

Metrology for Gene Expression: Measurement Batch Effects, Probe Sensitivity, Gene-List Reproducibility

Walter Liggett, NIST

Jean Lozach, Illumina

Anne Bergstrom Lucas, Agilent

Ron Peterson

Marc Salit, NIST

Danielle Thierry-Mieg, NCBI

Jean Thierry-Meig, NCBI

Russ Wolfinger, SAS

Outline

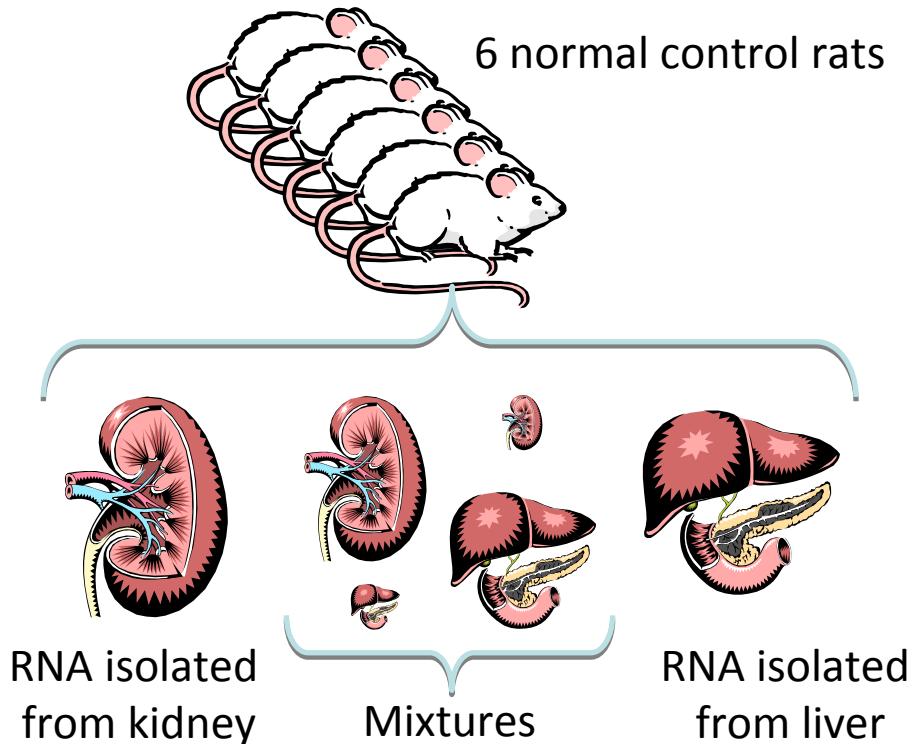
Replicate animals—measurements on the animals in a control group—allow three perspectives on gene expression metrology

- Comparison of measurement batch effects with biological variation
- Comparison of platforms in terms of probe sensitivity given in units of animal variation
- Performance of gene lists obtained from t statistics

RNA Titrations

“Titration” mixture fraction as surrogate for [RNA]

- a measure of “Truth”
- enables model-based analysis
 - each gene should follow a known, linear response
 - assess performance through analysis of residuals



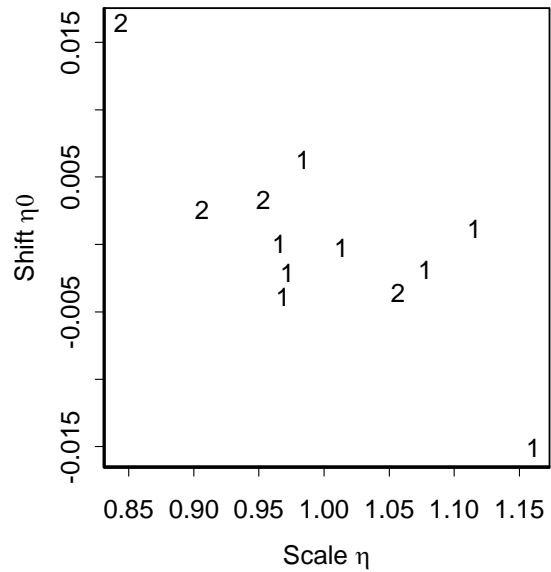
Normalization—A Brief Sketch

- Preliminary: For each array, divide intensities by their geometric mean (Add 30 for Agilent). Thus obtain for array i and probe p , the quantity y_{ip}
- Parametric normalization: Based on a parametric model of the titration
 - Kidney-to-liver ratio of mRNA concentrations
 - Proportions of materials A and D in the mixture samples depend on this ratio
 - Normalized intensities obtained from two array constants, η_{0i}, η_i : $(y_{ip} - \eta_{0i})/\eta_i$

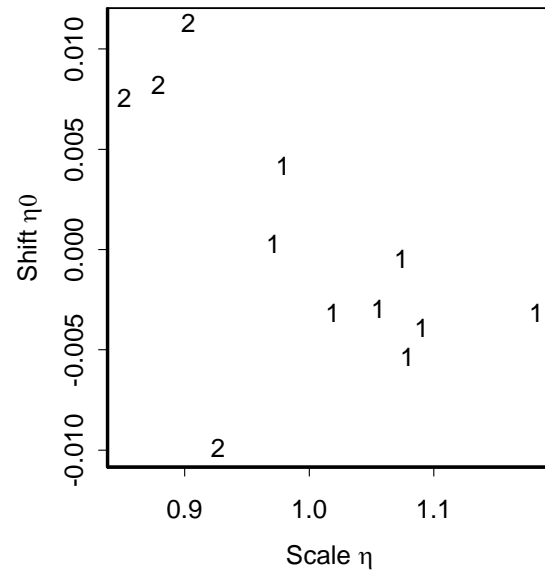
Agilent Normalization

- Normalization parameters η_0, η for each array
- User (protocol) indicated for each array
- Shows that the normalization removes, in part, the user batch effect

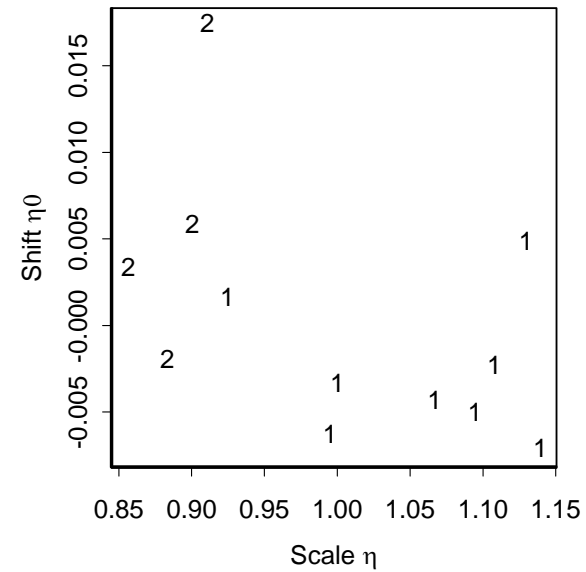
Animal 1



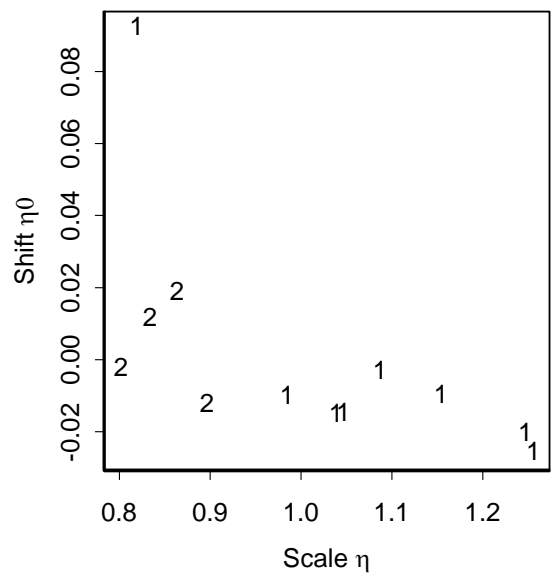
Animal 2



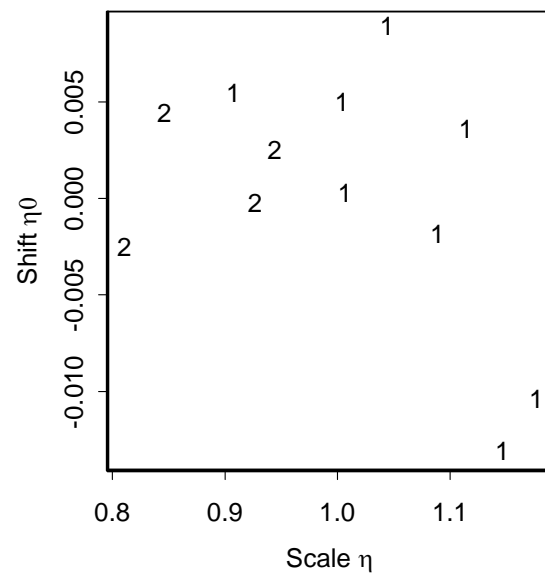
Animal 3



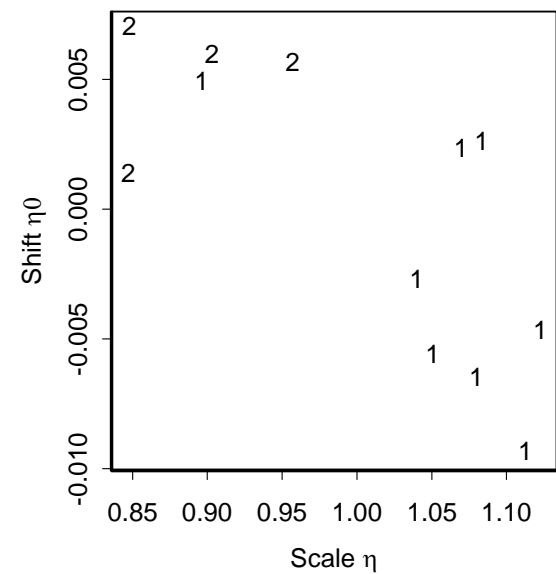
Animal 4



Animal 5



Animal 6



Normalization Output

- Normalized intensities
- Regression weights for modeling intensities
- Gene selection criteria based on 72—array, animal-average liver and kidney intensities values
 - Expressed genes
 - Differentially expressed genes

Batch Effects

- Affymetrix
 - Scanner
 - Fluidics machine
- Agilent
 - User (an unintended difference in protocol)
 - Substrate (4 arrays per substrate)
- Illumina
 - Operator
 - Chip (12 arrays per chip)

An Approach to Batch Effects Modeling

Animal-average intensity from normalization

$$\hat{x}_{Aij} \bar{\theta}_{Ag} + \hat{x}_{Dij} \bar{\theta}_{Dg}$$

Terms in order displayed

- Liver proportion in mixture
- Liver average for the particular probe
- Kidney proportion in mixture
- Kidney average for the particular probe

Linear Mixed-Effects Model

- Animal effects--fixed
 - Proportional to animal average intensity
 - Constant
- Scanner, user, operator effects--fixed
 - Proportional to animal average intensity
 - Constant (included only for Agilent)
- Fluidics machine, substrate, chip--random
 - Proportional to animal average intensity

Model (con't)

- Animal effects

$$\alpha_j \left(\bar{\theta}_{Ag} / 2 + \bar{\theta}_{Dg} / 2 \right) + \beta_j \left(\hat{x}_{Aij} \bar{\theta}_{Ag} + \hat{x}_{Dij} \bar{\theta}_{Dg} \right)$$

- Scanner, user, operator effects

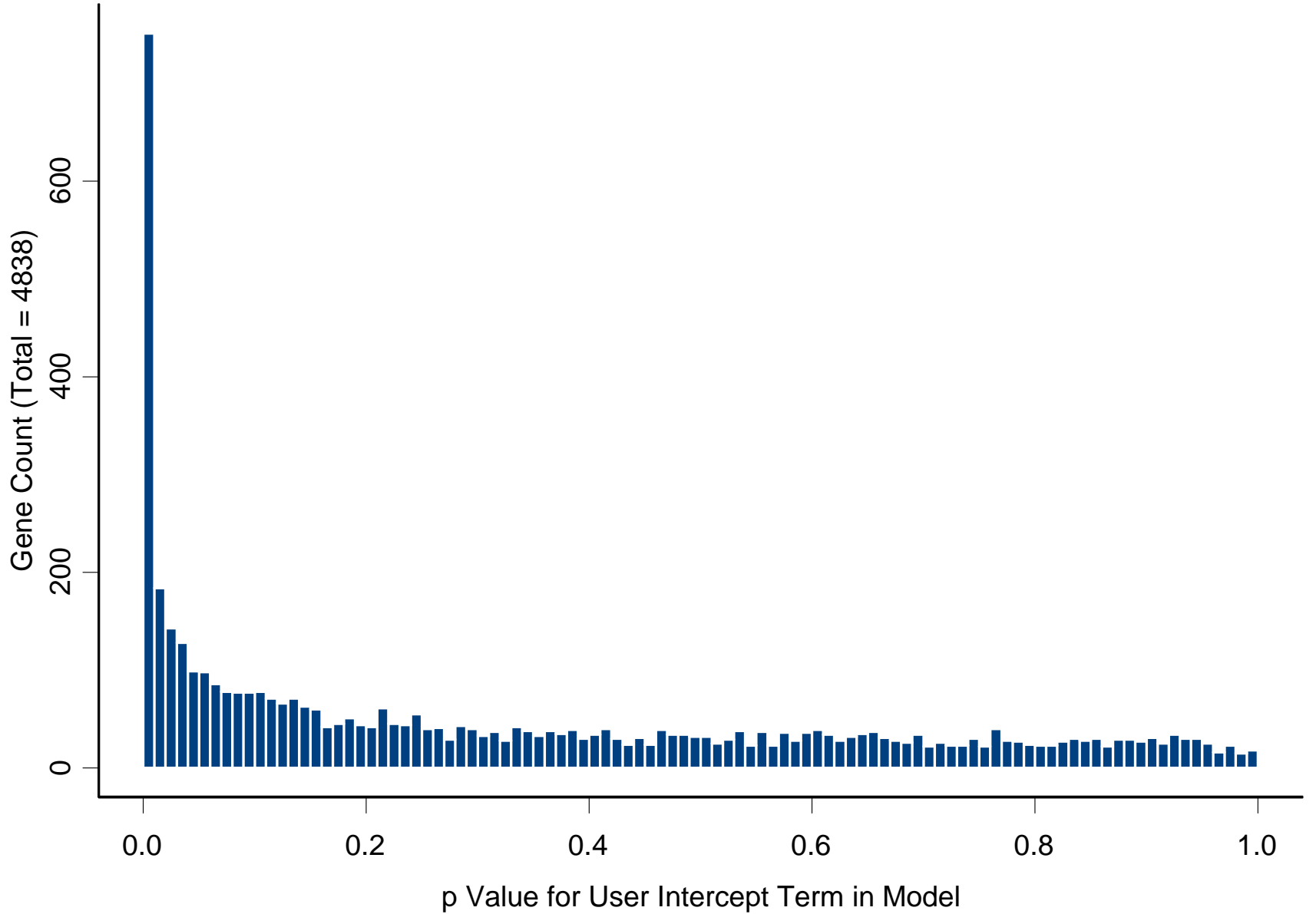
$$\delta_i \left(\bar{\theta}_{Ag} / 2 + \bar{\theta}_{Dg} / 2 \right) + \gamma_i \left(\hat{x}_{Aij} \bar{\theta}_{Ag} + \hat{x}_{Dij} \bar{\theta}_{Dg} \right)$$

