

郭 敏,李祥龙. 不同物种 *TYR* 基因编码蛋白结构及功能的生物信息学分析[J]. 江苏农业科学,2017,45(21):44-48.  
doi:10.15889/j.issn.1002-1302.2017.21.012

# 不同物种 *TYR* 基因编码蛋白结构及功能的生物信息学分析

郭 敏,李祥龙

(河北科技师范学院动物科技学院,河北秦皇岛 066004)

**摘要:**利用生物信息学方法对 13 个物种 *TYR* 基因编码蛋白的理化性质,以及一级、二级、三级结构和亚细胞定位进行预测和分析,同时拟构建 13 个物种 *TYR* 基因的系统发育树。结果表明:*TYR* 基因编码产物为不稳定蛋白(貉子除外),整条多肽链表现出亲水性。二级结构主要为  $\alpha$  螺旋、 $\beta$  折叠, $\beta$  转角区域分布少且散,均存在信号肽和跨膜区,可以判断 *TYR* 基因编码蛋白是定位于生物膜上的膜蛋白或分泌蛋白,主要在嘌呤和嘧啶、运输和结合、电压门控离子通道及转录中发挥作用。各物种 *TYR* 基因编码产物主要定位于内质网、高尔基体以及细胞质膜。除原鸡外,其他物种该蛋白均存在跨膜区,貉子有 2 个跨膜区域。各物种系统发育树与动物学进化观点基本一致。

**关键词:**物种;酪氨酸酶基因;毛色;蛋白质结构;生物信息学

**中图分类号:** Q754 **文献标志码:** A **文章编号:** 1002-1302(2017)21-0044-05

动物的毛色、羽色不但可以作为质量性状,也可作为一种经济性状被利用。目前,绿色、环保、优质的毛皮产品已成为广大消费者追求的目标。毛色表型主要受黑色素的影响,是黑色素在毛皮质和髓质中沉淀的种类和数量不同造成的。研究表明,动物毛色的形成主要与黑色素细胞中合成的 2 种色素(真黑色素、褐黑色素)比例有关。真黑色素通常可使皮肤或毛皮表现为褐色、黑色,褐黑色素能使皮肤或毛皮表现为黄

色、红棕色<sup>[1]</sup>,黑色素沉淀取决于这 2 种色素的相对含量。对于动物黑色素细胞而言,酪氨酸酶(tyrosinase,简称 TYR)作为关键酶参与黑色素的合成<sup>[2]</sup>,具有加氧酶、脱氢酶双重功能<sup>[3]</sup>,动物黑色素的形成主要遵循多巴(DOPA)途径,指在动物体内,酪氨酸在 TYR 的催化作用下,经由 DOPA、多巴醌、吲哚醌等形式,最后生成黑色素。酪氨酸酶相关蛋白酶 1(TYRP1)、酪氨酸酶相关蛋白酶 2(TYRP2)与酪氨酸酶(TYR)构成了酪氨酸酶基因家族,TYRP1、TYRP2 各自催化特异的反应参与黑色素的合成,这 3 种成员协同调控黑色素的形成<sup>[4]</sup>,并且与其他毛色相关基因共同作用调控毛色的形成。引起黑色素代谢紊乱的主要原因是 *TYR* 基因突变或酶的失活,最终将导致机体组织异常表达<sup>[5]</sup>。有研究表明,哺乳动物 *TYR* 基因包含 5 个外显子和 4 个内含子,序列长度约为 1.6 kb<sup>[6]</sup>。本研究选用 GenBank 中已提交物种的 *TYR* 基因编码蛋白序列,利用生物信息学方法对 *TYR* 基因编码蛋白的结构与功能进行预测和分析,以期研究 *TYR* 基因及其编码

收稿日期:2016-06-01

基金项目:河北省高校创新团队领军人才培养计划(编号:LJRC004);河北省应用基础研究计划重点基础研究项目(编号:15962901D);河北省自然科学基金重点项目(编号:C2016407114)。

