Supporting Information

Supplementary Figures

A simple method to quantify protein abundances from one thousand cells

Burcu Vitrinel[1,2], Dylan E. Iannitelli[2], Esteban O. Mazzoni[2,3], Lionel Christiaen*[2], and Christine Vogel*[1$]

1 Center for Genomics and Systems Biology, Department of Biology, New York University, New York, NY, USA

2 Center for Developmental Genetics, Department of Biology, New York University, New York, NY, USA

3 NYU Neuroscience Institute,NYU Langone Medical Center, New York, NY, USA

*equally contributing authors

[$]corresponding author: cvogel@nyu.edu, +1212-998-3976

Figure S1: We optimized the buffer composition and cell lysis method; the graph shows the numbers of Protein groups (a.) and Peptides (b.) detected by label-free mass spectrometry using 5,000 embryonic stem cells. The error bars show standard variation.

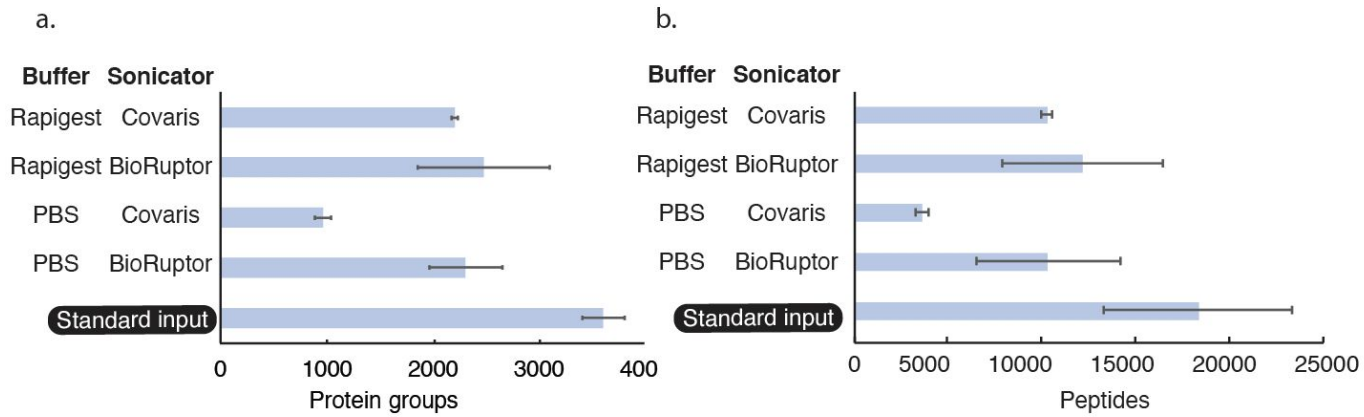


Figure S2. Different sample preparation techniques correlate between each other and are highly reproducible (Spearman correlation coefficient). ESC - embryonic stem cells, MN - motor neurons

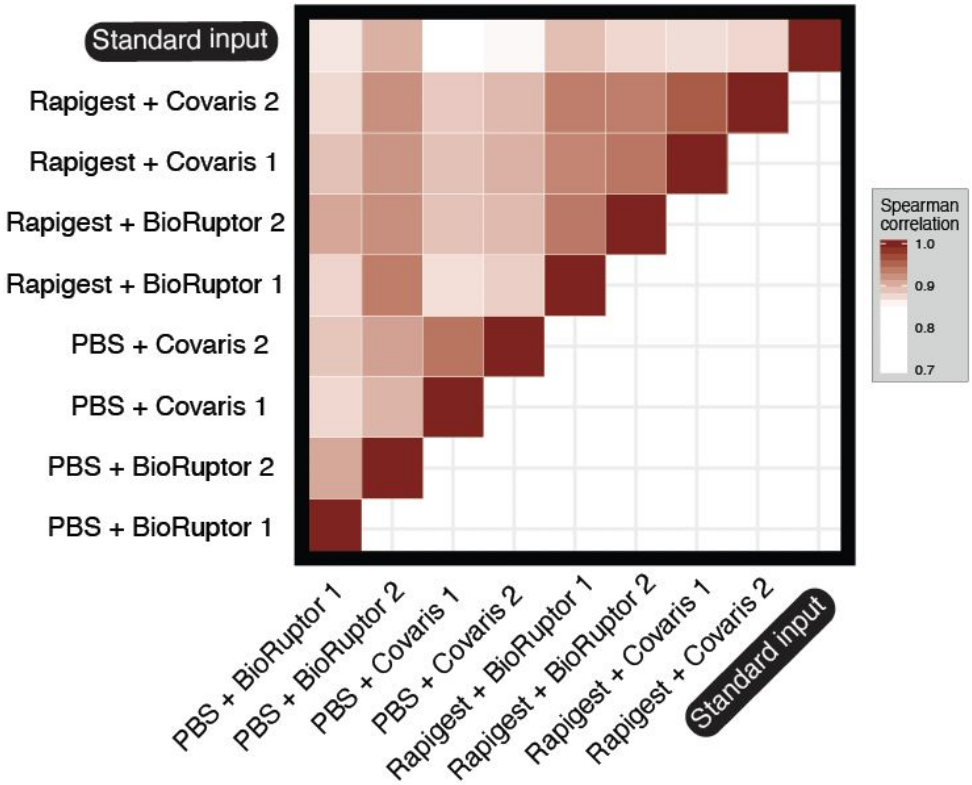


Figure S3. We compared the number of protein groups identified using 1,000 and 5,000 embryonic stem cells using label-free quantification and observed no significant differences. The data were acquired on the same run sequence and using the same conditions. They both have similar numbers of unique peptides and similar numbers of protein groups identified from unique peptides (5,000 and 1,000 cells: 900 and 777 vs 863 and 874, respectively). Increasing column loading is a common method to increase protein identification. In our case, using 5 times the amount of sample did not increase identifications. We speculate based on the fact that both 1,000 and 5,000 cell samples are very small, the number of protein identifications may have plateaued.

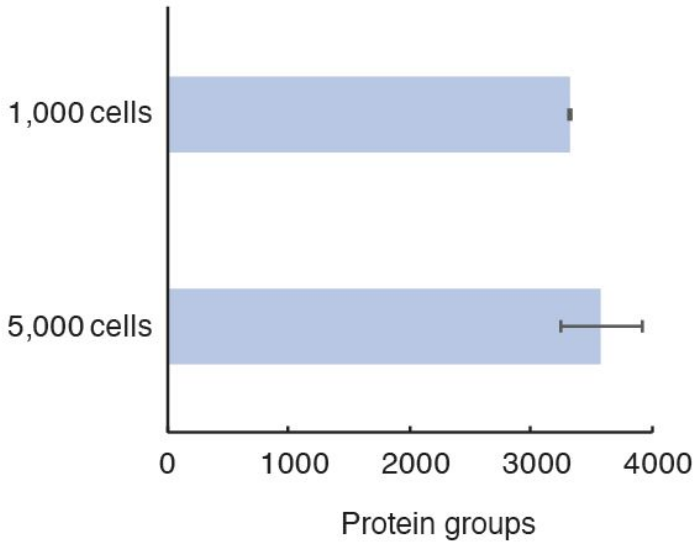


Figure S4. We optimized the size of the carrier channel, and to do so we used samples from 60,000, 20,000, and 10,000 cells as carrier and 1,000 in the sample channels. The graphs show that the experiment with the 10,000 cell carrier channel has the most consistent quantification (Spearman correlation) at similar numbers of total protein identification (see main text). ESC - embryonic stem cells, MN - motor neurons

