

学校编码: 10384  
学号: 33320130153905

密级\_\_\_\_\_

廈門大學

博士学位论文

**多聚腺苷化因子 CPSF100 决定拟南芥表型  
和环境响应能力的机制研究**

**Mechanism of Cleavage and Polyadenylation Specificity  
Factor 100 on Determining Phenotypes and Ability of  
Environmental Response of Arabidopsis**

林 俊 城

指导教师姓名: 李庆顺 教授  
专业名称: 环境科学  
论文提交日期: 2017年3月  
论文答辩时间: 2017年4月

2017年3月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为( )课题(组)的研究成果,获得( )课题(组)经费或实验室的资助,在( )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1.经厦门大学保密委员会审查核定的保密学位论文，  
于 年 月 日解密，解密后适用上述授权。

2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

## 目 录

摘要.....	I
Abstract.....	IV
<b>第 1 章 绪论</b> .....	<b>1</b>
1.1 真核生物基因多聚腺苷化与分子机制.....	1
1.2 选择性多聚腺苷化.....	3
1.3 多聚腺苷化因子调控 mRNA 生命周期过程.....	4
1.3.1 多聚腺苷化因子与 RNA 转录酶 Pol II 活性相关.....	5
1.3.2 多聚腺苷化因子调控转录起始.....	6
1.3.3 多聚腺苷化因子调控转录延伸和转录本剪接.....	6
1.3.4 多聚腺苷化因子调控转录终止.....	6
1.3.5 多聚腺苷化因子参与 mRNA 运输.....	7
1.3.6 多聚腺苷化因子调控 mRNA 降解.....	7
1.4 多聚腺苷化在生物发育中的生物学功能及其分子机制.....	8
1.5 多聚腺苷化在生物对环境响应中的生物学功能及其分子机制.....	11
1.6 多聚腺苷化轮廓的研究方法.....	13
1.7 研究背景、内容和研究意义.....	14
1.7.1 研究背景.....	14
1.7.2 研究内容.....	15
1.7.3 研究意义.....	15
<b>第 2 章 CPSF100 调控植物表型和环境响应能力</b> .....	<b>16</b>
2.1 前言.....	16
2.2 材料与方法.....	17
2.2.1 拟南芥材料与培养.....	17
2.2.2 结实性状及遗传特性测定和显微镜观察.....	17
2.2.3 根系发育表型测定.....	17

2.2.4 抗旱、抗病能力测定.....	17
2.3 结果.....	18
2.3.1 AtCPSF100 调控拟南芥结实遗传 .....	18
2.3.2 AtCPSF100 影响根系发育 .....	19
2.3.3 AtCPSF100 调控拟南芥的环境响应能力 .....	20
2.4 讨论.....	21
<b>第 3 章 AtCPSF100 在拟南芥基因多聚腺苷化中的作用 .....</b>	<b>23</b>
3.1 前言.....	23
3.2 材料与方法.....	24
3.2.1 材料培养.....	24
3.2.2 突变的 AtCPSF100 蛋白结构模拟 .....	24
3.2.3 总 RNA 提取和质量控制 .....	25
3.2.4 PAT-seq 高通量测序文库的构建.....	25
3.2.5 高通量测序数据的质量检验和控制.....	29
3.2.6 合格测序数据的数据分析.....	29
3.3 结果.....	30
3.3.1 三维蛋白质结构模拟.....	30
3.3.2 PAT-seq 文库电泳分析 .....	30
3.3.3 原始测序数据的质量控制.....	31
3.3.4 Poly(A) 位点在基因组上的分布.....	32
3.3.5 Poly(A)位点在基因不同区间上的分布.....	33
3.3.6 AtCPSF100 调控不同基因类型的多聚腺苷化 .....	36
3.3.7 AtCPSF100 不影响近端 NUE 的 poly(A)信号选择 .....	37
3.3.8 AtCPSF100 影响远端 FUE 的 poly(A)信号选择.....	41
3.4 讨论.....	42
<b>第 4 章 CPSF100 调控多聚腺苷化和基因表达的机制 .....</b>	<b>45</b>
4.1 前言.....	45
4.2 材料与方法.....	45

