

DOI: 10.3724/SP.J.1005.2009.00180

大腹园蛛大壶状腺丝蛋白-1 基因部分 cDNA 的克隆和序列分析

任洪林¹, 张文惠³, 潘风光², 卢士英¹, 柳增善¹, 张俊辉¹

1. 人兽共患病教育部重点实验室, 吉林大学人兽共患病研究所, 长春 130062;
2. 吉林大学军需科技学院, 长春 130062;
3. 福建省亚热带植物研究所, 厦门 361006

摘要: 大腹园蛛大壶状腺表达拖丝蛋白新基因的克隆, 为进一步研究蛛丝蛋白基因以及人工表达蛛丝蛋白提供参考依据。文章利用“通用方法”即反转录—置换法构建大腹园蛛(*Araneus ventricosus*)大壶状腺(Major ampullate gland) cDNA 文库, 并筛选出具有典型重复结构的大腹园蛛大壶状腺丝蛋白-1 部分 cDNA 序列 *AvMaSp1* (GenBank 登录号: AY177203)。该部分序列大小为 1 408 bp, 编码区为 1 288 bp, 编码氨基酸 429 个, 预测分子量为 34.07 kDa, 典型的重复结构为 (GA)_nA_m(GA)_N, 与十字园蛛(*Araneus diadematus*)丝蛋白基因 *ADF-1* (GenBank 登录号: ADU47853)同源关系最近, 一致性为 75.0%。

关键词: 大腹园蛛; 大壶状腺; 蜘蛛丝; 拖丝; 蛛丝蛋白

Cloning and sequence analysis of major ampullate spidroin-1 partial cDNA from *Araneus ventricosus*

REN Hong-Lin¹, ZHANG Wen-Hui³, PAN Feng-Guang², LU Shi-Ying¹, LIU Zeng-Shan¹, ZHANG Jun-Hui¹

1. Key Laboratory of Zoonosis, Ministry of Education, Institute of Zoonosis, Jilin University, Changchun 130062, China;
2. College of Light Industry and Economic, Jilin University, Changchun 130062, China;
3. Fujian Institute of Subtropical Botany, Xiamen 361006, China

Abstract: In this paper, we successfully constructed the cDNA library of major ampullate gland of *Araneus ventricosus* using pUC18 vector and cloned the partial cDNA (*AvMaSp1*, GenBank accession number AY177203) encoding spider major ampullate gland spidroin-1 by means of picking colony randomly (Bird gun). The partial cDNA sequence of *AvMaSp1* was 1 408 bp and its region in encoding was 1 288 bp. The protein deduced from *AvMaSp1* contained 429 amino acid residues and molecular weight was 34.07 kDa. The repetitive motif of this cDNA sequence was (GA)_nA_m(GA)_N and the identity with *A. diadematus* fibroin-1 mRNA (*ADF-1*, GenBank accession number ADU47853) was 75.0%.

Keywords: *Araneus ventricosus*; major ampullate gland; spider silk; dragline silk; spidroin

收稿日期: 2008-05-06; 修回日期: 2008-09-04

基金项目: 国家自然科学基金项目(编号: 30170718)资助

作者简介: 任洪林(1974-), 男, 讲师, 博士, 研究方向: 分子生物学与人兽共患病。Tel: 0431-87836716; E-mail: rainbow86330@163.com

通讯作者: 柳增善(1959-), 男, 博士, 教授, 博士生导师, 研究方向: 食品安全与人兽共患细菌病。Tel: 0431-87836703; E-mail: zslu1959@163.com

卢士英(1973-), 女, 博士, 讲师, 研究方向: 生物技术与人兽共患病。Tel: 0431-87836716; E-mail: lushiyi1129@163.com

