



家驴 TGF- β 基因家族成员的鉴定、表达及多态分析

黄恩泽, 赵舒悦, 王雨数, 张 欢, 殷 杰, 张翔麟, 党瑞华*

(西北农林科技大学 动物科技学院, 陕西杨凌 712100)

摘 要: 转化生长因子- β (TGF- β)基因家族在细胞生长、分化、凋亡以及细胞外基质合成与沉积等多种生物学过程中发挥着关键作用。为了探究该基因家族在驴上的表达情况,本研究对驴 TGF- β 基因家族进行了系统的调查分析,包括基因鉴定、表达分析及其功能和进化关系的探讨。使用 HMMER 软件包和 BLAST 工具对驴 TGF- β 基因家族成员进行鉴定,并结合多种生物信息学工具对其序列进行了深入分析。此外,运用转录组数据分析该基因家族在不同组织中的表达模式。结果表明,家驴 TGF- β 基因家族共包含 33 个成员,编码蛋白质的氨基酸数目为 153~507 aa,分子质量在 16.19~55.67 kD 之间,等电点分布为 4.84~11.66,所有成员均为亲水性蛋白。蛋白质二级结构预测表明,TGF- β 基因家族成员二级结构以无规则卷曲为主,其次是 α -螺旋,最后是 β -折叠。亚细胞定位结果显示,31 个成员定位于细胞外基质、BMP2 位于内质网上、BMP10 位于细胞质膜上、BMP7 位于线粒体。系统发育分析表明,TGF- β 基因家族分为两个亚类,其中驴和人的基因家族成员分别聚集,表明其具有较近的进化关系。转录组数据进一步显示,大多数 TGF- β 基因家族成员在成年驴的 13 种组织中均有表达,其中 NODAL 基因在所有组织中均表现出较高的表达水平。该研究结果为深入解析驴 TGF- β 基因家族的生物学功能奠定了基础。

关键词: 驴; TGF- β 基因家族; 系统进化; 表达分析

[中图分类号] S813.3 [文献标志码] A [文章编号] 1004-6704(2025)-03-0010-10

Identification, Expression and Polymorphism Analysis of TGF- β Gene Family Members in Donkeys

HUANG Enze, ZHAO Shuyue, WANG Yushu, ZHANG Huan,

YIN Jie, ZHANG Xianglin, DANG Ruihua*

(College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100)

Abstract: Transforming growth factor- β (TGF- β) gene family plays a key role in a variety of biological processes such as cell growth, differentiation, apoptosis, and extracellular matrix synthesis and deposition. To investigate the expression of TGF- β gene family in donkeys, we performed a systematic survey of the donkey TGF- β gene family, including gene identification, expression analysis, and discussion of their functional and evolutionary relationships. The members of the donkey TGF- β gene family were identified using the HMMER software package and BLAST tools, and the sequences were analyzed in depth in combination with a variety of bioinformatics tools. Furthermore, the expression patterns of this gene family in different tissues were analyzed using transcriptome data. The results showed that the TGF- β gene family consisted of 33 members, encoding proteins with amino

acids ranging from 153 to 507 aa, molecular weights ranging from 16.19 to 55.67 kD, and isoelectric point distributions ranging from 4.84 to 11.66. All members were hydrophilic proteins. Protein secondary structure prediction showed that the secondary structure of TGF- β gene family members was

[收稿日期] 2025-01-11

[基金项目] 大学生创新创业训练计划项目(X202410712243)

[第一作者] 黄恩泽(2004-),男,本科,主要从事动物遗传育种与繁殖方面的研究。E-mail:hez4188@163.com

*[通信作者] 党瑞华, E-mail: dangruihua@nwsuaf.edu.cn

dominated by random coil, followed by α -helix, and finally β -fold. The results of subcellular localization showed that 31 members were located in the extracellular matrix, BMP2 was located in the endoplasmic reticulum, BMP10 was located in the plasma membrane, and BMP7 was located in the mitochondria. Phylogenetic analysis showed that the TGF- β gene family was divided into two subclasses, in which the donkey and human gene family members were clustered separately, indicating a close evolutionary relationship. The transcriptome data further showed that most TGF- β gene family members were expressed in 13 tissues of the adult donkey, with the NODAL gene showing high expression levels in all tissues. These results lay a foundation for further understanding the biological functions of TGF- β gene family in donkeys.

Key words: donkey; TGF- β gene family; phylogenetic evolution; analysis of expression

转录因子是一类高度保守的蛋白,通过识别特定 DNA 序列来调控染色质结构和转录过程,进而形成一个指导基因组表达的复杂系统,参与生物体内各种生理生化反应^[1]。转化生长因子 β (Transforming growth factor β , TGF- β) 作为一种转录因子,主要功能是抑制或刺激细胞增殖,是一种双功能因子。它在细胞增殖与分化、伤口愈合及免疫系统调控中发挥着重要作用,并在多个病理过程中扮演关键角色,如骨骼疾病^[2]、纤维化和癌症^[3]。TGF- β 家族包含多种功能不同的分泌蛋白,这些蛋白在发育模式、组织分化以及维持体内平衡方面有着多种作用。例如:骨骼肌的负调控因子 GDF-8,引发异位骨形成和影响骨骼形态发生的 BMP 类蛋白,由睾丸中的支持细胞分泌,通过抑制米勒管发育在早期性腺发育和性别分化中发挥重要作用的 AMH 等^[4-7]。TGF- β 基因家族广泛参与多种生物学过程^[8],是开发药物的理想靶点。

TGF- β 是一种 25 kD 二硫键连接的二聚体,其前体通常为折叠的环状复合物,其中二硫键、 α -螺旋与 β -折叠在其结构稳定性方面起着重要作用^[9]。近年来,对 TGF- β 的研究不断深入,尤其是在动物遗传育种领域,学者们越来越重视 TGF- β 基因家族在多种经济性状遗传调控中的作用机制。然而,关于驴的 TGF- β 基因家族在全基因组范围内的筛选、进化分析、表达规律研究及家族成员的基因多态性分析,尚未见系统性的研究和报道。本研究旨在全面揭示家驴 TGF- β 基因家族的特征与功能,为深入解析其生物学功能及分子标记开发奠定基础。

1 材料和方法

1.1 TGF- β 基因家族鉴定及特征分析

从 NCBI(<https://www.ncbi.nlm.nih.gov>) 数据库下载了驴的基因组序列、蛋白序列和基因注释文件。并从 Pfam(<https://pfam.xfam.org>) 数据库获取了 TGF- β (PF00019) 基因家族特有的结构域信息。使用 HMMER 软件包鉴定候选的 TGF- β 基因家族成员^[10]。运用 TBtools 工具将 TGF- β 结构域和驴蛋白序列文件进行 BLAST 比对,获得 E 值 $< 10^{-5}$ 的所有蛋白质序列。通过 CDD 网站 (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) 提交候选基因家族蛋白序列,对 TGF- β 全长蛋白序列进行保守结构域进行验证,并删除不含 TGF- β 结构域的序列。在 TBtools 软件中进行可视化,获得最终的基因家族成员。