- Fluidics machine, substrate, chip effects

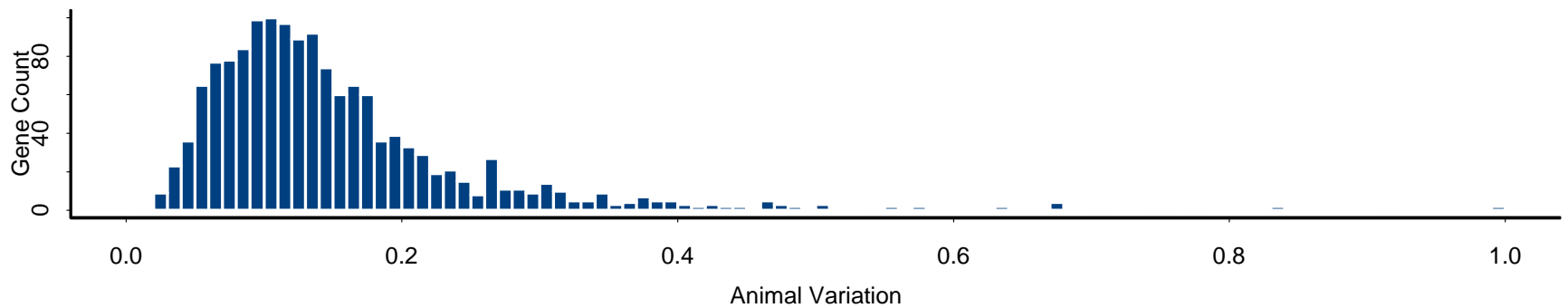
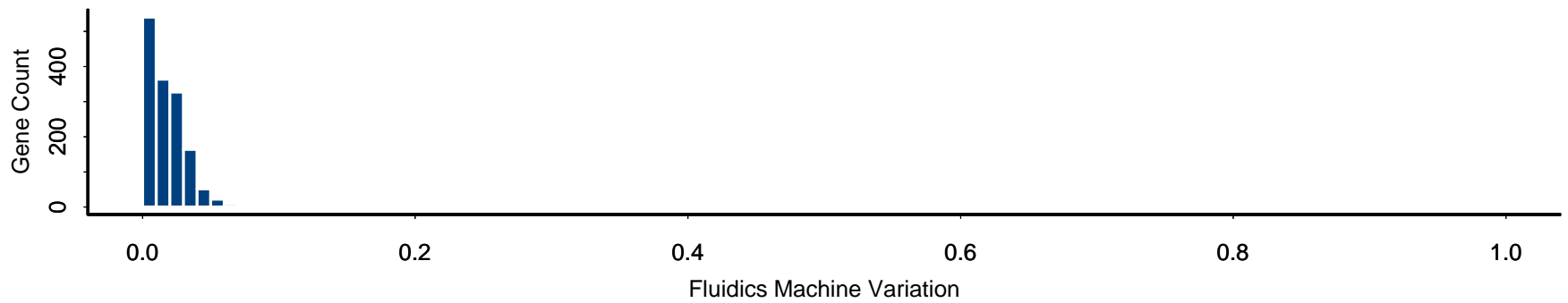
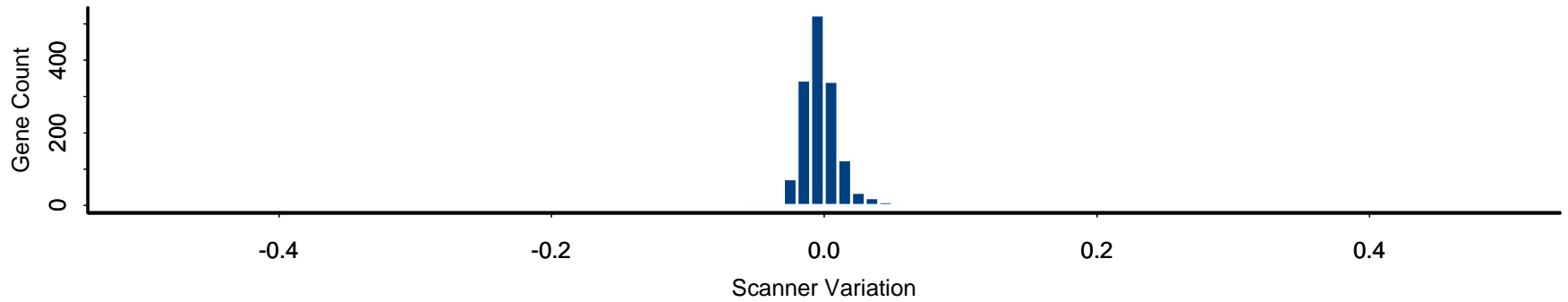
$$\zeta \left(\hat{x}_{Aij} \bar{\theta}_{Ag} + \hat{x}_{Dij} \bar{\theta}_{Dg} \right)$$

- Random error—standard deviation proportional to intensity
- Apply to selected probes

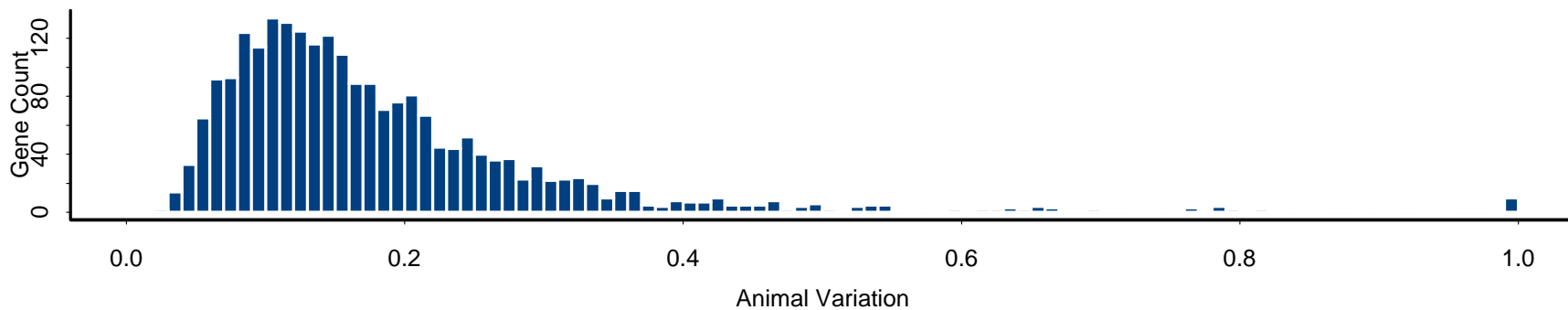
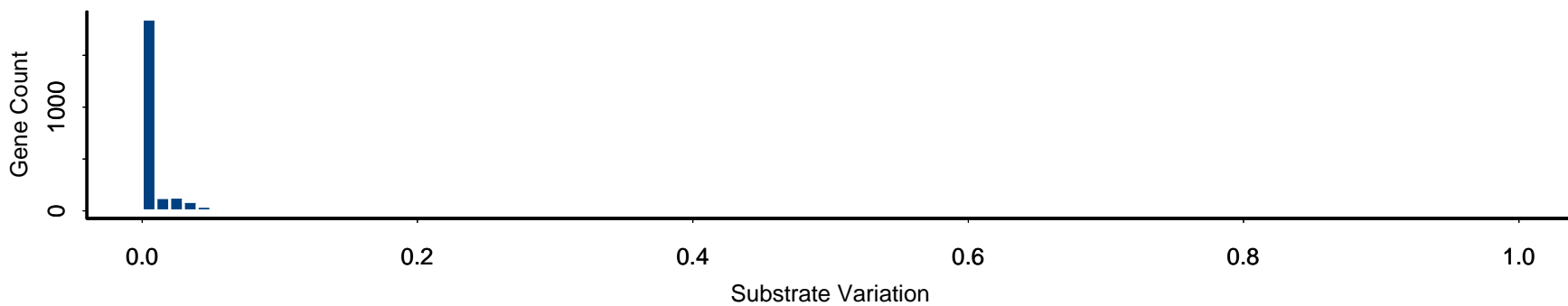
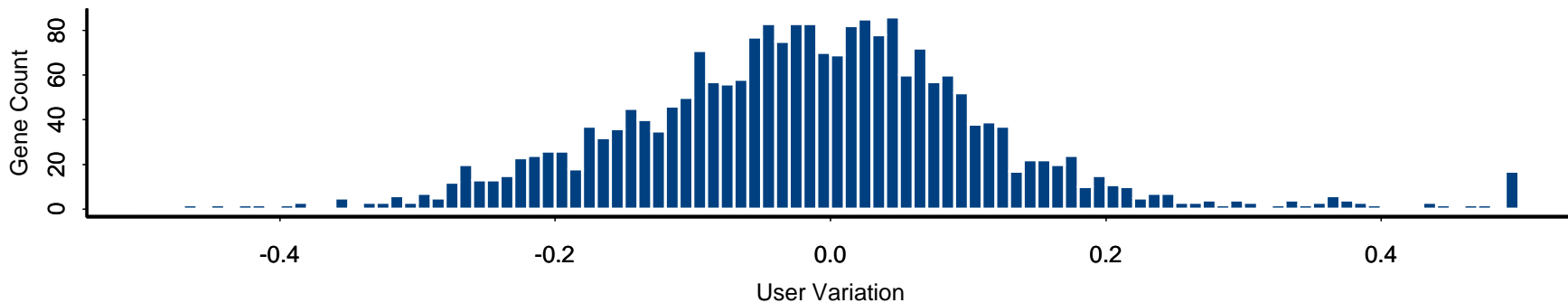
Agilent Model Comparison



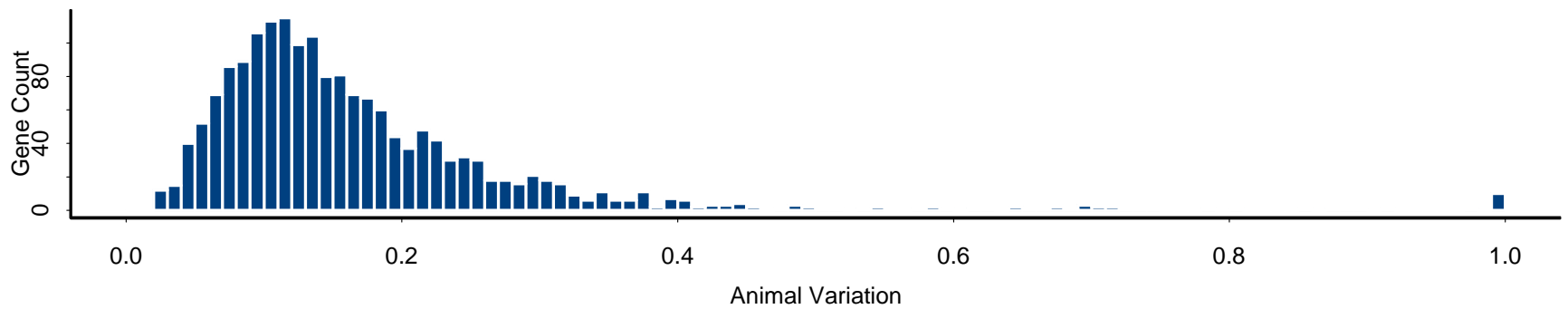
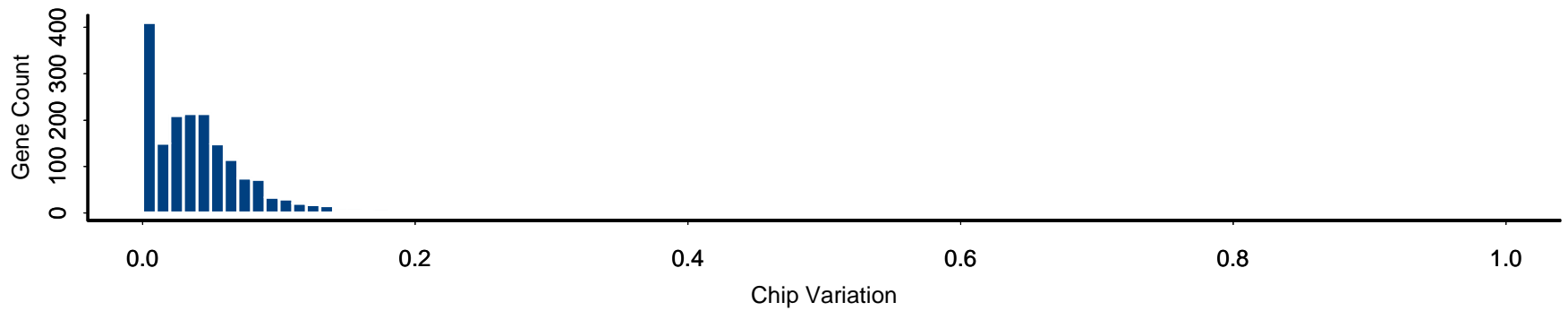
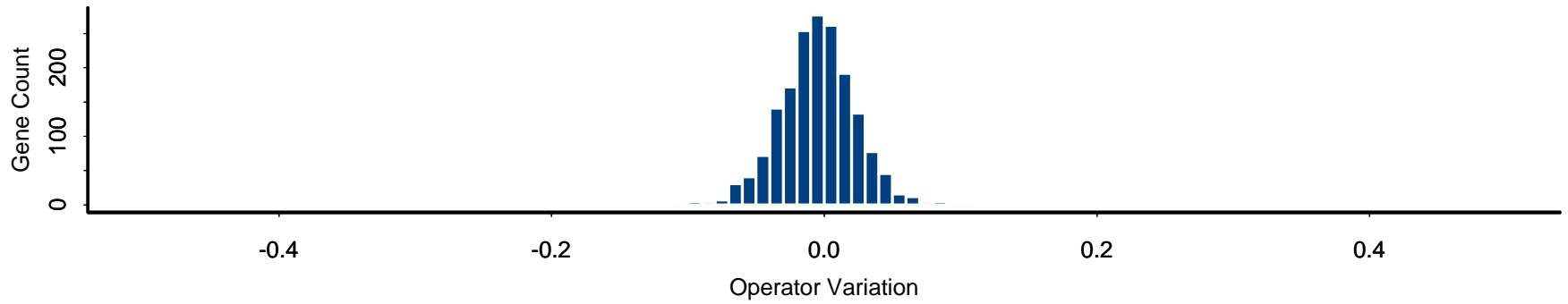
Affymetrix Batch Effects



