

DOI: 10.3724/SP.J.1005.2013.00867

## 植物基因组 DNase I 超敏感位点的研究进展

张韬, 杨足君

电子科技大学生命科学与技术学院, 成都 610054

**摘要:** 真核生物的基因表达与调控是一系列顺式作用元件(*cis*-acting element)与反式作用因子(*trans*-acting factor)相互作用的结果。顺式作用元件本身不编码蛋白, 只是基因组中一些特定的区域, 它们需与反式作用因子结合才能发挥其生理作用。目前已知, 顺式作用元件往往与染色质的结构密切相关, 能与反式作用因子结合发挥生物学功能的顺式作用元件, 通常存在于脱氧核糖核酸酶 I 超敏感位点 (DNase I hypersensitive site) 的内部或周围。随着越来越多的植物基因组测序的进展, 大规模高通量鉴定植物基因组的顺式作用元件正在受到关注, 通过鉴定脱氧核糖核酸酶 I 超敏感位点可以有效确定顺式作用元件的富集区域, 将对功能基因组学研究提供的数据支持。文章综述了近期高通量鉴定和分析植物基因组脱氧核糖核酸酶 I 超敏感位点研究进展。

**关键词:** 基因表达调控; 脱氧核糖核酸酶 I 超敏感位点; 植物基因组

## Progress on identification and analysis of DNase I hypersensitive sites in plant genomes

ZHANG Tao, YANG Zu-Jun

School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China

**Abstract:** Eukaryotes's gene expression and regulation relies on the interaction of their *cis*-acting elements and *trans*-acting factors. The *cis*-elements are specific DNA sequences in the genome and frequently located in the untranslated regions. The *trans*-acting factors are usually considered to be proteins that bind to the *cis*-acting sequences to regulate gene expression. It is well known that the *cis*-elements are always associated with DNase I hypersensitive site, which is a signature of open chromatin. The identification of all the functional *cis*-element using high-throughput method in plant genome has not been initiated in plant genomes. With the rapid achievement of genomics studies, an increasing number of plant genomes have been sequenced. Genome-wide identification of DNase I hypersensitive sites will be a considerably efficient method to locate *cis*-element in plants, which will provide a vital potential for further plant functional genomics. The present review is to reveal the recent progresses on identification and analysis of DNase I hypersensitive site in plant genomes.

**Keywords:** gene expression and regulation; DNase I hypersensitive site; plant genome

收稿日期: 2012-12-20; 修回日期: 2013-02-04

基金项目: 国家自然科学基金项目(编号: 31171542 和 31101143)和电子科技大学中央高校业务费项目资助

作者简介: 张韬, 博士研究生, 研究方向: 植物生物信息学。Tel: 028-83206556; E-mail: forrestzhang1982@gmail.com

通讯作者: 杨足君, 博士, 教授, 研究方向: 植物分子细胞遗传学。E-mail: yangzujun@uestc.edu.cn

网络出版时间: 2013-2-25 16:38:06

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20130225.1638.002.html>

随着模式生物基因组学的推进,大量物种的全基因组序列已经顺利完成,为人们全面认识基因组提供了非常有利的条件。然而真核生物基因的转录水平受多种因素影响,基因组中的功能性调控元件对基因的表达起着至关重要的作用。对于调控元件的研究,常规的研究方法,都是对与某类基因相关的调控元件逐一进行分析,因而可以具体阐释生物体内这些特定基因表达调控的机理,但是该手段只适合于对某些已知调控蛋白的结合序列进行研究,对基因组内调控元件的大规模筛选和鉴定并不适用。早在 20 世纪 80 年代初期,研究人员发现染色质对脱氧核糖核酸酶 (DNase) 的敏感性与基因的转录有关,即当一个基因处于转录活性状态时,含有这个基因的染色质区域对 DNase 降解的敏感性要比无转录活性区域高出 100 倍以上,被称为 DNase 超敏感位点(DNase hypersensitive site, DHS),当时认为是真核生物某部分染色质处于开链或松散状态(Open chromatin),容易被 DNase 非特异性降解<sup>[1-3]</sup>。经过进一步研究发现,具有功能的顺式作用元件,如启动子、增强子、抑制因子和绝缘子等都与 DNase 超敏感位相偶联<sup>[4]</sup>。因此,鉴定 DNase 超敏感位点成为近年来国内外基因组学研究的热点。但是由于技术的限制,用传统的凝胶电泳方法鉴定 DNase 超敏感位点的工作量特别巨大<sup>[2,5,6]</sup>。随着高通量技术尤其是芯片和大规模测序技术的发展,目前在酵母、果蝇以及人类基因组研究中,通过对 DNase 降解后的片段进行芯片杂交或大规模测序,对其结果开展系统的生物信息学分析,成功地从全基因组水平上准确鉴定了 DNase 超敏感位点<sup>[7-12]</sup>。在植物基因组研究领域,美国威斯康辛大学 Jiming Jiang 实验室率先通过使用回收 DNase 降解片段,进行大规模测序技术,在水稻和拟南芥中鉴定了大量 DNase 超敏感位点<sup>[13,14]</sup>,为进一步探索 DNase 超敏感位点在植物基因表达调控中的作用奠定了基础。本文着重以植物为例,介绍如何通过高通量手段鉴定 DNase 超敏感位点,阐释 DNase 超敏感位点与植物基因表达调控、组蛋白修饰以及甲基化等之间的关系。

