DOI:10.12113/202002003

毛竹 *SWEET* 基因家族的全基因组鉴定与分析

李新然[1,2],张智俊[1,2*],喻珮瑶[1,2],陈家璐[1,2],马瑞芳[1,2],刘笑雨[1,2]
(1.浙江农林大学 省部共建亚热带森林培育国家重点实验室,杭州 311300;
2.浙江农林大学 林业与生物技术学院,杭州 311300)

摘　要:糖外排转运蛋白(Sugars will eventually be exported transporters, SWEETs)在植物生理活动和发育过程中起重要作用。为探究 *SWEET* 基因家族在毛竹(*Phyllostachys edulis*)的生长发育过程中起的作用,基于毛竹基因组数据,通过生物信息学方法对 *SWEET* 基因家族成员进行鉴定,并对其编码的蛋白质理化性质、系统进化及共线性关系、基因结构、启动子元件及表达模式、蛋白互作网路分析、GO 注释等进行细致分析。研究结果表明:该家族基因结构、基序和结构域相对保守,所有成员均含有 *MtN3_slv* 结构域。上游启动子序列中含有多个同非生物胁迫以及生长发育相关元件,结合转录组表达量分析显示,多个家族成员在毛竹不同组织器官均有表达。共线性分析揭示毛竹 *SWEET* 家族在演化过程中存在全基因组多倍化事件。蛋白互作网路分析挖掘出2个重要核心家族成员,GO 注释分析进一步证实毛竹 *SWEET* 主要负责体内糖类物质的转运。以上结果为毛竹 *SWEET* 基因功能鉴定提供了重要参考,对于毛竹快速生长分子机制研究打下了基础。

关键词:毛竹;*SWEET* 基因家族;基因分析;功能鉴定
中图分类号:Q343.1　　**文献标志码**:A　　**文章编号**:1672-5565(2020)04-236-11

Genomic identification and expression analysis of *SWEET* gene family in moso bamboo (*Phyllostachys edulis*)

LI Xinran[1,2], ZHANG Zhijun[1,2*], YU Peiyao[1,2], CHEN Jialu[1,2], MA Ruifang[1,2], LIU Xiaoyu[1,2]
(1.*State Key Laboratory of Subtropical Forest Cultivation*, *Zhejiang A & F University*, *Hangzhou* 311300, *China*
2.*College of Forestry and Biotechnology*, *Zhejiang A & F University*, *Hangzhou* 311300, *China*)

Abstract: Sugars will eventually be exported transporters (*SWEETs*) play an important role in plant growth and development. In order to explore the function of *SWEET* gene family in the growth and development of moso bamboo (*Phyllostachys edulis*), the physicochemical properties, phylogenetic relationship, genes structure and expression pattern protein-protein interaction, and GO annotation of the *SWEET* family were analyzed in detail by bioinformatics methods based on genome and related transcriptome data from moso bamboo (*Phyllostachys edulis*). Results show that the gene structure, motif, and domain of this family were relatively conserved, and all members contained the *MtN3_slv* domain. The upstream promoter sequences contained multiple cis-acting elements related to growth and development. Multiple *SWEET* family members were expressed in different tissues and organs of moso bamboo by transcriptome expression analysis. Collinear analysis shows that there were genome-wide polyploidy events in the *SWEET* family of bamboo. Two important core family members were disclosed by protein-protein interaction (PPI) network analysis. The analysis of GO annotation further proves that *SWEET* was mainly responsible for the transport of carbohydrate in the body. The results of this study will help to reveal the biological functions of the *SWEET* gene family and lay a foundation for the study of the molecular mechanism of the rapid growth of moso bamboo.

Keywords: *Phyllostachys edulis*; *SWEET*; Gene analysis; Functional identification

---

收稿日期:2020-02-07;**修回日期**:2020-03-11.
基金项目:国家自然科学基金项目(No.31770721).
作者简介:李新然,男,本科生,研究方向:生物信息学.E-mail:861068481@qq.com.
*通信作者:张智俊,男,副教授,硕士生导师,研究方向:生物信息学.E-mail:zjzhang@zafu.edu.cn.

毛竹(*Phyllostachys edulis*)是我国非常重要的经济类植物,具有适应环境性强、产量高、生长周期短的特点,是世界上高生长最快的禾本科植物[1]。近年来,竹类植物的高生长机理一直是研究的热点问题。已有研究表明,竹笋生长发育过程中的许多生物过程如植物激素、细胞的生长与发育、代谢与能量、光合作用等都影响着竹类植物的高生长[1]。作为光合作用的主要产物,糖类为竹笋高生长过程中细胞的分裂和伸长提供大量能量。因此,对糖类转运相关基因的研究,将有助于深入探究竹类植物高生长的分子机理。

SWEET 蛋白是近年发现的一类不依赖环境 pH 值,顺浓度梯度双向跨膜运输的低亲和糖转运蛋白[2]。*SWEET* 家族成员分布非常广泛,高等真核生物中普遍存在。另外,在原生动物、后生动物、真菌、细菌,古生菌中也有分布[3-4]。由 *SWEET* 基因编码的膜蛋白均具有一定数量的保守跨膜结构域,被命名为 *MtN3/saliva* [3]。该结构域最早在苜蓿(*Medicago Sativa*)根部结瘤素(*nodulin*)中被发现[5]。目前已有研究对主要的 *SWEET* 蛋白做系统发育分析,并分为三大分支。第 1 分支为植物类 *SWEET* 基因,大部分蛋白含有 2 个 *MtN3/saliva* 跨膜结构域水;第 2 分支为动物类 *SWEET* 基因,具有 2 个特定 *MtN3/saliva* 结构域;第 3 分支是从细菌到古生菌(球菌属)及线虫的某些 *SWEET* 蛋白, 该分支的 *SWEET* 蛋白由具有 3 个跨膜阿尔法螺旋的一个 *MtN3/saliva* 结构域组成[2]。自 *SWEET* 基因家族发现以来,相继在拟南芥(*Arabidopsis thaliana*)、葡萄(*Vitis vinifera*)、番茄(*Solanum lycopersicum*)、木薯(*Manihot esculenta*)[6-9]等植物中发现了数量不等的 *SWEET* 基因家族成员。*SWEET* 基因家族成员参与许多生理过程。有研究表明,拟南芥中的 RPG1(*ATSWEET8/AT5g40260*)编码的膜蛋白对拟南芥小孢子外壁的形成是必须的[10];在水稻(*Oryza sativa*)中该基因家族成员也涉及到生殖发育[11];拟南芥位于质膜上的蛋白质 *SAG29*(*AtSWEETl5*)蛋白,能在高盐环境下调节细胞活性[12];*SWEET* 基因也参与衰老过程的调控,水稻中 *OsSWEET5* 的过表达会使植株在幼苗阶段表现出生长延迟和过早衰老, 而该基因敲除植株并没有引起表型的变化[13]。Chu 等研究发现水稻 *OsSWEET11*(Xa13)突变的植株能够抵抗白粉病[14],这表明 *SWEET* 基因在行使糖转运蛋白功能的同时也参与了寄主与病原菌的互作,在促进植物免疫中也起到重要作用。但目前尚未有 *SWEET* 家族在竹类植物中的研究报道。

