ChIP-seq: A Powerful Tool for Studying Protein–DNA Interactions in Plants

Xifeng Chen[1†], Vijai Bhadauria[2†] and Bojun Ma[1*]

[1]College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua, China.
[2]Crop Development Centre and Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada.

*Correspondence: mbj@zjnu.cn
†These authors contributed equally

https://doi.org/10.21775/cimb.027.171

Abstract

DNA-binding proteins, including transcription factors, epigenetic and chromatin modifiers, control gene expression in plants. To map the binding sites of DNA-binding proteins in the genome is crucial for decoding gene regulatory networks. Chromatin immunoprecipitation (ChIP) followed by high-throughput DNA sequencing (ChIP-seq) is a widely used approach to identify the DNA regions bound by a specific protein *in vivo*. The information generated from ChIP-seq has tremendously advanced our understanding of the mechanisms of transcription factors, cofactors and histone modifications in regulating gene expression. In this review, we reviewed the recent research advances of ChIP-seq in plants, including a description of the ChIP-seq workflow and its various applications in plants, and in addition, provided a perspective for the potential advances of ChIP-seq.

Introduction

Genome-wide mapping of protein–DNA interactions and epigenetic markers is essential for full understanding of transcriptional regulation. Chromatin immunoprecipitation (ChIP) is an effective method for isolating the DNA fragments binding by a specific protein *in vivo* (Solomon *et al.*, 1988). This technology can co-precipitate proteins of interest with DNAs fragments, which are physically associated with the proteins. Previously, ChIP assay was often followed by a hybridization of the DNA fragments isolated to microarray (called ChIP-chip), in order to identify the sequences of DNA fragments in genome wide. However, along with the rapid development of DNA-sequencing technology, the DNA fragments obtained from ChIP have been sequenced directly by high-throughput sequencing (called ChIP-seq). Due to the greater resolution, the coverage and specificity of ChIP-seq are significantly improved compared with ChIP-chip (Luo and Lam, 2014). Thus, ChIP-seq has become a popular method for profiling chromatin modifications and transcription factor (TF)-binding sites in eukaryote genomes. In this review, we describe the workflow of ChIP-seq optimized for plants, summarize the applications of ChIP-seq in plants, and provide perspective regarding its potential advances.

ChIP-seq workflow for plants

ChIP assay

The protocol of ChIP assay for model plants (rice and *Arabidopsis*) was developed by Zhu *et al.* (2012). First, the DNA-binding protein is crosslinked to DNA *in vivo* by treating cells with formaldehyde. Next, the cells are lysed and chromatin is sheared by sonication into small fragments followed by immunoprecipitation of protein–DNA complexes using the specific protein antibodies. To

ensure high-quality data, the choices of antibody and negative control are critical. Many antibodies are commercially available against the epitope tags, such as GFP, MYC or FLAG, which are usually fused to interest proteins in transgenic plants (Table 10.1). The implementation of epitope tags removes the muddled process in the synthesis of a certain antibody, and non-transgenic plants or plants only expressing the tags are the perfect negative controls. Certainly, ChIP can also be performed using a specific antibody against the interest protein itself in wild-type plants, which avoids the need to generate transgenic plants with complex genetic backgrounds. However, it is difficult to design a perfect negative control and thus increases the chance of false positives. The last step is reversing the crosslinks to release the DNA fragments from proteins followed by purification of the released DNA fragments. The quality of purified DNA can be tested by quantitative real-time PCR (ChIP-qPCR) to determine the enrichment of genomic regions of interest.

Sequencing and data analysis

For high-throughput sequencing, a library of the purified DNA fragments are created and sequenced by multiple platforms (Chaitankar *et al.*, 2016). Presently, Illumina platforms (Genome Analyzer II or HiSeq X) are mostly used for ChIP-seq (Table 10.1). After sequencing, converting the raw ChIP-Seq data into biological insight requires sequential data analysis (Fig. 10.1). Generally, the raw reads from sequencing are firstly run quality control by program Fastqc (www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) or Fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) to remove the adapt and other contaminative sequences, and then mapped to the reference genome using a short-read mapping program, such as Bowtie (http://bowtie-bio.sourceforge.net/index.shtml). Following successful mapping reads to genome, the next task is genome-wide identification of the putative binding sites of the interest protein, or called peaks calling. This process can be performed by various algorithms (Chaitankar *et al.*, 2016), and the program MACS (http://liulab.dfci.harvard.edu/MACS/), QuEST (http://mendel.stanford.edu/sidowlab/downloads/quest/) and CisGenome (www.biostat.jhsph.edu/~hji/cisgenome/) are mostly used for plants (Table 10.1). Lastly, the discovery of binding sequence motifs can be performed by programs MEME, DREME or MEME-ChIP (http://meme-suite.org/).

Applications of ChIP-seq in plants

Genome-wide identification of TF's binding sites and target genes

Regulation of gene expression is very complex in plant cells. The binding of TFs to specific DNA target sequences (or *cis*-regulatory elements) is the fundamental basis of gene regulatory networks. ChIP-seq has enabled precise determination of genome-wide distribution of TF binding sites *in vivo*. This has been illustrated by recent studies of many plant TFs involved in regulation of development, circadian clock, flowering time, hormone signalling, abiotic or biotic stress in different tissues or at different developmental stages of both monocotyledonous and dicotyledonous plants (Table 10.1).

