

Whirly转录因子研究进展

孔凡英, 邓永胜, 周斌, 孟庆伟*

山东农业大学生命科学学院, 作物生物学国家重点实验室, 山东泰安271018

摘要: Whirly蛋白是广泛存在于植物细胞内的一种转录因子。它既能与单链DNA结合, 也能与RNA结合, 无论在细胞核还是在质体内都有着广泛而复杂的生物学功能。本文概述了Whirly蛋白的结构、种类、分布及其作用机制, 并重点讨论了其在细胞核及质体内的功能, 最后对Whirly蛋白研究中需要解决的问题做了展望。

关键词: 转录因子; Whirly蛋白; 单链DNA结合蛋白

Research Advancement of Whirly Transcription Factors

KONG Fan-Ying, DENG Yong-Sheng, ZHOU Bin, MENG Qing-Wei*

State Key Laboratory of Crop Biology, College of Life Sciences, Shandong Agricultural University, Tai'an, Shandong 271018, China

Abstract: Whirly transcription factors which are ubiquitous in the cell were shown to bind to both ssDNA and RNA and have different functions both in nucleolus and plastids. This review summarizes the crystal structure, distribution, mechanism of Whirly proteins binding to ssDNA and focuses on their functions in nucleolus and plastids. The problems remain to be determined are also prospected.

Key words: transcription factor; Whirly protein; single-strand DNA binding protein

Whirly转录因子是一类植物所特有的高度保守的单链DNA结合(single-strand DNA binding, SSB)蛋白, 迄今为止在动物(包括线虫、果蝇)和酵母中都没有发现。第一个Whirly家族成员是2000年Desveaux等从马铃薯(*Solanum tuberosum*)中分离出来的核转录因子PBF-2 (PR-10a binding factor 2), 它可以与激发子应答元件ERE (elicitor response element)以单链形式结合。随后, Whirly家族成员在众多的物种中陆续被报道, 如在拟南芥中发现了3个该家族蛋白成员, 马铃薯中发现了2个(Krause等2005; Marechal等2008)。此外, 在大豆、小麦、水稻、玉米、百合等几十种植物中也都可以找到该蛋白的氨基酸序列。广泛的分布预示着它们在植物生长发育和各种生理过程中可能扮演着十分重要的角色(Grabowski等2008)。下面从Whirly蛋白家族的结构、分布和功能等几个方面全面回顾目前相关研究的概况。

1 Whirly蛋白的结构

蛋白质的结构决定功能, 研究转录因子的结构可以解释它们是如何与靶DNA序列结合, 如何激活转录, 以及它们的活性是如何调节的。学者们对Whirly蛋白的自由形式及Whirly-ssDNA复合体等的结构进行了大量研究。

1.1 Whirly蛋白一级结构的保守性

Whirly蛋白一般有3个结构域: Whirly结构域、N末端结构域和C末端多变区。Whirly结构域是最重要的结构域, 是ssDNA结合的区域; N末端结构域可能有叶绿体或线粒体信号肽以及转录激活区; C末端多变区具有自我调节区, 可以调节ssDNA结合活性(Desveaux等2002) (图1-A)。

Whirly蛋白在植物中是高度保守的, 将番茄*LeWhy1*基因编码的氨基酸序列与GenBank中所有的氨基酸序列进行比对, 结果显示: 无论双子叶植物还是单子叶植物的Whirly蛋白都是高度同源的(图1-B)。Whirly结构域是Whirly蛋白最为保守的区域, 尤其是氨基酸残基KGKAAL、YDW、K (图1-B中*标出), 说明该家族成员具有类似的ssDNA结合机制(Desveaux等2005)。这个结构域一般位于蛋白质的中间部分, 含有大约160个氨基酸, 由8个 β -折叠(β 1-8)和1个 α -螺旋(α 2)组成。在 β -折叠6(β 6)附近还有一个高度保守的区域是核定位信

