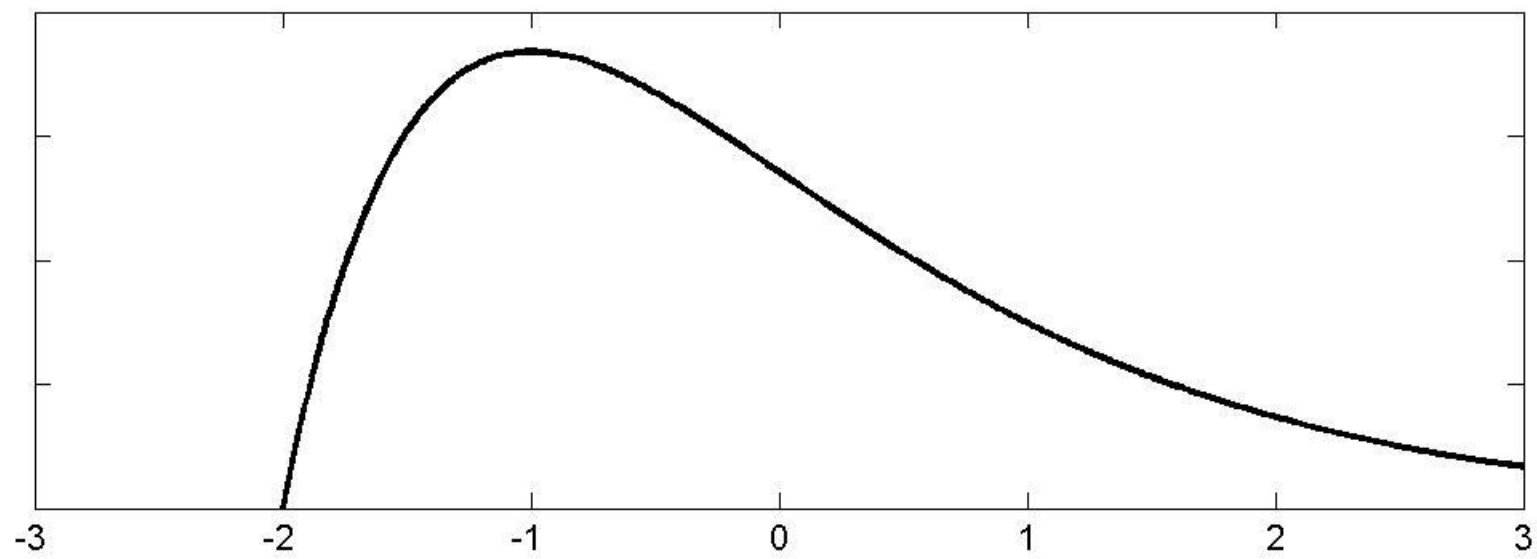
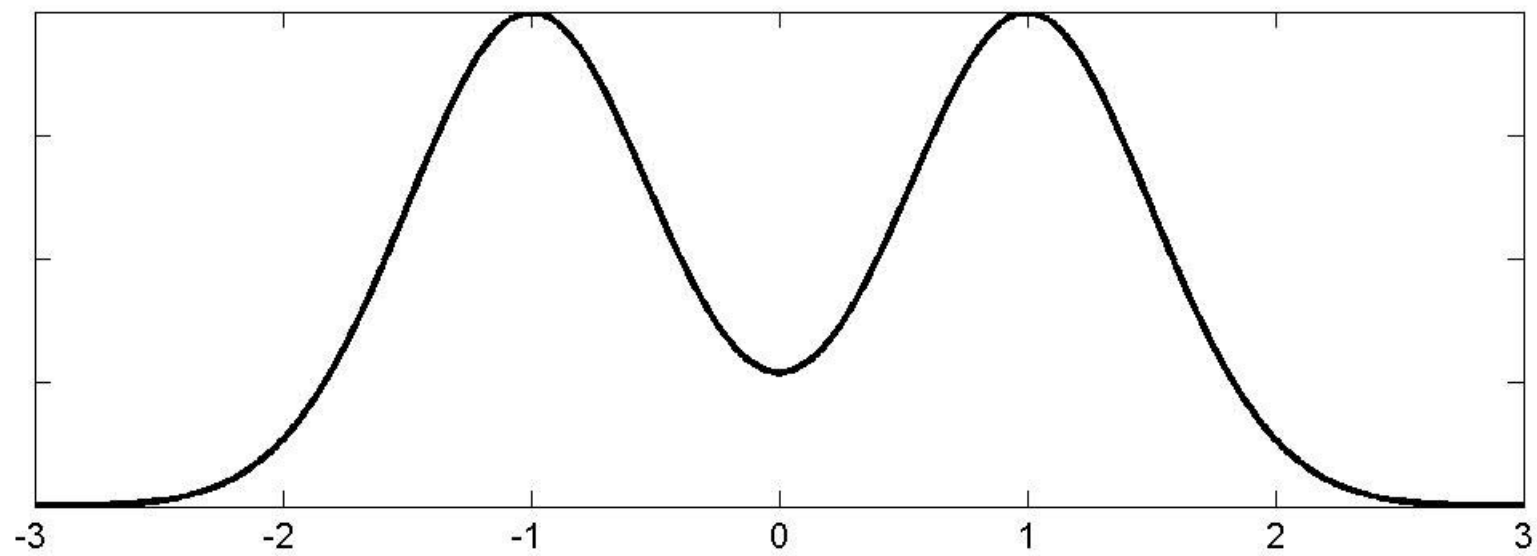
Appropriate

If the object of interest is the difference in means.

Inappropriate

In general, biological variability is important. For many analyses such as finding predictors of disease and discrimination analyses, understanding this variability is crucial.





Replication

Pool RNAs?

Pooling RNAs reduces biological variability. It also eliminates the ability to see that variability.

Appropriate

If the object of interest is the difference in means.

