



茶小绿叶蝉粘样蛋白 EfMLP 基因的克隆、 鉴定及功能分析

何雪怡, 雷雨欢, 宋仕月, 夏露霞, 王世宇, 马成文, 魏可欣,
王梦馨, 潘 钺*, 韩宝瑜*

(中国计量大学生命科学学院, 浙江省生物计量及检验检疫技术重点实验室, 杭州 310018)

摘要:【目的】探明茶小绿叶蝉 *Empoasca flavescens* 粘样蛋白 EfMLP 基因的分子特征、表达模式和生物学功能。【方法】基于茶小绿叶蝉转录组数据, PCR 克隆获得 4 条 EfMLP 基因的全长 cDNA 序列并进行生物信息学分析; qRT-PCR 检测 EfMLP 基因在茶小绿叶蝉不同发育阶段(卵、1–5 龄若虫和初羽化雌雄成虫)及初羽化成虫不同组织(体壁、脂肪体、唾液腺、肠道、卵巢和精巢)中的表达量; 通过饲喂法利用 RNAi 沉默茶小绿叶蝉 5 龄若虫 *EfMLP2* 和 *EfMLP4*, 生物测定分析 EfMLP 基因沉默后茶小绿叶蝉的存活率。【结果】获得 4 条茶小绿叶蝉 EfMLP 基因全长 cDNA 序列, 分别命名为 *EfMLP1* (GenBank 登录号: OR504428), *EfMLP2* (GenBank 登录号: OR504429), *EfMLP3* (GenBank 登录号: OR504430) 和 *EfMLP4* (GenBank 登录号: OR504431), 这 4 条 EfMLP 基因编码蛋白均具有 O-联糖基化位点形成粘蛋白结构域(mucin domain, MD)且该结构域包含高度重复的串联重复序列, 其中 EfMLP3 和 EfMLP4 的氨基酸序列含有保守的 2 型几丁质结合域(chitin binding domain, CBD)。系统发育分析结果表明, EfMLP 被分到两个不同的分支上, 属于两种不同的 MLP 类型, 该结果与昆虫的分类地位无相关性, 猜测可能与其功能分化相关。*EfMLP1* 和 *EfMLP2* 在茶小绿叶蝉初羽化雌雄成虫和初羽化成虫唾液腺中均特异性高表达, 而 *EfMLP3* 和 *EfMLP4* 的表达在卵、若虫和成虫等不同发育阶段和初羽化成虫脂肪体等多个组织中均可被检测到。与饲喂 dsGFP 对照组比较, 饲喂 ds*EfMLP2* 和 ds*EfMLP4* 可分别有效抑制茶小绿叶蝉体内 *EfMLP2* 和 *EfMLP4* 的表达量, 并可显著降低茶小绿叶蝉的存活率。【结论】EfMLP 在茶小绿叶蝉取食中扮演重要的角色, 可作为开发基于 RNAi 的叶蝉防治技术的潜在靶标。

关键词: 茶小绿叶蝉; 粘样蛋白; 基因克隆; 表达模式; RNA 干扰

中图分类号: Q966 文献标识码: A 文章编号: 0454-6296(2024)02-0151-12

Cloning, identification and functional analysis of the mucin-like protein EfMLP genes in *Empoasca flavescens* (Hemiptera: Cicadellidae)

HE Xue-Yi, LEI Yu-Huan, SONG Shi-Yue, XIA Lu-Xia, WANG Shi-Yu, MA Cheng-Wen, WEI Ke-Xin, WANG Meng-Xin, PAN Cheng*, HAN Bao-Yu* (Zhejiang Provincial Key Laboratory of Biometrology and Inspection and Quarantine, College of Life Sciences, China Jiliang University, Hangzhou 310018, China)

Abstract: 【Aim】This study aims to investigate the molecular characteristics, expression patterns, and biological functions of the mucin-like protein EfMLP genes of *Empoasca flavescens*. 【Methods】Based on

基金项目: 国家自然科学基金项目(32001910); 浙江省重点研发计划项目(2020C02026); 浙江省基础公益研究-计划项目(LGN20C140005); 浙江省基本科研业务费资助项目(2022YW10)

作者简介: 何雪怡, 女, 1998 年 10 月生, 浙江绍兴人, 硕士研究生, 研究方向为生物化学与分子生物学, E-mail: hexy1031@163.com

* 通讯作者 Corresponding authors, E-mail: pancheng@cjl. edu. cn; hanbaoyu@cjl. edu. cn

收稿日期 Received: 2023-09-24; 接受日期 Accepted: 2023-11-14

the transcriptome data of *E. flavescens*, the full-length cDNA sequences of four EfMLP genes were cloned by PCR and analyzed by bioinformatics. qRT-PCR was used to detect the expression levels of EfMLP genes across different developmental stages (egg, 1st – 5th instar nymphs, and newly emerged female and male adults), and in different tissues (integument, fat body, salivary gland, gut, ovary, and testis) of the newly emerged adults. *EfMLP2* and *EfMLP4* in the 5th instar nymph were silenced by RNAi through feeding method, and the survival rates of *E. flavescens* after silencing the EfMLP genes by RNAi were determined by bioassay. 【Results】 The full-length cDNA sequences of four EfMLP genes of *E. flavescens* were obtained, and named *EfMLP1*, *EfMLP2*, *EfMLP3* and *EfMLP4* with the GenBank accession numbers of OR504428, OR504429, OR504430 and OR504431, respectively. The obtained four EfMLPs all contain highly repetitive tandem repeat sequences, which are rich in O-linked glycosylation sites, forming the mucin domain (MD). Among them, both *EfMLP3* and *EfMLP4* contain a conserved type-2 chitin binding domain (CBD). Phylogenetic analysis result revealed that EfMLPs were divided into two different branches belonging to two different MLP types, which showed no correlation with insect taxonomy, but might be considered to be related to their functions. *EfMLP1* and *EfMLP2* exhibited specifically high expression in the newly emerged female and male adults and the salivary glands of the newly emerged adults. In contrast, the expression of *EfMLP3* and *EfMLP4* was identified in various developmental stages, including egg, nymphal and adult stages, as well as in diverse tissues such as the fat body of the newly emerged adult. Inhibition of the expression of *EfMLP2* and *EfMLP4* in *E. flavescens* by feeding ds*EfMLP2* and ds*EfMLP4* significantly reduced the survival rate of *E. flavescens* compared with the control group fed with ds*GFP*. 【Conclusion】 EfMLPs play an important role in the feeding of *E. flavescens* and can be used as a potential target in control of this pest insect based on RNAi strategies.

Key words: *Empoasca flavescens*; mucin-like protein; gene cloning; expression pattern; RNA interference

昆虫粘样蛋白 (mucin-like protein, MLP) 的成员最初从粉纹夜蛾 *Trichoplusia ni* 的围食膜中被鉴定出来, 由于氨基酸序列组成与人类粘蛋白 MUC2 相似而被命名 (Wang and Granados, 1997a, 1997b; 刘小民等, 2011)。该蛋白属于高分子量糖蛋白, 具有高度的 O-联糖基化位点形成粘蛋白结构域 (mucin domain, MD) (Fass and Thornton, 2023), 该结构域的序列中苏氨酸、丝氨酸和脯氨酸的残基比例比较高, 且通常含有相同或高度相似的串联重复序列 (Dhanisha *et al.*, 2018)。MLP 通常可分为分泌型和膜结合型两类 (Perez-Vilar and Hill, 1999; Hegedus *et al.*, 2009; Fass and Thornton, 2023), 其特殊序列和重复区的数量在非同源蛋白和同源粘蛋白的不同位置之间呈现出多样化。分泌型粘蛋白是粘液的重要组成部分, 在各种组织、粘膜的表面形成一层凝胶状物质, 可以有效防止有害物质的入侵, 其还参与胞内信号传导, 被包装成分泌物进入囊泡中。膜结合型粘蛋白的羧基末端有 1 个能连接到质膜的跨膜结构域, 促进两者的结合以及信号传导和蛋白质的稳定 (Fass and Thornton, 2023)。昆虫同样含

有分泌型粘蛋白 (围食膜相关) 和膜结合型粘蛋白 (微绒毛相关) (Perez-Vilar and Hill, 1999; Hegedus *et al.*, 2009)。粘蛋白的高度 O-联糖基化可以增强蛋白质应对水解酶降解的稳定性, 对细胞和组织的正常发育、生长和分化至关重要 (Loukas *et al.*, 2000; Cummings, 2009; Schjoldager and Clausen, 2012), 同时保护细胞免受寄生物、细菌、病毒的感染 (Tellam *et al.*, 1999; Rose and Voynow, 2006)。可见, MLP 在昆虫取食、互作和抵抗病菌入侵中也可能具有一定的功能。