随着毛竹基因组测序的完成,使得在基因组水平上研究毛竹 *SWEET* 家族的信息成为可能。本研究拟鉴定出毛竹 *SWEET* 基因家族成员,通过蛋白质理化性质、系统进化关系、基因结构、启动子元件、蛋白互作网络、GO 注释分析,并结合相关转录组数据挖掘,发现 *SWEET* 家族在毛竹不同组织以及不同激素水平下的表达情况,为进一步研究 *SWEET* 基因家族的结构与功能、揭示毛竹快速生长的分子机制打下基础。

1　材料与方法

1.1　毛竹 *SWEET* 家族成员鉴定及理化性质分析

在 ftp://parrot.genomics.cn/gigadb/pub/10.5524/100001_101000/100498/站点下载毛竹的基因组序列、基因编码区序列(Coding sequence, CDS)、蛋白序列以及基因组注释(General feature format, GFF)在内的数据文件。在 Pfam 数据库(http://pfam.xfam.org/)中以 *SWEET* 隐马尔可夫模型(HMM)PF03083 为模板[15]下载结构域数据,以此为种子模型利用 HMMER3 软件(http://hmmer.janelia.org/)检索本地毛竹蛋白数据库,筛选 E 值设置为小于等于 1×10^{-20},对检索结果进行排序整理,去重处理,初步获得候选基因家族成员。进一步通过 TBtools 软件从毛竹全基因组数据库中,提取得到毛竹 *SWEET* 基因家族成员的基因、CDS、蛋白 fasta 序列以及基因基因结构和位置信息。利用在线工具 ProtParam (https://web.expasy.org/protparam/)、ProtScale (https://web.expasy.org/protscale/) 分析 *SWEET* 基因家族各成员理化性质。利用 TMHMM (http://www.cbs.dtu.dk/services/TMHMM/)在线工具对基因家族各成员跨膜区进行分析。

1.2　*SWEET* 基因家族进化分析

拟南芥(*Arabidopsis thaliana*)和水稻(*Oryza sativa*)全基因组信息分别从拟南芥数据库 TAIR10 (http://www.arabidopsis.org/index.jsp)以及水稻数据库(http://rice.plantbiology.msu.edu)中下载,并用 HMMER3 软件对相应本地数据库进行检索,经过筛选后得到对应的拟南芥和水稻的 *SWEET* 基因家族蛋白序列。对毛竹、拟南芥、水稻 *SWEET* 基因家族成员蛋白序列进行 ClustalW 多重比对,用 MEGA 7.0 (http://www.megasoftware.net)软件邻位连接法构建种内和种间系统进化树,自检值取 1 000 次抽样。

1.3　*SWEET* 基因家族染色体分布及种内、种间共线性分析

利用 Tbtools 软件(v0.66836)中的 BLAST 模块

将毛竹基因组所有蛋白序列自身比对及毛竹与水稻、毛竹与拟南芥基因组蛋白序列两两比对,结合三个物种全基因组染色体位置信息,利用 MCScanX[16]获取 *SWEET* 家族染色体分布和种间共线性关系,并用 Circos 和 Multipe synteny plot 分别对种内和种间的共线性结果进行可视化。

1.4 *SWEET* 基因家族基序和保守结构域预测

通过在线网站 NCBI Conserved Domain (https://www.ncbi.nlm.nih.gov/cdd/) 和 MEME (http://meme.sdsc.edu/meme/meme-intro.html) 对 *SWEET* 家族成员的保守结构域(domain)和基序(Motif)进行预测[17]。

1.5 *SWEET* 基因蛋白质互作网络预测分析

在 Search Tool for the Retrieval of Interacting Genes/ Proteins(STRING)数据库(https://string-db.org/)中上传 *SWEET* 基因家族蛋白序列,选择拟南芥蛋白库进行比对,根据已知的蛋白质互作关系,*SWEET* 基因家族的蛋白互作信息进行评估和预测。

1.6 *SWEET* 基因家族 GO 富集分析

对毛竹 *SWEET* 基因家族的 30 个基因构建基因集,利用在线软件(https://www.i-sanger.com/)对毛竹的差异表达基因集 *SWEET* 参与的生物功能进行 GO(Gene ontology)注释富集预测分析。

1.7 *SWEET* 基因家族成员基因结构及启动子特征分析

依据毛竹全基因组的 GFF 注释文件基因位置信息,分析毛竹 *SWEET* 家族的基因结构,同时绘制基因结构图;提取毛竹 *SWEET* 基因上游 1 500 bp 序列作为启动子序列信息,利用在线工具 PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)[18]预测毛竹 *SWEET* 基因的顺式作用元件,并将预测结果整理筛选出顺式作用元件,在 TBtools 上进行启动子位置可视化。

1.8 *SWEET* 基因家族表达分析

选取 Ensemble 数据库中毛竹不同部位(叶、花、竹鞭、根)的 4 组转录组数据(登录号 ERR105067, ERR105069, ERR105073, ERR105075),以及选取 NCBI SRA 数据库中毛竹幼苗根组织经过赤霉素(Gibberellin, GA)、萘乙酸(Naphthalene acetic acid, NAA)处理的共 12 组转录组数据(登录号 SRR613113, SRR613114, SRR613115, SRR613116, SRR613117, SRR613118; SRR5710702, SRR5710701, SRR5710700, SRR5710699, SRR5710698, SRR5710697),分别计算毛竹 *SWEET* 基因的表达丰度 TPM (Transcripts per million reads)值,以 log2 TPM 值利用 TBtools Amazing Heatmap 绘制基因表达热图。

2 实验结果与分析

2.1 *SWEET* 基因家族的鉴定

通过植物 *SWEET* Pfam (PF03083)模型搜索出 68 个家族成员,通过 E 值≤1×10^{-20}筛选初步得到 34 个成员,再结合保守结构域、基因结构、染色体定位等特征,进一步筛选去除重复转录本,最终得到 30 个 *SWEET* 基因家族成员(见表 1)。按照这些基因在染色体图中出现的位置顺序,依次命名为 *PeSWEET*1-*PeSWEET*30。*SWEET* 基因家族的 30 个成员的氨基酸序列及理化性质分析如表 1 所示。家族成员中蛋白分子量最大的为63.85 kD,蛋白分子量最小的为25.11 kD。氨基酸序列长度最长为 229 aa, 最短的长度为 582 aa。等电点位于 4.96~9.39 之间,其中 23 个家族蛋白为碱性蛋白(理论等电点>7),有 6 个为酸性蛋白(理论等电点<7)。所有家族成员蛋白的平均疏水指数都大于零,均为疏水蛋白。跨膜结构域预测结果显示所以家族成员均具有跨膜结构域,且有 21 个家族成员蛋白含有 7 个跨膜结构域。

2.2 *SWEET* 基因家族分类与进化树构建

选取 30 个水稻 *SWEET* 家族成员与 15 个拟南芥和 20 个水稻 *SWEET* 家族成员的氨基酸序列,进行 NJ 法系统聚类分析,以分析毛竹 *SWEET* 基因家族的进化分类,参考 *SWEET* 家族在其它物种之间的分类)[9],毛竹 *SWEET* 家族可分为四个亚家族(I-IV)(见图 1)。其中 I 亚家族的毛竹 *SWEET* 家族成员数量最多,为 17 个。各个亚族中拟南芥 *SWEET* 基因的氨基酸序列自成一个支,水稻和毛竹的 ZF-HD 基因的氨基酸序列聚成一支,表明 *SWEET* 亚族中毛竹与水稻的亲缘关系较近,与拟南芥亲缘关系较远。

