Overview of the NC State Approach to Genomic Dose-Response Modeling

Fred Wright
Director, Bioinformatics Research Center
Departments of Statistics and Biological Sciences
Overview/preview of statistical procedures

Quality Control
• For sequence-based transcriptomic technologies, threshold individual genes based on expression level
• Outlier checks
• Compare control samples to all other control samples

Normalization
• Currently done per-experiment, e.g. using DESeq2 for sequence-based transcriptomics
Overview/preview of statistical procedures

Testing
• For statistical flags, we use simple rank-based procedures (see later)
• For differential expression analysis, we use shrinkage-based methods (for example, DESeq2, limma)

Multiple testing
• False discovery control

Dose-response curve fitting
• Highly reliant on 4-parameter (Hill) logistic fitting, or 3-parameter if that makes more sense in the context. With more data, gain-loss modeling
A series of choices and tests

• When dealing with gene expression dose-response data, natural tension among statistical flags, testing, and modeling

• Some of the pipeline reflects a specific sequencing technology

• Potential concern that controls may differ from dosed conditions for technical reasons
Pipeline overview

https://github.com/jshousephd/HT-CBA

Somewhat platform-specific
Display total read counts per well in a matrix

Low average correlation with other control samples
Inspection and analysis of control samples
Why differential expression packages provide shrunken estimates of variance to boost power

You can estimate variability from the mean

Much of this spread is sampling variation

Zhou et al., *Bioinformatics*, 2011, 27 (19), 2672–2678
The effects of count thresholds per gene

- Set criteria for analysis to be those genes with average counts $\geq 5$ (threshold) across samples
- Keep features (genes) with at least $\frac{1}{2}$ treatments that meet criteria
- $qvalue$ package to calculate $\hat{\pi}_0$ (estimated proportion of true null genes) by count (%tile)
Differential Gene Expression Assessment – 4 chemicals/drugs and treatment of iPSC cardiomyocytes (Rusyn Lab). Analysis by DESeq2
Statistical flag generation and dose-response decision chart

Comparing treatment groups via rank tests and Moment-Corrected Correlation

IV. Concentration Response Modeling and Point of Departure Assessment

Data Files
- Processes
- Decision
- Output

Normalized Counts

Calculate p and q values for:
- i. Spearman’s Rho - Controls + Treatments
- ii. Spearman’s Rho - Treatments Only
- iii. Wilcoxon Test - Controls vs. All Treatments

Is Overall Trend Sig?
- NO
  - Assign PCD = Max Dose
- YES
  - Does Vehicle = Treatments?
    - NO
      - Flagged Report
    - YES
      - Is Treatment Trend Significant?
        - NO
          - Flagged Report
        - YES
          - Point of Departure Assessment
            - i. Constant Model
            - ii. 3P-Hill Function
            - iii. Gain-Loss Model
            - iv. 4P-Hill Function
            - v. Assess best fit

Visualization:
- Heatmaps/Dendrograms
- PCA/Dispersion
A few remarks on dose-response curve-fitting

• With lots of data, one can explore a large number of models

• With few data points, may need to reduce the number of models explored

• Nonparametric smoothing methods may work okay, but finding appropriate bandwidths may be tricky with little data

• Most points-of-departure involve interpolation, so different reasonable models often agree

• For gene expression, need to handle testing as well as estimation
The 4-parameter logistic model is sigmoidal, has a “floor,” a “ceiling,” and parameters that govern when it rises, and how steeply.

However, depending on the range of doses, the model may offer a reasonable fit to data that might have been modeled more simply.
Benchmark dose typically uses variability to determine points of departure.

Deviations of fit from control mean in terms of variability of measurement (e.g. 1 SD departure, percentage change, etc.)

Note: Variability depends on technology.

Mean and variance can be computed from controls.
$\text{EC}_{10}$, $\text{EC}_{50}$ depend on the fit alone

Deviations of fit from control fit in terms of a pre-defined change

Note: For large sample sizes, should not depend on the variability

response

control

dose

ceiling

10% change

floor
Dose response fitting for the cardiomyocyte data using \textit{tcpl} and \textit{drm} 4P Hill, choose winner based on lowest AIC.

POD based on 1 SD departure from control mean.
“Significant” cardiac-associated toxicology pathways for dofetilide, based on fold-change

- IPA Analysis, using non-significant S1500+ genes as background (not the whole transcriptome)

- We do not yet using any kind of whole-transcriptome extrapolation
Discovery vs. predictive pathway analysis

• We use simple enrichment approaches (like everyone else), IPA, DAVID/EASE, etc.
• The simple tools provide easy results and some insight
• We and others have critiqued these methods as not providing accurate p-values per pathway, preferring full resampling approaches (e.g. SAFE, GSEA)
• **Final pathway-based PODs are based on minimum median pathway PODs, much like BMDExpress**
• Data on large numbers of chemicals will enable deeper investigations of pathway perturbations, and new methods to fully exploit the data
Use of points-of-departure for pathway-based determinations of overall transcriptional POD

• The uncertainty in pathway-based transcriptional points of departure could use further development

• We have been experimenting with bootstrapping to quantify this uncertainty at the per-gene level

• Also, bootstrapping may be very useful to quantify uncertainty for median pathway POD, because the constituent genes are correlated
Summary
• We have described a pipeline for handling gene expression dose-response data
• Much of the effort concerns the practicalities of QC and handling samples of small to medium size
• Once the foundation is laid, interesting comparisons can be made across multiple chemicals and chemical classes
• For example, I didn’t even discuss comparison to databases such as LINCS
Going further 1: ToxPi evaluations of pathway activity

- “slices” are composed similar measured features of possible concern
- Overall ToxPi score reflects weighted sum of slice sizes

Grim et al., *Green Chem.*, 2016, 18, 4407-4419
Going further 1: ToxPi evaluations of pathway activity (ToxPi 2.0)

Updated interface

Slices and uncertainty

Clustering by ToxPi profile

PODs of genes in expression pathway 1

PODs of genes in expression pathway 2
Going further 2: Evaluating population variability in pathway response (human cell line studies, mouse studies)

Workflow for estimating underlying variability when \( n \) samples measure a quantity with error (Chiu et al., ALTEX. 2017; 34(3): 377–388. doi:10.14573/altex.1608251)
Going further 2: Evaluating population variability in pathway response (human cell line studies, mouse studies)

- Can we (should we) be doing this analysis for gene expression pathway PODs?
- (Otherwise, when hundreds/thousands of chemicals are calculated, the most extreme-appearing will be over/under-estimated)
- How to approach it?
Acknowledgments

NC State
John House
Dereje Jima
Skylar Marvel
David Reif
Yi-Hui Zhou

Texas A&M
Ivan Rusyn
Fabian Grimm
Abhi Venkatratnam

BioSpyder
Pete Shephard

Funding: EPA STAR grants RD83516602 and R83580201