使用 ExPASy 在线工具 (<https://www.expasy.org/>) 分析 TGF- β 家族蛋白的氨基酸数量、等电点、分子质量、不稳定指数、脂溶性系数和总平均亲水性^[11-13]。通过 WoLF PSORT 在线网站 (<https://wolfpsort.hgc.jp/>) 进行亚细胞定位^[14]。使用 SOPMA 在线分析网站 (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa%20_sopma.html) 对 TGF- β 基因家族成员蛋白质的二级结构进行预测^[15]。

1.2 系统进化树分析

从 NCBI 数据库下载驴、人和小鼠的蛋白序列文件,使用 MEGA11 软件通过邻接法 (neighbor-joining, NJ) 构建驴、人和小鼠的 TGF- β 基因家族的系统发育树,校验参数中 Bootstrap 值设为 1 000, Site Coverage Cutoff 设为 50,其他参数保持默认值^[16]。随后,利用在线软件 Evolview (<https://evolgenius.info/evolview-v2/#login>) 对进化树进行美化^[17]。

1.3 染色体定位及基因结构

使用 TBtools 软件的“Gene Location Visualize”功能插件,结合驴 TGF- β 家族成员 ID 与驴 gff3 基因注释文件 (ASM1607732v2),绘制驴 TGF- β 基因染色体定位图谱。然后使用 TBtools 软件的“Visualize Gene Structure”插件,结合驴 TGF- β 家族 ID 和驴 gff3 基因注释文件,对基因结构进行可视化。最后,借助 Adobe Illustrator 软件进行图形美化^[18]。

1.4 TGF- β 基因家族蛋白保守基序分析

首先,通过在线预测分析软件 MEME (<http://meme-suite.org/index.html>) 对驴 TGF- β 基因家族

保守的蛋白基序进行分析^[19]。接着,将 TGF- β 基因家族的蛋白序列文件上传至 NCBI 的 Batch CD-Search 工具进行预测^[20]。最后,利用 TBtools 软件“Gene Structure View”插件将上述两次获得的结果进行可视化。

1.5 TGF- β 基因家族成员在驴不同组织中表达特征分析

从 NCBI 官网的 SRA 数据库^[21]下载成年驴 13 种组织(皮肤、肌肉、血液、大脑、心脏、肺、胃、脾脏、肝脏、肾脏、盲肠、睾丸、附睾)的 RNA-seq 原始测序数据(PRJNA431818)^[22]。使用 Trimmomatic 切除^[23]低质量的碱基和接头序列。随后,利用 Hisat2 软件建立索引,并将质控后高质量数据比对到驴参考基因组^[24]。接着,使用 Stringtie 对转录本进行组装并估计基因的表达丰度(FPKM)^[25]。最后,使用 TBtools 软件中的“HeatMap”功能绘制 TGF- β 家族基因在驴各组织中的表达热图。

1.6 功能富集分析

为了深入研究驴 TGF- β 基因家族的生物学功能,采用在线工具 DAVID(<https://david.ncifcrf.gov/>)分别从生物学过程(biological process)、分子功能(molecular function)和细胞组分(cellular component)等方面对 TGF- β 基因家族开展基因功能富集分析。

2 结果与分析

2.1 TGF- β 基因家族鉴定及基本性质分析

使用 HMMER 和 BLAST 工具检索基因组和蛋白组数据库,并通过 CDD 检查序列是否存在完整的 TGF- β 结构域,从驴基因组中鉴定出 34 个 TGF- β 基因家族成员,并分析了它们的特征和理化性质(表 1)。驴 TGF- β 基因家族成员氨基酸数量的变化范围较大,从 153 个氨基酸(PSPN)到 507 个氨基酸(GDF5);蛋白序列的相对分子质量在 16.19~55.67 kD,与氨基酸数量变化趋势一致;等电点变化范围从 4.84(BMP10)到 11.66(ARTN);不稳定系数从 36.79(PSPN)到 84.20(ARTN);脂溶性系数为 64.40(ARTN)到 97.25(GDF3)。TGF- β 家族成员总平均亲水性(grand average of hydropathicity)均为负值,表明它们是亲水性蛋白。

蛋白质二级结构预测结果显示(表 2),驴 TGF- β 基因家族成员蛋白组成主要由无规则卷曲构成,其次是 α -螺旋, β -折叠的比例最小。不同成员蛋白质的无规则卷曲、 α -螺旋、 β -折叠和延伸链等二级结构比例差异较大,其中 α -螺旋的占比为 14.17%到

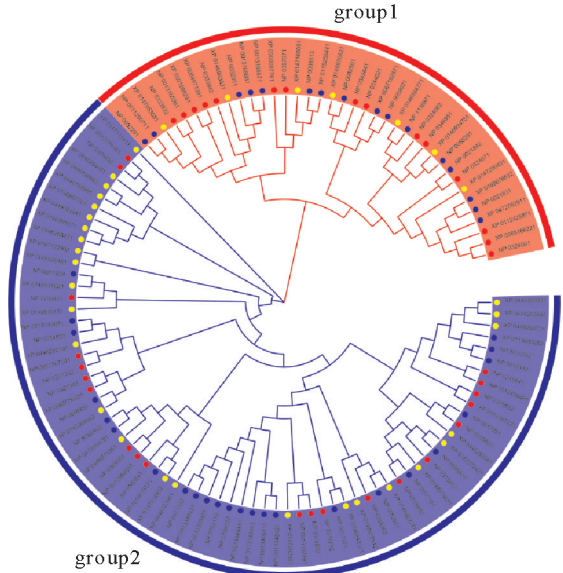
43.79%,延伸链占比为 8.53%到 23.93%, β -折叠占比为 1.46%到 9.76%,无规则卷曲占比为 40.52%到 65.12%。亚细胞定位结果显示,TGF- β 基因家族 30 个成员位于细胞外基质,BMP2 位于内质网上,BMP10 位于细胞质膜上,BMP7 位于线粒体。