Agilent Batch Effects



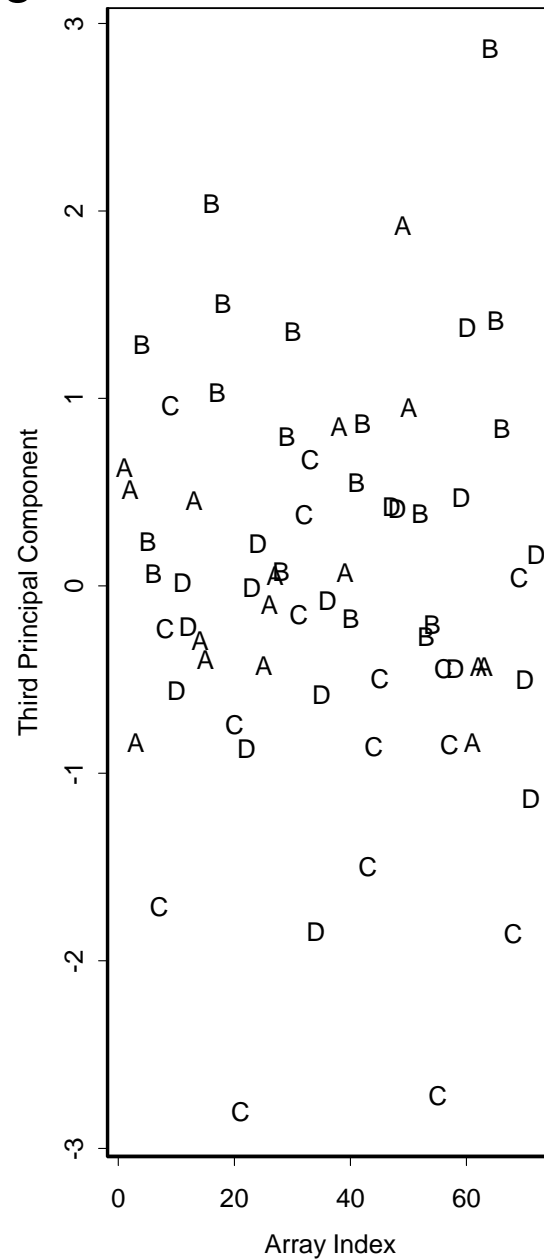
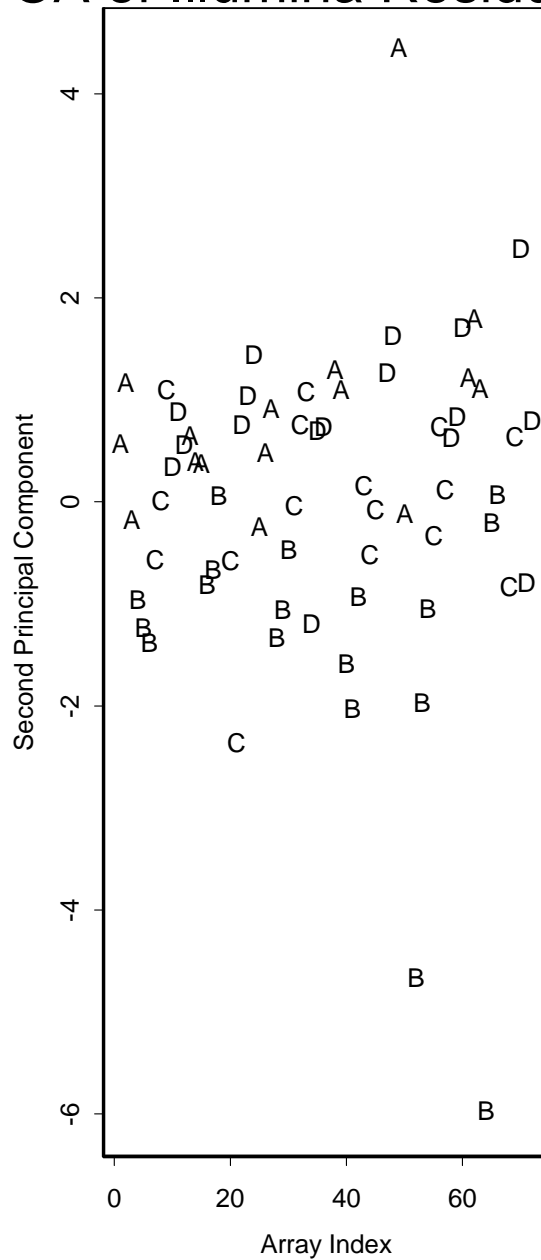
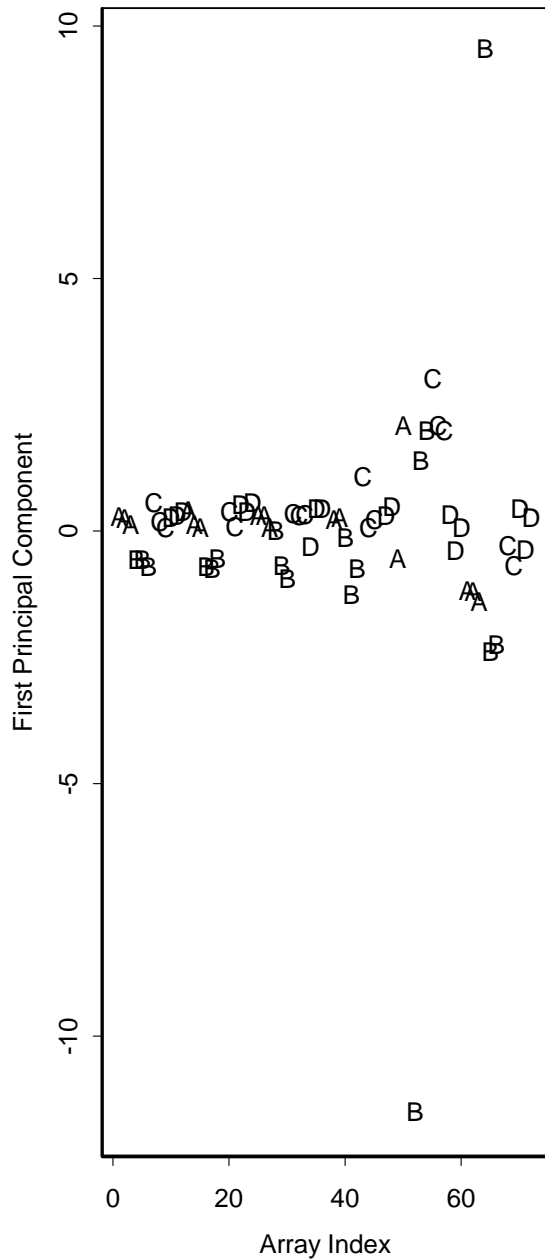
Illumina Batch Effects



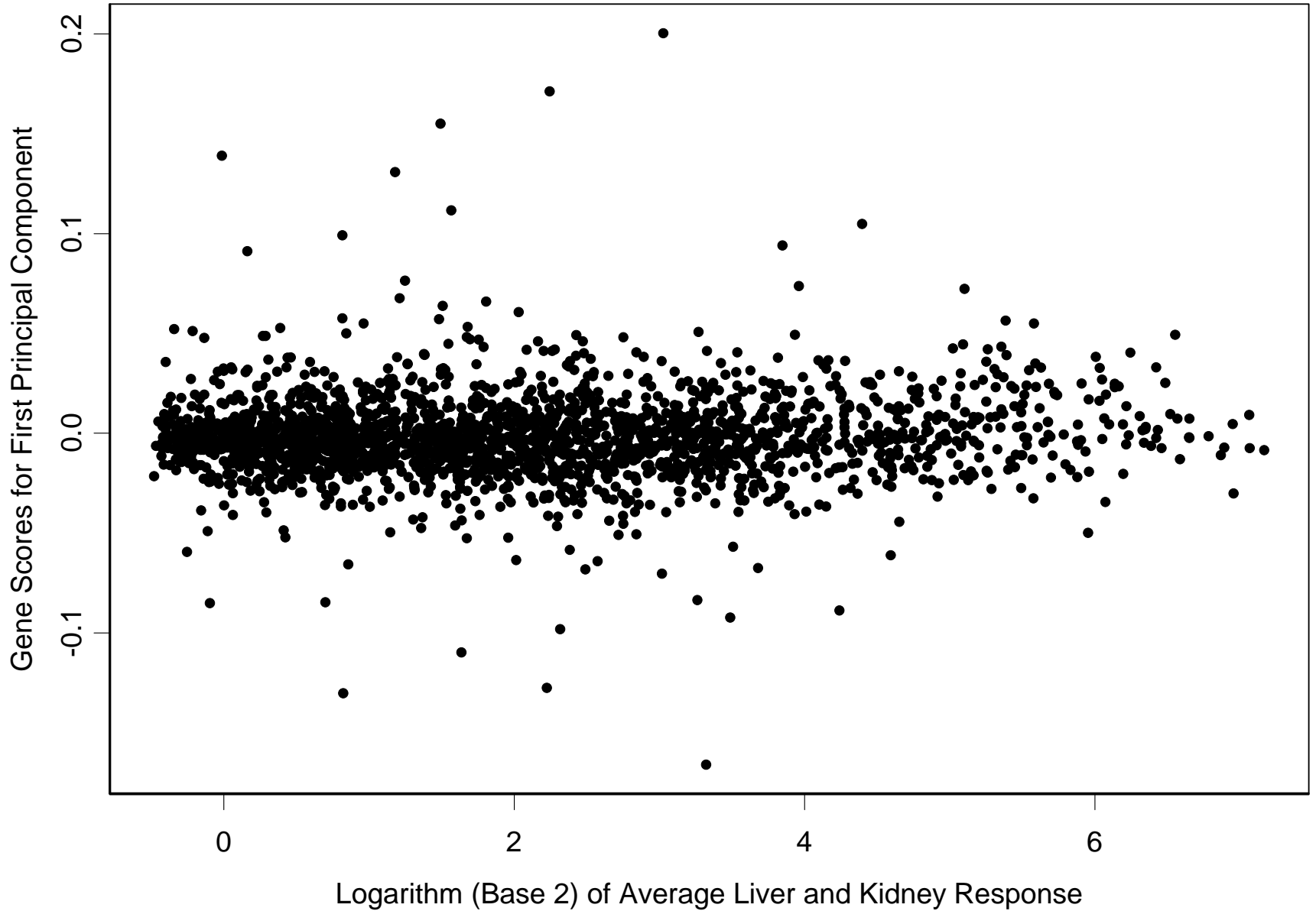
Checking Residuals

- Residuals normalized using the regression weights
- Matrix of residuals: number of probes by number of arrays
- Singular value decomposition
- Graph the largest components

PCA of Illumina Residuals



PCA of Illumina Residuals



Group-by-Group Platform Comparison

- Corresponding to each mRNA transcript is a group of probes that target this transcript
- We compare platforms by comparing corresponding probes, those in a group
- Two comparisons
 - Liver-kidney differential expression
 - Probe sensitivity

