Variance Stabilizing Transformations for image-based compound profiling features

LEonard Wafula

August 31, 2017

Introduction

☐ Proxy biological method for distinguishing compounds using a range of features extracted from image-based assays

☐ The features provide information on

i. intracellular biomarkers: texture, intensity, spatial distribution ...
ii. cells: shape, geometry, quantity .. .

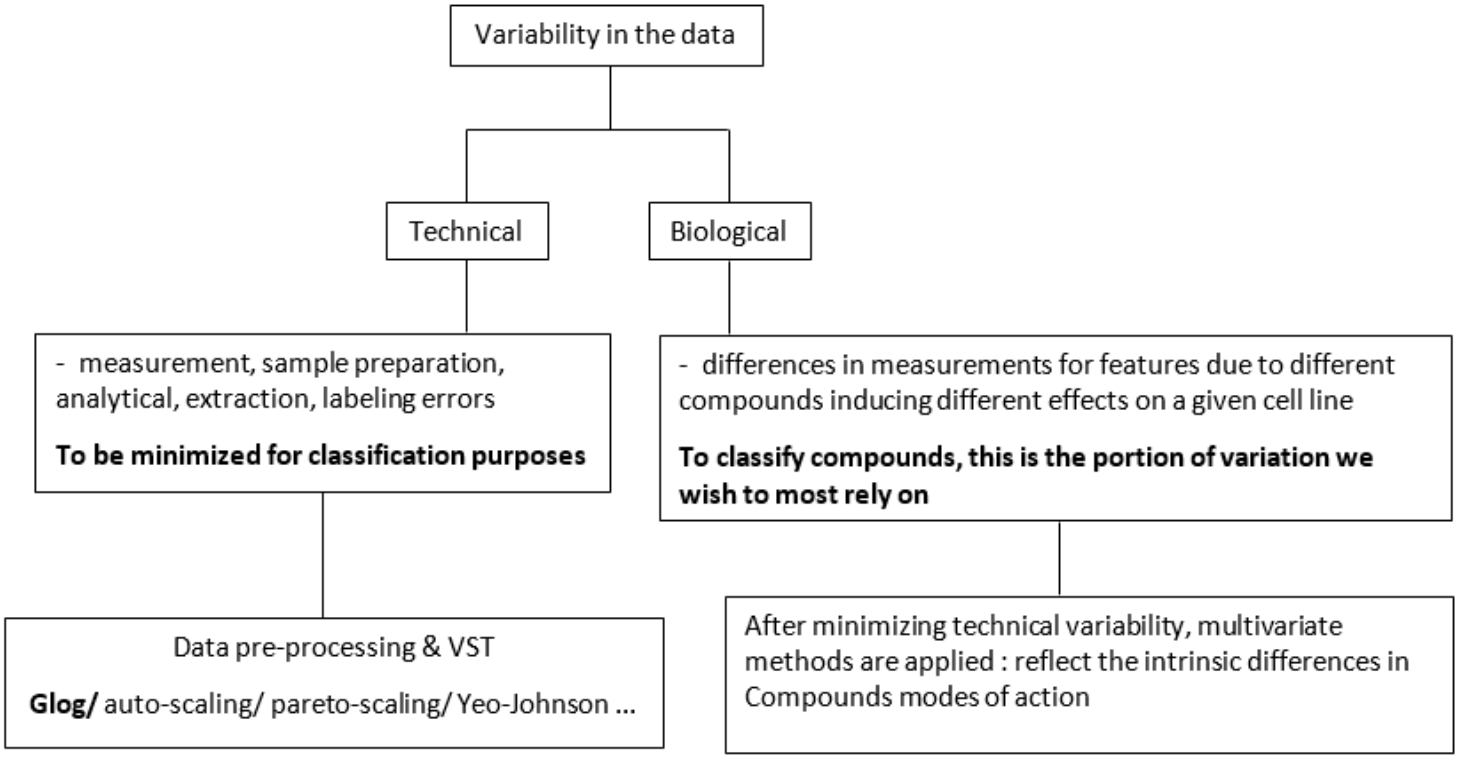
☐ Why

i. describe & predict a compound's mechanism of action
ii. preferentially identify highly specific compounds having desirable effect(s) on a given biological target
iii. early detection of undesired compound effects on cells + cellular activity: toxicity