## 1 高通量鉴定 DNase I 超敏感位点

目前,通过大规模测序技术鉴定 DNase 超敏感位点的主要路线是:(1)细胞核提取;(2)DNase 消

化降解;(3)脉冲场凝胶电泳分离大片段;(4)通过 T4 DNA 连接酶连接一个含有 *Mem* 限制性内切酶识别位点并在末端标记了生物素的接头;(5)利用 *Mem* 限制性内切酶酶切连接产物;(6)在 *Mem* 切口处连接另一个接头,随后进行 PCR 扩增连接产物;(7)PCR 产物回收,建库,并使用 Illumina 高通量测序仪测序;(8)生物信息学分析测序结果,鉴定 DNase 超敏感位点(图 1)<sup>[13,14]</sup>。由于 *Mem* I 限制性内切酶的特性是在生物素标记位置的 20 bp 处,因而获得片段长度非常适宜于 Illumina 高通量测序,可以较同类研究获得更高通量的数据,利于准确鉴定和分析基因组中 DNase 超敏感位点,适用于较大的植物基因组的 DNase 超敏感位点研究。

对基因组 DNase 超敏感位点的高通量测序数据的分析流程如图 2 所示。首先,需要除去建库时的接头序列,然后将除去接头的数据和目标基因组进行比对。利用 F-Seq 对对比结果进行分析,输出 wig 格式文件以供后续 GBrowse 使用。如果 F-Seq<sup>[15]</sup>鉴定出的区域结果满足  $FDR < 0.01$  的条件,则该位点为 DNase 超敏感位点。再将 DNase 超敏感位点与表达谱、甲基化以及组蛋白数据等进行关联性分析。最后,将上述分析所得出的结果使用基因组可视化工具 GBrowse 来进行显示,这样建立的 DNase 超敏感位点就可以与公共数据库的基因组信息进行整合,开展深入的比较分析。

通过比较分析发现,使用高通量数据鉴定的基因组 DNase 超敏感位点,与使用传统的 Southern blot 方法鉴定的超敏感位点具有高度的一致性(图 3)。如图所示,该区域为拟南芥第 5 号染色体长臂的一个 80 kb 区域为例。在这段区域中共含有 34 个基因,Kodama 等<sup>[6]</sup>通过 Southern 杂交的方法鉴定了 40 个 DNase 超敏感位点,其中 29 个超敏感位点与高通量数据完全吻合,有 9 个 Southern 杂交的超敏感位点非常接近高通量数据的结果,但是由于在高通量数据分析中使用了比较严谨的参数( $FDR < 0.01$ ),在最终的结果中将这部分超敏感位点进行了过滤。另外,通过高通量数据分析,发现了 13 个以前未发现的超敏感位点<sup>[13,14]</sup>。使用高通量 Single-Ends 数据可以大规模的识别和鉴定 DNase 超敏感位点。但是,由于使用的 Illumina 测序结果是短的序列,在进行序列比对(Sequence mapping)时要求较高的基因