Inappropriate

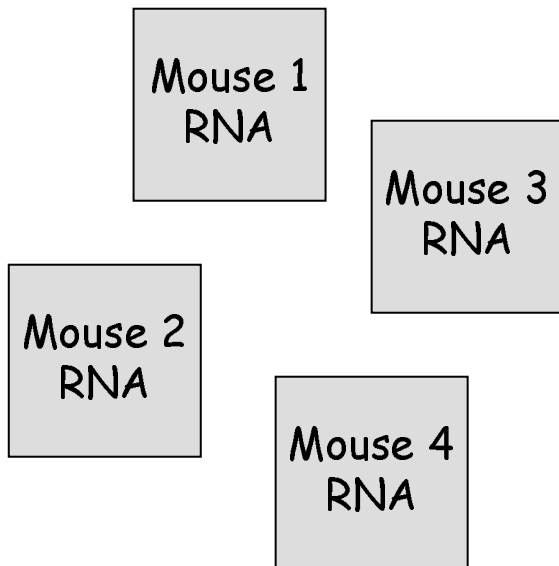
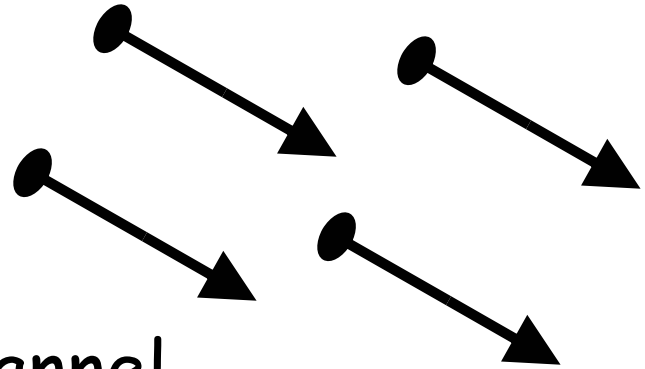
In general, biological variability is important. For many analyses such as finding predictors of disease and discrimination analyses, understanding this variability is crucial.

The “design” of an experiment is

1. The samples selected for comparison
2. The specification of the units to which the samples will be applied
3. The rules by which the samples are allocated to the experimental units
4. The specifications of the measurements to be made

Representation of Microarray Designs

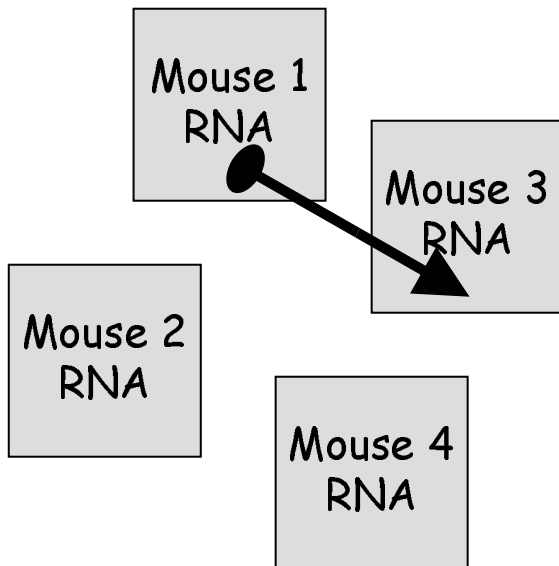
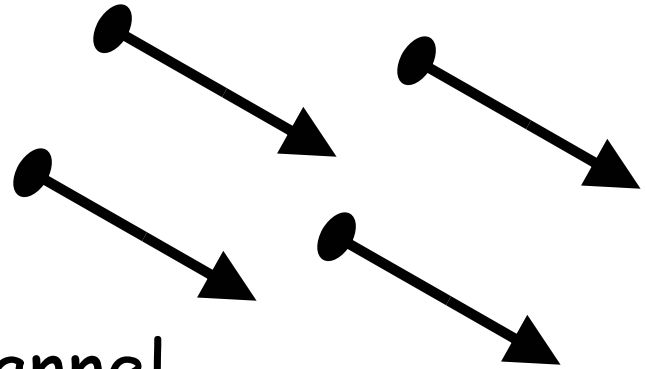
Microarrays are represented by arrows, where one end is the "red" channel and the opposite end is the "green" channel



RNA samples are represented as rectangles or circles

Representation of Microarray Designs

Microarrays are represented by arrows, where one end is the "red" channel and the opposite end is the "green" channel



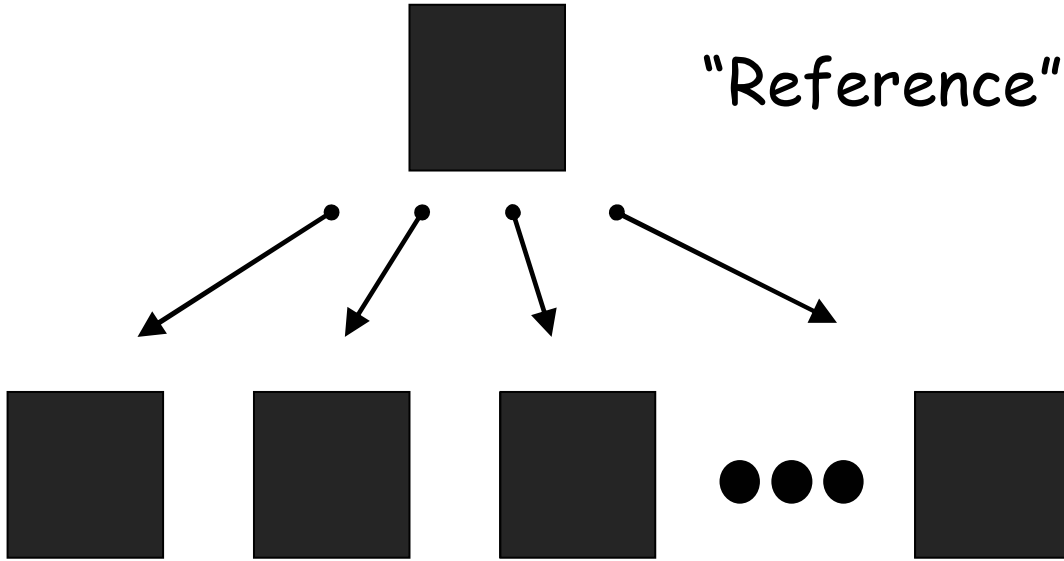
RNA samples are represented as rectangles or circles

Microarrays compare samples assayed on the same array. Samples not assayed together can be compared indirectly if and only if the design is “connected*.” In statistical design, this makes microarray designs *incomplete block designs*.