目前对于昆虫 MLP 家族的鉴定大多为分泌型且集中在不含保护性角质层的器官, 如唾液腺、中肠和马氏管等。中肠粘蛋白的研究较多集中于鳞翅目和双翅目如小菜蛾 *Plutella xylostella* (Sarauer *et al.*, 2003)、甜菜夜蛾 *Spodoptera exigua* (张霞等, 2008; 李雯等, 2010)、蓀带夜蛾 *Mamestra configurata* (Shi *et al.*, 2004)、粉纹夜蛾 (Wang and Granados, 1997a, 1997b)、棉铃虫 *Helicoverpa armigera* (Fang *et al.*, 2009) 和埃及伊蚊 *Aedes aegypti* (Wu *et al.*, 2019) 等。上述研究表明, 肠道上皮中 MLP 会被粉纹夜蛾颗粒

体病毒(*Trichoplusia ni* granulosis virus, TnGV)或粘质沙雷氏菌 *Serratia marcescens* 等分泌的 Enhancin 蛋白消化,导致围食膜被破坏,促进虫媒病毒等的感染。唾液腺的相关研究显示,MLP 是唾液腺分泌物的重要组成部分。如黑腹果蝇 *Drosophila melanogaster* 在变态过程中从唾液腺中分泌出果蝇胶涂遍全身,用来保护果蝇蛹免受外界环境的影响从而提高存活率,该物质被鉴定属于分泌型 MLP (Monier and Courtier-Orgogozo, 2022)。MLP 被认为是褐飞虱 *Nilaparvata lugens* 和灰飞虱 *Laodelphax striatellus* 唾液鞘的重要组成部分,帮助其在水稻上进行取食,且与植物防御反应相关(Huang *et al.*, 2016, 2018, 2019; Liu *et al.*, 2016)。Miao 等(2018)在白背飞虱 *Sogatella furcifera* 水状唾液中发现 MLP 且在唾液腺中特异性表达,猜测其与植物诱导防御反应相关。综上所述,MLP 可能是粘液重要组成(胶状唾液),也参与昆虫中肠防御或植物诱导防御等过程,因此昆虫 MLP 的鉴定可为害虫生物防治潜在的新靶标的研究提供前提和基础。

茶小绿叶蝉 *Empoasca flavescens* 隶属于半翅目(Hemiptera)叶蝉科(Cicadellidae),是我国茶园发生最普遍、危害最严重的害虫之一(杨子银等, 2022)。该害虫具有个体较小、种群密度较大、繁殖能力强和极难防治等特点。探寻茶小绿叶蝉新型防治靶标和途径,从而有效延缓茶小绿叶蝉抗药性的发展是茶树防控面临的长期问题。目前,茶小绿叶蝉的 MLP 尚未鉴定,也未见文献报道,在 GenBank 数据库中也未找到相关基因片段,对 MLP 的研究尚属空白。为探寻茶小绿叶蝉防治新靶标,本研究对叶蝉 MLP 的编码基因 *EfMLP1* - 4 进行克隆、鉴定及序列分析,及其不同发育阶段、初羽化成虫不同组织的表达量差异性,并通过 RNAi 研究 *EfMLP2* 和 *EfMLP4* 在生长发育中的作用,以期利用 MLP 开发叶蝉绿色防控新技术提供科学依据。

1 材料与方法

1.1 供试昆虫

本研究所用茶小绿叶蝉成虫及若虫虫源均收集自浙江省杭州市西湖区上茅家埠茶园(30°14'40.812"N, 120°7'54.552"E),采集的龙井43 品种茶枝饲养。虫源采集后在人工气候箱中进行饲养,环境条件设置为温度(26 ± 1) °C,相对湿度 65% ± 5%,光周期 14L: 10D,用新鲜茶枝在实验室条件下

培养 48 h 以上。

解剖镜下用显微镊剥开被为害的茶梢,获得茶小绿叶蝉的卵。分别收集 1 - 5 龄若虫、初羽化雌雄成虫等各个发育时期的样本。在解剖镜下滴加冷藏的 PBS 缓冲液于冰面上,用显微镊解剖初羽化成虫(24 h 内)得到体壁、脂肪体、唾液腺、肠道、卵巢和精巢。以上样品分别置于 1.5 mL 无 RNA 的离心管中,经液氮冷冻后,保存于 -80 °C。

1.2 总 RNA 提取及 cDNA 合成

1.1 节样品经液氮研磨后,使用 Ultrapure RNA Kit(DNase I)(康为世纪,泰州)试剂盒,参照说明书提取叶蝉的总 RNA。RNA 的完整性、浓度和纯度分别通过 1% 琼脂糖凝胶电泳检测和微量核酸检测仪测定,RNA 溶液保存于 -80 °C 备用。根据反转录试剂盒 PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa, 大连)说明书合成 cDNA 第 1 链,储存于 -20 °C 备用。

1.3 EfMLP 基因克隆

根据本实验室测得的茶小绿叶蝉转录组数据库,根据注释信息筛选疑似粘蛋白 EfMLP 基因的序列片段,通过 RT-PCR 和 SMART-RACE-PCR 技术进行片段验证和全长扩增,最终获得 4 条 EfMLP 基因的全长 cDNA 序列。使用 Primer Premier 5.0 软件(PREMIER Biosoft, California)进行序列分析并设计引物(表 1),由杭州有康生物科技有限公司合成引物并完成后续的测序工作。

以 SMART™ RACE cDNA Amplification Kit (Clontech, California)合成 3'和 5'cDNA 第 1 链为模板,进行第 1 轮和第 2 轮扩增,具体反应体系和 PCR 程序参数参照刘亚芹等(2017)方法修改。获得全长序列后,跨开放阅读框设计引物,验证序列完整性。PCR 反应体系(50 μL): cDNA 模板 1 μL, 上下游引物(10 mmol/L)各 1 μL, PCR 酶 PrimeSTAR Max DNA Polymerase(TaKaRa, 大连) 25 μL, 加水至 50 μL。反应程序: 98 °C 10 s; 55 °C 5 s, 72 °C 25 s, 35 个循环;72 °C 10 min。所有的 PCR 产物经 1% 琼脂糖凝胶电泳检测鉴定,交由杭州有康生物科技有限公司进行回收纯化及测序,根据测序结果拼接获得全长序列或序列验证。

1.4 序列生物信息学分析

首先使用 DNAMAN 软件对氨基酸序列进行分析,将测序所得的基因序列 5'端序列和 3'端序列片段进行拼接。使用 ORFfinder 工具网站(<https://www.ncbi.nlm.nih.gov/orffinder/>)对氨基酸的开放

阅读框进行预测,并且根据 DNA 序列得到氨基酸序列。使用 Compute pI/Mw tool 网站 (https://web.expasy.org/compute_pi/) 预测氨基酸的分子量和等电点(Bjellqvist *et al.*, 1994; Gasteiger *et al.*, 2005)。使用 SignalP 6.0 网站 (<https://services.healthtech.dtu.dk/services/SignalP-6.0/>) 对氨基酸序列中的信号肽序列进行预测(Nielsen *et al.*, 2019)。使用 DeepTMHMM 网站 (<https://dtu.biolib.com/DeepTMHMM>) 对氨基酸序列的跨膜区进行预测(Hallgren *et al.*, 2022)。使用 SecretomeP 2.0 网站 (<https://services.healthtech.dtu.dk/services/SecretomeP-2.0/>) 对非经典分泌蛋白进行预测(Bendtsen *et al.*, 2004)。使用 PROSITE 网站 (<https://prosite.expasy.org/>) 对氨基酸序列的结构域和氨基酸富集区进行预测(Sigrist *et al.*, 2002, 2005)。使用 ProtParam tool 分析网站(<https://web.expasy.org/protparam/>) 对各类氨基酸含量进行分析(Gasteiger *et al.*, 2005)。使用 Cell-PLoc 2.0(<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/>) 预测亚细胞定位(Chou and Shen, 2008)。使用 NetNGlyc 1.0 (<https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/>) 网站对氨基酸序列 N 糖基化位点进行预测(Gupta and Brunak *et al.*, 2002)。使用 NetOGlyc 4.0 网站(<https://services.healthtech.dtu.dk/services/NetOGlyc-4.0/>) (Steenfot *et al.*, 2013) 以及 YinOYang 1.2 网站 (<https://services.healthtech.dtu.dk/services/YinOYang-1.2/>) (Gupta and Brunak *et al.*, 2002) 对氨基酸序列 O-联糖基化位点和 O-联糖基化富集区进行预测。使用 BLAST 查找工具 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) 进行序列分析,查找相似序列,比较序列之间的同源性(Johnson *et al.*, 2008)。基于氨基酸序列,使用 MEGA 11 软件的 ClustalW 对比序列后,删除两端冗余序列,调整保留保守区,使用最大似然法(maximum likelihood)进行系统进化树的构建,并且设置 1 000 次的 bootstraps 进行检验(Tamura *et al.*, 2013)。

1.5 EfMLP 基因时空表达分析

以 1.2 节合成的 cDNA 为模板,使用 Primer Premier 5.0 软件设计 RT-qPCR 特异性引物(表 1),分析 EfMLP 基因在茶小绿叶蝉不同发育阶段和初羽化成虫组织表达的特异性。采用 TB Green® PremixEx Taq™ II (Tli RNaseH Plus) (TaKaRa, 大连)试剂盒在实时荧光定量 PCR 仪 Step One Plus

