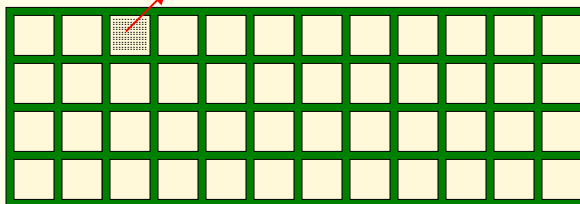


## Center for Animal Functional Genomics Michigan State University

- cDNA microarray BOTL-3
- 709 bovine EST clone inserts
- 345 amplicons (known genes)

Slides: 4 x 12 patches  
(9 x 9 spots /each)



### Controls

β actin  
RPL19  
GAPDH  
Lambda Q  
DMSO  
Blank  
etc.

## Different types of "replication" in microarray experiments

(Churchill, 2002)

### 1. Assessing measurement error

- ⇒ Multiple spots per gene on an array
- ⇒ Multiple arrays to study the same sample

\* Useful for quality control ...

But not too useful for inference on treatment effects!

### 2. Assessing biological variability

- ⇒ True replication

But... Should use appropriate error terms for testing treatments !

## Example 1

- ⇒ Effect of diet on blood pressure



- ① Two individuals (one vegetarian; another non-vegetarian)
- ② Same gender, same age, nonsmoking, etc.
- ③ The diastolic pressure is measured 3 times (consecutive days, for example)

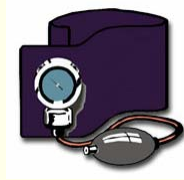
Vegan	Meat
80	80
95	100
110	90

Can we have a baseline and compare both groups?



## Example 2

⇒ Multiple individuals per group



Vegan				Meat			
$V_1$	$V_2$	...	$V_n$	$M_1$	$M_2$	...	$M_n$
80	88	...	73	110	93	...	115

$$y_{ij} = \mu + G_i + \varepsilon_{ij}$$

$$\begin{cases} i = 1, 2 \\ j = 1, 2, \dots, n \end{cases}$$

ANOVA TABLE

SV	DF	E[MS]
<b>Group</b>	1	$\sigma_R^2 + n\phi_G$
<b>Residual</b>	$2(n - 1)$	$\sigma_R^2$
<b>Total</b>	$2n - 1$	

## Example 3

⇒ Repeated measurements in each individual

Vegan				Meat			
$V_1$	$V_2$	...	$V_n$	$M_1$	$M_2$	...	$M_n$
80	88	...	73	110	93	...	115
105	89	...	79	106	105	...	90
90	100	...	75	105	89	...	97

$$y_{ij} = \mu + G_i + I_{ij} + \varepsilon_{ijk}$$

$$\begin{cases} i = 1, 2 \\ j = 1, 2, \dots, n \\ k = 1, 2, \dots, r \end{cases}$$

ANOVA TABLE

SV	DF	E[MS]
<b>Group</b>	1	$\sigma_S^2 + n\sigma_I^2 + nr\phi_G$
<b>Indiv.   Group</b>	$2(n - 1)$	$\sigma_S^2 + r\sigma_I^2$
<b>Residual</b>	$2n(r - 1)$	$\sigma_S^2$
<b>Total</b>	$2nr - 1$	

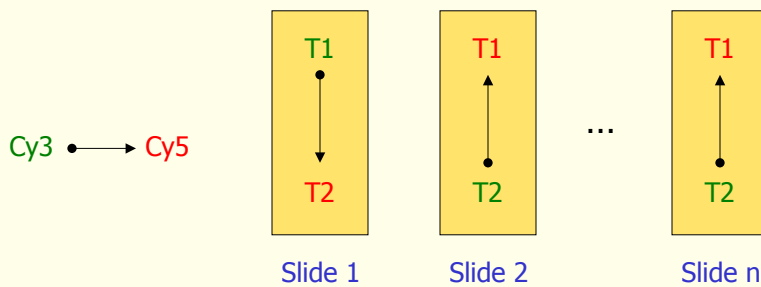


Back to microarray experiments ...