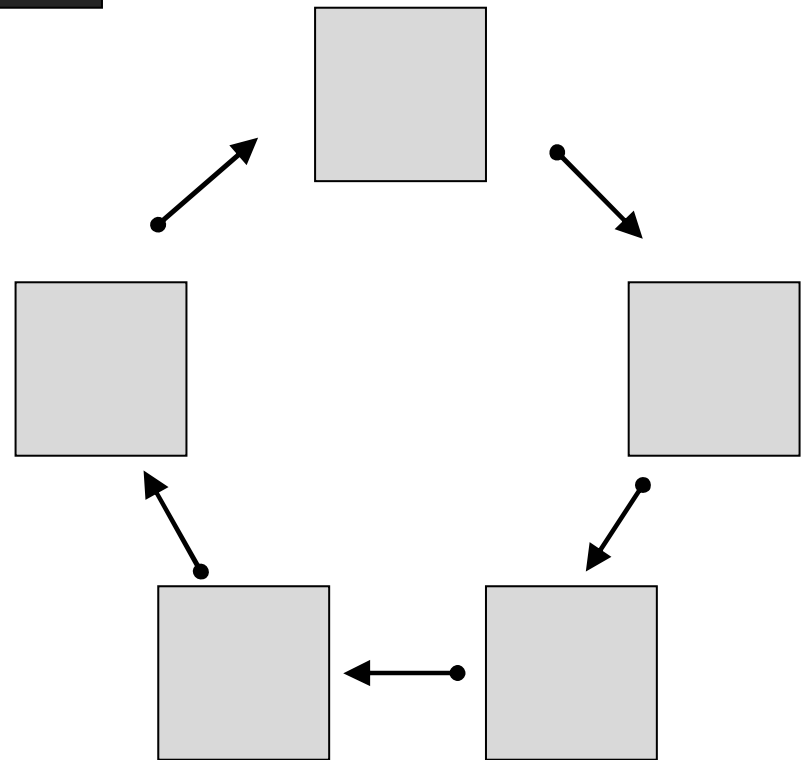
*Exception: Unless spot variability is modeled. See Jin, Riley, Wolfinger, White, Passador-Gurgel, Gibson, *Nature* 2001.

Kerr & Churchill, *Biostatistics*, 2001.

"Reference" Design



"Loop" Design



A-Optimality

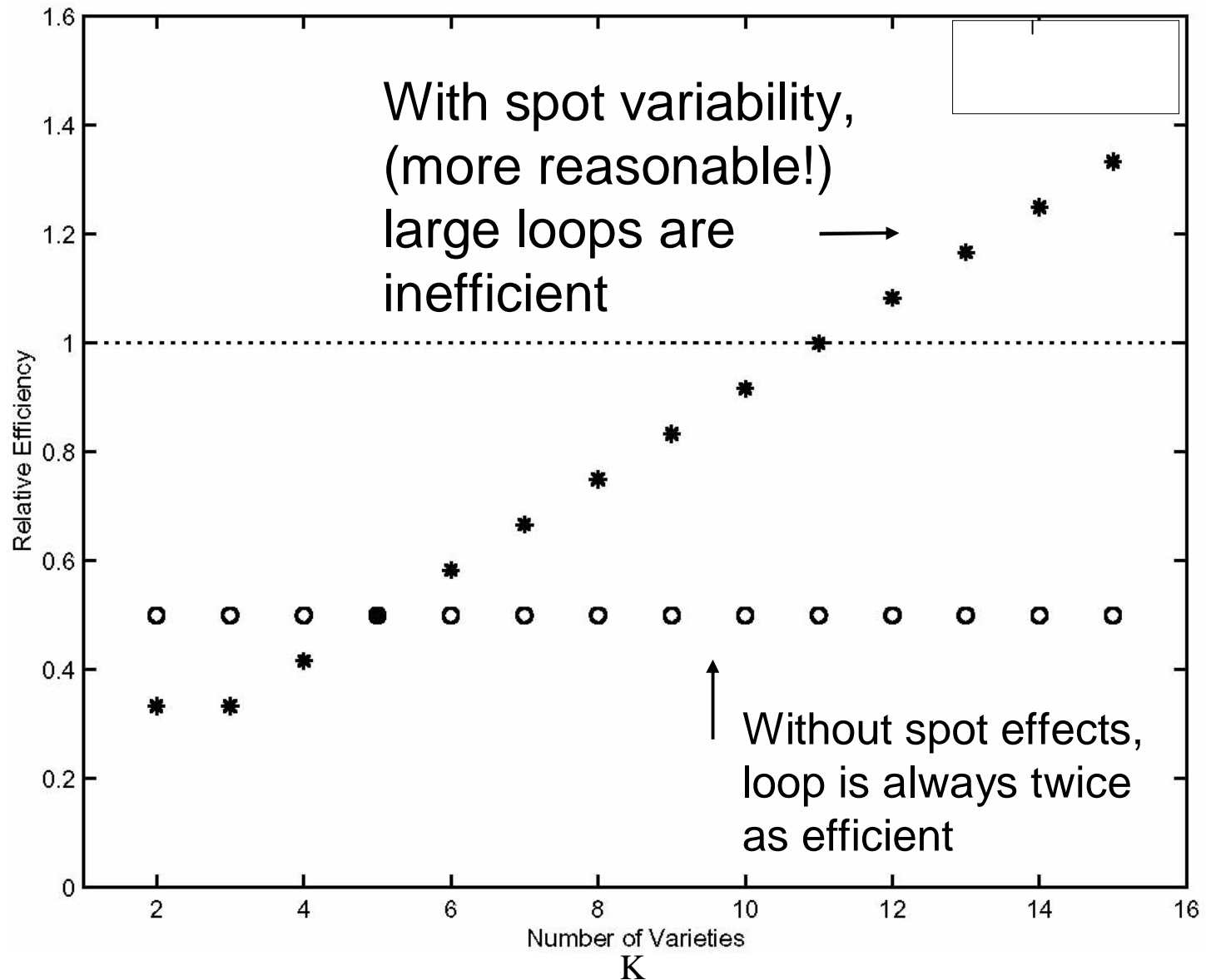
For a given number of RNAs and a given number of arrays, the A-optimal design minimizes the average variance over all pairwise comparisons of RNAs.

(Just one possible criterion, may or may not be appropriate.)

For K samples of interest, the loop and reference designs each use K arrays.

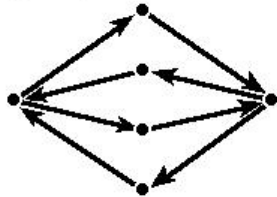
How do these two design strategies compare according to the A -optimality criterion?

