Sequence Analysis and Phylogenetic Studies of Hypoxia-Inducible Factor-1a

Jagadeesha Poyya¹, Chandrashekhar G Joshi¹, D Jagadeesha Kumar² and HG Nagendra²

¹Department of Biochemistry, Mangalore University, Post Graduate Centre, Chikka Aluvara, Kodagu, Karnataka, India. ²Department of Biotechnology, Sir M. Visvesvaraya Institute of Technology, Bangalore, India.

ABSTRACT: Hypoxia-inducible factors (HIF) belong to the basic helix loop helix-PER ARNT SIM (bHLH-PAS) family of transcription factors that induce metabolic reprogramming under hypoxic condition. The phylogenetic studies of hypoxia-inducible factor-1a (HIF-1a) sequences across different organisms/species may leave a clue on the evolutionary relationships and its probable correlation to tumorigenesis and adaptation to low oxygen environments. In this study, we have aimed at the evolutionary investigation of the protein HIF-1α across different species to decipher their sequence variations/mutations and look into the probable causes and abnormal behaviour of this molecule under exotic conditions. In total, 16 homologous sequences for HIF-1a were retrieved from the National Center for Biotechnology Information. Sequence identity was performed using the Needle program. Multiple aligned sequences were used to construct the phylogeny using the neighbour-joining method. Most of the changes were observed in oxygen-dependent degradation domain and inhibitory domain. Sixteen sequences were clustered into 5 groups. The phylogenetic analysis clearly highlighted the variations that were observed at the sequence level. Comparisons of the HIF-1a sequence among cancer-prone and cancer-resistant animals enable us to find out the probable clues towards potential risk factors in the development of cancer.

KEYWORDS: Hypoxia-inducible factor-1a (HIF-1a), hypoxia, phylogeny, oxygen-dependent degradation domain, cancer resistance

RECEIVED: January 30, 2017. ACCEPTED: May 3, 2017.

PEER REVIEW: Five peer reviewers contributed to the peer review report. Reviewers reports totalled 1411 words, excluding any confidential comments to the academic editor.

TYPE: Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article

Introduction

High proliferation and less vasculature are features under hypoxia in most of the solid tumours. Hypoxic phenotype is the aggressiveness of the tumour with poor prognostic effect.¹ Tumour tissues are adapted to this condition with metabolic reprogramming with the expression of hypoxia-inducible factors (HIFs). Hypoxia-inducible factors belong to the bHLH-PAS family of transcription factors that induce metabolic reprogramming under hypoxic condition.² Hypoxia-inducible factor is a heterodimer of HIF- α (containing 826aa) and HIF- β (containing 824aa) subunits. Hypoxia-inducible factor-1 α exists in 3 isoforms, namely, HIF-1 α , HIF-2 α , and HIF-3 α . The HIF-1 β subunits are also known as aryl hydrocarbon receptor nuclear transporters (ARNTs). The ARNTs are ubiquitously expressed in cells and are stable under normoxic conditions. Recent studies have shown that ARNT level is influenced by hypoxia, and it depends on the cell type.³ However, the HIF-1 α subunits are sensitive to oxygen and undergo oxygen-dependent proteasomal degradation, which is regulated by 2 different mechanisms. The first mechanism involves prolyl hydroxylase 2 (PHD2) that adds hydroxyl groups to proline residues (at 402 and 564) present on the oxygen-dependent degradation domain (ODD) of HIF-1a. Von Hippel-Lindau tumour suppressor protein (pVHL) recognizes and binds to the hydroxylated proline residues and directs HIF-1 α for proteasomal degradation by interacting with proteasomal E3 ligase complex. Another complementary

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article

CORRESPONDING AUTHOR: Chandrashekhar G Joshi, Department of Biochemistry, Mangalore University, Post Graduate Centre, Chikka Aluvara, Kodagu, Karnataka 571232, India. Email: josheejoshee@gmail.com

mechanism which regulates HIF-1 α level is factor inhibiting HIF (FIH). It is an asparagine hydroxylase enzyme that adds hydroxyl group to asparagine residue (803) present in C-terminal transactivation domain (C-TAD) of HIF-1a and prevents the interaction with CBP (CREB-binding protein) and p300 transcriptional coactivators, which also promotes proteasomal degradation of HIF-1a.4-6

Hypoxia-inducible factor- 1α has a half-life of less than 5 minutes under normoxic conditions⁴ and is involved in foetal and postnatal physiology.⁷ However, under hypoxic conditions, HIF-1 and HIF-2 α help in developing resistance to chemotherapy and radiotherapy due to downstream effects.^{5,6,8–10} The HIF-1 α plays a major part in the adaptation of solid tumour to low oxygen levels¹ that exists at high altitude¹¹ and deep aquatic and subterranean environments,¹² as its lifetime increases to 8 minutes and develops stability.¹¹ Hence, in this study, we have aimed at the evolutionary investigation of the protein HIF-1 α across different species to decipher their sequence variations/ mutations and investigate the probable causes and abnormal behaviour of this molecule under exotic conditions.

The phylogenetic studies of HIF-1 α sequences across different organisms/species may throw light on the evolutionary correlations to tumorigenesis and adaptation to low oxygen environments. This understanding may also aid in developing novel strategies to combat cancer while exploring approaches to enhance the efficacy of radiotherapy and chemotherapy.

 \bigcirc

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Cancer Informatics Volume 16: 1-8 © The Author(s) 2017 Reprints and permissions: sagepub.co.uk/journalsPermissions.nav DOI: 10.1177/1176935117712242



Materials and Methods

Sequencing analysis

The protein sequence of human HIF-1 α (hHIF-1 α ; isoform 1) was retrieved from Universal Protein Resource (UniProt),13 which contained 826 amino acids. To examine the occurrence and evolutionary correlations of HIF-1 α in 16 different species, namely, chimpanzee (Pan troglodytes: JAA40616.1), Sumatran orangutan (Pongo abelii: NP_001126975.1), rhesus macaque (Macaca mulatta: H9ET32), crab-eating macaque (Macaca fascicularis: NP_001270825.1), dog (Canis lupus familiaris: NP_001274092.1), yak (Bos grunniens: Q0PGG7.1), cow (Bos taurus: NP_776764.2), Tibetan antelope (Pantholops hodgsonii: AAX89137.1), goat (Capra aegagrus hircus: NP_001272657.1), blind mole-rat (BMR; Nannospalax judaei: CAG29396.1), rat (Rattus norvegicus: NP_077335.1), mouse (Mus musculus: AAC53461.1), naked mole-rat (NMR; Heterocephalus glaber: NP_001297196.1), naked carp (Gymnocypris przewalskii: AAW69834.1), and zebrafish (Danio rerio: AAQ91619.1), respectively. Global alignment (Needle) with BLOSUM62 (Gap open penalty 10, Gap extend penalty 2)14-16 and multiple alignment via the tool MultAlin¹⁷ were performed. These sequences were selected on the basis of hypoxic-adapted and normoxic animals.

