Figure S1  Density curves of synonymous mutations to assess false orthologs for the pairwise comparisons of *H. carbonarium* (A), *H. flavolineatum* (B) and *H. macrostomum* (C). Histograms indicate overall distribution of pair-wise $dS$ values for each pairwise comparison, and bars indicate the density of genes assigned to each Gaussian component (Component1= blue; 2= red; 3= green; all genes= black). Genes assigned to the third component (green) were excluded from the analyses, as they had the highest $dS$ values and were assumed to be inaccurate calls.
Figure S2  Histograms show the distribution of pairwise $1.0<dN/dS<4.0$ for *H. carbonarium* (A), *H. flavolineatum* (B), *H. macrostomum* (C). Colored lines represent the density of four-component mixture models fitted to all $dN/dS$ (between 0 and 20). The fourth component (teal) was assumed to capture false ortholog calls and was removed. For all three species, the transition between the third and fourth component (vertical black dotted line) occurred at a $dN/dS$ value of 3.
Figure S3  Bar graphs reflecting the completeness of the transcriptome in reference to the BUSCO ortholog database for vertebrates (A) and ray-finned fish (B) for the species *H. carbonarium* (HCAR), *H. flavolineatum* (HFLA) and *H. macrostomum* (HMAC).
**Figure S4**  Gene Ontology (GO) terms significantly enriched for the \( d_N/d_S \) comparisons between *H. carbonarium* and *H. flavolineatum*, for the categories Biological Process (A), Cellular Component (B) and Molecular Function (C). These results include only genes under positive selection, and size and brightness of GO terms depends on the level of significance. Numbers that precede the category name are the genes present in this particular analysis over the total number of genes that belong to that category.

**A**
- 74/111 immune response
- 123/205 immune system process
- 19/21 humoral immune response
- 84/127 defense response
- 69/109 positive regulation of immune system process

**B**
- 91/137 extracellular space
- 83/136 extracellular region
- 131/234 plasma membrane part
- 53/87 integral component of plasma membrane
- 6/6 transcriptionally active chromatin

**C**
- 33/46 carbohydrate binding
- 8/8 scavenger receptor
- 16/20 glycoprotein binding
Gene Ontology (GO) terms significantly enriched for the $d_N/d_S$ comparisons between *H. carbonarium* and *H. macrostomum*, for the categories Biological Process (A), Cellular Component (B) and Molecular Function (C). These results include only genes under positive selection, and size and brightness of GO terms depends on the level of significance. Numbers that precede the category name are the genes present in this particular analysis over the total number of genes that belong to that category.
Figure S6  Gene Ontology (GO) terms significantly enriched for the $d_N/d_S$ comparisons between *H. flavolineatum* and *H. macrostomum*, for the categories Biological Process (A), Cellular Component (B) and Molecular Function (C). These results include only genes under positive selection, and size and brightness of GO terms depends on the level of significance. Numbers that precede the category name are the genes present in this particular analysis over the total number of genes that belong to that category.

A  
- 24/31 response to bacterium  
- 92/135 defense response  
- 47/67 response to other organism  
- 30/42 positive regulation of cytokine production  
- 137/219 immune system process  
- 60/69 immune effector process  
- 8/8 complement activation, alternative pathway  
- 20/23 humoral immune response  
- 8/8 innate immune response  
- 81/118 immune response  
- 17/20 positive regulation of response to stimulus  
- 20/23 response to glucocorticoid  
- 50/81 response to organic cyclic compound  
- 27/42 cellular response to organic cyclic compound  
- 13/16 cellular response to steroid hormone stimulus  
- 55/83 response to lipid  
- 32/46 cellular response to lipid  
- 8/8 response to ammonium ion  
- 6/6 complement activation, alternative pathway  
- 52/72 innate immune response  
- 70/116 positive regulation of immune system process  
- 34/49 regulation of immune effector process  
- 101/159 regulation of immune system process  
- 8/8 regulation of B cell mediated immunity  
- 20/29 leukocyte migration  
- 23/33 taxis  
- 16/20 zymogen activation  
- 8/9 regulation of telomere maintenance  
- 8/9 negative regulation of peptidase  
- 5/5 epoxygenase P450 pathway  
- 8/9 long-chain fatty acid metabolic process  
- 124/190 lipid metabolic process  