Relative Efficiency

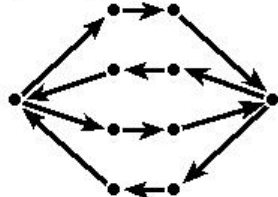


A-Optimal Designs

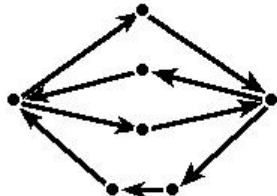
$v = 6$



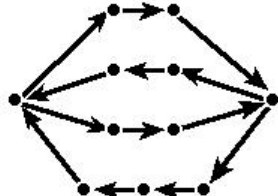
$v = 10$



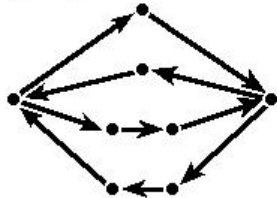
$v = 7$



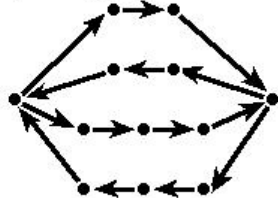
$v = 11$



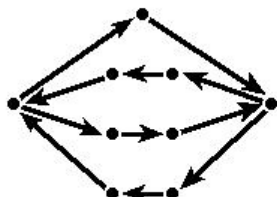
$v = 8$



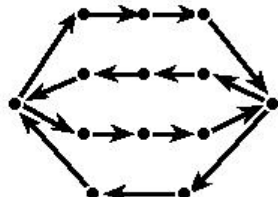
$v = 12$



$v = 9$



$v = 13$

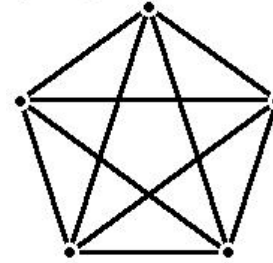


For v varieties, these designs are A-optimal among all even designs using $v+2$ arrays.

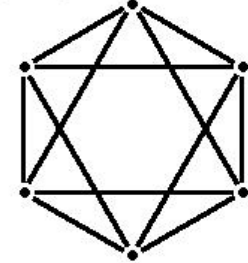
A-Optimal Designs

For v varieties, these designs are A-optimal among all designs using $2*v$ arrays.

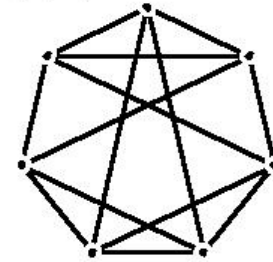
$v = 5$



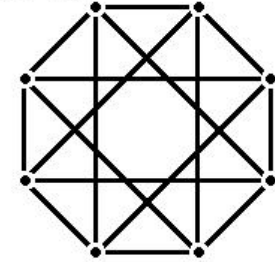
$v = 6$



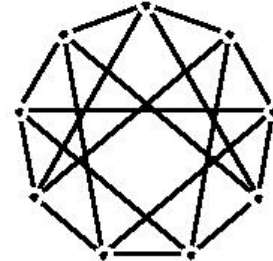
$v = 7$



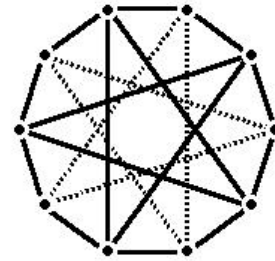
$v = 8$



$v = 9$



$v = 10$



Generally, we want efficient designs for an appropriate criterion but we should consider other aspects of good design:

- Robust properties
- Extendibility
- Simplicity of execution
- Useful sub-designs

- Robust properties

What happens to the efficiency if we lose an array? If we lose some spots on every array?

- Extendibility

What if we decide to add more samples to the study later?

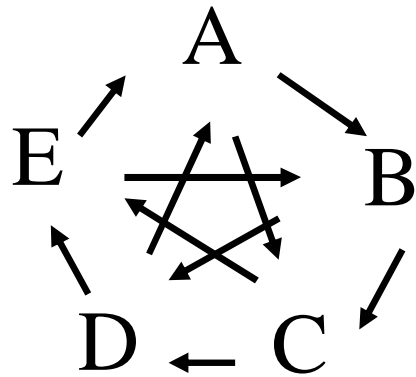
- Simplicity of execution

Will we be able to keep track of the assays we need to do?

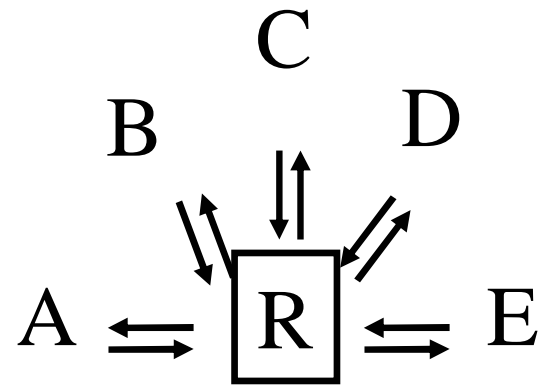
- Useful sub-designs

What if we want to analyze the data on just a subset of samples?

Robustness



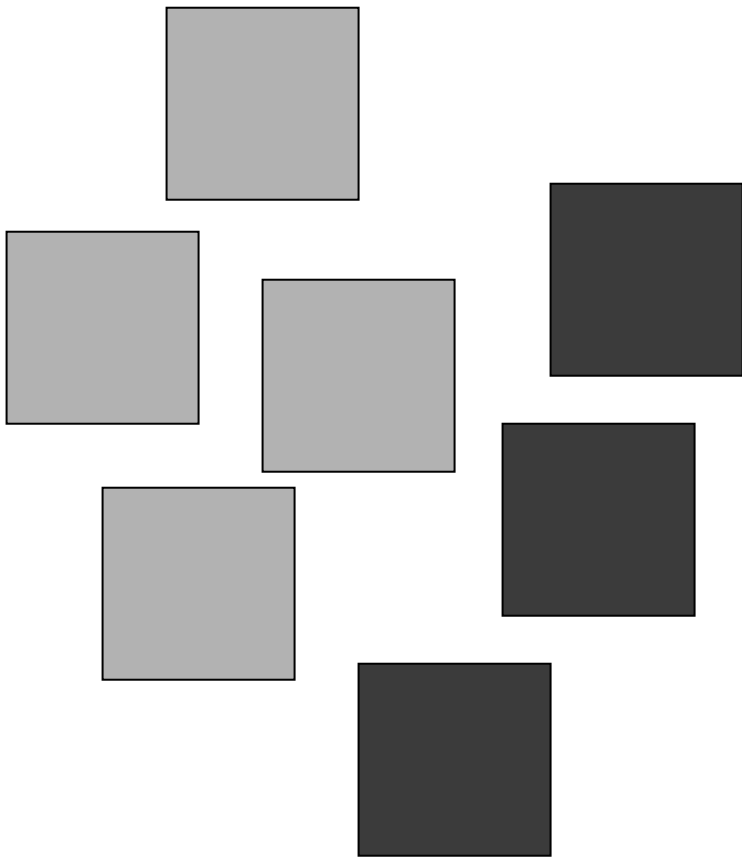
"Double Loop"



