Sequence and functional characterization of hypoxia inducible factors, HIF1α, HIF2αa, and HIF3α, from the estuarine fish, *Fundulus heteroclitus*

Ian K. Townley[1,2], Sibel I. Karchner[3], Elena Skripnikova[1,2], Thomas E. Wiese[2], Mark E. Hahn[3], and Bernard B. Rees[1]

[1]Department of Biological Sciences, University of New Orleans, New Orleans, LA 70148

[2]College of Pharmacy, Xavier University of New Orleans, New Orleans, LA 70125

[3]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543

Address correspondence to Bernard B. Rees, brees@uno.edu

Running Title: Hypoxia inducible factors in killifish

ABSTRACT

The hypoxia inducible factor (HIF) family of transcription factors plays central roles in the development, physiology, pathology, and environmental adaptation of animals. Because many aquatic habitats are characterized by episodes of low dissolved oxygen, fish represent ideal models to study the roles of HIF in the response to aquatic hypoxia. The estuarine fish *Fundulus heteroclitus* occurs in habitats prone to hypoxia, it responds to low oxygen via behavioral, physiological, and molecular changes, and one member of the HIF family, HIF2α, has been previously described. Herein, cDNA sequencing, phylogenetic analyses, and genomic approaches were used to determine other members of the HIFα family from *F. heteroclitus* and their relationships to HIFα subunits from other vertebrates. *In vitro* and cellular approaches demonstrated that full-length forms of HIF1α, 2α, and 3α independently formed complexes with the β subunit (ARNT) to bind to hypoxia response elements and activate reporter gene expression. Quantitative PCR showed that HIFα mRNA abundance varied among organs of normoxic fish in an isoform-specific fashion. Analysis of the *F. heteroclitus* genome revealed a locus encoding a second HIF2α, HIF2αb, a predicted protein lacking oxygen sensing and transactivation domains. Finally, sequence analyses demonstrated polymorphism in the coding sequence of each *F. heteroclitus* HIFα subunit, suggesting that genetic variation in these transcription factors may play a role in the variation in hypoxia responses among individuals or populations.

Keywords: Environmental adaptation, oxygen, gene expression

INTRODUCTION

The hypoxia inducible transcription factors (HIFs) play key roles in the regulation of gene expression in animals during normal development and physiology, as well as in several human pathologies associated with low tissue oxygen (20, 23, 46). These heterodimeric transcription factors are composed of α and β subunits, both of which are basic-helix-loop-helix PER-ARNT-SIM (bHLH-PAS) transcription factors. During normal oxygen levels (normoxia), the cellular abundance of the α subunit is kept low, primarily due to oxygen-dependent proteolysis signaled for by the modification of specific proline residues by prolyl hydroxylase domain proteins [PHDs; (20)]. The ability of HIF to activate gene expression is also suppressed during normoxia by hydroxylation of an asparagine residue of the α subunit (31). Both proline and asparagine hydroxylation are inhibited by low oxygen, resulting in an increase in the cellular abundance of the α subunit. The β subunit, previously identified as the aryl hydrocarbon receptor nuclear translocator (ARNT), is present during both reduced and normal oxygen levels and serves as a dimerization partner for other transcription factors, namely the aryl hydrocarbon receptor (37). Thus, at low oxygen tensions, HIFα accumulates, dimerizes with ARNT, binds to specific DNA sequences in target genes (hypoxia-response elements, HRE), and together with other accessory proteins, activates gene expression (20, 46).

The regulation, tissue expression, and gene targets of HIF are most thoroughly described in mammals, which have three genes encoding different forms of HIFα. Semenza and Wang (62) originally described HIF1α as a regulator of erythropoietin (*EPO*) gene expression during hypoxia in mammalian cell culture. Subsequently, HIF1α was found to be widely expressed in different mammalian cell types and tissues and to affect the expression of dozens, if not hundreds, of genes (20, 46). HIF2α, also known as endothelial PAS-domain protein-1 (EPAS-1),

has a more restricted distribution, but also plays a central role in the molecular response to low oxygen (23). HIF3α was originally described by Gu et al. (16) as an oxygen-dependent transcriptional activator, although truncated forms resulting from differential splicing repress the activity of HIF1α and HIF2α during hypoxia (7, 32, 33, 36). There are target genes regulated specifically by HIF1α, 2α, or 3α, as well as some genes whose expression is altered by multiple forms of HIF (41, 60, 75). Together, HIFs regulate the expression of genes involved in multiple cellular processes including erythropoiesis, angiogenesis, carbohydrate transport and metabolism, iron metabolism, and mitochondrial metabolism and autophagy (7, 20, 23, 46).

Fish encounter marked reductions in ambient oxygen in a variety of habitats (48), and various species respond to low oxygen through a suite of morphological, behavioral, and physiological adjustments. Because changes in gene expression may underlie some of these responses (9, 12, 35, 53, 67), it is of interest to understand the distribution of HIFs in fish and their potential roles in regulating gene expression in response to aquatic hypoxia. As in other vertebrates, there are multiple forms of HIFα in fish. Although fish HIF1α and HIF2α appear to be orthologous to their mammalian counterparts, the origins of HIF3α are less clear, leading some studies to refer to the third fish α subunit as HIF4α or HIF1α-like (29, 53). Moreover, carps and related species have duplicate copies of all three HIFα genes, a result, potentially, of an ancient teleost-specific genome duplication (57). Considerable variability exists among fish in the oxygen-dependence, tissue-specificity, and regulatory mechanism (transcriptional or post-translational) of HIFα accumulation, likely reflecting the ecological and phylogenetic diversity of the species studied, as well as differences in experimental design (e.g., life history stage, conditions of hypoxic exposure).

In the current study, we use the mummichog or Atlantic killifish, *Fundulus heteroclitus*, to further explore HIFα in fish. This species is widely distributed throughout estuaries along the Atlantic coast of North America and an ideal system for the study of teleost responses to physio-chemical stressors, including low oxygen (3, 61). Habitats occupied by *F. heteroclitus* may become hypoxic on daily, tidal, or seasonal time scales (64) and this species tolerates lower levels of oxygen than many other common marsh fishes (70). Exposure to low oxygen leads to increased blood oxygen transport (14, 65), altered tissue enzyme activities (13), restricted growth (51, 65), and changes in aerobic and anaerobic metabolism (2, 5, 6). A full-length form of HIF2α (hereafter referred to as HIF2αa; see below) has been sequenced from *F. heteroclitus* (44), and the promoter of the lactate dehydrogenase-B (*Ldh-B*) gene of *F. heteroclitus* contains a novel, non-canonical HRE (50). In addition, there is a draft genome sequence for this species, allowing genomic analyses that are not possible with many other species (52). Finally, *F. heteroclitus* belongs to the euteleostei, a group that comprises about two-thirds of the approximately 24,000 teleost fishes that diversified after the split leading to the Otocephala [herrings, carps, tetras, catfish, and related species (43)]. Hence, study of *F. heteroclitus* may provide insights into the biology of fishes that might differ from conclusions based upon fish models that have duplicated HIFα genes (zebrafish, catfish, and carp).

The specific objectives of this study were (1) to sequence HIF1α and HIF3α from *F. heteroclitus*; (2) to query the *F. heteroclitus* genome for other *HIFα* genes; (3) to assess DNA binding and transactivation of gene expression by HIF1α, HIF3α, and the previously described HIF2αa proteins; and (4) to measure the transcript abundance for these HIFα subunits in tissues of normoxic fish. In addressing these objectives, we also documented sequence variation in the

HIFα subunits of *F. heteroclitus* and identified a short-form of HIF2α, HIF2αb, in the *F. heteroclitus* genome.

MATERIALS AND METHODS

Animals. Fundulus heteroclitus were collected with minnow traps from the salt marshes surrounding Scorton Creek, MA (41° 45' N, 70° 26' W) and transported to Woods Hole Oceanographic Institution, Woods Hole, MA. Fish were kept in aerated, filtered sea water at ambient temperature (ca. 21°C) and fed once a day. Fish were euthanized with an overdose of MS-222 (1 g/l) buffered with sodium bicarbonate (4 g/l). Tissues were rapidly dissected, snap frozen in liquid nitrogen, and stored at – 80°C. Animal care and handling were approved by the IACUCs at the University of New Orleans and Woods Hole Oceanographic Institution.

Cloning and sequencing of HIF1α and HIF3α. The liver from a single *F. heteroclitus* was homogenized in RNA STAT-60 (Tel-Test, Inc) and total RNA was prepared according to the manufacturer's directions. Messenger RNA was purified from 400 μg total RNA with MicroPoly(A) Purist (Ambion) and 1 μg mRNA was used as template for cDNA synthesis and rapid amplification of cDNA ends (RACE) using a Clontech Marathon cDNA-Amplification kit (BD Biosciences). All PCR primers are given in Table 1. For HIF1α, gene-specific primers for RACE were based upon an internal HIF1α fragment of approximately 920 bp amplified using primers (HIF1-Forward and HIF1-Reverse) derived from rainbow trout HIF1α (AF304864). For HIF3α, gene-specific RACE primers were designed based upon a partial HIF-like sequence (AF433668).

Both 5' and 3' RACE used two rounds of "nested" PCR. Briefly, the first round PCR used a gene-specific (gs) primer (5' gs outer, 5' gs alt, 3' gs outer, or 3' gs alt) and adaptor

primer 1 (AP 1) in a "touchdown" PCR protocol [30 s at 94°C; 5 cycles of 5 s at 94°C and 4 min at 72°C; 5 cycles of 5 s at 94°C and 4 min at 70°C; 25 cycles of 5 s at 94°C and 4 min at 68°C; 7 min at 72°C (HIF1α) or 7 min at 68°C (HIF3α)]. Second round PCR used a gene-specific primer (5' gs inner, 5' gs alt, 3' gs inner, or 3' gs alt) with adaptor primer 2 (AP 2) in a PCR program of 30 s at 94°C; 20 cycles of 5 s at 94°C and 2 min at 68°C; 7 min at 68°C. For 3'RACE of HIF1α, the second round PCR program was 30 s at 94°C; 20 cycles of 5 s at 94°C, 30 sec at 65°C, 2 min at 68°C; 7 min at 68°C. RACE products were gel purified, ligated into pGemT-Easy (Promega), and transformed into *E. coli* JM109 high-efficiency competent cells. Multiple positive clones of each product were sequenced by the University of Maine Sequencing Center using primers against vector sequences. The resulting sequences were aligned and used to design primers specific to the 5'- and 3' untranslated regions of HIF1α and HIF3α (Table 1).

The full-length *F. heteroclitus* HIF1α cDNA was amplified from the original cDNA using HIF1 5' UTR and HIF1 3'UTR primers and a PCR program of 30 s at 94°C; 35 cycles of 10 s at 94°C, 30 sec at 62°C, 3 min at 68°C; 7 min at 68°C. Full-length HIF3α cDNA was amplified using HIF3 5' UTR and HIF3 3'UTR primers and a PCR program of 30 s at 94°C; 35 cycles of 10 s at 94°C, 30 sec at 65°C, 3 min at 68°C; 7 min at 68°C. Advantage 2 DNA polymerase (BD Biosciences) was used for all RACE and full length PCR. PCR products were gel-purified, cloned and sequenced as stated above for RACE products. Additional sequencing primers were synthesized as needed to sequence the entire insert. A minimum of eight clones, representing multiple PCR runs, were sequenced in both directions for each HIFα.

Sequence alignment and phylogenetic analysis. F. heteroclitus HIF1α and HIF3α sequences were assembled and inspected for sequence quality and potential PCR errors. When a single clone differed from all other sequences at a single nucleotide position, this was considered

to be PCR error, and that position was manually edited to represent the consensus nucleotide at that position. This process resulted in "correcting" 71 nucleotides in a total of more than 68,000 positions. When two or more clones had the same nucleotide at a given position and differed from the rest of the sequences, this was considered to be a polymorphism, and these sequences were retained as putative HIF1α and HIF3α variants.

For phylogenetic analyses, HIF1α, HIF2α, and HIF3α sequences were downloaded from NCBI (www.ncbi.nlm.nih.gov/Genbank) and Ensemble (www.ensembl.org). Multiple alignment of deduced amino acid sequences was performed with ClustalX (www.ebi.ac.uk/Tools/clustalw2), and phylogenetic analyses were conducted in MEGA6 (66). The relationship among HIF sequences was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (19). Initial trees for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using a JTT model. The analysis involved 33 amino acid sequences. The region corresponding to residues 1-360 of human HIF1α from each HIF sequence was used for the alignment and all positions containing gaps and missing data were eliminated, resulting in a total of 227 positions in the final dataset.

Antibody production. Regions of HIFα cDNAs corresponding to peptides of approximately 10 kDa in the C-terminal half of *F. heteroclitus* HIF1α (amino acids 407-510), HIF2αa (501-636), and HIF3α (530-647) were either cut from the pcDNA construct with *Pst*I and *Sal*I or amplified by PCR (see Table 1 for primer sequences). Products were ligated into pET-Duet vector (Novagen), and constructs were sequenced to confirm in-frame fusion. *E. coli* BL21 (DE3) cells (Novagen) were transformed, induced with 1 mM IPTG for 1.5 h at 37°C, collected by centrifugation, and lysed by three passages through a French press at 138 MPa in 20

mM Tris, pH 8.0; 50 mM NaCl; 50 mM KCl plus 1% Triton X-100. Lysates were centrifuged at 10,000 x *g* for 15 min at 4°C and the HIFα peptides were purified by chromatography on Ni-NTA agarose (Qiagen). Antibodies were generated in chickens using 1 mg of each purified fusion protein as immunogens (Aves Labs). The resulting IgYs were purified from eggs.

