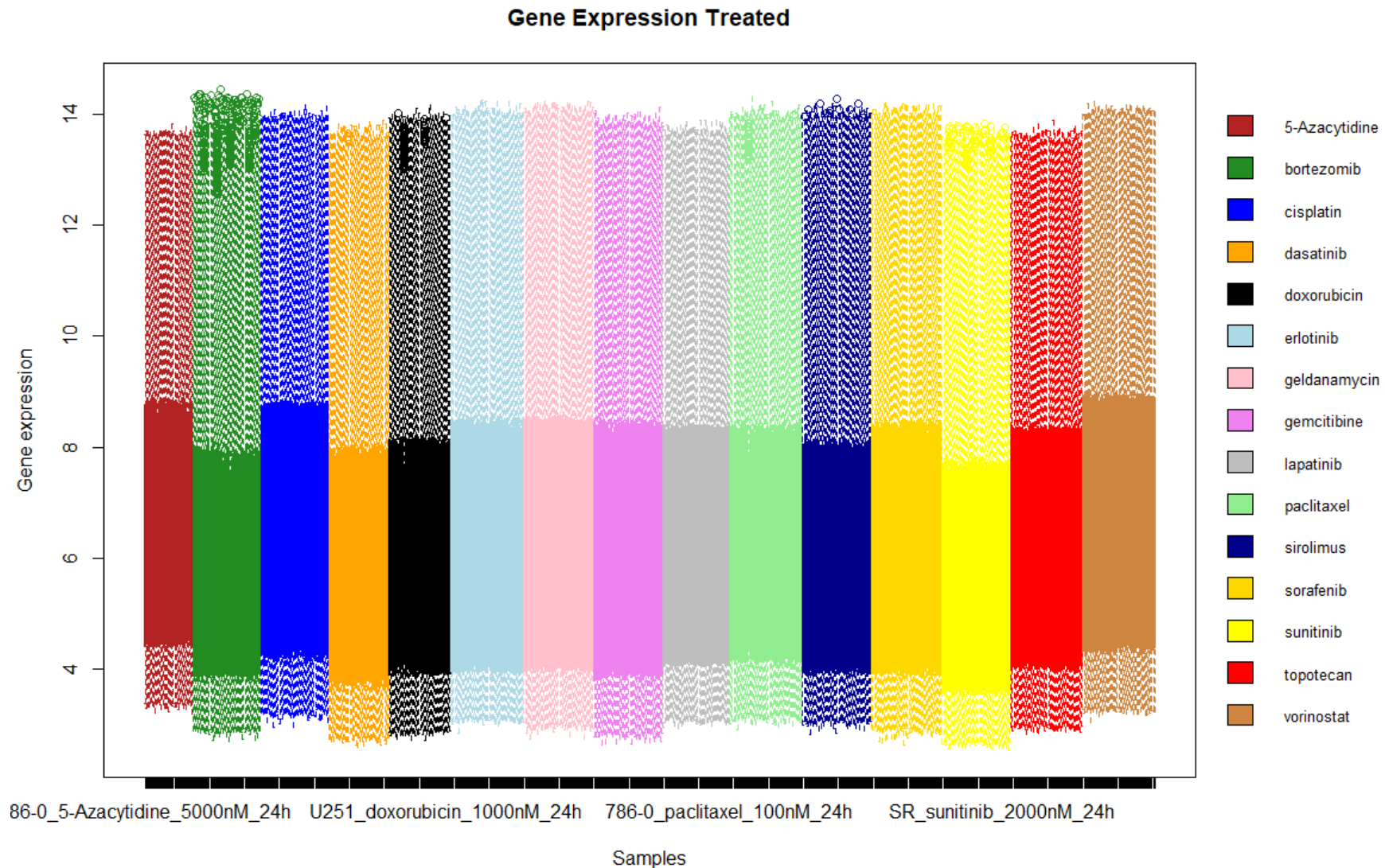


# Plots

Teresa von Linde, Dorothe Mersch, Amelie Vogt, Johanna Bauer

# Broad Analysis

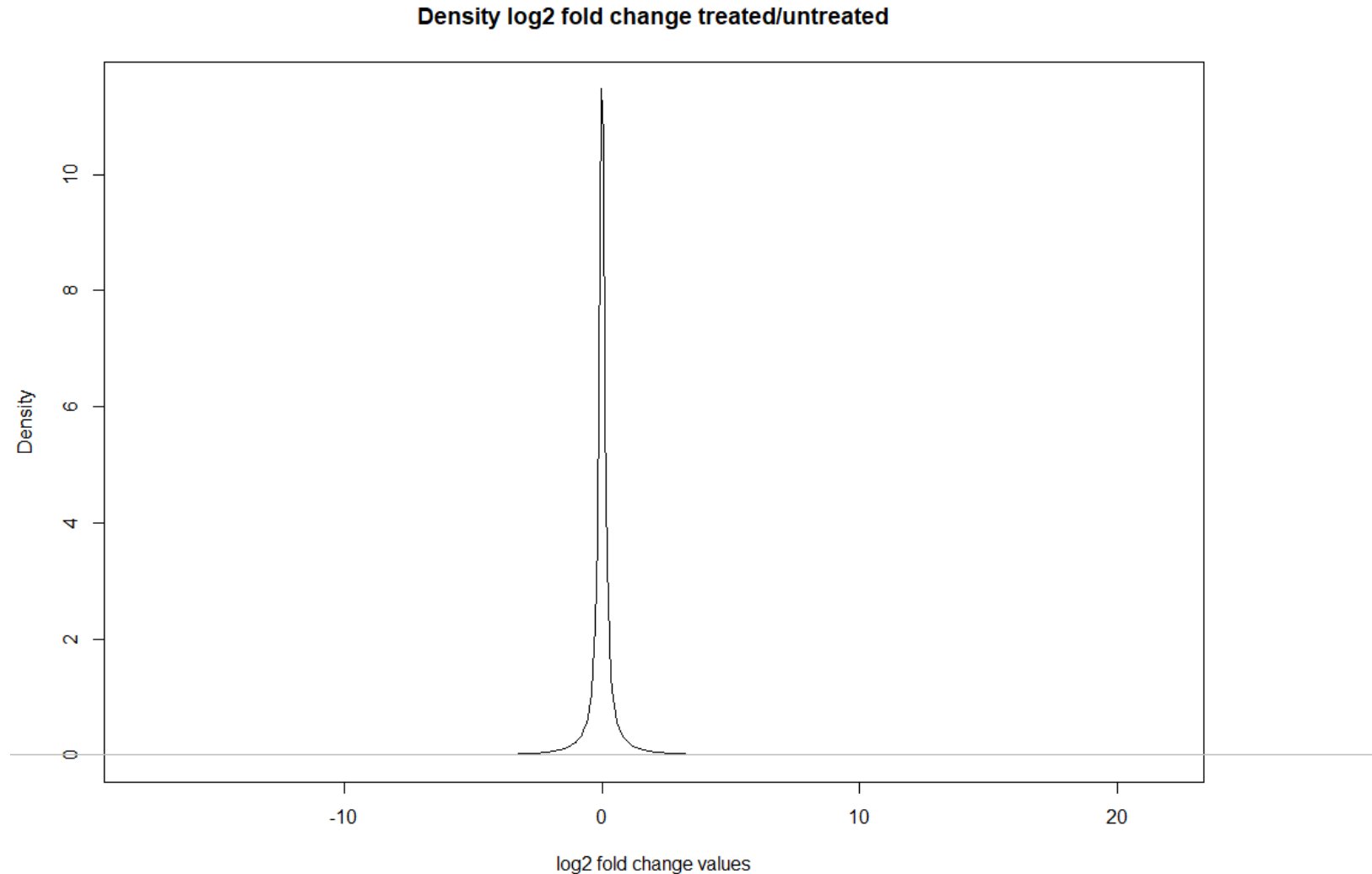
In order to get an overview a boxplot of the treated data was computed. There was a visible difference between the 15 batches.



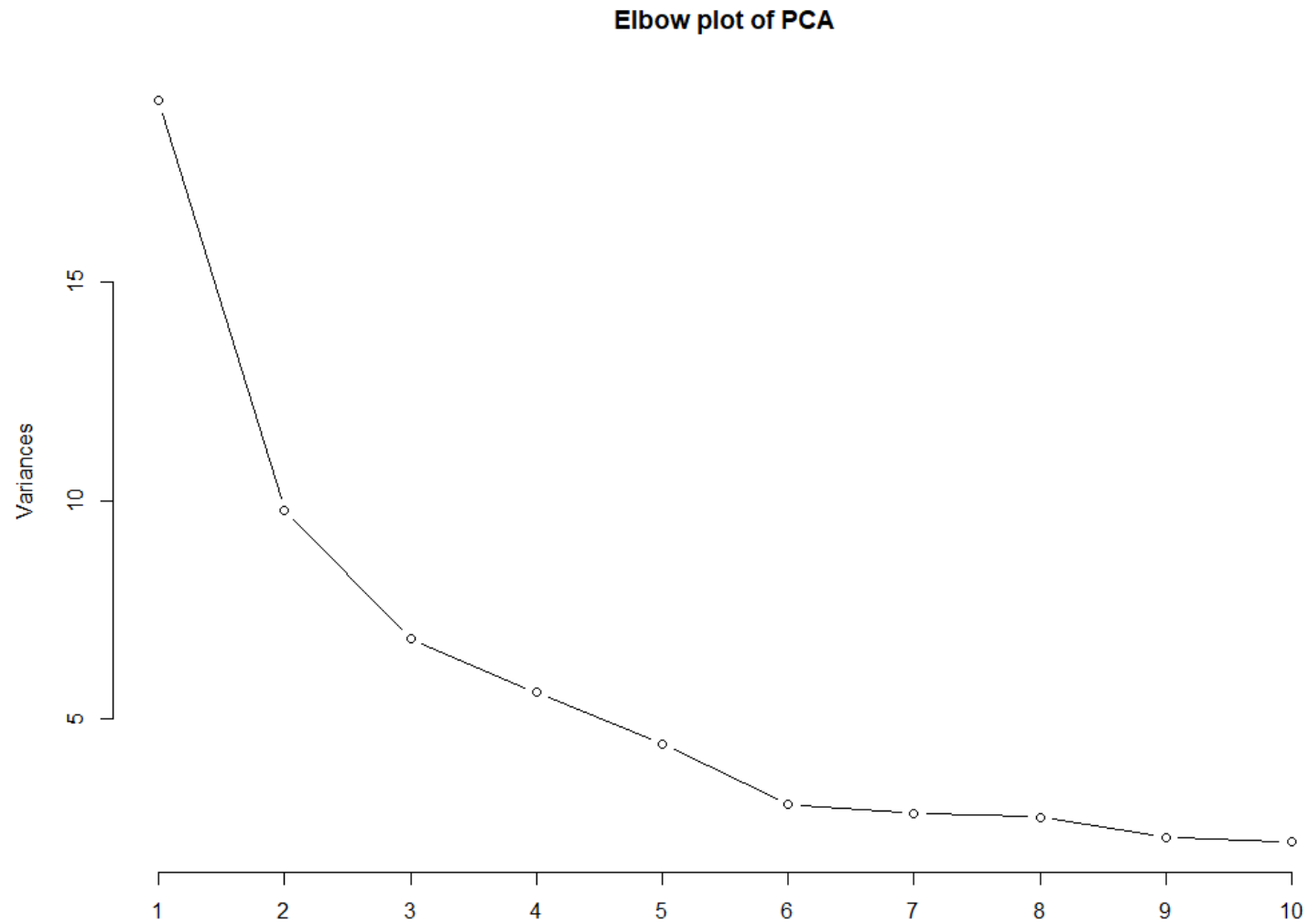
Through scaling the data the different batches disappeared.