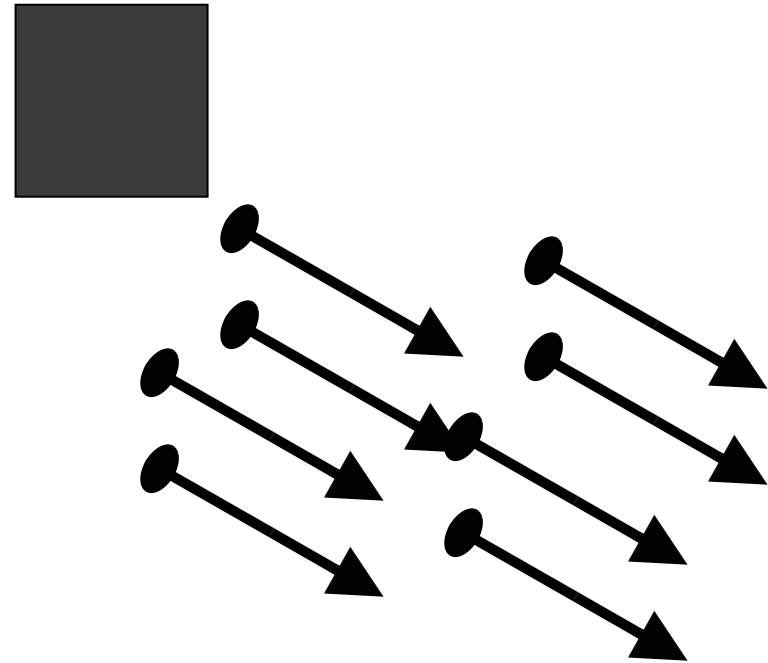
"Double Reference"

The relative efficiency of the double reference design compared to the double loop design for comparing A,B,C,D,E is 40%. The double-loop also has an advantage in robustness because there are so many more "connections" between any pair of samples.

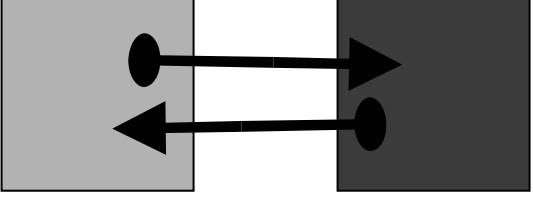
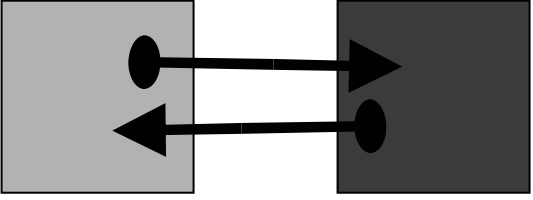
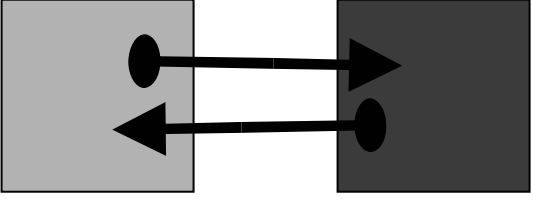
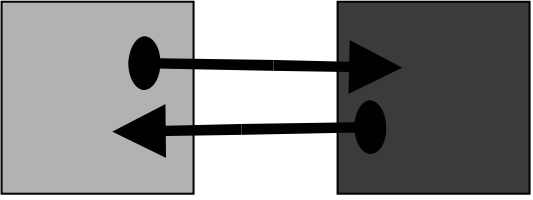
Example: Comparing
2 “Treatments”

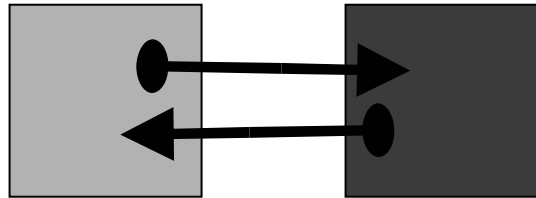
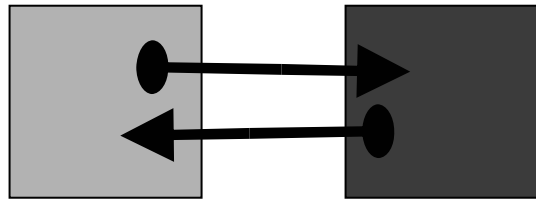
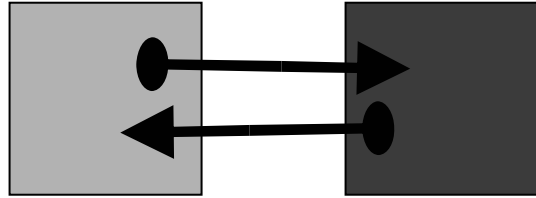
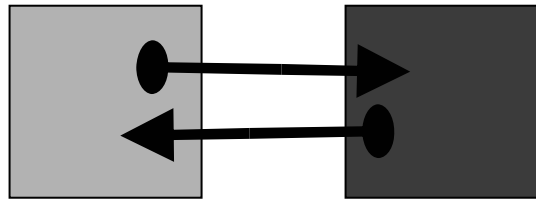