2.2 TGF- β 基因家族系统进化树分析

为了研究驴 TGF- β 基因家族成员与人、小鼠之间的进化关系,用驴、人和小鼠的 TGF- β 基因家族成员蛋白序列进行多序列比对,并在 MEGA 软件中用 NJ 方法构建系统进化树。驴有 34 个 TGF- β 家族蛋白成员,人有 36 个 TGF- β 家族成员,小鼠有 40 个 TGF- β 家族成员,共计 110 个蛋白成员。系统发育树显示,TGF- β 基因蛋白家族被分为了 group1 和 group2 两大亚类(图 1),其中 group1 含有 38 个同源基因蛋白,包括 7 个驴 TGF- β 家族成员,16 个人 TGF- β 家族成员和 15 个小鼠 TGF- β 家族成员;group2 含有 72 个同源基因蛋白,包括 27 个驴 TGF- β 家族成员,20 个人 TGF- β 家族成员和 25 个小鼠 TGF- β 家族成员。在进化树上,3 个物种的 TGF- β 基因蛋白家族成员相互聚集,相间排布,表明它们在进化上较为接近,并且这些基因家族成员可能具有功能上的相似性。

2.3 驴 TGF- β 基因家族成员染色体定位及基因结构分析

从驴基因组的 gff3 注释文件中提取各个 TGF- β 基因家族成员的位置信息,利用 TBtools 生成染



人. 蓝色;驴. 黄色;鼠. 红色

图 1 驴、人和小鼠 TGF- β 基因家族蛋白系统发育进化树
Fig. 1 Phylogenetic tree of donkey, human and mouse TGF- β gene family proteins

染色体分布图(图 2)。结果表明,TGF-β 基因家族成员非均匀分布在 1、2、3、4、5、6、7、8、10、12、15、19、20、22、26、30 号染色体和 X 染色体上。其中在 1 号、2 号、5 号、7 号、22 号和 30 号染色体上都定位到 3 个驴 TGF-β 基因家族的成员,6 号、12 号、15 号和 20 号染色体都定位到 2 个驴 TGF-β 基因家族的成员。3 号、4 号、8 号、9 号、19 和 X 染色体均定位到 1 个驴 TGF-β 基因家族的成员。

基因结构分析结果显示(图 3),驴 TGF-β 基因家族成员含有 1~9 个外显子。其中,INHBC 含有 1 个外显子,14 个成员含有 2 个外显子,5 个成员含

有 3 个外显子,3 个成员含有 4 个外显子,2 个成员含有 5 个外显子,5 个成员含有 7 个外显子,2 个成员含有 8 个外显子,BMP8A 含有 9 个外显子。驴 TGF-β 基因家族绝大多数成员都含有内含子,其中,含有 1 个内含子的成员有 13 个,含有 2 个内含子的成员有 7 个,含有 3 个内含子的成员有 5 个,含有 6 个内含子的成员有 7 个,TGFB2 含有的内含子数最多,包含 7 个内含子。驴 TGF-β 基因家族成员的外显子和内含子的个数差异都较大,说明驴 TGF-β 基因家族成员的基因结构具有多样性。

表 1 驴 TGF-β 基因家族蛋白理化性质
Table 1 Physicochemical properties of donkey TGF-β gene family proteins

基因名称	基因号	氨基酸 数量	蛋白质相对 分子质量/kD	等电点	不稳定 指数	脂溶性 系数	总平均 亲水性
GDF5	XP_044629204.1	507.00	55.67	8.20	58.63	71.83	-0.440
BMP5	XP_014707242.1	454.00	51.47	9.00	47.99	77.97	-0.439
BMP7	XP_014687467.1	431.00	49.30	7.36	53.46	76.26	-0.411
BMP2	XP_014684231.1	395.00	44.89	8.85	53.99	79.24	-0.473
BMP8B	XP_044626372.1	402.00	44.78	8.90	60.58	82.04	-0.363
BMP8A	XP_044626354.1	404.00	45.00	8.79	60.65	82.60	-0.358
BMP4	XP_014720200.1	409.00	46.60	8.57	57.89	79.80	-0.532
GDF6	XP_014703240.2	461.00	51.43	9.19	69.87	69.87	-0.610
GDF7	XP_044627936.1	452.00	46.47	9.29	55.16	79.51	-0.118
GDF3	XP_014706486.2	364.00	40.76	7.11	52.09	97.25	-0.087
BMP3	XP_014691802.1	475.00	53.46	9.62	59.67	79.05	-0.519
INHBA	XP_014720043.1	426.00	47.70	7.56	62.09	79.44	-0.522
BMP10	XP_014692224.1	424.00	48.02	4.84	47.26	86.93	-0.349
INHBE	XP_014699477.1	351.00	38.52	9.00	63.28	86.78	-0.255
INHBC	XP_014699475.1	351.00	38.37	8.27	46.82	88.12	-0.049
GDF1	XP_044629219.1	369.00	39.18	11.20	69.92	89.32	-0.059
GDF10	XP_014697015.1	477.00	52.54	9.52	56.44	76.77	-0.451
GDF2	XP_014697014.2	427.00	47.43	6.34	50.24	75.57	-0.426
BMP15	XP_014682648.1	394.00	45.23	9.67	52.86	96.19	-0.309
GDF9	XP_014712290.2	451.00	50.63	8.90	55.54	79.18	-0.350
MSTN	XP_014693062.1	375.00	42.84	7.46	42.79	83.95	-0.406
TGFB3	XP_014696363.1	412.00	47.17	8.13	48.59	82.06	-0.499
TGFB2	XP_014705358.1	442.00	50.53	8.74	53.00	80.52	-0.396
NODAL	XP_044612398.1	350.00	39.92	5.81	63.07	85.23	-0.374
TGFB1	XP_014716505.1	391.00	44.07	8.83	46.29	89.77	-0.255
GDF15	XP_044629495.1	300.00	32.98	10.72	63.53	92.20	-0.298
INHA	XP_014711541.1	367.00	39.45	7.62	68.96	90.71	-0.042
NRTN	XP_044609194.1	197.00	22.37	11.12	68.83	85.79	-0.434
ARTN	XP_044626547.1	225.00	23.11	11.66	84.20	64.40	-0.385
GDNF	XP_014686778.1	211.00	23.81	9.25	64.20	78.20	-0.547
PSPN	XP_014699167.1	153.00	16.20	9.48	36.79	86.93	-0.099
LOC106844730	XP_044618896.1	367.00	40.79	8.40	44.02	87.14	-0.334
LOC106844732	XP_044618866.1	367.00	40.77	8.61	46.35	87.14	-0.334

表 2 驴 TGF-β 基因家族蛋白二级结构预测及亚细胞定位

Table 2 Secondary structure prediction and subcellular localization of donkey TGF-β gene family proteins

