**Final report Group 4**

1. **Introduction**

RNA-binding proteins are proteins that bind to the double or single stranded RNA in cells and by this interact with transcripts of RNA-driven processes. This large group of proteins play key roles in RNA processing and modification (Alternative splicing, RNA editing and polyadenylation), export, mRNA localization and translation. The malfunction of RBP´s underlies the origin of many diseases from neurogenerative disorders to cancer. Therefore it is necessary to increase the number of recognized RBP´s and the understanding of their molecular mechanisms.

For that reason, M.Caudron-Herger et al. created R-Deep with the aim to detect RBP´s automatically. In general, R-Deep is a proteome-wide, unbiased, and enrichment-free screen based on density gradient ultracentrifugation.

1. **Dataset exploration**

Our given dataset consists of mass-spectometry data from non-synchronized A549 cells which are human lung carcinoma cells of a caucasian male. In general, the data show the protein amount of each of our 3680 human proteins per fraction. The RDeep screen has been repeated three times so it comprises three replicates for each sample. All in all, we got 3680 rows with one human proteins per row and 150 columns for our Ctrl and RNase group for 25 fractions and three replicates each.

**2.1 Restructuring the data**

We created two seperate dataframes (3\*3680x25), one for our control group and one for our RNase group. The columns represent the name of the human protein and replicate 1-3. The rows show fraction 1-25 below each other.

1. **Normalization**