Sequence and functional characterization of hypoxia inducible factors, HIF1α, HIF2α, and HIF3α, from the estuarine fish, *Fundulus heteroclitus*

Ian K. Townley\textsuperscript{1,2}, Sibel I. Karchner\textsuperscript{3}, Elena Skripnikova\textsuperscript{1,2}, Thomas E. Wiese\textsuperscript{2}, Mark E. Hahn\textsuperscript{3}, and Bernard B. Rees\textsuperscript{1}

\textsuperscript{1}Department of Biological Sciences, University of New Orleans, New Orleans, LA 70148
\textsuperscript{2}College of Pharmacy, Xavier University of New Orleans, New Orleans, LA 70125
\textsuperscript{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), HIF2α (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 MgCl₂, 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-<sup>32</sup>P]ATP (PerkinElmer), annealing with equimolar amounts of complementary strand, and desalting on ProbeQuant G-50 microcolumns (GE Healthcare). Incorporation of [<sup>32</sup>P] was assessed by liquid scintillation counting (BeckmanCoulter). Labeled probe was included at 40
fmol and approximately $10^5$ cpm in each reaction. When used, double stranded competitor DNA at 100 to 800-fold molar excess or antibodies against HIFα subunits (see above) or ARNT (MA-515; ThermoFisher) were pre-incubated in the reaction mixture for 15 min at room temperature prior to the addition of probe. After probe addition, reactions were incubated a further 15 min, transferred to 4°C, and electrophoresed on 5% polyacrylamide gels (37.5:1, acrylamide:bis-acrylamide) in 0.3 x Tris-borate-EDTA buffer (TBE) at 15 mA for 1.5 h at 4°C. Gels were dried and imaged by phosphor imaging (Bio-Rad). Band density was quantified using Quantity One (Bio-Rad).

**HIFα over-expression and reporter gene expression.** The ability of HIF proteins to activate transcription was determined by transient transfection into COS-7 monkey kidney cells, an established cell line that has been used previously to assess HIF function (21). COS-7 cells (American Type Culture Collection) were grown in high-glucose Dulbecco’s modified Eagle medium (DMEM; GIBCO) containing 10% fetal bovine serum (FBS; HyClone) in 5% CO$_2$ at 37°C as described earlier (22). Transient transfection of cells was used to assess the effect of increasing doses of plasmids encoding each *F. heteroclitus* HIFα on the expression of firefly luciferase reporter plasmids, N443pGL3 (50), which has the promoter of *F. heteroclitus Ldh-B* including a characterized HRE upstream of the luciferase gene, and p3XHifREluc (11), which has three copies of the HRE from the mouse heme oxygenase gene upstream of the luciferase gene. All experiments included a plasmid encoding *Renilla* luciferase under the control of the thymidylate kinase promoter (pRL-TK; Promega) to control for transfection efficiency.

One day prior to transfection, cells were plated at $10^5$ cells per well in 12-well plates. Replicate wells were transfected in 1 ml serum-free DMEM containing Lipofectamine 2000 (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, HIF2αα, 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
(Promea). 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 + 2αα; 1α + 3α; 2αα + 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, HIF2αα, 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).
Quantitative PCR. Total RNA was isolated from frozen tissues (liver, white skeletal muscle, intestine, gill, kidney, brain, heart, spleen, ovary, and testes) of six normoxic *F. heteroclitus* (three females, three males) as described above. Gill and intestine were first allowed to thaw in RNA later-ICE (Ambion) to allow dissection of gill lamella away from the gill arch or to flush the contents of the intestine. After isolation, RNA quality was assessed by Experion capillary electrophoresis (Bio-Rad) and RNA quantity was determined by spectrophotometry. Two RNA samples from gill and one from kidney were removed from this analysis due to extensive RNA degradation.

