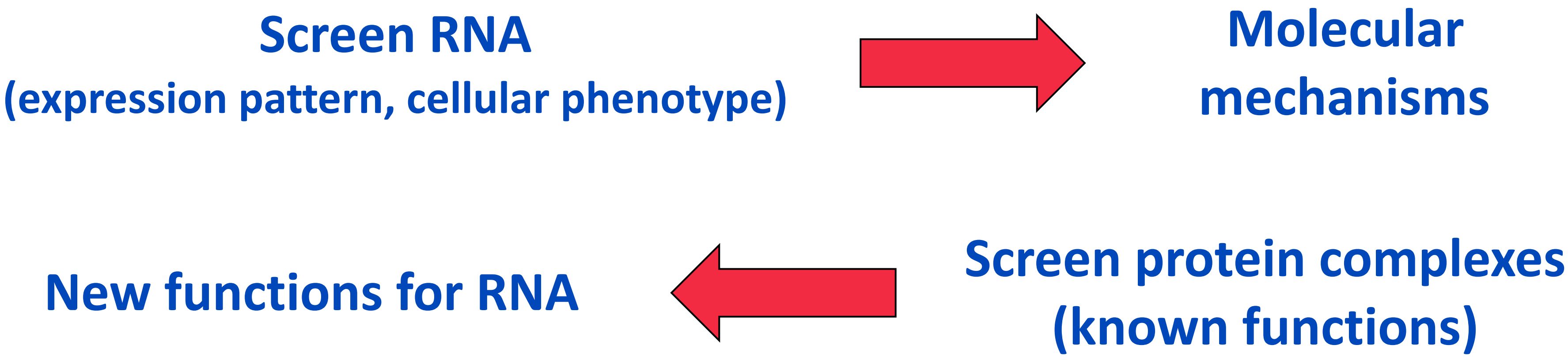
# Project 03: Proteome-wide screen for RNA-dependent proteins

**Dr. Maïwen Caudron-Herger / Fabio Rauscher**
Division "RNA Biology & Cancer”
German Cancer Research Center, Heidelberg

# RNA & Proteins: Challenges

## 1. Molecular mechanisms = biggest challenge in (nc)RNA research

**Idea: Reverse the screening process !**

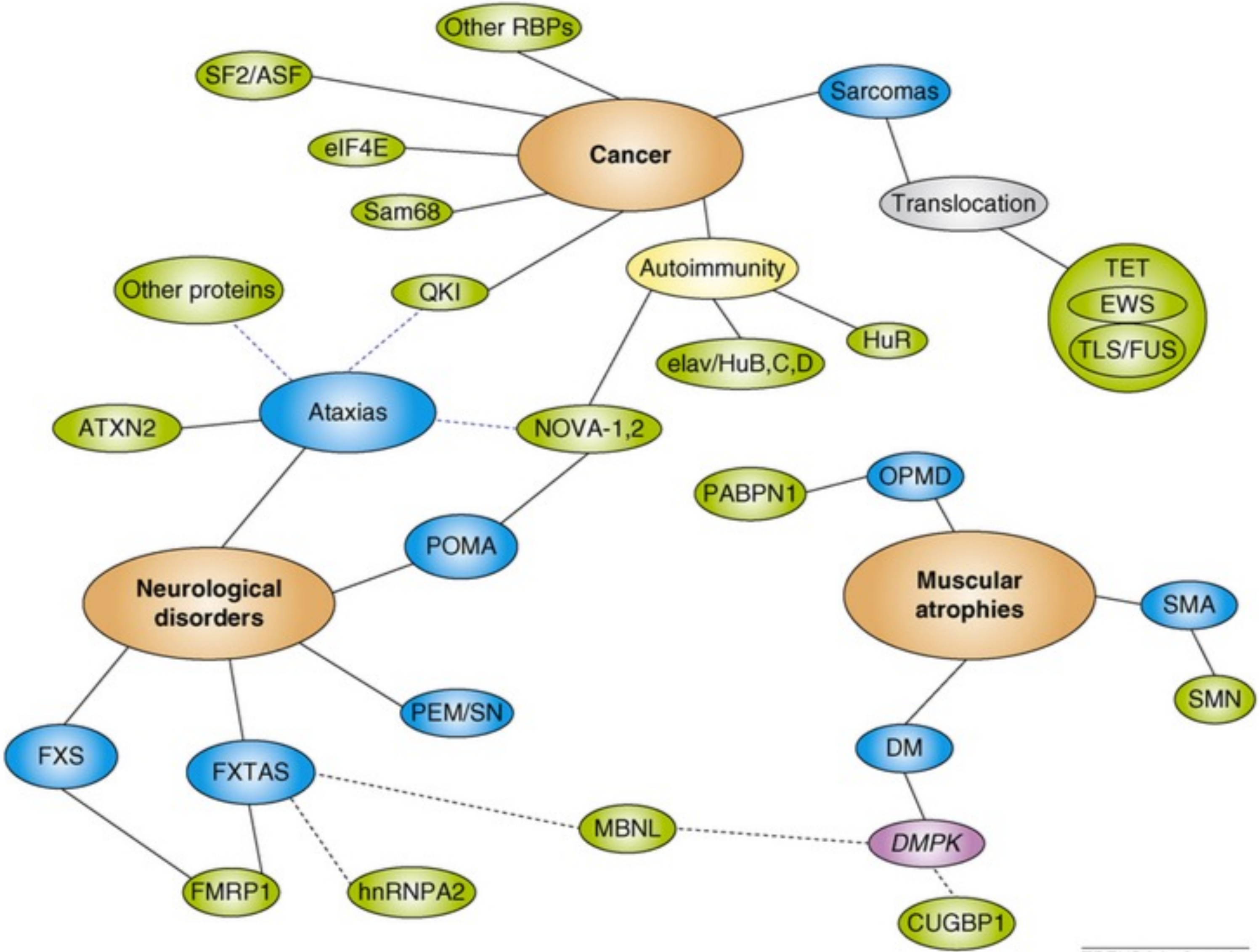


Screen RNA
(expression pattern, cellular phenotype)

Molecular
mechanisms

Screen protein complexes
(known functions)

New functions for RNA



TRENDS in Genetics

- Lukong et al. (2008). RNA-binding proteins in human genetic disease. Trends in Genetics *24*, 416-425.

