Supporting Information

Rabinowitz et al. 10.1073/pnas.1620755114

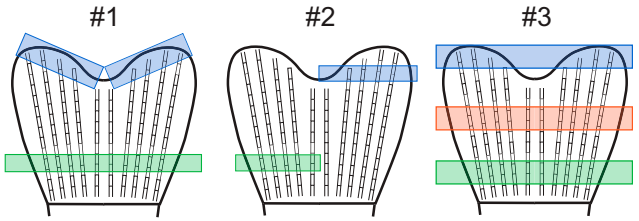

Fig. S1. Illustrations showing regions harvested for the proximodistal proteomics screens. Proximal (green) and distal (blue) regions were collected from three separate biological experiments. Experiment 1 regions were pooled from eight fish, whereas experiments 2 and 3 were pooled from 20 fish each. The middle (orange) region was only collected during experiment 3 from 20 pooled fish.

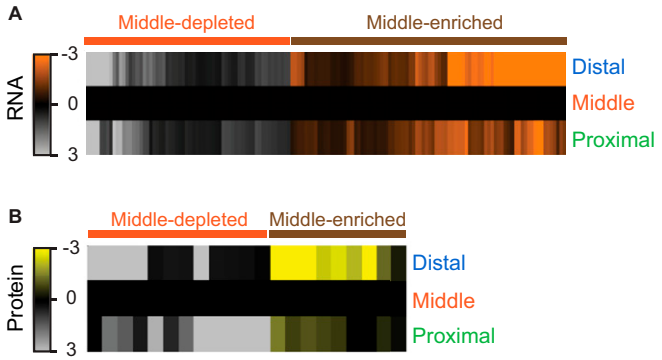

Fig. S2. Middle-enriched and middle-depleted gradients were also identified by RNA-seq and LFQ proteomics. (*A* and *B*) Heat maps of 147 transcripts (*A*) or 45 proteins (*B*) with middle-depleted (red) or middle-enriched (brown) gradients. In *A*, transcript values shown are RPM normalized to middle and then $\log_2$ transformed. Each transcript was differentially expressed (FDR < 1%) between proximal and middle regions and also between middle and distal regions. In *B*, the average protein abundance from all technical replicates is shown as $\log_2$ abundance normalized to the middle region. All protein values for the heat map are derived from experiment 3 only (Fig. S1). Each protein was differentially expressed (FDR < 5%) between proximal and distal regions (see Dataset S21 description for more details).

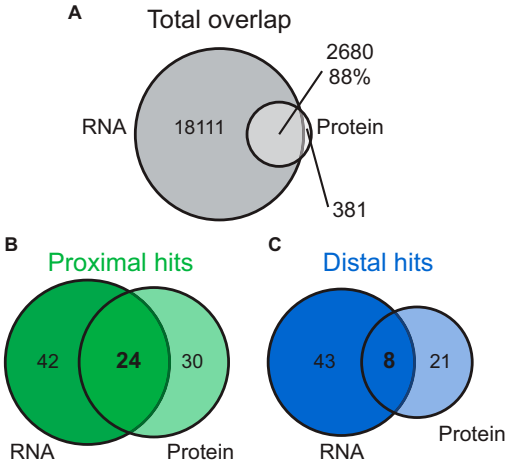

Fig. S3. Comparison between proximal and distal RNA-seq and LFQ proteomics reveals extensive posttranscriptional regulation. (*A*) RNA-seq identified and quantified 20,791 transcripts, whereas LFQ proteomics identified and quantified 3,061 proteins. There were 2,680 genes measured by both screens. (*B* and *C*) Comparing the proximal and distal differentially expressed transcripts and proteins, restricted to the 2,680 commonly measured by both screens, we found 24 genes were proximally enriched for both mRNA and protein (*B*) and 8 genes were distally enriched for both mRNA and protein (*C*). In both the proximal and distal regions, fewer than 50% of the differentially expressed transcripts were found to have similar differentially expressed protein.

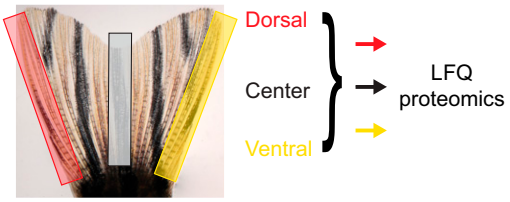

Fig. S4. Representative caudal fin showing regions harvested for dorsal, central, and ventral proteomics.