.. introduction

It's all good, but .... Most of these features often

- are highly correlated: need to limit features used for analysis
- have non-normal distributions: mean-variance relationship present
  - multivariate classification methods hugely depend on variance



Aim of the analysis

To assess the:

I. effect of glog transformation on separation of treatment replicates from non-replicates

II. *effect of glog transformation on proportion of actively-called treatments*

III. *performance of a glog transformation on treatments separation when applied at cell- or well-level*

♣ Treatment: a compound at a given concentration (a compound can have 4 or 5 concentration levels - 1 μM (microMolar), 3 μM, 3.34 μM, 9 μM and 11.1 μM)

Data, the glog transformation and data pre-processing

Data

♢ cancer cell lines: Liver & Colon

♢ 3 batches @ 18 plates

♢ Number of cells

a. plate: btwn 134909 & 281177 (152679 & 330117) in $1^{st}(2^{nd})$

b. well: 73 & 1812 (95 & 2120) in $1^{st}(2^{nd})$

♢ 311 compounds including control

♢ total of 1253 treatments

♢ 462 features extracted from each cell

..Data, the glog transformation and data pre-processing

Glog transformation

- formula

$$z = \text{Log}(y - \alpha + \sqrt{(y - \alpha)^2 + \lambda})$$

- where
  - α: feature mean across controls
  - λ: transformation parameter

Data pre-processing

- Aggregation - calculating mean for each feature per well
- Normalization -

$$\frac{\text{feature}_{value} - \text{mean.feature}_{DMSO}}{\text{pooled.SD.feature}_{across.plates}}$$

- Feature selection
  - MRMR [1]: identify set of features with low pairwise correlation & high reproducibility among replicates.
  - AUC value for btwn 2-75 features
  - optimal feature: maximizes separation of treatment replicates within 1 std error of AUC
- Active calling: treatments with $\geq 50\%$ active replicates

Methodology

⋆ Hotelling's T^2 method

⇝ measures difference in 2 multivariate means

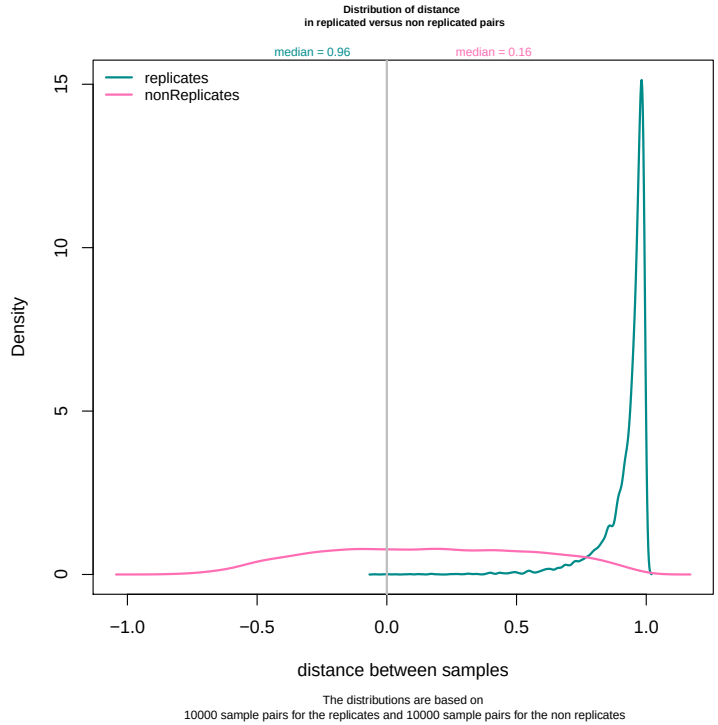
⇝ formula

$$T^2 = \frac{(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)\prime(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)}{\boldsymbol{S}_p(\frac{1}{n_1} + \frac{1}{n_2})}$$