# RNA & Proteins: Challenges

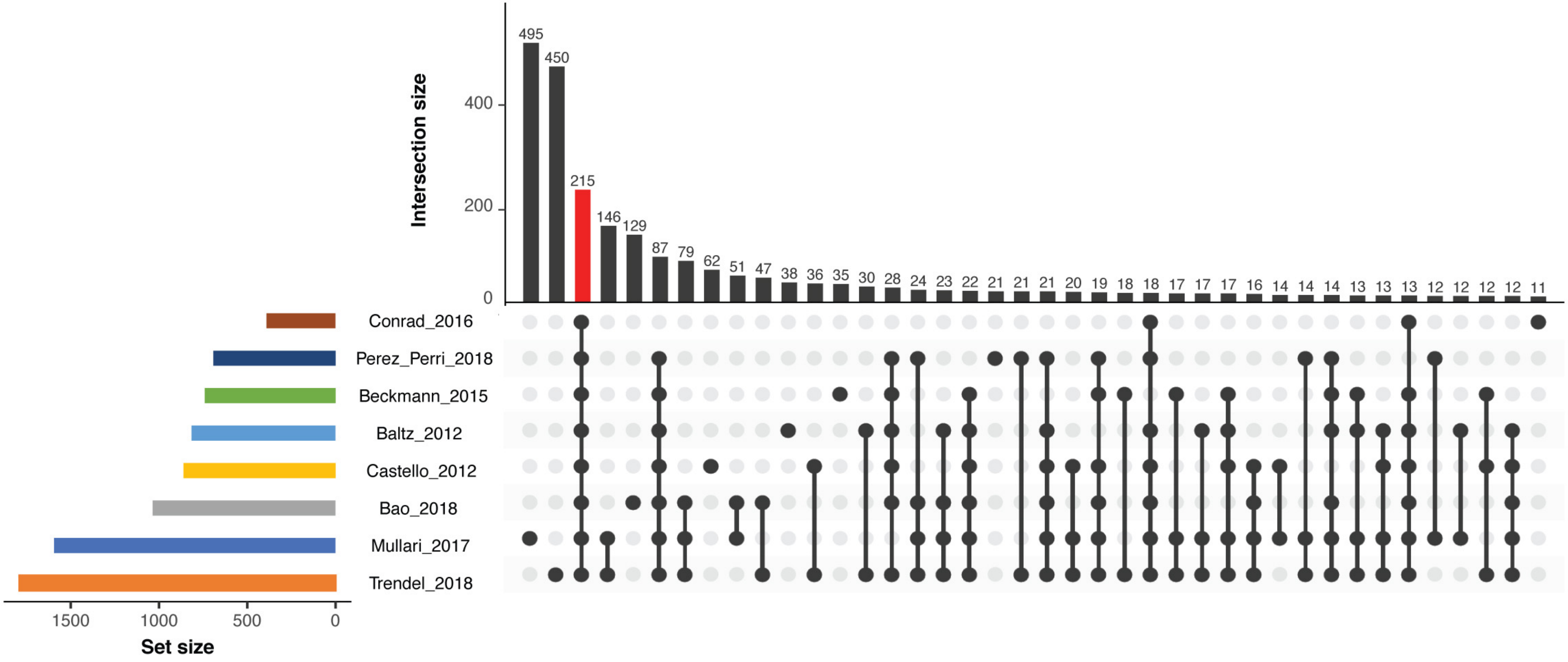
1. Molecular mechanisms = biggest challenge in (nc)RNA research

**Idea: Reverse the screening process !**

New functions for RNA ← Screen protein complexes (known functions)

2. RBPs mainly identified by RNA pulldowns

# Little overlap between RBP studies



# RNA & Proteins: Challenges

## 1. Molecular mechanisms = biggest challenge in (nc)RNA research

**Idea: Reverse the screening process !**

New functions for RNA ← Screen protein complexes (known functions)

## 2. RBPs mainly identified by RNA pulldowns

**Need for orthogonal methods to identify RBPs!**

**Idea: identify proteins and complexes affected by RNA**

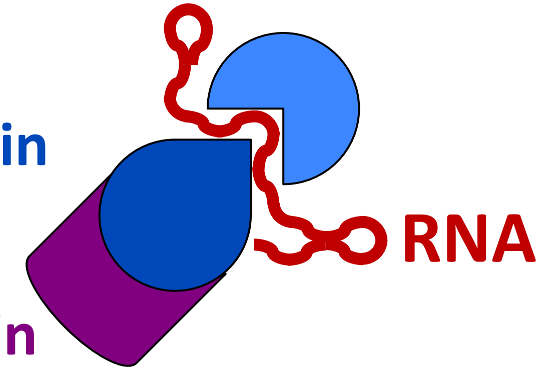
## The concept of “RNA dependence”