4.2.1 材料培养.....	45
4.2.2 总 RNA 抽提、纯化与逆转录.....	46
4.2.3 数据分析.....	46
4.2.4 基因甲基化分析.....	46
4.2.5 荧光定量 PCR 验证基因表达.....	47
4.3 结果.....	48
4.3.1 CPSF100 调控基因多聚腺苷化和基因表达.....	48
4.3.2 AtCPSF100 在生物学过程中的作用.....	51
4.3.3 AtCPSF100 在分子水平上的功能.....	52
4.3.4 AtCPSF100 通过 APA 调控基因沉默的分子机制.....	53
4.4 讨论.....	56
<b>第 5 章 AtCPSF100 通过调控 Pol II 的磷酸化调整转录终止.....</b>	<b>58</b>
5.1 前言.....	58
5.2 材料与方法.....	59
5.2.1 材料培养和蛋白质的提取.....	59
5.2.2 数据分析与筛选.....	59
5.2.3 基因相对定量分析.....	59
5.2.4 Western Blot 分析.....	60
5.2.5 ChIP-qPCR 分析.....	60
5.3 结果.....	61
5.3.1 AtCPSF100 调控基因间区 poly(A)位点的表达.....	61
5.3.2 基因间区转录本延长范围和表达水平.....	62
5.3.3 CPSF100 调控 Pol II 去磷酸化基因的表达.....	63
5.3.4 CPSF100 调控 Pol II 上 CTD 的 Ser2 的磷酸化水平.....	64
5.3.5 AtCPSF100 调控 Pol II 在基因上的停留位置.....	66
5.4 讨论.....	67
<b>第 6 章 AtCPSF100 可能调控基因转录表达的其他过程.....</b>	<b>69</b>
6.1 前言.....	69

6.2 材料与方法.....	69
6.2.1 材料培养.....	69
6.2.2 RNA 提取纯化与 cDNA 合成.....	69
6.2.3 3'RACE 验证.....	70
6.2.4 基因表达的相对定量.....	71
6.2.5 数据分析.....	71
6.3 结果.....	71
6.3.1 AtCPSF100 调控 DNA 甲基化因子的表达 .....	71
6.3.2 AtCPSF100 调控多聚腺苷化因子的 poly(A)轮廓 .....	72
6.3.3 AtCPSF100 调控 RNA 剪接因子的多聚腺苷化和表达.....	73
6.3.4 转录本的延长没有明显增加 miRNA 的靶标.....	75
6.4 讨论.....	76
<b>第 7 章 结论与展望.....</b>	<b>78</b>
7.1 结论.....	78
7.2 特色与创新点.....	79
7.3 不足与展望.....	80
<b>参考文献.....</b>	<b>81</b>
<b>附录.....</b>	<b>90</b>
附图 2-1 <i>emb1265</i> 中胚胎的发育情况.....	90
附图 3-1 AtCPSF100 定位在拟南芥细胞核内.....	91
附图 3-2 AtCPSF100 与 AtCPSF73-I 和 AtCPSF73-II 的 C 末端互作.....	92
附图 6-1 AtCPSF100 与 RNA 聚合酶 Pol II 的 CTD 互作 .....	93
附表 2-1 <i>emb1265</i> 与 <i>Ler</i> 的测交试验 .....	94
附表 4-1 DE-PAC gene 和 DE gene 列表 ( <i>p</i> <sub>adj</sub> <0.05) .....	95
附表 6-1 <i>AtCPSF100</i> 和 <i>PCFS4</i> 的 poly(A).....	117
附表 6-2 含有 DE-PAC 的 RNA 剪接因子的 poly(A).....	118
附表 6-3 预测 DE-PAC 基因之间的 miRNA 靶标.....	119
博士期间的成果.....	120