⇝ normality assumptions for optimal results

⋆ AUC method

2-steps involved in AUC-calculation



⇝ separation btwn distributions quantified [ROC curve + AUC]

Results: EDA

No. of replicates (% of sample treatments)		
9 (%)	36 (%)	45 (%)
1192 (95.208)	28 (2.236)	32 (2.556)

★ DMSO control replicated across 1512 wells

★ Implications

- ⋆ For calculation of Hotelling's T^2, a limited number of selected features was used to maximize its power
- ⋆ 10 highest ranked features from MRMR used to calculate T^2

Results: Transformations effects on treatments separation

Prologue

◇ Only glog transformations [λ =0.1 &, 0.5 - 25 at 0.5 interval]

◇ Each transformed compared to its corresponding untransformed data defined by overlapping actively-called treatments

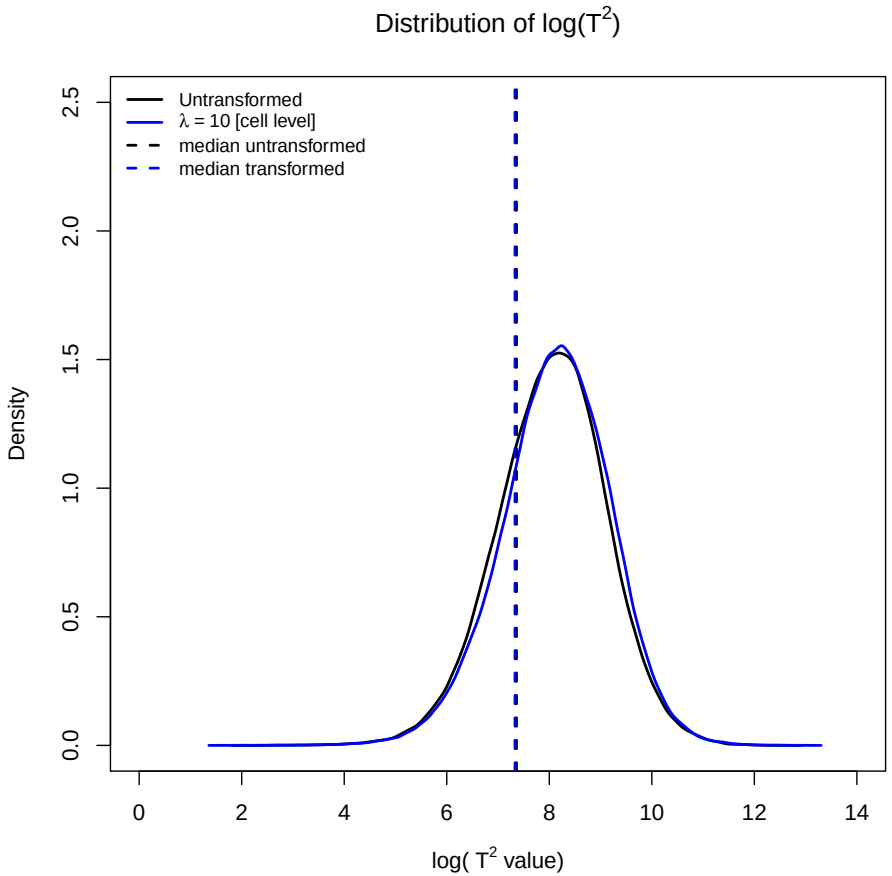
◇ Improved separation: +ve shifts distribution (and/or higher) of T^2 (AUC) for transformed compared to untransformed means