Liver-Kidney Differential Expression

- Consider a simple model of the response to a transcript concentration

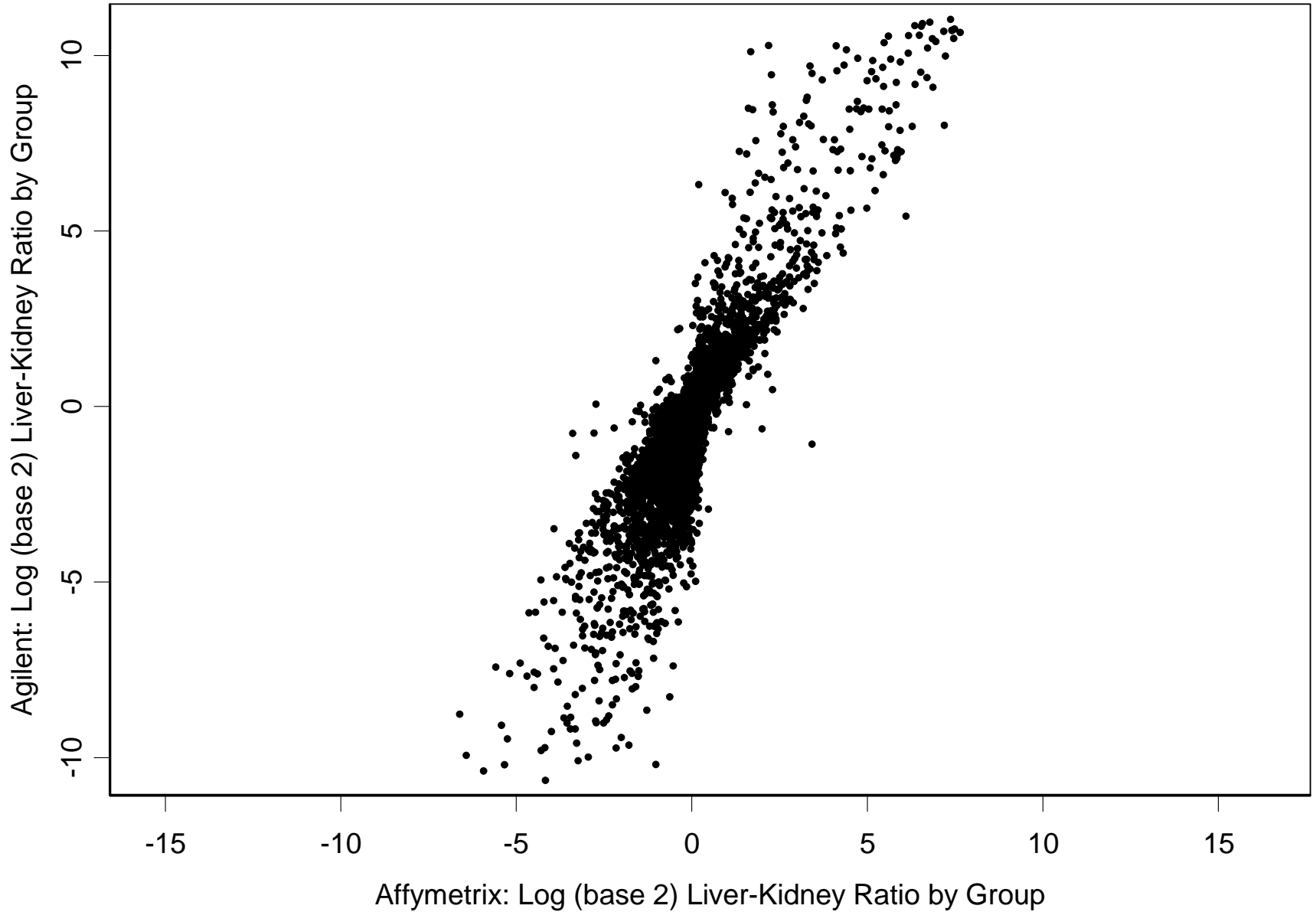
$$\beta(x + \alpha)$$

- Differential expression

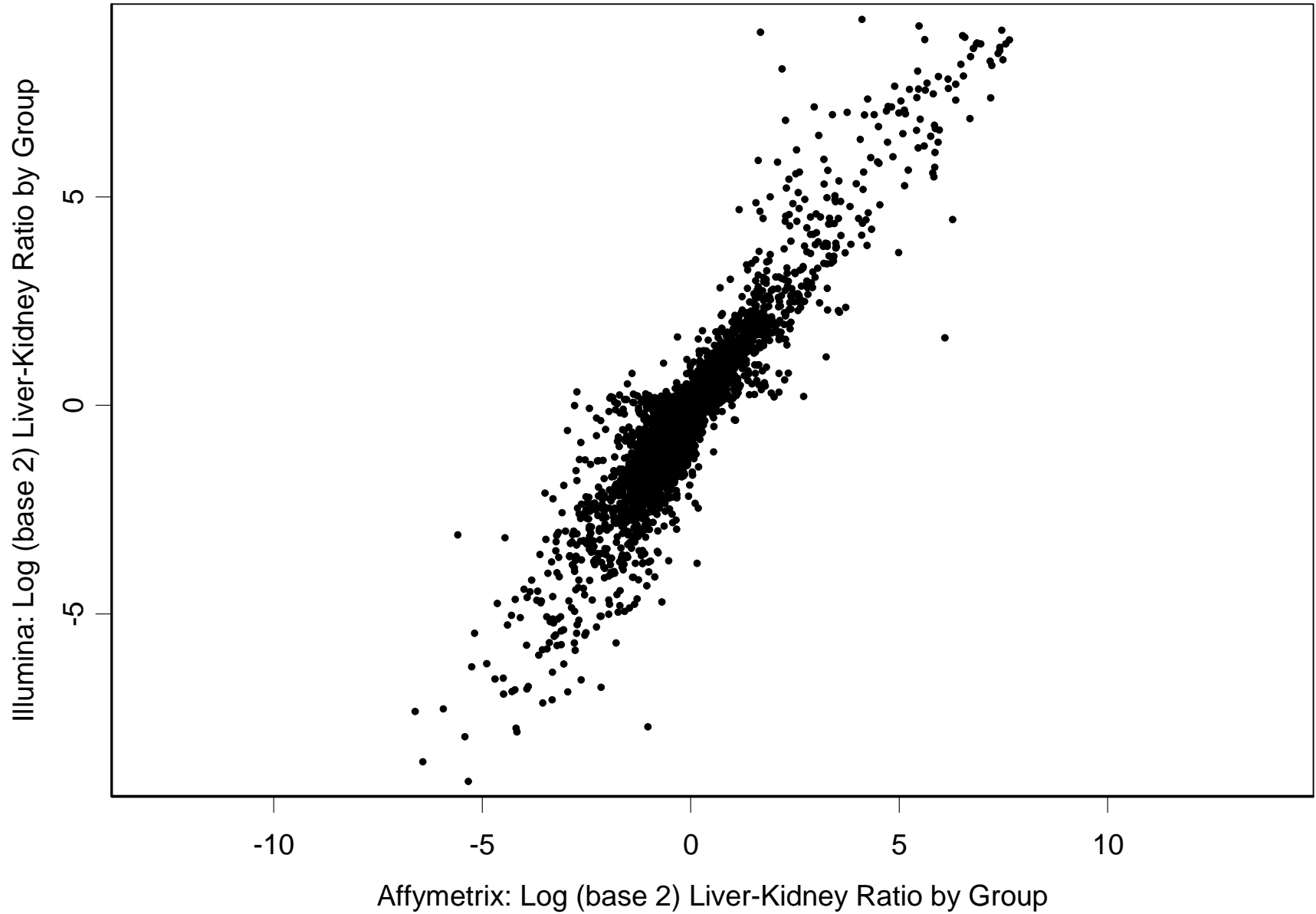
$$\frac{\beta(x_1 + \alpha)}{\beta(x_2 + \alpha)} \approx \frac{x_1}{x_2} \left(1 + \frac{\alpha}{x_1} - \frac{\alpha}{x_2} \right)$$

- Platform comparison depends on the difference in concentration and the difference in the model intercept

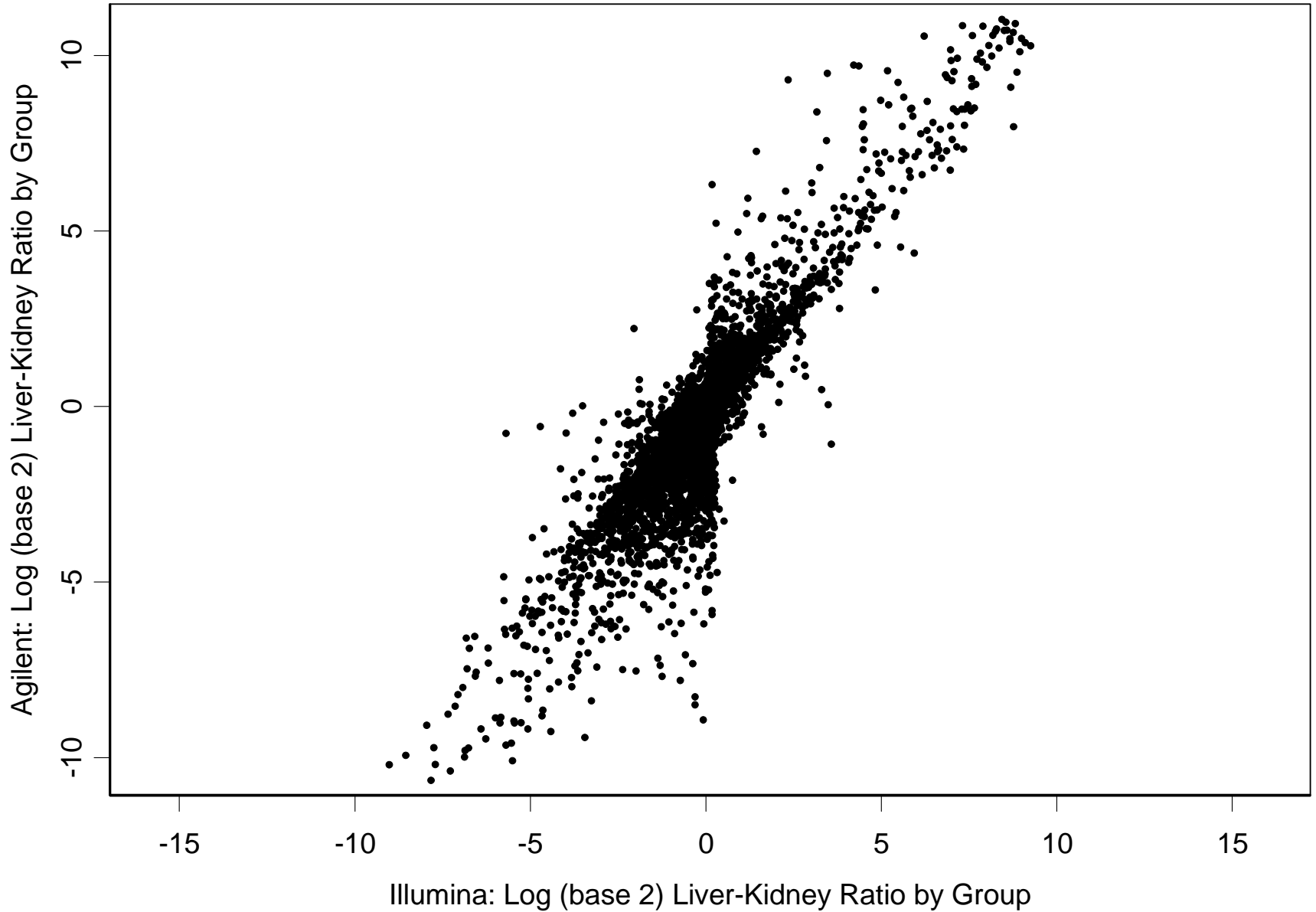
Liver-Kidney Differential Expression



Liver-Kidney Differential Expression



Liver-Kidney Differential Expression



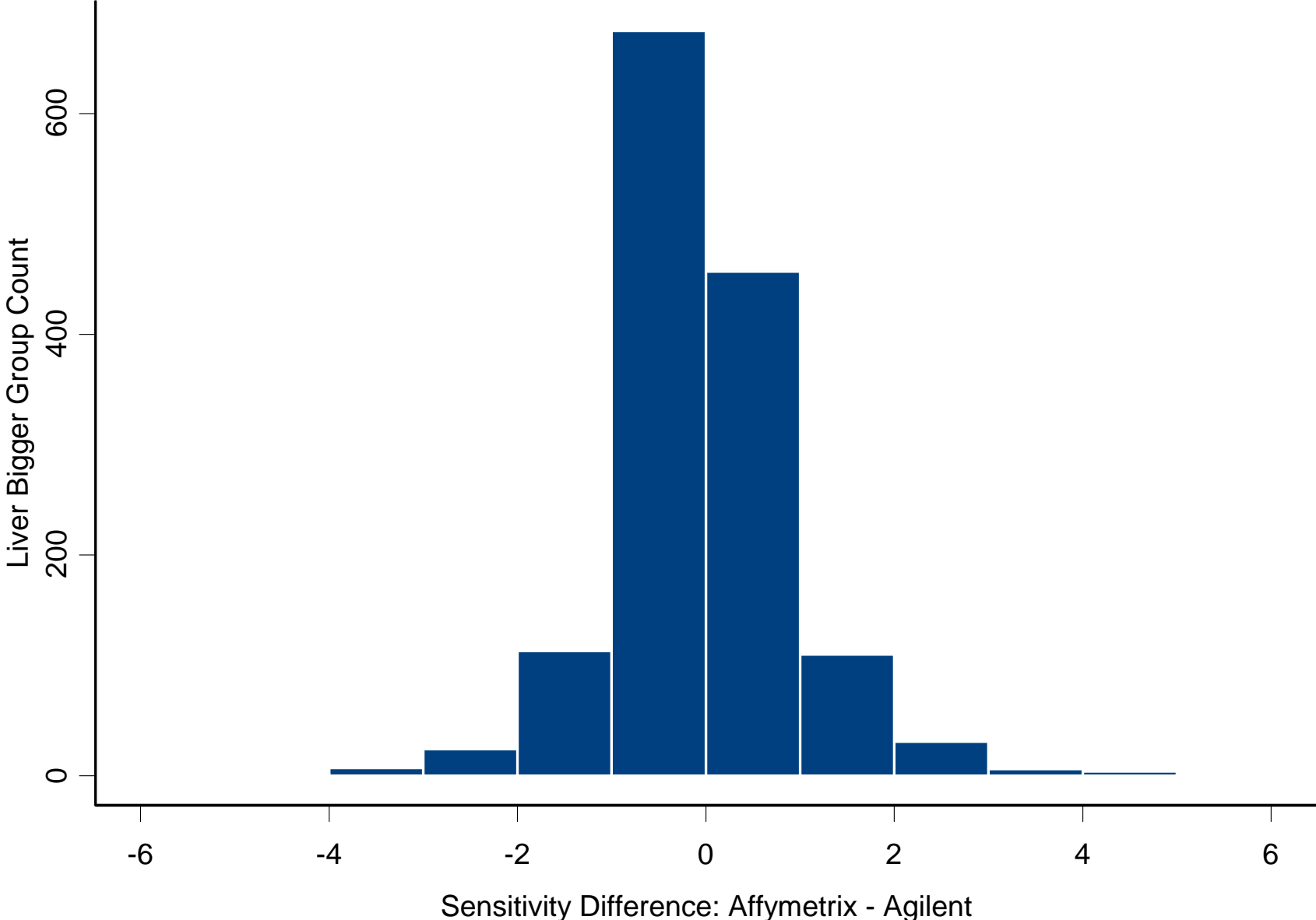
Probe Sensitivity

- Sensitivity is the intensity increase with a unit increase in the transcript concentration
- Need to make the sensitivity dimensionless
- A possibility: Divide by the standard deviation of the animal variation
- Obtain the sensitivity in units of animal variation
- Platform comparison is possible because samples with the same transcript concentration are measured

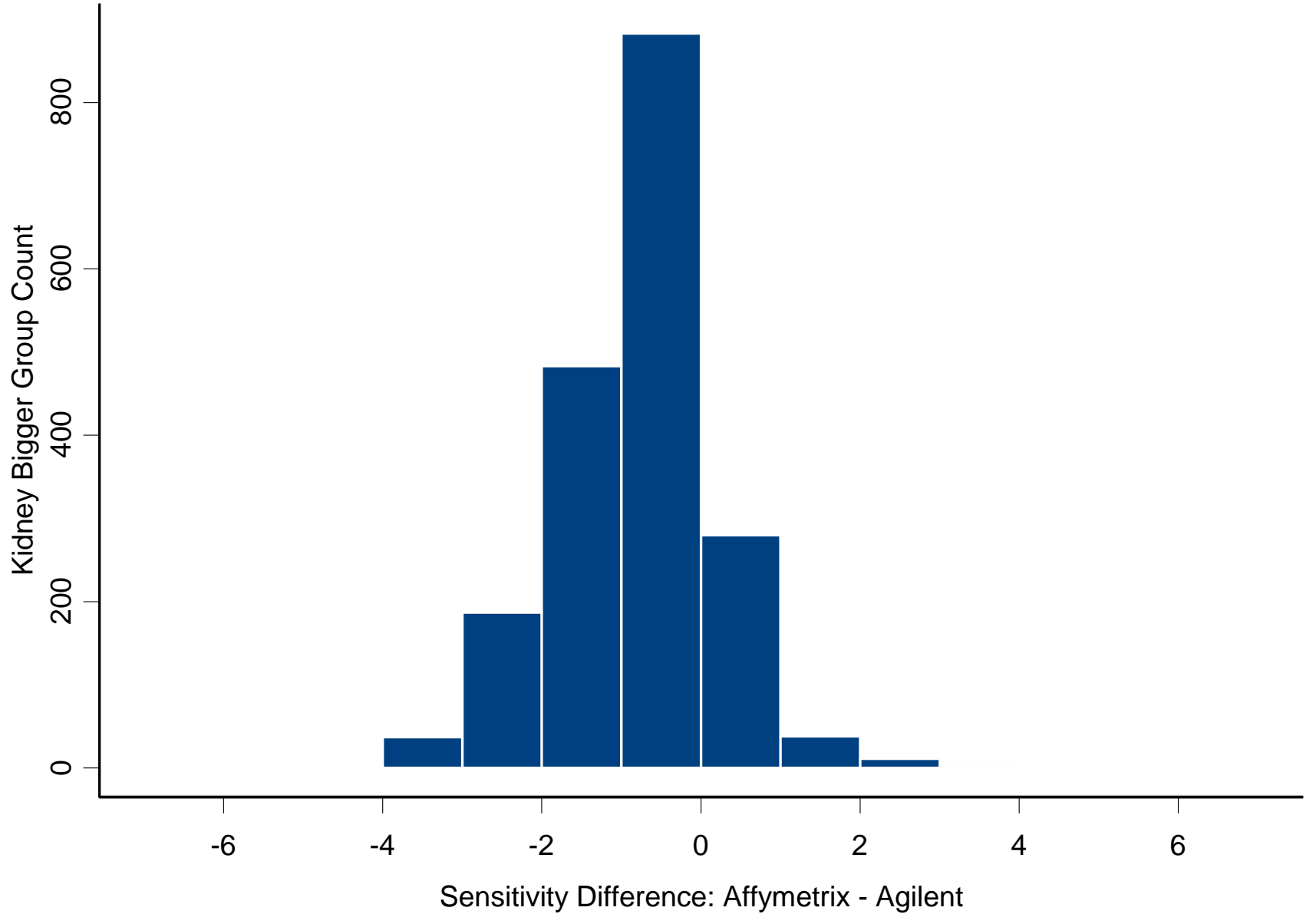
Probe Sensitivity (con't)