蜘蛛丝以其重量最轻、弹性和抗拉性能最强被誉为“生物钢”,其织物防弹性能越来越受到人们的关注。特别是拖丝所能承受的拉力和弹性高于目前西方专用的防弹纤维凯夫拉尔^[1,2](拖丝:抗拉力 5×10^9 N/m²、弹力 35%~50%; Kevlar: 抗拉力 4×10^9 N/m²、弹性 2%~5%),具有更为广阔的应用前景。在军事上用于制作防弹背心、高性能头盔等;在医疗上用于制作手术缝合线、人造血管和肌腱;还可用于制造航天飞机材料以及为极小型电路传输数据的纳米管等,我国传统中医还利用蜘蛛丝的赔干粉沫治疗脱肛。

拖丝由蜘蛛大壶状腺产生,本实验通过构建大腹园蛛(*Araneus ventricosus*)大壶状腺(Major ampullate gland) cDNA 文库的方法,从文库中随机筛选重组克隆子测序分析,获得大壶状腺拖丝蛋白部分 cDNA 序列 *AvMaSp1*,为研究蛛丝基因和蛋白的结构功能特点以及利用生物技术手段表达蛛丝蛋白奠定了基础。

1 材料和方法

1.1 材料

1.1.1 菌株及载体

受体菌 *E.coli* DH5 α 购自 Promega, pUC18 载体购自 TaKaRa。

1.1.2 主要试剂盒及试剂

TRIZOL^R Reagent Total RNA Isolation Reagent 购自 Invitrogen; PolyATtract^R mRNA Isolation Systems 和 Universal RiboClone^R cDNA Synthesis System 购自 Promega; 限制性内切酶 *EcoR*、*CIAP*、DL-2000 Marker、溴化乙啶(TaKaRa); T4 DNA 连接酶(Promega); 琼脂糖(西班牙); DEPC (Serve); RNase A(华美); 氨苄青霉素(Sigma); 其他常用试剂均为国产分析纯。

1.2 方法

解剖大腹园蛛分离大壶状腺,按照 TRIZOL^R Reagent Total RNA Isolation Reagent 操作手册分离总 RNA,按照 PolyATtract^R mRNA Isolation Systems 操作手册分离 mRNA,按照 Universal RiboClone^R cDNA Synthesis System 操作手册合成 cDNA,使用

T4 DNA 连接酶与用 *CIAP* 处理的 pUC18 在 4℃ 连接过夜,转化至 *E.coli* DH5 α 感受态中,进行氨苄(100 μ g/mL)抗性 LB 琼脂平板筛选,将涂有转化液的平板 37℃ 恒温培养 20 h,随机挑取生长菌落在含有氨苄的液体 LB 中 37℃ 200 r/min 培养 18 h,使用传统碱裂解法^[3]小提 37℃ 培养的菌体生长管的质粒,电泳与标准 pUC18 质粒比较大小,大于 pUC18 的质粒进行 *EcoR* 酶切鉴定,选择插入片段 1 500 bp 左右的重组克隆子送样,由上海生工生物工程技术服务有限公司完成基因序列测定。根据蛛丝基因及蛛丝蛋白重复性高的结构特点,分析测序结果,筛选出大腹园蛛大壶状腺丝蛋白-1 部分 cDNA 序列 *AvMaSp1*。使用 DNASTAR 7.0 分析核苷酸和氨基酸序列,在美国国家生物信息中心 NCBI (<http://www.ncbi.nlm.nih.gov/>) 利用 BlastN 和 BlastX 工具进行同源序列搜索比对^[4],使用 Mega 4.0 软件 Neighbor-Joining (NJ) 算法绘制同源蛋白系统进化树,申请 GenBank 新基因登录号。

2 结果与分析

2.1 *AvMaSp1* cDNA 核苷酸和预测氨基酸序列

测序核苷酸序列及预测氨基酸序列见图 1。*AvMaSp1* cDNA 长 1 408 bp,该部分编码区 1 288 bp,编码 429 个氨基酸,预测编码蛋白分子量为 34.07 kDa。在 GenBank 申请新基因登录号为 AY177203,相应的蛋白登录号为 AAN85281。

