# **BIOINFORMATION** Discovery at the interface of physical and biological sciences

open access

www.bioinformation.net

**Hypothesis** 

**Volume 8(25)** 

# Phylogenetic analysis of uroporphyrinogen III synthase (UROS) gene

### Abjal Pasha Shaik<sup>1,\$,\*</sup>, Abbas H Alsaeed<sup>1\$</sup> & Asma Sultana<sup>2\$</sup>

<sup>1</sup>Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia; <sup>2</sup>Janaki Nagar Colony, No: 45, Tolichowki, Hyderabad, India; Abjal Pasha Shaik – Email: afzalshaik@gmail.com; Phone: +966544023766; Fax: +96614684995; \*Corresponding author \$ All authors contribute equally

Received November 25, 2012; Accepted November 26, 2012; Published December 19, 2012

#### Abstract:

The uroporphyrinogen III synthase (UROS) enzyme (also known as hydroxymethylbilane hydrolyase) catalyzes the cyclization of hydroxymethylbilane to uroporphyrinogen III during heme biosynthesis. A deficiency of this enzyme is associated with the very rare Gunther's disease or congenital erythropoietic porphyria, an autosomal recessive inborn error of metabolism. The current study investigated the possible role of UROS (Homo sapiens [EC: 4.2.1.75; 265 aa; 1371 bp mRNA; Entrez Pubmed ref NP\_000366.1, NM\_000375.2]) in evolution by studying the phylogenetic relationship and divergence of this gene using computational methods. The UROS protein sequences from various taxa were retrieved from GenBank database and were compared using Clustal-W (multiple sequence alignment) with defaults and a first-pass phylogenetic tree was built using neighbor-joining method as in DELTA BLAST 2.2.27+ version. A total of 163 BLAST hits were found for the uroporphyrinogen III synthase query sequence and these hits showed putative conserved domain, HemD superfamily (as on 14<sup>th</sup> Nov 2012). We then narrowed down the search by manually deleting the proteins which were not UROS sequences and sequences belonging to phyla other than Chordata were deleted. A repeat phylogenetic analysis of 39 taxa was performed using PhyML and TreeDyn software to confirm that UROS is a highly conserved protein with approximately 85% conserved sequences in almost all chordate taxons emphasizing its importance in heme synthesis.

Key words: Uroporphyrinogen III synthase, Protein sequences, Phylogeny, Sequence alignment

#### Background:

Congenital erythropoietic porphyria (CEP) or Gunther's disease is a rare inherited autosomal recessive trait and is caused due to the deficiency of uroporphyrinogen III synthase (UROS), the fourth enzyme in heme biosynthesis. Only about 150 cases of CEP have been reported to date [1-4]. CEP symptoms are heterogeneous, ranging from severe hemolytic anemia in utero to milder, later onset forms, with skin lesions due to cutaneous photosensitivity in adult life [4]. The deficiency of functional UROS causes buildup of porphyrins to toxic levels in red blood cells. The excess porphyrins can then accumulate in the skin causing oversensitivity to sunlight [4, 5]. UROS is the central point for the synthesis of heme, a complex organic molecule, is

a cyclic tetrapyrrole that contains a centrally chelated Fe and functions in the transport oxygen. The human UROS gene located on chromosome 10 (10q26.2; Cytogenetic Location: 10q25.2-q26.3; EC 4.2.1.75; 1371 bp mRNA; Entrez Pubmed ref NP\_000366.1, NM\_000375.2) and expressed as a 265 amino acid UROS enzyme catalyzes the cyclization of the linear tetrapyrrole, hydroxymethylbilane, to the macrocyclic uroporphyrinogen III which is eventually converted to heme. Most organisms have the ability to synthesize their own heme molecules by this largely conserved metabolic pathway **[4, 6, 7]**. UROS enzyme is localized in the cytosol and plays a critical part in production of heme which is an essential prosthetic group in many cellular reactions in prokaryotes and eukaryotes.

The presence of this pathway across Bacteria, Archaea, and Eukarya suggests that heme performs a central function in the evolution of life. The acquisition by eukaryotes of the heme is one of the most important events in cellular evolution and any interference to its synthesis can thus have dire consequences to the survival of these organisms. Because of the inherent role of UROS in the synthesis of heme, and considering its role in many other inherent metabolic pathways in the cell, we aimed to elucidate if variance in this gene exists in various species during evolution by using a phylogenetic analysis of published protein sequences of these genes.



**Figure 1:** Putative sequence of UROS in super families; UROS (HemD) catalyzes the production of uroporphyrinogen-III, the fourth step in the biosynthesis of heme. This ubiquitous enzyme is present in eukaryotes, bacteria and archaea. Cd Length: 239 Bit Score: 171.72 E-value: 5.15e-52; http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?RID=B6X0K5J901R&mode=all



**Figure 2:** First pass phylogenetic tree constructed by multiple alignment using BLAST pair wise alignments: Results presented using Taxonomic name [163 hits]

#### Methodology:

### Data Set, Sequence Alignment and Construction of Phylogenetic Tree

The GenBank database **[8]** was queried to retrieve all available protein sequences of the UROS protein. These sequences were retrieved and saved in FASTA sequence format. These sequences were then aligned using Clustal W **[9]** algorithm using default parameters. The initial first-pass phylogenetic tree was constructed using Neighbour Joining **[10]** method (maximum sequence difference of 0.85) using Domain