作者简介:郭 敏(1991—),女,内蒙古呼和浩特人,硕士研究生,研究方向为功能基因组学。E-mail:1908384796@qq.com。

通信作者:李祥龙,博士,教授,博士生导师,研究方向为动物遗传育种。E-mail:lixianlongen@yahoo.com。

中,具有很好的理论研究和实际应用价值。

## 参考文献:

- [1]祝仲纯,吴海珊. 从萝卜的下胚轴诱导植株成功[J]. 遗传,1980,2(3):36.
- [2]崔群香,汪隆植,娄 平,等. 从萝卜下胚轴再生完整植株[J]. 南京农专学报,1999,15(4):29-36.
- [3]Jeong W J, Min S R, Liu J R. Somatic embryogenesis and plant regeneration in tissue culture of radish (*Raphanus sativus* L.) [J]. Plant Cell Reports,1995,14(10):648-651.
- [4]Pua E C, Sim G E, Chi G L, et al. Synergistic effect of ethylene inhibitors and putrescine on shoot regeneration from hypocotyl explants of Chinese radish (*Raphanus sativus* L. var. *longipinnatus* Bailey) in vitro[J]. Plant Cell Reports,1996,15(9):685-690.
- [5]武 剑,龚义勤,邓 波,等. 萝卜雄性不育组织培养的研究[J]. 贵州农业科学,2003,31(5):8-11.
- [6]武 剑,龚义勤,邓 波,等. 萝卜离体再生的影响因素[J]. 中国蔬菜,2003(6):6-8.
- [7]李海萍,张鲁刚,张 静,等. 萝卜带柄子叶高频再生体系的建立[J]. 植物学报,2011,46(3):331-337.
- [8]毛倩卓. 侵染萝卜的 dsRNA 病毒研究——带病毒萝卜的组织培养、细胞和分子生物学研究[D]. 杭州:浙江理工大学,2011:12-15.
- [9]熊秋芳,张雪清,骆海波,等. 萝卜组织培养的研究与应用[J]. 长江蔬菜,2006(5):36-38.
- [10]吴丽艳,柏柯帆,李石开,等. 利用花序轴组培快繁青花菜萝卜胞质雄性不育系的研究[J]. 云南农业大学学报(自然科学版),2009,24(5):712-716.
- [11]赖正峰,林加耕,李华东,等. 日本高秆青花菜组织培养初探(简报)[J]. 亚热带植物科学,2005,34(1):64.

蛋白的功能,进而为该基因调控的动物毛色机制研究奠定基础。

1 材料与方法

1.1 序列来源

从美国国立生物技术信息中心(NCBI)网站(<http://www.ncbi.nlm.nih.gov/>)的 GenBank 中下载包括斑马鱼、犬、东北虎、家兔、家猫、绵羊、貉子、家牛、人、狨猴、原鸡、中央蝙蝠、家猪等 13 个物种的 *TYR* 基因编码蛋白序列<sup>[7]</sup>。序列来源见表 1。

表 1 不同物种 *TYR* 基因编码蛋白序列

物种	登录号
犬( <i>Canis lupus familiaris</i> )	AAQ17535.1
家猫( <i>Felis catus</i> )	XP_003992691.1
家牛( <i>Bos taurus</i> )	NP_851344.1
人( <i>Homo sapiens</i> )	NP_000363.1
家猪( <i>Sus scrofa</i> )	NP_001020383.1
家兔( <i>Oryctolagus cuniculus</i> )	NP_001075546.1
绵羊( <i>Ovis aries</i> )	NP_001123499.1
貉子( <i>Nyctereutes procyonoides procyonoides</i> )	AIQ86602.1
狨猴( <i>Callithrix jacchus</i> )	XP_002754751.1
原鸡( <i>Gallus gallus</i> )	XP_015135298.1
斑马鱼( <i>Danio rerio</i> )	NP_571088.2
东北虎( <i>Panthera tigris altaica</i> )	XP_007079066.1
中央蝙蝠( <i>Pteropus alecto</i> )	XP_006926765.1

1.2 蛋白结构分析

1.2.1 一级结构 利用 Expasy 开发的针对蛋白质基本理化性质的 Protparam 工具(<http://www.expasy.org/tools/protparam.html>)分析 *TYR* 基因编码蛋白序列的相对分子质量、氨基酸组成、等电点(pI)、消光系数、半衰期、不稳定系

数、总平均亲水性等。同时通过 Expasy 开发的 Protscale 工具对 *TYR* 基因编码的蛋白质进行亲水性、疏水性分析。利用 TMHMM Server v. 2.0 在线软件(<http://www.cbs.dtu.dk/services/TMHMM-2.0/>)预测跨膜区<sup>[8-9]</sup>。

1.2.2 二级结构 利用 Chou & Fasman Secondary Structure Prediction Server 软件预测 *TYR* 基因所编码蛋白质的  $\alpha$  螺旋、 $\beta$  折叠和  $\beta$  转角。同时,分别使用 Protfun 2.2 Server<sup>[10-11]</sup>、Signal P 4.1 Server<sup>[12]</sup>在线软件分析预测 *TYR* 基因编码产物的功能分类和信号肽。