Electrophoretic mobility shift assays (EMSA). To generate full-length proteins for EMSA, inserts encoding *F. heteroclitus* HIF1α and HIF3α were released from pGEM-T Easy by digestion with *Pst*I and *Apa*I or *Nsi*I and *Apa*I, respectively, and independently ligated into pcDNA 3.1/Zeo (+) (Life Technologies) digested with *Pst*I and *Apa*I. The resulting plasmids were used along with plasmids encoding *F. heteroclitus* HIF2αa and ARNT2 (45) to make *in vitro* transcribed and translated (IVTT) proteins using a rabbit reticulocyte lysate system (Promega). The EMSA protocol was modified from (62). Each reaction (20 µl final volume) consisted of 10 mM Tris (pH 7.5), 50mM KCl, 50 mM NaCl, 5% glycerol, 5 mM DTT, 1 mM $MgCl_2$, 1 mM EDTA, and 100 ng calf thymus DNA. Binding reactions usually included 1 µl IVTT *F. heteroclitus* ARNT2 and 1 µl IVTT *F. heteroclitus* HIF1α, HIF2αa, or HIF3α. Some reactions contained 1 µl IVTT ARNT2 and 3 µl IVTT HIF2αa because of low yields of HIF2αa in IVTT. Reactions to assess non-specific binding included 2 - 4 µl of IVTT ARNT2, IVTT luciferase (Promega), or IVTT reactions primed with empty pcDNA 3.1/Zeo (+). Competitor and probe DNA sequences were based upon the human *EPO* 3' enhancer (62) or a hypoxia response element in the promoter of the *F. heteroclitus Ldh-B* gene (50). Probe DNA was made by end-labeling single-stranded oligonucleotides using T4 Polynucleotide kinase (Promega) and [γ-^{32}P]ATP (PerkinElmer), annealing with equimolar amounts of complementary strand, and desalting on ProbeQuant G-50 microcolumns (GE Healthcare). Incorporation of [^{32}P] was assessed by liquid scintillation counting (BeckmanCoulter). Labeled probe was included at 40

196 fmol and approximately 10^5 cpm in each reaction. When used, double stranded competitor DNA
197 at 100 to 800-fold molar excess or antibodies against HIFα subunits (see above) or ARNT (MA-
198 515; ThermoFisher) were pre-incubated in the reaction mixture for 15 min at room temperature
199 prior to the addition of probe. After probe addition, reactions were incubated a further 15 min,
200 transferred to 4°C, and electrophoresed on 5% polyacrylamide gels (37.5:1, acrylamide:bis-
201 acrylamide) in 0.3 x Tris-borate-EDTA buffer (TBE) at 15 mA for 1.5 h at 4°C. Gels were dried
202 and imaged by phosphor imaging (Bio-Rad). Band density was quantified using Quantity One
203 (Bio-Rad).

204 *HIFα over-expression and reporter gene expression.* The ability of HIF proteins to
205 activate transcription was determined by transient transfection into COS-7 monkey kidney cells,
206 an established cell line that has been used previously to assess HIF function (21). COS-7 cells
207 (American Type Culture Collection) were grown in high-glucose Dulbecco's modified Eagle
208 medium (DMEM; GIBCO) containing 10% fetal bovine serum (FBS; HyClone) in 5% CO_2 at
209 37°C as described earlier (22). Transient transfection of cells was used to assess the effect of
210 increasing doses of plasmids encoding each *F. heteroclitus* HIFα on the expression of firefly
211 luciferase reporter plasmids, N443pGL3 (50), which has the promoter of *F. heteroclitus Ldh-B*
212 including a characterized HRE upstream of the luciferase gene, and p3XHifREluc (11), which
213 has three copies of the HRE from the mouse heme oxygenase gene upstream of the luciferase
214 gene. All experiments included a plasmid encoding *Renilla* luciferase under the control of the
215 thymidylate kinase promoter (pRL-TK; Promega) to control for transfection efficiency.

216 One day prior to transfection, cells were plated at 10^5 cells per well in 12-well plates.
217 Replicate wells were transfected in 1 ml serum-free DMEM containing Lipofectamine 2000
218 (3.3:1 ratio of µl reagent to µg plasmid; Invitrogen), 500 ng reporter (N443pGL3 or

p3XHifREluc), 10 ng TK-pRL, 400 ng *F. heteroclitus* ARNT2 plasmid, and 0 to 160 ng of *F. heteroclitus* HIF1α, HIF2αa, or HIF3α plasmid (pcDNA constructs from above). Empty pcDNA (up to 160 ng) was included to balance the total plasmid DNA concentration in each transfection. Transfection mixes were removed after 4 h and replaced by DMEM containing 10% FBS. Cells were grown for an additional 24 h at which point they were harvested in Passive Lysis Buffer (Promega). Firefly and *Renilla* luciferase activities were measured in cell lysates using the Dual-Luciferase Reporter Assay (Promega). The relative luciferase for each well was calculated as firefly luciferase divided by *Renilla* luciferase. Fold-induction was then calculated as the ratio of relative luciferase of wells transfected with HIFα plasmids to wells transfected with empty pcDNA. Replicate wells were averaged and treated as a single determination for statistical purposes, and the entire experiment was repeated three times (n = 3). Other experiments were done in which cells were transfected with lower amounts of HIFα plasmids (0 – 8 ng) or with 80 ng each of two HIFα plasmids (1α + 2αa; 1α + 3α; 2αa + 3α).

In one experiment, COS-7 cells transfected with 0-160 ng of each *F. heteroclitus* HIFα plasmid were lysed in 100 µl SDS-PAGE sample buffer (28) and used in western blots for HIF1α, HIF2αa, HIF3α, or ARNT2. Samples (40 µl) were electrophoresed on 10% polyacrylamide gels (28) and transferred to PVDF membranes (68). Positive controls (IVTT proteins from *F. heteroclitus*) were included in each gel. Chicken polyclonal antibodies against *F. heteroclitus* HIF 1α, 2α, and 3α and mouse monoclonal antibody against human ARNT (MA-515, ThermoScientific) were used at a dilution of 1:1000. Horseradish peroxidase-conjugated secondary antibodies (Thermo Scientific) were used at 1:5000. Blots were developed by chemiluminescence and imaged in a ChemiDoc XRS (Bio-Rad).

241 *Quantitative PCR.* Total RNA was isolated from frozen tissues (liver, white skeletal
242 muscle, intestine, gill, kidney, brain, heart, spleen, ovary, and testes) of six normoxic *F.*
243 *heteroclitus* (three females, three males) as described above. Gill and intestine were first allowed
244 to thaw in RNA later-ICE (Ambion) to allow dissection of gill lamella away from the gill arch or
245 to flush the contents of the intestine. After isolation, RNA quality was assessed by Experion
246 capillary electrophoresis (Bio-Rad) and RNA quantity was determined by spectrophotometry.
247 Two RNA samples from gill and one from kidney were removed from this analysis due to
248 extensive RNA degradation.
249 For all samples, 1 µg of RNA was reverse transcribed into cDNA using iScript reverse
250 transcriptase (Bio-Rad) in a final volume of 20 µl. One microliter of the product (equivalent to
251 50 ng of total RNA) was used for quantitative PCR using iQ SYBR Green Supermix (Bio-Rad)
252 and primers specific for HIF1α, HIF2αa, or HIF3α (Table 1) in a reaction volume of 20 µl. One
253 primer of each primer pair was designed to span a putative exon-exon boundary (determined by
254 comparison of *F. heteroclitus* HIFα cDNAs with their homologs in the *Fugu* genome), and each
255 primer pair was designed to amplify a product of 100 to 150 bp. PCR was conducted in a 96-well
256 format in a MyiQ or iQ5 thermocylcer (Bio-Rad) using the following program: 3 min 95°C; 40
257 cycles of 15 s at 95°C and 1 min at 68°C; 10 min at 72°C. Melt curves were done for each
258 sample (81 30-s steps of 0.5°C from 55 to 95°C) to ensure the presence of a single amplification
259 product in each reaction. Selected reactions were sequenced and shown to be the anticipated
260 HIFα amplimers. No amplification was observed in PCR of samples that were not reverse
261 transcribed or in reactions lacking template.
262 All samples were analyzed in duplicate. Each plate also contained duplicate wells of four
263 serial dilutions of plasmid encoding the appropriate HIFα from *F. heteroclitus* in order to assess

the efficiency of PCR, which ranged from 90-105%. Furthermore, a reference cDNA sample was made by pooling an equal volume of all cDNA samples and included in duplicate wells in every plate in order to correct for run-to-run variation in PCR amplification. For every reaction, the threshold cycle (C_t) was determined as fluorescence increased above a user-determined background (set to 160). Transcript copy number was determined from standard curves of known concentrations of plasmids encoding each HIFα.

Survey of HIFα sequence variation. An RNA-seq data set was searched to identify single nucleotide polymorphisms (SNPs). The data set, which will be described in detail in a separate publication, was obtained from 200 *F. heteroclitus* F2 embryos from multiple F1 parents originating (F0) from Scorton Creek, MA. The RNA was sequenced using the Illumina HiSeq 2500 platform (100 bp paired-end reads), and sequence data (representing 692 x 10^6 reads) were aligned to the draft *F. heteroclitus* genome (version 2b) using Tophat version 2.0.9 (69). Alignments were visualized using IGV (55) and *HIFα* loci were analyzed for variants; only variants with a frequency ≥ 2% of reads were included.

Statistical analyses. Analysis of variance (ANOVA) was used to determine effects of oligonucleotide sequence on protein binding in EMSA, the effects of HIFα plasmid type and concentration on reporter gene expression, and the effects of organ and HIFα gene on mRNA expression. mRNA copy number was log10 transformed prior to analysis to meet assumptions of ANOVA. Significant ANOVA results were followed by Bonferroni's post-hoc comparisons. *P* values of 0.05 or less were considered to be statistically significant. Statistical analyses were done in Prism version 5.0a (Graphpad Software Inc) or SYSTAT13 (SYSTAT Software, Inc).

RESULTS AND DISCUSSION

Sequence and domain analysis of F. heteroclitus HIF1α and HIF3α. cDNA sequencing revealed two *F. heteroclitus* HIF1α variants, which were 2442 or 2448 bp in length and included 129 bp of 5'-untranslated region, open reading frames of 2259 or 2265 bp, and 54 bp of 3'-untranslated region. The predicted proteins are 753 or 755 amino acids in length (calculated molecular masses of 84.0 and 84.2 kDa) and contain specific functional and structural domains that are conserved among vertebrate HIF1α proteins (Fig. 1A). The predicted proteins have an N-terminal basic helix-loop-helix (bHLH) domain, PAS A and PAS B domains, a PAS-associated C-terminal (PAC) motif, and an oxygen-dependent degradation domain (ODD). Two putative transactivation domains (TADs) were found, one overlapping with the last 50 amino acids of the ODD (TAD-N) and one at the extreme C-terminus of the deduced protein (TAD-C). Over the entire length, the deduced amino acid sequence of *F. heteroclitus* HIF1α shows high identity with HIF1α from other fishes (e.g., 72% amino acid identity with tilapia HIF1α) and tetrapod vertebrates (e.g., 47% identical to human HIF1α).

Two *F. heteroclitus* HIF3α variants, differing by a single synonymous nucleotide substitution, were also sequenced. These were 2178 bp in length and included 35 bp of 5'-untranslated region, an open reading frame of 1986 bp, and 157 bp of 3'-untranslated region. The predicted protein is 662 amino acids (calculated molecular mass of 73.7 kDa) and has a bHLH domain, PAS A and PAS B domains, a PAC domain, an ODD, and one TAD overlapping with the ODD (Fig. 1C). The deduced amino acid sequence of *F. heteroclitus* HIF3α ranges from 71% identical to tilapia HIF3α to 36% identical to human HIF3α. Four DNA sequences (two variants each of *F. heteroclitus* HIF1α and HIF3α) have been submitted to GenBank (accession numbers KR703588, KR703589, KR703597, and KR703598). A previously-described HIFα-like partial

cDNA (AF433668) is now identified as being the 3' end of *F. heteroclitus* HIF3α. The remaining characterization focused on one variant of each subunit, HIF1α*1 (Accession KR703588) and HIF3α*1 (Accession KR703597), hereafter referred to FhHIF1α and FhHIF3α, respectively.