The antibodies mostly used to immunoprecipitate the DNA-binding protein are those commercially available ones raised against the epitomic tags such as GFP, HA, FlAG or MYC (Table 10.1). Accordingly, the transgenic plants that ectopically express tagged protein of interest are used for ChIP-seq. This is predominantly used for the ChIP-seq analysis in *Arabidopsis* because of its high efficient transformation system. For the plant species that are recalcitrant to transformation, the specific monoclonal or polyclonal antibodies are often used for ChIP-seq (Table 10.1).

Illumina GA II or HiSeq X is the mostly used deep sequencing platform for ChIP-seq (Table 10.1). After mapping reads to the reference genomes, the algorithm to identify the peaks enriched with reads in genomes were mostly performed by the program MACS, or by program QuEST, Cisgenome, ERANGE in few cases (Table 10.1). The number of peaks identified varies from hundreds to tens of thousands according to different interest proteins (Table 10.1). According to the genomic features, location of the peaks can be categorized into upstream sequence from transcription start sites (TSS), 5′ UTR, 3′ UTR, exon or intron. The regions distributed near TSS are promoter's site of genes, which are consider to be the targets for

the TFs. Moreover, the expression of these putative target genes can be further validated by RNA-seq, microarray or qPCR (Table 10.1). Searching the common motifs in binding sites from ChIP-seq data likely leads to identify the cis regulatory elements of TFs, such as G-box (CACGTG) for PRRs and PIFs, EE (AAATATCT) for CCA1 (Table 10.1).

Genome-wide discovery of histone marks

Histone marks (e.g. methylation and acetylation) play a substantial role in regulating gene expression throughout plant growth and development as well as in responses to different environmental stresses. ChIP-seq can perform genome-wide evaluation of the distribution of modified histones *in vivo*. Du *et al.* (2013) employed ChIP-seq to explore the di- and tri-methylation of H3K4 (H3K4me2 and H3K4me3) and acylation of H3K9 and H3K27 (H3K9ac and H3K27ac) in rice seedlings. They discovered that the histone marks were enriched around the TSSs. Using ChIP-seq, Brusslan *et al.* (2012) discovered that the enrichment degree of H3K4me3 mark were corresponded to the up- or down-regulation of genes throughout senescence in *Arabidopsis*. Hussey *et al.* (2015) used ChIP-seq to determine the role of the histone modification H3K4me3 in the epigenomic regulation of xylogenesis in *Eucalyptus grandis* tree. They discovered that H3K4me3 enrichment can activate the transcription of genes involved in developing xylem.

Hu *et al.* (2012) characterized the genome-wide distribution and alteration of H3K9ac in maize throughout cold stress with ChIP-seq. Their outcomes suggested that the genic regions showed a substantial complete decrease of H3K9ac due to cold stimulus. In *Arabidopsis*, the H3K4me3 abundance changed robustly located on the genes responding to dehydration stress, and these genes showed broader H3K4me3 distribution profiles from ChIP-seq (van Dijk *et al.*, 2010). Similarly in rice, the alterations in H3K4me3, determined with ChIP-seq, were positively correlated with transcript alterations in response to drought stress, and H3K4me3 modification levels were increased mainly for genes with low expression levels and decreased for genes with high expression levels (Zong *et al.*, 2013). Using ChIP-seq, Ayyappan *et al.* (2015) found H3K9me2 and H4K12ac marks impacted a great number of stress-responsive genes in common bean under rust (*Uromyces appendiculatus*) stress.

Genome-wide analysis of the distribution of RNA polymerase

All eukaryotes have three DNA-dependent RNA polymerases (Pol I, II and III) that are necessary for the transcription. Liu *et al* (2014) identified 6499 ChIP-seq peaks for RNA pol II in poplar, and discovered the distribution of peaks was highly coincident with genic areas and greatly corresponded with transcript abundance. Moreover, plants have two additional polymerases (Pol IV and Pol V) that function in the RNA-directed DNA methylation pathway. Recently, the ChIP-seq protocols for studying of the genome-wide distribution of Pol II (de Lorenzo, 2015) and Pol V (Rowley *et al.*, 2013) were described in *Arabidopsis*. Zhong *et al.* (2012) pinpointed the targets of Pol V in *Arabidopsis* on a genome-wide scale with ChIP-seq, and discovered that Pol V is enriched at promoters and evolutionarily recent transposons. This localization pattern is greatly corresponded with Pol V-dependent DNA methylation and small RNA accumulation.