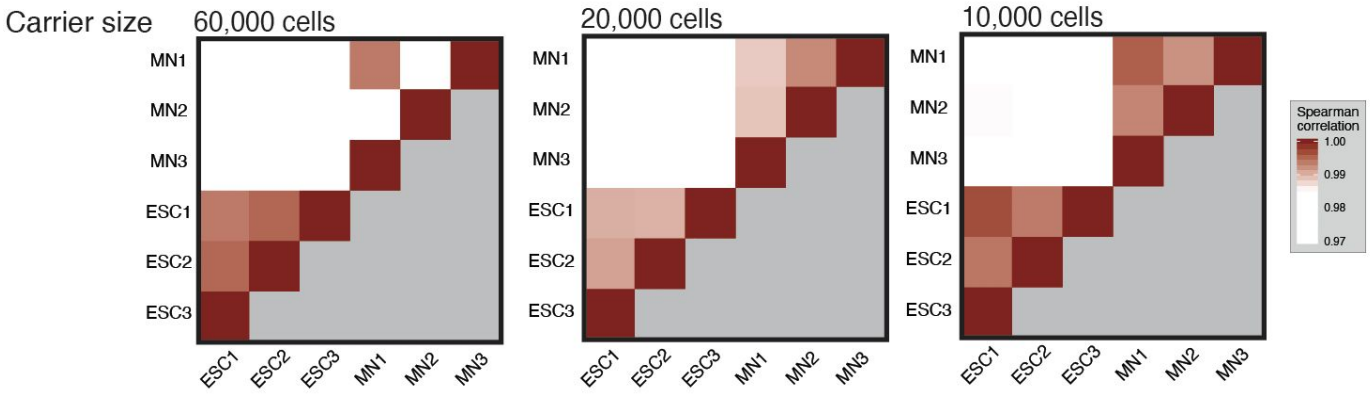


Figure S5. The replicates of minimal input MN (left) and ESC (right) show high Spearman correlation coefficients of 0.99 and 0.98, respectively. ESC - embryonic stem cells, MN - motor neurons

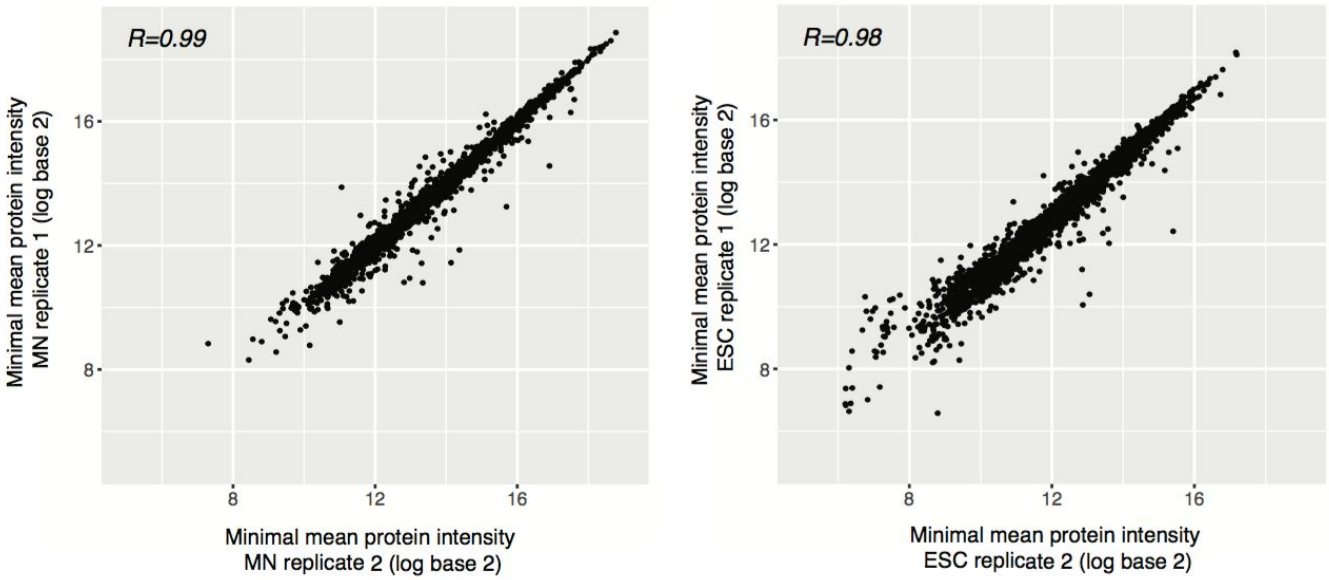


Figure S6. Correlations coefficients (Spearman) of quadruplicates resulting from minimal and standard input embryonic stem cells (ESCs).

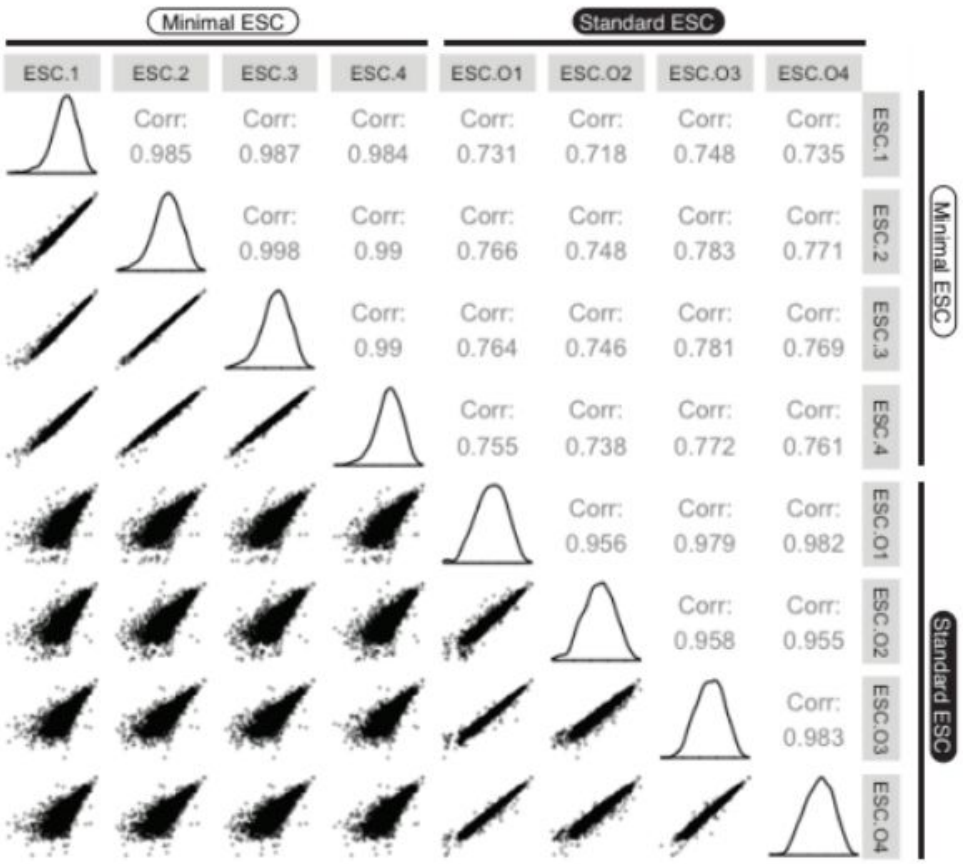


Figure S7. Correlations coefficients (Spearman) of quadruplicates resulting from minimal and standard input motor neurons (MNs).