1.2.3 亚细胞定位 利用 PSORT II Prediction 在线软件(<http://psort.hgc.jp/form<sup>2</sup>.html>)对 *TYR* 基因编码蛋白在细胞内的位置进行定位与分析<sup>[13]</sup>。

1.3 构建系统发育树

应用 DNAMAN 软件对不同物种 *TYR* 基因编码蛋白序列进行多序列比对并构建系统发育树<sup>[14]</sup>。

2 结果与分析

2.1 不同物种 *TYR* 基因编码蛋白的一级结构

2.1.1 蛋白质理化性质 由表 2 可知,13 个物种 *TYR* 基因编码蛋白的氨基酸组成相同,且亮氨酸(Leu)、丝氨酸(Ser)含量明显高于其他氨基酸。表 3 为 13 个物种 *TYR* 基因编码蛋白的理化参数,包括氨基酸数量、相对分子质量、pI 值、分子式、原子总数、正电荷残基数/负电荷残基数、不稳定指数、脂肪系数和亲水性总平均值等。可以看出,脂肪系数均高达 70 以上,其中貉子的脂肪系数最高,达到 81.39;亲水性总平均值表现为负值,说明这 13 个物种 *TYR* 基因编码的蛋白均为亲水蛋白;除貉子(不稳定指数 < 40,稳定)外,其他物种 *TYR* 基因编码的蛋白均为不稳定蛋白质。

表 2 不同物种 *TYR* 基因编码蛋白的氨基酸组成

物种	氨基酸种类及比例(%)																			
	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
犬	5.7	5.3	5.5	5.7	3.6	3.8	5.3	6.8	3.0	4.3	10.4	3.2	2.5	5.8	6.4	8.1	4.7	2.6	4.0	3.4
家猫	6.2	5.7	5.1	5.8	3.4	4.0	5.1	6.4	3.0	4.5	10.4	3.2	2.6	5.8	6.2	8.1	4.5	2.6	4.0	3.2
家牛	6.0	5.7	4.9	6.0	3.2	3.8	5.1	6.6	2.8	3.6	10.4	2.8	2.3	5.8	6.8	8.5	4.7	2.5	4.5	4.0
人	5.7	5.1	5.1	5.7	3.2	4.3	5.1	6.4	3.2	4.2	10.4	3.2	2.8	5.9	6.2	9.6	3.6	2.6	4.3	3.2
家猪	5.3	5.6	5.1	5.5	3.4	4.0	5.5	7.2	2.6	4.3	10.4	3.0	2.4	6.2	6.8	8.7	4.3	2.4	4.1	3.2
家兔	5.5	5.5	5.3	5.5	3.2	4.0	5.3	6.8	3.0	3.6	10.6	2.8	2.6	5.8	6.2	8.9	4.7	2.6	4.2	4.0
原鸡	5.0	6.6	6.6	4.4	4.2	4.2	5.4	7.0	2.2	5.0	8.8	2.0	3.2	6.0	6.0	8.2	4.4	2.0	4.4	4.2
绵羊	5.8	5.8	4.9	5.7	3.2	3.6	5.3	6.8	2.8	3.6	10.6	2.8	2.3	5.8	6.8	8.5	4.7	2.5	4.5	4.0
貉子	5.7	4.2	5.6	5.2	2.5	2.7	6.0	6.0	2.5	4.9	8.4	5.9	2.9	4.6	4.9	8.8	5.8	1.4	3.9	8.1
狨猴	6.0	4.5	4.9	5.7	3.2	4.3	5.1	6.4	3.0	4.5	10.4	3.6	2.8	5.8	6.2	9.8	3.6	2.6	4.2	3.2
斑马鱼	6.6	7.1	4.5	5.1	2.8	5.4	4.7	7.1	2.6	4.1	9.2	1.3	2.2	6.0	5.8	8.6	4.5	2.2	4.3	5.8
东北虎	6.2	5.6	5.1	5.6	3.4	4.0	5.1	6.4	3.2	4.1	10.4	3.2	2.8	5.8	6.2	8.3	4.5	2.6	4.0	3.4
中央蝙蝠	4.9	5.3	5.8	5.3	3.0	4.1	5.5	7.0	3.0	4.7	9.8	3.0	2.4	6.0	6.6	8.3	4.9	2.6	4.3	3.4