Enhanced Lookup Time Accelerated Basic Local Alignment Search Tool [DELTA BLAST] pairwise alignments between a query and the database sequences searched **[11]**. The evolutionary distance between two retrieved sequences modeled as expected fraction of amino acid substitutions per site given the fraction of mismatched amino acids in the aligned region was taken by the software using Grishin computation **[12]**. Using the results from DELTA BLAST, we created a firstpass phylogenetic tree after which we used a purpose-built computational phylogenetic method using Phylogeny.fr

software. Sequences were aligned with MUSCLE (v3.7) **[13, 14]** configured for highest accuracy (MUSCLE with default settings). After alignment, ambiguous regions (i.e. containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) using the following parameters: minimum length of a block after gap cleaning: 10; no gap positions were allowed in the final alignment; all segments with contiguous nonconserved positions bigger than 4 were rejected and 85% minimum number of sequences for a flank position **[15-17]**. The phylogenetic tree was reconstructed using the maximum

likelihood method implemented in the PhyML program (v3.0 aLRT) **[13, 18].** The Jones-Taylor-Thornton (JTT) substitution model was selected assuming an estimated proportion of invariant sites (of 0.197) and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma=1.199). Reliability for internal branch was assessed using the aLRT test (SH-Like). The graphical representation and edition of the phylogenetic tree were performed with TreeDyn (v198.3) **[19].** 