Dataset S1. Total distal versus proximal RNA-seq data. All RNA-seq data came from five biological replicates per region (proximal, middle, and distal), with each replicate comprised of a pool of two male and two female tissue samples. Measurements are sequencing-depth adjusted raw counts for each transcript. These adjusted raw counts were used to calculate fold change, *P* value, and FDR

Dataset S1

Dataset S2. Total middle versus proximal RNA-seq data. Measurements are sequencing-depth adjusted raw counts for each transcript. These adjusted raw counts were used to calculate fold change, *P* value, and FDR

Dataset S2

Dataset S3. Total distal versus middle RNA-seq data. Measurements are sequencing-depth adjusted raw counts for each transcript. These adjusted raw counts were used to calculate fold change, *P* value, and FDR

Dataset S3

Dataset S4. Distal versus proximal differentially expressed hits. Genes were considered to have differential expression between proximal and distal regions with an FDR < 5% and a fold change > 3

Dataset S4

Dataset S5. Unidirectional gradients between proximal-middle-distal regions. Using an FDR < 1%, we identified genes with differential expression between proximal and middle regions, as well as differential expression between middle and distal regions, in the same direction. Additionally, a "score" is provided that aims to give lower priority to genes that have significant *P* value but small fold change, and higher priority to genes with both significant *P* value and large fold change. This score is calculated as: score = $-\log_{10}$(*P* value) × $\log_2$(region 1/region 2). The final score for each gene is the sum of the scores from each pairwise comparison

Dataset S5

Dataset S6. Middle-enriched and middle-depleted patterns between proximal-middle-distal regions. Using an FDR < 1%, we identified genes with differential expression between proximal and middle regions, as well as differential expression between middle and distal regions, in opposite directions. Middle-enriched RNAs are highlighted in red, and middle-depleted RNAs are highlighted in brown

Dataset S6

Dataset S7. Transcription factors differentially expressed between proximal and distal regions. For reference, we used the total transcription factor database downloaded from www.transcriptionfactor.org/index.cgi?Download on February 10, 2016. Of the 693 transcription factors measured, 53 were differentially expressed between proximal and distal regions (FDR < 1% and fold change > 3). Proximally enriched RNAs are highlighted in green and distally enriched are highlighted in blue

Dataset S7

Dataset S8. Transmembrane receptors differentially expressed between proximal and distal regions. A list of 964 receptors was manually curated from all transcripts measured by RNA-sEq. Eighty RNAs were differentially expressed between proximal and distal regions (FDR < 1% and fold change > 3). Proximally enriched RNAs are highlighted in green and distally enriched are highlighted in blue

Dataset S8

Dataset S9. Ion channel genes differentially expressed between proximal and distal regions. A list of 222 receptors was manually curated from all transcripts measured by RNA-sEq. Fifty-five RNAs were differentially expressed between proximal and distal regions (FDR < 1% and fold change > 3). Proximally enriched RNAs are highlighted in green and distally enriched are highlighted in blue

Dataset S9

Dataset S10. RA pathway genes differentially expressed between proximal and distal regions. A list of 55 RA pathway genes was manually curated from all transcripts measured by RNA-sEq. Twelve RNAs were differentially expressed between proximal and distal regions (FDR < 1% and fold change > 3). Proximally enriched RNAs are highlighted in green and distally enriched are highlighted in blue

Dataset S10

Dataset S11. WNT pathway genes differentially expressed between proximal and distal regions. A list of 52 WNT pathway genes was manually curated from all transcripts measured by RNA-sEq. Thirteen RNAs were differentially expressed between proximal and distal regions (FDR < 1% and fold change > 3). Proximally enriched RNAs are highlighted in green and distally enriched are highlighted in blue

Dataset S11

Dataset S12. FGF pathway genes differentially expressed between proximal and distal regions. A list of 41 FGF pathway genes was manually curated from all transcripts measured by RNA-sEq. Ten RNAs were differentially expressed between proximal and distal regions (FDR < 1% and fold change > 3). Proximally enriched RNAs are highlighted in green and distally enriched are highlighted in blue

Dataset S12

Dataset S13. All proteins identified and quantified from all proteomics screens. This dataset shows $\log_2$ transformed values for protein quantification for the technical replicates and averages from all proximal, middle, distal, regenerating, dorsal, central, and ventral regions examined. Three biological experiments were done by using uninjured proximal and distal regions (Bio#1, Bio#2, and Bio#3; Fig. S1). Bio#2 (fusion) are the same samples from Bio#2 run on a different machine. Bio#1 regions were pooled from eight age-matched fish. Bio#2 regions were from 20 age- and gender-matched fish. Bio#3 was from 20 age- and gender-matched fish. Each sample was run with at least three technical replicates. Regenerating proximal and distal samples were from one experiment with respective regions pooled from 20 age- and gender-matched fish (Fig. 2*A*). Dorsal, central, and ventral samples were from one experiment with respective regions pooled from 20 age-matched and gender-matched fish (Fig. S4)