2.3 毛竹 *SWEET* 家族染色体定位及共线性分析

毛竹 *SWEET* 基因家族的染色体分布状况如图 2a 所示,30 个家族成员分布在 15 条染色体上,不同的染色体基因分布密度不同,并且有 4 对基因发生了并联复制。从单条染色体看,scaffold_14 上分布的家族成员最多,出现了 4 个 *SWEET* 基因。毛竹与拟南芥、水稻之间的共线性关系如图 2b 所示,仅有 5 个 *SWEET* 同源蛋白基因出现在拟南芥染色体中,有 25 个毛竹 *SWEET* 基因可以在 7 条水稻染色体上找到对应的旁系同源基因。由此可见,毛竹与水稻之间的亲缘关系较近。此外,有 9 个水稻中的 *SWEET* 家族成员在毛竹中有对应的两个同源拷贝,个别家族成员有 3 个拷贝,且仅有 3 个 *SWEET* 基因

在毛竹中没有发现相对应的拷贝。由此推断，　　化事件。
SWEET 家族在演化过程中可能存在全基因组多倍

表 1　毛竹 *SWEET* 基因编码蛋白序列理化性质

Table 1　Physicochemical properties of proteins encoded by *SWEET* in *Phyllostachys edulis*

ID 号 ID number	基因名称 Gene name	等电点 Theoretical pI	相对分子量(kD) Molecular weight (kD)	氨基酸数量(aa) Number of amino acid (aa)	脂肪族氨基酸指数 Aliphatic index	平均疏水指数 GRAVY	跨膜 α-螺旋 Number of predicted TMHs
PH02Gene47308.t1	*PeSWEET*01	8.25	32.53	297	114.81	0.493	7
PH02Gene19617.t1	*PeSWEET*02	9.30	31.18	284	110.49	0.837	7
PH02Gene06361.t1	*PeSWEET*03	9.10	26.64	242	114.71	0.697	7
PH02Gene15530.t1	*PeSWEET*04	9.36	27.42	254	108.66	0.615	6
PH02Gene16827.t1	*PeSWEET*05	9.30	27.51	245	106.94	0.649	6
PH02Gene16828.t1	*PeSWEET*06	9.39	27.34	245	106.53	0.633	6
PH02Gene39291.t1	*PeSWEET*07	8.86	31.91	293	118.81	0.709	7
PH02Gene21722.t1	*PeSWEET*08	9.10	63.85	582	102.25	0.358	8
PH02Gene08624.t1	*PeSWEET*09	9.02	33.03	302	114.34	0.769	7
PH02Gene31774.t1	*PeSWEET*10	6.32	29.48	267	101.16	0.225	5
PH02Gene31775.t1	*PeSWEET*11	6.32	32.70	291	97.83	0.272	5
PH02Gene48050.t1	*PeSWEET*12	6.32	33.89	310	111.65	0.447	7
PH02Gene06545.t1	*PeSWEET*13	8.68	25.11	229	121.88	0.914	7
PH02Gene07867.t1	*PeSWEET*14	9.27	27.90	256	116.13	0.647	6
PH02Gene31772.t1	*PeSWEET*15	6.28	27.07	249	111.61	0.681	7
PH02Gene36743.t1	*PeSWEET*16	9.18	26.98	246	120.33	0.725	7
PH02Gene26486.t1	*PeSWEET*17	9.38	33.69	308	102.86	0.337	7
PH02Gene48092.t1	*PeSWEET*18	5.70	25.49	280	118.61	0.684	5
PH02Gene10697.t1	*PeSWEET*19	9.17	28.99	270	115.96	0.689	6
PH02Gene23249.t1	*PeSWEET*20	4.96	27.30	249	112.01	0.705	7
PH02Gene28963.t1	*PeSWEET*21	9.18	27.11	246	120.33	0.722	7
PH02Gene23481.t1	*PeSWEET*22	9.26	27.60	254	119.02	0.714	7
PH02Gene38775.t1	*PeSWEET*23	9.06	34.13	312	117.18	0.654	7
PH02Gene30953.t1	*PeSWEET*24	9.30	27.54	246	112.44	0.704	7
PH02Gene36245.t2	*PeSWEET*25	9.08	29.66	270	111.22	0.854	7
PH02Gene39928.t1	*PeSWEET*26	9.30	27.54	246	112.44	0.704	7
PH02Gene37155.t1	*PeSWEET*27	9.28	27.52	254	119.37	0.728	7
PH02Gene44746.t1	*PeSWEET*28	7.00	32.10	294	114.76	0.711	7
PH02Gene41640.t1	*PeSWEET*29	9.08	33.78	233	107.23	0.427	7
PH02Gene43386.t1	*PeSWEET*30	9.02	31.73	290	115.31	0.188	7