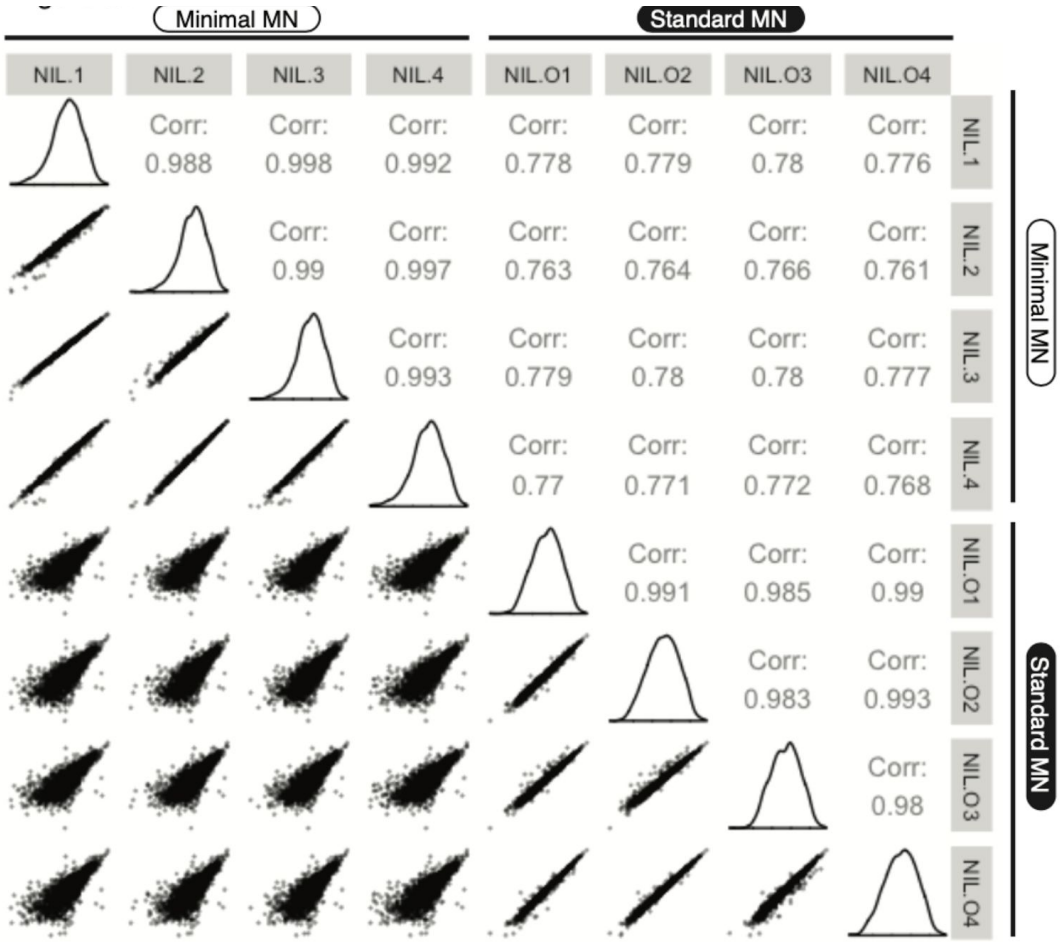


Figure S8. The distributions of the number of peptides per protein (left) and the number of unique peptides per protein (right) are similar for standard (top) and minimal input (bottom) sample preparations. In both preparations, the percentation of proteins quantified based on one peptide is similar: minimal input: 26%, standard input: 27%.

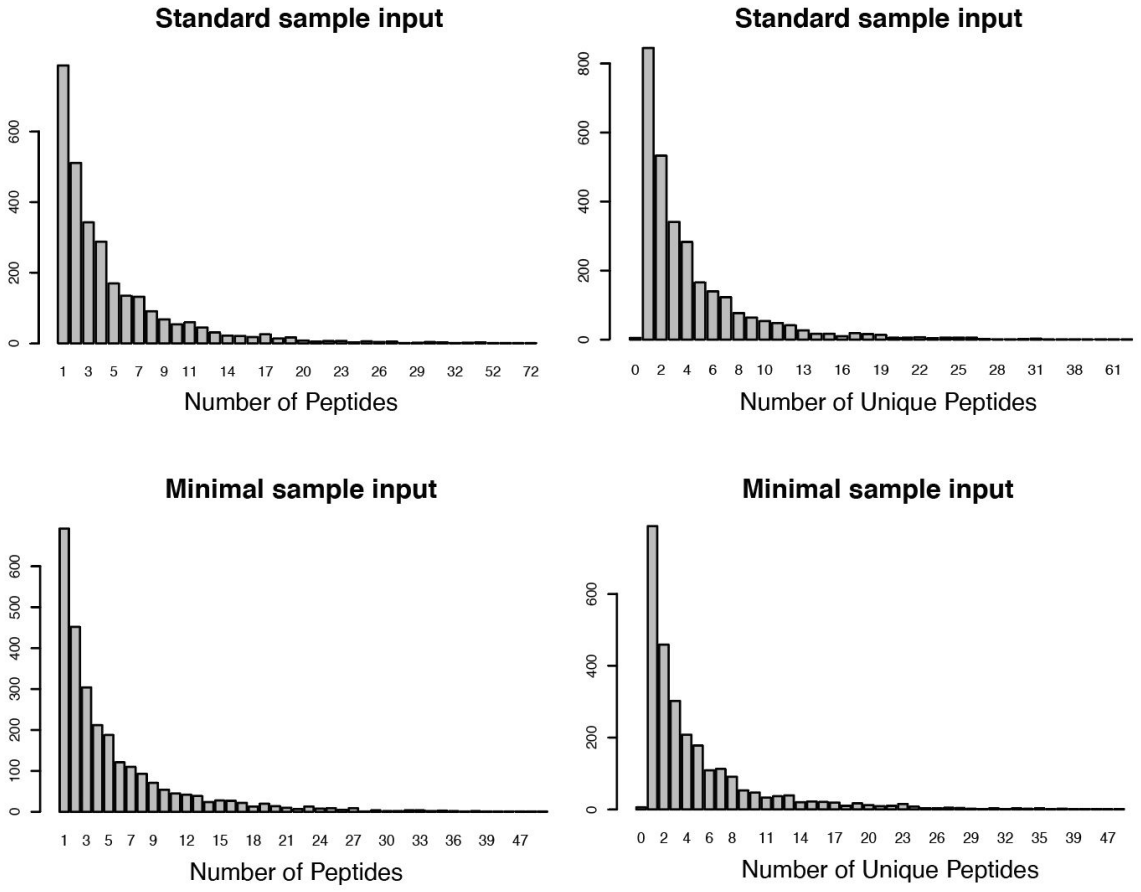


Figure S9. Proteins quantified by a single peptide are overall comparable to the proteins detected by 2 or more peptides regarding abundance correlations.

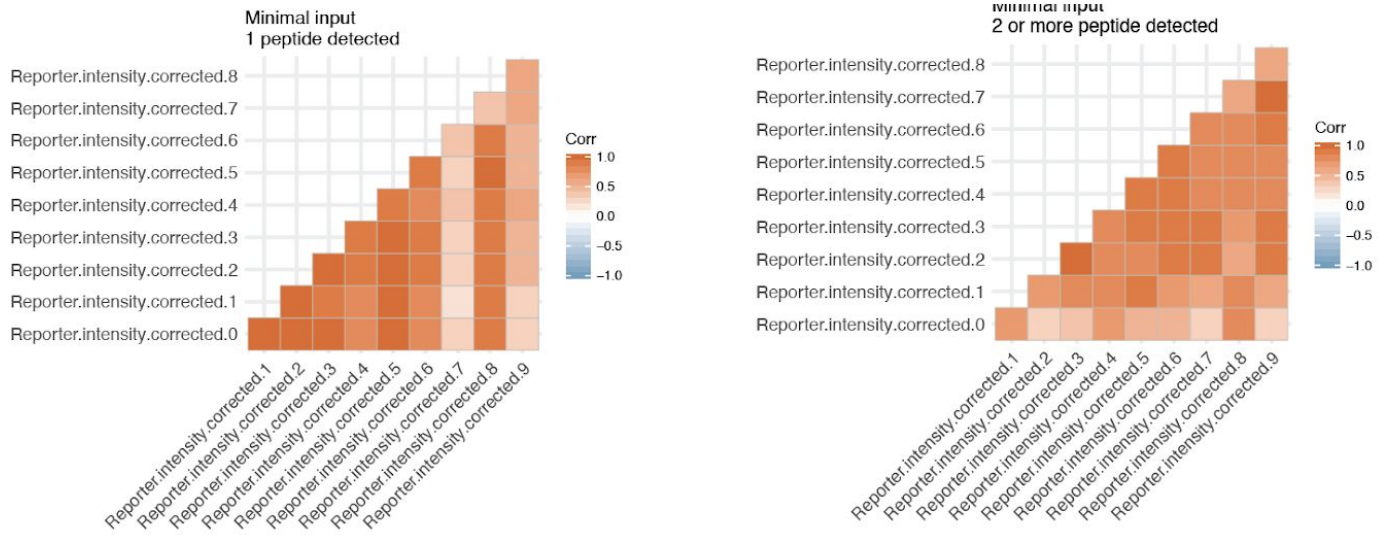


Figure S10. The graph shows the second and third principal components (without the outlier ESC-2 from the standard prep). The analysis was done using 1,763 protein groups that were identified in both minimal and standard preparations.

