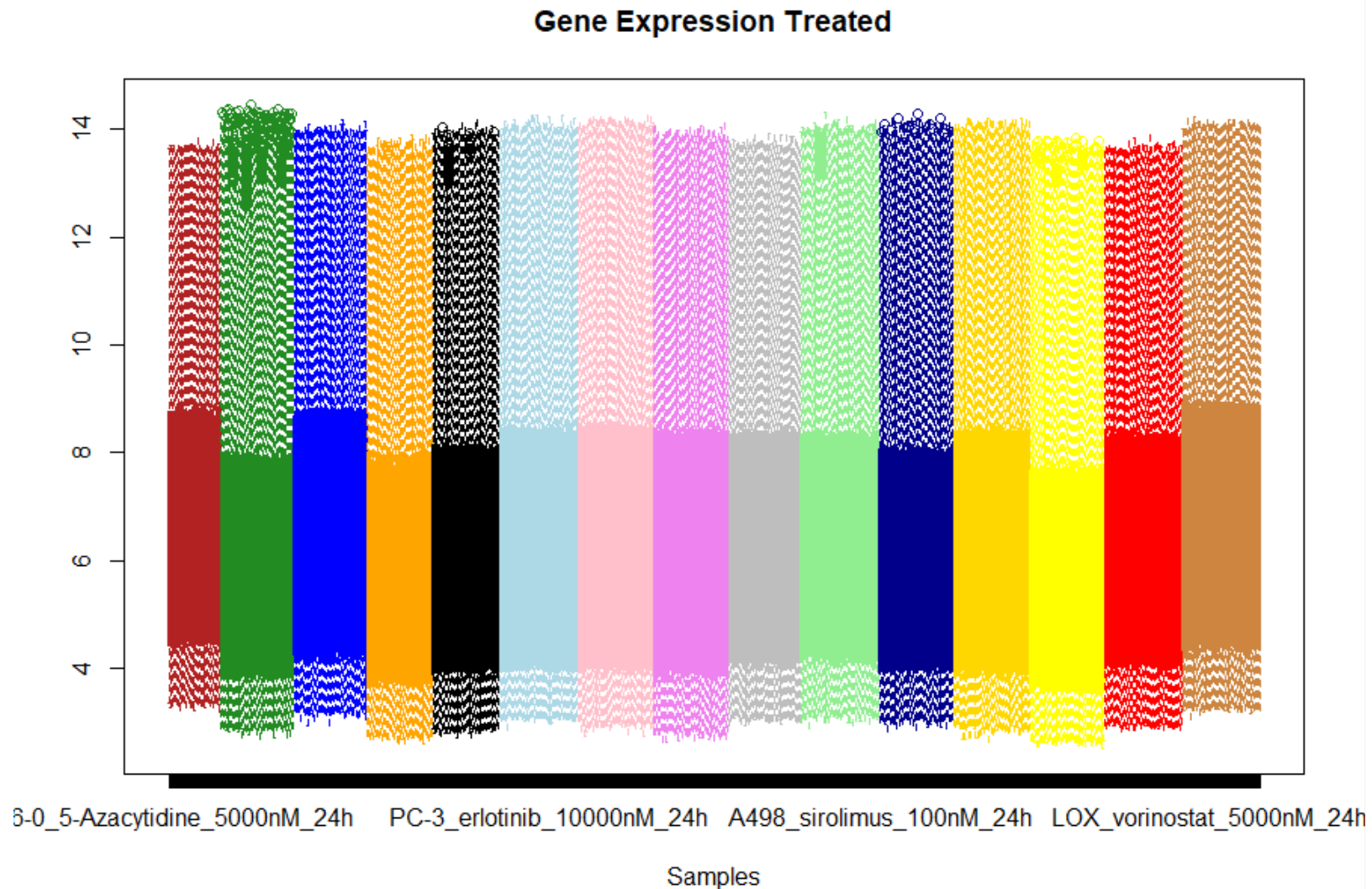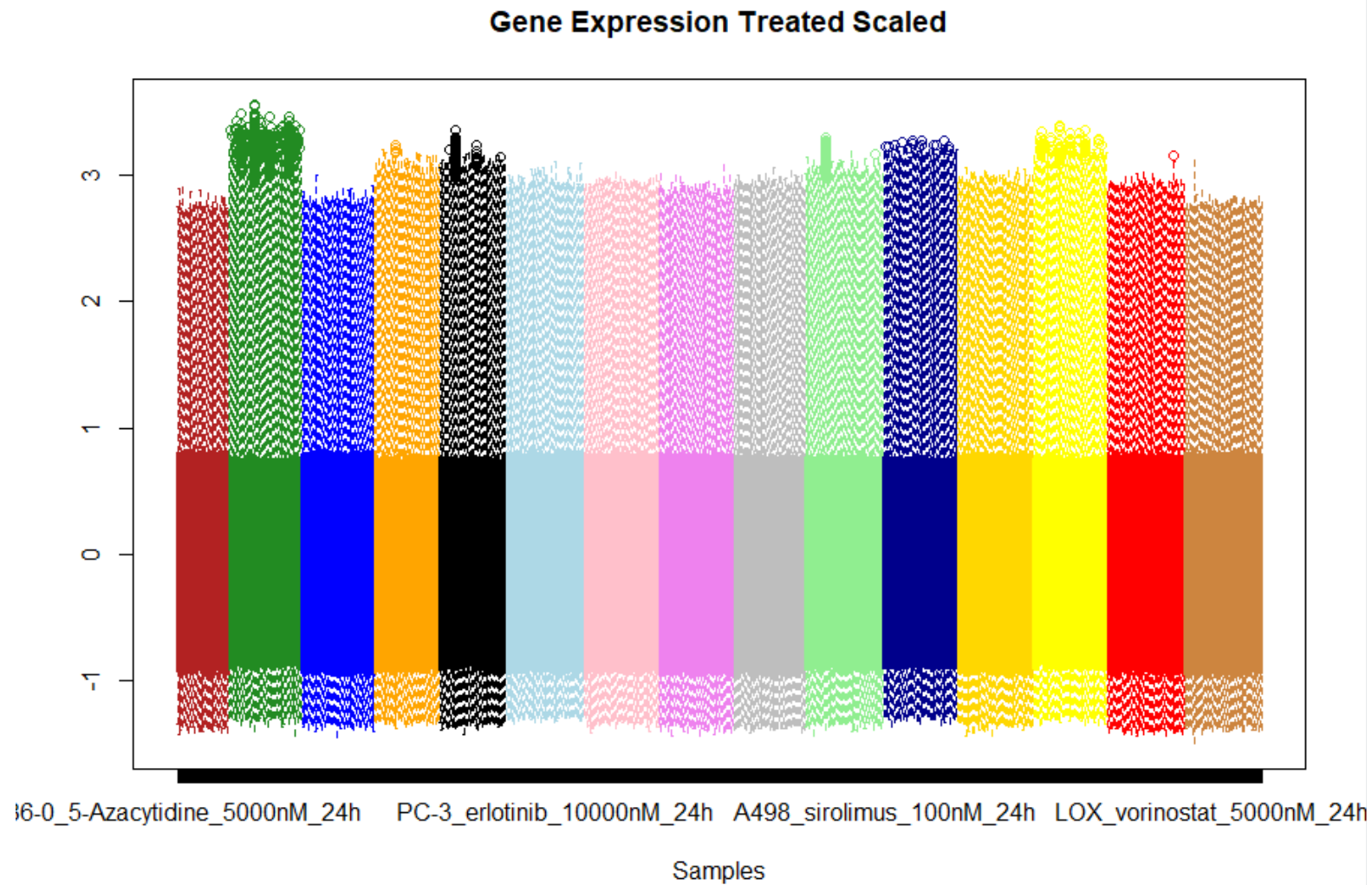# Plots

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.

Gene Expression Treated

6-0_5-Azacytidine_5000nM_24h   PC-3_erlotinib_10000nM_24h   A498_sirolimus_100nM_24h   LOX_vorinostat_5000nM_24h

Samples

Through scaling the data the different batches disappeared.


Gene Expression Treated Scaled

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.

