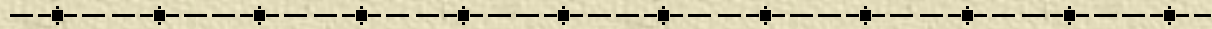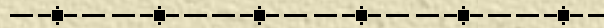# Educational Initiatives and Data Analysis in the Microarray Core

Danny Park

Bioinformatics (Sidney St)

Lipid Metabolism Unit (Freeman)

Hats I Wear

- Software development, maintenance, support for Microarray Core

- Learning & recommending analysis techniques, algorithms, and software tools

- Training, teaching, educating researchers

- Misc. bioinformatics

Today's Presentation

* Demonstrate the most basic *analysis techniques*
* Using our most commonly used *software (BASE)*
* For the most common kind of *experiments*
* While pretending you're a typical *audience* (biologists, inexperienced with microarrays)

Work Flow

Work Flow, distilled

Researcher

training / feedback

Me
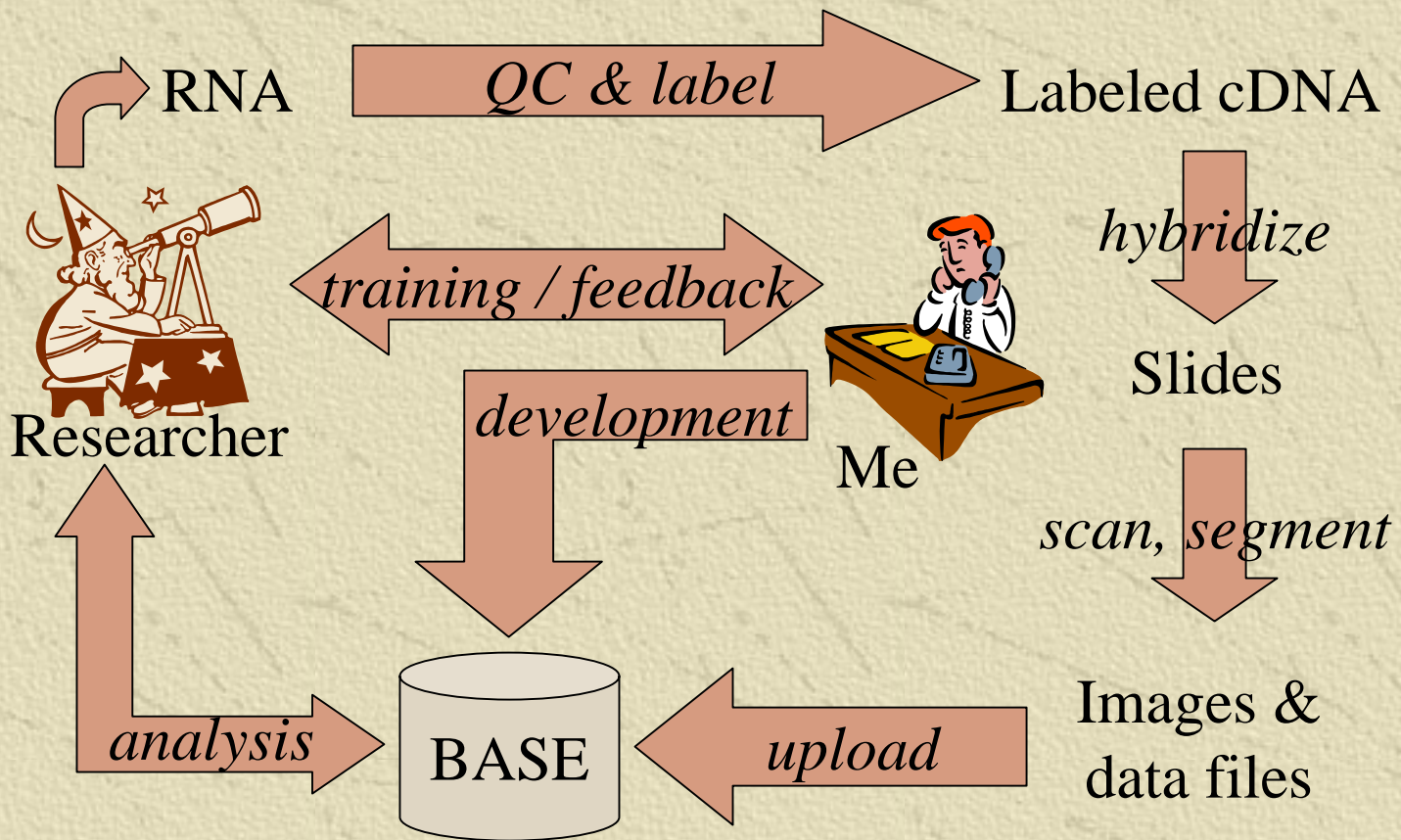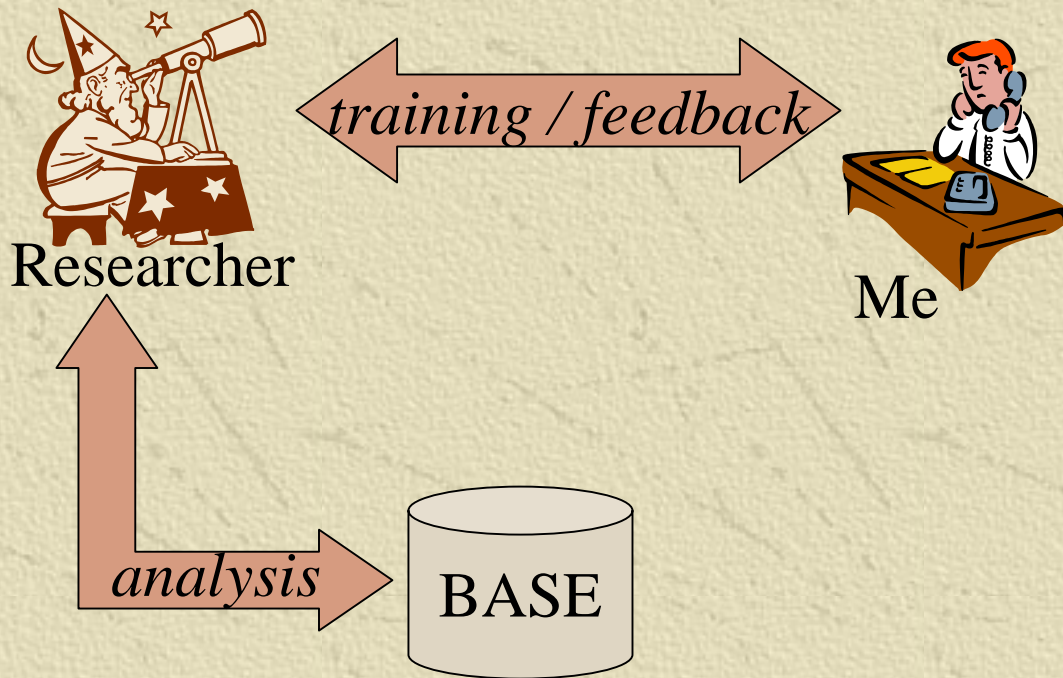
analysis

