

文章编号: 1674-5566(2009)04-0403-06

奥利亚罗非鱼 DMRT1 基因 推导蛋白的结构和功能预测

曹谨玲¹, 陈剑杰¹, 吴婷婷²

(¹ 山西农业大学动物科技学院山西省生态畜牧与环境兽医重点实验室, 山西 太谷 030801;

² 中国水产科学研究院淡水渔业研究中心农业部水生动物遗传育种
和养殖生物学重点开放实验室, 江苏 无锡 214081)

摘要: 利用生物信息学的方法对 DMRT1 基因推导的氨基酸序列进行结构特征和功能域预测分析, 探讨了 DMRT1 的信号肽、亲疏水性、跨膜拓扑结构、卷曲螺旋结构、基序、功能域及高级结构。结果表明: DMRT1 基因推导蛋白不含螺旋卷曲区域, 没有信号肽, 是一个非跨膜的亲水性稳定蛋白, 该蛋白以游离形式存在于细胞质内, 不会发生跨膜运动。DMRT1 基因推导蛋白包含两个相同的保守的功能结构域, 分别行使性别调控, 使 DNA 形成二聚体和结合回纹结构的功能。DMRT1 基因推导蛋白含有多个磷酸化位点, 推测它可能在细胞信号传导中发挥作用, 且其生物活性可能接受信号途径中多种信号的调控。DMRT1 的高级结构中含有两个 α -螺旋区域。以上结论为实验室研究奥利亚罗非鱼 DMRT1 基因编码蛋白的功能, 为明确 DMRT1 基因与性别调控的关系奠定了基础。

关键词: DMRT1; 生物信息学; 蛋白质结构; 蛋白质功能; 性别调控

中图分类号: S 917 **文献标识码:** A

Structural and functional prediction of DMRT1 deduced protein in *Oreochromis aureus*

CAO Jin-ling¹, CHEN Jian-jie¹, WU Ting-ting²

(¹ Key Lab of Ecological Animal Husbandry and Environmental Veterinary Medicine of
Shanxi Province Shanxi Agricultural University Taigu 030801, China;

² Key Open Laboratory for Genetic Breeding of Aquatic Animals and Aquaculture Biology Ministry of
Agriculture Freshwater Fishery Research Center Chinese Academy of Fishery Sciences Wuxi 214081, China)

Abstract: DMRT1 gene was analyzed by bioinformatics methods on amino acid level and the structural characteristics and function domains of its encoded protein were analyzed and predicted in *Oreochromis aureus*. Meantime the signal peptide, hydrophilicity, hydrophobicity, transmembrane topological structure, wined helical structure, motif, functional domains and advanced structure of DMRT1 were also discussed. The results showed that DMRT1 had no helical segment and did not contain signal peptide. It was a non-transmembrane and hydrophilic protein. DMRT1 included two same functional domains which played the roles of sex control, dimerising and binding palindromic DNA, respectively. DMRT1 included several

收稿日期: 2008-09-09

基金项目: 国家“八六三”高科技研究发展计划 (2001AA243061); 国家自然科学基金 (30371116); 无锡市自然科学基金 (CK030001)

作者简介: 曹谨玲 (1979-), 女, 山西太谷人, 讲师, 博士, 主要从事水生动物遗传育种与繁殖方面的研究。Tel: 15935664795,

E-mail: caojinling7928@163.com

phosphorylation sites implicating that they could play some roles during cellular signal conduct and their activities might be related to the regulation of many signals during signal route. DMRT¹ had advanced structures including two α -helix regions. This conclusions are helpful for studying the functions of DMRT¹ protein in *Oreochromis aureus* and the relations between sex control and them.

Key words: DMRT¹; bioinformatics; protein structure; protein function; sex control

DMRT¹ (DM-related transcription factor 1) 基因被认为是一种性别决定和性别分化基因, 存在于鸡的 Z 染色体^[1], 人的常染色体 9p24.7^[2] 以及鱼类的常染色体上^[3-4]。国外有关鱼类 DMRT 基因研究工作主要集中在青鳉^[3, 5-8], 尼罗罗非鱼^[9], 虹鳟^[4, 10] 以及河豚上^[11-12]。罗非鱼是原产于非洲的热带鱼类, 现已成为世界性的主要养殖鱼类。罗非鱼雄性比雌性生长快 40%~50%, 因此提高罗非鱼的雄性率具有重要的理论与实践意义。本研究在已获得的奥利亚罗非鱼 (*Oreochromis aureus*) DMRT¹ 氨基酸序列的基础上, 运用生物信息学方法对其蛋白质的结构和功能进行了预测和分析, 为实验室研究奥利亚罗非鱼 DMRT¹ 基因编码蛋白的功能, 明确它们与性别调控的关系奠定了基础。