The consensus prolyl hydroxylation motif is the hexameric sequence LXXLAP (8), and within the putative ODDs of each predicted FhHIFα subunit there are multiple potential sites for prolyl hydroxylation (Figs. 1B and 1D). In FhHIF1α, P429 occurs in an LXXLAP motif that is highly conserved among vertebrates; P561 aligns with putative prolyl hydroxylation sites in other vertebrates, but occurs in the motif LXXFAP; and P576 occurs in the canonical LXXLAP sequence in FhHIF1α but aligns with a region of HIF1α that is poorly conserved among vertebrates (Fig 1B). In FhHIF3α, there are two potential prolyl hydroxylation sites, P414 and P522 (Fig. 1D). The more N-terminal of these two sites is only found in fishes: in zebrafish, grass carp, and pufferfish, this proline occurs in the motif LXXLAP, but in FhHIF3α, P414 occurs in LXXLEP. The second proline, P522, is conserved across vertebrate species where it occurs in the LXXLAP motif, except in FhHIF3α, where it is found in an LXXYAP motif.

In vitro studies of the substrate specificity of human PHD enzymes have shown that peptides having a range of amino acid substitutions in and around the LXXLAP motif are hydroxylated, albeit at variable rates (18, 30). Specifically, peptides with the L to F substitution predicted in FhHIF1α and the L to Y substitution predicted in FhHIF3α were hydroxylated as efficiently as the consensus sequence, while a peptide with the A to E substitution predicted in the first putative hydroxylation site in FhHIF3α was a poor substrate (30). In a recent study of zebrafish HIF3α, Zhang et al. (75) showed that the second of these sites, rather than the first, was critical in determining the stability of zebrafish HIF3α. In addition, the predicted amino acid sequence immediately following this second putative hydroxylation site (YISMDDDFQL) is

identical in *F. heteroclitus* and zebrafish, including a leucine (L532 in FhHIF3α and L503 in zebrafish HIF3α) that also contributes to HIF3α stability in zebrafish (75). While these results strongly suggest that FhHIF1α and FhHIF3α are targets of hydroxylation, which proline residue(s) are hydroxylated, and the conditions under which they are modified, have yet to be experimentally determined in any fish HIFα subunit.

There is also a putative target of asparaginyl hydroxylation, N730, in FhHIF1α (Fig 1A), which occurs in the CEVN motif conserved in the N-terminus of vertebrate HIF1α. In contrast, FhHIF3α lacks a second TAD and an asparaginyl hydroxylation site near the C-terminus (Fig 1C), in accordance with other vertebrate HIF3α subunits (7).

Phylogenetic analysis of F. heteroclitus HIFα proteins. Phylogenetic analyses of the deduced amino acid sequences of FhHIF1α and FhHIF3α, along with the previously described FhHIF2αa (44), grouped the *F. heteroclitus* HIFα proteins with their orthologs from other vertebrates (Fig. 2). Bootstrap analyses strongly support three clades, as well as the branching patterns within each clade (e.g., HIF1α) that are concordant with vertebrate phylogeny (e.g., tetrapod vertebrates forming a separate group from fishes). Rytkönen et al (57) described two forms of each *HIFα* gene in cyprinid fishes, which include zebrafish, grass carp, and asp, among others. They speculated that the three *HIFα* genes were duplicated as part of the ancestral teleost-specific whole genome duplication (39), followed by loss of one of each duplicate in most, but not all, non-cyprinid lineages. The phylogenetic tree shown in Fig. 2 clearly places the *F. heteroclitus* HIFα proteins with those common to all euteleosts rather than with the cyprinid-specific forms.

In addition to HIF1α, 2α, and 3α, searching the *F. heteroclitus* genome revealed a second *HIF2α* locus with a shorter coding region. The deduced amino acid sequence of this HIF2α

aligns with the "relic" duplicated HIF2α found in several non-cyprinid fishes (57), although bootstrap support for the phylogenetic placement of these short forms is low due to their divergent sequences. This predicted protein, previously deposited to GenBank with the description endothelial PAS domain containing protein-1 (XP_012709515.1), is 367 amino acids in length, contains two PAS domains, but has a poorly conserved bHLH domain and lacks other domains associated with HIFα subunits (i.e., ODD and TAD). This deduced protein is hereafter referred to as FhHIF2αb, with the original HIF2α (44) corresponding to FhHIF2αa.

Genomic organization of F. heteroclitus HIFα loci. All four *HIFα* loci were characterized in the *F. heteroclitus* genome. The identity and order of genes flanking the killifish HIFα genes (Fig. 3A) are similar to those found adjacent to the orthologous *HIFα* genes in the other fish genomes (10, 57, 59, 74), supporting the groupings based upon amino acid sequence analysis. The genes immediately flanking *FhHIF1α* are perfectly conserved in three euteleost genomes searched (green pufferfish, stickleback, tilapia), whereas gene order is less well conserved for *FhHIF2αa*, *FhHIF2αb*, and *FhHIF3α*. As predicted by their phylogeny, the gene order is more similar among euteleost genomes than between any of these and the zebrafish genome.

The *FhHIF1α* gene has 15 exons distributed across approximately 15 kb (Fig. 3B). The *FhHIF2αa* gene has 16 exons over 72 kb, while the *FhHIF2αb* gene has 10 exons over 13kb. The *FhHIF3α* gene has 15 exons, but the length of the *FhHIF3α* gene could not be deduced due to the presence of three introns of unknown length. The number of exons and exon junctions of the killifish *HIF* genes are conserved when compared to those of *HIFα* genes in humans and other fishes (10, 74). No other *HIFα* genes or pseudogenes were found in the *F. heteroclitus* genome.

DNA binding by F. heteroclitus HIF. FhHIF1α and FhHIF3α proteins made by IVTT, along with IVTT HIF2αa from *F. heteroclitus* (44), were used in EMSA in the presence of *F. heteroclitus* ARNT2 to assess their ability to bind specific HREs (Figs. 4 and 5). The first oligonucleotide probe used was the HRE present in the human *EPO* 3' enhancer, EPO-w18 (62) (Fig. 4A). A band (arrows in Fig. 4B-D) was observed in EMSA using this probe when each of the three *F. heteroclitus* HIFα proteins was included in the reactions [see (-) and (+) lanes in Fig. 4B-D]. The intensity of this band was reduced in a dose-dependent fashion when 100 to 800 molar excess of unlabeled wild-type oligonucleotide (EPO-w18) was used as a competitor. An oligonucleotide mutated at three positions in the HRE (EPO-m18) was not as effective a competitor at the same molar ratios. Fig. 4E shows the proportion of labeled probe bound by FhHIFs in the presence of 400-fold molar excess of unlabeled competitor oligonucleotide, relative to the band intensity in the absence of competitor DNA (n = 4-7). For all three *F. heteroclitus* HIFα proteins, DNA binding was reduced by a statistically greater extent by EPO-w18 compared to EPO-m18 (black and cross-hatched bars, Fig. 4E) indicating that each forms a sequence-specific complex with the canonical, mammalian HRE.

The *F. heteroclitus Ldh-B* promoter has a putative HRE that is characterized by two potential HIF-binding sites (HBSs) arranged as an inverted repeat spaced by eight nucleotides (50). In addition, both sites have the sequence A<u>T</u>GTG rather than the consensus sequence of A<u>C</u>GTG, differing at one position thought to be critical for HIF binding (71). Unlabeled oligonucleotides corresponding to the wild-type sequence of each HBS (HBS1-wt and HBS2-wt; Fig. 4A) were assessed for their ability to compete with EPO-w18 for binding to *F. heteroclitus* HIF. Both oligonucleotides reduced the intensity of the specific HIF-DNA band, but not as effectively as unlabeled EPO-w18 at the same molar ratios (Fig. 4B-D). When competitor DNA

was used in 400-fold molar excess of labeled probe, DNA binding by FhHIF1α was reduced by HBS1-wt and HBS2-wt to a level intermediate to that observed by EPO-w18 and EPO-m18, while HBS1-wt and HBS2-wt were similar to EPO-m18 in their effect on FhHIF2αa and FhHIF3α DNA-binding (Fig. 4E). These results suggest that *F. heteroclitus* HIFs bind better to the consensus HRE found in the human *EPO* 3' enhancer than they do to either HBS from the *F. heteroclitus Ldh-B* promoter.

The interaction between *F. heteroclitus* HIFs and the putative HRE from the *F. heteroclitus Ldh-B* promoter was further explored using an oligonucleotide probe containing both HBS1 and HBS2 (Fig. 5A). The three *F. heteroclitus* HIFα proteins independently formed a complex with *F. heteroclitus* ARNT2 to bind to this oligonucleotide (arrows in Fig. 5B-D). EMSA reactions with FhHIF3α showed two bands suggesting different oligomeric states. These protein-DNA complexes required the α subunit [compare (-) and (+) lanes], and they were competed by 100- to 800-fold molar excess of unlabeled oligonucleotide having the wild-type sequence at both HBS1 and HBS2 (wt/wt). When both HBSs were mutated, the resulting double mutant (mut/mut) was not as effective a competitor. When HBS1 and HBS2 were individually mutated, the resulting oligonucleotides (wt/mut and mut/wt) competed as effectively as the wild-type oligonucleotide for FhHIF1α (Fig. 5B) and FhHIF3α (Fig. 5D). For FhHIF2αa, the mut/wt oligonucleotide was intermediate between the wt/wt and mut/mut oligonucleotides in reducing the band intensity, whereas wt/mut appeared to be as effective as wt/wt probe (Fig. 5C). Results of replicate experiments at 400-fold molar excess (Fig. 5E; n = 3-4) support the conclusion that the two binding sites are equally effective in binding to *F. heteroclitus* HIF complexes, with the possible exception of FhHIF2αa, which may bind to HBS1 better than HBS2.

To ensure that the bands observed in EMSA were the result of heterodimers of HIFα and ARNT2, super-shift reactions were carried out with antibodies generated against each FhHIFα subunit or ARNT (Fig. 6). For EMSA reactions with IVTT FhHIF1α, FhHIF2αa, and FhHIF3α, the specific band was either slowed or abolished by pre-incubation with chicken polyclonal antibodies against the respective α subunit or by a mouse monoclonal antibody against human ARNT that recognizes *F. heteroclitus* ARNT2 (44). Pre-immune IgY from the same hens used for HIFα antibody production had no effect on the protein-DNA complexes. Similar results were generated when either the human *EPO* HRE or the *F. heteroclitus Ldh-B* HRE were used as probes. The super-shift experiments confirm that the protein-DNA complexes include the respective FhHIFα subunit and ARNT.

The above results demonstrate that, in association with *F. heteroclitus* ARNT2, the three FhHIFα proteins bind to a consensus HRE, extending earlier observations with *F. heteroclitus* HIF2α (44). In addition, all three are also capable of binding to a non-canonical HRE, previously described in the promoter of *F. heteroclitus Ldh-B* (50). HIF complexes appeared to bind to the non-canonical HRE, having the core sequence A<u>T</u>GTG, with lower affinity than they did to a consensus HRE (A<u>C</u>GTG), as evidenced by these oligonucleotides being less effective competitors than EPO-w18 oligonucleotide (Fig. 4), but they nevertheless bound specifically (Fig. 5). While the physiological role of this non-canonical sequence remains to be elucidated, a recent analysis of DNA sequences immunoprecipitated with a HIF1α antibody indicated that a considerable fraction of HREs in zebrafish have the sequence A<u>T</u>GTG (15). That observation, combined with the present results and those of Rees et al. (50), support the idea that, at least in fish, HIF may regulate the expression of genes that contain HREs that differ from the mammalian consensus HRE. Importantly, this non-canonical HRE is characterized by two

elements that exist in a perfect inverted repeat, spaced by 8 nucleotides, nearly identical to the arrangement of enhancers described for several mammalian genes that are bound by HIF (24). In these mammalian genes, one binding site binds to HIF and the other site binds to another nuclear factor, but both are required for hypoxic gene expression. The present results argue that *F. heteroclitus* HIF1α and 3α bind to both sites equally well, but *F. heteroclitus* HIF2αa binds more strongly to HBS1.

FhHIFα proteins drive reporter gene expression in COS-7 cells. Mammalian COS-7 cells were transiently transfected with increasing amounts of plasmids (0 – 160 ng) encoding FhHIF1α, FhHIF2αa, or FhHIF3α, along with *F. heteroclitus* ARNT2 to assess the ability of their protein products to drive expression of a luciferase gene under the control of the *F. heteroclitus Ldh-B* promoter (50). All FhHIFα proteins drove reporter gene expression in a dose-dependent fashion (Fig. 7A). The fold-increase in luciferase expression compared to cells transfected with empty pcDNA was significantly greater when cells were transfected with FhHIF1α than when they were transfected with FhHIF2αa or FhHIF3α, at all amounts of plasmid (2-way ANOVA, $P < 0.001$). To assess reporter gene expression at lower concentrations of the HIFα plasmids, cells were transfected with 0 – 8 ng of each HIFα plasmid. As observed at higher amounts of plasmid, FhHIF1α induced greater reporter gene expression at these plasmid concentrations than did FhHIF2αa or FhHIF3α (Fig. 7B).