Improvement for sequencing of repetitive genomic sequences

The heterochromatin or microsatellites with greatly replicated sequences are difficult to sequence by traditional techniques. Centromeric sequences are greatly replicated with tandem repeats, such as satellite DNA and retrotransposons, and greatly changeable in various species, even among tightly related organisms (Smith *et al.*, 2011). Nevertheless, a highly conserved function of the centromere is correlated with its epigenetic characteristics, including the histone H3 variant CENH3 in plants (Birchler *et al.*, 2011). Guo *et al.* (2016) performed ChIP-seq with CENH3 specific antibodies, and identified the ectopic genomic sequences present at the new centromeres of the 4DS chromosome in wheat (*Triticum aestivum*). Similarly, Nagaki *et al.* (2015) isolated the centromeric DNA sequences of sunflower (*Helianthus annuus*) by CENH3-ChIP-seq, and identified two different types of repetitive DNA sequences: the tandem repeat sequences and the long-interspersed-nuclear-element-like sequences. Additionally, Kowar *et al.* (2016) implemented CENH3-ChIP-seq analysis revealed that the centromeres of sugar beet (*Beta*

Table 10.1 ChIP-seq analysis of transcription factors in plants

TF	Species	Biological process	Tissue	Anti-body	Sequencing platform	Algorithm for peaks	Binding regions	Expression analysis	Target genes	Common cis-element	Reference
PRR1	*Arabidopsis*	Circadian clock	Seedling	Anti-GFP	Illumina GA II	ERANGE	867	Microarray	301	CACGTG or AATATCT	Huang *et al*., 2012
PRR5	*Arabidopsis*	Circadian clock	Plant	Anti-GFP	Illumina GA II	MACS	542	Microarray	64	CACGTG	Nakamichi *et al*., 2012
PRR7	*Arabidopsis*	Circadian clock	Seedling	Anti-HA	Illumina	QuEST	73	Microarray	83	CACGTG	Liu *et al*., 2013
PRR9	*Arabidopsis*	Circadian clock	Plant	Anti-HA	Illumina HiSeq 2500	MACS	–	Microarray	150	CACGTG	Liu *et al*., 2016
CCA1	*Arabidopsis*	Circadian clock	Plant	Anti-FLAG	Illumina GA II	MACS	449	Microarray	113	AAAATATCT or AAAAATCT	Kamioka *et al*., 2016
FHY3	*Arabidopsis*	Circadian clock.	Seedling	Anti-FLAG	Illumina GA II	MACS	785	Microarray	71	CACGCGC	Ouyang *et al*., 2011
PCFS4	*Arabidopsis*	Flowering time	Seedling	Anti-POD	Illumina GA II	CisGenome	892	cDNA/EST	7	–	Xing *et al*., 2013
SWP73B	*Arabidopsis*	Development	Seedling	Anti-GFP	Illumina HiSeq 2500	MACS	–	Microarray	377	–	Sacharowski *et al*., 2015
PIF1	*Arabidopsis*	Development	Seedling	Anti-MYC	Illumina HiSeq	MACS	3027	RNA-seq	304	CACGTG or CACATG	Pfeiffer *et al*., 2014
PIF3	*Arabidopsis*	Development	Seedling	Anti-MYC	Illumina GA II	MACS	1064	RNA-seq	136	CACGTG or CACATG	Pfeiffer *et al*., 2014
PIF4	*Arabidopsis*	Development	Seedling	Anti-MYC	Illumina HiSeq	MACS	2710	RNA-seq	198	CACGTG or CACATG	Pfeiffer *et al*., 2014
PIF5	*Arabidopsis*	Development	Seedling	Anti-HA	Illumina	MACS	2058	RNA-seq	236	CACGTG or CACATG	Pfeiffer *et al*., 2014
FHY3	*Arabidopsis*	Development	Inflorescence	Anti-FLAG	Illumina	MACS	1885	RNA-seq	238	CACGCGC	Li *et al*., 2016
WRKY33	*Arabidopsis*	Hormone	Plant	Anti-HA	Illumina HiSeq 2500	QuEST	1680	RNA-seq	318	TTGACT/C	Liu *et al*., 2015c
AL5	*Arabidopsis*	Abiotic stress	Seedling	Anti-AL5	Illumina GA II	MACS	193	qPCR	8	–	Wei *et al*., 2015

