



研究报告

Research Report

蒺藜苜蓿 PEPC 蛋白生物信息学分析及在碱胁迫中功能的预测

任永晶 , 宋婷婷 , 姜柳 , 许慧慧 , 才华 

东北农业大学生命科学学院, 哈尔滨, 150030

 通讯作者, caihuaneau@sohu.com;  作者

豆科基因组学与遗传学, 2016年, 第7卷, 第6篇 doi: 10.5376/lgg.cn.2016.07.0006

收稿日期: 2016年08月22日 接受日期: 2016年08月22日 发表日期: 2016年08月22日






© 2016 BioPublisher 生命科学中文期刊出版平台

本文首次发表在《分子植物育种》2016年, 第14卷, 第2期上。现依据版权所有人授权的许可协议, 采用 [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) 协议对其进行授权, 再次发表与传播。只要对原作有恰当的引用, 版权所有人允许并同意第三方无条件的使用与传播。



摘要 磷酸烯醇式丙酮酸羧化酶(*phosphoenolpyruvate carboxylase, PEPC*)是一种广泛存在于自然界中的代谢酶, 在高等植物中参与光合作用的碳固定、生物合成前体的供应、pH的调控等多种生物学过程。本研究通过生物信息学方法, 识别、筛选、获得蒺藜苜蓿(*Medicago truncatula*)PEPC(*MtPEPC*)的氨基酸序列 Medtr2g076670.2, 并对与 Medtr2g076670.2 相似的 25 个豆科 PEPC 基因进行系统进化树分析。着重对 4 个 *MtPEPCs* 和 6 个 *GmPEPCs* 基因进行功能域分析、蛋白二级结构域预测及组织表达特异性分析, 以推测蒺藜苜蓿 PEPC 基因的功能。结果表明, 该 10 个基因分为两个 Group 分支, 两个分支中存在不同的亚细胞定位信号, 并且两个分支中大豆基因在地上和地下组织之间存在差异。通过蛋白相互作用的预测分析, PEPC 蛋白可与苹果酸脱氢酶(Malate dehydrogenase, MDH)、丙酮酸激酶(Pyruvate kinase, PK)及 2 个未知蛋白存在相互作用; 在碱胁迫下, *GsPEPCs* 基因与苹果酸脱氢酶基因是“共表达”基因。可以推测, PEPC 对碱胁迫的反应, 可能是通过调节有机酸含量实现的。通过对 *MtPEPC* 蛋白和 *GmPEPC* 蛋白的生物信息学分析, 获得了其相应的分子生物学特征, 为 PEPC 蛋白在碱胁迫反应中的功能提供理论依据。

关键词 蒺藜苜蓿; 大豆; PEPC; 生物信息学; 碱胁迫

Bioinformatics Analysis of PEPC Proteins from *Medicago truncatula* and Functional Predictions under Alkaline Stress

Ren Yongjing , Song Tingting , Jiang Liu , Xu Huihui , Cai Hua 

College of Life Sciences, Northeast Agricultural University, Harbin, 150030

 Corresponding author, caihuaneau@sohu.com;  Authors

Abstract Phosphoenolpyruvate carboxylase (*PEPC*) is a widespread metabolic enzyme in nature, and it participates in carbon fixation of photosynthesis in higher plants, supply of biosynthetic precursors, regulates of pH. This study had identified, screened and obtained the protein sequence Medtr2g076670.2 of *PEPC* gene in *Medicago truncatula* through bioinformatics methods, phylogenetic tree was constructed with 25 homologous genes with Medtr2g076670.2. Functional domains, secondary structure and tissue transcript-leveled expression of 4 *MtPEPCs* and 6 *GmPEPCs* were analysed to predict the biological function of *PEPCs* in *Medicago truncatula*. These results showed that the 10 genes were divided into two Group, and there were different subcellular localization signals in two Groups, and also there were different expression characteristics in ground and underground tissues between the two Groups of soybean genes. Through predicting the interaction network proteins, we found the *PEPC* protein can interact with malate dehydrogenase (MDH), pyruvate kinase(PK) and 2 unknown functional proteins. In addition, under alkali stress, *GsPEPCs* and malic acid dehydrogenase gene were co-expression. It could be speculated that the reaction of PEPC responding to alkali stress might be achieved by regulating the organic acid content. The molecular biological characteristics of *MtPEPC* proteins and *GmPEPC* proteins could be obtained by bioinformatics analysis. This study would provide a theoretical basis for identifying the functional of PEPC protein under the alkali stress response.

Keywords *Medicago truncatula*, Soybean (*Glycine max* (L) Merr.), *PEPC*, Bioinformatics, Alkali stress

