Electronic Supplementary Information (ESM)

Evolutionary genetics of hypoxia tolerance in cetaceans during diving

Ran Tian, Zhengfei Wang, Xu Niu, Kaiya Zhou, Shixia Xu[*], Guang Yang[*]

Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210023, China

Table S1 The 17 candidate genes and their protein products

Gene	**Ensembl ID**	**Protein Product**	**UniProtKB/Swiss-Prot Summary of Gene Function**	**GeneCard Summary of Disorders / Diseases**	**Sources of 3D Structure**	**Amino Acid Length**
HBA	ENSG00000188536	Hemoglobin alpha	Involved in oxygen transport from the lung to the various peripheral tissues.	Alpha-thalassemia, beta-thalassemia, thalassemia, hemoglobinopathies, h disease	PDB: 1SI4	141
HBB	ENSG00000244734	Hemoglobin beta	Involved in oxygen transport from the lung to the various peripheral tissues.	Beta-thalassemia, thalassemia, hemoglobinopathies, intermedia thalassemia, erythroleukemia	PDB: 1HHO	146
MB	ENSG00000198125	Myoglobin	Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.	Rhabdomyolysis, acute myocardial infarction, rhabdomyolysis exertional acute, renal failure acute, rhabdomyosarcoma	PDB: 1U7S	153
NGB	ENSG00000165553	Neuroglobin	Involved in oxygen transport in the brain. Involved in neuroprotection during oxidative stress.	Ischemia, stroke	-	150
EDN1	ENSG00000078401	Endothelin-1	Endothelins are endothelium-derived vasoconstrictor peptides.	Hypertension pulmonary, vasospasm, hypertension arterial, atherosclerosis, essential hypertension	I-TASSER	212
EDN2	ENSG00000127129	Endothelin-2	Endothelins are endothelium-derived vasoconstrictor peptides	Essential hypertension, carcinoma renal cell, cardiovascular diseases, hypertension pulmonary, vascular diseases	I-TASSER	178
EDN3	ENSG00000124205	Endothelin-3	Endothelins are endothelium-derived vasoconstrictor peptides.	Shah-waardenburg syndrome, hirschsprung disease, waardenburgs	I-TASSER	238

				syndrome, colonic aganglionosis, central hypoventilation syndrome, congenital		
EDNRA	ENSG00000151617	Endothelin receptor type A	Receptor for endothelin-1. Mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system.	Staphylococcal scalded skin syndrome, hypertension arterial, staphylococcal scarlet fever, hypertension pulmonary, bullous impetigo	I-TASSER	426
EDNRB	ENSG00000136160	Endothelin receptor type B	Non-specific receptor for endothelin 1, 2, and 3. Mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system.	Hirschsprung disease, shah-waardenburg syndrome, waardenburgs syndrome, staphylococcal scalded skin syndrome, hirschsprung disease 2	I-TASSER	442
AGTR1	ENSG00000144891	Angiotensin II receptor type 1	Receptor for angiotensin II. Mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system	Essential hypertension, diabetic nephropathies, renal disease, cardiovascular diseases, heart failure	-	358
AGTR2	ENSG00000179772	Angiotensin II receptor type 2	Receptor for angiotensin II. Cooperates with MTUS1 to inhibit ERK2 activation and cell proliferation.	Diabetic nephropathies, cardiac hypertrophy, ventricular hypertrophy, renal disease, hypertrophy	-	120
ADRA1A	ENSG00000120907	Adrenergic receptor alpha-1 A	This alpha-adrenergic receptor mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system. Nuclear	Benign prostatic hyperplasia, attention-deficit hyperactivity, disorder	-	466

			ADRA1A-ADRA1B heterooligomers regulate phenylephrine (PE)-stimulated ERK signaling in cardiac myocytes.			
ADRA1B	ENSG00000170214	Adrenergic receptor alpha-1 B	This alpha-adrenergic receptor mediates its action by association with G proteins that activate a phosphatidylinositol-calcium second messenger system. Nuclear ADRA1A-ADRA1B heterooligomers regulate phenylephrine (PE)-stimulated ERK signaling in cardiac myocytes.	Hyperfunction, cardiomyopathy, heart failure	-	520
ADRA1D	ENSG00000171873	Adrenergic receptor alpha-1 D	This alpha-adrenergic receptor mediates its effect through the influx of extracellular calcium.	Benign prostatic hyperplasia, essential hypertension, obstruction	I-TASSER	572
AVP	ENSG00000101200	Arginine vasopressin	Vasopressin has a direct antidiuretic action on the kidney, it also causes vasoconstriction of the peripheral vessels.	Central diabetes insipidus, hyponatremia, siadh, diabetes insipidus nephrogenic, carcinoma oat cell	I-TASSER	
AVPR1A	ENSG00000166148	Arginine vasopressin receptor 1 A	Receptor for arginine vasopressin. The activity of this receptor is mediated by G proteins which activate a phosphatidyl-inositol-calcium second messenger system.	Adrenal hyperplasia, dysmenorrheal, tumors, ischemia, cancer lung	-	418
AVPR1B	ENSG00000198049	Arginine vasopressin receptor 1 B	Receptor for arginine vasopressin. The activity of this receptor is mediated by G proteins which activate a	Cushing syndrome, carcinoid, major depression, adenoma, tumors	-	424

	phosphatidyl-inositol-calcium second messenger system	

Note: The Protein Data Bank (PDB) is a database containing experimentally determined three-dimensional structures of proteins, nucleic acids and other biological macromolecules, with approximately 8000 entries (http://www.rcsb.org/pdb/). I-TASSER is a state-of the-art hierarchical protein structure modeling approach that is based on the secondary-structure enhanced profile-profile threading alignment (Zhang 2008). The “-” indicated 3D structure did not be used in this study.

Tables S2 Statistics for amplified exons of 11 genes for each species.

Note: tick represents successfully amplified, while no sign in any genes indicate no amplification despite numerous attempts.

Species	HBA		
	Exon1	Exon2	Exon3
Tursiops truncatus	✔	✔	✔
Delphinus capensis	✔	✔	✔
Sousa chinensis	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔
Delphinapterus leucas			
Lipotes vexillifer	✔	✔	✔
Mesoplodon densirostris	✔	✔	✔
Kogia sima	✔	✔	✔
Physeter macrocephalus	✔	✔	✔
Balaenoptera acutorostrata	✔	✔	✔
Banaenoptera omurai			

Species	HBB		
	Exon1	Exon2	Exon3
Tursiops truncatus	✔	✔	✔
Delphinus capensis	✔	✔	✔
Sousa chinensis	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔
Delphinapterus leucas	✔	✔	✔
Lipotes vexillifer	✔	✔	✔
Mesoplodon densirostris	✔	✔	✔
Kogia sima	✔	✔	✔
Physeter macrocephalus	✔	✔	✔
Balaenoptera acutorostrata	✔	✔	✔
Banaenoptera omurai	✔	✔	✔

Species	MB		
	Exon1	Exon2	Exon3
Tursiops truncatus	✔	✔	✔
Delphinus capensis	✔	✔	✔
Sousa chinensis	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔
Delphinapterus leucas	✔	✔	✔
Lipotes vexillifer	✔	✔	✔
Mesoplodon densirostris	✔	✔	
Kogia sima	✔	✔	✔
Physeter macrocephalus	✔	✔	
Balaenoptera acutorostrata	✔	✔	✔
Banaenoptera omurai	✔	✔	✔

Species	NGB			
	Exon1	Exon2	Exon3	Exon4
Tursiops truncatus		✔	✔	✔
Delphinus capensis		✔	✔	✔
Sousa chinensis		✔	✔	✔
Neophocaena phocaenoides		✔	✔	✔
Delphinapterus leucas		✔	✔	✔
Lipotes vexillifer		✔	✔	✔
Mesoplodon densirostris		✔	✔	✔
Kogia sima		✔	✔	✔
Physeter macrocephalus		✔	✔	✔
Balaenoptera acutorostrata		✔	✔	✔
Banaenoptera omurai		✔	✔	✔

Species	EDN1				
	Exon1	Exon2	Exon3	Exon4	Exon5
Tursiops truncatus	✔	✔	✔	✔	✔
Delphinus capensis	✔	✔	✔	✔	✔
Sousa chinensis	✔	✔	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔	✔	✔
Delphinapterus leucas	✔	✔	✔	✔	✔
Lipotes vexillifer	✔	✔	✔	✔	✔
Mesoplodon densirostris	✔	✔	✔	✔	✔
Kogia sima	✔	✔	✔	✔	✔
Physeter macrocephalus	✔	✔	✔	✔	✔
Balaenoptera acutorostrata	✔	✔			✔
Banaenoptera omurai	✔	✔	✔	✔	✔

