









研究报告

Research Report

基于生物信息学方法的FOX基因家族对比分析

常凯 , 熊怡淞 , 曲远青 , 陈清红 , 吴艾霖 , 吴丽娟 

中国人民解放军成都军区总医院实验医学中心高湿医学全军重点实验室, 成都, 610083

 通讯作者, wulijuan1638@126.com;  作者

计算分子生物学, 2015 年, 第 4 卷, 第 6 篇 doi: 10.5376/cmb.cn.2015.04.0006

收稿日期: 2015 年 08 月 20 日 接受日期: 2015 年 08 月 20 日 发表日期: 2015 年 08 月 20 日

© 2015 BioPublisher 生命科学中文期刊出版平台

本文首次发表在《基因组学与应用生物学》印刷版 2015 年, 第 34 卷, 第 09 期上。现依据版权所有人授权的许可协议, 采用 [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) 协议对其进行授权, 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。

摘要 以 FOXA, FOXO 和 FOXP 为核心, 利用生物信息学方法对比分析 FOX 家族成员的分子特征、启动子、蛋白理化性质、蛋白结构、亚细胞定位, 构建分析蛋白三维模型并绘制家族系统进化树。结果表明: FOX 家族的核酸序列相似性较低, FOXP 和 FOXN 序列内含较多内含子, 但仍具有溯源性。蛋白的脂肪族氨基酸指数介于 49.71 到 76.11, 分子量介于 29084 到 79586.1, 较为离散。GRAVY 值介于 -0.764 到 -0.179, 蛋白均表现为亲水性。pI 值在亚家族间非常稳定, 可作为蛋白分类的辅助指标。对比统计 FOX 蛋白高级结构各亚家族的保守结构域和功能位点, 可见功能域的差异导致了其调节途径的多样。分析结果可为 FOX 蛋白的研究提供有价值信息, 为进一步研究 FOX 在人体内的疾病发生途径与作用机理提供依据。



关键词 叉头框基因家族; 生物信息学; 对比分析; 分子特征; 蛋白特征

Comparative Analysis of FOX Family Genes Based on the Bioinformatics

Methods

Chang Kai , Xiong Yisong , Qu Yuanqing , Wu Ailin , Chen Qinghong , Wu Lijuan 

KeyLaboratoryofHighHumidityMedicine,CenterofLaboratoryMedicine,ChengduMilitaryGeneralHospital,Chengdu,610083

 Corresponding author, wulijuan1638@126.com;  Authors

Abstract The molecular characteristics, promoter sequence, protein properties, amino acid sequence and the tertiary structure were analysed via bioinformatics methods. The phylogenetic tree of the FOX family was build. The result showed that the similarity of nucleic acid sequence was low, but the homology was not low. Aliphatic amino acid index of protein between 29084 to 79586.1; GRAVY values ranged from -0.764 to -0.179, and the protein was hydrophilic. The isoelectric point was stable in the FOX subfamily and can be used as the assistant index of protein classification. The statistics and analysis of the FOX structural domain conserved domains and functional sites revealed the differences of protein function. The results of the analysis can provide valuable information for the study of FOX protein, and provide the basis for further study of the mechanism of the disease in the human body and the mechanism of the disease about FOX.

Keywords Forkhead box family, Bioinformation, Comparative analysis, Molecular characteristics, Protein properties

转录因子是真核表达调控中非常重要的一类反式作用因子, 能通过多元化分子调节机制作用于多条通路参与整个生物进程。叉头框家族(Forkhead box family, FOX)为转录因子大家族之一, 以分子结构上有一个明显的叉头 DNA 结合区域为特征, 该叉头 DNA 结合区由 100 个氨基酸组成, 称“Winged helix”或“Forkhead”的 DNA 结合结构域

(DNA-binding domain, DBD)(Nandymazumdar, 2015)。人类叉头框基因家族命名从 FOXA 到 FOXS, 生物学功能涉及细胞周期、细胞分化、胚胎发育和免疫调节等。是肿瘤、糖尿病和心血管疾病等相关研究的热点。

目前发现的FOX蛋白均具高保守的DBD区域, 研究人员用核磁共振和X射线结晶性的方法分析出

多种FOX蛋白DBD的空间三维结构。FOX蛋白呈现出一个似蝴蝶结状的翼型螺旋区，核心部位由3个 α 螺旋依次排列，中间通过3个 β 折叠连接2个大环组成，形成以 $\alpha 1-\beta 1-\alpha 2-\alpha 3-\beta 2-W 1-\beta 3-W 2$ 的基序。本文基于生物信息学方法对已克隆出的多组FOX大家族成员进行比对分析，为进一步研究FOX奠定基础。

本文在NCBI数据库和PDB数据库中选取近十年来发布的FOX家族mRNA，genomic DNA和氨基酸序列100余个，涵盖不同物种、不同亚家族(Brown, 2015)，分别进行编码序列(coding sequence, CDS)分析、基因组序列分析和蛋白序列的分析。