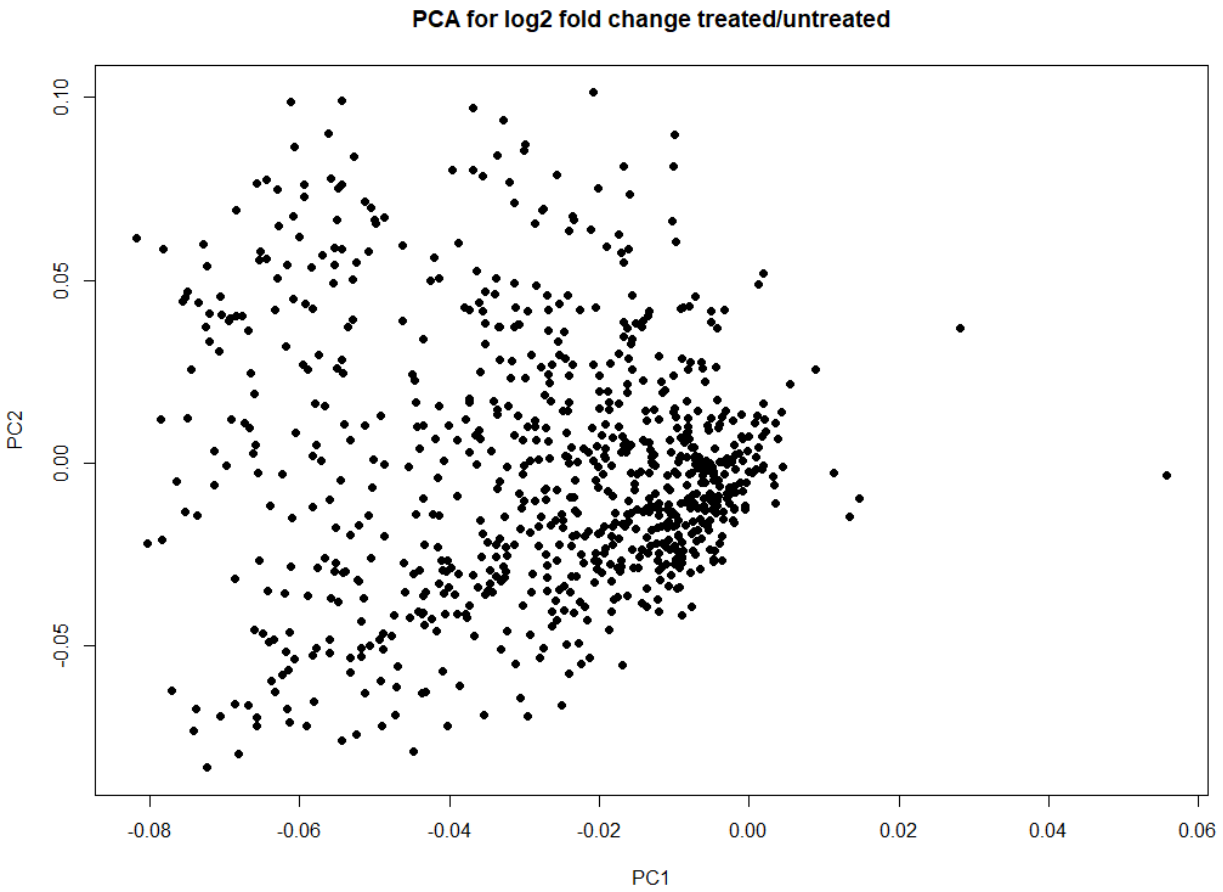
In order to analyze the gene expression change from untreated to treated, log2 fold change values were computed. The distribution was checked through a density plot.



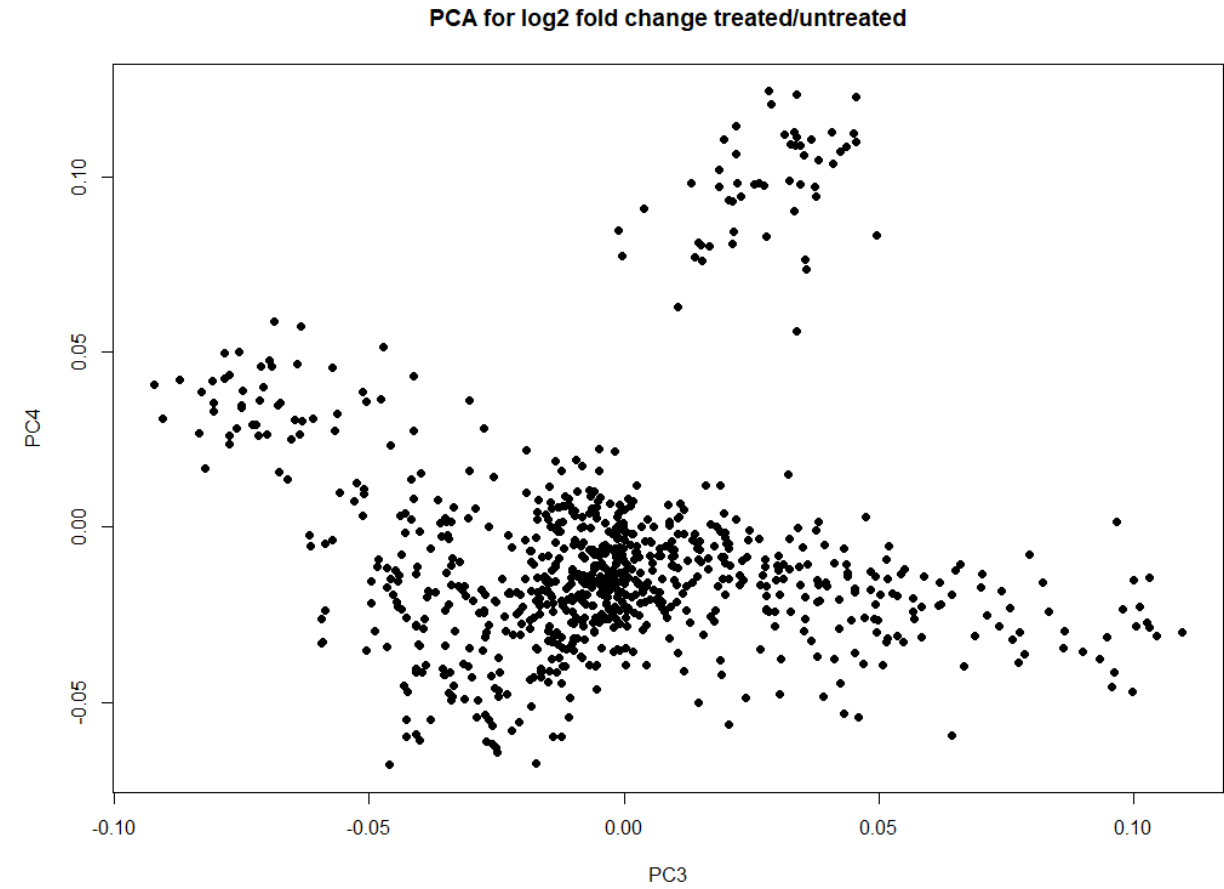
For a dimensional reduction, a principal component analysis was performed.