Human HIF-1a has 10 domains, namely, bHLH (17-70), PAS-A (85-158), PAS-B (228-298), PAS-C (302-345), ODD (401-603), N-terminal VHL recognition site (380-417), C-terminal VHL recognition site (556-572), inhibitory domain (ID; 576-785), N-terminal transactivation domain (N-TAD; 531-575), and C-TAD (786-826), respectively. Hence, domain-wise comparisons of hHIF-1 α with all other organisms were also performed. The alignments were manually curated to remove poorly aligned regions. The output was studied for the position-wise conservation and variations/ mutations of amino acids across all the 10 domains of HIF-1 α , which are depicted in Figures 1 to 3. Multiple sequence alignment of all 16 protein sequences was deduced using blosum62 with default parameters to observe the various conserved domains across evolution (Corpet1988). The same is also illustrated in Figures 1 to 3.

Phylogenetic analysis

The 16 sequences were then used for phylogenetic inferences and evolutionary trees for respective domains of HIF-1 α . The output was constructed with MEGA 6.0¹⁹ and is highlighted in Figure 4. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analysed.²⁰ The neighbour-joining method was used to infer the evolutionary history.²¹ The bootstrap (value 100) was repeated 1000 times to generate consensus. The branches that are not reproduced less than 50% of the time during bootstrap were collapsed. Poisson correction method was used to compute the evolutionary distances.²² All positions containing gaps and missing data were eliminated.

Results

This study was aimed at deciphering the amino acid variation and phylogenetic relations across the 16 orthologs of HIF-1 α . The identities between the sequences were calculated with reference to hHIF-1 α . The tabulated results are presented in Supplementary Table 1. Most species exhibit lower identity values for the ODD and IDs, respectively. The values in Supplementary Table 1 highlight that the variations across all domains are maximum for aquatic fishes, naked carp, and zebrafish, respectively. Similar results were exhibited during domain-wise comparison via multiple sequence alignment. The conservation of functional sites in the bHLH domain of HIF-1 α was observed across all species considered for the study. Apart from the DNA-binding residues, which were highly conserved, the changes were observed as in Figure 1A. The DNA-interacting basic amino acids, namely, K21, R29, K19, R30, R27, D55, and K56,¹⁸ appear to be conserved across the 16 species, indicating that DNA-binding function is intact. Maximum difference was observed in naked carp, which exhibits the following polymorphisms with respect to the hHIF-1 α . The classical changes S31C, F34Y, H42Q, S51T, V59I, L67M, and V69T were, respectively, observed in the bHLH domain of naked carp. Similarly, in the PAS-A domain, permissible substitutions such as K85E, A86S, V100L, T104S, D105E, I112L, D114E, N117S, Y119S, E126D, V132I, D134E, T136S, M144L, T149V, N152T, L154S, V155K, and K157T were observed in both naked carp and zebrafish, with respect to hHIF-1 α . The heterodimerization of HIF-1 α with ARNT is crucial for its transactivity. PAS domains of these 2 proteins play key roles in heterodimerization. As Bersten et al²³ have highlighted in his review, the PAS-A domain is less accessible to small and other sensory molecules than the PAS-B. These polymorphisms found in PAS-A may affect the overall activity of HIF-1 α . It is well known that PAS-B is the key domain that establishes protein-protein interaction with ARNT and HSP90.^{24,25} Residues interacting with ARNT were conserved across the 16 selected sequences. However, with PAS-B exhibiting changes, it might affect HSP90 interaction.

Most of the changes in PAS-B domain were observed in naked carp and zebrafish. I235V, S247T, E268D, E269D, G272N, I275V, M294L, and T296A substitutions were found in both naked carp and zebrafish. PAS-C showed the following changes, namely, R311K polymorphism in naked carp, zebrafish, and yak, respectively. Substitutions such as Y314F, T327P, and V342L were observed in naked carp and zebrafish. N-terminal VHL recognition site of hHIF-1 α also shows the following changes with respect to naked carp, which denote interesting substitutions, namely, S380P, E381M, D382K, T383N, S384L, S385D, K392E, D395E, and L399V. Similarly, the changes E381L, D382E, T383S, S385T, E393D, P394S, and T407A are seen in zebrafish. Similarly, interesting

