

## 研究报告

### Research Report



# 血红蛋白表征的 *In silico* 分析引领糖基化敏感性研究迈向新领域

Reza , Ahmad , Fazlolah , Seyed 

1 动物遗传学系, 农学部, Bu-Ali Sina University, 哈马丹, 伊朗

2 动物遗传学系, 伊朗农业生物技术研究所, 雷什特, 伊朗

3 动物遗传学系, 农学部, University of Guilan, 雷什特, 伊朗

 通讯作者, Talebi\_re@yahoo.com;  作者

计算分子生物学, 2015 年, 第 4 卷, 第 7 篇 doi: 10.5376/cmb.cn.2015.04.0007

收稿日期: 2015 年 09 月 07 日 接受日期: 2015 年 09 月 07 日 发表日期: 2015 年 09 月 07 日



© 2015 BioPublisher 生命科学中文期刊出版平台

本文首次以英文发表在 *Computational Molecular Biology* 2015, Vol. 5, No. 1, 1-13 上。现依据版权所有人授权的许可协议, 采用 [Creative Commons Attribution License](#) 协议对其进行授权, 用中文再次发表与传播。只要对原作有恰当的引用, 版权所有人允许并同意第三方无条件的使用与传播。如果读者对中文含义理解有歧义, 请以英文原文为准。

**摘要** 应用 *In silico* 分析不同物种的血红蛋白表征特别是糖基化的抗性。多重序列比对结果显示骆驼、人类、牛和马中都具有 L-K-V-K-E-G-E-A-L-R-P-T-F-F-D-L-S-A-V-K-H-G-K-V-H-D-L-S-L-H-K-L-V-D-P-N-F-L-L-L-A-F-T-P-A-K-V-L-Y 的 2 个血红蛋白亚基序列。血红蛋白序列在进化过程中和不同物种之间是非常保守的。双峰驼家庭 HBA 存在于相同的群体(91% bootstrapping), 然而 HBB 在同一个群体(95% bootstrapping)。骆驼血红蛋白似乎比其他物种更加基础。基于等电点, 骆驼 HBB 在电泳中具有高于 HBA 的迁移率。骆驼中血红蛋白的带电荷和亲水性氨基酸比人类多。葡萄糖与蛋白质或 HbA1c 的非酶结合被揭示, 人类中的两个亚基血红蛋白都具有对 N-糖基化的抗性, 但骆驼 HBB 仅对 O-糖基化具有抗性。此外, HBB 在骆驼中具有关键作用, 例如对糖基化和之后糖尿病的抗性。

**关键词** *In silico*; 多重比对; 血红蛋白; 糖基化; HbA1c



## *In Silico* investigation for hemoglobin characterizations leads to new aspects in susceptibility to glycosylation

Reza Talebi , Ahmad Ahmadi , Fazlolah Afraz , Seyed Zeyaeedin Mirhoseini 

1. Department of Animal Genetics, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

2. Department of Animal Genetics, Agricultural Biotechnology Research Institute of Iran, Rasht, Iran

3. Department of Animal Genetics, Faculty of Agriculture, University of Guilan, Rasht, Iran

 Corresponding author, Talebi\_re@yahoo.com;  Authors

**Abstract** *In silico* analysis were performed between some species for hemoglobin characterizations specially resistivity to glycosylation. Multiple alignments of both hemoglobin subunits were revealed L-K-V-K-E-G-E-A-L-R-P-T-F-F-D-L-S-A-V-K-H-G-K-V-H-D-L-S-L-H-K-L-V-D-P-N-F-L-L-L-A-F-T-P-A-K-V-L-Y conserved in camel, human, cow and horse entirely. Hemoglobin sequences are well conserved in evolution and between species. *Camelus* families HBA were in same cluster (91% bootstrapping). However for HBB they were in same cluster (95% bootstrapping). Camel hemoglobin appeared to be more basic than either that of other species. Based on isoelectric points, camel HBB has high mobility than HBA in electrophoresis. The charged and hydrophilic amino acids of hemoglobin in camel were more than human. The non-enzymatic binding of glucose to the protein or HbA1c is revealed both subunits hemoglobin in human is resistance to N-glycosylation but camel HBB is only resistant to O-glycosylation. HBB is less susceptible to glycosylation than HBA. Also HBB has critical roles in camel for instance resistivity to glycosylation and diabetes subsequently.

**Abbreviations** HB, Hemoglobin; HBA,  $\alpha$ -chain of hemoglobin; HBB,  $\beta$ -chain of hemoglobin; GHB, glycosylated hemoglobin; AA, Amino Acid; NA, nucleic acid; Vs., versus; p I, Isoelectric Point.

**Keywords** *In silico*; Multiple alignments; Hemoglobin; Glycosylation; HbA1c

球蛋白是可以通过身体组织传输氧的独特分子。血红蛋白和肌红蛋白是分别在血液和肌肉中负责的两类球蛋白。它们存在于所有脊椎动物和一些种类的无脊椎动物中。脊椎动物中的血红蛋白结构为球蛋白折叠与一个血红素基团结合的典型结构,该血红素基团是一个可逆地将氧结合到其中心亚铁原子化合物的基团。成人血红蛋白是红细胞中每条链的分子量为 15 kDa 的 $[\alpha(2):\beta(2)]$ 四聚体血红蛋白,其中它负责结合肺中的氧并将结合的氧运输到整个身体以用于有氧代谢途径(Dickerson and Geiss, 1983; Paoli and Nagai, 2004; Pairet and Jaenicke, 2010)。除了骆驼科以外的哺乳动物中的红细胞的形状是球形和凹形的,在骆驼科中是椭圆形,并且核较少(Goniakowska-Witalinska and Witalinski, 1976; Azwai et al., 2007)。骆驼科的红细胞对渗透性溶血具有高度抗性,因此能够扩展高达其原始体积的 240%而在低渗溶液中不会破裂(Oyewale et al., 2011)。

糖基化血红蛋白(GHB)或美拉德反应是由葡萄糖和 N-末端缬氨酸之间的非酶相互作用和 HB 的  $\beta$  链的赖氨酸残基形成的,可能引起与 HbA1c-a 型血红蛋白相关的糖尿病(Rohlfing et al., 2002; Wautier and Schmidt, 2004; Soranzo, 2011)。由于 HbA1c 长期存在并且没有分解,几十年来已经使用 HbA1c 浓度来评估长期血糖控制对糖尿病患者的影响。10 种 GHB 的形成是不可逆的,其在红细胞中的水平取决于血糖浓度。GHB 的测定方法在 20 世纪 70 年代首次引入,提供长期血糖控制的指数,其已经被证明引起糖尿病治疗的改变,导致改善的代谢控制。HbA1c 程度直接与人类的血糖水平有关,也可能与其他动物的血糖水平有关(Nathan et al., 2008; Shahbazkia et al., 2010; Bazzi et al., 2013)。基于氨基酸和葡萄糖之间的相关性,真核生物中糖基化的过程可以分为五种类型: N-连接, O-连接, C-连接, P-连接和 G-连接,其中 N-和 O-连接最常见的糖基化形式(Chauhan et al., 2013)。

骆驼和牛中的血糖浓度比为 $(9.7\pm 2.8 \text{ mM})$ (高水平)和 $(5.7\pm 0.73 \text{ mM})$ (最低水平)。骆驼中高水平

的血糖可能是由于其强大的胰岛素抵抗能力引起的(Jirimutu et al., 2012; Bazzi et al., 2013)。骆驼和牛中糖基化血红蛋白(GHB)水平分别为  $3.4\pm 0.23\%$  和  $3.2\pm 0.11\%$ 。在较高的葡萄糖浓度下,骆驼 Hb 的低糖化表明某些因素在高葡萄糖浓度下保护 Hb 免于糖基化(Bazzi et al., 2013)。根据 Oyewale 等人(2011)的研究,单峰骆驼血红蛋白为研究沙漠极高温条件下的生存条件提供了一个有趣的案例,骆驼 Hb 也表现出比人或牛的正常血红蛋白更高的电泳迁移率(Oyewale et al., 2011)。