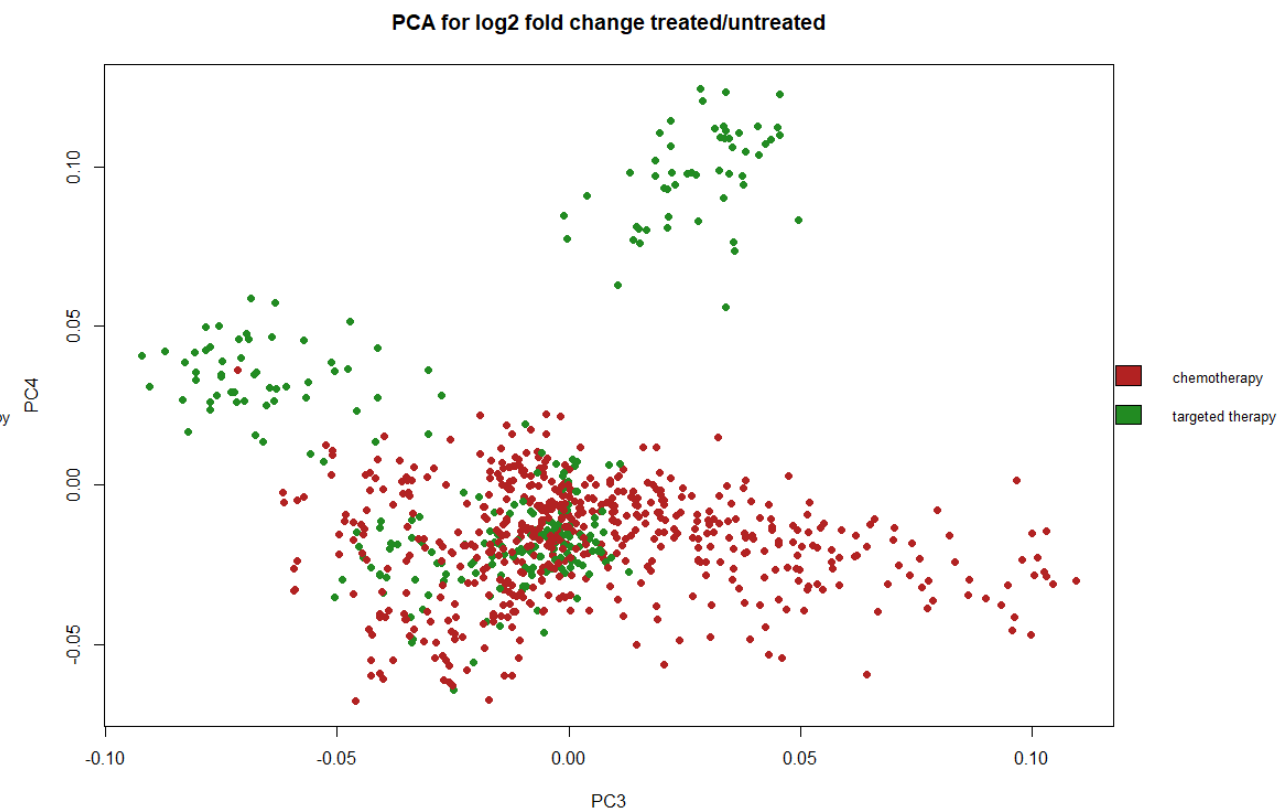
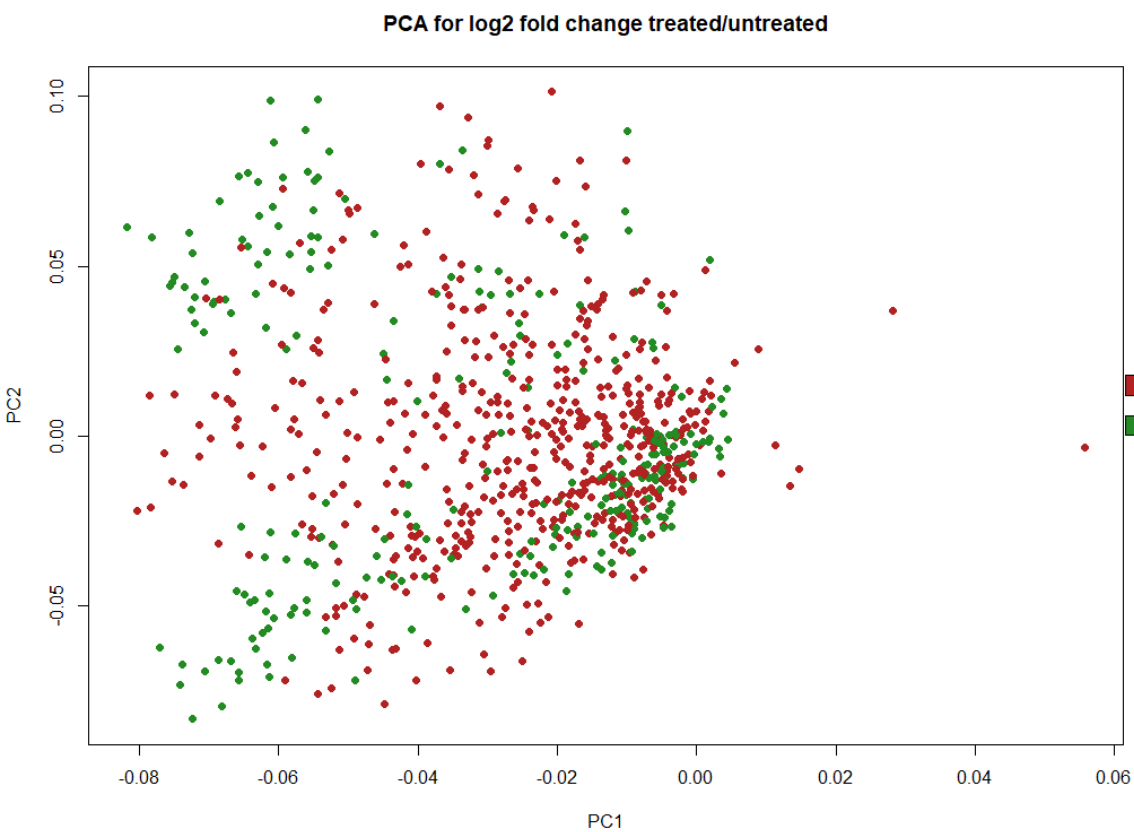
Plotting: PC1 against PC2 and



PC3 against PC4

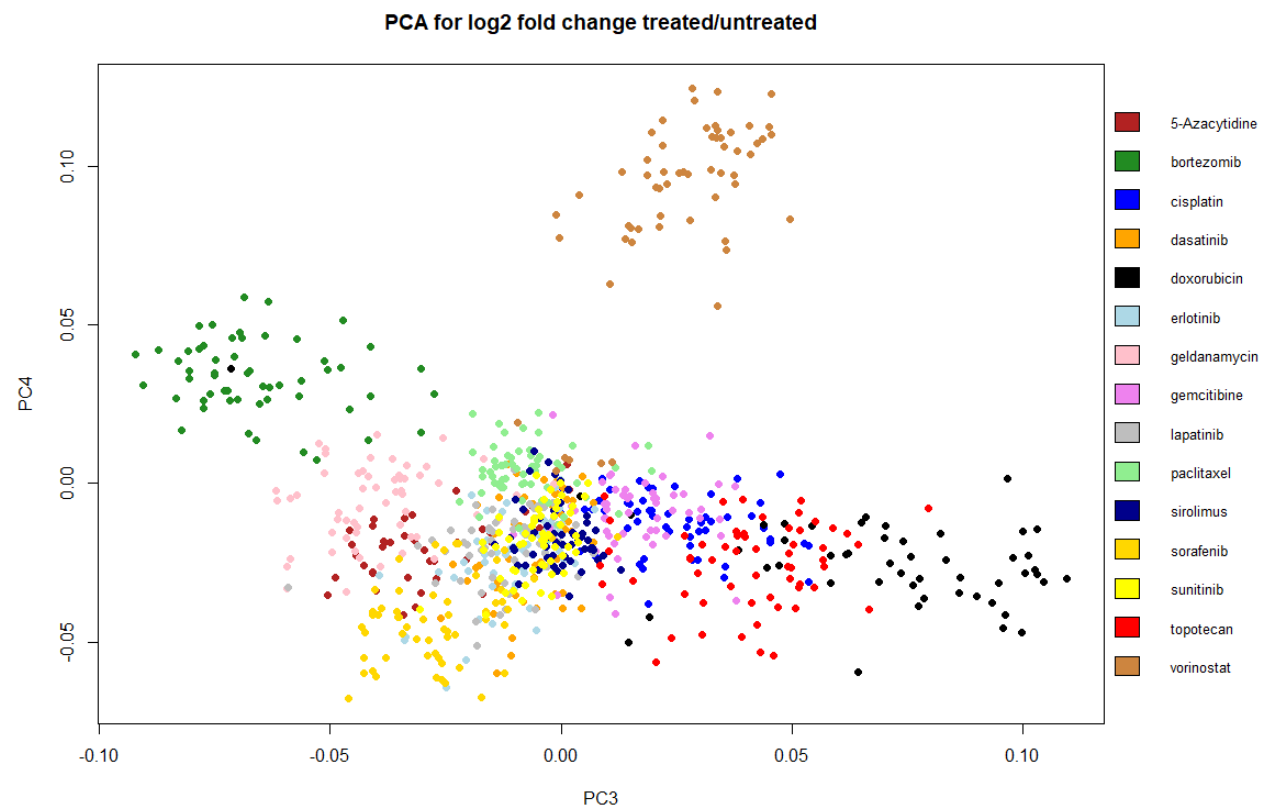
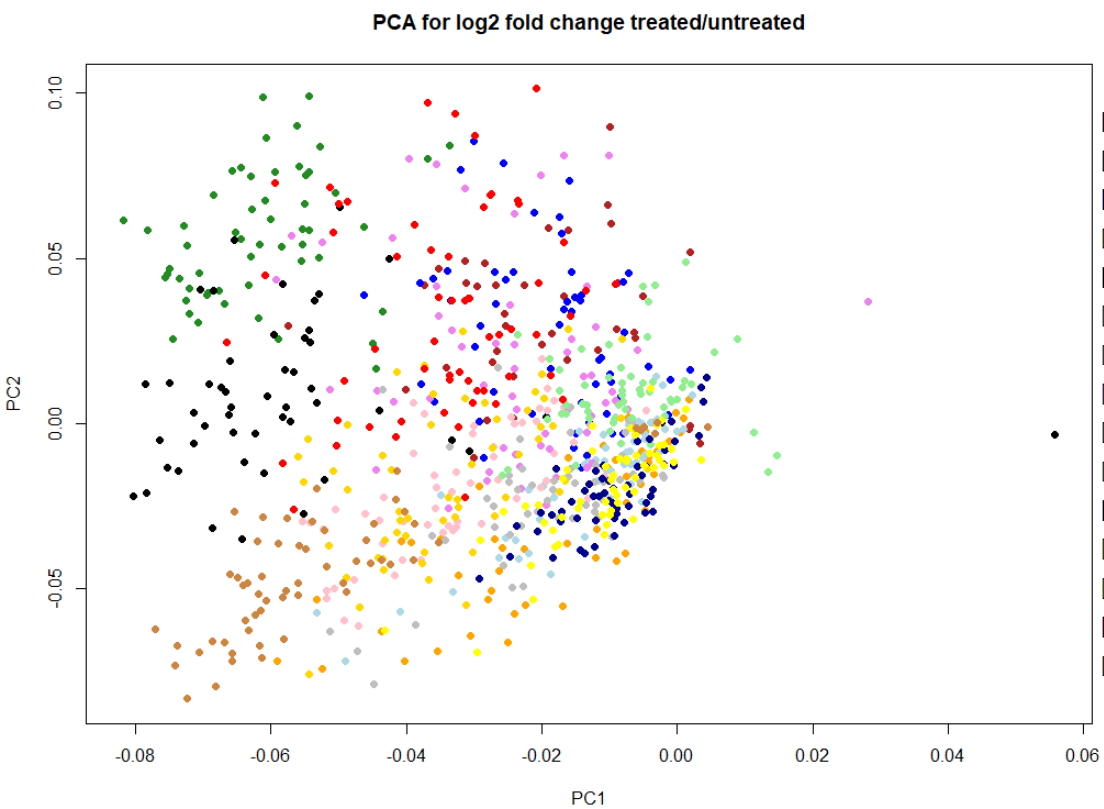


Color chemotherapy agents (firebrick) against targeted therapy agents (forestgreen) in those two plots.