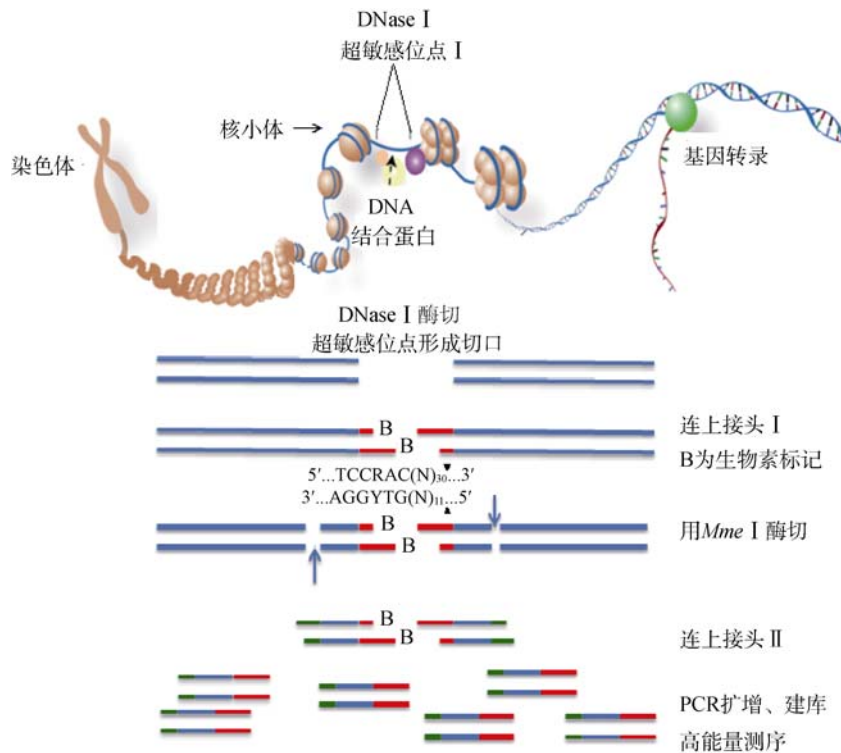


图 1 DNase 超敏感建库流程

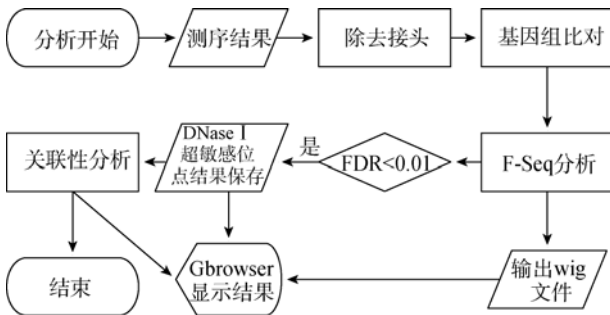


图 2 DNase 超敏感位点生物信息学分析流程

组序列的测序质量, 因而对只有草图或者含有大量重复序列的基因组, 进行 DNase 超敏感位点进行分析时, 常常就需要大幅度提高测序深度, 从而导致相应的测序费用也会相应增加, 这也是短序列分析 DNase 超敏感位点的缺点, 因而改进实验方法从而获得更长的片段, 值得进一步完善。

## 2 DNase 超敏感位点的分布

在真核生物基因组中, 由于 DNase 超敏感位点与功能性的顺式作用元件密切相关, DNase 超敏感位点在基因组中的分布研究至关重要。研究表明, 由于基因组的结构不同, DNase 超敏感位点的分布

也不尽相同<sup>[7,13,14]</sup>。从表 1 中可以发现, DNase 超敏感位点的分布与基因组大小和基因长度密切相关。对于植物基因组来说, 基因组越大超敏感位点就越倾向于分布在基因间区(Intergenic region)。我们在水稻叶片中总共鉴定出 97 974 个超敏感位点, 而在人类中有 94 925 个超敏感位点。在基因间区中, 人类有 39%的超敏感位点位于该区域内, 而在水稻中则是 42%, 两者的结果非常相似。但是, 在水稻和人类的基因中超敏感位点的分布却不尽相同。超敏感位点在人类基因组的外显子中的比例只有水稻的大约 1/2, 而在内含子中却是水稻的 4 倍, 这是由于水稻和人类基因组结构的差异造成的, 因为人类基因组中内含子的平均长度为 6 155 bp, 是水稻平均长度的 15 倍, 而且在人类基因组中内含子长度占基因总长度的比例约为 78%, 在水稻中则是 38%。因而通过大量物种测序基因组的分析, 可以深入理解 DNase I 超敏感位点的分布与基因组进化的内在联系。

## 3 DNase 超敏感位点与基因表达的关系

通过大规模 RNA 测序(RNA-Seq)结果与 DNase 超敏感位点的关联分析, 研究发现高表达基因内部或上下游往往存在一个或多个超敏感位点, 而沉

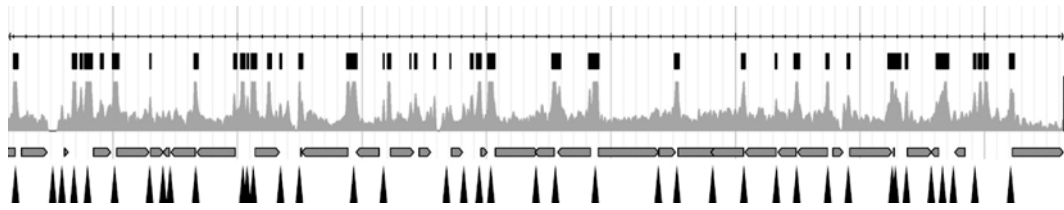


图 3 传统 Southern 杂交与高通量测序鉴定 DNase 超敏感位点的比较<sup>[14]</sup>

图中黑色方块为高通量测序方法鉴定的 DNase 超敏感位点, 峰图为测序结果, 横向箭头为基因, 三角箭头为传统 Southern 杂交鉴定的 DNase 超敏感位点。

默或者低表达的基因或基因周围却很少存在超敏感位点<sup>[16-18]</sup>。例如, 在人类 CD4<sup>+</sup>T 细胞中, 2 000 个高表达基因中有 1 933 个与超敏感位点相关; 在水稻的叶片中, 有 93% 的高表达基因与超敏感位点相关, 而在低表达基因中只有 23% 的基因与超敏感位点有关, 这说明 DNase 超敏感位点的存在与基因表达高度相关<sup>[13]</sup>。在拟南芥中, 根据 RNA-Seq 结果将基因从表达量由高到低分为 5 类, 发现位于转录起始位点(TSS)上游和转录终止位点(TTS)下游的脱氧核糖核酸酶 I 的敏感程度与基因的表达水平呈正相关, 即表达水平越高 DNase 越敏感<sup>[14]</sup>, 这也证实了在有转录因子与顺式作用元件结合的区域, 染色质的结构处于开链或者松散状态, 即更容易被 DNase 所降解(图 4)。