目前,相关学者已经对骆驼物种的红细胞和血红蛋白进行了大量的研究(Lin et al., 1976; Braunitzer et al., 1980; Farooq et al., 2011; Oyewale et al., 2011)。Oyewale 等人(2011)比较了在不同环境情况下单峰骆驼(*Camelus dromedarius*)和驴(*Equus asinus*)之间的红细胞脆性。他们提出除了血液储存时间之外,红细胞环境的温度和 pI 的变化可以在骆驼和驴红细胞的渗透行为中各自发挥重要作用(Oyewale et al., 2011)。Bazzi 等人(2013)发现骆驼中的血糖浓度 $(9.7\pm 2.8 \text{ mM})$ 较高并且糖基化血红蛋白含量较低 $(3.4\pm 0.23\%)$ ,但是牛血液样品在葡萄糖浓度 $(5.7\pm 0.73 \text{ mM})$ 或糖基化血红蛋白 $(3.2\pm 0.11\%)$ 中没有显著差异(Bazzi et al., 2013)。骆驼血液中高葡萄糖浓度中的低糖基化显示高葡萄糖浓度可以保护 Hb 不发生糖基化。糖基化 Hb 的水平不仅可以反映对血糖水平,红细胞寿命和红细胞膜渗透性的依赖性,而且还显示在食品体系方面(Ardia, 2006)。Borai 等人(2011)建议 HbA1c 可以作为胰岛素抗性的一种简单可靠的标志物,以用来标记对胰岛素相对敏感的正常成人的葡萄糖耐受性(Borai et al., 2011)。根据以往关于血红蛋白在糖尿病中的作用的研究,我们决定研究骆驼属单峰骆驼,双峰驼,野骆驼,牛,人类和马的血红蛋白的结构、化学特征和组成。尽管骆驼具有比人类更高的血糖水平,但是骆驼血液中的糖基化程度远小于我们人类。因此,HB 亚基的 DNA 和蛋白质序列在所提及的物种之间进行比较以研究糖基化的抗性和随后的糖尿病。

## 1 材料与amp;方法

### 1.1 数据收集

从生物数据库例如 Uniprot 和 KEGG 获得用于单峰骆驼、双峰驼、野骆驼、牛、人类和马的 AA 和 HB( $\alpha$  和  $\beta$  亚基)核酸(NA)序列(表 1)。

### 1.2 分析

序列已经通过 Lasergene 软件转化为氨基酸(\*.pro)。同时相关文件更改为 FASTA 格式(\*.fas), 这种形式常用于基因组数据分析。我们利用 MEGA 5.01 (Tamura et al., 2011) 和 CLC Genomics Workbench 7.5 软件使用 ClustalW 对每个物种中的蛋白质进行多重序列比对。进化史使用基于 Kimura 蛋白距离标准的 UPGMA 进行推断。UPGMA 假定不变的进化速率。

### 1.3 蛋白质组成

全蛋白结构类型使用 Alpha Deléage 和 Roux 对 Nishikawa 和 Ooi(1987)修改后的方法进行预测 (Deléage and Roux, 1987)。利用 Protean 软件 (DNASTAR Inc., Madison, WI. USA)测定蛋白质结构、滴定曲线、氨基酸组成和每种蛋白质的频率。

### 1.4 蛋白质结构预测

我们使用在 ExPASy 网站 (<http://swissmodel.expasy.org>) 中提供的 SWISS-MODEL 预测基于同源性建模的血红蛋白亚基的三级结构。该模型可使蛋白质建模可访问世界各地的所有生物化学家和分子生物学家。

### 1.5 N-和 O-糖基化位点的预测

在这项研究中, 我们利用一个新的网络服务器 GlycoEP([http://www.imtech.res.in/raghava/glycoep/su\\_bmit.htm](http://www.imtech.res.in/raghava/glycoep/su_bmit.htm))更准确地预测骆驼物种和人类中 HBA、HBB 或葡萄糖与蛋白质的非酶结合(如在 HbA1c 的情况下)的 N 链接、O 链接和 C 链接的糖基化位点。

### 1.6 比较建模

为了计算比较建模, 我们使用 BlastP 在所提到的物种中的一些标准, 例如同源性, 保守性, 一致性和 E 值, 并比较 NCBI HomoloGene 数据库以将它们分配到基因家族。最后, 进行 PSI-BLAST(位置特异性迭代 BLAST)以便在所选数据库中发现骆驼血红蛋白和各种蛋白质之间的同源基因(Altschul et al., 1997)。

## 2 结果与amp;讨论

### 2.1 血红蛋白亚基的多重序列比对

多重序列比对显示 Leu, Lys, Val, Lys, Glu, Gly, Glu, Ala, Leu, Arg, Pro, Thr, Phe, Phe, Asp, Leu, Ser, Ala, Val, Lys, His, Gly, Lys, Val, His, Asp, Leu, Ser, Leu, His, Lys, Leu, Val, Asp, Pro, Asn, Phe, Leu, Leu, Leu, Ala, Phe, Thr, Pro, Ala, Lys, Val, Leu 和 Tyr 在所有血红蛋白亚基中完全保守(图 1)。由于保守序列, 可以推测在所述种类的这些 AA 中发生最小突变, 或者一些 aa 因为通过突变消除了这些保守序列的血红蛋白。根据 Binder 等(2013)研究, 在保守序列的基础上, Leu 具有主要贡献, 同时, 亮氨酸对肠糖异生和胰岛生理学的积极效果可能有助于预防 II 型糖尿病(Binder et al., 2013)。同时, 在图 1 的基础上, 在所提及的物种中分别保留了大量的 AA 用于血红蛋白的  $\alpha$  和  $\beta$  亚基, 因此这表明在不同物种血红蛋白结构中发生最小的突变。因此, 我们可以得知, 血红蛋白的结构可以保护其免受诱变剂的威胁, 由于  $\alpha$  和  $\beta$  亚基在骆驼、牛、马和人类中高度保守, 因此骆驼或其他物种中血红蛋白的最独特的效力与  $\alpha$ HBA 和 HBB 的结构、组成和特征有关。如图 3 所示, 本研究中每个血红蛋白链仅具有一个结构域 (domain 1)。

具有高频共有序列的 AAs 属于 L-K-V-K-E-G-E-A-L-R-P-T-F-F-D-L-S-V-K-H-G-K-V-H-L-D-L-L-S-L-H-K-L-V-D-P-N-F-L-L-L-A-F-T-P-A-K-V-L-Y 这种最常见的模式。基于共有序列可以鉴定受突变影响较小的核苷酸。例如, 根据图 1, 在牛的血红蛋白的  $\alpha$  亚基部分中, A-A-G-A-W-V-S-G-A-A-A 的 AA(氨基酸)序列分别被一个双峰驼中的 S-K-T-T-F-I-G-A-K-G-D 取代。此外, 在骆驼和野骆驼的血红蛋白  $\alpha$  亚基部分, S-K-T-F-I-G-A-G 的氨基酸序列分别被人类中的 P-A-A-W-V-A-G-S 取代, 这也显示在了系统发育图的结果中(图 2)。在两个亚基的骆驼科之间没有发现氨基酸的取代, 但是对于其他物种显然发现了氨基酸的取代(图 1)。

因此, 它可能涉及血红蛋白序列在进化和物种之间是保守的。在 *Camelus* 家族(*Camelus bactrianus*, *Camelus dromedarius* 和 *Camelus ferus*)的  $\beta$  链序列的