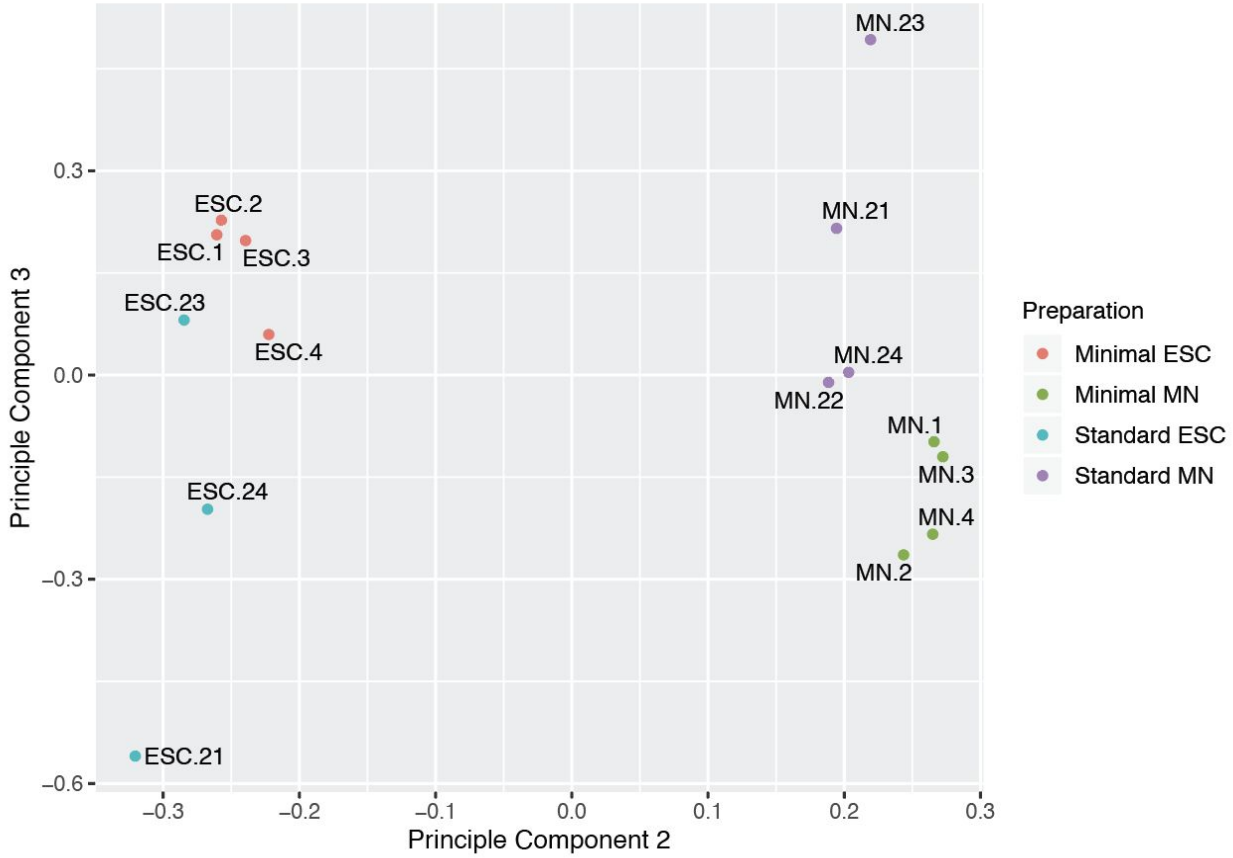


Figure S11. Minimal (top) and standard (bottom) input preparations show similar correlation with corresponding transcript abundances, with Spearman coefficients ranging from 0.39 to 0.43. ESC - embryonic stem cells, MN - motor neurons

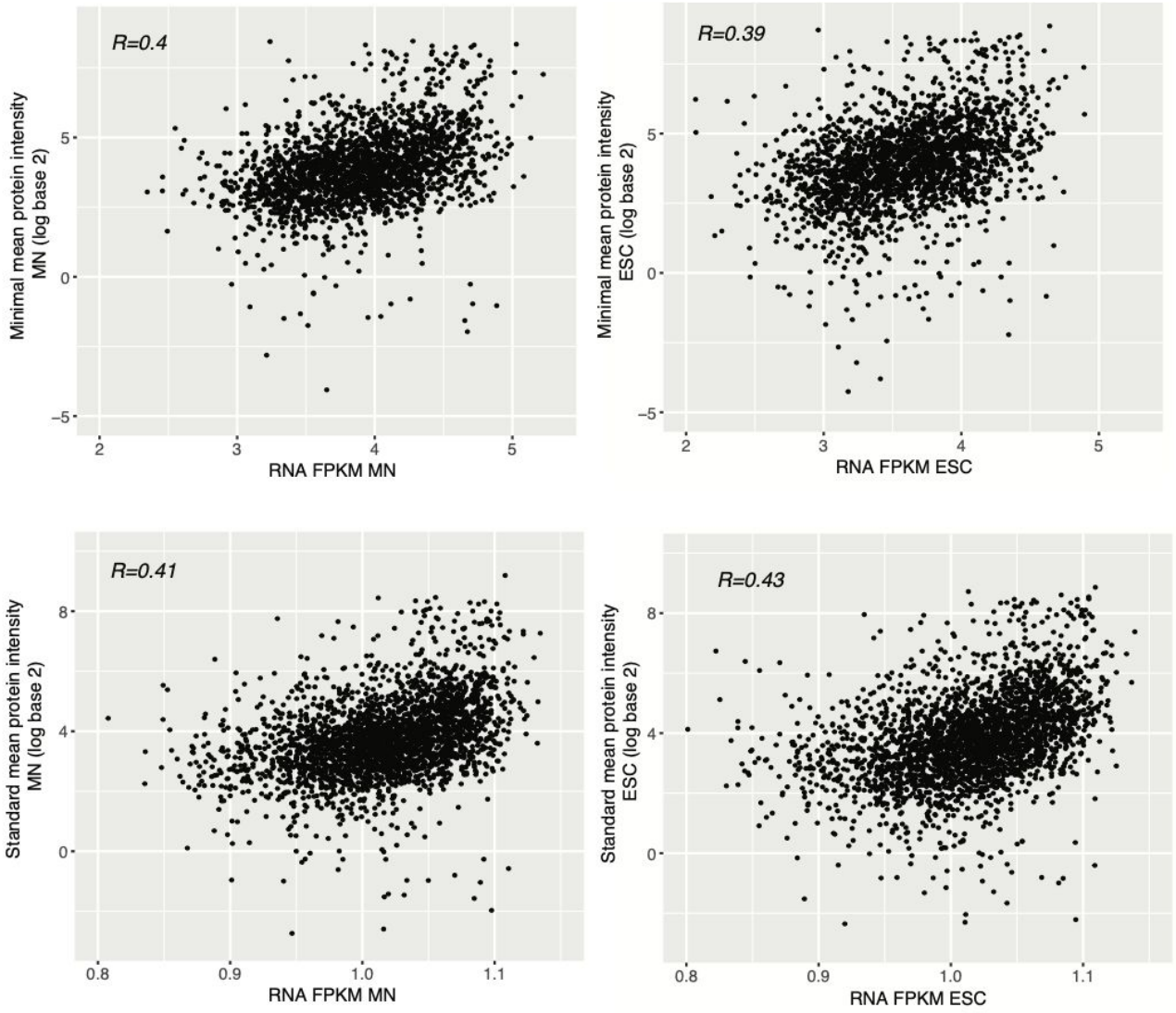


Figure S12. Differentially expressed proteins have similar function enrichments (GO Slim) between the minimal input and standard protocol (p-value<0.05). ESC - embryonic stem cells, MN - motor neurons