#### 4 DNase 超敏感位点与表观遗传学标记的关系

研究表明, DNase 超敏感位点与植物基因组的甲基化程度密切相关。例如在水稻基因组中无论是基因区(基因内部、启动子区、基因下游)还是基因

间区的超敏感位点的甲基化水平都明显的低于周围的非超敏感区域(图 4)。DDM1 是维持正常甲基化的关键蛋白, 拟南芥的 *DDM1* 基因突变后, 其基因组的甲基化水平会显著降低<sup>[19-23]</sup>。研究表明, 拟南芥 *ddm1* (Decreased DNA methylation 1)突变植株中, 位于着丝粒和近着丝粒区中的在野生型中高甲基化区域, 会变得比野生型更容易被 DNase 降解(图 5), 而在这段区域中, *ddm1* 突变株的甲基化水平相对于野生型来说明显的降低<sup>[24]</sup>。基于上述两方面的证据, 可以得出植物基因组中甲基化水平与 DNase 的敏感程度密切相关, 具有 DNase 超敏感位点的区域往往倾向于低甲基化。根据 DNase 超敏感位点与 7 种组蛋白修饰以及 LHP1 的关联性分析我们发现<sup>[25-28]</sup>, DNase 超敏感位点与基因激活相关的组蛋白修饰 (H3K9ac<sup>[26]</sup>、H2A.Z<sup>[27]</sup>、H3K4me3<sup>[28]</sup>)呈正相关, 而与基因沉默的组蛋白修饰(H3K9me2<sup>[26]</sup>)以及 LHP1<sup>[25]</sup> (异染色质相关蛋白)负相关 (图 6)。因此, 进一步开展植物表观遗传突变体的 DNase 超敏感位点的分布与变异规律, 可以推动对植物表观遗传修饰对基因表达调控分子机制的理解。

表 1 植物与人类基因组中 DNase 超敏感位点的分布比较(%)

基因组位置	水稻叶片	水稻愈伤	拟南芥叶片	拟南芥花	人类 CD4 <sup>+</sup> T 细胞
基因上游 200 bp	13.5	10.0	28.0	26.0	-
基因上游 200 bp~1 kb	13.9	13.9	20.0	20.0	-
基因上游 2 kb	-	-	-	-	13.0
5'UTR	5.1	3.8	6.0	7.0	-
外显子	16.6	17.3	9.0	9.0	6.0
内含子	10.6	11.1	6.0	5.0	39.0
3'UTR	5.5	6.8	5.0	5.0	-
基因下游 200 bp	4.3	4.4	4.0	4.0	-
基因下游 200 bp~1 kb	-	-	8.0	8.2	-
基因下游 2 kb	-	-	-	-	3.0
转座子相关	-	-	1.0	1.0	-
基因间区	42.3	44.5	15.0	15.2	39.0

注: 表中 UTR 为非翻译区, - 为未统计。

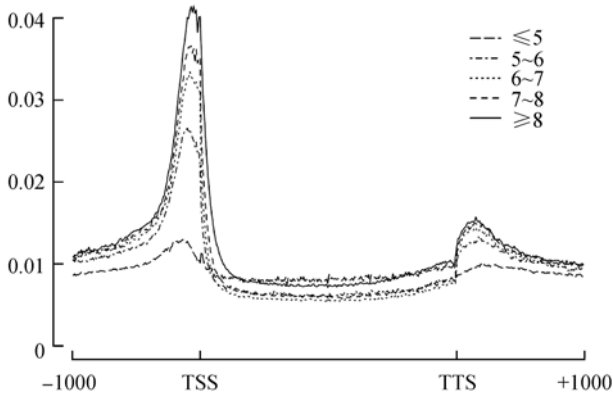


图 4 基因表达与 DNase 敏感程度的关系  
 图中将基因按表达水平分类, 用不同的曲线表示该类基因在基因上下游 1 kb 和基因内的 DNase 的敏感程度(DNase Reads 标准化后的均值)。图中实线为 Log<sub>2</sub> 表达值基因表达值大于 8 的基因<sup>[16-18]</sup>, 虚线分别为表达值 7~8、6~7、5~6 以及小于 5 的基因。由图中可以得出, 基因表达量越高, DNase 的敏感程度在 TSS(转录起始位点)上游和 TTS(转录终止位点)下游越高。