◇ 1^{st} cell line

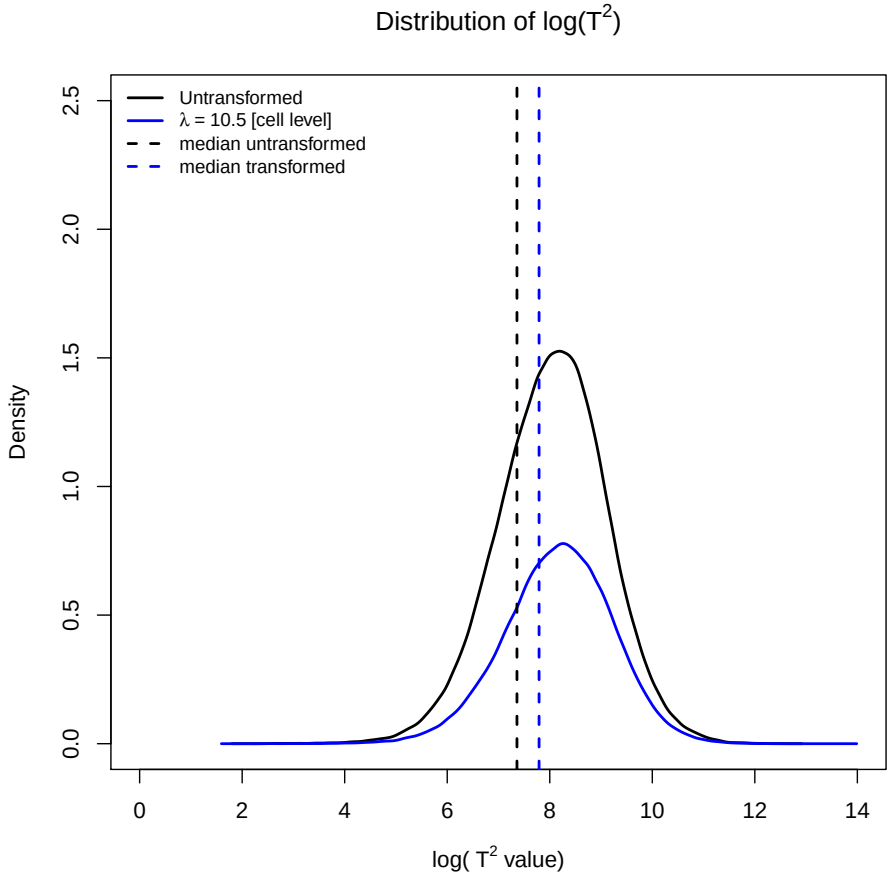
Transformations effects on treatments separation - T^2