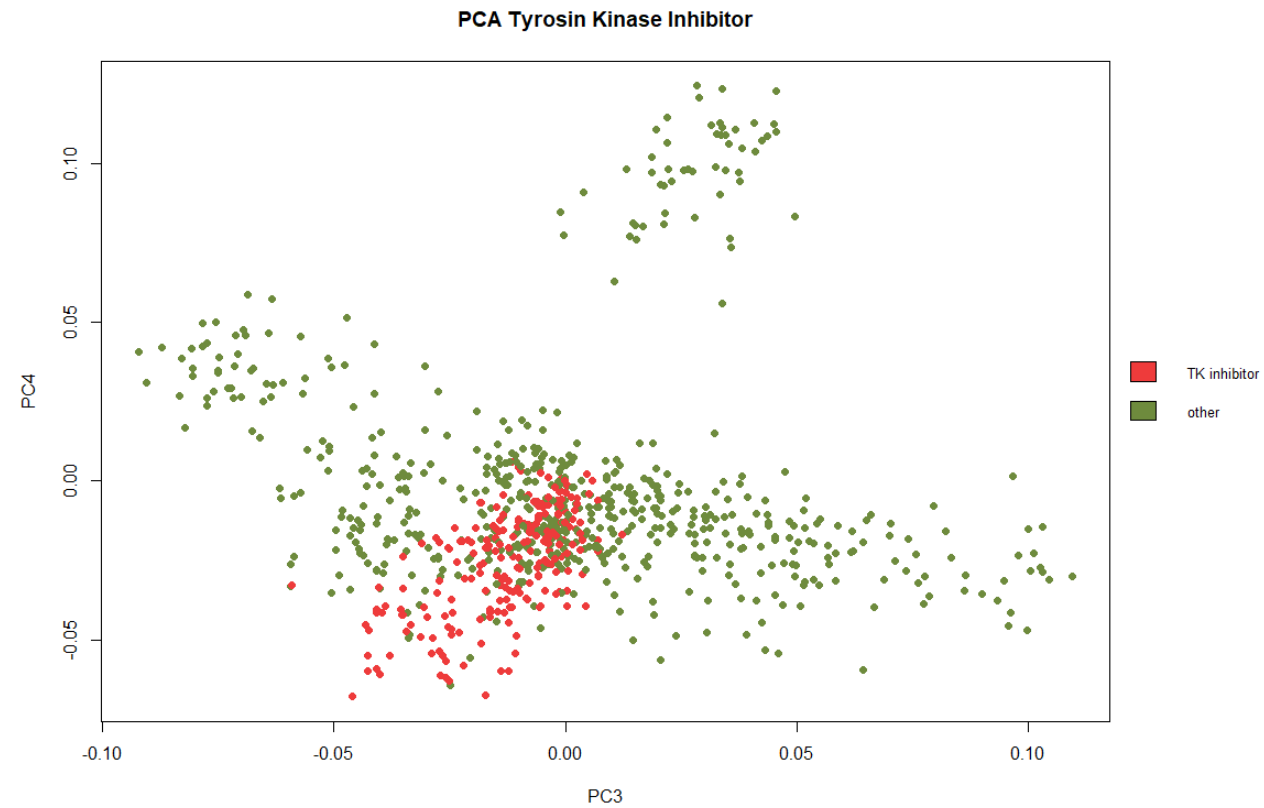
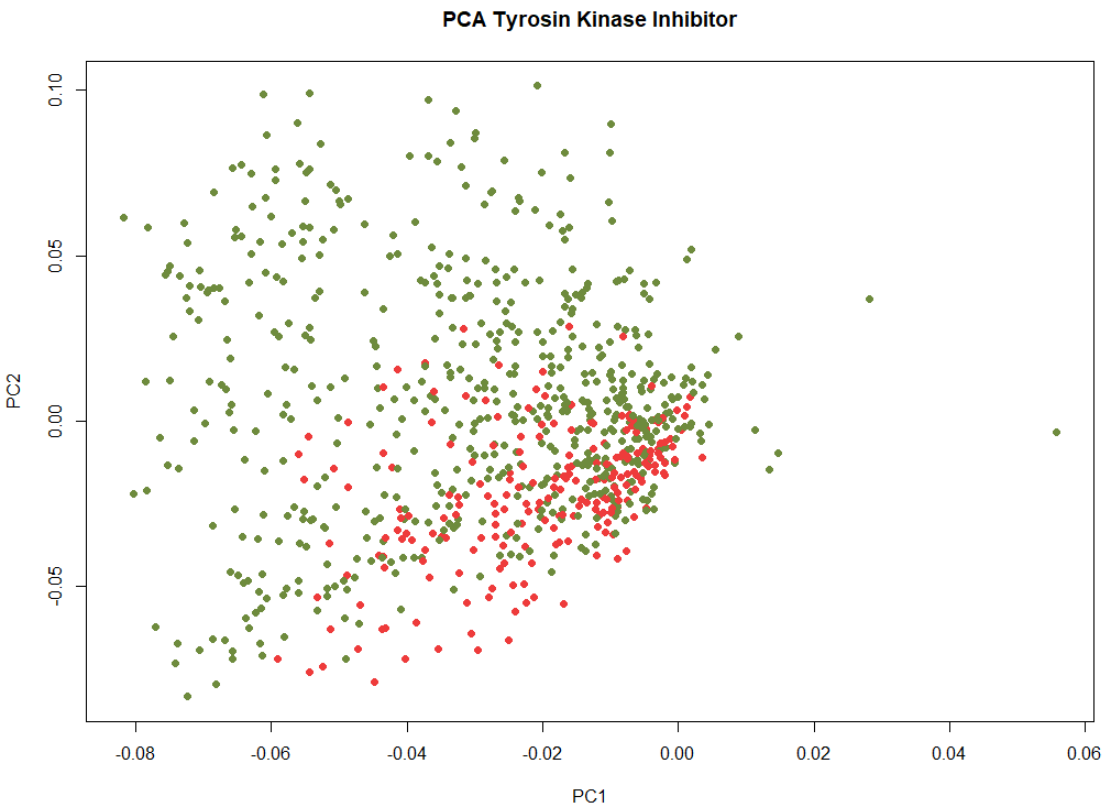




Color every drug in a different color.



Color tyrosine kinase inhibitors red (others: green) in those two plots.