参加会议.....	120
致谢.....	121

厦门大学博硕士论文摘要库



**Contents**

<b>Abstract (in Chinese)</b> .....	I
<b>Abstract (in English)</b> .....	IV
<b>Chapter 1 Introduction</b> .....	1
<b>1.1 Molecular mechanism of polyadenylation in eukaryotes</b> .....	1
<b>1.2 Alternative polyadenylation</b> .....	3
<b>1.3 The function of poly(A) factors in mRNA life circle</b> .....	4
1.3.1 Poly(A) factors affect the activity of Pol II .....	5
1.3.2 Poly(A) factors regulate transcription initiation .....	6
1.3.3 Poly(A) factors regulate transcription elongation and splicing .....	6
1.3.4 Poly(A) factors regulate transcription termination .....	6
1.3.5 Poly(A) factors are involved in RNA export .....	7
1.3.6 Poly(A) factors regulate transcription degradation .....	7
<b>1.4 The biological function and molecular mechanism of poly(A) factors in development</b> .....	8
<b>1.5 The biological function and molecular mechanism of poly(A) factors in environmental adapting</b> .....	11
<b>1.6 Methods for studying polyadenylation profile</b> .....	13
<b>1.7 Study background, content, method and significance</b> .....	14
1.7.1 Background .....	14
1.7.2 Content .....	15
1.7.3 Significance .....	15
<b>Chapter 2 CPSF100 regulates plant phenotypes and ability of environmental response</b> .....	16
<b>2.1 Intronduction</b> .....	16
<b>2.2 Materials and methods</b> .....	17

2.2.1 Materials and growth condition .....	17
2.2.2 Seeds phenotypic study .....	17
2.2.3 Root phenotypic study .....	17
2.2.4 Drought and Bacterial resistance test.....	17
<b>2.3 Results</b> .....	18
2.3.1 AtCPSF100 regulates seed setting rate .....	18
2.3.2 AtCPSF100 affects root development .....	19
2.3.3 AtCPSF100 coordinates plant environmental resistance.....	20
<b>2.4 Discussion</b> .....	21
<b>Chapter 3 The role of AtCPSF100 in polyadenylation of Arabidopsis</b> .....	23
<b>3.1 Introduction</b> .....	23
<b>3.2 Materials and methods</b> .....	24
3.2.1 Materials .....	24
3.2.2 Build 3-D model .....	24
3.2.3 RNA extraction and quality control .....	25
3.2.4 PAT-seq library preparation and sequencing .....	25
3.2.5 Data quality control .....	29
3.2.6 Data mining .....	29
<b>3.3 Results</b> .....	30
3.3.1 3-D structure of Protein .....	30
3.3.2 PAT-seq library.....	30
3.3.3 Data quality .....	31
3.3.4 The distribution of Poly(A) site on chromosome .....	32
3.3.5 The distribution of Poly(A) site on genes .....	33
3.3.6 AtCPSF100 regulates polyadenylation of genes .....	36
3.3.7 AtCPSF100 cannot affect NUE poly(A) signal choice .....	37
3.3.8 AtCPSF100 affects FUE poly(A) signal choice .....	41
<b>3.4 Discussion</b> .....	42

<b>Chapter 4 Mechanism of CPSF100 regulating polyadenylation and gene expression</b>	45
<b>4.1 Introduction</b>	45
<b>4.2 Materials and methods</b>	45
4.2.1 Materials	45
4.2.2 RNA extraction and RT	46
4.2.3 Data mining	46
4.2.4 Methylation analysis	46
4.2.5 Real-time PCR	47
<b>4.3 Results</b>	48
4.3.1 CPSF100 coordinates gene polyadenylation and expression	48
4.3.2 The role of AtCPSF100 in biological process	51
4.3.3 The molecular function of AtCPSF100	52
4.3.4 AtCPSF100 mediates gene silencing by APA	53
<b>4.4 Discussion</b>	56
<b>Chapter 5 AtCPSF100 coordinates transcription termination by modulating phosphorylation of Pol II</b>	58
<b>5.1 Introduction</b>	58
<b>5.2 Materials and methods</b>	59
5.2.1 Materials and protein extraction	59
5.2.2 Data analysis	59
5.2.3 Real-time pCR	59
5.2.4 Western Blot	60
5.2.5 ChIP-qPCR	60
<b>5.3 Results</b>	61
5.3.1 AtCPSF100 determines the expression level of intergenic poly(A) sites	61
5.3.2 The correlation of distance of elongation and expression level	62
5.3.3 CPSF100 affects the expression level of dephosphorylation gene of Pol II	63