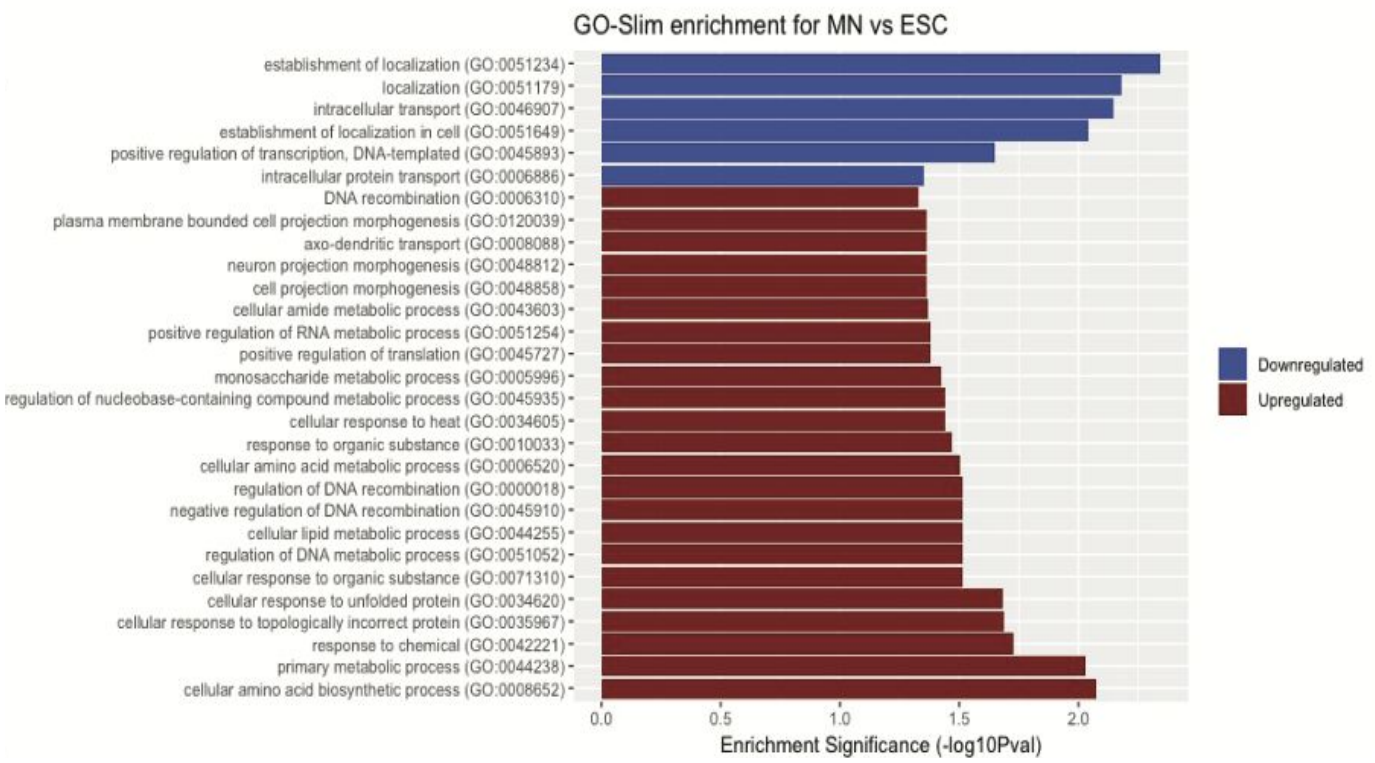


Figure S13. Volcanos plot showing in red the significantly different genes in experimental conditions (Student t-test, q-value<0.05). ESC - embryonic stem cells, MN - motor neurons

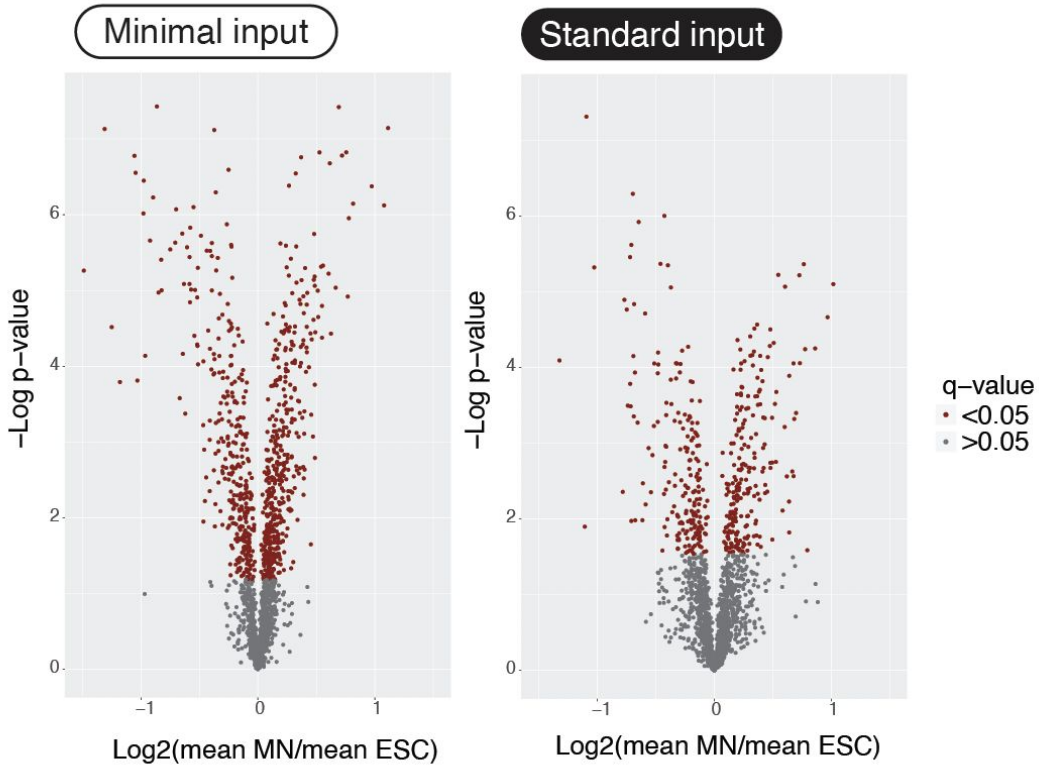


Figure S14. Differentially expressed proteins identified by the minimal input preparation agree with results from standard input preparations. The differentially expressed protein groups from the minimal input data are plotted as a volcano plot (top) and intensity scatter plot (bottom) showing the protein groups that are specific to the minimal input prep in grey, identified and significantly differentially expressed in both minimal and standard input preparations in red and not significantly differentially expressed in the minimal input preparation in blue. We find that 552, 265, and 135 proteins are significantly differentially expressed in both, specifically in the minimal input, and specifically in the standard input sample, respectively. ESC - embryonic stem cells, MN - motor neurons