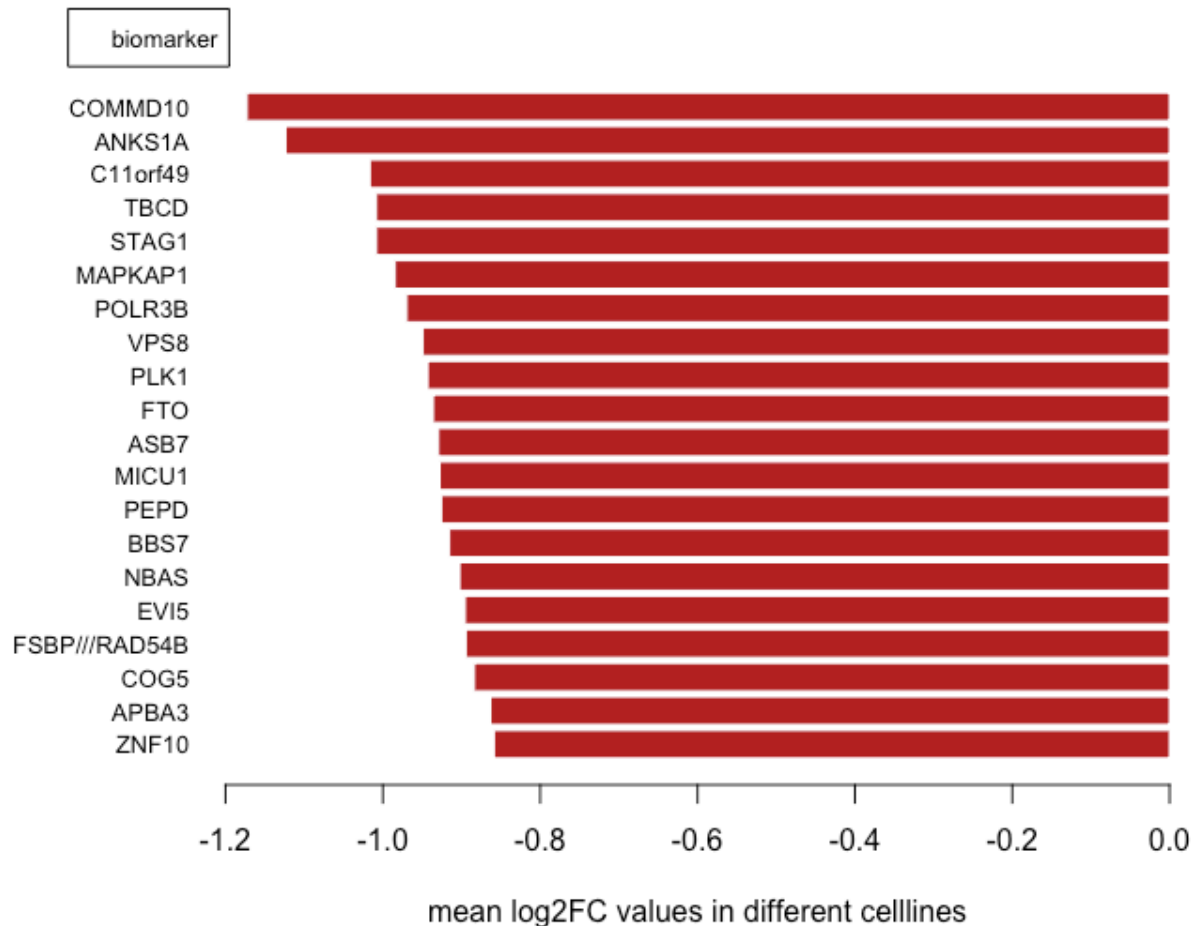


# Specific Analysis

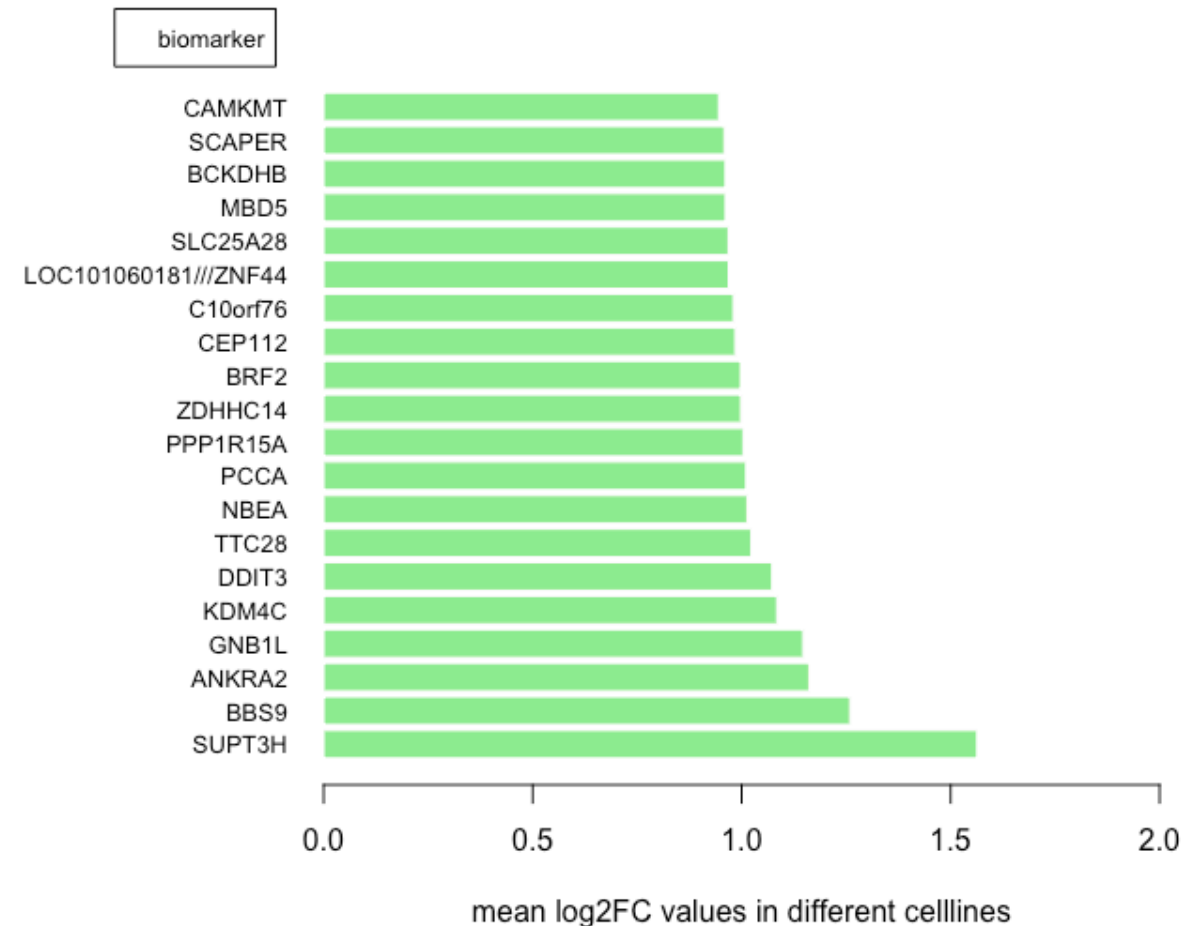
## Part I:

Identification of genes as biomarkers for cisplatin through extreme log2 FC-values.  
Visualizing the most extreme FC-values for cisplatin.