Log2 FC Treated/Untreated

N = 10891881   Bandwidth = 0.006492
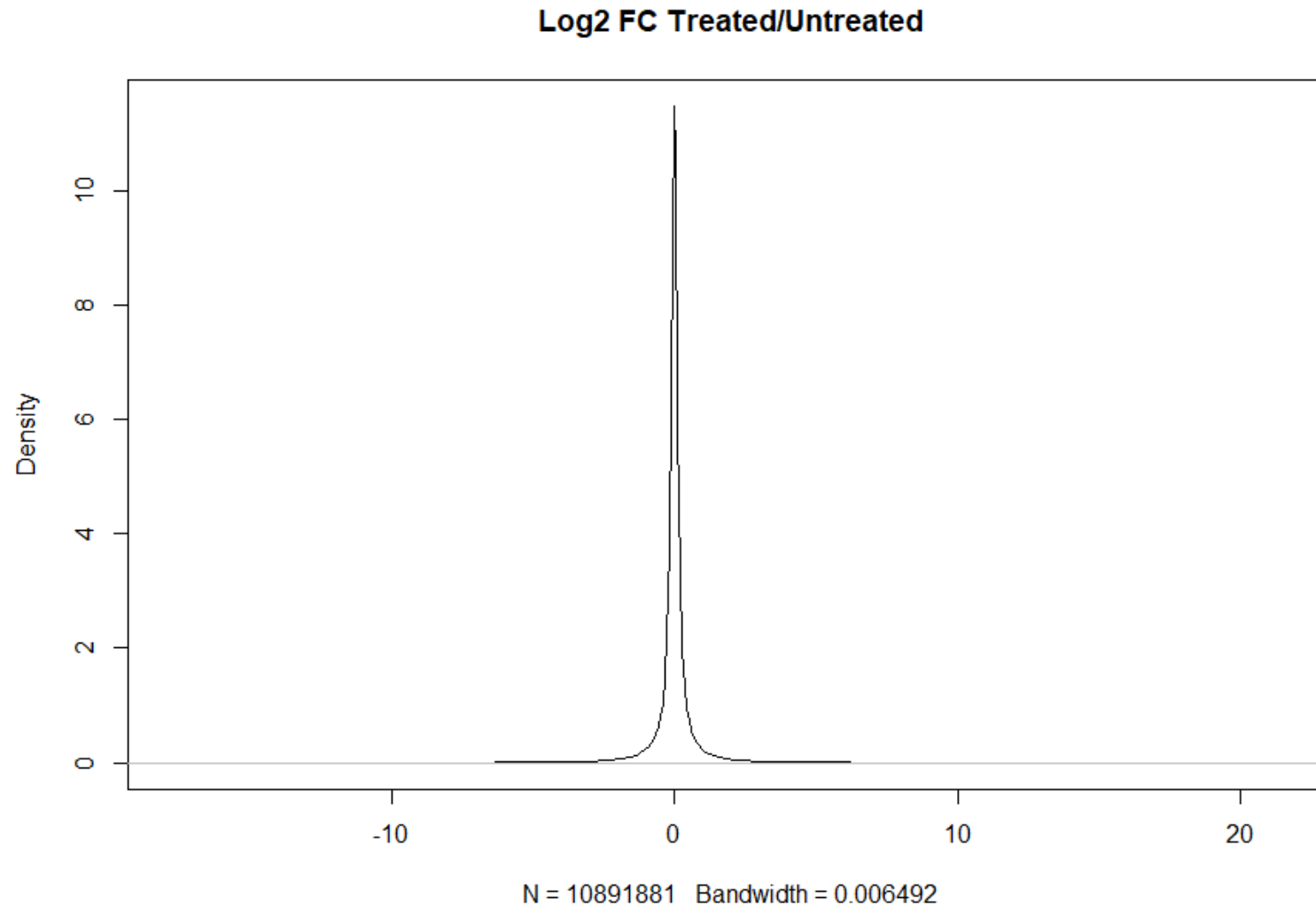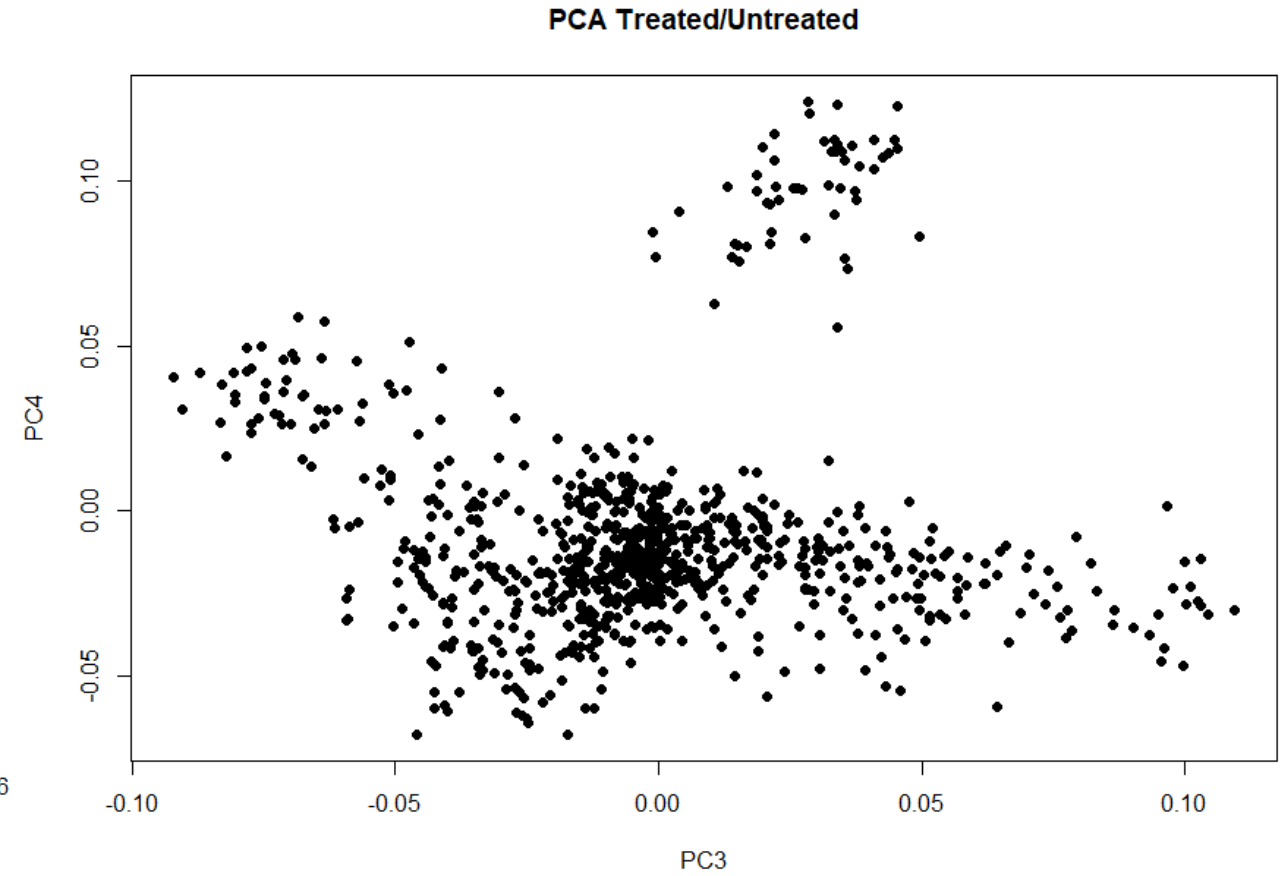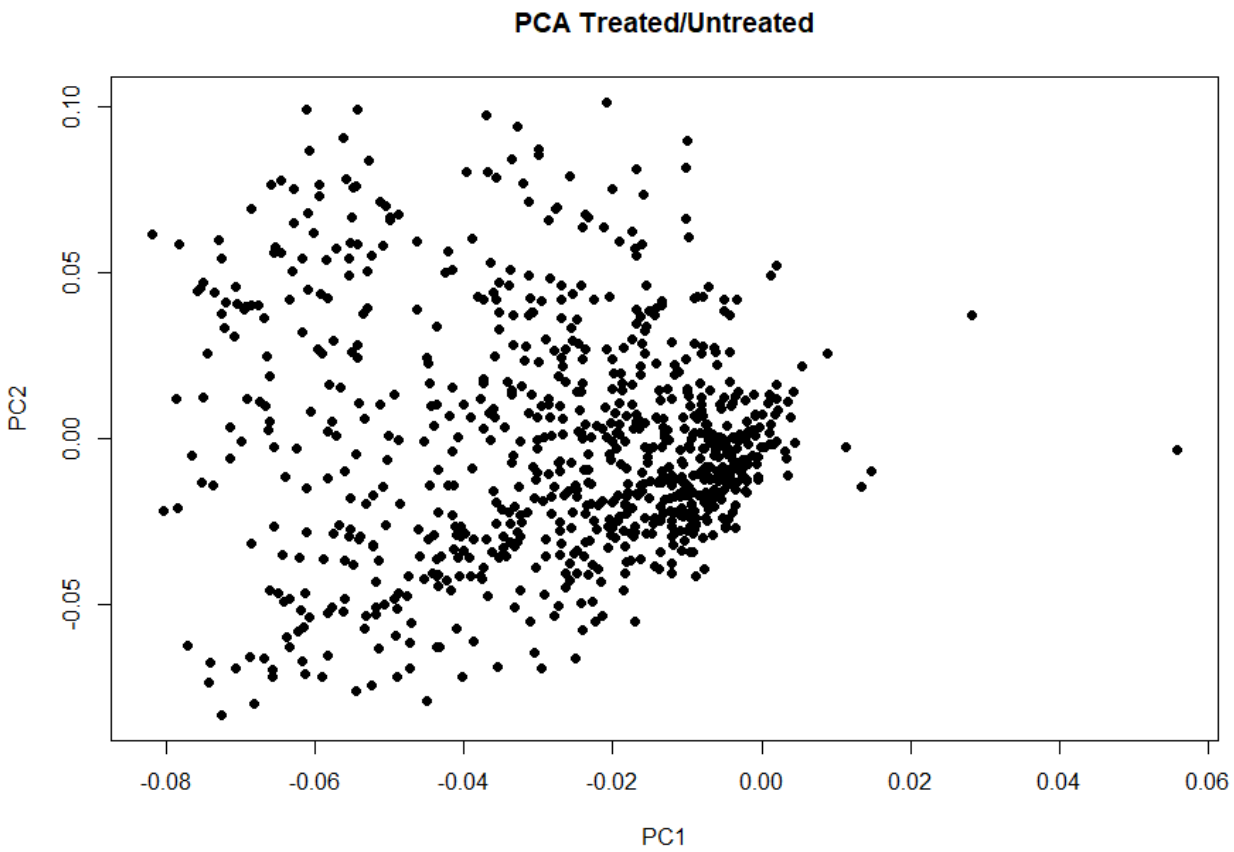
x-Achse: log2FC-values