基因名称	二级结构分布比例				亚细胞定位
	α-螺旋	延伸链	β-折叠	无规则卷曲	
GDF5	22.88	17.95	3.35	55.82	extracellular
BMP5	23.79	16.52	2.86	56.83	extracellular
BMP7	24.59	17.87	3.02	54.52	mitochondrial
BMP2	25.06	17.72	2.03	55.19	endoplasmic reticulum
BMP8B	20.40	17.66	2.49	59.45	extracellular
BMP8A	22.03	18.07	2.48	57.43	extracellular
BMP4	23.47	19.56	1.96	55.01	extracellular
GDF6	22.56	10.20	2.60	64.64	extracellular
GDF7	38.27	11.50	8.19	42.04	extracellular
GDF3	26.37	17.03	5.77	50.82	extracellular
BMP3	25.89	14.95	3.79	55.37	extracellular
INHBA	25.35	19.48	4.69	50.47	extracellular
BMP10	28.54	17.92	2.59	50.94	plasma membrane
INHBE	23.93	19.94	3.70	52.42	extracellular
INHBC	24.50	23.93	4.56	47.01	extracellular
GDF1	29.00	12.74	9.76	48.51	extracellular
GDF10	22.22	15.09	3.14	59.54	extracellular
GDF2	25.06	18.03	3.28	53.63	extracellular
BMP15	30.46	11.93	1.78	55.84	extracellular
GDF9	26.39	15.96	2.44	55.21	extracellular
MSTN	20.80	23.73	3.47	52.00	extracellular
TGFB3	34.95	15.53	1.46	48.06	extracellular
TGFB2	34.39	13.12	2.49	50.00	extracellular
NODAL	21.71	16.29	4.00	58.00	extracellular
TGFB1	32.74	17.14	3.58	46.55	extracellular
GDF15	30.33	12.33	4.00	53.33	extracellular
INHA	14.17	17.44	3.27	65.12	extracellular
NRTN	41.12	10.15	7.61	41.12	extracellular
ARTN	22.67	8.89	4.44	64.00	extracellular
GDNF	30.33	8.53	5.69	55.45	extracellular
PSPN	43.79	10.46	5.23	40.52	extracellular
LOC106844730	27.52	18.80	3.27	50.41	extracellular
LOC106844732	28.88	17.71	3.54	49.86	extracellular