### 5 DNase 超敏感位点与转录因子的结合

基因表达调控是调控蛋白与调控序列(*cis*-element)相互作用的结果。早在 20 世纪 70 年代, 研究者们就发现蛋白质在与 DNA 结合后, 能保护结合部位不会被 DNase 降解, 而通过确定脱氧核糖核酸酶降解后留下的 DNA 片段就能确定蛋白与 DNA 结合的区域, 这个方法称为脱氧核糖核酸酶足迹法(DNase footprint)<sup>[29, 30]</sup>。而且具有行使调控功能的 DNA 序列或与调控蛋白的结合区域, 与 DNase 超敏感位点重叠<sup>[4, 31]</sup>。基于以上分析, 可以进一步通过收集 DNase 超敏感位点序列, 通过高通量测序的方法来鉴定具有调控蛋白的结合区域<sup>[12, 14, 32, 33]</sup>。例如, 拟南芥中 *SUPERMAN* 基因(AT3G23130)与花发育调控至关重要基因<sup>[34]</sup>, 该基因上游启动子区域具有 MADS box motif。根据 RNA-Seq 结果发现该基因只在花中表达。同时, 我们的研究鉴定出一个花

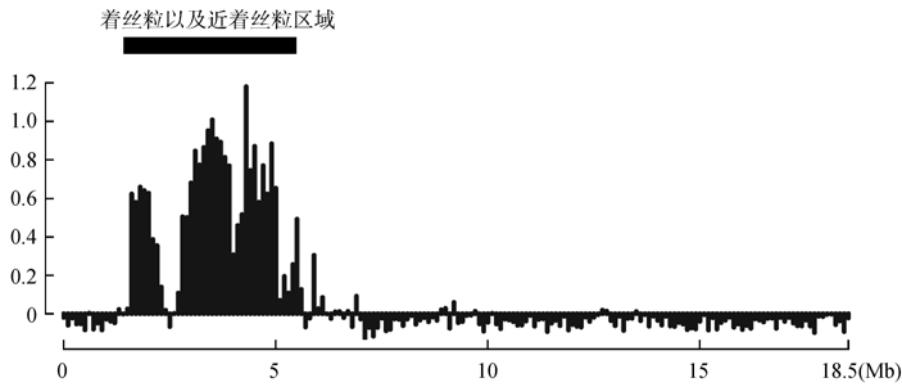


图 5 拟南芥低甲基化突变株 DNase 敏感程度变化<sup>[14]</sup>  
 图为拟南芥低甲基化突变株 *ddm1* 与野生型拟南芥 DNase I 敏感程度的比较。图中柱状图高度为 Log<sub>2</sub>(100 kb *ddm1* 突变株标准化 Reads 总数/100 kb 野生型标准化 Reads 总数)。由图可以获知, 原本高甲基化的拟南芥 4 号染色体的着丝粒以及近着丝粒区域在去甲基化后 DNase 的敏感程度相对于野生型拟南芥来说有明显的升高, 这也说明甲基化水平与 DNase 的敏感程度负相关。

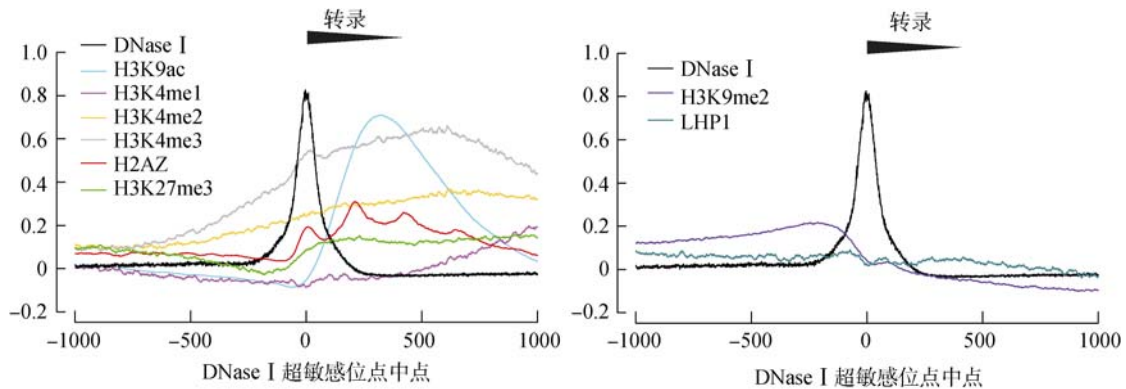


图 6 DNase 超敏感位点与组蛋白及其他表观遗传学标记的关系  
 横坐标 0 点为 DNase 超敏感位点的中点, 纵坐标为 DNase、组蛋白修饰以及其他表观遗传学标记的相对值(标准化 z-score)。图中所有的 DNase I 超敏感位点均来自基因上游 200 bp。