表 1 血红蛋白的  $\alpha$  和  $\beta$  链通过访问生物学数据库获得, 蛋白名称、登录号和 GenBank 如下所示

Table 1  $\alpha$  and  $\beta$  chains of hemoglobin are obtained from accessing biological database. Protein names, Access number and GeneBank following is shown

| Protein names           | Access Number | GeneBank  |
|-------------------------|---------------|---|
| HBA_Camelus bactrianus  | P63105        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBA_Camelus dromedarius | P63106        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBA_Camelus ferus       | 102513641     | <a href="http://www.genome.jp/kegg/">http://www.genome.jp/kegg/</a>           |
| HBA_Homo sapiens        | P69905        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBA_Bos taurus          | P01966        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBA_Equus caballus      | P01958        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBB_Camelus bactrianus  | P68230        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBB_Camelus dromedarius | P68231        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBB_Camelus ferus       | S9WA55        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBB_Homo sapiens        | P68871        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBB_Bos taurus          | D4QBB3        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |
| HBB_Equus caballus      | F6RDD3        | <a href="http://www.uniprot.org/uniprot/">http://www.uniprot.org/uniprot/</a> |

Abbreviations: HBA:  $\alpha$ -chain of hemoglobin; HBB:  $\beta$ -chain of hemoglobin

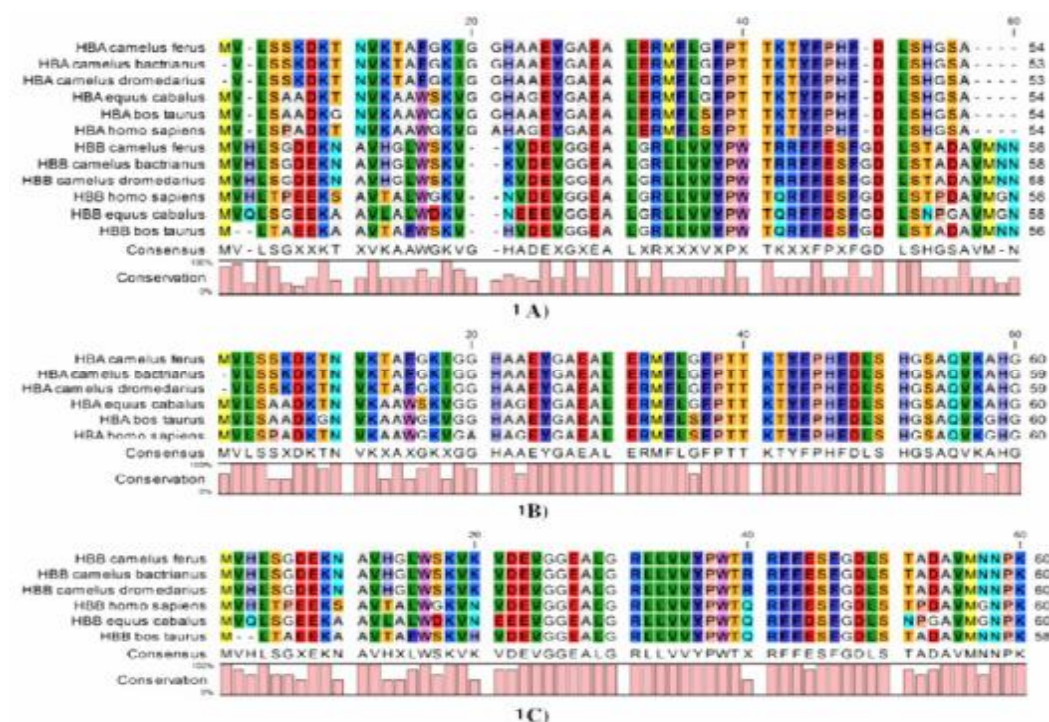


图 1 HB 亚基的多重序列比对. 1A) 血红蛋白的  $\alpha$  和  $\beta$  链在人, 家养单峰驼和双峰驼, 野生双峰驼, 牛和马之间的序列比对. 1B) 所列物种之间血红蛋白 A 链序列比对. 1C) 所列物种之间血红蛋白 B 链序列比对. 相同的缺失和保守序列分别用虚线(-) 和粉红色柱表示

Figure 1 Multiple sequence alignment of HB subunits. 1A)  $\alpha$  and  $\beta$  chains of hemoglobin were aligned between human, domestic one and two-humped camel, wild two humped camel, cow and horse. 1B) A chain of hemoglobin was aligned between mentioned species. 1C) B chain of hemoglobin was aligned between mentioned species. Identical gaps and conserved sequences are indicated by dashes (-) and pink columns respectively



一些部分中, H, D, N, H, G, S, K, V 和 D 的 AA 被 Q, E, A, L, A, D, N, E 和 E 取代, 或者骆驼科中 S, G, D, N, H, G, S, K 和 R 的 AA 被 T, P, E, S, T, A, G, N 和 Q 取代。因此, 在突变期间, 物种中氨基酸的分化发生在同一多肽中, 这是与生物中血红蛋白效率有关的主要因素。因此, 突变在使动物适应不同环境情况方面发挥了重要或独特的作用。Lin 等(1976)年研究显示, 在成年骆驼血红蛋白和成年人血红蛋白之间, 在  $\alpha$  链的 N-末端 20 个氨基酸残基中发现 6 个氨基酸差异, 残基: 4,5,12,14,17,19; 在  $\beta$  链中, 在位置 4,5,6,9,12,13,16 和 19 处发现了 8 个氨基酸取代。在  $\alpha$ 5Ala $\rightarrow$ Lys 和  $\beta$ 19Asn $\rightarrow$ Lys 处的取代增加了骆驼的净正电荷血红蛋白两倍, 而其他取代没有显著差异 (Lin et al., 1976)。Pieragostini 等(2010)在反刍动物品种中提出了血红蛋白变体与血液学模式的适应性意义(Pieragostini et al., 2010)。Braunitzer 等(1980)表示, 野骆驼 HBA 中 5 种 AAs 和 HBB 中 2 中 AAs 分别与羊驼属发生交换。同时, 他们推导出野骆驼和单峰骆驼的 AAs 序列是相同的(Braunitzer et al., 1980)。然而, 我们假设通过蛋白质或基因工程将 AA 替代为多肽结构, 将是如何增加血红蛋白对糖基化和随后糖尿病的电阻率的适当方法。

## 2.2 系统发育分析

如图 2 圆图所示, 双峰驼和野骆驼中的 HBA 在 100% bootstrapping 时属于同一亚群, 与单峰驼  $\alpha$  亚基相邻。Camelus 家族的 HBB 与此类似(图 2)。由于双峰驼和单峰驼  $\alpha$ -链与人类的系统发育距离大于牛和马(图 2)。系统发育树揭示了双峰驼、单峰驼和野骆驼的  $\beta$  链(HBB)在同一亚群中 95% bootstrapping, 55% bootstrapping 接近马属的 HBB。基于图 2 中 HBA 在马和牛在相同的亚群中具有 39% bootstrapping, 但是远离人类的 HBA。系统发育分析证明, 每个血红蛋白亚基在物种中是保守的, 并且它取决于蛋白质的特定功能和性能。

## 2.3 全蛋白组成

HBA 和 HBB 的全蛋白组成在骆驼科中是相同的, 但是完全不同于其他物种(表 2)。基于表 2, 骆驼属中血红蛋白的等电点(pI)似乎超过了人类, 牛和马(表 2)。事实上, 骆驼中的血红蛋白亚基的 pI 是碱性的, 而人类并不是。骆驼中的血红蛋白亚基还包含比人类更多的带正电和带负电的氨基酸。pII

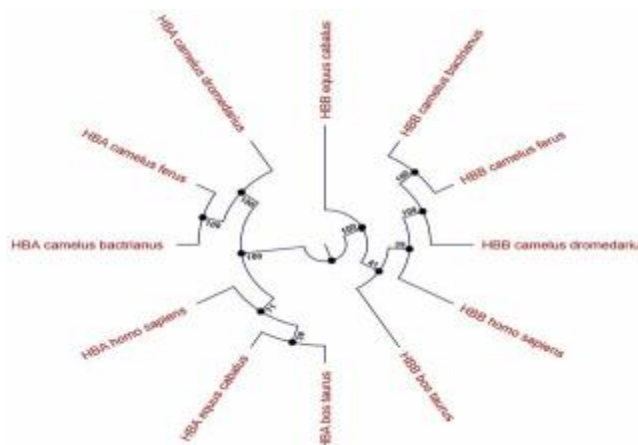


图 2 UPGMA 的圆形剖面图. 通过使用基于 JTT 矩阵模型的 UPGMA 方法推断进化历史. 1000 个重复预测的 bootstrap consensus tree 表示类群分析进化历史. 启发式搜索的初始树自动获得如下. 该分析涉及 12 个氨基酸序列