## TWO-SAMPLE EXPERIMENTS

\* Slides as blocks

- Two treatments (different experimental conditions)
- J 'biological' replications per treatment
- J slides total



## TWO-SAMPLE EXPERIMENTS

★ One slide per experimental unity (individual or pooled mRNA)

★ Model:

$$\log(y_{gijkm}) = \mu_g + T_{gi} + B_{gj} + D_{gk} + TB_{gij} + \text{interactions} + S | B_{g(j)m} + \epsilon_{ijkm}$$

★ ANOVA:

Biological replication

SV	DF
Trt (T)	1
Animal (B)	J - 1
Dye (D)	1
T*B	J - 1
Others	$\phi_1$
Spot (S B)	J(s - 1)
Residual	$\phi_2$
Total	2sJ - 1

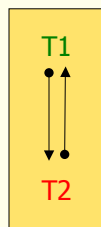
$$F_T = \frac{MS(T)}{MS(T * B)}$$

## TWO-SAMPLE EXPERIMENTS

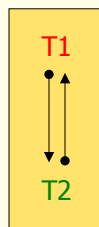
★ Slides as blocks

- Two treatments (different experimental conditions)
- J 'biological' replications per treatment
- 2J slides total (dye swap)

Cy3 → Cy5

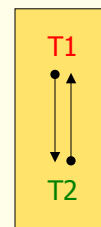


Slide 1



Slide 2

...



Slide n

## TWO-SAMPLE EXPERIMENTS

★ Two slides (dye swap) per experimental unity

★ Model:

$$\log(y_{ijklm}) = \mu_g + T_{gi} + B_{gj} + D_{gk} + TB_{gij} + A | B_{g(j)l} + \text{interactions} + S | A_{g(l)m} + \varepsilon_{ijklm}$$

★ ANOVA:

Biological replication

SV	DF
Trt (T)	1
Animal (B)	J - 1
Dye (D)	1
T*B	J - 1
Array (A B)	J
Others	$\phi_1$
Spot (S B)	2J(s - 1)
Residual	$\phi_2$
Total	4sJ - 1

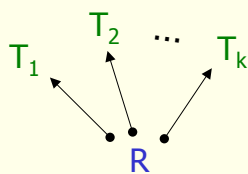
$$F_T = \frac{MS(T)}{MS(T * B)}$$

## MULTIPLE SAMPLES (TREATMENTS)

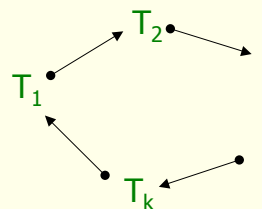
★ Incomplete block structure

- { k treatments (different experimental conditions)
- { n 'biological' replications per treatment
- { k × n slides total

Reference Design



Loop Design

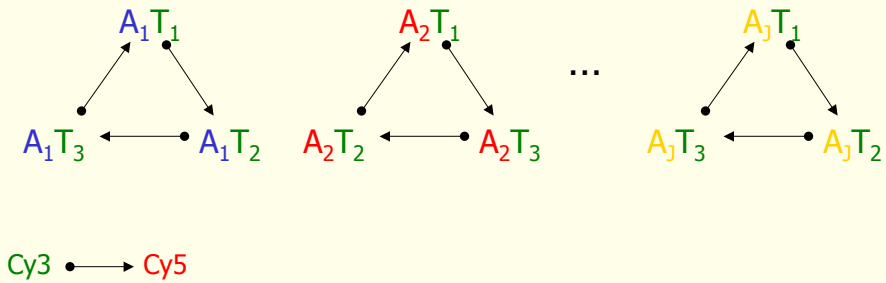


Cy3 → Cy5

## LOOP EXPERIMENTS

- \* Each animal under all trts (within a loop)