Replication: N mice from
each of two groups



Limited number
of arrays, say 2N



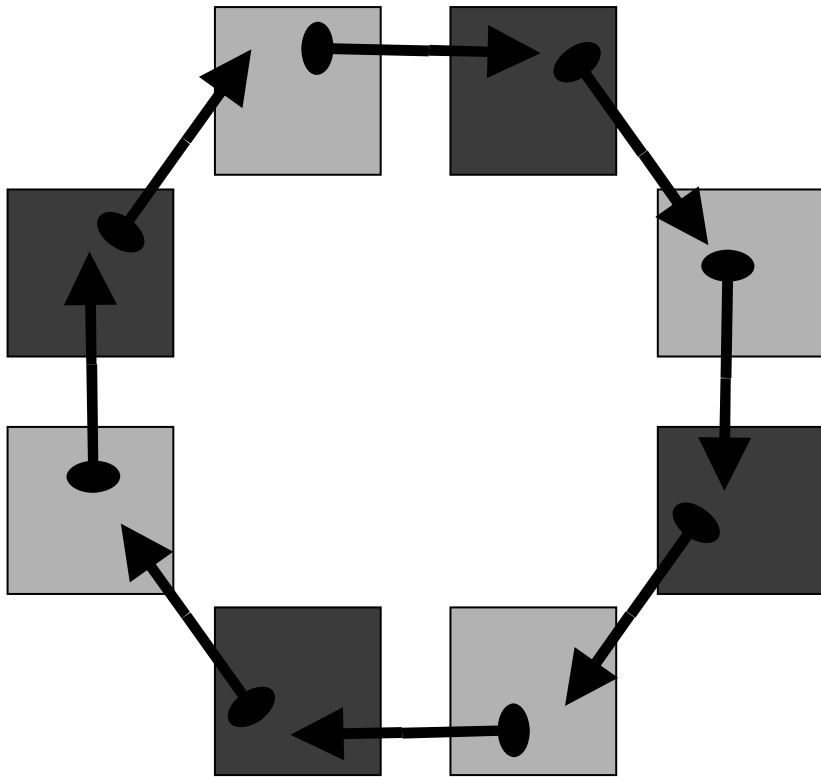


N mice in each group

N “dye-swap” assays

2N arrays

$$\text{Var}(\overline{T_{mt}} - \overline{C_{tl}}) = \sigma^2/N$$

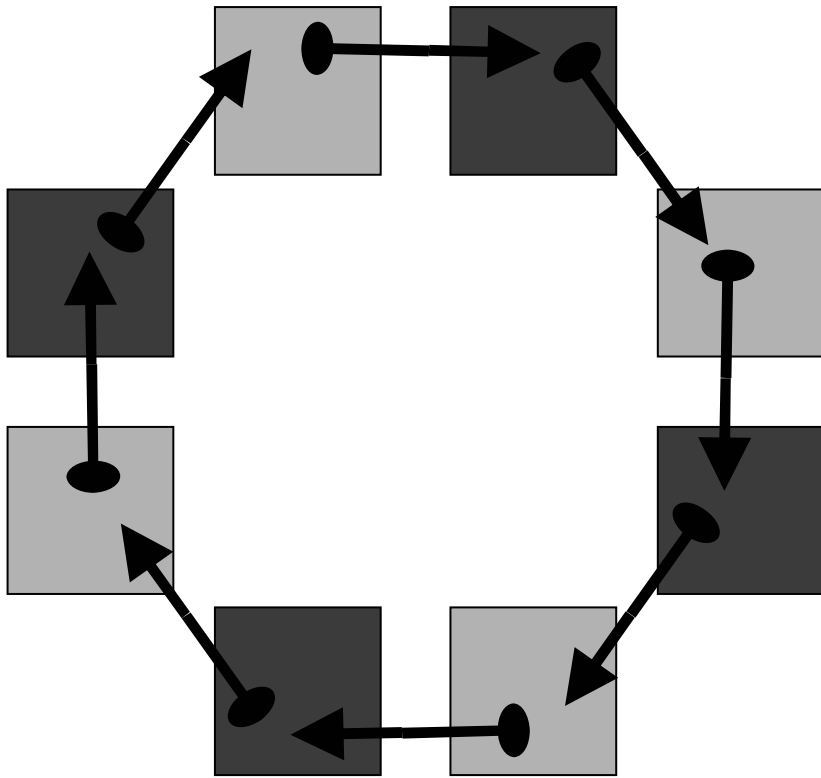


N mice in each group

2N “looped” assays

2N arrays

$$\text{Var}(\overline{T_{mt}} - \overline{C_{tl}}) = \sigma^2/N$$



The same precision for
comparing the groups.
Is anything gained?

The ability to measure
biological variability.

Case Example

A type of transgenic mouse lives ~30% longer than wild-type (wt) mice.

An investigator wants to study gene expression differences in mutant mice from two founder lines, using young and old mice and looking at four different tissues.

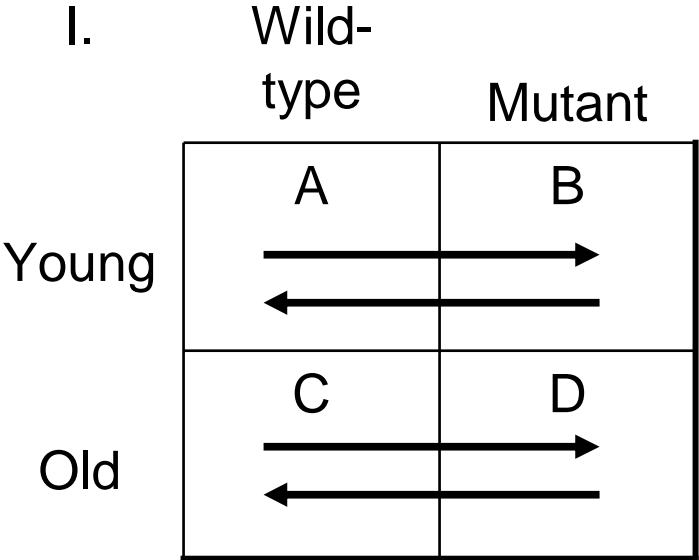
- Genotype (2) – mutant and wild-type
- Age (2) – young and old
- Founder (2)
- Tissue (4)

The most interesting comparison is between genotypes, but this is not the only interesting comparison.

