Electronic Supplementary Material (ESI) for Molecular Omics.

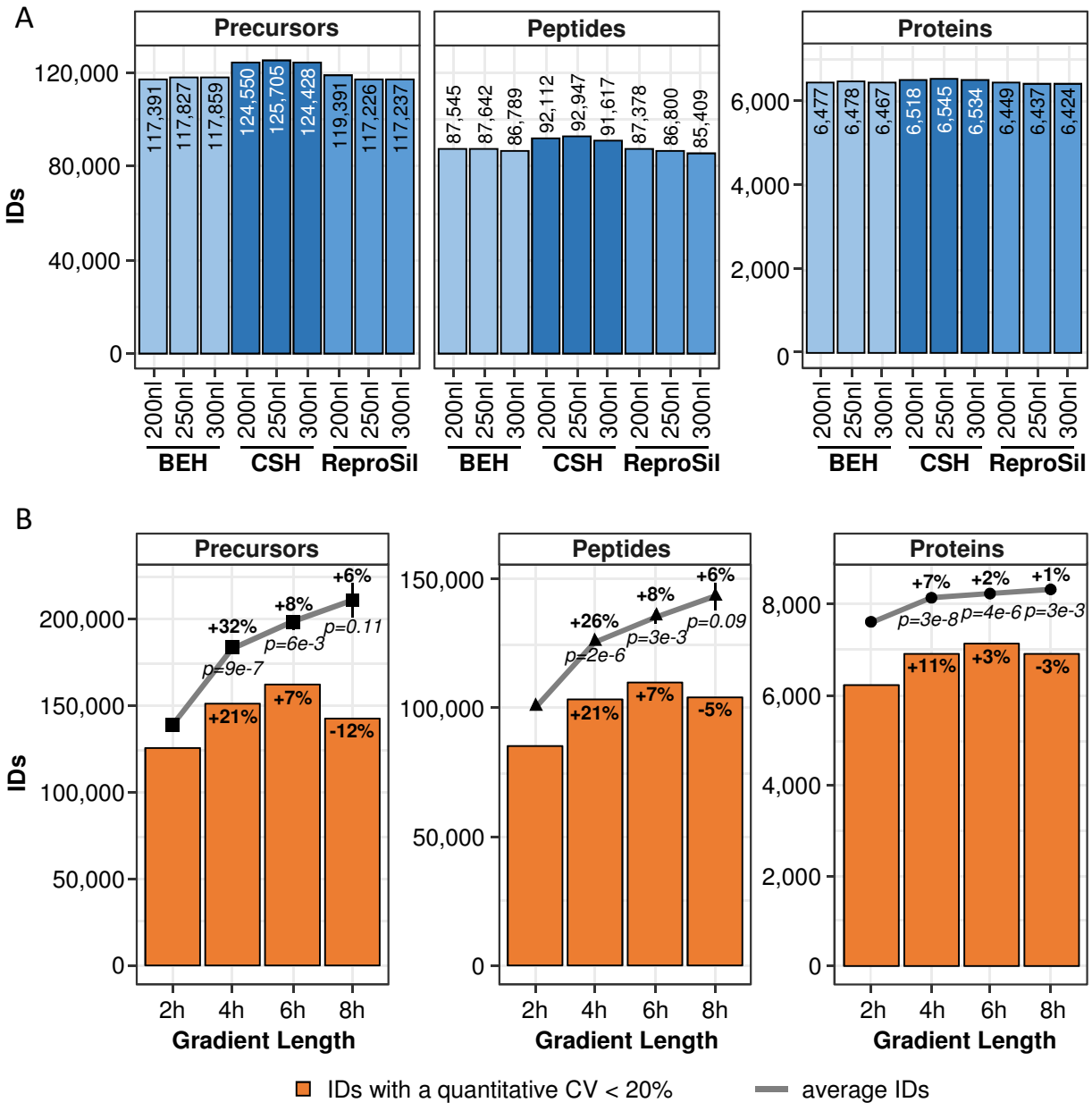
Supplementary Figures

Surpassing 10,000 identified and quantified proteins in a single run by optimizing current LC-MS instrumentation and data analysis strategy