ASR5	Rice	Metal stress	Root	Anti-ASR5	Illumina HiSeq 2000	MACS	649	RNA-seq	104	GGCCCA(T/A)	Arenhart *et al*., 2014
MADS58	Rice	Development	Panicle	Anti-MADS58	Illumina	MACS	520	qPCR	13	–	Chen *et al*., 2015a
CCA1	Maize	Circadian clock	Aerial tissue	Anti-CCA1	Illumina HiSeq 2500	MACS	10136	qPCR	3	AAAATA or AAGAAA	Ko *et al*., 2016
O2	Maize	Development	Endosperm	Anti-O2	Illumina GA II	MACS	1686	RNA-seq	35	TGACGTGG	Li *et al*., 2015
NAC	Soybean	Development	Cotyledon	Anti-NAC	Illumina HiSeq 2000	MACS	8246	RNA-seq	72	G[AT]G[AG]G[AG]GA, C[AC]C[GA][TC][GA]CC or TGGGCC	Shamimuzzaman *et al*., 2013
YABBY	Soybean	Development	Cotyledon	Anti-YABBY	Illumina HiSeq 2000	MACS	18064	RNA-seq	96	CC[CA][TC]C[TA][CT]C, GA[AG]AGAAA or CCCCAC	Shamimuzzaman *et al*., 2013
ASR1	Tomato	Drought stress	Leaf	Anti-ASR1	Illumina HiSeq 2000	MACS	225	qPCR	3	(A/T)(A/G)AGCCCA	Ricardi *et al*., 2014
ERF68	Tomato	Defence	Leaf	Anti-HA	Illumina	MACS	2997	qPCR	8	–	Liu and Cheng, 2016
VRN1	Barley	Vernalization	Seedling	Anti-HA	Illumina HiSeq 2000	MACS	514	RNA-seq	37	$CC[A/T]_6GG$ or CATGCATG	Deng *et al*., 2015
ARK1	Populus	Woody growth,	Vascular	Anti-ARK1	Illumina GA II	MACS	14463	RNA-seq	866	AAACCCT/A	Liu *et al*., 2015b
MaASR	Banana	Drought stress	Plant	Anti-MaASR	Illumina	MACS	410	Microarray	10	–	Zhang *et al*., 2015

Algorithm for peaks, the algorithm method used in genome-wide peaks calling; Binding regions, the number of TF's binding regions from peaks-calling analysis; Biological process, the biological process that the TF involved; Common *cis*-element, the Common *cis*-elements which may be bound by the TF; Expression analysis, the methods used for analysing the expression of TF's target genes; Sequencing Platform, the Platform used for high-throughput DNA sequencing; Species, the species that the TF was from; Target genes, the number of TF's target genes validated by expression; TF, transcription factor; Tissue, the tissue that used in ChIP-seq assay.
–, information was not described in the published article.

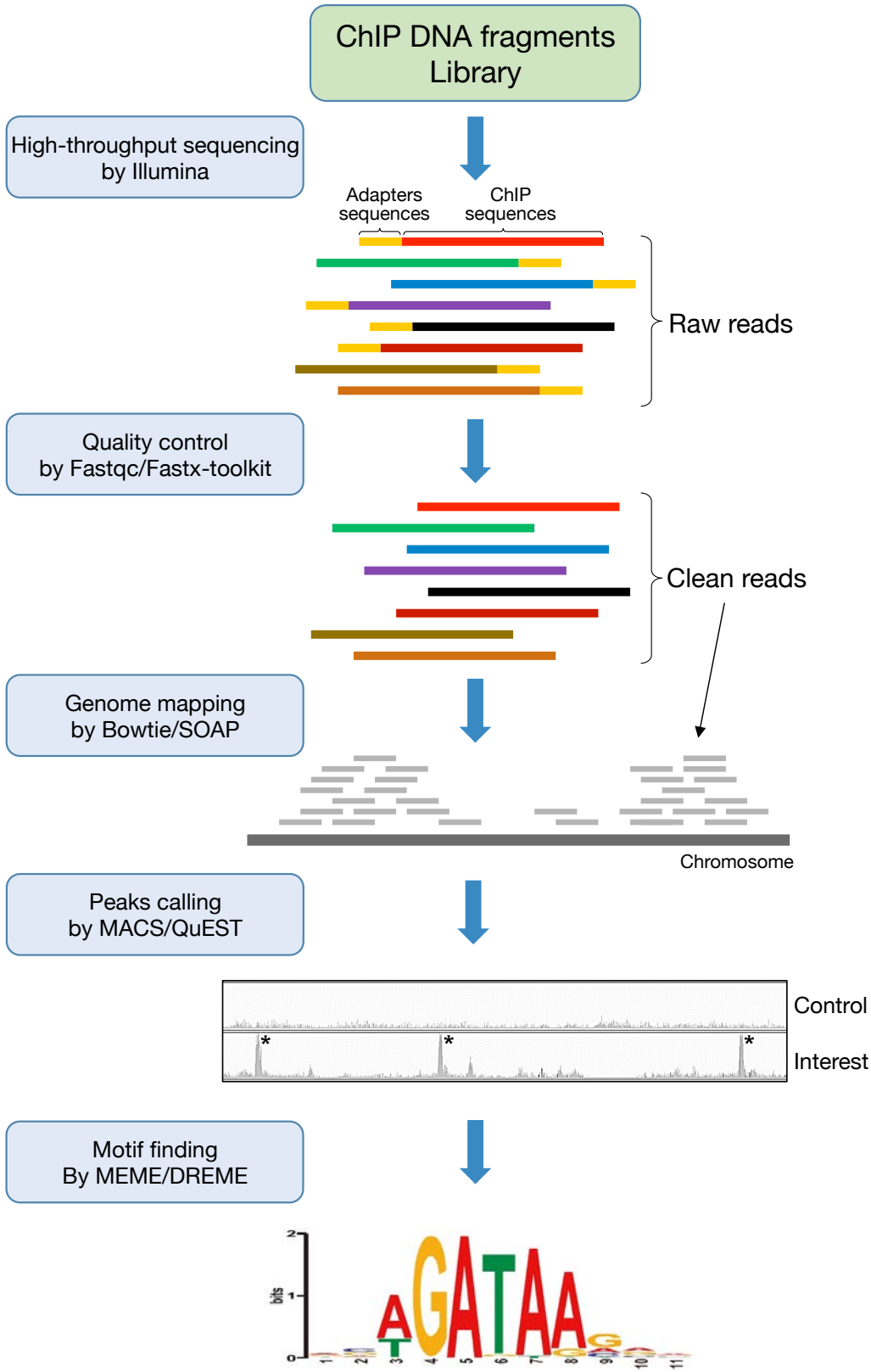

Figure 10.1 Overview of workflow for ChIP-seq data analysis.