	Wild-type	Mutant
Young		
Old		

x {Founder 1,2} x {Tissue A,B,C,D},

Original Design

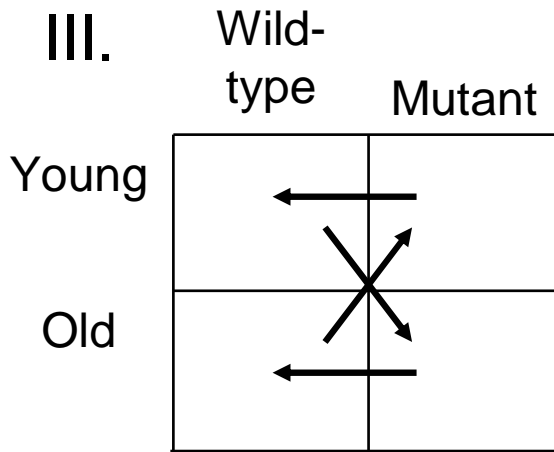
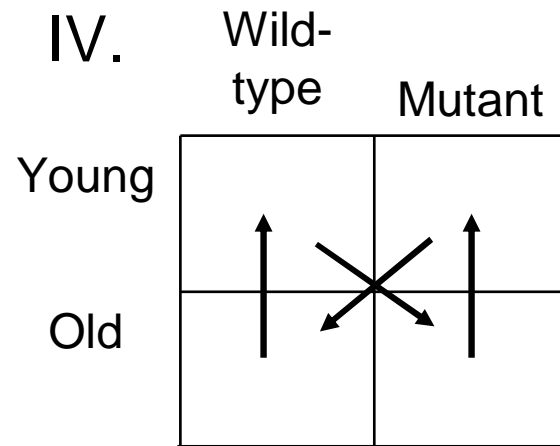
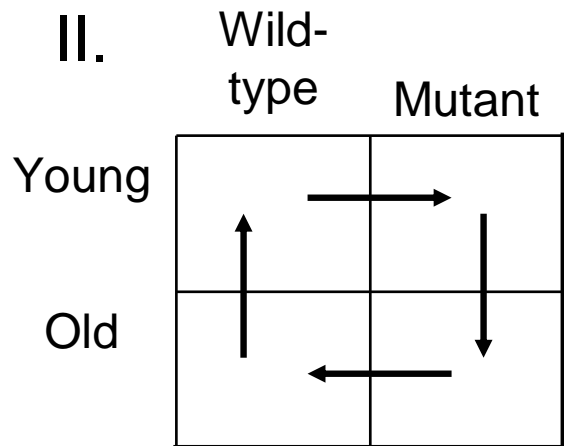


X {Founder 1,2} X {Tissue A,B,C,D},

Variance of Comparisons*			
GEN within AGE A-B or C-D	GEN across AGE $\frac{1}{2}(A+C)-$ $\frac{1}{2}(B+D)$	AGE within GEN A-C or B-D	AGE across GEN $\frac{1}{2}(A+B)-$ $\frac{1}{2}(C+D)$
1	$\frac{1}{2}$	Not possible	Not possible

*The numbers are only relative, to be used to compare designs.

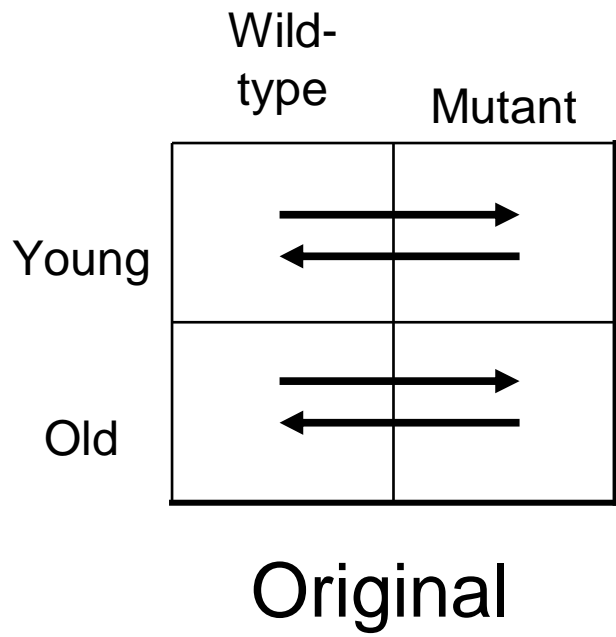
Alternative Designs

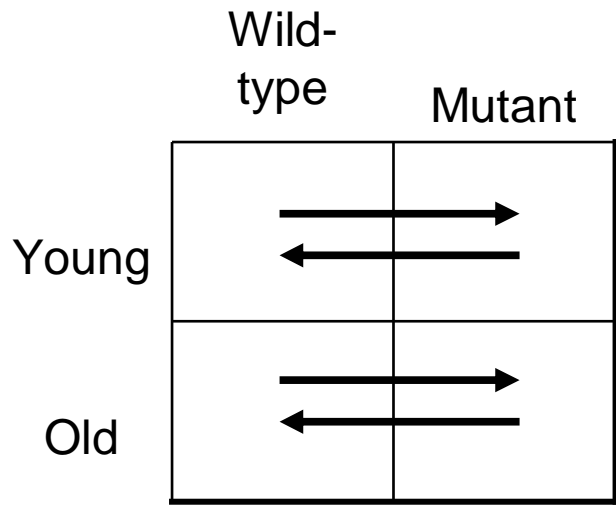


x {Founder 1,2} x {Tissue A,B,C,D},

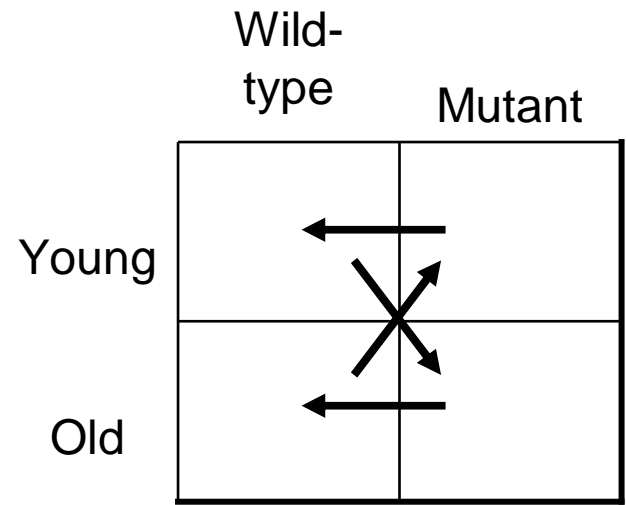
Variance Comparisons

Design	GEN within AGE	GEN across AGE	AGE within GEN	AGE across GEN
I.	1	$\frac{1}{2}$	Not possible	Not possible
II.	1.5	1	1.5	1
III.	1.5	$\frac{1}{2}$	2	1
IV.	2	1	1.5	$\frac{1}{2}$