收稿 2012-04-19 修定 2012-05-24

资助 国家重点基础研究发展计划(2009CB118505)和国家自然科学基金(31071338, 31171474)。

* 通讯作者(E-mail: qwmeng@sdau.edu.cn; Tel: 0538-8249606)。

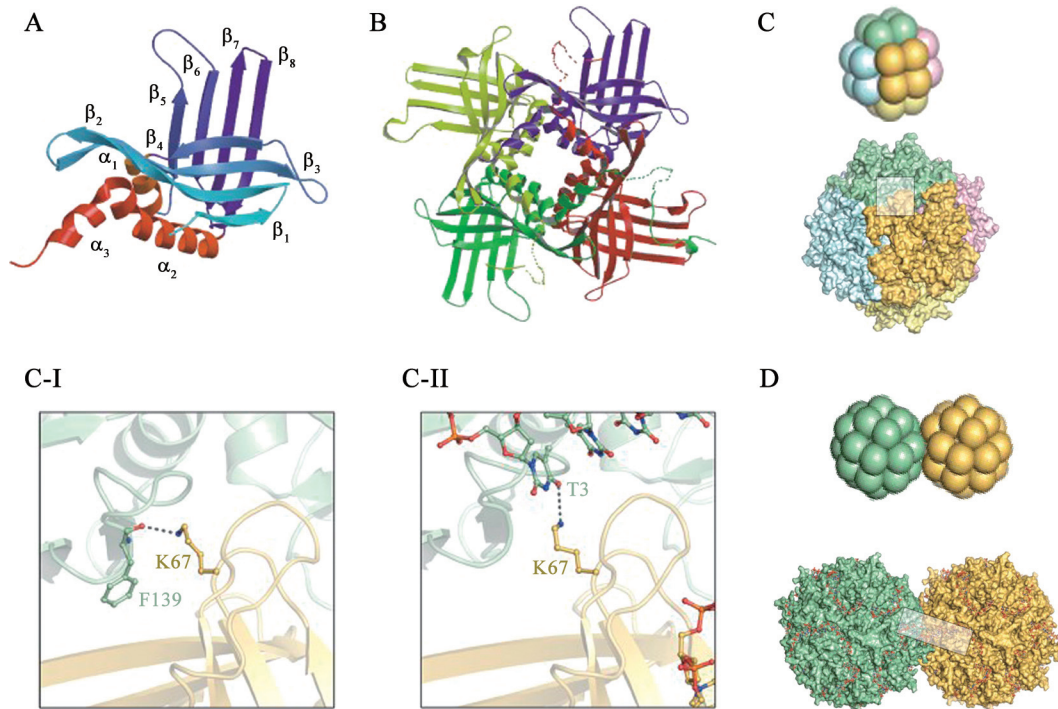


图3 Whirly蛋白的晶体结构

Fig.3 Crystal structure of Whirly proteins

A: Whirly蛋白单体的晶体结构(改自Desveaux等2002); B: Whirly蛋白四聚体的晶体结构(改自Desveaux等2002); C: Whirly蛋白二十四聚体的晶体结构(改自Cappadocia等2011); C-I: StWhy2的晶体结构(改自Cappadocia等2011); C-II: StWhy2-DNA的晶体结构(改自Cappadocia等2011); D: Whirly蛋白四十八聚体的晶体结构(改自Cappadocia等2011)。

2008)。据此,我们推测Whirly蛋白也可能利用中心空穴存储有害的物质,保护DNA免受胁迫的伤害。酵母双杂交实验显示Whirly蛋白也可以和不同的Whirly蛋白结合,形成异源四聚体(Desveaux等2005)。

1.2.3 Whirly蛋白二十四聚体的晶体结构 Cappadocia等(2011)发现Whirly蛋白的四聚体可以进一步组合形成四聚体的六聚体,即二十四聚体(图3-C)。二十四聚体是一个内空的432对称的球壳蛋白,外直径为12 nm,内直径为5 nm。Cappadocia等(2011)发现K67,即KGKAAL里的第2个赖氨酸,对于二十四聚体的形成至关重要。在马铃薯StWhy2的晶体结构中,它可以与F138位C骨架形成氢键(图3-C-I);在StWhy2-DNA晶体结构中,它也可以与碱基T3边缘形成氢键(图3-C-II)。突变体实验证明K67的突变不影响StWhy2形成四聚体及与ssDNA结合的能力,但突变的Whirly蛋白不能形成二十四聚体。Cappadocia等(2011)还得到了K67A(将67位的K突变为A)突变的Whirly蛋

白自由形式及其与dT32结合形成的复合体的晶体结构,发现该氨基酸突变对Whirly蛋白四聚体晶体结构的影响不大,只是139-145环处有微小重排,新的环和对称的四聚体结合,抑制二十四聚体的形成。此外,只有当Whirly蛋白浓度很高或与长的ssDNA结合时才可能形成二十四聚体,说明二十四聚体对于长的ssDNA的结合具有重要作用。

两个相邻的二十四聚体可以通过 β -折叠或外表面的ssDNA的相互作用形成四十八聚体(图3-D)。以此类推,四十八聚体可以进一步结合形成更大的蛋白质聚合体,从而可以与大分子的ssDNA结合(Cappadocia等2011)。

2 Whirly蛋白与ssDNA结合

ssDNA结合蛋白在核酸代谢中起着非常重要的作用,它们参与DNA复制、修复、重组、转录及端粒保护等过程。Whirly蛋白是广泛存在于植物中的一类ssDNA结合蛋白,它的功能依赖于与ssDNA的结合,因此了解Whirly蛋白与ssDNA结合

的机制具有重要意义。

2.1 Whirly蛋白与ssDNA结合特征

2.1.1 无序列特异性 Whirly蛋白曾被报道与3种核酸序列结合: ERE元件(Desveaux等2000, 2004)、端粒重复序列(Yoo等2007)、拟南芥*AtKPI*的上游元件(Xiong等2009), 这3种序列没有相似性。在叶绿体或线粒体内, Whirly蛋白也可以和叶绿体或线粒体基因组的许多区域结合, 这些区域也没有序列相似性(Cappadocia等2008; Prikryl等2008; Marechal等2009)。这意味着Whirly蛋白与ssDNA的结合并没有序列特异性。我们知道植物质体基因组是富含A/T的, Cappadocia等(2010)分析了St-Why2与其他富含A/T的ssDNA的互作, 试验显示ERE32、cERE32、rcERE32和dT32都可以和St-Why2结合, 并且它们的晶体结构也与StWhy2-ERE32的晶体结构一致。既然如此, 为什么报道的细胞核内的Whirly蛋白却与特定的序列结合呢? 我们猜测可能是因为Whirly蛋白主要定位于质体, 在细胞核内含量很少, 所以在细胞核内, Whirly蛋白只能和极少数具有明显单链结构的位置结合, 如转录区富含A/T的区域或端粒重复结构; 而在质体内, 丰富的Whirly蛋白使其能与各种ssDNA结合, 从而不具有序列特异性。在质体内, Whirly蛋白能否与特定的ssDNA序列结合, 调控质体基因的表达还有待进一步的研究。