The above results were from experiments conducted with a reporter gene under the control of the *F. heteroclitus Ldh-B* promoter. In one experiment, the reporter plasmid encoded the luciferase gene under the control of three copies of the HRE from the mouse heme oxygenase gene [p3XHifREluc; (11)]. Transfection with 0 – 160 ng of FhHIF1α plasmid resulted in greater induction of luciferase expression compared to equivalent amounts of FhHIF2αa or FhHIF3α

plasmid, the same result as seen when the reporter plasmid was under the control of the *F. heteroclitus Ldh-B* promoter (Fig. 7C). These results further illustrate that *F. heteroclitus* HIF binds to both the consensus and non-canonical HRE to drive reporter gene expression, with the amount of reporter gene expression depending upon the HIFα isoform.

Two possible mechanisms of the greater induction observed for FhHIF1α compared to FhHIF2αa or FhHIF3α are (a) greater protein expression from the FhHIF1α plasmid or (b) greater transactivation of reporter gene expression at equivalent HIFα protein levels. To evaluate these two possibilities, FhHIFα protein amounts in transiently transfected COS-7 cells were determined by western blotting (Fig. 7D). While protein levels of all three HIFα subunits increased with increasing amounts of plasmid transfected, the immunoreactivity observed using HIF1α or HIF3α antibodies was greater than observed with the HIF2αa antibodies. Assuming equal affinities of the antibodies for their respective antigens, this implies that more FhHIF1α and FhHIF3α protein were expressed at these levels of transfected plasmid compared to FhHIF2αa. Direct visualization of IVTT products *in vitro* also showed less protein expression from the FhHIF2αa plasmid (data not shown). Hence, lower FhHIF2αa protein levels might explain the lower reporter gene expression in cells transfected with FhHIF2αa plasmid (Figs. 7A-C). Because protein levels of FhHIF1α and FhHIF3α appear to be more similar (Fig. 7D), the lower gene expression in cells transfected with FhHIF3α plasmid might be due to poorer transactivation by FhHIF3α compared to FhHIF1α.

The present results show that all three *F. heteroclitus* HIFα proteins act as transcriptional activators, but potentially with differing effectiveness. The present data support earlier observations that HIF1α and HIF3α from grass carp (29) and zebrafish (74) are able to drive reporter gene expression when expressed in mammalian cells. In those studies, HIF1α drove

491 greater reporter gene expression than HIF3α, similar to the pattern described here for *F.*
492 *heteroclitus* HIF1α and HIF3α. Together, these results support the conclusion that HIF1α is a
493 more effective activator of reporter gene expression than HIF3α, which is consistent with
494 deduced protein structure of FhHIF1α having two TADs compared to a single TAD in FhHIF3α
495 (Fig. 1). Although the lower reporter gene expression driven by FhHIF2αa in comparison to
496 FhHIF1α may be the result of relatively poorer protein expression from the HIF2αa plasmid (Fig.
497 7D), differences in transcriptional activity by HIF1α and HIF2α have been reported in bichir and
498 naked carp (4). In those species, which α subunit leads to higher reporter gene expression
499 depends upon the species and whether cells were incubated at normal or reduced oxygen levels.
500 In mammals, shorter forms of HIF3α, resulting from differential splicing, act as
501 repressors of HIF1α mediated gene expression (32, 33, 36). In one experiment, 80 ng of two
502 HIFα plasmids were co-transfected into COS-7 cells (1α + 2αa; 1α + 3α; 2αa + 3α). The
503 simultaneous expression of two HIFα plasmids neither augmented nor diminished luciferase
504 gene expression (Fig. 7E). Under these conditions, therefore, there was no evidence that
505 FhHIF3α reduced the level of reporter gene expression by either FhHIF1α or FhHIF2αa.
506 Importantly, the current experiments were done in the presence of excess ARNT2: there is
507 evidence that HIFα subunits compete for ARNT only when the latter is present in limiting
508 quantities (7, 29).
509 *FhHIFα mRNA in normoxic F. heteroclitus*. The abundance of each HIFα transcript was
510 determined by quantitative RT-PCR of total RNA isolated from organs of normoxic *F.*
511 *heteroclitus* (Fig. 8). One-way ANOVA demonstrated pronounced differences in transcript
512 abundance among organs for all three HIFα transcripts ($P < 0.001$). The abundance of FhHIF1α
513 transcripts was greatest in testes, spleen, skeletal muscle, and kidney, lowest in brain and heart,

514 and intermediate in gill, liver, intestine, and ovary (Fig. 8A). The abundance of FhHIF2αa
515 transcripts was greatest in gill, liver, and intestine, lowest in brain and muscle, and intermediate
516 in other organs (Fig. 8B). The abundance of FhHIF3α transcripts was greatest in spleen and
517 ovary, lowest in brain and heart, and intermediate in other organs (Fig. 8C).

518 For 6 of 10 organs, transcript abundance was significantly different among HIF1α,
519 HIF2αa, and HIF3α (one-way ANOVA; all $P < 0.05$), but the pattern of variation depended upon
520 the organ (compare across panels A-C, Fig. 8). In 4 of 6 organs that showed significant
521 differences among HIFα transcripts, HIF1α, alone or in combination with either HIF2αa or
522 HIF3α, was the predominant transcript. This pattern has been observed in HIFα mRNA and
523 protein levels in other fishes and supports the idea that HIF1α plays an important role in oxygen-
524 regulated gene expression across a broad range of organs, whereas HIF2αa and HIF3α may have
525 more anatomically restricted functions (26, 40).

526 One striking feature of the present data is that *F. heteroclitus* brain and heart have very
527 low levels of all HIFα transcripts. This result contrasts with the relatively high HIFα mRNA
528 levels in brain and heart measured under normoxia in several other fishes (10, 49, 54, 56, 58) and
529 does not support the idea that the levels of HIFα transcripts are highest in organs that have strict
530 aerobic requirements (58). Another intriguing result was the high level of HIF2αa seen in *F.*
531 *heteroclitus* gill. This observation confirms earlier results from semi-quantitative PCR (44),
532 agrees with observations from other species (10, 56, 58), and suggests that HIF2αa may play an
533 important role in gill, an organ responsible for gas exchange, as well as ionic and osmotic
534 regulation. Measures of HIFα protein abundance will be necessary to clarify whether these
535 organ-dependent mRNA levels are coupled to protein abundance and, hence, functional
536 differences.

Intraspecific sequence variation in HIF alphas. As mentioned above, sequencing of *F. heteroclitus* HIF1α and HIF3α cDNAs from a single individual yielded two variants of each. These variants were recovered in multiple clones obtained from multiple PCR runs (see Methods) and, therefore, they were considered to represent two allelic variants of each gene. The *FhHIF1α*1* allele has an A at nucleotide position 108 vs C for *FhHIF1α*2*, and the sequence GAGGAG at positions 1099-1104 in *FhHIF1α*1* is a six nucleotide gap in *FhHIF1α*2*. Both polymorphisms resulted in amino acid differences: L or F at amino acid position 36; and EE or a two amino acid gap at positions 367-368. The *F. heteroclitus HIF3α* alleles differed by a single synonymous substitution at nucleotide position 417 (C in *FhHIF3α*1* or T in *FhHIF3α*2*). Because liver RNA from a single individual served as the template for cDNA sequencing, that individual must have been heterozygous at both loci.

An analysis of the transcriptome of *F. heteroclitus* embryos identified additional single nucleotide polymorphisms (SNPs). Using high-throughput sequencing of pooled mRNA from 200 embryos, 16 SNPs were identified in *HIF1α*, four of which are non-synonymous; 18 SNPs were identified in *HIF2αa*, three of which are non-synonymous; and 10 SNPs were identified in *HIF3α*, five of which are non-synonymous. The locations of the predicted amino acid changes in the corresponding proteins are shown in Table 2. Although much of the variation in amino acid sequence occurs outside of predicted functional domains, at least one substitution in each HIFα occurs in either the bHLH domain or ODD, including one in a potential prolyl hydroxylation motif in HIF3α (position 413; see Fig. 1D). Five out of the 12 substitutions are non-conservative (negative BLOSUM62 scores; Table 2), including examples within predicted functional domains.

559 Although the physiological significance of polymorphism in *F. heteroclitus HIFα* genes,
560 if any, is unknown, it is important to point out that genetic variation in mammalian *HIFα* genes is
561 associated with several physiological traits or pathological conditions. Variation at the *HIF1α*
562 locus has been associated with variation in maximum oxygen consumption of elite athletes (38,
563 47), diabetes prevalence (42), and cancer risk and progression (25, 27, 76). High altitude
564 adaptation in Tibetans is associated with genetic variation in *HIF2α*, with specific SNPs
565 correlated with reduced hematocrit in this population (1, 63, 72). Among the largest allele
566 differences between Tibetan and non-Tibetans were non-coding SNPs (1, 72), suggesting that
567 regulatory changes, rather than structural changes, in *HIF2α* may be important in adaptation to
568 high altitude hypoxia. The current observation of multiple SNPs in *F. heteroclitus HIFα* genes,
569 combined with the recent reports of polymorphism in FIH-1 from the cyprinid *Megalobrama*
570 *amblycephala* (73) and evidence of selection acting on the HIF2αa locus in a pollution-tolerant
571 population of *F. heteroclitus* (52), suggests that genetic variation in elements of the HIF
572 signaling pathway could underlie differences among individuals or populations in their responses
573 to aquatic hypoxia or other stressors.
574 *Perspectives and Significance.* The HIF family of transcription factors plays important
575 roles in development, physiology, pathology, and environmental adaptation of animals. Herein,
576 HIF1α and 3α were sequenced and characterized, along with the previously described HIF2αa, in
577 *F. heteroclitus*, an important model for environmental biology (3, 61). Genomic and
578 phylogenetic analyses clearly placed the *F. heteroclitus* HIFs with their euteleost paralogs. DNA
579 binding and functional analyses showed that these HIFα subunits independently formed
580 complexes with *F. heteroclitus* ARNT2 to bind to canonical and non-canonical HREs and drive
581 reporter gene expression. Analysis of the *F. heteroclitus* genome uncovered a second HIF2α,

582 HIF2αb, which may be a "relic" of a teleost-specific genome duplication maintained in certain
583 fish species (57). Although not functionally characterized here, the predicted protein lacks
584 domains needed for oxygen sensing and activation of gene expression, suggesting that it may act
585 to repress the transcriptional activity of other HIFα subunits, analogous to the role of mammalian
586 HIF3α splice variants (32, 33, 36). Finally, cDNA sequencing and SNP analysis revealed
587 polymorphism in *F. heteroclitus* HIF1α, HIF2αa, and HIF3α. Future research is needed to
588 explore the whether this sequence variation is related to HIF expression, stability, function, and,
589 ultimately, the organismal response to hypoxia. Genetic variation in these transcription factors
590 could influence the capacity of individual killifish to tolerate low levels of oxygen and thereby
591 affect the ability of populations to adapt to a changing aquatic environment.
592

ACKNOWLEDGEMENTS

We thank Jawed Alam for the p3XHifREluc reporter plasmid. This research was supported in part by the National Science Foundation (IBN-0236494 and DEB-1120263) and by National Institute of Environmental Health Sciences (NIEHS) grant P42ES007381 (Superfund Basic Research Program at Boston University). Data interpretation was aided by reference to a preliminary draft of the *F. heteroclitus* genome sequence, which was supported by funding from the National Science Foundation (collaborative research grants DEB-1120512, DEB-1265282, DEB-1120013, DEB-1120263, DEB-1120333, DEB-1120398). The funding agencies were not involved in study design or performance or in the decision to publish the manuscript. The U.S. Government is authorized to produce and distribute reprints for governmental purposes notwithstanding any copyright notation that may appear hereon.

REFERENCES

1. **Beall CM, Cavalleri GL, Deng LB, Elston RC, Gao Y, Knight J, Li CH, Li JC, Liang Y, McCormack M, Montgomery HE, Pan H, Robbins PA, Shianna KV, Tam SC, Tsering N, Veeramah KR, Wang W, Wangdui PC, Weale ME, Xu YM, Xu Z, Yang L, Zaman MJ, Zeng CQ, Zhang L, Zhang XL, Zhaxi PC, Zheng YT**. Natural selection on EPAS1 (HIF2 alpha) associated with low hemoglobin concentration in Tibetan highlanders. *Proc Natl Acad Sci USA* 107: 11459-11464, 2010.

2. **Borowiec BG, Darcy KL, Gillette DM, Scott GR**. Distinct physiological strategies are used to cope with constant hypoxia and intermittent hypoxia in killifish (*Fundulus heteroclitus*). *J Exp Biol* 218: 1198-1211, 2015.

3. **Burnett KG, Bain LJ, Baldwin WS, Callard GV, Cohen S, Di Giulio RT, Evans DH, Gomez-Chiarri M, Hahn ME, Hoover CA, Karchner SI, Katoh F, MacLatchy DL, Marshall WS, Meyer JN, Nacci DE, Oleksiak MF, Rees BB, Singer TD, Stegeman JJ, Towle DW, Van Veld PA, Vogelbein WK, Whitehead A, Winn RN, Crawford DL**. *Fundulus* as the premier teleost model in environmental biology: opportunities for new insights using genomics. *Comp Biochem Physiol D-Genomics Proteomics* 2: 257-286, 2007.

4. **Chi W, Gan X, Xiao W, Wang W, He S**. Different evolutionary patterns of hypoxia-inducible factor alpha (HIF-alpha) isoforms in the basal branches of Actinopterygii and Sarcopterygii. *FEBS Open Bio* 3: 479-483, 2013.