in Überschrift: density of... schreiben

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

PCA.FC

x-Achse: PC
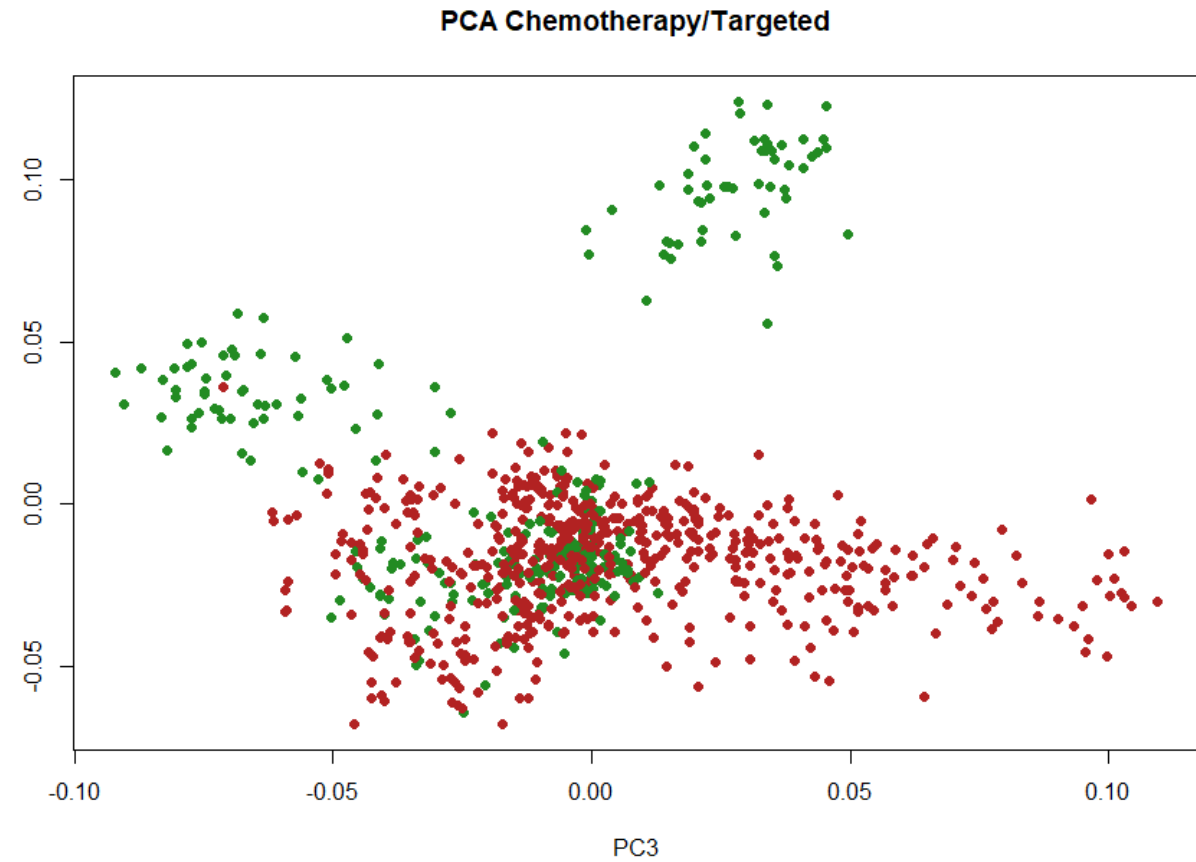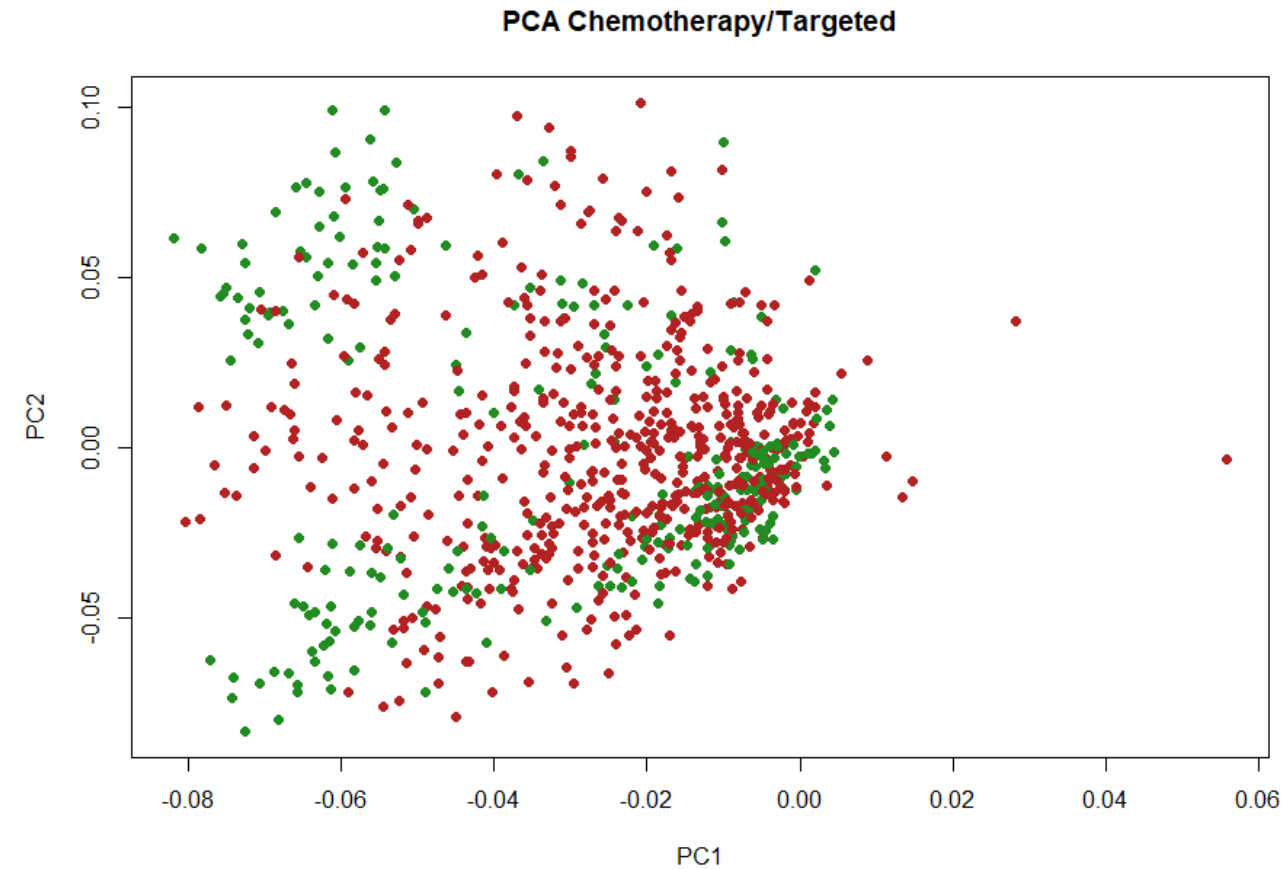y-Achsenbeschriftung verlängern (0 bis 20 oderso)
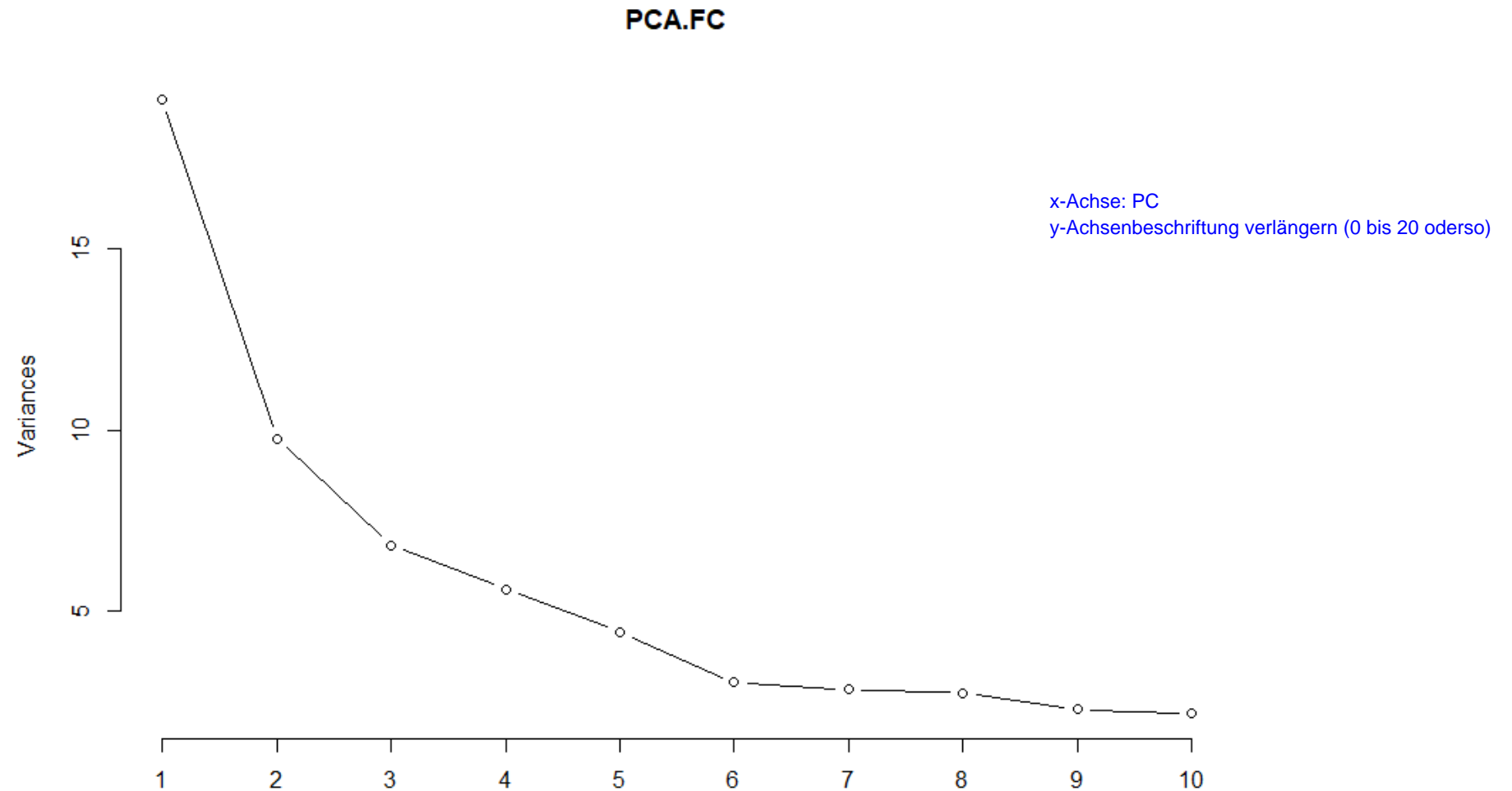
Plotting: PC1 against PC2 and