⁂ very high [very different] + very low [highly similar] values present

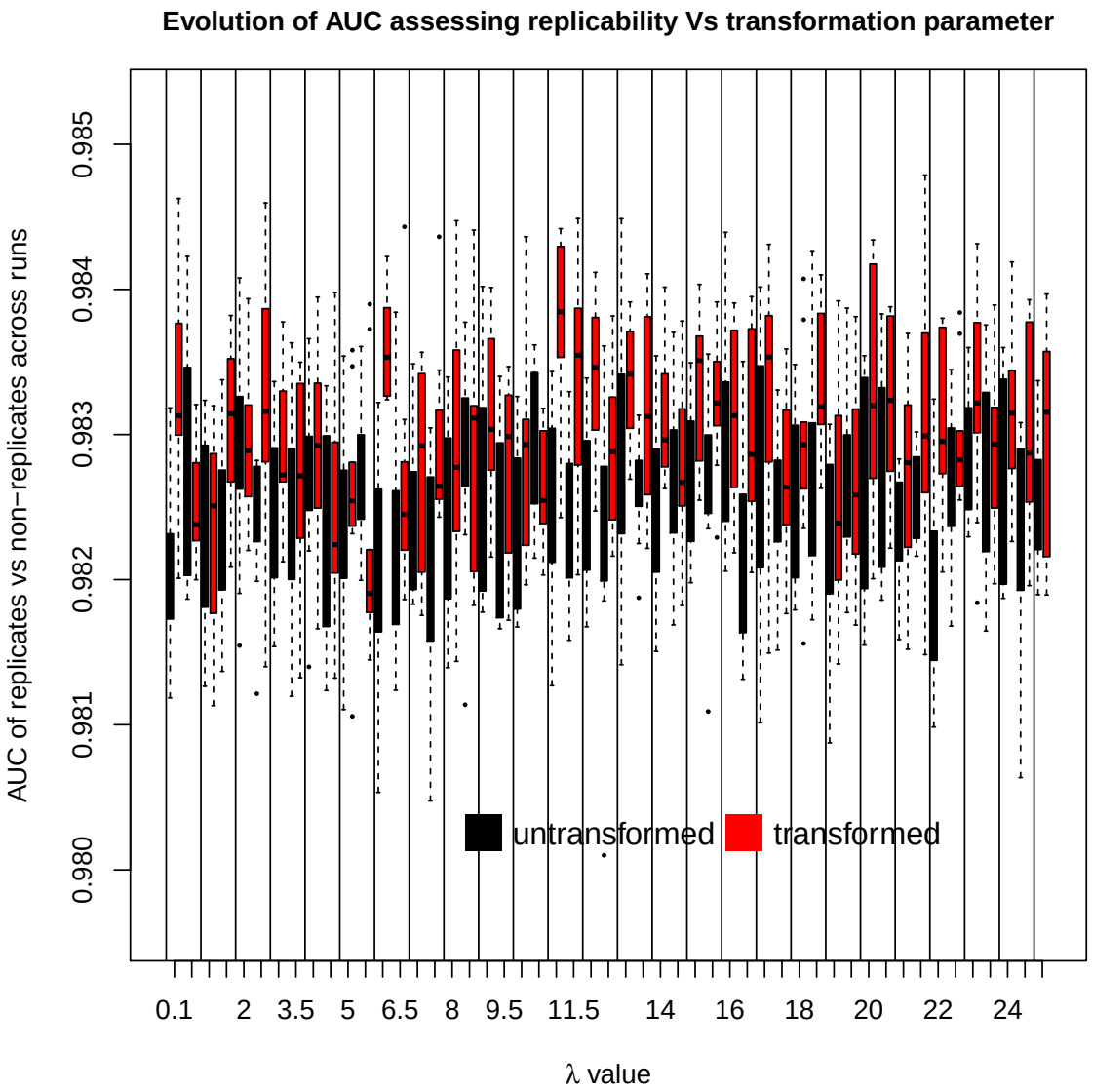
⁂ others led to no improvements (e.g $\lambda = 10$)