组织特异性的 DNase 超敏感位点, 该位点正好与 MADS box motif 偶联<sup>[4]</sup>。经过精细分析, 发现在花

组织中的 MADS box motif 正好被某些蛋白保护, 从而无法被 DNase 降解(图 7)。因而 AT3G23130 基

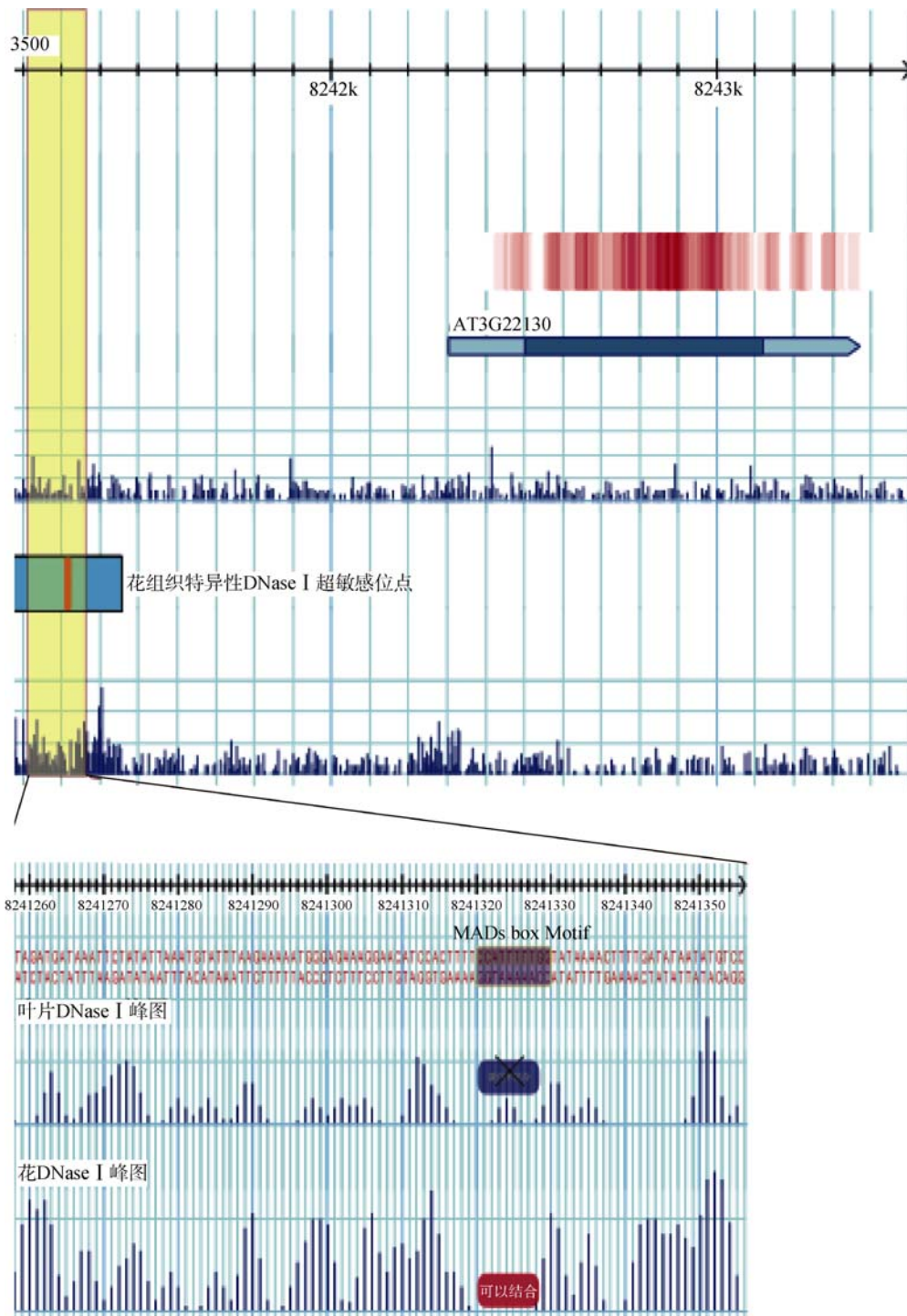


图 7 DNase 足迹法示例<sup>[4]</sup>

图中红色的热点图(Heatmap)为 RNA 表达谱数据, 可以发现 *SUPERMAN* 基因(AT3G23130)为花组织特性表达基因。而在该基因上游, 具有花组织特异性的 DNase 超敏感位点。在花组织特异性的 DNase 超敏感位点内部存在一个花组织特异性的脱氧核糖核酸酶足迹, 该位点正是调控蛋白的结合位点。而在叶片中, 却不存在脱氧核糖核酸酶足迹。

因上游序列具有花组织特异性的 DNase 超敏感位点, 说明该区域在花组织中染色质结构变得比叶片松散, 能更好的与调控蛋白结合从而发挥生物学功能。同时, 在花特异性的 DNase 超敏感位点中, 具有一个不容易被降解的区域, 则说明该区域被特定的蛋白所保护, 该蛋白很有可能是调控 *SUPERMAN* 基因表达的蛋白<sup>[35]</sup>。

通过 DNase 足迹法可以精确的鉴定调控蛋白的结合区域, 优于传统的仅用染色质免疫共沉淀 (Chromatin immunoprecipitation, ChIP) 方法, 因为后者只能确定某一较大的区域是蛋白与 DNA 的结合区域。因此, 将 DNase 足迹法与染色质免疫共沉淀产物测序 (ChIP-Seq) 相结合的办法, 可以非常精确地鉴定调控蛋白的结合位点。