1 结果与分析

1.1 基因序列的对比分析

应用Vector NTI对FOX家族的核酸序列进行对比分析，可见3'-UTR和5'-UTR的长短差异和碱基结构差异都很大，但CDS区域的同源性与保守性都很高，是区分亚家族的主要标志。应用RNA structure软件分别对FOX基因的3'-UTR进行结构预测，结果并未发现特殊的发卡与高级结构(Ponugoti, 2013)。对FOX家族基因进行内含子与外显子分析，可见FOXD、FOXQ、FOXL和FOXK无内含子，FOXO、FOXA家族有两个外显子，而FOXP和FOXN的外显子均大于5个。内含子能够调节基因转录的速率，为结合蛋白提供结合位点，这证明了FOXP具有复杂的功能结构和较远的遗传距离。染色体分布随机，并未集中于特定的染色体区带。取CDS上游2000bp的序列进行启动子预测，启动子序列由小写字母表示，大写字母为转录起始位点。

1.2 氨基酸序列的聚类对比分析

将来源于人、小鼠等不同物种的多个FOX亚群的氨基酸序列在CLUSTALX和MEGA5.0软件中进行分子进化树分析，基于Neighbor-Joining的原理建树(Lan, 2013)，结果显示绿色标记的分支为FOXP家族，红色标记的分支为FOXA家族，蓝色分支为FOXO家族，黄色分支标记为FOXL家族，紫色分支为其他FOX亚家族。由此可见，进化树的聚类分析不是由物种进行分类，而是由FOX的各个亚家族的差异进行聚类，进而说明虽然物种间差异巨大，但FOX亚家族内部的氨基酸序列相似性极高且足以稀释物种间的差异。各分支内部的不同物种聚类进

化距离分析发现，人类的FOX基因均处于进化树最顶端，可见人类FOX的进化发生最晚，功能最复杂(Csuros, 2013)。

1.3 FOX蛋白特性对比分析

利用SignalP 3.0 Server和 TMHMM Server v. 2.0网络服务器对FOX进行跨膜区分析，结果发现在FOX氨基酸序列中不存在跨膜区。应用Wolf-psort软件进行蛋白的亚细胞定位分析均显示FOX，蛋白定位于细胞核与细胞质中，这与文献报道的FOX因子发生磷酸化定位于细胞质中，在应激状态下可定位于细胞核中的研究结果一致。

在ExPASy Proteomics Server网站上用ProtScale软件对FOX蛋白进行预测，统计Aliphatic index值、GRAVY值、pI值、分子量等一系列数据，分析后发现Aliphatic index、GRAVY值和分子量在不同物种、不同FOX亚群间没有典型的规律性分布；但不同物种间的同一FOX亚群的pI值非常稳定，不会因为分子量的大小和氨基酸的分布而改变。如图2所示，不同的FOX亚群间和亚群内会有非常显著地pI差异(应用SPSS的Duncan's统计方法分析， $P < 0.01$)，说明pI值可以作为区分不同亚群的标志(Chang, 2014)。介于此种特性，应用等电聚焦电泳可以快速有效的区分FOX隶属于哪一亚群。GRAVY值均为负值，表明FOX蛋白均为亲水性蛋白，均有助于与靶DNA的结合。脂肪族氨基酸指数最低49.71，最高76.11，分布均匀(Chang, 2014)。

1.4 氨基酸序列比对与高级结构比对

利用Vector NTI、NCBI BLAST软件对FOX蛋白进行对比分析(Lu, 2004)，进行FOX全家族氨基酸的同源性为7.2%，相似性为0.1%。FOX亚家族成员内的比对分析，发现FOXA同源性为51.6%，相似性为17.5%；FOXP同源性为58.6%，相似性为8.8%；FOXO同源性为100%，相似性为8.5%；FOXQ同源性为33.6%，相似性为8.4%；FOXL同源性为54%，相似性为13.6%；FOXN同源性为37.5%，相似性为15.2%；FOXJ同源性为46.3%，相似性为8.8%。上述比对结果提示，FOX家族间的序列相似性和同源性均非常低，氨基酸序列差异非常大。由于FOX亚家族内的同源性在50%左右，相似性只有8%左右，因此即便是同一亚家族的基因相似性也并不高。