Species	EDN2				
	Exon1	Exon2	Exon3	Exon4	Exon5
Tursiops truncatus	✔	✔	✔	✔	✔
Delphinus capensis	✔	✔	✔	✔	✔
Sousa chinensis	✔	✔	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔	✔	✔
Delphinapterus leucas	✔	✔	✔	✔	✔
Lipotes vexillifer	✔	✔	✔	✔	✔
Mesoplodon densirostris	✔	✔	✔	✔	✔
Kogia sima	✔	✔	✔	✔	✔
Physeter macrocephalus	✔	✔	✔	✔	✔
Balaenoptera acutorostrata	✔	✔	✔	✔	✔
Banaenoptera omurai	✔	✔	✔	✔	✔

Species	EDN3				
	Exon1	Exon2	Exon3	Exon4	Exon5
Tursiops truncatus	✔	✔	✔	✔	✔
Delphinus capensis	✔	✔	✔	✔	✔
Sousa chinensis	✔	✔	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔	✔	✔
Delphinapterus leucas	✔	✔	✔	✔	✔
Lipotes vexillifer	✔	✔	✔	✔	✔
Mesoplodon densirostris	✔	✔	✔	✔	✔
Kogia sima	✔	✔	✔	✔	✔
Physeter macrocephalus	✔	✔	✔	✔	✔
Balaenoptera acutorostrata	✔	✔	✔	✔	✔
Banaenoptera omurai	✔	✔	✔	✔	✔

Species	EDNRA							
	Exon1	Exon2	Exon3	Exon4	Exon5	Exon6	Exon7	Exon8
Tursiops truncatus	✔	✔	✔	✔	✔	✔	✔	✔
Delphinus capensis	✔	✔	✔	✔	✔	✔	✔	✔
Sousa chinensis	✔	✔	✔	✔	✔	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔	✔	✔	✔	✔	✔
Delphinapterus leucas	✔	✔	✔	✔	✔	✔	✔	✔
Lipotes vexillifer	✔	✔	✔	✔	✔	✔	✔	✔
Mesoplodon densirostris	✔	✔	✔	✔	✔	✔	✔	✔
Kogia sima	✔	✔	✔	✔	✔	✔	✔	✔
Physeter macrocephalus	✔	✔	✔	✔	✔	✔	✔	✔
Balaenoptera acutorostrata	✔	✔	✔	✔	✔	✔	✔	✔
Banaenoptera omurai	✔	✔	✔	✔	✔	✔	✔	✔

Species	EDNRB							
	Exon1	Exon2	Exon3	Exon4	Exon5	Exon6	Exon7	Exon8
Tursiops truncatus	✔	✔	✔	✔	✔	✔	✔	✔
Delphinus capensis	✔	✔	✔	✔	✔	✔	✔	✔
Sousa chinensis	✔	✔	✔	✔	✔	✔	✔	✔
Neophocaena phocaenoides	✔	✔	✔	✔	✔	✔	✔	✔
Delphinapterus leucas	✔	✔	✔	✔	✔	✔	✔	✔
Lipotes vexillifer	✔	✔	✔	✔	✔	✔	✔	✔
Mesoplodon densirostris	✔	✔	✔	✔	✔	✔	✔	✔
Kogia sima	✔	✔	✔	✔	✔	✔	✔	✔
Physeter macrocephalus	✔	✔	✔	✔	✔	✔	✔	✔
Balaenoptera acutorostrata	✔	✔	✔	✔	✔	✔	✔	✔
Banaenoptera omurai	✔	✔	✔	✔	✔	✔	✔	✔

Species	AGTR1
	Exon1
Tursiops truncatus	✔
Delphinus capensis	✔
Sousa chinensis	✔
Neophocaena phocaenoides	✔
Delphinapterus leucas	✔
Lipotes vexillifer	✔
Mesoplodon densirostris	✔
Kogia sima	✔
Physeter macrocephalus	✔
Balaenoptera acutorostrata	✔
Banaenoptera omurai	✔

Species	AGTR2
	Exon1
Tursiops truncatus	✔ Partial sequences
Delphinus capensis	✔ Partial sequences
Sousa chinensis	✔ Partial sequences
Neophocaena phocaenoides	✔ Partial sequences
Delphinapterus leucas	✔ Partial sequences
Lipotes vexillifer	✔ Partial sequences
Mesoplodon densirostris	✔ Partial sequences
Kogia sima	✔ Partial sequences
Physeter macrocephalus	✔ Partial sequences
Balaenoptera acutorostrata	✔ Partial sequences
Banaenoptera omurai	✔ Partial sequences

Table S3 List of primers used to amplify the coding regions of hypoxia-tolerance-related genes in this study.

Gene	Sense Primer	Antisense Primer	Melting Temp (°C)	Product Size (bp)
HBA	GGACCAGGATAGCACTTC	GGACCAGGATAGCACTTC	50-55	1384
HBB	GAAGGGAACTGAGGAACA	GTGAGCAAACAGGGTAAA	50-55	820
	GGACCAGACAATCACAGA	AGACAACCAGCAGCCTAA	50-55	451
	YAGCKRCTGCTTACACTTGC	CTCTTCCATTCTAAACTGC	50-55	610
	TGCTTCTTGAAATGGTCCTA	GAATATGGTCTTCCGATGGT	50-55	412
MB	GCCTTGCTCTTCTTCTCT	CTGACTCCAATCTCTGACA	50-55	302
	CTTGCTCTTCTTCTCTTCTTC	TCCTGACTCCAATCTCTGA	50-55	302
	GCTCTTCTTCTCTTCTTCAG	GCTCTAGGCATACAGGATT	50-55	375
	CATTTAATCCTCACGCTTCT	CACTTCCTCCTTGGTAACA	50-55	648
	TGCCTGGTAGATAGTAAGTG	TAGGTGAAAGGAGAGTTGAG	50-55	557
NGB	CTTCCTGTCCCAGTCCA	GAGTCGCCAGCCTCTAA	50-55	731
	TCCTATGCCAGGTCGTATGAG	GCTCAAAGCCCTAAACACAC	50-55	752
	GGTTCTGGATTATGGTCGTG	AGGCTGTTCTGGGATGG	50-55	496
	TCTCTTCCTTCTGGTTGG	GGCTCCTTCACTTCCTAT	50-55	459
	CTCAGGTTCATCTCAGGTT	GGCTACAATGGACACAGA	50-55	586
EDN1	TTCGGGTTTAGTTCGCAGTG	ATTCCCTTGACATCCTTTCT	50-55	866
	CCAAATGTCTAACAGGCAGTAT	TCTAGGTCTCGCTCTACCATAA	50-55	915
	CCAAATGTCTAACAGGCAGTAT	AGCCACGTATGACCTCAGCAAC	50-55	755
	CGAAACAGATACCAAAGC	CTCCCATCAAGGCAAGA	50-55	801
	TCACTTGAGGTTAGCAGGGTA	TGGAGGCAAATGAAGACG	50-55	483
EDN2	AGGGCACCGAGGCAATG	GCAGTGCTTGGCGTAGGGA	50-55	1110
	GGTGGGACTTCTTGAGGTGAG	GGGCAGTGAGATGGGTTGG	50-55	484
	CGGTCCTGGTTTATCTAATCCTC	TTTCTACTTTCCCAAATGCTCC	50-55	777
	GGAGCATTTGGGAAAGTAGAA	TCTGGTGGTAACAAAGAGCATAG	50-55	654