蒺藜苜蓿(*Medicago truncatula*)是一种一年生的二倍体苜蓿,作为一种优良的牧草在全世界范围内被广泛种植。蒺藜苜蓿具有基因组小、生长期短、自花授粉、根瘤固氮、遗传转化效率高等特点,已成为豆科生物学以及逆境响应研究的模式植物(李紫薇等, 2014; 陈爱民等, 2006; 宋辉等, 2014; 姜格格等, 2013)。大豆与蒺藜苜蓿同属于豆科植物,在遗传上有较高的相似性,大豆基因组及转录组方面的研究可为蒺藜苜蓿、紫花苜蓿、豌豆、三叶草等豆科植物的研究提供更多的参考(刘志敏等, 2015)。磷酸烯醇式丙酮酸羧化酶(*phosphoenolpyruvate carboxylase*, *PEPC*)是一种广泛存在于自然界中的代谢酶,分布在古菌,细菌,藻类和维管植物中。在高等植物中参与光合作用的碳固定、生物合成前体的供应、对 pH 的调控及豆科作物根瘤固氮过程信号级联反应(Xu et al., 2007)。有研究报道,在植物面对干旱(周宝元等, 2011; 方立锋, 2008)、强光(李霞等, 2005)、盐(Echevarría et al., 2001)、低磷胁迫(Wu et al., 2003)和冷害(Jane, 2011)等逆境情况下, *PEPC* 蛋白的活性有所提高。由此表明, *PEPC* 在植物非生物胁迫反应中也发挥一定的作用。然而, *PEPC* 在碱胁迫反应中的作用尚无明确的报道。本研究以 *MtPEPC* 蛋白序列 Medtr2g076670.2 为目的基因,借助蒺藜苜蓿基因组数据库、大豆基因组数据库及公共数据库(NCBI, GEO, Phytozome),对蒺藜苜蓿 *PEPC* 蛋白及与其同源的大豆 *PEPC* 蛋白 Glyma.12G229400.1 进行生物信息学分析,根据大豆 *PEPC* 蛋白的结构域、组织表达差异、碱胁迫表达谱及蛋白质相互作用预测蒺藜苜蓿 *PEPC* 蛋白在碱胁迫反应中的生物学功能,为蒺藜苜蓿该类基因的功能研究提供参考。

1 结果与分析

1.1 序列获取与功能域分析

以蒺藜苜蓿基因 Medtr2g076670.2 为种子序列,共获得具有相似保守结构域的首蓿 *PEPC* 蛋白 4 个,大豆 *PEPC* 蛋白 6 个(表 1)。首蓿 *PEPC* 基因分别定位于第 2、8、4 号染色体上,大豆 *PEPC* 基因分别定位于第 6、12、13 号染色体上。

1.2 结构功能域分析与同源序列比对

蒺藜苜蓿 Medtr2g076670.2 在 161~966 处为由 806 个氨基酸组成的 *PEPcase* 结构域(图 1);与之最

相近的大豆蛋白为 Glyma.12G229400.1 (*GmPEPC1*),相似性为 94.30%,也具有完整的 *PEPcase* 结构域,由此推测这两个基因可能具有相似的生物学功能。

1.3 蛋白质序列系统进化树分析

根据 Medtr2g076670.2 的氨基酸序列,在 Phytozome 数据库中检索 6 个豆科物种中其他的同源蛋白,经过筛选得到含有完整 *PEPC* 的结构域的氨基酸序列 25 个,利用 MEGA5.0 软件并采用 NJ 法构建系统进化树(图 2)。25 个 *PEPC* 蛋白一共被分成了 3 个分支, Group I 分支中,蒺藜苜蓿蛋白 Medtr2g076670.2、Medtr8g463920.2 (“■”所示)与大豆 Glyma.12G229400.1 (“●”所示)相似性最高, Glyma.12G161300.1、Glyma.06G229900.1 也属于此分支;在 Group II 中, Medtr2g092930.1 与大豆中的 Glyma.12G210600.1、Glyma.13G290700.1 两个蛋白相似性最高,而 Medtr4g079860.1 则与大豆的 Glyma.06G277500.1 具有更高的相似性,并与四季豆的 *PEPC* 蛋白聚为一组;在 Group III 中,黄瓜、草莓和苹果中的几个蛋白聚为一组。此 6 个大豆 *PEPC* 基因在非生物胁迫反应中的作用均未见报道,并且均不是根瘤增强型的基因(Hata et al., 1998)。但值得注意的是,已经报道盐胁迫响应的野生大豆 *GsPEPC* 和 *GsPEPC1* (Qi et al, 2014)与大豆 Glyma.12G161300.1、Glyma.06G229900.1 具有 98%~99%的相似性。由此暗示, *PEPC* 蛋白在参与基本的生物代谢外,还通过其它的生物学过程响应非生物胁迫。

1.4 *PEPC* 蛋白二级结构分析

信号肽预测显示,蒺藜苜蓿和大豆 *PEPC* 蛋白都没有信号肽序列和跨膜区;均含有 Ser/Thr 和 Tyr 蛋白激酶磷酸化活性位点;并且蛋白主要由 α -螺旋、不规则卷曲和延伸链组成。然而亚细胞定位预测发现, Group I 中与 Group II 中的蛋白亚细胞定位信号存在差异,蒺藜苜蓿和大豆表现同样的特性。Group I 中蒺藜苜蓿和大豆 *PEPC* 蛋白主要定位于核、微体、叶绿体类囊体膜、叶绿体基质、线粒体(仅 Medtr2g076670.2 无线粒体定位信号)等几个组织,而 Group II 中的 5 个蛋白却没有叶绿体类囊体基质的定位信号。亚细胞定位信号的差异暗示,同是 *PEPC* 类蛋白,在不同的组织部位和亚细胞结构

表1 蒺藜苜蓿和大豆 PEPC 蛋白的获取与定位

Table 1 Obtaining and location of PEPC proteins in *Medicago truncatula* and soybean

转录子名称	基因长度(bp)	染色体定位	蛋白质长度(aa)	同源性(%)	S 值
Name of transcript	Length of gene (bp)	Chromosome location	Length of protein (aa)	Homology (%)	S value
Medtr2g076670.2	3 570	chr2:32093734..32099214	966	100.00	1 969.5
Medtr8g463920.2	3 391	chr8:22564514..22569401	967	95.4	1 818.5
Medtr2g092930.1	3 358	chr2:39535830..39541285	966	94.1	1 696.4
Medtr4g079860.1	3 129	chr4:30905150..30912077	966	92.3	1 693.7
Glyma.12G229400.1 (GmPEPC1)	3 446	Chr12:38938314..38944919	967	94.3	1 826.2
Glyma.12G161300.1 (GmPEPC1)	3 540	Chr12:29898332..29906466	967	95.6	1 815.1
Glyma.06G229900.1 (GmPEPC16)	3 289	Chr06:35345701..35354129	967	95.6	1 813.1
Glyma.12G210600.1 (GmPEPC4)	3 443	Chr12:36952133..36958451	966	93.4	1 718.8
Glyma.06G277500.1 (GmPEPC-L)	3 675	Chr06:46935641..46944254	966	93.6	1 710.3
Glyma.13G290700.1 (GmPEPC4)	2 937	Chr13:39106079..39112900	978	91.6	1 688.7