在PBIL网站上使用SOPMA软件对FOX进行二级结构预测表明：FOXO1包含25.8%的 α -螺旋、

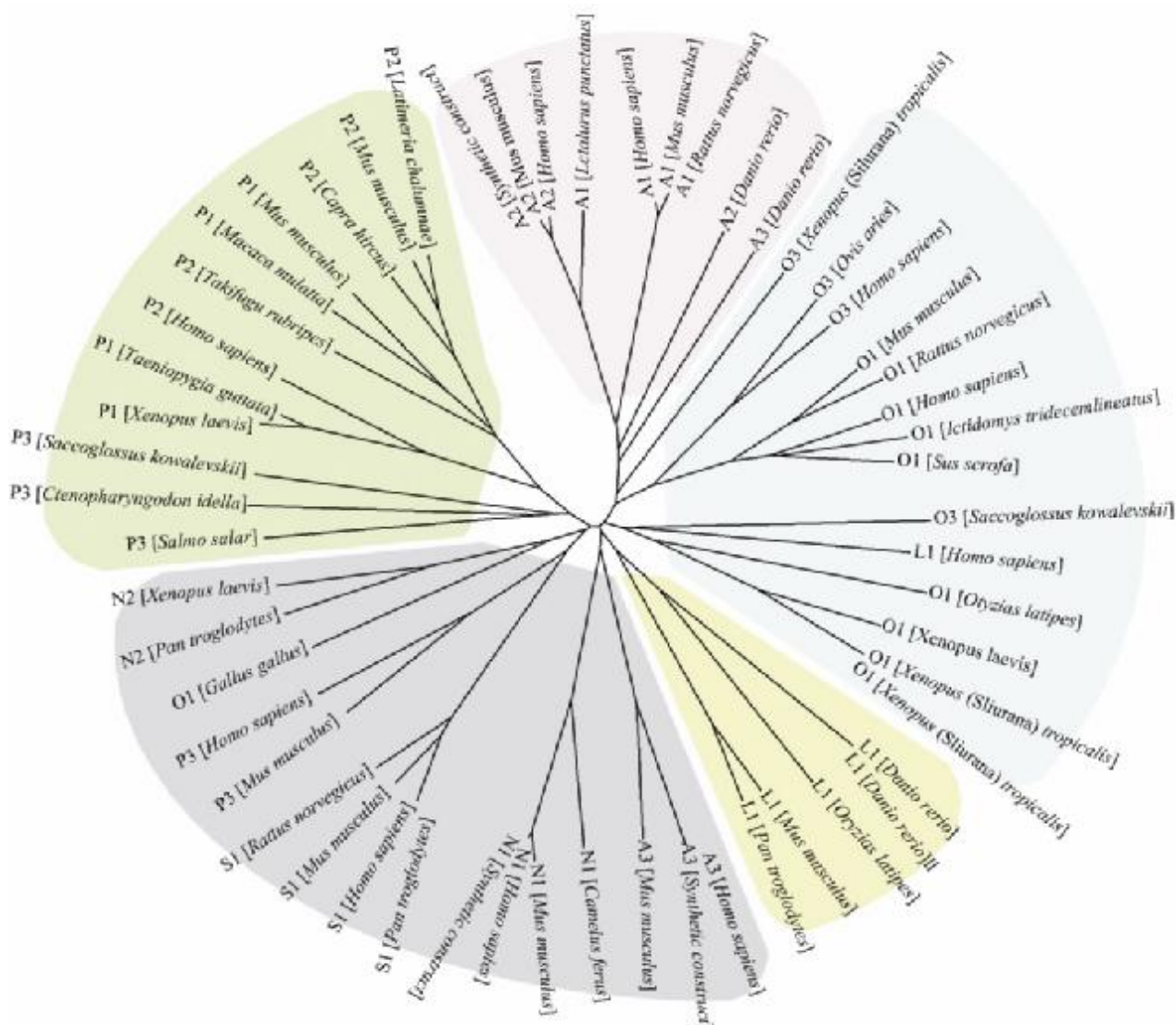


图 1 基于 MEGA5.0 软件平台用 Neighbor-Joining 方法构建的 FOX 进化树

注: 来源于 FOXF 家族的一支在绿色框内; FOXA 家族的一支在红色框内; FOXO 家族的一支在蓝色框内; FOXL 家族的一支在黄色框内

Figure 1 A phylogenetic tree of FOX from different organisms constructed by Neighbor-Joining method on MEGA 5.0

Note: FOXF showed in green box; FOXA showed in red box; FOXO showed in blue box; FOXL showed in yellow box

11.91%延伸带、6.56%β转角和55.73%随意卷曲构成; FOXF3包含34.34%的α-螺旋、14.14%延伸带、7.07%β转角和44.44%随意卷曲构成; FOXA1包含27.12%的α-螺旋、12.71%延伸带、7.84%β转角和52.33%随意卷曲构成; FOXA2包含32.6%的α-螺旋、9.63%延伸带、10.07%β转角和47.7%随意卷曲构成; FOXQ包含13.65%的α-螺旋、11.91%延伸带、5.71%β转角和68.73%随意卷曲构成(Geourjon, 1995。上述

数据提示: FOX家族的氨基酸序列均可预测出典型的α1-β1-α2-α3-β2-W1-β3-W2结构。

基于日益庞大的蛋白质结构数据库, 利用信息学的方法对蛋白质的高级结构进行预测和分析, 对理解其结构及功能具有一定意义。该研究用同源建模法, 对蛋白质分子的三维结构进行初步模拟。应用SWISS-MODLE和WebLab ViewerLite软件构建FOX三级结构进行功能位点与结构进行比较分析,