5. **Cochran RE, Burnett LE**. Respiratory response of the salt marsh animals, *Fundulus heteroclitus*, *Leiostomus xanthurus*, and *Palaemonetes pugio* to environmental hypoxia and hypercapnia and to the organophosphate pesticide, azinphosmethyl. *J Exp Mar Biol Ecol* 195: 125-144, 1996.

6. **Du SNN, Mahalingam S, Borowiec BG, Scott GR**. Mitochondrial physiology and reactive oxygen species production are altered by hypoxia acclimation in killifish (*Fundulus heteroclitus*). *J Exp Biol* 219: 1130-1138, 2016.

7. **Duan CM**. Hypoxia-inducible factor 3 biology: complexities and emerging themes. *Am J Physiol-Cell Physiol* 310: C260-C269, 2016.

8. **Epstein ACR, Gleadle JM, McNeill LA, Hewitson KS, O'Rourke J, Mole DR, Mukherji M, Metzen E, Wilson MI, Dhanda A, Tian YM, Masson N, Hamilton DL, Jaakkola P, Barstead R, Hodgkin J, Maxwell PH, Pugh CW, Schofield CJ, Ratcliffe PJ**. *C-elegans* EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* 107: 43-54, 2001.

9. **Everett MV, Antal CE, Crawford DL**. The effect of short-term hypoxic exposure on metabolic gene expression. *J Exp Zool A-Ecol Genet Physiol* 317A: 9-23, 2012.

10. **Geng X, Feng JB, Liu SK, Wang YP, Arias C, Liu ZJ**. Transcriptional regulation of hypoxia inducible factors alpha (HIF-alpha) and their inhibiting factor (FIH-1) of channel catfish (*Ictalurus punctatus*) under hypoxia. *Comp Biochem Physiol B* 169: 38-50, 2014.

11. **Gong PF, Hu B, Stewart D, Ellerbe M, Figueroa YG, Blank V, Beckman BS, Alam J**. Cobalt induces heme oxygenase-1 expression by a hypoxia-inducible factor-independent

mechanism in Chinese hamster ovary cells - Regulation by Nrf2 and MafG transcription factors. *J Biol Chem* 276: 27018-27025, 2001.

12. **Gracey AY, Troll JV, Somero GN**. Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. *Proc Natl Acad Sci USA* 98: 1993-1998, 2001.

13. **Greaney GS, Place AR, Cashon RE, Smith G, Powers DA**. Time course of changes in enzyme activities and blood respiratory properties of killifish during long-term acclimation to hypoxia. *Physiol Zool* 53: 136-144, 1980.

14. **Greaney GS, Powers DA**. Cellular regulation of an allosteric modifier of fish hemoglobin. *Nature* 270: 73-74, 1977.

15. **Greenald D, Jeyakani J, Pelster B, Sealy I, Mathavan S, van Eeden FJ**. Genome-wide mapping of Hif-1alpha binding sites in zebrafish. *BMC Genomics* 16: 923, 2015.

16. **Gu YZ, Moran SM, Hogenesch JB, Wartman L, Bradfield CA**. Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3 alpha. *Gene Expr* 7: 205-213, 1998.

17. **Henikoff S, Henikoff JG**. Amino-acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89: 10915-10919, 1992.

18. **Huang JH, Zhao Q, Mooney SM, Lee FS**. Sequence determinants in hypoxia-inducible factor-1 alpha for hydroxylation by the prolyl hydroxylases PHD1, PHD2, and PHD3. *J Biol Chem* 277: 39792-39800, 2002.

19. **Jones DT, Taylor WR, Thornton JM**. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8: 275-282, 1992.

20. **Kaelin WG, Ratcliffe PJ**. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Mol Cell* 30: 393-402, 2008.

21. **Kallio PJ, Okamoto K, O'Brien S, Carrero P, Makino Y, Tanaka H, Poellinger L**. Signal transduction in hypoxic cells: inducible nuclear translocation and recruitment of the CBP/p300 coactivator by the hypoxia-inducible factor-1alpha. *EMBO J* 17: 6573-6586, 1998.

22. **Karchner SI, Jenny MJ, Tarrant AM, Evans BR, Kang HJ, Bae I, Sherr DH, Hahn ME**. The active form of human aryl hydrocarbon receptor (AHR) repressor lacks exon 8, and its Pro 185 and Ala 185 variants repress both AHR and hypoxia-inducible factor. *Mol Cell Biol* 29: 3465-3477, 2009.

23. **Keith B, Johnson RS, Simon MC**. HIF1 alpha and HIF2 alpha: sibling rivalry in hypoxic tumour growth and progression. *Nat Rev Cancer* 12: 9-22, 2012.

24. **Kimura H, Weisz A, Ogura T, Hitomi Y, Kurashima Y, Hashimoto K, D'Acquisto F, Makuuchi M, Esumi H**. Identification of hypoxia-inducible factor 1 ancillary sequence and its function in vascular endothelial growth factor gene induction by hypoxia and nitric oxide. *J Biol Chem* 276: 2292-2298, 2001.

25. **Knechtel G, Szkandera J, Stotz M, Hofmann G, Langsenlehner U, Krippl P, Samonigg H, Renner W, Langner C, Dehchamani D, Gerger A**. Single nucleotide

polymorphisms in the hypoxia-inducible factor-1 gene and colorectal cancer risk. *Mol Carcinogen* 49: 805-809, 2010

26. **Köblitz L, Fiechtner B, Baus K, Lussnig R,P elster B**. Developmental Expression and Hypoxic Induction of Hypoxia Inducible Transcription Factors in the Zebrafish. *PLoS One* 10: e0128938, 2015.

27. **Kuo WH, Shih CM, Lin CW, Cheng WE, Chen SC, Chen W, Lee YL**. Association of hypoxia inducible factor-1 alpha polymorphisms with susceptibility to non-small-cell lung cancer. *Transl Res* 159: 42-50, 2012.

28. **Laemmli UK**. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685, 1970.

29. **Law SHW, Wu RSS, Ng PKS, Yu RMK, Kong RYC**. Cloning and expression analysis of two distinct HIF-alpha isoforms - gcHIF-1alpha and gcHIF-4alpha - from the hypoxia-tolerant grass carp, *Ctenopharyngodon idellus*. *BMC Mol Biol* 7: 2006.

30. **Li DX, Hirsila M, Koivunen P, Brenner MC, Xu L, Yang C, Kivirikko KI, Myllyharju J**. Many amino acid substitutions in a hypoxia-inducible transcription factor (HIF)-1 alpha-like peptide cause only minor changes in its hydroxylation by the HIF prolyl 4-hydroxylases - Substitution of 3,4-dehydroproline or azetidine-2-carboxylic acid for the proline leads to a high rate of uncoupled 2-oxoglutarate decarboxylation. *J Biol Chem* 279: 55051-55059, 2004.

31. **Lisy K, Peet DJ**. Turn me on: regulating HIF transcriptional activity. *Cell Death Differ* 15: 642-649, 2008.

32. **Makino Y, Cao RH, Svensson K, Bertilsson GR, Asman M, Tanaka H, Cao YH, Berkenstam A, Poellinger L**. Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression. *Nature* 414: 550-554, 2001.

33. **Makino Y, Kanopka A, Wilson WJ, Tanaka H, Poellinger L**. Inhibitory PAS domain protein (IPAS) is a hypoxia-inducible splicing variant of the hypoxia-inducible factor-3 alpha locus. *J Biol Chem* 277: 32405-32408, 2002.

34. **Marchler-Bauer A, Bryant SH**. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* 32: W327-W331, 2004.

35. **Marques IJ, Leito JTD, Spaink HP, Testerink J, Jaspers RT, Witte F, van Den Berg S, Bagowski CP**. Transcriptome analysis of the response to chronic constant hypoxia in zebrafish hearts. *J Comp Physiol B* 178: 77-92, 2008.

36. **Maynard MA, Evans AJ, Hosomi T, Hara S, Jewett MAS, Ohh M**. Human HIF-3 alpha 4 is a dominant-negative regulator of HIF-1 and is down-regulated in renal cell carcinoma. *FASEB J* 19: 1396-1406, 2005.

37. **McIntosh BE, Hogenesch JB, Bradfield CA**. Mammalian Per-Arnt-Sim proteins in environmental adaptation. *Annu Rev Physiol* 72: 625-645, 2010.

38. **McPhee JS, Perez-Schindler J, Degens H, Tomlinson D, Hennis P, Baar K, Williams AG**. HIF1A P582S gene association with endurance training responses in young women. *Eur J Appl Physiol* 111: 2339-2347, 2011.

724 39. **Meyer A, Van de Peer Y**. From 2R to 3R: evidence for a fish-specific genome
725 duplication (FSGD). *Bioessays* 27: 937-945, 2005.

726 40. **Mohindra V, Tripathi RK, Singh RK, Lal KK**. Molecular characterization and
727 expression analysis of three hypoxia-inducible factor alpha subunits, HIF-1alpha, -2alpha
728 and -3alpha in hypoxia-tolerant Indian catfish, *Clarias batrachus* [Linnaeus, 1758]. *Mol Biol*
729 *Rep* 40: 5805-5815, 2013.

730 41. **Mole DR, Blancher C, Copley RR, Pollard PJ, Gleadle JM, Ragoussis J, Ratcliffe PJ**.
731 Genome-wide association of hypoxia-inducible factor (HIF)-1 alpha and HIF-2 alpha DNA
732 binding with expression profiling of hypoxia-inducible transcripts. *J Biol Chem* 284: 16767-
733 16775, 2009.

734 42. **Nagy G, Kovacs-Nagy R, Kereszturi E, Somogyi A, Szekely A, Nemeth N,**
735 **Hosszufalusi N, Panczel P, Ronai Z, Sasvari-Szekely M**. Association of hypoxia inducible
736 factor-1 alpha gene polymorphism with both type 1 and type 2 diabetes in a Caucasian
737 (Hungarian) sample. *BMC Med Genet* 10: 2009.

738 43. **Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC,**
739 **Friedman M, Smith WL**. Resolution of ray-finned fish phylogeny and timing of
740 diversification. *Proc Natl Acad Sci USA* 109: 13698-13703, 2012.

741 44. **Powell WH, Hahn ME**. Identification and functional characterization of hypoxia-
742 inducible factor 2 alpha from the estuarine teleost, *Fundulus heteroclitus*: Interaction of
743 HIF-2 alpha with two ARNT2 splice variants. *J Exp Zool* 294: 17-29, 2002.

744 45. **Powell WH, Karchner SI, Bright R, Hahn ME**. Functional diversity of vertebrate
745 ARNT proteins: Identification of ARNT2 as the predominant form of ARNT in the marine
746 teleost, *Fundulus heteroclitus*. *Arch Biochem Biophys* 361: 156-163, 1999.

747 46. **Prabhakar NR, Semenza GL**. Adaptive and maladaptive cardiorespiratory
748 responses to continuous and intermittent hypoxia mediated by hypoxia-inducible factors 1
749 and 2. *Physiol Rev* 92: 967-1003, 2012.

750 47. **Prior SJ, Hagberg JM, Phares DA, Brown MD, Fairfull L, Ferrell RE, Roth SM**.
751 Sequence variation in hypoxia-inducible factor 1alpha (HIF1A): association with maximal
752 oxygen consumption. *Physiol Genomics* 15: 20-26, 2003.

753 48. **Rabalais NN, Diaz RJ, Levin LA, Turner RE, Gilbert D, Zhang J**. Dynamics and
754 distribution of natural and human-caused hypoxia. *Biogeosciences* 7: 585-619, 2010.

755 49. **Rahman MS, Thomas P**. Molecular cloning, characterization and expression of two
756 hypoxia-inducible factor alpha subunits, HIF-1 alpha and HIF-2 alpha, in a hypoxia-tolerant
757 marine teleost, Atlantic croaker (Micropogonias undulatus). *Gene* 396: 273-282, 2007.

758 50. **Rees BB, Figueroa YG, Wiese TE, Beckman BS, Schulte PM**. A novel hypoxia-
759 response element in the lactate dehydrogenase-B gene of the killifish *Fundulus heteroclitus*.
760 *Comp Biochem Physiol A Mol Integr Physiol* 154: 70-77, 2009.

761 51. **Rees BB, Targett TE, Ciotti BJ, Tolman CA, Akkina SS, Gallaty AM**. Temporal
762 dynamics in growth and white skeletal muscle composition of the mummichog *Fundulus*
763 *heteroclitus* during chronic hypoxia and hyperoxia. *J Fish Biol* 81: 148-164, 2012.

52. **Reid NM, Proestou DA, Clark BW, Warren WC, Colbourne JK, Shaw JR, Karchner SI, Hahn ME, Nacci D, Oleksiak MF, Crawford DL, Whitehead A**. The landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* in press, 2016.

53. **Richards JG**. Metabolic and molecular responses of fish to hypoxia. In: *Hypoxia*, edited by Richards JG, Farrell AP, and Brauner CJ. London: Academic Press, 2009, p. 443-485.

