

# Using Galaxy for Molecular Assay Design

2012 Galaxy Community Conference

J Ireland, Principal Bioinformatics Scientist

5AM

# Galaxy at 5AM Solutions

- 5AM: Custom software and bioinformatics services company
- Galaxy converts for one year
  - <http://info.5amsolutions.com/blog/bid/86705/DIY-Bioinformatics-A-Whole-New-Galaxy>
- Galaxy in action
  - Internal bioinformatics resource
  - Client 1: data source integration
  - Client 2: custom workflows for assay design
  - Client 3 (in progress): deploy Galaxy for assay design



# What is Molecular Assay Design (MAD)?

- Design

- In-silico process to identify oligos (probes, primers) to detect and/or quantify nucleic acid targets

- Balance

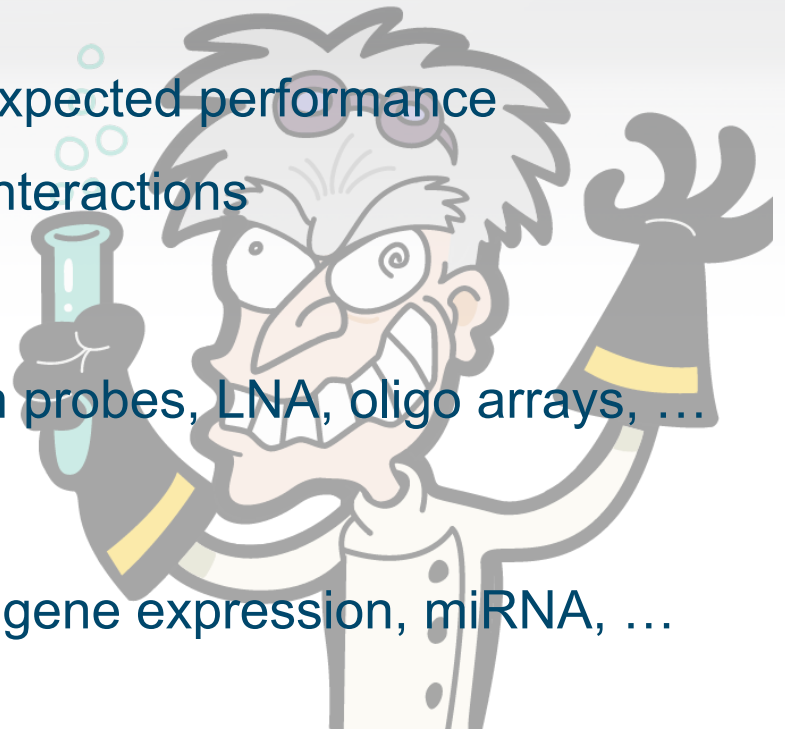
- Target Coverage: oligo location, expected performance
- Cost: \$\$\$, oligo real-estate, oligo interactions

- Technologies

- PCR, TaqMan, molecular inversion probes, LNA, oligo arrays, ...

- Applications

- Targeted sequencing, genotyping, gene expression, miRNA, ...