Figure 2 Circular cladogram by UPGMA. The evolutionary history was inferred by using the UPGMA method based on the JTT matrix-based model. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. Initial tree(s) for the heuristic search were obtained automatically as follows. The analysis involved 12 amino acid sequences

在骆驼 HBB 中的基本效应大于 HBA(表 2), 因此我们认为它与骆驼 HBB 的高迁移率相比 HBA 在电泳和随后的糖基化中有关。然而, 我们认为骆驼家族的 HBB 中的等电点比人和牛多, 但稍微小于马(表 2)。换句话说, 在骆驼和人类中的  $\beta$  链(HBB)具有碱性和中性特征, 或者它可以影响糖基化。它在电泳期间影响蛋白质的迁移率, 因此骆驼和人 HBB 在电泳期间分别具有高和低的迁移率。除了 Bazzi 等(2013)年宣布骆驼血红蛋白具有比人或牛的正常血红蛋白更高的电泳迁移率(Bazzi et al., 2013)。基于表 2, 双峰驼和骆驼骆驼的血红蛋白(HBA)的  $\alpha$  链的 pI 值略高于骆驼属(9.33)。两个血红蛋白亚基的骆驼 pI 比人类多(表 2)。根据研究结果和表 2, 我们注意到人类和骆驼之间的 HBA 的整个蛋白质组成的显著差异不存在。因此, 我们设想其他情况受到骆驼血红蛋白的最小糖基化的影响, 例如, 与人类相比, 骆驼红细胞可能更限制葡萄糖转运。Oyewale 等人(2011)发现骆驼红细胞比驴红细胞更耐溶血(或更少渗透脆性)(Oyewale et al.,

2011)。在这项研究中,如图 2 所示,骆驼家族血红蛋白与人类相比具有更高比例的亲水性氨基酸(极性)(例如人和骆驼的 HBA 分别为 20.42 和 21.28),并且这使得骆驼血红蛋白具有亲水性在骆驼家族中 HBA 和 HBB 的带电荷 A(氨基酸)比人类多(表 2)。根据 Bazzi 等人(2013)年的研究,也可能的是骆驼血的氧化还原状态或红细胞的内部环境作为血液 pH 的变化可能有助于 HbA1c 水平的降低。1976 年被描述骆驼羚血红蛋白比人类 HBA 在纸张电泳上拥有更多的正电荷,并且表明仅由单一成分组成(Lin et al., 1976)。一些研究人员宣称,与牛、绵羊、山羊、猪、小鼠、人和其他动物相比,骆驼的红细胞对溶血更具抗性(Al-Qarawi et al., 2004; Oyewale et al., 2011)。由于来自骆驼的血红蛋白在极高的温度下提供了适应沙漠中的生活条件的有趣的情况(Balasubramanian et al., 2009)。在渗透不可除去的骆驼红细胞中的水的比例几乎是人类的三倍,这似乎导致骆驼红细胞对渗透变化的极好的抗性(Bognera et al., 2005)。正如食物对鸟类中血红蛋白(GHB)的糖基化的影响被证明一样(Miksik and Hodny, 1992)。

## 2.4 同源性的测量

基于表 3 中显示的信息,双峰驼和单峰驼的 HBA 和 HBB 的相似性为 100%,并且具有 100%覆盖率(表 3)。然而,在骆驼家族中 HB(HBB)的  $\beta$  链的相似性是 100%和 100%查询覆盖。Hb 的  $\alpha$  和  $\beta$  链均为 84%的相似性,在骆驼科和人类之间具有 100%的覆盖率。骆驼家族 Hb 的  $\alpha$  链与牛和马的相似度分别为 85%和 86%。骆驼家族与牛和马的  $\beta$  链的相似度分别为 83%和 82%(表 3)。因此,血红蛋白的  $\alpha$  和  $\beta$  链在所列物种中是保守的。所以,亚铁转移氧取决于血红蛋白的基本功能。

## 2.5 蛋白质结构建模

基于使用 ExPASy 网络服务器的同源性建模的  $\alpha$  和  $\beta$  链(HBA 和 HBB)的三级结构的预测如图 3 所示。HBA 的三级结构在骆驼科家族中是非常相似的,并且 N-连接的糖基化是 Asn,这可能是由每个突变发生在 Hb 的  $\alpha$  链中引起的。因此,通过我们的研究结果,,例如氧运输等血红蛋白的最关键的功能不是由其他类型的亚基如  $\gamma$ 、 $\delta$  或  $\zeta$  完成。事实上  $\alpha$  和  $\beta$  链可能控制血红蛋白的最关键功能。如图 3 所示, HBB 和 HBA 的三级结构在骆驼科家族中

都是相同的,但它不同于人类(图 3)。如图 4 所示,所述物种中的 HBA 和 HBB 的配体是血红素。血红素的作用与氧有关,两个血红蛋白亚单位都受氧运输的影响。我们在本文中已经看到,基于同源性建模,在骆驼科家族中 HBB 的三级结构很可能是人类的  $\alpha$ -链(图 3)。尽管基于图 3 的事实,骆驼家族中的 HBB 对糖基化不具有易感性位置。如图 3 所示,在家养和野生骆驼的 Asn131 和 Asn132 对于 HBA 中的 N-糖基化是易感氨基酸,然而人 HBA 中的所有 Asn 都具有对 N-糖基化的抗性。根据人类 HBA 和骆驼 HBB 之间的相似结构,我们得出结论,人类中的 HBA 具有对 N-糖基化的抗性。而如图 3 所示, HBB 具有比 HBA 更少的 N-糖基化位置;因此, HBB 比 HBA 更多地参与关键运动。如前所述, HBB 的序列在所有考虑的物种中是高度保守的,也许这解释了血红蛋白在氧运输中的关键作用。

## 2.6 糖基化位点

葡萄糖与蛋白质的非酶结合(如在 HbA1c 的情况下)的糖基化预测是基于二元谱(BPP)模式进行。 $\alpha$  和  $\beta$  链的糖基化位点的结果显示在表 4 中。在该研究中,在所提及的物种中没有发现任何 HBA 和 HBB 的 C 连接的糖基化。C 连接的糖基化是相对罕见的,并且在其中发现聚糖连接到共有序列 W-X-X-W 或 W-X-X-C 或 W-X-X-F(其中 X 是任何 AAs)中的第一 Trp 残基的碳上(Krieg et al., 1998)。基于 HBA 中的 N-糖基化的表 4,我们发现了双峰驼和单峰驼中 Asn131 的潜在的糖基化位点,而野骆驼的 Asn132,但是人类对于两个血红蛋白亚基没有显示潜在的 N-糖基化位点(表 4)。通过向氮原子添加葡萄糖来识别 N-连接的糖基化, Asn 的 N4 是特异性识别共有序列 Asn-X-Ser/Thr, 其中 X 是除脯氨酸之外的氨基酸(Gavel and von Heijne, 1990)。在这项研究中,我们发现大部分共有序列比如 HBA 潜在 O-糖基化对于 O-糖基化是敏感的,例如如家骆驼的 Ser3-Ser4-Thr8 和人的 Ser4-Thr42(表 4)。基于表 4,  $\alpha$  血红蛋白可能易受 O 糖基化,特别是野骆驼。如表 3 中所示的 HBB 的 O-糖基化,我们发现在人类中仅对 Thr5 具有潜在的糖基化(表 4)。没有位置对骆驼 HBB 中的 O-糖基化敏感。O-糖基化通过向游离的含羟基的 AAs 残基加入葡萄糖而形成,所述残基包括 Ser、Thr 以及在一定程度上羟基脯氨酸和羟赖氨酸(Christlet et al., 2001)。因此,我们得出结论,

Table 2 Composition of hemoglobin. Predicted structural class of the whole protein: Alpha Deleage & Roux Modification of Nishikawa & Ooi 1987