1 材料与方法

1.1 材料

DMRT¹ 的氨基酸序列测定参见笔者所在实验室的研究^[13]。

1.2 方法

按照 <http://www.us.expasy.org>, <http://www.ch.embnet.org>, DNASTar 及 NCBI 上提供的核酸、蛋白质在线工具进行分析。氨基酸组成分析利用 DNASTar 的 EDITSEQ 工具; 二级结构预测利用 SOPMA 分析工具; 三级结构预测利用 SWISS-MODEL 和 CPHmodels-2.0 Server 工具; 信号肽预测利用 signalP 分析工具; 跨膜结构分析利用 TMPRED; 蛋白质的亲疏水性分析利用 DNASTar 的 Protean 工具; 卷曲螺旋结构利用 COILS 工具; 基序分析利用 PROSITE 工具; 功能结构域分析利用 NCBI CD-Search service 工具。

2 结果

2.1 DMRT¹ 的氨基酸序列结构组成特征分析

采用 DNASTar 程序对 DMRT¹ 基因编码蛋白质结构进行分析, 结果表明: DMRT¹ 基因编码的蛋白质由 292 个氨基酸组成 (图 1), 分子量为 31.58 kDa, 理论等电点 pI 为 8.04, 极性氨基酸 (N, C, Q, S, T, Y) 占 38.36%, 酸性氨基酸 (K, R) 占 9.59%, 碱性氨基酸 (D, E) 占 8.21%, 疏水性氨基酸 (A, I, L, F, W, V) 占 23.28%。

1	MSQDKQSKQV	PDCSGPMSPT	KAQKSPRMPK	CSRQRNHGYV	SPLKGHKRFC	NWRDCQCPKC
61	KLIAERQRVM	AAQVALRRQQ	AQEEELGICS	PVSLSGSEMM	VKNEVGADCL	FSVEGRSPTP
121	TSHPTSAVTG	TRSASSPSPS	AAARAHTEGP	SDLLLETPYY	NFYQPSRYPT	YYGNLYNYSQ
181	YQMPHGDGRL	PSHSVSSQYR	MHSYYPAAATY	LTQQLGSTSC	VPPFFSLDDN	NNSCSETMAA
241	SFSPSSISA	GHDSTMVCRS	ISSLVNGDAK	AECEASSQAA	GFTVDAIEGG	ATK

图 1 DMRT¹ 基因的氨基酸序列

Fig. 1 amino acid sequence of DMRT¹ gene

2.2 DMRT¹ 信号肽预测

信号肽本身由氨基酸组成, 通常位于蛋白 N 一端, 由可被剪切的 15~30 氨基酸的前导序列组成, 在

信号肽的 N-端或靠近 N-端处通常有 2~3 个极性氨基酸,而在信号肽的中部都是一个唯一的疏水核心或者由很多的疏水氨基酸构成。信号肽没有其它保守序列,也没有酸性基团。当被转运的蛋白质肽链的合成在 mRNA 起始密码上启动后,首先合成一段信号肽,它被内质网膜上的受体蛋白所识别并相互结合,此时受体蛋白可聚合而产生膜内通道。此外,内质网膜上核糖体受体蛋白还可与核糖体结合形成延伸出去的通道。这样,当蛋白质肽链合成继续进行的同时,信号肽率先经过通道进入内质网,随后信号肽被膜内侧的信号肽酶水解掉,蛋白质再变成成熟蛋白质^[14]。故预测和分析信号肽有助于蛋白质功能域的划分和蛋白质细胞定位。由图 2 可知,DMRT1 在第 23、30 和 35 氨基酸残基处存在信号肽酶切位点的几率仅仅分别为 0.585、0.055 和 0.040,说明该蛋白不存在信号肽酶切位点,无信号肽,是一种非分泌蛋白。从而可以推测 DMRT1 在细胞质中合成并以可溶性状态保留在原处,直接作用于胞液中可溶性的代谢物。