表1 人类 FOX 基因统计分析
Table 1 The statistical analysis of *Homo sapiens* FOX gene

基因名 Gene name	登录号 Access number	外显子 Exon region	染色体 Chromosome	启动子预测 Promoter predictor
H sapiens fox O1	NG_023244	1-630; 105 351-106 690	Chr 13	tctagcattttaaanaaggtcaagatacagtgatccggAtgggggctg
H sapiens fox O3	NC_000006.12	1-621; 102 245-103 647	Chr 6	cgcgcccgccttaagggcgggccgcccccccgggcGccctcctcc
H sapiens fox A1	NC_000014	1-73; 2 260-3 608	Chr 14	tcataagatataaaccggctgtgtactaccctgctcttAgcgcgaagg
H sapiens fox A2	NC_000020.11	1-88; 1 119-2 429 156 848-157 054; 220 353-220 561; 1-185; 232 322-232 493; 151 285-151 445; 238 991-239 135; 144 603-144 746; 225 704-225 827; 85 742-85 848; 182 727-182 833; 221 337-221 441; 145 754-145 845; 197 323-197 410; 210 302-210 388; 220 656-220 740; 227 576-227 647	Chr 20	ctgactctttttaaanaagggctgggggtggggatGctgtgctgt
H sapiens fox P1	NC_000003.12	215 897-216 112; 203 291-205 494; 205 015-205 194; 1-169; 202 029-202 166; 227 411-227 531; 218 170-218 279; 108 104-108 196; 225 682-225 783; 144 294-144 372	Chr 3	aaaatacacaanaattagccggcagtggtgcatgcacCtgtatcct
H sapiens fox P2	NC_000007.14	1-210; 4 438-4 589; 937-1 080; 7 021-7 159; 735-858; 1 648-1 756; 6 740-6 842; 1 487-1 581; 2 707-2 791; 2 994-3 078; 5 301-5 381	Chr 7	agtgatattttaaanaagggctgtatagtcctcctggtGtcaaaagta
H sapiens fox P3	NC_000023	10 735-11 250; 534-999; 13 147-13 467; 10 560-10 569; 5 122-5 255;	Chr 11	aggctcactgtaaaaagtgccagtactgggtctgtaagtcAgtcaagat
H sapiens fox N1	NC_000017.11	1-123; 3 281-3 392; 6 767-6 876	Chr 17	tgatgacggcattttaggggcacggactgtagccaggcAcctgtgctga
H sapiens fox N2	NC_000002.12	1-539; 28 706-29 229; 12 831-12 933; 27 076-27 148; 16 429-16 493	Chr 2	acctggcctataagctgcactatttattacagcatttAactatata
H sapiens fox D1	NC_000005.10	1-1 398	Chr 5	ggcagtggttatanaagtcggcggcagtgaccgccAcccgcagc
H sapiens fox Q1	NC_000006	1-1 212	Chr 6	gccgaggggtatattggccggccgccgagggggagggcGcgcagccg
H sapiens fox L1	NC_000016.10	1-1 038	Chr 16	tcagtcggttatatatgggtgccggaaatttccaaAtaacacag
H sapiens fox L2	NC_000003.12	1-1 131	Chr 3	agtcactttttagaaagcccccgcacagctgggcccggcTggcccgcc
H sapiens fox C1	NC_000006.12	1-1 662	Chr 6	taactcgtgtaanaagtcctcctcctgagtgacggatgctcAaaggtcag

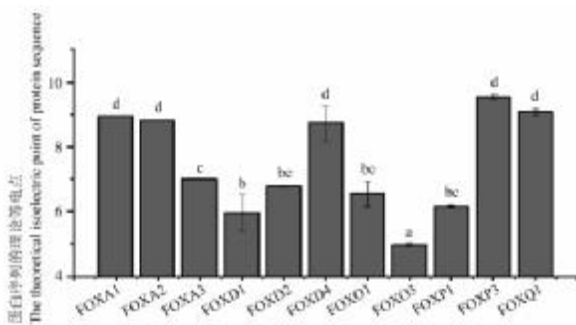


图2 FOX 蛋白序列的等电点统计分析
注: 统计分析应用邓肯氏复极差测验; 显著性差异由字母标示 ($p < 0.01$)
Figure 2 The pI point of FOX protein sequence
Note: The statistical analysis was performed by Duncan's multiple range tests; The different letters indicate significant difference ($p < 0.01$)

结合Swiss-PdbViewer拉式构象图进行蛋白构想合理性评估(Biasini, 2014)。分析Psi和Phi角分布如图3所示, 超过90%的氨基酸残基位于理想Psi和Phi角分布区域内(蓝色框), 可见预测结构的排布稳定, 同源建模结果相对可靠(Johansson, 2012)。

如图4所示, 各FOX亚家族均具有典型的DBD结构和保守的功能位点。FOXO的DBD结构与16碱基DNA序列(5'-TTGTTTAC-3')结合或与胰岛素响应元件IRE(5'-TTGTTTGTG-3')结合, 保守功能位点为Ans 211和His 211(Brent, 2008)。FOXP的单