PC3 against PC4

PC Werte von FC in Überschrift

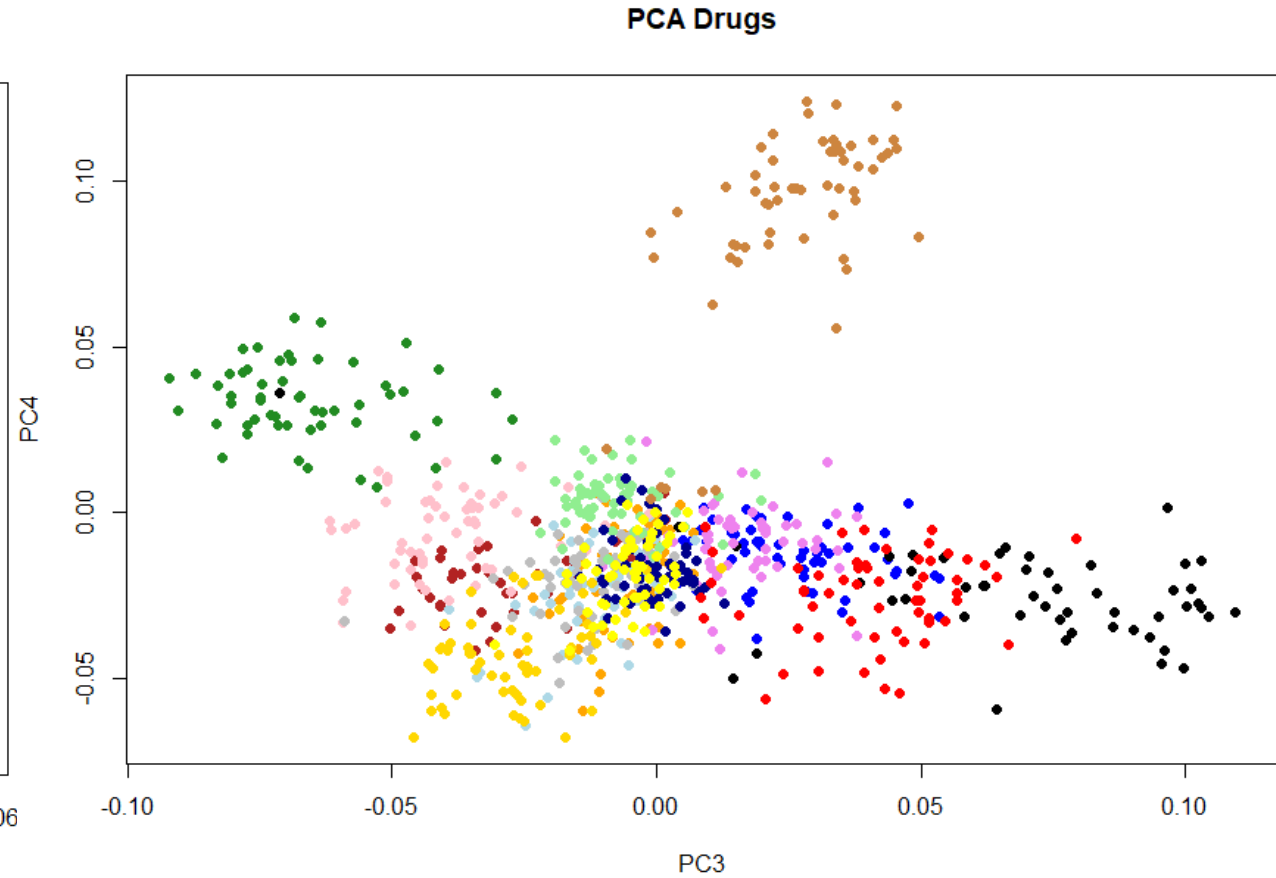# Color chemotherapy agents (firebrick) against targeted therapy agents (forestgreen) in those two plots.
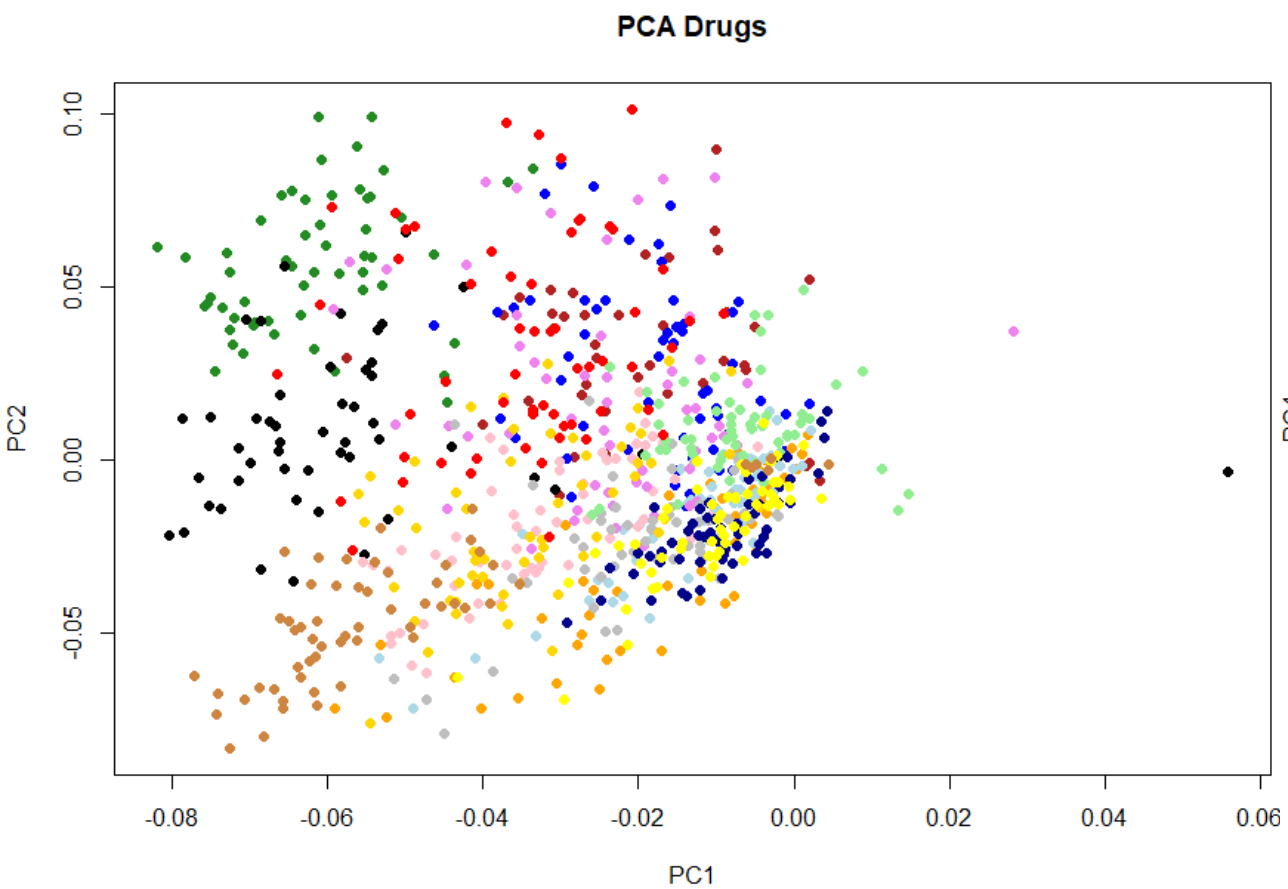

PCA Chemotherapy/Targeted

PCA Chemotherapy/Targeted

Farbe als Legende dazu

und siehe letzte Folie

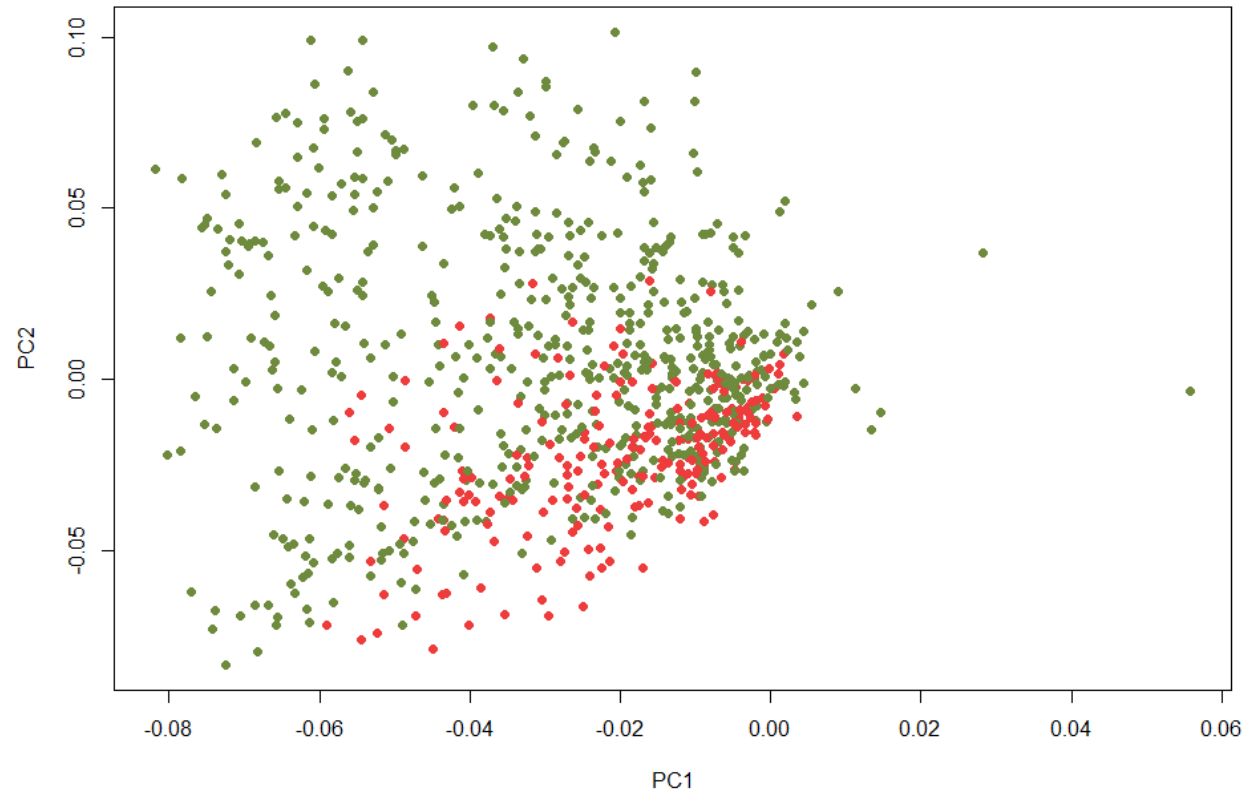PC4 wurde abgeschnitten

Color every drug in a different color.