	GCACAGGGTGGAAATGGA	CCTTGGGCAAACACCTTCA	50-55	315
	TTGGGTGACGGAAGGTAAA	ACGCCGTTCCTTTCAGC	50-55	713
	TTGGGTGACGGAAGGTAAA	ACCCATCCAGCCTCTACGG	50-55	833
EDN3	CAGCGGGCACAGATA	CCCACAGGAGGTAAGGA	50-55	1075
	CCACAGTTGGGTCTTGA	CACATCGGGTGGAACA	50-55	680
	TTCGAGTGGGACATGC	CATACGGCAGCGACT	50-55	751
	AGGTGACATTGGCATAAAG	TGAACCCAGCAGAAGC	50-55	857
EDNRA	ACTGCCATGTCCTCATAT	CTCCCACTCAGGCTCTAA	50-55	1491
	CTACTCAAAGGAAAAACTGCC	TCTTACCCTGGAAATGGA	50-55	1139
	GTCTCTTGCTACCACACTAA	TTCGTTGCCGTAAGGATG	50-55	526
	TAAGCCGTCATCCAAACC	GTTGCCTGAACCAAAGAT	50-55	644
	AACTCCAAATGGGTAGAA	ATAAGCATCAGCGTGTCT	50-55	753
	CACCTCCAAGCCCATTAC	CTGGCACAGTTCCAACCC	50-55	745
	TTGAAGAGGTGGAGGCAT	GAGAATGTAGAGTAGGTGGTT	50-55	379
	GCTGACATAGCCCACAAC	CCTGCACCCTTGAACTCT	50-55	1287
	TTCTGACTGCCACCTATAAG	CCTCTGAATATGTTGTTGTCTG	50-55	444
	TTACAAACTCCCTACCCT	AAACCAGTCAACCCTTCA	50-55	517
	CCAGCCCTTTCCACTCAT	CCGCCAACTCCCACCATT	50-55	767
	GTTCTGCCTCTTATTATCC	GGTGACCAGTTCTTACGG	50-55	1256
EDNRB	CATCCTTTCCTCCCTGTC	GGTTCCAATGATTACTGACTT	50-55	952
	TCTCCCTCTGGGCACAAC	GGTTCCAATGATTACTGACTT	50-55	998
	CCACTTCCTTGTTCCAGA	CAGAGCCCAGTTTTATTG	50-55	401
	CCACTTCCTTGTTCCAGA	GGCAGAGCCCAGTTTTAT	50-55	403
	GCCCTTTGGAGTTGAGAT	TCTGTATTCCTGAGCACC	50-55	660
	AGGCGATGACAGGGACAG	TGGTTGAGGCAGTAGGGA	50-55	885
	AATAAAACTGGGCTCTGC	GGGCGACTTCTACTTGAT	50-55	635

	TGAGGCGATGACAGGGAC	CCGCTCACCGCAACTATA	50-55	547
	TAACTTGCCTATGTGAAGAC	CTACTCGCTCATTTGGAC	50-55	1242
	GCTTCGTGTCTCCTCAAC	TGCTGTGCTGCTTCATAC	50-55	804
AGTR1	GCTGTCTACACTGCTAT	ATCTCCTGTTGCTCCTCT	50-55	1217
	CCCCACGTCTACTTGTCT	AGCCTTCTTTAGGGTCTT	50-55	962
AGTR2	TTACCAGCAGTCTTCACG	CAAGTAATCGGAACCCTA	50-55	973

Table S5 Previously reported maximum dive durations for cetaceans

Cetacean species	**Family**	**Max. dive time (min)**	**References**
Odontocetes			
Delphinus delphis	Delphinidae	5	Evans 1971
Tursiops truncatus	Delphinidae	8	Ridgway and Harrison 1986
Delphinus capensis	Delphinidae	5	Evans 1971
Sousa chinensis	Delphinidae	5	Parra et al. 2004
Orcinus orca	Delphinidae	10.4	Baird et al.2005
Neophocaena phocaenoides	Phocoenidae	2.48	Akamatsu et al. 2002
Delphinapterus leucas	Delphinidae	18.3	Martin et al. 1993
Lipotes vexillifer	Lipotidae	2	Parra et al. 2004
Mesoplodon densirostris	Ziphiidae	57.4	Tyack et al. 2006
Physeter macrocephalus	Physeteridae	73	Watkins et al. 1993
Kogia sima	Kogiidae	43	Breese and Tershy 1993
Mysticetes			
Balaenoptera acutorostrata	Balaenopteridae	13.4	Croll et al. 2001

Table S6 One-ratio model analysis in 17 genes

Genes	**Models**	**-lnL**	**2 ΔLnL**	**P value**	**ω values**
HBA	M0(one-ratio)	3088.622			ω0= 0.200
	M0(constrained)	3224.482	271.72	<0.001	ω0= 1
HBB	M0(one-ratio)	3528.038			ω0= 0.283
	M0(constrained)	3627.432	198.788	<0.001	ω0= 1
MB	M0(one-ratio)	3945.824			ω0= 0.097
	M0(constrained)	4331.292	770.936	<0.001	ω0= 1
NGB	M0(one-ratio)	1987.589			ω0= 0.073
	M0(constrained)	2220.366	465.554	<0.001	ω0= 1
AGTR1	M0(one-ratio)	5917.725			ω0= 0.081
	M0(constrained)	6514.433	1193.416	<0.001	ω0= 1
AGTR2	M0(one-ratio)	3569.236			ω0= 0.080
	M0(constrained)	3869.325	600.178	<0.001	ω0= 1
EDN1	M0(one-ratio)	6094.134			ω0= 0.337
	M0(constrained)	6227.731	267.194	<0.001	ω0= 1
EDN2	M0(one-ratio)	3507.497			ω0= 0.386
	M0(constrained)	3563.845	112.696	<0.001	ω0= 1
EDN3	M0(one-ratio)	6548.742			ω0= 0.460
	M0(constrained)	6624.958	152.432	<0.001	ω0= 1
EDNRA	M0(one-ratio)	6112.603			ω0= 0.089
	M0(constrained)	6633.350	1041.494	<0.001	ω0= 1
EDNRB	M0(one-ratio)	7887.625			ω0= 0.214
	M0(constrained)	8229.283	683.316	<0.001	ω0= 1
ADRA1A	M0(one-ratio)	7962.800			ω0= 0.053
	M0(constrained)	9146.433	2367.266	<0.001	ω0= 1
ADRA1B	M0(one-ratio)	7389.238			ω0= 0.044
	M0(constrained)	8504.499	2230.522	<0.001	ω0= 1
ADRA1D	M0(one-ratio)	10490.310			ω0= 0.075
	M0(constrained)	11694.867	2409.114	<0.001	ω0= 1
AVP	M0(one-ratio)	3365.857			ω0= 0.092
	M0(constrained)	3688.287	644.860	<0.001	ω0= 1
AVPR1A	M0(one-ratio)	8878.038			ω0= 0.128
	M0(constrained)	9567.991	1379.906	<0.001	ω0= 1
AVPR1B	M0(one-ratio)	9980.951			ω0= 0.166
	M0(constrained)	10601.350	1240.798	<0.001	ω0= 1

Table S7 Selective pressure analyses in cetaceans-only dataset using site model in PAML.

Genes	**Site Model**	**-ln *L***	**2 Δ(ln *L*)**	***P* value**	**ω values**
HBA	M8a	976.937			ω= 1
	M8	973.857	6.16	0.013	ω= 2.459
HBB	M8a	995.882			ω= 1
	M8	991.361	9.042	0.003	ω= 4.860
MB	M8a	1194.849			
	M8	1192.871	3.956	0.047	ω= 1.796
EDN2	M8a	1201.265			ω= 1
	M8	1197.744	7.042	0.008	ω= 14.891
ADRA1D	M8a	2767.301			ω= 1
	M8	2762.365	9.872	0.002	ω= 26.247

Table S8 All Amino acid sites under positive selection identified by PAML, DATAMONKET and TreeSAAP

Gene	**Site**	**PAML**		**Datamonkey**						**TreeSAAP properties**	
	Position	**Site Model (M8)**^{a}	**Branch-Site Model**^{a}	**FEL** $p<0.2$ ^{b}	**REL BF>50**^{b}	**FUBAR pp>0.8**^{b}	**AA Changes**	**Parallel Changes**	**Clade**^{c}	**Radical Changes in AA Properties**^{d}	**Total**
HBA	16	0.912		0.106	450.038	0.820	G-A		a	P_α, P_c, P	3
							A-S		h	P_α, P_c, P	3
							A-G		c1	P_α, P_c, P	3
	18	0.897		0.134	406	0.805	V-I		a	pK'	1
							I-V	Yes	b1, n	pK'	1
	20	0.989			149.32	0.901	G-S		a		
							S-N	Yes	d, m1		
							N-S		v		
							S-G	Yes	c1, i1		
	22	0.981		0.084	475.336	0.836	A-G	Yes	d, m1	P_α, P_c, P	3
							G-S		e		
	25	0.920		0.116	416.941	0.839	Y-F	Yes	n,b1	F	1
	52	0.843			176.704	0.859	G-N		N		
							G-D		m1	α_c, P_r, E_t,	3
	74	0.802			134.561		L-M		a		
							M-I		d	pK'	1
							M-L		n		