vulgaris) containing the satellite pBV and the Ty3-gypsy retrotransposon Beetle7. Consequently, ChIP-seq could be more efficient in sequencing of repeat regions, and the distinctive sequences flanking repeats are helpful in aligning the reads to the genome.

Annotation of putative missing genes in genome

Since the histone modifications have a high concurrence possibility with active transcript loci, ChIP-seq may be implemented in predicting unannotated genes. For example, comparing with the rice genome annotation resources MSU/TIGR (http://rice.plantbiology.msu.edu/) and RAP–DB (http://rapdb.dna.affrc.go.jp/), Du *et al.* (2013) discovered 6,629 RAP–DB genes which were lost annotation in MSU/TIGR. Of these unannotated genes, they found that 471 genes had both H3K4me3 and H3K9ac enrichment inside their TSS regions, and proved the existence of these genes by expression analysis. Though the mRNA-seq could trace missing genes systematically, ChIP-seq is one of the best approaches for characterizing lost genes in organisms whose genomes have yet to be annotated to a large degree.

Perspective

Genome-wide evaluations of epigenomic and transcriptomic profiles provide extensive resources for determining epigenetic regulatory mechanisms. ChIP-seq has been established as a main tool for comprehending these mechanisms. Even with the remarkable advancement in high-throughput sequencing platforms in the recent years, the DNA sequencing capacity increased rapidly with the lower cost. However, the greater challenges for biologists are very complicated bioinformatic processing of immense amount of ChIP-seq data using different programs (Fig. 10.1). This will require development of user-friendly and durable software to evaluate data and closer interaction between experimentalists and bioinformaticians. For example, ENCODE (Encyclopedia of DNA Elements; Landt *et al.*, 2012), offered a group of working procedures and guidelines for ChIP-seq, and address antibody validation, experimental replication, sequencing depth, and data quality assessment experiments. Unfortunately, ENCODE is designed only for human and animals so far. ChIPseek (Chen *et al.*, 2014) is a web-based analysis tool of ChIP data for animals and plants (*Arabidopsis* and rice), and provides a statistical summary of the dataset including a histogram of peak length distribution and of distances to the nearest TSS, and offers charts of genomic areas to get a complete perspective on the dataset. This is appropriate for biologists to analyse ChIP data without needing any programming expertise. Moreover, there are additional powerful ChIP-seq tools for specific plant spices, such as PRI-CAT (Muiño *et al.*, 2011) and *CressInt* (Chen *et al.*, 2015b) for *Arabidopsis*, and PopGenIE (Liu *et al.*, 2015a) for *Populus*. ChIP-seq knowledge is being rapidly added to public repositories that will culminate in the inception of databases to analyse gene expression and protein–protein networks. Such comprehensive analyses facilitate the systematic elucidation of diverse biological activities intricately cooperating in the genome.

Acknowledgements

We are grateful to Dr Jianzhong Liu for critical reading of this manuscript and his valuable comments.

Funding

This work was supported by the National Key Research and Development Program of China (2016YFD0100405); the Ministry of Agriculture of China (2016ZX08009003-001); the National Natural Science Foundation of China (31371593, 31671650); and the Natural Science Foundation of Zhejiang Province (LZ14C130001).

References

Arenhart, R.A., Bai, Y., de Oliveira, L.F., Neto, L.B., Schunemann, M., Maraschin Fdos, S., Mariath, J., Silverio, A., Sachetto-Martins, G., Margis, R., *et al.* (2014). New insights into aluminum tolerance in rice: the ASR5 protein binds the STAR1 promoter and other aluminum-responsive genes. Mol. Plant *7*, 709–721. http://dx.doi.org/10.1093/mp/sst160

Ayyappan, V., Kalavacharla, V., Thimmapuram, J., Bhide, K.P., Sripathi, V.R., Smolinski, T.G., Manoharan, M., Thurston, Y., Todd, A., and Kingham, B. (2015). Genome-wide profiling of histone modifications (H3K9me2 and H4K12ac) and gene expression in Rust (*Uromyces appendiculatus*) inoculated common bean (*Phaseolus vulgaris* L.). PLOS ONE *10*, e0132176. http://dx.doi.org/10.1371/journal.pone.0132176

Birchler, J.A., Gao, Z., Sharma, A., Presting, G.G., and Han, F. (2011). Epigenetic aspects of centromere function in plants. Curr. Opin. Plant Biol. *14*, 217–222. http://dx.doi.org/10.1016/j.pbi.2011.02.004

Brusslan, J.A., Rus Alvarez-Canterbury, A.M., Nair, N.U., Rice, J.C., Hitchler, M.J., and Pellegrini, M. (2012). Genome-wide evaluation of histone methylation changes associated with leaf senescence in *Arabidopsis*. PLOS ONE *7*, e33151. http://dx.doi.org/10.1371/journal.pone.0033151

Chaitankar, V., Karakülah, G., Ratnapriya, R., Giuste, F.O., Brooks, M.J., and Swaroop, A. (2016). Next generation sequencing technology and genomewide data analysis: Perspectives for retinal research. Prog. Retin. Eye Res. *55*, 1–31.