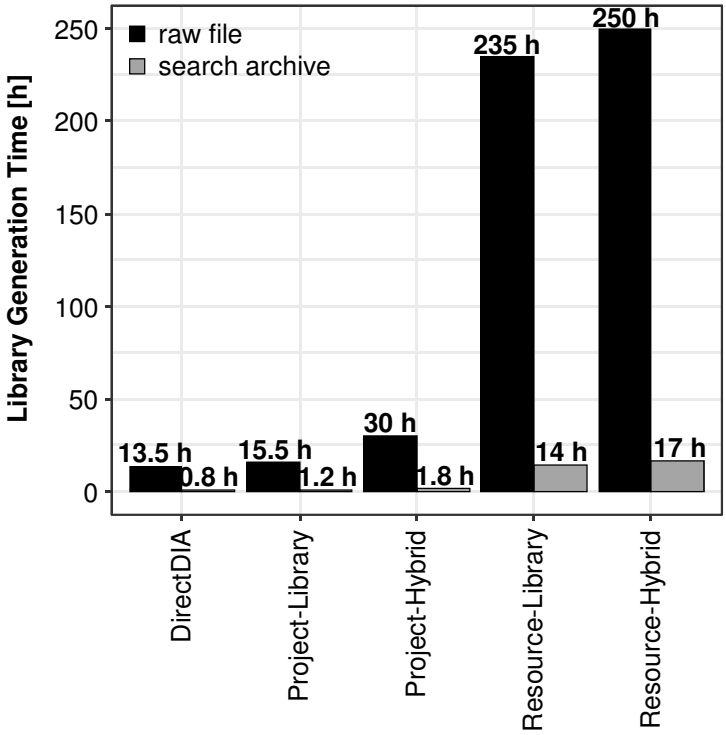
Jan Muntel, Tejas Gandhi, Lynn Verbeke, Oliver M. Bernhardt, Tobias Treiber, Roland Bruderer and Lukas Reiter‡

‡contact: lukas.reiter@biognosys.com

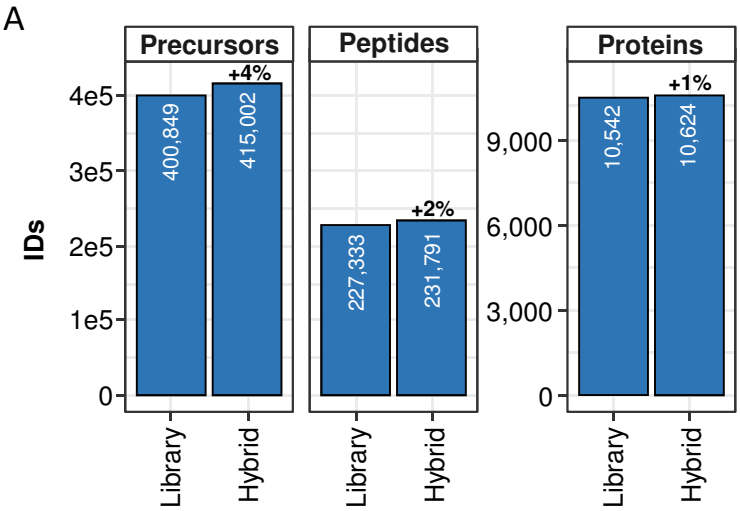
Supplementary Figure 1: Optimization of Chromatography. (A) Comparison of 3 solid phases using 2h gradients and a HeLa sample. The DIA method was optimised to the chromatographic peak width. DIA data were analysed using a project-specific HeLa library. (B) Comparison of gradient lengths using a HeLa sample. The DIA method was optimised to the chromatographic peak width. DIA data were analysed using a project-specific HeLa library. Three run replicates were acquired to calculate average identifications per run including standard deviation and identification with a CV < 20%.

