成骨细胞特异因子 2 基因家族的全基因组鉴定、分类及进化分析

https://doi.org/10.3969/j.issn.2095-4344.2348
投稿日期：2020-03-09
送审日期：2020-03-26
采用日期：2020-04-16
在线日期：2020-08-18
中图分类号：
R459.9；R336；R34
文章编号：
2095-4344(2021)08-01243-06
文献标识码：B

谭婧玉，刘海文

文章快速阅读：

文章特点—

△首次通过生物信息学方法对成骨细胞特异因子基因家族进行了系统的比较分析；

△以公布的 8 个物种的基因组数据为基础，对成骨细胞特异因子基因保守结构、蛋白特性及系统发育关系进行系统的比较分析，揭示该基因家族在生长发育中的作用和物种进化中的变化。

由 8 个物种成骨细胞特异因子基因组数据 → 方法：依据 HMMER 构建成骨细胞特异因子基因家族的 HMM 模型。 → 观察结果：(1) 基因结构；(2) 结构域特征；(3) 生化特征；(4) 进化关系；(5) 表达特性。

文题释义：

成骨细胞特异性因子2：即Periostin，是由成骨细胞及其前体细胞分泌产生的一种细胞外基质蛋白，参与细胞外基质的重建，对基质蓄积、细胞基质相互作用及纤维化进行调节，具有促进细胞黏附及细胞迁移等功能。

保守结构域：是指在生物进化中一个蛋白家族中具有不变或相同的结构域，具有重要的功能，不能被改变。结构域是蛋白质中由不同二级结构和超二级结构组合而成的独立的稳定结构区域，是蛋白质功能单元。

摘要

背景：目前，Fasciclin基因家族的保守结构、蛋白特性、系统发育关系尚无系统的研究。

目的：探讨成骨细胞特异因子基因家族的进化历史，对成骨细胞特异因子基因家族的保守结构、蛋白特性、系统发育关系进行系统的比较分析。

方法：实验基于已公布的人、黑猩猩、斑马鱼等8个物种的基因组数据，依据HMMER软件构建成骨细胞特异因子基因家族的HMM模型，分析了其基因结构、结构域特征、生化特征、进化关系和表达特性。

结果与结论：①在鱼类(斑马鱼)、两栖类(非洲爪蟾)、鸟类(鸡)、爬行类(中华鳖)、哺乳类(大鼠和小鼠)和灵长类(黑猩猩和人)等8个物种中，各检索出5，5，4，5，5，4，4，4个成骨细胞特异因子候选基因；②通过对成骨细胞特异因子基因在8个物种中的系统发育分析，发现成骨细胞特异因子基因家族整个进化树可以分为4个进化分支，其中2个进化分支的基因具有更强的保守性，可能是主要的成骨细胞特异因子；其中一个分支则可能是两栖类、哺乳类和灵长类特有的成骨细胞特异因子；另外一个分支也具有较好的保守性，但是非洲爪蟾和小鼠在该进化分支中有2个拷贝基因，这个进化分支可能对这两类物种起着重要的作用；③上述数据说明，在物种进化过程中，该家族基因并没有发生较大变化，这类基因可能是重要的生长发育相关基因，推测这类基因的缺失可能会导致个体的死亡或导致物种的灭绝。

关键词：骨；动物；实验；基因；基因组；细胞；结构；蛋白

Genome-wide identification, classification and phylogenetic analysis of Fasciclin gene family for osteoblast specific factor 2

Tan Jingyu, Liu Haiwen
The First Affiliated Hospital of Jinzhou Medical University, Jinzhou 121001, Liaoning Province, China
Tan Jingyu, Master, Physician, The First Affiliated Hospital of Jinzhou Medical University, Jinzhou 121001, Liaoning Province, China
Corresponding author: Liu Haiwen, MD, Associate research fellow, The First Affiliated Hospital of Jinzhou Medical University, Jinzhou 121001, Liaoning Province, China

Abstract

BACKGROUND: There is no systematical report on the conserved structure, protein characteristics and phylogenetic relationship of the Fasciclin gene family.

锦州医科大学附属第一医院，辽宁省锦州市 121001
第一作者：谭婧玉，女，1990 年生，吉林省人，汉族，2016 年锦州医科大学毕业，硕士，医师。
通讯作者：刘海文，博士，副研究员。锦州医科大学附属第一医院，辽宁省锦州市 121001
https://orcid.org/0000-0002-0556-7837 (谭婧玉)
基金资助：辽宁省自然科学基金资助项目 (2019JH8/10300032)，项目负责人：刘海文
引用本文：谭婧玉，刘海文 . 成骨细胞特异因子 2 基因家族的全基因组鉴定、分类及进化分析 [J]. 中国组织工程研究，2021，25(8):1243-1248.

