Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Plant Biology

Identification of *cis*-regulatory elements by chromatin structure

Zefu Lu[1], William A Ricci[2], Robert J Schmitz[1] and Xiaoyu Zhang[2]

The systematic identification of *cis-regulatory elements* (CREs) in plant genomes is critically important in understanding transcriptional regulation during development and in response to environmental cues. Several genome-wide structure-based methods have been successfully applied to plant genomes in the past few years. Here, we review recent results on the identification and characterization of CREs in multiple plant species and in different biological processes and discuss future applications of chromatin accessibility data to understand the mechanism, function and evolution of transcriptional regulation networks.

Addresses
[1] Department of Genetics, University of Georgia, Athens, GA 30602, USA
[2] Department of Plant Biology, University of Georgia, Athens, GA 30602, USA

Corresponding author: Zhang, Xiaoyu (xiaoyu@uga.edu)

Current Opinion in Plant Biology 2018, **42**:90–94

This review comes from a themed issue on **Genome studies and molecular genetics**

Edited by **Yves Van de Peer** and **J Chris Pires**

For a complete overview see the Issue and the Editorial

Available online 25th April 2018

https://doi.org/10.1016/j.pbi.2018.04.004

Introduction

A key aspect in the understanding of transcriptional regulation in plants is the genome-wide identification and functional characterization of *cis-regulatory elements* (CREs) involved in protein-DNA interactions. Approaches dependent on sequence conservation within short distances from target genes have produced a number of important results. However, there are several limitations: sequence conservation can be restricted to only a few base pairs (bps) involved in protein-DNA interactions, CREs can be located rather far from their target genes (10 000s–100 000s of bps), and sequence conservation alone provides little information regarding the tissue specificity and the functionality (activating versus repressive) of CREs or the potential involvement of epigenetic pathways. To minimize the impact of these limitations, several recent studies aimed at identifying CREs based on chromatin structure and modifications have been successfully carried out in plants. The major findings and potential implications are summarized in this review.

The basic unit of eukaryotic chromosomes is the nucleosome [1], which contains approximately 147 bps of DNA wrapped around a histone octamer (two copies of each of the core histones H2A, H2B, H3, and H4) [2]. Nucleosome positioning and occupancy can have major effects on transcription and other DNA-dependent processes because nucleosomal DNA involved in histone-DNA interaction is often a poor binding substrate for non-histone proteins, including transcription factors (TFs) [3,4]. Consequently, CREs are preferentially located in *accessible chromatin regions* (ACRs) — discrete sites with low nucleosome occupancy due to DNA sequence content, competitive binding between TFs and histones, or through the actions of chromatin remodeling factors [5–7]. ACRs can be identified based on their elevated sensitivity to enzymes such as micrococcal nuclease (MNase), DNase I or the bacterial transposase Tn5.

The coupling of chromatin accessibility assays with massively parallel sequencing in MNase-seq, DNase-seq and Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) represents an important technological development and has enabled the identification of ACRs on a genome-wide scale [8,9••,10–12]. Additionally, integration with genome-wide datasets on TF binding, cytosine methylation, and histone covalent modifications has enabled high-throughput identification of CREs in plant, metazoan, and yeast genomes [13–15,16••].

CREs are enriched in ACRs

Several lines of evidence indicate that CREs primarily reside within ACRs. First, ACRs are highly enriched near the promoter regions of genes, and the degree of accessibility positively correlates with the expression level of associated genes [17••,18••,19••]. In addition, known TF binding motifs are highly enriched in ACRs, particularly within 'footprints' which are small regions within ACRs resistant to DNase I digestion because they are bound by TFs [18••]. Furthermore, intraspecific sequence polymorphisms within ACRs are highly correlated with transcriptional variations and phenotypic differences [20]. For example, although ACRs account for only ~1% of the maize genome, sequence polymorphisms within ACRs can explain 40% of the heritable phenotypic variations across a diverse panel of maize inbred lines [17••]. Similarly, a significantly higher fraction of trait-associated

sequence polymorphisms in *Arabidopsis thaliana* reside within ACRs than other intergenic regions [18••]. However, it should be noted that many sequence polymorphisms within ACRs do not appear to affect the expression of nearby genes, underlying the notion that only small fractions within individual ACRs are likely CREs [21,22]. This is further supported by the analysis of conserved non-coding sequences and transcriptional networks within and between related plant species [23,24].