⁂ some led to slight but non-significant improvements



Transformations effects on treatments separation - AUC

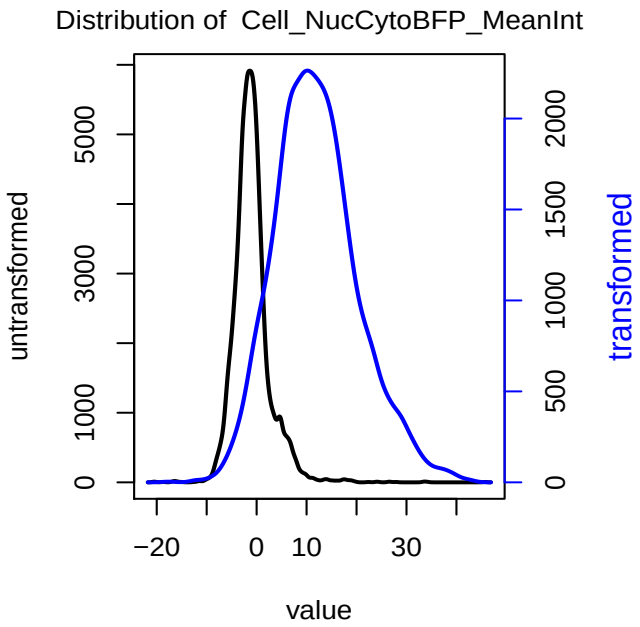


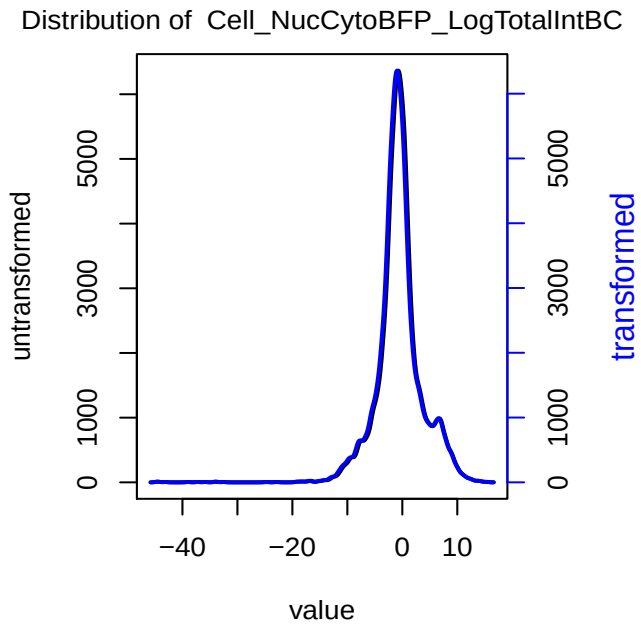
- high AUC b4-transformation
- some (e.g $\lambda = 5.5$) separated slightly poorer
- others (e.g $\lambda = 0.1$) led to marginal increases
- differences were minimal & non-significant

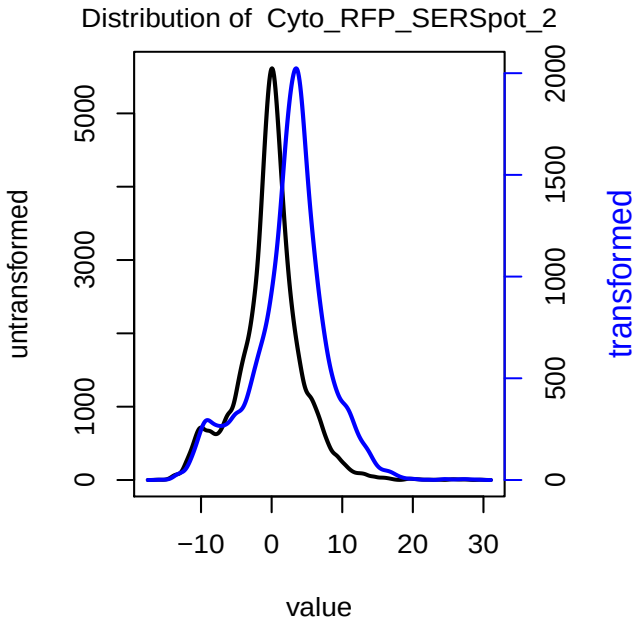
Transformations effects on treatments separation- Epilogue

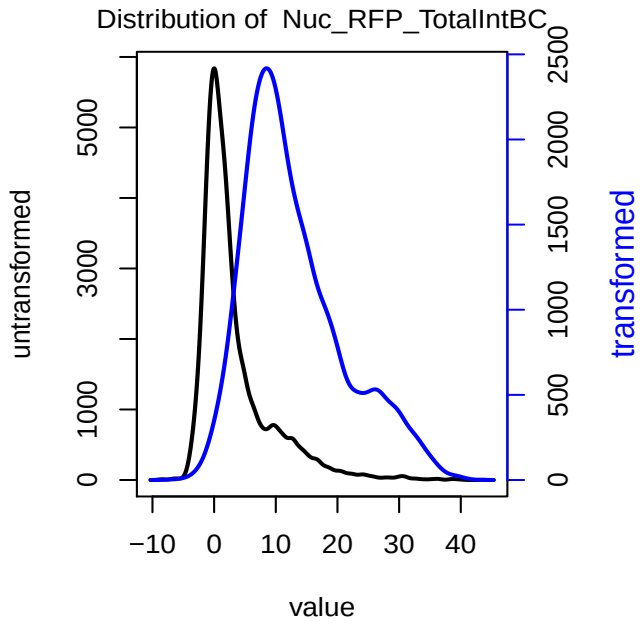
- ◇ In both methods, minimal & non-significant differences were observed
- ◇ Why?