OBJECTIVE: To investigate the evolutionary history of the Fasciclin gene family, and to systematically analyze the conservative structure, protein characteristics, and phylogenetic relationships of this gene family
METHODS: Based on the published genome data of eight species such as people, chimpanzee and zebrafish, HMM models were established using the HMMER software to analyze the characteristics and genetic structure, structure domain and biochemical characteristics, evolutionary relationships and expression characteristics of the Fasciclin gene family.
RESULTS AND CONCLUSION: In the eight species of fish (zebrafish), amphibian (xenopus), bird (chicken), reptile (*Trionyx sinensis*), mammal (rat, mouse) and primate (chimpanzee, human), 5, 5, 4, 5, 5, 4, 4, and 4 fasciclin candidate genes were retrieved respectively. Through the phylogenetic analysis of the fasciclin gene in these eight species, the evolutionary tree of fasciclin gene families can be divided into four evolution branch. The two branches of all are more conservative, likely to be the main osteoblast specific factor; the third branch may be osteoblast specific factor of amphibians, mammals and primate; and the final branch also has a good conservative, but the Xenopus and mouse in the branch has two copies of the genes, and this branch may play an important role in the evolution of these two species. Overall, in the evolution of species, the gene family has no larger changes, and such genes may be important for the growth and development. The lack of such genes may lead to the death of the individual or the extinction of species.
Key words: bone; animal; experiment; gene; genome; cell; structure; protein

Funding: *the Natural Science Foundation of Liaoning Province, No. 2019JH8/10300032 (to LHW)*
How to cite this article: *TAN JY, LIU HW. Genome-wide identification, classification and phylogenetic analysis of Fasciclin gene family for osteoblast specific factor 2. Zhongguo Zuzhi Gongcheng Yanjiu. 2021;25(8):1243-1248.*

0 引言 Introduction

成骨细胞 (osteoblast，OB) 主要由骨膜和骨髓基质的间质细胞分化而来 [1]，具有特异的分泌多种生物活性物质的功能，与破骨细胞协同作用，对骨的形成和重建过程进行调节，维持骨骼的代谢平衡 [2-3]。成骨细胞特异性因子 2 是成骨细胞及其前成骨细胞分泌的一种细胞外基质蛋白，具有促进细胞黏附和细胞迁移等功能 [4-5]。细胞外基质蛋白除了参与细胞外基质的重建，还对基质蓄积及细胞基质相互作用及纤维化进行调节 [6]。成骨细胞特异性因子 2 分子包含一个特有的信号序列，结构分析显示属 Fasciclin 家族 [7]。目前，Fasciclin 基因家族的保守结构、蛋白特性及系统发育关系尚无系统的研究。

为探讨成骨细胞特异因子 Fasciclin 基因家族的进化历史，实验对鱼类 (斑马鱼)、两栖类 (非洲爪蟾)、鸟类 (鸡)、爬行类 (中华鳖)、哺乳类 (大鼠及小鼠) 和灵长类 (黑猩猩及人)8 个物种 Fasciclin 基因家族的起源、物种进化过程中的演化历史进行深入的研究，从 Fasciclin 基因家族的保守结构、蛋白特性及系统发育关系进行系统的比较分析和归纳总结，为研究动物中特定的组织或器官的形成和疾病产生相关基因家族的起源和进化提供了一个独特的视角。

1 资料和方法 Data and methods

1.1 设计 基因组学分析。

1.2 时间及地点 于 2019 年 2 月至 8 月在锦州医科大学附属第一医院临床生物信息学研究所完成。

1.3 资料来源 从 Ensembl Genome Browser 数据库 (http://asia.ensembl.org/) 下载得到斑马鱼、非洲爪蟾、鸡、中华鳖、小鼠、大鼠、黑猩猩及人类的基因组数据信息，见**表 1**，这些数据信息包括基因编码区序列、基因位置信息和基因蛋白序列。从 Pfam (http://pfam.sanger.ac.uk/) 获得 Fasciclin 基因家族保守结构域的 Profile HMM (PF02469.22) 文件信息。

表 1 | 各种动物的数据信息
Table 1 | Information of animals

物种中文名称	物种英文名称	物种拉丁名称	基因组大小 (bp)	染色体和质粒数目	GenBank ID
斑马鱼	Zebrafish	*Danio rerio*	1 412 464 843	26	GCA_000002035.2
非洲爪蟾	Xenopus	*Xenopus tropicalis*	1 511 717 716	/	GCA_000004195.1
鸡	Chicken	*Gallus gallus domesticus*	1 046 932 099	34	GCA_000002315.2
中华鳖	Chinese soft shell turtle	*Pelodiscus sinensis*	2 202 466 388	/	GCA_000230535.1
小鼠	Mouse	*Mus musculus*	2 796 638 596	22	GCA_000001635.3
大鼠	Rat	*Rattus norregicus*	2 909 698 938	22	GCA_000001895.3
黑猩猩	Chimpanzee	*Pan troglodytes*	3 323 267 922	26	GCA_000001515.4
人类	Human	*Homo sapiens*	3 228 143 073	25	GCA_000001405.12