Interspecific comparative analysis of ACRs

Rapid and genome-wide identification of ACRs also makes it possible to explore the variations in CREs across species. A recent study of ACRs in *A. thaliana*, *Solanum lycopersicum*, *Medicago truncatula*, and *Oryza sativa* revealed that chromatin accessibility associated with 373 syntenic orthologs or 52 expressologs was not frequently conserved across species [19••]. Interestingly, ACRs associated with genes under the control of specific combinations of TFs (e.g. ELONGATED HYPOCOTYL 5 (HY5), ABSCISIC ACID RESPONSIVE ELEMENTS-BINDING FACTOR 3 (ABF3), C-REPEAT/DRE BINDING FACTOR 2 (CBF2), and MYB DOMAIN PROTEIN 77 (MYB77)) appear to be more conserved over evolutionary time scales [19••]. These results suggest that the regulation of these genes is dependent on a set of conserved TF-binding sites within ACRs rather than the entire ACR sequences.

Chromatin accessibility dynamics during plant development and in response to environmental stimuli

Differential TF-DNA interactions play a major role in determining distinct transcription profiles in different developmental stages and in response to external stimuli. As such, chromatin accessibilities at CREs associated with expressed genes are expected to be dynamic [25]. DNase-seq experiments on chromatin isolated from *A. thaliana* root hair and non-root hair cells using the INTACT method ('Isolation of Nuclei TAgged in specific Cell Types') indeed identified cell-type-specific ACRs [26]. Interestingly, despite the differences in cell identity, the overall ACR landscapes in root hair and non-root hair cells are remarkably similar. These results are in stark contrast to the differences in chromatin accessibility observed in animal cell types. Among other possibilities, the relatively stable accessibility observed in *A. thaliana* could be due to the fact that spatial clustering of CREs is common in *A. thaliana*, and therefore activating and repressive TFs can bind different CREs within the same ACR. Future studies in plants with larger genomes should determine whether longer intergenic space leads to the spatial separation of CREs and larger differences in CRE chromatin accessibility through development.

In addition to cell differentiation, ACRs can also be dynamic in response to external stimuli. For example, a set of 734 ACRs showed strong quantitative changes throughout photomorphogenesis [18••]. The dynamic ACRs could be clustered into five temporal patterns, and each cluster was enriched for footprints of specific TFs, suggesting that the distinct temporal patterns of ACR dynamics are driven by the binding of different TFs. The identification of specific TF footprints for each cluster may provide a lead for determining the causative trans-acting factors that drive the dynamic accessibility. Interestingly, the densities of binding sites for TFs involved in light response are similar in light-induced ACRs and static ACRs, indicating many ACRs are poised for stimuli responses and are accessible even in the absence of the stimuli.

Gene-distal CREs in large plant genomes

The human genome hosts an abundance of gene-distal CREs [13,27,28], many of which form complex interactions with cognate genes by forming dynamic, developmentally labile chromatin loops [27,29]. Some observed forms of interactions include: CREs skipping nearby genes to interact with more distal genes, multiple CREs interacting with a single gene, and the promoters of genes acting as distal CREs for other genes [27]. A combination of chromatin attributes — accessible chromatin, cytosine hypomethylation, enrichment of H3K4me1 H3K27ac or H3K27me3, and absence of H3K4me3 — is used to locate and discern the biological status of gene-distal transcriptional enhancers in the human genome [28].