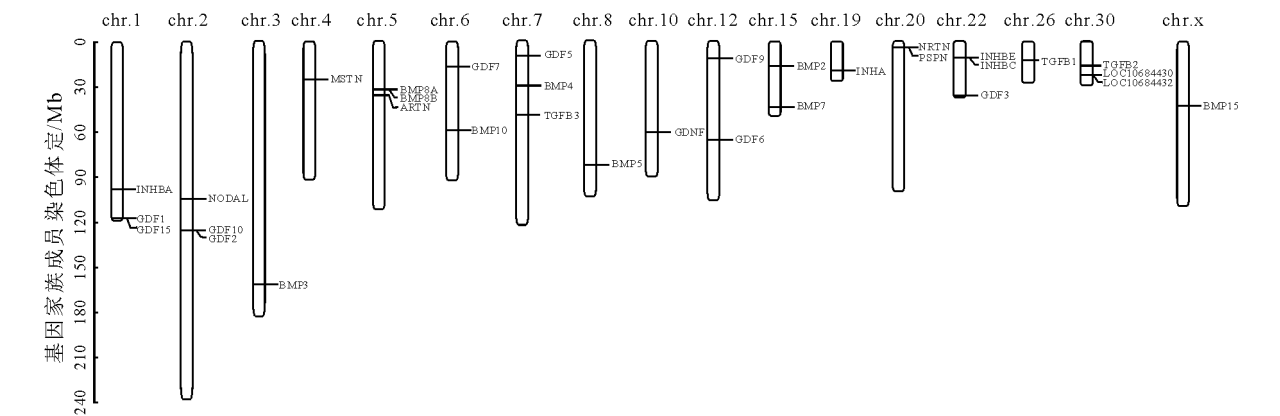


图 2 33 个驴 TGF-β 基因家族成员染色体定位图

Fig. 2 Chromosome mapping of 34 donkey TGF-β gene family members

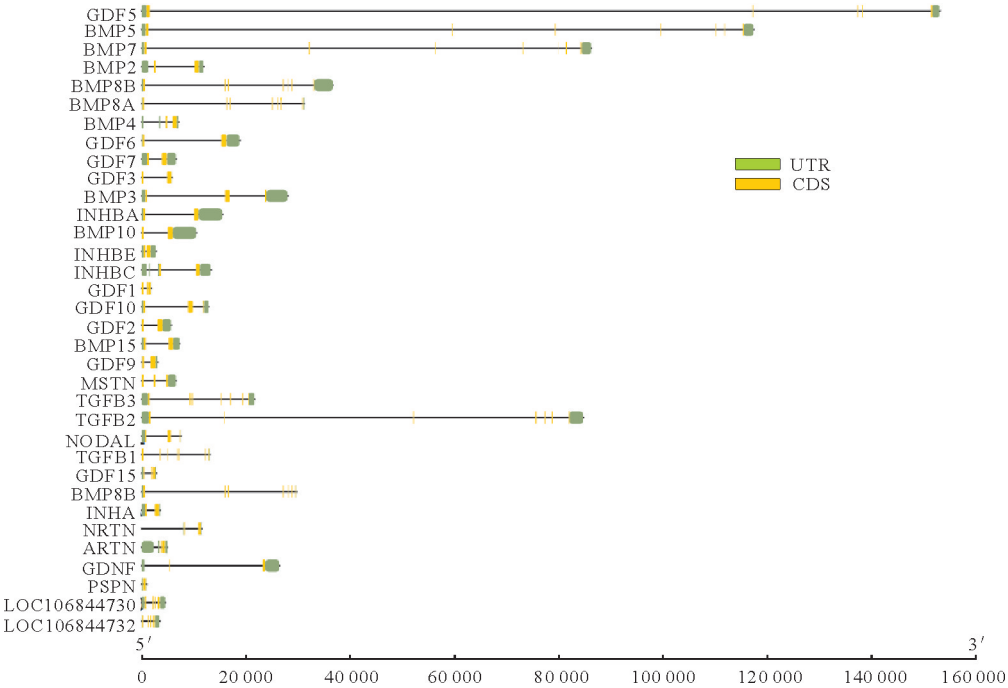
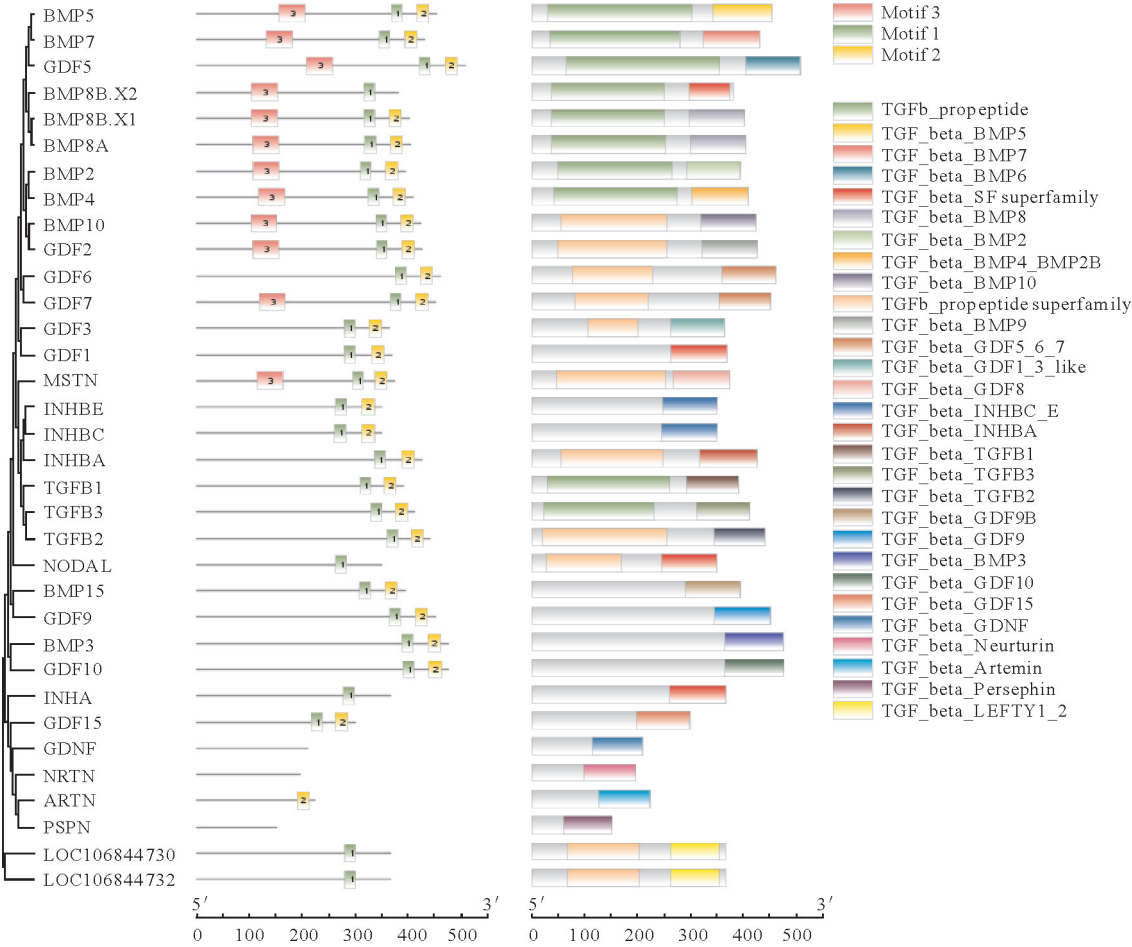


图 3 驴 TGF-β 基因家族成员结构分析

Fig. 3 Structural analysis of donkey TGF-β gene family members



左侧为 TGF-β 的邻接系统发育树;中间为 TGF-β 蛋白中的 3 个保守基序;
右侧为 TGF-β 的保守结构域;基序大小可以通过底部的比例估计

图 4 TGF-β 基因家族蛋白保守基序分布

Fig. 4 Distribution of conserved motif of TGF-β gene family proteins

2.4 驴 TGF-β 基因家族蛋白保守基序分析

为进一步了解驴 TGF-β 基因家族蛋白的结构特征,对该基因家族所编码蛋白序列的保守基序进行了预测。结果显示,在驴 TGF-β 基因家族中共鉴定出 3 种保守基序,其中 TGF-β 结构域由 motif 1 和 motif 2 组成,这两个保守基序在家族中分布也最为广泛。由图 4 可知,25 个驴 TGF-β 基因家族蛋白含有保守基序 1 和保守基序 2,其中 11 个驴 TGF-β 基因家族蛋白序列 3 个保守基序都含有。其余 9 个中的 BMP8B 含有保守基序 1 和保守基序 3,其余 4 个只含有保守基序 1,ARTN 只含有保守基序 2,GDNF、NRTN 和 PSPN 没有这 3 种保守基序。说明大多数驴 TGF-β 基因家族蛋白有着较高的保守性。

2.5 TGF-β 基因家族成员在驴背长肌中表达特征分析

为研究 TGF-β 基因家族在驴不同组织中的表达情况,本研究从 NCBI 公共数据库下载了驴转录组测序数据,获得了 33 个 TGF-β 基因家族成员在 13 种不同组织中的表达情况(图 5)。结果显示,BMP5、GDF6、GDF3、BMP3、BMP15、GDF9、GDNF 基因在所有组织中表达丰度<0.5(FPKM<0.5);GDF5、BMP8B、TGFB、NODAL、ARTN、PSPN 基因在所有组织中表达丰度>1(FPKM>1)。由图 5 可知,血液中 TGFB1 基因表达量最高,睾丸中 INHA 基因表达量最高,肺中 GDF5 基因表达量最高,肝中 PSPN 基因高表达量最高,NODAL 基因在 13 个组织中均高表达。

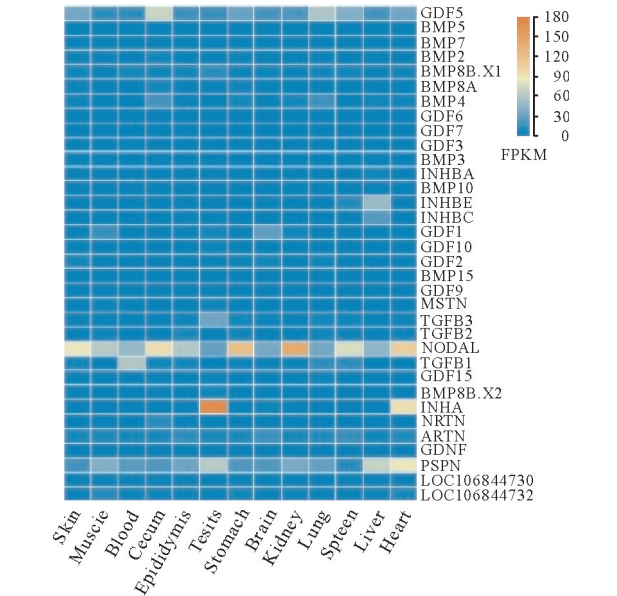


图 5 驴 TGF-β 家族成员在不同组织中表达情况
Fig. 5 Expression of donkey TGF-β family members in different tissues