PCA Drugs

PCA Drugs
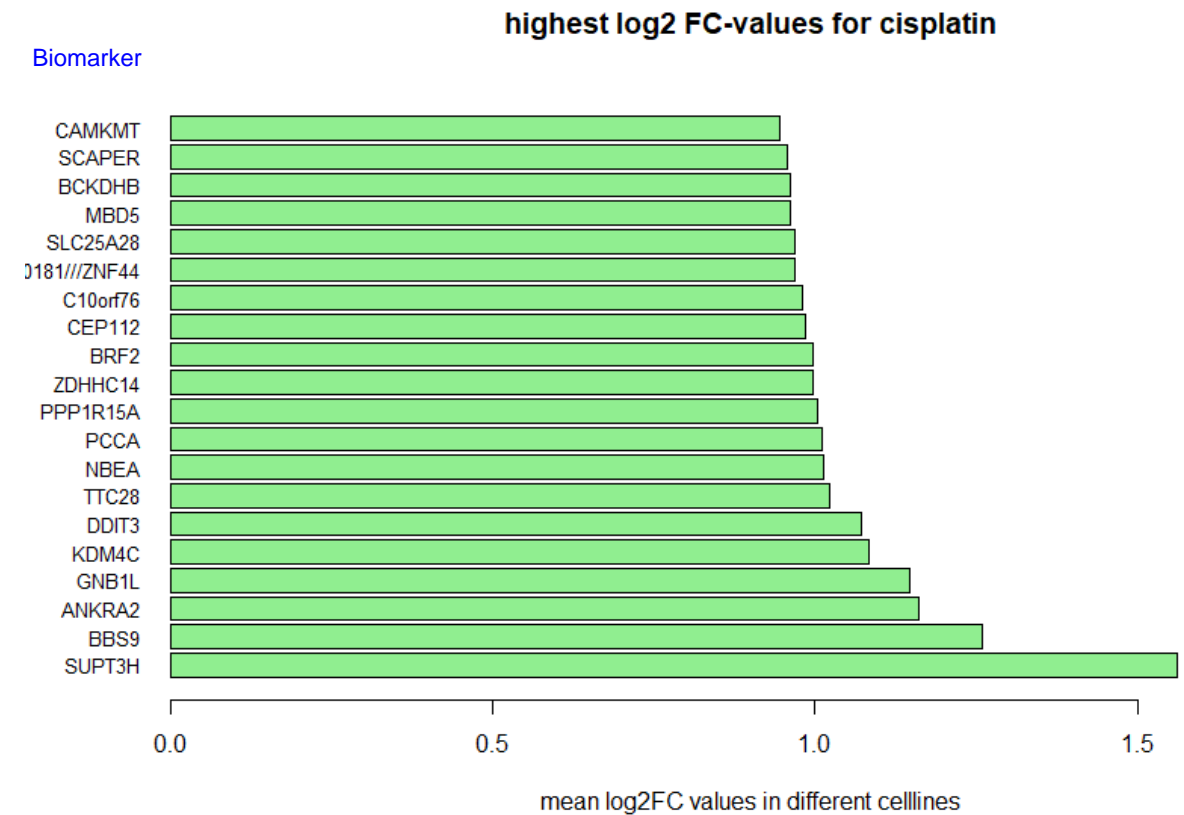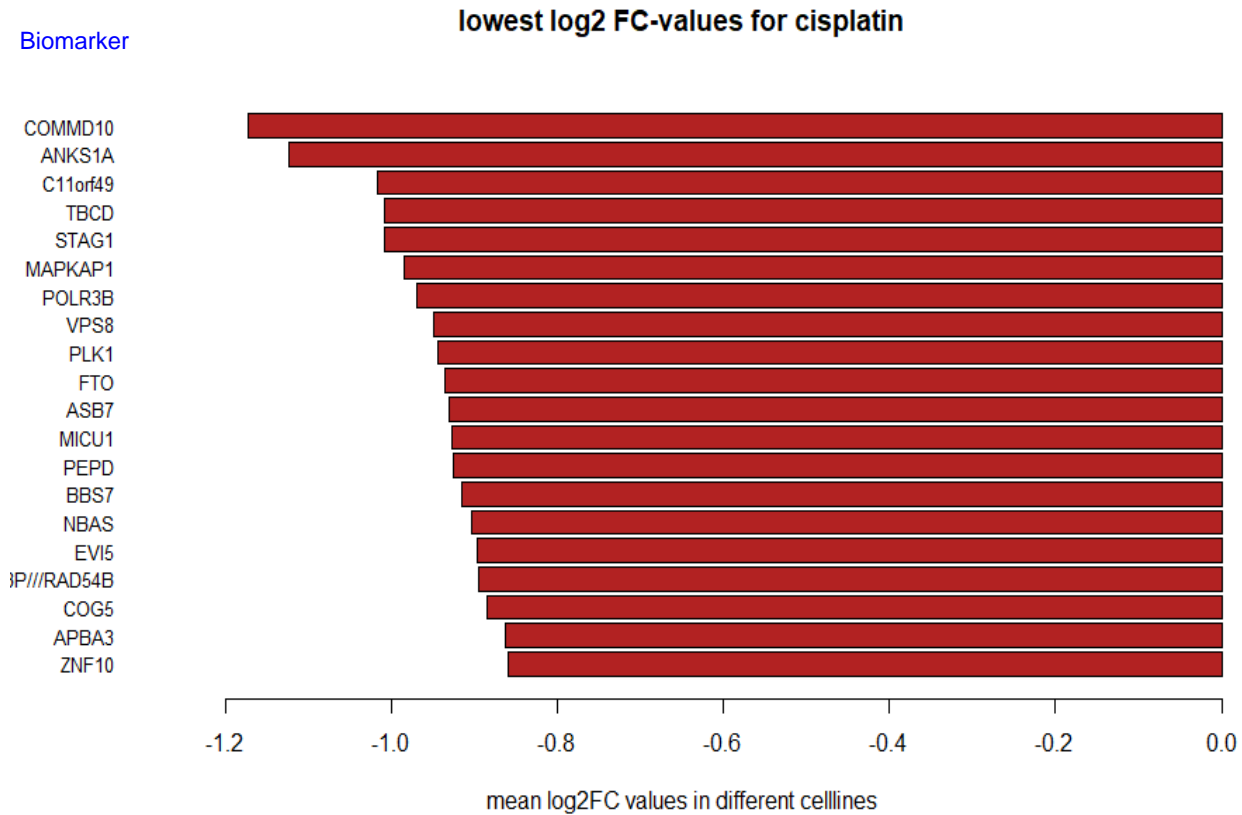
Color tyrosine kinase inhibitors red (others: green) in those two plots.
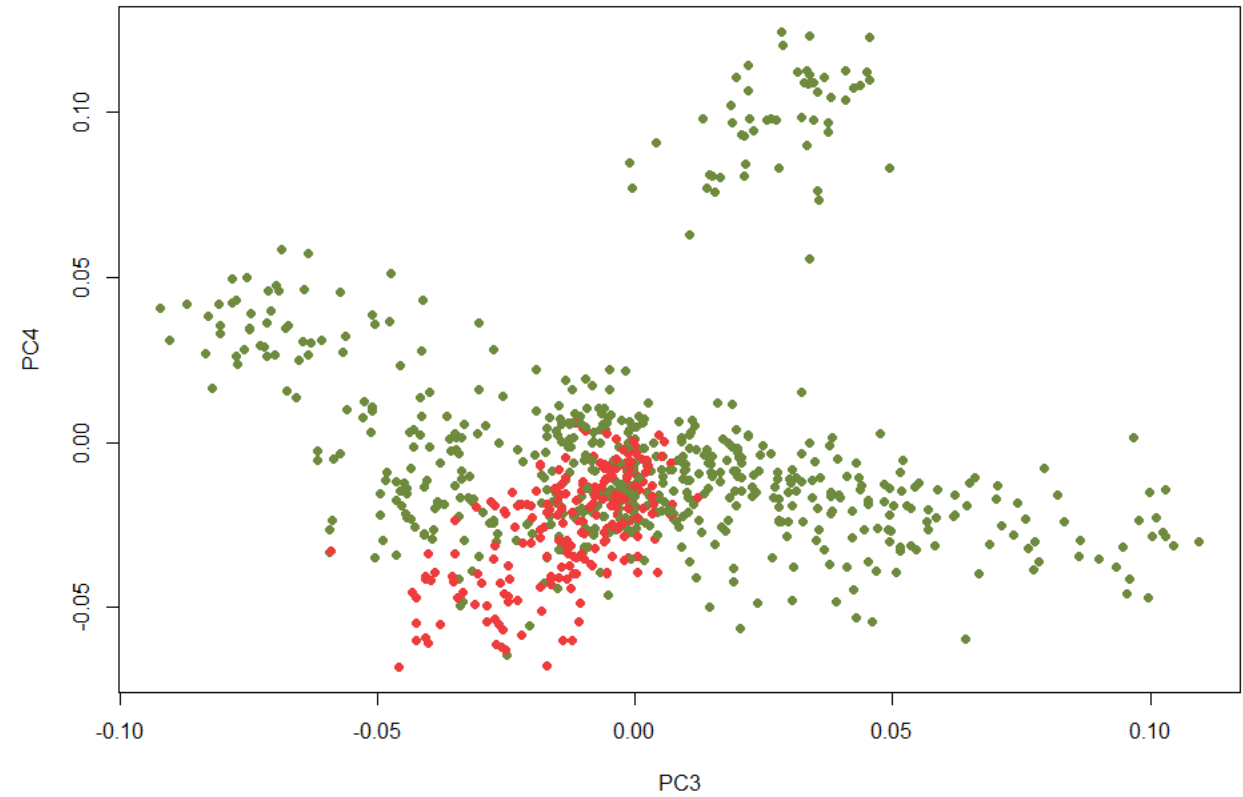
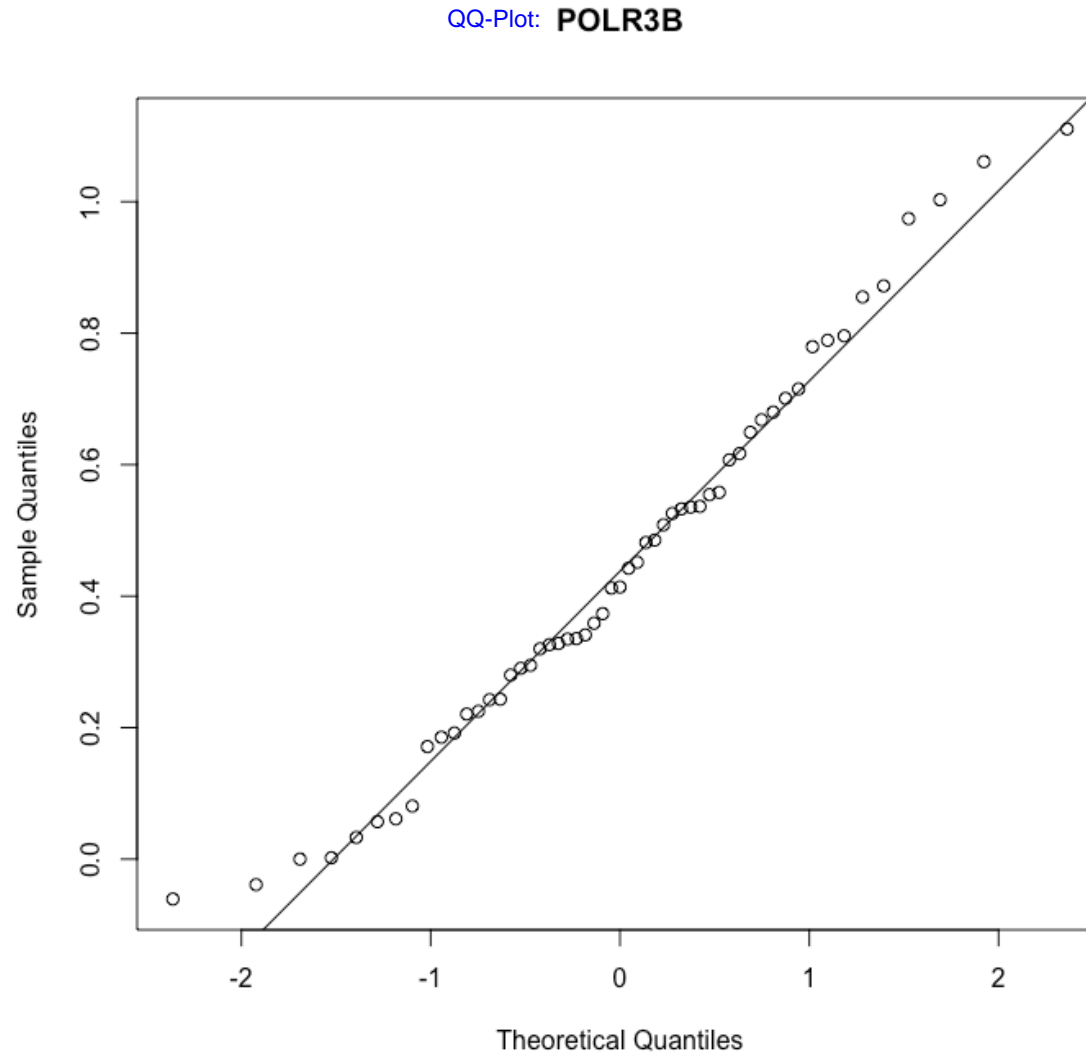# Specific Analysis