54. **Rimoldi S, Terova G, Ceccuzzi P, Marelli S, Antonini M, Saroglia M**. HIF-1 alpha mRNA levels in Eurasian perch (*Perca fluviatilis*) exposed to acute and chronic hypoxia. *Mol Biol Rep* 39: 4009-4015, 2012.

55. **Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP**. Integrative genomics viewer. *Nat Biotechnol* 29: 24-26, 2011.

56. **Rojas DA, Perez-Munizaga DA, Centanin L, Antonelli M, Wappner P, Allende ML, Reyes AE**. Cloning of hif-1 alpha and hif-2 alpha and mRNA expression pattern during development in zebrafish. *Gene Expr Patterns* 7: 339-345, 2007.

57. **Rytkönen KT, Akbarzadeh A, Miandare HK, Kamei H, Duan CM, Leder EH, Williams TA, Nikinmaa M**. Subfunctionalization of Cyprinid hypoxia-inducible factors for roles in development and oxygen sensing. *Evolution* 67: 873-882, 2013.

58. **Rytkönen KT, Prokkola JM, Salonen V, Nikinmaa M**. Transcriptional divergence of the duplicated hypoxia-inducible factor alpha genes in zebrafish. *Gene* 541: 60-66, 2014.

59. **Rytkönen KT, Williams TA, Renshaw GM, Primmer CR, Nikinmaa M**. Molecular evolution of the metazoan PHD-HIF oxygen-sensing system. *Mol Biol Evol* 28: 1913-1926, 2011.

60. **Schödel J, Mole DR, Ratcliffe PJ**. Pan-genomic binding of hypoxia-inducible transcription factors. *Biol Chem* 394: 507-517, 2013.

61. **Schulte PM**. What is environmental stress? Insights from fish living in a variable environment. *J Exp Biol* 217: 23-34, 2014.

62. **Semenza GL, Wang GL**. A nuclear factor induced by hypoxia via *de novo* protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. *Mol Cell Biol* 12: 5447-5454, 1992.

63. **Simonson TS, Yang YZ, Huff CD, Yun HX, Qin G, Witherspoon DJ, Bai ZZ, Lorenzo FR, Xing JC, Jorde LB, Prchal JT, Ge RL**. Genetic evidence for high-altitude adaptation in Tibet. *Science* 329: 72-75, 2010.

64. **Smith KJ, Able KW**. Dissolved oxygen dynamics in salt marsh pools and its potential impacts on fish assemblages. *Mar Ecol Prog Ser* 258: 223-232, 2003.

65. **Stierhoff KL, Targett TE, Grecay PA**. Hypoxia tolerance of the mummichog: the role of access to the water surface. *J Fish Biol* 63: 580-592, 2003.

66. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S**. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30: 2725-2729, 2013.

67. **Ton C, Stamatiou D, Liew CC**. Gene expression profile of zebrafish exposed to hypoxia during development. *Physiol Genomics* 13: 97-106, 2003.

68. **Towbin H, Staehelin T, Gordon J**. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci USA* 76: 4350-4354, 1979.

69. **Trapnell C, Pachter L, Salzberg SL**. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105-1111, 2009.

70. **Wannamaker CM, Rice JA**. Effects of hypoxia on movements and behavior of selected estuarine organisms from the southeastern United States. *J Exp Mar Biol Ecol* 249: 145-163, 2000.

71. **Wenger RH, Stiehl DP, Camenisch G**. Integration of oxygen signaling at the consensus HRE. *Sci STKE* 2005: re12, 2005.

72. **Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng HC, Liu T, He WM, Li K, Luo RB, Nie XF, Wu HL, Zhao MR, Cao HZ, Zou J, Shan Y, Li SZ, Yang Q, Asan, Ni PX, Tian G, Xu JM, Liu XA, Jiang T, Wu RH, Zhou GY, Tang MF, Qin JJ, Wang T, Feng SJ, Li GH, Huasang, Luosang JB, Wang W, Chen F, Wang YD, Zheng XG, Li Z, Bianba ZM, Yang G, Wang XP, Tang SH, Gao GY, Chen Y, Luo Z, Gusang L, Cao Z, Zhang QH, Ouyang WH, Ren XL, Liang HQ, Zheng HS, Huang YB, Li JX, Bolund L, Kristiansen K, Li YR, Zhang Y, Zhang XQ, Li RQ, Li SG, Yang HM, Nielsen R, Wang J, Wang JA**. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329: 75-78, 2010.

73. **Zhang B, Chen N, Huang CH, Huang CX, Chen BX, Liu H, Wang WM, Gul Y, Wang HL**. Molecular response and association analysis of *Megalobrama amblycephala* fih-1 with hypoxia. *Mol Genet Genomics* 291: 1615-1624, 2016.

74. **Zhang P, Lu L, Yao Q, Li Y, Zhou J, Liu Y, Duan C**. Molecular, functional, and gene expression analysis of zebrafish hypoxia-inducible factor-3alpha. *Am J Physiol Regul Integr Comp Physiol* 303: R1165-1174, 2012.

75. **Zhang P, Yao Q, Lu L, Li Y, Chen PJ, Duan CM**. Hypoxia-inducible factor 3 is an oxygen-dependent transcription activator and regulates a distinct transcriptional response to hypoxia. *Cell Rep* 6: 1110-1121, 2014.

76. **Zhou YQ, Lin L, Wang Y, Jin X, Zhao X, Liu DJ, Hu T, Jiang L, Dan HX, Zeng X, Li J, Wang JY, Chen QM**. The association between hypoxia-inducible factor-1 alpha gene G1790A polymorphism and cancer risk: a meta-analysis of 28 case-control studies. *Cancer Cell Int* 14: 2014.

837 Table 1. Sequences of primers used in this study.
838

Primer Name	Use	Primer Sequence (5' – 3')
HIF1-Forward	PCR	AGGAGGGGGAAGGAATCTGAGGT
HIF1-Reverse	PCR	ACAACACACTGGGGCTGGGAG
AP-1	RACE	CCATCCTAATACGACTCACTATAGGGC
AP-2	RACE	ACTCACTATAGGGCTCGAGCGGC
HIF1 5'gs outer	RACE	GTCTCCGTCTTCAGACAGCACCAG
HIF1 5'gs inner	RACE	AGCTGCATGTCCAGAGCTGGTTTG
HIF1 5'gs alt	RACE	CTCTAAAGCCTTCAGGTAGGAGCC
HIF1 3'gs outer	RACE	CGGAGGACCTGCTGAACCGGTCTG
HIF1 3'gs inner	RACE	CAAAGACTCACCATCACTTGTTTG
HIF1 3'gs alt	RACE	ACCACGCTCTGGACTCGGACTATC
HIF3 5'gs outer	RACE	ACAGGGGTTCGGTTTTGGGTAG
HIF3 5'gs inner	RACE	TGTTGTTGTAGAGGACGGTGGC
HIF3 3'gs outer	RACE	TGACCTTCCTCAGCGATTTGC
HIF3 3'gs inner	RACE	AAGCAAGAGCCGTCCTCCATAG
HIF1 5'UTR	PCR	GAACCCAGGAGGAACTCTTATGTG
HIF1 3'UTR	PCR	TAAATAAGTGGCAGTGGGGTC
HIF3 5'UTR	PCR	AACAGAGAGCCTTGGATTTGGTGTTC
HIF3 3'UTR	PCR	GGCGAGGATCTTTTGTGCGTACAG
HIF2 5'subclone	Ab production	*TTCACTGCAG*AGCCCAGGTGATTACTACAGC
HIF2 3'subclone	Ab production	*TTCAGTCGAC*CCTCTGTCGGGTGTCCACAG
HIF3 5'subclone	Ab production	*TTCACTGCAG*TTCCAGCTGACCTTCCTCAGC
HIF3 3'subclone	Ab production	*TTCAGTCGAC*GAAGATGTCCCTCATCAGAGCTG
HIF1-F1	QPCR	CAAGTCGGCTACGTGGAAGGTG
HIF1-R1	QPCR	CAGATCAGGACCAGATAGGGGAC
HIF2-F1	QPCR	GGCTTCATCACCGTGGTAACATC
HIF2-R1	QPCR	CTGTGACCTGTGAGCTCCACCTG
HIF3-F1	QPCR	GTCAACAAGCACATCGGCATCACG
HIF3-R1	QPCR	CCATCAGTTTCTTACTCAGACCTGG

839
840 Italicized bases indicate restriction enzyme sites included for sub-cloning.

Table 2. Location of predicted amino acid substitutions in HIFα proteins based upon single nucleotide polymorphisms found by sequencing mRNA pooled from 200 *F. heteroclitus* embryos.

Subunit	Position	Substitution[a]	Score[b]	Domain(s)[c]
HIF1α	36	F/L	0	bHLH
HIF1α	361	E/A	-1	Between PAC and ODD
HIF1α	370	E/Q	2	Between PAC and ODD
HIF1α	630	T/A	0	Between ODD and TAD-C
HIF2αa	87	A/T	0	Between bHLH and PAS-A
HIF2αa	165	K/R	2	Between PAS-A and PAS-B
HIF2αa	441	H/P	-2	ODD
HIF3α	37	V/G	-3	bHLH
HIF3α	300	T/P	-1	Between PAS-B and PAC
HIF3α	413	E/D	2	Putative prolyl hydroxylation motif
HIF3α	422	S/P	-1	Between PAC and ODD
HIF3α	502	T/A	0	ODD

[a]The first amino acid shown for each substitution is from HIF1α*1 (ALL26120.1; this study), HIF2αa [ALL95711.1; (44)], or HIF3α*1 (ALL26129.1; this study).

[b]Values are from BLOSUM62 alignment score matrix (17). Positive scores are conservative substitutions and negative scores are non-conservative substitutions.

[c]Domain abbreviations are: bHLH, basic helix-loop-helix; PAS, PER-ARNT-SIM domain; PAC, motif C-terminal to PAS domain; ODD, oxygen-dependent degradation domain; TAD-C, C-terminal transactivation domain.

FIGURE LEGENDS

Fig. 1. *Fundulus heteroclitus* HIF1α and HIF3α protein models contain conserved functional domains. Protein domains were identified by Conserved Domain Search (34) in FhHIF1α (A) and FhHIF3α (C). Abbreviations are: bHLH, basic helix-loop-helix; PAS, PER-ARNT-SIM domain; PAC, PAS-associated C-terminal motif; ODD, oxygen-dependent degradation domain; TAD, transactivation domain (N for more N-terminal; C for more C-terminal in FhHIF1α). The numbers along the top of each diagram represent amino acid number and potential sites of prolyl or asparaginyl hydroxylation. Deduced full-length proteins are 753-755 aa for FhHIF1α and 662 aa for FhHIF3α. Putative prolyl hydroxylation sites were compared among vertebrate species for FhHIF1α (B) and FhHIF3α (D) through multiple amino acid alignments. For each potential LXXLAP motif, the numbers refer to amino acid residue number in the intact polypeptide and a dot indicates a residue identical to the human sequence, except for the more N-terminal HIF3α site, which was compared to the zebrafish sequence. Species abbreviations are Hs, *Homo sapiens*; Mm, *Mus musculus* (mouse); Gg, *Gallus gallus* (chicken); Ac, *Anolis carolinensis* (green lizard); Mu, *Micropogonias undulatus* (Atlantic croaker); Dr, *Danio rerio* (zebrafish); Ci, *Ctenopharyngodon idella* (grass carp); Fh, *Fundulus heteroclitus* (Atlantic killifish).

Fig. 2. Phylogenetic analyses reveal relationships among HIFα forms in *F. heteroclitus* and other vertebrates. Selected HIFα amino acid sequences were analyzed using the Maximum Likelihood method based on the JTT matrix-based model (19). The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The region corresponding to residues 1-360 of human HIF1α from each

878 HIF sequence was aligned using ClustalX. All positions containing gaps or missing data were
879 eliminated. Killifish sequences are shown in grey boxes. Species and accession numbers are:
880 human (*Homo sapiens*) HIF1α NP_001521, HIF2α NP_001421, HIF3α NP_690007; mouse
881 (*Mus musculus*) HIF1α NP_001300848, HIF2α NP_034267, HIF3α NP_058564; lizard (*Anolis*
882 *carolinensis*) HIF1α XP_008121253; HIF2α XP_003225226, HIF3α XP_016853234; killifish
883 (*Fundulus heteroclitus*) HIF1α ALL26120, HIF2αa NP_001296843, HIF2αb XP_012709515,
884 HIF3α ALL26129; tilapia (*Oreochromis niloticus*) HIF1α XP_005477095, HIF2α
885 XP_003438301, HIF2αb XP_003441929, HIF3α XP_005461198; croaker (*Micropogonias*
886 *undulatus*) HIF1α ABD32158, HIF2α ABD32159; zebrafish (*Danio rerio*) HIF1αa
887 NP_001295488, HIF1αb NP_001296971, HIF2αa NP_001034895, HIF2αb XP_695262, HIF3αa
888 ADF58783, HIF3αb NP_001012371; grass carp (*Ctenopharyngodon idella*) HIF1α AAR95697,
889 HIF2α AAT76668, HIF3α AAR95698; asp (*Aspius aspius*) HIF1αa AFD32323, HIF1αb
890 ABO26713, HIF2α AFD32324. Ensembl Release 85 predicted proteins: green puffer (*Tetraodon*
891 *nigroviridis*) HIF2αb ENSTNIP00000008013 and stickleback (*Gasterosteus aculeatus*) HIF2αb
892 ENSGACP00000003681.
893
894 Fig. 3. Genomic organization of *F. heteroclitus HIFα* genes is conserved. The flanking genes
895 (Fig. 3A) and gene models (Fig. 3B) of killifish *HIF1α*, *HIF2αa*, *HIF2αb*, and *HIF3α* genes are
896 shown. Flanking genes were obtained from the *F. heteroclitus* genome sequence database v2b
897 and compared to green pufferfish, stickleback, tilapia, and zebrafish genomes using Genomicus
898 version 85.01 (http://www.genomicus.biologie.ens.fr). The orientation of the flanking genes is
899 shown by arrows above the gene abbreviations. Italic letters below gene names indicate shared
900 synteny among orthologous genes in one or more euteleost genome (a) or in the zebrafish

genome (b). An asterisk indicates that a nominally different paralog (e.g., *PRKCE* instead of *PRKCH*) is found at that location in the other genomes. Exons and introns (Fig 3B) were predicted by alignment of cDNA sequences to the *Fundulus heteroclitus* genome sequence, v2b. Exons are shown in black boxes, with white portions of the first and last exons representing untranslated regions. Lines connecting exons represent introns. Exons and introns are drawn to scale, except for introns longer than 500 bp, in which cases the length of the intron is indicated. Question marks indicate introns of unknown lengths.