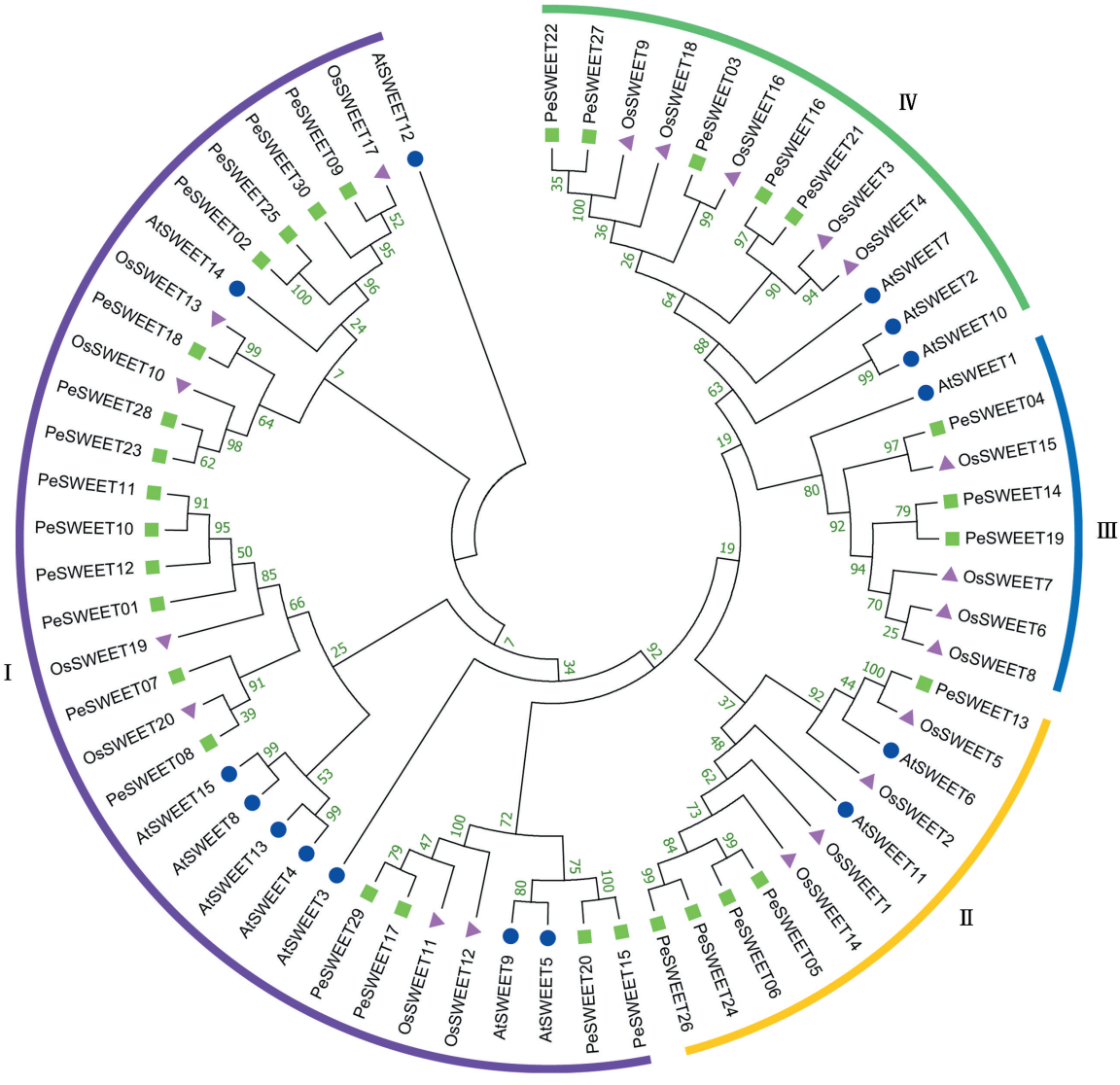


图 1 毛竹(*Phyllostachys edulis*),拟南芥(*Arabidopsis thaliana*),水稻(*Oryza sative*) *SWEET* 家族系统进化树

Fig.1 Phylogentic tree of *SWEET* gene family from *Phyllostachys edulis*(Pe), *Arabidopsis thaliana*(At), and *Oryza sative*(Os)

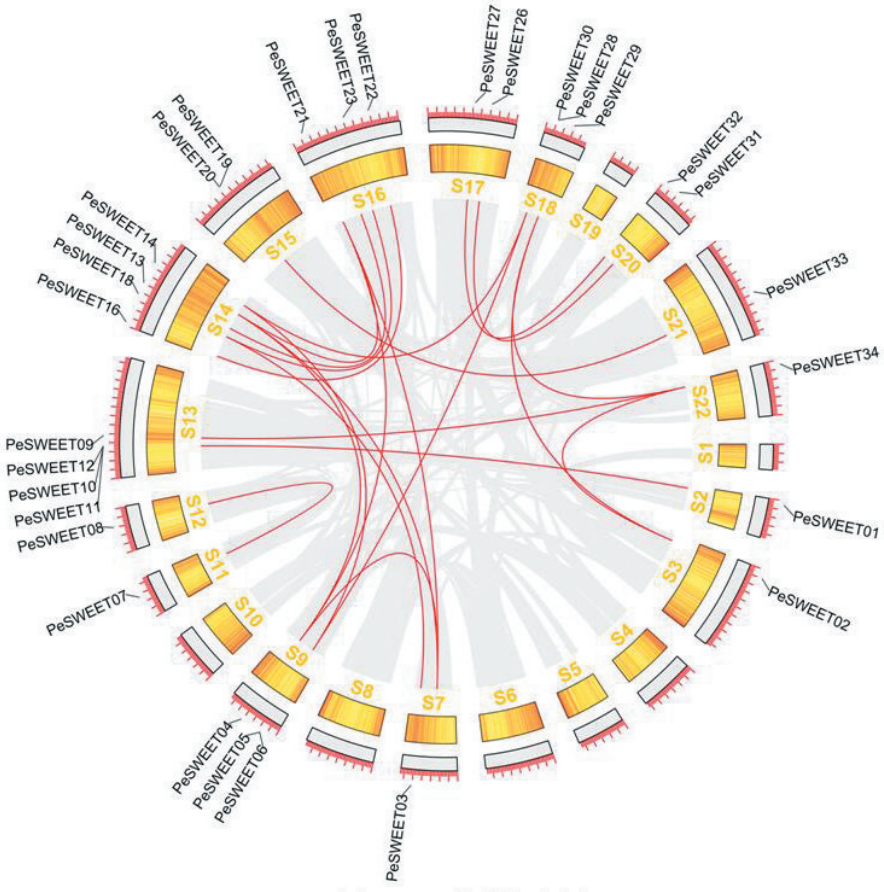
2.4 毛竹 *SWEET* 家族基序及保守结构域分析

依据 NCBI Conserved Domain 进一步分析毛竹 *SWEET* 基因家族的保守结构域如图 3a,该家族所有成员均含有 *MtN*3_*slv* 结构域。*PeSWEET*4、5、6、13、14、18、19、22、24、26 除了具有典型的 *MtN*3_*slv* 结构域外,还含有 *PQ-loop superfamily* 结构域,且多位于 N 端。*PQ-loop superfamily* 与 *MtN*3_*slv* 在氨基酸序列上十分接近,都具有糖运输功能[19]。其余的基因家族成员均含有两个 *MtN*3_*slv* 结构域。由此可知,毛竹 *SWEET* 基因家族 *MtN*3_*slv* 结构域具有一定的保守性,但该结构域的分布位置和数量具有一定分化。

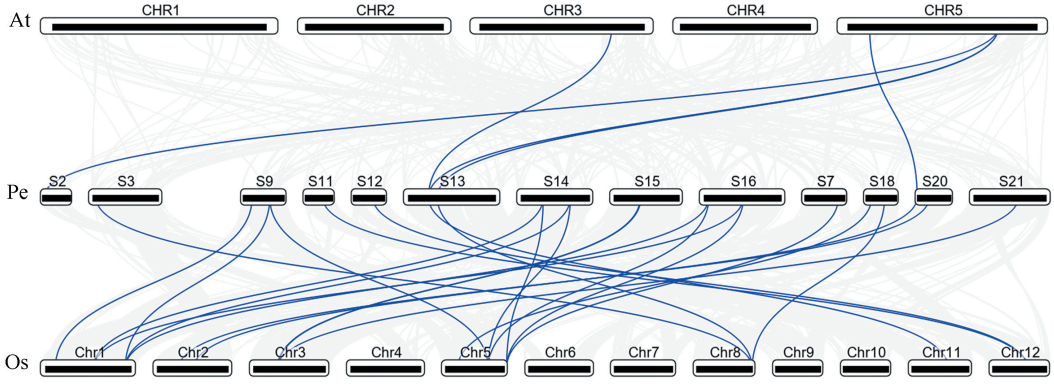
利用在线工具 MEME 分析 30 个毛竹 *SWEET* 基因家族成员(见图 3b),发现该家族成员含有 3~7 个数量不等的基序。其中 *PeSWEET*18 只含有 3 个基序,只有 *PeSWEET*18 不含有 Motif 4。Motif 1、Motif 2、Motif 3 在 30 个 *SWEET* 基因家族成员中均有出现,说明这三个基序为高度保守结构。