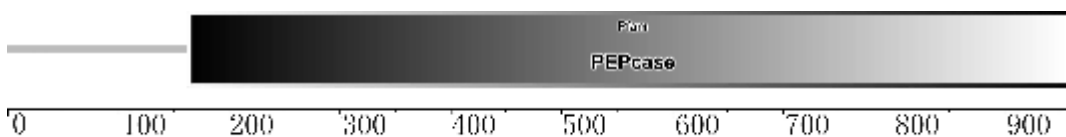


图1 蒺藜苜蓿和大豆 PEPC 结构域

Figure 1 Domain of PEPC in *Medicago truncatula* and soybean

中, 参与不同的功能; 并且, 二级结构的差异决定了 PEPC 蛋白具有多种生物学功能。

1.5 PEPC 蛋白组织特异性分析

由于蒺藜苜蓿数据库中没有基因表达谱数据, 4 个 *MtPEPCs* 基因序列及功能域与 6 个 *GmPEPCs* 基因具有较高的相似性, 借助 *GmPEPCs* 基因组织特异性分析数据推测 *MtPEPCs* 基因的功能。根据进化树的分支, 分别获得 GroupI 分支的 3 个 PEPC 蛋白 (Glyma.12G229400.1, Glyma.12G161300.1, Glyma.06G229900.1) 和 GroupII 分支的 3 个 PEPC 蛋

白 (Glyma.12G210600.1, Glyma.13G290700.1, Glyma.06G277500.1) 根、根毛、叶片、根瘤、花等 9 个组织表达特性图谱(图 3)。由此可见, GroupI 分支的 3 个 *GmPEPCs* 基因在地下组织(根毛、根中及根瘤)中表达量高于叶片, 茎等地上组织; 相反 GroupII 分支的 *GmPEPCs* 基因主要在天上组织中表达。基因组织表达的特性与定位信号的分析、系统进化树分析具有类似的结果。*PEPCs* 基因的生物学功能在地上部位和地下组织中存在不同的功能分工。