Fig. 4. *F. heteroclitus* HIFs bind the human *EPO* 3' enhancer. DNA sequences (A) of oligonucleotides corresponding to the wild type or mutated versions of the human *EPO* 3' enhancer [EPO-w18 or EPO-m18 (62)] or one of two putative HIF-binding sites (HBSs) from the *F. heteroclitus Ldh-B* gene [HBS1-wt or HBS2-wt (50)]). The known or putative binding sites are shown in bold type and the mutated bases are shown in lower case. Representative gels from EMSA containing IVTT *F heteroclitus* HIF1α (B), HIF2αa (C), or HIF3α (D) in combination with *F. heteroclitus* ARNT2. Radiolabeled, double-stranded EPO-w18 was included as the probe in every lane of every gel. The leftmost lanes indicate the absence (-) or presence (+) of the appropriate IVTT HIFα subunit. The remaining lanes included unlabeled competitor DNA at the indicated molar excess relative to labeled probe. The arrows indicate bands that were specifically competed by EPO-w18 oligonucleotide. Part E shows the average band intensities (± one standard deviation) of replicate EMSA at 400-fold molar excess with each competitor DNA relative to the band intensity in the same gel when no competitor DNA was included ($n = 4 – 7$). For each HIFα subunit, competitors that differed in their ability to compete for HIF binding are represented by different letters (Bonferroni's post-hoc comparisons, $P < 0.05$).

924

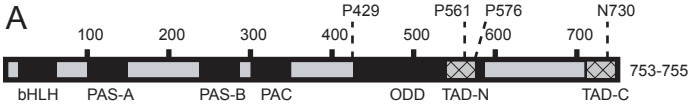
925 Fig. 5. *F. heteroclitus* HIFs bind a putative HRE from the *F. heteroclitus Ldh-B* promoter. DNA
926 sequences (A) of oligonucleotides corresponding to the wild type or mutated versions of a
927 putative HRE from the *F. heteroclitus Ldh-B* promoter (50). The wild type HRE (wt/wt) contains
928 two HIF binding sites (HBSs, bold type), whereas the mutated versions (mut/mut; mut/wt;
929 wt/mut) have changes at one or both HBS (lower case). Representative gels from EMSA
930 containing IVTT *F heteroclitus* HIF1α (B), HIF2αa (C), or HIF3α (D) in combination with *F.*
931 *heteroclitus* ARNT2. Radiolabeled, double-stranded wt/wt was included as the probe in every
932 lane of every gel. The leftmost lanes indicate the absence (-) or presence (+) of the appropriate
933 IVTT HIFα subunit. The remaining lanes included unlabeled competitor DNA at the indicated
934 molar excess relative to labeled probe. The arrows indicate bands that were specifically
935 competed by wt/wt oligonucleotide. Part E shows the average band intensities (± one standard
936 deviation) of replicate EMSA at 400-fold molar excess with each competitor DNA relative to the
937 band intensity in the same gel when no competitor DNA was included ($n = 3 – 4$). For each HIFα
938 subunit, competitors that differed in their ability to compete for HIF binding are represented by
939 different letters (Bonferroni's post-hoc comparisons, $P < 0.05$).

940

941 Fig. 6. Antibodies against *F. heteroclitus* HIF subunits supershift specific EMSA bands. All
942 reactions included one of three IVTT HIFα subunits and *F. heteroclitus* ARNT2 (+). Additional
943 components were chicken polyclonal antibodies against each α subunit (lanes labeled anti-
944 HIF1α, anti-HIF2αa, and anti-HIF3α, respectively), or a mouse monoclonal antibody against
945 human ARNT (anti-ARNT). Control reactions included pre-immune IgY. Arrows indicate the
946 bands whose mobility is specifically shifted by antibodies against HIFα subunits or ARNT, but

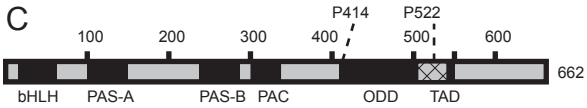
947 not by pre-immune IgY. The labeled DNA probe for HIF1α and HIF3α was 18 bp from the
948 human *EPO* 3' enhancer (EPO-w18), and for HIF2αa the DNA probe was 32 bp from the *F.*
949 *heteroclitus Ldh-B* promoter (see Figs. 4A and 5A for sequences).
950
951 Fig. 7. *F. heteroclitus* HIFs induce HRE-dependent reporter gene expression in a dose-dependent
952 fashion. COS-7 cells were co-transfected with a plasmid encoding firefly luciferase under the
953 control of the *F. heteroclitus Ldh-B* promoter containing a putative HRE and increasing amounts
954 of plasmids encoding each HIFα subunit. Part (A) shows the increase in luciferase expression
955 (mean ± one standard deviation of three experiments) in COS-7 cells transfected with 10 – 160
956 ng of each FhHIFα plasmid relative to COS-7 cells transfected with empty pcDNA (0 ng). Part
957 (B) shows results of a single experiment in which COS-7 cells were co-transfected with lower
958 amounts of each FhHIFα plasmid (0 – 8 ng), showing greater expression driven by FhHIF1α
959 even at low plasmid doses. Part (C) shows results when the reporter plasmid encoded firefly
960 luciferase under the control of the HRE from mouse heme oxygenase gene and COS-7 cells were
961 co-transfected with increasing amounts of FhHIFα plasmids (0 – 160 ng). Part (D) shows
962 FhHIFα and ARNT2 protein levels determined in parallel by western blotting. *F. heteroclitus*
963 ARNT2 plasmid was included in all transfections at a single dose (400 ng). The (+) lane
964 indicates IVTT of each respective protein included in one lane of the gel as a positive control.
965 Part (E) shows results of a single experiment in which COS-7 cells were co-transfected with a
966 plasmid encoding firefly luciferase under the control of the *F. heteroclitus Ldh-B* promoter and
967 either 80 ng of one FhHIFα plasmid or 80 ng each of two FhHIFα plasmids.
968

969 Fig. 8. HIFα transcripts are expressed in multiple organs of normoxic *F. heteroclitus*. The levels
970 of HIF1α (A), HIF2αa (B), and HIF3α (C) mRNA were determined in organs of adult *F.*
971 *heteroclitus* by quantitative PCR. For a given HIFα subunit, the effects of organ were determined
972 by one-way ANOVA. Organs with significantly different abundance of a given HIFα mRNA are
973 indicated by having different lower case letters (Bonferroni's post-hoc comparisons, $P < 0.05$).
974 In each organ, differences in transcript abundance among the HIFα forms were assessed by a
975 one-way ANOVA and differences among HIFα forms within a given organ are indicated by
976 uppercase italic letters (comparing across panels A-C). Error bars are one standard deviation (n =
977 6, except for kidney where $n = 5$ and gill where $n = 4$).



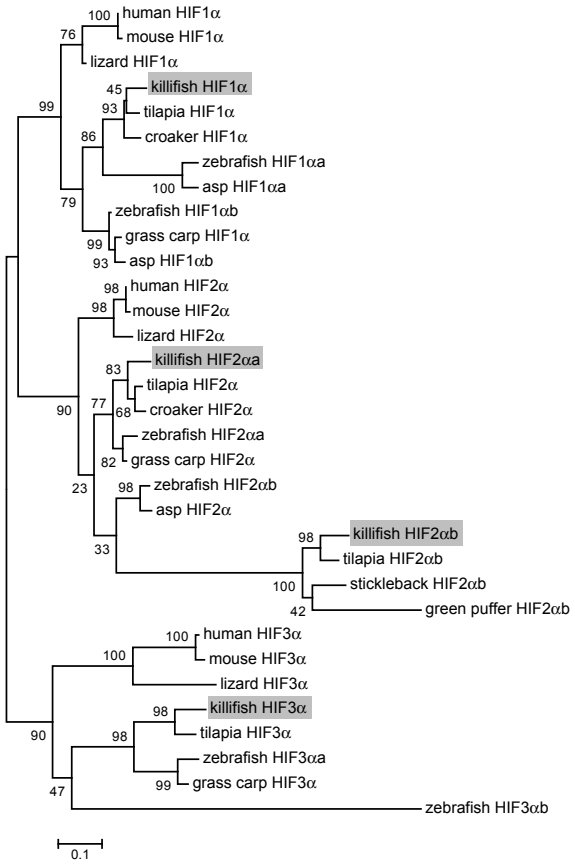
B

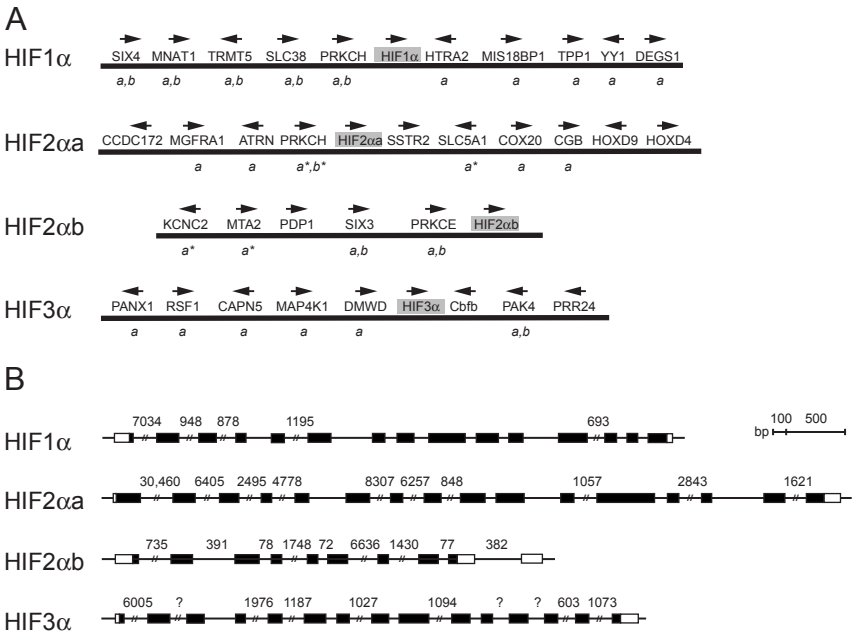
Hs	397 **LTLLAP** 402	559 **LEMLAP** 564	574 **LRSFDQ** 579
Mm	397 402	572 577	587 592
Gg	397 ..V... 402	557 562	572 577
Ac	400 ..Y... 405	561 566	576 581
Mu	428 433	557 562	572 ...LP. 577
Dr	409 ..V... 414	552 557	567 ..IPSP 572
Ci	407 ..V... 412	550 555	565 ..VPSP 570
Fh	424 429	556 ...F.. 561	571 .PGLAP 576