(ABI, California) 上进行检测。反应体系(20 μ L): TB Green 10 μ L, 正反向引物(10 mmol/L)各 0.8 μ L, ROX 0.4 μ L, cDNA 2 μ L, ddH₂O 6 μ L。反应程序: 95 $^{\circ}$ C 30 s; 95 $^{\circ}$ C 5 s, 60 $^{\circ}$ C 60 s,共 40 个循环。以 *EfTubulin* 作为内参基因(Paim *et al.*, 2012),每组样品设置 3 次生物学和技术重复。采用 $2^{-\Delta\Delta C_t}$ 相对定量法计算基因的表达量(Livak and Schmittgen, 2001),不同发育阶段和组织特异性的表达分别以卵和唾液腺的基因表达量作为 1 计算。

1.6 EfMLP 基因 dsRNA 合成和 RNAi 分析

基于目标基因和 *GFP* (对照)的 cDNA 全序列,设计基因的 dsRNA 合成引物(表 1),参照 T7 RiboMAX™ Express RNAi System (Promega, Wisconsin) 试剂盒说明书合成 ds*EfMLP2*, ds*EfMLP4* 和 ds*GFP*,并检测质量和浓度后备用。

本研究使用饲喂方式进行目的基因的 RNAi 处理,改良 Parafilm 膜夹营养液法(潘铖等, 2021)的取食装置为饲喂装置,使用 5% 无菌蔗糖为载体溶液,dsRNA 的终浓度为 300 ng/ μ L,以等量 ds*GFP* 作为阴性对照组,并等量的 Buffer 溶液作为空白对照组(CK)。选择发育进度一致的茶小绿叶蝉 5 龄若虫,30 头为 1 个处理组,经过 1 h 饥饿处理后,装入饲喂装置,以湿润的纱布封口,排列于透气网盒中,确保饲喂装置位于上端,略微倾斜,转移至人工气候箱(温度等参数同 1.1 节)。每个处理组均设置 6 个独立的生物学重复,用于检测 RNAi 的基因沉默效率和存活率统计。以饲喂 ds*GFP* 的茶小绿叶蝉作为对照,分别于饲喂后 12, 24, 48 和 72 h 时收集茶小绿叶蝉,按照 1.5 节方法进行 qRT-PCR 检测目标基因的表达量。校正存活率(%) = (处理组试虫存活率/空白对照组试虫存活率) $\times$ 100。

1.7 数据分析

使用 Graphpad Prism 7.0 软件制图。数据处理与分析采用 DPS 数据处理系统(唐启义, 2020)。使用 *t* 检验法分析两组样本间的差异显著性,使用单因素方差分析、Tukey 氏 HSD 法分析 3 组分及以上样本间的差异显著性,差异显著性水平为 $P < 0.05$ 。

2 结果

2.1 EfMLP 基因的克隆及理化特征

共获得 4 条符合预期的茶小绿叶蝉 EfMLP 基因 cDNA 全长序列,分别将其命名为 *EfMLP1* (GenBank

表 1 本研究所用引物信息
Table 1 Primer information used in this study

引物 Primers		引物序列(5′ – 3′) Primer sequences	用途 Use
EfMLP1-3′	1; TCAAGTGAGGAAACAACAGTTGT	2; GATTCAAGTGAATCTTCAACAACAAC	RACE
EfMLP1-5′	1; TCAGATTTCATCTCGTGATACAACGT	2; AGTTGACTCACTAGTGCTGCTAGAC	
EfMLP2-3′	1; TTGTATCAGCAAAAGGATGGTGAAG	2; CAAATATGGGATAATCCAGATTACCTG	
EfMLP2-5′	1; TGATTCTGAAACGACAGTCGTGGT	2; CACTTGAGTCTTCAGATGATGATTCT	
EfMLP3-3′	1; TGGTGGCCAGCAACCTGGTGGATTAC	2; CCAGGGCCCATCGGCAGGATACCCAG	
EfMLP3-5′	1; ATTGTGCCATTAGGACAAAGGAAATC	2; CCTTGGCTTCCAGGAATCCTCCGG	全长 cDNA 序列扩增 Amplification of full-length cDNA sequence
EfMLP4-3′	1; GGTGCTGGATACACAAAATACGACT	2; TCAAGTCAGCCAGAACCCAACCGCC	
EfMLP4-5′	1; TAGTGAAGAAGTAGTTGAGTGTTAC	2; GAATTTTTTACAGTCACCTTTAACCT	
UPM	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT		
EfMLP1	F: ATGAATCTGATTCAAGTGAATC	R: GTACTTGATTCACTGAATCAG	
EfMLP2	F: GAGATGAATCTGATTCAAGTGA	R: GCCTTAGTCTGAAATGCTGG	qRT-PCR
EfMLP3	F: CAAGATGACGCAGCTCCACAA	R: CGATTATCTTCTGTAAGGAG	
EfMLP4	F: GCTGATGACGTCGCGATGGCTCG	R: TCATCGATGTCCCAGCTCGTT	
qEfMLP1	F: CCACAGAGTCTGAAGATTCACCAC	R: TCTGAAACGACAGTTGTGTTGAGG	
qEfMLP2	F: CGCAAAGTGTTCCCTACCC	R: ATGCTGGAGTCCGACGATAAG	
qEfMLP3	F: AACCTCTTTTCAGATGCGACC	R: TGGGGCAGAGTTACAGTCG	RNAi
qEfMLP4	F: ACAAAGTTCCAGCACAAACGG	R: TAAAGGCTTTGCGGTTGTGG	
EfTubulin	F: GTGGTGCCAGGAGGTGACTT	R: ACCCTCTCCGACGTACCACT	
dsEfMLP2	F: <u>TAATACGACTCACTATAGGG</u> AGATAGTTATGGCTCTTTTCCCTG	R: <u>TAATACGACTCACTATAGGG</u> GGCTTGATGATTTTGAGTCGG	
dsEfMLP4	F: <u>TAATACGACTCACTATAGGG</u> GCCTGTAAACCCTGCCTCTG	R: <u>TAATACGACTCACTATAGGG</u> TACTACTGGATGTGCTATTGGACC	
dsGFP	F: <u>TAATACGACTCACTATAGGG</u> AGAATGGTGAGCAAGGGCGAGGAGC	R: <u>TAATACGACTCACTATAGGG</u> AGATTACTTGACAGCTCGTCCATGC	

下划线表示 T7 启动子区域；引物序列前的数字 1 和 2 分别表示第 1 和第 2 轮扩增引物。The underline position represents the T7 promoter region; the numerals 1 and 2 before primer sequences indicate the primers in the 1st and 2nd rounds of amplification, respectively.

登录号：OR504428），*EfMLP2*（GenBank 登录号：OR504429），*EfMLP3*（GenBank 登录号：OR504430）和 *EfMLP4*（GenBank 登录号：OR504431），cDNA 全长 3.3 ~ 6.1 kb，其中 *EfMLP4* 的 cDNA 序列最长，预测的开放阅读框长 6 033 bp，共编码 2 010 个氨基酸；最短的是 *EfMLP3*，预测的开放阅读框长 3 024 bp，编码 1 007 个氨基酸。EfMLP4 相对分子量最大，为 208.6 kD；EfMLP3 的最小，为 97.5 kD。预测的理论等电点分别为 4.64、5.12、5.03 和 4.7，表明 4 个 EfMLP 均属于酸性蛋白质。极性氨基酸在序列中占比 50% 以上，极性氨基酸以丝氨酸、苏氨酸为主，符合 MLP 的序列特征。所有 EfMLP 均有定位在细胞外，均可能为分泌蛋白。EfMLP3 蛋白序列的 5′端预测到含有 1 个 21 aa 的信号肽编码序列，编码的序列为 MTQLHKILYITAAFSIYIAKA；EfMLP4 蛋白序列含有 17 aa 的信号肽编码序列，编码的序列为 MTSRWLVLCACAVLVQT。4 个粘样蛋白均不含

有跨膜结构，同时 EfMLP1 和 EfMLP2 蛋白序列的 NN-score 均为 0.6 以下。因此，我们猜测 EfMLP3 和 EfMLP4 为经典分泌蛋白，属于分泌型粘样蛋白；而 EfMLP1 和 EfMLP2 不含信号肽序列，猜测可能为膜结合型粘样蛋白，但也不排除通过非经典蛋白分泌途径运输到细胞外（Zheng and Ge, 2022）。