5.3.4 The phosphorylation level of Ser2 on CTD of Pol II	64
5.3.5 AtCPSF100 affects Pol II retention on genes	66
<b>5.4 Discussion</b>	<b>67</b>
<b>Chapter 6 AtCPSF100 regulates other processes in RNA biogenesis</b>	<b>69</b>
<b>6.1 Introduction</b>	<b>69</b>
<b>6.2 Materials and methods</b>	<b>69</b>
6.2.1 Materials	69
6.2.2 RNA extraction and cDNA synthesis	69
6.2.3 3'RACE	70
6.2.4 Real-time PCR	71
6.2.5 Data mining	71
<b>6.3 Results</b>	<b>71</b>
6.3.1 AtCPSF100 is essential for methylase expression	71
6.3.2 AtCPSF100 affects the poly(A) profiles of poly(A) factors	72
6.3.3 AtCPSF100 affects the poly(A) profiles of splicing factors	73
6.3.4 AtCPSF100 cannot affect the miRNA target site	75
<b>6.4 Discussion</b>	<b>76</b>
<b>Chapter 7 Conclusion and prospect</b>	<b>78</b>
<b>7.1 Conclusion</b>	<b>78</b>
<b>7.2 Innovation of this study</b>	<b>79</b>
<b>7.3 Defects and prospects</b>	<b>80</b>
<b>References</b>	<b>81</b>
<b>Appendix</b>	<b>90</b>
<b>Fig. S2-1 embryo development of <i>emb1265</i></b>	<b>90</b>
<b>Fig. S 3-1 AtCPSF100 is located in nucleus</b>	<b>91</b>
<b>Fig. S 3-2 The interaction of AtCPSF100, AtCPSF73-I and AtCPSF73-II</b>	<b>92</b>
<b>Fig. S 6-1 AtCPSF100 interacts with CTD of Pol II</b>	<b>93</b>

**Table S2-1 Cross test of *emb1265* and *Ler*** .....94

**Table S4-1 DE-PAC gene and DE gene list (*padj*<0.05)** .....95

**Table S6-1 poly(A) profiles of AtCPSF100 and PCFS4** .....117

**Table S6-2 poly(A) profiles of RNA splicing factors** .....118

**Table S6-3 Predict of miRNA target sites**.....119

**Outcomes during PhD study** .....120

**Attended meeting** .....120

**Acknowledgements** .....121

厦门大学博硕士学位论文摘要库

## 摘要

多聚腺苷化 (Polyadenylation, poly(A)) 是真核生物大多数信使核糖核酸 (mRNA) 加工成熟的必经过程, 这一过程包含了 poly(A) 信号的识别、3' 末端的剪切和添加 poly(A) 尾巴。真核生物中有 75% 以上的基因含有至少 2 个 poly(A) 位点, 并通过选择使用位点调整基因表达。因此, 选择性 poly(A) (Alternative polyadenylation, APA) 被认为是调控基因表达的一种重要方式, 与植物适应变化的环境关系密切。Cleavage and Polyadenylation Specificity Factors 100 (CPSF100) 是多聚腺苷化核心装置的蛋白之一, 能够调控基因转录本通读, 并且调控拟南芥基因表达沉默、开花时间和胚胎发育。生化研究表明, CPSF100 可能是多聚腺苷化装置核心复合体 CPSF 的骨架蛋白, 位于多聚腺苷化蛋白因子互作网络中心。另外, CPSF100 为多聚腺苷化剪切因子 CPSF73 和 RNA 运输因子 THOC5 发挥功能所必需, 表明了 CPSF100 对真核细胞内基因添加 poly(A) 的准确进行至关重要。它又可以和 RNA 剪接因子共沉淀, 暗示着 CPSF100 在 RNA 剪接中的潜在功能。然而, 我们对 CPSF100 在真核生物基因的多聚腺苷化和基因转录中的具体作用还知之甚少, 也还未见 CPSF100 是否参与 poly(A) 信号识别的报道。本研究利用模式植物拟南芥为研究材料, 利用 Poly(A) tags sequencing (PAT-seq) 技术和生物信息学方法对拟南芥 *AtCPSF100* 的突变体 *esp5* 进行了全基因组范围转录本的 3' 末端进行高通量测序研究, 并从多方面研究了 this 遗传突变体的表型, 主要研究结果如下:

