

RESEARCH

Open Access



# Inferring regulatory element landscapes and gene regulatory networks from integrated analysis in eight hulless barley varieties under abiotic stress

Qijun Xu<sup>1,2†</sup>, Shunmou Huang<sup>3†</sup>, Ganggang Guo<sup>4</sup>, Chunbao Yang<sup>1,2</sup>, Mu Wang<sup>1,2</sup>, Xingquan Zeng<sup>1,2</sup> and Yulin Wang<sup>1,2\*</sup>

## Abstract

**Background:** The cis-regulatory element became increasingly important for resistance breeding. There were many DNA variations identified by resequencing. To investigate the links between the DNA variations and cis-regulatory element was the fundamental work. DNA variations in cis-regulatory elements caused phenotype variations in general.

**Results:** We used WGBS, ChIP-seq and RNA-seq technology to decipher the regulatory element landscape from eight hulless barley varieties under four kinds of abiotic stresses. We discovered 231,440 lowly methylated regions (LMRs) from the methylome data of eight varieties. The LMRs mainly distributed in the intergenic regions. A total of 97,909 enhancer-gene pairs were identified from the correlation analysis between methylation degree and expression level. A lot of enriched motifs were recognized from the tolerant-specific LMRs. The key transcription factors were screened out and the transcription factor regulatory network was inferred from the enhancer-gene pairs data for drought stress. The NAC transcription factor was predicted to target to TCP, bHLH, bZIP transcription factor genes. We concluded that the H3K27me3 modification regions overlapped with the LMRs more than the H3K4me3. The variation of single nucleotide polymorphism was more abundant in LMRs than the remain regions of the genome.

**Conclusions:** Epigenetic regulation is an important mechanism for organisms to adapt to complex environments. Through the study of DNA methylation and histone modification, we found that many changes had taken place in enhancers and transcription factors in the abiotic stress of hulless barley. For example, transcription factors including NAC may play an important role. This enriched the molecular basis of highland barley stress response.

**Keywords:** BS-seq, ChIP-seq, RNA-seq, Transcription factor, Hulless barley, Lowly methylated regions, TFBS; abiotic stress

## Background

Barley (*Hordeum vulgare L.*) is one of the major cereals grown worldwide and among the oldest of domesticated crops [1]. Hulless barley has been the staple food of Tibetans in China for thousands of years [2]. The genome and variome of hulless barley have been examined in previous studies [3–5]. Thus far, details of the hulless barley regulome have not yet been elucidated.

<sup>†</sup>Qijun Xu and Shunmou Huang contributed equally to this work.

\*Correspondence: Wangyulin8609@126.com

<sup>1</sup> State Key Laboratory of Hulless Barley and Yak Germplasm Resources and Genetic Improvement, Lhasa 850002, China

Full list of author information is available at the end of the article



Projects similar to the Encyclopedia of DNA Elements (ENCODE) have emerged, with the goal of annotating the noncoding, functional genome of a given species by generating spatiotemporal maps of chromatin accessibility, TF occupancy, protein and DNA modifications, and gene expression [6, 7]. The availability of a hullless barley ENCODE would help facilitate hypothesis generation, cross-species comparisons, genome annotation, and understanding of epigenomic functions throughout plant evolution [8].

DNA methylation has been extensively studied as an epigenetic marker in mammals [9] and plants [10]. Epigenetic markers are involved in the activities of genes and transposon elements [11]. Recent studies indicate that DNA methylation can identify transcriptional enhancers, including lowly methylated regions (LMRs) [12]. It has been suggested that LMRs, non-CpG island loci that usually contains transcription factor binding sites (TFBS), act as regulatory elements that define cellular identity. This notion is further supported by the identification of LMRs—the localized CpG-poor distal regulatory regions exhibiting under average methylation of 30%, DNase I hypersensitivity, and the presence of enhancer chromatin marks—that tend to be occupied by cell-type-specific TFs [12].

