In the previous section, we firstly identified different expression pattern between proximal and distal region of the caudal fin in both transcripts, proteins and metabolites using heatmap. To capture the significant ones among them, volcano plot with FDR (or P-Value) and fold change as threshold are applied. In this process, 85 transcripts and 51 proteins with significance are identified which has 32 overlapped genes. Additionally, 42 metabolites with significance are also identified using the similar approach. In our following section, enrichment analysis will be conducted towards those genes using Go database with biological function, cellular component and molecular function annotated. Moreover, the metabolites’ metabolic processes will be annotated using MetaboAnalysis website.

Now let’s focus on the enrichment analysis toward transcriptomics and proteomics. Figure 1 demonstrated the GO terms enriched for overlapped genes from transcripts and proteins with significance. We picked top 2 terms from three Go terms categories respectively from Figure 1 to plot the bubble plot and demonstrated in Figure 2. For biological processes, skeletal system development and epithelial cell differentiation are outlined. It is explainable for its significant relation to regeneration since both skeleton and skin cells are both critical body parts for regeneration. For cellular components, external encapsulating structure played an important role in signal transduction and therefore, contributes greatly to the regeneration. As for the intermediate filament, it is also reasonable, since the most important step for regeneration, cell replication, requires it. For molecular functions, aldehyde dehydrogenase activity for NAD+ and NADP+ are both two key enzymes involved in cell respiration and therefore may impact regeneration process through energy supply.

To give you a more direct feeling for genes’ function, a chord plot for 6 top enriched functions with corresponding genes are plotted in Figure 3. Genes are listed on the left semicircle, while the different GO terms are listed in the right semicircle. Each chord represents an association between a gene and one or more GO terms. In the bottom right, a color gradient bar indicates log fold change goes from blue (indicating proximal enrichment) to red (indicating distal enrichment). From this figure, one key finding is that single biological function can involve both distally and proximally enriched genes. Taking skeleton system development as an example, it involves the distally enriched gene “and2” and the proximally enriched gene “zgc: 172244” which has been marked in red color in Figure 3.

Now let’s focus on the enrichment analysis toward metabolomics. The left figure is a bubble plot in which each bubble represents a specific metabolic process. The horizontal axis represents the minus logarithmic transformation of the p-value, indicating a measure of statistical significance. The larger the value, the higher the statistical significance. The vertical axis lists the names of different metabolisms. Moreover, the size of the bubbles represents the enrichment ratio, which shows the proportion of significantly expressed genes in a specific process. The right figure is the bar plot, much more obviously showing the statistical significance from low to high with yellow color to red color. From two figures demonstrated above, the detailed metabolism can be categorized into three types which has been marked in three different colors in both Figure 3 and Figure 4: metabolism like “purine and pyrimidine metabolism” for DNA and RNA synthesis; metabolism like “alanine, aspartate and glutamate metabolism” for protein synthesis; metabolism like “Citrate Cycle (TCA cycle)” for energy supply. These three types of metabolism all play critical role in regeneration.

To summarize the whole enrichment process, for transcriptomics and proteomics, we identified the top 2 terms from biological processes, cellular components and molecular functions and illustrated their relation to regeneration. Besides this, to give you a more direct feelings about genes’ function with Go terms, we plotted the chord plot, showing the relation between genes and go terms. Additionally, we find a single go term can involve both proximally enriched genes and distally enriched genes. For metabolomics, we categorized the enrichment results to three types: which are DNA and RNA synthesis, protein synthesis and energy supply. Due to this, the metabolites’ function with regeneration can be better understanded.