# RNA-Dependent Proteins (R-DeeP)

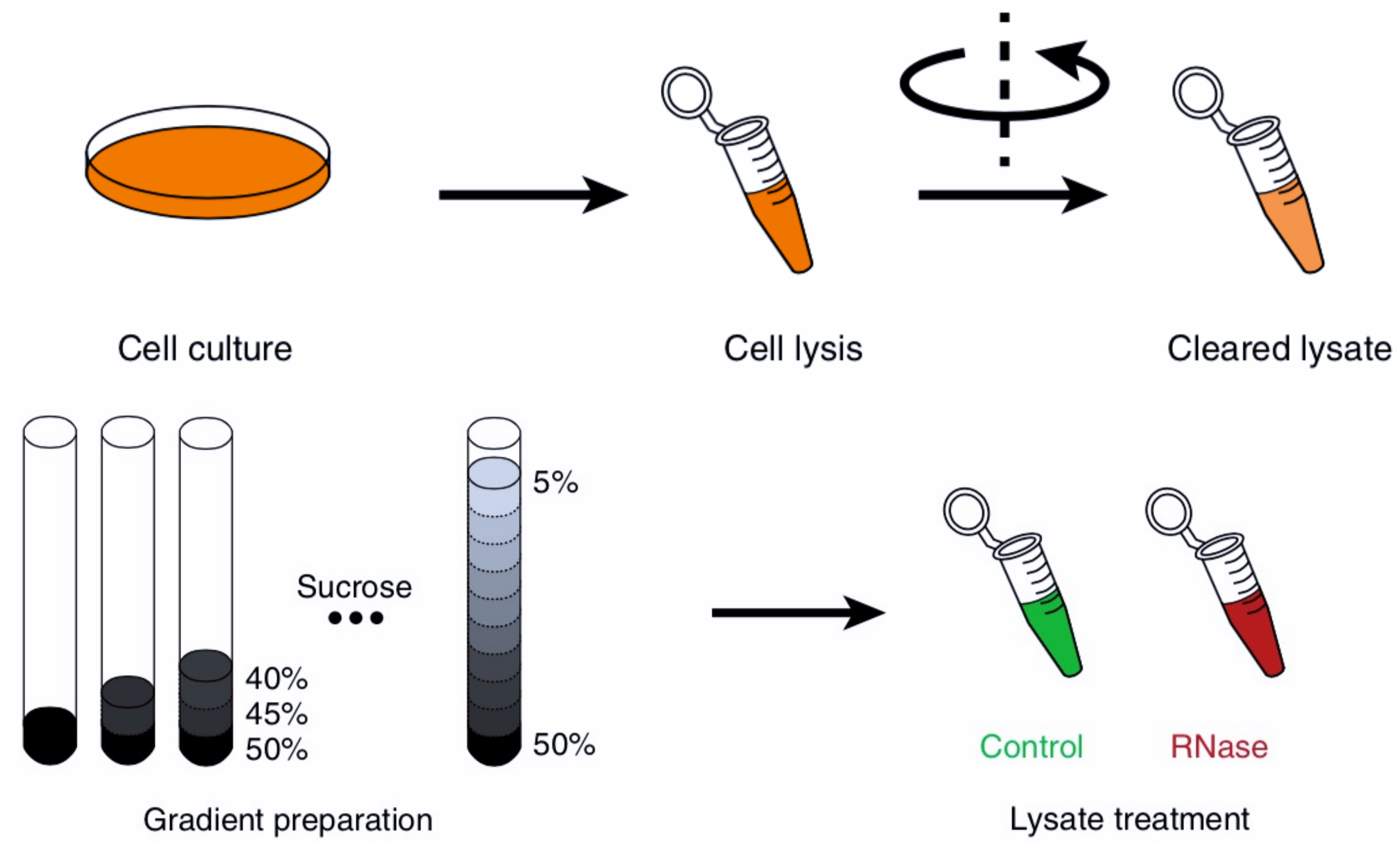
**RNA dependence:**
**Proteins and protein complexes,**
**whose molecular interactions depend on RNA**

**RNA-dependent Proteins**

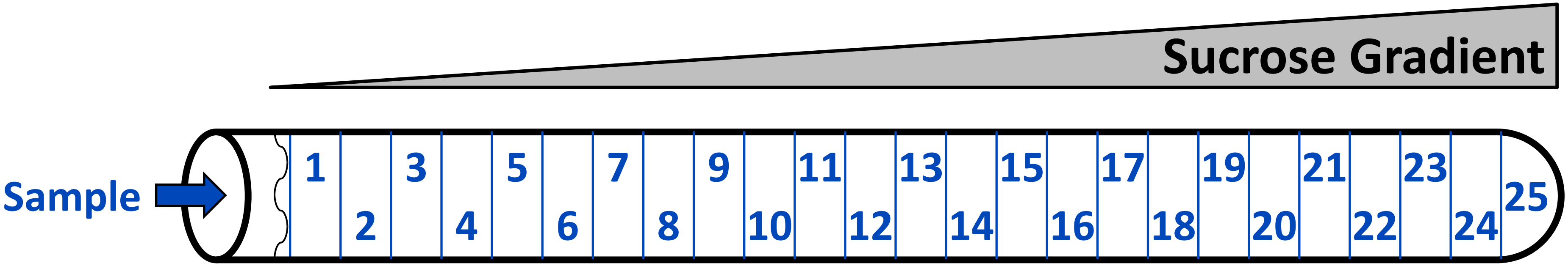
- **RNA-binding protein** (RBP)
- **RBP-binding protein** (if RBP bound to RNA)