	20 30	40	50	60	70	
Human	R R K E K S R D A A R <mark>S</mark> R R	SKESEVFYELAHQ	LPLPHNVSSHL	KASVMRLTIS	SYLRVR	
rhesus-monkey	R R K E K S R D A A R <mark>S</mark> R R	SKESEVFYELAHQ	LPLPHNVSSHLD	KASVMRLTIS	SYLRVR	
Cow	RRKEK SRDAARSRR	SKESEVFYELAHO	LPLPHNVSSHLD	KASVMRLTIS	SYLRVR Sylrvr	
Chimpanzee	RRKEKSRDAARSRR	SKESEVFYELAHO	LPLPHNVSSHLD	KASVMRLTIS	5 YLRVR	
Sumatran-orangutan	R R K E K S R D A A R <mark>S</mark> R R	SKESEVFYELAHQ	LPLPHNVSSHLD	KASVMRLTIS	SYLRVR	
NMR	RRKEKSRDAARSRR	SKESEVFYELAHQ	LPLPHNVSSHLD	KASVMRLTIS	SYLRVR	
Rat	RRKEKSRDAARSRR	SKESEVFYELAHO	LPLPHNVSSHLL	OKASVMRLTIS OKASVMRLTIS	SYLRVR	
BMR	R R K E K S R D A A R <mark>S</mark> R R	SKESEVFYELAHQ	LPLPHNVSSHL	KASVMRLTIS	YLRVR	
Tibetan-antelope	RRKEKSRDAARSRR	SKESEVFYELAHQ	LPLPHNVSSHLD	KASVMRLTIS	SYLRVR	
Yak Crab-eating-macaque	RRKEKSRDAARSRR	SKESEVFYELAHQ Skeseveyelaho	LPLPHNVSSHLL	KASVMRLTIS	SYLEVE	
Goat	RRKEKSRDAARPRR	SKESEVFYELAHQ	LPLPHNVSSHLD	KASVMRLTIS	SYLRVR	
Zebra-fish	R R K E K S R D A A R <mark>S</mark> R R	GKESEVFYELAHQ	LPLPHNVTSHL	KASIMRLTIS	SYLRMR	
Naked-carp	RRKEKSRDAARSRR	GKECEVYYELAQQ gKEgEV%YELADO	LPLPHNVISHLU LDLPHNVgSHI	KASIMRLTIS KASIMPLTIS	SYMRTR SV S B v B	
consensus/sv	KARERDRURARDAK	PKEPEVOLEDANY	A	ARD . MADI I.	JIQKVK	
			A			1.5.0
	90 10		120	130	140	150
Chimpanzee	KAQMNCFYLKALDG KAOMNCFYLKALDG	FVMVLTDDGDMIY	ISDNVNKYMGLT ISDNVNKYMGLT	QFELTGHSVF OFBLTGHSVF	DETHECOHEEMRE	MLTHRNGLVKKG
Dog	KAQMNCFYLKALDGI	FVMVLTDDGDMIY	ISDNVNKYMGLT	QFELTGHSVF	DFTHPCDHEEMR	MLTHRNGLVKKG
Cow	KAQMNCFYLKALDG	FVMVLTDDGDMIY	ISDNVNKYMGLT	QFELTGHSVF	DFTHPCDHEEMRE	MLTHRNGLVKKG
Goat	KAQMNCFYLKALDG	FVMVLTDDGDMIY	ISDNVNKYMGLT ISDNVNKYMGLT	QFELTGHSVF	DETHPCDHEEMRE	MLTHRNGLVKKG
Tibetan-antelope	KAOMNCFYLKALDGI	FVMVLTDDGDMIY	ISDNVNKYMGLT	QFELTGHSVF	DFTHPCDHEEMRI	MLTHRNGLVKKG
rhesus-monkey	KAQMNCFYLKALDG	FVMVLTDDGDMIY	ISDNVNKYMGLT	QFELTGHSVF	DFTHPCDHEEMR	EMLTHRNGPVKKG
Sumatran-orangutan	KAQMNCFYLKALDG	FVMVLTDDGDMIY	ISDNVNKYMGLT	QFELTGHSVF	DETHPCDHEEMRE	ML THRNGPVKKG
Rat	KAOMNCFYLKALDGI	FVMVLTDDGDMIY	ISDNVNKYMGLT	OFELTGHSVF	DFTHPCDHEEMRE	MLTHRNGPVRKG
Blind-Mole-rat	KAQMNCFYLKAL DGI	FVMVLTDDGDMIY	ISDNVNKYMGLT	QFELTGHSVF	DFTHPCDHEEMR	MLTHRNGPVKEG
Crab-eating-macaque	KAQMNCFYLKALDG]	FVMVLTDDGDMIY	ISDNVNKYMGLT	QFELTGHSVF	DFTHPCGHEEMRE	MLTHRNGPVKKG
Naked-carp	ESOLNSFYLKALEG	FLVVLSEDGDMVY	LSENVSKSMGLT	OFDLTGHSIF	EFSHPCDHEELRE	MLVHRTGSKKT
Zebra-fish	ESQLNGFYLKALEGI	FLMVLSEDGDMVY	LSENVSKSMGLT	QFDLTGHSIF	EFSHPCDHEELRE	MLVHRTGSKKT.
consensus>50	kaQ\$#cFYLKAL#G	FvmVLt#DGDM!Y	iS#NVnKyMGLT	QF#LTGHS!F	#FtHPCdHEE\$rH	MLtHRnG.vkkg
			R			
			В			
	230 240	250	B ¢ o	270	280	290
Human	230 240 Phpsnie <mark>i</mark> pldsk	250 TFLSRHSLDMKFS	B ⊿60 YCDERITELMG	270 YEPEELLGRS	280 IYEYYHALDSDH	290 LTKTHHDMFTKG
Human rhesus-monkey Corr	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS	∠ 60 YCDERITELMG YCDERITELMG	270 YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	∠ 60 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee	230 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	∠60 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan	230 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	∠ 60 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 YEPEELIGRS YEPEELIGRS YEPEELIGRS YEPEELIGRS YEPEELIGRS YEPEELIGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	2 6 0 Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat	230 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	2 6 0 Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG Y CDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	A 60 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	Z ⁴⁶⁰ YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	A 60 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	Z 6 0 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	Z 6 0 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 $Y = P = E \perp L G R S$	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH ISEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHMFTKG LTKTHHMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus 50	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	Z 6 0 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 $Y = P = E \perp L G R S$ $Y = P = E \perp L R R$ $Y = P = E \perp$	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH YEYYHALDSDH VYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	Z 6 0 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	270 Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = E L L G R S Y = P = D L L N R S Y = P = D L L N R S Y = P = H + L G R S	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK	2 5 0 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS	A 60 YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG YCDERITELMG	2 7 0 YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK	250 TFLSRHSLDMKFS T	A CDERITELMG YCDERITELMG	2 7 0 YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHMFTKG LTKTHHMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human	