2.2 *AvMaSp1* cDNA 核苷酸序列分析及同源性比对

AvMaSp1 cDNA 序列具有 mRNA 完整的 3' 端多聚腺苷酸尾部结构和末端加尾 AATAAA 信号(图 1),但不具有完整的 5' 端和起始密码子 ATG 结构。3' 端非编码区长 117 bp,终止密码子 TAA 下游 84 nt 出现 AATAAA 信号, poly(dA) 尾巴出现在 AATAAA 信号下游 15 nt (即终止密码子 TAA 下游 104 nt)。NCBI BlastN 同源搜索结果显示只与十字园蛛(*Araneus diadematus*)丝蛋白基因 *ADF-1* (GenBank 登录号: ADU47853) 同源关系最近。使用 DNASTAR 7.0 软件,根据 ClustalW 计算方法分析,基因序列一致性为 75.0%。

		一致性(%)																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
趋势性(%)	1		45.7	52.5	46.3	50.1	38.7	47.9	39.3	47.9	51.6	46.7	49.3	49.8	47.1	73.7	1	AAC47008
	2	92.0		45.0	53.3	56.8	53.2	56.2	51.1	70.4	72.4	68.7	72.3	82.7	63.9	48.6	2	AAK30595
	3	73.4	94.4		38.9	37.0	35.1	42.3	33.3	47.5	49.0	49.2	50.9	50.7	46.0	59.4	3	AAT08436
	4	90.3	71.3	115.7		85.0	48.2	76.0	46.3	55.2	55.7	54.2	58.8	52.6	52.7	49.1	4	CAM32249
	5	79.5	63.4	123.5	16.7		48.2	82.8	47.6	60.1	60.6	57.5	62.2	56.7	56.5	53.2	5	CAJ00428
	6	116.3	71.7	132.1	84.7	84.7		50.4	83.3	58.8	61.2	57.1	61.9	58.2	73.5	44.0	6	ABD61593
	7	85.8	64.6	103.2	29.0	19.5	78.8		48.7	58.0	57.7	56.3	61.0	55.3	52.8	50.4	7	CAM32250
	8	114.1	76.9	141.2	90.4	86.4	19.0	83.4		57.3	59.0	54.2	58.9	54.1	71.2	42.1	8	ABD61594
	9	85.8	37.6	86.6	66.9	56.3	59.1	60.8	62.2		93.5	80.2	83.3	70.9	67.5	50.1	9	AAT75310
	10	75.6	34.4	82.4	65.7	55.4	54.1	61.5	58.5	6.8		85.1	86.9	76.0	68.2	50.2	10	AAT75311
	11	89.2	40.4	82.0	69.2	61.8	62.7	64.5	69.4	23.0	16.6		89.9	73.9	69.0	52.5	11	ABR37275
	12	81.7	34.5	77.4	59.0	52.3	52.8	54.5	58.9	19.0	14.4	10.8		74.5	66.5	53.4	12	AAT75312
	13	80.5	19.8	78.0	73.1	63.5	60.2	66.8	69.5	36.9	28.9	32.0	31.2		67.0	50.6	13	AAC47009
	14	87.9	48.9	91.1	72.8	64.1	32.8	72.8	36.4	42.5	41.3	40.0	44.2	43.4		46.1	14	AAR13811
	15	32.4	83.8	57.7	82.4	71.6	97.5	78.8	104.0	79.5	79.4	73.3	71.1	78.2	91.1		15	AAN85281
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		