2.5 毛竹 *SWEET* 家族蛋白质互作 PPI 网络分析

利用分子互作网络检索工具 STRING 数据库预测 *SWEET* 基因家族的蛋白互作特性(见图 4)。整个蛋白互作网络共有 9 个节点,节点间存在 18 组蛋白互作关系。如图可知,每个蛋白节点与其他蛋白节点都有一定的互作关系,其中 *PeSWEET*12、*PeSWEET*15 位于整个 *SWEET* 基因家族蛋白互作网络的中心,分别与 7 个家族成员蛋白产生相互作用。



(a) 毛竹SWEET基因染色体分布



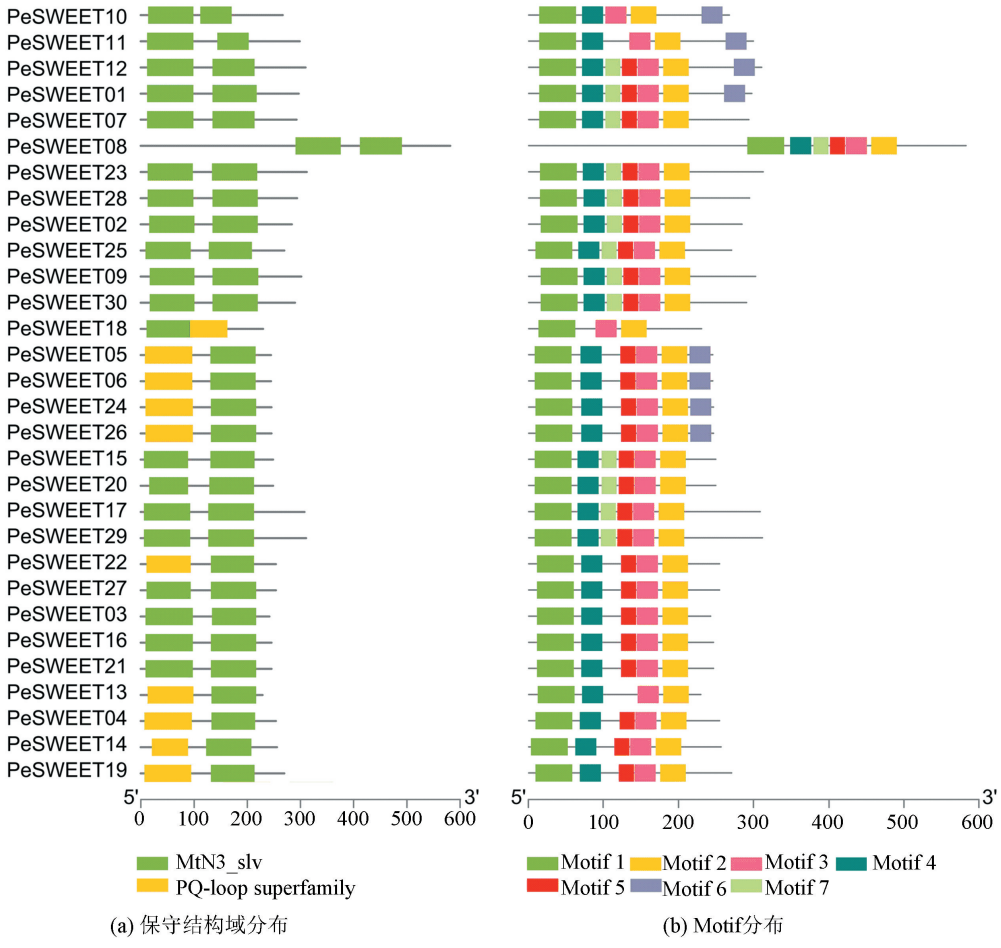
(b) 毛竹与拟南芥及水稻种间共线性分析

图 2　毛竹 *SWEET* 家族染色体分布及种间共线性分析

Fig.2　Chromosome distribution and inter-species collinearity analysis of *SWEET* family in *Phyllostachys edulis*

注:(a)毛竹 *SWEET* 基因染色体分布,灰色线条代表毛竹所有成员的共线性关系,红色线条代表 *SWEET* 家族成员间共线性关系,黄色区域代表基因密度,S 代表了染色体 Scaffold。(b)毛竹与拟南芥及水稻种间共线性分析,At 代表拟南芥,Pe 代表毛竹,Os 代表水稻;CHR 和 Chr 代表了 Chromosome ,S 代表了染色体 Scaffold ;基因名称如表 1 所示。

Notes:(a) Chromosomal distribution of *SWEET* gene in *Phyllostachys edulis*. The gray lines represent the collinear relationship of all the members of *Phyllostachys edulis*. The red line represents the collinear relationship between the members of the *SWEET* family. The yellow area represents the gene density and S represents the chromosome Scaffold. (b) Interlinear analysis of *Phyllostachys edulis* and *Arabidopsis thaliana* and *Oryza sativa*. At represents *Arabidopsis thaliana*, Pe represents *Phyllostachys edulis*, and Os represents *Oryza sativa*. CHR and Chr represent Chromosome, and S represents the chromosome Scaffold. The gene names are shown in Table 1.



(a) 保守结构域分布 (b) Motif分布

图 3 毛竹 *SWEET* 基因家族结构域分析和保守元件分析

Fig.3 Anaylsis of conserved domain and motif elements of *SWEET* gene family in *Phyllostachys edulis*

注:图中的刻度尺代表氨基酸酸序列长度;基因名称如表 1 所示.

Notes: The scale represents the length of the amino acid sequence. The gene names are shown in table 1.

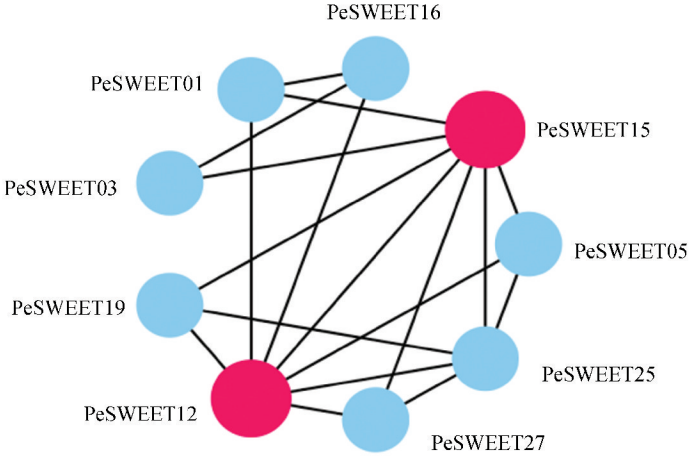


图 4 毛竹 *SWEET* 家族蛋白互作网络预测图

Fig.4 Protein-protein interaction network of *SWEET* gene family in *Phyllostachys edulis*

2.6 毛竹 *SWEET* 家族基因 GO 功能注释

利用 GO 对 *SWEET* 家族基因进行功能注释分析,GO 结构显示该家族基因的功能主要富集于生物进程(Biological process, BP)、细胞组分(Cellular component, CC)和分子功能(Molecular function, MF) 3 大类共 20 个小类上(见图 5)。在 BP 分类