# R-DeeP Screen: Density Gradient Fractionation

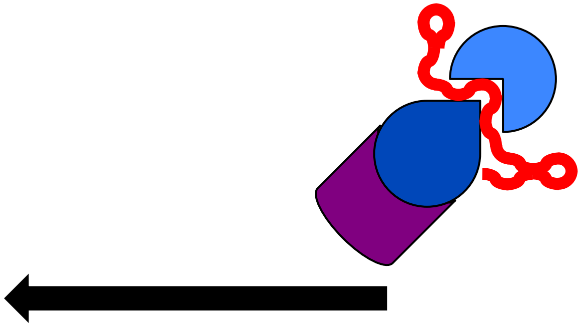


# R-DeeP Screen: Density Gradient Fractionation



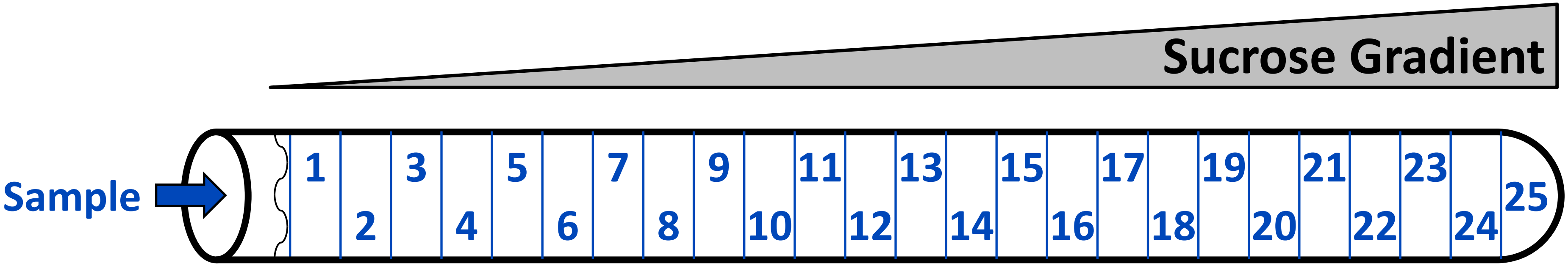
**Control**

**RNase**



**RNA-dependent shifts**

# R-DeeP Proteome-wide Screen



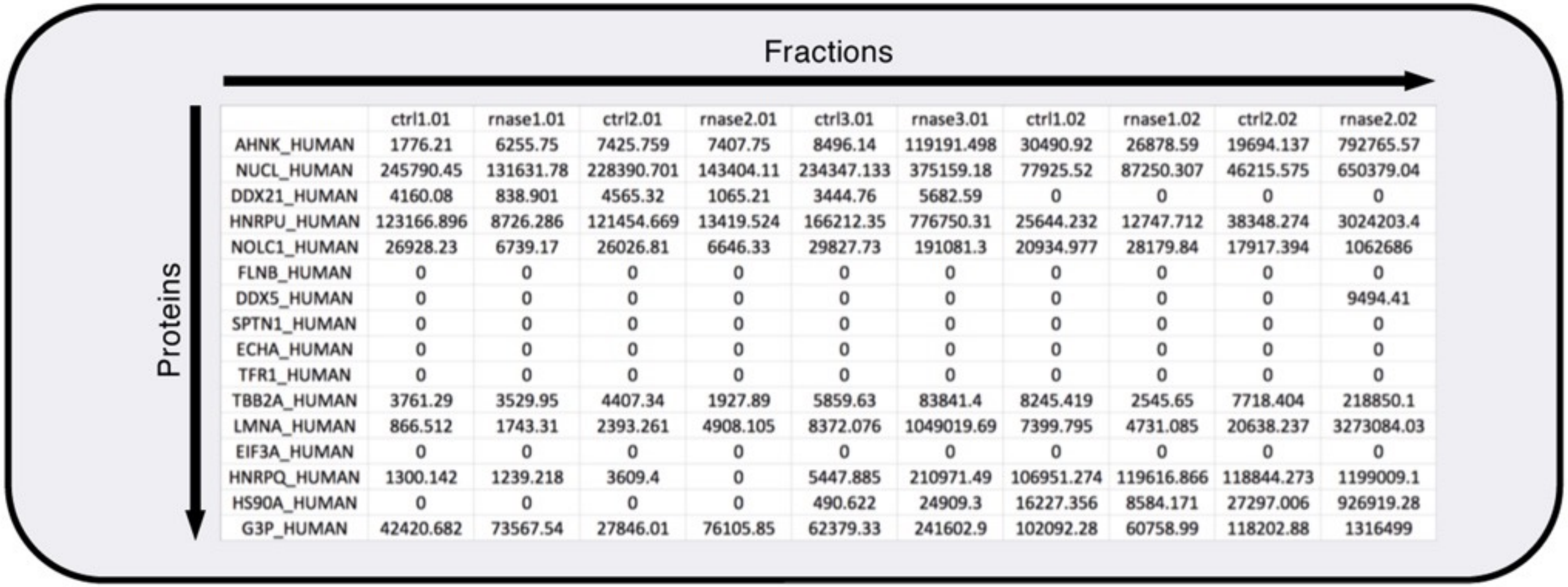
**Control**

**RNase**

**Triplicates (150 fractions)** → **Mass Spectrometry** → **5000 Proteins quantified**

# Analysis of the Mass Spectrometry Data

## Dataset

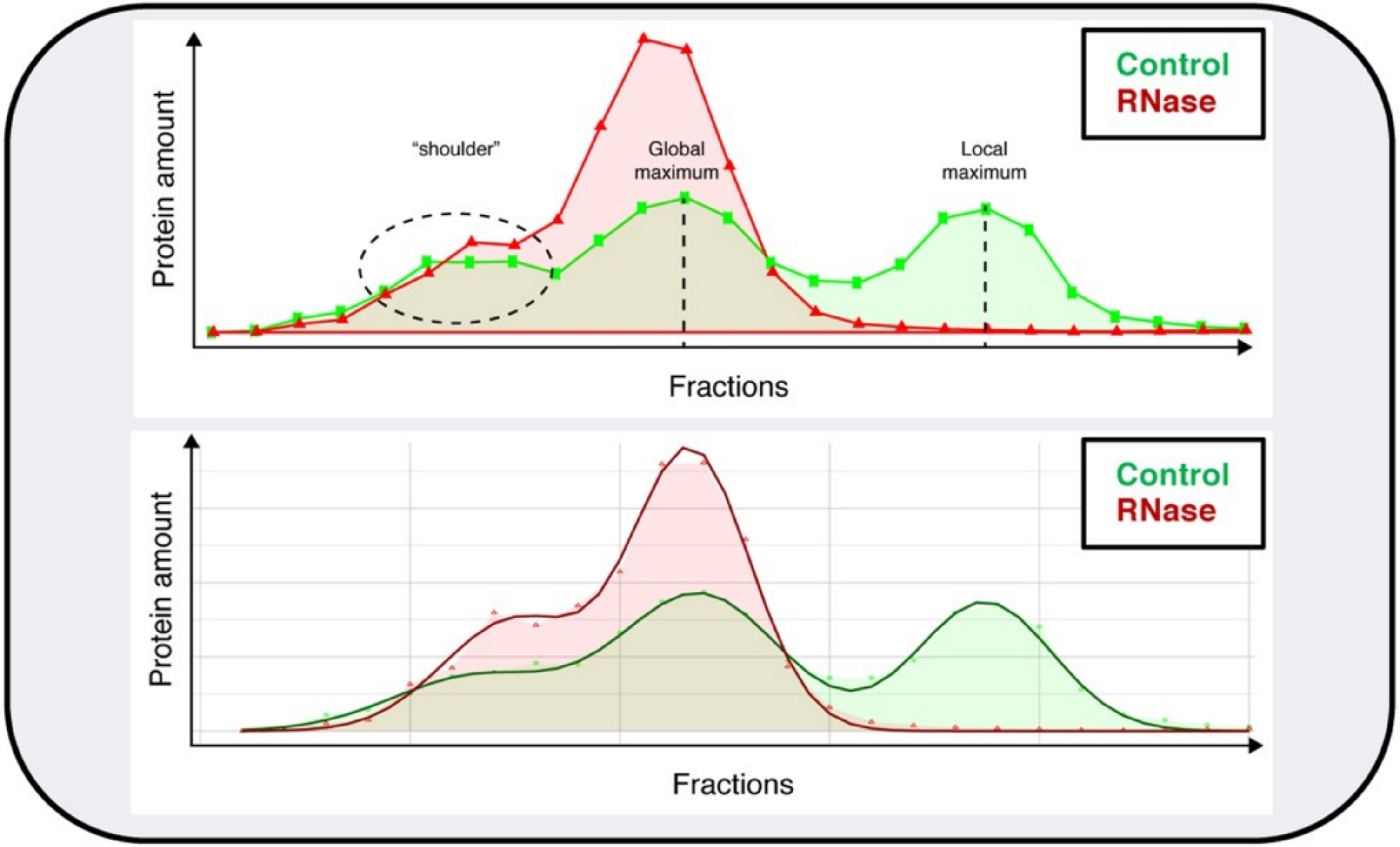


| | ctrl1.01 | rnase1.01 | ctrl2.01 | rnase2.01 | ctrl3.01 | rnase3.01 | ctrl1.02 | rnase1.02 | ctrl2.02 | rnase2.02 |
|---|---|---|---|---|---|---|---|---|---|---|
| AHNK_HUMAN | 1776.21 | 6255.75 | 7425.759 | 7407.75 | 8496.14 | 119191.498 | 30490.92 | 26878.59 | 19694.137 | 792765.57 |
| NUCL_HUMAN | 245790.45 | 131631.78 | 228390.701 | 143404.11 | 234347.133 | 375159.18 | 77925.52 | 87250.307 | 46215.575 | 650379.04 |
| DDX21_HUMAN | 4160.08 | 838.901 | 4565.32 | 1065.21 | 3444.76 | 5682.59 | 0 | 0 | 0 | 0 |
| HNRPU_HUMAN | 123166.896 | 8726.286 | 121454.669 | 13419.524 | 166212.35 | 776750.31 | 25644.232 | 12747.712 | 38348.274 | 3024203.4 |