1.6 蛋白质相互作用分析

同样借助大豆 PEPCs 蛋白的相互作用, 辅助了

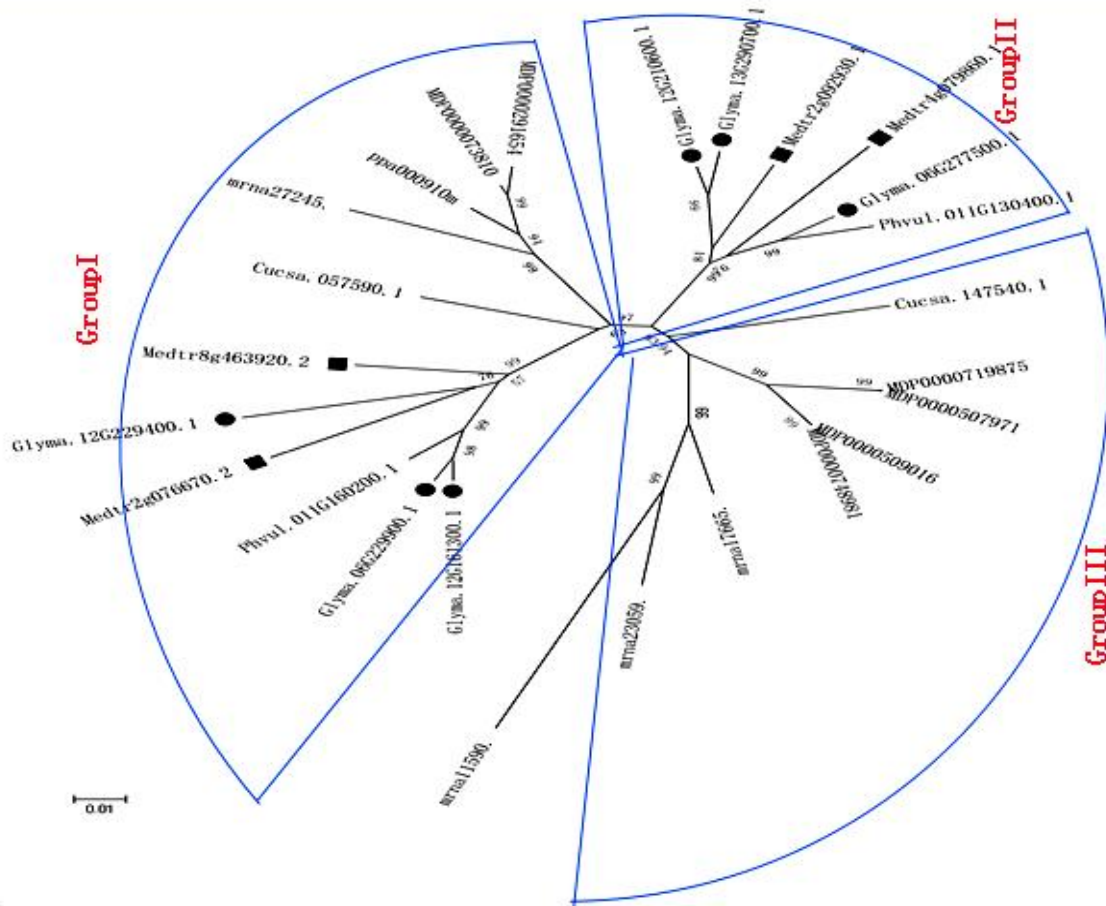


图 2 Medtr2g076670.2 基因系统进化树

Figure 2 Phylogenetic tree of Medtr2g076670.2 gene

解 MtPEPCs 蛋白的功能和调控机制。利用 STRING 交互式数据库搜索与 Glyma.12G229400.1 蛋白相互作用的蛋白, 构建蛋白相互作用网络(图 4)。与 *GmPEPC1*(Glyma.12G229400.1, Glyma.12G35840.1)互作的 10 个蛋白中(表 3), 有 5 个是苹果酸脱氢酶(Malate dehydrogenase, MDH)。MDH 可以催化苹果酸与草酰乙酸间的可逆转换, 主要参与 TCA 循环、光合作用、C4 循环等代谢途径(汪新颖等, 2009)。其中 Glyma.11G04720.1 具有乙醛酸循环体活性, 说明 *GmPEPC* 蛋白与乙醛酸循环代谢也存在关系。Glyma.20G02980.1、Glyma.07G35110.3 这两个蛋白都有丙酮酸激酶活性, 丙酮酸激酶参与糖酵解途径, 是糖无氧酵解途径的三个关键调节点之一。丙酮酸激酶在第三阶段的最后一步中催化磷酸烯醇式丙酮酸(PEP)转变为丙酮酸, 产生能量 ATP。在机体缺氧时糖酵解途径能迅速提供能量, 为细胞的生命活动提供部分能

量, 其生成的一些中间产物是脂类和氨基酸等合成的前体(欧艳, 2010)。根据与 *GmPEPC1* 蛋白互作的蛋白类型可以看出 *PEPC* 参与 TCA 循环、光合作用、C4 循环、乙醛酸循环、糖酵解等多种代谢途径。

1.7 碱胁迫基因表达分析

从 NCBI 的 GEO 中下载了 NaHCO₃ 胁迫下大豆根(GSE17883)和叶(GSE20323)中的基因表达数据, 分析了 4 个大豆 *PEPC* 基因及与之相互作用的 9 个蛋白在碱胁迫中的表达情况(图 5), 推测 MtPEPCs 基因在碱胁迫下可能的生物学功能。在 NaHCO₃ 胁迫下同一个基因在地上和地下组织中表达量及表达模式并不相同。Glyma.12G229400.1 在根中的表达量明显高于在叶中的表达量, 与之互作的蛋白 Glyma17G10880.1 在根中的表达量与在叶中的表达量也明显不同。除 Glyma.12G229400.1 基因外, Glyma06G277500.1 在根和叶中的表达量都比较低, 并且在碱胁迫下差异变化较少; Glyma12G161300.1 在根和叶中的表达量均较低,

表2 蒺藜苜蓿和大豆同源蛋白一、二级结构预测信息

Table 2 Predicted information of the primary and secondary structure of homologous protein in *Medicago truncatula* and soybean

转录子名称	信号肽和跨膜区 Transcript name	磷酸化位点预测 Signal peptide and transmembrane segment	二级结构预测 Phosphorylation predictions	亚细胞定位 Secondary structure prediction	亚细胞定位 Subcellular localization
Medtr2g076670.2	无 No	Ser: 2; Thr: 15; Tyr: 8	α -螺旋、不规则盘绕和延伸链 Alpha helix, random coil, extended strand	核、微体、叶绿体类囊体膜、 叶绿体基质 Nucleus, microbody (peroxisome), mitochondrial matrix space, lysosome (lumen)	
Medtr8g463920.2	无 No	Ser: 23; Thr: 13; Tyr: 9	α -螺旋、不规则盘绕和延伸链 Alpha helix, random coil, extended strand	核、线粒体基质、 微体、叶绿体类囊体膜 Nucleus, mitochondrial matrix space, microbody (peroxisome), mitochondrial inner membrane	
Medtr2g092930.1	无 No	Ser: 28; Thr: 14; Tyr: 9	α -螺旋、不规则盘绕和延伸链 Alpha helix, random coil, extended strand	线粒体基质、微体、 线粒体膜、线粒体膜间隙 Mitochondrial matrix space, microbody (peroxisome), mitochondrial inner membrane, mitochondrial intermembrane space	
Medtr4g079860.1	无 No	Ser: 27; Thr: 17; Tyr: 6	α -螺旋、不规则盘绕和延伸链 Alpha helix, random coil, extended strand	线粒体基质、微体、 线粒体膜、线粒体膜间隙 Mitochondrial matrix space, microbody (peroxisome) mitochondrial matrix space, lysosome (lumen)	
Glyma.12G229400.1	无 No	Ser: 24; Thr: 14; Tyr: 9	α -螺旋、不规则盘绕和延伸链 Alpha helix, random coil, extended strand	核、线粒体基质、 微体、叶绿体类囊体膜 Nucleus, mitochondrial matrix space, microbody (peroxisome), mitochondrial inner membrane	
Glyma.12G161300.1	无 No	Ser: 24; Thr: 14; Tyr: 9	α -螺旋、不规则盘绕和延伸链 Alpha helix, random coil, extended strand	线粒体基质、微体、核、叶绿体类囊体膜 Mitochondrial matrix space, microbody (peroxisome), nucleus, mitochondrial inner membrane	
Glyma.06G229900.1	无 No	Ser: 26; Thr: 15; Tyr: 10	α -螺旋、不规则盘绕和延伸链 Alpha helix, random coil, extended strand	线粒体基质、微体、核、叶绿体类囊体膜 Mitochondrial matrix space, microbody (peroxisome), nucleus, mitochondrial	