Danio reri	MEVLLLKePrEses-eaDPVTkRLascGhtATLTPaLSFrFwSL	Macaca_mul	PSFSEKLSHPEdYGGLIFTSPRAVEAaqLCLEKINKTEVWERSLKEKWNAKSVYVVGnAT	Equus caba	ESITVYQKIPHPGIQANLNSYYSQQGVPASITFFSPSGLthSLKHIQELSGDSIDDIKFA
Mahi fur:	Malifilly Trades a Constitution of the Constitution of the Constitution	Monodelphi	OSESEKLS: PEAYOGLIFTSPRAVEAVELCI. KKennil of WEKSI. KKKWI AKNOVUWCART	Saimiri bo	ESTUVIORI DEDCTOMUNSVYSOOCUDI STREESDOCT + VSI KUTORI SCHOTENTA
vakirugu_r	Revision	Concerning and the	DODODRI ONDE VOOL BERODENDEN IN 17 DE	Saimiri DU	Perial Karala Contraction Cont
Oreochromi	GsDPYIkELesrGhkqTLlPVLSFkFvSL	sarcopniiu	rsrsarishPlatoshVrtSrKavkalqisLERagqlEVWKKSLKEKWrlKpaTVVGqAT	Callithrix	ESITVYQTIPHPGIQqNLNSYYSQQGVPASITFFSPSGLtYSLKHIQELSGDnIDQIKFA
Inonlanama	MANTI I MADARA CADDUT LEI ARCHART THU OF LEVEL	Anolis car	gS1fEKLScPEgY1GLIFTSPRAVEAIkLCLadksKkEdWknyLrEKWNtKSIYVVGKAT	Homo sabie	ESTTWYOTVARDCTOWNLNSYYSOOGVPASITEESPECTYSLKHTORLSCONTDOLKEA
HIODIODOMA	"HIVEDERFLOYES GODFIIREDASUGURALELEVESTREVED	Mana i annui	- OL F-WI OUDE - WORL - EMODDAL BATH: OL H- H-WHE-W OL MWHI W- HUDBIOWAR	nomo_pupre	DETTINATION TO AN ADDRESS OF ADDRES
Dicentrarc	MdVLLLKePrdees-rpDFYIkELasYGhkATLIPVLSFkFvSL	таепторудт	GOTIGUROULACIGORALISLUMIBULUCIVEVSUVP942000000000000000000000000000000000000	Pan_panisc	PETLAIÖLABHAEIÖÄNPAELISÖÖRALVELLELEAGPELISPVHIÖPPSADUIDÖTVLV
Esox luciu	MhVLLLKePrEnns-GsDPYTkELashGhk&TLTPVLSEkEvSL	Meleagris	edlfEKLSHPErYGGLVFTSPRALEAIkiCLreksKnEaWsKSLKqqWNtKptYVVGKAT	Pongo abel	ESITVYQTIPHPGIQqNLNSYYSQQGVPASITFFSPSGLtYSLKHIQELSGnnIDQIKFA
BOON_IGGIG	HEVEBBECCIERGGO GODITIEBBECHONINE DITVEDENE VOB	Galling gal	egifFKI SHDReVCCI#FTSDRAIFATKiCI rengKnReVeKSIKgeWNAKhtVD/CKAT	Nomagone 1	DOLUMENT DEDCTOWER NOVACION OF THE DODOCE FY OF KUTOPE CONVENTION AND
Saimo_saia	MnvLLLKepreggs-GpDPTIKELasrGnKATLIPVLSFTFVSL	Garran Jar	editminout.perconstructurenteriorensummakenterioren	nomascus_1	PETILI XIIII DI KANDALI DI KANDALI DI
Xenopus Si	MKVLLLKDPKtkDl-asDPYvkELssYGLgATLIPVLSFkFvSL			Papio_anub	ESVTVYQTIPHPGIQQNLNSYTSQQGVPASITFFSPSGLtYSLKHIQELSGDSIDQIKFA
Yozanua la		Danio reri	ASLVvd1GLrp1GEDsGtAdaLA1111gREnpdigPLfFPCGsiKREVLPtaLrnshIP1	Macaca mul	ESVTVYOTIPHPGIQqNLNSYYSOOGVPASITFFSPSGLtYSLKHIOELSGDnIDDIKFA
venobna_ra	HEAPPERDENTERDE_SRIPPORTETEAPPEREAPP	Taki fuan a	hat the 1 GED + GE	Manaalalahi	POTORIVOLT DUD-TO-I-I-OVEC-OCEDIOT-EPODOCI-b-al-mutopEcoDeEDOTVEN
Oryctolagu	MKVLLLKDaKdDaC-GQDFYvRELGLhGLEATLIFVLaFEFLSL	.akiiudu_r	NarALalerchiernnerwarzeitekronniberrechertekraredsevet	Ronogerphi	DOLLAIONTALEELGAATIOLOSGOCEMOLALIOPOGRANGDSALOBIOGINGINERK
Cricetulus	MKULLKDaKdDDs-GODPYTGRLGLYGLSATLTPULSEEFLSL	Oreochromi	ASLVqn1GLNp1GEDtGtAEvLsr1IieREatNipPLfFPCGsiKREVLPtaLrenGvP1	Sarcophilu	ES1TVYQKVqHPrmQdsLkSYfStQGVPAgvTFFSPSGLtYSLpHIQELSGaSfDrIKFA
Debbooking	MULTING MODE CODDUCTIONS TO AN ADDR OF	Inonloroma	heaVrn ] GLNn ] GEDtGt &dvLee ] TieREdtNin DLfEDCGeiKEEVLDteLronCvD]	Anolis car	EilTVYOTtoHPdlOesLrnYfSDOGVPASITFFSPEGyKYcLKoImkLSGDlItDIKFA
Rattus_nor	HEVLLLKDBRODDS-GQDFTIGELTCSLEATLIPVLSFEEMSL	татортороны	Addythiobhytobbookdyborritekbackipibritebookdubirecalegidati		POINTROM - TRANSPORTED AND PRODOCINE - LANDAL ACCULATION
Mus muscul	MEVLLLKDaKEDDs-G1DFYIgELrLcGLEATLIFVLSFEFmSL	Dicentrarc	AaLVrs1GLNp1GEDtGtAqvLssvIidREdtNiLPLfFPCGsirREVLPKaLrDnGvP1	Taeniopygi	RELIAIOINGHEGIÕGEREELERÄÄREREELERERAAFICPOHIÖEPRONIUUÕIVER
Carria nara		Esox luciu	AaLVenlGLiplGEDtGtAdvLsrlTieREdtSiLPLfFPCGsiKREVLPtaLKknGvPl	Meleagris	ESITVYQTaqHehlQqsLsSYfSQQGIPASIvFFSPSGvKfcLqHIQkLSGDSanQIKFA
Cavia_porc		0-1	A - LUL- 1 GL - 1 GED+G+1 d- ( 1 T 1 - D D-H-U-1 DL GEDOG - : MDEUK D+-L G-D1	Callue dal	FS TWO TO ANAL SOLD STUDIES TO FREE STORES IN A HIGH SCORE AND THE STORES
Heteroceph	MKVLLLKDaKEDDC-GQDFYIKELGLYGLEATLIPVLSFEFqSL	saimo_saia	Agraeulerabiergaresillieurginithritheresiuwrarteueati	Garras_gar	PPI 111% adment#damanto#Kattwartstantorondur#knoopramerut#
Otolemur a	WEVILLIKDaKEDDG-GODBYTERIGLYGLEATSTEVISEEELSL	Xenopus Si	ASLVeeIGLssEGEqsGNAEKLAECICSkqlSySaPiLFPCGsLKREVLPKkLqeKkvPl		
Des kount		Xenonus la	ASLVeeTGLssEGEgsGNARKLARgTCSkg1SvSaPiLEPCGsLKKEVLPKkLgeKnyP1	Danio reri	sIGrTTAdAlqArGLtVSCcAEkPTakhLAiaItq-ALrdl
Bos_taurus	NKULLLKDPKEDDC-GQDPTVKELGLTGLEATLIPVLSFEFLSL	nemopub_ra	Abb/ccibbobobobgob/Abb/bolgibybalibrecobladb/bridbgchiti	Takifum r	ATCODDeeAmiliEct overeal Providing at a K-AI O