图2 预测的 *AvMaSp1* 编码大壶腹腺丝蛋白-1 氨基酸序列与同源蛛丝蛋白一致性分析

1~15 为蛛丝蛋白的 NCBI 数据库蛋白登录号, 分别为: AAC47008: *Araneus diadematus* 蛛丝纤维蛋白-1(ADF-1); AAK30595: *Argiope trifasciata* 大壶状腺丝蛋白-1; AAT08436: *Agelenopsis aperta* 大壶状腺丝蛋白; CAM32249: *Euprostenops australis* 大壶状腺丝蛋白-2; CAJ00428: *Euprostenops australis* 大壶状腺丝蛋白-1; ABD61593: *Deinopis spinosa* 大壶状腺丝蛋白-2a; CAM32250: *Euprostenops australis* 类大壶状腺丝蛋白; ABD61594: *Deinopis spinosa* 大壶状腺丝蛋白-2b; AAT75310: *Nephila clavipes* 大壶状腺丝蛋白-1; AAT75311: *Nephila clavipes* 大壶状腺丝蛋白-1; ABR37275: *Nephilengys cruentata* 类大壶状腺丝蛋白; AAT75312: *Nephila clavipes* 大壶状腺丝蛋白-1; AAC47009: *Araneus diadematus* 蛛丝纤维蛋白-2; AAR13811: *Argiope amoena* 大壶状腺丝蛋白-2; AAN85281: *Araneus ventricosus* 大壶状腺丝蛋白-1(MaSp1)。下图同。AvMaSp1 和 ADF-1 编码蛋白分别用双下划线和单下划线标出。

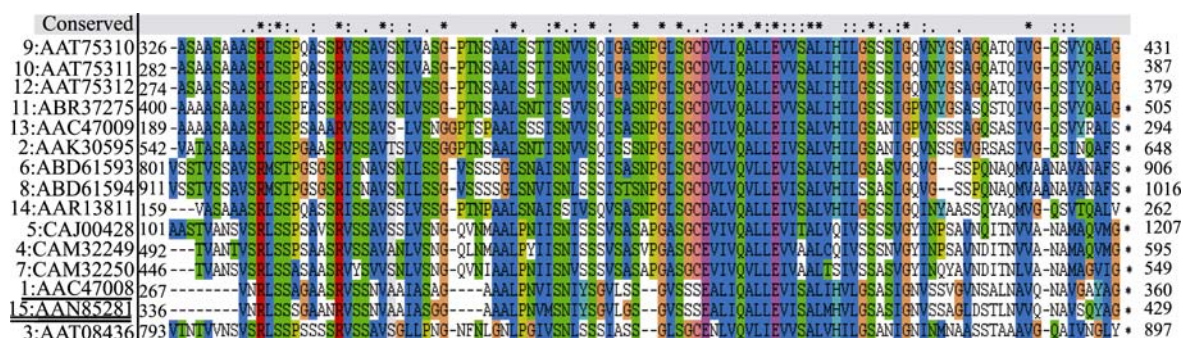


图3 预测的 *AvMaSp1* 编码大壶状腺丝蛋白-1 与同源蛛丝蛋白羧基端保守性分析

1~15 为蛛丝蛋白的 NCBI 数据库蛋白登录号, 同上图。*AvMaSp1* 和 *ADF-1* 编码蛋白分别用双下划线和单下划线标出。

列中重复出现 7 次, 富含丙氨酸区是证明该丝蛋白具有 β -折叠构型的关键证据^[6], 形成丝蛋白的结晶区。(GA)₂₋₆ 反复出现 17 次, 而且多聚丙氨酸前后总是与多聚 GA 相连。多聚 GA 和 A 可能为蛛丝蛋白的刚性结构区^[7]。GGYGQGY 反复出现 9 次, 可能构成非结晶区的 β -转角结构, 将重复区氨基酸序列分成 10 部分, 在其前端为 GGX (X: A、S), 在其后端为 GX(GA)₃₋₆(A)₄₋₆(GA)₂₋₃ (X: S、A、T、G)。序列中 GGXGGYGQGY (X: A、S)与(GA)₃₋₆(A)₄₋₆(GA)₂₋₃二者交替排列, 结晶区与非结晶区交替出现, 共同

构成了丝蛋白刚性与弹性的基础。在尾部序列中含有两个 SSS, 这些特点与 Lewis^[8]发现的小壶腹腺的 *Misp1* 基因相类似。由于多聚丙氨酸的多次出现, AvMaSp1 丝蛋白可能刚性更好些, 其典型重复结构为 $(GA)_nA_m(GA)_N$ (图 5)。C 端为不重复序列, 但相对保守(图 3)。Guerette 等^[9]也报道了十字园蛛 ADF1-4 与其他蜘蛛的某些蛛丝蛋白羧基端序列相似性较高。丝蛋白羧基端序列的相似, 可能对正确的将同用途蛛丝蛋白装配成一条蜘蛛丝而言是必要的^[10]。AvMaSp1 与 ADF-1 蛋白序列一致性最高, 为 73.7%