For all samples, 1 µg of RNA was reverse transcribed into cDNA using iScript reverse transcriptase (Bio-Rad) in a final volume of 20 µl. One microliter of the product (equivalent to 50 ng of total RNA) was used for quantitative PCR using iQ SYBR Green Supermix (Bio-Rad) and primers specific for HIF1α, HIF2α, or HIF3α (Table 1) in a reaction volume of 20 µl. One primer of each primer pair was designed to span a putative exon-exon boundary (determined by comparison of *F. heteroclitus* HIFα cDNAs with their homologs in the *Fugu* genome), and each primer pair was designed to amplify a product of 100 to 150 bp. PCR was conducted in a 96-well format in a MyiQ or iQ5 thermocycler (Bio-Rad) using the following program: 3 min 95°C; 40 cycles of 15 s at 95°C and 1 min at 68°C; 10 min at 72°C. Melt curves were done for each sample (81 30-s steps of 0.5°C from 55 to 95°C) to ensure the presence of a single amplification product in each reaction. Selected reactions were sequenced and shown to be the anticipated HIFα amplimers. No amplification was observed in PCR of samples that were not reverse transcribed or in reactions lacking template.

All samples were analyzed in duplicate. Each plate also contained duplicate wells of four 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 (Ct) 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 10⁶ 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 HIF2aa 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 ATGTG rather than the consensus sequence of ACGTG, 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 ATGTG, with lower affinity than they did to a consensus HRE (A CGTG), 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 ATGTG (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 \textit{F. heteroclitus} HIF1\(\alpha\) and 3\(\alpha\) bind to both sites equally well, but \textit{F. heteroclitus} HIF2\(\alpha\) binds more strongly to HBS1.

\textit{FhHIF\(\alpha\) 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 FhHIF1\(\alpha\), FhHIF2\(\alpha\)a, or FhHIF3\(\alpha\), along with \textit{F. heteroclitus} ARNT2 to assess the ability of their protein products to drive expression of a luciferase gene under the control of the \textit{F. heteroclitus} \textit{Ldh-B} promoter (50). All FhHIF\(\alpha\) 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 FhHIF1\(\alpha\) than when they were transfected with FhHIF2\(\alpha\)a or FhHIF3\(\alpha\), at all amounts of plasmid (2-way ANOVA, \(P < 0.001\)). To assess reporter gene expression at lower concentrations of the HIF\(\alpha\) plasmids, cells were transfected with 0 – 8 ng of each HIF\(\alpha\) plasmid. As observed at higher amounts of plasmid, FhHIF1\(\alpha\) induced greater reporter gene expression at these plasmid concentrations than did FhHIF2\(\alpha\)a or FhHIF3\(\alpha\) (Fig. 7B).

The above results were from experiments conducted with a reporter gene under the control of the \textit{F. heteroclitus} \textit{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 FhHIF1\(\alpha\) plasmid resulted in greater induction of luciferase expression compared to equivalent amounts of FhHIF2\(\alpha\)a or FhHIF3\(\alpha\)
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 FhHIF2α 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
greater reporter gene expression than HIF3α, similar to the pattern described here for *F. heteroclitus* HIF1α and HIF3α. Together, these results support the conclusion that HIF1α is a more effective activator of reporter gene expression than HIF3α, which is consistent with deduced protein structure of FhHIF1α having two TADs compared to a single TAD in FhHIF3α (Fig. 1). Although the lower reporter gene expression driven by FhHIF2αa in comparison to FhHIF1α may be the result of relatively poorer protein expression from the HIF2αa plasmid (Fig. 7D), differences in transcriptional activity by HIF1α and HIF2α have been reported in bichir and naked carp (4). In those species, which α subunit leads to higher reporter gene expression depends upon the species and whether cells were incubated at normal or reduced oxygen levels.

In mammals, shorter forms of HIF3α, resulting from differential splicing, act as repressors of HIF1α mediated gene expression (32, 33, 36). In one experiment, 80 ng of two HIFα plasmids were co-transfected into COS-7 cells (1α + 2αα; 1α + 3α; 2αα + 3α). The simultaneous expression of two HIFα plasmids neither augmented nor diminished luciferase gene expression (Fig. 7E). Under these conditions, therefore, there was no evidence that FhHIF3α reduced the level of reporter gene expression by either FhHIF1α or FhHIF2αa. Importantly, the current experiments were done in the presence of excess ARNT2: there is evidence that HIFα subunits compete for ARNT only when the latter is present in limiting quantities (7, 29).