体结合序列 5'-TGGAAAATTTGTTTCA-3', 在FOXP3中这段序列是活化T细胞的核因子(nuclear factor, NFAT), 该蛋白功能位点为Trp313和Met335(Bandukwala, 2011)。同样, FOXA的功能位点为Arg213和Arg219, FOXQ的功能位点为Asn165和His169(Jin, 1999)。可见, FOX家族的氨基酸序列相似性虽不高, 但有着及其相似的三级结构及功能位点, 由此可决定结合的靶向DNA序列及功能(Van, 2000)。可见三级结构预测的结果与二级结构预测分析较为一致(Marsden, 1998)。蛋白高级结构比对可知每个亚家族都有相对保守的功能域和保守位点, 这是进行靶DNA结合与基因调控的关键, 有助于通过查询FOX家族成员保守位点预测蛋白调节功能。

2讨论

FOX蛋白不仅能作为典型的转录因子通过激活目标蛋白调节基因转录, 还能直接同凝聚染色质结合参与其重构, 协同其他转录因子参与转录调节。生物信息学对比分析与既往研究证实, FOX蛋白参与TGF- β /Smads信号转导, 在肝纤维化和肿瘤的发生发展过程中起重要作用(Attisano, 2001)。FOX蛋白参与MAPK信号转导, 调节相关基因转录, 参与细胞生长、发育和分裂(Wu, 2005)。FOX蛋白也参与Akt/PKB信号转导, 可通过联合作用打破细

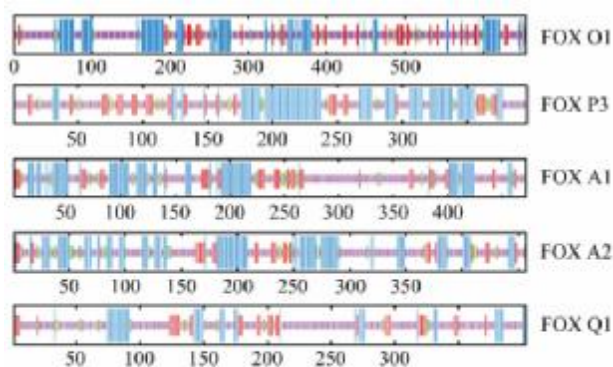


图 3 FOX 蛋白二级结构预测

Figure 3 FOX protein secondary structure prediction by consensus prediction from multiple alignments

胞增殖与凋亡之间的平衡, 促进细胞生长从而导致肿瘤发生(Hilgier, 2005)。在FOX蛋白参与Wnt信号转导过程中调节Wnt蛋白, 在调控胃肠道上皮细胞和间叶细胞的增殖中发挥关键作用(Kontny, 2006)。FOX家族的核苷酸序列差异较大, 但仍有相对保守的区域, 即编码DBD结构的核苷酸序列。氨基酸序列分析可见每一个亚家族的氨基酸序列均具有高度的同源性。跨膜区分析预测表明, FOX蛋白不具有理论跨膜区域, 但亚细胞定位与实验表明FOX出现在细胞质中和细胞核中, 由此推断FOX蛋白可能通过核孔复合体或其他途径跨越核膜(Wilhelmsen, 2007)。本研究发现等电点在不同FOX亚种间表现出的显著性差异, 但在亚种内非常稳定, 这将成为FOX蛋白亚群区分时的又一辅助指标。因此, 等电聚焦电泳可作为一种快速有效区分FOX亚群的方法(Maccarrone, 2015)。

3材料与方法

3.1数据来源

以FOXA, FOXO和FOXP基因为出发点, 在NCBI及PDB数据库中运行BLAST程序, 找出FOX家族核酸序列及蛋白序列各107个作为候选, 应用Vector NTI软件剔除序列过长或过短的干扰DNA与蛋白。最终确定涵盖FOXA、FOXO、FOXP及典型FOX亚家族的核酸及蛋白序列79个。

3.2 FOX亚家族的核酸及蛋白序列的生物信息学分析

应用Vector NTI对FOX家族的核酸序列进行对比分析, BDGP promoter软件进行FOX DNA序列进行启

动子预测(http://fruitfly.org/seq_tools/promoter.html)。应用 ExPasy 数据库中 ProtParam 软件 (<http://www.expasy.org/tools/ProtParam.html>) 对比分析并统计目的蛋白的理化特性包括: 等电点, 分子量, 氨基酸数目, 脂肪系数, 亲疏水系数等。在 ExPASy Proteomics Server网站上用ProtScale软件 (<http://web.expasy.org/protScale>)对FOX蛋白的氨基酸序列进行亲/疏水性预测。Vector NTI8.0软件对其进行保守序列分析比对。应用WoLF PSORT在线软件 (http://www.genscript.com/psort/wolf_psort.html)对FOX蛋白进行亚细胞定位预测。TMHMM Server v2.0(<http://www.cbs.dtu.dk/services/TMHMM/>) 和 Tmpred软件对蛋白的跨膜区域进行预测。在PBIL网站上使用MLRC软件对FOX进行二级结构预测。蛋白结构域建模比对分析和分析结果评估通过Clustal X、MEGA5.0、Swiss Modle、Swiss-PdbView和WebLab ViewLite完成。