	76	0.864			146.049		D-N		a	α_c,	1
							N-T	Yes	n, c1		
							T-A		i1	P_α,	1
	112	0.886			178.977		A-V		a	P_β	1
							V-L	Yes	d, m1		
							V-A		n	P_β	1
HBB	3		0.938				H-Q		f1		
	11		0.928				A-G		f1	P_α, P_c, P	3
	12		0.995				V-L		f1		
	45	0.953		0.092	442.775	0.963	S-A	Yes	q, c1	P_α, P_c, P	3
							S-H		n	F	1
	51		0.943				T-S		f1		
	60	0.917			56.360		K-N		b	pH_i	1
			0.988				N-S		i1		
			0.960				N-K		f1	pH_i	1
	62		0.997				K-Q		f1	pH_i	1
	63	0.994		0.200	384.186	0.958	A-S		i1	P_α , P_c , P	3
			1.000				A-K		c	K^o, B_r, E_{sm}, E_t, h	5
	73	0.943	0.936		179.09	0.927	S-G	Yes	e, f1		
							G-S		a1		

	88	0.891	0.933	0.169	308.649	0.926	T-Q		i1		
							T-A		c	P_α	1
	134		0.971				V-L		i1		
MB	6	0.861		0.140			G-A	Yes	l, q	P_α, P_c, P	3
	36	0.921		0.102		0.896	G-S		n		
							S-H		p	F	1
	52	0.872		0.141			T-S	Yes	l, p		
	75			0.159		0.875	G-A	Yes	d, n	P_α, P_c, P	3
							A-G		k	P_α, P_c, P	3
	130	0.856		0.151			G-A	Yes	r, e1	P_α, P_c, P	3
EDN1	125			0.199		0.850	Q-E		i		
EDN2	12			0.039	761.565	0.965	V-A		c	P_β	1
							V-I		i1	pK'	1
	114	0.942		0.023	190.831	0.984	R-Q	Yes	z, c1	pH_i	1
							R-L		l1	N_s, α_n, pH_i, B_r, R_F, h, H_{nc}, p, E_{sm}, E_t	10
	162	0.937		0.057	70.988	0.977	Q-R	Yes	f, c1, m1	pH_i	1
EDN3	34			0.144		0.838	T-A		f1	P_α	1
							T-R		m1	pH_i, H_{nc}, E_{sm},	3
	152			0.155		0.843	E-Q		a		0

							Q-L		z	N_s, B_r, h, F, p, E_t	6
							Q-R		m1	pH_i	1
	167			0.169		0.842	R-G		d	C_a, pH_i, M_v, M_w, H_{nc}, V^0, u, E_{sm}, E_t	9
EDNRA	19			0.196	63.872	0.912	I-V		f	Pk'	1
							I-F		m1		
	22			0.173	73.036	0.916	N-D		f1	α_c	1
							N-S		i1		
	161				54.314	0.878	F-L	Yes	z, b1		
EDNRB	80				10721.8	0.888	T-I	Yes	f, b1	N_s, B_r, Pk', R_a, H_p, H_t	6
							T-A		f1	P_α	1
	347			0.195	11076.6		L-R		f	N_s, B_r, R_F, h, pH_i, H_{nc}, p, α_n, $E_{sm,}$, E_t	10
							R-H		w	E_{sm}	1
ADRA1D	454	0.988		0.041	436.44	0.968	P-T		a	K^o, H_t	2
							T-A	Yes	f, g1	P_α	1
	475	0.867					G-R		b1	C_a, pH_i, M_v, M_w, H_{nc}, V^0,μ, E_{sm}, E_t	9
	483	0.909					Q-P		b1	P_c, K^o,α_M, H_t	3
AVP	159				1111.96	0.852	A-S		t	P_α, P_c, P	3
	162				1271.98	0.885	G-P		t	B_l, α_c, H_t	3

[a] Codons identified by PAML as under positive selection along with Bayesian (BEB) analysis PPs for sites with P>80% under M8 and branch-site models.

[b] Codons were estimated in DATAMONKEY.

[c] Amino acid substitution occurred along clades, with detailed information marked in Figure 1.

[d] Radical changes in amino acid properties under category 6-8 were detected in TreeSAAP. Physicochemical amino acid properties available in TreeSAAP are as following: $\boldsymbol{\alpha_c}$: Power to be - C-term., α-helix; $\boldsymbol{\alpha_n}$: Power to be in the N-terminal of an α-helix; $\boldsymbol{B_r}$: Buriedness; $\boldsymbol{C_a}$: Helical contact energy; $\boldsymbol{E_l}$: Long-range non-bonded energy; $\boldsymbol{E_{sm}}$: Short and medium range non-bonded energy; $\boldsymbol{E_t}$: Total non-bonding Energy; $\boldsymbol{F}$: Mean r.m.s. fluctuation displacement; $\boldsymbol{h}$:Hydropathy; $\boldsymbol{H_{nc}}$: Normal consensus hydrophobicity; $\boldsymbol{H_p}$: Surrounding hydrophobicity; $\boldsymbol{H_t}$: Thermodynamic transfer hydrophobicity; $\boldsymbol{K^o}$: Compressibility; $\boldsymbol{\mu}$: Refractive index; $\boldsymbol{M_v}$: Molecular volume; $\boldsymbol{M_w}$: Molecular weight; $\boldsymbol{N_s}$: Average number of surrounding residues; $\boldsymbol{P_\alpha}$: α- helical tendencies; $\boldsymbol{P_\beta}$: β-structure tendencies; $\boldsymbol{P_c}$: Coil tendencies; $\boldsymbol{P}$: Turn tendencies; $\boldsymbol{p}$: Polarity; $\boldsymbol{pH_i}$: Isoelectric point; $\boldsymbol{pK'}$: Equilibrium Constant of ionization for COOH; $\boldsymbol{P_r}$: Polar requirement; $\boldsymbol{R_a}$: Solvent accessible reduction ratio; $\boldsymbol{R_F}$: Chromatographic index; $\boldsymbol{V^0}$: Partial specific volume;

Table S9 Habitat information of species habitually exposed to hypoxia environment in this study

Species	Genus	Family	Order	Habitat information	Reference
Bos grunniens	Bos	Bovidae	Artiodactyla	High-altitude environments of Tibetan Plateau, a low partial pressure of oxygen and high ultraviolet radiation	Qiu et al. 2012
Bos mutus	Bos	Bovidae	Artiodactyla	High-altitude environments of Tibetan Plateau, a low partial pressure of oxygen and high ultraviolet radiation	Leslie and Schaller 2009
Pantholops hodgsonii	Pantholops	Bovidae	Artiodactyla	High-altitude environments of Tibetan Plateau, a low partial pressure of oxygen and high ultraviolet radiation	Ge et al. 2013
Leptonychotes weddellii	Leptonychotes	Phocidae	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009
Mirounga angustirostris	Mirounga	Phocidae	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009
Phoca groenlandica	Phoca	Phocidae	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009
Pusa hispida	Pusa	Phocidae	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009
Cystophora cristata	Cystophora	Phocidae	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009

Erignathus barbatus	Erignathus	Phocidae	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009
Callorhinus ursinus	Callorhinus	Otariidae;	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009
Odobenus rosmarus divergens	Odobenus	Odobenidae	Carnivora	Marine mammals, substantial parts of their lives moving within the water, aquatic foraging, performing extreme dives	Stewart 2009
Condylura cristata	Condylura	Soricomorpha	Insectivora	Semi-aquatic habit, adept burrower and excellent swimmer	Hamilton 1931; Hickman 1984
Spalax judaei	Nannospalax	Spalacidae	Rodentia	Subterranean, entire life in underground burrows, extremely hypoxic environments	Fang et al. 2014
Spalax carmeli	Nannospalax	Spalacidae	Rodentia	Subterranean, entire life in underground burrows, extremely hypoxic environments	Fang et al. 2014
Spalax golani	Nannospalax	Spalacidae	Rodentia	Subterranean, entire life in underground burrows, extremely hypoxic environments	Fang et al. 2014
Heterocephalus glaber	Heterocephalus	Bathyergidae	Rodentia	Inhabit subterranean burrows in extremely hard soils, colony	Sherman et al 2002
Ochotona curzoniae	Ochotona	Ochotonidae	Lagomorpha	High-altitude environments of Tibetan Plateau, a low partial pressure of oxygen and high ultraviolet radiation	Smith and Foggin. 1999
Trichechus manatus latirostris	Trichechus	Trichechidae	Sirenia	Marine mammals, aquatic herbivore, diving	Stewart 2009