Comparatively little is known about gene-distal CREs in plants. Results from genetic and transgenic assays suggest that most CREs in the compact genome of *A. thaliana* are located relatively close to their target genes. However, a handful of functionally validated examples demonstrates that gene-distal CREs do indeed exist in plants, and that they show some similarities in chromatin attributes to their mammalian counterparts [30]. The hepta-repeat enhancer in *Zea mays*, which resides 100 kb upstream of its cognate gene *booster1*, is characterized by accessible chromatin, enrichment of acetylated histone H3, and a physical association with *booster1* [31–33]. The Block C enhancer in *A. thaliana* resides 5 kb upstream of its cognate gene *FLOWERING LOCUS T* (*FT*) and is similarly characterized by accessible chromatin, acetylated H3, and a physical association with *FT* [34,35]. Additional examples of gene-distal *CREs*, such as the enhancer of *teosinte branched 1* (*tb1*) [36] and the *VEGETATIVE TO GENERATIVE1 (VGT1)* enhancer of *ZmRap2.7* [37], were first identified as agronomic quantitative trait loci (QTL) located 60-70 kb from their target genes and are implicated in the domestication of *Z. mays*.

An important line of evidence from QTL studies suggests that gene-distal CREs may be widespread in *Z. mays*. Wallace *et al.* [38•] found that SNPs associated with agronomically important traits are distributed throughout

the intergenic space in *Z. mays* and often located tens to hundreds of kb away from the nearest genes. Liu *et al.* [39] identified expression QTL associated with 18 000 *Z. mays* genes, and 30% of the leading expression QTL were intergenic and greater than 100 kb away from the corresponding genes.

Recently published epigenomic datasets revealed that gene-distal ACRs are indeed abundant in plant genomes [9••,10,16••,17••,18••,19••,40,41••]. The proportion of gene-distal ACRs scales with genome size and the amount of intergenic space [42]. In *A. thaliana* (135 Mb genome), approximately 80% of ACRs reside within 2000 bp of genes, whereas in *S. lycopersicum* (~980 Mb genome), only 50% of ACRs reside within 2000 bp of genes [19••]. In *Z. mays* (2.3 Gb genome), a significant proportion of ACRs are found tens to hundreds of kb away from genes [17••]. Collectively, these results suggest that the expansion of intergenic space by mechanisms such as transposon insertions may contribute to the localization of ACRs away from their target genes.

The generation of plant epigenomic datasets will be valuable for the annotation of functional genomic elements [43]. Epigenomic data have been used in combination with chromatin accessibility to identify transcriptional enhancers in *Z. mays* [41••]. Approximately 1500 putative intergenic enhancers were identified on the basis of chromatin attributes typically used for the identification of mammalian enhancers. Like mammalian enhancers, the putative *Z. mays* gene-distal enhancers show accessible chromatin, cytosine hypomethylation, and enrichment of H3K9ac, potentially in correspondence with the expression status of adjacent genes. In contrast to mammalian enhancers, H3K4me1 appears to be absent from the gene-distal CREs [41••]. Thus, H3K4me1 is unlikely to serve as a marker for CREs in plants. These results were further corroborated by a recently published Hi-C dataset that demonstrated enrichment of accessible chromatin and acetylated H3 at the boundaries of chromatin loops in the maize genome [44].

Conclusion and future perspectives

The recent applications of rapid and efficient genome-wide ACR profiling assays have produced an unprecedented amount of data for the identification of CREs and the understanding of how TF networks control gene expression during development and in response to environmental stimuli. However, a major limitation to these assays is the lack of cell-type resolution. ACRs from tissues containing mixed cell types provide little information regarding which specific cell types the CREs function in, and ACRs that are specific to uncommon cell types may produce insufficient signals for their identification. Although this can be overcome by fluorescence activated cell/nuclei sorting or INTACT, both assays require the generation of transgenic plants, which is time consuming and impractical in many important crops species. It is exciting to anticipate that the maturation of single-cell technologies might facilitate ACR identification in specific cell types. A second technical aspect that requires further optimization is the experimental and computational methods for defining DNA footprints that result from TF-binding within ACRs, which will facilitate the identification of bound TFs as well as the elucidation of the dynamics of TF-DNA interactions.