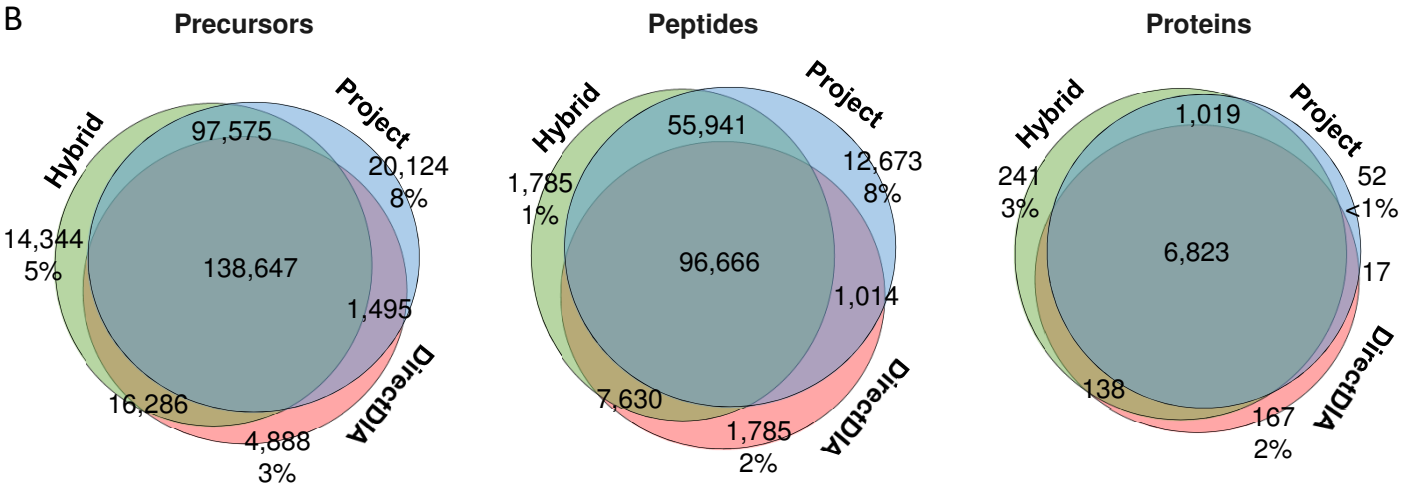


Supplementary Figure 2: Library Generation Times using Search Archives. The DirectDIA analysis comprised 6 runs, the project library 40 runs and the resource library 2,206 runs. The hybrid libraries were combination of the DirectDIA analysis with the respective library. The processing times can vary based on computer configuration and other processes running at the same time.