B  
- 42/63 cell surface  
- 90/140 extracellular space  
- 151/257 plasma membrane part  
- 17/18 external side of plasma membrane  
- 24/31 side of membrane  
- 95/145 extracellular region  

C  
- 12/12 cargo receptor  
- 8/8 scavenger receptor  
- 9/9 cytokine receptor  
- 53/85 receptor  
- 37/48 carbohydrate binding  
- 24/30 cysteine-type peptidase  
- 16/18 cysteine-type endopeptidase  
- 79/120 peptidase  
- 5/6 antigen binding  
- 7/7 aromatase  
- 17/21 monoxygenase  
- 9/10 protein binding involved in cell-cell adhesion

\[ p < 0.01 \]  
\[ p < 0.05 \]  
\[ p < 0.1 \]
Figure S7  
Example of a heatmap for transcripts of Vitellogenin and Zona Pellucida genes, which were highly upregulated for seven individuals of *Haemulon flavolineatum* (HFL; in bold). These individuals were removed from the analyses in order to get more accurate measures of differential expression and GO enrichment. The other individuals of *H. flavolineatum* have similar levels of expression to all the collected individuals of *H. macrostomum* (HMCR).
**Figure S8** Gene Ontology (GO) terms significantly enriched in the comparison of gene expression of *H. carbonarium* and *H. flavolineatum*, for the categories Cellular Component (A) and Molecular Function (B). Size and brightness of GO terms depends on the level of significance, and numbers that precede the category name are the genes present in this particular analysis over the total number of genes that belong to that category. GO terms in red are upregulated categories for *H. carbonarium*, while blue ones are downregulated. The category of Biological Process only had two significant terms: Translation (downregulated, 182 sequences) and Immune Responses (upregulated, 177 sequences).

A

- 19/69 extracellular matrix part
- 7/11 microfibril
- 53/233 extracellular matrix
- 245/812 extracellular space
- 22/68 collagen trimer
- 242/1314 plasma membrane part
- 5/9 immunoglobulin complex
- 17/130 receptor complex
- 127/731 endoplasmic reticulum
- 139/779 ribonucleoprotein complex
- 2/78 ribosome
- 12/32 cytosolic large ribosomal subunit
- 85/588 nucleolus

B

- 52/405 regulatory region nucleic acid binding
- 34/289 transcription regulatory region sequence-specific DNA binding
- 68/205 serine hydrolase
- 43/132 structural constituent of ribosome
- 77/7 nutrient reservoir
Figure S9  Gene Ontology (GO) terms significantly enriched with the MWU Test of gene expression for *H. carbonarium* and *H. macrostomum*, for the categories Biological Process (A), Cellular Component (B) and Molecular Function (C). Due to the large number of significant GO terms for BP (171), only those with a FDR<0.02 were included in the tree. Size and brightness of GO terms depends on the level of significance, and numbers that precede the category name are the genes present in this particular analysis over the total number of genes that belong to that category. GO terms in red are upregulated categories for *H. carbonarium*, while blue ones are downregulated.
**Figure S10**  Gene Ontology (GO) terms significantly enriched with the MWU Test of gene expression for *H. flavolineatum* and *H. macrostomum*, for the categories Biological Process (A), Cellular Component (B) and Molecular Function (C). Size and brightness of GO terms depends on the level of significance, and numbers that precede the category name are the genes present in this particular analysis over the total number of genes that belong to that category. GO terms in red are upregulated categories for *H. flavolineatum*, while blue ones are downregulated.