1.4 实验方法

1.4.1 Fasciclin 基因分析 根据从 Pfam 下载得到的 Fasciclin 基因家族保守结构域的的 Profile HMM 文件，利用 HMMER 3.0 软件的 HMMsearch 检索斑马鱼、非洲爪蟾、鸡、中华鳖、小鼠、大鼠、黑猩猩及人类全基因组的蛋白质序列文件，以“Trusted cutoff”作为筛选条件，并以 e 值 1e-05 做为阈值，获取 8 个物种中 Fasciclin 蛋白家族的候选基因。利用 InterProScan (https://www.ebi.ac.uk/interpro/about/interproscan/) 对候选基因进行分析 [8]，并结合 Fasciclin 蛋白家族保守结构域的 Profile HMM profile 信息 (PF02469.22)，验证其编码蛋白中存在 Fasciclin 相关结构域，确定 8 个物种的 Fasciclin 蛋白家族的目标基因。

1.4.2 Fasciclin 蛋白保守结构域和 Motif 结构分析 利用 InterProScan (https://www.ebi.ac.uk/interpro/about/interproscan/) 和 MEME(http://meme.nbcr.net/meme/cgi-bin/meme.cgi) 两个在线蛋白质序列分析工具 [9]，以默认值为阈值，对获取的 8 个物种中的 Fasciclin 基因的蛋白质序列进行保守结构域和保守的 Motif 分析，以确定 Fasciclin 基因的蛋白质序列的结构特征。

1.4.3 Fasciclin 蛋白信号肽和蛋白特性分析 利用蛋白质信号肽在线分析软件 TargetP 1.1(http://www.cbs.dtu.dk/services/TargetP/)[10]，以默认值为筛选条件，对 Fasciclin 蛋白质序列的信号肽区域进行分析，以获取 8 个物种 Fasciclin 基因的蛋白质序列的信号肽信息。利用在线分析软件 ProtParam(http://web.expasy.org/protparam/)[11]，以默认值为筛选条件，对 Fasciclin 蛋白的等电点 (Theoretical pI)、理论分子量 (molecular weight，Mw) 和 GRAVY(Grand average of hydropathy) 进行分析，以获取 8 个物种 Fasciclin 基因的蛋白质信息。

1.4.4 构建 Fasciclin 蛋白家族的系统发育树 利用 ClustalW[12]，以 Gap Opening 值为 10，Gap Extension 值为 10，以及 Protein

Weight Matrix 设置为 Gonnet series， Alignment 设置为 Do Complete Alignment 为参数，对获取的 8 个物种的 Fasciclin 蛋白序列进行多序列比对，利用 MEGA 5.0(http://megasoftware.net) 软件 [13]，Bootstrap 值设置为 1 000，采用 Neighbor-Joing 模型构建系统发育树。根据 Pfam 结构域类型和进化关系对各家族成员进行分类。

1.5　主要观察指标　各物种的基因结构、结构域特征、生化特征、进化关系和表达特性。

2　结果 Results

2.1　Fasciclin 基因的鉴别与分析　利用 Fasciclin 基因家族保守结构域的 HMM profile，从选取的 8 个物种中共鉴定出 36 个 Fasciclin 基因，见**表 2**，其中斑马鱼 5 个、非洲爪蟾 5 个、鸡 4 个、中华鳖 5 个、小鼠 5 个、大鼠 4 个、黑猩猩 4 个、人类 4 个；蛋白数目 36 628–58 074 个。

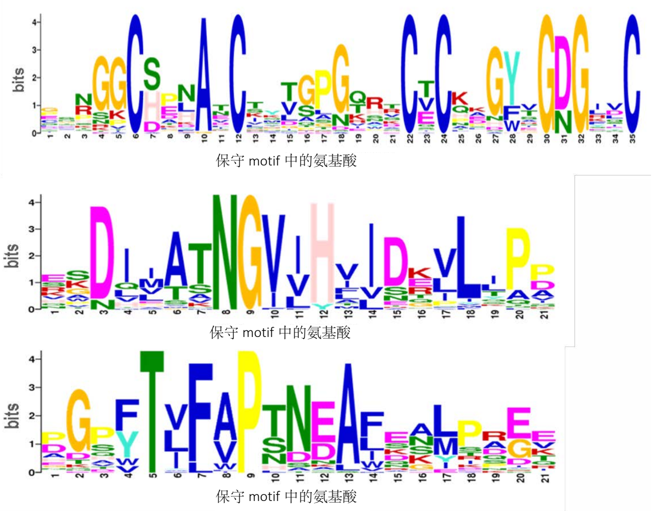
表 2｜各物种中 Fasciclin 基因的鉴别与分析
Table 2｜ Identification and analysis of Fasciclin genes in various species

物种名称	英文名称	Fasciclin 基因	cDNA 的数目	mRNA 的数目	Protein 数目	基因数占蛋白数百分比 (%)
斑马鱼	Zebrafish	5	36 628	4 714	36 628	0.013 651
非洲爪蟾	Xenopus	5	43 859	1 306	43 859	0.011 400
鸡	Chicken	4	40 571	1 558	40 571	0.009 859
中华鳖	Chinese soft shell turtle	5	48 910	1 042	48 910	0.010 223
小鼠	Mouse	5	56 845	8 036	56 845	0.008 796
大鼠	Rat	4	58 074	1 713	58 074	0.006 888
黑猩猩	Chimpanzee	4	56 687	8 681	56 687	0.007 056
人类	Human	4	48 418	20 168	48 418	0.008 261