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

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

QQ-Plot: **POLR3B**
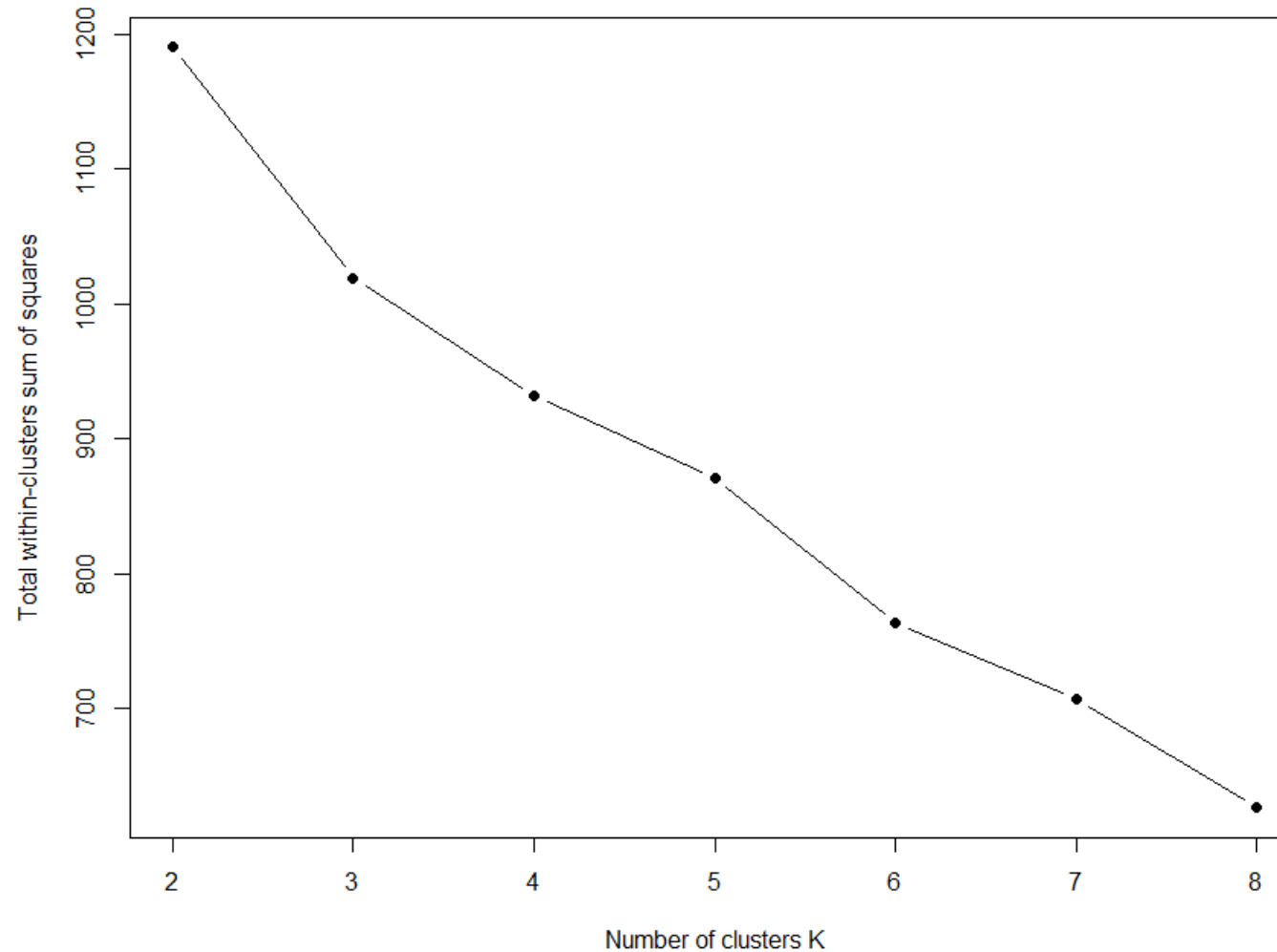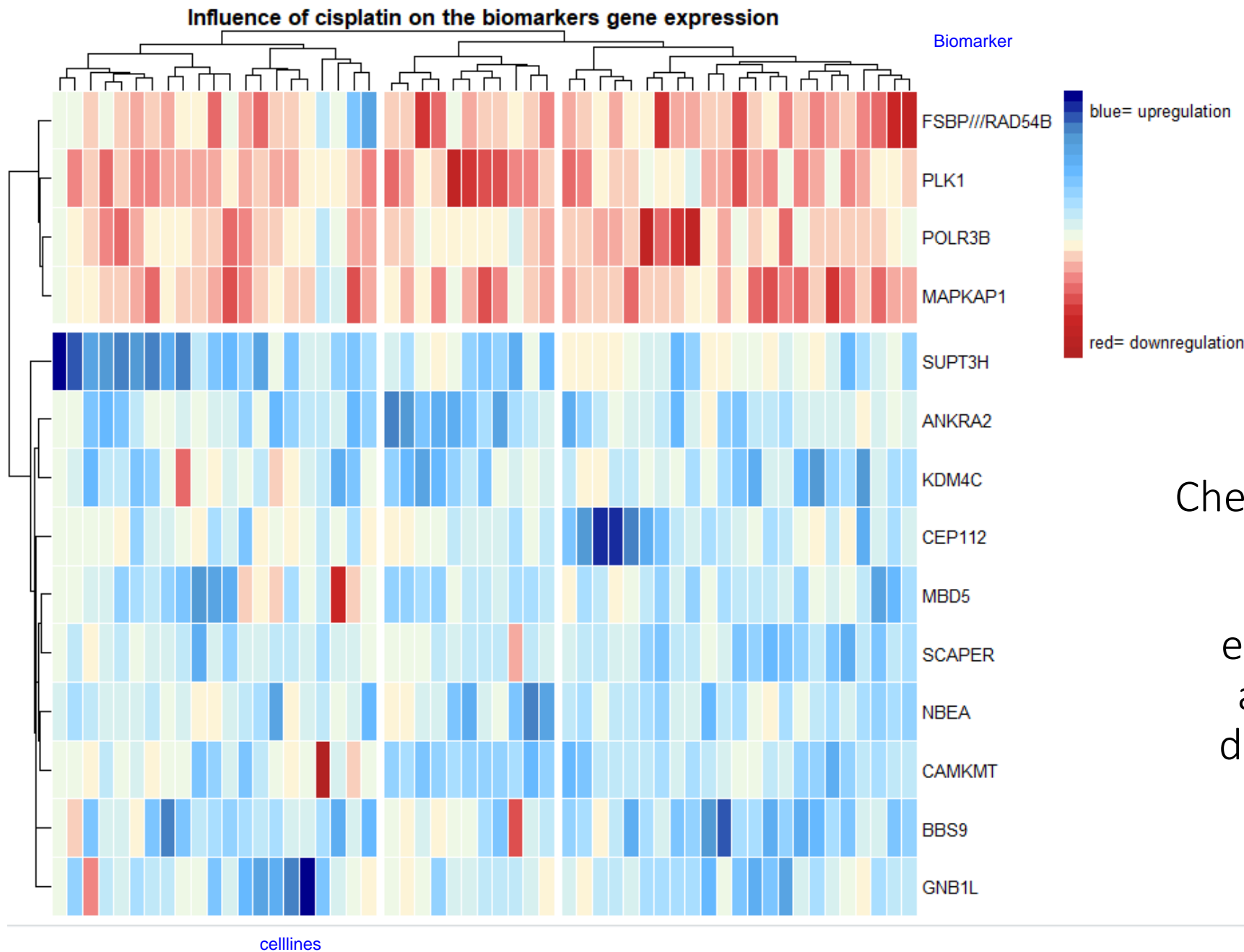
dazuschreiben welche Werte wir genommen haben: treated.cisplatin

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

Elbowplot kmeans clustering of celllines

Influence of cisplatin on the biomarkers gene expression

Biomarker

blue= upregulation

red= downregulation

FSBP///RAD54B
PLK1
POLR3B
MAPKAP1
SUPT3H
ANKRA2
KDM4C
CEP112
MBD5
SCAPER
NBEA
CAMKMT
BBS9
GNB1L
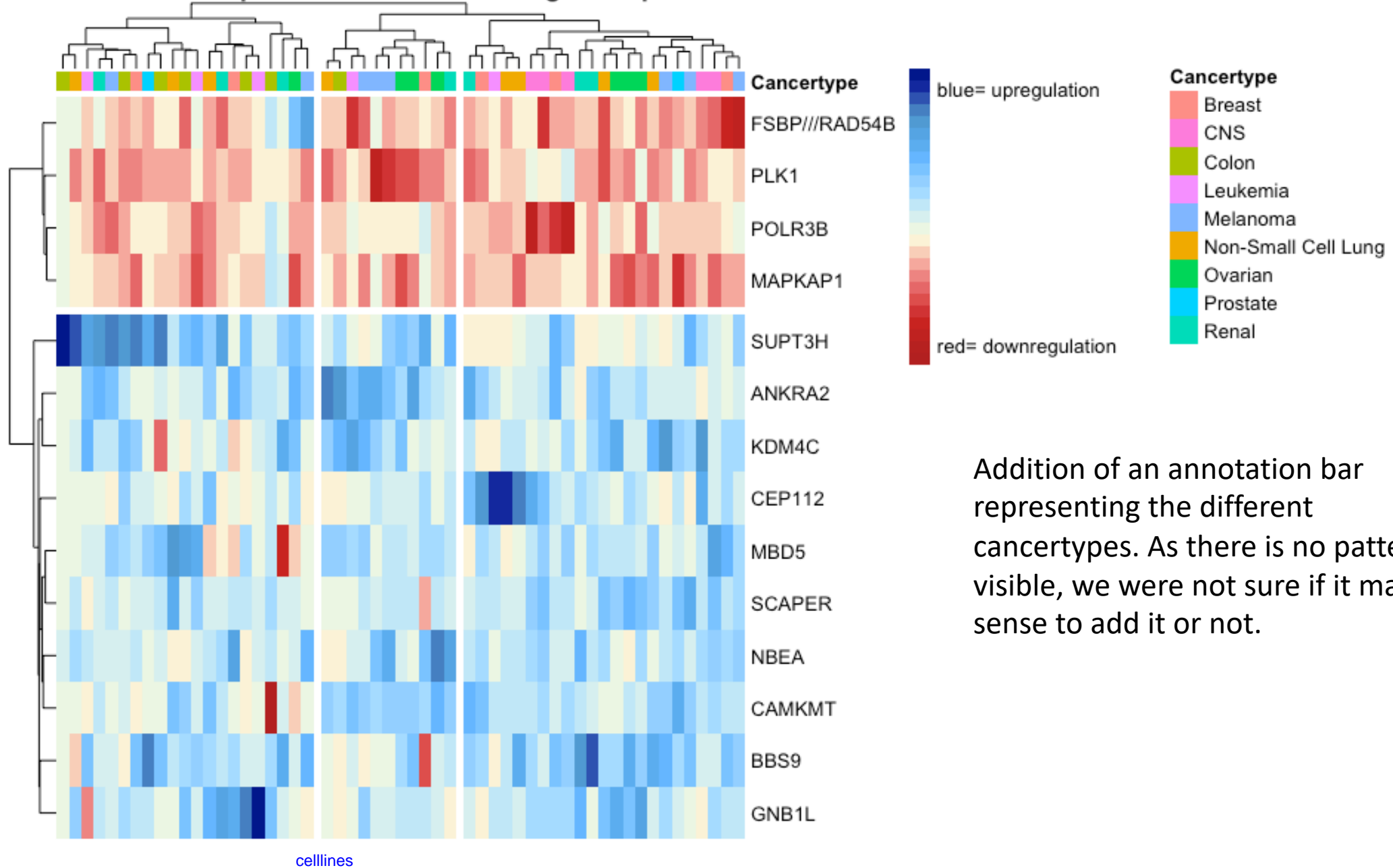
celllines