2.2 MD 结构域和特征序列

MLP 属于高分子量糖蛋白，具有高度的 O-联糖基化位点形成 MD 结构域（Fass and Thornton, 2023）。结构域分析结果表明，EfMLP1 的氨基酸序列中含有 2 个高度 O-联糖基化 MD 结构域，为 MD1（第 440 – 673 位氨基酸）和 MD2（第 813 – 1 256 位氨基酸），富含丝氨酸（30.7%）和苏氨酸（25.5%）。EfMLP2 同样预测到 1 个高度 O-联糖基化的 MD（第 220 – 466 位氨基酸），丝氨酸和苏氨酸的含量分别为 34.0% 和 26.7%。EfMLP3 含有 1 个富含 23.5% 脯氨酸的 MD（第 647 – 1 004 位氨基酸）。EfMLP4

则预测到了3个MD,分别为MD1(第255-341位氨基酸)、MD2(第441-833位氨基酸)和MD3(第1702-1837位氨基酸),其中丝氨酸含量最多,分别占到31.0%,39.4%和36.8%。除EfMLP2未预测到N联糖基化位点外,EfMLP1预测到1个在第814位氨基酸的N联糖基化位点,fMLP3预测到2个在第577和586位氨基酸的N联糖基化位点,EfMLP4共预测到5个N联糖基化位点,分别在第118,266,442,1701和1882位氨基酸。

MLP通常含有相同或高度相似的串联重复序列,本研究的4个EfMLP均发现了重复序列。EfMLP1有3段串联重复序列为分别为9段TTTSETTESSTTT、10段SSSTSES以及6段SSESSTTTT的重复序列。EfMLP2的MD中发现1段重复3次的序列为TEQSSSTSESTSETSTTTTVVSESSSEDSS。EfMLP3含有11段GPQGPSAQFPGQ的重复序列。EfMLP4的MD2内有8段QQSSSSG的重复序列,MD3中含有14段QPGQPGQPG的重复序列。除具

有已知昆虫MLP的氨基酸序列特征外,还发现EfMLP1和EfMLP2中均存在谷氨酰胺GLN富集区,重复序列QQQQPQSLKI在EfMLP1的2段GLN富集区分别重复6次和3次,该序列也在EfMLP2的GLN富集区中重复3次。EfMLP中发现了甘氨酸GLY富集区,该区域内序列GSQGGAPGFG重复5次。可见4个EfMLP的均含有MD结构域,也存在MLP特有的重复序列,但重复序列的长短、复杂程度和重复次数均存在一定的差异性。

2.3 MLP的CBD结构域

对克隆的4个EfMLP进行几丁质结合功能域(chitin-binding domain, CBD)预测(图1),其中EfMLP1和EfMLP2未鉴定到CBD结构域。EfMLP3的氨基酸序列含有1个保守的2型CBD结构域,几丁质结合功能域保守区的特征结构为CX₁₅CX₅CX₉CX₁₂CX₇C。EfMLP4含有10个保守的2型CBD结构域,特征结构为CX₁₃CX₅CX₁₃₋₁₇CX₁₂CX₈₋₁₁C。

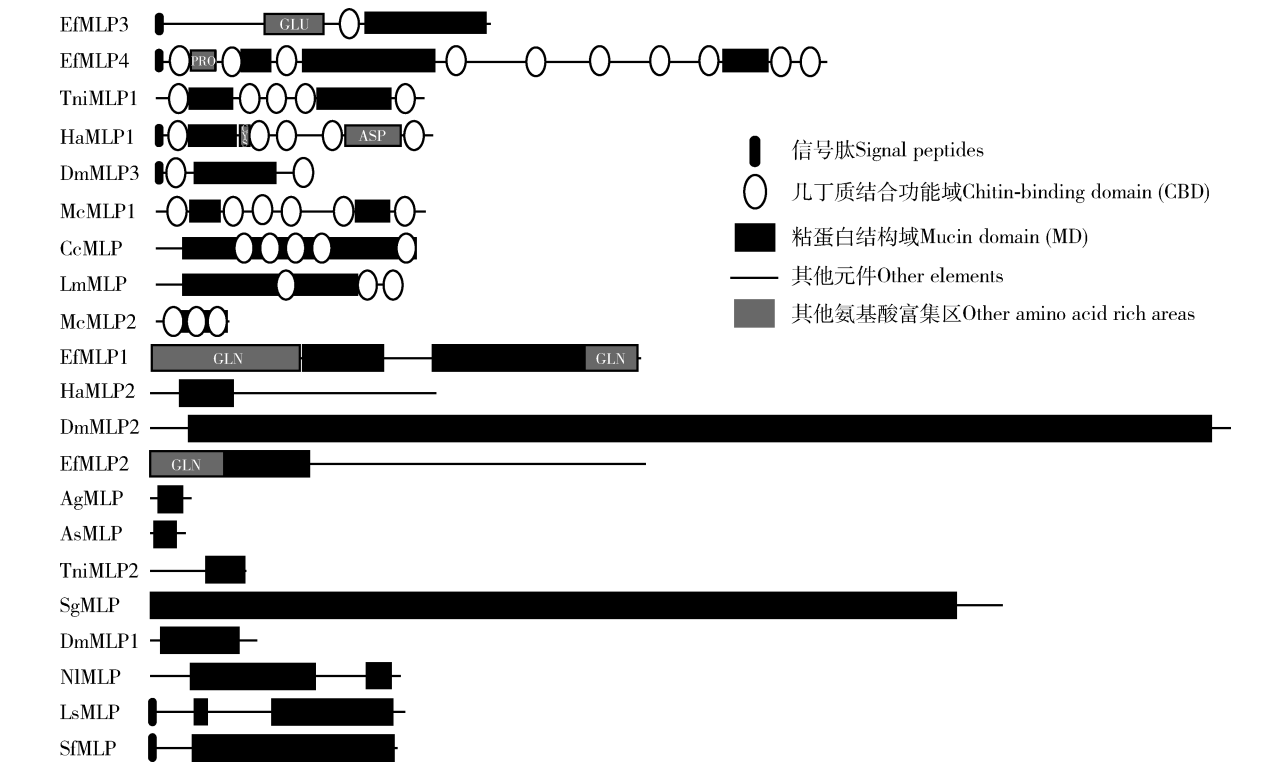


图1 茶小绿叶蝉EfMLP及其他昆虫MLP的氨基酸序列结构

Fig. 1 Structure of EfMLPs from *Empoasca flavescens* and MLPs from other insects based on amino acid sequence
蛋白来源物种和GenBank登录号 Protein origin species and GenBank accession numbers; EfMLP1-4: 茶小绿叶蝉 *Empoasca flavescens*, OR504428, OR504429, OR504430, OR504431; TniMLP1, TniMLP2: 粉纹夜蛾 *Trichoplusia ni*, AAC47557, XP_026747627; HaMLP1, HaMLP2: 棉铃虫 *Helicoverpa armigera*, ABW98670, XP_049693344; DmMLP1-3: 黑腹果蝇 *Drosophila melanogaster*, NP_524024, NP_996054, NP_652552; McMLP1, McMLP2: 蓓带夜蛾 *Mamestra configurata*, AAL17912, ACN52067; CeMLP: 麦秆蝇 *Cephus cinctus*, XP_015595092; LmMLP: 亚洲飞蝗 *Locusta migratoria*, QKV49919; AgMLP: 冈比亚按蚊 *Anopheles gambiae*, AAC08530; AsMLP: 斯氏按蚊 *Anopheles stephensi*, AEY77529; SgMLP: 沙漠蝗 *Schistocerca gregaria*, XP_049861009; NIMLP: 褐飞虱 *Nilaparvata lugens*, AQL58581; LsMLP: 灰飞虱 *Laodelphax striatellus*, UUG47375; SfMLP: 白背飞虱 *Sogatella furcifera*, AQP26312. 图2同。The same for Fig. 2.

2.4 系统发育

茶小绿叶蝉、粉纹夜蛾和黑腹果蝇等类似,种内和种间的 MLP 氨基酸序列一致性均较低(低于 40%)。系统进化树(图 2)表明,MLP 氨基酸序列不论是种内还是种间的差异性均较大,茶小绿叶蝉、粉纹夜蛾、棉铃虫等物种的 MLP 被分到不同的分支上,属于两种不同的 MLP 类型,该结果与昆虫的分类地位无相关性,猜测可能与其功能分化相关。其中 EfMLP1 和 EfMLP2 与褐飞虱等的同源性略高,该类蛋白均不含 CBD 结构域,多数被报道为唾液鞘蛋白或唾液腺蛋白。而 EfMLP3 和 EfMLP4 与棉铃虫 HaMLP1 (ABW98670) (Fang *et al.*, 2009; 刘小民等, 2011)、亚洲飞蝗 *Locusta migratoria* (Zhao *et al.*, 2020)、黑腹果蝇 DmMLP3 等的同源性较高,该类蛋白多数含有 CBD 结构域,被认定为昆虫肠粘蛋白(insect intestinal mucin, IMM)是围食膜的重要组成部分。

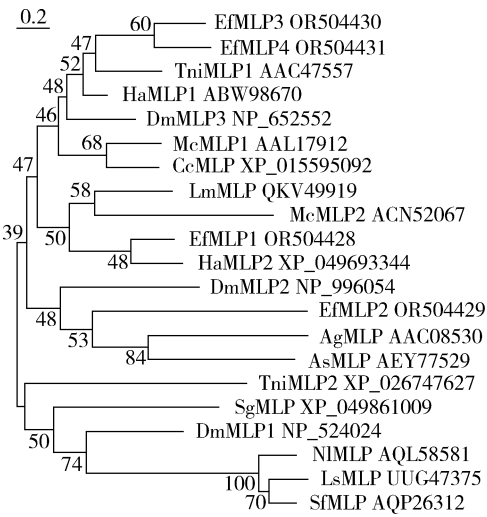


图 2 最大似然法构建的基于氨基酸序列的茶小绿叶蝉 EfMLP 及其他昆虫 MLP 的系统发育树
Fig. 2 Phylogenetic tree of EfMLPs from *Empoasca flavescens* and MLPs from other insects constructed by maximum likelihood method based on amino acid sequence
标尺示遗传距离。The scale bar indicates the genetic distance.