2.3 DMRT1 跨膜结构域的预测

跨膜结构域一般富含疏水性氨基酸残基,起着固系于细胞膜中的“抛锚”作用,具有跨膜结构域的蛋白属于跨膜蛋白类。跨膜结构域的预测对正确认识蛋白质的结构、功能与细胞定位具有重要意义。利用 Tmpred 对 DMRT1 跨膜域进行预测^[15]。结果表明(图 3),DMRT1 没有跨膜结构,属于非跨膜蛋白类,说明 DMRT1 不发生跨膜运动,而是完全在细胞内起作用。

2.4 DMRT1 亲疏水性分析预测

蛋白质二级和三级结构预测时必须分析蛋白质的亲疏水性,通过了解整条肽链中不同肽段的亲疏水性,可以对一些膜蛋白中跨膜的肽段进行预测。不仅可以为蛋白质次级结构提供参考,还可以为结构域和功能域的划分提供依据。在细胞质中合成的新生肽链,有相当一部分被传送并定位到细胞内的不同细胞器上,或被分泌到细胞外,折叠为特定空间构象的肽链,表明带有大量的亲水基团,虽然在细胞质中很容易被传送,但是不能通过脂质构成的细胞器膜。

由图 4 可知,DMRT1 在第 1~58, 75~87, 114~123, 129~150, 156~206, 226~236 区域和 265~277 区域为亲水性的,余下的少数氨基酸为疏水性的。由此可见,亲水性残基所占比例远大于疏水性残基,因此可以推测 DMRT1 是亲水性的。

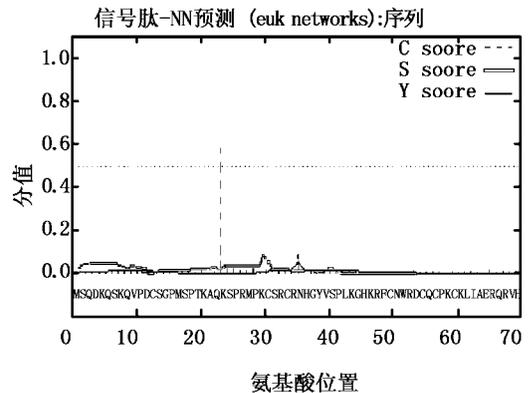


图 2 DMRT1 信号肽预测

Fig 2 Prediction of the signal peptide of DMRT1

C score: 原始剪切位点的分值; S score: 信号肽的分值; Y score: 综合剪切点的分值

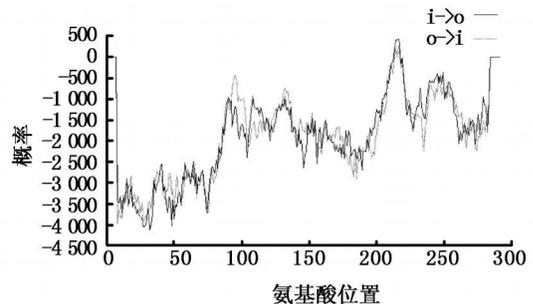


图 3 DMRT1 氨基酸跨膜区段的预测

Fig 3 Prediction of the transmembrane amino acid segment of DMRT1

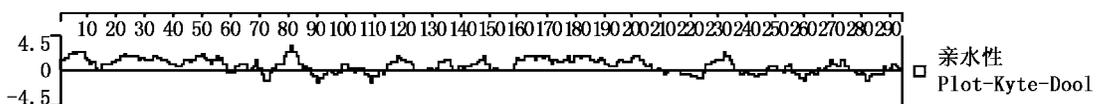


图 4 DMRT1 亲疏水性预测分析

Fig 4 Hydrophilicity/hydrophobicity analysis of DMRT1

2.5 DMRT1卷曲螺旋结构的预测

活性高,易溶解的氨基酸区段易形成螺旋结构,正确预测 DMRT1卷曲螺旋结构可为揭示蛋白质高级结构提供参考价值。Coils工具用于专一性分析左手螺旋形成的可能性。利用 Coils分析工具,对 DMRT1基因的氨基酸序列形成卷曲螺旋的倾向性进行预测,以 window = 14 为试验参数,按照几率 > 50%就可形成螺旋的规则,比较权重设定和不设定两种情况下的分析结果一致,得出结论:DMRT1不存在卷曲螺旋结构(图 5)。