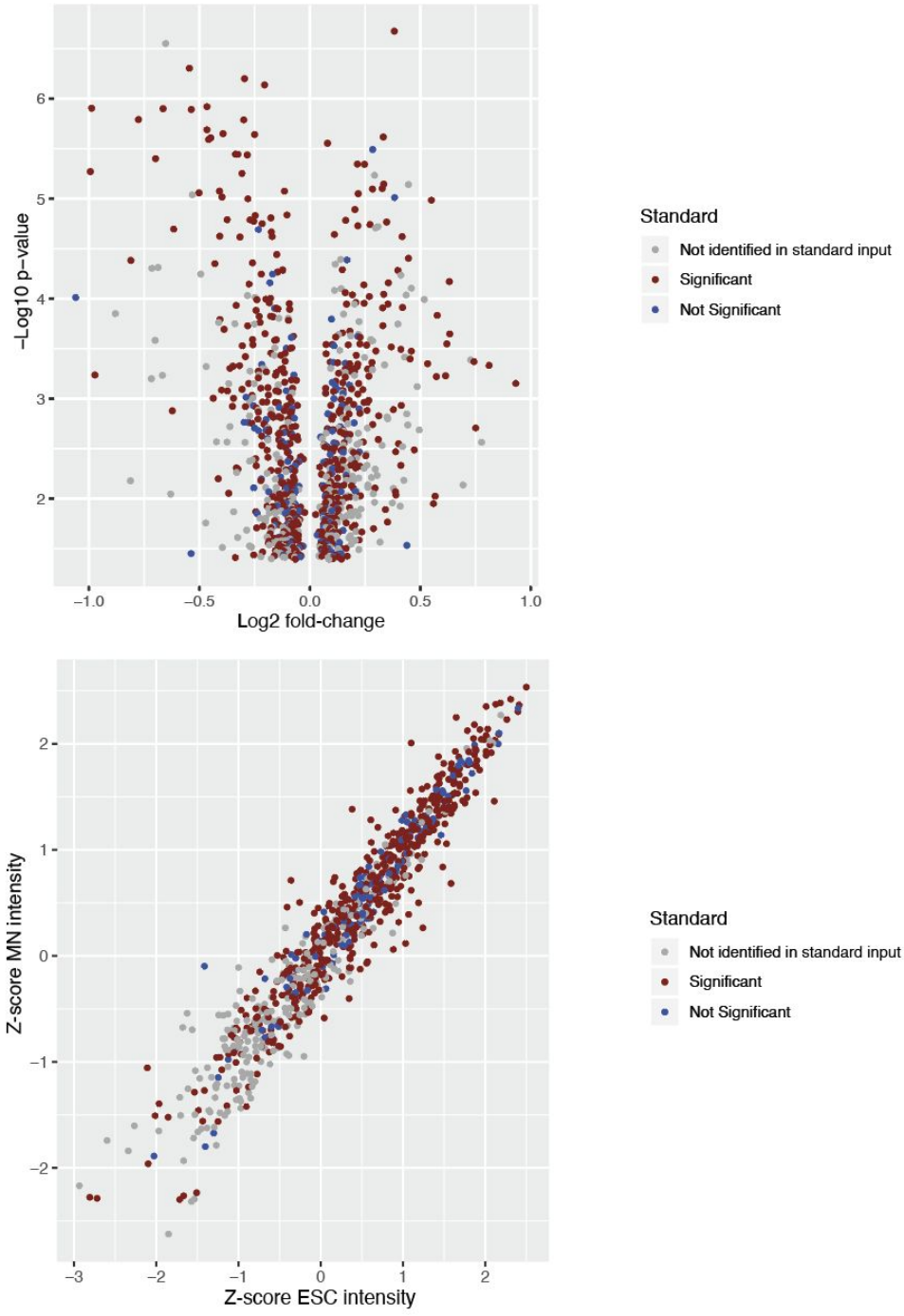
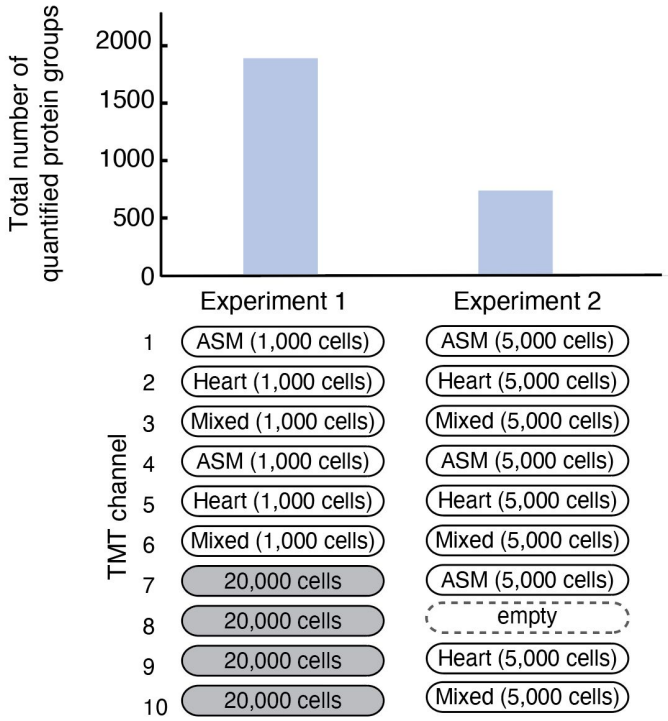
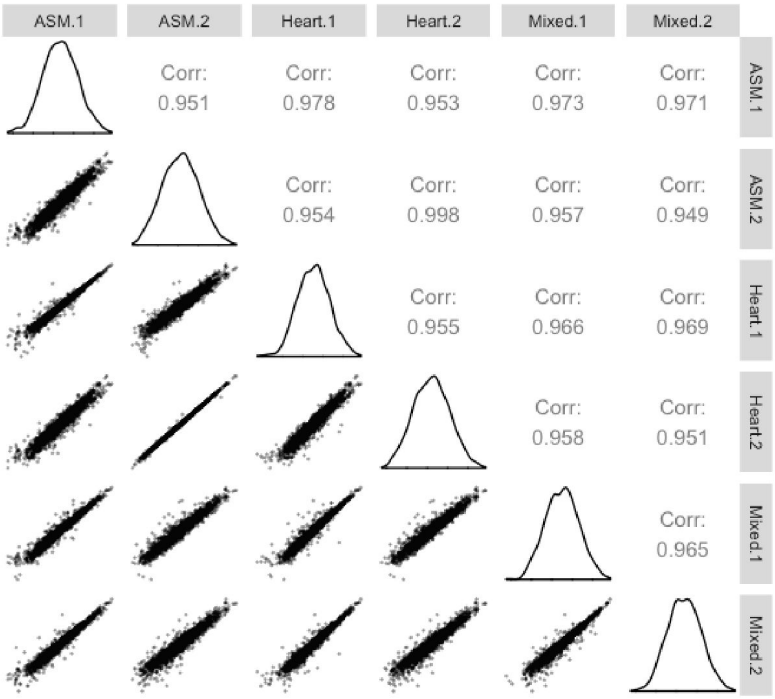


Figure S15. Proteomics data obtained from minimal input protocol using sorted *Ciona robusta* cardiopharyngeal lineage. a. We tested two experimental designs. Experiment 1 used 1,000 cells for each experimental channel with 4 carrier channels of 20,000 whole embryo cells. It identified 1,904 proteins. Experiment 2 used 5,000 cells for each experimental condition with leaving channel 8 empty. It identified 732 identified proteins. b. Experiment 1: reproducibility between protein intensity measurements across replicates. Pearson's correlation coefficients range between 0.95 and 0.98. c. Experiment 2: reproducibility between protein intensity measurements across replicates. Pearson's correlation coefficients range between 0.95 and 0.98. ASM: Atrial Siphon Muscle, condition electroporated *Mesp>Mek*^{S216D,S220E}. Heart: condition electroporated *Mesp>Fgfr*^{DN}. Mixed: condition electroporated *Mesp>LacZ*.

a.



b.


ASM.1
ASM.2
Heart.1
Heart.2
Mixed.1
Mixed.2
Corr: 0.951
Corr: 0.978
Corr: 0.953
Corr: 0.973
Corr: 0.971
Corr: 0.954
Corr: 0.998
Corr: 0.957
Corr: 0.949
Corr: 0.955
Corr: 0.966
Corr: 0.969
Corr: 0.958
Corr: 0.951
Corr: 0.965

c.