Loxodonta	MKVLLLKDPKEDDC-GODPYIRELGLYGLEATLIPVLSFEFLSL	oryctolagu	ASTVSTIGLTADGEHSGEAKKLAETICSKE-SSpLPLLKPCGTLKREVLPCVLKedGVPM	Taxilugu_L	RIGHT TECHNARISTIC BERNART TIGHTART TIC REC
Mustala nu		Cricetulus	AcLVNKIGLDaEGaHsGNAEKLAEVICSRp-SSALPLLFPCGTiKgdtLPKMLKDKGIPM	Oreochromi	AIGPTTqhA1AAEGLCVSCTAEKPTaEhLA1aIaK-tLQ
Muscera_pu	HEOLDEROBICEDIC GUDFTIKEDSETGERTETFVESFEFESE	Dattur nor	A DELIMENT OF DE ROSSERIES REFERENCE DE LE RECORDE VIER LE DER LE	Anoplopoma	AIGPTTodAmtAEGLcVSCaAEkPTaEhLAAaIaK-ALO
Ailuropoda	MKVLLLKDaKEDDC-GQDPYIRELGLYGEEATLIPVLSFEFLSL	Naccus IIVI	ADDARGEDADGAGGGGGGGGGGGGGGGGGGGGGGGGGGGG	Dicentrarc	ATCOPPERAINT AND AVECTATED TO THE AND AVERAL AND AVERAL
Canis lupu	MKVLLLKDPKKDDC-GODPYTRRLGLYGFRATLTPVLSERFLSL	Mus muscul	ASLVNKIGLDaEGagsGNAEKLAEYICSkp-SSeLPLLFPCGTiKgdtLPHMLKDKGIPM	DICENCIALC	HIGHTIQUHICADODCVDCARDAFTABILDARATAIC HDQ
Falia	MULTITADE CODATA CONTRACT AND A CONT	Cavia porc	ASLVIKIGLDFEGENCGNARKLAEVICSBE-SDALPLLEPCGTLKREVLPDMLKDKGIPM	ESOX_luciu	AIGPUTACIAMI CAGLSVSCSAERPTAENLATGI aK-ALQ
1 CI 1 S	MUALTERDIGTOR - COLLINERT LERATEI LATSERETZ		A DESCRIPTION OF A DESC	Salmo sala	ATGPTTAdAmtARGLsVSCsAdkPTaRhLAtGTaK-ALOrgtdetdbkpt
Felis catu	MKVLLLKDPKEDDC-GQDPYIRELGLYGLEATLIPVLSFEFLSL	neceroceph	ASLVMRIGLNIDGANCGNAKKLAKTICSRE-SpvLPLLFPCGTLKREaLPhMLKDKGIPM	Vananija Ci	ATCHER AND ACCONTRACT CONTRACT OF A STORE A STORE
Sug scrofe		Otolemur q	ASLVsKIGLDTkGENCGNAEKLAEYICSRE-SSALPLLFPCGnLKREILPKMLcaeGIPM	Achopus_51	RIGPTTRERMRENGIPVSUTRINPTPQQLAIGIKQ-AQDQ
PUD_POINTS	HIV DEDUCTIONALOOPTITADODIODERIELPYLOEBELOL	Dog taurun	ACIMETCI PUTCEFCONARXI ARVICODE POALDI I ROCCEI VERTI DANI VENCI DI	Xenopus la	AIGPTTAeAmAeEGiPVSCTAqnPTFQdLAiGIKqigmQ
Equus_caba	MKVLLLKDFKEDDr-GQDFYIRELGLYGfEATLIFVLSFEFLSL	BOS_CAULUS	ASLVHIIGSHI DOBUGWARALABIICSKE-DSALPLEFOUTLAKEILFALADKGIFI	Orvetolami	ATGUTTARALAROGL PVSCTAESUTDOALARG] BK-ALD] DGGC
Saimiri bo	MKVLLLKDaKEDDC-GODPYIRELGLYGIEATLIPVLSEEELSL	Loxodonta	ASLVrKIGLDTEGEROGNAEKLAEVICSRE-SpALPLLFPCGTLKgEILPHMLKDqGIPM		Top B D3 331 of Wroots Borness 1 2 mentating rappings
Callistanda	MARTIN PARTY CONTRACTOR CONTRACTO	Mustela nu	ASLVNKTGLOPEGANSONARKLARVTOSER-SpaleLEDCODyKGRTLDKNI KDyGTDM	uricetulus	ALGFSTUKAMAARGLFVSCIAESPTFQALAAGIRK-vLQPntCC
Callienrix	MRULLLKUAKEDDC-GQDFTIKELGLIGLEATLIPVLSEEELSL	nabeera_pa	Non weight bourgerander i ober bereget wight in installister in the	Rattus nor	AIGPsTURAmAAkGLPVSCTAESPTPQALAAGIRK-ALkPnsCC
Homo sapie	MKVLLLKDaKEDDC-GODFYIRELGLYGLEATLIPVLSFEFLSL	Alluropoda	ASLVNKIGLDTEGANCGNAKKLAETICSRq-SphLPLLFPCGTwKgEILPKMLKDTGIPM	Mus muscul	ATGP®TTRAMAAKGLPVSCTAESPTPOALAAGTRn-vLkPnhCC
Dan nanigo		Canis lupu	ASLVNKIGLDaEGaNOGNAEKLAEVICSRE-SpALPLLFPCGTvKgEILPgMLKDrGIPM	find the out	A CONTRACTOR OF A CONTRACT OF
Pail_pailise	HEOLDERDERDER - GUPTIKEDSLIGERIEI FVESTERES	Telie .	- OF TREAT OF DEPARTMENT A PART OF D. D. MINT I PROCESS FIT DIAL ADVICT DA	Cavia_porc	AIGPTTARALAREGLIVSCTAESPTPQALREGIRR-ALQPHChC
Pongo_abel	MEVLLLKDaKEDDC-GQDPYIRELeLYGLEATLIPVLSFEFLSL	Tells	SOFANDIGPDI NOPOCOAMBUTCOUP_OMATLECOLAUZEITELUATVOLEU	Heteroceph	AIGPTTAhAvAAEGL1VSCTAESPTPOALAvaIR1-vLOP1hhC
Nomascus	NKYLLLKDØKEDDC-GODEVIRELGLYGLEATLIEVLSEEELSL	Felis catu	SSLVNKIGLDTEGESOGNAEKLAEYICSRE-SPALPLLFPCGTvKgEILPKMLKDKGIPM	Otolomum or	TORREDUCTOR DUCCES FOR DOLLAR CTRY ALONG dCG
notified out _1		Sue errofa	A SUVARTICUM NORMOUNARY LARY LOSS F-SSALDLI ROCONLERED DRM. KOKOY DM	ororemur_g	HIGHTINGHINGGENVSCINESPIPONEKCGIGG-HEGPUCC
Papio_anub	mfpcsvtapykdcqaiMKvLLLKDaKEDDC-GQDPTIKELGLYGLEATLIPVLSFEFLSL	Sus_sciola	ADDVSICIOND/DSERVARENDARITICONE COMPRESSIONER DEFICIENTIALE AVEN	Bos taurus	AIGPTTAGALAAQGLPVSCTAESPTPrALAAGIRt-tLQPHaC
Macaca mul	MEVLLLKDaKEDDC-GODFYIRELGLYGLEATLIFVLSFEFLSL	Equus_caba	ASLVNKIGLDTEGENCGNAEKLAEVICSRE-SSALPLLFPCGTLKGEILPKMLKDKGIPM	Loxodonta	ALGPPTARALWAGGLPVSCTAESPTPGALAAGTRK-ALGAHSCC
Manadalahi	WEULLIKOTKADAR-CODDVERSICALCERATITUS SEELSI	Saimiri bo	ASVVsKIGLOPEGELCGNARKLARYICSBE-SSALPLLEPCGnLKBELLPnMLKDKGIaM	Mustal	TOD-BIDII-SOOL WOO-TROPDIIISAN TA STORE
nonouerphir	unannuerumda añoriadoren nitataren en	Gallikhnin	A DI MANT CI DEDCERACINA EVI A EVI CODE DONI DI L'EDCO L'UDELL'UDVOL AVI	Mustera_pu	AIGPSTARALLAQGLIVSUMALSPTPIALARGIQV-ALQUPS