注:Ala 为丙氨酸,Arg 为精氨酸,Asn 为天冬酰胺,Asp 为天冬氨酸,Cys 为半胱氨酸,Gln 为谷氨酰胺,Glu 为谷氨酸,Gly 为甘氨酸,His 为组氨酸,Ile 为异亮氨酸,Leu 为亮氨酸,Lys 为赖氨酸,Met 为甲硫氨酸,Phe 为苯丙氨酸,Pro 为脯氨酸,Ser 为丝氨酸,Thr 为苏氨酸,Trp 为色氨酸,Tyr 为酪氨酸,Val 为缬氨酸。下同。

2.1.2 跨膜区分析 通过软件分析可知:貉子 *TYR* 基因编码产物有 2 个跨膜区域,原鸡 *TYR* 基因编码产物无跨膜区域,第 1~499 aa 位于胞外。对 13 个物种进行跨膜区预测结果显示,除原鸡外,其他物种该蛋白均存在跨膜区(表 4),揭示该蛋白是定位于生物膜的跨膜蛋白。

2.2 不同物种 *TYR* 基因编码蛋白的二级结构

2.2.1  $\alpha$ -螺旋、 $\beta$ -折叠、 $\beta$ -转角的预测 经过软件分析可知:在各物种 *TYR* 基因编码蛋白的二级结构中, $\alpha$ -螺旋、 $\beta$ -折叠区域分布较广,而  $\beta$ -转角区域分布少且较分散(表 5)。

表 3 不同物种 *TYR* 基因编码蛋白的理化性质

物种	氨基酸数量(个)	相对分子质量	pI 值	分子式	原子总数(个)	正电荷残基数/ 负电荷残基数	不稳定指数	脂肪系数	亲水性总平均值	稳定性
犬	530	60 336.3	5.67	C <sub>2705</sub> H <sub>4060</sub> N <sub>726</sub> O <sub>785</sub> S <sub>32</sub>	8 308	(Asp + Glu)/(Arg + Lys) = 58/45	53.08	72.91	-0.335	不稳定
家猫	530	60 477.5	5.82	C <sub>2712</sub> H <sub>4079</sub> N <sub>731</sub> O <sub>783</sub> S <sub>32</sub>	8 337	(Asp + Glu)/(Arg + Lys) = 58/47	54.72	73.66	-0.330	不稳定
家牛	530	60 278.0	5.54	C <sub>2706</sub> H <sub>4050</sub> N <sub>724</sub> O <sub>789</sub> S <sub>29</sub>	8 298	(Asp + Glu)/(Arg + Lys) = 59/45	53.44	71.98	-0.348	不稳定
人	529	60 393.2	5.71	C <sub>2709</sub> H <sub>4051</sub> N <sub>725</sub> O <sub>787</sub> S <sub>32</sub>	8 304	(Asp + Glu)/(Arg + Lys) = 57/44	56.76	71.76	-0.356	不稳定
家猪	531	60 454.4	5.64	C <sub>2714</sub> H <sub>4070</sub> N <sub>726</sub> O <sub>787</sub> S <sub>31</sub>	8 328	(Asp + Glu)/(Arg + Lys) = 58/46	55.31	71.85	-0.344	不稳定
家兔	530	60 361.2	5.67	C <sub>2705</sub> H <sub>4055</sub> N <sub>727</sub> O <sub>788</sub> S <sub>31</sub>	8 306	(Asp + Glu)/(Arg + Lys) = 57/44	52.05	72.15	-0.340	不稳定
原鸡	499	56 940.6	6.03	C <sub>2530</sub> H <sub>3824</sub> N <sub>694</sub> O <sub>737</sub> S <sub>37</sub>	7 822	(Asp + Glu)/(Arg + Lys) = 49/43	57.18	71.14	-0.303	不稳定
绵羊	530	60 304.1	5.69	C <sub>2709</sub> H <sub>4060</sub> N <sub>726</sub> O <sub>786</sub> S <sub>29</sub>	8 310	(Asp + Glu)/(Arg + Lys) = 58/46	53.83	72.53	-0.341	不稳定
貉子	972	109 022.3	6.15	C <sub>4863</sub> H <sub>7546</sub> N <sub>1294</sub> O <sub>1451</sub> S <sub>52</sub>	15 206	(Asp + Glu)/(Arg + Lys) = 109/98	37.79	81.39	-0.224	稳定
狨猴	530	60 222.1	5.59	C <sub>2705</sub> H <sub>4054</sub> N <sub>716</sub> O <sub>787</sub> S <sub>32</sub>	8 294	(Asp + Glu)/(Arg + Lys) = 57/43	53.22	73.47	-0.308	不稳定
斑马鱼	534	60 584.4	6.07	C <sub>2711</sub> H <sub>4088</sub> N <sub>748</sub> O <sub>785</sub> S <sub>27</sub>	8 359	(Asp + Glu)/(Arg + Lys) = 52/45	57.55	75.24	-0.265	不稳定
东北虎	531	60 590.7	5.94	C <sub>2715</sub> H <sub>4082</sub> N <sub>734</sub> O <sub>783</sub> S <sub>33</sub>	8 347	(Asp + Glu)/(Arg + Lys) = 57/47	56.03	72.60	-0.336	不稳定
中央蝙蝠	531	60 700.5	5.67	C <sub>2728</sub> H <sub>4076</sub> N <sub>730</sub> O <sub>792</sub> S <sub>29</sub>	8 355	(Asp + Glu)/(Arg + Lys) = 57/44	51.37	71.28	-0.385	不稳定