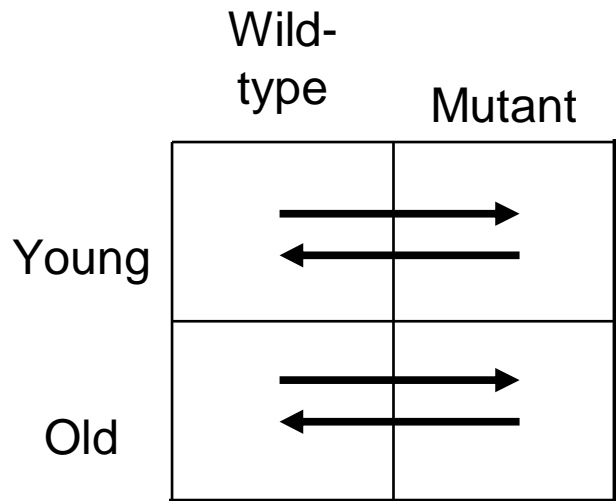




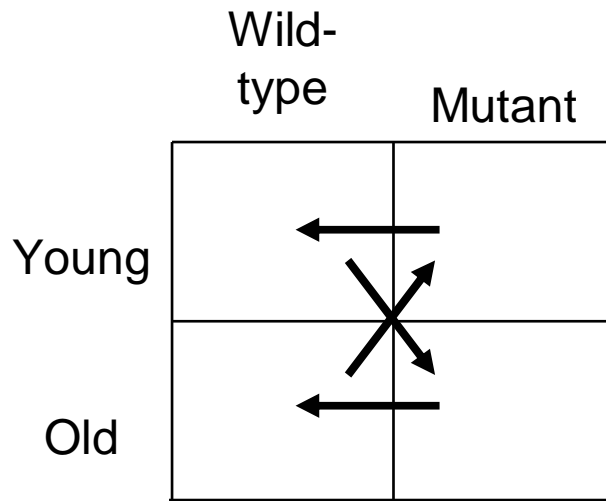
Original



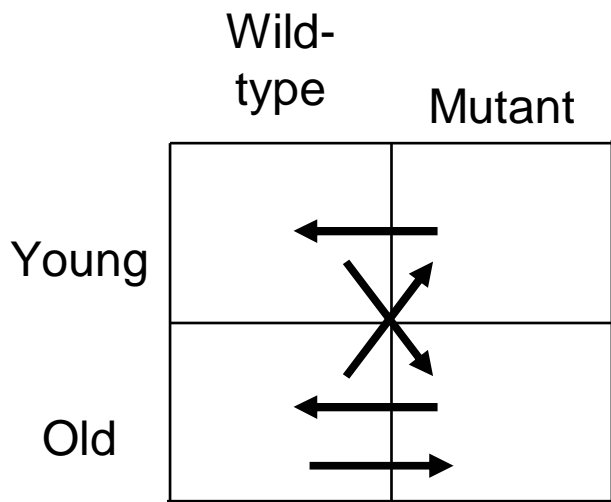
Proposed



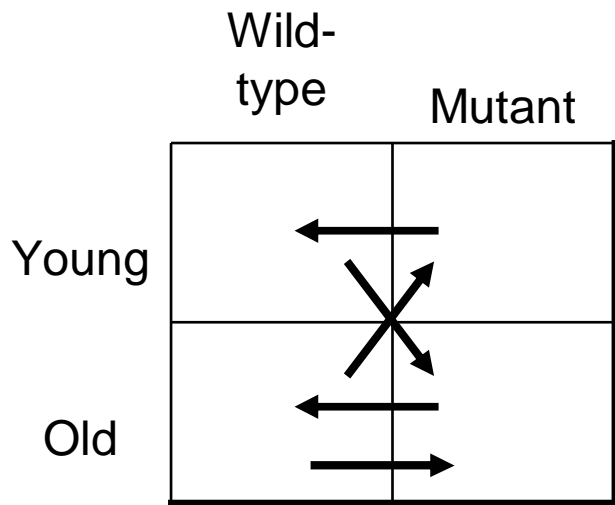
Original



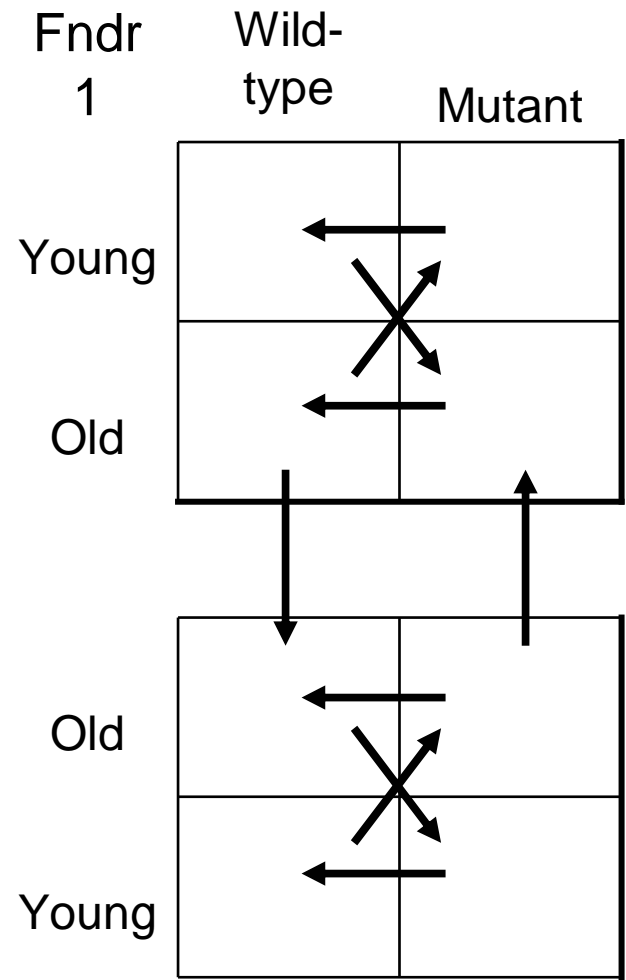
Proposed



Final



Final



Fndr 2
Wild-type Mutant

Alternative

Variance Comparisons

(Now across founders)

Design	GEN within AGE (w/n Old/Young)	GEN across AGE	AGE within GEN	AGE across GEN	Differences between Founders
Final	0.47 / 0.72	0.22	0.88	0.5	Not possible
Alter- native	0.47 / 0.72	0.22	0.88	0.5	Possible

Summary

- Replication is important at all levels – especially biological replication.
- The best microarray studies have a clear objective. The “design” of a microarray experiment is an important determinant of the success of the study
 - Efficiency
 - Robustness, etc.

katiek@u.washington.edu