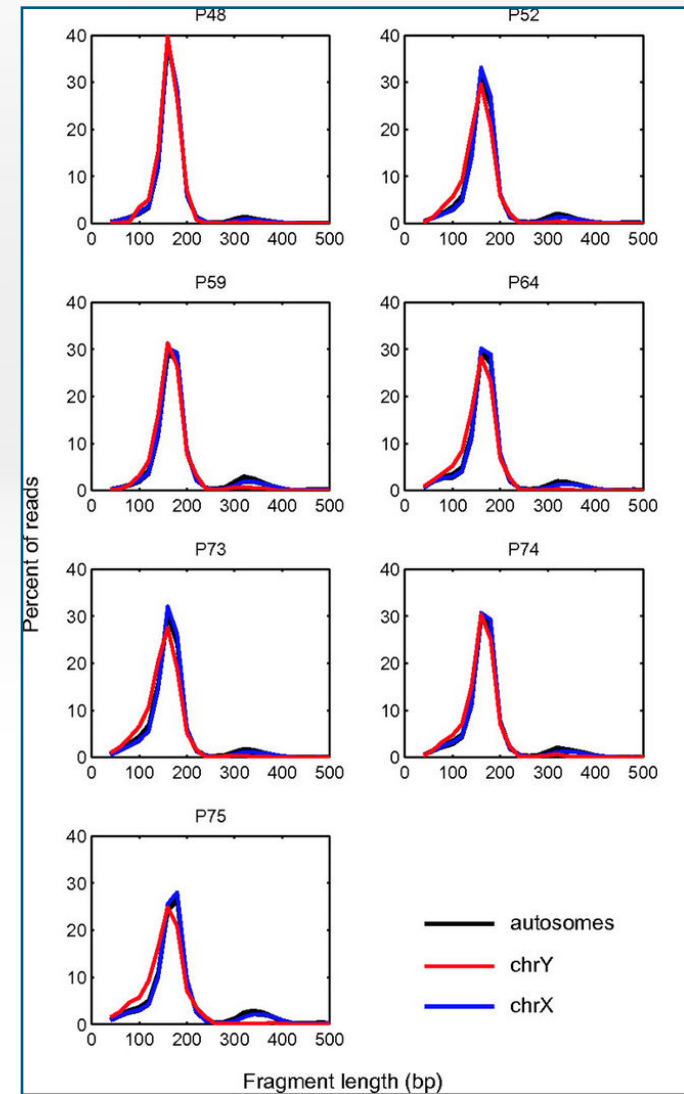


# The MAD Workflow

- Target Selection
  - Identify locations on genome/transcriptome of interest while avoiding trouble regions
- Assay Design
  - Select sets of primers and/or probes against targets
- Assay Prioritization/Filter
  - Filter, prioritize and annotate assay designs
- Ensemble Selection
  - Select final combination(s) of assays

