

Supplementary Figures

PREDICTED: *Danio rerio* zgc:161969 (zgc:161969), transcript variant X1, protein coding

NCBI Reference Sequence: XM_005168715.4

MERSRTAF **DHWMHKHYFTYRDTRGRNITVQC�LCLPKINILSTARDSTSNLKKHLECHS** PHLMRKHQHIEMSEP
EPLFLPPKRLKSAVWEYFGYYKDPGGTIIADGRPVCRMCRRKVSARSNTSNMQHHLRDHHPMEFAKLNKHAV
PCRKLREAAGIDGSSSFATNILENQTEESSEAKDDKNDSTIYYQTSQSRLNALVFNFIVEDVQPIISILEQPGF
RKLIEALSRGKRVMSPNFVTRLEMAFDKMKTELKAKLSKVQTLCTTADIWSVQDRSYFGMTCHWLEDNLERRS
AALACTRIPSNYTCETIIAKIQETHSSFNIESKVQATVTDNGNNFVKAFKEFSSDDEQEVEQSQFEDLGSILCD
GDIGGDFFLSYFLPPHQRCAGQTLTTIASKDLAEAVTKGPAGKVRSSAAGKCAKIWHKAESSMETAVAMESIAN
MKFTLPCLNQWSSEYYAINKMLTNDSQLNELTELLGVPHFTPDETSYLTEYTDVFKPVAFALDLLQGEEKCFL
GIVIPTLLTLKRKLEEKAAANTRLFSAIEETVKAIDTRFKQVFESSDARLATVTTTPQFRLWWLPEEEREVLRAQ
LITEVLQVDQGTETPEETNGGAVHEDEFFSYGPGSPGYKCGKREAAEEVWLYLQGTNKDLKCLNEFPGVKKVFIK
FNTTLPSSAPVQQLFNSSGNNVNPKNPLTDEQFEHVLLRLYNSKIISALE

Danio rerio zgc:161969 (zgc:161969), protein coding [\[in this study\]](#)

NCBI Reference Sequence: NM_001083849.1

MERSRTAF **DHWMHKHYFTYRDTRGRNITVQC�LCLPKINILSTARDSTSNLKKHLERKHAVPCRKLREAAGID**
GSSSFATNILENQTEESSEAKDDKNDSTIYYQTSQSRLNALVFNFIVEDVQPIISILEQPGFRKLIEALSRGKRV
MSPNAFVTRLEMAFDKMKTELKAKLSKVQTLCTTADIWSVQDRSYFGMTCHWLEDNLERRSAAALACTRIPSNY
CETIIAKIQETHSSFNIESKVQATVTDNGNNFVKAFKEFSSDDEQEVEQSQFEDLGSILCDGDIGGDFFLSYFL
PPHQRCAGQTLTTIASKDLAEAVTKGPAGKVRSSAAGKCAKIWHKAESSMETAVAMESIANMKFTLPCLNQWS
EYYAINKMLTNDSQLNELTELLGVPHFTPDETSYLTEYTDVFKPVAFALDLLQGEEKCFLGIVIPTLLTLKRK
LEEKAAANTRLFSAIEETVKAIDTRFKQVFESSDARLATVTTTPQFRLWWLPEEEREVLRAQLITEVLQVDQGT
EPETNGGAVHEDEFFSYGPGSPGYKCGKREAAEEVWLYLQGTNKDLKCLNEFPGVKKVFIKFNNTLPSSAPVQ
LFNSGNNVNPKNPLTDEQFEHVLLRLYNSKIISTLE

Danio rerio uncharacterized protein LOC569044 [\[in this study; the sequence same as NM_001083849.1 protein coding\]](#)

NCBI Reference Sequence: NP_001077318.1

MERSRTAF **DHWMHKHYFTYRDTRGRNITVQC�LCLPKINILSTARDSTSNLKKHLERKHAVPCRKLREAAGID**
GSSSFATNILENQTEESSEAKDDKNDSTIYYQTSQSRLNALVFNFIVEDVQPIISILEQPGFRKLIEALSRGKRV
MSPNAFVTRLEMAFDKMKTELKAKLSKVQTLCTTADIWSVQDRSYFGMTCHWLEDNLERRSAAALACTRIPSNY
CETIIAKIQETHSSFNIESKVQATVTDNGNNFVKAFKEFSSDDEQEVEQSQFEDLGSILCDGDIGGDFFLSYFL
PPHQRCAGQTLTTIASKDLAEAVTKGPAGKVRSSAAGKCAKIWHKAESSMETAVAMESIANMKFTLPCLNQWS
EYYAINKMLTNDSQLNELTELLGVPHFTPDETSYLTEYTDVFKPVAFALDLLQGEEKCFLGIVIPTLLTLKRK
LEEKAAANTRLFSAIEETVKAIDTRFKQVFESSDARLATVTTTPQFRLWWLPEEEREVLRAQLITEVLQVDQGT
EPETNGGAVHEDEFFSYGPGSPGYKCGKREAAEEVWLYLQGTNKDLKCLNEFPGVKKVFIKFNNTLPSSAPVQ
LFNSGNNVNPKNPLTDEQFEHVLLRLYNSKIISTLE

Figure S1. The deduced amino acid sequences of two potential transcripts and one uncharacterized protein LOC569044. The two potential transcripts XM_005168715 and NM_001083849.1 were indicated by NCBI (upper and middle panels). The below panel showed the uncharacterized protein LOC569044, which is focused in this study. Yellow color indicated an essential ZBED domain, while green color indicated the XM_005168715 transcript which lacks few amino acids compared with those of NM_001083849.1 and uncharacterized protein LOC569044.