- Modeling gives
 - Animal-average slope from liver and kidney fit
 - Standard deviation from the animal variation
- Consider only groups for which every probe (over all platforms) has either the liver bigger than the kidney or the kidney bigger than the liver
- Display differences of the \log_{10} sensitivities
- Note the large variation with platform

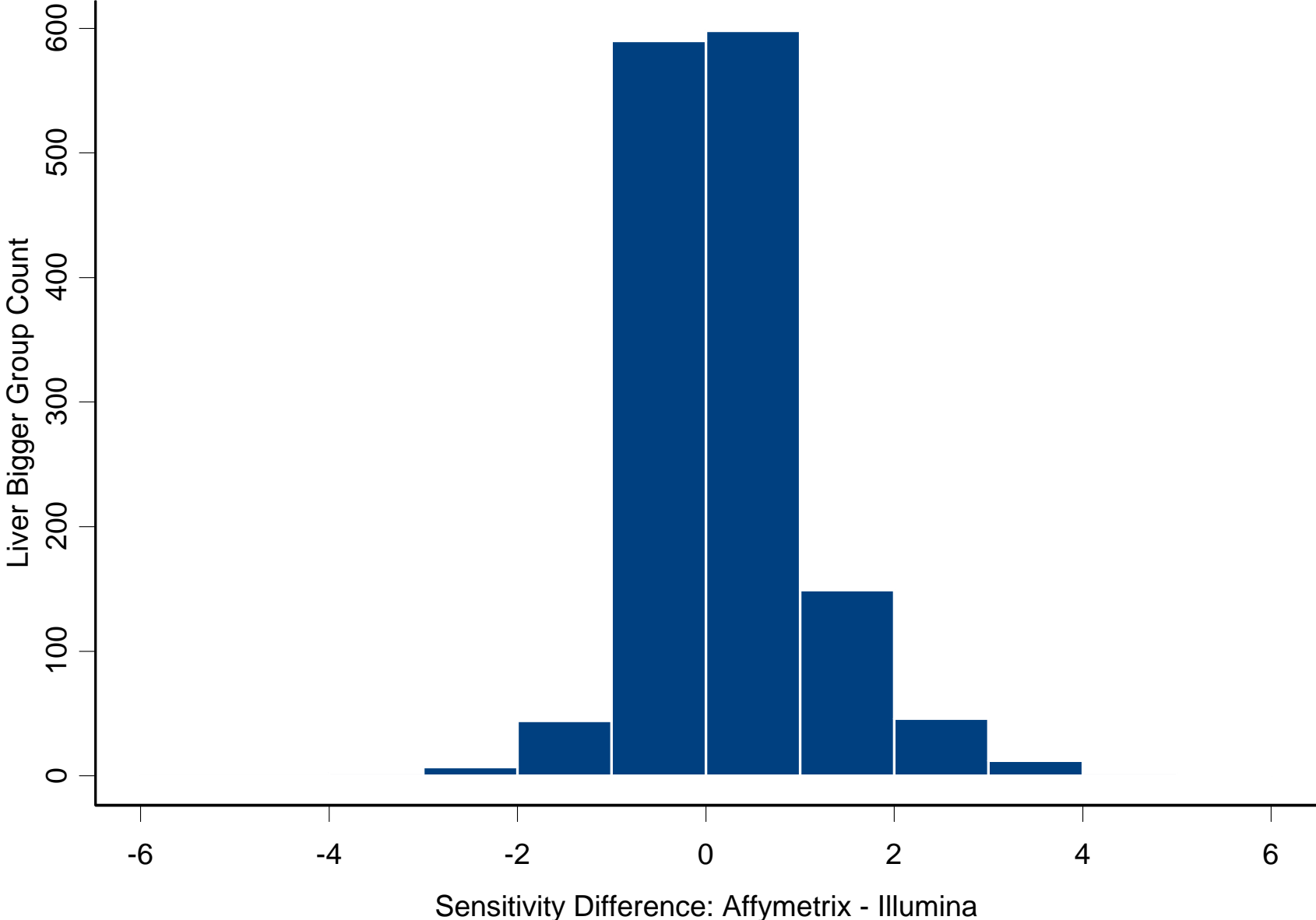
Group-by-Group Comparison of Measurement Sensitivity



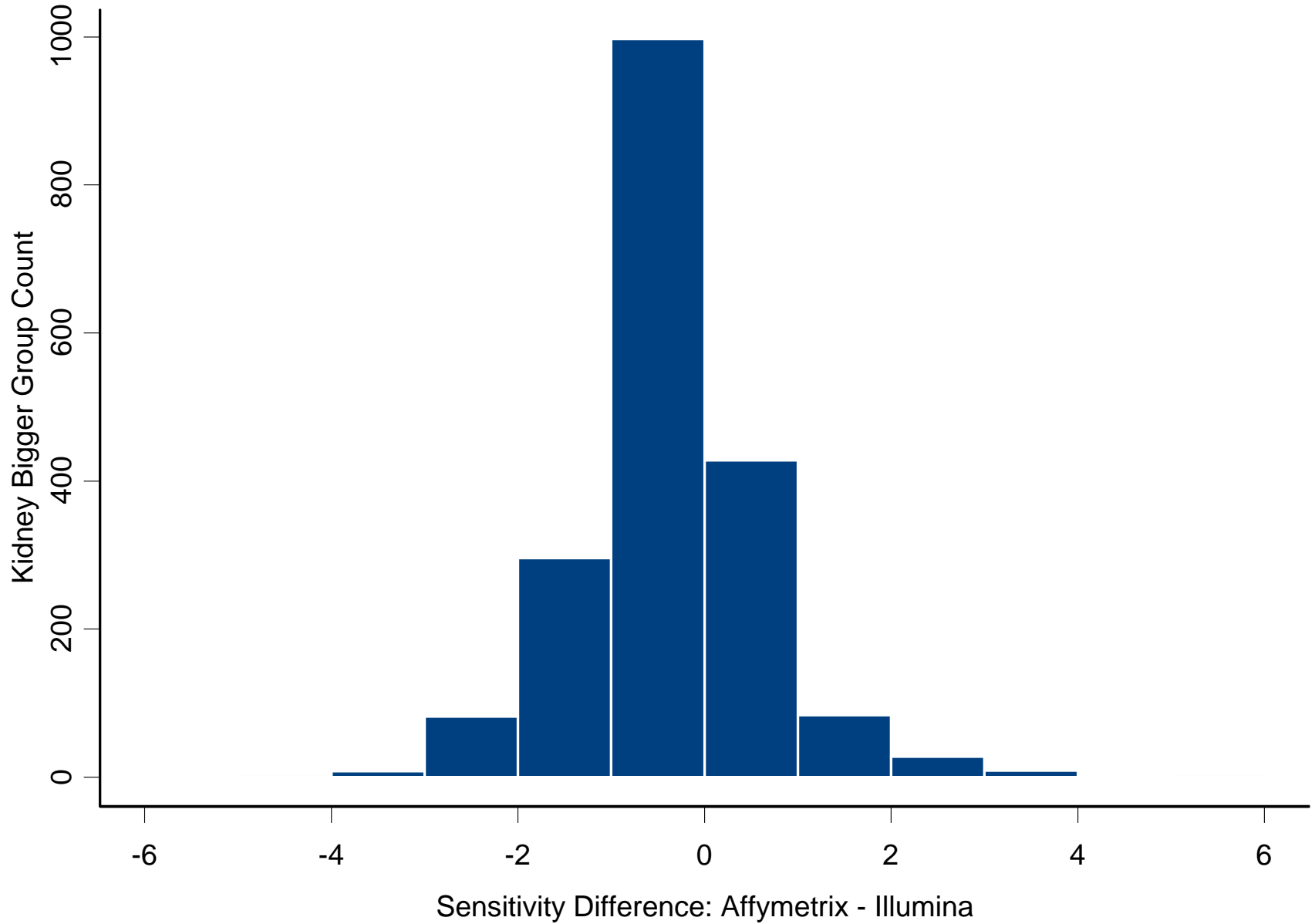
Group-by-Group Comparison of Measurement Sensitivity



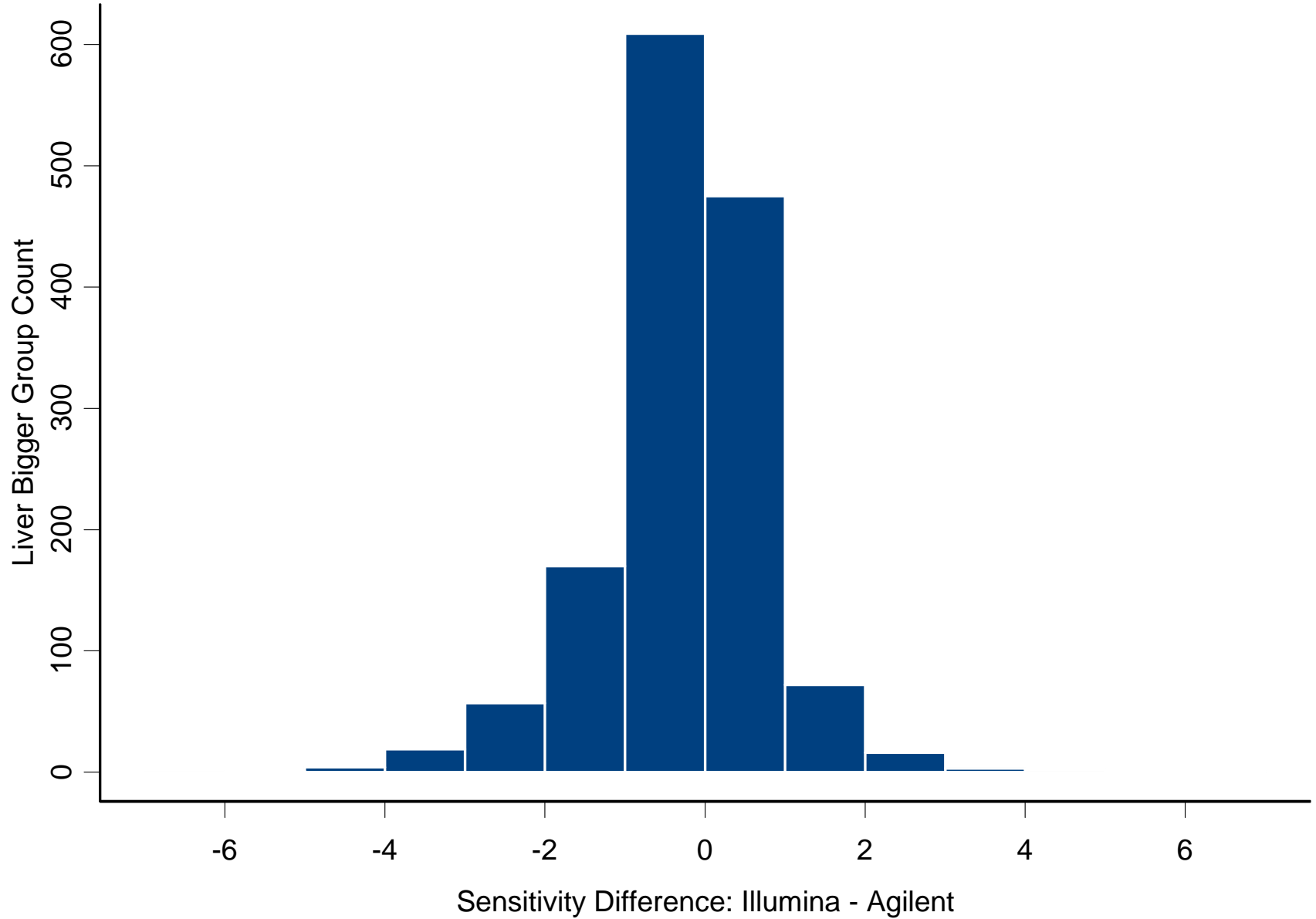
Group-by-Group Comparison of Measurement Sensitivity