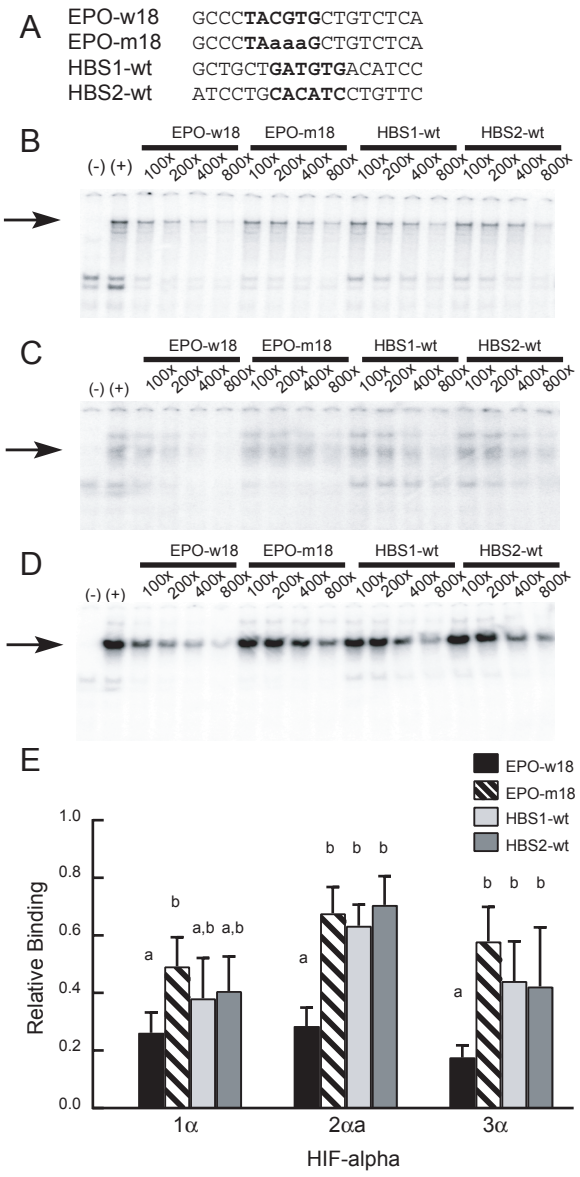


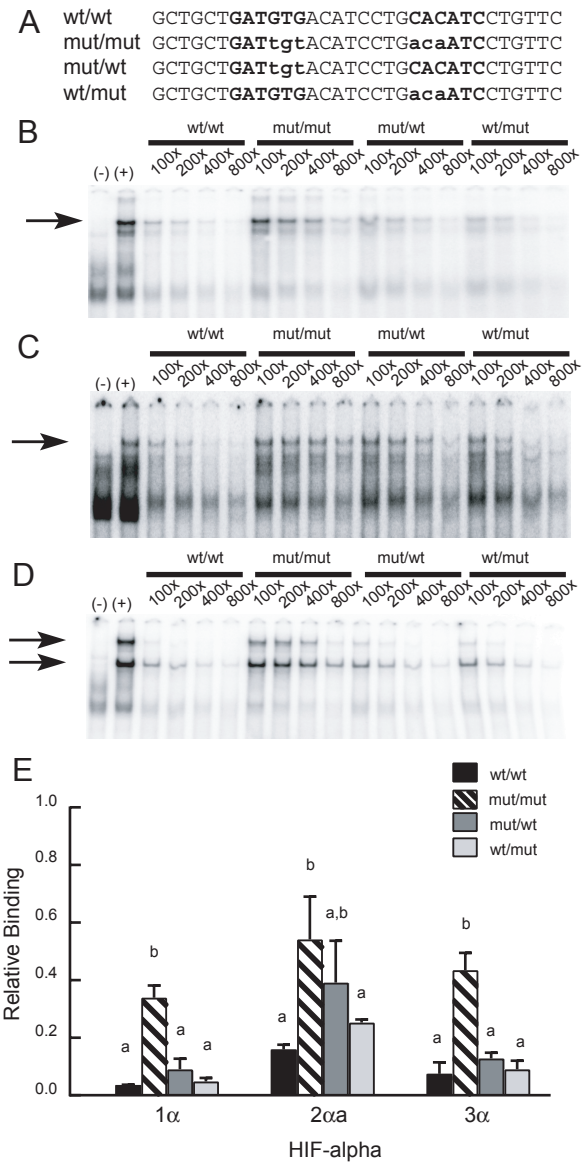
D

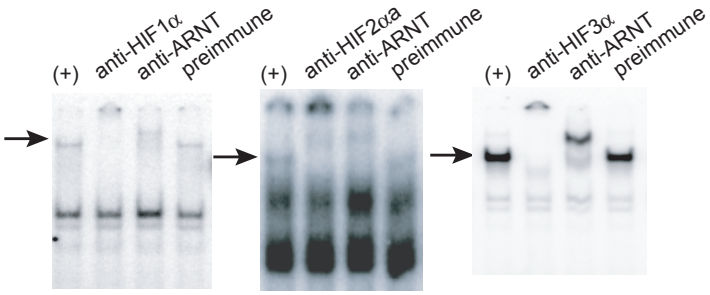
Hs		485 **LEMLAP** 490
Mm		482 487
Ac		498 503
Dr	333 **LLQLAP** 338	433 .D.... 438
Ci	398 .M.... 403	498 .D.... 503
Fh	409 ....E. 414	517 ...Y.. 522

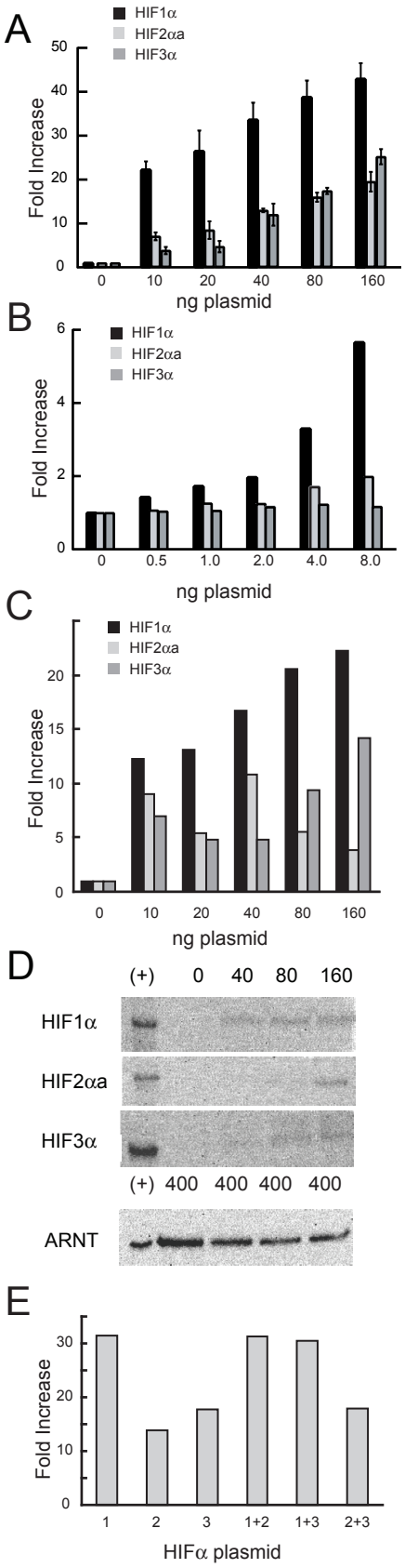

human HIF1α
mouse HIF1α
lizard HIF1α
killifish HIF1α
tilapia HIF1α
croaker HIF1α
zebrafish HIF1αa
asp HIF1αa
zebrafish HIF1αb
grass carp HIF1α
asp HIF1αb
human HIF2α
mouse HIF2α
lizard HIF2α
killifish HIF2αa
tilapia HIF2α
croaker HIF2α
zebrafish HIF2αa
grass carp HIF2α
zebrafish HIF2αb
asp HIF2α
killifish HIF2αb
tilapia HIF2αb
stickleback HIF2αb
green puffer HIF2αb
human HIF3α
mouse HIF3α
lizard HIF3α
killifish HIF3α
tilapia HIF3α
zebrafish HIF3αa
grass carp HIF3α
zebrafish HIF3αb
0.1

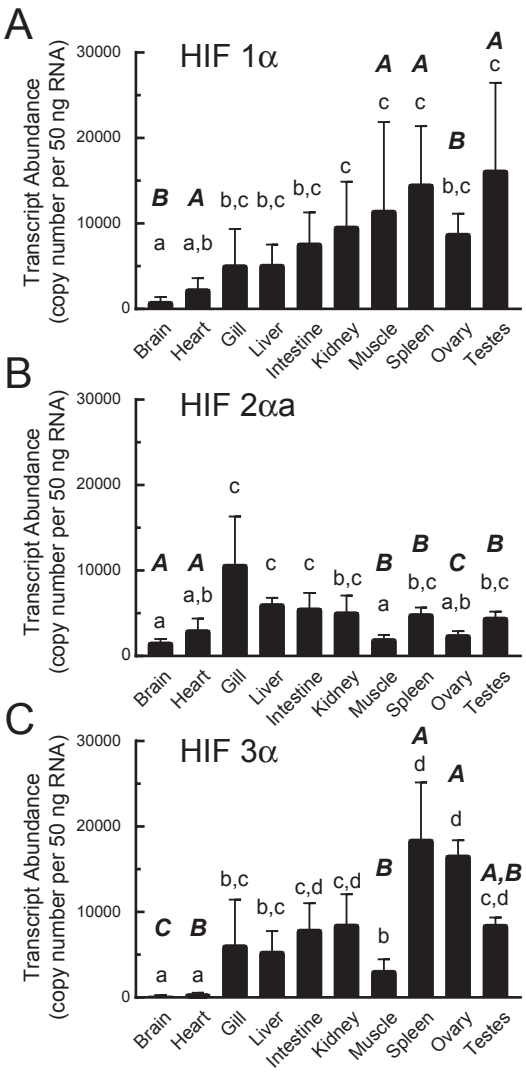

A
HIF1α
SIX4 MNAT1 TRMT5 SLC38 PRKCH HIF1α HTRA2 MIS18BP1 TPP1 YY1 DEGS1
a,b a,b a,b a,b a,b a a a a a
HIF2αa
CCDC172 MGFRA1 ATRN PRKCH HIF2αa SSTR2 SLC5A1 COX20 CGB HOXD9 HOXD4
a a a*,b* a* a a
HIF2αb
KCNC2 MTA2 PDP1 SIX3 PRKCE HIF2αb
a* a* a,b a,b
HIF3α
PANX1 RSF1 CAPN5 MAP4K1 DMWD HIF3α Cbfb PAK4 PRR24
a a a a a a,b
B
HIF1α
7034 948 878 1195 693
100 500
bp
HIF2αa
30,460 6405 2495 4778 8307 6257 848 1057 2843 1621
HIF2αb
735 391 78 1748 72 6636 1430 77 382
HIF3α
6005 ? 1976 1187 1027 1094 ? ? 603 1073


A
EPO-w18 GCCCTACGTGCTGTCTCA
EPO-m18 GCCCTAaaaGCTGTCTCA
HBS1-wt GCTGCTGATGTGACATCC
HBS2-wt ATCCTGCACATCCTGTTC
B
EPO-w18
EPO-m18
HBS1-wt
HBS2-wt
(-) (+)
100x
200x
400x
800x
C
D
E
EPO-w18
EPO-m18
HBS1-wt
HBS2-wt
Relative Binding
1.0
0.8
0.6
0.4
0.2
0.0
a
b
a,b
a,b
a
b
b
b
a
b
b
b
1α
2αa
3α
HIF-alpha


A
wt/wt GCTGCTGATGTGACATCCTGCACATCCTGTTC
mut/mut GCTGCTGATtgtACATCCTGacaATCCTGTTC
mut/wt GCTGCTGATtgtACATCCTGCACATCCTGTTC
wt/mut GCTGCTGATGTGACATCCTGacaATCCTGTTC
B
C
D
wt/wt
mut/mut
mut/wt
wt/mut
100x
200x
400x
800x
(-) (+)
E
wt/wt
mut/mut
mut/wt
wt/mut
Relative Binding
1.0
0.8
0.6
0.4
0.2
0.0
a
b
a
a
a
b
a,b
a
a
b
a
a
1α
2αa
3α
HIF-alpha


(+)
anti-HIF1α
anti-ARNT
preimmune
(+)
anti-HIF2αa
anti-ARNT
preimmune
(+)
anti-HIF3α
anti-ARNT
preimmune
HIF1α
HIF2αa
HIF3α


A
HIF1α
HIF2αa
HIF3α
Fold Increase
50
40
30
20
10
0
0
10
20
40
80
160
ng plasmid
B
HIF1α
HIF2αa
HIF3α
Fold Increase
6
4
2
0
0
0.5
1.0
2.0
4.0
8.0
ng plasmid
C
HIF1α
HIF2αa
HIF3α
Fold Increase
20
15
10
5
0
0
10
20
40
80
160
ng plasmid
D
(+)
0
40
80
160
HIF1α
HIF2αa
HIF3α
(+)
400
400
400
400
ARNT
E
Fold Increase
30
20
10
0
1
2
3
1+2
1+3
2+3
HIFα plasmid


A
HIF 1α
Transcript Abundance
(copy number per 50 ng RNA)
30000
20000
10000
0
B
a
a,b
b,c
b,c
b,c
c
A
c
A
c
B
b,c
A
c
Brain
Heart
Gill
Liver
Intestine
Kidney
Muscle
Spleen
Ovary
Testes
B
HIF 2αa
Transcript Abundance
(copy number per 50 ng RNA)
30000
20000
10000
0
A
a
A
a,b
c
c
c
b,c
B
a
B
b,c
C
a,b
B
b,c
Brain
Heart
Gill
Liver
Intestine
Kidney
Muscle
Spleen
Ovary
Testes
C
HIF 3α
Transcript Abundance
(copy number per 50 ng RNA)
30000
20000
10000
0
C
a
B
a
b,c
b,c
c,d
c,d
B
b
A
d
A
d
A,B
c,d
Brain
Heart
Gill
Liver
Intestine
Kidney
Muscle
Spleen
Ovary
Testes