

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 吴艾霖 陈清红 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

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 之间, 分子量介于 29 084 到 79 586.1 之间, 较为离散。GRAVY 值介于 -0.764 到 -0.179 之间, 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*

Key Laboratory of High Humidity Medicine, Center of Laboratory Medicine, Chengdu Military General Hospital, Chengdu, 610083

* Corresponding author, wulijuan1638@126.com

DOI: 10.13417/j.gab.034.001818

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 built. The result showed that the similarity of nucleic acid sequence was low. Aliphatic amino acid index of protein is between 29 084 to 79 586.1; GRAVY values ranged from -0.764 to -0.179, and the protein is hydrophilic. The isoelectric point was stable in the FOX subfamily and could 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 might 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 and Artsimovitch, 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 蛋白均具有高保守的 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

基金项目: 本研究由国家自然科学基金(81302596)资助

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。

本研究在 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 et al., 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(表 1),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 et al., 2013)。对 FOX 家族基因进行内含子与外显子分析,可见 FOXD、FOXQ、FOXL 和 FOXC 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 2 000 bp 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(图 1) 基于 Neighbor-Joining 的原理建树(Lan et al., 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

(Csürös, 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 et al., 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 et al., 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 蛋白进行比对分析(Lu and Moriyama, 2004) 进行 FOX

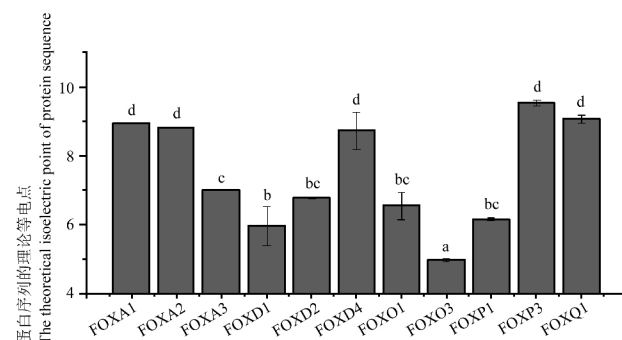


图 2 FOX 蛋白序列的等电点统计分析

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 ($p < 0.01$)

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

基因名	登录号	外显子	染色体	启动子预测
Gene name	Access number	Exon region	Chromosome	Promoter predictor
H sapiens fox O1	NC_023244	1~630; 105 351~106 690	Chr 13	tctagcattttaaaaaaggfcaagaataacagfgaatecggAtgsgggcgtg
H sapiens fox O3	NC_000006.12	1~621; 102 245~103 647	Chr 6	cgcTgcctcgtttaaaaggcggcgccctcccctccggcGccccctccc
H sapiens fox A1	NC_000014	1~73; 2 260~3 608	Chr 14	tcataaagataataaacgggtgctgactcacctctcttAgccgcagggt
H sapiens fox A2	NC_000020.11	1~88; 1 119~2 429	Chr 20	cTgacttttttaaaaaagggtgggggtgggggtgattGcttggtcgttt
		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		
H sapiens fox P1	NC_000003.12	215 897~216 112; 203 291~203 494; 205 015~205 194; 1~169; 202 029~	Chr 3	aaaaatacacaaaaaattagccggsgcatggtggcatgacCtgrtaatct
		202 166; 227 411~227 531; 218 170~218 279; 108 104~108 196; 225 682~		
		225 783; 144 294~144 372		
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~	Chr 7	agTgaattttaaaaaaggcTgtatagTcatccatggTtGfcaaaaggta
		6 842; 1 487~1 581; 2 707~2 791; 2 994~3 078; 5 301~5 381		
H sapiens fox P3	NC_000023	10 735~11 230; 534~999; 13 147~13 467; 10 360~10 569; 5 122~5 255;	Chr 11	aggTcaactgctaaaaaggTggcagTactggggatcTgaatgcAgTccaagat
H sapiens fox N1	NC_000017.11	1~123; 3 281~3 392; 6 767~6 876	Chr 17	TgagaactgTgcattttatTggggcaccggactTgTgaccaggcActgtgctga
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	accTggTcctataaaagcTgcactTttattacagccattTcAactaata
H sapiens fox D1	NC_000005.10	1~1 398	Chr 5	ggcagggattataaaagTcggcgcgcgagactccggccAcccggcagc
H sapiens fox Q1	NC_000006	1~1 212	Chr 6	ggcgcagggttataTggggcggccgaaaggTggaggcgaGacgcagccg
H sapiens fox L1	NC_000016.10	1~1 038	Chr 16	TcagTgctttataTtggTgcccggaaattTtccaaAataaacag
H sapiens fox L2	NC_000003.12	1~1 131	Chr 3	agTcactTtttataaacggccccgcacagccTgagaccggccTgccccgcc
H sapiens fox C1	NC_000006.12	1~1 662	Chr 6	TaacTgcgTaaaaaaagTcctcctgcctggggTgacggatgctAaaagTtcag