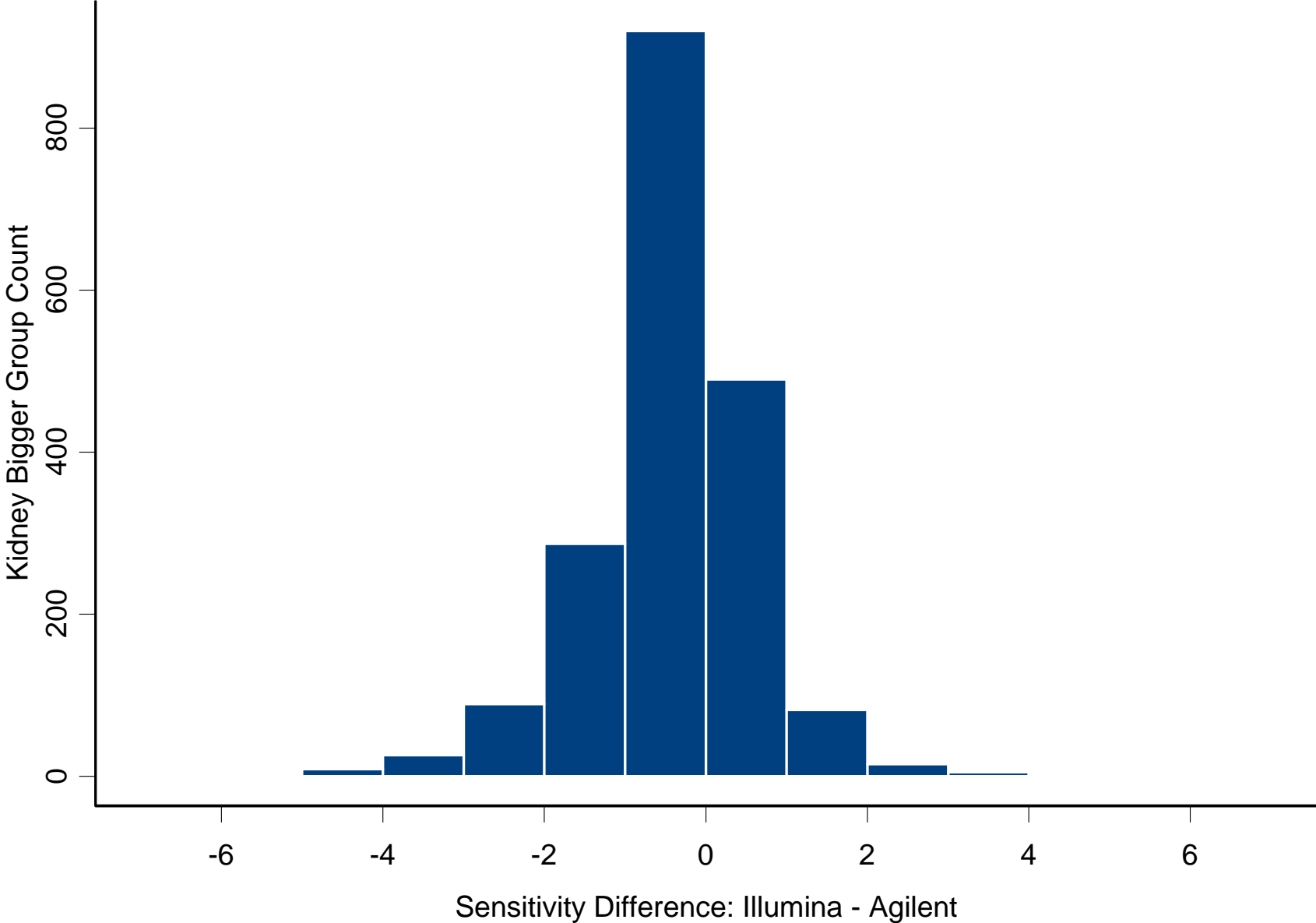
Group-by-Group Comparison of Measurement Sensitivity



Group-by-group Comparison of Measurement Sensitivity



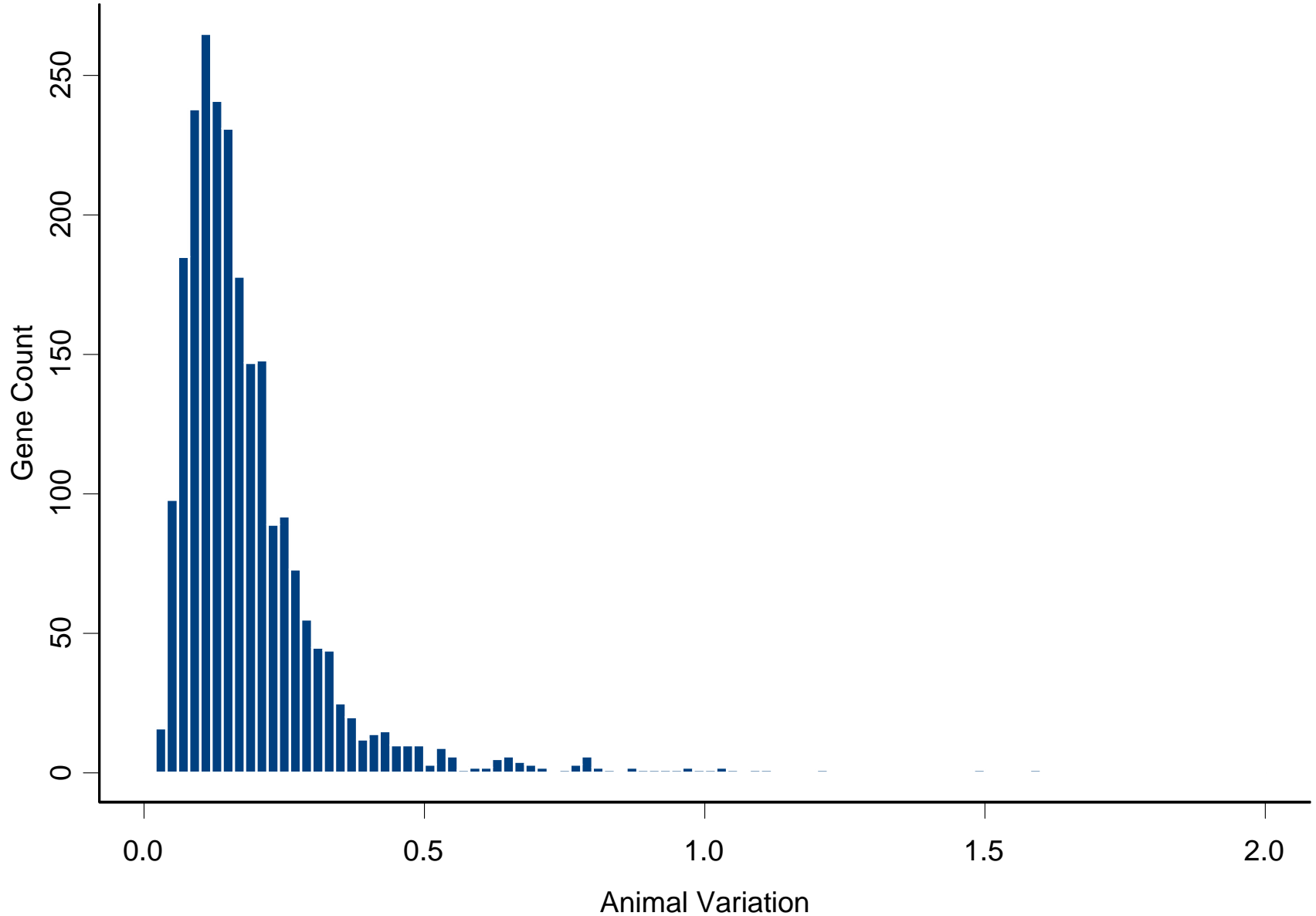
Group-by-Group Comparison of Measurement Sensitivity



Gene Lists

- Animal variation shows general need to use the t statistic in developing a gene list
- The validity of the t statistic can be checked through the animal replicates
- The reproducibility of gene lists can be assessed by adding various fractions of the kidney-liver difference

Agilent Animal Variation

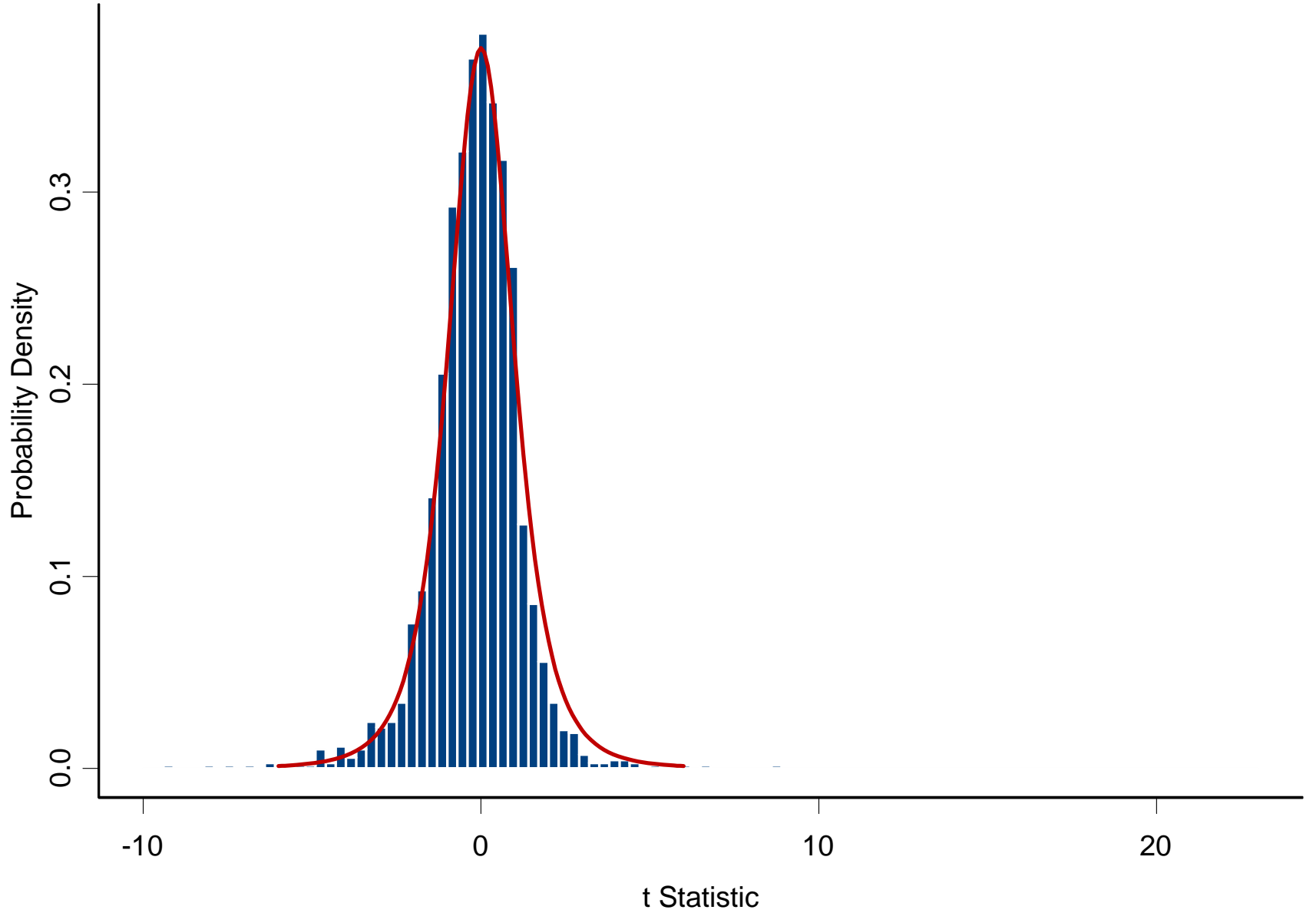


Large Differential Expression

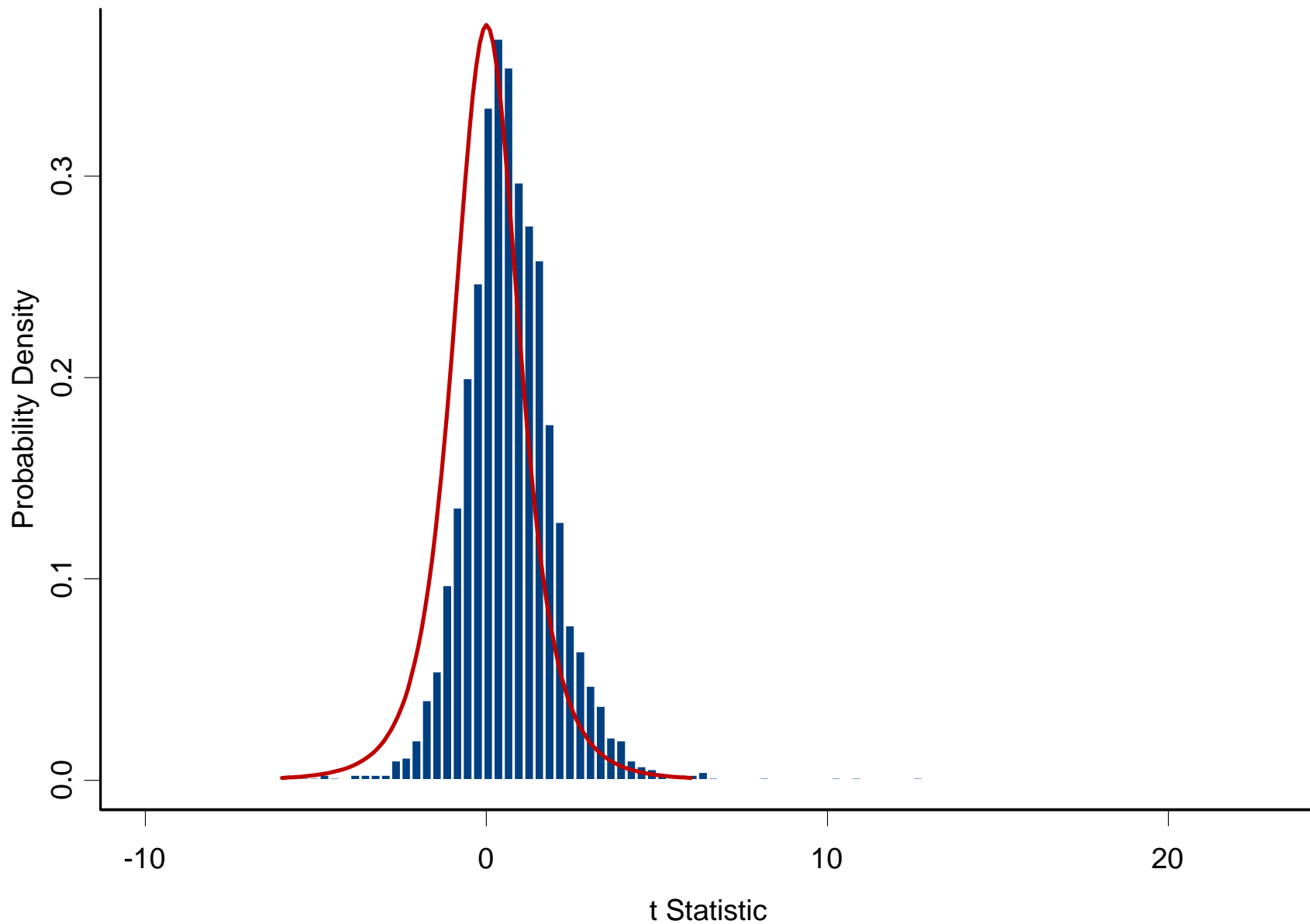
$$\log_2\left(\frac{\mu_1 + \delta_1}{\mu_2 + \delta_2}\right) \approx \log_2\left(\frac{\mu_1}{\mu_2}\right) + \frac{1}{\log_e(2)}\left(\frac{\delta_1}{\mu_1} - \frac{\delta_2}{\mu_2}\right)$$