inner membrane

Glyma.12G210600.1	无	Ser: 26; Thr: 14;	α -螺旋、不规则盘绕和延伸链	线粒体基质、微体、线粒体膜、线粒
	No	Tyr: 7	Alpha helix、random coil、 extended strand	体膜间隙 Mitochondrial matrix space、microbody (peroxisome)mitochondrial inner membrane、mitochondrial intermembrane space
Glyma.06G277500.1	无	Ser: 28; Thr: 13;	α -螺旋、不规则盘绕和延伸链	线粒体基质、微体、核、线粒体膜
	No	Tyr: 9	Alpha helix、random coil、 extended strand	Mitochondrial matrix space、microbody (peroxisome)mitochondrial inner membrane
Glyma.13G290700.1	无	Ser: 28; Thr: 13;	α -螺旋、不规则盘绕和延伸链	线粒体基质、微体、线粒体膜、线粒
	No	Tyr: 9	Alpha helix、random coil、 extended strand	体膜间隙 Mitochondrial matrix space、microbody (peroxisome)mitochondrial inner membrane、mitochondrial intermembrane space

为下调表达; Glyma12G210600.1 在根和叶中的表达量均较高,但是在根中,该基因为上调表达,而在叶中则为下调表达。另外,根据基因表达聚类发现,在根和叶中, Glyma.10G34150.1 和 Glyma.08G06820.1 基因表达模式与 Glyma.12G229400.1 相似,可确定该 3 个基因为“共表达”基因。该结果与相互作用网络构建的结果一致,说明 *GmPEPC1* 蛋白与该 2 个蛋白存在相互关系, Glyma.10G34150.1 为未知蛋白,而 Glyma.08G06820.1 注释为苹果酸脱氢酶。由此推测, *PEPC* 蛋白参与苹果酸脱氢酶的代谢途径,并通过该途径代谢的变化响应碱胁迫。

分别对大豆 *PEPC* 蛋白根和叶中的基因表达模式进行聚类分析(图 6),可以发现在叶中,3 h 时基因的表达量达到了最大值,而在根中基因表达变化差异最大的时间点为 6 h。但是也不难看出,根是最先响应碱胁迫的,胁迫 1 h 后,基因上调表达达到最高值。

2 讨论

据报道, *PEPC* 基因有不同的类型。在高等植物中 *PEPC* 有 2 个已知的功能形式,光合型的 *PEPC* 在 C4 光合作用和景天酸代谢(CAM)途径中催化大

气 CO₂ 初始固定形成 C4 二羧酸(Deng, et al., 2015); 非光型的 *PEPC* 已在 C3 植物的叶子和各种器官中发现,如:根发育中的种子和豆科植物根瘤中等。在根瘤中, *PEPC* 发挥共生固氮的作用,在碳氮代谢中起到重要的作用(Naima et al., 1997; Nomura et al., 2006)。本研究发现与 Glyma.12G229400.1 相似的 25 个 *PEPC* 蛋白一共被分成了 3 个分支,其中 Group I 和 Group II 2 个分支中的 *PEPC* 蛋白在亚细胞定位信号存在着差异。除此之外, Group I 的 3 个 *GmPEPCs* 基因在地下组织(根毛,根中及根瘤)中表达量高于叶片,茎等地上组织;而 Group II 分支的 *GmPEPCs* 基因主要在天上组织中表达。根据碱胁迫表达谱数据分析的结果,根最先响应碱胁迫,胁迫 1 h 后,基因上调表达达到最高值。综上所述, *PEPCs* 基因对非生物胁迫的反应除了要考虑其不同类型外,还应考虑地上和地下 2 个不同的组织区域,地下部分首先感受胁迫信号,其势必与地上组织存在不同的功能分工。

Hata S(1998)等发现, *GmPEPC7* 基因与 *GmPEPC1/15* 及 *GmPEPC4* 基因在编码区具有 98%~99%的相似性,但是 3'-UTR 的相似性仅有 37%。而正是由于其非编码区的差异,导致