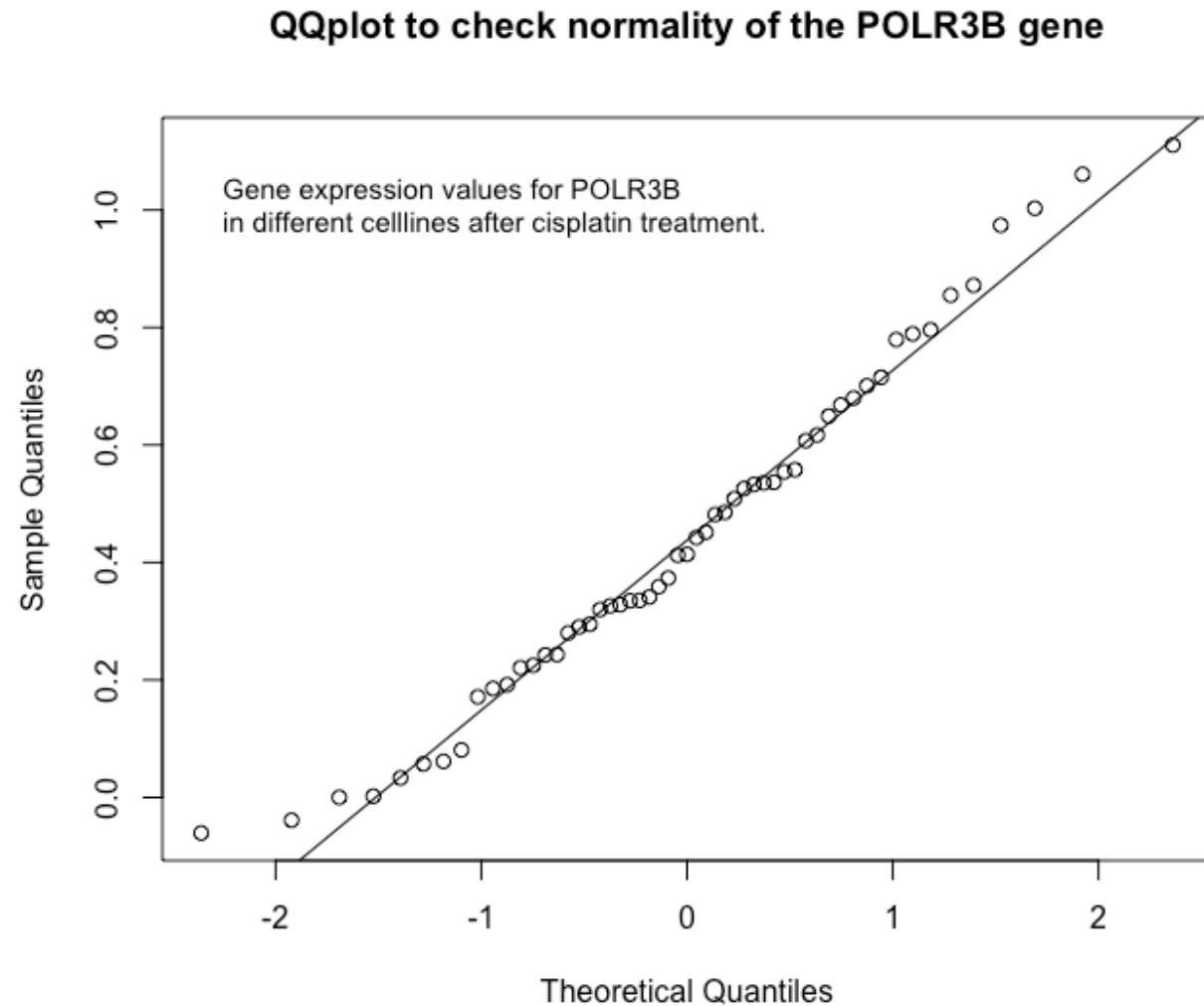
lowest log2 FC-values for cisplatin



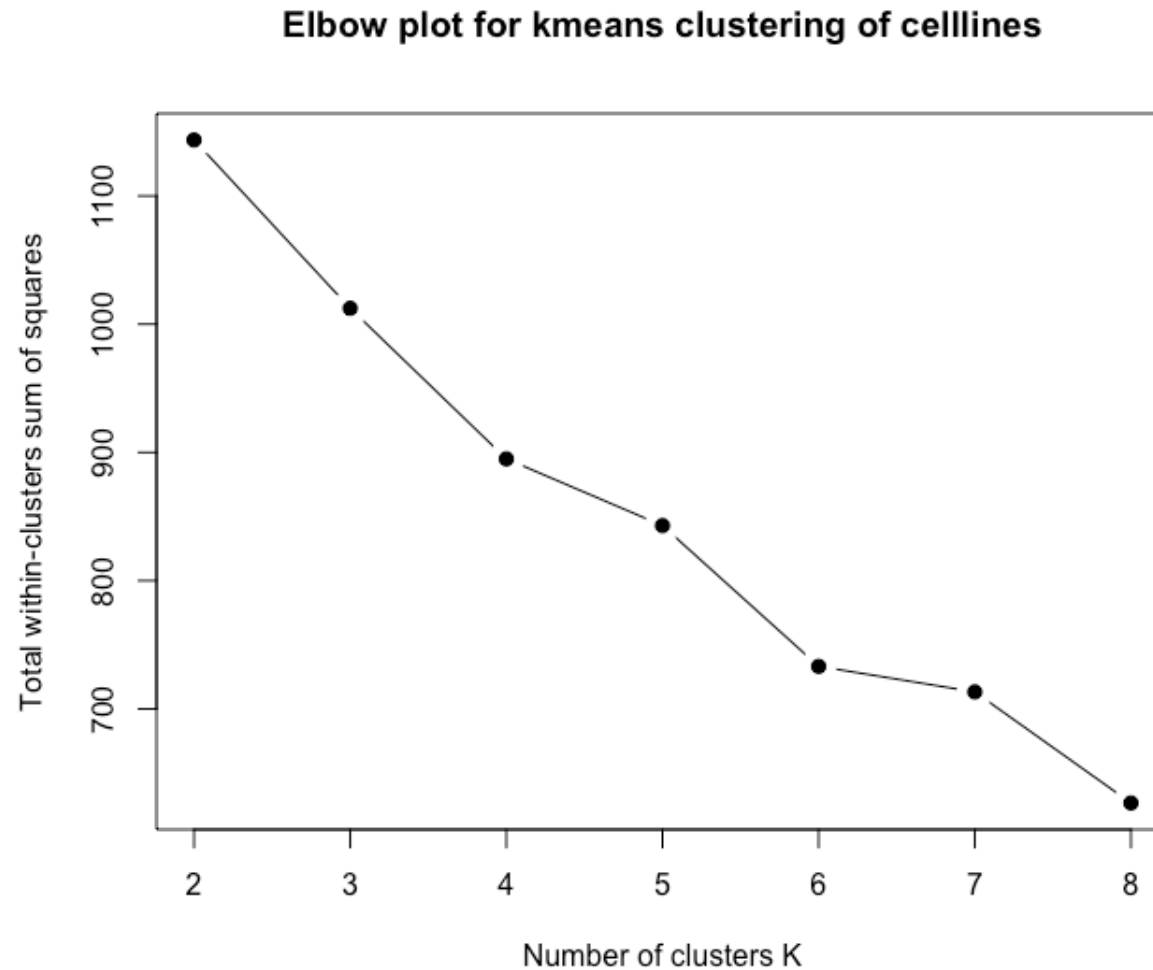
highest log2 FC-values for cisplatin



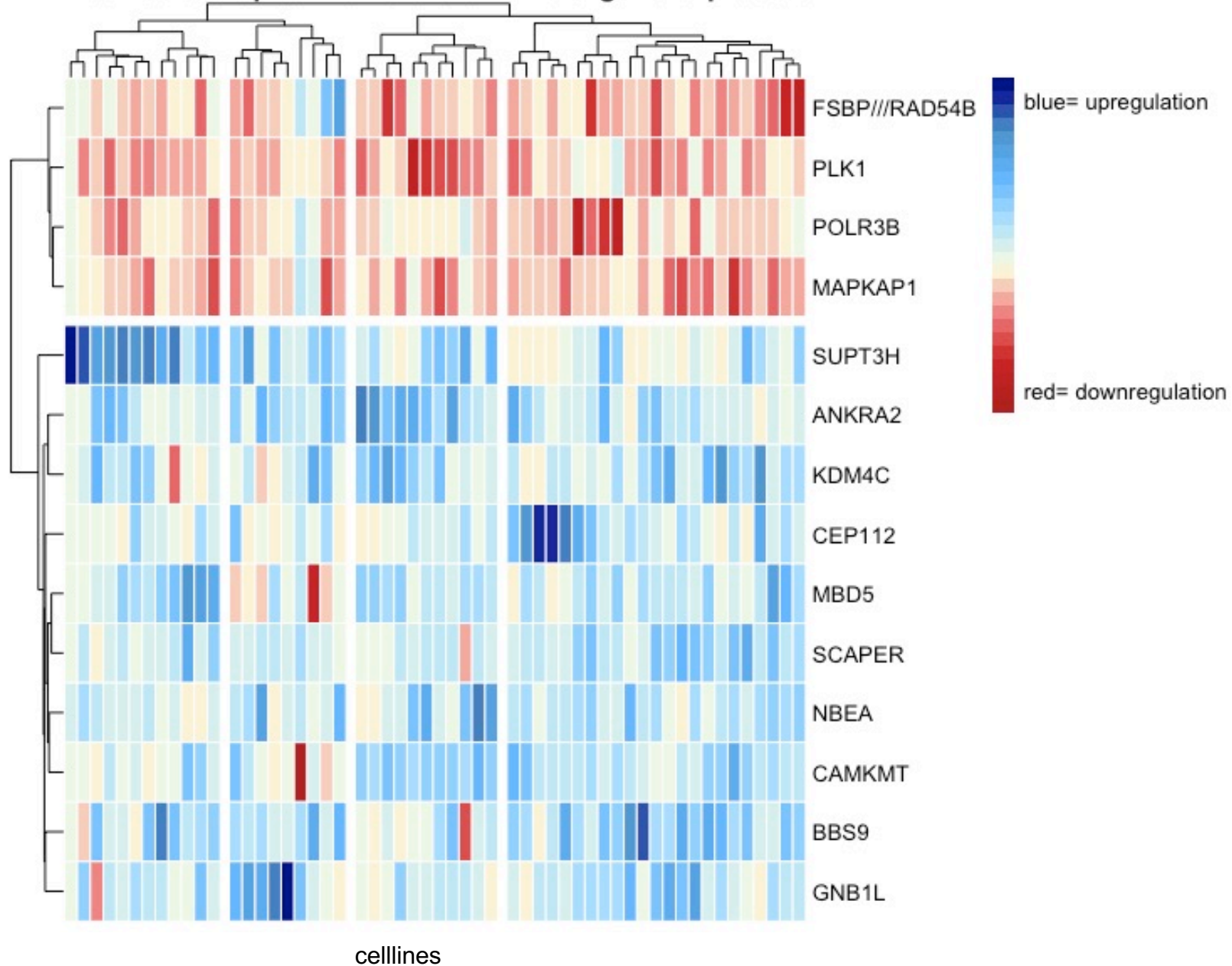
Part I (b): T-Test to check significance of the previously found biomarker. But before:  
Checking normality of the example gene“POLR3B” through a QQplot.



Part II: Influence of cisplatin on the biomarkers gene expression (up or down regulation). Checking for the optimal number of clusters for the following heatmap.

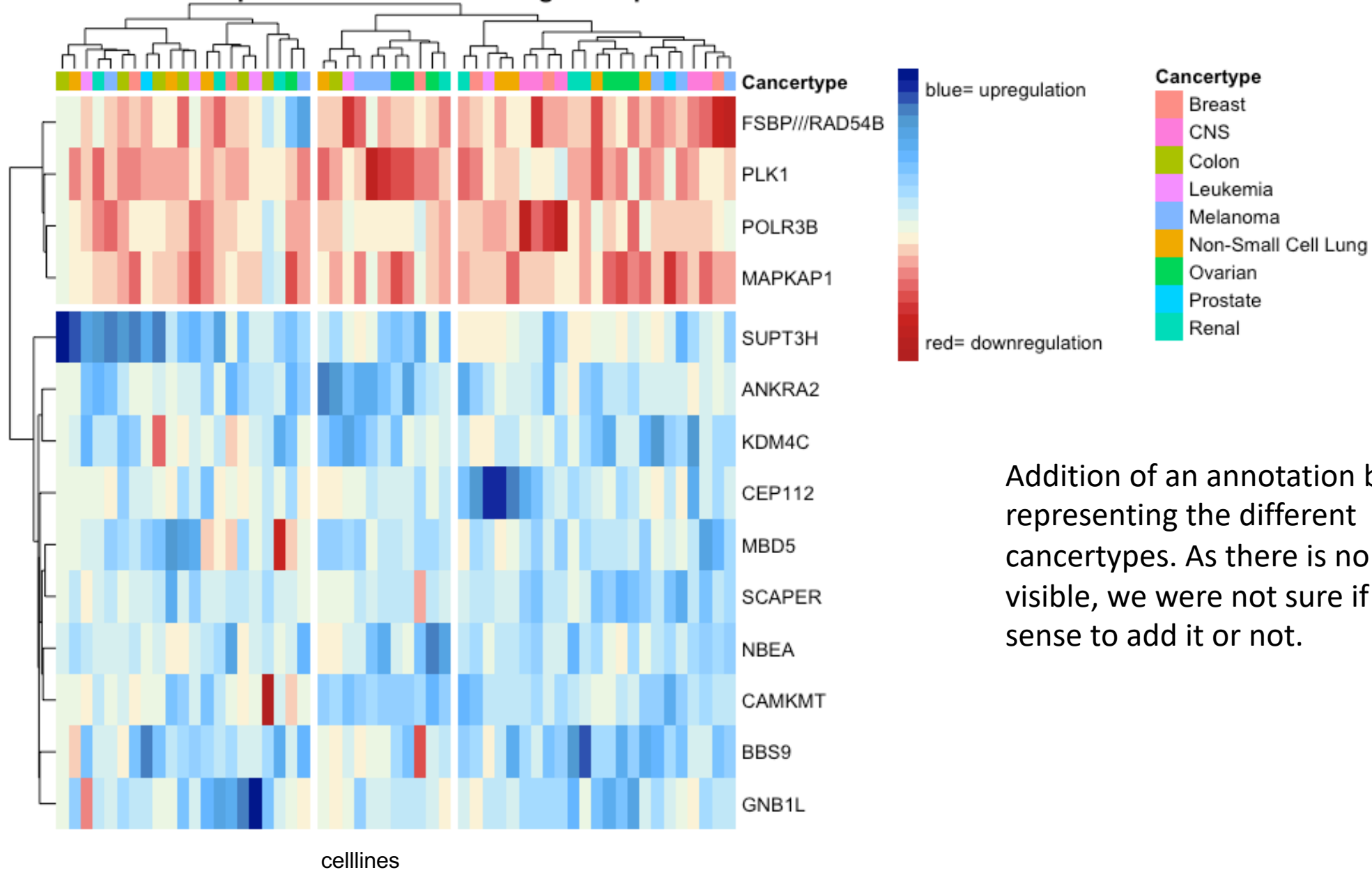


## Influence of cisplatin on the biomarkers gene expression



Checking the influence of cisplatin on the biomarkers gene expression using the amount of clusters determined with the elbow plot.