The ability to rapidly identify ACRs also makes it possible to study CRE variations in DNA sequence and cell-type specificity during domestication and evolution. Direct sequence comparison between ACRs associated with orthologous genes in different species is not always informative as only short stretches of DNA sequences within each ACR are expected to be functionally relevant. Additionally, the history of small-scale and large-scale gene/genome duplication events in different plant species, together with different models of gene regulatory evolution, interferes with the straightforward analysis and interpretation of ACRs and CREs across species. However, better identification of CREs should be achieved by improved footprinting assays as well as other data such as DNA affinity purification sequencing (DAP-seq) [45]. The results should not only assist in the understanding of CRE variation as the basis for phenotypic evolution, but also provide important information for bioengineering by modifying CRE sequences to improve agronomically important traits.

The widespread presence of distal CREs revealed by ACR profiling assays also raises a number of mechanistic questions. The long distance between distal CREs and their target genes necessitates 3D interactions. Recent Hi-C studies on *Z. mays*, *S. lycopersicum*, *Sorghum bicolor*, *Setaria italica*, and *O. sativa* demonstrated extensive intra-chromosomal interactions throughout the chromosome arms [44,46]. The large-genome plants (e.g. *Z. mays* and *S. lycopersicum*) had thousands of chromatin loops, the majority of which formed between islands of genic regions separated by repressive heterochromatin. A fraction of the interactions are formed between intergenic and genic regions. However, the limited resolution of the genome-wide dataset was inadequate to detect specific CRE-gene interactions. Future Hi-C studies with enhanced resolution (such as HiChIP [47] and ChIAPET [48]) will be necessary to elucidate specific CRE-gene interactions. Finally, most distal ACRs are associated with highly localized chromatin signatures, including DNA hypomethylation and certain histone modifications. Future studies of the establishment, maintenance and removal of these chromatin signatures should provide important information on how TFs interact with chromatin modification pathways to control gene expression.

Acknowledgements

This work was supported by the National Science Foundation (NSF) grants IOS-1546867 to RJS and XZ, and IOS-1238142 to XZ.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Kornberg RD: **Chromatin structure: a repeating unit of histones and DNA**. *Science* 1974, **184**:868-871.

2. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ: **Crystal structure of the nucleosome core particle at 2.8 A resolution**. *Nature* 1997, **389**:251-260.

3. Radman-Livaja M, Rando OJ: **Nucleosome positioning: how is it established, and why does it matter?** *Dev Biol* 2010, **339**:258-266.

4. Zaret KS, Carroll JS: **Pioneer transcription factors: establishing competence for gene expression**. *Genes Dev* 2011, **25**:2227-2241.

5. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B *et al.*: **The accessible chromatin landscape of the human genome**. *Nature* 2012, **489**:75-82.

6. Consortium EP, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M: **An integrated encyclopedia of DNA elements in the human genome**. *Nature* 2012, **489**:57-74.

7. He B, Chen C, Teng L, Tan K: **Global view of enhancer-promoter interactome in human cells**. *Proc Natl Acad Sci U S A* 2014, **111**: E2191-E2199.

8. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ: **ATAC-seq: a method for assaying chromatin accessibility genome-wide**. *Curr Protoc Mol Biol* 2015, **109**:21-29 21 29.

9. •• Lu Z, Hofmeister BT, Vollmers C, DuBois RM, Schmitz RJ: **Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes**. *Nucleic Acids Res* 2017, **45**:e41.
The first ATAC-seq study in plants. The ATAC method was optimized for plants and used to identify ACRs in *Arabidopsis* and maize.

10. Bajic M, Maher KA, Deal RB: **Identification of open chromatin regions in plant genomes using ATAC-Seq**. *Methods Mol Biol* 2018, **1675**:183-201.

11. Zhang W, Zhang T, Wu Y, Jiang J: **Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis***. *Plant Cell* 2012, **24**:2719-2731.

12. Zhang W, Wu Y, Schnable JC, Zeng Z, Freeling M, Crawford GE, Jiang J: **High-resolution mapping of open chromatin in the rice genome**. *Genome Res* 2012, **22**:151-162.

13. Liu YW, Yu S, Dhiman VK, Brunetti T, Eckart H, White KP: **Functional assessment of human enhancer activities using whole-genome STARR-sequencing**. *Genome Biol* 2017, **18**.