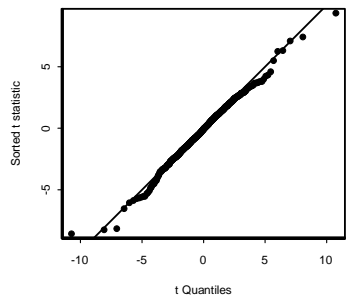
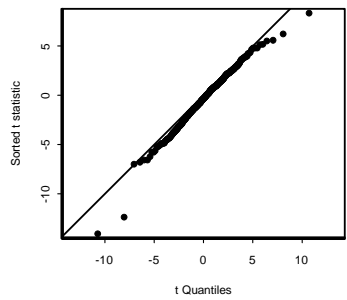
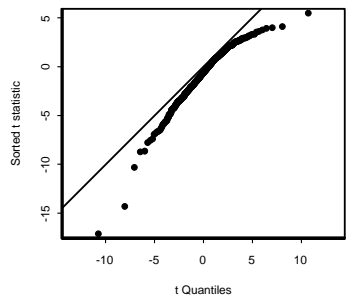
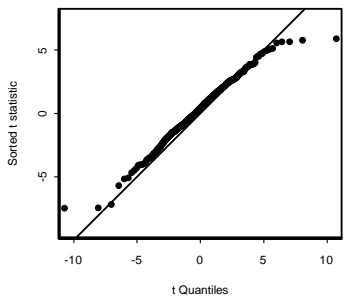
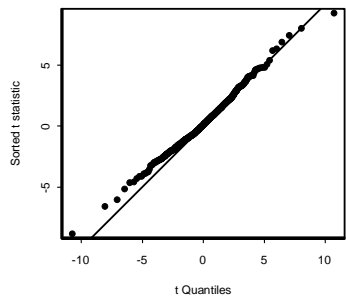
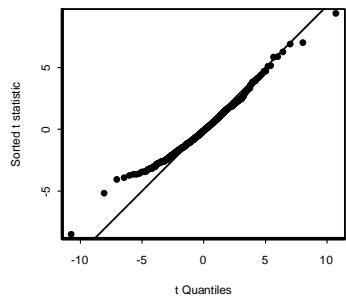
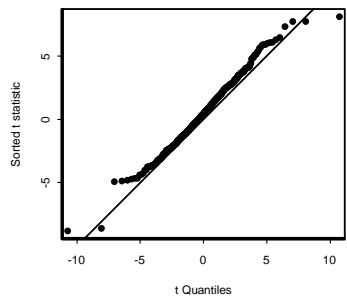
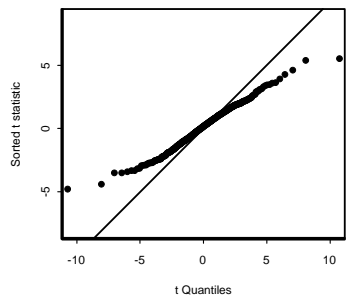
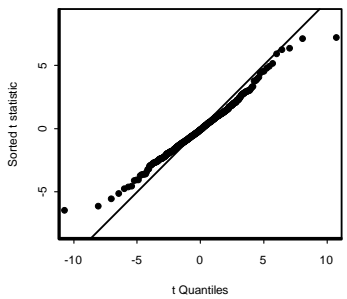
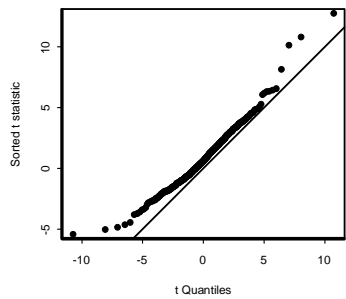
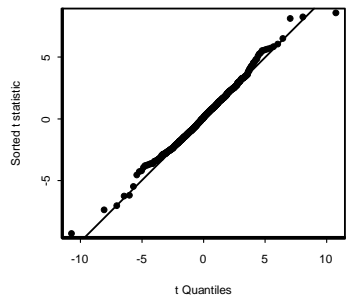
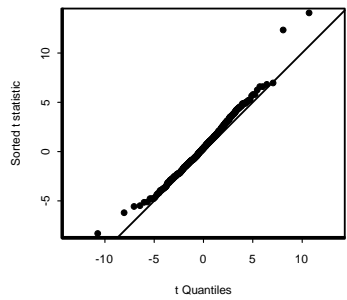
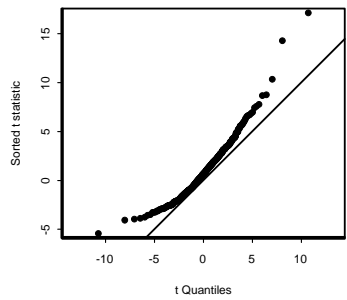
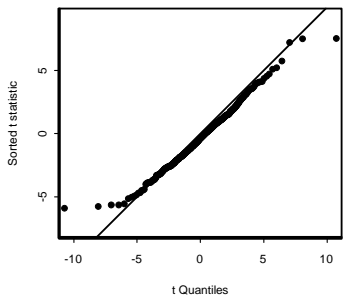
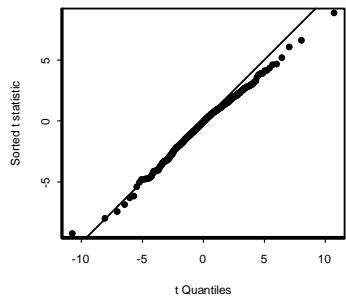
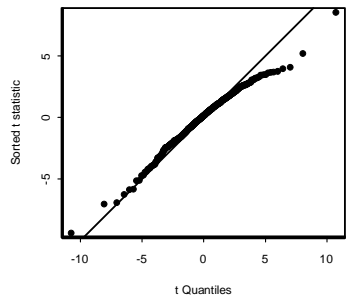
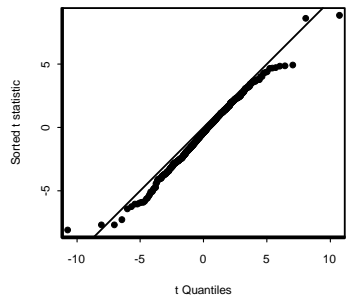
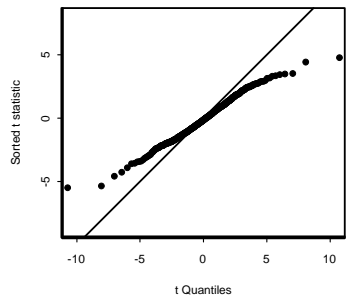
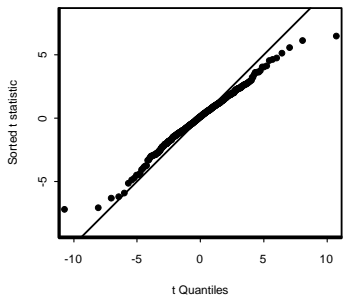
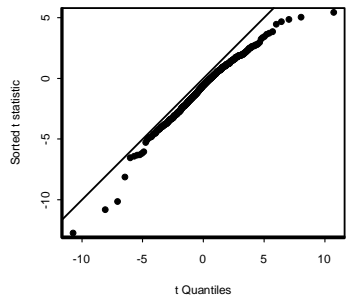
- Animal variation shows standard deviation of δ / μ
- Noise in differential expression depends on probe
- For 4-fold differential expression, one might not have to quantify the animal variation

t Statistic--t Distribution Comparison



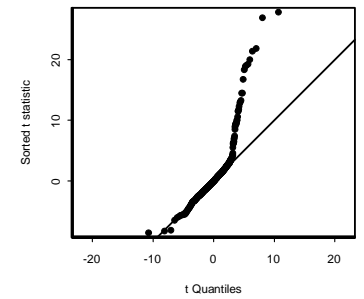
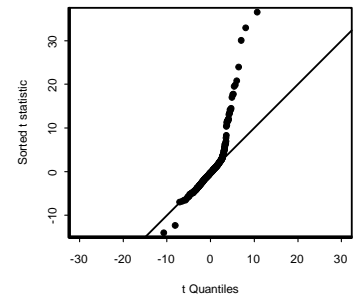
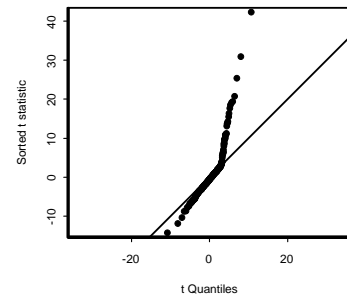
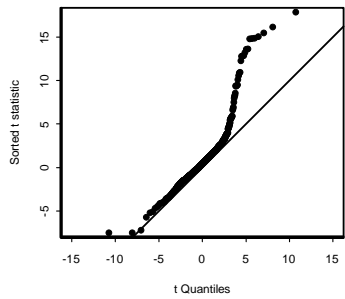
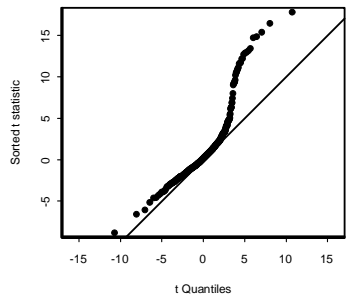
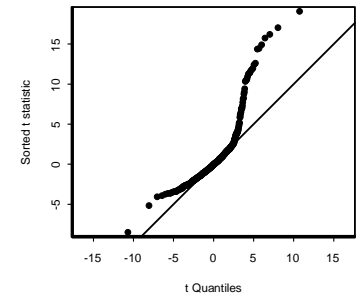
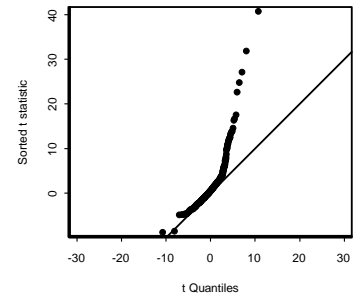
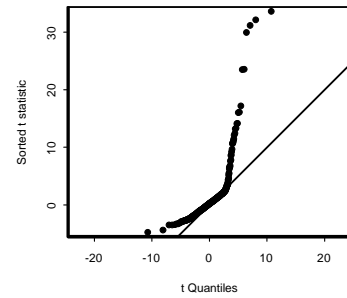
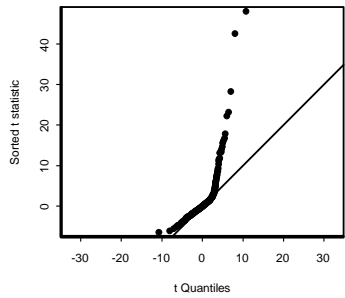
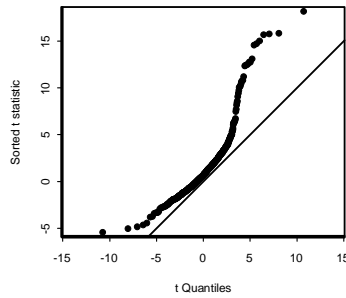
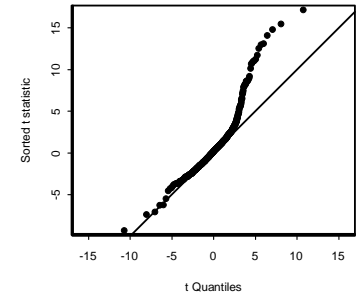
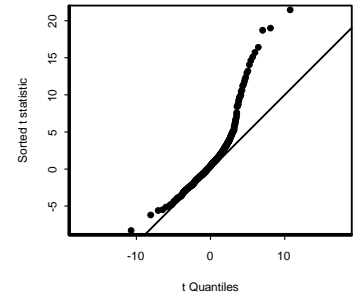
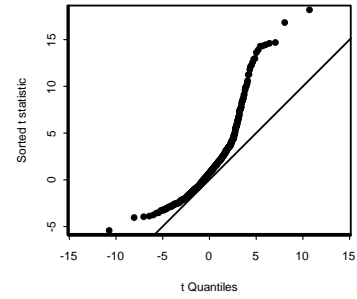
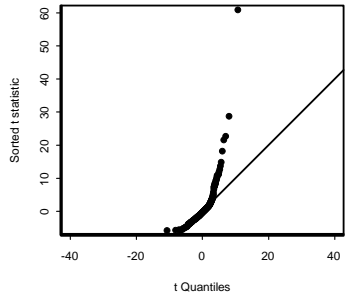
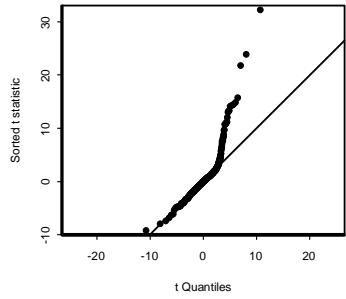
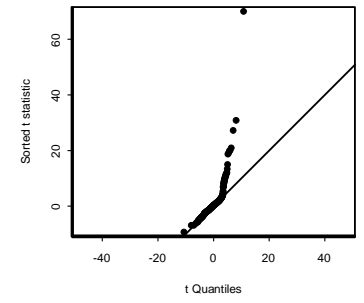
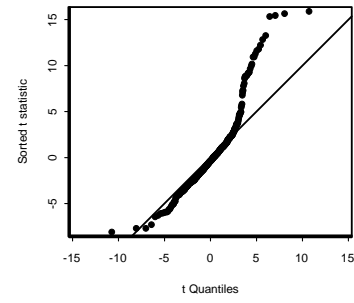
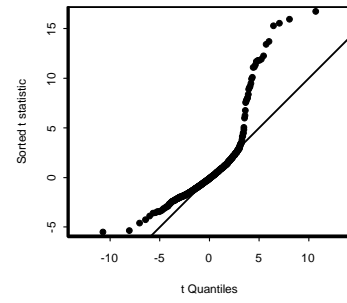
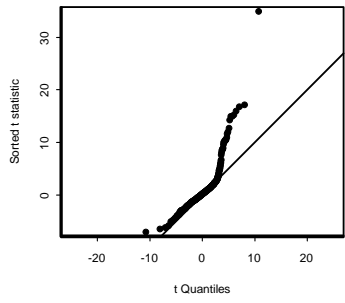
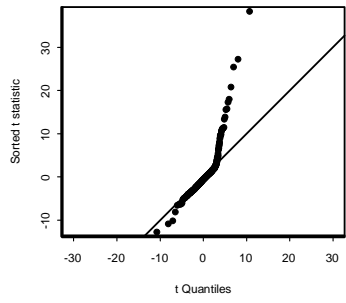
t Statistic--t Distribution Comparison