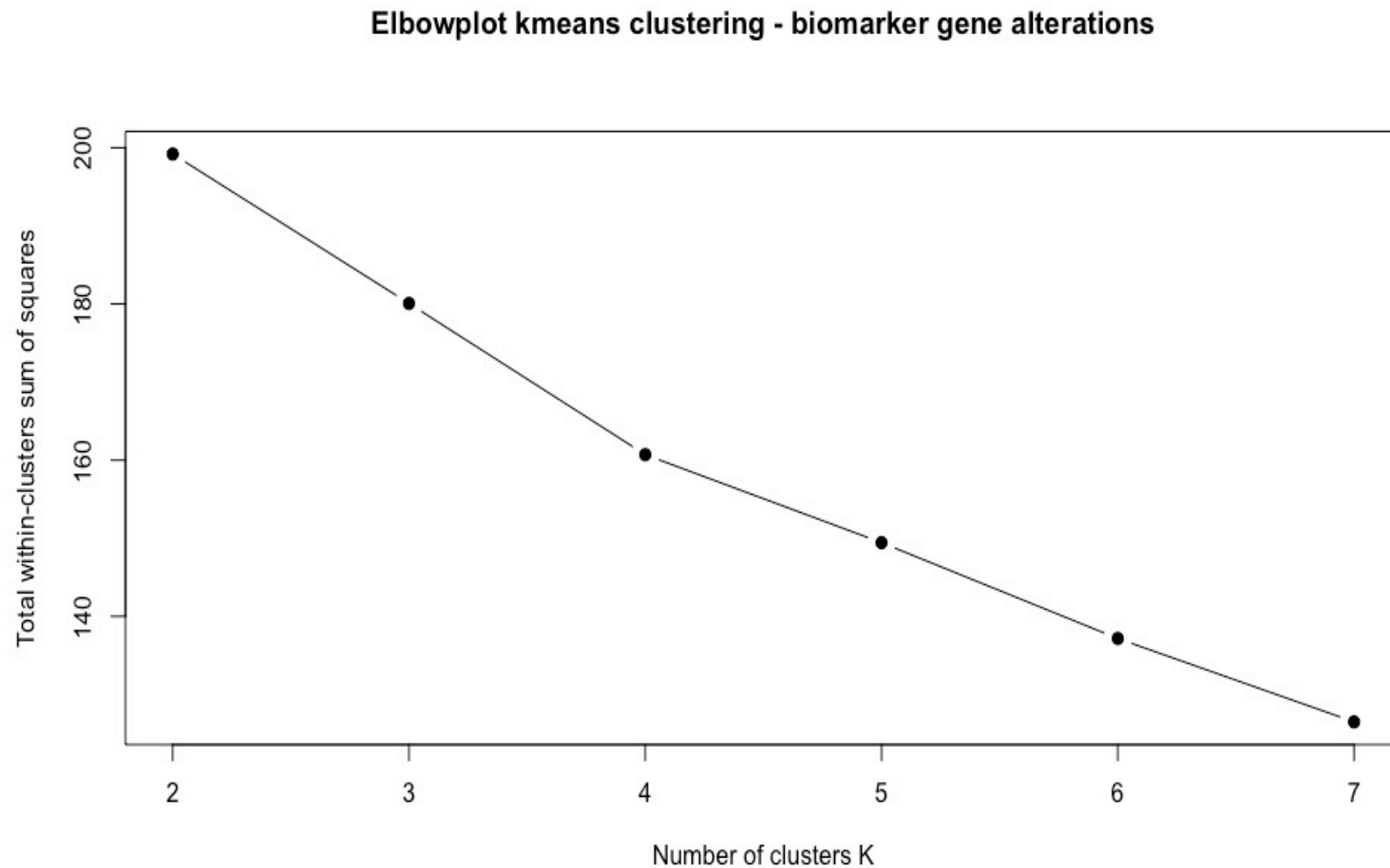
# Influence of cisplatin on the biomarkers gene expression



Addition of an annotation bar representing the different cancetypes. As there is no pattern visible, we were not sure if it makes sense to add it or not.

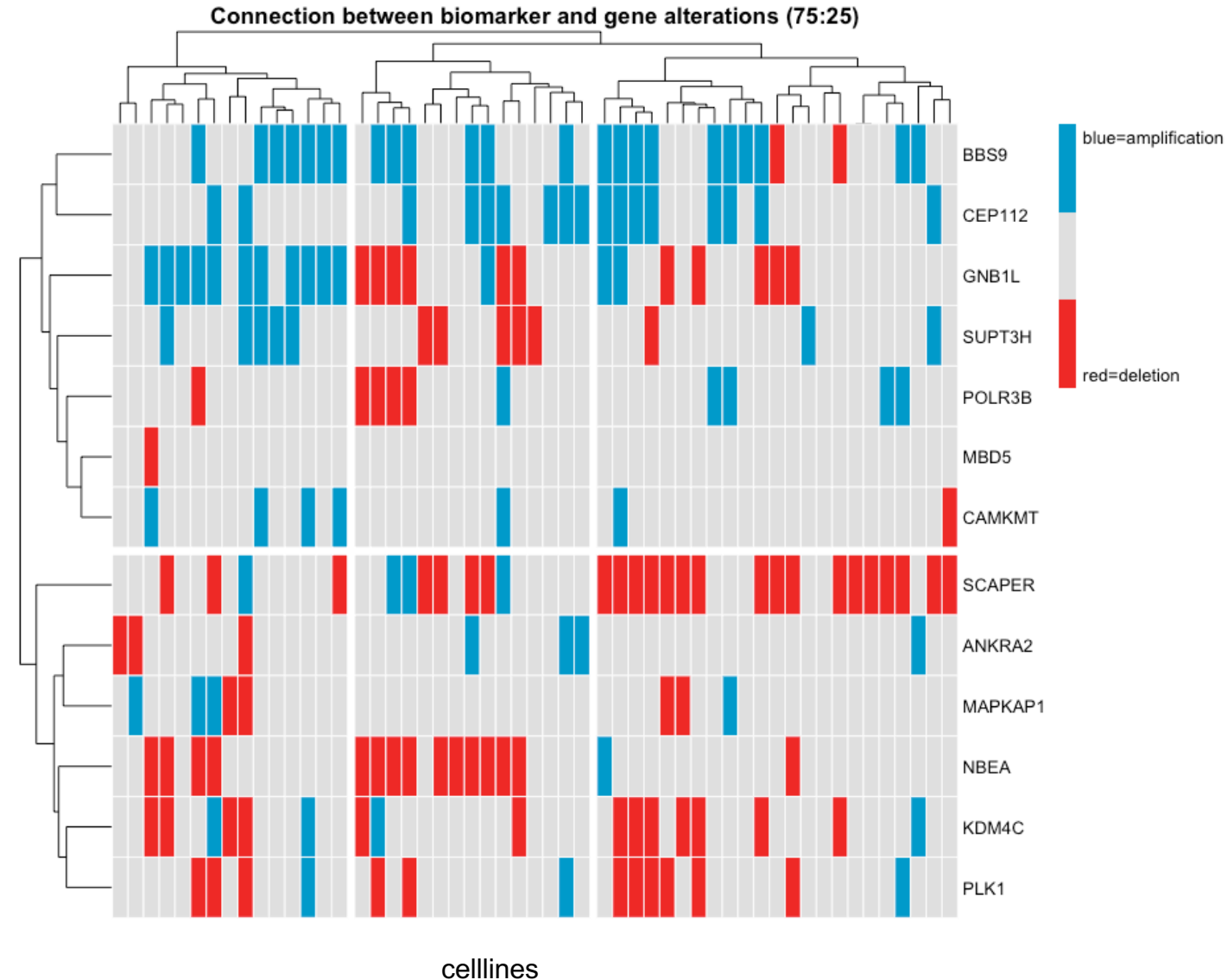
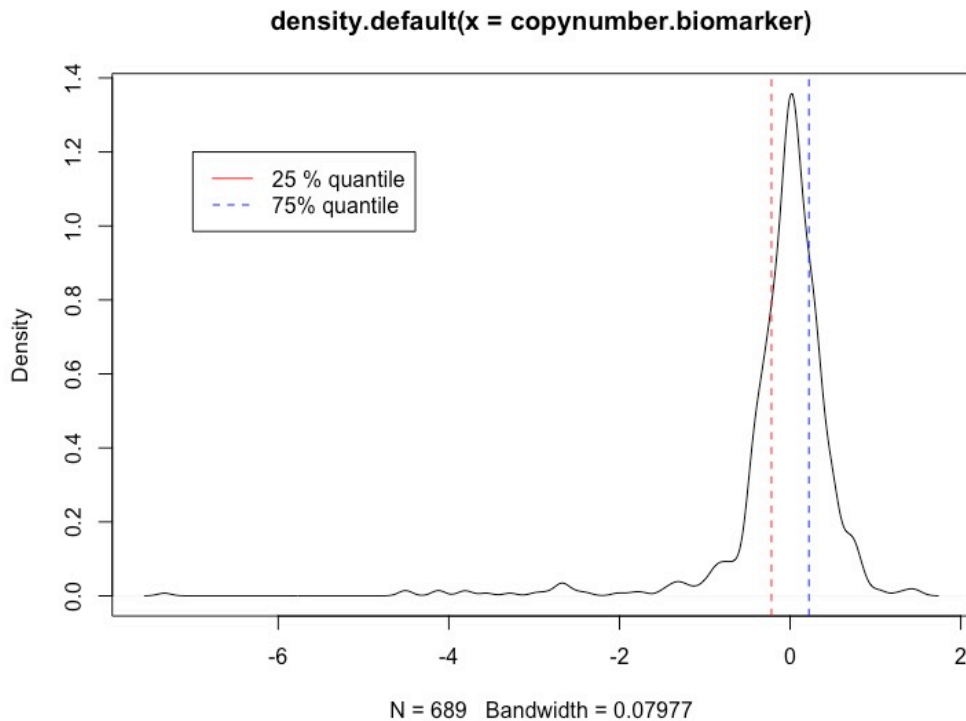


Part III: Further analysis of the biomarker for cisplatin.  
Checking for the optimal number of clusters for the following heatmap.