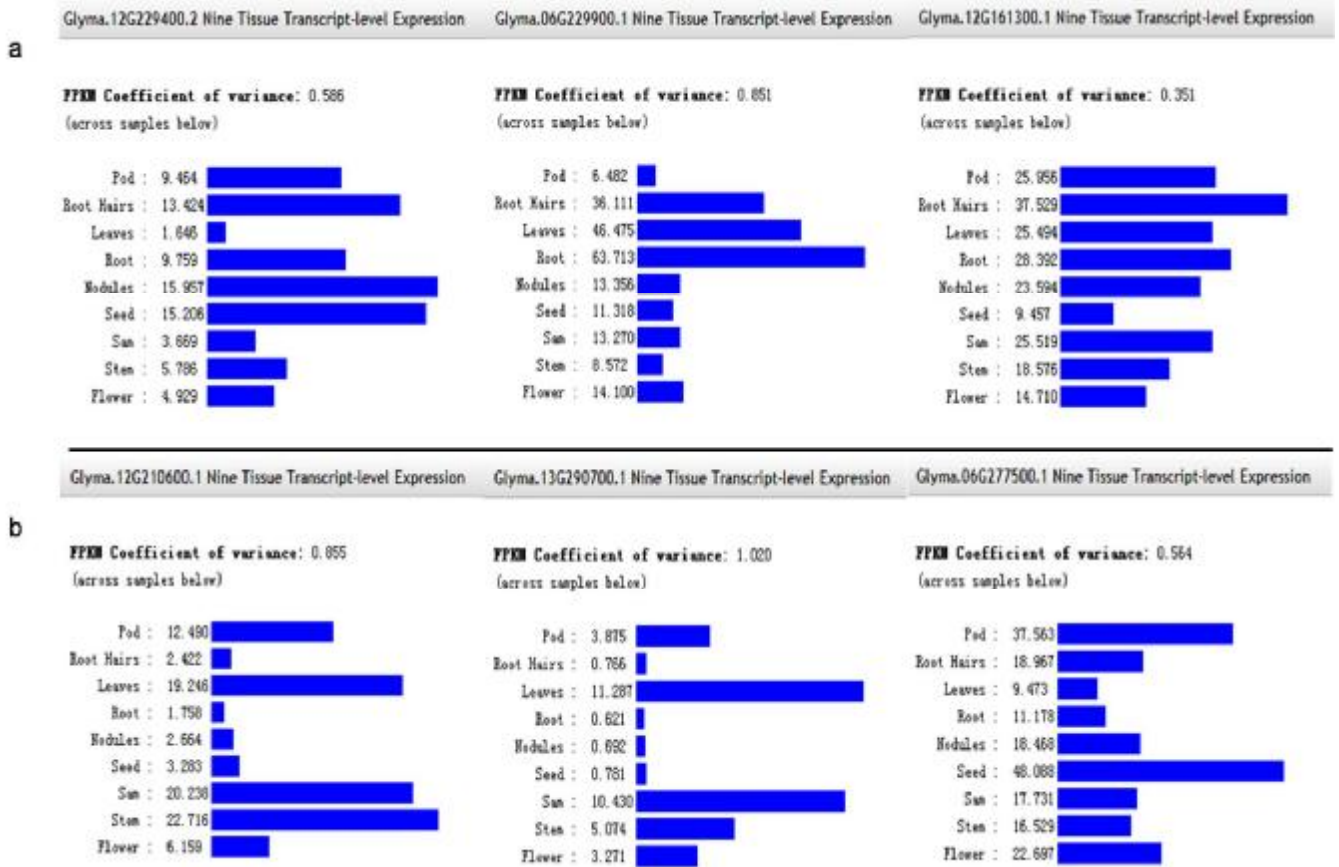


图 3 6 个大豆 PEPC 基因 9 组织的表达特性

注: a: Group I 分支中 3 个 *GmPEPCs* 基因组织表达分析 (Glyma.12G229400.1, Glyma.12G161300.1, Glyma.06G229900.1); b: Group I 分支中 3 个 *GmPEPCs* 基因组织表达分析(Glyma.12G210600.1, Glyma.13G290700.1, Glyma.06G277500.1)

Figure 3 6 *GmPEPCs* transcript-level expressions in nine tissue

Note: a: Expression analysis of 3 *GmPEPCs* from Group I transcript-level in different tissues (Glyma.12G229400.1, Glyma.12G161300.1, Glyma.06G229900.1); b: Expression analysis of 3 *GmPEPCs* from Group II transcript-level in different tissues (Glyma.12G210600.1, Glyma.13G290700.1, Glyma.06G277500.1)

GmPEPC7 基因是根瘤增强型的 PEPC, 另两个基因则为持家基因。本文中所述的 6 个大豆 PEPC 基因均不是根瘤增强型的基因, 但是不可忽视其非编码区及其启动子的差异。蒺藜苜蓿中 4 个 PEPC 基因在氨基酸序列上与大豆 6 个 PEPC 基因具有 91%~98% 的相似性, 蒺藜苜蓿中 PEPC 基因是否具有与其相似的大豆 PEPC 基因的功能, 还不能单纯从其序列的相似性推测, 还需要考虑其非编码区和启动子顺式作用元件的差异。

根据 pH 稳态调节假说, 高浓度 HCO_3^- 能激活 PEPC, 具有最佳活性, 其结果是, 产生苹果酸防止大量胞浆碱化。如果苹果酸的产生使胞浆过度酸化那么 PEP 的活性被抑制, 而 NADP-苹果酸酶

(NADP-ME)被激活, 增加苹果酸脱羧酶酶活以防止胞质酸化(Sakano et al., 1998)。蛋白互作分析显示, 与 PEPC 蛋白相互作用的 9 个蛋白中, 有 5 个蛋白为苹果酸脱氢酶; 另外, 根据基因表达聚类发现, 该 2 个基因为“共表达”基因, 由此可以推定, PEPC 可能是通过调控苹果酸脱氢酶的活性响应植物耐碱。另外, 在根中, 根瘤的发育也伴随着苹果酸的产生和分泌, 这个过程的调控也与 PEPC 蛋白相关。

3 材料与方法

3.1 序列获取与定位

利用 NCBI 网站 (<http://www.ncbi.nlm.nih.gov/genbank/>), 搜索根据序

列的相似性搜索 蒺藜苜蓿 *MtPEPC* (Medtr2g076670.2) 的同源蛋白。在 Phytozome(<http://www.phytozome.net/>) 中搜索与 Medtr2g076670.2 蛋白具有相同结构域的大豆同源蛋白。

3.2 结构功能域分析及同源基因系统进化树构建

运用在线分析工具 SMART(<http://smart.embl-heidelberg.de/>)进行结构功

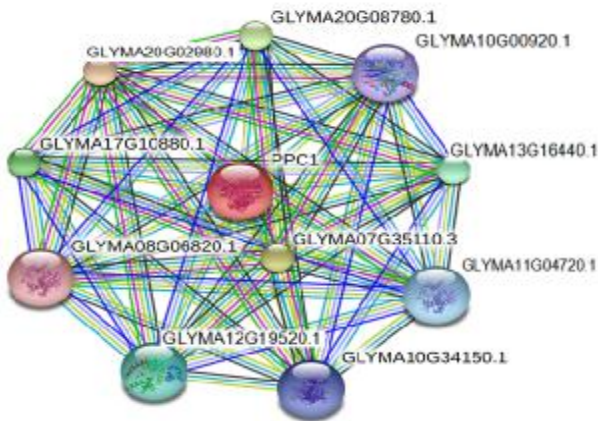


图4 Glyma.12G229400.1 蛋白与其他蛋白的相互作用网络
Figure 4 The interacted network of Glyma.12G229400.1 protein with other proteins