Sarcophilu	MKVLLLKDaKdDDs-G1DFY1qELGscGfEATLIFVLaFEFLSL	Californix	ASPASEIGEDI NOFICAMARUPARIICORP-SCHEREPERICAURURAIAM	Ailuropoda	AIGPsTARALtAQGLPVSCmAESPTPlALAkGlql-ALQgprC
Anolis car	MEVILLENDPERsen-GnDPYTeRLGshGLEATLIPVLaFEFiSL	Homo sapie	ASLVsKIGLDTEGEtCGNAEKLAEYICSRE-SSALPLLFPCGnLKREILPKaLKDKGIaM	Canis lupu	ALGPSTARALABOGL PVSCTAESPTPOrLAFG] B] -ALGAD TC
Taonionuci		Pan nanisc	ASLVSKIGLOTEGETCGNARKLARTICSER-SSALPLIEPCGDLERELPKALKDEGTAM	Falia.	ATCD-BADALAACCI WOODA FOODD ALAACIDI ALCOPTO
raenropygr	HEOLEDBO FRANDS - SPDF11KBD3618 DBR101F9 D3F0E L3D	Provenskal	A DI UL-WI CI DEDCE-COM EVI A ENTONE EL COM EL EDCC-1 METTI DV-1 MENOT-M	relis	HIGHSIHWHPHAGPEADCIHPOLLEDHPHKHTULHPAGAbàC
Meleagris	MKVLLLKDPKdkes-GpDPYvkELGLcGfEATLIPVLSFEFvSL	Pongo_aber	ASPASULOPDI NOVICOMARUPARI ICSUR-SSHELPPLARUPUSI TANDI SU	Felis catu	AIGPSTARALAAQGL PVSCTAESPTPpALAkG
Gallus dal	MKVLLLKDPKdkes-GoDPYvkELGLYGFEATLIPVLSFEFvSL	Nomascus l	ASLVsKIGLDTaGEtOGNAEKLAEYICSRE-SSArPLLFPCGnLKREILPKaLKDKGIAM	Sus scrofa	ALGETTARALAAGGLEVSCTAESETEYALAAGTEt - ALOPHYC
	·····	Danio anub	A CUMENT CUPPECE CONARXI A FUTO OD F., CONTRUCT DU LEDOCO LERET U DESLEVENCES M	bub_beroru	AIGHT INGREAGENT CONNECTIVE AND AND A AND
		rapio_anus	ASSYSTEDDI DOBUCOVABRUABI I COND-SCADE DEFECCIUDADE DE RADADIOI AN	Equus_caba	AIGPTTARALAAQGLPVSCTAESPTPSALAAGIRt-ALQPHSC
Danio reri	nelSErLfgPErfGGLIFTSPRAVEAVktsLElgklrhoWD-avKdKWNAKSVYVVGKAT	Macaca_mul	ASLVsKIGLDTEGEtCGNAEKLAEVICSRE-SSALPLLFPCGnLKREILPKaLKDKGIAM	Saimiri bo	AIGPTTARALAAOGL PISCTAESPTFOALALGIRK-ALOPHqCC
Takifum r	at 1 ody: for DELLOCI TETER DAVEAUENCI E - EL E-ON Konnologia KONNUCCIAT	Monodelphi	ASLVrK]GLATEGERCENGEKLARTICSGE-putLPLLEPCGFLKgEaLPnMLopKGvPM	Callithrig	ATCOPUADAL ADOLI DURCUA ESPEDAL ALCERIZ-AL ODI ACC-
Taviinăn_r	2 CIONNELLERUNGARIELDELMARMANNELRERUNGAERMANNEALAADUMI	Assesshils	ADISTRUCTURE DEDISOR STRUCTURE AND ADDRESS AND ADDRESS	Calification	RIGHTIRIGERAGSEFVSCIRESFIF@REASITECREGEIGCC
Oreochromi	ntlSdKLfqFEkYGGLIFTSPRAVEAVrmCLEaDerrEeWDKaaKdKWNtKSIYVVGKAT	Sarcophilu	ASPARAAPPPER AND WORKPARTICSKREDHEPPPLAAPPARAPPERADELI	Homo_sapie	AIGPTTARALAAQGLPVSCTAESPTPQALAtGIRK-ALQPHgCC
Anoplonoma	nt1SdKLfgPRkbGGLIFTSPRAVEAVkmCLRaDerrRei/nsSvHdKiWsKSTYV/GKAT	Anolis car	AvLVaelGLipQGENsGNAEKLsgYICSREspNSLPLLFPCGsLKREtLPtiLreKGIal	Pan panisc	AIGPaTARALAAOGLPVSCTAESPTPOALAtGIRK-ALOPHqCC
Discobusion		Teenionuri	A CI MAATCI UD/CEPACNARKI ARVICEPL-DUCALI PD/CALERRYI DEDI MAKCID]	Dongo ohol	ATCHERADAL ANOCH INFORMATION AND AND AND AND AND AND AND AND AND AN
Dicentrarc	NEISGRVEGPEKRGGLVFTSPRAVERVERVERUEBBEEFFEWRSSVEGEWRSSVEGEWRSSTIVVGERT	таспторудт	ADAACTORADAACTORELICOVV-buoographicographicatic	Pongo_abei	RIGFTTRRRBAUGLFVDCIREFTFURLAUGIRG-REUFRGCU
Esox luciu	nv1SdKLfrPdkhGGLIFTSPRAVEAVkmCLEehggmEeWnsnmrdKWNAKSVYVVGKAT	Meleagri <i>s</i>	ASLVeeIGLspQGEksGNAEKLAEVICSkk-phSssLLFPCGalKREVLPtaLKeKGIP1	Nomascus l	AIGPTTARALAAQGLPVSCTAESPTPQALALGIRK-ALQPHgCC
Salmo golo	nt]SdKLfaDRahooLTETSDRAVEAVmCLEdDargRowndiEdawaaverman	Gallus dal	ASLVeeIGLcpOGEksGNAEKLAEVICSREophSsPLLEPCGalKREVLPtalKoKGTP1	Papio anub	AIGPTTARALAAOGLPVSCTAESPTPOALAEGTRK-ALOPHoCC
varus sara	In Strong Prophysics For A strange of the strong of the st	"Aer	-r *	Magnen	ATCODERADAL AACT DUDCERADODEDOLIA -CTDU ALCODE
λenopus_Si	cnrtckushPEsYaGLIFTSPRAVEAVELCLqtpdhkEaWndrLcaEWNsEpVYVVGEAT			nacaca_mui	HIGHLIHUHDAAQGDFVGCIHDSFIFQADACGINK-ADQFNGCC
Xenopus la	dh FfdKLSHPEsYaGLI FTSPRAVEAVkLCLgKnahkEawkdhLcaKWNsKoVYVVGKAT	Danio reri	EtlTVYQTseHPdlQkNisdYftQQGVPAcvvFFSPSGvsfcldllkDLSGcqLEQIKFA	Monodelphi	AIGPTTARALtsaGLPliCTAEkPcPhdLAAGlqK-ALQP
America 1 mm	DAPARTY	Taki furu r	Em] TWYOT SHIDD SHILL SY FLOOCEDD ST SEESDSCHEET STREET SCHOLES THE	Sarconhilu	ATGPTTARALARAGI PVSCTAREPOPORLAGGI oK-ALOPsonses
otAccorgân	POPORING CLERGIGGELETOPRAVLARQUCLERSSIRAEWWERSDEERWNVRSVTVVGGAT	.antrugu_1	DELTA STACHTHEORETOTICOGOCHOTATATATATATATATATATATATATATATATATATATA	laslis	TOLDER'S TI'RE AND
Cricetulus	PS1SEKLSHPEqfGGLIFTSPRAVEAVELSLEQDNKTE@WEKSLKnKWNAKSVYVVGnAT	ureochrom	stitvigraehrolekNLknYfteQGIPASIaFFSPSGvKfcLevvrrLSaeqLtQIKFA	Anoiis_car	RIGHTTRERLEUBGISVSCIRISPTFQQLUUGIRR-SLQIKQC
Battus por	PS1SEKLSHPR#fGGLIFTSPRAVEAVELCLEKTNETE=WEKSLK++MUAKCUVUUC= htt	Anoplopoma	EtlTVYOTaeHPdlekNLNnYfteOGtPASTaFFSPSGvKfcLevvrrLSGehmaOTKFA	Taeniopygi	AIGPsTAeALrgaGvsVSCTAgSFTaQdLAAaIht-ALQPHpCsgpQ