230 240 PHPSNIEIPLDSK	2 5 0 TFLSRHSLDMKFS	A 6 0 Y CDERITELMG Y CDERITELMG	270 YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog	230 240 PHPSNIEIPLDSK	250 TFLSRHSLDMKFS T	A 6 0 YCDERITELMG	270 YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDS	250 TFLSRHSLDMKFS TFLSRHSLMS TFLSRHSLDMKFS TFLSRHSLMS T	A 60 YCDERITELMG	270 YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDS	250 TFLSRHSLDMKFS T	A 6 0 Y CDERITELMG Y CDERI Y CDERI Y CDERITELMG Y CDERI Y C	270 YEPELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH ISEYYHALDSDH YEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Pat	230 240 PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDS	250 TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLDMKFS STSLSRHTLDMKFS STSLSRHSLDMKFS STSLSRHSLDMKFS STSLSRHSLDMKFS STSLSRHSLDMKFS STSLSRHSLDMKFS STSLSRHSLDMKFS STSLSRHSLDMKFS STSLSRHSLDMKFS STSLSRHTLDMKFS STSLSRHTLDMKFS STSLSRHTLDMKFS STSLSRHTLDMKFS STSLSRHTLDMKFS STSLSRHTLDMKFS STSLSRHTLDMKFS STSLSRHSLDMKFS STSLSRHTLDMKFS STSLS	A 60 YCDERITELMG	270 YEPELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH ISEYYHALDSDH ISEYYHALDSDH YEYYHALDSDH ISEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Rat Blind-Mole-rat	230 240 PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK TGQYRMLAKRGGYVY TGQYRMLAKRGGYVY TGQYRMLAKRGGYVY	250 TFLSRHSLDMKFS TFLSRHSLMKFS NVETQATVIYNTK NVETQATVIYNTK VETQATVIYNTK VETQATVIYNTK VETQATVIYNTK VETQATVIYNTK VETQATVIYNTK VETQATVIYNTK	A 60 YCDERITELMG YCDERI YCDERITELMG YCDE	270 YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH YEYYHALDSDH YEYYHALDSDH YEYYHALDSDH YEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDFFKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Rat Blind-Mole-rat Sumatran-orangutan	230 240 PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDS	250 TFLSRHSLDMKFS TFLSRHSLM TFLSRHSLDMKFS TFLSRHSLDMKFS TFLSRHSLMKS TFLSRHSLM T	B 260 YCDERITELMG YCDERI YCDERITELMG YCDERITELMG YCDERITELMG YCDERI Y	270 YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH YYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Rat Blind-Mole-rat Sumatran-orangutan Crab-eating-macaque	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDS	250 TFLSRHSLDMKFS TFLSRHSLS TFLSRHSLDMKFS TFLSR	B 260 YCDERITELMG YCDERITELM	270 YEPELLGRS YEPEELS YEPEELGRS YEPEELGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELG	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHMFTKG LTKTHHMFFKG LTKTHHMFFKG LTKTHHMFFKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Sumatran-orangutan Crab-eating-macaque Tibetan-antelope Goat	230 240 PHPSNIEIPLDSK PHPSNIEVPLDSK S10 TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV	250 TFLSRHSLDMKFS TFLSRHSLS TFLSRHSLDMKFS TFLSR	B 260 YCDERITELMG YCDERITELM	270 YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNFFKG LTKTHHNFFKG LTKTHHNFFKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Rat Blind-Mole-rat Sumatran-orangutan Crab-eating-macaque Tibetan-antelope Goat Cow	230 240 PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV TGQYRMLAKRGGYVV	250 TFLSRHSLDMKFS TFLSRHSL TFLSRHSLDMKFS TFLSRH	A 6 0 YCDERITELMG	270 YEPELLGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Rat Blind-Mole-rat Sumatran-orangutan Crab-eating-macaque Tibetan-antelope Goat Cow Yak	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDS	250 TFLSRHSLDMKFS TFLSRHSLS TFLSRHSLDMKFS TFLSR	A 60 YCDERITELMG YCDERI YCDERI YCDERITELMG YCDERITELMG YCDERI YCDR YCDR YCDR YCDR YCDR YCDR Y	270 YEPELLGRS YEPEELS YEPEELGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEELLGRS YEPEEL	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNLFAKG LTKTHHNLFAKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Rat Blind-Mole-rat Sumatran-orangutan Crab-eating-macaque Tibetan-antelope Goat Cow Yak Naked-carp Zebra-fish	230 240 PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK TGQYRMLAKRGGYVV	250 TFLSRHSLDMKFS STFLSRHSLMKFS STFLSRHSLMKFS TFLSRHSLMKFS TFLSRHSLMKFS TFLSRHSLMKFS TFLSRHSLMKFS TFLSRHSLMKFS STFLSRHSLMKFS STFLSRHSLMKFS TFLSRHSLMKFS TFLSRHSLMKFS STFLSRHS	A 6 0 Y CDERITELMG Y CDERIT	270 YEPELLGRS YEPEELGRS YEPEELLGRS	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHNFFKG LTKTHHNFFKG LTKTHHNFFKG LTKTHHNFFKG
Human rhesus-monkey Cow Dog Chimpanzee Sumatran-orangutan Naked-mole-rat Mouse Rat Blind-Mole-rat Tibetan-antelope Yak Crab-eating-macaque Goat Naked-carp Zebra-fish consensus>50 Human rhesus-monkey Dog Chimpanzee Naked-mole-rat Mouse Rat Blind-Mole-rat Sumatran-orangutan Crab-eating-macaque Tibetan-antelope Goat Cow Yak Naked-carp Zebra-fish consensus>50	230 240 PHPSNIEIPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVPLDSK PHPSNIEVV TGQYRMLAKRGGYVV	250 TFLSRHSLDMKFS TFLSRHSLMKFS NUETQATVIYNTK NU	A 6 0 Y CDERITELMG Y CDERITE	2 7 0 Y E P E E L L G R S Y	280 IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH IYEYYHALDSDH	290 LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDMFTKG LTKTHHDFFKG