2.6 DMRT1功能基序分析

Prosite数据库查询结果表明:DMRT1可能含 1 个蛋白激酶 C磷酸化位点(25~27),5 个肉豆蔻酰化位点(130~135, 173~178, 214~219, 216~221, 245~250),3 个酪蛋白激酶 II 磷酸化位点(95~98, 226~229, 233~236),2 个 N-糖基化位点(177~180, 231~234)。提示该蛋白可能在细胞信号传导中发挥作用,它们的生物活性可能接受信号途径中多种信号的调控;它们都含有酪蛋白激酶 II 磷酸化位点,这个位点的磷酸化/去磷酸化可能是调节其活性的方式。

2.7 DMRT1蛋白质结构域分析

由图 6 可知,DMRT1 含有 DM-domain 而 DM-domain 是果蝇的性别决定基因 Doublesex (dsx)和线虫的性别决定基因 Mab-3 所编码的蛋白质都包含的一个新的具有 DNA 结合能力的保守基序,并由此通过相似的机制控制着性别的分化发育。因此,提示 DMRT1 与性别调控有关。

2.8 DMRT1的高级结构预测

2.8.1 DMRT1蛋白的二级结构预测

由图 7 可以看出,DMRT1 蛋白含有 19.18%的 α -螺旋,4.11%的 β -转角和 62.33%的无规卷曲。

2.8.2 DMRT1蛋白的三级结构预测

采用 CPHmodels-2.0 Server 蛋白质三级结构预测软件对 DMRT1 的三级结构进行预测(图 8),可以看到 DMRT1 蛋白质三级结构的基本轮廓及使其稳定的主要次级键。为了更清晰地展示 DMRT1 蛋白的三级结构,进一步采用 SWISS MODEL 对其三级结构进行预测(图 9),可以看到 DMRT1 蛋白具有两个 α -螺旋区域。由图 8、9 可知,构成 DMRT1 的长短轴相差不多,基本上呈球形,属于球状蛋白,疏水基多聚在分子的内部,而亲水基多分布在分子表面,因此说明该蛋白是亲水的,这进一步验证了 DMRT1 亲疏水性分析预测结果的正确性。

3 讨论

随着生命科学实验技术和生物信息学的飞速发展和完善,利用功能强大的生物分析软件,对基因和蛋白质的结构和功能进行分析、预测,反映了未来生命科学发展的趋势。蛋白质的组成、结构决定其功能。通过研究蛋白质序列结构特征来阐明蛋白质的功能和蛋白质细胞定位是科学的。利用核酸/蛋白质分析软件,从分子水平对 DMRT1 氨基酸序列的结构特征进行分析,预测结果表明:DMRT1 蛋白不含螺旋区域,没有信号肽,是一个非跨膜的亲水性稳定蛋白,该蛋白以游离形式存在于细胞质内,不会发生

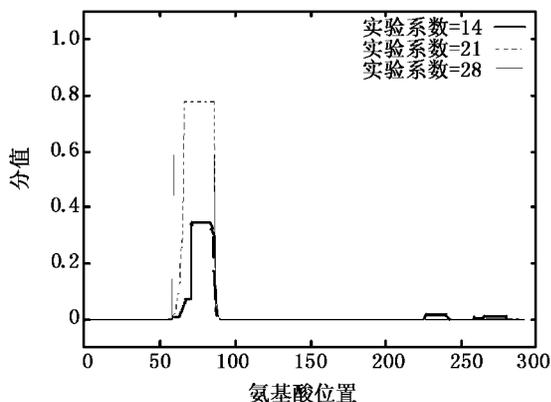


图 5 DMRT1基因的氨基酸序列卷曲螺旋的预测(不设定权重和设定权重)

Fig 5 Predicted amino acid sequence of the wound helix of DMRT1 in *O. rostrum* is aureus (not by weight and by weight)