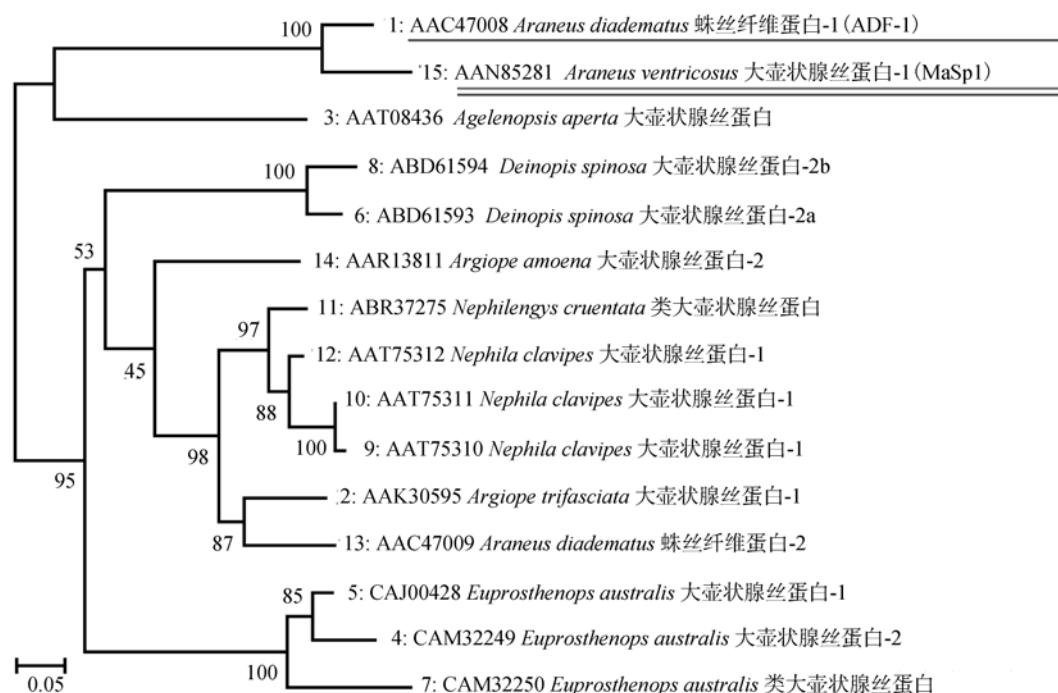


图 4 预测的 *AvMaSp1* 编码大壶状腺丝蛋白-1 与同源蛛丝蛋白系统进化树

分支上的数字表示重复 1 000 次的置信值。图中 1~15 为蛛丝蛋白的 NCBI 数据库蛋白登录号。*AvMaSp1* 和 *ADF-1* 编码蛋白分别用双下划线和单下划线标出。



图 5 预测的 *AvMaSp1* 编码大壶状腺丝蛋白-1 氨基酸序列重复结构特点

•: 非重复的氨基酸; -: 拉伸显示重复出现的氨基酸特征; 大腹园蛛大壶状腺丝蛋白-1 典型重复结构: $(GA)_nA_m(GA)_n$ 。

(图 2)。系统进化树分析也显示这两个园蛛属蜘蛛的蛛丝蛋白基因亲缘关系最近(图 4)。*AvMaSp1* 所编码的氨基酸多为甘氨酸和丙氨酸等分子量小的小侧链氨基酸, 甘氨酸含量为 38.69%, 丙氨酸含量为 31.7%, 小侧链氨基酸空间障碍小, 易于形成结构紧密有弹性和刚性的 β 转角和 β 片层结构^[11]。分子量小也正说明了蜘蛛丝是重量最轻的弹性纤维。