作者贡献

常凯完成论文构思、数据分析统计和写作; 熊怡淞、曲远青、陈清红和吴艾霖参与论文分析与修改; 吴丽娟完成论文的修改和最终定稿。

致谢

感谢国家自然科学基金(81302596)资助; 感谢成都军区总医院细胞分子科平台支持。

参考文献

- Attisano L., Silvestri C., Izzi L., and Labbé E., 2001, The transcriptional role of Smads and FAST (FoxH1) in TGFbeta and activin signalling, *Mol. Cell Endocrinol.*, 180 (1-2): 3-11
- Bandukwala H.S., Wu Y., Feuerer M., Chen Y., Barboza B., Ghosh S., Stroud J.C., Benoist C., Mathis D., Rao A., and Chen L., 2011, Structure of a domain-swapped FOXP3 dimer on DNA and its function in regulatory T cells, *Immunity*, 34(4): 479-491
- Biasini M., Bienert S., Waterhouse A., Arnold K., Studer G., Schmidt T., Kiefer F., Cassarino T.G., Bertoni M., Bordoli L., and Schwede T., 2014, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, *Nucleic Acids Res.* 42(Web Server issue): W252-W258
- Brent M.M., Anand R., and Marmorstein R., 2008, Structural

FOXO1	[<i>H. sapiens</i>]	RRNAWGNLSYADLITKAIESSAEKRLTLSQLYEWVMVKSVPYFKDKGDSNSSAGWKN SIRHNL SLHISKFRVQNEGTGKSSWMLNPEG
FOXO4	[<i>M. musculus</i>]	RRNAWGNQSYAELISQAIESAPEKRLTLAQIYEWVVRTVVPYFKDKGDSNSSAGWKN SIRHNL SLHISKFKVIINEATGKSSWMLNPDG
FOXO5	[<i>D. rerio</i>]	RRNAWGNLSYADLITKAIESSTPKRLTLSQLYEWVMVSSVPYFKDKGDSNSSAGWKN SIRHNL SLHSRFRVQNEGTGKSSWMLNPEG
FOXO3	[<i>M. musculus</i>]	RRNAWGNLSYADLITRAIESSPKRLTLSQLYEWVMVRCVVPYFKDKGDSNSSAGWKN SIRHNL SLHSRFRVQNEGTGKSSWMLNPDG
FOXO6	[<i>P. alecto</i>]	RRNAWGNLSYADLITKAIESAPDKRLTLSQLYDWMVRYVVPYFKDKGDSNSSAGWKN SIRHNL SLHTRFRVQNEGTGKSSWMLNPEG
FOXF3	[<i>H. sapiens</i>]	RPPFTYATLIRWAILEAPEKQRTLNEIYEWFRMFAPFRNHDPATWKN AIRHNL SLHKCFVRVSEKGA VVTVDELEFRKKR
FOXP1	[<i>P. alecto</i>]	RPPFTYASLIRQAILESPKQLTLNEIYNWFRMFAYFRNAATWKN AVRHNL SLHKCFVRVENVKGA VVTVDEVEFQKRR
FOXP1	[<i>R. norvegicus</i>]	RPPFTYASLIRQAILESPKQLTLNEIYNWFRMFAYFRNAATWKN AVRHNL SLHKCFVRVENVKGA VVTVDEVEFQKRR
FOXP4	[<i>R. norvegicus</i>]	RPPFTYASLIRQAILETPDRQLTLNEIYNWFRMFAYFRNTATWKN AVRHNL SLHKCFVRVENVKGA VVTVDEREYQKRR
FOXP3	[<i>R. rattus</i>]	RPPFTYATLIRWAILEAPEKQRTLNEIYEWFRMFAPFRNHDPATWKN AIRHNL SLHKCFVRVSEKGA VVTVDEFFFRKKR
FOXP3	[<i>P. alecto</i>]	RPPFTYATLIRWAILEAPEKQRTLNEIYHWFRMFAPFRNHDPATWKN AIRHNL SLHKCFVRVSEKGA VVTVDEFFFRKKR
FOXA1	[<i>H. sapiens</i>]	KMLTLSEIYQWIMDLFPYRQNRQQRWQNSIRHLSLNFDCPVKVARSPDKPGKGSYWTLHPDSGNMFENG CYLRQKRFK
FOXA1	[<i>M. musculus</i>]	KMLTLSEIYQWIMDLFPYRQNRQQRWQNSIRHLSLNFDCPVKVARSPDKPGKGSYWTLHPDSGNMFENG CYLRQKRFK
FOXA2	[<i>H. sapiens</i>]	KMLTLSEIYQWIMDLFPYRQNRQQRWQNSIRHLSLNFDCFLKVPSPDKPGKGSYWTLHPDSGNMFENG CYLRQKRFK
FOXA3	[<i>M. musculus</i>]	KMLTLSEIYQWIMDLFPYRQNRQQRWQNSIRHLSLNFDCPVKVARSPDKPGKGSYWTLHPDSGNMFENG CYLRQKRFK
FOXA1	[<i>L. pumilus</i>]	KMLTLSEIYQWIMDLFPYRQNRQQRWQNSIRHLSLNFDCPVKVARSPDKPGKGSYWTLHPDSGNMFENG CYLRQKRFK
FOXA1	[<i>X. tropicalis</i>]	KMLTLSEIYQWIMDLFPYRQNRQQRWQNSIRHLSLNFDCPVKVARSPDKPGKGSYWTLHPDSGNMFENG CYLRQKRFK
FOXQ1	[<i>H. sapiens</i>]	MAIRDSAGGRRLTAEINEYLMGKFPFRGSGYTGWRNSVRHNL SLNDCPVKVL RDPSPRWGKDN YWMLNPNSEYTFADGVFRRRRKR
FOXQ1	[<i>M. musculus</i>]	MAIRDSAGGRRLTAEINEYLMGKFPFRGSGYTGWRNSVRHNL SLNDCPVKVL RDPSPRWGKDN YWMLNPNSEYTFADGVFRRRRKR
FOXQ1	[<i>O. latipes</i>]	MAIRDSAGGRRLTADINIDYLMGKFPFRGSGYTGWRNSVRHNL SLNDCFLKVL RDPSPRWGKDN YWMLNPNSEYTFADGVFRRRRKR
FOXQ2	[<i>P. infanta</i>]	MAILASPEKRLLLCDIYQIQENYPYRNDRSWRNSIRHNL SLNECFIKYGRSDN--GRGNFWAIIHPANVEDFSRGDFIIRRRAR
FOXQ2	[<i>D. rerio</i>]	MAILDSDEKLLLCDIYQWIMDIYPTKSKDKNWRNSVRHNL SLNECFIKYGRSDN--GKGHFWAIIHPANTQDFNSGDIYHRRRAR
FOXQ2	[<i>S. faecalis</i>]	MAILASERKLLLCDIYQWIMDIYPTKSKDKNWRNSVRHNL SLNECFIKYGRSDN--GKGHFWAIIHPANIQDFSKGDIYHRRRAR