*FhHIFα mRNA in normoxic *F. heteroclitus*. The abundance of each HIFα transcript was determined by quantitative RT-PCR of total RNA isolated from organs of normoxic *F. heteroclitus* (Fig. 8). One-way ANOVA demonstrated pronounced differences in transcript abundance among organs for all three HIFα transcripts (*P* < 0.001). The abundance of FhHIF1α transcripts was greatest in testes, spleen, skeletal muscle, and kidney, lowest in brain and heart,
and intermediate in gill, liver, intestine, and ovary (Fig. 8A). The abundance of FhHIF2αa transcripts was greatest in gill, liver, and intestine, lowest in brain and muscle, and intermediate in other organs (Fig. 8B). The abundance of FhHIF3α transcripts was greatest in spleen and ovary, lowest in brain and heart, and intermediate in other organs (Fig. 8C).

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

One striking feature of the present data is that F. heteroclitus brain and heart have very low levels of all HIFα transcripts. This result contrasts with the relatively high HIFα mRNA levels in brain and heart measured under normoxia in several other fishes (10, 49, 54, 56, 58) and does not support the idea that the levels of HIFα transcripts are highest in organs that have strict aerobic requirements (58). Another intriguing result was the high level of HIF2αa seen in F. heteroclitus gill. This observation confirms earlier results from semi-quantitative PCR (44), agrees with observations from other species (10, 56, 58), and suggests that HIF2αa may play an important role in gill, an organ responsible for gas exchange, as well as ionic and osmotic regulation. Measures of HIFα protein abundance will be necessary to clarify whether these organ-dependent mRNA levels are coupled to protein abundance and, hence, functional 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.
Although the physiological significance of polymorphism in *F. heteroclitus* HIFα genes, if any, is unknown, it is important to point out that genetic variation in mammalian HIFα genes is associated with several physiological traits or pathological conditions. Variation at the HIF1α locus has been associated with variation in maximum oxygen consumption of elite athletes (38, 47), diabetes prevalence (42), and cancer risk and progression (25, 27, 76). High altitude adaptation in Tibetans is associated with genetic variation in HIF2α, with specific SNPs correlated with reduced hematocrit in this population (1, 63, 72). Among the largest allele differences between Tibetan and non-Tibetans were non-coding SNPs (1, 72), suggesting that regulatory changes, rather than structural changes, in HIF2α may be important in adaptation to high altitude hypoxia. The current observation of multiple SNPs in *F. heteroclitus* HIFα genes, combined with the recent reports of polymorphism in FIH-1 from the cyprinid *Megalobrama amblycephala* (73) and evidence of selection acting on the HIF2αa locus in a pollution-tolerant population of *F. heteroclitus* (52), suggests that genetic variation in elements of the HIF signaling pathway could underlie differences among individuals or populations in their responses to aquatic hypoxia or other stressors.

**Perspectives and Significance.** The HIF family of transcription factors plays important roles in development, physiology, pathology, and environmental adaptation of animals. Herein, HIF1α and 3α were sequenced and characterized, along with the previously described HIF2αa, in *F. heteroclitus*, an important model for environmental biology (3, 61). Genomic and phylogenetic analyses clearly placed the *F. heteroclitus* HIFs with their euteleost paralogs. DNA binding and functional analyses showed that these HIFα subunits independently formed complexes with *F. heteroclitus* ARNT2 to bind to canonical and non-canonical HREs and drive reporter gene expression. Analysis of the *F. heteroclitus* genome uncovered a second HIF2α,
HIF2αb, which may be a “relic” of a teleost-specific genome duplication maintained in certain fish species (57). Although not functionally characterized here, the predicted protein lacks domains needed for oxygen sensing and activation of gene expression, suggesting that it may act to repress the transcriptional activity of other HIFα subunits, analogous to the role of mammalian HIF3α splice variants (32, 33, 36). Finally, cDNA sequencing and SNP analysis revealed polymorphism in *F. heteroclitus* HIF1α, HIF2αa, and HIF3α. Future research is needed to explore the whether this sequence variation is related to HIF expression, stability, function, and, ultimately, the organismal response to hypoxia. Genetic variation in these transcription factors could influence the capacity of individual killifish to tolerate low levels of oxygen and thereby affect the ability of populations to adapt to a changing aquatic environment.
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