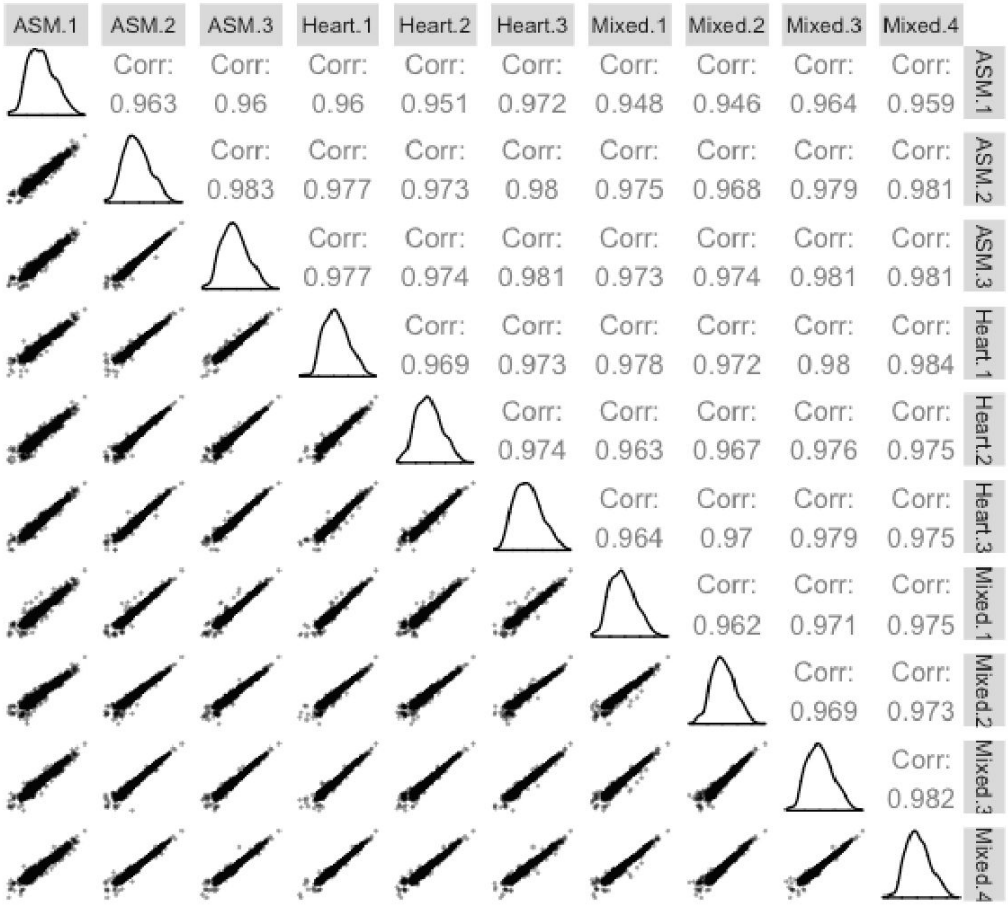
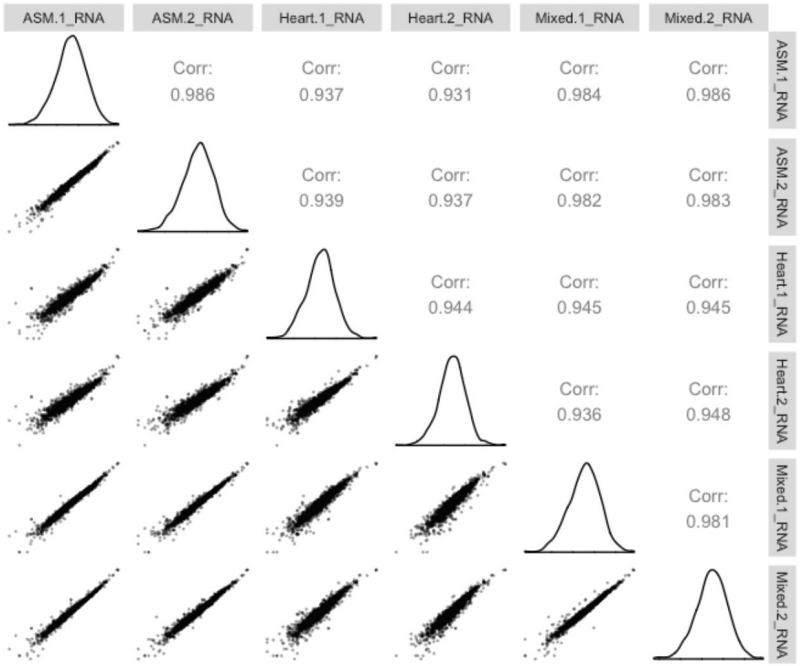

ASM.1
ASM.2
ASM.3
Heart.1
Heart.2
Heart.3
Mixed.1
Mixed.2
Mixed.3
Mixed.4
Corr: 0.963
Corr: 0.96
Corr: 0.96
Corr: 0.951
Corr: 0.972
Corr: 0.948
Corr: 0.946
Corr: 0.964
Corr: 0.959
Corr: 0.983
Corr: 0.977
Corr: 0.973
Corr: 0.98
Corr: 0.975
Corr: 0.968
Corr: 0.979
Corr: 0.981
Corr: 0.977
Corr: 0.974
Corr: 0.981
Corr: 0.973
Corr: 0.974
Corr: 0.981
Corr: 0.981
Corr: 0.969
Corr: 0.973
Corr: 0.978
Corr: 0.972
Corr: 0.98
Corr: 0.984
Corr: 0.974
Corr: 0.963
Corr: 0.967
Corr: 0.976
Corr: 0.975
Corr: 0.964
Corr: 0.97
Corr: 0.979
Corr: 0.975
Corr: 0.962
Corr: 0.971
Corr: 0.975
Corr: 0.969
Corr: 0.973
Corr: 0.982

Figure S16. Transcriptomic data obtained from the sorted *Ciona robusta* cardiopharyngeal lineage. Experimental design as in Figure S15a. a. Experiment 1: reproducibility between RNA count data measurements across replicates. Pearson's correlation coefficients range between 0.93 and 0.99. b. Experiment 2: reproducibility between RNA count data measurements across replicates. Pearson's correlation coefficients range between 0.94 and 0.99. ASM: Atrial Siphon Muscle, condition electroporated *Mesp>Mek*^{S216D,S220E}. Heart: condition electroporated *Mesp>Fgfr*^{DN}. Mixed: condition electroporated *Mesp>LacZ*.

a.



b.