中,所有家族成员均具有定位类别(Localization)和单有机体过程类别(Single-organism process)。在 CC 分类中,细胞膜(Membrane)以及膜组件类别(Membrane part)占比例最多,为所有成员共有,其次为细胞(Cell)以及细胞部分类别(Cell part)各占比百分之九十。在 MF 分类中 *SWEET* 家族基因的分子功能主要富集在转运活性(Transporter activity)功能。GO 富集结果表明,30 个毛竹 *SWEET* 基因对毛竹体内物质的转运起重要作用。

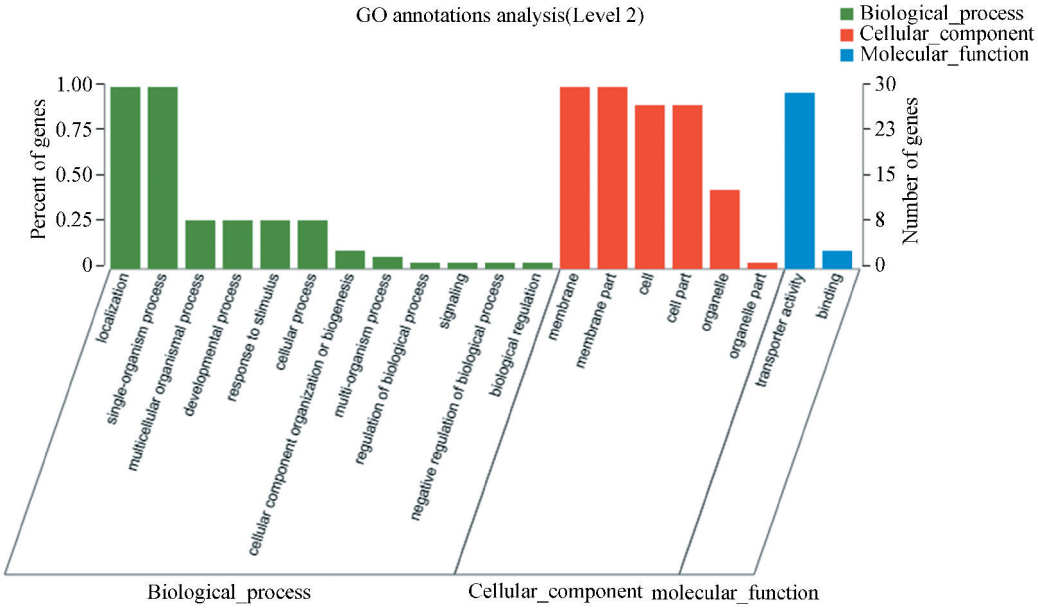


图 5　毛竹 *SWEET* 基因家族的 GO 注释分析

Fig.5　GO annotation analysis of *SWEET* familly in *phyllostachys edulis*

注:绿色部分 BP 表示生物进程,红色部分 CC 表示细胞组分,蓝色部分 MF 表示分子功能.

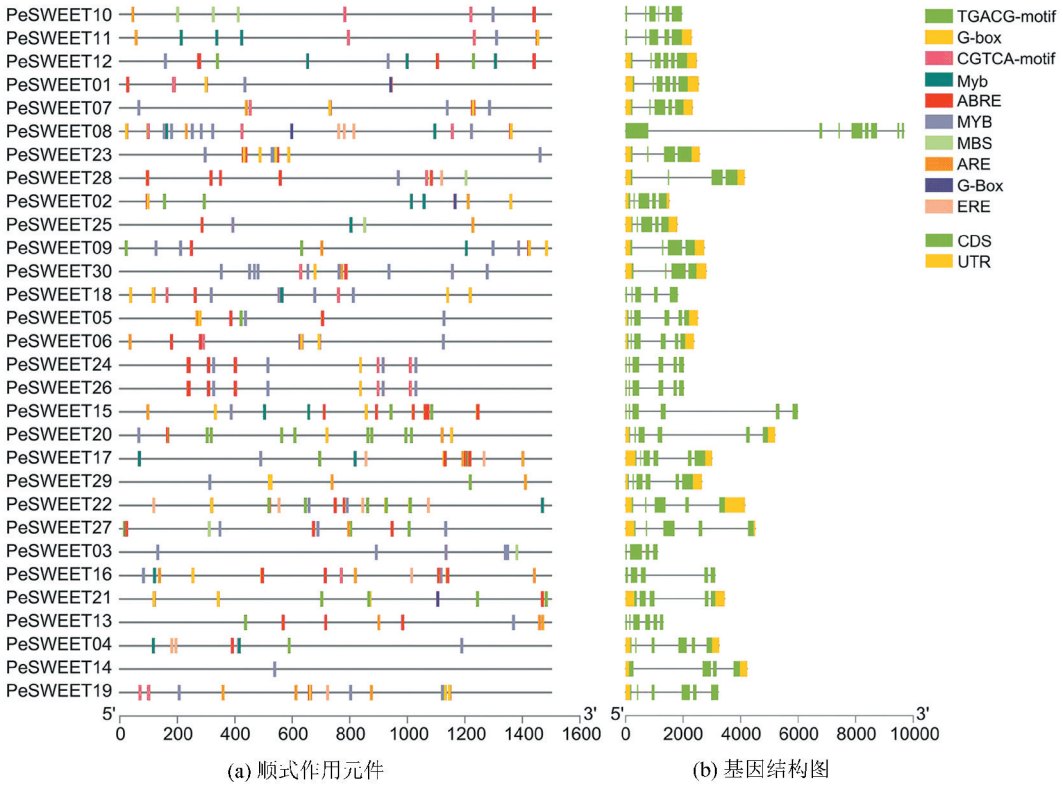
Notes: The green part BP indicates the biological process, the red part CC indicates the cellular component, and the blue part MF indicates the molecular function.

2.7　毛竹 *SWEET* 家族基因结构及启动子特征分析

提取 30 条毛竹 *SWEET* 基因上游1 500 bp的核苷酸序列,进行启动子顺势作用元件预测,结果表明除启动子核心元件 TATA-box、CAAT-box 以外,各家族成员上还存在多种作用元件(见图 6a)。如光响应相关元件(Light cycle-related element) G-box,环境胁迫相关元件(Environmental stress-related element) ARE 和 MBS,以及响应脱落酸元件 ABRE、响应乙烯元件 ERE 等激素应答相关元件(Hormonal responses-related element),同时还有生长发育相关元件(Development related element) CGTCA-motif 和 TGACG-motif。分析毛竹 *SWEET* 家族的基因结构(见图 6b),各毛竹 *SWEET* 基因具内含子(Intron)数量 3~7 个不等。其中多数基因含 5 个内含子。