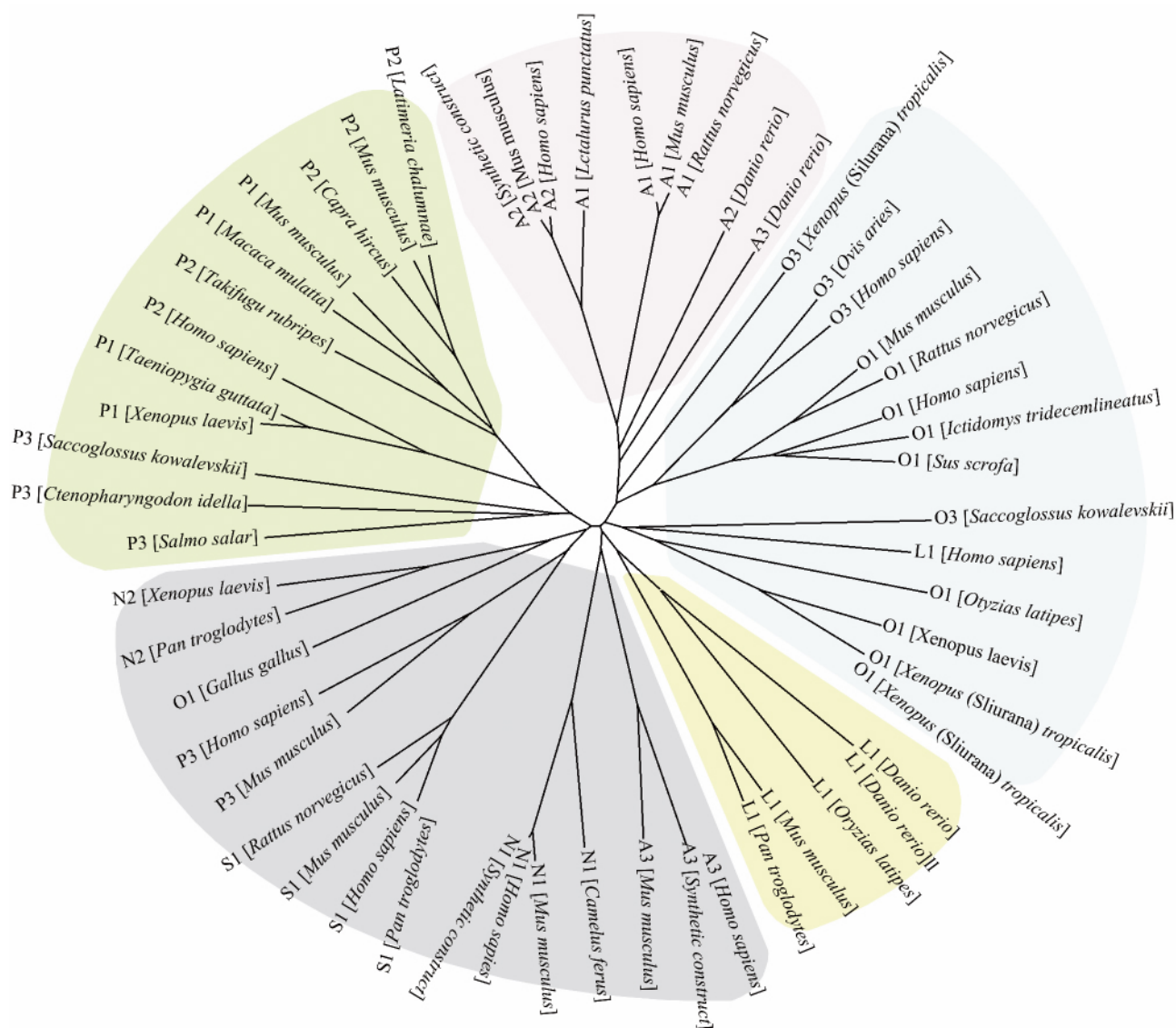


图 1 基于 MEGA5.0 软件平台用 Neighbor-Joining 方法构建的 FOX 进化树
注: 来源于 FOXP 家族的一支在绿色框内; FOXA 家族的一支在红色框内; FOXO 家族的一支在蓝色框内; FOXL 家族的一支在黄色框内