Mus mus 3	a constant by could for the bay About the brought and brought and any and a second sec	Digentrone	Et MERCHAAUDALANIE knyftangt DE CHERCHONNE CO-T FOTER	Neleacris	ALGPPTARAmeAaGi PVSCTAESPTPOdLAAGT/K-ALTD/moCiller
MUS_MUSCUl	PSISBRUSHPEGIGGLIFTSPRAVEAVELCLEKDNKTE&WEKSLKdrWNAKSVYVVG8AT	DICENCIAIC	scille Margareter and with the Anterna States and Anterna An	0-11	A COMPANY AND A
Cavia porc	cSFSEKLSHPEdYGGLIFTSPRAVEAVELCLEKDsKaEdwEKpLKEKWstKSVYVVGKAT	Esox_luciu	Et1TVYQTVqHPdlekNLknhfteQGVPASvaFFSPSGvKfcletvrrLSGerLnQIKFA	Gallus_gal	AIGFTTAEAMEAAG1FVSCTAESFTFQGLtAGIqK-ALrPqnCilSE
Heterocomb	an EFENI CHDE AVOOL I EPODENTELUTI OT ENTRIPED ANTWH I NEUKINA PODENTER	Salmo sala	dtlmWYOTagHPdlekNLknhfteOGVPASTaFFSPSGwKfclenwrrLSCogLbOTKFP		
recercebu	OUL FRICOULDATE OF MAN PRATE POLITICAL AND TRANSPORTED FROM THE AND TRANSPORTATION OF THE ADDRESS OF THE ADDRES	Vananus Of	The THE ROOM - DID - A CALL - REAL AND A CALL AND A CAL		
Otolemur_g	PSFSEKLSHPEnYGGLIFTSPRAVEAVELCLEKDNeaEWEKSLrEKWNAKpVYVVGKAT	Aenopus_S1	stitvigraphpavQdsLmnYitkeGvPASvvFFSPSGvKYtLKilkDLpnDqLnQIKFg		
Bos taurus	PS]SEKLSHPEGYGGLIFTSPRAVEAVETCLGEDTKAEUGEKSLEENWAKSVYVVGAAT	Xenopus la	EtITVYQTqPHPaIQvsLsdYftkeGVPASIvFFSPSGLKYcllflkDLpsDaLmOIKFA		
Longdor to	DORORAL CHIRACOLI DECODALIZATION OF DADAME AND	Oractolarm	RETUVE AND A CONTRACTOR OF A CONTRACT OF A C		
POXOGOULS	rorosnuonregiooliftorkavervelulekusktesweksikekwnakpvtvvGsAT	oryccorayu	TOTI		
Mustela pu	PSFSEKLSHPEqYGGLIFTSPRAVEAVELCLEKUNKTEWWKhSLrEKWNAKSVYVVGnAT	Cricetulus	ESmhVYQTVPHPGIQgNLeSfYenQGIPASITFFSPSGLKYSLeyIQvLSGDSfDQIKFv		
hilumonodo	DOPORTI CUDRAVICI TEPODDAUFAUFI OLEVINAVER-OPA OLAVINAVORNINA-AM	Rattus por	RSmbWYOTTPHPGTOgsLeSYYenOGTPASITERSPSGLKYSLewTOaLSC-SEPOTKRy		
ATTATOPOUG	I OF OBIGORIEGY I OF DEFINITION AND A DEFINITION OF A DEFINITI	Man man mil	no-havonumento and have dealer annual and a series and a series		
Canis_lupu	PSFSEKLSHPEGYGGLIFTSPRAVEAVELCLEKDNKTEWNASLKERWNAKUVYVVGNAT	MUR_WIRCHT	KSEDVIQIVPHPGIQGSLKSITEDQGIPASITEFSPSGLKISLEVIQALSGSSEDQIKF1		
Felis	SSESEKLSHPEGYGGLIFTSPRAVEAVELCLEKUNKTEWERSLKEKWNAKSVYVVC~DT	Cavia porc	ESITVYQTIPHPeIQgNLhSYYSkQGVPAcvTFFSPSGLRYSLKHIQELSGDSIDOIKFA		
Talia anto	OPOPUL CHIDE ALCOLI DEGODINICI DE DUCAMENTAL D'ALCOLINITATION CONTRACTORI	Heterocenh	REVENUED FOR MINE SAAR STORE OF COMPARISON OF A STORE CONTRACTOR OF A STORE S		
reiis_catu	SOFORNUORPEGIGGLIFTSPKAVAAVALULEKUNKTEWKRSLKEKWNAKSVIWGSRT	recercebu	and the second s		
Sus scrofa	PSFSEKLSHPEgYGGLIFTSPRAVEAVkLCLEKaNKTEaWktSLKEKWNsKSVYVVGnAT	Otolemur_g	KSITVYQTAAHFGIQGNLNSYYSKQGIPAqIAFFSPSGitYtLKHIQELSGDSIDQIKFA		
Emus caba	DSESSEN, SHDRAYCCLIFTSDRAVEAVELCIFKUMTENNIKSIEKNNAKOUVUUC-AT	Bos taurus	ESVTVYOTIPHPGIOGNLtSYYtOOGVPASTTFFSPSGLthslKHTORLSGDSTMOTKFW		
squus_cava	I ST SERVER AND A ST	Longdon to	EDITERYALING AND		
saimiri bo	PSFSEKLSHPEDYGGLIFTSPRAVEAaELCLEKDNKTEWERSLKEKWNAKSVYVVGRAT	BTHODOXGU	POTIAINVILHLOINGAPPOLIOVÄAANVILHELEOPPELÄINÄHINPPOLUUUUKH		
Callithrix	tSFSEKLSHPEdYGGLIFTSPRAVEAaELCLdgDNKTgWERSLKEKWNAKSVYVVGdAT	Mustela pu	ESITVYQkIPHPGIQvNLNSYYSkQGVPASITFFSPSGLmYSLKHIQELSGDSIDTIKFA		
Warne annia	DOPORTE CHIDE AVOID TEMODELLET A LE CHE CONTRACTOR DE LA CONTRACTION DE LA CONTRACTIONE DE LA CONTRACTICA DE LA CONTRA	Miluroroda	RETWYORT DEDGT OWNIN NEVY SHOCK DI STREESDOGLEVSI KETORI SCR. TROTKEN		
HOUD_Sapie	PSFSERLSHPEQIGGLIFTSPKAVLABELULEQINKTEWWERSLKEKWNAKSVIVVGRAT	ALLUIODOUG	POLIA I MALIULO I MARIA I POLICIO POLICO POLICIO POLICIO POLICO POLICIO POLICIO POLICIO POLICIO POLICIO POLICIO POLICI		
Pan panisc	PSFSEKLSHPEdYGGLIFTSPRAVEAaELCLdornKTEWERSLKERWNAKSVYVVGrAT	Canis_lupu	ESITVYQKVPHPGIQWNLNSYYakQGIPASITFFSPSGLaYSLKHIQELSGDnIDrIKFA		
Bongn shall	DODODKI CRDEMACCI I EMODDYALI * LI GI EN-SIAMERYSPO GI REARTY NORTHER * 20	Felis	RST I VYOKVPH PGTOWNLDSYYSOOGT PASTTERSPSGLEDS1KHTORLSGDDTGTTKEA		
rondo mer	FOFOBILION FOR OTHER FOF ANY DAREDUCTION AND A RODAL WARDALLAND AND A LANDARD AND AND A LANDARD AND AND AND AND AND AND AND AND AND AN	Talia astu	not involvented to avage of the second state o		
Nomascus_1	PSFSEKLSHPEdYGGLIFTSPRAVEAaELCLEKnNKTEWERSLKEKWNAKSVYVVGnAT	reils_cacu	FOITAIGEALHARTGAMPUALLAGGELAROILERABAGPCUATERIÖEPZGDUIGLIEEK		
Panio anub	SSESSKI, SHPROYOGLI FTSPRAVRA act. CLRKnNKTRWERSLKRKWARSVYVVGA AT	Sus scrofa	ESITVYQTIPHPGIQqNLNSYYSQQGIPASITFFSPSGLsYSLKHIOELSGDSIDOIKFv		
		_	······································		