The modification of histones influences the local structure of chromatin by altering its association with nucleosomal DNA or due to the recruitment of chromatin remodeling complexes [13, 14]. Dynamic shifts in the chromatin structure affect gene regulation by modifying the availability of cis-regulatory elements (CREs), which bind with combinations of TFs in a spatiotemporal manner [15]. Genetic variation of CREs contributing to alter the tissue specificity of regulatory factors can be used to develop improved phenotypes [16]. CREs can reside within the core or proximal promoters of genes and distal to their target genes, such as in enhancer sequences that influence gene expression at long range [15].

Histone modifications are linked to changes in chromatin structure. H3K4me3 marks TSS(Transcription Start Site) and promotes gene expression via recruitment of the NURF(Nucleosome Remodeling Factor) complex for chromatin remodeling [17, 18]. H3K27me3 recruits PRC1, which contributes to chromatin compaction and impedes gene expression [19, 20]. The H3K4me3 and H3K27me3 histone modifications have been broadly utilized in research on chromatin regulation [21, 22].

In our study, lowly methylated regions (LMRs) were detected using methylome data. Using ChIP-seq technology, the different histone modifications (H3K4me3 and H3K27me3) were detected in the LMRs of tolerant and susceptible varieties subjected to stress. The aim of this approach was to characterize 1) the regulatory landscape