2.1.2 对不同DNA的亲合力不同 Whirly蛋白可以与ssDNA、dsDNA的溶解部分、RNA结合, 但亲合力不同。StWhy2可以和16个碱基的黏性末端及16个碱基的ssDNA分子结合, 但对前者的亲合力明显高于后者(Yoo等2007)。这意味着解链的dsDNA对Whirly蛋白与ssDNA的结合也有作用。Whirly蛋白对ssDNA的亲合力高于dsDNA (Desveaux等2000; Prikryl等2008)。Cappadocia等(2010)用同位素标记dsDNA后用电泳的方法发现: 随着加入Whirly蛋白的增加, 产生了更多的ssDNA, 即Whirly蛋白可以诱导ssDNA的产生。这说明Whirly蛋白之所以更倾向于与ssDNA结合可能是因为Whirly蛋白可以促进dsDNA向单链状态转换。这也可以解释为什么Whirly蛋白对ssDNA的粘性末端有极高的亲合力。除了ssDNA外, Whirly蛋白还能与RNA结合, 但亲合力较低。

2.2 Whirly蛋白与ssDNA结合的分子机制

基于结构的序列分析显示无论叶绿体还是线粒体定位的Whirly蛋白ssDNA结合结构域的氨基酸是保守的, 在结构上也是一致的, 说明Whirly蛋白在2种细胞器中与ssDNA结合的机制是一致的, 并具有类似的功能。

2.2.1 Whirly蛋白结合ssDNA的位置 Whirly蛋白的高度保守的Whirly结构域是其行使ssDNA结合功能的结构域, 弯曲的 β -片层起着与ssDNA结合的作用。Desveaux等(2002)用分子筛凝胶层析和EMSA实验证明p24以四聚体的形式与ssDNA结合, 每个四聚体只能结合一分子的ssDNA, 并且推测ssDNA和 β -片层面结合, 并围绕p24四聚体折叠(图4-A)。Desveaux等(2002)用EMSA蛋白缺失突变实验进一步表明, 高度保守的KGKAAL是核心作用区域, 为了更好地结合, 还需要YDW和 β -折叠6位置的K。保守的KGKAAL基序也存在于其他ssDNA结合蛋白中, 如: DNA聚合酶, 端粒末端结合蛋白, RAD52及RecA同源蛋白, 说明该区域对于ssDNA的结合具有重要作用(Desveaux等2005)。为了进一步研究Whirly蛋白是如何与ssDNA结合的, Cappadocia等(2010)获得了StWhy2₄₈₋₂₁₆-ssDNA复合体的晶体结构(图4-B), 发现ssDNA主要结合在相邻原体的 β -片层边缘及之间, 而 β -片层的上部则没有, 这与其他ssDNA结合蛋白不同。事实上, 大部分ssDNA结合蛋白都以 β -片层的核心作为主要的结合平台。Cappadocia等(2010)推测在 β -片层上部之所以没有观察到DNA是因为在此区域的DNA形成了长的、极细的DNA分子, 或者进入了复合体内部。另外, Cappadocia等(2010)还发现尽管StWhy2可以与32个核苷酸长度的ssDNA形成晶体, 每个蛋白单体只可以结合9个核苷酸, 据此推测长的ssDNA链可以从四聚体的一个结合位置转移到另一个位置, 从而与长的ssDNA链结合。Cappadocia等(2011)又发现Whirly蛋白可以通过形成二十四聚体或更大的聚合体来和大分子的ssDNA结合。相互临近的二十四聚体结合的DNA片段的5'和3'端仅相距1.9 nm, 相当于3个核苷的距离, 所以长的ssDNA在晶体中有2种可能: 进入同一个四聚体的邻近的原体或进入邻近的二十四聚体。