$\left\{ \begin{array}{l} I \text{ treatments} \\ J \text{ animals (or pools)} \\ I \times J \text{ slides total (} J \text{ loops of } I \text{ slides)} \end{array} \right.$ 
2 (I × J) if dye swap



## LOOP EXPERIMENTS

- \* Animal  $\Leftrightarrow$  Loop

- \* Model:

$$\log(y_{ijklm}) = \mu_g + T_{gi} + B_{gj} + D_{gk} + TB_{gij} + A | B_{g(j)l} + \text{interactions} + S | A_{g(l)m} + \epsilon_{ijklm}$$

- \* ANOVA:

SV	DF
Trt (T)	I - 1
Animal (B)	J - 1
Dye (D)	1
T*B	(I - 1)(J - 1)
Array (A B)	J(I - 1)
Others	$\phi_1$
Spot (S B)	IJ(s - 1)
Residual	$\phi_2$
Total	2sIJ - 1

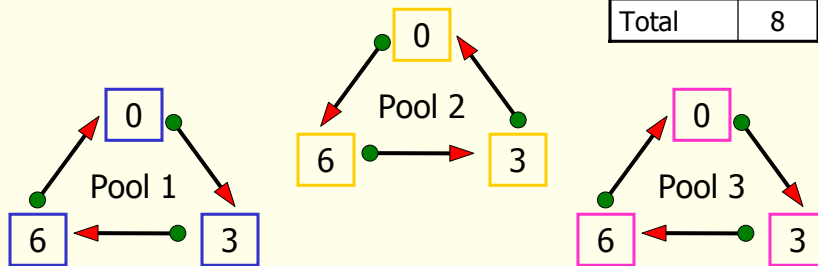
$$F_T = \frac{MS(T)}{MS(T*B)}$$

## Cartilage Impact

Pooi-See Chang and Dr. Mike Orth

- Time after impact (0, 3 and 6 h)
- Three pools of three animals each
- Multiple Loop Design

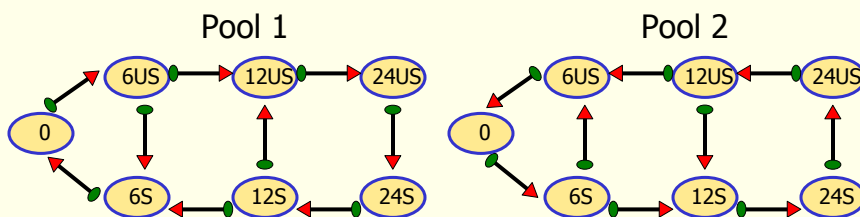
S.V.	d.f.
Pool	2
Time	2
Residual	4
Total	8



## *In Vitro* Bovine Immune Response to *Trypanosoma brucei* Infection

Dr. Emmeline Hill and Dr. Paul Coussens

- Time course (0, 6, 12 and 24 h)
- Stimulated (S) and Non Stimulated (NS)
- Two pools of two animals

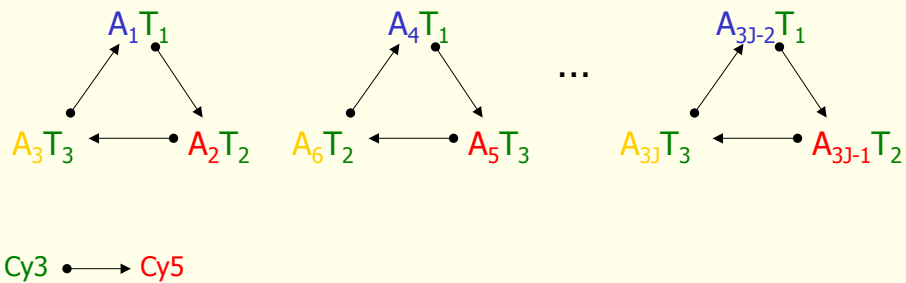


Error Term for testing Trt: Pool\*Treatment (6 d.f.)