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...


Table 1. Sequences of primers used in this study.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Use</th>
<th>Primer Sequence (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIF1-Forward</td>
<td>PCR</td>
<td>AGGAGGGGGAAGGAATCTGAGGT</td>
</tr>
<tr>
<td>HIF1-Reverse</td>
<td>PCR</td>
<td>ACAACACACTGGGGCTGGGAG</td>
</tr>
<tr>
<td>AP-1</td>
<td>RACE</td>
<td>CCATCCTAATACGACTCACATATAGGGCC</td>
</tr>
<tr>
<td>AP-2</td>
<td>RACE</td>
<td>ACTCACTATAAGGCCTGACAGGGCAG</td>
</tr>
<tr>
<td>HIF1 5’gs outer</td>
<td>RACE</td>
<td>GTCTCCGTCTTCAGACAGACCAG</td>
</tr>
<tr>
<td>HIF1 5’gs inner</td>
<td>RACE</td>
<td>AGCTGCATGTCCAGAGCTGTTTG</td>
</tr>
<tr>
<td>HIF1 5’gs alt</td>
<td>RACE</td>
<td>CTCTAAAGGCTTCAGGTAGGACC</td>
</tr>
<tr>
<td>HIF1 3’gs outer</td>
<td>RACE</td>
<td>CGGAGGACCTGCTGAACCGGTCTG</td>
</tr>
<tr>
<td>HIF1 3’gs inner</td>
<td>RACE</td>
<td>CAAAGACTCACCATCACCAGTTTGT</td>
</tr>
<tr>
<td>HIF1 3’gs alt</td>
<td>RACE</td>
<td>ACCACGCTGCTGGACTCGGACTATC</td>
</tr>
<tr>
<td>HIF3 5’gs outer</td>
<td>RACE</td>
<td>ACAGGGGTTCGGTTTTGGTAG</td>
</tr>
<tr>
<td>HIF3 5’gs inner</td>
<td>RACE</td>
<td>TGTTGTTGTAGGACCGGTGTC</td>
</tr>
<tr>
<td>HIF3 3’gs outer</td>
<td>RACE</td>
<td>TGACCTTCCTCAGCGATTGTC</td>
</tr>
<tr>
<td>HIF3 3’gs inner</td>
<td>RACE</td>
<td>AAGCAAGAGGCCTCCTCCTCAGAT</td>
</tr>
<tr>
<td>HIF1 3’UTR</td>
<td>PCR</td>
<td>GAAACCAAGGAGAACTCTCTATGT</td>
</tr>
<tr>
<td>HIF1 3’UTR</td>
<td>PCR</td>
<td>TAAATAAGTGCCAGTGGAACAG</td>
</tr>
<tr>
<td>HIF3 5’UTR</td>
<td>PCR</td>
<td>AACAGAGAGCACCTGGAATTGCT</td>
</tr>
<tr>
<td>HIF3 3’UTR</td>
<td>PCR</td>
<td>GGCGAGGCTTCTTGCTGATC</td>
</tr>
<tr>
<td>HIF2 5’subclone</td>
<td>Ab production</td>
<td>TTCACTGCAAGCCAGAGTGATTACTACAGC</td>
</tr>
<tr>
<td>HIF2 3’subclone</td>
<td>Ab production</td>
<td>TTCACTGCAACCCTCCTGGAATCACAG</td>
</tr>
<tr>
<td>HIF3 5’subclone</td>
<td>Ab production</td>
<td>TTCACTGCAAGGCTCCTGGAATGCT</td>
</tr>
<tr>
<td>HIF3 3’subclone</td>
<td>Ab production</td>
<td>TTCACTGCAAGGCTCCTGGAATGCT</td>
</tr>
<tr>
<td>HIF1-F1</td>
<td>QPCR</td>
<td>CAAGTCCGGCTACCTGGAAGGT</td>
</tr>
<tr>
<td>HIF1-R1</td>
<td>QPCR</td>
<td>CAGATCAGGACCAGATAGGGGAC</td>
</tr>
<tr>
<td>HIF2-F1</td>
<td>QPCR</td>
<td>GCCTGCACAGCTCCTGGAATAC</td>
</tr>
<tr>
<td>HIF2-R1</td>
<td>QPCR</td>
<td>CTGTGACCTGTGAGCTCCACCT</td>
</tr>
<tr>
<td>HIF3-F1</td>
<td>QPCR</td>
<td>GTCAGACCATCCGGCATCAG</td>
</tr>
<tr>
<td>HIF3-R1</td>
<td>QPCR</td>
<td>CCATCAGGTTTCTTACTAGACCT</td>
</tr>
</tbody>
</table>