2.5 茶小绿叶蝉 EfMLP 基因的时空表达特征

EfMLP1 和 *EfMLP2* 在孵化后表达量上调,若虫期在 3 龄若虫期表达量达到最高,与卵和若虫期相比,初羽化雌雄成虫中表现为特异性高表达,且初羽化雄成虫中的表达量显著高于初羽化雌成虫中的 ($P < 0.05$) (图 3: A, B); *EfMLP3* 在不同发育阶段的表达量变化不大,在 5 龄若虫和初羽化雄成虫中的表达量显著高于卵中的表达量 ($P < 0.05$),在 1 -

4 龄若虫和初羽化雌成虫中的表达量显著低于卵中的 ($P < 0.05$) (图 3: C); 而 *EfMLP4* 在 1 - 5 龄若虫中的表达量高于卵和初羽化雌雄成虫中的 (图 3: D)。

EfMLP1 和 *EfMLP2* 在茶小绿叶蝉的初羽化成虫各组织中具有相似的表达模式,都是仅在唾液腺中高表达,在其余组织体壁、脂肪体、精巢和卵巢中表达量极低(图 4: A, B)。 *EfMLP3* 和 *EfMLP4* 的表达模式相似,在唾液腺、体壁、脂肪体和肠道中广泛表达,且表达量较卵巢和精巢等组织中的更高;其中 *EfMLP3* 在脂肪体中表达量最高,为在唾液腺中表达量的 4 倍,而 *EfMLP4* 则在体壁中表达量最高(图 4: C, D)。

2.6 RNAi 效果

本研究获得的 4 条粘样蛋白基因分属于两种不同的 MLP 类型,根据 2.5 节的结果分别选取每种类型中表达差异更为显著的 *EfMLP2* 和 *EfMLP4* 进行 RNAi 后续实验。qRT-PCR 检测结果表明,饲喂 ds*EfMLP2* 和 ds*EfMLP4* 可分别有效抑制茶小绿叶蝉 5 龄若虫体内 *EfMLP2* 和 *EfMLP4* 的表达(图 5: A, B)。与阴性对照组(饲喂 ds*GFP*)相比,饲喂 ds*EfMLP2* 24, 48 和 72 h 时 *EfMLP2* 的表达量分别显著下降 68.1%, 60.9% 和 98.2% ($P < 0.05$) (图 5: A),饲喂 ds*EfMLP4* 24, 48 和 72 h 时 *EfMLP4* 的表达量分别显著下降 41.7%, 98.3% 和 98.0% ($P < 0.05$) (图 5: B)。

生物测定实验表明(图 5: C),在饲喂 24, 48 和 96 h 时 ds*EfMLP2* 处理组的校正存活率分别为 61.5%, 39.1% 和 19.6%, 而 ds*EfMLP4* 处理组的校正存活率分别为 50.0%, 42.0% 和 33.3%, 均低于阴性对照组(饲喂 ds*GFP*)的校正存活率。结果表明,通过 RNAi 抑制 *EfMLP2* 和 *EfMLP4* 表达均可显著降低茶小绿叶蝉的存活率。

3 讨论

为了进一步研究茶小绿叶蝉 MLP 的功能及其在开发叶蝉防控技术的应用潜力,本研究克隆鉴定了 4 条新的茶小绿叶蝉 MLP 编码基因,并分别命名为 *EfMLP1* - 4, 分析了其在时空特异性的表达规律,并初步分析了其在茶小绿叶蝉生长发育过程中的作用。亚细胞定位预测结果显示,所有 EfMLP 均定位在细胞外,均可能被分泌到体外。EfMLP3 和 EfMLP4 分别拥有 1 个 21 和 17 aa 的信号肽,无跨膜

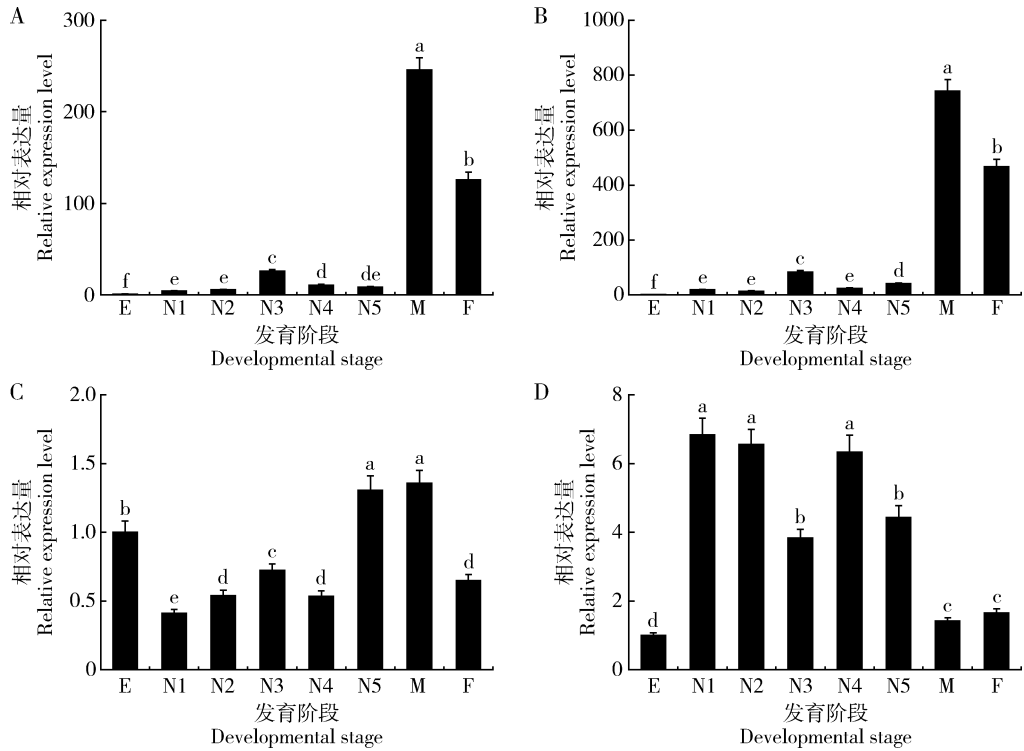


图3 茶小绿叶蝉 EfMLP 基因在不同发育阶段的表达量

Fig. 3 Expression levels of EfMLP genes in *Empoasca flavescens* at different developmental stages

A: *EfMLP1*; B: *EfMLP2*; C: *EfMLP3*; D: *EfMLP4*. E: 卵 Egg; N1 – N5: 分别为 1 – 5 龄若虫 1st – 5th instar nymphs, respectively; F: 初羽化雌成虫 Newly emerged female adult; M: 初羽化雄成虫 Newly emerged male adult. 图中数据为平均值 $\pm$ 标准误; 柱上不同字母表示基因表达量在不同发育阶段间差异显著 ($P < 0.05$, Tukey 氏 HSD 法)。Data in the figure are mean $\pm$ SE. Different letters above bars indicate significant difference in the gene expression level among different developmental stages ($P < 0.05$, Tukey's HSD method).