蜘蛛丝蛋白基因序列长, 克隆全长基因困难大, 目前仅有 8 个全长蛛丝蛋白基因的报道。Hayashi 等^[6, 12]报道了金纺蜘蛛和马达加斯加蜘蛛的鞭毛样

丝蛋白基因组 DNA, 其中金纺蜘蛛鞭毛样丝蛋白基因组 DNA 总长 36 kb, 中心区长约为 30 kb 包括 13 个外显子, 其蛋白重复区中包括 3 个重复单元, 从第 3 到第 12 内含子核酸数目都是 1 420 bp; 日本^[13]报道了横纹金蛛(*Argiope bruennichi*)卵袋丝蛋白 1 和 2 的全基因序列; 美国^[10, 14, 15]报道利用筛选基因组文库的方法, 获得黑寡妇蜘蛛(*Latrodectus hesperus*)的拖丝蛋白 1 和 2 及卵袋丝蛋白 1 和 2 的全基因序列。与鞭毛样丝蛋白基因不同, 黑寡妇蜘蛛拖丝蛋白 1 和 2 均是无内含子、单一超长外显子基因, ORF 框长

度分别为 9.4 kb 和 11.3 kb。

国内没有获得全长蛛丝蛋白基因的报道,多数有关蜘蛛丝的研究工作集中在蜘蛛丝力学特性以及人工表达蛛丝蛋白和成丝机理的研究。苏州大学潘志娟课题组研究表明拉断单位体积大腹园蛛蛛丝所需要的功远大于钢丝和 Kevlar^[16],并研究不同条件下再生蜘蛛丝蛋白的分子结构^[17]以及成丝条件对丝力学性能的影响^[18]。牛恒尧等^[19]将十字园蛛的 *ADF3* 丝蛋白基因在大肠杆菌中表达。郭廷晴等^[20]仿照金纺蜘蛛拖丝基因的一段序列人工合成了丝蛋白样基因单元,并利用 pGEX-KG 原核表达载体,在大肠杆菌中表达仿金纺蜘蛛拖丝蛋白并制备兔抗仿蜘蛛拖丝蛋白抗体。李敏等人尝试将美国 Tufts 大学提供的金纺蜘蛛拖丝蛋白合成基因单体进行多聚串联,并利用 pET30(a)表达载体在大肠杆菌中表达^[21],还高密度发酵大量表达 RGD-蛛丝融合蛋白^[22]。国内鲜有蜘蛛丝基因序列的原始创新性报道,大腹园蛛大壶状腺丝蛋白-1 部分 cDNA (*AvMaSp1*)的克隆不仅丰富了我国在蛛丝蛋白新基因领域的研究内容,而且也利用 *AvMaSp1* 设计探针获得全长丝蛋白基因序列,以及通过生物工程手段人工表达蛛丝蛋白提供科学依据。

参考文献(References):