## LOOP EXPERIMENTS

★ Each animal under one trt

$\left\{ \begin{array}{l} I \text{ treatments} \\ J \text{ animals per treatment (} I \times J \text{ animals in total)} \\ I \times J \text{ slides total (} J \text{ loops of } I \text{ slides)} \end{array} \right.$



## LOOP EXPERIMENTS

★ Model:

$$y_{gijklmn} = \mu_g + T_{gi} + B | T_{g(i)j} + D_{gk} + L_l + A | L_{g(l)m} + \text{interactions} + S | A_{g(m)n} + \epsilon_{ijklmn}$$

★ ANOVA:

SV	DF
Trt (T)	I - 1
Dye (D)	1
Animal (B T)	I(J - 1)
Loop (L)	J - 1
Array (A L)	J(I - 1)
Others	$\phi_1$
Spot (S B)	IJ(s - 1)
Residual	$\phi_2$
Total	2sIJ - 1

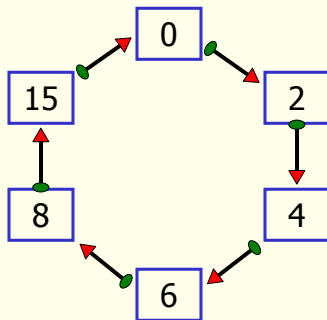
$$F_T = \frac{MS(T)}{MS(B|T)}$$

## mRNA Profiles Post Infection

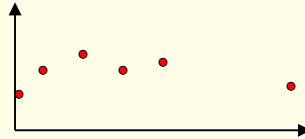
Dr. Line Morgills and Dr. Jeanne Burton

- Time after infection (0, 2, 4, 6, 8, 15 days)
- A different animal per time point
- Loop Design

Cy3 ● → Cy5



S.V.	d.f.
Linear	1
Quadratic	1
Lack of Fit	3
Total	5



## OTHER EXPERIMENTAL DESIGNS

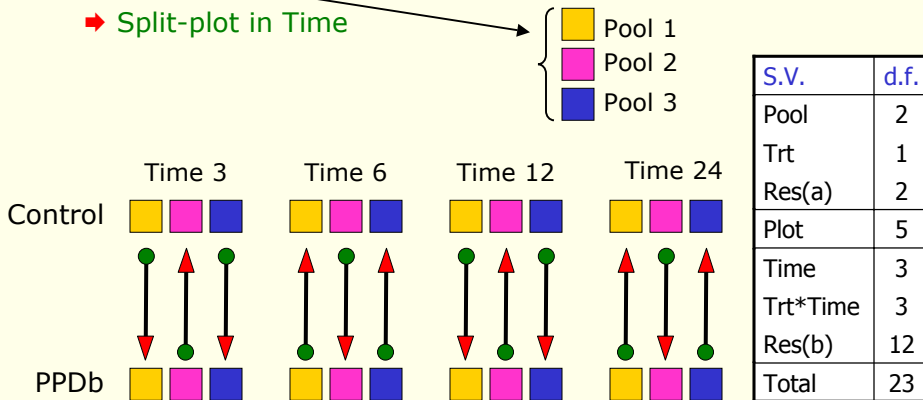
### \* SPLIT-PLOT

- ⇒ Within a CRD or RBD (on plots)
- ⇒ Split-Plot on Time

## Bovine TB Experiment

Kieran Meade and Dr. Paul Coussens

- PPDb stimulation over time (3, 6 12, and 24 h)
- Three pools with two infected animals each
- Split-plot in Time

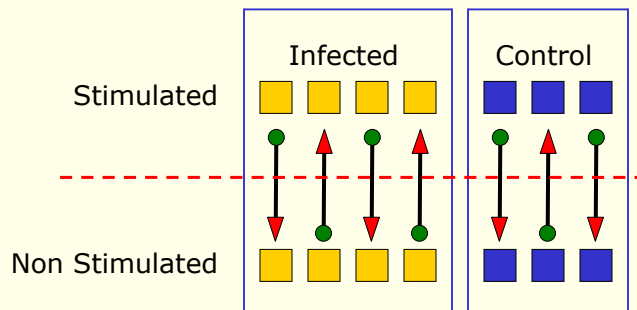


## Peripheral Blood Mononuclear Cells from MPTb Infected Cows

Dr. Paul Coussens

- Infected ( $n_1 = 4$ ) and Control ( $n_2 = 3$ ) Cows
- Two treatments: Stimulation or not
- Split-plot Design