Chen, R., Shen, L.P., Wang, D.H., Wang, F.G., Zeng, H.Y., Chen, Z.S., Peng, Y.B., Lin, Y.N., Tang, X., Deng, M.H., *et al.* (2015a). A gene expression profiling of early rice stamen development that reveals inhibition of photosynthetic genes by OsMADS58. Mol. Plant *8*, 1069–1089. http://dx.doi.org/10.1016/j.molp.2015.02.004

Chen, T.W., Li, H.P., Lee, C.C., Gan, R.C., Huang, P.J., Wu, T.H., Lee, C.Y., Chang, Y.F., and Tang, P. (2014). ChIPseek, a web-based analysis tool for ChIP data. BMC Genomics *15*, 539. http://dx.doi.org/10.1186/1471-2164-15-539

Chen, X., Ernst, K., Soman, F., Borowczak, M., and Weirauch, M.T. (2015b). *CressInt*: a user-friendly web resource for genome-scale exploration of gene regulation in *Arabidopsis thaliana*. Curr. Plant Biol. 3-4, 48–55. http://dx.doi.org/10.1016/j.cpb.2015.09.001

de Lorenzo, L. (2015). Genome-wide analysis of distribution of RNA polymerase II isoforms using ChIP-seq. Methods Mol. Biol. *1255*, 209–221. http://dx.doi.org/10.1007/978-1-4939-2175-1_18

Deng, W., Casao, M.C., Wang, P., Sato, K., Hayes, P.M., Finnegan, E.J., and Trevaskis, B. (2015). Direct links between the vernalization response and other key traits of cereal crops. Nat. Commun. *6*, 5882. http://dx.doi.org/10.1038/ncomms6882

Du, Z., Li, H., Wei, Q., Zhao, X., Wang, C., Zhu, Q., Yi, X., Xu, W., Liu, X.S., Jin, W., *et al.* (2013). Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in *Oryza sativa* L. *Japonica*. Mol. Plant *6*, 1463–1472. http://dx.doi.org/10.1093/mp/sst018

Guo, X., Su, H., Shi, Q., Fu, S., Wang, J., Zhang, X., Hu, Z., and Han, F. (2016). *De novo* centromere formation and centromeric sequence expansion in wheat and its wide hybrids. PLOS Genet. *12*, e1005997. http://dx.doi.org/10.1371/journal.pgen.1005997

Hu, Y., Zhang, L., He, S., Huang, M., Tan, J., Zhao, L., Yan, S., Li, H., Zhou, K., Liang, Y., *et al.* (2012). Cold stress selectively unsilences tandem repeats in heterochromatin associated with accumulation of H3K9ac. Plant Cell Environ. *35*, 2130–2142. http://dx.doi.org/10.1111/j.1365–3040.2012.02541.x

Huang, W., Pérez-García, P., Pokhilko, A., Millar, A.J., Antoshechkin, I., Riechmann, J.L., and Mas, P. (2012). Mapping the core of the *Arabidopsis* circadian clock defines the network structure of the oscillator. Science *336*, 75–79. http://dx.doi.org/10.1126/science.1219075

Hussey, S.G., Mizrachi, E., Groover, A., Berger, D.K., and Myburg, A.A. (2015). Genome-wide mapping of histone H3 lysine 4 trimethylation in *Eucalyptus grandis* developing xylem. BMC Plant Biol. *15*, 117. http://dx.doi.org/10.1186/s12870-015-0499-0

Kamioka, M., Takao, S., Suzuki, T., Taki, K., Higashiyama, T., Kinoshita, T., and Nakamichi, N. (2016). Direct repression of evening genes by CIRCADIAN CLOCK-ASSOCIATED1 in the *Arabidopsis* circadian clock. Plant Cell *28*, 696–711. http://dx.doi.org/10.1105/tpc.15.00737

Ko, D.K., Rohozinski, D., Song, Q., Taylor, S.H., Juenger, T.E., Harmon, F.G., and Chen, Z.J. (2016). Temporal shift of circadian-mediated gene expression and carbon fixation contributes to biomass heterosis in maize hybrids. PLOS Genet. *12*, e1006197. http://dx.doi.org/10.1371/journal.pgen.1006197