Figure 1. Multiple sequence alignment. (A) bHLH domain. The DNA-interacting basic amino acids K21, R29, K19, R30, R27, D55, and K56 are highlighted in red.¹⁸ (B) PAS-A, (C) PAS-B, and (D) PAS-C. PAS-B is involved in interaction with aryl hydrocarbon receptor nuclear transporter (ARNT) and HSP90. Residues interacting with PAS-B and ARNT were highlighted in red. PAS-C also includes residues interacting with ARNT as mentioned in crystal structure PDB ID (4H6J), and it is highlighted in red.

mutations, namely, T407I and G414S, were observed in NMR. S385C was the other interesting polymorphism reflected in mouse, rat, and BMR, respectively.

The ODD exhibits notable changes compared with other domains across all the 16 species. The posttranslational

modification residues such as P402, K532, and P564 are all conserved across the species (Figure 2). However, a critical change L559P has been observed in the binding site of HIF-1 α , which is crucial for PHD2 interaction, which is seen in naked carp. Remaining binding site residues, namely, M568,

	он	410	420		430	440	450	460	
Human Chimpanzee Sumatran-orangutan rhesus-monkey Crab-eating-macaque Yak Tibetan-antelope Goat Cow Dog Naked-mole-rat Blind-Mole-rat Rat Mouse Naked-carp Zebra-fish consensus>50	OH APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD APAAGD	410 TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG TIISLDFG	420 SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET SNDTET		430 VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM VPLYNDVM	440 IDS DNEKLQ DS DNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKLQ IDS SNEKL IDS SNE	450 NINLAMSPLP NINLAMSPLP NINLAMSPLP NINLAMSPLP NINLAMSPLP NINLAMSPLP NINLAMSPLP NINLAMSPLP NINLAMSPLP SINLAMSPLP NINLAMSPLP SIILSPL. PLSLSPL.	460 TAETPKP TAETPKP TAETPKP TSETPKP TSETPKP ASETPKP ASETPKP ASETPKP ASETPKP ASETPKP ASETPKP ASETPKP SETPKP TP 	LRSSADP LRSSAP LRSSADP
				~ ~ 1		•	•		
	470	4 8 0	4	٥٥	500	510	520	53	<u></u>
Human Chimpanzee Sumatran-orangutan rhesus-monkey Crab-eating-macaque Yak Tibetan-antelope Goat Cow Dog Naked-mole-rat Blind-Mole-rat Rat Mouse Naked-carp Zebra-fish consensus>50	ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV ALNQEV	ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLEPNF ALKLESSF KLKTGA KLETGG alKLEPNF	E = E LELS $E = E S LEE LS$ $S LEE LS$ $S LEE LS$ $S S LEE LS$ $E S S LEE S S LEE LS$ $E S L S S LEE S S LEE S$ $E S L S S LEE S S LEE S S LEE S LS$ $E S L S S LEE S S$	FTMPQI(FTMPQI(FTMPQI(FTMPQI(FTMPQI(FTMPQI(FTMPQI(FTMPQI(FTMPQI(FTMPQI(SASDLG) SASDLG SASDLG SASDLG	2DQTPSPSI 2DQTPSPSI 2DQPPSPSI 2DQPPSPSI 2DQPPSPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI 2DQPASPSI	OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE OGSTRQSSPE TSGLGSSGP. TSGLGSSGP.	PNSPSEYCFY PNSPSEYCFY PNSPSEYCFY PNSPSEYCFY PNSPSEYCFD PNSPSEYCFD PNSPSEYCFD PNSPSEYCFD PNSPSEYCFD PNSPSEYCFD PNSPSEYCFD PNSPSEYCFD PNSPSEYCFD - NSPMDYGFP PNSPSEYCFD - NSPMDYGFP	VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN VDSDMVN	E FK LELV E FK LELV E FK LELV E FK LELV E FK LELV E FK LELV E FK LELV V FK LELV V FK LELV V FK LELV V FK LELV V FK LELV E FK LDLV E FK LDLV E FK LL LV
	540	5 5	0	560	OH 570	58	0 5 S	0	600
Human Chimpanzee Sumatran-orangutan rhesus-monkey Crab-eating-macaque Yak Tibetan-antelope Goat Cow Dog Naked-mole-rat Blind-Mole-rat Rat Mouse Naked-carp Zebra-fish consensus>50	EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE EKLFAE	DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKNP DTEAKTP DTEAKTP DTEAKTP	H H	DLDLEMI DLDLEMI DLDLEMI DLDLEMI DLDLEMI DLDLEMI DLDLEMI DLDLEMI DLDLEMI DLDLEMI DLDLEMI LDLEMI .LDPEMI .LDLEMI	APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI APYIPMDI	DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ DDFQLRSFDQ	LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLENSST LSPLENSST LSPLENSST LSPLENSST LSPLENSST LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESSSA LSPLESA LSPLESSA LSPLESA	PESASPC PESASPC PESASPC PESASPC PESASPC PESASPC PQSAS. PQSAS. PQSAS. PQSAS. PQSAS. PQSAS. PQSAS. PQSVS. PPSVS. SPSWS. SLSAMS. VSAMS.	STVTVFC STVTVFC STVTVFC STVTVFC STVTVFC STVTVFC TNTVFC TNTVFC TNTVFC TNTVFC TVTVFC TVTVFC TVTVFC TVTVFC TVTVFC TVTVFC TVTGFC TVTGFC
				C-Terminal	VHL recognition site	c			

Figure 2. Multiple sequence alignment of oxygen-dependent degradation domain. Hydroxylation sites P402 and P564 are highlighted in yellow. The acetylation site (K532) is shown in orange red. The N-terminal transactivation domain is shown in green, and the C-terminal VHL recognition site is shown in red box.

D571, F572, P567, M561, E560, A563, Y565, P564, I566, L562, L574, Q573, and D570,²⁶ appear well conserved across rat, mouse, humans, and BMR. The N-TAD (which is an integral part of ODD) shows E534D, L539M, E542 N548T, T552S, L559P, and R575H polymorphisms in naked carp. Interesting changes such as A541T in goat, T555N in NMR, T552A in rat, and E534D, E542I, and N548I in zebrafish were observed.

The ID shows huge variations, and the same are depicted in Figure 3A. The C-TAD shows the following changes, namely, M787L, D788E, S797R, G808D, S809R, R810H, and N811H in zebrafish and naked carp, respectively. Furthermore, E789G is also found in both BMR and zebrafish.

The regulation and half-life of the protein are dependent on the hydroxylation of proline residues at 402 and 564 by the PHD2 under normoxic condition. The domain that includes these residues is called ODD. The hydroxylated residues were recognized by the pVHL and facilitated the ligation with proteasomal complex.¹ Thus, recognition by pVHL is also a major factor towards the HIF-1 α turnover. There are lots of changes in N-terminal VHL recognition site of zebrafish and naked carp. The rodents except NMR show S385C substitution. C-terminal VHL recognition site possesses 2 binding sites. Site 1 from 560 to 567 (EMLAHypYIP)⁴ is well conserved. But in site 2, containing residues (DFQLRSF)⁴ from 571 to 577, the following changes are observed. R575H, S576P, and F577S