- [1] Stauffer SL, Coguill SL, Lewis RV. Comparison of physical properties of three silks from *Nephila clavipes* and *Araneus gemmoides*. *J Arachnol*, 1994, 22(1): 5–11.
- [2] Scheibel T. Spider silks: recombinant synthesis, assembly, spinning, and engineering of synthetic proteins. *Microb Cell Fact*, 2004, 3(1): 14. [\[DOI\]](#)
- [3] J.萨姆布鲁克, E.F.费里奇, T.曼尼阿蒂斯著.分子克隆实验指南.金冬雁,黎孟枫译.第二版,北京:科学出版社,1999, 19–22.
- [4] Ren HL, Wang KJ, Zhou HL, Yang M. Cloning and organisation analysis of a hepcidin-like gene and cDNA from Japan sea bass, *Lateolabrax japonicus*. *Fish Shellfish Immunol*, 2006, 21(3): 221–227. [\[DOI\]](#)
- [5] Tassanakajon A, Klinbunga S, Paunglarp N, Rimphanitchayakit V, Udomkit A, Jitrapakdee S, Sritunyalucksana K, Phongdara A, Pongsomboon S, Supungul P, Tang S, Kuphanumart K, Pichyangkura R, Lursinsap C. *Penaeus monodon* gene discovery project: the generation of an EST collection and establishment of a database. *Gene*, 2006, 384: 104–112. [\[DOI\]](#)
- [6] Hayashi CY, Lewis RV. Evidence from Flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silk. *J Mol Biol*, 1998, 275(5): 773–784. [\[DOI\]](#)
- [7] Gatesy J, Hayashi C, Motriuk D, Woods J, Lewis R. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science*, 2001, 291(5513): 2603–2605. [\[DOI\]](#)
- [8] Colgin MA, Lewis RV. Spider Minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like "spacer regions". *Protein Sci*, 1998, 7(3): 667–672.
- [9] Guerette PA, Ginzinger DG, Weber BH, Gosline JM. Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science*, 1996, 272(5258): 112–115. [\[DOI\]](#)
- [10] Ayoub NA, Garb JE, Tinghitella RM, Collin MA, Hayashi CY. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS ONE*, 2007, 2(6): e514. [\[DOI\]](#)
- [11] Winkler S, Kaplan DL. Molecular biology of spider silk. *J Biotechnol*, 2000, 74(2): 85–93. [\[DOI\]](#)
- [12] Hayashi CY, Lewis RV. Molecular architecture and evolution of a modular spider silk protein gene. *Science*, 2000, 287(5457): 1477–1479.
- [13] Zhao AC, Zhao TF, Nakagaki K, Zhang YS, Sima YH, Miao YG, Shiomi K, Kajiura Z, Nagata Y, Takadera M, Nakagaki M. Novel molecular and mechanical properties of egg case silk from wasp spider, *Argiope bruennichi*. *Biochemistry*, 2006, 45(10): 3348–3356. [\[DOI\]](#)
- [14] Hu X, Kohler K, Falick AM, Moore AM, Jones PR, Sparkman OD, Vierra C. Egg case protein-1. A new class of silk proteins with fibroin-like properties from the spider *Latrodectus hesperus*. *J Biol Chem*, 2005, 280(22): 21220–21230. [\[DOI\]](#)
- [15] Hu X, Kohler K, Falick AM, Moore AM, Jones PR, Vierra C. Spider egg case core fibers: trimeric complexes assembled from TuSp1, ECP-1, and ECP-2. *Biochemistry*, 2006, 45(11): 3506–3516. [\[DOI\]](#)
- [16] 李春萍, 潘志娟, 刘敏. 大腹圆蛛丝的拉伸机械性能. 丝绸, 2002, (9): 46–48.
- [17] 许箫, 潘志娟, 钱丹娜, 夏艳杰. 不同工艺条件下再生蜘蛛丝蛋白的分子结构. 苏州大学学报(工科版), 2005, 25(2): 10–14.
- [18] 刘敏, 潘志娟, 李春萍. 自控能力及成丝条件与蜘蛛丝力学性能的关系. 纺织学报, 2005, 26(2): 8–11.
- [19] 牛恒尧, 王义琴, 萨其拉, 李文彬, 孙勇如. 在大肠杆菌中表达蜘蛛拖丝融合蛋白. 高技术通讯, 2001, 11(11): 5–7.
- [20] 郭廷晴, 赵昀, 汪生鹏, 董良, 陆长德. 兔抗仿蜘蛛牵丝蛋白抗体的制备及应用. 生物化学与生物物理学报, 2003, 35(8): 756–760.
- [21] 李敏, 章文贤, 黄智华, 黄建坤. 蜘蛛拖丝蛋白基因的构建及在大肠杆菌中的表达. 生物工程学报, 2002, 18(3): 331–334.
- [22] 阮超然, 黄晶星, 魏梅红, 李敏. 高分子量 RGD-蛛丝蛋白重组体的构建、高密度发酵及纯化. 生物工程学报, 2007, 23(5): 858–861.