表 2 血红蛋白的组成. 全蛋白质预测的结构分类: Alpha Deleage & Roux Modification of Nishikawa & Ooi 1987

|         | HBA    |        |        |        |        |        | HBB    |        |        |        |        |        |
|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|         | C. b   | C. d   | C. f   | H. s   | B. t   | E. c   | C. b   | C. d   | C. f   | H.s    | B.t    | E.c    |
| MW(kDa) | 15.172 | 15.172 | 15.303 | 15.257 | 15.184 | 15.245 | 16.245 | 16.245 | 16.364 | 15.998 | 15.979 | 16.139 |
| Length  | 141aa  | 141aa  | 142aa  | 142aa  | 142aa  | 142aa  | 147aa  | 147aa  | 148aa  | 147aa  | 145aa  | 147aa  |
| pI      | 9.39   | 9.39   | 9.33   | 9.12   | 8.87   | 9.12   | 8.71   | 8.71   | 9.05   | 7.11   | 6.64   | 6.77   |
| A.I     | 87.234 | 87.234 | 86.62  | 90.775 | 93.521 | 93.451 | 94.762 | 94.762 | 94.517 | 93.469 | 92.759 | 95.442 |
| Ala (A) | 0.121  | 0.121  | 0.120  | 0.148  | 0.141  | 0.113  | 0.088  | 0.088  | 0.088  | 0.102  | 0.110  | 0.095  |
| Cys (C) | 0.007  | 0.007  | 0.007  | 0.007  | 0.000  | 0.007  | 0.007  | 0.007  | 0.007  | 0.014  | 0.007  | 0.007  |
| Asp (D) | 0.071  | 0.071  | 0.070  | 0.056  | 0.056  | 0.063  | 0.061  | 0.061  | 0.061  | 0.048  | 0.062  | 0.048  |
| Glu (E) | 0.021  | 0.021  | 0.021  | 0.028  | 0.035  | 0.021  | 0.048  | 0.048  | 0.047  | 0.054  | 0.055  | 0.068  |
| Phe (F) | 0.057  | 0.057  | 0.056  | 0.049  | 0.049  | 0.049  | 0.048  | 0.048  | 0.047  | 0.054  | 0.069  | 0.054  |
| Gly (G) | 0.064  | 0.064  | 0.063  | 0.049  | 0.063  | 0.070  | 0.088  | 0.088  | 0.088  | 0.088  | 0.069  | 0.095  |
| His (H) | 0.078  | 0.078  | 0.077  | 0.070  | 0.070  | 0.070  | 0.061  | 0.061  | 0.061  | 0.061  | 0.048  | 0.061  |
| Ile (I) | 0.007  | 0.007  | 0.007  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  | 0.000  |
| Lys (K) | 0.092  | 0.092  | 0.092  | 0.077  | 0.077  | 0.077  | 0.075  | 0.075  | 0.074  | 0.075  | 0.076  | 0.075  |
| Leu (L) | 0.128  | 0.128  | 0.127  | 0.127  | 0.141  | 0.148  | 0.129  | 0.129  | 0.128  | 0.122  | 0.117  | 0.129  |
| Met (M) | 0.007  | 0.007  | 0.014  | 0.021  | 0.014  | 0.014  | 0.014  | 0.014  | 0.014  | 0.014  | 0.021  | 0.014  |
| Asn (N) | 0.021  | 0.021  | 0.021  | 0.028  | 0.021  | 0.028  | 0.061  | 0.061  | 0.061  | 0.041  | 0.055  | 0.048  |
| Pro (P) | 0.043  | 0.043  | 0.042  | 0.049  | 0.042  | 0.042  | 0.027  | 0.027  | 0.027  | 0.048  | 0.028  | 0.034  |
| Glu (Q) | 0.007  | 0.007  | 0.007  | 0.007  | 0.007  | 0.007  | 0.014  | 0.014  | 0.014  | 0.020  | 0.021  | 0.027  |
| Arg (R) | 0.021  | 0.021  | 0.021  | 0.021  | 0.021  | 0.021  | 0.041  | 0.041  | 0.041  | 0.020  | 0.028  | 0.027  |
| Ser (S) | 0.085  | 0.085  | 0.085  | 0.077  | 0.092  | 0.092  | 0.048  | 0.048  | 0.047  | 0.034  | 0.041  | 0.041  |
| Thr (T) | 0.071  | 0.071  | 0.070  | 0.063  | 0.056  | 0.063  | 0.027  | 0.027  | 0.027  | 0.048  | 0.041  | 0.020  |
| Val (V) | 0.078  | 0.078  | 0.077  | 0.092  | 0.085  | 0.085  | 0.122  | 0.122  | 0.122  | 0.122  | 0.124  | 0.122  |
| Trp (W) | 0.000  | 0.000  | 0.000  | 0.007  | 0.007  | 0.007  | 0.014  | 0.014  | 0.014  | 0.014  | 0.014  | 0.014  |
| Tyr (Y) | 0.021  | 0.021  | 0.021  | 0.021  | 0.021  | 0.021  | 0.027  | 0.027  | 0.027  | 0.020  | 0.014  | 0.020  |

Abbreviation: aa: amino acids; MW: molecular weight; p I : isoelectric point; AI: aliphatic index; HBA:  $\alpha$ -chain of hemoglobin; HBB:  $\beta$ -chain of hemoglobin; C. d: Camelus dromedarius; C. b: Camelus bactrainus; C. d: Camelus ferrus; B. t: Bos taurus; E. c: Equus cabalus; H. s: Homo sapiens.

HBA 中的 O-连接的糖基化相比于 N-, C-连接的糖基化发生得更多, 而 HBB 特别是在骆驼中耐 N, O-糖基化, 尽管  $\alpha$  血红蛋白最易于 O-连接的糖基化, 这也许取决于酒精因素在骆驼和人类氨基酸。

### 2.7 PSI-BLAST 分析(位置特异性迭代 BLAST)

家养骆驼(HBA 和 HBB)的  $\alpha$  和  $\beta$  链血红蛋白的 PSI-BLAST(位置特异性迭代 BLAST)的结果如图 4 所示。我们发现骆驼 HB 与其他物种如蝙蝠、大熊猫、普通猴等之间有很多 HomoloGene。当我们发现家养骆驼的血红蛋白与其他哺乳动物如黑狐或蝙蝠有很多 HomoloGene 不同时是非常令人兴奋的。

### 3 结论

本研究的结果使我们得出结论, 骆驼 HBB 似乎不比人类更易受 N-糖基化。基于等电点, 骆驼 HBB 和 HBA 具有高迁移率, 因此影响糖基化。我们知道血红蛋白的  $\alpha$  和  $\beta$  亚基具有如生物过程、分子功能和参与疾病的很多功能, 但最重要的是, 参与从肺到外周组织的氧运输。尽管如此, 血红素是所有物种中 HBA 和 HBB 的常见配体; 它可以通过亚铁与氧连接。以及, O-连接糖基化比 N-, C-连接糖基化容易发生; 然而骆驼 HBB 对 O 连接的糖基化具有抗性。虽然, 血红蛋白的  $\alpha$  链更易于 O-糖基化。这两个血红蛋白亚基对 N-糖基化均具有抗性。最终, 这些因素, 连同其他翻译后修饰, 可能负责

表 3 所列物种的血红蛋白  $\alpha$  链 (HBA) 和  $\beta$  链 (HBB) 的基因组同源性. *E-value* 表示物种同源性标准, query coverage 表示物种间比较