14. Yao L, Berman BP, Farnham PJ: **Demystifying the secret mission of enhancers: linking distal regulatory elements to target genes**. *Crit Rev Biochem Mol Biol* 2015, **50**:550-573.

15. Narlikar L, Ovcharenko I: **Identifying regulatory elements in eukaryotic genomes**. *Brief Funct Genomic Proteomic* 2009, **8**:215-230.

16. •• Zhu B, Zhang W, Zhang T, Liu B, Jiang J: **Genome-wide prediction and validation of intergenic enhancers in *Arabidopsis* using open chromatin signatures**. *Plant Cell* 2015, **27**:2415-2426.
Identification of putative enhancers by DNase-seq and functional validation using reporter assays.

17. •• Rodgers-Melnick E, Vera DL, Bass HW, Buckler ES: **Open chromatin reveals the functional maize genome**. *Proc Natl Acad Sci U S A* 2016, **113**:E3177-E3184.
Differential MNase-seq study in maize showing that 1% of the genome in ACRs can explain 40% of the phenotypic variations.

18. •• Sullivan AM, Arsovski AA, Lempe J, Bubb KL, Weirauch MT, Sabo PJ, Sandstrom R, Thurman RE, Neph S, Reynolds AP *et al.*: **Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana***. *Cell Rep* 2014, **8**:2015-2030.
A detailed study of ACRs in *Arabidopsis* using DNase-seq, in multiple cell types, during photomorphogenesis and in response to heat stress.

19. •• Maher KA, Bajic M, Kajala K, Reynoso M, Pauluzzi G, West DA, Zumstein K, Woodhouse M, Bubb K, Dorrity MW *et al.*: **Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules**. *Plant Cell* 2018, **30**:15-36.
A comparison study of ACRs and regulatory networks in several plant species using ATAC-seq.

20. Wittkopp PJ, Kalay G: **Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence**. *Nat Rev Genet* 2011, **13**:59-69.

21. Alexandre CM, Urton JR, Jean-Baptiste K, Huddleston JL, Dorrity MW, Cuperus JT, Sullivan AM, Bemm F, Jolic D, Arsovski AA *et al.*: **Complex relationships between chromatin accessibility, sequence divergence, and gene expression in *A. thaliana***. *Mol Biol Evol* 2017.

22. Korkuc P, Schippers JH, Walther D: **Characterization and identification of cis-regulatory elements in *Arabidopsis* based on single-nucleotide polymorphism information**. *Plant Physiol* 2014, **164**:181-200.

23. Haudry A, Platts AE, Vello E, Hoen DR, Leclercq M, Williamson RJ, Forczek E, Joly-Lopez Z, Steffen JG, Hazzouri KM *et al.*: **An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions**. *Nat Genet* 2013, **45**:891-898.

24. Van de Velde J, Heyndrickx KS, Vandepoele K: **Inference of transcriptional networks in *Arabidopsis* through conserved noncoding sequence analysis**. *Plant Cell* 2014, **26**:2729-2730.

25. Tsompana M, Buck MJ: **Chromatin accessibility: a window into the genome**. *Epigenet Chromat* 2014, **7**:33.

26. Sijacic P, Bajic M, McKinney EC, Meagher RB, Deal RB: **Chromatin accessibility changes between *Arabidopsis* stem cells and mesophyll cells illuminate cell type-specific transcription factor networks**. *Plant J* 2018.

27. Mumbach MR, Satpathy AT, Boyle EA, Dai C, Gowen BG, Cho SW, Nguyen ML, Rubin AJ, Granja JM, Kazane KR *et al.*: **Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements**. *Nat Genet* 2017, **49**:1602-1610.

28. Shlyueva D, Stampfel G, Stark A: **Transcriptional enhancers: from properties to genome-wide predictions**. *Nat Rev Genet* 2014, **15**:272-286.

29. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping**. *Cell* 2014, **159**:1665-1670.

30. Weber B, Zicola J, Oka R, Stam M: **Plant enhancers: a call for discovery**. *Trends Plant Sci* 2016, **21**:974-987.

31. Stam M, Belele C, Dorweiler JE, Chandler VL: **Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation**. *Genes Dev* 2002, **16**:1906-1910.