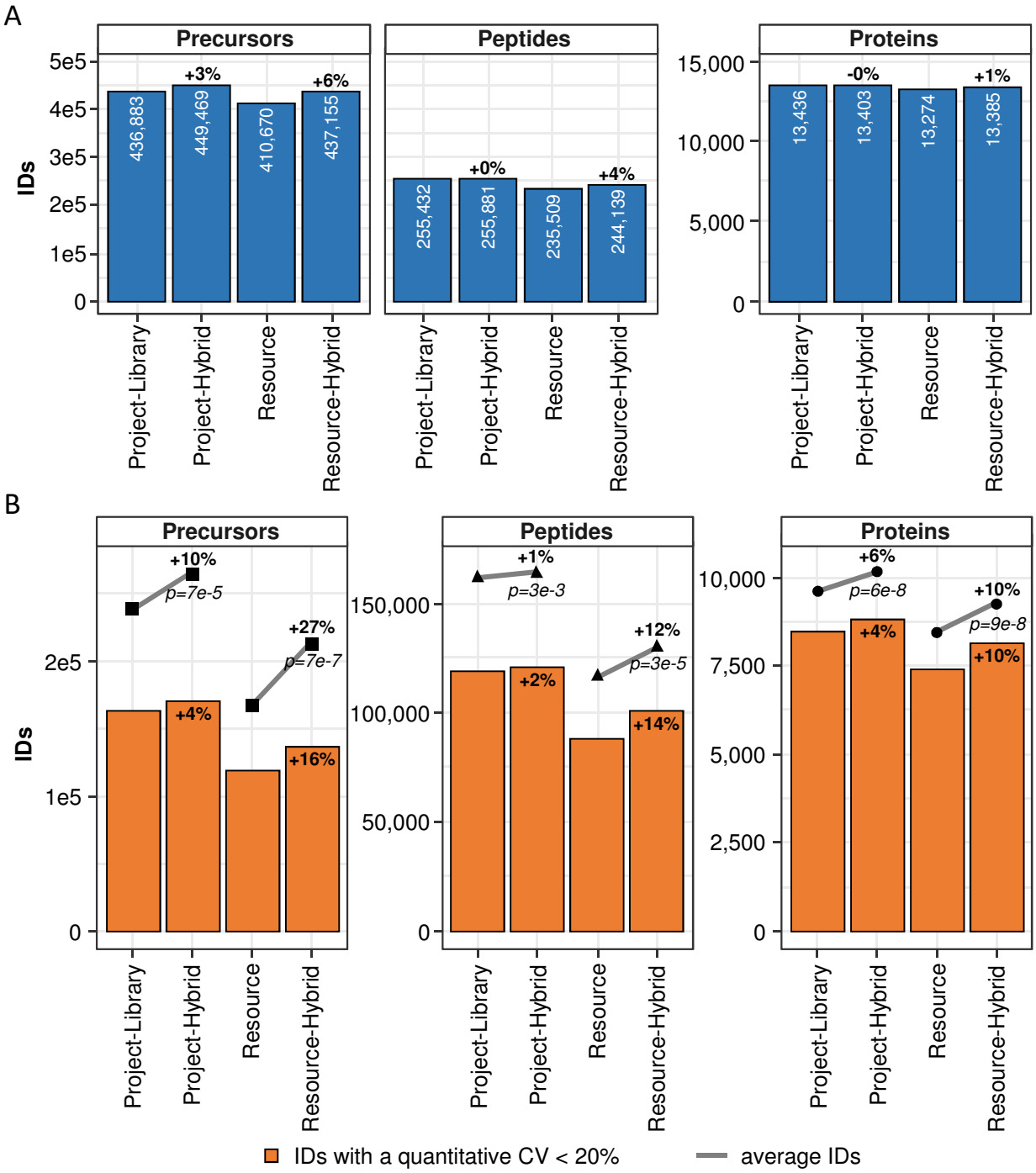


Supplementary Figure 3: Overview of Hela Libraries. (A) Library size of the project-specific and the hybrid HeLa library. DDA and DIA data were searched using Pulsar (SpectroMine). FDR on precursor, peptide and protein level was set to 1%. Libraries were generated in SpectroMine using search archives. (B) Overlap in quantified precursors, peptides and proteins from DirectDIA, project-library and project-hybrid library-based analysis of a 6h HeLa triplicate run. For protein groups only the first entry was considered.