| NOLC1_HUMAN | 26928.23 | 6739.17 | 26026.81 | 6646.33 | 29827.73 | 191081.3 | 20934.977 | 28179.84 | 17917.394 | 1062686 |
| FLNB_HUMAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| DDX5_HUMAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9494.41 |
| SPTN1_HUMAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ECHA_HUMAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TFR1_HUMAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TBB2A_HUMAN | 3761.29 | 3529.95 | 4407.34 | 1927.89 | 5859.63 | 83841.4 | 8245.419 | 2545.65 | 7718.404 | 218850.1 |
| LMNA_HUMAN | 866.512 | 1743.31 | 2393.261 | 4908.105 | 8372.076 | 1049019.69 | 7399.795 | 4731.085 | 20638.237 | 3273084.03 |
| EIF3A_HUMAN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HNRPQ_HUMAN | 1300.142 | 1239.218 | 3609.4 | 0 | 5447.885 | 210971.49 | 106951.274 | 119616.866 | 118844.273 | 1199009.1 |
| HS90A_HUMAN | 0 | 0 | 0 | 0 | 490.622 | 24909.3 | 16227.356 | 8584.171 | 27297.006 | 926919.28 |
| G3P_HUMAN | 42420.682 | 73567.54 | 27846.01 | 76105.85 | 62379.33 | 241602.9 | 102092.28 | 60758.99 | 118202.88 | 1316499 |

## Part 1:

- **Description of the dataset**
- **Reproducibility**
- **Normalization**

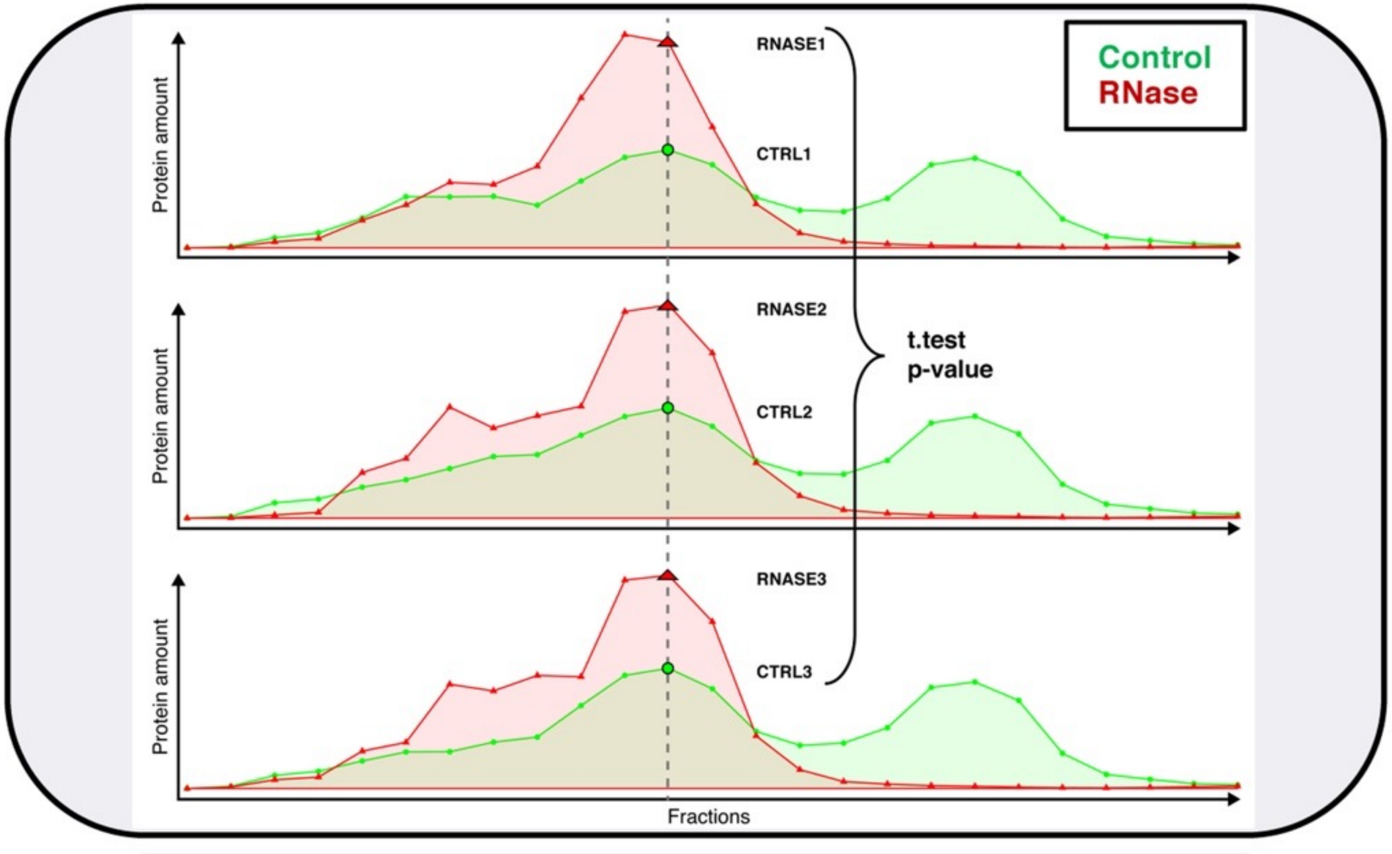
# Analysis of the Mass Spectrometry Data