32. Louwers M, Bader R, Haring M, van Driel R, de Laat W, Stam M: **Tissue- and expression level-specific chromatin looping at maize b1 epialleles**. *Plant Cell* 2009, **21**:832-842.

33. Belele CL, Sidorenko L, Stam M, Bader R, Arteaga-Vazquez MA, Chandler VL: **Specific tandem repeats are sufficient for paramutation-induced trans-generational silencing**. *PLoS Genet* 2013, **9**:e1003773.

34. Liu L, Adrian J, Pankin A, Hu J, Dong X, von Korff M, Turck F: **Induced and natural variation of promoter length modulates the photoperiodic response of FLOWERING LOCUS T**. *Nat Commun* 2014, **5**:4558.

35. Cao S, Kumimoto RW, Gnesutta N, Calogero AM, Mantovani R, Holt BF: **A distal CCAAT/NUCLEAR FACTOR Y complex promotes chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in *Arabidopsis***. *Plant Cell* 2014, **26**:1009-1010.

36. Vann L, Kono T, Pyhajarvi T, Hufford MB, Ross-Ibarra J: **Natural variation in teosinte at the domestication locus teosinte branched1 (tb1)**. *PeerJ* 2015, **3**.

37. Salvi S, Sponza G, Morgante M, Tomes D, Niu X, Fengler KA, Meeley R, Ananiev EV, Svitashev S, Bruggemann E *et al.*: **Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize**. *Proc Natl Acad Sci U S A* 2007, **104**:11376-11380.

38. • Wallace JG, Bradbury PJ, Zhang N, Gibon Y, Stitt M, Buckler ES: **Association mapping across numerous traits reveals patterns of functional variation in maize**. *PLoS Genet* 2014, **10**:e1004845.
Association mapping study in maize revealing the presence of numerous regulationary elements in intergenic regions.

39. Liu H, Luo X, Niu L, Xiao Y, Chen L, Liu J, Wang X, Jin M, Li W, Zhang Q *et al.*: **Distant eQTLs and non-coding sequences play critical roles in regulating gene expression and quantitative trait variation in maize**. *Mol Plant* 2017, **10**:414-426.

40. Silva E, Batista S, Caetano L, Cerejeira MJ, Chaves M, Jacobsen SE: **Integrated approach for the quality assessment of freshwater resources in a vineyard area (South Portugal)**. *Environ Monit Assess* 2011, **176**:331-341.

41. •• Oka R, Zicola J, Weber B, Anderson SN, Hodgman C, Gent JI, Wesselink JJ, Springer NM, Hoefsloot HCJ, Turck F *et al.*: **Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize**. *Genome Biol* 2017, **18**:137.
Identification of intergenic transcriptional enhancers in maize based on MNase hypersensitivity and the presence of histone 3 lysine 9 acetylation.

42. Mei W, Stetter MG, Gates DJ, Stitzer MC, Ross-Ibarra J: **Adaptation in plant genomes: bigger is different**. *Am J Bot* 2018, **105**:16-19.

43. Lane AK, Niederhuth CE, Ji L, Schmitz RJ: **pENCODE: a plant encyclopedia of DNA elements**. *Annu Rev Genet* 2014, **48**:49-70.

44. Dong P, Tu X, Chu PY, Lu P, Zhu N, Grierson D, Du B, Li P, Zhong S: **3D chromatin architecture of large plant genomes determined by local A/B compartments**. *Mol Plant* 2017, **10**:1497-1509.

45. O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR: **Cistrome and epicistrome features shape the regulatory DNA landscape**. *Cell* 2016, **166**:1598.

46. Liu C, Cheng YJ, Wang JW, Weigel D: **Prominent topologically associated domains differentiate global chromatin packing in rice from *Arabidopsis***. *Nat Plants* 2017, **3**:742-748.

47. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, Chang HY: **HiChIP: efficient and sensitive analysis of protein-directed genome architecture**. *Nat Methods* 2016, **13**:919-922.

48. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH *et al.*: **An oestrogen-receptor-alpha-bound human chromatin interactome**. *Nature* 2009, **462**:58-64.