结构,是经典分泌蛋白。而 EfMLP1 和 EfMLP2 不含信号肽序列,但组织特异性表达显示均在唾液腺中特异性高表达。课题组已利用 Parafilm 膜夹营养液法和 LC-MS/MS 完成茶小绿叶蝉水状唾液蛋白的系统鉴定,鉴定发现水状唾液中含有 EfMLP2 和 EfMLP4 的多肽片段(未发表),我们与 Zheng 和 Ge (2022) 等的综述观点一致,认为目前已有越来越多的证据表明,非经典分泌蛋白参与发育、免疫和脂质代谢,因此 EfMLP1 和 EfMLP2 是唾液中破碎唾液腺细胞的膜蛋白还是被分泌到体外,从而参与叶蝉与茶树之间的互作,仍需进一步研究。

生物信息学分析显示,本研究中的 EfMLP 虽然同源性都较低,但均含有至少 1 个高度 O-糖基化的且富含丝氨酸、苏氨酸或脯氨酸残基的 MD 结构域,且序列中含有高度重复的串联重复序列。有学者认为粉纹夜蛾 MLP 中 MD 的重复序列为 TTQAPT(重复 15 次),甜菜夜蛾 MLP 中的重复序列为 TTTQAPITTT(重复 7 次),黑腹果蝇的重复序列为 PTTTK(重复 26 次),因此猜测部分 MLP 的重复序列多为较短氨基酸序列(可能为 4 ~ 8 个),而且在

结构上简单重复多次(Wang and Granados, 1997b; 张霞等, 2008; Monier and Courtier-Orgogozo, 2022)。在本研究中,符合上述规律的有 EfMLP1 中 SSSTSES(重复 10 次)、EfMLP4 的 MD2 中 QQSSSSG(重复 8 次)和 MD3 中 QPGQPGQPG(重复 14 次)的 3 段重复序列。也有 11 ~ 13 个氨基酸的重复序列,如 EfMLP1 中 2 段重复序列 TTTSETTESSTTT(重复 9 次)、SSESSTSTTTTT(重复 6 次)和 EfMLP3 中的重复序列 GPQGPSAQFPGQ(重复 11 次),结果与前人报道一致,序列较短且重复频率较高。与之相反的是,本研究中 EfMLP2 中含有 29 个氨基酸的重复序列(重复 3 次)。当然也有学者发现小菜蛾 MLP 的重复序列较长(27 个氨基酸),序列相对复杂且重复频率低,仅发生 2 ~ 11 次(Saraauer *et al.*, 2003)。除具有已知昆虫 MLP 的氨基酸序列特征(MD 结构域)外,张霞等(2008)在甜菜夜蛾的 MLP 还含有 1 个独特的天冬氨酸 ASN 富集区,该区域内序列 DDDKPGCNGNCPE 重复达 12 次,刘小民等(2011)在棉铃虫 MLP 中也发现了甘氨酸-天冬氨酸富集区,在本研究中 EfMLP1 和 EfMLP2 发现了 GLN 富

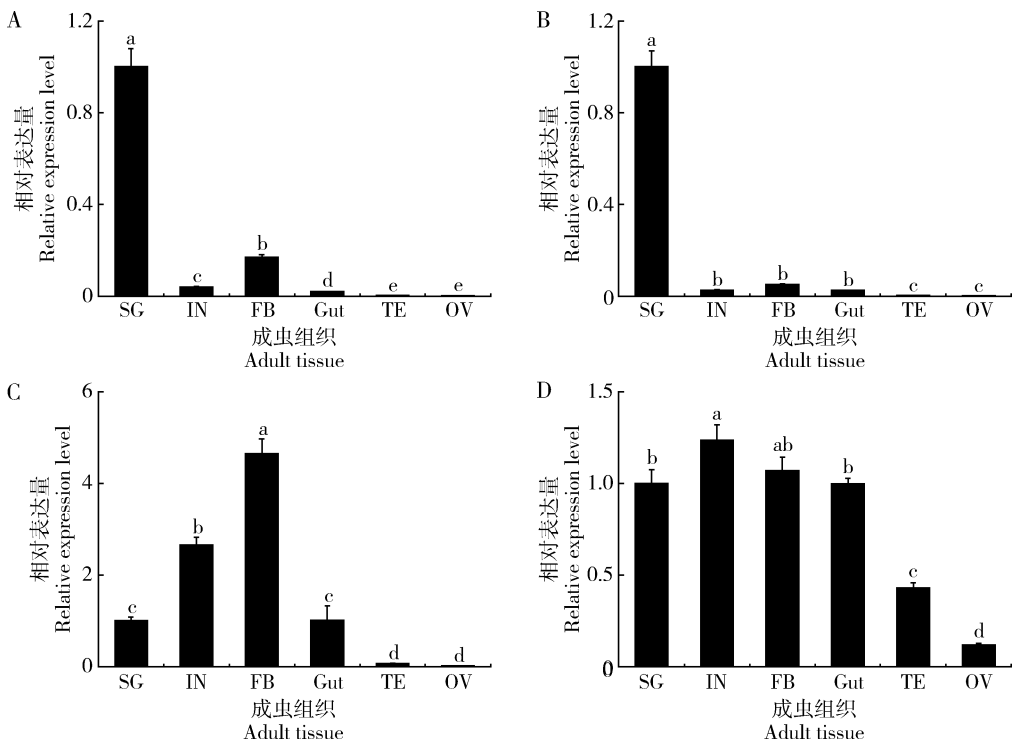


图4 茶小绿叶蝉 EfMLP 基因在初羽化成虫不同组织中的表达量

Fig. 4 Expression levels of EfMLP genes in different tissues of the newly emerged adults of *Empoasca flavescens* A: *EfMLP1*; B: *EfMLP2*; C: *EfMLP3*; D: *EfMLP4*. SG: 唾液腺 Salivary gland; IN: 体壁 Integument; FB: 脂肪体 Fatty body; Gut: 肠道; TE: 精巢 Testis; OV: 卵巢 Ovary. 图中数据为平均值 $\pm$ 标准误; 柱上不同字母表示基因表达量在不同组织间差异显著 ($P < 0.05$, Tukey 氏 HSD 法)。Data in the figure are mean $\pm$ SE. Different letters above bars indicate significant difference in the gene expression level among different tissues ($P < 0.05$, Tukey's HSD method).

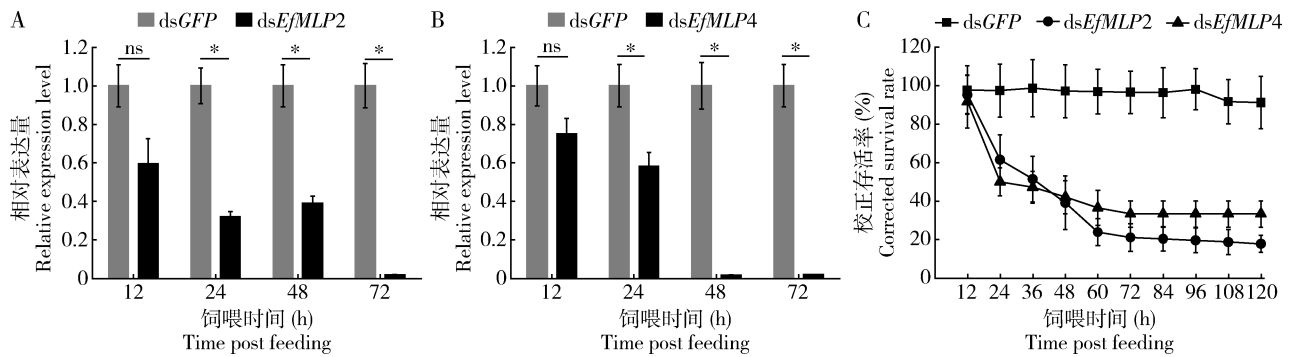


图5 饲喂茶小绿叶蝉 5 龄若虫 dsRNA 后 EfMLP2 (A) 和 EfMLP4 (B) 的相对表达量和校正存活率 (C)

Fig. 5 Relative expression levels of *EfMLP2* (A) and *EfMLP4* (B) and corrected survival rates (C) after feeding the 5th instar nymphs of *Empoasca flavescens* with dsRNA

dsGFP: 阴性对照 Negative control. 图中数据为平均值 $\pm$ 标准误; 柱上星号和 ns 分别表示两组间差异显著 ($P < 0.05$) 和不显著 ($P > 0.05$) (t 检验)。Data in the figure are mean $\pm$ SE. Asterisk and ns above bars indicate significant difference ($P < 0.05$) and no significant difference ($P > 0.05$), respectively, between two groups (t -test).

集区, EfMLP3 发现了 GLY 富集区, 且均存在重复序列, 以上富集区为首次报道, 其作用未知。综上所述, 我们猜测 MLP 序列中存在重复序列是寻常现象, 但是对于重复序列的长短和重复频率, 存在的意义和功能需还有待进一步的研究。

MLP 的成员中昆虫肠粘蛋白 IMM 是研究最多

的类型, 其被认为是围食膜的重要组成部分, 多数研究报道其含有 CBD 结构域。无脊椎动物中 MLP 的功能也被认为同几丁质的结合有关, 这些 CBD 结构域包括 6 ~ 10 个保守的半胱氨酸残基和少数芳香族氨基酸残基, 形成分子内和分子间二硫键, 以维持围食膜蛋白的稳定性, 有助于围食膜结构的正常形成