能域分析, 删除不含有完整 *PEPC* 保守结构域的蛋白序列, 只保留存在完整 *PEPC* 保守结构域的并与 Medtr2g076670.2 具有 80% 以上相似性的蛋白。采用邻近相连算法(Neighbor-Joining), 利用 MEGA5.0 软件构建系统进化树。

3.4 蒺藜苜蓿及大豆 *PEPC* 蛋白的二级结构预测

蛋白质序列信号肽及跨膜结构域预测分别采用 CBS (<http://www.cbs.dtu.dk/services/SignalP/>) 服务器中蛋白预测工具 signal peptide 和 TMHMM Serverv.2.0。亲疏水性分析 (<http://www.cbs.dtu.dk/services/NetPhos/>)、磷酸化位点分析(<http://www.expasy.ch/egi-bin/protscal.pl>)及二级结构的预测 (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html)均利用在线工具进行分析。

3.6 蛋白质相互作用分析

应用 STRING 交互式数据库 (<http://string-db.org/>) 搜索与 Glyma.12G229400.1(Glyma.12G35840.1, *PPC1*) 蛋

白相互作用的其他蛋白质信息, 搜索条件限制在 10 个蛋白内, 构建的蛋白相互作用网络。

3.7 *PEPC* 基因及其相互作用蛋白基因碱胁迫表达差异分析

在 NCBI 的 GEO(<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9730>) 数据库中下载大豆的根和叶在碱 (NaHCO_3) 胁迫下的表达谱数据, 利用 MeV4.9 软件进行表达模式和层次聚类分析。

作者贡献

才华指导实验设计, 进行论文写作与修改。任永晶是本研究实验工作的具体执行人, 完成数据分析和论文初稿的写作。宋婷婷、姜柳和和许慧慧完成基因表达谱数据、进化树分析及数据整理、文章校稿工作。全体作者都阅读并同意最终的文本。

致谢

本研究由国家自然科学基金(31302022)和黑龙江省自然科学基金(C2015018)共同资助。感谢哈尔滨工程大学王洋和鲍佩华博士对本文序列分析方法指导。

参考文献

- Chen A.M., Lian R.L., Sun J., and Wang Y.Z., 2006, Leguminous model plant-*Medicago truncatula*, Zhiwu Shenglixue Tongxun (Plant Physiology Communications), 42(5): 997-1003 (陈爱民, 连瑞丽, 孙杰, 王彦章, 2006, 豆科模式植物—蒺藜苜蓿, 植物生理学通讯, 42(5): 997-1003)
- Echevarría C., Garcia-Mauriño S., and Alvarez R., 2001, Salt stress increases the Ca^{2+} -independent phosphoenolpyruvate carboxylase kinase activity in *Sorghum* leaves, Planta, 214(2): 283-287
- Fang L.F., 2008, Physiological impact of *PEPC* gene overexpressed rice plants under drought stress, Thesis for M.S., Chinese Academy of Agricultural Sciences, Supervisor: Zhao M., pp.06-01 (方立锋, 2008, 转 *PEPC* 基因水稻干旱逆境下的生理反应, 硕士学位论文, 中国农业科学院, 导师: 赵明, pp.06-01)
- Jane A.L., 2011, C_4 Cycles: Past, Present, and Future Research on C_4 Photosynthesis, Plant Cell, 23(11): 3879-3892 Hata S., Izui K., and Kouchi H., 1998, Expression of a soybean nodule-enhanced phosphoenolpyruvate carboxylase gene that shows striking similarity to another gene for a house-keeping isoform, The Plant Journal, 13(2): 267-273