2.6 驴 TGF-β 基因家族功能富集分析

为进一步了解驴 TGF-β 基因生物学家族功能,利用 DAVID 在线分析网站对 TGF-β 家族基因进行 GO 注释显示三个方面(细胞成分、生物过程和分子功能)的功能富集。GO 分析发现(图 6),TGF-β 家族参与多个与生长发育相关的生物学功能。其中 TGF-β 转录因子富集到 2 个细胞组分注释条目,细胞胞外空间(GO:0005615)和细胞外区(GO:0005576);分子功能中主要富集于生长因子活性(GO:0008083)、细胞因子活性(GO:0005125)、转化生长因子 β 受体结合(GO:0005160)、II 型转化生长因子 β 受体结合(GO:0005114)、III 型转化生长因子 β 受体结合(GO:0034714)、受体酪氨酸激酶结合(GO:0030971)、激素活性(GO:0005179)、胶质细胞衍生的神经营养因子受体结合(GO:0030116)、BMP 受体结合(GO:0070700)。生物过程主要富集于 BMP 信号通路(GO:0030509)、SMAD 蛋白信号传导(GO:0060395)、细胞分裂的正向调控(GO:0051781)、TGF-β 受体信号通路(GO:0007179)、细胞群体增殖的调控(GO:0042127)、前后轴特化(GO:0009948)、成骨细胞分化(GO:0001649)、外周神经系统发育(GO:0007422)和软骨发育(GO:0051216)。

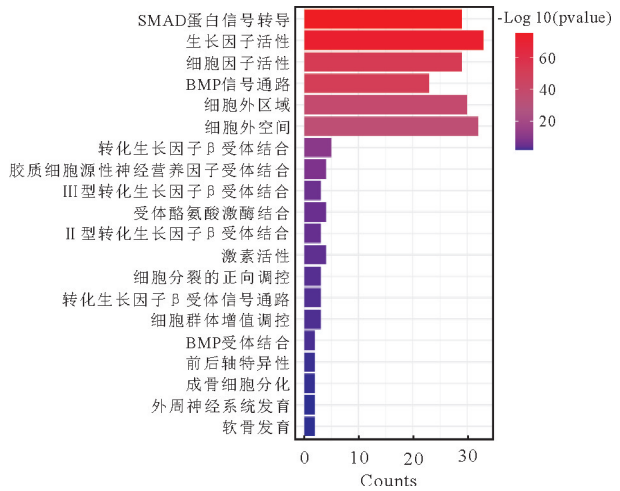


图 6 驴 TGF-β 家族成员 GO 富集分析图
Fig. 6 GO enrichment analysis of donkey TGF-β family members

3 讨论

TGF-β 是一类具有广泛生物活性的多功能细胞因子,在细胞增殖、分化和凋亡等过程中发挥关键作用。TGF-β 不仅参与胞外基质合成、胚胎发育、器官形成和骨骼再生,而且在机体免疫应答中扮演重要角色^[26],此外,TGF-β 信号通路与各种疾病的

联系和在干细胞的自我更新和多能性维持中的作用也受到广泛关注^[27]。本研究共鉴定出 33 个驴 TGF- β 基因家族的成员,数量低于人和小鼠鉴定出的 36 个 TGF- β 基因家族成员和 40 个基因家族成员。基因染色体的定位显示,33 个驴 TGF- β 基因家族成员非均匀分布在 1、2、3、4、5、6、7、8、10、12、15、19、20、22、26、30 号染色体以及 X 染色体上。

在驴 TGF- β 基因家族蛋白中,34 个成员的氨基酸数目在 153~507 aa,其中 88.29% 的蛋白含氨基酸数量介于 300~500 个,相对分子质量为 16.19~55.67 kD,与氨基酸数量变化趋势一致。驴 TGF- β 基因家族蛋白序列的等电点在 4.84~11.66,其中 28 个蛋白质的等电点>7.5,表明大多数驴 TGF- β 基因家族蛋白属于碱性蛋白^[28],其不稳定系数介于 36.79~84.20 之间,只有 PSPN 的不稳定系数<40,属于稳定蛋白,其余的均为不稳定蛋白^[29]。总平均亲水性为序列中所有氨基酸亲水值总和与氨基酸总数比值,负值越大表示亲水性越好,正值越大表示疏水性越强,驴 TGF- β 基因家族蛋白亲水性均为负值,说明均为亲水性蛋白^[30]。由此可见,驴 TGF- β 基因家族蛋白在理化性质方面具有较明显的差异。

蛋白二级结构分析结果表明,驴 TGF- β 基因家族蛋白结构主要是无规则卷曲,无规则卷曲占比在 50% 及以上的 TGF- β 基因家族蛋白有 26 个,其次为 α -螺旋,最后是 β -折叠。蛋白二级结构中无规则卷曲所含比例较高,可能是因为无规则卷曲连接其他 3 类结构,所以致使所占百分比高^[31]。此外,不同成员蛋白质无规则卷曲、延伸链、 α -螺旋和 β -折叠等二级结构分布比例差异明显,这可能是造成驴 TGF- β 基因家族成员多样性的原因之一。亚细胞定位分析结果显示,驴 TGF- β 基因家族蛋白大部分定位于细胞外基质,少数定位于细胞膜系统中,说明驴 TGF- β 基因可能是多数通过细胞信号转导系统参与生物学进程^[32]。驴 TGF- β 基因家族系统进化树分析也表明,TGF- β 基因家族成员在进化过程中发生了分化。

驴 TGF- β 基因家族成员基因结构分析表明该基因家族成员外显子和内含子都呈现多样性,外显子数变异较大,1 个成员有一个外显子,21.21% 的成员含有 2 个外显子,其他成员则含有 3~9 个外显子,内含子个数变异也较大,其中,有 39.40% 的成员含有 1 个内含子,其他成员含有 2~7 个。驴 TGF- β 基因家族成员的外显子和内含子的个数差异都较大,说明驴 TGF- β 基因家族成员具有多样

性,在驴的物种进化过程中发挥重要作用^[33]。

序列特异性蛋白的结合位点是基序,这些基序在基因表达调控中起着至关重要的作用。本研究中驴 TGF- β 基因家族具有相同保守基序的成员在系统发育树中聚类在一起,这表明它们可能具有相似的生物学功能;而不同分支的成员往往具有不同的生物学功能,这主要是由于它们所含基序的差异所致;可以观察到同一进化分支的成员间具有相似的保守基序组成,而不同分支的成员则具有不同的基序组成,如 *TGFB1* 和 *TGFB2* 在同一分支且有着相同基序,而 *MSTN* 处在不同分支基序与两者相比也有差异。这些发现强调了基序在 TGF- β 基因家族成员的功能分化和进化过程中的关键作用^[34]。