of the hullless barley genome; 2) the enhancers that are differentially modified when comparing tolerant and susceptible varieties under stress; 3) the links between the enhancers and target genes; and 4) the gene regulatory network involved in the response to stress.

## Results

### Genome-wide identification of LMRs in hullless barley

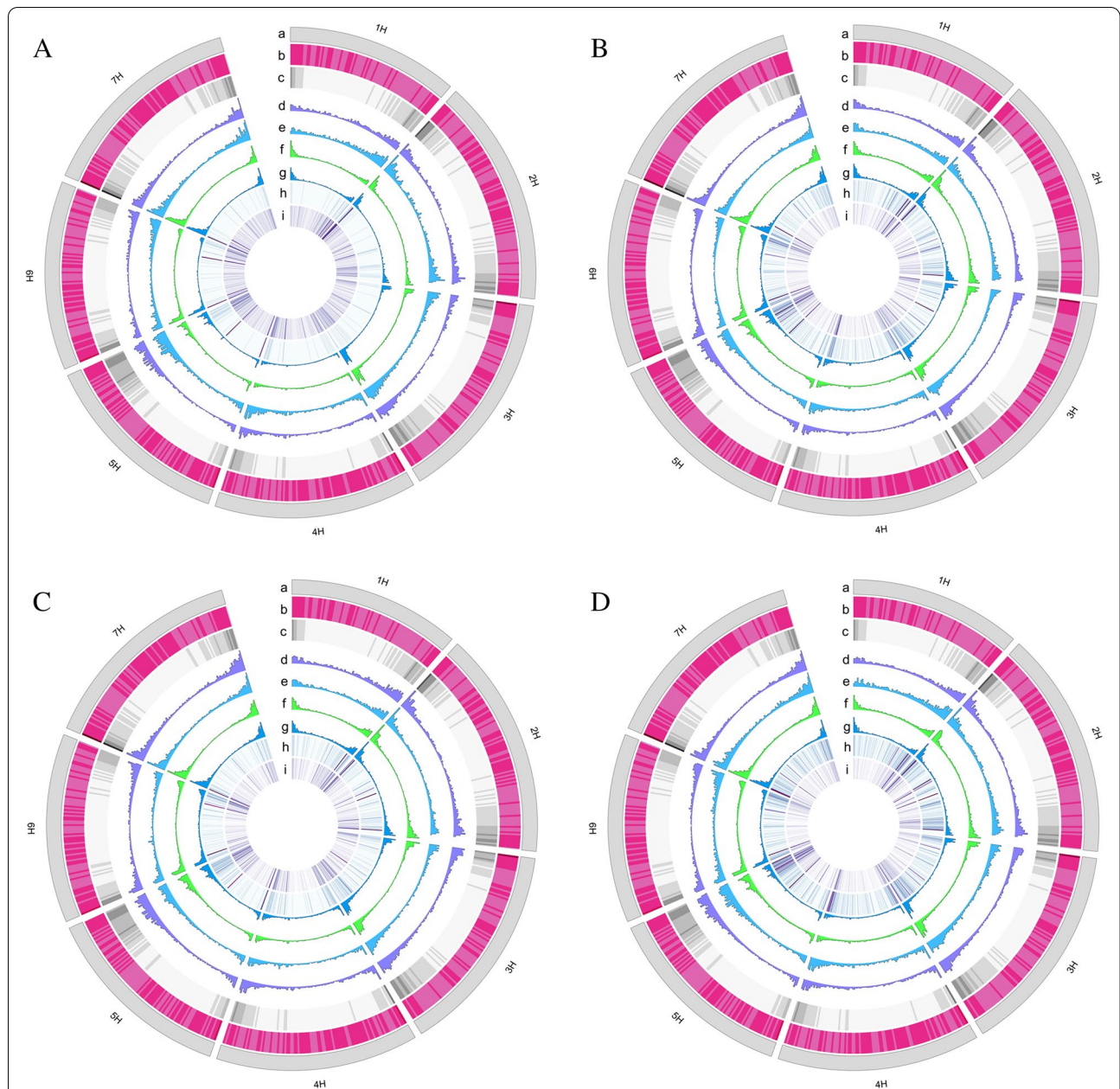
To investigate the relevant regulatory elements in hullless barley under abiotic stress, we examined eight whole-genome bisulfite sequencing (WGBS) libraries (see Methods) for the genome-wide identification of LMRs that are correlated with enhancers. The abiotic stresses included in our study were drought, salinity, cold, and low nitrogen. The libraries were sequenced using Illumina HiSeq X Ten. A total of 5.1 billion 150 bp paired-end reads were generated from sodium bisulfite-treated DNA with a 99% bisulfite conversion rate for each library. On average, about 637 million sequencing reads with an average of 26 × coverage were aligned to the hullless barley reference genome for each sample (Table 1). We identified 231,440 LMRs in total and 73,249 LMRs on average for each sample. The LMRs distribution of eight hullless barley varieties under four stress conditions was consistent, which the LMRs mainly distributed in the terminal of the chromosomes at a frequency similar to the distribution of gene density (Fig. 1).