2.2　Fasciclin 蛋白保守结构域和 Motif 的分析　利用 InterProScan (http://www.ebi.ac.uk/Tools/services/web_iprscan/toolform.ebi) 对获取的蛋白序列进行保守结构域的检索，并对保守结构域数目及位点进行详细的分析，见**表 3**。

斑马鱼的 Fasciclin 蛋白长度 566–2 025，共有 24 个保守结构域；鸡的 Fasciclin 蛋白长度 2 044–3 652，共有 20 个保守结构域；人类的 Fasciclin 蛋白长度 710–2 587，共有 21 个保守结构域；小鼠的 Fasciclin 蛋白长度 271–2 062，共有 19 个保守结构域；中华鳖的 Fasciclin 蛋白长度 924–2 608，共有 19 个保守结构域；黑猩猩的 Fasciclin 蛋白长度 686–2 677，共有 20 个保守结构域；大鼠的 Fasciclin 蛋白长度 861–2 847，共有 19 个保守结构域；非洲爪蟾的 Fasciclin 蛋白长度 529–1 655，共有 15 个保守结构域。

利用在线软件 MEME(http://meme.nbcr.net/meme/) 对 Fasciclin 蛋白的 Motif 结构进行分析，分析结果如**图 1**(Logo)。结果发现 Fasciclin 蛋白家族有 3 个潜在的 Motif，分别含有 35，21 和 21 个氨基酸残基，Motif 1 序列为 xx[NR] GGC[SH][PE][NHL]ATCTY[TV] [GS]PG[QT]RxC[TVE]C[KQ] xG Y[VT] G[DN]IVC，Motif 2 序列为 E[SK]DI[IM]ATNGV [IV]HVID[KE][VL] L[IL]P[PD]，Motif 3 序列为 [PD]GP[FY]T[VLI]F[AV]P [TS]N[ED] A[FL] E[AN][LM]PR[EG]E。Fasciclin 基因家族总共有 3 个高度保守的 Motif，推测可能是主要的执行功能的结构域。



图注：此 3 个图分别代表通过多序列比对获取的保守的氨基酸中 Motif1、Motif2 和 Motif3；纵坐标：Bits 或者氨基酸数目；单位：AA；大字母代表较高的保守性，小字母代表较低的保守性。

图 1 ｜ Fasciclin 蛋白的 Motif 结构分析
Figure 1 ｜ Motif structure analysis of Fasciclin proteins

2.3　Fasciclin 蛋白信号肽分析　利用在线分析软件 TargetP 1.1(http://www.cbs.dtu.dk/services/TargetP/) 对 Fasciclin 蛋白的信号肽区域进行预测分析，结果显示在小鼠中有 2 个信号肽，一个可能为分泌蛋白，一个可能为线粒体蛋白；在黑猩猩中有 2 个信号肽，可能为分泌蛋白；大鼠有 4 个信号肽，3 个可能为分泌蛋白，一个可能为线粒体蛋白；非洲爪蟾中有 1 个信号肽，可能为线粒体蛋白；在斑马鱼、人和鸡中总共有 8 个信号肽，可能为分泌蛋白，其他物种未知，见**表 4**。

2.4　Fasciclin 蛋白的生化特性　利用在线分析软件 ProtParam(http://web.expasy.org/protparam/) 对 Fasciclin 蛋白的等电点 (Theoretical pI)、理论分子量及 GRAVY(Grand average of hydropathy) 进行分析，结果显示 Fasciclin 蛋白家族的成员为偏酸性的亲水性蛋白，且结构稳定，同时该类蛋白的长度变化较大，范围在 529–2 847 个氨基酸，其理论分子质量差异较大，在 57.301–308.514 kD，见**表 5**。

2.5　Fasciclin 蛋白序列比对和系统发育学分析　采用 ClustalW 对获取的 Fasciclin 蛋白序列进行多序列比对，并进行保守结构域的分析。

实验截取了 Fasciclin 蛋白序列多序列比对结果中的 3 段比较保守的序列区域进行分析，以小鼠基因 MmGENSCAN00000039557 为例，3 段区域分别为：第 1 段序列保守区域位于 87–187 个氨基酸残基之间；第 2 段序列保守区域位于 285–385 个氨基酸残基之间；第 2 段序列保守区域位于 482–582 个氨基酸残基之间。分析发现，Ala113，Asn115，Gly116，His119 和 Asn 172 在第 1 段 Fasciclin 蛋白质序列的多序列比对中具有较好的保守性，其中 Asn115 和 Gly116 是 2 个氨基酸残基连续的 1 个小肽段；G287，H327，L338，C362 和 G375 共 5 个氨基酸残基在第 2 段 Fasciclin 蛋白质序列的多序列比对中具有较好的保守性；而在第 3 段多序列比

