

# R/mpMap Workshop

## Part 1: Overview

Emma Huang

TAMU, 2 Sep. 2015

# Plan

## Wednesday

- Part 1: Overview (20 min)
  - Introduction
  - Background
  - Installation
  - Documentation
  - Importing Data
- Exercises (5 min)
- Break/Questions (5 min)

# Plan (Thursday)

8:30-9:30

- Part 2: Linkage Map Construction (45 min)
  - Estimating recombination fractions
  - Grouping
  - Ordering
  - Refinement
- Exercises (10 min)
- Break/Questions (5 min)

# Plan

9:30-10:30

- Part 3: QTL Mapping (45 min)
  - Full Model
  - Association mapping
  - Meta-alleles
  - Mixed models
- Exercises (10 min)
- Break/Questions (5 min)

# Plan

10:30-11:30

- Part 4: Advanced Topics (45 min)
  - Imputation
  - Selective phenotyping
  - Simulation
  - Haplotype Mosaics/Recombination
  - Visualization
- Exercises (10 min)
- Break/Questions (5 min)

What your research supposedly looks like:

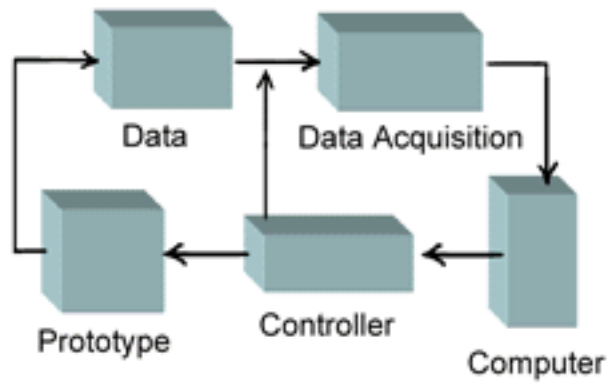


Figure 1. Experimental Diagram

What your research *actually* looks like:

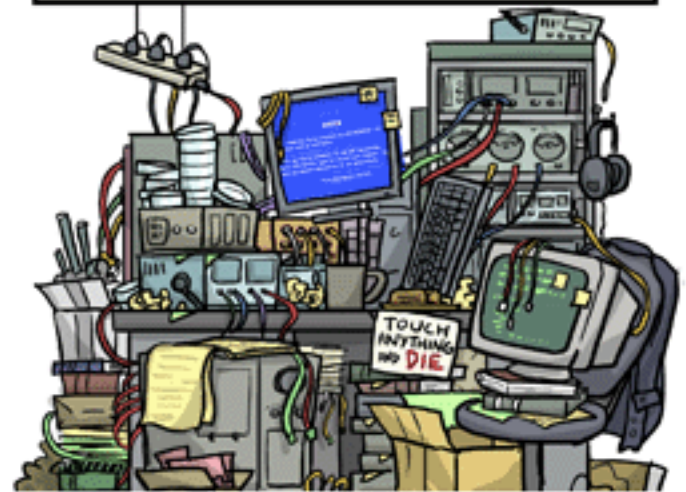


Figure 2. Experimental Mess

# Part 1: Overview

# Introduction



1998-2002  
B.S. With Honors  
Mathematics



UNC  
GILLINGS SCHOOL OF  
GLOBAL PUBLIC HEALTH

2002-2007  
Ph.D.  
Biostatistics



2007-2015  
Postdoc  
Research Scie  
Senior Research S  
Statistical Genetics

# At CSIRO

- 2009: Start of MAGIC linkage map construction
- 2011: Published R/mpMap
- 2013: Workshop @ NIAB on MAGIC-Type Populations
  - <http://openwetware.org/wiki/MAGIC:Slides>
- 2014: Genetics/G3 special collection
  - <http://www.genetics.org/cgi/collection/multiparental>
- 2015: Workshop @ PAG Asia on Multiparental Populations



# Collaborators

- Wheat MAGIC4/8, CSIRO
- Durum wheat MAGIC4, U. Bologna
- Barley MAGIC8, U. Bonn
- Tomato MAGIC8, INRA
- Arabidopsis MAGIC19, Cambridge, Oxford
- Chickpea MAGIC8, ICRISAT
- Rice MAGIC8/16, IRRI
- ...