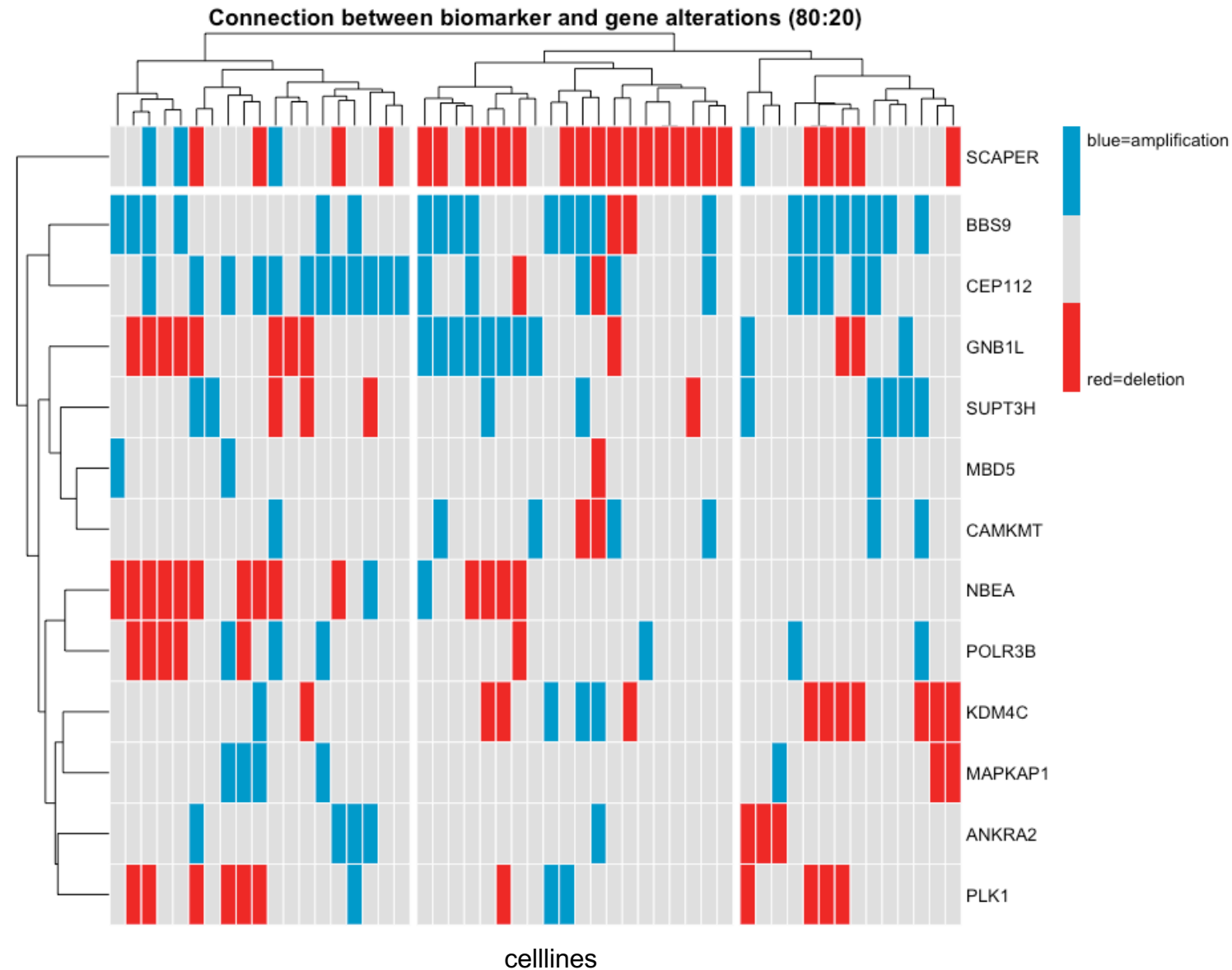
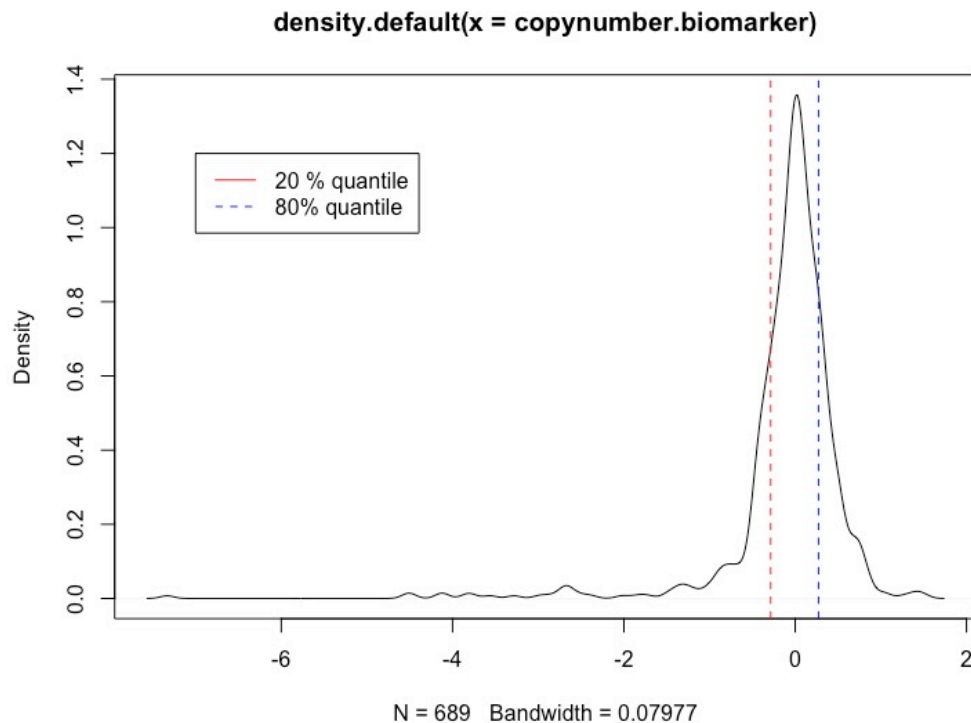
# Part III: Further analysis of biomarker – amplification, deletion or neutral gene copy number?

Genes were identified as amplified, if their gene copynumber is higher than the **75%-quantile** and as deleted if the value is lower than the **25%-quantile**.

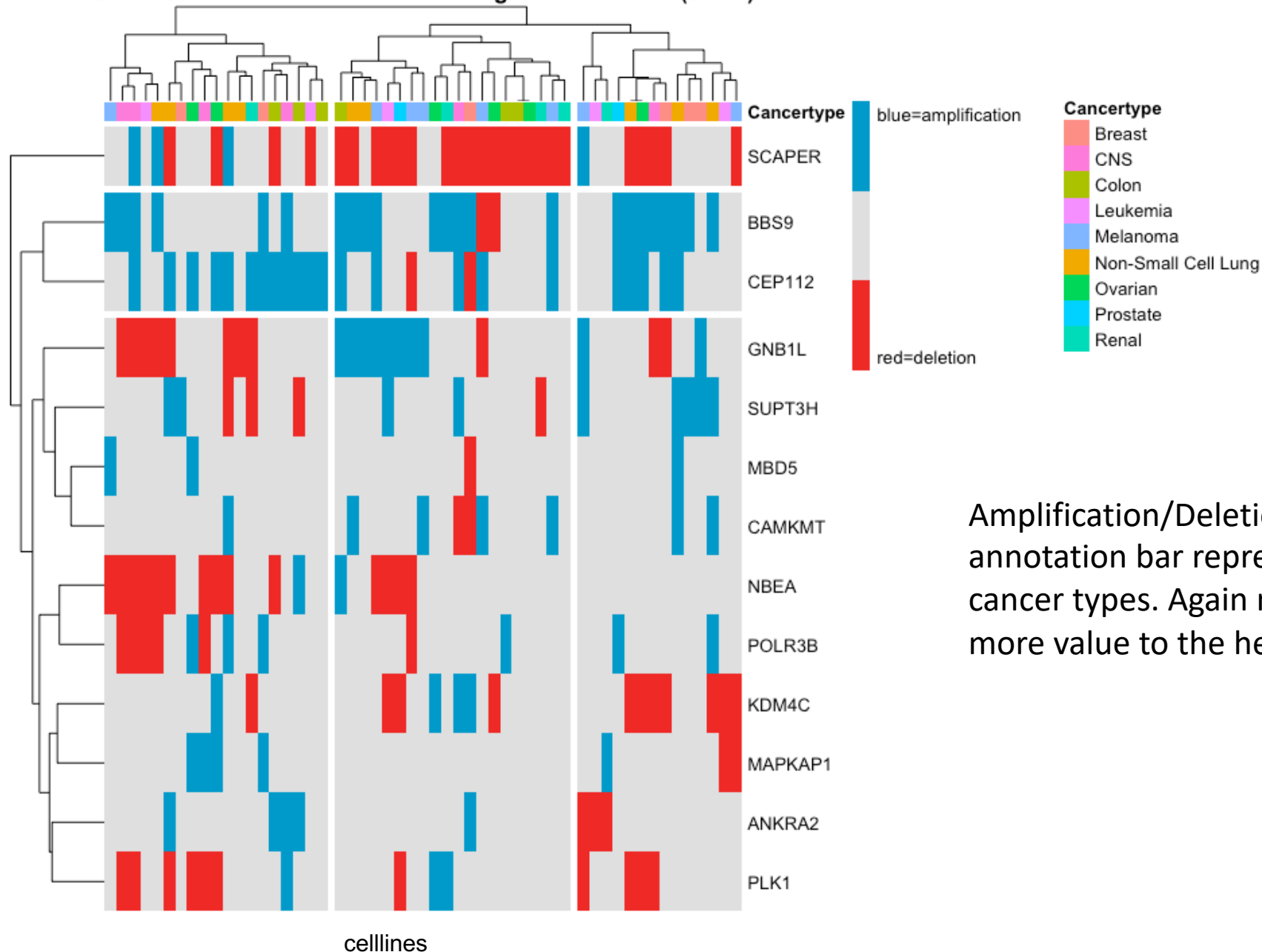


# Part III: Further analysis of biomarker – amplification, deletion or neutral gene copy number?

Genes were identified as amplified, if their gene copynumber is higher than the **80%-quantile** and as deleted if the value is lower than the **20%-quantile**.



# Connection between biomarker and gene alterations (80:20)



Amplification/Deletion heatmap with annotation bar representing the different cancer types. Again not sure if it really adds more value to the heatmap.