表 3 | Fasciclin 基因家族保守结构域分析统计表
Table 3 | Conserved domain analysis of the Fasciclin gene family

物种名称	基因名称	蛋白长度	保守结构域数目	保守结构域位点
斑马鱼	DrGENSCAN00000005997	2 025	5	[308，412] [423，521] [797，945] [1 103，1 206] [1 234， 1 280]
	DrGENSCAN00000013514	1 008	4	[112，230] [246，367] [383，494] [508，630]
	DrGENSCAN00000016182	969	4	[84，203] [219，340] [355，435] [452，574]
	DrGENSCAN00000021703	566	4	[109，229] [245，366] [381，493] [507，554]
	DrGENSCAN00000032200	1 849	7	[236， 339] [363，451] [623，727] [748，864] [945，1 023] [1 036， 1 156] [1 635， 1 741]
鸡	GgGENSCAN00000023786	3 652	5	[1629，1 728] [1755，1 879] [2 211，2 312] [2 821，2 931] [2 957，3 086]
	GgGENSCAN00000024417	2 044	4	[656，738] [823，940] [955，1 067] [1 081，1 201]
	GgGENSCAN00000028901	2 645	7	[574，688] [695，816] [1 147，1 266] [1 267，1 435] [1 749，1 849] [1 875，2 002] [2 440，2540]
	GgGENSCAN00000034897	2 536	4	[1 728，1 848] [1 863，1 984] [1 998，2 111] [2 125，2 240]
人类	HsGENSCAN00000009131	2 587	6	[380，486] [514，639] [996，1 116] [1 136，1 226] [1 647，1 751] [1 811，1 866]
	HsGENSCAN00000027083	2 347	7	[393，477] [487，589] [906，1 015] [1 016，1 122] [1 460，1 537] [1 549，1 677] [2 115，2 206]
	HsGENSCAN00000032421	710	4	[123，236] [252，373] [387，500] [514，634]
	HsGENSCAN00000039628	1 082	4	[111，228] [243，364] [378，491] [505，627]
小鼠	MmGENSCAN00000004871	271	4	[461，574] [931，1 051] [1 512，1 616] [1 673，1 729]
	MmGENSCAN00000012601	2 062	4	[3，119] [650，757] [783， 913] [1 439，1 529]
	MmGENSCAN00000012636	1 067	3	[296，379] [391，509] [924，1 032]
	MmGENSCAN00000039557	597	4	[14，127] [143，264] [278，391] [405，525]
	MmGENSCAN00000054729	774	4	[107， 187] [202， 323] [337， 450] [466， 582]
中华鳖	PsGENSCAN00000019001	924	4	[192，311] [328，448] [462，558] [562，680]
	PsGENSCAN00000022691	1 878	1	[1 743，1824]
	PsGENSCAN00000033251	2 608	6	[713，833] [1 240，1 349] [1 350，1 456] [1 639，1 743] [1 772，1 902] [2 397，2 522]
	PsGENSCAN00000040902	2 355	4	[579，645]
	PsGENSCAN00000042302	1 016	4	[356，431] [445，566] [580，693] [707，829]
黑猩猩	PtGENSCAN00000002705	2 677	6	[470，576] [604，729] [1 086，1 206] [1 226，1 316] [1 737，1 841] [1 907，1956]
	PtGENSCAN00000031176	779	4	[113，199] [214，335] [349，462] [476，598]
	PtGENSCAN00000034467	2 343	6	[393，477] [879，988] [989，1 095] [1 456， 1 533] [1 545，1 673] [2 111，2 202]
	PtGENSCAN00000044748	686	4	[46，159] [175，296] [310，423] [437，542]
大鼠	RnGENSCAN00000012813	2 847	5	[749，855] [874，987] [1 338，1 458] [1 916，2 020] [2 046，2 111]
	RnGENSCAN00000019244	2 216	6	[225，346] [352，442] [940，1 051] [1 053，1 158] [1 430，1 537] [1 563，1 693]
	RnGENSCAN00000043944	940	4	[113，233] [248，369] [438，521] [658，774]
	RnGENSCAN00000049708	861	4	[329，412] [428，549] [563，676] [690，810]
非洲爪蟾	XtGENSCAN00000013074	1 655	5	[211，310] [322，419] [904，1 008] [1 034，1 165] [1 423，1 525]
	XtGENSCAN00000021180	834	1	[57，164]
	XtGENSCAN00000021616	723	4	[102，219] [236，357] [371，484] [498，618]
	XtGENSCAN00000032296	529	1	[6，66]
	XtGENSCAN00000037793	971	4	[162，280] [302，358] [372，485] [499，622]

表 4 | Fasciclin 蛋白信号肽分析
Table 4 | Fasciclin protein signal peptide analysis

物种名称	信号肽数量	分泌蛋白数量	线粒体蛋白数量
小鼠	2	1	1
黑猩猩	2	2	0
大鼠	4	3	1
非洲爪蟾	1	0	1
斑马鱼	3	3	0
人类	3	3	0
鸡	2	2	0
中华鳖	未知	未知	未知