S.V.	d.f.
Group	1
Res(a)	5
Plot	6
Trt	1
Group*Trt	1
Res(b)	5
Total	13



# Leptnin and IGF-I Intramammary Infusion

Brett Etchebarne and Dr. Mike VandeHaar

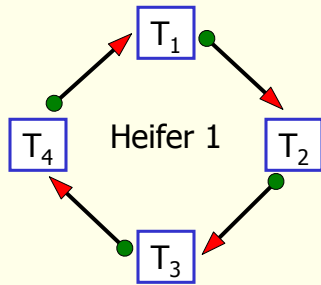
➔ Four (2 × 2) treatments, one in each quarter

➔ Six Heifers, Loop Design

➔ Treatments:

Trt	IGF	Lept
1	N	N
2	N	Y
3	Y	N
4	Y	Y

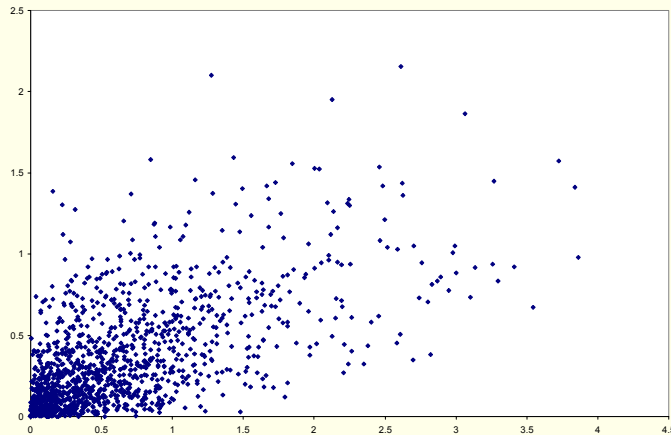
S.V.	d.f.
Dye	1
Trt	3
Heifer	5
Trt*Heifer	15
Array Heifer	18
Spot Array	48
Residual	53
Total	133



# Leptnin and IGF-I Intramammary Infusion

Brett Etchebarne and Dr. Mike VandeHaar

➔ P-values (log scale) from fixed and mixed effects models



## FINAL REMARKS

- ⇒ Experimental Design
- ⇒ Technical Replication:
  - Measurement error assessment
  - Future allocation of sampling efforts
- ⇒ Biological replication & Error variance
- ⇒ Borrowing information across genes (variance)

## REFERENCES

- Churchill, G. A. (2002) Fundamental of experimental design for cDNA microarrays. *Nature Genetics* 32: 490-495.
- Cleveland, W. S. (1979) Robust locally-weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74: 829-836.
- Cleveland, W. S., Grosse, E. (1991) Computational methods for local regression. *Statistics and Computing* 1: 47-62.
- Efron, B., Storey, J. D., Tibshirani, R. (2001) Microarrays, empirical Bayes methods, and false discovery rates. Stanford Technical Report # 216.
- Kerr, M. K., Churchill, G. A. (2001) Statistical design and the analysis of gene expression microarray data. *Genet. Res. Camb.* 77: 123-128.
- Lee, M.-L. T., Kuo, F. C., Whitmore, G. A., Sklar, J. (2000) Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl. Acad. Sci. USA* 97: 9834-9839.
- Wolfinfer, R. D., Gibson, G., Wolfinfer, E. D., Bennett, L., Hamadeh, H., Bushel, P. Afshari, C, Paules, R. S. (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comp. Biol.* 8: 625-637.
- Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., Speed, T. P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nuclei Acids Research* 30, No. 4 e15..