Dataset S13

Dataset S14. Bio#1, proximal versus distal differential expression. Technical replicates were used to identify differentially expressed proteins between the proximal and distal regions by using *P* values (*t* test) and FDR. Proteins were considered differentially expressed with an FDR < 5%

Dataset S14

Dataset S15. Bio#2, proximal versus distal differential expression. Differentially expressed proteins were identified as in Dataset S14

Dataset S15

Dataset S16. Bio#2 (fusion), proximal versus distal differential expression. Differentially expressed proteins were identified as in Dataset S14

Dataset S16

Dataset S17. Bio#3, proximal versus distal differential expression. Differentially expressed proteins were identified as in Dataset S14

Dataset S17

Dataset S18. Combined proximal versus distal differential expression from all biological experiments. All technical replicates from Datasets S14–S17 were averaged together to identify differentially expressed proteins between the proximal and distal regions by using *P* values (*t* test) and FDR. Proteins were considered differentially expressed with an FDR < 5%

Dataset S18

Dataset S19. Total hit list for proximal/distal differential expression. This dataset shows the 432 proteins that were considered differentially expressed (FDR < 5%) in any single biological experiment or in the combined data (Dataset S18)

Dataset S19

Dataset S20. Strict hit list for proximal/distal differential expression. This dataset shows the final list of 113 differentially expressed proteins, between proximal and distal regions, used for this study. The list includes proteins that were differentially expressed (FDR < 5%) in the same direction, in at least two of the five datasets (Datasets S14–S18)

Dataset S20

Dataset S21. Proximal-middle-distal gradients. Because middle regions were only measured in Bio#3, protein expression gradients were strictly based on values from this dataset. Eight proteins from Bio#3 were differentially expressed (FDR < 5%) between each of the three sampled regions, in a unidirectional manner. To generate a more inclusive list of potential gradients, we first identified proteins that were differentially expressed (FDR < 5%) between proximal and distal regions (Dataset S17). From this filtered list of 283 proteins, we then compared the average values for proximal, middle, and distal regions to find instances where the middle value was between the proximal and distal values, regardless of statistical significance. Based on these criteria, we identified 126 proximally enriched gradients (green), 112 distally enriched gradients (blue), 26 middle-depleted (red), and 19 middle-enriched (brown) gradients

Dataset S21

Dataset S22. Proximal and distal regenerating fins at 1 dpa. Differentially expressed proteins were identified as in Dataset S14

Dataset S22

Dataset S23. Proximal and distal regenerating fins at 3 dpa. Differentially expressed proteins were identified as in Dataset S14

Dataset S23

Dataset S24. Dorsal versus ventral differential expression. Technical replicates were used to identify differentially expressed proteins between the dorsal and ventral regions by using *P* values (*t* test) and FDR. Proteins were considered differentially expressed with an FDR < 5%. No proteins satisfied this FDR requirement

Dataset S24

Dataset S25. Dorsal and ventral combined versus the center differential expression. Differentially expressed proteins were identified as in Dataset S24. Because no proteins were differentially expressed between dorsal and ventral edge, they were averaged together to compare against the center region. Thirty-two proteins were considered differentially expressed (FDR < 5%). This differential expression is represented by a "yes" in the "Regulated" column

Dataset S25

Dataset S26. Metabolomics raw data. Concentrations for 199 metabolites and 4 standards were measured by using LC-MS/MS for proximal, middle, and distal regions of the caudal fin (similar to regions used in Fig. 1*A*). Six biological replicates were used for each region. Replicates were individual age- and gender-matched fish. Metabolite concentrations for each region from each fish are shown. We were able to measure 125 of the 199 metabolites screened

Dataset S26

Dataset S27. Normalized metabolite concentrations. To normalize the raw data, we divided each metabolite's raw data value by the average value from all 125 metabolites measured within that individual sample

Dataset S27

Dataset S28. Proximal and distal differential expression. Using normalized values calculated in Dataset S21 and MetaboAnalyst, we identified 42 metabolites differentially expressed between proximal and distal regions of the fin (FDR < 5%)

Dataset S28

Dataset S29. Proximal-middle-distal gradients. For the 42 metabolites differentially expressed between proximal and distal regions, we used the normalized concentrations (Dataset S21), further normalized to the middle region, to identify metabolite abundance gradients in the fin. Twenty-six metabolites were found to be proximally enriched (green highlight) and 16 were found to be distally enriched (blue highlight)

Dataset S29