Table S10 Convergent/parallel amino acid changes detected in hypoxia-exposed species

Gene	**Parallel Substitution**	**Branch Pair**	**Observed Number**	**Expected Number**	**P value**
ADRA1D	P50A A53G	*Heterocephalus glaber* .VS. ancestor of whales	2	0	0
AVPR1B	I/V270T	*Heterocephalus glaber* .VS. ancestor of whales	1	0	0

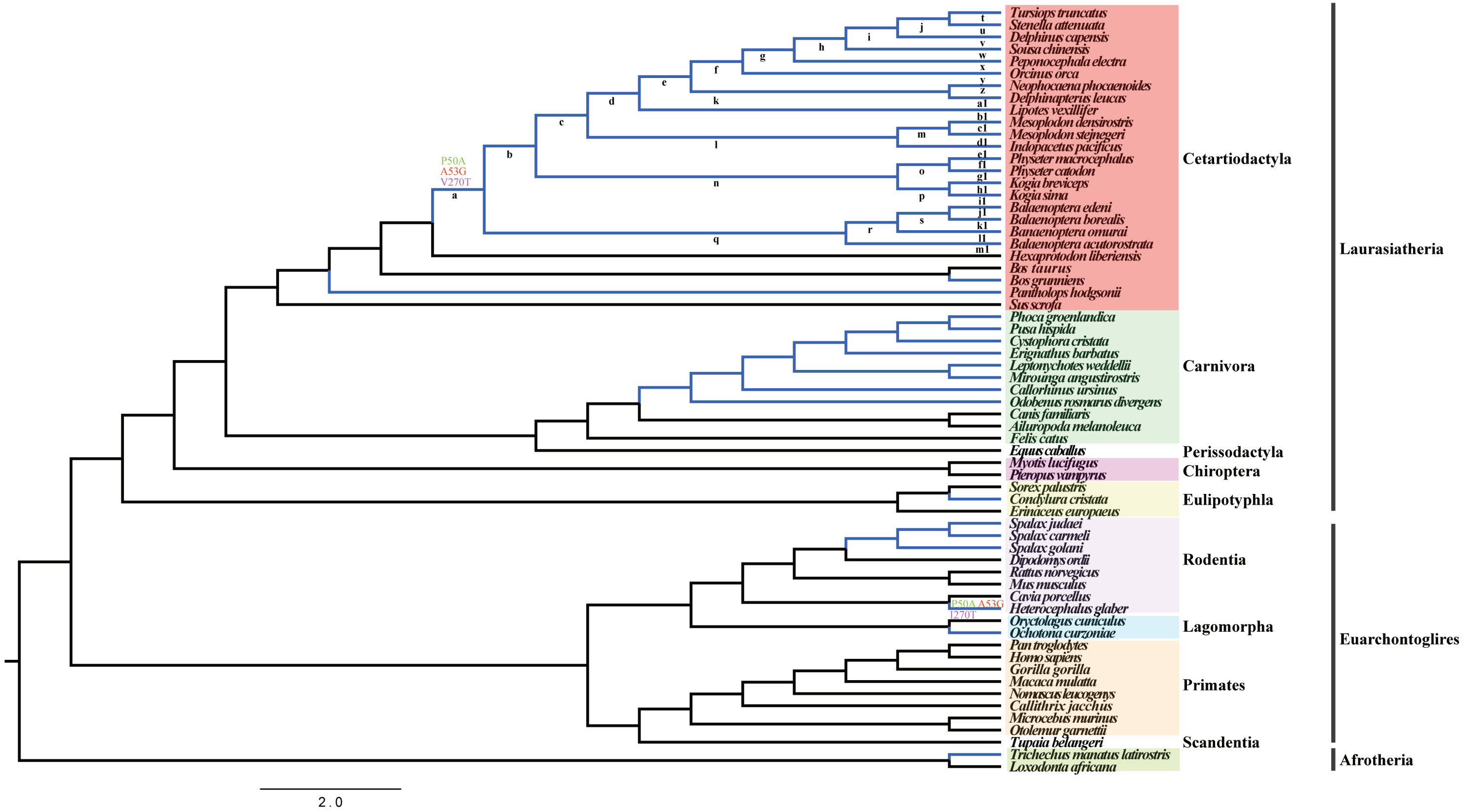
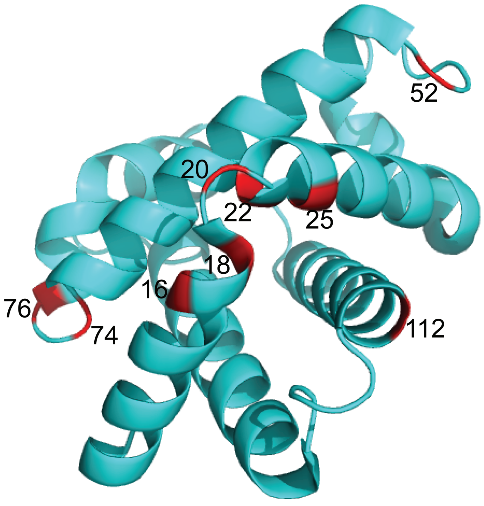
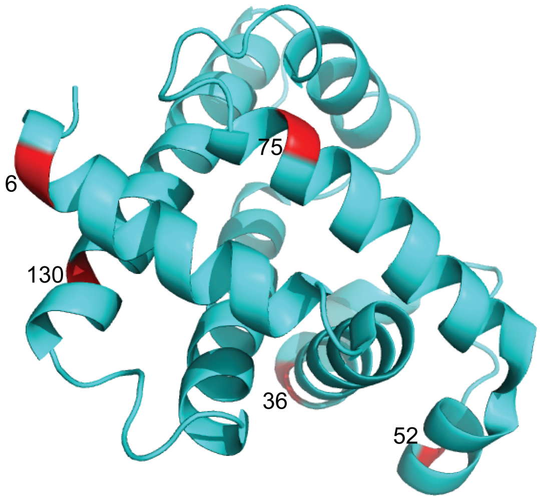


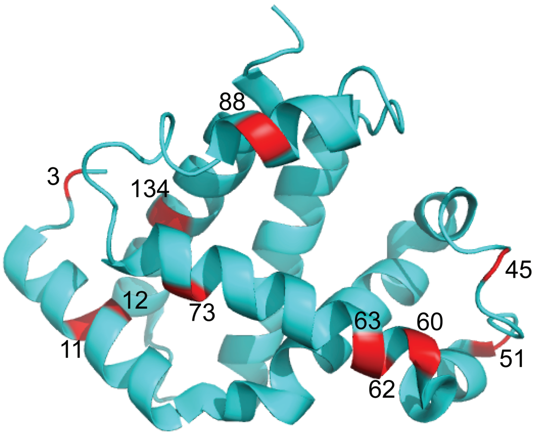
Figure. S1 The tree topology used to conduct the selective pressure analysis in PAML. This tree is based on OrthoMaM (Ranwez et al. 2007). Different orders of mammals were marked with different colors. Blue color branches in the tree represented cetaceans (branch: a-m1) (fig. 1) and other hypoxia species. They are used in the branch-site models tests, and results listed in table 2 and fig. 1. Convergent/parallel amino acid changes were identified the blue lineages leading to hypoxia-exposed animals. The symbols above the branches are the amino acid replacements (ADRA1D: P50A, A53G**,** AVPR1B: I/V270T).