2.8　毛竹 *SWEET* 家族基因表达模式分析

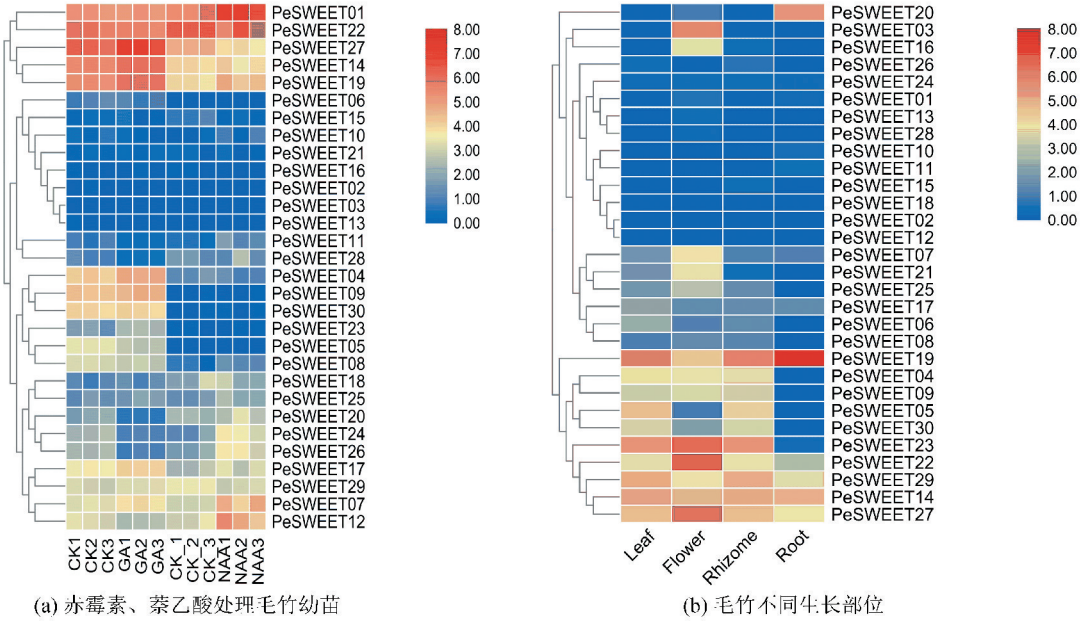
SWEET 基因家族对植物糖分的转运起着重要的作用,根据相关转录组数据,使用 TBtools 绘制表达热图(见图 7),以探究毛竹 *SWEET* 基因的潜在功能。研究毛竹 *SWEET* 在幼苗不同激素(赤霉素、萘乙酸)水平下以及毛竹不同部位(叶、花、鞭、根)的表达情况。图 7a 所示,对毛竹幼苗赤霉素处理后 *PeSWEET*20、*PeSWEET*24、*PeSWEET*26 表达量明显下调,*PeSWEET*23 表达量明显上调,其余家族基因表达量变化不明显。类似的,幼苗施加萘乙酸处理后 *PeSWEET*07、*PeSWEET*12 表达量明显上调,*PeSWEET*24、*PeSWEET*26 表达量明显下调。说明了不同基因对不同种类激素的响应不同,推测这些基因可能受激素调控。并且 *PeSWEET*02、*PeSWEET*03、*PeSWEET*13、*PeSWEET*16、*PeSWEET*21 对两种激素处理均无反应。由图 7b 可知,*PeSWEET*03、*PeSWEET*07、*PeSWEET*16、*PeSWEET*21 在毛竹的花中特异性表达,推测这些基因与毛竹的生殖活动有关。*PeSWEET*20 在根中特异性表达,此外 *PeSWEET*14、*PeSWEET*19、*PeSWEET*22、*PeSWEET*27、*PeSWEET*29 在毛竹的四个部位均有一定量的表达,其中 *PeSWEET*19 在根中具有极高的表达量,表达丰度值高达 232.8。



(a) 顺式作用元件 (b) 基因结构图

图 6 毛竹 *SWEET* 基因家族启动子区顺式作用元件及基因结构图

Fig.6 Structures analysis and cis-acting regulatory elements identified in the promoter region of *SWEET* gene family in *Phyllostachys edulis*



(a) 赤霉素、萘乙酸处理毛竹幼苗 (b) 毛竹不同生长部位

图 7 毛竹 *SWEET* 基因家族表达分析

Fig.7 Gene expression analysis of *SWEET* gene family in *Phyllostachys edulis*

注:CK 代表了 3 组对照组重复试验;GA1、GA2、GA3 代表了赤霉素的 3 组重复试验;NAA1、NAA2、NAA3 代表了萘乙酸的 3 组重复试验;右上方的颜色柱表示基因丰度;蓝色表示低表达,红色表示高表达。基因名称如表 1 所示.

Notes: CK represents repeated tests of three control groups; GA1, GA2, GA3 represent three groups of repeated tests of gibberellin; NAA1, NAA2, NAA3 represent three groups of repeated tests of naphthalene acetic acid. Color bar at the top right represents gene expression abundance value. where blue indicates low expression, and red indicates high expression. The gene names are shown in Table 1.

3 讨 论

SWEET 基因家族成员普遍存在于各种生物中,可涉及生物体生长发育、生殖、寄主与病原菌的互作、非生物应激反应等生理过程[2]。不同植物中的 *SWEET* 基因数量不同,本研究通过生物信息学的数据分析,从毛竹转录因子数据库中找到了 30 个候选的 *SWEET* 基因,而 Han 等(2015)统计的拟南芥、水稻以及番茄中的 *SWEET* 基因家族成员数量分别为 17、21 和 29 个[8],均少于毛竹。但毛竹中 *SWEET* 基因总数明显少于大豆中的 52 个[20],这可能与不同物种间基因组序列差异有关。本研究提取的 *SWEET* 基因家族与 Chen 等(2012)分类方法相一致[21],共分为 4 个亚族。结合 GO 功能注释分析,表明该家族所以成员均具有糖类转运的功能。种间共线性分析结果显示,毛竹的 *SWEET* 家族成员与同为单子叶植物的水稻之间存在着大量的共线性,有染色体水平的基因加倍现象,而拟南芥作为双子叶植物与毛竹的 *SWEET* 家族成员之间共线性不显著。表明了在系统进化的过程中,双子叶植物与单子叶植物的进化分支存在差异。每个家族成员都含有 *MtN3/saliva* 结构域,表明 *SWEET* 家族在进化中相对保守。

基因的表达大部分与上游启动子有关。经启动子分析显示,*SWEET* 基因启动子存在许多生长以及逆境响应相关顺式作用元件,预示其可能参与毛竹生长发育以及抗逆过程;*SWEET* 基因表达模式分析结果显示,*SWEET* 基因可响应激素调控,这与 *SWEET* 基因启动子上存在有相应激素调控元件有关;另外,某些基因在花中特异性高效表达,表明这些基因在毛竹生殖过程中起一定作用。综上说明 *SWEET* 蛋白除了具有的糖转运功能外,该家族基因还可能参与毛竹的生长发育,花的生殖发育以及胁迫反应过程。

通过生物信息学分析,确定了毛竹基因组中 *SWEET* 基因家族成员及染色体分布情况,获得了 *SWEET* 家族结构、功能、进化等特征。结合转录组数据分析,揭示了 *SWEET* 基因家族的表达特征,证实毛竹 *SWEET* 家族可响应激素调控并参与毛竹的生长发育与生殖发育过程。这些研究提供了毛竹 *SWEET* 的重要信息,为今后 *SWEET* 基因家族在毛竹快速高生长过程中的分子功能和作用机制研究奠定了坚实基础。

参考文献(References)

[1]陶贵耘,傅鹰,周明兵.竹类植物快速生长的机理研究进展[J].农业生物技术学报,2018,26(5):871-887. DOI:10.3969/j.issn.1674-7968.2018.05.015.
TAO Guiyun, FU Ying, ZHOU Mingbing. Advances in studies on molecular mechanisms of rapid growth of bamboo species[J]. Journal of Agricultural Biotechnology, 2018, 26(5):871-887. DOI:10.3969/j.issn.1674-7968.2018.05.015.