基因表达结果表明,除极少数驴 TGF- β 基因家族成员在成年驴个别组织中不能表达外,大多数的驴 TGF- β 基因家族成员在成年驴 13 种不同的组织中均有表达,结果显示某些 TGF- β 家族基因在某些特定的组织或细胞中表达量较高,血中 *TGFB1* 基因表达量最高,睾丸中 *INHA* 基因表达量最高,肺中 *GDF5* 基因表达量最高,肝中 *PSPN* 基因高表达量最高,*NODAL* 基因在 13 个组织中均高表达。之前的研究发现 *TGFB1* 在细胞增殖、分化、免疫调节以及疾病发展中有着重要功能^[35],*INHA* 基因在绵羊睾丸中的高表达^[36],*PSPN* 在人体免疫系统的高度相关^[37],*NODAL* 基因调控人体左右体轴对称性^[38]、在多种器官(包括心脏、肝脏、胰腺和肠道等)的发生和发展^[39]、调控干细胞的分化方向和维持其多能性等方面的重要作用也均有报道^[40]。这些研究显示这些基因在这些组织或细胞中具有显著的高表达相关性。

在本次针对驴 TGF- β 基因家族的研究工作中,我们采用了 GO 功能富集分析技术,研究结果清晰地表明,这一基因家族在生物体的生长、分化、发育、信号转导、免疫应答以及癌症的发生路径中,有着重要功能,这对于深入理解驴的生理机制和疾病防治具有重要意义。该结论与前人的研究结果相一致^[41-45],进一步证实了 TGF- β 基因家族在众多生物学及病理学过程中具有重要的核心作用。

参考文献:

- [1] LAMBERT S A, JOLMA A, CAMPITELLI L F, et al. The human transcription factors [J]. Cell, 2018, 172 (4): 650-665.
- [2] TANG Y, WU X W, LEI W Q, et al. TGF-beta1-induced migration of bone mesenchymal stem cells couples bone resorption with formation [J]. Nature Medicine, 2009,

- 15(7):757-765.
- [3] HU H H, CHEN D Q, WANG Y N, et al. New insights into TGF- β /Smad signaling in tissue fibrosis[J]. *Chemico-Biological Interactions*, 2018, 292: 76-83.
 - [4] ROSCHEL H, UGRINOWISTCH C, SANTOS A R, et al. Effect of eccentric action velocity on expression of genes related to myostatin signaling pathway in human skeletal muscle[J]. *Biology of Sport*, 2018, 35(2): 111-119.
 - [5] DURLINGER A L L, VISSER J A, THEMME A P N. Regulation of ovarian function: The role of anti-Müllerian hormone[J]. *Reproduction*, 2002, 124(5): 601-609.
 - [6] WU M R, CHEN G Q, LI Y P. TGF- β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease[J]. *Bone Research*, 2016, 4: 16 009.
 - [7] MORIKAWA M, DERYNCK R, MIYAZONO K. TGF- β and the TGF- β family: Context-dependent roles in cell and tissue physiology[J]. *Cold Spring Harbor Perspectives in Biology*, 2016, 8(5): 021 873.
 - [8] 鲁翠云, 匡友谊, 郑先虎, 等. 转化生长因子- β (TGF- β) 基因家族在水产养殖中的潜在应用价值[J]. *水产学杂志*, 2020, 33(6): 89-96.
LU C Y, KUANG Y Y, ZHENG X H, et al. Potential application of transforming growth factor- β (TGF- β) gene family in aquaculture[J]. *Chinese Journal of Fisheries*, 2020, 33(6): 89-96.
 - [9] SHI M L, ZHU J H, WANG R, et al. Latent TGF- β structure and activation[J]. *Nature*, 2011, 474(7 351): 343-349.
 - [10] EDDY S R. Profile hidden Markov models[J]. *Bioinformatics*, 1998, 14(9): 755-763.
 - [11] LI H, LIANG Z Y, CHAO Y H, et al. Exploring the GRAS gene family in *Taraxacum kok-saghyz* Rodin: Characterization, evolutionary relationships, and expression analyses in response to abiotic stresses[J]. *Biochemical and Biophysical Research Communications*, 2024, 733: 150 693.
 - [12] REN W C, KONG L Y, JIANG S, et al. Genome-wide identification, evolution, and characterization of GATA gene family and GATA gene expression analysis post-MeJA treatment in *Platycodon grandiflorum* [J]. *Journal of Plant Growth Regulation*, 2025, 44(1): 155-167.
 - [13] 李俊杰, 于 朋, 李 昀, 等. 花鲈 Claudins 基因家族的鉴定、进化分析及对环境盐度的表达响应[J]. *海洋科学*, 2022, 46(8): 65-78.
LI J J, YU P, LI Y, et al. Identification and evolutionary analysis of Claudins gene family in spotted sea bass and its expression characteristics in gill tissue under different salinity [J]. *Marine Sciences*, 2022, 46(8): 65-78.
 - [14] HORTON P, PARK K J, OBAYASHI T, et al. WoLF PSORT: Protein localization predictor [J]. *Nucleic Acids Research*, 2007, 35(Web Server issue): 585-587.
 - [15] GEORJON C, DELÉAGE G. SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments [J]. *Computer Applications in the Biosciences*, 1995, 11(6): 681-684.
 - [16] 刘京鸽, 王帅威, 吴结革, 等. 猪 TGF- β 基因家族成员的鉴定、进化及表达分析[J]. *畜牧与兽医*, 2022, 54(11): 12-22.
LIU J G, WANG SH W, WU J G, et al. Identification, evolutionary and expression analysis of the TGF- β gene family in pig [J]. *Animal Husbandry & Veterinary Medicine*, 2022, 54(11): 12-22.
 - [17] SUBRAMANIAN B, GAO S H, LERCHER M J, et al. Evolvview v3: A webserver for visualization, annotation, and management of phylogenetic trees[J]. *Nucleic Acids Research*, 2019, 47(W1): 270-275.
 - [18] 王金磊, 赵玉鑫, 穆红梅, 葫芦 bHLH 基因家族鉴定及生物信息学分析[J]. *分子植物育种*. : 1-14.
 - [19] 赵佳莹, 应宇鑫, 许张婷, 等. 掌叶覆盆子 TIFY 基因家族鉴定及表达分析 [J]. *核农学报*, 2024, 38(10): 1 823-1 835.
ZHAO J Y, YING Y X, XU ZH T, et al. Identification and expression profiles of TIFY gene family in *rcbus chingii hu* [J]. *Journal of Nuclear Agricultural Sciences*, 2024, 38(10): 1 823-1 835.
 - [20] MARCHLER-BAUER A, BRYANT S H. CD-Search: Protein domain annotations on the fly [J]. *Nucleic Acids Research*, 2004, 32(Web Server issue): 327-331.
 - [21] ZHAO P J, ZHENG X R, FENG W, et al. Profiling long noncoding RNA of multi-tissue transcriptome enhances porcine noncoding genome annotation[J]. *Epigenomics*, 2018, 10(3): 301-320.
 - [22] WANG Y N, MIAO X Y, ZHAO Z C, et al. Transcriptome atlas of 16 donkey tissues [J]. *Frontiers in Genetics*, 2021, 12: 682 734.
 - [23] BOLGER A M, LOHSE M, USADEL B. Trimmomatic: A flexible trimmer for Illumina sequence data [J]. *Bioinformatics*, 2014, 30(15): 2 114-2 120.
 - [24] KIM D, LANGMEAD B, SALZBERG S L. HISAT: A fast spliced aligner with low memory requirements [J]. *Nature Methods*, 2015, 12(4): 357-360.
 - [25] PERTEA M, PERTEA G M, ANTONESCU C M, et al. StringTie enables improved reconstruction of a