Kowar, T., Zakrzewski, F., Macas, J., Koblížková, A., Viehoever, P., Weisshaar, B., and Schmidt, T. (2016). Repeat composition of CenH3-chromatin and H3K9me2-marked heterochromatin in Sugar Beet (*Beta vulgaris*). BMC Plant Biol. *16*, 120. http://dx.doi.org/10.1186/s12870-016-0805-5

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., *et al.* (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. *22*, 1813–1831. http://dx.doi.org/10.1101/gr.136184.111

Li, C., Qiao, Z., Qi, W., Wang, Q., Yuan, Y., Yang, X., Tang, Y., Mei, B., Lv, Y., Zhao, H., *et al.* (2015). Genome-wide characterization of *cis*-acting DNA targets reveals the transcriptional regulatory framework of opaque2 in maize. Plant Cell *27*, 532–545. http://dx.doi.org/10.1105/tpc.114.134858

Li, D., Fu, X., Guo, L., Huang, Z., Li, Y., Liu, Y., He, Z., Cao, X., Ma, X., Zhao, M., *et al.* (2016). *FAR-RED ELONGATED HYPOCOTYL3* activates *SEPALLATA2* but inhibits *CLAVATA3* to regulate meristem determinacy and maintenance in *Arabidopsis*. Proc. Natl. Acad. Sci. U.S.A. *113*, 9375–9380. http://dx.doi.org/10.1073/pnas.1602960113

Liu, A.C., and Cheng, C.P. (2016). Pathogen-induced ERF68 regulates hypersensitive cell death in tomato. Mol. plant Pathol. Online. http://dx.doi.org/10.1111/mpp.12460

Liu, L., Missirian, V., Zinkgraf, M., Groover, A., and Filkov, V. (2014). Evaluation of experimental design and computational parameter choices affecting analyses of ChIP-seq and RNA-seq data in undomesticated poplar trees. BMC Genomics *15* (Suppl. 5), S3. http://dx.doi.org/10.1186/1471-2164-15-S5-S3

Liu, L., Ramsay, T., Zinkgraf, M., Sundell, D., Street, N.R., Filkov, V., and Groover, A. (2015a). A resource for characterizing genome-wide binding and putative target genes of transcription factors expressed during secondary growth and wood formation in *Populus*. Plant J *82*, 887–898. http://dx.doi.org/10.1111/tpj.12850

Liu, L., Zinkgraf, M., Petzold, H.E., Beers, E.P., Filkov, V., and Groover, A. (2015b). The *Populus* ARBORKNOX1 homeodomain transcription factor regulates woody growth through binding to evolutionarily conserved target genes of diverse function. New Phytol. *205*, 682–694. http://dx.doi.org/10.1111/nph.13151

Liu, S., Kracher, B., Ziegler, J., Birkenbihl, R.P., and Somssich, I.E. (2015c). Negative regulation of ABA signaling by

WRKY33 is critical for *Arabidopsis* immunity towards Botrytis cinerea 2100. Elife *4*, e07295. http://dx.doi.org/10.7554/eLife.07295

Liu, T., Carlsson, J., Takeuchi, T., Newton, L., and Farré, E.M. (2013). Direct regulation of abiotic responses by the *Arabidopsis* circadian clock component PRR7. Plant J. *76*, 101–114. http://dx.doi.org/10.1111/tpj.12276

Liu, T.L., Newton, L., Liu, M.J., Shiu, S.H., and Farre, E.M. (2016). A G-box-like motif is necessary for transcriptional regulation by circadian pseudo-response regulators in *Arabidopsis*. Plant Physiol. *170*, 528–539. http://dx.doi.org/10.1104/pp.15.01562

Luo, C., and Lam, E. (2014). Quantitatively profiling genome-wide patterns of histone modifications in *Arabidopsis thaliana* using ChIP-seq. Methods Mol. Biol. *1112*, 177–193. http://dx.doi.org/10.1007/978-1-62703-773-0_12

Muiño, J.M., Hoogstraat, M., van Ham, R.C., and van Dijk, A.D. (2011). PRI-CAT: a web-tool for the analysis, storage and visualization of plant ChIP-seq experiments. Nucleic Acids Res. *39*, W524–7. http://dx.doi.org/10.1093/nar/gkr373

Nagaki, K., Tanaka, K., Yamaji, N., Kobayashi, H., and Murata, M. (2015). Sunflower centromeres consist of a centromere-specific LINE and a chromosome-specific tandem repeat. Front Plant Sci. *6*, 912. http://dx.doi.org/10.3389/fpls.2015.00912

Nakamichi, N., Kiba, T., Kamioka, M., Suzuki, T., Yamashino, T., Higashiyama, T., Sakakibara, H., and Mizuno, T. (2012). Transcriptional repressor PRR5 directly regulates clock-output pathways. Proc. Natl. Acad. Sci. U.S.A. *109*, 17123–17128. http://dx.doi.org/10.1073/pnas.1205156109

Ouyang, X., Li, J., Li, G., Li, B., Chen, B., Shen, H., Huang, X., Mo, X., Wan, X., Lin, R., *et al.* (2011). Genome-wide binding site analysis of FAR-RED ELONGATED HYPOCOTYL3 reveals its novel function in *Arabidopsis* development. Plant Cell *23*, 2514–2535. http://dx.doi.org/10.1105/tpc.111.085126