表 4 不同物种 *TYR* 基因编码蛋白的跨膜区域预测结果

物种	跨膜区数量(个)	跨膜区氨基酸位点范围	胞内区氨基酸位点范围	胞外区氨基酸位点范围
斑马鱼	1	478 ~ 500	501 ~ 534	1 ~ 477
东北虎	1	474 ~ 496	497 ~ 531	1 ~ 473
犬	1	474 ~ 496	497 ~ 530	1 ~ 473
家兔	1	477 ~ 499	500 ~ 530	1 ~ 476
家猫	1	474 ~ 496	497 ~ 530	1 ~ 473
绵羊	1	474 ~ 496	497 ~ 530	1 ~ 473
貉子	2	517 ~ 539;649 ~ 671	540 ~ 648	1 ~ 516;672 ~ 972
家牛	1	474 ~ 496	497 ~ 530	1 ~ 473
人	1	474 ~ 496	497 ~ 529	1 ~ 473
狨猴	1	474 ~ 496	497 ~ 530	1 ~ 473
原鸡	0			1 ~ 499
中央蝙蝠	1	477 ~ 499	500 ~ 531	1 ~ 476
家猪	1	474 ~ 496	497 ~ 531	1 ~ 473

表 5 不同物种 *TYR* 基因编码蛋白的  $\alpha$ -螺旋、 $\beta$ -折叠、 $\beta$ -转角的预测结果

物种	$\alpha$ -螺旋区氨基酸位点范围	$\beta$ -折叠区氨基酸位点范围	$\beta$ -转角区	总残留量(aa)
斑马鱼	102 ~ 144;195 ~ 225;378 ~ 403;440 ~ 518	1 ~ 23;386 ~ 403;468 ~ 484;490 ~ 505	分布少且散	H;328;E;247;T;69
东北虎	114 ~ 142;265 ~ 289;374 ~ 417;454 ~ 523	80 ~ 107;210 ~ 225;385 ~ 401;474 ~ 500	分布少且散	H;317;E;230;T;68
犬	114 ~ 142;232 ~ 250;316 ~ 339;452 ~ 499	80 ~ 111;210 ~ 225;385 ~ 401;474 ~ 500	分布少且散	H;282;E;227;T;69
家兔	114 ~ 142;232 ~ 250;374 ~ 417;454 ~ 522	1 ~ 14;80 ~ 90;366 ~ 377;474 ~ 499	分布少且散	H;287;E;224;T;68
家猫	114 ~ 144;232 ~ 250;374 ~ 417;454 ~ 522	80 ~ 107;210 ~ 225;385 ~ 401;474 ~ 499	分布少且散	H;309;E;232;T;68
绵羊	114 ~ 142;187 ~ 217;165 ~ 280;383 ~ 417	80 ~ 98;162 ~ 182;210 ~ 225;385 ~ 410	分布少且散	H;284;E;235;T;69
貉子	184 ~ 256;305 ~ 350;502 ~ 558;588 ~ 642; 673 ~ 720;873 ~ 925	190 ~ 222;301 ~ 324;389 ~ 432;503 ~ 559; 633 ~ 677;758 ~ 778	分布少且散	H;617;E;622;T;133
家牛	114 ~ 142;187 ~ 207;265 ~ 280;383 ~ 417	80 ~ 98;210 ~ 225;385 ~ 410;459 ~ 499	分布少且散	H;279;E;236;T;70
人	114 ~ 145;232 ~ 250;374 ~ 417;460 ~ 521	170 ~ 184;265 ~ 276;385 ~ 401;474 ~ 499	分布少且散	H;297;E;223;T;67
狨猴	114 ~ 145;168 ~ 188;374 ~ 417;460 ~ 522	80 ~ 107;366 ~ 377;385 ~ 401;485 ~ 501	分布少且散	H;289;E;231;T;67
原鸡	124 ~ 151;273 ~ 299;382 ~ 425;458 ~ 485	88 ~ 120;174 ~ 189;215 ~ 233;433 ~ 486	分布少且散	H;258;E;322;T;70
中央蝙蝠	114 ~ 142;265 ~ 280;364 ~ 417;454 ~ 530	80 ~ 107;385 ~ 400;459 ~ 567;574 ~ 599	分布少且散	H;289;E;225;T;71
家猪	114 ~ 142;265 ~ 280;374 ~ 417;454 ~ 523	80 ~ 106;210 ~ 225;385 ~ 400;474 ~ 500	分布少且散	H;276;E;230;T;75