Human Sumatran-orangutan	580 SFDQLSPL SFDOLSPL	590 ESSSASPESASPESASPESASPESASPESASPESASPES	600 OSTVTVFO	610 QTQIQEPTANA OTOIOEPTANA	620 TTTTATTDEL TTTTATTDEL	630 KTVTKDRMEI	640 DIKILIASPSI DIKILIASPSI	650 PTHIHKETTSA PTHIHKETTSA	660 TSSPYRD
Chimpanzee rhesus-monkey Crab-eating-macaque	SFDQLSPL SFDQLSPL	ESSSASPESAS ESSSASPESAS ESSSASPESAS	OSTVTVFO	QTQIQEPTANA QTQIQEPTANA OTQIQEPTANA	TTTTTATTDEL TTTTTATTDEL	KTVTKDCMEI KTVTKDRMEI	DIKILIASPS	PTHIHKETTSA STHIHKETTSA STHIHKETTSA	TSSPYRD TSSPYRD
Yak Tibetan-antelope	SFDQLSPL	ENSSTSPOSAS. ENSSTSPOSAS.	TNTVFQ	PTQMQEPPIAT PTQMQEPPIAT	VTTTATSDEL	KTVTKDGMKI KTVTKDGMEI	DIKILIAFPSI	P PHVPKEPPCA P PHVPKEPPCA	TTSPYSD
Cow Dog	SFDQLSPL	ENSSTSPOSAS. ENSSTSPOSAS. ESNSTSPOSAS.	TNTVFQ TITVFQ	PTOMOEPPIAT PTOMOEPPIAT PTPMOEPPLTT	TSTTATTOEL	KTVTKDGIEI KTVTKDGIEI	DIKILIAFPSI DIKILIAFPSI DIKILIAAPSI	PPHVPKEPPCA PPHVPKEPPCA PTHVPKVTTSA	TTSPYSD TTSPYSD TTSPYSD
Blind-Mole-rat Naked-mole-rat Rat	LFDQLSPL SFDQLSPL SFDQLSPL	ESSSPSPPGVN. ESSSASPQSVS. ESNSPSPPSVS.	TATAFQ TTTVFQ TVTGFQ	QTQLQEPTIST QTQMQEPTVTT QTQLQKPTITV	TTTTTTTTNBL TTTTPPTTEBL ATATATTDBS	K SVTKDSMEI KSVTKDSMEI KAVTKDNIEI	DIKILIASPS DIQILIVSPS DIKILIASP	STHTPKETTTA PTHLLKETAHA STQVPQEMTTA	TTSPYSG TTSPYSN KASAYSG
Mouse Naked-carp Zebra-fish	SFDQLSPL PSPL PSPL	ESNSPSPPSMS. DPLPSGPHSLS. DPLPSATHSVS.	TVTGFQ .AMSSLFQ .AMSSLFQ	QTQLQKPTITA PLPS.PASPAS PLPSSPASPAS	ATTTTATTDES SSSSSAVMQEL SSTSSTVKQEA	KTETKD NK EI PS	DIKILIASPS: RAPSI RAPSI	STQVPQETTTA PLHLLQEVCNA PLHLLQEVCSA	KASAYSG PVSPFSG PVSPFSG
consensus>50	sfdqlSPL	#sss.spqsas.	tvtvFQ	qtqmqepti	ttttatt#El	ktvtkd.med	dikiliasPsp	pthvpkett.A	t.Sp%sd
Human	TOSRTASP	0 680 NRAGKGVIEQTE	69 KSHPRSPN	0 700 VLSVALSQRTT	710 VPEEELNPKI	720	730	740 AVGIGTLLQQ	PDDHAAT
Chimpanzee rhesus-monkey	TQSRTASPI	NRAGKGVIEQTE NRAGKGVIEQTE NRAGKGVIEQTE	KSHPRSPN KSHPRSPN KSHPRSPN	VLSVALSQRT1 VLSVALSQRT1 VLSVTLSQRT1	VPEEBLNPKI VPEEBLNPKI VPEEBLNPKI	LALQNAQRKE	RKMEHDGSLF(RKMEHDGSLF(AVGIGILLQQ AVGIGTLLQQ	PDDHAAT PDDHAAT
Crab-eating-macaque Yak Tibetan-antelope	TGSRTASP TGSRTASP TGSRTASP	NRAGKGVIEQTE SRAGKGVIEQTE N <mark>R</mark> AGKGVIEQTE	KSHPRSPN KSHPRSPN KSHPRSPN	VLSVALSQRTI VLSVALSQRTI VLSVALSQRTI	TVPEEBLNPKI TAPEEBLNPKI TAPEEBLNPKI	LALQNAQRKE LALQNAQRKE LALQNAQRKE	RKMEHDGSLF(RKIEHDGSLF(RKIEHDG <mark>SL</mark> F(QAVGIGTLLQQ QAVGIGTLLQQ QAVGIGTLLQQ	PDDHAAT PDDRATT PDDRATT
Goat Cow Dog	TGSRTASPI TGSRTASPI TGSRTASPI	NRAGKGVIEQTE NRAGKGVIEQTE NRAGKGVIEQTE	KSHPRSPN KSHPRSPN KSHPRSPN	VLSVALSQRTI VLSVALSQRTI VLSVTLSQRTI	CAPEEBLNPKI CAPEEBLNPKI CIPEEBLNPKI	LALQNAQRKE LALQNAQRKE LALQNAQRKE	RKIEHDGSLF(RKIEHDGSLF(RKIEHDGSLF(QAVGIGTLLQQ QAVGIGTLLQQ QAVGIGTLLQQ	PDDRATT PDDRATT PDDRATT
Blind-Mole-rat Naked-mole-rat Rat	TQSRTASP TPSRTTSP THSRTASP	NRAGQGVIEQTE NRAGKGAIEQTE DRAGKRVIEKTI	KSHPRSPN KSHPRSPN KAHPRSLN	LLSVSLSQRNT VVPVTLSQRNT .LSVTLNQRNT	IVPEEELNPXI IIPEEELNPXI IVPEEELNPXT	IALQNAQRKE QALQNAQRKE IALQNAQRKE	RKMEHDGSLF(RKMEHDGSLF(RKMEHDGSLF(AAGIGTLLQQ AVGIGTLLQQ AAGIGTLLQQ	PDDRAPA SDDRAAT PGDRAPT
Mouse Naked-carp Zebra-fish	THSRTASPI T. RDVSP	DRAGKRVIEQTI ARS	. PIPQSSN . STPOSSS	LSATLNQRNT	VPEEELNPKT NRELSPKM NREMSPKM	IASQNAQRKE LAIQNAQRKE LAFONIORKE	RKMEHDGSLF(RKLEEVTSLSI RKLNEVTSLSI	AAGIGTLLQQ BAVGLSALLQS BAVGLGALLHS	PGDCAPT VDSAIEP VDSAIDP
consensus>50	t.sRtaSP	nRagkgvieqte	kshPrSpn	vlsv.lsqrtt	vp#eE\$nPKi	1A1QNaQRKF	RKm#hdgSLf	AvGigtLLqq	pddra.t
Human	750 TSLSWKR	760 KGCKSSEQNGME	770 QKTIILIP	780 SDLACRLLGO					
Sumatran-orangutan Chimpanzee rhesus-monkey	TSLSWKRV TSLSWKRV TSLSWKRV	KGCKSSEQNGME KGCKSSEQNGME KGCKS <mark>S</mark> EQN <mark>G</mark> ME	QKTIILIP QKTIILIP QKTIILIP	SDLACRLLGQ SDLACRLLGQ SDLACRLLGQ					
Crab-eating-macaque Yak Tibetan-antelope	TSLSWKRVI TSLSWKRVI TSLSWKRVI	KGCKSSEQNGME KGCKSSEQNGME KGCKSSEQN <mark>G</mark> ME	QKTIILIP QKTIILIP QKTIILIP	SDLACRLLGQ SDLVCRLLGQ SDLACRLLGQ					
Goat Cow Dog	TSLSWKRV TSLSWKRV TSLSWKRV	KGCKSSEQNGME KGCKSSEQNGME KGCKSSEQNGME	OKTIILIP OKTIILIP OKTIILIP	SDLACRLLGQ SDLACRLLGQ SDLACRLLGQ					
Blind-Mole-rat Naked-mole-rat Rat	TSLSWKRV TSLSWKRV MSLSWKRV	KGCKSSGQNGME KGCKSSEQNGTE KGYISSEODGME	OKTIILIP OKTIILIP OKTIFLIP	SDLACRLLGO SDLACRLLGO SDLACRLLGO					
Mouse Naked-carp Zebra-fish	MSLSWKRV GKRAKV	KGFISSEONGTE FELKGSSLLG.G	OKTILLP NKTILLP	SDLACRLLGO SDVASRLLCS					
consensus>50	tslswkrV	kgcksSeqnGme	#KTIiliP	SDlacRLLgq					
		A							
Human rhesus-monkey	SMDESGL	POLTSYDCEVN POLTSYDCEVN	APIQGSRN APIQGSRN	LLQGEELLRA LLQGEELLRA	LDQVN				
Cow Dog Chimpanzee	SMDESGL SMDESGL SMDESGL	POLTSYDCEVN POLTSYDCEVN POLTSYDCEVN	APIQGSRN APIQGSRN APIQGSRN	LLQGEELLRA	LDQVN				
Goat Naked-mole-rat	SMDESGL	PQLTSYDCEVN PQLTSYDCEVN	APIOGSRN APIOGSRN	LLQGEBLLRA	LDQVN				
Mouse Rat Sumatran-orangutan	SMDESGL SMDESGL SMDESGL	PQLTSYDCEVN PQLTSYDCEVN PQLTSYDCEVN	APIQGSRN APIQGSRN APIQGSRN	LLQGEBLLRA LLQGEBLLRA LLQGEBLLRA	LDQVN				
Tibetan-antelope Yak Crab-eating-macaque	SMDESGL SMDESGL SMDESGL	POLTSYDCEVN POLTSYDCEVN POLTSYDCEVN	APIOGSRN APIOGSRN APIOGSRN	ILLQGEELLRA ILLQGEELLRA ILLQGEELLRA	LDQVN				
Blind-Mole-rat Naked-carp Zebra-fish	SMDGSGL SLENSGL	POLTSYDCEVN POLTRYDCEVN POLTRYDCEVN	APIQGSRN APVODRHH	LLQGEELLRA	LDQVN				
consensus>50	S\$#eSGL	PQLTSYDCEVN Site 1(795-8	AP!Qgsri 06)	Site 2 (812-8	LDQVN 23)				
		В	1991) 1991		5				