(Shen and Jacobs-Lorena, 1999; Wang and Granados, 2001)。根据进化树结果(图 2),进一步分析了进化树选取的 MLP 氨基酸序列结构中 MD 结构域、氨基酸富集区、CBD 结构域等组成。结合进化树的分析,可发现本研究中的 MLP 也可根据是否含有 CBD 结构域分为两种不同的类型。EfMLP1 和 EfMLP2 在唾液腺中特异性高表达,其不含有 CBD 结构域,与褐飞虱、白背飞虱、灰飞虱等被报道来源于唾液腺或唾液鞘蛋白的研究结果(Huang *et al.*, 2016, 2018; Liu *et al.*, 2016; Miao *et al.*, 2018)一致,因此我们猜测 EfMLP1 和 EfMLP2 具有参与唾液鞘形成或取食行为等。EfMLP3 和 EfMLP4 中分别含有 1 和 10 个 CBD 结构域,结构域分析表明 11 个 CBD 结构域均属于 II 型几丁质结合域(图 1),保守区序列分别为 CX₁₅CX₅CX₉CX₁₂CX₇C 和 CX₁₃CX₅CX₁₃₋₁₇CX₁₂CX₈₋₁₁C,这与张霞等(2008)报道粉纹夜蛾、小菜蛾、甜菜夜蛾和蓥带夜蛾的 CBD 结构域的保守序列符合 CX₁₃₋₁₅CX₅CX₉CX₁₂CX₇C 结构基本一致。从图 1 中可发现存在 CBD 结构域的 MLP 蛋白中,未发现 MD 结构域和 CBD 结构域的数量和分布具有一致性或保守性。如粉纹夜蛾 TniMLP1、棉铃虫 HaMLP1 含有 5 个 CBD 结构域,蓥带夜蛾含有 6 个 CBD 结构域,均分布在 MD 结构域两端,但是麦秆蝇 CcMLP 的 4 个 CBD 结构域和蓥带夜蛾 McMLP2 的 3 个 CBD 结构域分布在 MD 结构域内部,而亚洲飞蝗 LmMLP 的 3 个 CBD 结构域有 1 个在 MD 内部,2 个在外部(Wang and Granados, 1997b; Shi *et al.*, 2004; Fang *et al.*, 2009; 刘小民等, 2011; Zhao *et al.*, 2020),MLP 的功能与其各元件的结构组成关系需更多的数据支撑解析。

EfMLP 基因在叶蝉体内的表达呈现出明显的时空特异性。不含 CBD 结构域的蛋白及其对应基因 *EfMLP1* 和 *EfMLP2* 在雄成虫的表达量高于雌成虫的(图 3: A, B),可能是这种蛋白对雄虫更加重要。*EfMLP1* 和 *EfMLP2* 在茶小绿叶蝉的唾液腺中特异表达(图 4: A, B),与在褐飞虱、白背飞虱中的组织表达模式(Huang *et al.*, 2016, 2018; Liu *et al.*, 2016; Miao *et al.*, 2018)相同,因此猜测其可能参与了唾液鞘形成及寄主植物的互作。相比较 *EfMLP1* 和 *EfMLP2* 在卵期几乎无表达,有 CBD 结构域的 EfMLP3 和 EfMLP4 蛋白对应的 *EfMLP3* 和 *EfMLP4* 在每个发育阶段均被检测到,与甜菜夜蛾 MLP 表达模式相同,可见该类蛋白贯穿叶蝉的整个生命周期可能具有重要的作用(李雯等, 2010)(图 3: C,

D); *EfMLP3* 和 *EfMLP4* 在茶小绿叶蝉中组织中,除精巢和卵巢外的唾液腺、体壁、脂肪体和肠道中均为广泛表达(图 4: C, D),与棉铃虫、甜菜夜蛾、粉纹夜蛾中发现含有 CBD 结构域的 MLP 蛋白仅在中肠和整个围食膜表现为高表达而其他组织未检测到的结果(Wang and Granados, 1997b; 李雯等, 2010; 刘小民等, 2011)不一致,猜测可能与物种或检测手段差异有关。

刺吸式在取食前会分泌少量胶状唾液形成唾液鞘,唾液鞘具有提供机械稳定性、润滑口针和抵御化学防御等功能,取食完成后唾液鞘残留在植物组织内(Tjallingii, 2006; Will and Vilcinskis, 2015; 杨子银等, 2022)。RNAi 抑制 MLP 表达后将白背飞虱接种于感性水稻品种 TN1 上存活率从第 1 天开始下降,而人工饲料上的存活率没有显著差异(苗雨桐, 2018)。RNAi 抑制 MLP 表达后将褐飞虱接种于抗性水稻 ASD7 上呈现致死效应,而在 TN1 上存活率与对照无差异(Huang *et al.*, 2018)。本研究中,使用饲喂法进行 RNAi 抑制 *EfMLP2* 和 *EfMLP4* 表达后,茶小绿叶蝉在人工饲料饲养下也呈现存活率下降现象(图 5: C)。苗雨桐(2018)认为白背飞虱和褐飞虱接种于 TN1 和人工饲料后的存活率差异是由于取食人工饲料时所需要克服的阻碍较小,本研究中我们发现叶蝉转至人工饲料饲养相对茶苗饲养存活率和最长存活时间均一定程度上降低,而进行 RNAi 抑制后存活率和最长存活时间进一步显著降低,因此仍然可认为 MLP 是叶蝉成功取食的关键蛋白,但对唾液鞘形成和存活率等多方面影响的仍需进一步探究。

RNAi 是一种通过 dsRNA 来高效特异性抑制目标基因表达的技术,广泛应用于基因功能解析和害虫生物防治研究领域(Liu *et al.*, 2020)。本研究通过基因克隆技术获得 4 条 MLP 基因序列, RNA 干扰茶小绿叶蝉 *EfMLP2* 和 *EfMLP4* 后,叶蝉的死亡率显著增加。可见, EfMLP 在茶小绿叶蝉取食中扮演重要的角色,可作为开发 RNAi 的叶蝉防治技术中的潜在靶标。

参考文献 (References)

- Bendtsen JD, Jensen LJ, Blom N, von Heijne G, Brunak S, 2004. Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.*, 17(4): 349–356.
- Bjellqvist B, Basse B, Olsen E, Celis JE, 1994. Reference points for comparisons of two-dimensional maps of proteins from different

human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis*, 15(3-4): 529-539.

Chou KC, Shen HB, 2008. Cell-PLoc: A package of Web servers for predicting subcellular localization of proteins in various organisms. *Nat. Protoc.*, 3(2): 153-162.

Cummings RD, 2009. The repertoire of glycan determinants in the human glycome. *Mol. BioSyst.*, 5(10): 1087-1104.

Dhanisha SS, Guruvayoorappan C, Drishya S, Abeesh P, 2018. Mucins: Structural diversity, biosynthesis, its role in pathogenesis and as possible therapeutic targets. *Crit. Rev. Oncol. Hematol.*, 122: 98-122.

Fang SL, Wang L, Guo W, Zhang X, Peng DH, Luo CP, Yu ZN, Sun M, 2009. *Bacillus thuringiensis* Bel protein enhances the toxicity of Cry1Ac protein to *Helicoverpa armigera* larvae by degrading insect intestinal mucin. *Appl. Environ. Microbiol.*, 75(16): 5237-5243.

Fass D, Thornton DJ, 2023. Mucin networks: Dynamic structural assemblies controlling mucus function. *Curr. Opin. Struct. Biol.*, 79: 102524.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A, 2005. Protein identification and analysis tools on the ExPASy server. In: Walker JM ed. The Proteomics Protocols Handbook. Humana Press, Totowa, NJ. 571-607.

Gupta R, Brunak S, 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.*, 7(3): 310-322.

Hallgren J, Tsigirgos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O, 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *BioRxiv*, <https://doi.org/10.1101/2022.04.08.487609>.

Hegedus D, Erlandson M, Gillott C, Toprak U, 2009. New insights into peritrophic matrix synthesis, architecture, and function. *Annu. Rev. Entomol.*, 54: 285-302.

Huang HJ, Liu CW, Huang XH, Zhou X, Zhuo JC, Zhang CX, Bao YY, 2016. Screening and functional analyses of *Nilaparvata lugens* salivary proteome. *J. Proteome Res.*, 15(6): 1883-1896.

Huang HJ, Lu JB, Li Q, Bao YY, Zhang CX, 2018. Combined transcriptomic/proteomic analysis of salivary gland and secreted saliva in three planthopper species. *J. Proteom.*, 172: 25-35.

Huang HJ, Zhang CX, Hong XY, 2019. How does saliva function in planthopper-host interactions? *Arch. Insect Biochem. Physiol.*, 100(4): e21537.

Johnson M, Zaretskaya I, Raytselis Y, Merezukh Y, McGinnis S, Madden TL, 2008. NCBI BLAST: A better web interface. *Nucleic Acids Res*, 36(Web Server issue): W5-W9.

Li W, Guo W, Zhang X, Li J, Sun WM, 2010. Cloning, sequence analysis and detection of tissue-specific expression of the intestinal protein gene *SeM-8* from *Spodoptera exigua* (Lepidoptera: Noctuidae) peritrophic membrane. *Acta Entomol. Sin.*, 53(7): 727-733. [李雯, 郭巍, 张霞, 李杰, 孙伟明, 2010. 甜菜夜蛾围食膜肠粘蛋白基因 *SeM-8* 的克隆、序列分析及在不同组织中的表达检测. 昆虫学报, 53(7): 727-733]

Liu S, Jaouannet M, Dempsey DA, Imani J, Kogel KH, 2020. RNA-based technologies for insect control in plant production. *Biotechnol. Adv.*, 39: 107463.