对序列中含有 Asp506 和 Leu521 两个氨基端残基具有保守性，以及 1 个连续的氨基酸 Asn511 和 Gly512 肽段具有保守性。从物种的系统发育关系分析结果显示，系统进化树可以分为 4 个进化分支：Group Ⅰ，Ⅱ，Ⅲ，Ⅳ，见图 **2**。Group Ⅰ和 Group Ⅱ进化分支的基因具有更强的保守性，推测可能是主要的成骨细胞特异因子；Group Ⅲ推测是两栖类、哺乳类和灵长类特有的成骨细胞特异因子；Group Ⅳ也具有较好的保守性，但非洲爪蟾和小鼠在该进化分支中有 2 个拷贝基因，推测这个进化分支对这 2 类物种起着重要的作用。

3 讨论 Discussion

成骨细胞特异性因子 2 是一种高度保守的基质细胞蛋白 [14]，分子中包含一个特有的信号序列，结构分析显示属 Fasciclin 家族。实验从比较基因组学角度出发，对成骨细胞特异因子 2 Fasciclin 基因家族进行了详细的比较分析，并从 Fasciclin 基因家族的保守结构、蛋白特性及系统发育关系进行系统的比较分析和归纳总结，为揭示该基因家族在生长发育中的功能和进化历史或演化规律提供重要参考。

通过 Fasciclin 基因家族的保守结构来比较分析发现，该家族总共有 3 个高度保守的 motif，这 3 个保守的 motif 推测可能是主要的执行功能的结构域。

实验通过对 Fasciclin 基因家族的蛋白特性分析发现，发现 Fasciclin 蛋白家族的成员为偏酸性的亲水性蛋白，且结构稳定，同时该类蛋白的长度变化较大，范围为 529–2 847 个氨基酸残基，相应的理论分子量也差异较大，分子质量 57.301–308.514 kD。通过对 Fasciclin 基因家族信号肽的比较分析，发现在小鼠中有一个可能为分泌蛋白，一个可能为线粒体蛋白，在黑猩猩中有 2 个可能为分泌蛋白，大鼠有 3 个可能为分泌蛋白，一个可能为线粒体蛋白，蟾中有一个可能为线粒体蛋白，在斑马鱼、人和鸡中，预测的分泌蛋白占

表 **5** | Fasciclin 蛋白的生化特性

Table 5 | **Biochemical properties of Fasciclin proteins**

物种名称	基因名称	长度	分子质量 (kD)	等电点	疏水性	不稳定指数	定位
斑马鱼	DrGENSCAN00000005997	2 025	220.698	6.15	−0.289	44.44	Secretory pathway
	DrGENSCAN00000013514	1 008	111.575	7.31	−0.179	35.22	Secretory pathway
	DrGENSCAN00000016182	969	107.247	7.00	−0.074	35.62	Secretory pathway
	DrGENSCAN00000021703	566	61.997	6.96	−0.056	32.67	Secretory pathway
	DrGENSCAN00000032200	1 849	202.359	5.77	−0.189	41.42	Other location
鸡	GgGENSCAN00000023786	3 652	401.127	6.17	−0.246	45.04	Other location
	GgGENSCAN00000024417	2 044	225.241	8.33	−0.196	46.70	Other location
	GgGENSCAN00000028901	2 645	288.956	6.04	−0.292	37.17	Secretory pathway
	GgGENSCAN00000034897	2 536	281.903	7.11	−0.163	45.10	Other location
人类	HsGENSCAN00000009131	2 587	277.368	5.99	−0.196	45.83	Secretory pathway
	HsGENSCAN00000027083	2 347	254.155	6.02	−0.275	41.69	Other location
	HsGENSCAN00000032421	710	77.511	6.48	−0.051	32.66	Secretory pathway
	HsGENSCAN00000039628	1 082	121.599	8.69	−0.273	41.00	Secretory pathway
小鼠	MmGENSCAN00000004871	2 371	254.860	6.22	−0.111	41.89	Mitochondrion
	MmGENSCAN00000012601	2 062	224.820	6.53	−0.343	44.17	Other location
	MmGENSCAN00000012636	1 067	445.680	5.96	−0.300	39.63	Other location
	MmGENSCAN00000039557	597	64.919	5.64	−0.043	33.43	Other location
	MmGENSCAN00000054729	774	86.050	7.02	−0.134	33.85	Secretory pathway
中华鳖	PsGENSCAN00000019001	924	100.708	6.41	−0.202	42.63	Other location
	PsGENSCAN00000022691	1 878	207.866	6.20	−0.217	44.52	Other location
	PsGENSCAN00000033251	2 608	285.136	6.75	−0.318	39.19	Other location
	PsGENSCAN00000040902	2 355	252.630	6.91	−0.169	43.34	Other location
	PsGENSCAN00000042302	1 016	113.51	5.59	−0.335	41.90	Other location
黑猩猩	PtGENSCAN00000002705	2 677	287.475	5.95	−0.203	45.93	Secretory pathway
	PtGENSCAN00000031176	779	87.296	7.85	−0.173	36.06	Secretory pathway
	PtGENSCAN00000034467	2 343	253.457	5.96	−0.271	40.69	Other location
	PtGENSCAN00000044748	686	75.170	6.03	−0.184	39.44	Other location
大鼠	RnGENSCAN00000012813	2 847	308.514	6.04	−0.106	46.63	Mitochondrion
	RnGENSCAN00000019244	2 216	241.154	5.98	−0.320	40.44	Secretory pathway
	RnGENSCAN00000043944	940	104.946	6.67	−0.200	35.44	Secretory pathway
	RnGENSCAN00000049708	861	94.590	7.31	−0.186	43.54	Secretory pathway
非洲爪蟾	XtGENSCAN00000013074	1 655	182.386	5.54	−0.144	34.71	Other location
	XtGENSCAN00000021180	834	88.749	5.12	−0.107	39.92	Other location
	XtGENSCAN00000021616	723	80.029	6.26	−0.107	33.71	Other location
	XtGENSCAN00000032296	529	57.301	6.15	−0.372	48.01	Mitochondrion
	XtGENSCAN00000037793	971	107.819	6.05	−0.169	30.45	Other location