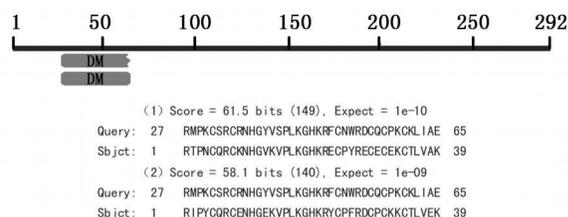
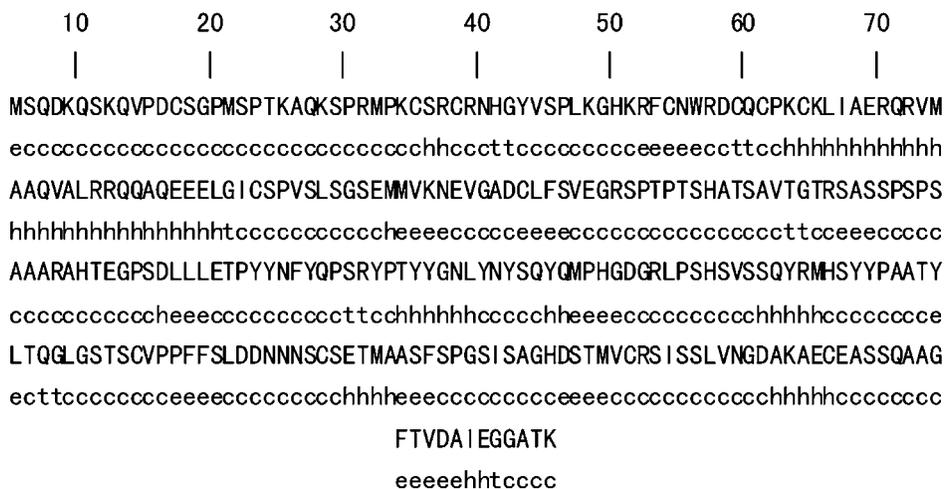


图 6 DMRT1结构域分析

Fig 6 Protein domain analysis of DMRT1

跨膜运动;包含两个保守的功能结构域,分别行使性别调控,使 DNA 形成二聚体和结合回纹结构的功能。利用 SCANP_{rosite}软件对该蛋白质可能的修饰位点进行分析,发现其含有多个位点可以发生磷酸化修饰,个别位点可以发生糖基化修饰。蛋白质的磷酸化和脱磷酸化在细胞信号传导中起重要作用。磷酸化、糖基化位点可能是 DMRT1蛋白完成其生理功能的结构基础。特别是该蛋白质序列中有多个酪蛋白激酶 II 磷酸化位点,表明酪蛋白激酶 II (CKII)对其生物功能的发挥具有重要的调节作用。



Alpha helix (Hh): 19.18%; Extended strand (Ee): 14.38%; Beta turn (Tt): 4.11%; Random coil (Cc): 62.33%

图 7 DMRT1蛋白二级结构预测

Fig 7 prediction on secondary structure of DMRT1

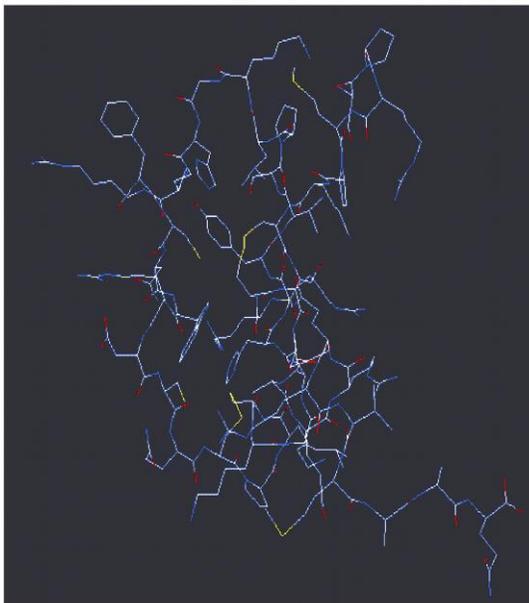


图 8 DMRT1的三级结构 (CPHmodels-2.0 Server)

Fig 8 Tertiary structure of DMRT1 (CPHmodels-2.0 Server)

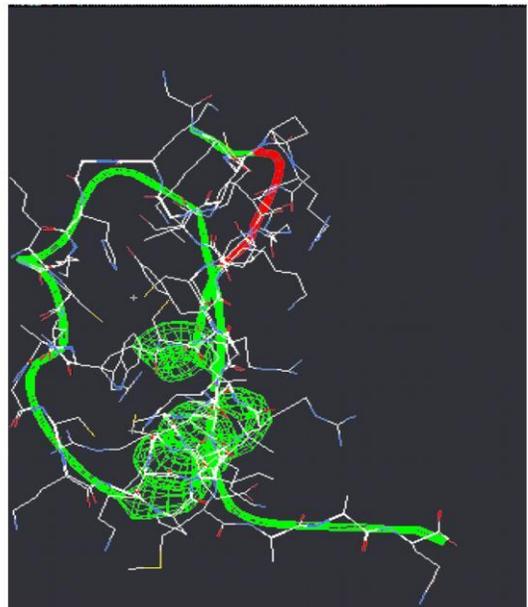


图 9 DMRT1的三级结构 (SWISS-MODEL)