Figure 3. Multiple sequence analysis of (A) inhibitory domain and (B) C-terminal transactivation domain (C-TAD). There are 2 binding sites in C-TAD. Site 1 in human hypoxia-inducible factor- 1α carboxy-terminal activation domain encompasses residues 795 to 806 and contains the hydroxylated asparagine (N803), and site 2 includes residues 812 to 823 and shows only weak binding independent of site 1. These sites are highlighted in green.

substitutions are observed in naked carp. S576I and F577P are found in zebrafish. In BMR, S576L is the change seen. These variations in VHL binding region may not render the VHLmediated degradation of HIF-1 α and increase the stability of protein even under normoxic condition. This may be the probable reason for BMR having a higher level of HIF-1 α under hypoxic conditions than normal rats.¹² The main regulatory domain is exhibiting good conservation of critical hydroxylation sites (P402 and P564). The binding site (L559, M568, D571, F572, P567, M561, E560, A563, Y565, P564, I566, L562, L574, Q573, and D570)²⁶ of ODD to PHD2 is also well conserved. Acetylation of K532 facilitates the interaction of HIF-1 α with VHL, and this site is also well conserved. Interestingly, only the naked carp showed changes in L559P. Naked carp is a hypoxic-adapted fish; this change may increase the stability of HIF-1 α by deviating interaction of ODD with PHD2. The C-TAD is a major domain involved in transactivity. The hydroxylation at N803 by FIH prevents its interaction with coactivator. So, any



Figure 4. Evolutionary relationships of taxa for hypoxia-inducible factor-1 α of selected organisms.



Figure 5. Phylogeny of the oxygen-dependent degradation domain. The tree shows similar trend in group formation like the tree for entire hypoxiainducible factor- 1α . Shuffle is only observed in the group of even-toed ungulates.



Figure 6. Correlation of changes in human hypoxia-inducible factor-1α oxygen-dependent degradation domain with respect to blind-mole rat. The critical residues such as P402, P564, and K532 are well conserved.

change in the interaction site of C-TAD will affect the regulation of transactivity. There are 2 binding sites in C-TAD. Site 1 in hHIF-1 α carboxy-terminal activation domain encompasses residues 795 to 806 and contains the hydroxylated asparagine. Naked carp and zebrafish show S797R I806V mutations in this site. Site 2 includes residues 812 to 823 and shows only weak binding independent of site 1. E817D substitution is found in site 2 of naked carp. The ID is a stretch domain (576-785) that represses the transactivity of HIF-1 α . There are lots of changes observed in this domain as illustrated in Figure 3A, which depicts close to 50% variations among naked carp, BMR, zebrafish, and hHIF. The removal of this particular domain increases the transactivity of HIF-1 α .²⁷

The phylogenetic analysis clearly highlighted the variations that were observed at the sequence level of HIF- α . Figure 4 indicates that the 16 sequences are clustered into 5 groups. Naked carp and zebrafish standout as an independent cluster, exhibiting significant differences across the set under study. The aquatic creatures, rodents, carnivore (dog), even-toed ungulate, and higher primates are grouped separately. The formation of each clade in the phylogenetic tree (Figure 4) may be due to adaptations by different organisms to their habitats. Even though humans and chimpanzee share the same ancestral branch, the ODD is crucial for determining the half-life of HIF-1a; different organisms may have diverged according to the natural selection (Figure 5). Alignment of sequences showed critical replacements in the ODD. This is evidenced by the phylogenetic distance between the normoxic animals such as human beings and hypoxic animals such as Judean Mountains blindmole rat. The taxa result infers that there is a clear divergence in HIF-1a protein as it gets evolved. All organisms were grouped out from rodents, zebrafish, and naked carp. From the phylogenetic analysis, we conclude that the variations in the different domains of HIF- α provide vital clues for plausible stabilization of this key molecule under hypoxic and normoxic conditions.