- (1) *AtCPSF100* 可以调控拟南芥的结实率和千粒重、主根长度、根毛长度和密度。它还能够调控对细菌感染的抵抗能力;
- (2) PAT-seq 表明 *AtCPSF100* 对 3'UTR 区域的 poly(A) 位点 (poly(A) tags cluster, PAC) 和表达水平 (poly(A) tags, PAT) 具有显著的正调控作用, 对基因间区的 poly(A) 位点和表达水平具有显著的负调控作用;
- (3) *AtCPSF100* 不仅参与编码基因的多聚腺苷化过程, 还能够参与非编码基因的多聚腺苷化过程, 体现了它对全基因组范围基因的调控作用;

(4) AtCPSF100 调控内含子区域 poly(A)位点上游近端区域的 poly(A)信号选择,表明了 AtCPSF100 可能与基因的转录本剪接有关;它参与调控 poly(A)位点上游远端区域 G-rich 的 poly(A)信号识别;

(5) 差异表达 (Differential expression, DE) 分析表明, AtCPSF100 至少显著调控 1385 个基因的 poly(A)位点选择 (DE-PAC gene), 还显著调控了 507 个基因 (DE gene) 的总体表达水平。这两类基因之间具有 398 个基因重叠。表明大多数差异表达的基因都是由差异的多聚腺苷化引起的; Gene Ontology (GO) 功能富集分析表明, DE-PAC gene 可以被显著富集到和拟南芥生长发育和抗性等相关的 14 个基因类型条目中, 而 DE gene 仅被显著富集到个别条目中, 表明了与植物生长发育和抗性相关的基因更容易受到 APA 的调控, 而 AtCPSF100 对基因表达的调控可能没有明显的偏好性;

(6) AtCPSF100 主要调控了编码具有结合活性蛋白质的基因的多聚腺苷化, 并通过这些蛋白质因子的表达来实现对下游基因的进一步调控。

(7) AtCPSF100 可能通过调控 RNA 聚合酶 II (Pol II) 去磷酸化的 CTD PHOSPHATASE-LIKE 3 (CPL3) 基因的表达水平从而调整 Pol II 的 C-terminal domain (CTD) 上 Ser2 的磷酸化水平, 从而促使基因转录能够正常终止, 防止发生通读导致基因沉默;

(8) AtCPSF100 为 *de novo* DNA 甲基化酶 *DRM1* 的表达所必需, 并能够调控自身、*PCFS4* (另一多聚腺苷化因子基因) 和 RNA 剪接因子的 poly(A)轮廓, 暗示了它可能参与基因转录加工的其他过程的调控; *esp5* 中 AtCPSF100 的突变可能没有引起 miRNA 靶标位点的增加。

本研究详细阐述了 AtCPSF100 在拟南芥进行准确多聚腺苷化过程中的作用, 发现多聚腺苷化因子调控拟南芥表型, 从而影响拟南芥的环境响应能力 (如根系表型的变化可能与植物抗旱有关, 抗病能力与环境生物胁迫有关)。本研究首次发现了植物多聚腺苷化因子能够同时控制编码和非编码基因的多聚腺苷化, 还第一次报道了植物多聚腺苷化因子参与远端 poly(A)信号选择和参与转录终止的调控。这些研究结果丰富了植物多聚腺苷化的分子机制的理论, 也为我们深入理解和研究真核生物 3'末端加工的工作机制奠定了基础。

关键词：选择性多聚腺苷化；多聚腺苷化信号；转录终止；转录后调控；3'末端形成

厦门大学博硕士学位论文摘要库



Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士论文摘要库