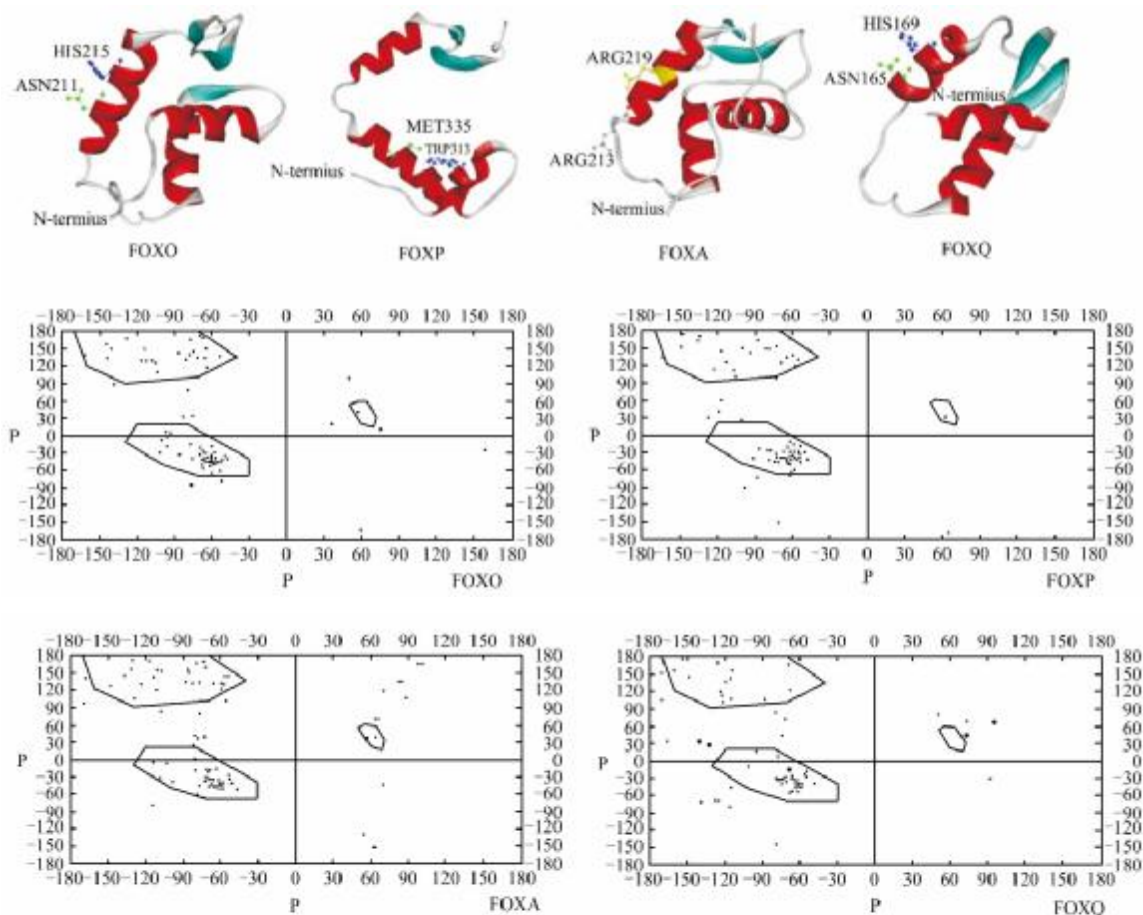


图4 FOX家族保守区域比对及三级结构建模

注: 上图为氨基酸保守序列比对; 中图为保守序列三级结构建模; 下图为拉式构象分布图

Figure 4 The Multi-alignment of amino acid sequences of FOX and 3-D structure of FOX established by homology-based modeling