Pfeiffer, A., Shi, H., Tepperman, J.M., Zhang, Y., and Quail, P.H. (2014). Combinatorial complexity in a transcriptionally centered signaling hub in *Arabidopsis*. Mol. Plant *7*, 1598–1618. http://dx.doi.org/10.1093/mp/ssu087

Ricardi, M.M., Gonzalez, R.M., Zhong, S., Dominguez, P.G., Duffy, T., Turjanski, P.G., Salgado Salter, J.D., Alleva, K., Carrari, F., Giovannoni, J.J., *et al.* (2014). Genome-wide data (ChIP-Seq) enabled identification of cell wall-related and aquaporin genes as targets of tomato ASR1, a drought stress-responsive transcription factor. BMC Plant Biol. *14*, 29. http://dx.doi.org/10.1186/1471–2229–14–29

Rowley, M.J., Böhmdorfer, G., and Wierzbicki, A.T. (2013). Analysis of long non-coding RNAs produced by a specialized RNA polymerase in *Arabidopsis thaliana*. Methods *63*, 160–169. http://dx.doi.org/10.1016/j.ymeth.2013.05.006

Sacharowski, S.P., Gratkowska, D.M., Sarnowska, E.A., Kondrak, P., Jancewicz, I., Porri, A., Bucior, E., Rolicka, A.T., Franzen, R., Kowalczyk, J., *et al.* (2015). SWP73 subunits of *Arabidopsis* SWI/SNF chromatin remodeling complexes play distinct roles in leaf and flower development. Plant Cell *27*, 1889–1906. http://dx.doi.org/10.1105/tpc.15.00233

Shamimuzzaman, M., and Vodkin, L. (2013). Genome-wide identification of binding sites for NAC and YABBY transcription factors and co-regulated genes during soybean seedling development by ChIP-Seq and RNA-Seq. BMC Genomics *14*, 477. http://dx.doi.org/10.1186/1471–2164–14–477

Smith, K.M., Phatale, P.A., Sullivan, C.M., Pomraning, K.R., and Freitag, M. (2011). Heterochromatin is required for normal distribution of *Neurospora crassa* CenH3. Mol. Cell. Biol. *31*, 2528–2542. http://dx.doi.org/10.1128/MCB.01285-10

Solomon, M.J., Larsen, P.L., and Varshavsky, A. (1988). Mapping protein-DNA interactions *in vivo* with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell *53*, 937–947.

van Dijk, K., Ding, Y., Malkaram, S., Riethoven, J.J., Liu, R., Yang, J., Laczko, P., Chen, H., Xia, Y., Ladunga, I., *et al.* (2010). Dynamic changes in genome-wide histone H3 lysine 4 methylation patterns in response to dehydration stress in *Arabidopsis thaliana*. BMC Plant Biol. *10*, 238. http://dx.doi.org/10.1186/1471-2229-10-238

Wei, W., Zhang, Y.Q., Tao, J.J., Chen, H.W., Li, Q.T., Zhang, W.K., Ma, B., Lin, Q., Zhang, J.S., and Chen, S.Y. (2015). The Alfin-like homeodomain finger protein AL5 suppresses multiple negative factors to confer abiotic stress tolerance in *Arabidopsis*. Plant J *81*, 871–883. http://dx.doi.org/10.1111/tpj.12773

Xing, D., Wang, Y., Xu, R., Ye, X., Yang, D., and Li, Q.Q. (2013). The regulatory role of Pcf11-similar-4 (PCFS4) in *Arabidopsis* development by genome-wide physical interactions with target loci. BMC Genomics *14*, 598. http://dx.doi.org/10.1186/1471-2164-14-598

Zhang, L., Hu, W., Wang, Y., Feng, R., Zhang, Y., Liu, J., Jia, C., Miao, H., Zhang, J., Xu, B., *et al.* (2015). The MaASR gene as a crucial component in multiple drought stress response pathways in *Arabidopsis*. Funct. Integr. Genomics. *15*, 247–260. http://dx.doi.org/10.1007/s10142-014-0415-y

Zhong, X., Hale, C.J., Law, J.A., Johnson, L.M., Feng, S., Tu, A., and Jacobsen, S.E. (2012). DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. Nat. Struct. Mol. Biol. *19*, 870–875. http://dx.doi.org/10.1038/nsmb.2354

Zhu, J., Sun, Y., and Wang, Z. (2012). Genome-wide identification of transcription factor-binding sites in plants using chromatin immunoprecipitation followed by microarray (ChIP-chip) or sequencing (ChIP-seq). Methods Mol. Biol. *876*, 173–188. http://dx.doi.org/10.1007/978–1-61779–809–2_14

Zong, W., Zhong, X., You, J., and Xiong, L. (2013). Genome-wide profiling of histone H3K4-tri-methylation and gene expression in rice under drought stress. Plant Mol. Biol. *81*, 175–188. http://dx.doi.org/10.1007/s11103-012-9990-2