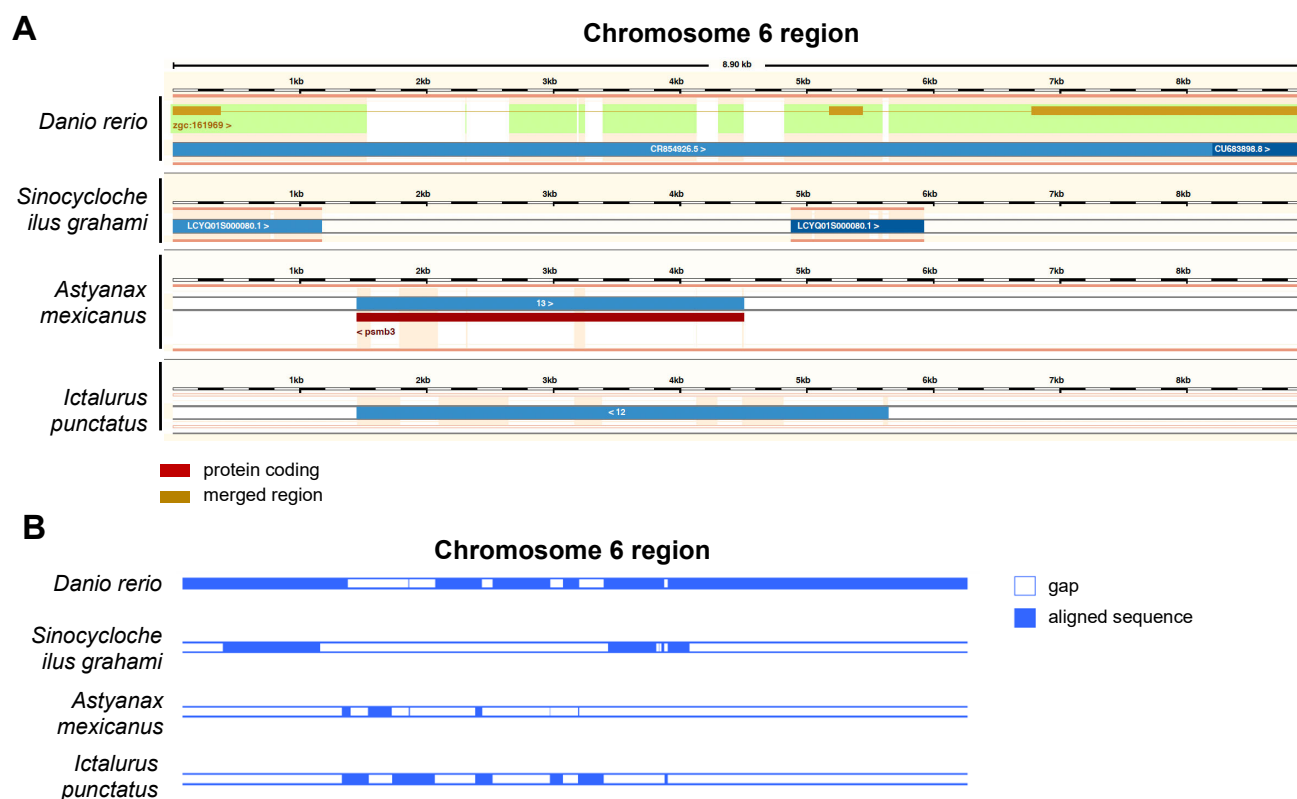


Figure S2. Comparison of the gene maps encoding zebrafish LOC569044 protein and those of multiple bony fishes. (A) *Sinocyclocheilus grahami*, *Astyanax mexicanus* and *Ictalurus punctatus* were found that they were similar alignment sequences located in chromosome 6 region compared to that of zebrafish *zgc:161969*. Horizontal lines indicated chromosome fragments, while dark brown color indicated protein encoded region in the genome. Aligned sequence were indicated in the light brown color and their positions were indicated above in kb. Synteny blocks in fragmented assembly defined by the *Ensembl*. **(B)** The schematic diagram showed the genomic alignments from panel A. The empty blue box indicated non-aligned sequence region, while the blue box indicated aligned sequence region in the genome. This model of synteny conservation is based on the *Ensembl*.