Table 3 Genome homology of  $\alpha$ -chain (HBA) and  $\beta$ -chain (HBB) of hemoglobin between mentioned species. The criterion of identity as homology with *E-value* and query coverage is shown for each comparison between species

| Hemoglobin chain Sequence   | Ident | <i>E-value</i> | Query coverage |
|---|-------|----------------|----------------|
| <i>Camelus bactrianus</i> Vs. <i>dromedarius</i> (HBA)  | 100%  | 7e-103         | 100%           |
| <i>Camelus bactrianus</i> and <i>dromedarius</i> Vs. <i>ferus</i> (HBA)                         | 100%  | 6e-102         | 98%            |
| <i>Camelus bactrianus</i> and <i>dromedarius</i> Vs. <i>Bos taurus</i> (HBA)                    | 85%   | 3e-87          | 100%           |
| <i>Camelus bactrianus</i> and <i>dromedarius</i> Vs. <i>Equus caballus</i> (HBA)                | 86%   | 2e-88          | 100%           |
| <i>Camelus bactrianus</i> and <i>dromedarius</i> Vs. <i>Homo sapiens</i> (HBA)                  | 84%   | 4e-88          | 100%           |
| <i>Camelus bactrianus</i> Vs. <i>dromedarius</i> (HBB)  | 100%  | 5e-110         | 100%           |
| <i>Camelus bactrianus</i> and <i>dromedarius</i> Vs. <i>ferus</i> (HBB)                         | 100%  | 5e-110         | 100%           |
| <i>Camelus bactrianus</i> , <i>dromedarius</i> and <i>ferus</i> Vs. <i>Bos taurus</i> (HBB)     | 83%   | 9e-82          | 97%            |
| <i>Camelus bactrianus</i> , <i>dromedarius</i> and <i>ferus</i> Vs. <i>Equus caballus</i> (HBB) | 82%   | 4e-82          | 100%           |
| <i>Camelus bactrianus</i> , <i>dromedarius</i> and <i>ferus</i> Vs. <i>Homo sapiens</i> (HBB)   | 84%   | 1e-92          | 100%           |
| <i>Camelus ferus</i> (HBA) Vs. <i>Camelus bactrianus</i> and <i>dromedarius</i> (HBB)           | 39%   | 3e-25          | 97%            |
| <i>Camelus ferus</i> (HBB) Vs. <i>Camelus bactrianus</i> and <i>dromedarius</i> (HBA)           | 39%   | 9e-30          | 97%            |
| <i>Camelus ferus</i> (HBA) Vs. <i>Camelus ferus</i> (HBB)                                       | 39%   | 3e-25          | 97%            |

Abbreviations: HBA:  $\alpha$ -chain of hemoglobin; HBB:  $\beta$ -chain of hemoglobin; Vs: versus

表 4 骆驼家族和人类中血红蛋白 ( $\alpha$  和  $\beta$  亚基) 中 N-, O-和 C-糖基化位点的预测. 糖基化类型、共有序列中糖基化的位置和糖基化的预测如下所示

Table 4 The predicted N-, O- and C-Glycosylation sites of hemoglobin (Alpha and Beta subunits) in *Camelus* family and human (*Homo sapiens*). Glycosylation type, position of glycosylation in consensus sequence and prediction of glycosylation are illustrated following

| Protein  | Glycosylation Type | Position               | Sequence | Prediction             |
|--|--------------------|------------------------|----------|------------------------|
| HBA <i>Camelus bactrianus</i> and <i>dromedarius</i> | N- glycosylation   | 9                      | NVK      | Non-glycosylated       |
|  |                    | 97                     | NFK      | Non-glycosylated       |
|  |                    | 131                    | NVS      | Potential glycosylated |
|  | O- glycosylation   | 3                      | S        | Potential glycosylated |
|  |                    | 4                      | S        | Potential glycosylated |
|  |                    | 8                      | T        | Potential glycosylated |
|  |                    | 12                     | T        | Non-glycosylated       |
|  |                    | 38                     | T        | Non-glycosylated       |
|  |                    | 39                     | T        | Non-glycosylated       |
|  |                    | 41                     | T        | Potential glycosylated |
|  |                    | 49                     | S        | Non-glycosylated       |
|  |                    | 52                     | S        | Non-glycosylated       |
|  |                    | 10                     | NVK      | Non-glycosylated       |
| 98   | NFK                | Non-glycosylated       |          |                        |
| 132  | NVS                | Potential glycosylated |          |                        |
| HBA <i>Camelus ferus</i>                             | O- glycosylation   | 4                      | S        | Potential glycosylated |
|  |                    | 5                      | S        | Potential glycosylated |
|  |                    | 9                      | T        | Non-glycosylated       |
|  |                    | 13                     | T        | Non-glycosylated       |
|  |                    | 39                     | T        | Potential glycosylated |
|  |                    | 40                     | T        | Potential glycosylated |
|  |                    | 42                     | T        | Potential glycosylated |
|  |                    | 50                     | S        | Potential glycosylated |
| 53   | S                  | Non-glycosylated       |          |                        |
| 68   | T                  | Non-glycosylated       |          |                        |



|                                 |                  |     |     |                        |
|---------------------------------|------------------|-----|-----|------------------------|
|                                 |                  | 79  | S   | Non-glycosylated       |
|                                 |                  | 82  | S   | Potential glycosylated |
|                                 |                  | 85  | S   | Potential glycosylated |
|                                 |                  | 103 | S   | Non-glycosylated       |
|                                 |                  | 109 | T   | Non-glycosylated       |
|                                 |                  | 119 | T   | Potential glycosylated |
|                                 |                  | 121 | S   | Non-glycosylated       |
|                                 |                  | 125 | S   | Potential glycosylated |
|                                 |                  | 134 | S   | Non-glycosylated       |
|                                 |                  | 135 | T   | Non-glycosylated       |
|                                 |                  | 138 | T   | Potential glycosylated |
|                                 |                  | 139 | S   | Potential glycosylated |
|                                 |                  | 10  | NVK | Non-glycosylated       |
|                                 |                  | 69  | NAV | Non-glycosylated       |
|                                 | N- glycosylation | 79  | NAL | Non-glycosylated       |
|                                 |                  | 98  | NFK | Non-glycosylated       |
|                                 |                  | 4   | S   | Potential glycosylated |
|                                 |                  | 9   | T   | Non-glycosylated       |
| HBA <i>Homo sapiens</i>         |                  | 36  | S   | Non-glycosylated       |
|                                 | O- glycosylation | 39  | T   | Non-glycosylated       |
|                                 |                  | 40  | T   | Non-glycosylated       |
|                                 |                  | 42  | T   | Potential glycosylated |
|                                 |                  | 50  | S   | Non-glycosylated       |
|                                 |                  | 53  | S   | Non-glycosylated       |
|                                 |                  | 68  | T   | Non-glycosylated       |
|                                 |                  | 10  | NAV | Non-glycosylated       |
|                                 |                  | 57  | NNP | Non-glycosylated       |
|                                 |                  | 58  | NPK | Non-glycosylated       |
|                                 | N- glycosylation | 70  | NSF | Non-glycosylated       |
|                                 |                  | 77  | NHL | Non-glycosylated       |
|                                 |                  | 81  | NLK | Non-glycosylated       |
| HBB <i>Camelus bactrianus</i> , |                  | 103 | NFR | Non-glycosylated       |
|                                 |                  | 109 | NVL | Non-glycosylated       |
|                                 |                  | 140 | NAL | Non-glycosylated       |
|                                 |                  | 5   | S   | Non-glycosylated       |
|                                 |                  | 17  | S   | Non-glycosylated       |
|                                 |                  | 39  | T   | Non-glycosylated       |
|                                 |                  | 45  | S   | Non-glycosylated       |
|                                 | O- glycosylation | 50  | S   | Non-glycosylated       |
|                                 |                  | 51  | T   | Non-glycosylated       |
|                                 |                  | 66  | S   | Non-glycosylated       |
|                                 |                  | 71  | S   | Non-glycosylated       |
|                                 |                  | 85  | T   | Non-glycosylated       |
|                                 |                  | 90  | S   | Non-glycosylated       |
|                                 |                  | 124 | T   | Non-glycosylated       |
|                                 |                  | 20  | NVD | Non-glycosylated       |
|                                 |                  | 58  | NPK | Non-glycosylated       |
|                                 | N- glycosylation | 81  | NLK | Non-glycosylated       |
|                                 |                  | 103 | NFR | Non-glycosylated       |
| HBB <i>Homo sapiens</i>         |                  | 109 | NVL | Non-glycosylated       |
|                                 |                  | 140 | NAL | Non-glycosylated       |
|                                 |                  | 5   | T   | Potential glycosylated |
|                                 |                  | 10  | S   | Non-glycosylated       |
|                                 |                  | 13  | T   | Non-glycosylated       |
|                                 |                  | 39  | T   | Non-glycosylated       |
|                                 | O- glycosylation | 45  | S   | Non-glycosylated       |
|                                 |                  | 50  | S   | Non-glycosylated       |