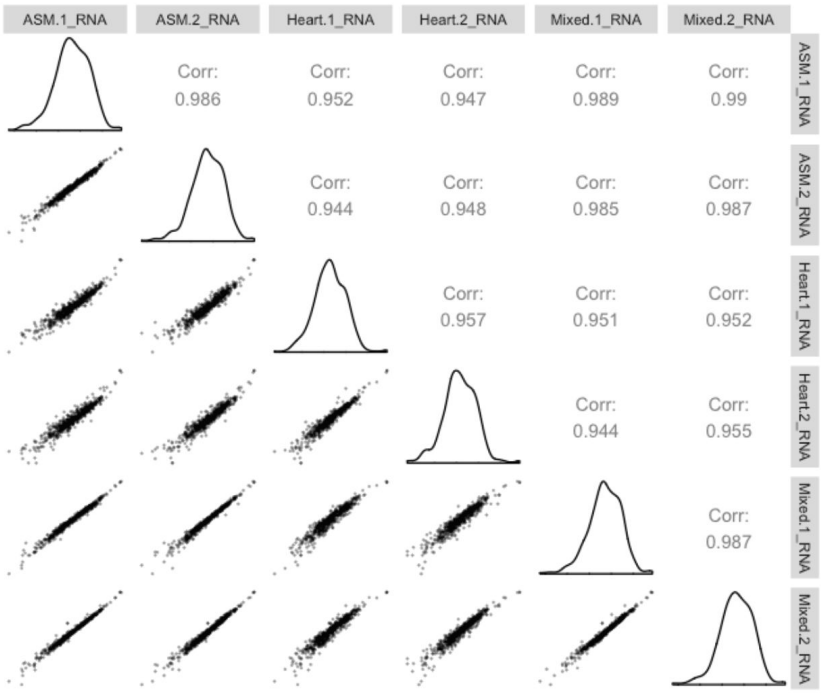
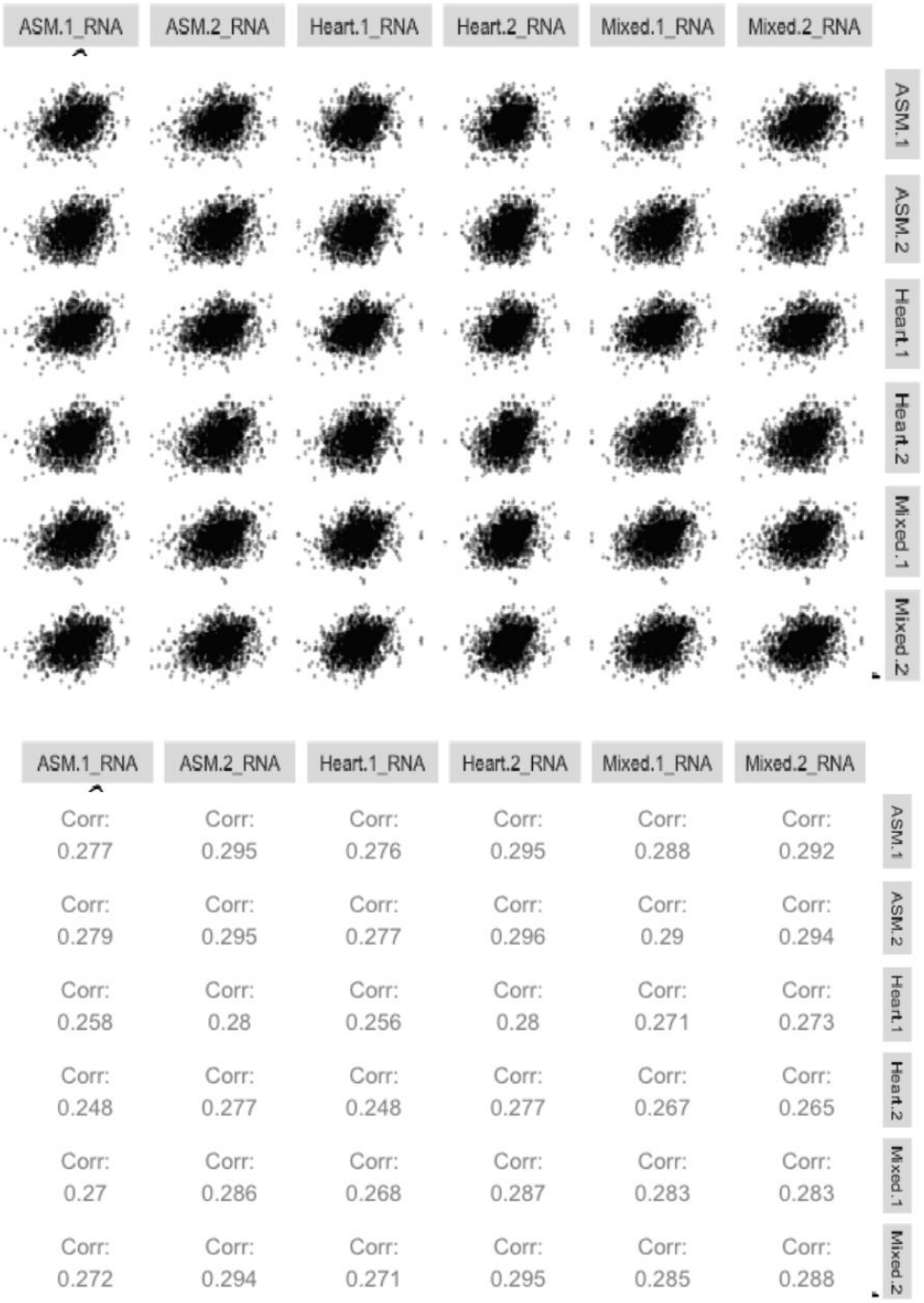
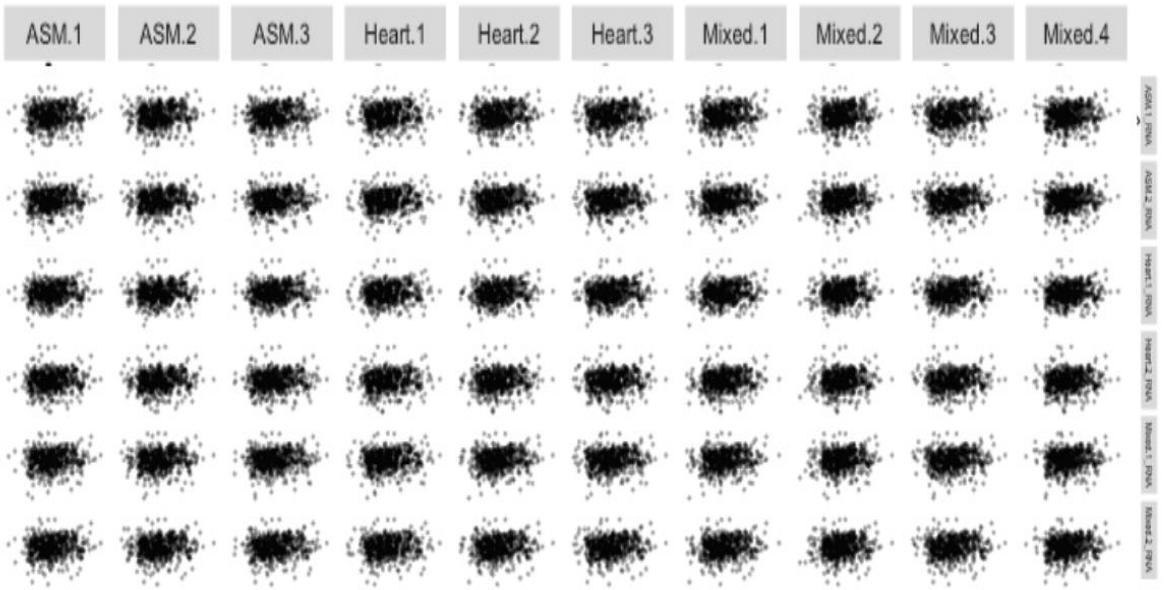


Figure S17. Comparison of transcriptomic and proteomic data obtained from sorted *Ciona robusta* cardiopharyngeal lineage (Spearman correlation coefficients between log base 2 transformed measurements). Experimental design as in Figure S15a. a. Experiment 1: Protein on x-axis, RNA on y-axis. All correlations are significant at p-value<0.01. b. Experiment 2: protein on x-axis, RNA on y-axis. All correlations are positive. ASM: Atrial Siphon Muscle, condition electroporated *Mesp>Mek*^{S216D,S220E}. Heart: condition electroporated *Mesp>Fgfr*^{DN}. Mixed: condition electroporated *Mesp>LacZ*.

a.



b.



RNA

ASM.1_RNA	ASM.2_RNA	Heart.1_RNA	Heart.2_RNA	Mixed.1_RNA	Mixed.2_RNA	
Corr: 0.11	Corr: 0.108	Corr: 0.103	Corr: 0.0953	Corr: 0.108	Corr: 0.11	ASM.1
Corr: 0.107	Corr: 0.105	Corr: 0.101	Corr: 0.0912	Corr: 0.104	Corr: 0.106	ASM.2
Corr: 0.107	Corr: 0.105	Corr: 0.103	Corr: 0.0937	Corr: 0.106	Corr: 0.106	ASM.3
Corr: 0.0978	Corr: 0.0967	Corr: 0.0978	Corr: 0.0899	Corr: 0.0954	Corr: 0.0989	Heart.1
Corr: 0.119	Corr: 0.119	Corr: 0.115	Corr: 0.109	Corr: 0.119	Corr: 0.121	Heart.2
Corr: 0.121	Corr: 0.119	Corr: 0.114	Corr: 0.107	Corr: 0.117	Corr: 0.122	Heart.3
Corr: 0.137	Corr: 0.135	Corr: 0.128	Corr: 0.117	Corr: 0.134	Corr: 0.138	Mixed.1
Corr: 0.12	Corr: 0.117	Corr: 0.106	Corr: 0.098	Corr: 0.116	Corr: 0.116	Mixed.2
Corr: 0.0964	Corr: 0.0928	Corr: 0.0901	Corr: 0.0796	Corr: 0.0938	Corr: 0.0951	Mixed.3
Corr: 0.114	Corr: 0.111	Corr: 0.107	Corr: 0.101	Corr: 0.112	Corr: 0.114	Mixed.4

protein