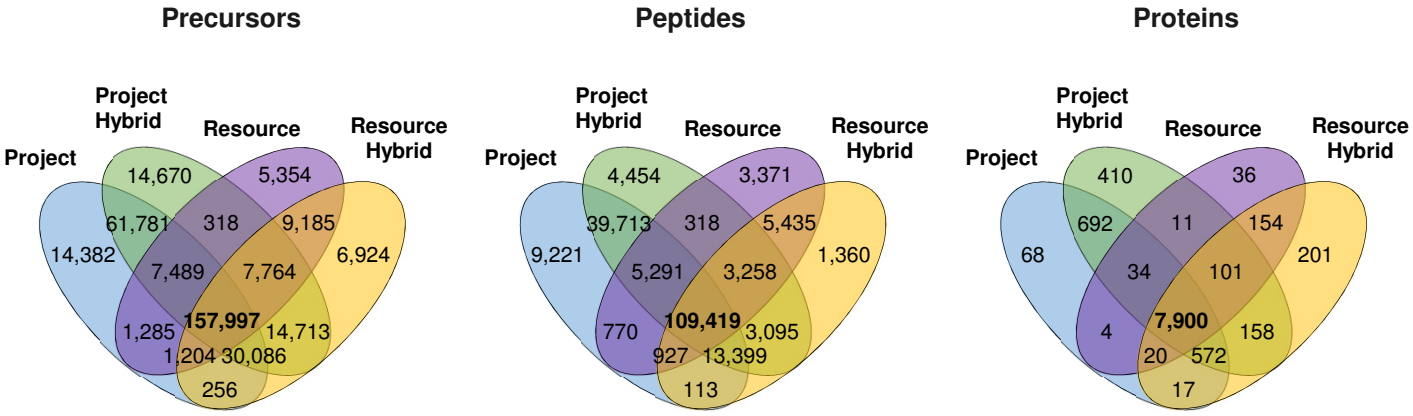




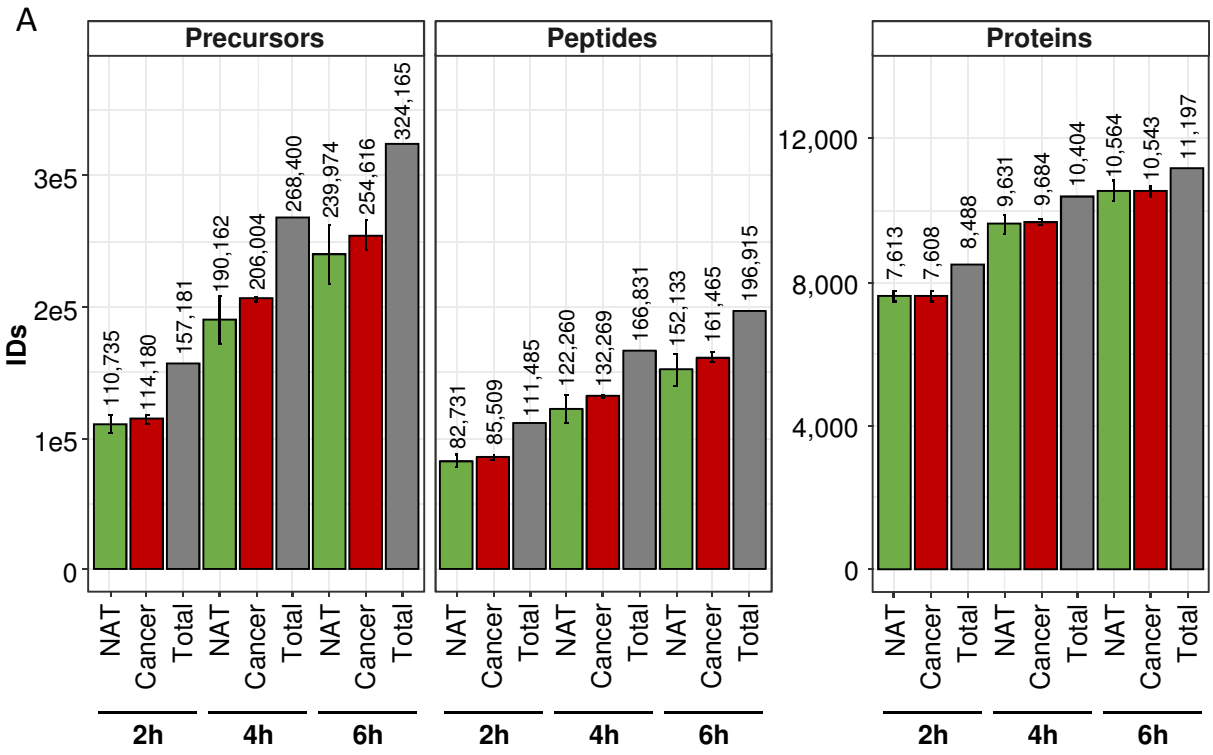
Supplementary Figure 4: Application of Hybrid Approach to Testis Tissue. (A) Library overview. DDA and DIA data were searched using Pulsar (SpectroMine). FDR on precursor, peptide and protein level was set to 1%. Libraries were generated in SpectroMine using search archives. (B) Application of libraries to injection triplicate of one testis tissue sample. DIA data were analysed in Spectronaut and precursor and protein FDR was set to 1%. Two-sample t-test was applied to test for statistical significance of identification differences between the DIA data analysis using the different libraries. Average run identifications and CVs < 20% were calculated based on the triplicate injection of the sample. (C) Overlap in quantified precursors, peptides and proteins from the analysis using the four different libraries (6h Testis triplicate). For protein groups only the first entry was considered.