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.

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

*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).

bValues are from BLOSUM62 alignment score matrix (17). Positive scores are conservative substitutions and negative scores are non-conservative substitutions.

cDomain 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.
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
HIF sequence was aligned using ClustalX. All positions containing gaps or missing data were eliminated. Killifish sequences are shown in grey boxes. Species and accession numbers are:

- human (*Homo sapiens*) HIF1α NP_001521, HIF2α NP_001421, HIF3α NP_690007; mouse (*Mus musculus*) HIF1α NP_001300848, HIF2α NP_034267, HIF3α NP_058564; lizard (*Anolis carolinensis*) HIF1α XP_008121253; HIF2α XP_003225226, HIF3α XP_016853234; killifish (*Fundulus heteroclitus*) HIF1α ALL26120, HIF2αa NP_001296843, HIF2αb XP_012709515, HIF3α ALL26129; tilapia (*Oreochromis niloticus*) HIF1α XP_005477095, HIF2α XP_003438301, HIF2αa NP_001296843, HIF2αb XP_003441929, HIF3α XP_005461198; croaker (*Micropogonias undulatus*) HIF1α ABD32158, HIF2α ABD32159; zebrafish (*Danio rerio*) HIF1α a NP_001295488, HIF1α b NP_001296971, HIF2αa NP_001034895, HIF2αb XP_695262, HIF3αa AAR95697, HIF3αb AAR95698; asp (*Aspius aspius*) HIF1αa AFD32323, HIF1αb ABO26713, HIF2α AFD32324. Ensembl Release 85 predicted proteins: green pufferfish (*Tetraodon nigroviridis*) HIF2αb ENSTNIP00000008013 and stickleback (*Gasterosteus aculeatus*) HIF2αb ENSGACP00000003681.

Fig. 3. Genomic organization of *F. heteroclitus* HIFα genes is conserved. The flanking genes (Fig. 3A) and gene models (Fig. 3B) of killifish HIF1α, HIF2αa, HIF2αb, and HIF3α genes are shown. Flanking genes were obtained from the *F. heteroclitus* genome sequence database v2b and compared to green pufferfish, stickleback, tilapia, and zebrafish genomes using Genomicus version 85.01 (http://www.genomicus.biologie.ens.fr). The orientation of the flanking genes is shown by arrows above the gene abbreviations. Italic letters below gene names indicate shared 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).
Fig. 5. F. heteroclitus HIFs bind a putative HRE from the F. heteroclitus Ldh-B promoter. DNA sequences (A) of oligonucleotides corresponding to the wild type or mutated versions of a putative HRE from the F. heteroclitus Ldh-B promoter (50). The wild type HRE (wt/wt) contains two HIF binding sites (HBSs, bold type), whereas the mutated versions (mut/mut; mut/wt; wt/mut) have changes at one or both HBS (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 wt/wt 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 wt/wt 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 = 3 – 4). 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).