Fig 9 Tertiary structure of DMRT1 (SWISS-MODEL)

利用以上生物信息学的序列分析方法,对奥利亚罗非鱼 DMRT1基因进行了分析,提供了 DMRT1基因功能预测信息,为其高级结构、功能研究提供了有益的指导,为系统研究奥利亚罗非鱼 DMRT1基因的性别调控机理提供参考。

参考文献:

- [1] Ellegren H. Hens cocks and avian sex determination: a quest for genes on Z or W [J]. *EMBO reports* 2001, 2(3): 192—196.
- [2] Flejter W L, Ferguson J, Gorski I, et al. A gene involved in XY sex reversal is located on chromosome 9, distal to marker D9S1779 [J]. *Hum Genet* 1998, 63: 794—802.
- [3] Kondo M, Froschauer A, Kitano A, et al. A molecular cloning and characterization of DMRT genes from the medaka *Oryzias latipes* and the platyfish *Xiphophorus maculatus* [J]. *Genes Dev* 2002, 16(18): 2390—2392.
- [4] Sakamoto T, Danzmann R G, Ghabib K, et al. A microsatellite linkage map of rainbow trout (*Oncorhynchus tshawytscha*) [J]. *Cytogenet Cell Genet* 2001, 92: 108—110.
- [5] Nanda L, Kondo M, Homung U, et al. A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka *Oryzias latipes* [J]. *Proc Natl Acad Sci USA* 2002, 99(18): 11778—11783.
- [6] Ohmuro-Matsuyama Y, Matsuda M, Kobayashi T, et al. Expression of DMY and DMRT1 in Various Tissues of the Medaka (*Oryzias latipes*) [J]. *Zool Sci* 2003, 20(12): 1395—1398.
- [7] Kobayashi T, Matsuda M, Kajima-Kobayashi H, et al. Two DM domain genes, DMY and DMRT1, involved in testicular differentiation and development in the medaka *Oryzias latipes* [J]. *Dev Dyn* 2004, 231(3): 518—526.
- [8] Winkler C, Homung U, Kondo M, et al. Developmentally regulated and non-sex-specific expression of autosomal dmrt genes in embryos of the Medaka fish (*Oryzias latipes*) [J]. *Mechanisms of Development* 2004, 121(7—8): 997—1005.
- [9] Guan G J, Kobayashi T, and Nagahama Y. Sexually dimorphic expression of two types of DM (Doublesex/Mab-3) domain gene in a teleost fish, the tilapia (*Oreochromis niloticus*) [J]. *Biochemical and Biophysical Research in Communications* 2000, 273(3): 662—666.
- [10] Marchand O, Gouvarou M, Cotta H, et al. DMRT1 expression during gonadal differentiation and spermatogenesis in the rainbow trout *Oncorhynchus mykiss* [J]. *Biochim et Biophys Acta* 2000, 1493(1—2): 180—187.
- [11] Brunner B, Homung U, Shan Z, et al. Genomic Organization and Expression of the Doublesex-Related Gene Cluster in Vertebrates and Detection of Putative Regulatory Regions for DMRT1 [J]. *Genomics* 2001, 77(1—2): 8—17.
- [12] Yamaguchi A, Lee K H, Fujimoto H, et al. Expression of the DMRT gene and its roles in early gonadal development of the Japanese pufferfish *Takifugu rubripes* [J]. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 2006, 1(1): 59—68.
- [13] 曹谨玲. 奥利亚罗非鱼 DMO 和 DMT 基因的克隆及分子生物学特征和功能分析 [D]. 无锡: 南京农业大学博士论文, 2007: 67—75.
- [14] Nielsen H, Engelbrecht J, Brunak S, et al. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites [J]. *Protein Eng* 1997, 10: 1—6.
- [15] Emanuelsson O, Nielsen H, Brunak S, et al. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence [J]. *Mol Biol* 2000, 300(4): 1005—1016.