# Going forward - less involvement with MAGIC

Associate Director in Population Analytics @ Janssen

- Focus on precision (personalized) medicine

Sep. 9-11 Wageningen, Netherlands

- EUCARPIA Biometrics in Plant Breeding

Nov. 16-18 China

- Nature Agrigenomics Conference

# Background: why mpMap?

- Alternatives:
  - HAPPY - applicable to general designs
  - R/qtl - more general
- Interfaces with both
- More specific: focus on MAGIC
- More general: broad range of tools

# Installation

```
install.packages("devtools")  
library(devtools)  
install_github("behuang/mpMap")  
library(mpMap)
```

```
## Loading required package: Rcpp
```

# Documentation

- Should be the first stop for
  - inputs
  - outputs
  - related pages
  - examples

```
library(mpMap)
```

```
?mpMap
```

```
?read.mpcross
```

```
?mpcross
```

# Importing Data

- Can be created using files or objects
- Necessary components
  - finals - progeny genotypes
  - founders - parental genotypes
  - pedigree - record of all crosses used in design
  - id - indices of rows in pedigree of genotyped finals
  - fid - indices of parental rows in pedigree
- Optional components
  - map
  - phenotypes

# Common issues importing data

- Numeric genotypes
- Pedigree structure
- Indices

$[((A \times B) \times (C \times D)) \times ((E \times F) \times (G \times H))]-1-1-1-1-1-$

BP-BP-BP

A/B//C/D///E/F//G/H

ABCDEFGH

# Simulated Data

```
map <- sim.map(len=100, n.mar=51, eq.spacing=T, include.x=F)
ped <- sim.mpped(4, 1, 2, 3, 1)
dat <- sim.mpcross(map, ped)
```

```
names(dat)
```

```
## [1] "founders" "finals"    "ibd"        "pheno"      "map"
## [7] "id"        "fid"        "qtlgeno"
```

```
dat$founders[,1:5]
```

```
##      D1M1 D1M2 D1M3 D1M4 D1M5
## A      1    0    1    1    1
## B      1    1    0    1    1
## C      0    1    0    0    0
## D      1    0    0    1    1
```



# Simulated Pedigree

ped

##		id	Male	Female	Observed
##	1	1	0	0	0
##	2	2	0	0	0
##	3	3	0	0	0
##	4	4	0	0	0
##	5	5	1	2	0
##	6	6	3	4	0
##	7	7	5	6	0
##	8	8	5	6	0
##	9	9	7	7	0
##	10	10	9	9	0
##	11	11	10	10	1
##	12	12	8	8	0
##	13	13	12	12	0
##	14	14	13	13	1

# Exercises

# Installation and Documentation

- What version number is it?
- When was it last updated?
- Name three functions used for linkage map construction
- Name three visualization functions
- Name two other software packages to which mpMap can output files

# Importing Data - Founders

- Which will work?
- What will cause problems?

##		D1M1	D1M2	D1M3	D1M4	D1M5
##	A	0	0	1	0	0
##	B	1	0	0	1	1
##	C	0	0	0	1	0
##	D	0	1	1	1	0

##		D1M1	D1M2	D1M3	D1M4	D1M5
##	A	AA	AA	BB	AA	AA
##	B	BB	AA	AA	BB	BB
##	C	AA	AA	AA	BB	AA
##	D	AA	BB	BB	BB	AA

# Importing Data - Finals

- Which will work?
- What will cause problems?

##		D1M1	D1M2	D1M3	D1M4	D1M5
##	L1	0	0	0	1	1
##	L2	1	2	0	1	1
##	L3	1	0	0	1	1
##	L4	0	0	0	0	0
##	L5	0	1	0	1	0
##	L6	0	0	0	1	0

##		D1M1	D1M2	D1M3	D1M4	D1M5
##	L1	0	0	0	1	1
##	L2	1	0	0	1	1
##	L3	1	0	0	1	1
##	L4	0	0	0	0	0
##	L5	0	1	1	1	0
##	L6	0	0	0	1	0

# Importing Data - Pedigree

- Which will work?
- What will cause problems?