Figure 1 A phylogenetic tree of FOX from different organisms constructed by Neighbor-Joining method on MEGA 5.0
Note: FOXP showed in green box; FOXA showed in red box; FOXO showed in blue box; FOXL showed in yellow box

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 33.6%，相似性为 8.4%；FOXQ 同源性为 54%，相似性为 13.6%；FOXD 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(图 2)表明 FOXO1 包含 25.8%的 α -螺旋、11.91%延伸带、6.56% β 转角和 55.73%随意卷曲构成；FOXO3 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 and Deléage,

1995)。上述数据提示 FOX 家族的氨基酸序列均可预测出典型的 $\alpha 1-\beta 1-\alpha 2-\alpha 3-\beta 2-W 1-\beta 3-W 2$ 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-MODEL 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,结合 Swiss-PdbViewer 拉式构象图进行蛋白构象合理性评估(Biasini et al., 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 et al., 2012)。

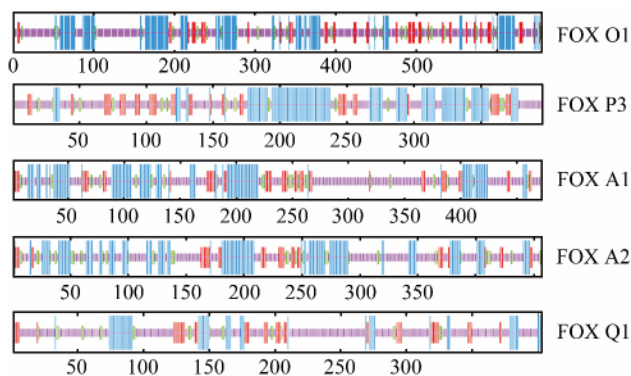


图 3 FOX 蛋白二级结构预测

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

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'-TTGTTTTG-3')结合,保守功能位点为 Ans 211 和 His 215 (Brent et al., 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 et al., 2011)。同样,FOXA 的功能位点为 Arg213 和 Arg219,FOXQ 的功能位点为 Asn165 和 His169 (Jin et al., 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 Dongen et al., 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 et al., 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 et al., 2005)。FOX 蛋白也参与 Akt/PKB 信号转导,可通过联合作用打破细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 et al., 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 et al., 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 et al., 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 and Filiou, 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。

FOXO1 [*H. sapiens*] RRNAWGNLSYADLITKAIESSAEKRLTLSQIYEWVMKSVYPYFKDKGDSNSSAGWKNRHSIRHNSLHSHKCFIRVQNEGTGKSSWWMLNPEG
 FOXO4 [*M. musculus*] RRNAWGNQSYAELISQAIESAPEKRLTLAQIYEWVVRTVPYFKDKGDSNSSAGWKNRHSIRHNSLHSHKCFIKVHNEATGKSSWWMLNPDG
 FOXO3 [*D. rerio*] RRNAWGNLSYADLITKAIESTPDKRLTLSQIYDWMVSSVPYFKDKGDSNSSAGWKNRHSIRHNSLHSHFRIRVQNEGTGKSSWWMLNPEG
 FOXO3 [*M. musculus*] RRNAWGNLSYADLITRAIESSPDKRLTLSQIYEWVMVRCVPYFKDKGDSNSSAGWKNRHSIRHNSLHSHFRMRVQNEGTGKSSWWMLNPDG
 FOXO6 [*P. alecto*] RRNAWGNLSYADLITKAIESAAPDKRLTLSQIYDWMVRYVPYFKDKGDSNSSAGWKNRHSIRHNSLHSHFRIRVQNEGTGKSSWWMLNPEG