- Deng H., Zhang L.S., Zhang G.Q., Zheng B.Q., Liu Z.J., and Wang Y., 2015, Evolutionary history of PEPC genes in green plants: Implications for the evolution of CAM in orchids, *Molecular Phylogenetics and Evolution*, available online, doi:10.1016/j.ympev.2015.10.007
- Jiang G.G., Song L.L., Guo D.L., Cai H.S., Guo C.H., and Shu Y.J., 2013, Genome-wide association mapping of aluminum tolerance in *Medicago truncatula*, *Caoye Xuebao (Acta Prataculturae Sinica)*, 22(4):170-178 (姜格格, 宋丽莉, 郭东林, 蔡洪生, 郭长虹, 束永俊, 2013, 蒺藜苜蓿耐酸铝性状的全基因组关联分析, *草业学报*, 22(4): 170-178)
- Li X, Jiao D.M., and Dai C.C., 2005, The Response to Photooxidation in Leaves of PEPC Transgenic Rice Plant (*Oryza sativa* L.), *Zuowu Xuebao (Acta Agronomica Sinica)*, 31(4): 408-413 (李霞, 焦德茂, 戴传超, 2005, 转 PEPC 基因水稻对光氧化逆境的响应, *作物学报*, 31(4): 408-413)
- Li Z.W., Ma T., Liang G.T., and Guo J., 2014, Photosynthesis in Leaves of *Medicago truncatula* under Salt Stress, *Xibei Zhiwu Xuebao (Acta Botanica Boreali-occidentalia Sinica)*, 34(10): 2070-2077 (李紫薇, 马天, 梁国婷, 国静, 2014, 蒺藜苜蓿叶片光合作用对盐胁迫的响应, *西北植物学报*, 34(10): 2070-2077)
- Liu Z.M., Liu W.X., Jia X.T., Zhang Z.S., and Wang Y.R., 2015, Genome-wide analysis of LEA gene family in *Medicago truncatula*, *Caoye Kexue (Pratacultural Science)*, 32(3): 382-391 (刘志敏, 刘文献, 贾喜涛, 张正社, 王彦荣, 2015, 蒺藜苜蓿 LEA 基因家族全基因组分析, *草业科学*, 32(3): 382-391)
- Nomura M., Mai H.T., Fujii M., Hata S., Izui K., and Tajima S., 2006, Phosphoenolpyruvate carboxylase plays a crucial role in limiting nitrogen fixation in *Lotus japonicus* nodules, *Plant and Cell Physiol.*, 47(5): 613-621
- Ou Y., 2010, Study on the mechanism of the change of the conformation of pyruvate kinase in the activation of ion and substrate, Thesis for M.S., FuDan University, Supervisor: Yu S.N., pp.05-10 (欧艳, 2010, 丙酮酸激酶在活化离子和底物诱导下的构象变化机理研究, 硕士学位论文, 复旦大学, 导师: 余绍宁, pp.05-10)
- Sakano K., 1998, Revision of biochemical pH-stat: involvement of alternative pathway metabolisms, *Plant Cell Physiol.*, 39(5):467-473
- Song H, and Nan Z.B., 2014, Genome-wide identification and analysis of *WRKY* transcription factors in *Medicago truncatula*, *Yichuan (Hereditas (Beijing))*, 36(2): 152-168 (宋辉, 南志标, 2014, 蒺藜苜蓿全基因组中 *WRKY* 转录因子的鉴定与分析, *遗传*, 36(2): 152-168)
- Naima B., Jean L. P., Eliane D., Jean P. R., Martine A. D., Jean V., Pierre G and Raymond C., 1993, 1997, Regulatory phosphorylation of C4 Phosphoenolpyruvate carboxylase. *Plant Physiol.* 101:891-897
- Wang X.Y., Wang B., Hou S.T., and Zhu G.P., 2009, Structure and function of malic dehydrogenase, *Shengwuxue Zazhi (Journal of Biology)*, 26(4): 69-72 (汪新颖, 王波, 侯松涛, 朱国萍, 2009, 苹果酸脱氢酶的结构及功能, *生物学杂志*, 26(4): 69-72)
- Wu P, Ma L. G., Hou X. L., Wang M. Y., Wu Y.R., Liu F.Y., and Deng X.W., 2003, Phosphate starvation triggers distinct alterations of genome expression in Arabidopsis roots and leaves, *Plant Physiology*, 132(3):1260-1271
- Qi X., Li M., Xie M., Liu X., Ni M., Shao G., Song C., Yim A. K., Tao Y., Wong F., Isobe S., Wong C., Wong K., Xu C., Li C., Wang Y., Guan R., Sun F., Fan G., Xiao Z., Zhou F., Phang T., Liu X., Tong S., Chan T. Yiu S., Tabata S., Wang J., Xu X., and Lam H., 2014, Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing, *Nature Communications*, 5
- Xu W., Sato S.J., Clemente T.E., and Chollet R., 2007, The PEP-carboxylase kinase gene family in *Glycine max (GmPpcK1-4)*: an indepth molecular analysis with nodulated, non-transgenic and transgenic plants, *The Plant Journal*, 49(5): 910-923
- Zhou B.Y., Ding Z.S. and Zhao M., 2011, Alleviation of drought stress inhibition on photosynthesis by overexpression of PEPC gene in rice, *Zuowu Xuebao (Acta Agronomica Sinica)*, 37(1): 112-118 (周宝元, 丁在松, 赵明, 2011, PEPC 过表达可以减轻干旱胁迫对水稻光合的抑制作用, *作物学报*, 37(1): 112-118)

表3 互做蛋白的注释名称和参与的生物学过程

Table 3 The annotation name of proteins in the interacted network and their involved biological processes

互做蛋白名称 Names of proteins in the interaction network	互做蛋白的注释名称 Annotation name of protein in the interaction network	参与的生物学过程 Involved biological processes
Glyma.20G02980.1	丙酮酸激酶 Pyruvate kinase	糖酵解途径 glycolytic pathway
Glyma.07G35110.3	丙酮酸激酶 Pyruvate kinase	糖酵解途径 glycolytic pathway
Glyma.20G08780.1	未知蛋白 Uncharacterized protein	未知途径 Unknown pathway
Glyma.17G10880.1	ARP2/3 复合体 ARPC3 ARP2/3 complex ARPC3	未知途径 Unknown pathway
Glyma.13G16440.1	苹果酸脱氢酶 Malate dehydrogenase	TCA 循环, 光合作用, C4 途径 TCA cycle, Photosynthesis, C4 pathway
Glyma.12G19520.1	苹果酸脱氢酶 Malate dehydrogenase	TCA 循环, 光合作用, C4 循环 TCA cycle, Photosynthesis, C4 pathway
Glyma.11G04720.1	苹果酸脱氢酶, 乙醛酸循环体 Malate dehydrogenase, glyoxysomal	TCA 循环, 光合作用, C4 循环 TCA cycle, Photosynthesis, C4 pathway
Glyma.10G34150.1	未知蛋白 Uncharacterized protein	未知途径 Unknown pathway
Glyma.10G00920.1	苹果酸脱氢酶 Malate dehydrogenase	TCA 循环, 光合作用, C4 循环 TCA cycle, Photosynthesis, C4 pathway
Glyma.08G06820.1	苹果酸脱氢酶 Malate dehydrogenase	TCA 循环, 光合作用, C4 循环 TCA cycle, Photosynthesis, C4 pathway