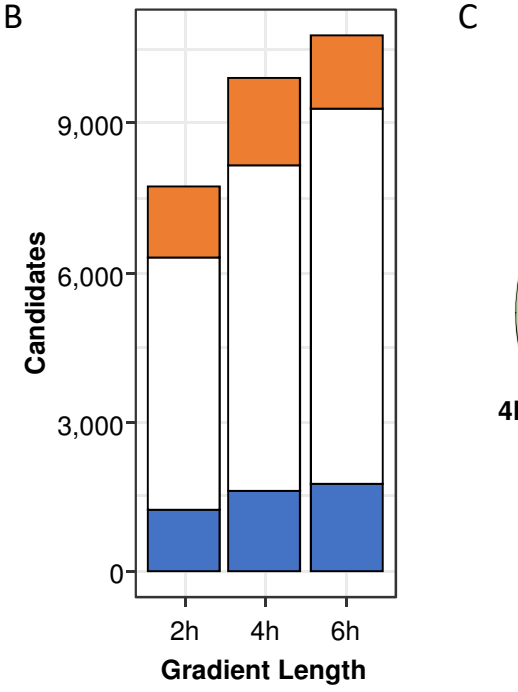


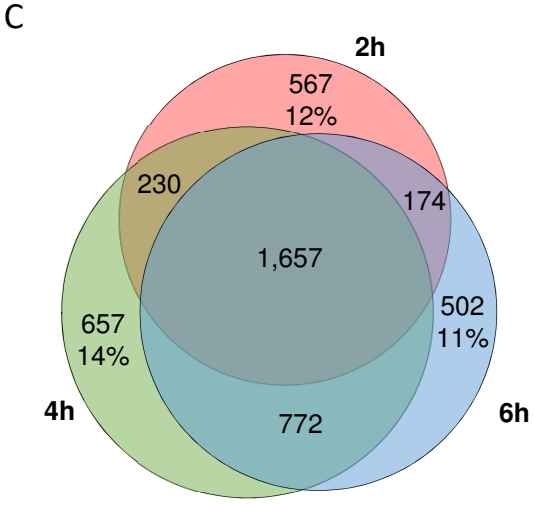
C


Precursors
Project
Project Hybrid
Resource
Resource Hybrid
14,670
5,354
61,781
318
9,185
14,382
7,489
7,764
6,924
1,285
157,997
14,713
1,204
30,086
256
Peptides
Project
Project Hybrid
Resource
Resource Hybrid
4,454
3,371
39,713
318
5,435
9,221
5,291
3,258
1,360
770
109,419
3,095
927
13,399
113
Proteins
Project
Project Hybrid
Resource
Resource Hybrid
410
36
692
11
154
68
34
101
201
4
7,900
158
20
572
17

Supplementary Figure 5: Analysis of Testis Cancer Set with 2h and 4h DIA Method. The hybrid libraries for DIA data analysis were generated by combining the testis project-specific library with the search results of the respective 2h, 4h or 6h DIA runs. (A) Overview of identifications. Average identifications and standard deviations are depicted for the NAT (green) and cancer (red) cohorts based on three biological replicates. (B) Overview of statistical analysis based on the Spectronaut pipeline. Proteins were considered significantly differentially abundant with an absolute fold-change > 1 and a Q value <0.01. (C) Overlap of significantly changed proteins. For protein groups only the first entry was considered.