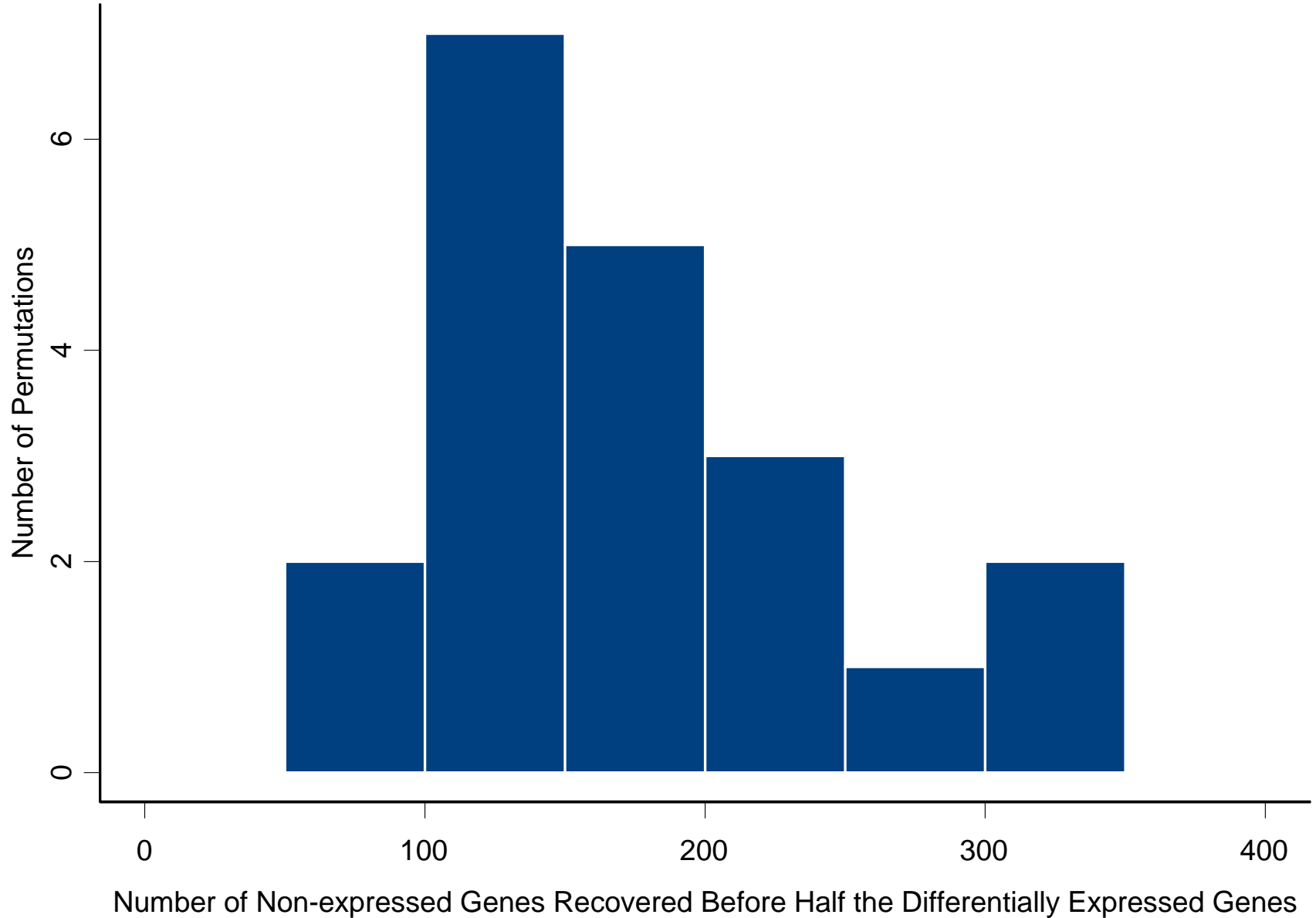


Gene List Reproducibility

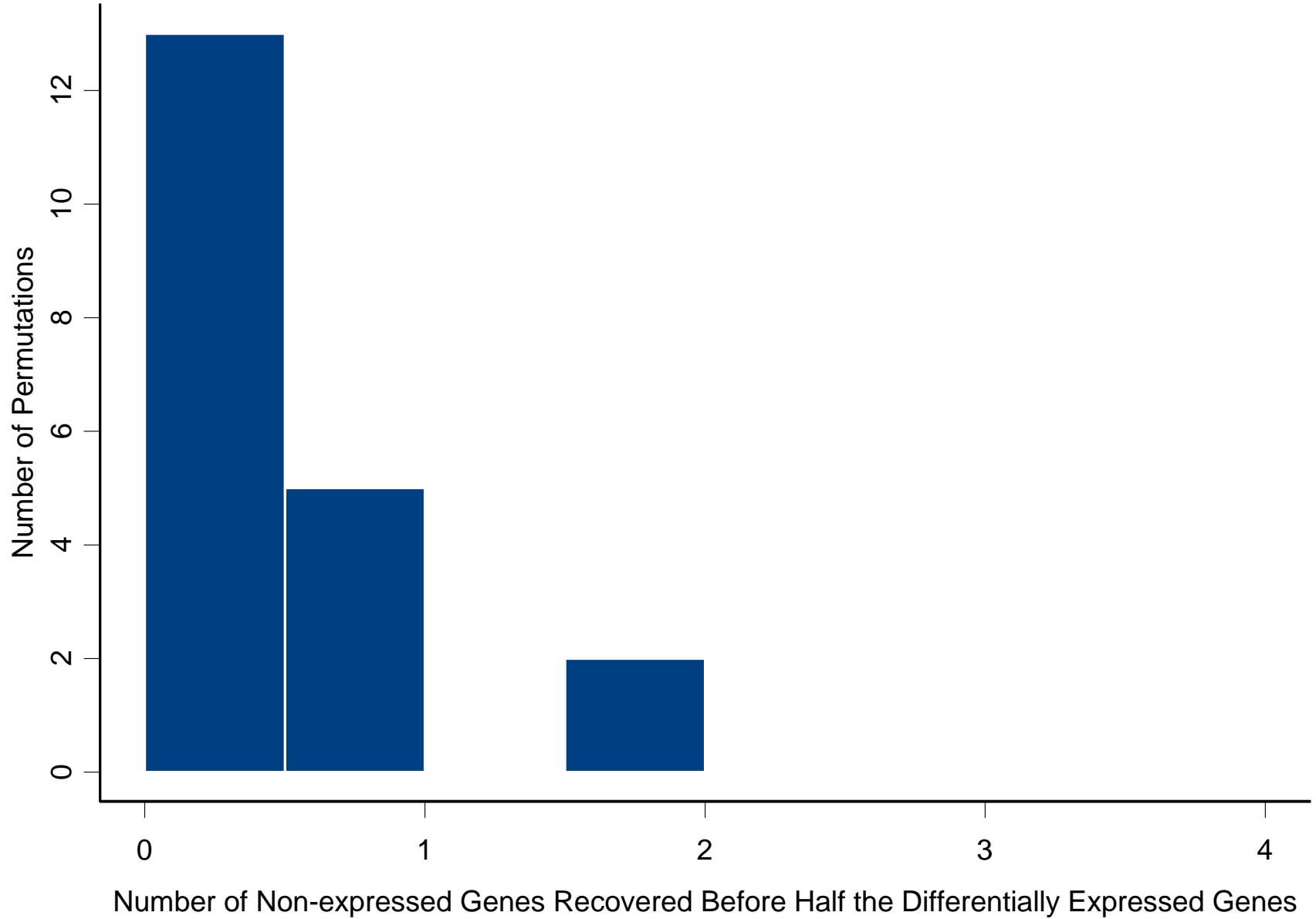
- Choose 100 probes with largest kidney – liver difference
- For these probes, replace 1% of liver with 1% of kidney
- This gives us measurements with 100 differentially expressed probes
- Consider the location of these probes in gene lists



Gene List Reproducibility--1 Percent Kidney



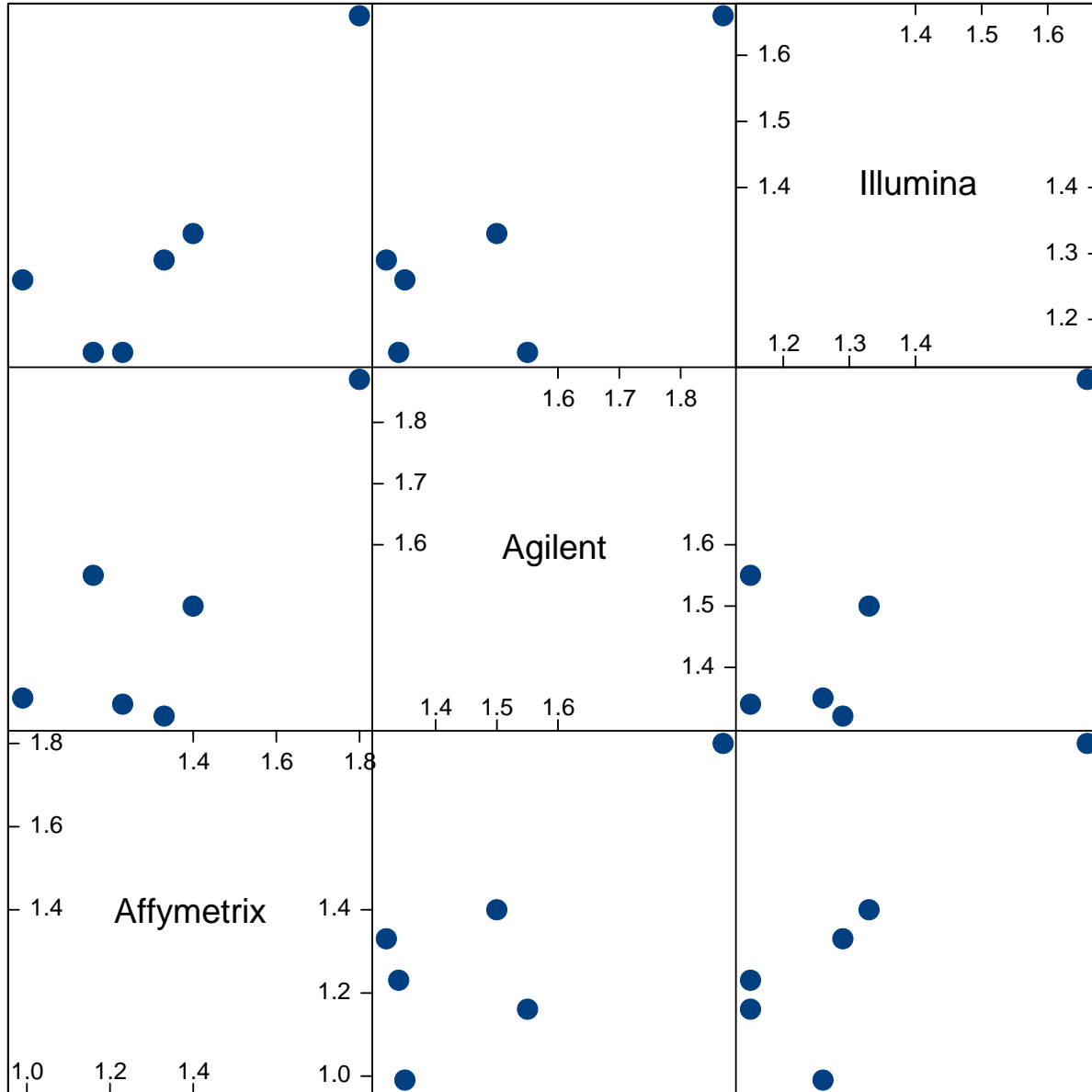
Gene List Reproducibility--10 Percent Kidney



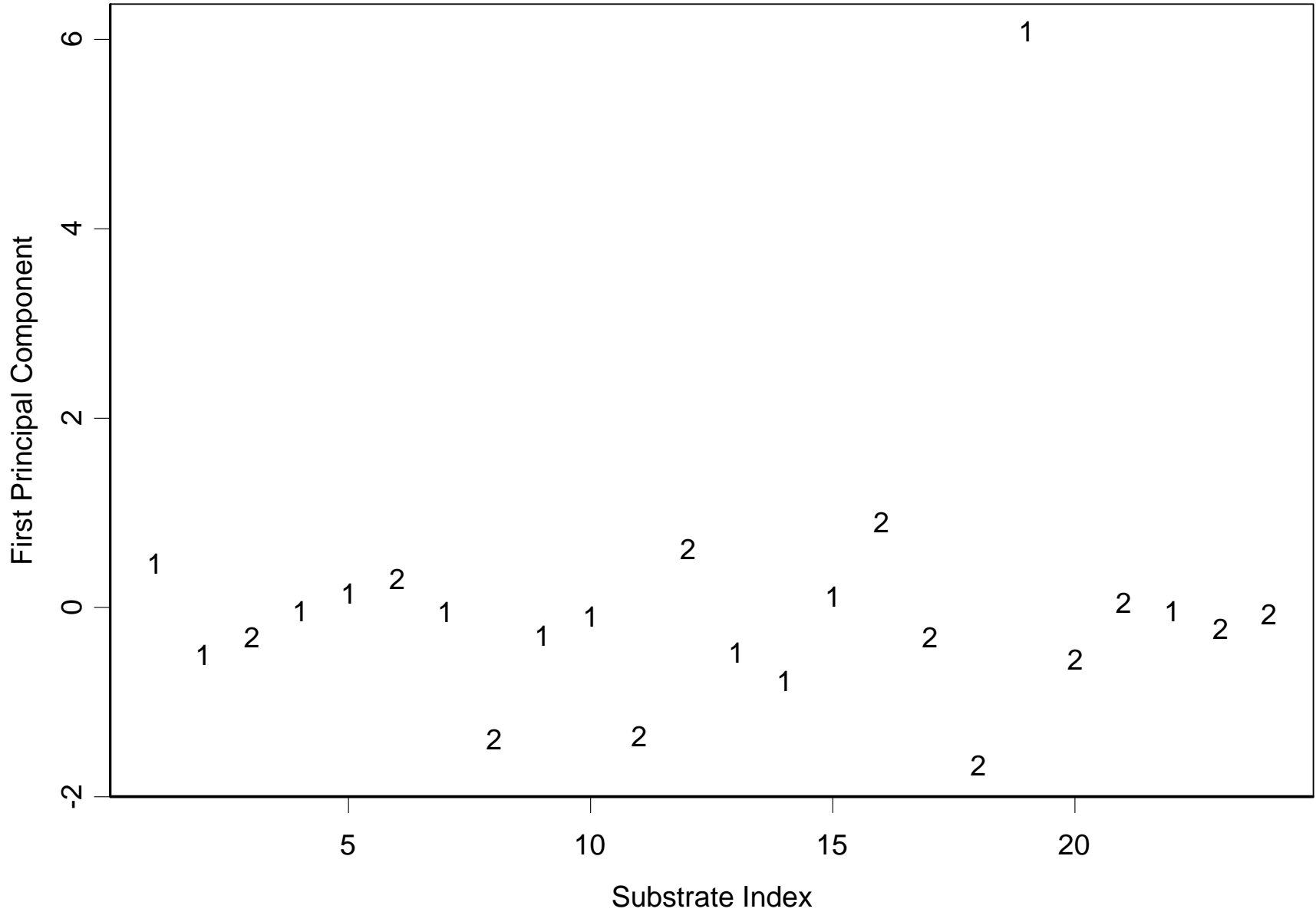
Conclusions

- Our experiment might be repeated to gain insight into gene expression measurement systems, for example, to qualify a measurement system for a study
- Extension of our experiment
 - More animals to provide more insight into gene-list reproducibility
 - More animal organs (brain, for example) to provide better coverage of the probes

mRNA Kidney-Liver Concentration Ratio



Principal Components Analysis of Agilent Substrate (Random) Effect



Principal Components Analysis of Agilent Substrate (Random) Effect