##		id	Male	Female	Observed
##	L1	1	0	0	1
##	L2	2	0	0	1
##	L3	3	0	0	1
##	L4	4	0	0	1
##	L5	5	1	2	0
##	L6	6	3	4	0

##		id	Male	Female	Observed
##	L1	1	0	0	0
##	L2	2	0	0	0
##	L3	3	0	0	0
##	L4	4	0	0	0
##	L5	5	1	2	0
##	L6	6	3	4	0

# Importing Data - Map

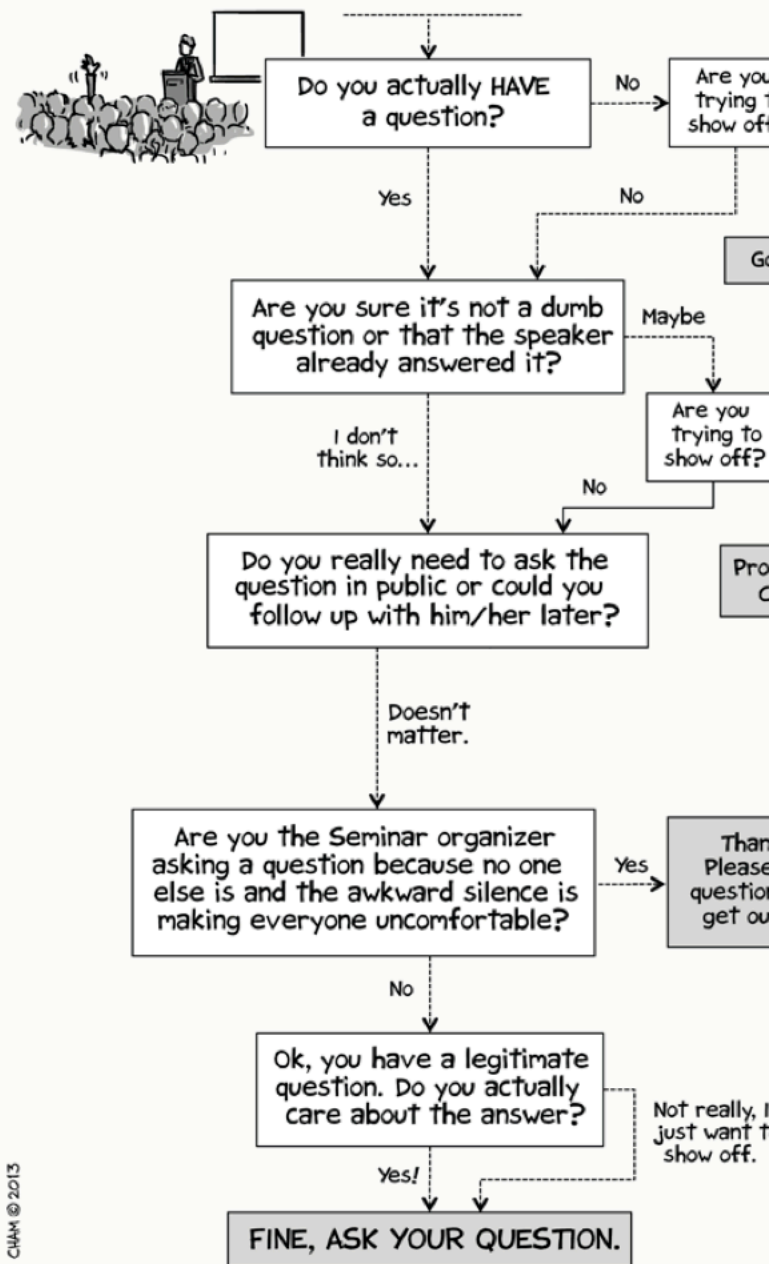
- Which will work?
- What will cause problems?

##		Marker	Chr	Pos
##	11	D1M1	1	0
##	12	D1M2	1	2
##	13	D1M3	1	4
##	14	D1M4	1	6
##	15	D1M5	1	8

##		Marker	Chr	Pos
##	12	D1M2	1	2
##	14	D1M4	1	6
##	15	D1M5	1	8
##	13	D1M3	1	4
##	11	D1M1	1	0

# Questions

## Should you ask a Question during Seminar?



JORGE CHAM © 2013

WWW.PHDCOMICS.COM