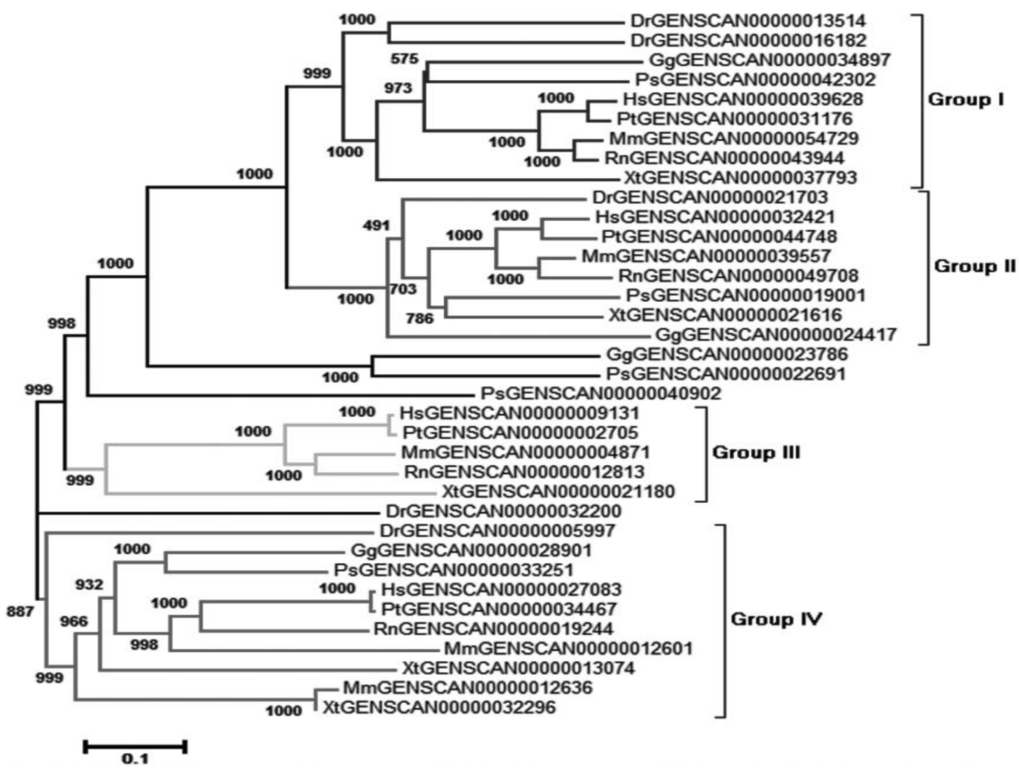


图 **2** | Fasciclin 蛋白的系统进化树

Figure 2 | **Phylogenetic tree of Fasciclin proteins**

大多数，3 个物种中共有 8 个可能为分泌蛋白。这就意味着 Fasciclin 蛋白家族的成员在行使功能的过程中有其组织和器官的特异性。

从 Fasciclin 蛋白家族的系统发育关系分析发现，Fasciclin 蛋白基因家族的基因具有严格的保守性，是主要的功能基因。整个进化树可以分为 4 个进化分支，其中的 2 个进化分支的基因具有更强的保守性，推测可能是主要的成骨细胞特异因子；其中一类则可能是两栖类、哺乳类和灵长类特有的成骨细胞特异因子；另外一个分支也具有较好的保守性，但是非洲爪蟾和小鼠在该进化分支中有 2 个拷贝基因，推测这个进化分支对这两类物种起着重要的作用。在物种进化的过程中，该家族的基因并没有发生较大的变化，推测这类基因可能是重要的生长发育相关基因，这类基因的缺失可能会导致个体的死亡或导致物种的灭绝。