|     |   |                  |
|-----|---|------------------|
| 51  | T | Non-glycosylated |
| 73  | S | Non-glycosylated |
| 85  | T | Non-glycosylated |
| 88  | T | Non-glycosylated |
| 90  | S | Non-glycosylated |
| 124 | T | Non-glycosylated |

| Description   | Max score | Total score | Query cover | E value | Ident |
|---|-----------|-------------|-------------|---------|-------|
| <a href="#">Hemoglobin subunit beta [Pteropus alecto]</a>   | 345       | 593         | 99%         | 6e-114  | 52%   |
| <a href="#">PREDICTED: hypothetical protein LOC100482027 [Alluropoda melanoleuca]</a>   | 338       | 553         | 99%         | 4e-111  | 53%   |
| <a href="#">Chain A, Deoxy Rhb1.2 (Recombinant Hemoglobin) [Homo sapiens]</a>   | 324       | 570         | 99%         | 1e-106  | 56%   |
| <a href="#">Chain A, Cyanomet Rhb1.1 (Recombinant Hemoglobin) [Homo sapiens]</a>  | 324       | 570         | 99%         | 1e-106  | 56%   |
| <a href="#">PREDICTED: uncharacterized protein LOC100390052 [Callithrix jacchus]</a>  | 324       | 324         | 100%        | 8e-106  | 51%   |
| <a href="#">Chain A, Deoxy Hemoglobin (A-Glyqlqlv-Cv1m.L29w.B.D.v1m) [Homo sapiens]</a>   | 320       | 563         | 99%         | 6e-105  | 55%   |
| <a href="#">Chain A, Deoxy Hemoglobin (a-qlvlqlv-cv1m.I29f.h58g.B.d.v1m.v67w) [Homo sapiens]</a>  | 317       | 558         | 99%         | 9e-104  | 54%   |
| <a href="#">Chain A, Deoxy Hemoglobin (a-qlv-cv1m.I29f.h58g.B.d.v1m.I106w) [Homo sapiens]</a>   | 316       | 558         | 99%         | 1e-103  | 54%   |
| <a href="#">Chain A, Deoxy Hemoglobin (A-Gly-Cv1m.L29w.H58g.B.D.v1m) [Homo sapiens]</a>   | 315       | 557         | 99%         | 3e-103  | 54%   |
| <a href="#">hypothetical protein [Sorex araneus]</a>  | 315       | 545         | 99%         | 6e-103  | 52%   |
| <a href="#">PREDICTED: hemoglobin subunit beta [Camelus ferus]</a>  | 303       | 385         | 91%         | 1e-100  | 100%  |
| <a href="#">Chain B, Crystal Structure Determination Of Camel(Camelus Dromedarius)hemoglobin At 2 Angstrom Resolution [Camelus dromedarius]</a> | 301       | 382         | 90%         | 2e-99   | 100%  |
| <a href="#">PREDICTED: hemoglobin subunit beta [Vicugna pacos]</a>  | 300       | 381         | 91%         | 4e-99   | 99%   |
| <a href="#">olfactory receptor 52Z1-like protein [Cricetulus criseus]</a>   | 310       | 593         | 98%         | 7e-96   | 52%   |
| <a href="#">PREDICTED: hemoglobin subunit alpha-like [Camelus ferus]</a>  | 286       | 379         | 90%         | 8e-94   | 100%  |
| <a href="#">RefName: Full=Hemoglobin subunit alpha; AltName: Full=Alpha-globin; AltName: Full=Hemoglobin alpha chain [Camelus bactrianus]</a>   | 285       | 379         | 90%         | 1e-93   | 100%  |
| <a href="#">PREDICTED: hemoglobin subunit alpha-like [Camelus ferus]</a>  | 281       | 370         | 90%         | 5e-92   | 99%   |
| <a href="#">RefName: Full=Hemoglobin subunit alpha; AltName: Full=Alpha-globin; AltName: Full=Hemoglobin alpha chain [Lama guanicoe]</a>        | 280       | 370         | 90%         | 1e-91   | 97%   |
| <a href="#">RefName: Full=Hemoglobin subunit alpha; AltName: Full=Alpha-globin; AltName: Full=Hemoglobin alpha chain [Vicugna vicugna]</a>      | 279       | 368         | 90%         | 5e-91   | 97%   |

图 4 在来自 GenBank 的不同物种的野骆驼, 单峰驼和双峰驼中  $\alpha$  链 (HBA) 和  $\beta$  链 (HBB) 之间的 PSI-BLAST (位置特异性迭代 BLAST)。我们发现了很多来自不同物种的同源基因, 它们各自与野骆驼, 单峰驼和双峰驼 (HBA) 的  $\alpha$  链和  $\beta$  链血红蛋白具有同一性。最大得分, 总得分, 查询覆盖, E 值和同一性显示为基因组同源性

Figure 4 PSI-BLAST (Position-Specific Iterated BLAST) between  $\alpha$ -chain (HBA) and  $\beta$ -chain (HBB) *Camelus ferus*, *dromedarius* and *bactrianus* with different species from GenBank. We found a lot of HomoloGene from different species that each them have identity with  $\alpha$ -chain and  $\beta$ -chain hemoglobin of *Camelus ferus*, *dromedarius* and *bactrianus* (HBA). Maximum score, total score, query cover, E-value and ident are shown as genome homolo

保护血红蛋白抗糖基化。最终我们认为骆驼的特定特征取决于不同类型的蛋白质。并且, 我们建议关于蛋白质的独特特征的研究将适合于鉴定多肽的功能和效率, 以及随后通过蛋白质工程生产针对糖尿病、帕金森病、阿尔茨海默病和其他的合成药物。

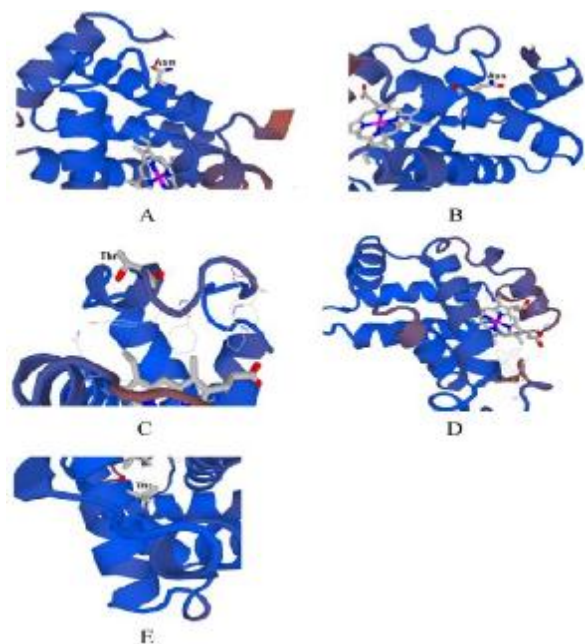


图 3 整个序列预测模型的三级结构. 3A) HBA 野骆驼. 3B) HBA 双峰驼和单峰驼. 3C) HBA 人类. 3D) HBB 野骆驼, 双峰驼和单峰驼 3E) HBB 人类

Figure 3 The tertiary structure of predicted model for the whole sequences. 3A) HBA *Camelus ferus*. 3B) HBA *Camelus bactrianus* and *dromedarius*. 3C) HBA *Homo sapiens*. 3D) HBB *Camelus ferus*, *Camelus bactrianus* and *dromedarius*. 3E) HBB *Homo sapiens*