Figure 3: MUSCLE (3.7) multiple sequence alignment [CLUSTAL format] of short-listed sequences [39 taxons] Similar residues are colored as the most conserved one (according to BLOSUM62).

#### **Results:**

From the NCBI GenBank database, 163 sequences of UROS covering the putative conserved HemD domain were used to construct a first-pass phylogenetic tree. The tree however had

some of the repetitive and unrelated sequences which were deleted. The short-listed sequences belonged to Porifera, Arthropoda, Echinodermata, Tunicates, Cephalochordates, Bacteria, Cnidaria, Fishes, Amphibia, Reptiles, Aves and

Mammalia. The accession information for these sequences is available in Table 1 (see supplementary material). Analysis of the sequences revealed that there is a high degree of sequence similarity of UROS enzyme in many of the selected sequences used for the phylogeny reconstruction. Putative conserved domains were observed in many taxa at the HemD region (Figure 1) [20]. The actual alignment was detected with superfamily member pfam02602 (E-value: 3.02e-41). BLAST produced 163 hits (Figure 2); these sequences were screened manually and only those related to the sequence in guestion (HPSE-1) from different taxa were retained for further analyses. This produced a total of 39 sequences from 31 taxa. Multiple sequence alignment results of these short-listed sequences is presented in (Figure 3). Using the PhyML program a tree was constructed for these sequences, the results of which are presented in (Figure 4).