## 6 结语

现有研究表明, DNase 超敏感位点在基因转录调控中起着重要的作用。基因启动子区的 DNase 敏感程度与基因表达正相关, 而 DNase 超敏感位点的甲基化水平总体偏低, 其原因在于该位点的染色质处于开链状态, 而开链状态更有利于调控蛋白(如转录因子等)的结合。通过对 DNase 超敏感位点的研究, 不仅从基因组层面上揭示了染色质结构与基因表达调控有着密切的关系, 而且可以精确地确定单个调控蛋白的结合位点。特别是最近美国国家人类基因组研究所公布了关于人类基因组调节和结构组成的 ENCODE(DNA 元件百科全书, Encyclopedia of DNA Elements)计划的研究成果<sup>[36]</sup>, 大量数据揭示了人类转录因子结合位点、核小体的 DNase 超敏感位点与基因表达的内在联系。随着植物基因组大规模测序技术的发展, 将人类 ENCODE 计划的高通量分析物种基因组的染色质开链状态的研究技术和分析方法, 应用在模式植物和栽培作物的基因组研究之中, 使高通量绘制不同植物物种之间或物种内不同个体之间, 甚至在不同组织或生理状态之间的 DNase 超敏感位点图谱成为可能, 进一步通过 DNase 超敏感位点与其它组学数据的关联整合分析, 可以帮助我们理解植物基因组中染色质结构与基因表达调控的关系, 阐明基因表达调控的分子机理, 无疑将全面推动植物功能基因组学的发展, 为揭示植物发育、进化及其环境适应性的机

制, 应用于植物遗传与育种改良奠定坚实的基础。

## 参考文献(References):

- [1] Keene MA, Corces V, Lowenhaupt K, Elgin SC. DNase I hypersensitive sites in *Drosophila* chromatin occur at the 5' ends of regions of transcription. *Proc Natl Acad Sci USA*, 1981, 78(1): 143–146. DOI
- [2] Wu C. The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature*, 1980, 286(5776): 854–860. DOI
- [3] Mcghee JD, Wood WI, Dolan M, Engel JD, Felsenfeld G. A 200-base pair region at the 5' end of the chicken adult  $\beta$ -globin gene is accessible to nuclease digestion. *Cell*, 1981, 27(1): 45–55. DOI
- [4] Gross DS, Garrard WT. Nuclease hypersensitive sites in chromatin. *Annu Rev Biochem*, 1988, 57(1): 159–197. DOI
- [5] Nedospasov SA, Georgiev GP. Non-random cleavage of SV40 DNA in the compact minichromosome and free in solution by micrococcal nuclease. *Biochem Biophys Res Commun*, 1980, 92(2): 532–539. DOI
- [6] Kodama Y, Nagaya S, Shinmyo A, Kato K. Mapping and characterization of DNase I hypersensitive sites in *Arabidopsis* chromatin. *Plant Cell Physiol*, 2007, 48(3): 459–470. DOI
- [7] Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng ZP, Furey TS, Crawford GE. High-resolution mapping and characterization of open chromatin across the genome. *Cell*, 2008, 132(2): 311–322. DOI
- [8] Fransz P, de Jong H. From nucleosome to chromosome: a dynamic organization of genetic information. *Plant J*, 2011, 66(1): 4–17. DOI
- [9] Segal E, Widom J. What controls nucleosome positions? *Trends Genet*, 2009, 25(8): 335–343. DOI
- [10] Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu TT, Linder-Basso D, Plachetka A, Shanower G, Tolstorukov MY, Luquette LJ, Xi RB, Jung YL, Park RW, Bishop EP, Canfield TK, Sandstrom R, Thurman RE, MacAlpine DM, Stamatoyannopoulos JA, Kellis M, Elgin SCR, Kuroda MI, Pirrotta V, Karpen GH, Park PJ. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature*, 2011, 471(7339): 480–485. DOI
- [11] Sabo PJ, Humbert R, Hawrylycz M, Wallace JC, Dorschner MO, McArthur M, Stamatoyannopoulos JA. Genome-wide identification of DNaseI hypersensitive sites using active chromatin sequence libraries. *Proc Natl Acad Sci USA*, 2004, 101(13): 4537–4542. DOI
- [12] Hesselberth JR, Chen XY, Zhang ZH, Sabo PJ, Sandstrom R, Reynolds AP, Thurman RE, Neph S, Kuehn MS, Noble WS, Fields S, Stamatoyannopoulos JA. Global mapping of