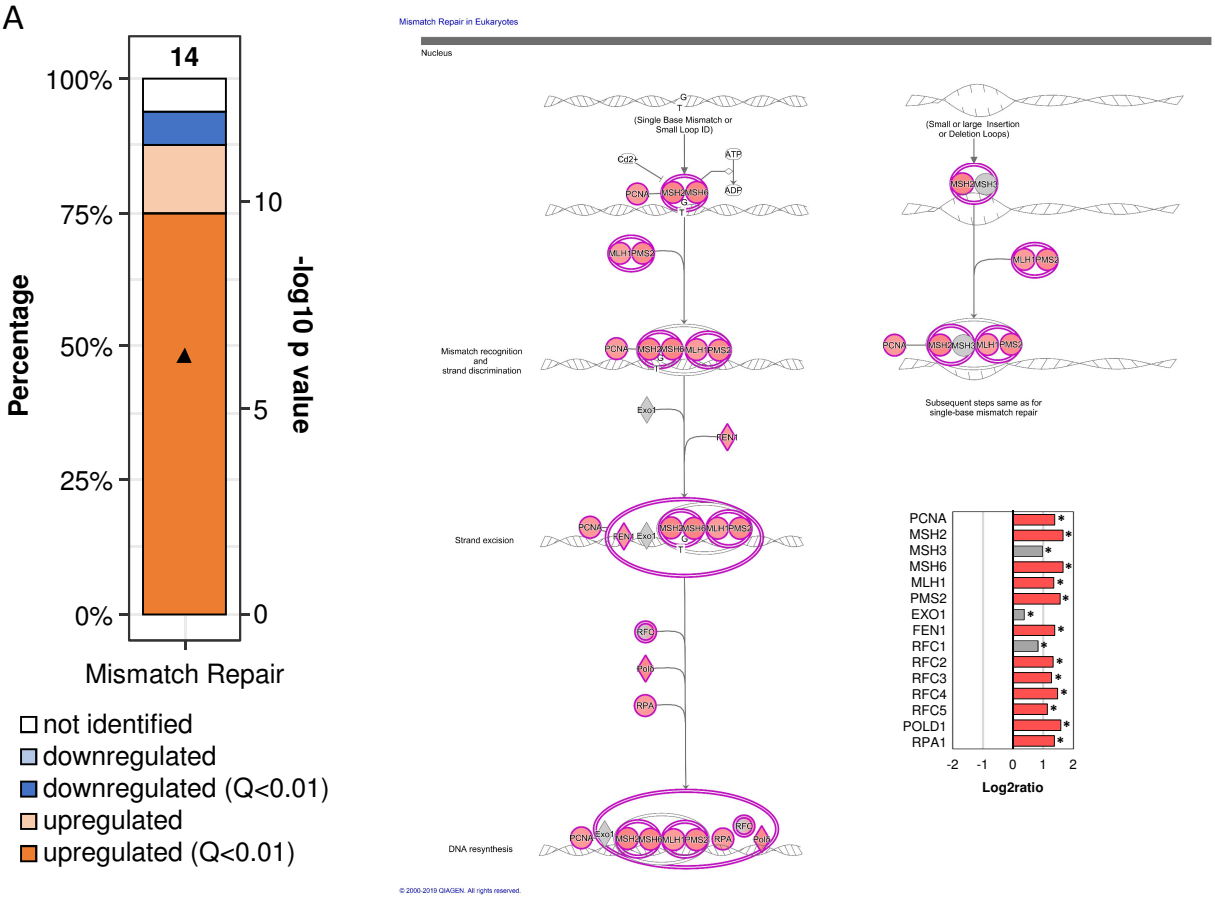






Supplementary Figure 6: Additional Analysis. (A) IPA (Qiagen) analysis of mismatch repair pathway. Pathway was exported from IPA and quantitative data of the proteins from this pathway are depicted as bar chart. Proteins quantified with a Q value < 0.01 are marked with a star. (B) Abundance changes between cancer and NAT cohort of the DNA-directed RNA polymerase proteins. Stars indicate proteins quantified with a Q value < 0.01.

A



B