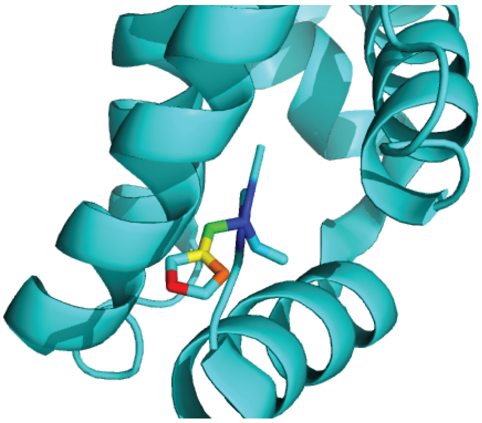
HBA



MB

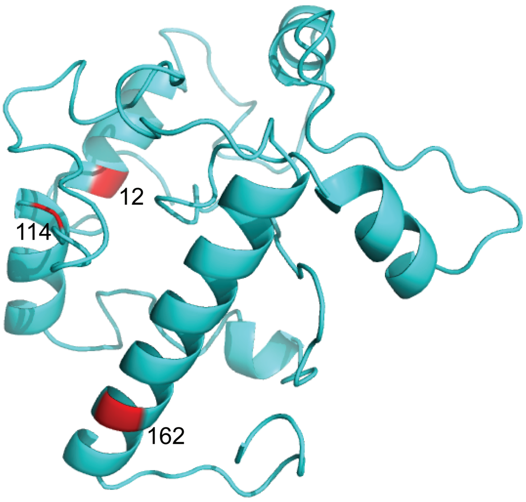


HBB

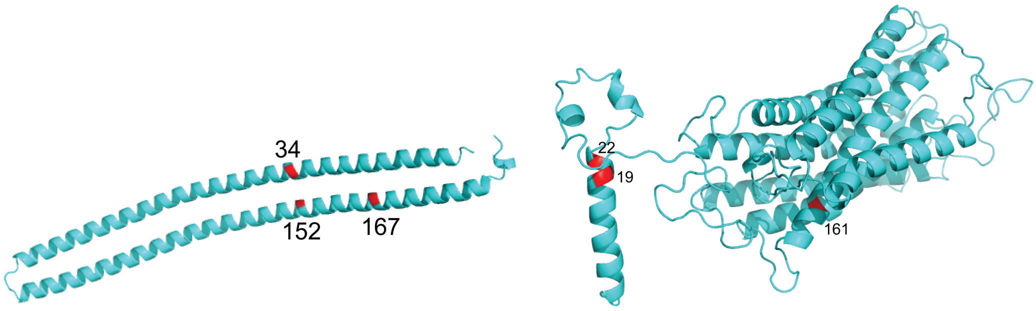


HBB residue 3

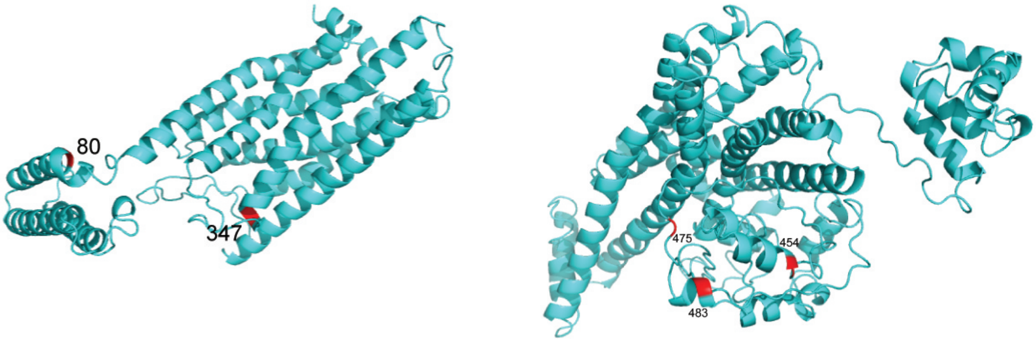
EDN1



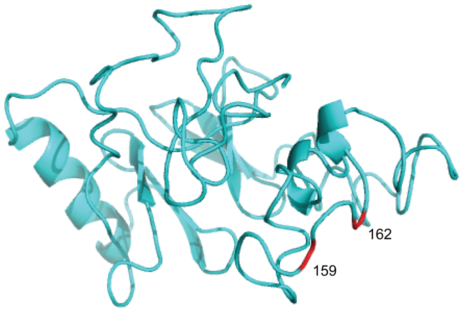
EDN2



EDN3 EDNRA

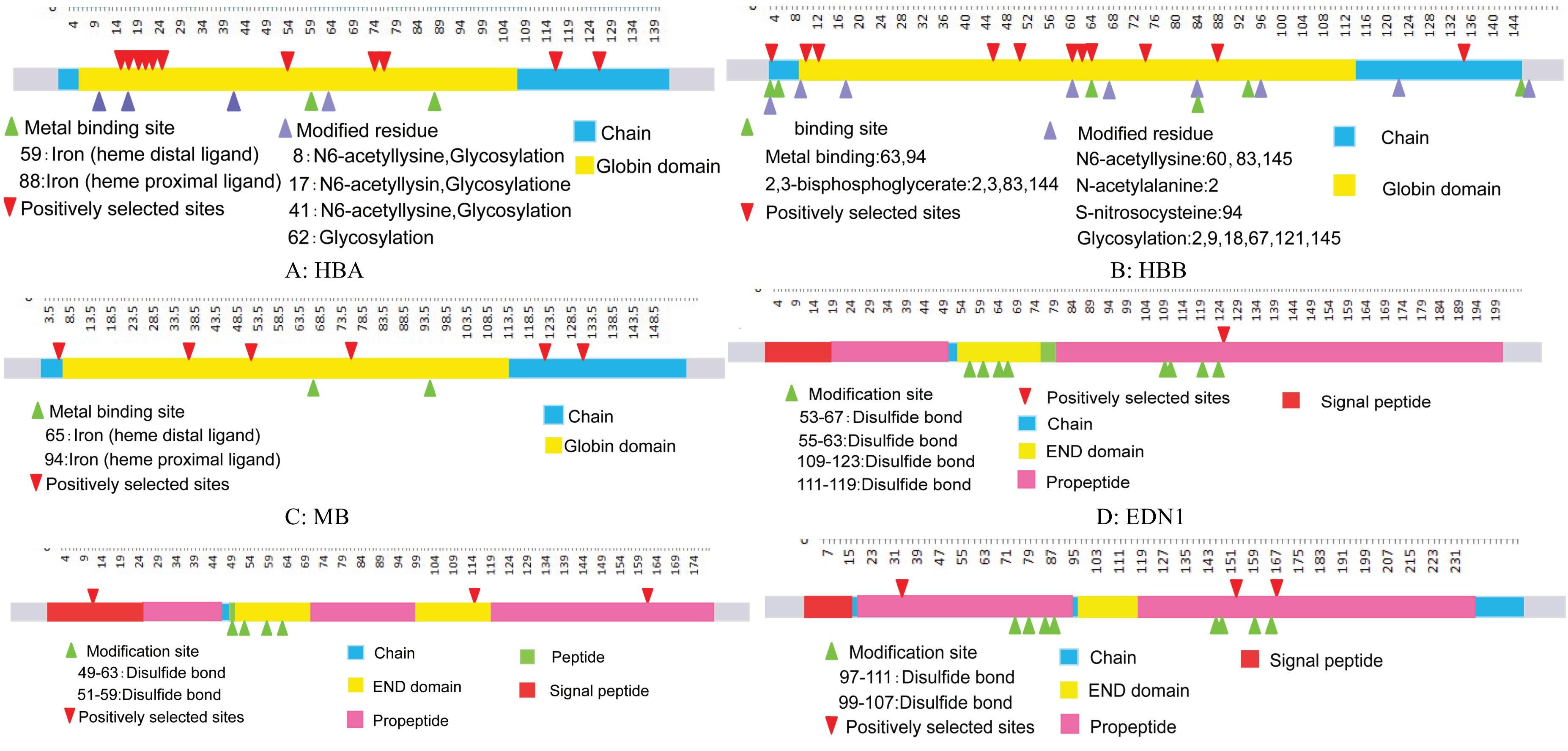


EDNRB ADRA1D



AVP

Figure S2 Distribution of mutations in the three-dimensional structure of HBA, HBB, MB, EDN1, EDN2, EDN3, EDNRA, EDNRB, ADRA1D, and AVP genes for cetacean. The figures were created using PyMOL (http://www.pymol.org).