2.2.2 信号肽的预测 通过分析蛋白序列有无信号肽,可以判断此蛋白是否为分泌蛋白<sup>[15]</sup>。表 6 的预测结果显示,原始剪切位点分值(raw cleavage site score,简称 C 值)的最大值与被结合的剪切位点的分值(combined cleavage site score,简称 Y 值)的最大值均趋向于 + 1,信号肽分值(signal peptide score,简称 S 值)的最大值在切割位点前高,在切割位点后

低,符合信号肽的标准,可以判断 *TYR* 基因编码产物中存在信号肽。

2.2.3 亲水性与疏水性的预测 亲水性氨基酸一般位于蛋白分子表面,而疏水性氨基酸位于蛋白质分子的内部<sup>[16]</sup>。氨基酸的亲水、疏水性可以用来判断蛋白质折叠的大致趋势。对 13 个物种 *TYR* 氨基酸序列的亲水、疏水性进行预测,正值

表 6 不同物种 *TYR* 基因编码蛋白的信号肽预测结果

物种	C 值的 最大值	S 值的 最大值	Y 值的 最大值	切割位点	是否存在 信号肽
斑马鱼	0.564	0.966	0.727	18 ~ 19	是
东北虎	0.683	0.932	0.773	18 ~ 19	是
犬	0.708	0.923	0.775	18 ~ 19	是
家兔	0.376	0.920	0.550	17 ~ 18	是
家猫	0.682	0.933	0.774	18 ~ 19	是
绵羊	0.599	0.878	0.705	17 ~ 18	是
貉子	0.599	0.878	0.705	17 ~ 18	是
家牛	0.811	0.971	0.865	25 ~ 26	是
人	0.658	0.919	0.760	17 ~ 18	是
狨猴	0.658	0.921	0.760	17 ~ 18	是
原鸡	0.801	0.986	0.864	26 ~ 27	是
中央蝙蝠	0.625	0.882	0.714	18 ~ 19	是
家猪	0.661	0.901	0.756	17 ~ 18	是

表疏水,负值表亲水;正值越大表示疏水性越强,负值越小表示亲水性越强,表 7 结果显示,13 个物种整条多肽链均表现亲水性,因此认为,*TYR* 基因编码产物为亲水蛋白,且位于蛋白质分子表面。

表 7 不同物种 *TYR* 基因编码蛋白的疏/亲水性分析结果

物种	氨基酸 最低得分	氨基酸 最高得分	多肽链 疏水性
斑马鱼	-3.122	2.756	亲水性
东北虎	-3.222	3.067	亲水性
犬	-3.033	3.067	亲水性
家兔	-2.756	2.544	亲水性
家猫	-3.222	3.067	亲水性
绵羊	-2.856	2.644	亲水性
貉子	-2.789	3.589	亲水性
家牛	-2.856	2.644	亲水性
人	-3.178	2.911	亲水性
狨猴	-3.033	2.911	亲水性
原鸡	-2.944	3.022	亲水性
中央蝙蝠	-3.189	2.989	亲水性
家猪	-3.011	2.644	亲水性

2.2.4 功能分类的预测 由表 8 可以看出,不同物种 *TYR* 基因编码蛋白的功能一致,并且在嘌呤和嘧啶、运输和结合、电压门控离子通道及转录中发挥作用的概率较大,分别为 0.331、0.773、0.279、0.219。

