**Supplementary Table S1.** Summary for the BLASTp searches of *Clonorchis sinensis* proteomic database in GenBank with human bHLH-PAS family proteins.

|  |  |  |
| --- | --- | --- |
| **Human query protein** |  | **Matched *C. sinensis* protein** |
| **Subfamily name** | **Protein name** | **Accession No.** |  | **Protein namea** | **Length (aa)b** | **Accession No.** | **E-value** |
| NCOA (SRC) | HsNCOA2 | Q15596.2 |  | CsSIM | 838 | GAA50744.1 | 5e-10 |
|  |  |  |  | CsHIF-1β | 718\* | GAA50153.1 | 6e-05 |
| ARNT | HsARNT2 | NP\_055677.3 |  | CsHIF-1β | 718\* | GAA50153.1 | 1e-53 |
|  |  |  |  | CsSIM | 838 | GAA50744.1 | 3e-26 |
| Orphan | HsMMAL1e | NP\_001284653.1 |  | CsHIF-1β | 718\* | GAA50153.1 | 7e-22 |
|  |  |  |  | CsSIM | 838 | GAA50744.1 | 2e-15 |
| Clock | HsCLOCK | NP\_004889.1 |  | CsSIM | 838\* | GAA50744.1 | 7e-23 |
|  |  |  |  | CsHIF-1β | 718\* | GAA50153.1 | 2e-10 |
|  |  |  |  | CsTFEB | 860  | GAA48933.1 | 7e-06 |
| Bmal/cycle | HsBmal1 | NP\_001284648.1 |  | CsHIF-1β | 718\* | GAA50153.1 | 9e-31 |
|  |  |  |  | CsSIM | 838 | GAA50744.1 | 1e-26 |
| AHR | HsAHR | NP\_001612.1 |  | CsAHR | 1514\* | GAA50018.1 | 1e-37 |
|  |  |  |  | CsSIM | 838 | GAA50744.1 | 2e-31 |
| NPAS4 | HsNPAS4 | NP\_849195.1 |  | CsSIM | 838 | GAA50744.1 | 3e-22 |
|  |  |  |  | CsAHR | 1514\* | GAA50018.1 | 1e-05 |
| SIM | HsSIM1 | NP\_005059.2 |  | CsSIM | 838 | GAA50744.1 | 2e-147 |
|  |  |  |  | CsHIF-1α | 250 | GAA56007.1 | 2e-10 |
| Trh | HsNPAS3 | NP\_071406.1 |  | CsSIM | 838 | GAA50744.1 | 3e-84 |
|  |  |  |  | CsHIF-1α | 250 | GAA56007.1 | 3e-13 |
|  |  |  |  | CsAHR | 1514\* | GAA50018.1 | 3e-10 |
|  |  |  |  | CsHIF-1β | 718\* | GAA50153.1 | 5e-07 |
| HIF | HsHIF-1α | NP\_001521.1 |  | CsSIM | 838 | GAA50744.1 | 2e-78 |
|  |  |  |  | CsHIF-1α | 250 | GAA56007.1 | 6e-12 |
|  |  |  |  | CsAHR | 1514\* | GAA50018.1 | 1e-08 |
|  |  |  |  | CsHIF-1β | 718\* | GAA50153.1 | 6e-06 |

aAbbreviations for *C. sinensis* proteins: SIM, single-minded homolog; AHR, aryl hydrocarbon receptor; TFEB, transcription factor EB; HIF, hypoxia-inducible factor.

bPartial sequences defined by annotators in GenBank are marked with an asterisk.

**Supplementary Table S2.** Primers used in the quantitative real-time RT-PCR (qRT-PCR).

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Gene name** | **Oligo name** | **Sequence (5′ → 3′)** |
| 1 | *CsHIF-1α* | CsHIF1α-q-F | GACAGCTTTCTGATTATCTTGGCA |
|  | CsHIF1α-q-R | AAATCGCTCTCTCCAGCACA |
| 2 | *CsSIM* | CsSIM-q-F | CGCTCATCCCGTCCAAACTG |
|  | CsSIM-q-R | TGACTGTTTTTCATTGCAGCCAC |
| 3 | *CsHIF-1β* | CsHIF1β-q-F | CACCTACAAACAGCCACCGACA |
|  | CsHIF1β-q-R | AGTGGGGTGTGTTAAAGCGGTA |
| 4 | *CsAHR* | CsAHR-q-F | TTGGAGAAGTTTCAGCGAATGGA |
|  | CsAHR-q-F | TCAGAGAGCGTTATTTGCCCA |
|  5 | *β-actin* | CsbAct-q-F | CAAAGCAGGTTTTGCTGGTG |
|  | CsbAct-q-R | ATACCTCGTTTGGATTGGGC |

**Supplementary Fig. S1.** Schematic procedure for the determination of the full-length *bHLH-PAS* genes in *Clonorchis sinensis*. The multiple sequences homologous to each of the *bHLH-PAS* genes were retrieved either from the GenBank proteomic (coding DNA sequences, thin black lines) and expressed sequence tags (blue lines) databases or from the transcriptome shotgun assembly database of the liver fluke. After separately aligning the retrieved sequences, a contig sequence (thick black line) was assembled to represent the probable full-length coding DNA sequence of the corresponding gene. Gaps observed in the alignments (dotted black lines) were filled with nucleotides, which were empirically determined from complementary DNAs by performing PCR (green lines; primers used in the PCR amplifications are shown in the parentheses). Numerals in the parentheses indicate the nucleotide positions matched to those in the contig sequences. The nucleotide positions upstream and downstream of the initiation and stop codons are marked as ‘-’ and ‘+’, respectively.



**Supplementary Fig. S2.** Comparison of primary structures between the *Clonorchis sinensis* bHLH-PAS proteins and their corresponding human orthologs. Amino acids solely from the basic helix-loop-helix (bHLH) to PAS-associated C-terminal (PAC) domains were isolated for the alignment. Dashes in the alignment were introduced to increase the identity values. Various shades of gray indicate the degrees of similarity among the proteins. Open circles in bHLH domain region mark amino acid sequences selected for the syntheses of CsHIF-1α and CsSIM oligopeptides, which were used in the preparation of specific antisera.