2.2.2 Whirly蛋白结合ssDNA的作用力 Whirly蛋

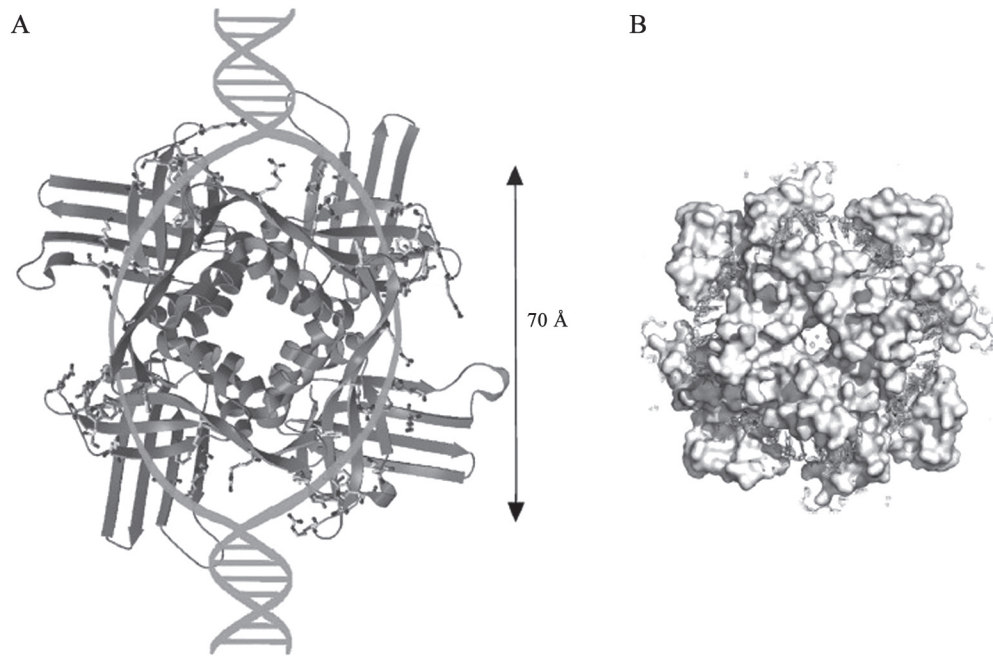


图4 Whirly蛋白与ssDNA结合的位置

Fig.4 The position of Whirly proteins binding to ssDNA

A: 改自Desveaux等(2002); B: 改自Cappadocia等(2010)。

白与ssDNA结合主要依赖于临近碱基及碱基与疏水氨基酸残基的堆叠和疏水键作用,还依赖其四倍体对称结构。ssDNA是以延伸的构象与Whirly蛋白结合的,dsDNA在结合过程中存在原子位阻,所以Whirly蛋白优先与ssDNA结合。另外,dsDNA骨架存在突然地扭曲也是其不能与Whirly蛋白结合的一个原因。Whirly蛋白主要通过与ssDNA的磷酸集团形成氢键,碱基在相邻亚基的 β -片层间堆叠来与ssDNA结合,只有很少的序列特异性互作,所以它们的结合没有序列特异性。Whirly蛋白的KGKAAL等氨基酸是ssDNA结合所必需的,是高度保守的,但这并不意味着Whirly蛋白与ssDNA是序列特异性互作,因为这些保守的氨基酸并不参与与碱基的结合,而是参与Whirly蛋白特定的空间结构的维持,如KGKAAL可以维持 β -片层的结构,它的第2个K是形成二十四聚体所需要的。

3 Whirly蛋白在植物细胞中的分布

Whirly蛋白在高等植物中分布广泛,大量存在于单子叶植物、双子叶植物和裸子植物中,并且每种植物中至少有2个该家族成员,分别定位于叶绿体和线粒体。生物信息学分析显示拟南芥的3个Whirly蛋白都有叶绿体或线粒体信号肽。拟南

芥原生质体及洋葱表皮GFP瞬时表达显示它们分别定位于叶绿体和线粒体(Krause等2005)。Grabowski等(2008)发现百合(*Hordeum vulgare*)的HvWhy1在同一个细胞里同时定位于细胞核和叶绿体,并且细胞核和叶绿体里的Whirly蛋白的分子量大小一致,只是在数量上叶绿体里的比细胞核里的多很多。而将Whirly蛋白预测的叶绿体信号肽去除后,HvWhy1主要定位于细胞核。这是迄今为止发现的第一个在同一个细胞内既定位于细胞核又定位于叶绿体的蛋白质。HvWhy1蛋白在叶绿体里主要定位于基质里,在类囊体膜上也有少许。在叶绿体的不同发育时期,Whirly蛋白的含量也不一样,在叶尖部分比叶柄部分多。通过双分子荧光技术,发现在细胞核里的Whirly蛋白主要以同源多聚体形式存在,而在叶绿体里的则不是(Grabowski等2008)。

4 Whirly蛋白在植物细胞内的生理功能

Whirly蛋白同时定位于细胞核和质体意味着它们在细胞核和叶绿体或线粒体中都具有重要的作用。目前对Whirly蛋白生理功能的研究主要集中在拟南芥中。人们通过过表达、RNAi、基因突变等功能基因组学研究手段获得了许多研究Whirly

蛋白生物功能的植物材料, 并利用这些植物材料研究了Whirly蛋白在抗病信号转导、维持端粒结构稳定、质体基因组修复、质体基因组表达等方面的作用(表1)。

4.1 Whirly蛋白在细胞核内的生理功能

4.1.1 抗病信号转导 植物细胞感受病菌感染后会激活一系列的信号转导, 引起基因表达的变化, *PR*基因就是可以被病原入侵激活表达的基因(Dangl和Jones 2001; Ryals等1996)。目前研究发现, *StWhy1*能特异地与ERE元件结合并诱导*PR-10a*基因的表达(Despres等1995)。体外研究发现非侵染条件下, *StWhy1*和*AtWhy1*保持非活性状态, 当受到适当的防卫信号刺激时才被激活(Desveaux等2000)。Desveaux等(2004)发现水杨酸可以激活*AtWhy1*的表达并参与水杨酸-依赖的抗病反应。水杨酸(salicylic acid, SA)是一种重要的植物激素, 在植物防卫反应中发挥重要功能, 它可以诱导抗病基因表达, 也能使植物获得系统获得性免疫(systemic acquired resistance, SAR)。非诱导免疫基因(non-expressor of PR genes1, NPR1)产物是SA诱导的SAR反应中一个十分重要的正调控因子, 位于SA积累的下游和*PR*基因表达的上游(Cao等1994; Delaney等1995; Mou等2003)。植物中似乎含有至少两条截然不同的SA调节防御反应的信号转导途

径: 一条依赖于NPR1, 调节*PR*基因的表达及对细菌和真菌病原体的抗性; 另一条可能不依赖于NPR1, 能激活对病毒的抗性。拟南芥*AtWhy1*本身的激活及发挥作用均不依赖于NPR1, 因此, *AtWhy1*可能是不依赖于NPR1的SA信号途径的一个重要元件, 其具体作用机制还需要进一步研究证实。*AtWhy1*也能与NPR1协同作用诱导*PR-1*基因的表达。此外, 姚沁涛等(2008)用RNA干扰技术发现水稻*OsWhirly*基因沉默会导致超敏反应(Hypersensitive response, HR)增强, 初步揭示Whirly转录因子可能是调控水稻HR反应的负调控因子。