BASE

# BASE for Bioinformatics People
"BioArray Software Environment"

- Typical web/client-server model:
  - Clients: javascript-enabled browser
  - Web server: Apache/PHP, some C++, Perl, R code for analysis
  - DB server: MySQL for storage
- Open source, headed by thep.lu.se
- Functions:
  - Data storage, archival
  - Basic analysis: filters, transformations, normalizations, graphs, etc
  - Not so good: clustering, visualization

Researchers

BASE

storage   analysis
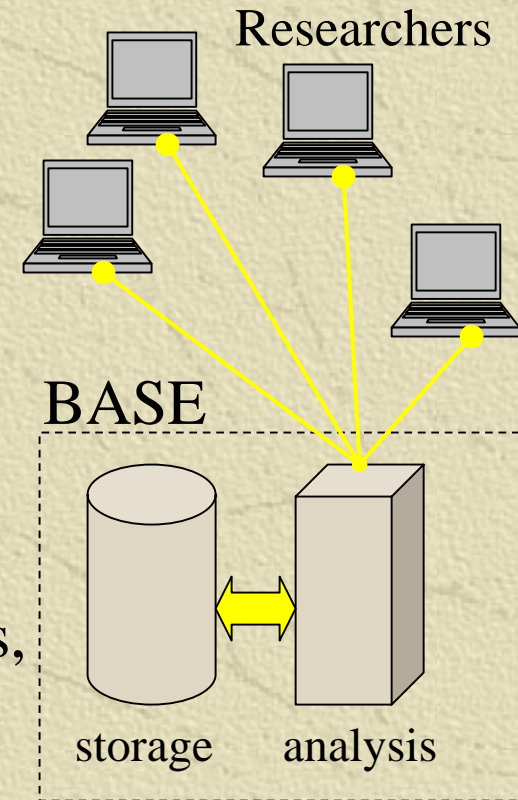
The Most Common *experiment*