表 8 不同物种 *TYR* 基因编码蛋白功能分类的预测结果

功能分类	概率												
	斑马鱼	东北虎	狗	家兔	绵羊	猫	貉子	牛	人	狨猴	原鸡	中央蝙蝠	猪
氨基酸生物合成	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
生物合成辅酶因子	0.210	0.210	0.210	0.210	0.210	0.210	0.210	0.210	0.210	0.210	0.210	0.210	0.210
细胞被膜	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033	0.033
细胞加工	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030
中央中间代谢	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.048
能量代谢	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035
脂肪酸代谢	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017
嘌呤和嘧啶	0.331	0.331	0.331	0.331	0.331	0.331	0.331	0.331	0.331	0.331	0.331	0.331	0.331
监管职能	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034
复制和转录	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020
翻译	0.071	0.071	0.071	0.071	0.071	0.071	0.071	0.071	0.071	0.071	0.071	0.071	0.071
运输和结合	0.773	0.773	0.773	0.773	0.773	0.773	0.773	0.773	0.773	0.773	0.773	0.773	0.773
信号转导	0.205	0.205	0.205	0.205	0.205	0.205	0.205	0.205	0.205	0.205	0.205	0.205	0.205
受体	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
结构蛋白	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
转运载体	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025
离子通道	0.169	0.169	0.169	0.169	0.169	0.169	0.169	0.169	0.169	0.169	0.169	0.169	0.169
电压门控离子通道	0.279	0.279	0.279	0.279	0.279	0.279	0.279	0.279	0.279	0.279	0.279	0.279	0.279
阳离子通道	0.146	0.146	0.146	0.146	0.146	0.146	0.146	0.146	0.146	0.146	0.146	0.146	0.146
转录	0.219	0.219	0.219	0.219	0.219	0.219	0.219	0.219	0.219	0.219	0.219	0.219	0.219
转录调控	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111
胁迫应答	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073
免疫应答	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
金属离子转移	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018

2.3 不同物种 *TYR* 基因编码蛋白的亚细胞定位

如表 9 所示,各物种 *TYR* 基因编码产物主要定位于内质网、高尔基体以及细胞质膜,其中人的该基因编码产物有小部分定位于细胞外(11.1%),而原鸡的该基因编码产物主要定位于细胞外,比例高达 66.7%,小部分定位于内质网、高尔基体及液泡中,且比例均为 11.1%。

2.4 不同物种 *TYR* 基因编码蛋白的三级结构

对不同物种 *TYR* 基因编码蛋白的三维结构进行同源建模,结果表明,各物种蛋白结构丰富多样,存在大量折叠、螺旋扭曲,这些特点对物种的生物学功能具有重大意义。

2.5 系统发育树

由图 1 可知,东北虎与家猫的亲缘关系最近,其次是狗、

表 9 不同物种 *TYR* 基因编码蛋白的亚细胞定位预测结果

物种	各亚细胞定位的比例(%)				
	内质网	高尔基体	细胞质膜	细胞外 (包括细胞壁)	液泡
犬	33.3	33.3	33.3	—	—
家猫	33.3	33.3	33.3	—	—
家牛	33.3	33.3	33.3	—	—
人	22.2	33.3	33.3	11.1	—
家猪	33.3	33.3	33.3	—	—
家兔	44.4	33.3	22.2	—	—
原鸡	11.1	11.1	—	66.7	11.1
绵羊	33.3	33.3	33.3	—	—
貉子	33.3	33.3	33.3	—	—
狨猴	33.3	33.3	33.3	—	—
斑马鱼	44.4	33.3	22.2	—	—
东北虎	33.3	33.3	33.3	—	—
中央蝙蝠	33.3	33.3	33.3	—	—

注:表中数据为 k 近邻法(k - NN)预测结果。

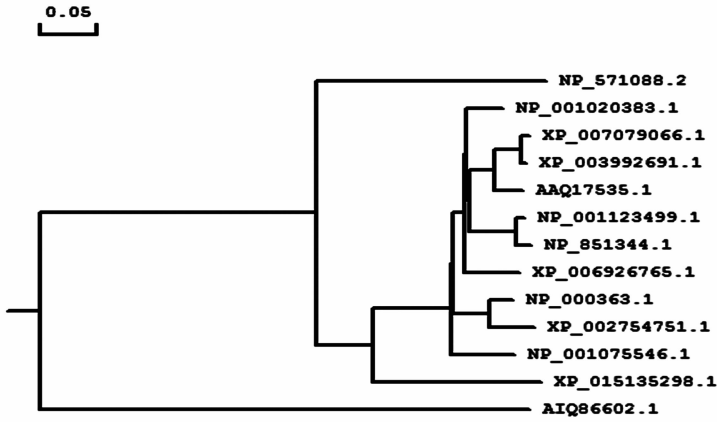


图1 不同物种 *TYR* 基因编码蛋白的系统发育树

参考文献:

[1] Ito S, Wakamatsu K, Oeki H. Chemical analysis of melanins and its application to the study of the regulation of melanogenesis [J]. Pigment Cell Research, 2000, 13(8): 103 - 109.