Fig. 6. Antibodies against F. heteroclitus HIF subunits supershift specific EMSA bands. All reactions included one of three IVTT HIFα subunits and F. heteroclitus ARNT2 (+). Additional components were chicken polyclonal antibodies against each α subunit (lanes labeled anti-HIF1α, anti-HIF2αa, and anti-HIF3α, respectively), or a mouse monoclonal antibody against human ARNT (anti-ARNT). Control reactions included pre-immune IgY. Arrows indicate the bands whose mobility is specifically shifted by antibodies against HIFα subunits or ARNT, but
not by pre-immune IgY. The labeled DNA probe for HIF1α and HIF3α was 18 bp from the human EPO 3’ enhancer (EPO-w18), and for HIF2αa the DNA probe was 32 bp from the F. heteroclitus Ldh-B promoter (see Figs. 4A and 5A for sequences).

Fig. 7. *F. heteroclitus* HIFs induce HRE-dependent reporter gene expression in a dose-dependent fashion. COS-7 cells were co-transfected with a plasmid encoding firefly luciferase under the control of the *F. heteroclitus* Ldh-B promoter containing a putative HRE and increasing amounts of plasmids encoding each HIFα subunit. Part (A) shows the increase in luciferase expression (mean ± one standard deviation of three experiments) in COS-7 cells transfected with 10 – 160 ng of each FhHIFα plasmid relative to COS-7 cells transfected with empty pcDNA (0 ng). Part (B) shows results of a single experiment in which COS-7 cells were co-transfected with lower amounts of each FhHIFα plasmid (0 – 8 ng), showing greater expression driven by FhHIF1α even at low plasmid doses. Part (C) shows results when the reporter plasmid encoded firefly luciferase under the control of the HRE from mouse heme oxygenase gene and COS-7 cells were co-transfected with increasing amounts of FhHIFα plasmids (0 – 160 ng). Part (D) shows FhHIFα and ARNT2 protein levels determined in parallel by western blotting. *F. heteroclitus* ARNT2 plasmid was included in all transfections at a single dose (400 ng). The (+) lane indicates IVTT of each respective protein included in one lane of the gel as a positive control. Part (E) shows results of a single experiment in which COS-7 cells were co-transfected with a plasmid encoding firefly luciferase under the control of the *F. heteroclitus* Ldh-B promoter and either 80 ng of one FhHIFα plasmid or 80 ng each of two FhHIFα plasmids.
Fig. 8. HIFα transcripts are expressed in multiple organs of normoxic *F. heteroclitus*. The levels of HIF1α (A), HIF2αa (B), and HIF3α (C) mRNA were determined in organs of adult *F. heteroclitus* by quantitative PCR. For a given HIFα subunit, the effects of organ were determined by one-way ANOVA. Organs with significantly different abundance of a given HIFα mRNA are indicated by having different lower case letters (Bonferroni’s post-hoc comparisons, *P* < 0.05).

In each organ, differences in transcript abundance among the HIFα forms were assessed by a one-way ANOVA and differences among HIFα forms within a given organ are indicated by uppercase italic letters (comparing across panels A-C). Error bars are one standard deviation (*n* = 6, except for kidney where *n* = 5 and gill where *n* = 4).
A

<table>
<thead>
<tr>
<th>EPO-w18</th>
<th>GCCCTAAGTGCTGCTTCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPO-m18</td>
<td>GCCCTAAAGTGCTTCA</td>
</tr>
<tr>
<td>HBS1-wt</td>
<td>GCTGTGAGTGACATCC</td>
</tr>
<tr>
<td>HBS2-wt</td>
<td>ATCCCTGCACATCTGTTCA</td>
</tr>
</tbody>
</table>

B

C

D

E