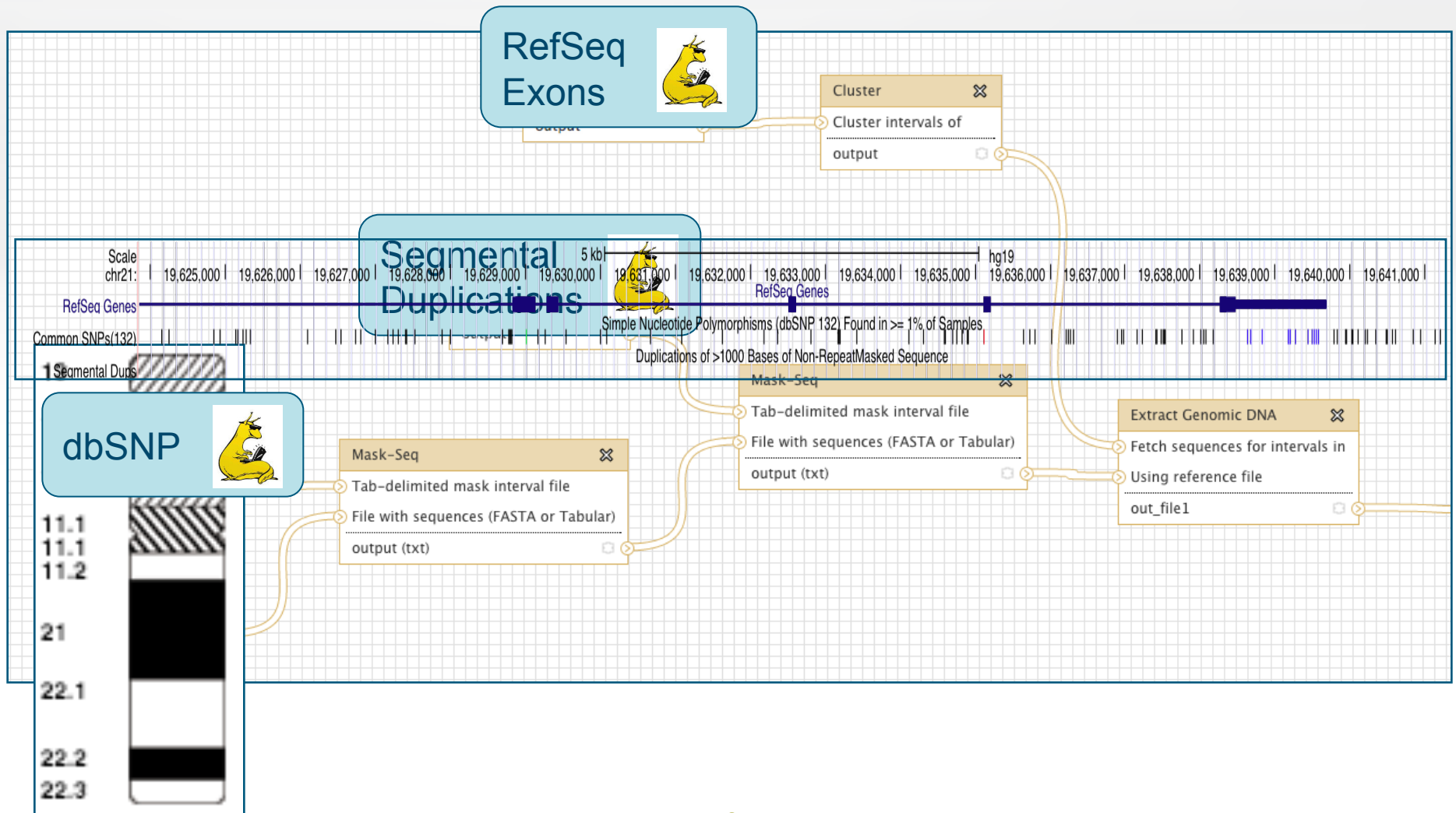
# Cell-Free DNA Testing – A MAD Example

- Client developing a multiplexed PCR assay for detection of trisomies in cfDNA
- Requires sensitive assay to detect fetal cfDNA in high background of maternal cfDNA
- Short assay lengths needed because cfDNA highly fragmented
- Genetic variation (SNPs, CNV) can confound detection