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 cancertypes. 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.
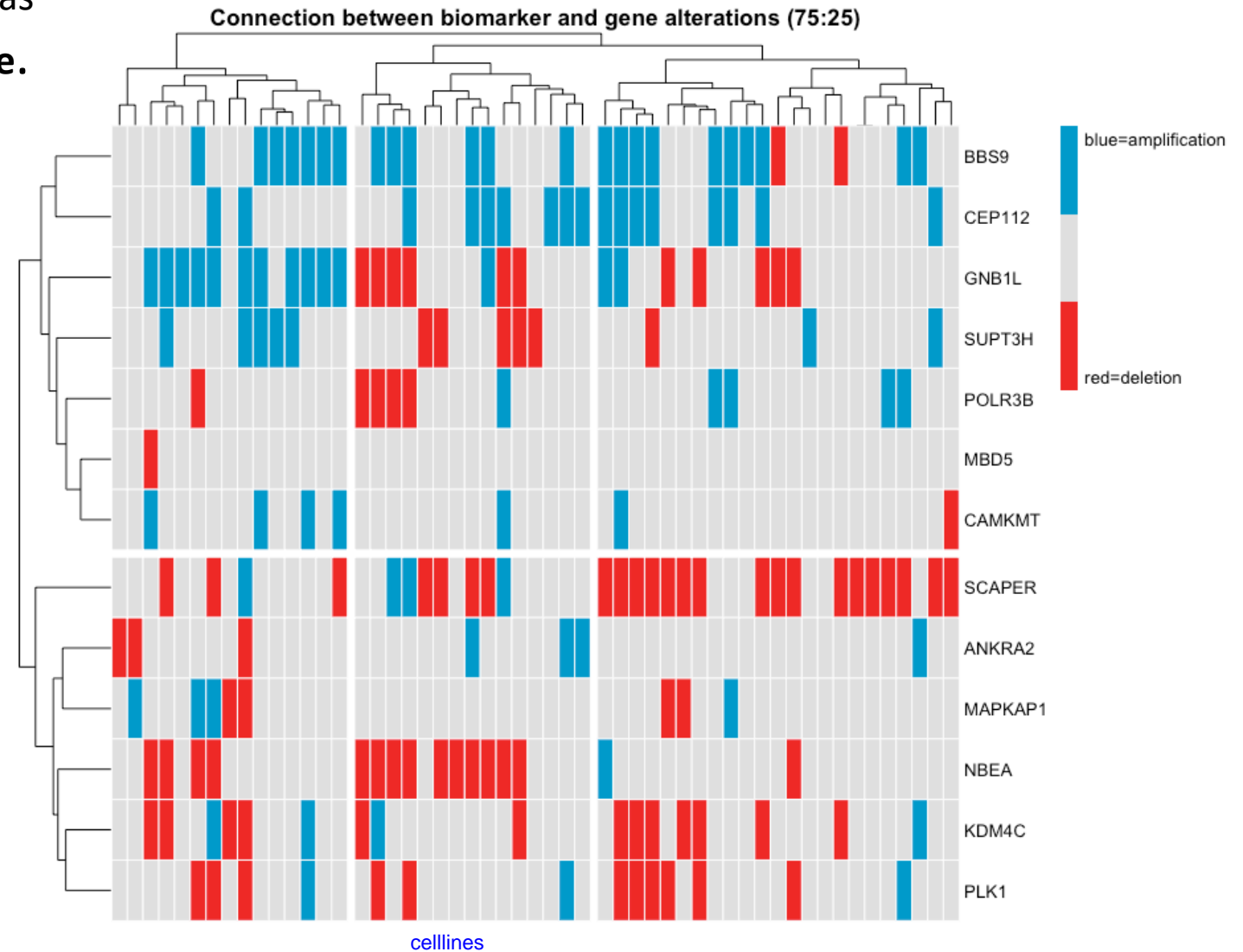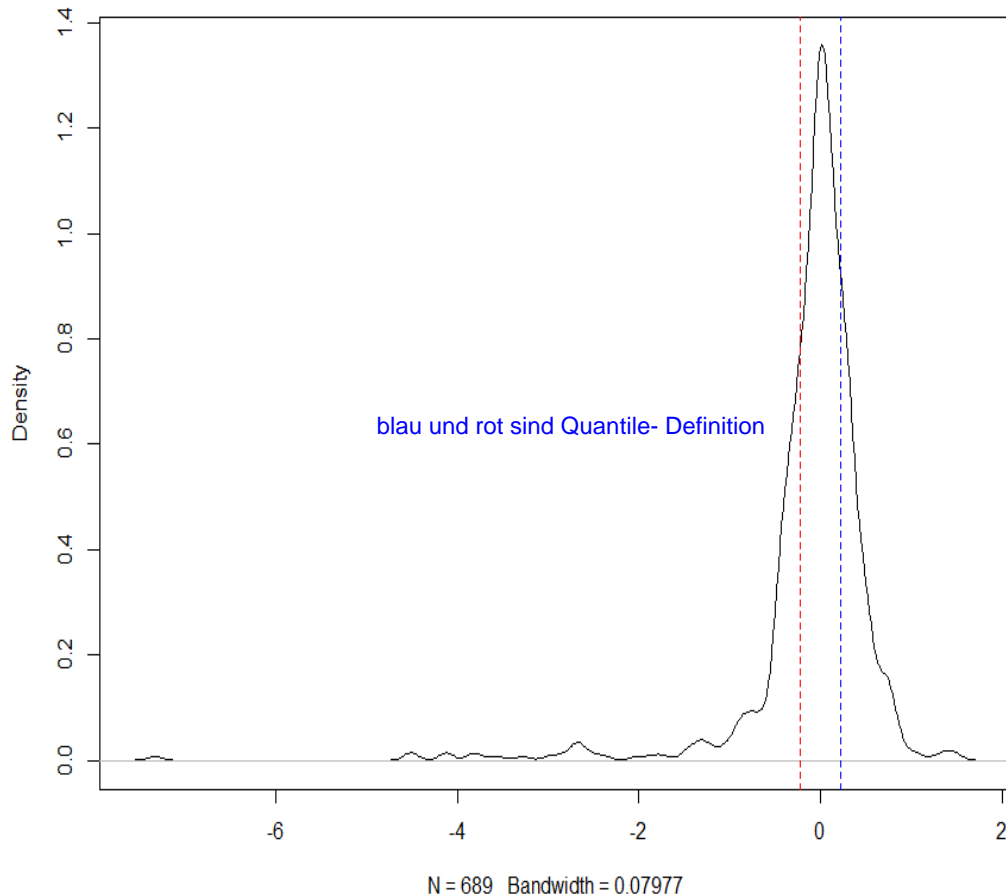## Checking for the optimal number of clusters. for the following heatmap

Elbowplot kmeans clustering - biomarker gene alterations
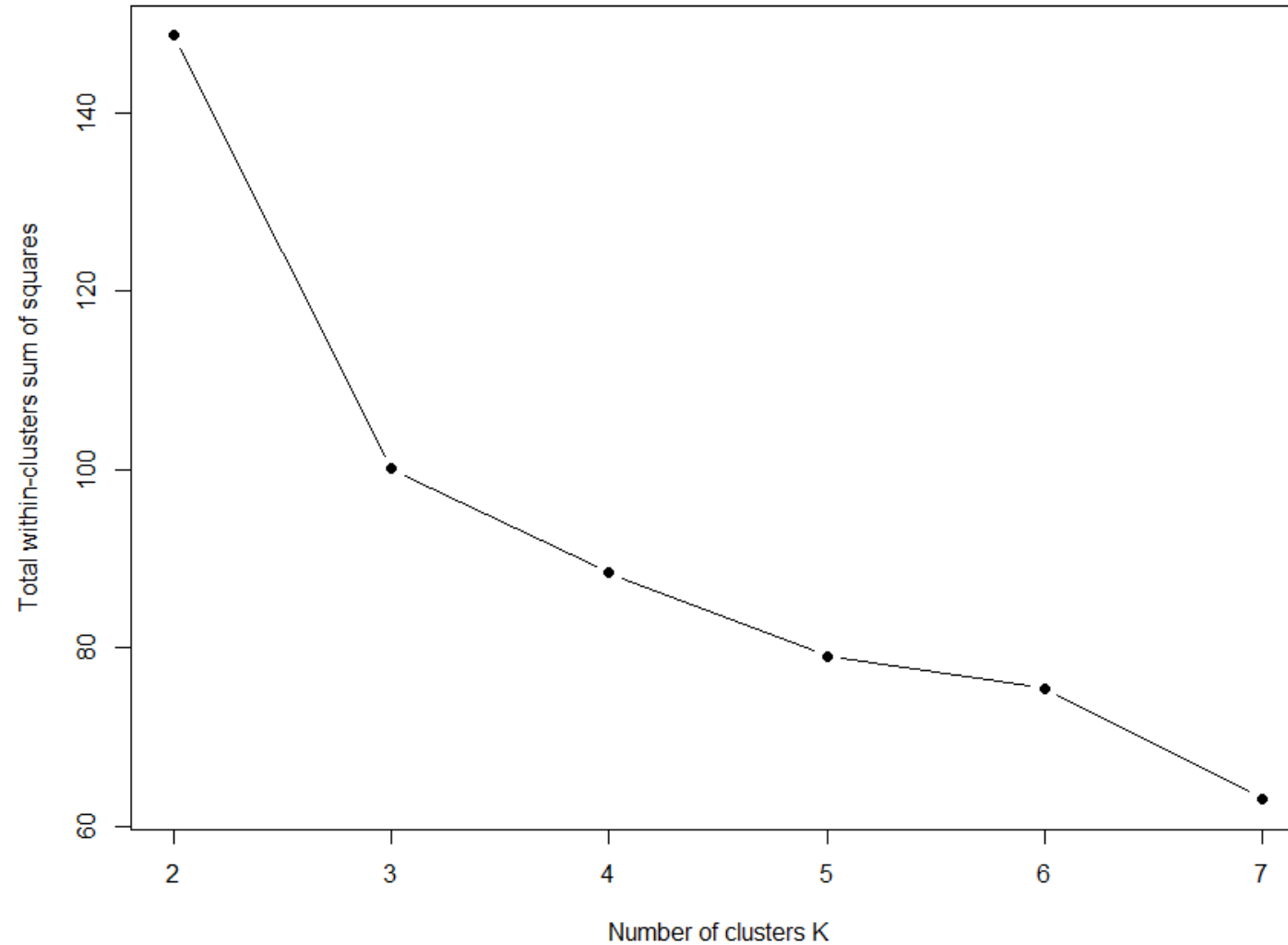
Part III: Further analysis of biomarker – amplification, deletion or neutral gene copy number?