4.1.2 维持端粒结构稳定 端粒是存在于真核细胞线状染色体末端的一小段DNA-蛋白质复合体, 它与端粒结合蛋白一起构成了特殊的“帽子”结构, 能够维持染色体的完整。端粒DNA主要功能有: 保护染色体不被核酸酶降解; 防止染色体相互融合; 为端粒酶提供底物, 解决DNA复制的末端隐缩, 保证染色体的完全复制(Greider 1996)。端粒动态平衡是维持染色体完整的重要过程, 需要端粒酶和其他相关蛋白的共同调节。在植物中, 端粒DNA末端存在一个大约20~30个碱基的富含G的粘性末端(Riha等2000), ssDNA结合蛋白通过与该区域结合调节端粒酶的活性, 如ssDNA结合蛋白STEP1就是这样调节端粒酶活性的(Kwon和Chung 2004)。

表1 Whirly蛋白的生物功能及其研究中用到的植物材料

Table 1 Plant materials used for researching the functions of Whirly proteins

来源	基因名称	材料名称	亚细胞定位	获得株系方法	Whirly生理功能	表型	参考文献
拟南芥	<i>AtWhy1</i>	KO-1 KO-2	叶绿体	T-DNA插入突变	维持端粒稳定	端粒变长, 端粒酶活性提高	Yoo等2007
拟南芥	<i>AtWhy1</i>	OE	叶绿体	35S启动子	维持端粒稳定	端粒变短, 端粒酶活性降低	Yoo等2007
拟南芥	<i>AtWhy2</i>	OEX	线粒体	35S启动子	调节线粒体基因表达	植株矮小, 角果短, 种子少; 叶片扭曲, 成熟叶早衰	Marechal等2008
拟南芥	<i>AtWhy1</i> <i>AtWhy3</i> <i>AtWhy1/3</i>	KO1 KO3 KO1/3	叶绿体	T-DNA插入突变	调节叶绿体基因表达	无 无 叶片卷曲, 叶绿体发育失常	Marechal等2009
拟南芥	<i>AtWhy1</i> <i>AtWhy3</i>	WHY1-OE WHY3-OE	叶绿体	35S启动子	调节细胞核基因表达	<i>AtKPI</i> 基因的表达被抑制	Xiong等2009
拟南芥	<i>AtWhy2</i>	K67A	线粒体	定点突变	抗病信号转导	对ciprofloxacin敏感	Cappadocia等2011
水稻	<i>OsWhy</i>	<i>OsWhy</i>	未知	RNAi	调控HR反应	病菌侵染后死亡率增高	姚沁涛等2008
玉米	<i>ZmWhy1</i>	<i>zmwhy1-1</i> <i>zmwhy1-2</i> <i>zmwhy1-1/-2</i>	叶绿体	转座子插入突变	调节叶绿体发育	叶片黄化, 植株矮小, 没有成熟核糖体	Prikryl等2008

材料名称为原文中使用的名称, KO表明基因敲除, OE、OEX表明过表达, K67A表明将67位的K突变为A。

研究发现, Myb类转录因子也可以与端粒末端重复序列结合(Karamysheva等2004; Schrupfova等2004; Yu等2000)。Yoo等(2007)通过质谱分析发现拟南芥AtWhy1是新的端粒结合蛋白,它可以特异地与植物单链端粒DNA序列结合。拟南芥*atwhy1*突变体的端粒长度明显增加,并且端粒酶活性也大大提高。与之对应的,在拟南芥中过表达该基因则使端粒变短,端粒酶活性变小。这说明AtWhy1在调节端粒长度和端粒酶活性,维持端粒动态平衡中具有重要作用。

4.2 Whirly蛋白在质体中的生理功能

4.2.1 调控质体基因组DNA-RRR途径

外界环境和生物体内部的因素经常会导致DNA分子的损伤或改变。DNA双链断裂(double-strands break, DSBs)是细胞最严重的损伤之一,它能引起遗传信息的缺失或重排,导致细胞死亡或突变。叶绿体或线粒体DNA要面对电子传递链所产生的大量的ROS,因此对其DNA复制、重组和修复(DNA replication, recombination, and repair, DNA-RRR)效率具有更高的要求。最近,科学家们鉴定出了一些细胞核编码的参与质体DNA-RRR途径的ssDNA结合蛋白,如线粒体定位的OSB1、在哺乳动物DNA-RRR过程起重要作用的复制蛋白A的同源蛋白等(Ishibashi等2006; Zaegel等2006)。Marechal等(2009)发现质体定位的Whirly蛋白参与维持拟南芥质体基因组的稳定。拟南芥*AtWhy1*和*AtWhy3*基因的双突变体不能形成有功能的叶绿体,质体基因组的重组区域也变大。Cappadocia等(2010)进一步研究发现Whirly蛋白对叶绿体和线粒体DNA双链断裂的修复是以错误倾向修复的方式进行的。Cappadocia等(2010)推测在这个过程中,Whirly蛋白通过与ssDNA的结合来阻止基因组错误的切割造成的退火。Cappadocia等(2011)进一步补充了这个模型:Whirly蛋白通过高度保守的氨基酸K进一步组装成二十四聚体,二十四聚体对质体基因组修复具有非常重要的作用。在突变体中过表达该氨基酸突变的基因所得到的植株依然对ciprofloxacin敏感(Cappadocia等2011)。