- protein-DNA interactions *in vivo* by digital genomic footprinting. *Nat Methods*, 2009, 6(4): 283–289. [DOI](#)
- [13] Zhang WL, Wu YF, Schnable JC, Zeng ZX, Freeling M, Crawford GE, Jiang JM. High-resolution mapping of open chromatin in the rice genome. *Genome Res*, 2012, 22(1): 151–162. [DOI](#)
- [14] Zhang WL, Zhang T, Wu YF, Jiang JM. Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *Plant Cell*, 2012, 24(7): 2719–2731. [DOI](#)
- [15] Boyle AP, Guinney J, Crawford GE, Furey TS. F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics*, 2008, 24(21): 2537–2538. [DOI](#)
- [16] Iida K, Kawaguchi S, Kobayashi N, Yoshida Y, Ishii M, Harada E, Hanada K, Matsui A, Okamoto M, Ishida J, Tanaka M, Morosawa T, Toyoda T. ARTADE2DB: Improved statistical inferences for *Arabidopsis* gene functions and structure predictions by dynamic structure-based dynamic expression (DSDE) analyses. *Plant Cell Physiol*, 2011, 52(2): 254–264. [DOI](#)
- [17] Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo TA, Okamoto M, Nambara E, Nakajima M, Kawashima M, Satou M, Kim JM, Kobayashi N, Toyoda T, Shinozaki K, Seki M. Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant Cell Physiol*, 2008, 49(8): 1135–1149. [DOI](#)
- [18] Okamoto M, Tatematsu K, Matsui A, Morosawa T, Ishida J, Tanaka M, Endo TA, Mochizuki Y, Toyoda T, Kamiya Y, Shinozaki K, Nambara E, Seki M. Genome-wide analysis of endogenous abscisic acid-mediated transcription in dry and imbibed seeds of *Arabidopsis* using tiling arrays. *Plant J*, 2010, 62(1): 39–51. [DOI](#)
- [19] Jeddleloh JA, Bender J, Richards EJ. The DNA methylation locus *DDM1* is required for maintenance of gene silencing in *Arabidopsis*. *Genes Dev*, 1998, 12(11): 1714–1725. [DOI](#)
- [20] Kakutani T, Munakata K, Richards EJ, Hirochika H. Meiotically and mitotically stable inheritance of DNA hypomethylation induced by *ddm1* mutation of *Arabidopsis thaliana*. *Genetics*, 1999, 151(2): 831–838. [DOI](#)
- [21] Jeddleloh JA, Stokes TL, Richards EJ. Maintenance of genomic methylation requires a SW12/SNF2-like protein. *Nat Genet*, 1999, 22(1): 94–97. [DOI](#)
- [22] Kakutani T, Jeddleloh JA, Richards EJ. Characterization of an *Arabidopsis thaliana* DNA hypomethylation mutant. *Nucleic Acids Res*, 1995, 23(1): 130–137. [DOI](#)
- [23] Vongs A, Kakutani T, Martienssen RA, Richards EJ. *Arabidopsis thaliana* DNA methylation mutants. *Science*, 1993, 260(5116): 1926–1928. [DOI](#)
- [24] Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, Carrington JC, Doerge RW, Colot V, Martienssen R. Role of transposable elements in heterochromatin and epigenetic control. *Nature*, 2004, 430(6998): 471–476. [DOI](#)
- [25] Zhang XY, Germann S, Blus BJ, Khorasanizadeh S, Gaudin V, Jacobsen SE. The *Arabidopsis* LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat Struct Mol Biol*, 2007, 14(9): 869–871. [DOI](#)
- [26] Zhou JL, Wang XF, He K, Charron JB, Elling AA, Deng XW. Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in *Arabidopsis* reveals correlation between multiple histone marks and gene expression. *Plant Mol Biol*, 2010, 72(6): 585–595. [DOI](#)
- [27] Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A. Z and DNA methylation are mutually antagonistic chromatin marks. *Nature*, 2008, 456(7218): 125–129. [DOI](#)
- [28] Zhang XY, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol*, 2009, 10(6): R62. [DOI](#)
- [29] Maniatis T, Ptashne M. Structure of the  $\lambda$  operators. *Nature*, 1973, 246(5429): 133–136. [DOI](#)
- [30] Galas DJ, Schmitz A. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res*, 1978, 5(9): 3157–3170. [DOI](#)
- [31] Cockerill PN. Structure and function of active chromatin and DNase I hypersensitive sites. *Febs J*, 2011, 278(13): 2182–2210. [DOI](#)
- [32] Boyle AP, Song LY, Lee BK, London D, Keefe D, Birney E, Iyer VR, Crawford GE, Furey TS. High-resolution genome-wide *in vivo* footprinting of diverse transcription factors in human cells. *Genome Res*, 2011, 21(3): 456–464. [DOI](#)
- [33] Song LY, Zhang ZC, Gräfer LL, Boyle AP, Giresi PG, Lee BK, Sheffield NC, Gräf S, Huss M, Keefe D, Liu Z, London D, McDaniell RM, Shibata Y, Showers KA, Simon JM, Vales T, Wang TY, Winter D, Zhang ZZ, Clarke ND, Birney E, Iyer VR, Crawford GE, Lieb JD, Furey TS. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res*, 2011, 21(10): 1757–1767. [DOI](#)
- [34] Sakai H, Medrano LJ, Meyerowitz EM. Role of SUPERMAN in maintaining *Arabidopsis* floral whorl boundaries. *Nature*, 1995, 378(6553): 199–203. [DOI](#)
- [35] Riechmann JL, Krizek BA, Meyerowitz EM. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proc Natl Acad Sci USA*, 1996, 93(10): 4793–4798. [DOI](#)
- [36] The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 2012, 489(7414): 57–74. [DOI](#)