Discussion

We have found significant changes in residues belonging to specific domains of HIF-1 α protein and vital clues to appreciate the evolutionary variations. The oxygen-dependent hydroxylation of ODD and its recognition by VHL are crucial steps in HIF-1 α half-life and activity. Changes found in this domain may enhance or decrease the activity of HIF-1 α . These variations in VHL binding region may not render the VHLmediated degradation of HIF-1 α and increase the stability of protein even under normoxic condition. This may be the possible reason for BMR having a higher level of HIF-1 α under hypoxic conditions than normal rats¹²; several changes were observed in ODD with respect to BMR (Figure 6). Furthermore, the effect of observed changes in N-terminal VHL recognition site and C-terminal VHL recognition site is not known. Particularly in site 2 from 571 to 577 (DFQLRSF), changes observed may be restraining the ODD binding to VHL in naked carp and zebrafish. But these changes are not prominent in either cancer-resistant animals such as BMR and NMR or cancer-prone animals such as humans, mouse, and rat.

Conclusions

This study is an attempt to appreciate the evolutionary changes of protein HIF-1 α and the detailed sequence variations across the set of species considered. Key substitutions in the oxygendependent hydroxylation of the ODD and the consequence of the same towards its recognition to VHL enable us to explain its criticality with respect to HIF-1 α half-life and its activity. Comparisons of the HIF-1 α sequence among cancer-resistant animals, such as BMR and NMR, and cancer-prone animals, such as human, mouse, and rat, suggest derivation of probable clues towards potential risk factors for cancer.

Acknowledgements

The authors thank Sir M Visvesvaraya Institute of Technology for kind support during the work and also thanks to Vijay Radhakrishan.

Author Contributions

JP, CGJ, DJK, and HGN conceived and designed the experiments. DJK and HGN analysed the data. CGJ and JP wrote the first draft of the manuscript. JP, CGJ, DJK, and HGN contributed to the writing of the manuscript. CGJ, DJK, and HGN agree with manuscript results and conclusions. JP, CGJ, DJK, and HGN jointly developed the structure and arguments for the paper. CGJ, DJK, and HGN made critical revisions and approved the final version. All authors reviewed and approved of the final manuscript

REFERENCES

- Brahimi-Horn MC, Chiche J, Pouysségur J. Hypoxia and cancer. J Mol Med. 2007;85:1301–1307.
- Wang GL, Jiang BH, Rue EA, Semenza GL. Hypoxia-inducible factor 1 is a basichelix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc Natl Acad Sci U S A*. 1995;92:5510–5514.
- Mandl M, Depping R. Hypoxia-inducible aryl hydrocarbon receptor nuclear translocator (ARNT) (HIF-1β): is it a rare exception? *Mol Med.* 2014; 20:215–220.
- Lisy K, Peet DJ. Turn me on: regulating HIF transcriptional activity. *Cell Death Differ*. 2008;15:642–649.
- Rob A, Cairns, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nat Rev Cancer. 2011;11:85–95.
- Staab A, Loeffler J, Said HM, et al. Effects of HIF-1 inhibition by chetomin on hypoxia-related transcription and radiosensitivity in HT 1080 human fibrosarcoma cells. *BMC Cancer*. 2007;7:1471–2407.
- Semenza GL. HIF-1: mediator of physiological and pathophysiological responses to hypoxia. J Appl Physiol. 2000;88:1474.
- Baumann R, Depping R, Delaperriere M, Dunst J. Targeting hypoxia to overcome radiation resistance in head & neck cancers: real challenge or clinical fairytale? *Expert Rev Anticancer Ther.* 2016;16:751–758.
- Bertout JA, Majmundar AJ, Gordan JD, et al. HIF2α inhibition promotes p53 pathway activity, tumor cell death, and radiation responses. *Proc Natl Acad Sci U* SA. 2009;106:14391–14396.
- Bhatt RS, Landis DM, Zimmer M, et al. Hypoxia-inducible factor-2α: effect on radiation sensitivity and differential regulation by an mTOR inhibitor. *BJU Int*. 2008;102:358–363.
- Qiu Q., Zhang G., Ma T, et al. The yak genome and adaptation to life at high altitude. *Nat Genet.* 2012;44:946–949.

- Shams I, Avivi A, Nevo E. Hypoxic stress tolerance of the blind subterranean mole rat: expression of erythropoietin and hypoxia-inducible factor 1α. Proc Natl Acad Sci U S A. 2004;101:9698–9703.
- Apweiler R, Bairoch A, Wu CH, et al. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 2017;45:D158–D169.
- Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*. 2000;16:276–277.
- Li W, Cowley A, Uludag M, et al. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 2015;43:W580–W584.
- McWilliam H, Li W, Uludag M, et al. Analysis tool web services from the EMBL-EBI. Nucleic Acids Res. 2013;41:W597–W600.
- Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 1988;16:10881–10890.
- Wu D, Potluri N, Lu J, Kim Y, Rastinejad F. Structural integration in hypoxiainducible factors. *Nature*. 2015;524:303–308.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30:2725–2729.
- Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985;39:783–791.
- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4:406–425.
- Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, eds. *Evolving Genes and Proteins*. Amsterdam, The Netherlands: Elsevier; 1995:97–166.
- Bersten DC, Sullivan AE, Peet DJ, Whitelaw ML. bHLH-PAS proteins in cancer. Nat Rev Cancer. 2013;13:827–841.
- Cardoso R, Love R, Nilsson CL, et al. Identification of Cys255 in HIF-1α as a novel site for development of covalent inhibitors of HIF-1α/ARNT PasB domain protein-protein interaction. *Protein Sci.* 2012;211885–1896.
- Katschinski DM, Le L, Schindler SG, Thomas T, Voss AK, Wenger RH. Interaction of the PAS B domain with HSP90 accelerates hypoxia-inducible factor-1alpha stabilization. *Cell Physiol Biochem*. 2004;14:351–360.
- Chowdhury R, McDonough MA, Mecinović J, et al. Structural basis for binding of hypoxia-inducible factor to the oxygen-sensing prolyl hydroxylases. *Structure*. 2009;17:981–989.
- Jiang B-H, Zheng JZ, Leung SW, Roe R, Semenza GL. Transactivation and inhibitory domains of hypoxia-inducible factor 1alpha: modulation of transcriptional activity by oxygen tension. J Biol Chem. 1997;272:19253–19260.