4.2.2 调节质体基因的表达

植物线粒体和叶绿体基因组分别编码大约70及100个基因,大部分基因产物参与质体基因表达或能量传递(Unseld等1997)。大量研究表明,许多细胞核基因编码的质

体定位的核酸结合蛋白可以调控质体基因的表达(Nickelsen 2003; Lurin等2004; Barkan等2007; Schwacke等2007)。Pfalz等(2006)通过凝胶过滤、亲和纯化等步骤分离了拟南芥和芥菜叶绿体转录活跃的染色质(transcriptionally active chromosome, TAC),从中鉴定出了35个组分,其中就有叶绿体定位的Whirly1蛋白,这说明Whirly蛋白可以与质体DNA结合。Prikryl等(2008)和Cappadocia等(2010)分别用免疫共沉淀的方法进一步证实了这个结论。玉米的ZmWhy1可以与CRS1互作,共同调节*atpF*内含子的剪切过程(Jenkins等1997; Asakura和Barkan 2006)。CRS1在体内和体外都可以和叶绿体基因*atpF*内含子结合,参与该基因内含子的剪切(Till等2001; Ostheimer等2003; Osterseizer等2005)。RNA酶或DNA酶可以打破CRS1与Zm-Why1的互作,说明它们是通过RNA或DNA互作的。*ZmWhy1*基因突变后*atpF*内含子剪切效率明显降低,成熟的23S、4.5S和16S rRNAs含量也下降,但rRNA前体增加,说明Whirly蛋白可能促进编码一种可以调节核糖体大亚基装配的基因的表达(Barkan 1993)。与叶绿体类似,线粒体DNA免疫共沉淀试验显示拟南芥线粒体定位的AtWhy2也可以和线粒体DNA结合。过表达*AtWhy2*基因的植株矮小,角果短,只能产生相当于WT一半的种子;叶片扭曲,黑绿色,成熟叶早衰,并且早衰相关基因SAGs大量表达(Gepstein等2003; Yoshida等2001)。Marechal等(2008)还用BN-PAGE的方法检测了过表达植株呼吸链复合体的活性,发现明显缺乏线粒体基因组编码的复合体I和IV,而细胞核基因编码的NADH脱氢酶和复合体II则没有变化,说明过表达该基因影响了线粒体基因的表达,而对细胞核基因的表达没有影响。进一步检测发现,线粒体基因组表达的基因如*nad3*、*nad4*、*nad7*、*cox1*、*cox2*、*cox3*和*atp8*、*atp9*、*rps3*、*rpl16*都明显下调。另外,Marechal等(2008)还发现代表成熟转录RNA的小RNA比代表初始转录物的大RNA变化小,这可能是因为转录后编辑过程补偿了初始转录物的下降,也说明Whirly蛋白的过量表达影响的是线粒体初始转录物的产生,这两个可能:抑制了转录或使线粒体DNA含量减少。检测线粒体DNA的含量发现,过表达*AtWhy2*基因的植株线粒体DNA含量明显下降(Marechal等2008)。

5 展望

植物有3个具有遗传信息DNA的细胞器——细胞核、线粒体和叶绿体。保护DNA, 维持遗传信息有序表达对于植物生存和发育至关重要。Whirly蛋白是迄今为止发现的第一个在同一个细胞内既定位于细胞核又定位于质体的蛋白质。它们在细胞核和质体内都具有非常重要的作用: 在细胞核内, Whirly蛋白能够调节基因表达, 参与SA依赖的抗病信号转导, 维持端粒稳定; 在质体内, 也能调节质体基因的表达, 维持质体基因组的稳定, 参与质体基因组修复(图5)。Whirly蛋白的结构及生理功能非常复杂, 虽然目前这方面的研究已经取得了很大的进展, 但还有许多问题值得进一步探讨。Whirly蛋白能够与ERE元件结合, 调节基因表达。事实上, 拟南芥基因组的很多基因都

含有ERE元件的核心序列GTCAAAA/T, 如过氧化物酶PRXCB (peroxidase)基因、酪氨酸转氨酶TAT3 (tyrosine aminotransferase3)、具有电子传递和氧结合能力的细胞色素P450家族蛋白基因及参与生长素合成的ACC合酶基因等, Whirly蛋白对于这些基因的表达是否具有调控作用, 有待进一步研究。迄今发现了许多可能的质体信号如ROS、光合电子传递链的氧化还原状态等; 也发现了许多核质信号转导组分如GUN, Executer1蛋白等, 但核质信号转导途径还不清楚, Whirly蛋白既能与细胞核DNA结合, 也能与质体DNA结合, 极有可能是核质信号转导组分, 它是否通过一种未知的机制调节核质基因表达的平衡, 参与核质信号转导尚待阐明。我们克隆了番茄的叶绿体定位的LeWhy1基因, 发现它被低温、高温、NaCl等非生物胁迫诱导表