**Part 2:**

- **Find the maxima (and fit the curves with Gaussians)**

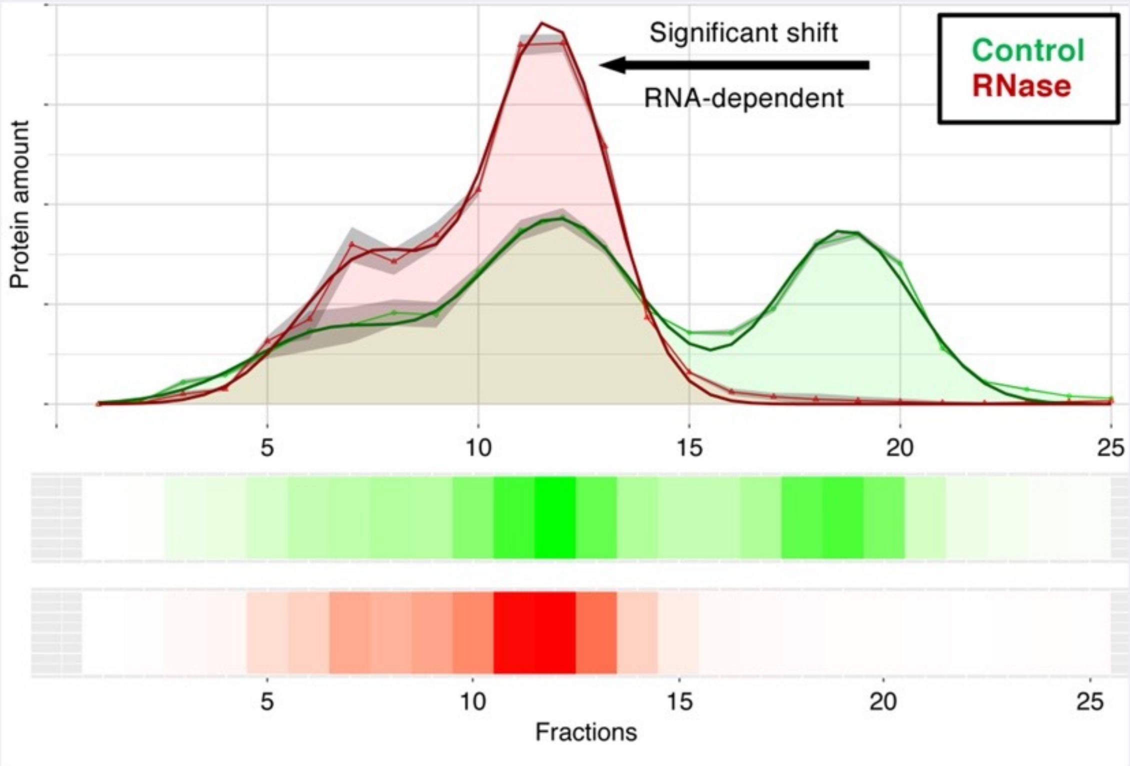
# Analysis of the Mass Spectrometry Data



## Part 3:

- **Evaluate the differences at the maxima**
- **Define selection criteria for RNA-dependent proteins**

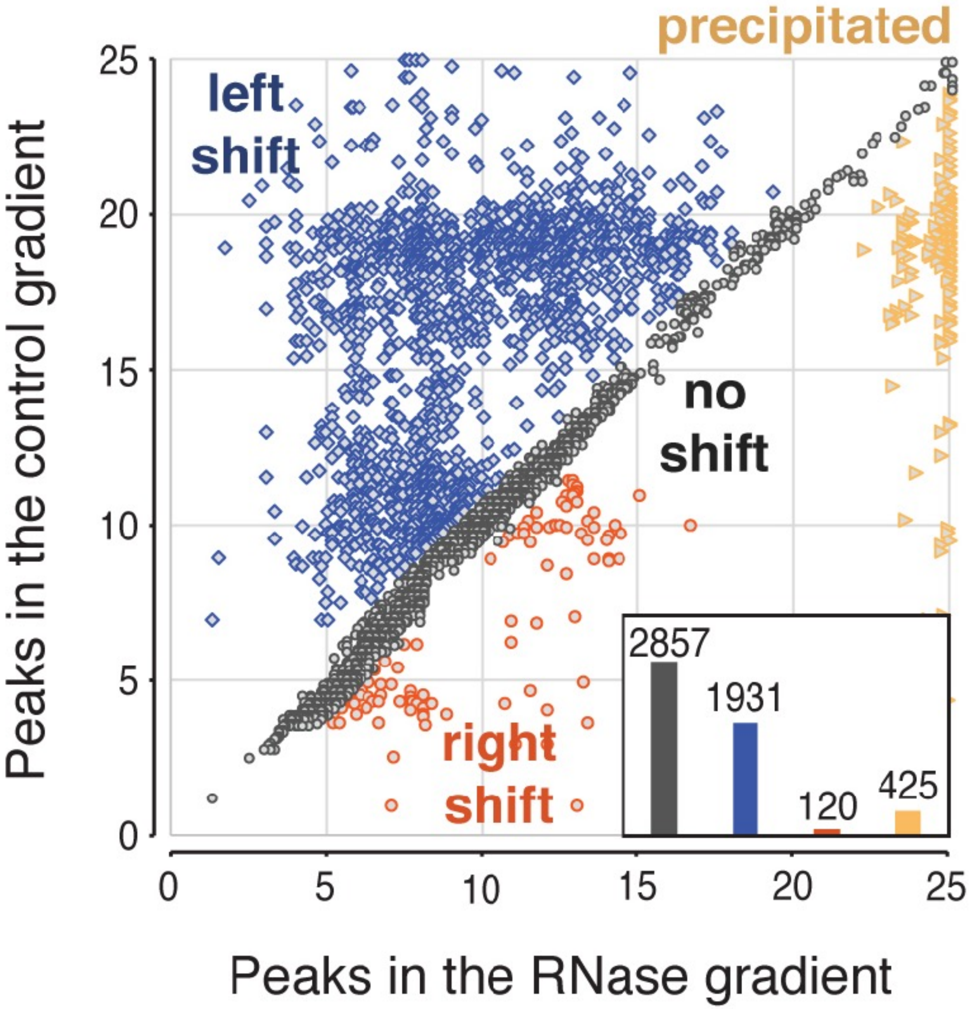
# Analysis of the Mass Spectrometry Data