**Figure 4:** Phylogenetic tree using short-listed sequences [cladogram] built using the PhyML software [39 taxons].

#### Discussion:

The evolutionary relationship of UROS enzyme in various species, taxa and phyla was evaluated using computational phylogenetics to identify similar genes in the short-listed organisms. A mix of algorithms and programs were used to construct a phylogenetic tree **[21, 22]**; the neighbor-joining method was used to calculate genetic distance and ClustalW to create trees based on multiple sequences. The JTT matrix method was used to generate mutation data matrices from ISSN 0973-2063 (online) 0973-8894 (print) Bioinformation 8(25): 1265-1270 (2012)

protein sequences and the set sequences were clustered at the 85% identity level. In addition, the sequences were aligned, and the observed exchanges amino acids were tallied [23]. Following this, the final phylogenetic tree was constructed [18]. We have used MUSCLE (multiple sequence comparison by logexpectation) method to achieve the highest scores. Close to four benchmarks showed that MUSCLE achieved the highest ranking of any method available at the time of publication. Using Gblocks, poorly aligned positions and divergent regions of sequences could be eliminated – these positions may not be homologous or may have been saturated by multiple substitutions [24, 25]. In addition, this program helped to reduce the necessity of manually editing multiple alignments with very fast processing. The PhyML software was used since it was shown to be at least as accurate and slightly faster as other existing phylogeny programs using simulated data. The DELTA BLAST algorithm which uses a heuristic method to identify homologous sequences helped to produce high scoring sequence alignment between the query and database sequences. We have used BLAST as a first pass sequence alignment tool to narrow down the target and most relevant sequences.

The phylogenetic tree indicated that UROS protein is conserved and plays an important role in organismal evolution (Figures 3 & 4). It is interesting to note that the conserved regions as shown in Homo sapiens are similar to those found in some other organisms that have this conserved gene. Presence of UROS in the major organisms indicates that it is crucial for the development of the physiology of cells. Its high conservation at certain domains indicates that its function is preserved. In conclusion, the evolutionary relationship of UROS gene was established based on the sequence alignment, conserved sequences and phylogenetic trees. The results of the published data on protein sequences of the above genes showed that the sequences are highly conserved especially at certain domains. Human sequences consistently clustered with their mammal orthologs within these genes clearly indicate the importance of these genes in evolution [26]. The phylogenetic reconstruction of the metabolic pathways of many organisms is one of the major goals of genomics. Reconstructions made through comparative genomics, and results from experiments on model systems help in understanding the biochemical diversity of life. Thus, analysis of phyletic patterns re-emphasizes the importance of certain metabolic enzymes in evolution.

#### Acknowledgement:

Authors would like to extend their appreciation to the Research Centre, College of Applied Medical Sciences and Deanship of Scientific Research at King Saud University for funding this research.

#### References:

- [1] Gross U, J Inherit Metab Dis. 2000 23: 641 [PMID: 11117426]
- [2] Anderson KE et al. The Molecular and Metabolic Bases of Inherited Disease. CS Scriver, AL
- [3] Beaudet, WS & Slyand D, Br JHaematol. 2002 117: 779 [PMID: 12060112]
- [4] Bickers DR & Frank J, Dermatology in general medicine. 2003 pp 1435.
- [5] Kauppinen R, *Porphyrias Lancet.* 2005 **365**: 241 [PMID: 15652607]
- [6] http://ghr.nlm.nih.gov/gene/UROS

- [7] van Dooren GG & Kennedy AT, Antioxidants Redox signal. DOI: 10.1089/ars.2012.4539 [PMID: 22320355]
- [8] Benson DA et al. Nucleic Acids Res. 2011 39: D32 [PMID: 21071399]
- [9] Thomson JD et al. Nucleic Acids Res. 1994 22: 4673 [PMID: 7984417]
- [10] Saitou N & Nei M, *Mol Biol Evol.* 1987 4: 406 [PMID: 3447015]
- [11] Boratyn GM et al. Biol Direct. 2012 7: 12 [PMID: 22510480]
- [12] Grishin NV, J Mol Evol. 1995 41: 67 [PMID: 18345592]
- [13] Dereeper A et al. Nucleic Acids Res. 2008 36: W465 [PMID: 18424797]
- [14] Edgar RC, Nucleic Acids Res. 2004 32: 1792 [PMID: 15034147]
- [15] Castresana J, Mol Biol Evol. 2000 17: 540 [PMID: 10742046]
- [16] Guindon S & Gascuel O, Syst Biol. 2003 52: 696 [PMID: 14530136]
- [17] Anisimova M & Gascuel O, Syst Biol. 2006 55: 539 [PMID: 16785212]

- [18] Dereeper A et al. BMC Evol Biol. 2010 10: 8 [PMID: 20067610]
- [19] Chevenet F et al. BMC Bioinformatics. 2006 7: 439 [PMID: 17032440]
- [20] Marchler-Bauer A et al. Nucleic Acids Res. 2011 39: 225 [PMID: 21109532]
- [21] Wu CH et al. Nucleic Acids Res. 2006 34: D187 [PMID: 16381842]
- [22] Strait DS & Grine FE, J Hum Evol. 2004 47: 399 [PMID: 15566946]
- [23] Jones DT et al. Comput Appl Biosci. 1992 8: 275 [PMID: 1633570]
- [24] Talavera, G & Castresana J, Syst Biol. 2007 56: 564 [PMID: 17654362]
- [25] Castresana J, Mol Biol Evol. 2000 17: 540 [PMID: 10742046]
- [26] http://www.ncbi.nlm.nih.gov/books/NBK20266/ [PMID: 21089240]

#### Edited by P Kangueane

Citation: Shaik et al. Bioinformation 8(25): 1265-1270 (2012)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

### Supplementary material:

. .

Table 1: Accession details of short-listed uroporphyrinogen-III synthase (UROS) sequences

Accession no:
VD 002021204 1
AP_002021304.1 ND_001252711.1
NF_001255711.1
XP_003904442.1
XP_002/56/46.1
NP_000366.1
XP_003277827.1
XP_003922155.1
XP_003816334.1
XP_003419757.1
NP_001178298.1
NP_001012068.1
XP_003208282.1
XP_003794158.1
XP_002721286.1
XP_003479664.1
NP_001231296.1
XP_003513922.1
NP_033505.1
XP_423886.2
XP_854192.1
XP_003994563.1
XP_002916164.1
EHB18080.1
XP_001490208.1
XP_002193805.1
NP_001107324.1
NP_001087355.1
XP_001362178.1
ACO13405.1
AES09336.1
ACI66186.1
ACQ58390.1
XP 003218642.1
XP_003441824.1
XP 003755055.1
XP 003961197.1
AAO49476.1
CBN81562.1