1. Transformation effect on features distributions ($\lambda = 10.5$)

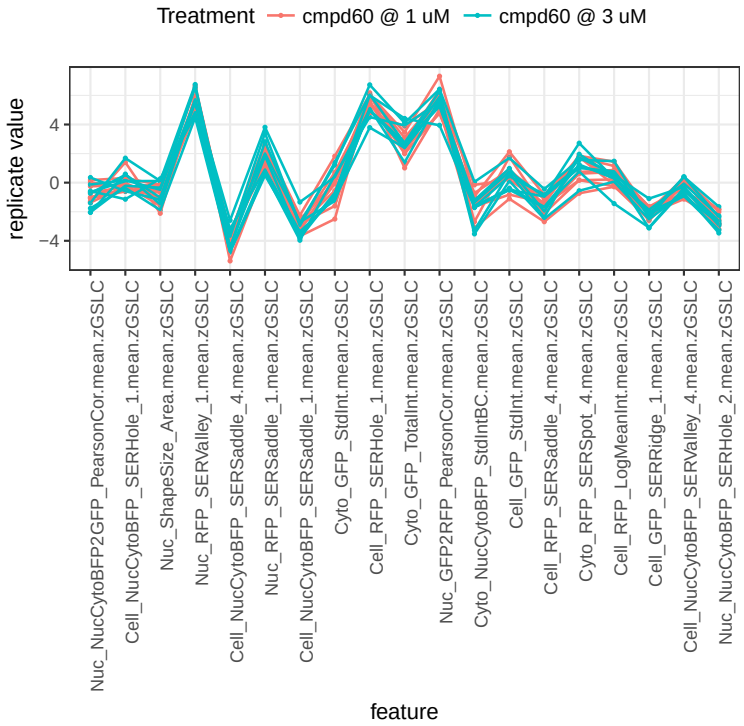


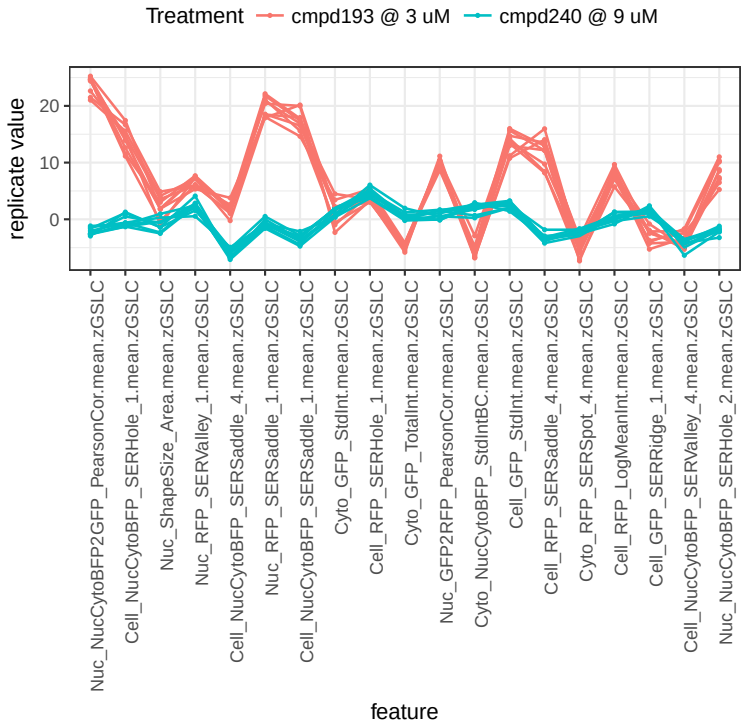






2. Differentiating ability of features selected (before transformation)





Conclusion

$\sim$ Transformations did not improve treatments separation beyond what was seen pre-transformation

Acknowledgement

Supervisors

- Prof. Dr. Ziv SHKEDY
- Prof. Dr. Nolen Joy PERUALILA
- Dr. Marjolein CRABBE
- Dr. Steffen JAENSCH

References

H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.