FOXP3 [*H. sapiens*] RPPFTYATLIRWAILEAPEKQRTLNEIYHWFTRMFAFFRNHPATWKNAIRHNSLHSHKCFVRVESEKGAVVTVDEFEFRKKR
 FOXP1 [*P. alecto*] RPPFTYASLIRQAILEPEKQRTLNEIYNWFRMFAFFRNAAATWKNAVRHNSLHSHKCFVRVENVKGAVVTVDEFEFQKRR
 FOXP1 [*R. norvegicus*] RPPFTYASLIRQAILEPEKQRTLNEIYNWFRMFAFFRNAAATWKNAVRHNSLHSHKCFVRVENVKGAVVTVDEFEFQKRR
 FOXP4 [*R. norvegicus*] RPPFTYASLIRQAILETPDRQLTNEIYNWFRMFAFFRNATWKNAVRHNSLHSHKCFVRVENVKGAVVTVDEREYQKRR
 FOXP3 [*B. taurus*] RPPFTYATLIRWAILEAPEKQRTLNEIYHWFTRMFAFFRNHPATWKNAIRHNSLHSHKCFVRVESEKGVVTVDEFEFRKKR
 FOXP3 [*P. alecto*] RPPFTYATLIRWAILEAPEKQRTLNEIYHWFTRMFAFFRNHPATWKNAIRHNSLHSHKCFVRVESEKGAVVTVDEFEFRKKR

FOXA1 [*H. sapiens*] KMLTLSEIQWIMDLFPYYRQNRQQRWQNSIRHLSFNDFCVKVARSPDKPGKGSYWTLHPDSGNMFENGCYLRRQKRFK
 FOXA1 [*M. musculus*] KMLTLSEIQWIMDLFPYYRQNRQQRWQNSIRHLSFNDFCVKVARSPDKPGKGSYWTLHPDSGNMFENGCYLRRQKRFK
 FOXA2 [*H. sapiens*] KMLTLSEIQWIMDLFPYYRQNRQQRWQNSIRHLSFNDFCVKVARSPDKPGKGSYWTLHPDSGNMFENGCYLRRQKRFK
 FOXA3 [*M. musculus*] KMLTLSEIQWIMDLFPYYRENQNRQQRWQNSIRHLSFNDFCVKVARSPDKPGKGSYWTLHPDSGNMFENGCYLRRQKRFK
 FOXA1 [*P. punctatus*] KMLTLSEIQWIMDLFPYYRQNRQQRWQNSIRHLSFNDFCVKVARSPDKPGKGSYWTLHPDSGNMFENGCYLRRQKRFK
 FOXA1 [*X. tropicalis*] KMLTLSEIQWIMDLFLYYRQNRQQRWQNSIRHLSFNDFCVKVARSPDKPGKGSYWTLHPDSGNMFENGCYLRRQKRFK

FOXQ1 [*H. sapiens*] MAIRDSAGRLTLAEINEYLMGKFPFFRGSYTGWRNSVRHNSLNDKCFVKVLRDPSRPWGKDNWMLNPSEYTFADGVFRRRRKR
 FOXQ1 [*M. musculus*] MAIRDSAGRLTLAEINEYLMGKFPFFRGSYTGWRNSVRHNSLNDKCFVKVLRDPSRPWGKDNWMLNPSEYTFADGVFRRRRKR
 FOXQ1 [*O. latipes*] MAIRDSAGRLTLADINDYLMKFPFFRGSYTGWRNSVRHNSLNDKCFVKVLRDPSRPWGKDNWMLNPQSEYTFADGVFRRRRKR
 FOXQ2 [*P. miniata*] MAILASPEKRLLLCDIYQIENYPYRNNDRS WRNSIRHNSLNECFIKYGRSGD---GRGNFWAVHPANVEDFSRQDFHRRRR
 FOXQ2 [*D. rerio*] MAILDSDEKLLLCDIYQWIMDHYPFYKSKDKNWRNSVRHNSLNECFIKAGRSDN---GKGHFVAIHPANFQDFSNQDYHRRRRR
 FOXQ2 [*S. formosus*] MAILASEEKLLLCDIYQWIMDHYPFYKSKDKNWKNSVRHNSLNECFIKAGRSDN---GKGHFVAIHPANIQDFSKGQDYHRRRRR