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.**

density plot - copynumber biomarker
density.default(x = copynumber.biomarker)

blau und rot sind Quantile- Definition

N = 689   Bandwidth = 0.07977

Connection between biomarker and gene alterations (75:25)

blue=amplification

red=deletion

celllines

Part III: Further analysis of biomarker – amplification, deletion or neutral gene copy number?
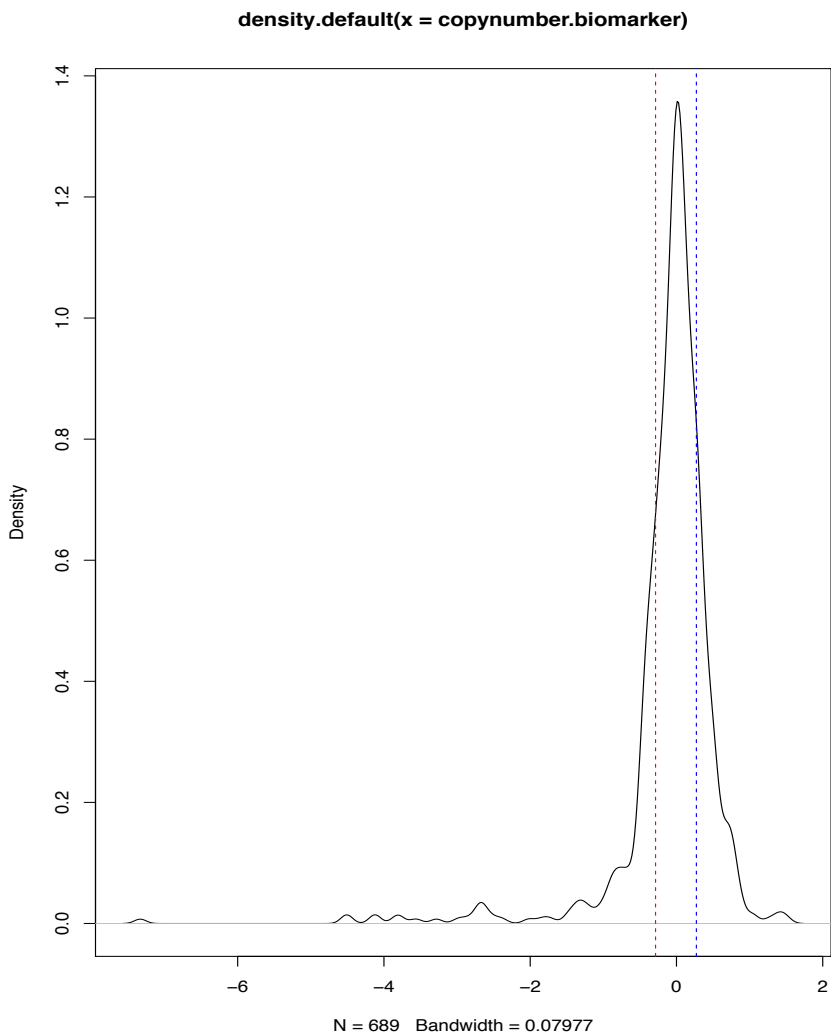
Genes were identified as amplified, if their gene copynumber is higher than the **80%-quantile** and as deleted if the value is lower than the **20%-quantile.**

density.default(x = copynumber.biomarker)

N = 689   Bandwidth = 0.07977
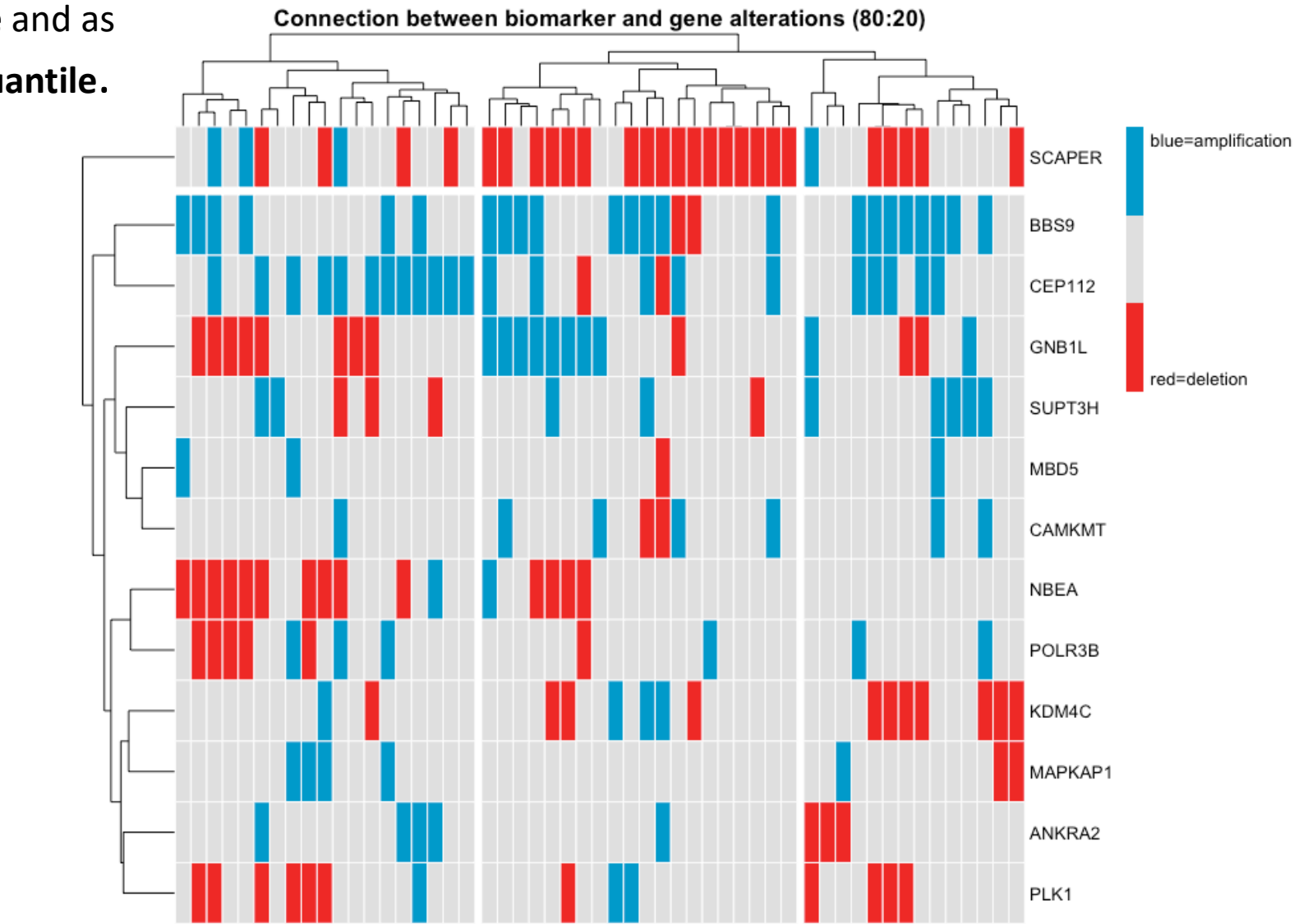
siehe letzte Folie

Connection between biomarker and gene alterations (80:20)

blue=amplification

red=deletion

Connection between biomarker and gene alterations (80:20)
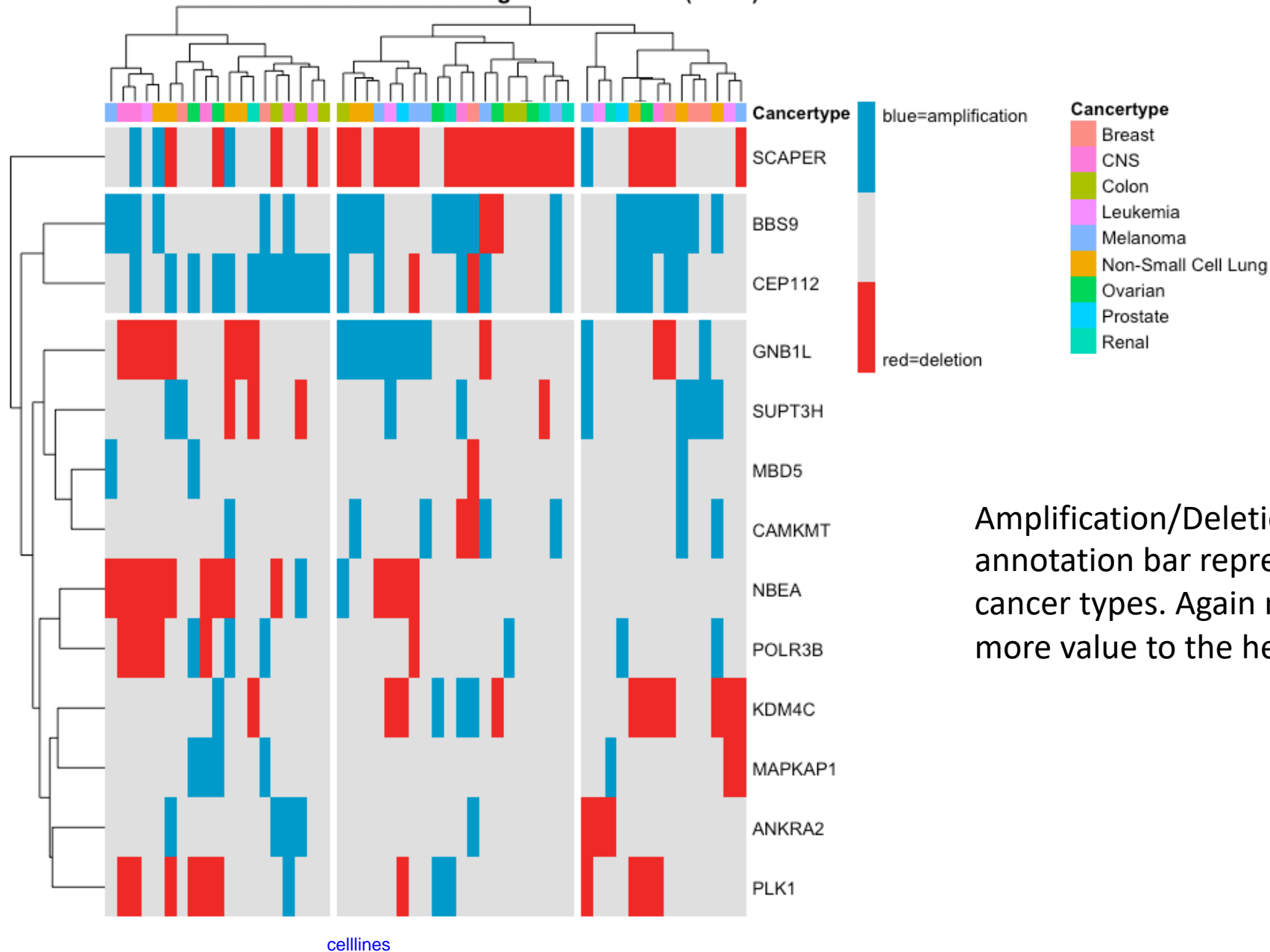
Amplification/Deletion heatmap with annotation bar representing the different cancer types. Again not sure if it really adds more value to the heatmap.