成骨细胞特异性因子 2 在长骨的骨膜中表达，也在心脏、肾脏、皮肤和肺等许多其他组织和器官中表达 [15-16]，可作为临床生物标记物进行癌症、呼吸系统疾病、肾衰竭、肾损伤和心肌梗死等疾病的诊治 [17-18]。研究该基因家族的进化历史或演化规律，对于研究物种的形成和演化，更好的服务于人类健康和研究具有重要的意义。

作者贡献：实验设计和评估为为通讯作者。资料收集和数据分析为第一作者。

经费支持：该文章接受了“辽宁省自然科学基金资助项目 (2019JH8/10300032)”的资助。所有作者声明，经费支持没有影响文章观点和对研究数据客观结果的统计分析及其报道。

利益冲突：文章的全部作者声明，在课题研究和文章撰写过程，没有因其岗位角色影响文章观点和对数据结果的报道，不存在利益冲突。

写作指南：该研究遵守国际医学期刊编辑委员会《学术研究实验与报告和医学期刊编辑与发表的推荐规范》。

文章查重：文章出版前已经过专业反剽窃文献检测系统进行 3 次查重。

文章外审：文章经小同行外审专家双盲外审，同行评议认为文章符合期刊发稿宗旨。

生物统计学声明：文章统计学方法已经通过锦州医科大学生物统计学专家审核。

4 参考文献 References

[1] SHI C, WU T, HE Y, et al. Recent advances in bone-targeted therapy. Pharmacol Ther. 2020;207:107473.

[2] 靳慧 , 葛娅娜 , 张成仁 , 等 . 成骨细胞中基因表达的调控 [J]. 中国组织化学与细胞化学杂志 , 2011,20(5):499-501.

[3] 程余婷 , 伍超 , 黄晓林 , 等 . 低剂量唑来膦酸对去势拔牙大鼠破骨及成骨细胞的影响 [J]. 中国组织工程研究 ,2020,24(17): 2686-2693.

[4] 王春杰 , 刘茂东 , 王斯 , 等 . 成骨细胞特异性因子 2 的研究进展 [J]. 中华临床医师杂志 (电子版),2013,7(9):4020-4022.

[5] XU D, MA HM, XIAO DF. Molecular characterization, tissue expression profile, and single nucleotide polymorphism analysis of the periostin gene in swine. Genet Mol Res. 2016;15(1):15017187.

[6] TAKARADA T, HINOI E, NAKAZATO R, et al. An analysis of skeletal development in osteoblast-specific and chondrocyte-specific runt-related transcription factor-2 (Runx2) knockout mice. J Bone Miner Res. 2013;28(10):2064-2069.

[7] SUGIURA T, TAKAMATSU H, KUDO A, et al. Expression and characterization of murine osteoblast-specific factor 2 (OSF-2) in a baculovirus expression system. Protein Expr Purif. 1995;6(3):305-11.

[8] SWATI S, VINITA H. Identification of coding sequence and its use for functional and structural characterization of catalase from Phyllanthusemblica. Bioinformation. 2018;14(1):8-14.

[9] 黎帮勇 , 胡尚连 , 曹颖 , 等 . 毛竹 NAC 转录因子家族生物信息学分析 [J]. 基因组学与应用生物学 ,2015,34(8):1769-1777.

[10] LEHNEN M, EBERT BE, BLANK LM. A comprehensive evaluation of constraining amino acid biosynthesis in compartmented models for metabolic flux analysis. Metab Eng Commun. 2017;5:34-44.

[11] GARG VK, AVASHTHI H, TIWARI A, et al. MFPPI-Multi FASTA ProtParam Interface. Bioinformation. 2016;12(2):74-77.

[12] 邓小燕 , 徐胜超 . 基于 Yarn 云平台的生物基因多序列比对并行算法 [J]. 基因组学与应用生物学 ,2019,38(7):3009-3015.

[13] 冀志蕊 , 董庆龙 , 谷彦冰 , 等 . 砀山酥梨 TCP 基因家族全基因组鉴定与分析 [J]. 基因组学与应用生物学 ,2017,36(2):766-773.

[14] BONNET N, GARNERO P, FERRARI S. Periostin action in bone. Mol Cell Endocrinol. 2016;432:75-82.

[15] MUROTA H, LINGLI Y, KATAYAMA I. Periostin in the pathogenesis of skin diseases. Cell Mol Life Sci. 2017;74(23):4321-4328.

[16] DAVID N, BETHANY BM. The role of periostin in lung fibrosis and airway remodeling. Cell Mol Life Sci. 2017;74(23):4305-4314.

[17] IDOLAZZI L, RIDOLO E, FASSIO A, et al. Periostin: The bone and beyond. Eur J Intern Med. 2017;38:12-16.

[18] GONZÁLEZ-GONZÁLEZ L, ALONSO J. Periostin: a matricellular protein with multiple functions in cancer development and progression. Front Oncol. 2018;8:225.