[2]YUAN Meng, WANG Shiping. Rice MtN3 / Saliva/ SWEET family genes and their homologs in cellular organisms[J]. Journal of Plant Physiology and Molecular Biology, 2013,(3):665-674. DOI:10.1093/mp/sst035.

[3]HAMADA M, WADA S, KOBAYASHI K, et al. Ci-Rga, a gene encoding an MtN3/saliva family transmembrane protein, is essential for tissue differentiation during embryogenesis of the ascidian Ciona intestinalis [J]. Differentiation, 2005, 73(7):364-376. DOI:10.1111/j.1432-0436.2005.00037.x.

[4]SAIER M H, TRAN C V, BARABOTE R D. TCDB: The Transporter Classification Database for membrane transport protein analyses and information [J]. Nucleic Acids Research, 2006, 34(Database issue):D181-D186. DOI:10.1093/nar/gkj001.

[5]GAMAS P, NIEBEL F D C, LESCURE N, et al. Use of a subtractive hybridization approach to identify new Medicago truncatula genes induced during root nodule development[J]. Molecular Plant-Microbe Interactions: MPMI, 1996, 9(4):233-242. DOI: 10.1094/mpmi-9-0233.

[6]何佳,黄学勇,关跃峰, 等. 拟南芥 MtN3/saliva 基因家族分析[J]. 上海师范大学学报(自然科学版),2010,39(3):290-298. DOI:10.3969/j.issn.1000-5137.2010.03.013.
HE Jia, HUANG Xueyong, GUAN Yuefeng, et al. Genome-wide analysis of MtN3/saliva protein family in Arabidopsis thaliana[J]. Journal of Shanghai Normal University(Natural Sciences), 2010, 39(3):290-298. DOI:10.3969/j.issn.1000-5137.2010.03.013.

[7]王梓然,匡柳青,陈尚武, 等. 在葡萄果实发育Ⅰ、Ⅲ期差异表达的 SWEET 基因家族成员的生物信息学分析[J]. 中国农业大学学报,2016,21(11):24-33. DOI:10.11841/j.issn.1007-4333.2016.11.04.
WANG Ziran, KUANG Liuqing, CHEN Shangwu, et al. Bioinformatics analysis of differentially expressed SWEET genes between grape berry development stage Ⅰ and Ⅲ[J]. Journal of China Agricultural University, 2016,21(11):24-33. DOI:10.11841/j.issn.1007-4333.2016.11.04.

[8]韩佳轩,姜晶.拟南芥、水稻和番茄 SWEET/MtN3/saliva 基因家族的分析[J]. 分子植物育种,2015,13(3):581-588. DOI:10.13271/j.mpb.013.000581.
HAN Jiaxuan, JIANG Jing. Genome-wide analysis of *SWEET* gene family in *Arabidopsis thaliana*, *Oryza sativa* and *Lycopersicum esculentum* [J]. Molecular Plant Breeding,

2015,13(3):581-588. DOI:10.13271/j.mpb.013.000581.

[9]薛蓓蓓, 覃丽芳, 董明右, 等. 木薯 SWEETs 基因家族生物信息学及表达特性研究[J]. 基因组学与应用生物学, 2019,38(1):260-268. DOI:10.13417/j.gab.038.000260.
XUE Beibei, QIN Lifang, DONG Mingyou, et al. Bioinformatics analysis and expressional characteristics of cassava SWEETs gene family[J]. Genomics and Applied Biology, 2019,38(1):260-268. DOI:10.13417/j.gab.038.000260.

[10]GUAN Yuefeng, HUANG Xueyong, ZHU Jungao, et al. RUPTURED POLLEN GRAIN1, a member of the MtN3 / saliva gene family, is crucial for exine pattern formation and cell integrity of microspores in Arabidopsis[J]. Plant Physiology, 2008 ,147 (2) : 852 - 863 . DOI:10.1104/pp.108.118026.

[11]WANG Lei, XIE Weibo, CHEN Ying, et al. A dynamic gene expression atlas covering the entire life cycle of rice [J]. The Plant Journal,2010,61(5):752-766. DOI:10.1111/j.1365-313X.2009.04100.x.

[12]SEO P J, PARK J M, KANG S K, et al. An Arabidopsis senescence-associated protein SAG29 regulates cell viability under high salinity[J]. Planta:An International Journal of Plant Biology, 2011, 233(1): 189-200. DOI:10.1007/s00425-010-1293-8.

[13]ZHOU Y, LIU L, HUANG W, et al. Overexpression of OsSWEET5 in rice causes growth retardation and precocious senescence [J]. PLoS One, 2014, 9 (4): e94210. DOI:10.1371/journal.pone.0094210.

[14]CHU Zhaohui, YUAN Meng, YAO Jialing, et al. Promoter mutations of an essential gene for pollen development result in disease resistance in rice[J]. Genes & Development, 2006, 20 (10) : 1250 - 1255. DOI: 10. 1101/gad.1416306.

[15]FINN R D, PENELOPE C, EBERHARDT R Y, et al. The Pfam protein families database: Towards a more sustainable future[J]. Nucleic Acids Research,2016,44(D1):279-285. DOI:10.1093/nar/gkv1344.

[16]WANG Yupeng, TANG Haibao, DEBARRY J D, et al. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity [J]. Nucleic Acids Research,2012,40(7):e49. DOI:10.1093/nar/gkr1293.

[17]BAILEY T L, MIKAEL B, BUSKE F A, et al. MEME Suite: Tools for motif discovery and searching[J]. Nucleic Acids Research, 2009, 37 (Web Server issue): W202-W208. DOI:10.1093/nar/gkp335.

[18]MAGALI L, PATRICE D, GERT T, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences[J]. Nucleic Acids Research,2002,30(1):325-327. DOI:10.1093/nar/30.1.325.

[19]玄元虎,朱毅勇,胡一兵. SWEET 蛋白家族研究进展[J]. 中国科学(生命科学),2014,44(7):676-684. DOI:10.1360/N052013-00068.
XUAN Yuanhu, ZHU Yiyong, HU Yibing. Research advances of the *SWEET* proteins family[J]. Science in China(Series C),2014,44(7):676-684. DOI:10.1360/N052013-00068.

[20]PATIL G, VALLIYODAN B, DESHMUKH R, et al. Soybean (*Glycine max*) *SWEET* gene family: Insights through comparative genomics, transcriptome profiling and whole genome re-sequence analysis[J]. BMC Genomics, 2015,16(1):520. DOI:10.1186/s12864-015-1730-y.

[21]CHEN Liqing, QU Xiaoqing, HOU Bihuei, et al. Sucrose efflux mediated by *SWEET* proteins as a key step for phloem transport[J]. Science,2012,335(6065):207-211. DOI: 10.1126/science.1213351.

[责任编辑:吴永英]