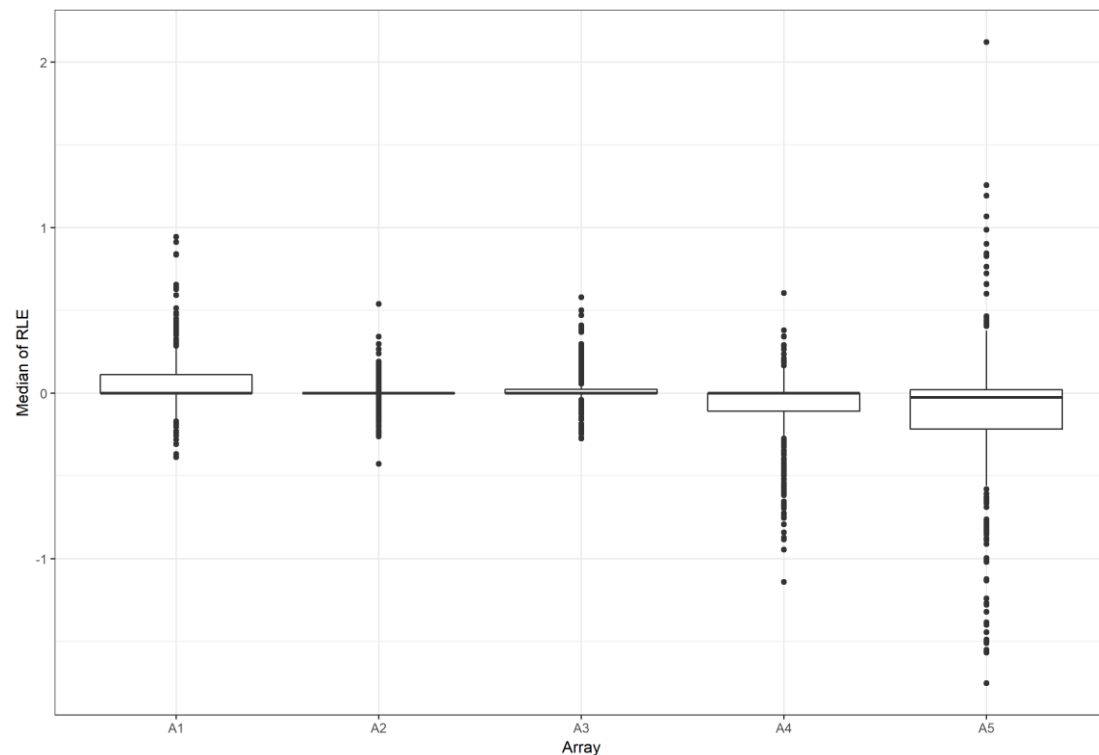


Name: Huajie Chen
Matrikelnr: 4199962

Boxplot of RLE from different Arrays



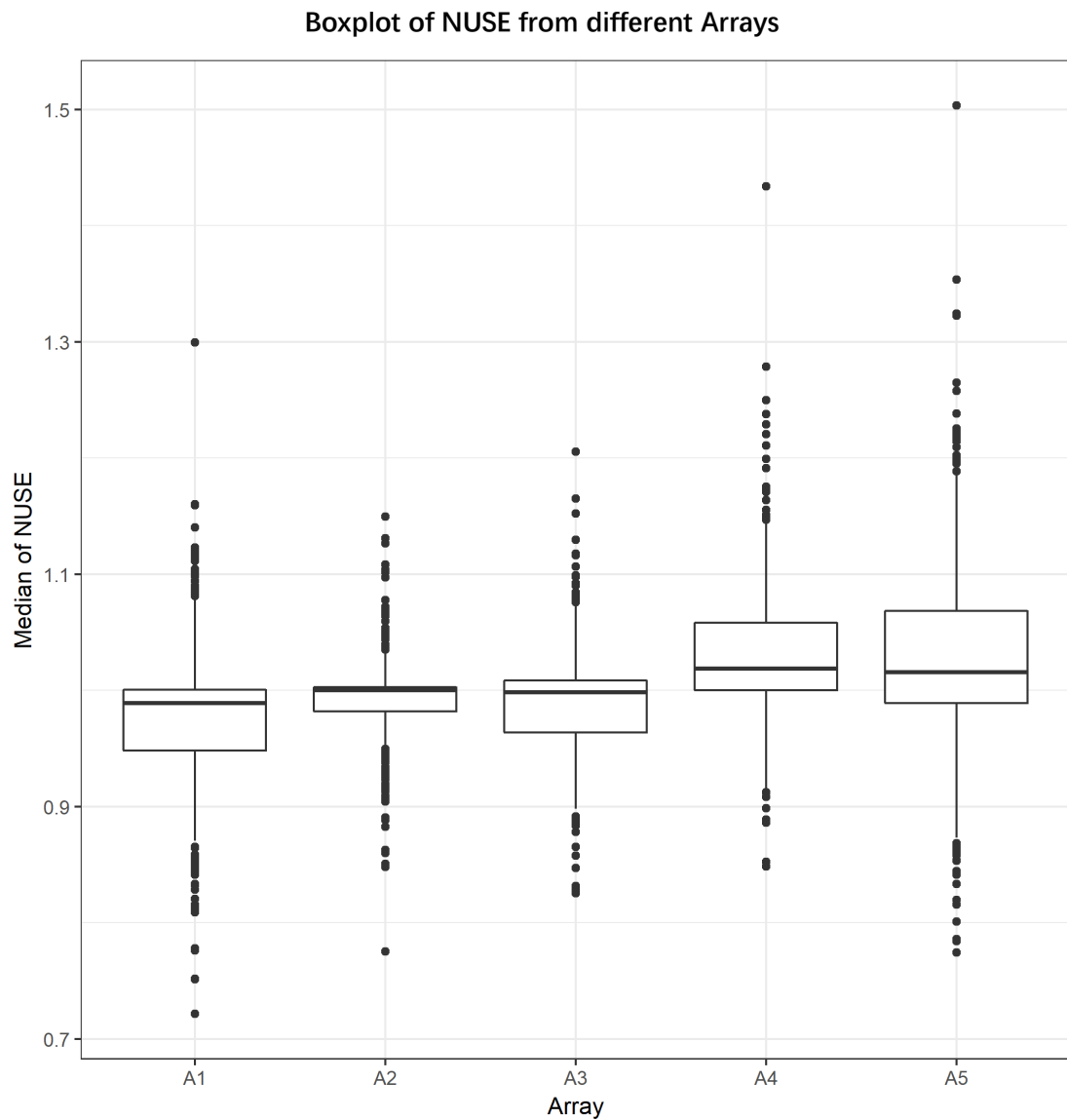
Discuss what you would expect from arrays with a median RLE not close to zero? What does an array with a large RLE IQR indicate?

If an array is with a median RLE not close to zero, then this array is in general with a low/high logarithmized (primary) expression values of the genes, which means this array is on average worse/better than the other arrays.

A large RLE IQR indicates that the fluctuation of expression value on this array is large, which means it is not stable as the ones with lower RLE IQR.

Draw boxplots of the respective distributions for each array and compare. What do you conclude?

According to the boxplot, A1 tends to show a higher expression value in general. A2 and A3 are with lower fluctuation. The Q1 of A4 is beneath zero, which states that its expression level is generally lower. A5's median of RLE is not close to zero and it is with a large IQR, meanwhile the Q1 is far beneath 0. Therefore A5 is not suitable for this experiment.



From this boxplot a conclusion may be drawn that the median of NUSE from A1 is generally lower than 1, which means its fluctuation of gene expression value is lower than the others. That makes it a better Array for this experiment than the rest. The median of NUSE from A2 and A3 is close to 1, which are acceptable. As for A4 and A5, their medians are far from 1(> 1) and their IQRs are also too large. In this case, they are not suitable for this experiment.