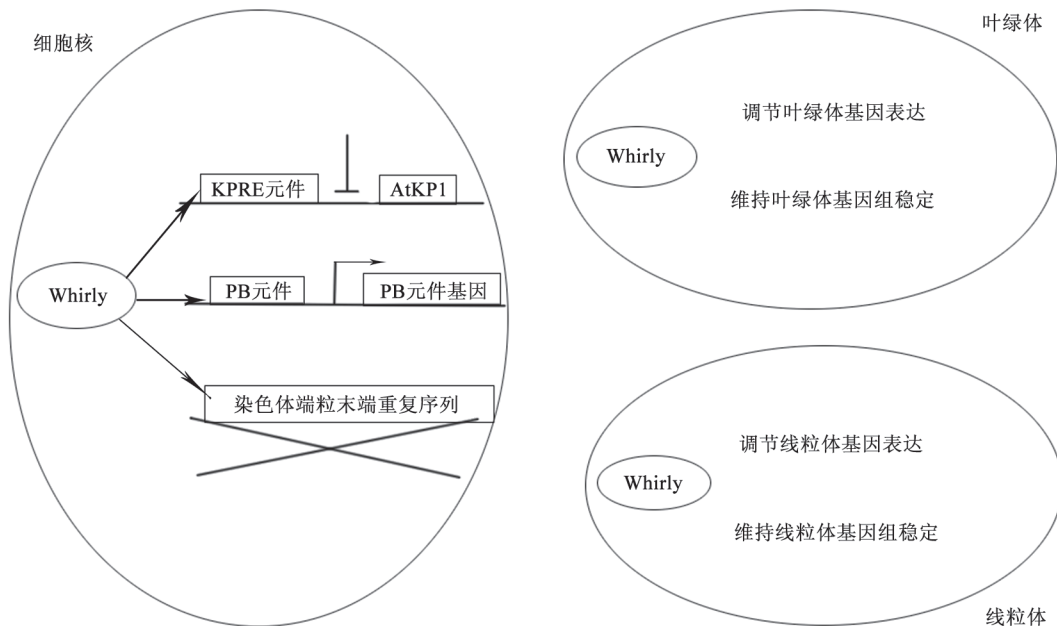


图5 Whirly蛋白在植物细胞中的生理功能
Fig.5 Functions of Whirly proteins in plant cells

达, 它在非生物胁迫下是否对植物具有保护作用, 及其具体的作用机理等问题仍需要解决。

参考文献

- 姚沁涛, 张文蔚, 刘莉, 韩榕, 简桂良, 齐放军(2008). Whirly转录因子对非寄主菌诱导水稻HR反应的负调控作用. 中国农业科技导报, 10: 53~58
- Asakura Y, Barkan A (2006). *Arabidopsis* orthologs of maize chlo-

- roplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant Physiol*, 142: 1656~1663
- Barkan A (1993). Nuclear mutants of maize with defects in chloroplast polysome assembly have altered chloroplast RNA metabolism. *Plant Cell*, 5: 389~402
- Barkan A, Klipcan L, Ostersetzer O, Kawamura T, Asakura Y, Watkins KP (2007). The CRM domain: an RNA binding module derived from an ancient ribosome-associated protein. *RNA*, 13: 55~64
- Cao HC, Bowling SA, Gordon AS, Dong X (1994). Characterization

- of an *Arabidopsis* mutant that is nonresponsive to inducers of systemic acquired resistance. *Plant Cell*, 6: 1583–1592
- Cappadocia L, Marechal A, Parent JS, Lepage E, Sygusch J, Brisson N (2010). Crystal structures of DNA-Whirly complexes and their role in *Arabidopsis* organelle genome repair. *Plant Cell*, 22: 1849–1867
- Cappadocia L, Parent JS, Zampini E, Lepage E, Sygusch J, Brisson N (2011). A conserved lysine residue of plant Whirly proteins is necessary for higher order protein assembly and protection against DNA damage. *Nucl Acids Res*, 40: 258–269
- Cappadocia L, Sygusch J, Brisson N (2008). Purification, crystallization and preliminary X-ray diffraction analysis of the Whirly domain of StWhy2 in complex with single-stranded DNA. *Cryst Commun*, 64: 1056–1059
- Dangl JL, Jones JDG (2001). Plant pathogens and integrated defence responses to infection. *Nature*, 411: 826–833
- Delaney T, Friedrich L, Ryals J (1995). *Arabidopsis* signal transduction mutant defective in chemically and biologically induced disease resistance. *Proc Natl Acad Sci USA*, 92: 6602–6606
- Despres C, Subramaniam R, Matton DP, Brisson N (1995). The activation of the potato PR-10a gene requires the phosphorylation of the nuclear factor PBF-1. *Plant Cell*, 7: 589–598
- Desveaux D, Allard J, Brisson N, Sygusch J (2002). A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nat Struct Biol*, 9: 512–517
- Desveaux D, Despres C, Joyeux A, Subramaniam R, Brisson N (2000). PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *Plant Cell*, 12: 1477–1490
- Desveaux D, Marechal A, Brisson N (2005). Whirly transcription factors: defense gene regulation and beyond. *Trends Plant Sci*, 10: 95–102
- Desveaux D, Subramaniam R, Despres C, Mess JN, Levesque C, Fobert PR, Dangl JL, Brisson N (2004). A “Whirly” transcription factor is required for salicylic acid-dependent disease resistance in *Arabidopsis*. *Dev Cell*, 6: 229–240
- Gepstein S, Sabehi G, Carp MJ, Hajouj T, Neshor MF, Yariv I, Dor C, Bassani M (2003). Large-scale identification of leaf senescence-associated genes. *Plant J*, 36: 629–642
- Grabowski E, Miao Y, Mulisch M, Krupinska K (2008). Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant Physiol*, 147: 1800–1804
- Grant RA, Filman DJ, Finkel SE, Kolter R, Hogle JM (1998). The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat Struct Biol*, 5: 294–303
- Greider CW (1996). Telomere length regulation. *Annu Rev Biochem*, 65: 337–365
- Ishibashi T, Kimura S, Sakaguchi K (2006). A higher plant has three different types of RPA heterotrimeric complex. *J Biochem*, 139: 99–104
- Jenkins B, Kulhanek D, Barkan A (1997). Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell*, 9: 283–296
- Karamysheva ZN, Surovtseva YV, Vespa L, Shakirov EV, Shippen DE (2004). A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in *Arabidopsis*. *J Biol Chem*, 279: 47799–47807
- Krause K, Kilbiński I, Mulisch M, Rodiger A, Schafer A, Krupinska K (2005). DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Lett*, 579: 3707–3712
- Krojer T, Sawa J, Schafer E, Saibil HR, Ehrmann M, Clausen T (2008). Structural basis for the regulated protease and chaperone function of DegP. *Nature*, 453: 885–890
- Kwon C, Chung IK (2004). Interaction of an *Arabidopsis* RNA-binding protein with plant single-stranded telomeric DNA modulates telomerase activity. *J Biol Chem*, 279: 12812–12818
- Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B (2004). Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, 16: 2089–2193
- Marechal A, Parent JS, Sabar M, Veronneau-Lafortune F, Abou-Rached C, Brisson N (2008). Overexpression of mtDNA-associated At-Why2 compromises mitochondrial function. *BMC Plant Biol*, 8: 42
- Marechal A, Parent JS, Veronneau-Lafortune F, Joyeux A, Lang BF, Brisson N (2009). Whirly proteins maintain plastid genome stability in *Arabidopsis*. *Proc Natl Acad Sci USA*, 106: 14693–14698
- Mou Z, Fan W, Dong X (2003). Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell*, 113: 935–944
- Nickelsen J (2003). Chloroplast RNA-binding proteins. *Curr Genet*, 43: 392–399
- Ostersetzer O, Watkins K, Cooke A, Barkan A (2005). CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. *Plant Cell*, 17: 241–255
- Ostheimer G, Williams-Carrier R, Belcher S, Osborne E, Gierke J, Barkan A (2003). Group II intron splicing factors derived by diversification of an ancient RNA binding module. *EMBO J*, 22: 3919–3929
- Pfalz J, Liere K, Kandlbinder A, Dietz KJ, Oelmüller R (2006). pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell*, 18: 176–197
- Prikryl J, Watkins KP, Friso G, van Wijk KJ, Barkan A (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucl Acids Res*, 36: 5152–5165
- Ragunathan S, Ricard CS, Lohman TM, Waksman G (1997). Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-Å resolution. *Proc Natl Acad Sci USA*, 94: 6652–6657
- Riha K, McKnight TD, Fajkus J, Vyskot B, Shippen DE (2000). Anal-