Liu SS, Jaouannet M, Dempsey DA, Imani J, Coustau C, Kogel KH, 2020. RNA-based technologies for insect control in plant production. *Biotechnol. Adv.*, 39: 107463.

Liu XM, Li J, Guo W, Xu DQ, Zhang X, 2011. Cloning, expression and histolocalization of an intestinal mucin gene *HM72* from *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae). *Acta Entomol. Sin.*, 54(2): 127-135. [刘小民, 李杰, 郭巍, 徐大庆, 张霞, 2011. 棉铃虫围食膜肠粘蛋白基因 *HM72* 的克隆、表达及其蛋白组织定位. 昆虫学报, 54(2): 127-135]

Liu XQ, Zhou HY, Zhao J, Hua HX, He YP, 2016. Identification of the secreted watery saliva proteins of the rice brown planthopper, *Nilaparvata lugens* (Stål) by transcriptome and shotgun LC-MS/MS approach. *J. Insect Physiol.*, 89: 60-69.

Liu YQ, Tian KH, Sun QL, Pan C, Li YY, Jiang JY, Jiang CJ, 2017. Cloning and expression analysis of miR156a-targeted genes *SPL6* and *SPL9* in *Camellia sinensis*. *J. Tea Sci.*, 37(6): 551-564. [刘亚芹, 田坤红, 孙琪璐, 潘铖, 李叶云, 蒋家月, 江昌俊, 2017. 茶树 miR156a 靶基因 *SPL6* 和 *SPL9* 的克隆及表达分析. 茶叶科学, 37(6): 551-564]

Livak KJ, Schmittgen TD, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C (T))} method. *Methods*, 25(4): 402-408.

Loukas A, Hintz M, Linder D, Mullin NP, Parkinson J, Tetteh KK, Maizels RM, 2000. A family of secreted mucins from the parasitic nematode *Toxocara canis* bears diverse mucin domains but shares similar flanking six-cysteine repeat motifs. *J. Biol. Chem.*, 275(50): 39600-39607.

Miao YT, 2018. The Proteomic Analysis of Watery Saliva Secreted by *Sogatella furcifera* and Functional Analysis of *Mucin*. MSc Thesis, Chinese Academy of Agricultural Sciences, Beijing. [苗雨桐, 2018. 白背飞虱水状唾液蛋白组学分析及 *Mucin* 基因功能研究. 北京: 中国农业科学院硕士学位论文]

Miao YT, Deng Y, Jia HK, Liu YD, Hou ML, 2018. Proteomic analysis of watery saliva secreted by white-backed planthopper, *Sogatella furcifera*. *PLoS ONE*, 13(5): e0193831.

Monier M, Courtier-Ordogozo V, 2022. *Drosophila* glue: A promising model for bioadhesion. *Insects*, 13(8): 734.

Nielsen H, Tsigirgos KD, Brunak S, von Heijne G, 2019. A brief history of protein sorting prediction. *Protein J.*, 38(3): 200-216.

Paim RM, Pereira MH, Pereira MH, Di Ponzio R, Rodrigues JO, Guarneri AA, Gontijo NF, Araújo RN, 2012. Validation of reference genes for expression analysis in the salivary gland and the intestine of *Rhodnius prolixus* (Hemiptera, Reduviidae) under different experimental conditions by quantitative real-time PCR. *BMC Res. Notes*, 5: 128.

Pan C, Niu YQ, Xia LX, Wu CF, Jing KT, Cheng QY, Wang MX, Han BY, 2021. Identification of bacterial proteins in the saliva of adults of the tea green leafhopper, *Empoasca flavescens* (Hemiptera:

- Cicadellidae). *Acta Entomol. Sin.*, 64(8): 929–942. [潘铖, 钮羽群, 夏露霞, 伍春芳, 景凯婷, 程墙义, 王梦馨, 韩宝瑜, 2021. 茶小绿叶蝉成虫唾液细菌蛋白的鉴定. 昆虫学报, 64(8): 929–942]
- Perez-Vilar J, Hill RL, 1999. The structure and assembly of secreted mucins. *J. Biol. Chem.*, 274(45): 31751–31754.
- Rose MC, Voynow JA, 2006. Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol. Rev.*, 86(1): 245–278.
- Sarauer BL, Gillott C, Hegedus D, 2003. Characterization of an intestinal mucin from the peritrophic matrix of the diamondback moth, *Plutella xylostella*. *Insect Mol. Biol.*, 12(4): 333–343.
- Schjoldager KTBG, Clausen H, 2012. Site-specific protein O-glycosylation modulates proprotein processing? Deciphering specific functions of the large polypeptide GalNAc-transferase gene family. *Biochim. Biophys. Acta*, 1820(12): 2079–2094.
- Shen ZC, Jacobs-Lorena M, 1999. Evolution of chitin-binding proteins in invertebrates. *J. Mol. Evol.*, 48(3): 341–347.
- Shi XZ, Chamankhah M, Visal-Shah S, Hemmingsen SM, Erlandson M, Braun L, Alting-Mees M, Khachatourians GG, O'Grady M, Hegedus DD, 2004. Modeling the structure of the type I peritrophic matrix; Characterization of a *Mamestra configurata* intestinal mucin and a novel peritrophin containing 19 chitin binding domains. *Insect Biochem. Mol. Biol.*, 34(10): 1101–1115.
- Sigrist CJ, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P, 2002. PROSITE: A documented database using patterns and profiles as motif descriptors. *Brief. Bioinform.*, 3(3): 265–274.
- Sigrist CJA, De Castro E, Langendijk-Genevaux PS, Le Saux V, Bairoch A, Hulo N, 2005. ProRule: A new database containing functional and structural information on PROSITE profiles. *Bioinformatics*, 21(21): 4060–4066.
- Steenftoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Levery SB, Clausen H, 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.*, 32(10): 1478–1488.
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S, 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.*, 30(12): 2725–2729.
- Tang QY, 2020. DPS Data Processing System, Vol. III. Specialized Statistics and Miscellanea. 5th ed. Science Press, Beijing. 12. [唐启义, 2020. DPS 数据处理系统, 第3卷. 专业统计及其他(第5版). 北京: 科学出版社. 12]
- Tellam RL, Wijffels G, Willadsen P, 1999. Peritrophic matrix proteins. *Insect Biochem. Mol. Biol.*, 29(2): 87–101.
- Tjallingii WF, 2006. Salivary secretions by aphids interacting with proteins of phloem wound responses. *J. Exp. Bot.*, 57(4): 739–745.
- Wang P, Granados RR, 1997a. An intestinal mucin is the target substrate for a baculovirus enhancer. *Proc. Natl. Acad. Sci. USA*, 94(13): 6977–6982.
- Wang P, Granados RR, 1997b. Molecular cloning and sequencing of a novel invertebrate intestinal mucin cDNA. *J. Biol. Chem.*, 272(26): 16663–16669.
- Wang P, Granados RR, 2001. Molecular structure of the peritrophic membrane (PM): Identification of potential PM target sites for insect control. *Arch. Insect Biochem. Physiol.*, 47(2): 110–118.
- Will T, Vilcinskis A, 2015. The structural sheath protein of aphids is required for phloem feeding. *Insect Biochem. Mol. Biol.*, 57: 34–40.
- Wu P, Sun P, Nie KX, Zhu YB, Shi MY, Xiao CG, Liu H, Liu QY, Zhao TY, Chen XG, Zhou HN, Wang PH, Cheng G, 2019. A gut commensal bacterium promotes mosquito permissiveness to arboviruses. *Cell Host Microbe*, 25(1): 101–112. e5.
- Yang ZY, Wu SH, Gu DC, 2022. Research progress on effects of tea green leafhopper infestation on tea plant growth and tea leaf quality. *J. Tea Commun.*, 49(1): 1–11. [杨子银, 吴淑华, 辜大川, 2022. 茶小绿叶蝉侵害对茶树生长和茶叶品质影响的研究进展. 茶叶通讯, 49(1): 1–11]
- Zhang X, Li GX, Guo W, 2008. Molecular cloning and sequencing of an insect intestinal mucin cDNA from *Spodoptera exigua*. *Sci. Agric. Sin.*, 41(11): 3898–3904. [张霞, 李国勋, 郭巍, 2008. 甜菜夜蛾肠粘蛋白 cDNA 基因的分子克隆与序列分析. 中国农业科学, 41(11): 3898–3904]
- Zhao XM, Zhang J, Yang JP, Niu N, Zhang JZ, Yang Q, 2020. Mucin family genes are essential for the growth and development of the migratory locust, *Locusta migratoria*. *Insect Biochem. Mol. Biol.*, 123: 103404.
- Zheng JF, Ge L, 2022. Diverse cellular strategies for the export of leaderless proteins. *Natl. Sci. Open*, 1(3): 20220018.

(责任编辑: 马丽萍)