## Part 4:

- Apply selection criteria
- Produce some graphics for specific proteins or groups
- Perform a linear regression

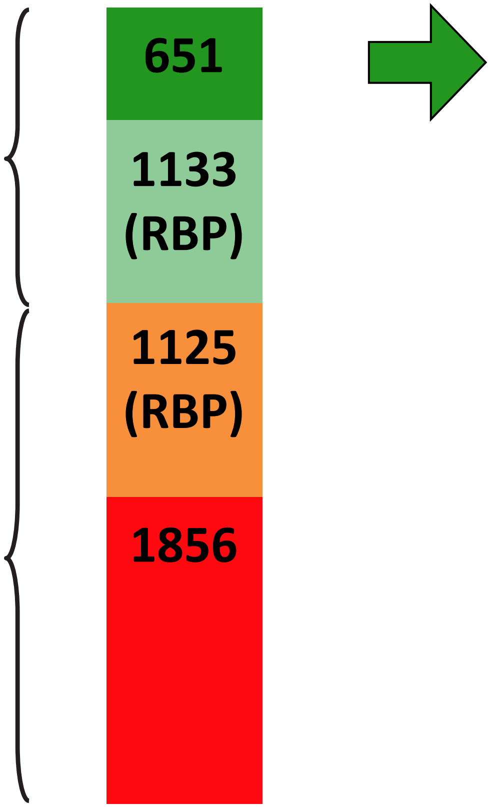
# Overview of different shift categories



# Part 5: further analysis of the results

**1784**
**RNA-dependent proteins**

**2981**
**RNA-independent proteins**



Discovery of
new RNA functions
(Gene Ontology)

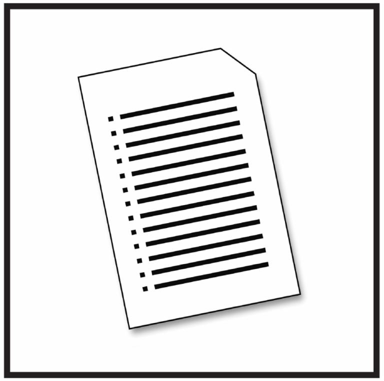
# R-DeeP: Database for RNA-dependent Proteins

**R-DeeP**

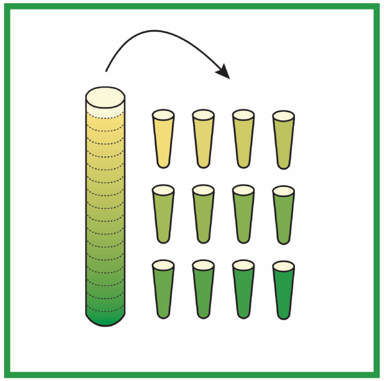
**Database for RNA-dependent Proteins**

**http://R-DeeP.dkfz.de**

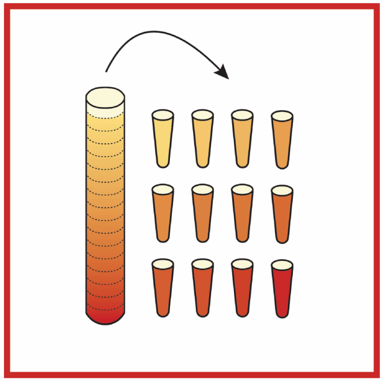
**Search by protein**



**Search by list**



**Search by fraction Control**



**Search by fraction RNase**

# R-DeeP:  Database for RNA-dependent Proteins

## Analysis at the protein level

**Peaks and shift information**

**Download options**

**Graphical view**

Position of HNRPC_HUMAN in the sucrose density gradients.

| | Fraction of the Maximum | Amount of Protein (%) |
|---|---|---|
| **Control Peaks** | 19.2 | 96.8 |
| **RNase Peaks** | 9.9 | 93.7 |

| Significant Shifts | Control Peak Max | Amplitude Loss (%) | RNase Peak Max | Amplitude Gain (%) | Shift Distance |
|---|---|---|---|---|---|
| | 19.2 | 100 | 9.9 | 99.6 | -9.3 |

Heterogeneous nuclear ribonucleoproteins C1/C2

| | |
|---|---|
| Entry Name: | HNRPC_HUMAN |
| Gene Name: | HNRNPC HNRPC |
| Protein Accession Number: | P07910 |
| Molecular Weight: | 34 KDa |
| Length: | 306 AA |

Links to protein databases:

Link to UniProt
Link to String database

This protein is listed as RNA-binding protein in published RBP studies. For more details, click on the link below.