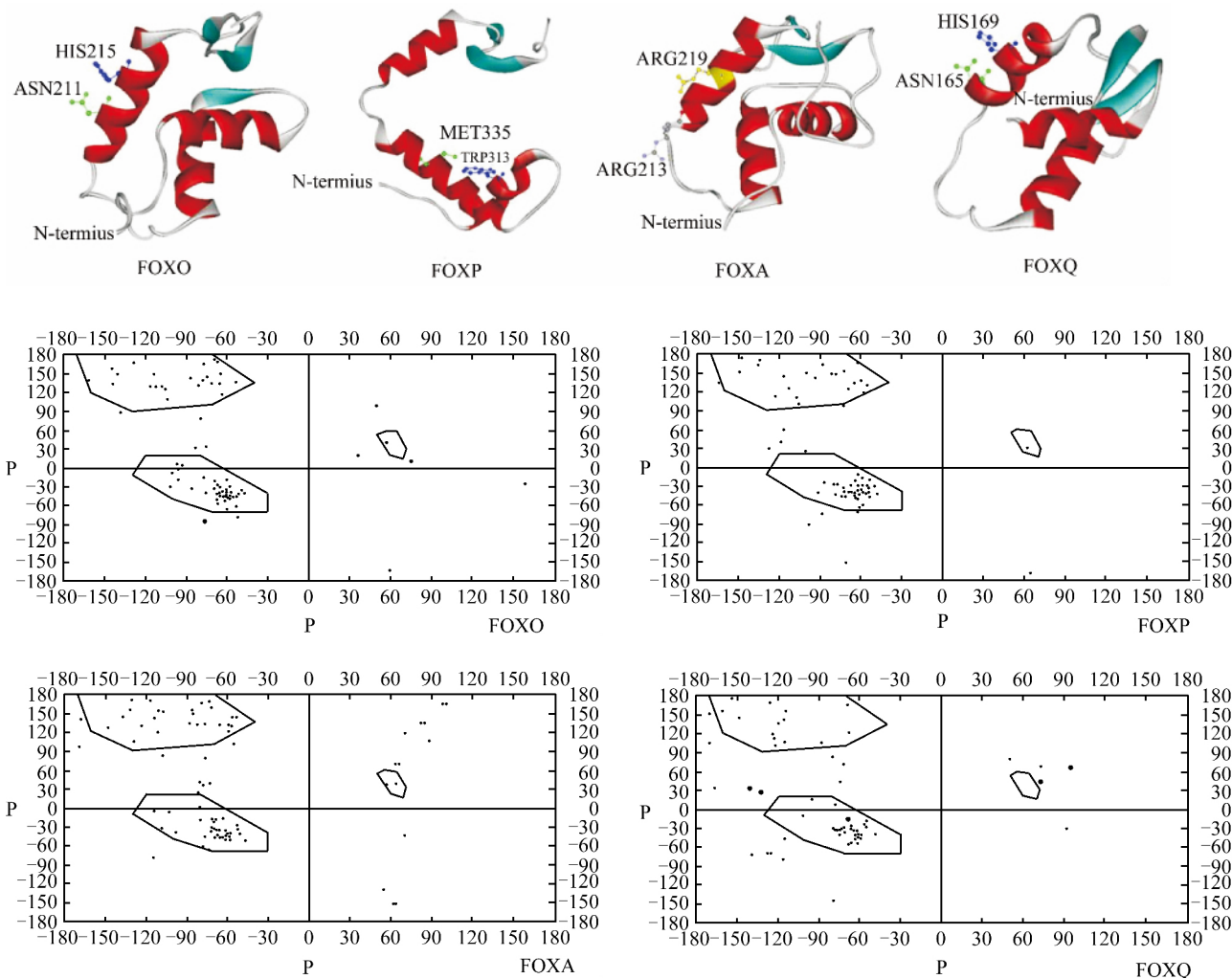


图4 FOX 家族保守区域比对及三级结构建模

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

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 TGFβ 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 basis 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 *Rauwolfia 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

Csürös 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., ChoraXzy-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 Wikström 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

=====

Journal of Mosquito Research (JMR)



Journal of Mosquito Research (ISSN 1927-646X) is an open access, peer reviewed journal published online by BioPublisher. The journal is publishing high quality original research on all aspects of mosquito and its affecting the living organisms, as well as environmental risk and public policy relevant to mosquito modified organisms. Topics include (but are not limited to) the research at molecular or protein level of mosquito, impact on the ecosystem, containing positive and negative information, natural history of mosquito, also publishing innovative research findings in the basic and applied fields of mosquito and novel techniques for improvement, as well as the significant evaluation of its related application field. Authors from low-income countries and areas or having financial difficulties can apply the discount for article processing charges (APC) upon

submission, and if the authors are qualified to be waived they don't necessarily pay the APC. We do not want fees to prevent the publication of worthy work.

Email: edit@jmr.biopublisher.ca

Web: <http://jmr.biopublisher.ca>