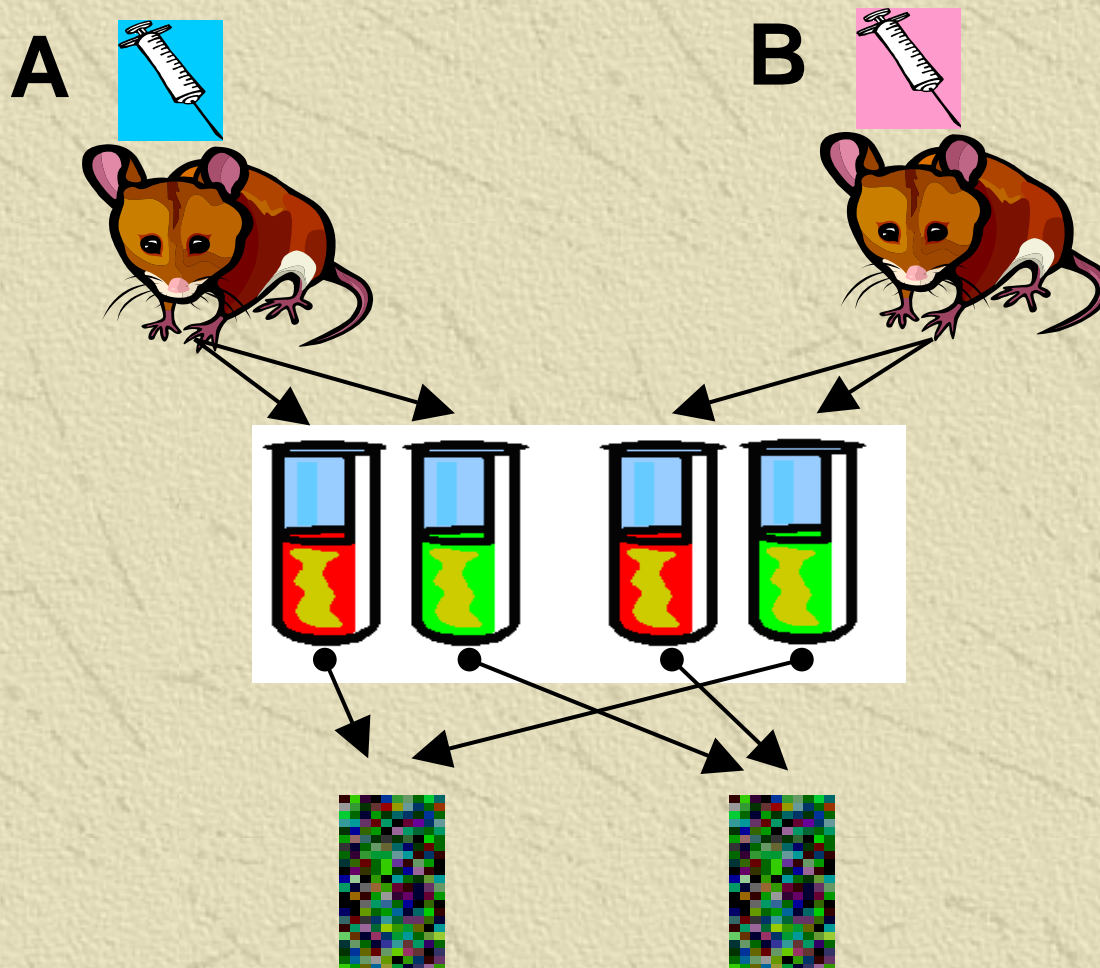
- Two-sample comparison w/N replicates
  - KO vs. WT
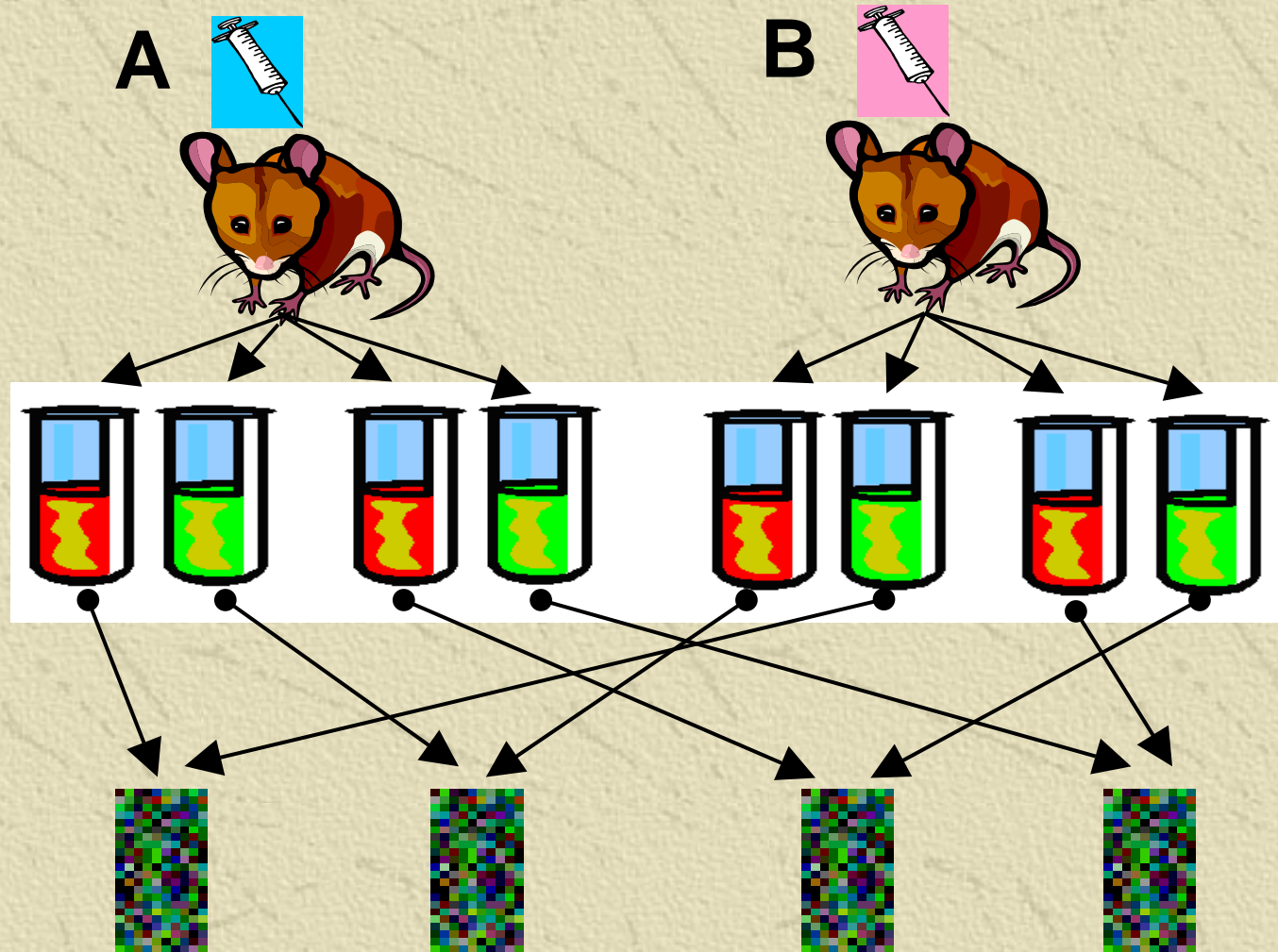 - Treated vs. untreated
  - Diseased vs. normal
  - Etc
- Question of interest: which genes are (most) differentially expressed?
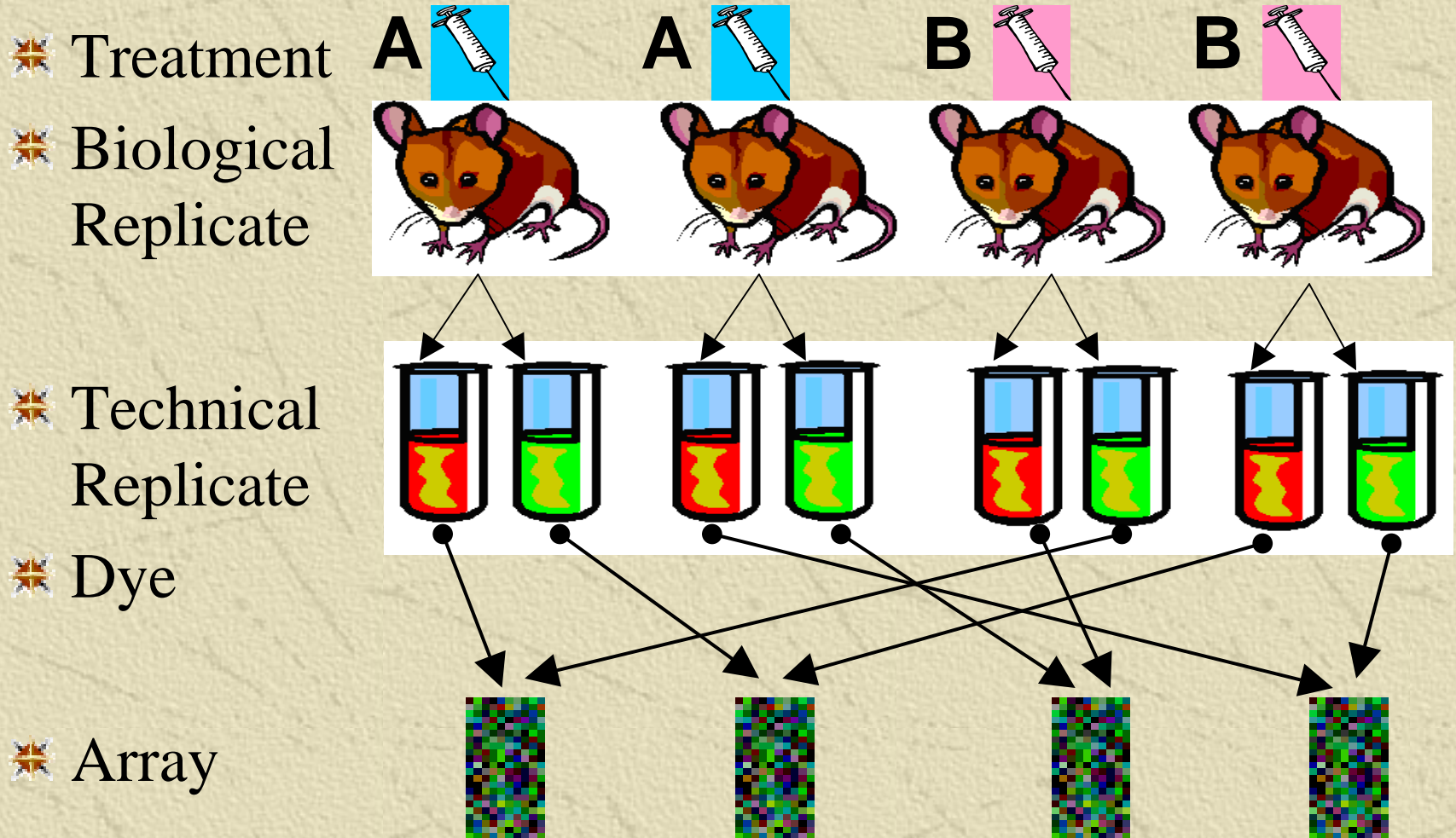
# Experimental Design – naïve

Experimental Design – tech repl

Experimental Design – bio repl

Treatment

Biological Replicate

Technical Replicate

Dye

Array

The Most Common *Analysis*

- Filter out bad spots
- Adjust low intensities
- Normalize – correct for non-linearities and dye inconsistencies
- Filter out dim spots
- Calculate average fold ratios and p-values per gene
- Rank, sort, filter, squint, sift data
- Export to other software

Live Demo
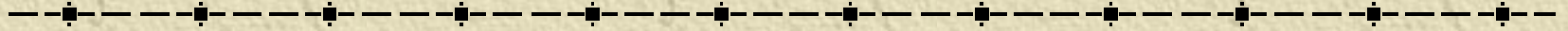
this space intentionally left blank

Acknowledgements

- Mason Freeman
- Harry Bjorkbacka
- Glenn Short
- Jocelyn Burke
- Najib El Messadi
- Jing Wang

- Chuck Cooper
- Xiaowei Wang
- Xiaoman Li (HSPH)

Stolen powerpoint figures from Gary Churchill
(Jackson Labs)