To look for potential interaction partners shifting from or to the same fraction you can use the following search modus:

Protein complexes with HNRPC_HUMAN

There is some information about this protein in the CORUM database:

Link to CORUM database

**Protein / Gene details**

**Link to RBP resources**

**List of proteins found in the same fractions**

**Link to protein complexes**

# RBP2GO: a new pan-species RBP database

RBP2GO

RBP2GO.dkfz.de

**13**
species

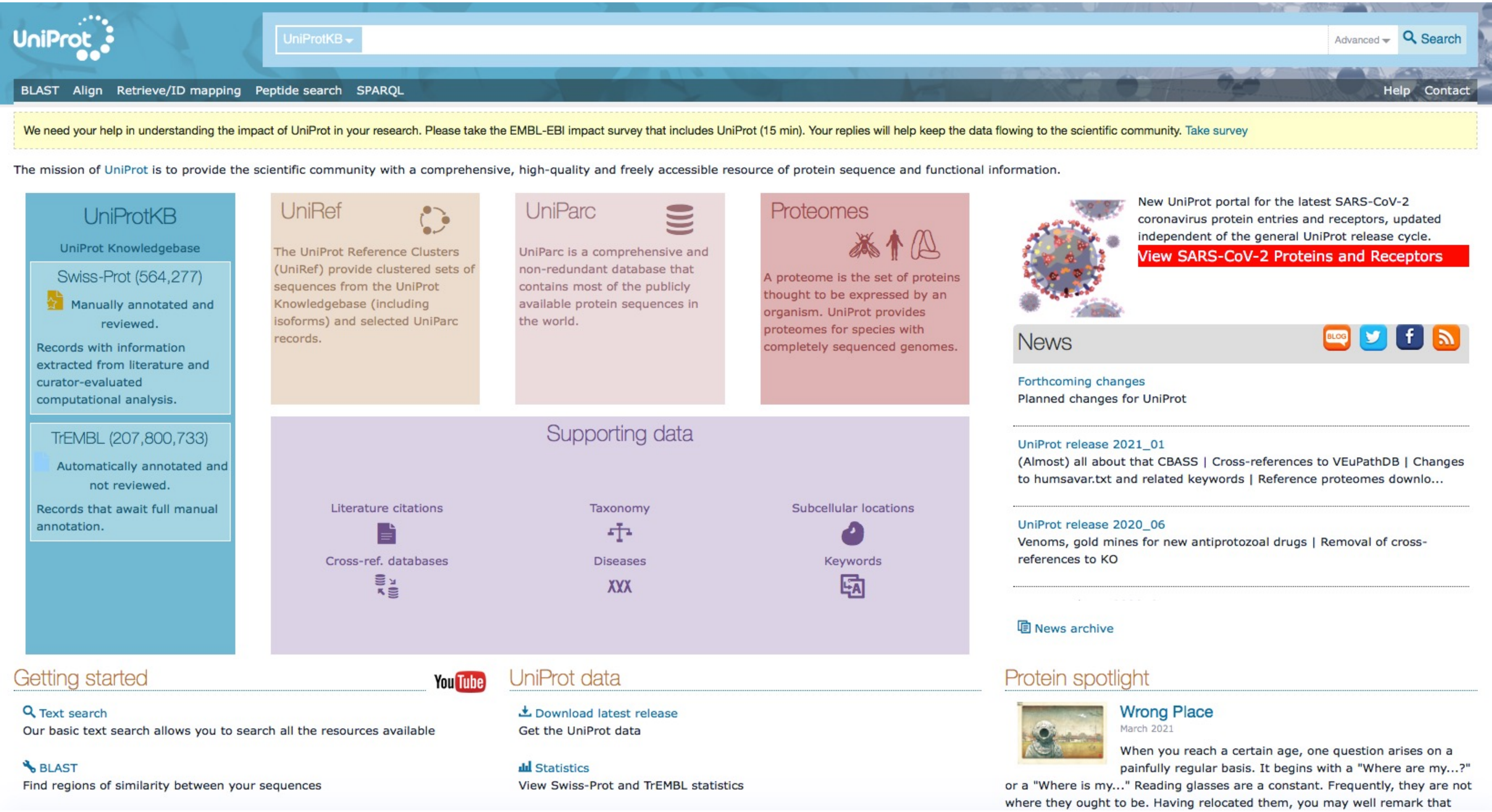
**105**
RBP datasets

**22552**
RBP candidates

## Gene Ontology

biological process
molecular function
cellular component

Properties
Cancer
Interactions
Domains
Homologs



We need your help in understanding the impact of UniProt in your research. Please take the EMBL-EBI impact survey that includes UniProt (15 min). Your replies will help keep the data flowing to the scientific community. Take survey

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

**UniProtKB**

UniProt Knowledgebase

**Swiss-Prot (564,277)**

Manually annotated and reviewed.

Records with information extracted from literature and curator-evaluated computational analysis.

**TrEMBL (207,800,733)**

Automatically annotated and not reviewed.

Records that await full manual annotation.

**UniRef**

The UniProt Reference Clusters (UniRef) provide clustered sets of sequences from the UniProt Knowledgebase (including isoforms) and selected UniParc records.

**UniParc**

UniParc is a comprehensive and non-redundant database that contains most of the publicly available protein sequences in the world.

**Proteomes**

A proteome is the set of proteins thought to be expressed by an organism. UniProt provides proteomes for species with completely sequenced genomes.

**Supporting data**

New UniProt portal for the latest SARS-CoV-2 coronavirus protein entries and receptors, updated independent of the general UniProt release cycle.

View SARS-CoV-2 Proteins and Receptors

**News**

Forthcoming changes
Planned changes for UniProt

UniProt release 2021_01
(Almost) all about that CBASS | Cross-references to VEuPathDB | Changes to humsavar.txt and related keywords | Reference proteomes downlo...

UniProt release 2020_06
Venoms, gold mines for new antiprotozoal drugs | Removal of cross-references to KO

News archive

**Getting started**

Text search
Our basic text search allows you to search all the resources available

BLAST
Find regions of similarity between your sequences

**UniProt data**

Download latest release
Get the UniProt data

Statistics
View Swiss-Prot and TrEMBL statistics

**Protein spotlight**

Wrong Place
March 2021

When you reach a certain age, one question arises on a painfully regular basis. It begins with a "Where are my...?" or a "Where is my..." Reading glasses are a constant. Frequently, they are not where they ought to be. Having relocated them, you may well remark that