- transcriptome from RNA-seq reads[J]. *Nature Biotechnology*, 2015, 33(3): 290-295.
- [26] 李莹莹, 崔东遥, 程 操, 等. 水产动物 TGF- β 超家族基因系统进化及生物学功能研究进展[J]. *大连海洋大学学报*, 2019, 34(6): 853-862.
- LI Y Y, CUI D Y, CHENG CH, et al. A review of TGF- β superfamily genes in aquatic animals: Phylogeny and biological functions[J]. *Journal of Dalian Ocean University*, 2019, 34(6): 853-862.
- [27] VALLIER L, TOUBOUL T, BROWN S, et al. Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells[J]. *Stem Cells*, 2009, 27(11): 2 655-2 666.
- [28] GUO S F, ZHU X Y, JAŃCZEWSKI D, et al. Measuring protein isoelectric points by AFM-based force spectroscopy using trace amounts of sample[J]. *Nature Nanotechnology*, 2016, 11(9): 817-823.
- [29] MAGLIERY T J. Protein stability: Computation, sequence statistics, and new experimental methods[J]. *Current Opinion in Structural Biology*, 2015, 33: 161-168.
- [30] KYTE J, DOOLITTLE R F. A simple method for displaying the hydropathic character of a protein[J]. *Journal of Molecular Biology*, 1982, 157(1): 105-132.
- [31] CHOU P Y, FASMAN G D. Prediction of protein conformation[J]. *Biochemistry*, 1974, 13(2): 222-245.
- [32] YANG Y Z. Wnt signaling in development and disease[J]. *Cell & Bioscience*, 2012, 2(1): 14.
- [33] ROY S W, GILBERT W. The evolution of spliceosomal introns: Patterns, puzzles and progress[J]. *Nature Reviews Genetics*, 2006, 7(3): 211-221.
- [34] 张晓旭, 马媛媛, 罗新锐, 等. 芸豆 GH3 基因家族鉴定及进化与表达分析[J]. *西北植物学报*, 2024, 44(9): 1 433-1 444.
- [35] 孙昭恒, 丁 樱, 郭 婷, 等. 转化生长因子 β 1 通路在 IgA 肾病中作用机制的研究进展[J]. *广西医学*, 2023, 45(7): 843-849.
- SUN ZH H, DING Y, GUO T, et al. Mechanism of transforming growth factor β 1 pathway in IgA nephropathy: A research progress[J]. *Guangxi Medical Journal*, 2023, 45(7): 843-849.
- [36] 采复拉·大木拉, 田志龙, 储明星. INHA 基因在公绵羊繁殖相关组织的表达研究[J]. *中国草食动物科学*, 2018, 38(4): 14-17.
- CAIFULA · DAMULA, TIAN ZH L, CHU M X. Expression of INHA gene in ram reproduction-related tissues[J]. *China Herbivore Science*, 2018, 38(4): 14-17.
- [37] 曲 伟, 商军锋, 邹慧琳, 等. PCT 和 PSPN 对急性髓系白血病化疗期间发热性中性粒细胞减少症患者细菌感染的鉴别价值[J]. *现代肿瘤医学*, 2024, 32(9): 1 703-1 708.
- QU W, SHANG J F, ZOU H L, et al. The differential value of PCT and PSPN in infectious febrile neutropenia during chemotherapy for acute myeloid leukemia[J]. *Journal of Modern Oncology*, 2024, 32(9): 1 703-1 708.
- [38] NONAKA S, SHIRATORI H, SAIJOH Y, et al. Determination of left-right patterning of the mouse embryo by artificial nodal flow[J]. *Nature*, 2002, 418(6 893): 96-99.
- [39] SOUKUP V, YONG L W, LU T M, et al. The Nodal signaling pathway controls left-right asymmetric development in amphioxus[J]. *Evo Devo*, 2015, 6: 5.
- [40] SHEN M M. Nodal signaling: Developmental roles and regulation[J]. *Development*, 2007, 134(6): 1 023-1 034.
- [41] COLAK S, TEN DIJKE P. Targeting TGF- β signaling in cancer[J]. *Trends in Cancer*, 2017, 3(1): 56-71.
- [42] RAHMAN M S, AKHTAR N, JAMIL H M, et al. TGF- β /BMP signaling and other molecular events: Regulation of osteoblastogenesis and bone formation[J]. *Bone Research*, 2015, 3: 15 005.
- [43] 付玲珠, 郑 婷, 张永生. TGF- β /Smad 信号转导通路及肝纤维化研究进展[J]. *中国临床药理学与治疗学*, 2014, 19(10): 1 189-1 195.
- FU L ZH, ZHENG T, ZHANG Y SH. Advances in understanding the role of TGF- β /Smad signalling pathways in the pathogenesis of liver fibrosis[J]. *Chinese Journal of Clinical Pharmacology and Therapeutics*, 2014, 19(10): 1 189-1 195.
- [44] 刘 镭, 赵琴平, 董惠芬, 等. TGF- β 信号传导通路及其生物学功能[J]. *中国病原生物学杂志*, 2014, 9(1): 77-83.
- LIU R, ZHAO Q P, DONG H F, et al. The TGF- β signaling pathways and their biological functions[J]. *Journal of Pathogen Biology*, 2014, 9(1): 77-83.
- [45] 孙昭恒, 丁 樱, 郭 婷, 等. 转化生长因子 β 1 通路在 IgA 肾病中作用机制的研究进展[J]. *广西医学*, 2023, 45(7): 843-849.
- SUN ZH H, DING Y, GUO T, et al. Mechanism of transforming growth factor β 1 pathway in IgA nephropathy: A research progress[J]. *Guangxi Medical Journal*, 2023, 45(7): 843-849.