A: HBA

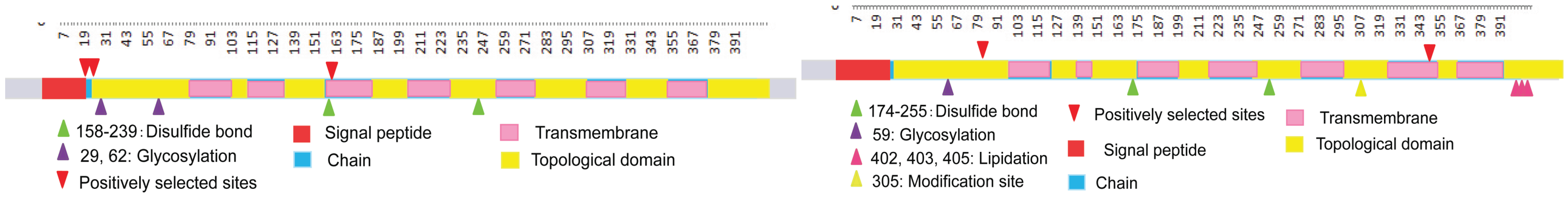
B: HBB

C: MB

D: EDN1

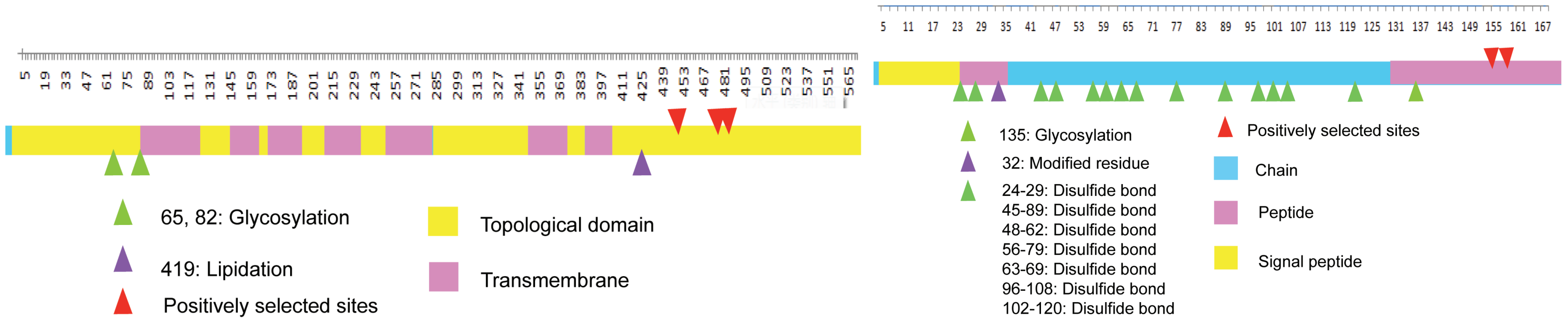
E: EDN2

F: EDN3



G: EDNRA

H: EDNRB



I: ADRA1D

J: AVP

Figure S3 Secondary structure and distribution of positively selected sites in relation to the functionally important regions for gene HBA, HBB, MB, EDN1, EDN2, EDN3, EDNRA, EDNRB, ADRA1D, and AVP.

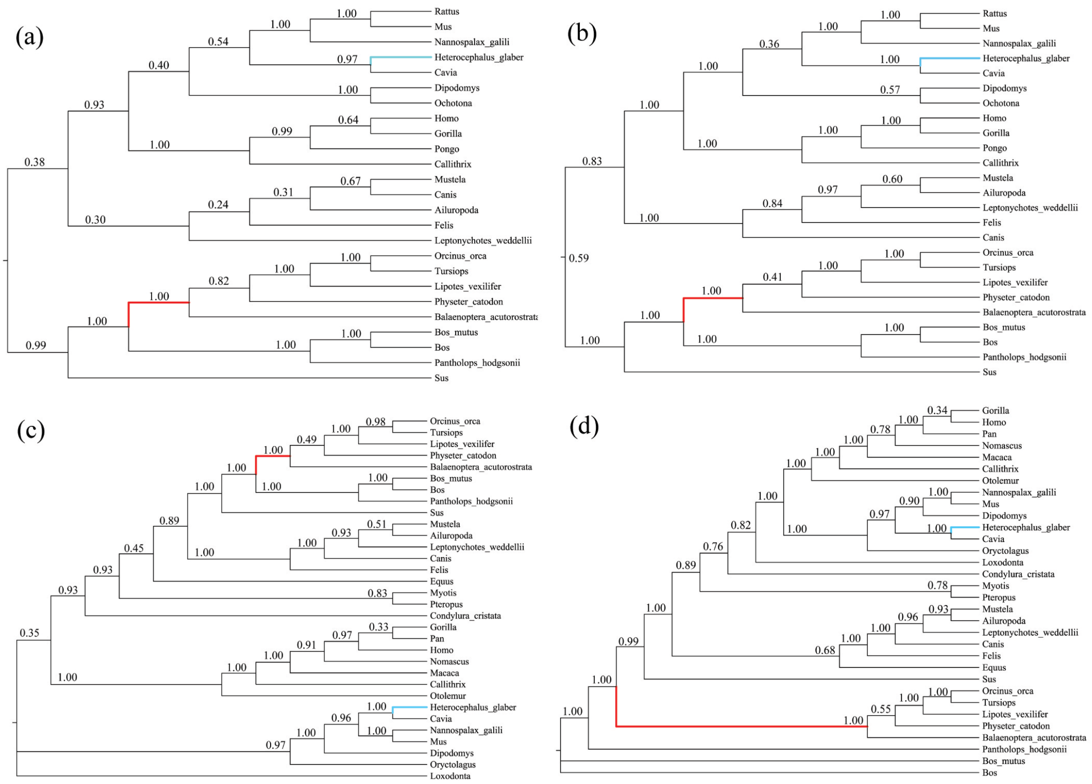
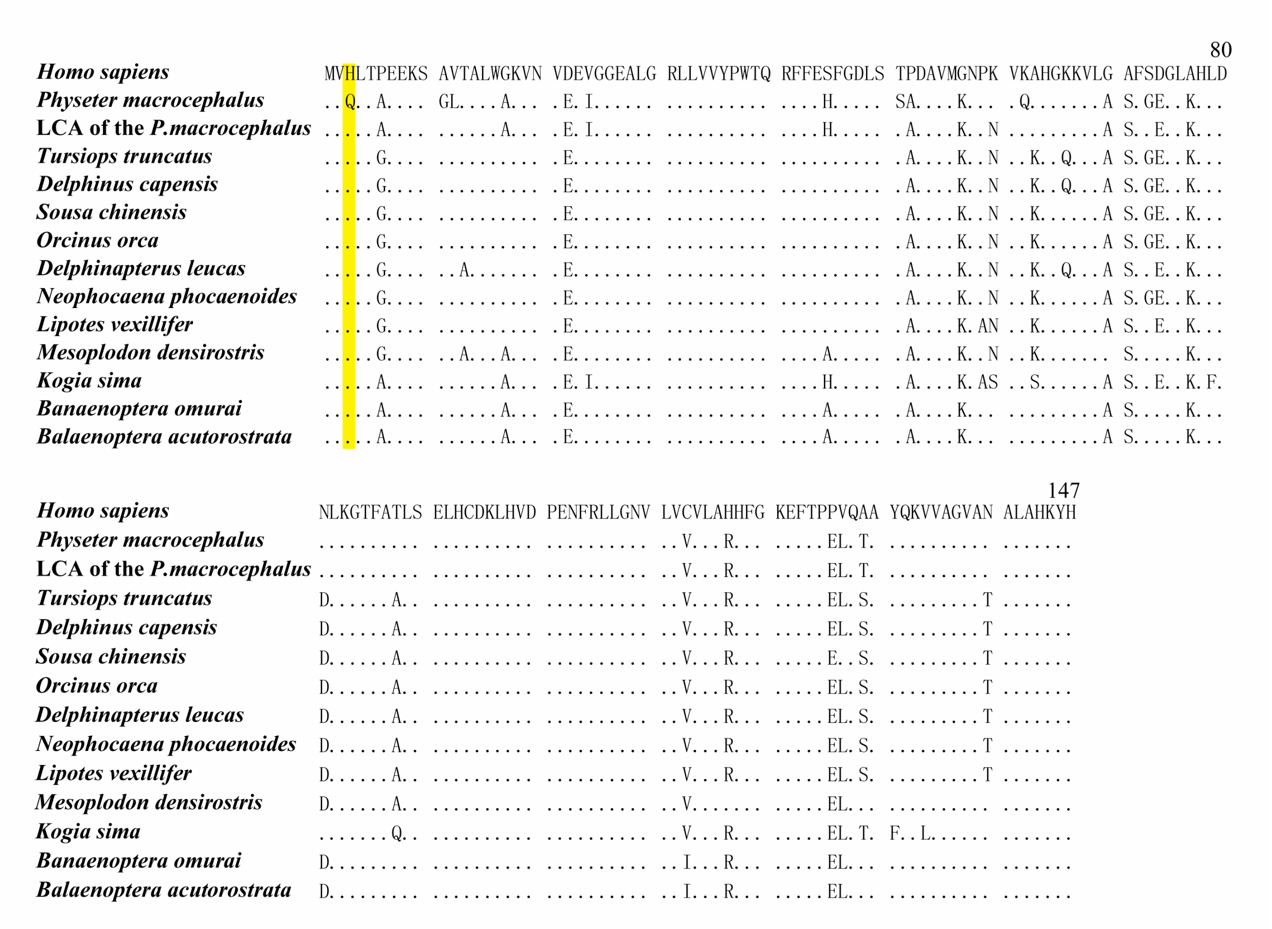