- ysis of the G-overhang structures on plant telomeres: evidence for two distinct telomere architectures. *Plant J*, 23: 633~641
- Ryals JA, Neuenschwander UH, Willits MG, Molina A, Steiner HY, Hunt MD (1996). Systemic acquired resistance. *Plant Cell*, 8: 1809~1819
- Schrumpfova P, Kuchar M, Mikova G, Skrisovska L, Kubiarova T, Fajkus J (2004). Characterization of two *Arabidopsis thaliana* myb-like proteins showing affinity to telomeric DNA sequence. *Genome*, 47: 316~324
- Schwacke R, Fischer K, Ketelsen B, Krupinska K, Krause K (2007). Comparative survey of plastid and mitochondrial targeting properties of transcription factors in *Arabidopsis* and rice. *Mol Gen Genet*, 277: 631~646
- Sutter M, Boehringer D, Gutmann S, Gunther S, Prangishvili D, Loessner MJ, Stetter KO, Weber-Ban E, Ban N (2008). Structural basis of enzyme encapsulation into a bacterial nanocompartment. *Nat Struct Mol Biol*, 15: 939~947
- Theil EC (1987). Ferritin: structure, gene regulation, and cellular function in animals, plants, and microorganisms. *Annu Rev Biochem*, 56: 289~315
- Till B, Schmitz-Linneweber C, Williams-Carrier R, Barkan A (2001). CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin. *RNA*, 7: 1227~1238
- Unsel M, Marienfeld JR, Brandt P, Brennicke A (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat Genet*, 15: 57~61
- Urich T, Gomes CM, Kletzin A, Frazao C (2006). X-ray structure of a self-compartmentalizing sulfur cycle metalloenzyme. *Science*, 311: 996~1000
- Xiong JY, Lai CX, Qu Z, Yang XY, Qin XH, Liu GQ (2009). Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene *AtKPI1* to mediate transcriptional repression. *Plant Mol Biol*, 71: 437~449
- Yang C, Curth U, Urbanke C, Kang C (1997). Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. *Nat Struct Biol*, 4: 153~157
- Yoo HH, Kwon C, Lee MM, Chung IK (2007). Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in *Arabidopsis*. *Plant J*, 49: 442~451
- Yoshida S, Ito M, Nishida I, Watanabe A (2001). Isolation and RNA gel blot analysis of genes that could serve as potential molecular markers for leaf senescence in *Arabidopsis thaliana*. *Plant Cell Physiol*, 42: 170~178
- Yu EY, Kim SE, Kim JH, Ko JH, Cho MH, Chung IK (2000). Sequence-specific DNA recognition by the Myb-like domain of plant telomeric protein RTBP1. *J Biol Chem*, 275: 24208~24214
- Zaegel V, Guermann B, Le Ret M, Andres C, Meyer D, Erhardt M, Canaday J, Gualberto JM, Imbault P (2006). The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in *Arabidopsis*. *Plant Cell*, 18: 3548~3563