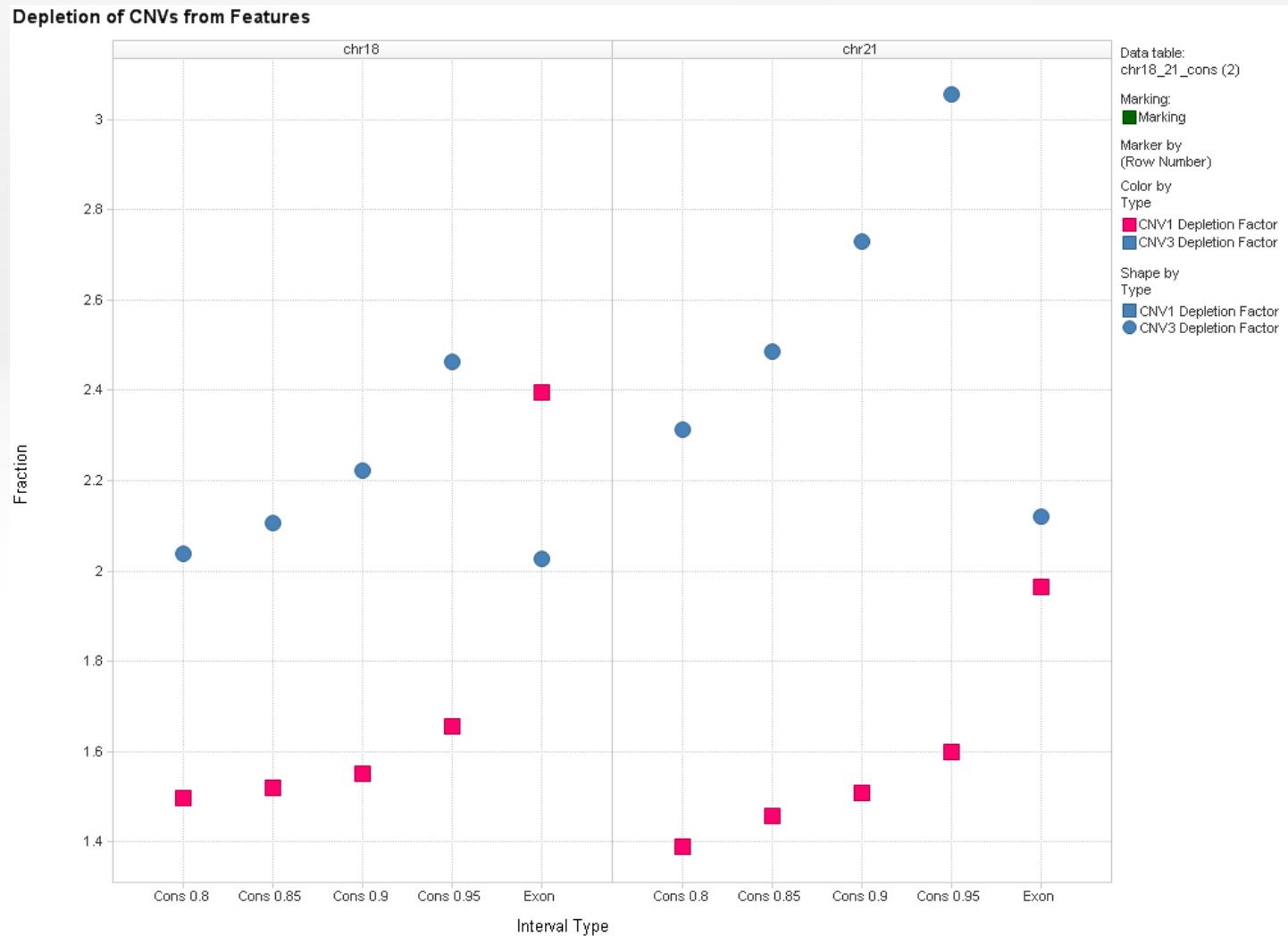


Fan, et al. Clin. Chem, 2010

# Targets for cfDNA Testing



# An Aside: Galaxy Powered MAD Analysis





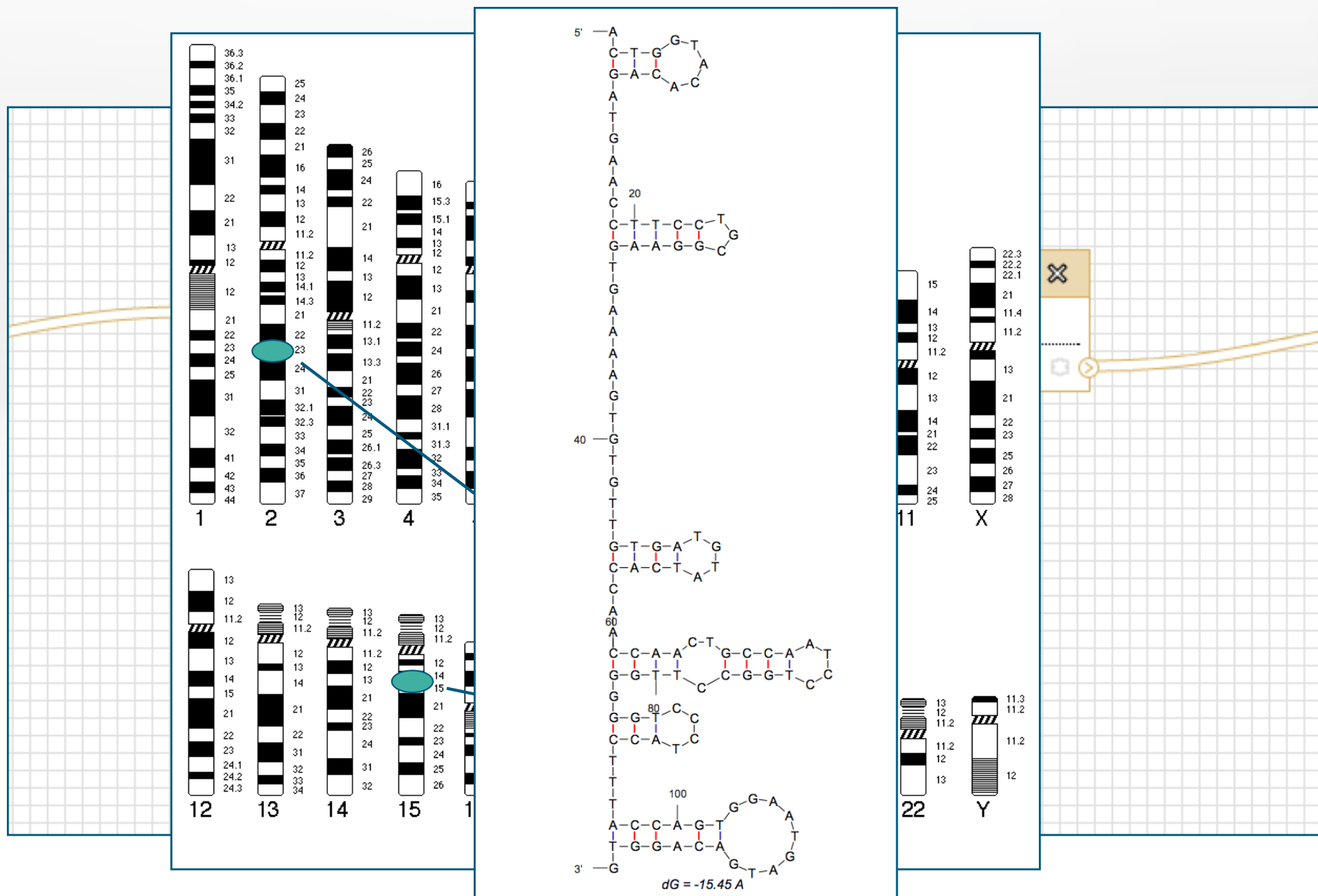
# Assay Design

[illegible]

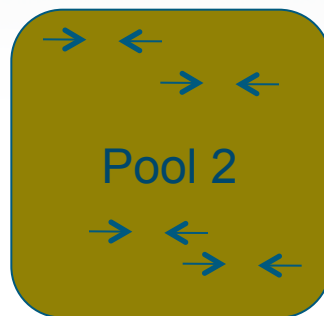
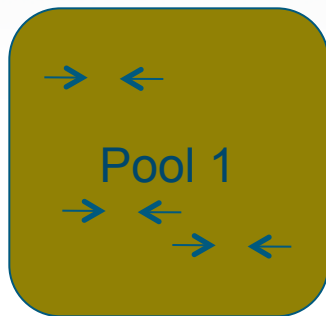
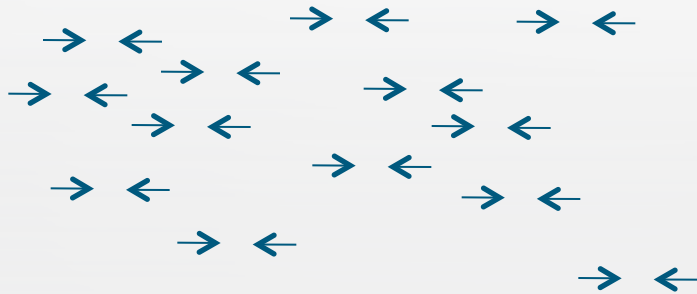
```
PRIMER_INTERNAL_MAX_SIZE=36      # Max probe length
PRIMER_INTERNAL_OPT_SIZE=24      # Opt probe length
PRIMER_INTERNAL_MIN_TM=64.0      # Min probe TM (Santa Lucia)
PRIMER_INTERNAL_OPT_TM=70.0      # Opt probe TM (Santa Lucia)
PRIMER_INTERNAL_MAX_TM=80.0      # Max probe TM (Santa Lucia)
INTERNAL_SPRIME_BASE_EXCLUDE=G   # No G at 5' end
PRIMER_MAX_SELF_ANY=8.0          # Primer annealing (anywhere) to itself
PRIMER_PAIR_MAX_COMPL_ANY=8.0    # Pair compl (anywhere)
PRIMER_INTERNAL_MAX_SELF_ANY=8.0 # Probe annealing (anywhere) to itself
```



# Assay Prioritization and Filtering



# Assay Ensembles



- Combine multiplexed pools
  - Avoid misamplification
  - Avoid oligo interactions
- Select “optimal” subset
  - Coverage
  - Expected performance
  - Best multiplex
- Programs to help
  - MultiPLX

# Looking Back



- No more Perl soup
- Data sets, parameters, steps preserved
- Reproducible design
- Sharable with client

# Looking Ahead

- Tool Shed submission
- Extend workflows to ensemble selection
- Automate report generation
  - Sweave for Galaxy?



# Thanks!



Andy Evans



Will FitzHugh

## The Galaxy Community!