Figure S4 The phylogenetic topology of the ADRA1D and AVPR1B gene based on nucleotide sequences and amino acid sequences. Numbers above the branches are the Bayesian posterior probabilities. (a) amino acid tree of ADRA1D; (b) nucleotide tree of ADRA1D; (c) amino acid tree of AVPR1B; (d) nucleotide tree of AVPR1B. Red and blue branches indicate the lineage along the ancestor of cetaceans and the naked mole rat, respectively.



```
                                                                                                         80
Homo sapiens                 MVHLTPEEKS AVTALWGKVN VDEVGGEALG RLLVVYPWTQ RFFESFGDLS TPDAVMGNPK VKAHGKKVLG AFSDGLAHLD
Physeter macrocephalus       ..Q..A.... GL....A... .E.I...... .......... ....H..... SA....K... .Q.......A S.GE..K...
LCA of the P.macrocephalus   .....A.... ......A... .E.I...... .......... ....H..... .A....K..N .........A S..E..K...
Tursiops truncatus           .....G.... .......... .E........ .......... .......... .A....K..N ..K..Q...A S.GE..K...
Delphinus capensis           .....G.... .......... .E........ .......... .......... .A....K..N ..K..Q...A S.GE..K...
Sousa chinensis              .....G.... .......... .E........ .......... .......... .A....K..N ..K......A S.GE..K...
Orcinus orca                 .....G.... .......... .E........ .......... .......... .A....K..N ..K......A S.GE..K...
Delphinapterus leucas        .....G.... ..A....... .E........ .......... .......... .A....K..N ..K..Q...A S..E..K...
Neophocaena phocaenoides     .....G.... .......... .E........ .......... .......... .A....K..N ..K......A S.GE..K...
Lipotes vexillifer           .....G.... .......... .E........ .......... .......... .A....K.AN ..K......A S..E..K...
Mesoplodon densirostris      .....G.... ..A...A... .E........ .......... ....A..... .A....K..N ..K.......  S.....K...
Kogia sima                   .....A.... ......A... .E.I...... .......... ....H..... .A....K.AS ..S......A S..E..K.F.
Banaenoptera omurai          .....A.... ......A... .E........ .......... ....A..... .A....K... .........A S.....K...
Balaenoptera acutorostrata   .....A.... ......A... .E........ .......... ....A..... .A....K... .........A S.....K...

                                                                                        147
Homo sapiens                 NLKGTFATLS ELHCDKLHVD PENFRLLGNV LVCVLAHHFG KEFTPPVQAA YQKVVAGVAN ALAHKYH
Physeter macrocephalus       .......... .......... .......... ..V...R... .....EL.T. .......... .......
LCA of the P.macrocephalus   .......... .......... .......... ..V...R... .....EL.T. .......... .......
Tursiops truncatus           D......A.. .......... .......... ..V...R... .....EL.S. .........T .......
Delphinus capensis           D......A.. .......... .......... ..V...R... .....EL.S. .........T .......
Sousa chinensis              D......A.. .......... .......... ..V...R... .....E..S. .........T .......
Orcinus orca                 D......A.. .......... .......... ..V...R... .....EL.S. .........T .......
Delphinapterus leucas        D......A.. .......... .......... ..V...R... .....EL.S. .........T .......
Neophocaena phocaenoides     D......A.. .......... .......... ..V...R... .....EL.S. .........T .......
Lipotes vexillifer           D......A.. .......... .......... ..V...R... .....EL.S. .........T .......
Mesoplodon densirostris      D......A.. .......... .......... ..V....... .....EL... .......... .......
Kogia sima                   .......Q.. .......... .......... ..V...R... .....EL.T. F..L...... .......
Banaenoptera omurai          D......... .......... .......... ..I...R... .....EL... .......... .......
Balaenoptera acutorostrata   D......... .......... .......... ..I...R... .....EL... .......... .......
```

Figure S5. An alignment of HBB gene in cetaceans. The sequence of human on the top, and amino acids identical to human HBB are indicated by dots. The important site 3 discussed in the text was shaded by yellow. The alignment of sperm whale (*Physeter macrocephalusdouc*) and its last common ancestor (LCA) show the species-specific amino acid change (H3Q).

References:

Akamatsu T, et al. 2002. Diving behaviour of freshwater finless porpoises (*Neophocaena phocaenoides*) in an oxbow of the Yangtze River, China. ICES J Mar Sci. 59:438–443.

Baird RW, Hanson MB, Dill LM. 2005. Factors influencing the diving behaviour of fish-eating killer whales: sex differences and diel and interannual variation in diving rates. Can J Zool. 83:257–267.

Breese D, Tershy BR. 1993. Relative abundance of Cetacea in the Canal de Ballenas, Gulf of California. Mar Mammal Sci. 9:319–324.

Croll DA, Acevedo-Gutiérrez A, Tershy BR, Urbán-Ramírez J. 2001. The diving behavior of blue and fin whales: is dive duration shorter than expected based on oxygen stores? Comp Biochem Physiol A Mol Integr Physiol. 129:797–809.

Evans WE. 1971. Orientation behavior of delphinids: radio telemetric studies. Ann N Y Acad Sci. 188:142–160.

Fang X, et al. 2014. Genome-wide adaptive complexes to underground stresses in blind mole rats Spalax. Nat Commun. 5:3966–3966.

Ge RL, et al. 2013. Draft genome sequence of the Tibetan antelope. Nat Commun. 4:54–56.

Hickman GC. 1984. Swimming ability of talpid moles, with particular reference to the semi-aquatic *Condylura cristata*. Mammalia. 48:505–513.

Hamilton WJ.1931. Habits of the star-nosed mole, *Condylura cristata*. J Mammal. 4:345–355.

Leslie DM, Schaller GB. 2009. *Bos grunniens* and *Bos mutus* (Artiodactyla: Bovidae). Mammalian Species, 836:1-17.

Martin AR, Smith TG, Cox OP. 1993. Studying the behaviour and movements of high Arctic belugas with satellite telemetry. Symp Zool Soc Lond. 66: 195–210.

Parra GJ, Corkeron PJ, Marsh H. 2004. The Indo-Pacific humpback dolphin, *Sousa chinensis* (Osbeck, 1765), in Australian waters: a summary of current knowledge. Aquatic Mammals 30:197–206.

Qiu Q, et al. 2012. The yak genome and adaptation to life at high altitude. Nat Genet. 44: 946–949.

Ridgway SH, Harrison RJ. 1986. Diving Dolphins. In: Bryden MM, Harrison RJ, editors. *Research on Dolphins*. Oxford: Clarendon Press. p. 33–58.

Sherman PW, Jarvis JUM. 2002. Extraordinary life spans of naked mole-rats (*Heterocephalus glaber*). J Zool. 258:307–311.

Smith AT, Foggin JM. 1999. The plateau pika (*Ochotona curzoniae*) is a keystone species for biodiversity on the Tibetan plateau. Anim Conserv. 2:235–240.

Stewart B. 2009. Diving Behavior. In: Perrin WF, Würsig B, Thewissen JGM, editors. *Encyclopedia of Marine Mammals*. San Diego: Academic Press. p. 321–329.

Tyack PL, Johnson M, Soto NA, Sturlese A, Madsen PT. 2006. Extreme diving of beaked whales. J Exp Biol. 209:4238–4253.

Watkins WA, et al. 1993. Sperm whales tagged with transponders and tracked underwater by sonar. Mar Mammal Sci. 9:55–67.