#### 致谢

感谢来自 Guilan university 的博士生 Dr. Hassani-Bafrani 对本文英文语法方面的修改。

#### 参考文献

Al-Qarawi A.A., Mousa H.M., 2004, Lipid concentrations in erythrocyte membranes in normal, starved, dehydrated and rehydrated camels (*Camelus dromedarius*), and in normal sheep (*Ovis aries*) and goats (*Capra hircus*). *J. Arid Environ*, 59: 675-683  
<http://dx.doi.org/10.1016/j.jaridenv.2004.02.004>  
Altschul S.F., Madden T.L., Schaffer A.A., et al., 1997,

Gapped Blast and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25: 3389-3402

<http://dx.doi.org/10.1093/nar/25.17.3389>

Ardia D.R., 2006, Glycated hemoglobin and albumin reflect nestling growth and condition in American kestrels. *Comp. Biochem. Physiol. A*, 143: 62-66

<http://dx.doi.org/10.1016/j.cbpa.2005.10.024>

Azwai S.M., Abdousslam O.E., AL-Bassam L.S., et al., 2007, Morphological characteristics of blood cells in clinically normal adult Llamas (*Lama glama*). *Vet. Arhiv*, 77(1): 69-79

Balasubramanian M., Moorthy P.S., Neelagandan K., Ponnuswamy M.N., 2009, Purification, crystallization and preliminary crystallographic study of haemoglobin from camel (*Camelus dromedarius*): a high oxygen-affinity lowland species. *Acta Cryst*, 65: 773-775

Bazzi M.D., Quazani M., Rabbani N., et al., 2013, Glycated hemoglobin in camel: minimal correlation with blood glucose level. *Arch. Biol. Sci., Belgrade*, 65 (3): 911-917

Binder E., Bermudez-Silva F.J., Andre C., et al., 2013, Leucine supplementation protects from insulin resistance by regulating adiposity levels. *PLoS ONE*. 8(9): e74705

<http://dx.doi.org/10.1371/journal.pone.0074705>

Bognera P., Misetab A., Berentec Z., et al., 2005, Osmotic and diffusive properties of intracellular water in camel erythrocytes: Effect of hemoglobin crowdedness. *Cell. Biol. Int*, 29: 731-736

<http://dx.doi.org/10.1016/j.cellbi.2005.04.008>

Borai A., Livingstone C., Abdelaal F., et al., 2011, The relationship between glycosylated haemoglobin (HbA1c) and measures of insulin resistance across a range of glucose tolerance. *Scand. J. Clin. Lab. Invest*, 71: 168-172

Braunitzer G., Schrank B., Stangl A., Wiesner H., 1980, Respiration at high altitudes phosphate-protein-interaction: Sequence of the

- hemoglobins of the hamster (*Mesocricetus aureatus*) and the camel (*Camelus ferus*, Camelidae). *J. Chem. Soc. Pak*, (1):1-7 (In Germany)
- Chauhan J.S., Rao A., Raghava G.P.S., 2013, In silico Platform for Prediction of N-, O- and C-Glycosites in Eukaryotic Protein Sequences. *PLoS ONE*, 8(6): e67008  
<http://dx.doi.org/10.1371/journal.pone.0067008>
- Christlet T.H.T., Veluraja K., 2001, Database analysis of O-glycosylation sites in proteins. *Biophys. J*, 80: 952-960  
[http://dx.doi.org/10.1016/S0006-3495\(01\)76074-2](http://dx.doi.org/10.1016/S0006-3495(01)76074-2)
- Deléage G., Roux B., 1987, An algorithm for protein secondary structure prediction based on class prediction. *Protein. Eng. Des. Sel*, 5: 87-91
- Dickerson R., Geiss I., 1983, Hemoglobin: Structure, Function, Evolution and Pathology. Benjamin/Cummings: Menlo Park, CA
- Farooq U., Samad H.A., Khurshid A., Sajjad S., 2011, Normal reference hematological values of one-humped (*Camelus dromedarius*) kept in Cholistan desert. *J. Anim. Plant. Sci*, 21(2): 157-160
- Gavel Y., von Heijne G., 1990, Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein. Eng*, 3(5): 433-442  
<http://dx.doi.org/10.1093/protein/3.5.433>
- Goniakowska-Witalinska L., Witalinski W., 1976, Evidence for a correlation between the number of marginal band microtubules and the size of vertebrate erythrocytes. *J. Cell Sci*, 22: 397-401
- Jirimutu., Zhen W., Guohui D., et al. 2012, Genome sequences of wild and domestic Bactrian camels. *Nat. Commun*, 3:1202
- Krieg J., Hartmann S., Vicentini A., et al., 1998, Recognition Signal for C-Mannosylation of Trp-7 in RNase 2 Consists of Sequence Trp-x-x-Trp. *Mol. Biol. Cell*, 9: 301-309  
<http://dx.doi.org/10.1091/mbc.9.2.301>
- Lin K.T.D., Bhowan A.S., Chernoff A.I., 1976, Studies on camel hemoglobin: physicochemical properties and some structural aspects of camel hemoglobin (*Camelus dromedarius*). *Biochim. Biophys. Acta*, 434 (1): 110-117  
[http://dx.doi.org/10.1016/0005-2795\(76\)90040-4](http://dx.doi.org/10.1016/0005-2795(76)90040-4)
- Miksik I., Hodny Z., 1992, Glycated hemoglobin in mute swan (*Cygnus olor*) and rook (*Corvus Frugilegus*). *Comp. Biochem. Physiol*, 3: 553-555
- Nathan D.M., Kuenen J., Borg R., et al., 2008, Translating the A1C assay into estimated average glucose values. *Diabetes Care*, 31:1473-1478  
<http://dx.doi.org/10.2337/dc08-0545>
- Oyewale J., Dzenda T., Yaqub L., et al., 2011, Alterations in the osmotic fragility of camel and donkey erythrocytes caused by temperature, pH and blood storage. *Veterinarski arhiv*, 81: 459-470
- Pairet B., Jaenicke E., 2010, Structure of the Altitude Adapted Hemoglobin of Guinea Pig in the R2-State. *PLoS ONE*, 5(8): e12389  
<http://dx.doi.org/10.1371/journal.pone.0012389>
- Paoli M., Nagai K., 2004, Hemoglobin. In: Messerschmidt A, ed. *Handbook of Metalloproteins*. Chichester: Wiley-VCH, pp 1-15
- Pieragostini E., Alloggio I., Petazzi F., 2010, Insights into Hemoglobin Polymorphism and Related Functional Effects on Hematological Pattern in Mediterranean Cattle, Goat and Sheep. *Diversity*, 2: 679-700  
<http://dx.doi.org/10.3390/d2040679>
- Rohlfing C., Hsiao-Mei W., Randie L., et al., 2002, Biological Variation of Glycohemoglobin. *Clin. Chem*, 48 (7): 1116-1118
- Shahbazkia H.R., Shareghi B., Aminlari M., 2010, Glycated hemoglobin is an indicator of blood glucose status in horses: Preliminary study. *J. Equine. Vet. Sci*, 30: 47-49  
<http://dx.doi.org/10.1016/j.jevs.2009.11.007>
- Soranzo N., 2011, Genetic Determinants of Variability in Glycated Hemoglobin (HbA1c) in Humans:



Review of Recent Progress and Prospects for Use  
in Diabetes Care. *Curr. Diab. Rep.*, 11: 562-569  
<http://dx.doi.org/10.1007/s11892-011-0232-9>

Tamura K., Peterson D., Peterson N., et al., 2011,  
MEGA5: Molecular Evolutionary Genetics  
Analysis using Maximum Likelihood,  
Evolutionary Distance, and Maximum Parsimony

*Methods. Mol. Biol. Evol.*, 28 (10): 2731-9

<http://dx.doi.org/10.1093/molbev/msr121>

Wautier J.L., Schmidt A.M., 2004, Protein Glycation:  
A Firm Link to Endothelial Cell Dysfunction.  
*Circ. Res.*, 95: 233-238

[http://dx.doi.org/10.1161/01.RES.0000137876.28454.  
64](http://dx.doi.org/10.1161/01.RES.0000137876.28454.64)