[2] Hearing V J, Tsukamoto K. Enzymatic control of pigmentation in mammals [J]. FASEB J, 1991, 5(14): 2902 - 2909.

[3] Crippa R, Horak V, Prota G, et al. Chemistry of melanins [J]. The Alkaloids, 1989, 36: 253 - 323.

[4] Jimbow K. Biological role of tyrosinase - related protein and its relevance to pigmentary disorders (*Vitiligo vulgaris*) [J]. Dermatol, 1999, 26(11): 734 - 737.

[5] Aigner B, Besenfelder U, Seregi J, et al. Expression of the murine wild - type tyrosinase gene in transgenic rabbits [J]. Transgenic Res, 1996, 5(6): 405 - 411.

[6] Oetting W S. The tyrosinase gene and oculocutaneous albinism type 1 (OCA1): a model for understanding the molecular biology of melanin formation [J]. Pigment Cell Res, 2000, 13(5): 320 - 325.

[7] 付鑫, 李祥龙, 周荣艳, 等. 14 个物种 *RPSA* 基因编码区生物信息学分析 [J]. 湖北农业科学, 2012, 51(2): 389 - 390.

[8] 杨洪一, 孙立娜, 张杰, 等. 黄瓜绿斑驳花叶病毒抗原表位预测

绵羊、家牛、野猪、中央蝙蝠、人、狨猴、家兔、原鸡、斑马鱼、貉子。其中绵羊与家牛关系较近, 人与狨猴关系较近。

3 结论

不同物种 *TYR* 基因编码蛋白脂肪系数均高达 70 以上, 其中貉子最高, 达到 81.39。除貉子外, 其他物种 *TYR* 基因编码蛋白均为不稳定蛋白质, 整条多肽链表现亲水性。除原鸡外, 其他物种该蛋白均存在跨膜区, 貉子有 2 个跨膜区域, 揭示该蛋白是定位于生物膜的跨膜蛋白。

*TYR* 基因编码蛋白的二级结构主要为  $\alpha$  螺旋、 $\beta$  折叠、 $\beta$  转角区域分布少且散, 均存在信号肽、跨膜区, 主要在嘌呤和嘧啶、运输和结合、电压门控离子通道及转录中发挥作用。各物种 *TYR* 基因编码产物主要定位于内质网、高尔基体以及细胞质膜, 人有小部分定位于细胞外 (11.1%), 而原鸡主要定位于细胞外 (66.7%), 小部分定位于内质网、高尔基体及液泡中 (11.1%)。系统发育树分析结果符合动物学进化观点。

[J]. 河南农业科学, 2013, 42(8): 67 - 70.

[9] 胡慧艳, 贾青, 陶隽, 等. 家猪 TBP 蛋白结构与理化性质的生物信息学分析 [J]. 河南农业科学, 2014, 43(3): 128 - 132.

[10] Jensen L J, Gupta R, Blom N, et al. Prediction of human protein function from post - translational modification and localization features [J]. Journal of Molecular Biology, 2002, 319(5): 1257 - 1265.

[11] Jensen L J, Gupta R, Strfeldt H H, et al. Prediction of human protein function according to Gene Ontology categorise [J]. Bioinformatics, 2003, 19(5): 635 - 642.

[12] 李玩生, 刘磊, 冯海燕, 等. 猪 T 细胞受体  $\beta$  链基因的克隆及生物信息学分析 [J]. 甘肃农业大学学报, 2010, 45(3): 1 - 6.

[13] Nakai K, Horton P. PSORT: A program for detecting sorting signals in protein and predicting their subcellular localization [J]. Trends Biochem Sci, 1999, 24(1): 34 - 36.

[14] 徐小刚, 刘雅婷, 李永忠, 等. 番茄斑萎病毒属病毒 SRNA 序列比对 [J]. 湖南农业大学学报 (自然科学版), 2010, 36(2): 142 - 146.

[15] Thorton J M. From genome to function [J]. Science, 2001, 292(5524): 2095 - 2097.

[16] 阎隆飞, 孙之荣. 蛋白质分子结构 [M]. 北京: 清华大学出版社, 1999: 221 - 229.