Note: The amino acid sequence alignment is on the top; 3-D structure of FOX is in the middle; The Ramachandran Diagram is on the below

- ba- sis for DNA recognition by FoxO1 and its regulation by posttranslational modification, *Structure*, 16(9): 1407-1416
- Brown G.R., Hem V., Katz K.S., Ovetsky M., Wallin C., Ermolaeva O., Tolstoy I., Tatusova T., Pruitt K.D., Maglott D.R., and Murphy T.D., 2015, Gene: a gene-centered information resource at NCBI, *Nucleic Acids Res.*, 43(Database issue): D36-D42
- Chang K., Chen M., Zeng L., Lan X., Wang Q., and Liao Z., 2014, Abscisic acid enhanced ajmalicine biosynthesis in hairy roots of *Rauvolfia verticillata* by upregulating expression of the MEP pathway genes, *Russian Journal of Plant Physiology*, 61(1): 136-140
- Chang K., Qiu F., Chen M., Zeng L., Liu X., Yang C., Lan X., Wang Q., and Liao Z., 2014, Engineering the MEP pathway enhanced ajmalicine biosynthesis, *Biotechnol. Appl. Biochem.*, 61(3): 249-255
- Csuros M., 2013, How to infer ancestral genome features by parsimony: dynamic programming over an evolutionary tree, *Computational Biology*, pp.29-45
- Geourjon C., and Deléage G., 1995, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, *Comput. Appl. Biosci.*, 11(6): 681-684
- Hilgier W., Oja S.S., Saransaari P., and Albrecht J., 2005, Taurine prevents ammonia-induced accumulation of cyclic GMP in rat striatum by interaction with GABAA and glycine receptors, *Brain Res.*, 1043(1-2): 242-246
- Jin C., Marsden I., Chen X., and Liao X., 1999, Dynamic DNA contacts observed in the NMR structure of winged helix protein-DNA complex, *J. Mol. Biol.*, 289(4): 683-690
- Johansson M.U., Zoete V., Michielin O., and Guex N., 2012, Defining and searching for structural motifs using DeepView/Swiss-PdbViewer, *BMC Bioinformatics*, 13: 173
- Kontny E., Rudnicka W., Chora Xzy-Massalska M., Marcinkiewicz J., and Maslinski W., 2006, Taurine chloramine inhibits proliferation of rheumatoid arthritis synoviocytes by triggering a p53-dependent pathway, *Inflamm. Res.*, 55(10): 446-455
- Lan X., Chang K., Zeng L., Liu X., Qiu F., Zheng W., Quan H., Liao Z., Chen M., Huang W., Liu W., and Wang Q., 2013, Engineering salidroside biosynthetic pathway in hairy root cultures of *Rhodiola crenulata* based on metabolic characterization of tyrosine decarboxylase, *PLoS One*, 8(10): e75459
- Lu G., and Moriyama E.N., 2004, Vector NTI, a balanced all-in-one sequence analysis suite, *Brief Bioinform.*, 5(4): 378-388
- Maccarrone G., and Filiou M.D., 2015, Protein profiling and phosphoprotein analysis by isoelectric focusing, *Methods in Molecular Biology*, 1295: 293-303
- Marsden I., Jin C., and Liao X., 1998, Structural changes in the region directly adjacent to the DNA-binding helix highlight a possible mechanism to explain the observed changes in the sequence-specific binding of winged helix proteins, *J. Mol. Biol.*, 278(2): 293-299
- Nandymazumdar M., and Artsimovitch I., 2015, Ubiquitous transcription factors display structural plasticity and diverse functions, *Bioessays*, 37(3): 324-334
- Ponugoti B., Xu F., Zhang C., Tian C., Pacios S., and Graves D.T., 2013, FOXO1 promotes wound healing through the up-regulation of TGF-beta1 and prevention of oxidative stress, *J. Cell Biol.*, 203(2): 327-343
- van Dongen M.J., Cederberg A., Carlsson P., Enerback S., and Wikstrom M., 2000, Solution structure and dynamics of the DNA-binding domain of the adipocyte-transcription factor FREAC-11, *J. Mol. Biol.*, 296(2): 351-359
- Wilhelmsen K., Litjens S.H., Kuikman I., Tshimbalanga N., Janssen H., van den Bout I., and Raymond K., and Sonnenberg A., 2007, Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin, *Journal of Cell Biology*, 171(5): 799-810
- Wu H., Jin Y., Wei J., Jin H., Sha D., and Wu J.Y., 2005, Mode of action of taurine as a neuroprotector, *Brain Res.*, 1038(2): 123-131