### Histone modification in LMRs

To explore the location distribution of LMR on the genome, it was found that seventy-five percent of LMRs were located in the intergenic regions of the genome (Fig. 2a) In order to examine the location of LMRs relative to sites of histone modification, 2.3 billion clean ChIP-seq reads were generated from the eight samples (mean of 94 ± 19 SD million reads per sample) (Table 2). For the H3K4me3 modification, 37,000 peaks were

**Table 1** The sequencing results from BS-seq

Sample	RawReads	CleanReads	CleanBases	Coverage
DQ	594,788,152	590,823,228	83,104,198,973	24
XL	533,428,824	530,497,464	74,392,650,272	21
BQ3	756,520,682	748,641,934	105,687,219,276	30
GND7	737,454,714	726,820,566	100,823,741,986	29
NC3	646,283,200	641,575,332	91,217,851,918	26
NC6	681,142,904	676,647,870	96,284,006,901	28
Z0119	606,827,696	604,717,660	86,834,819,607	25
Z0226	580,123,472	577,851,246	82,416,533,291	24
Total	5,136,569,644	5,097,575,300	720,761,022,224	207
Average	642,071,206	637,196,913	90,095,127,778	26

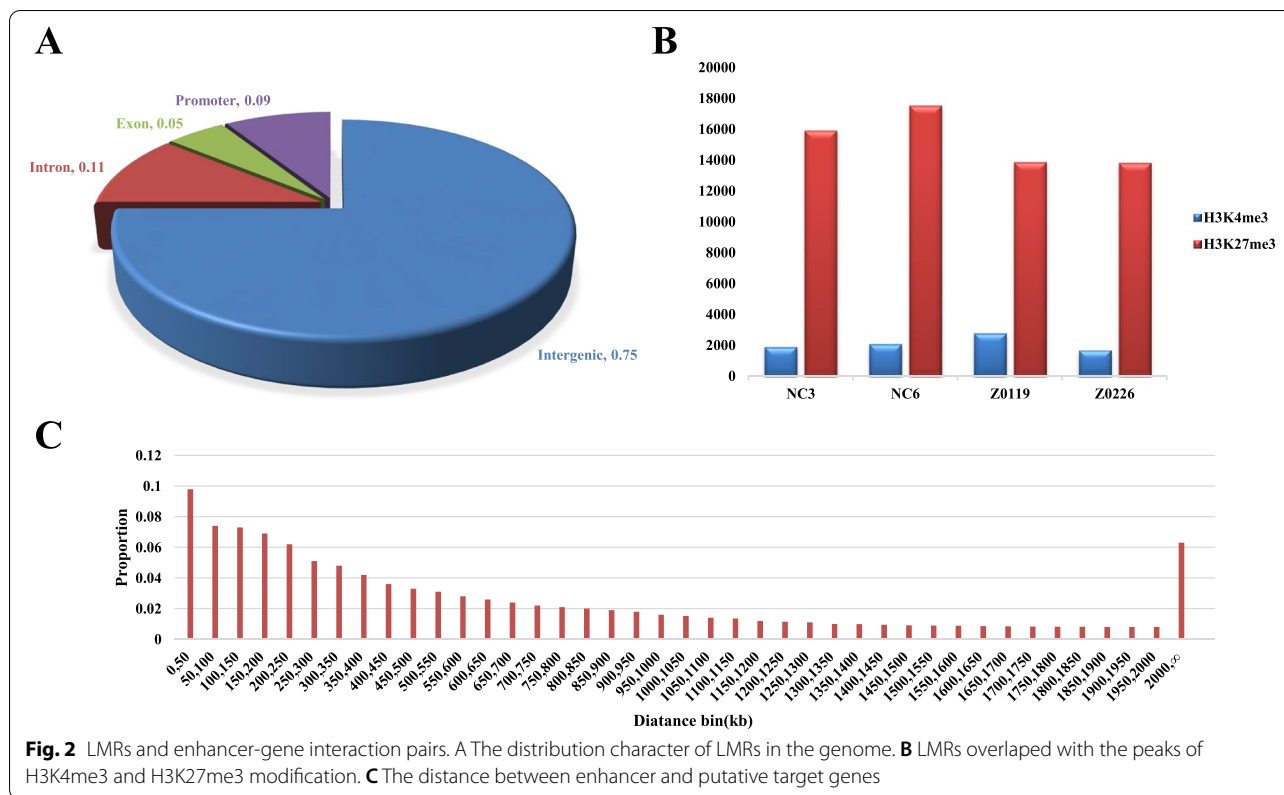


**Fig. 1** The genome-wide distribution of LMRs and histone modification. **A** The genome-wide distribution of LMRs and histone modification for XL and DQ. **B** The genome-wide distribution of LMRs and histone modification for NC3 and NC6. **C** The genome-wide distribution of LMRs and histone modification for Z0119 and Z0226. **D** The genome-wide distribution of LMRs and histone modification for BQ 3 and GND 7. a- Chromosome ideograms, b- Repeat elements, c—Gene density, d-LMRs density from XL,NC,Z0119,BQ3, respectively e- LMRs density from DQ,NC6,Z0226,GND7, respectively;f-H3K27me3 peaks density from XL,NC,Z0119,BQ3, respectively g- H3K27me3 peaks density from DQ,NC6,Z0226,GND7, respectively, h- Density of gene expression from RNA-seq in DQ,NC6,Z0226,GND7, respectively i- Density of gene expression from RNA-seq in DQ,NC6,Z0226,GND 7, respectively

detected on average, whereas 43,000 peaks were detected in the case of H3K27me3 modification (Table 3). A seven-fold higher number of LMRs were found to intersect with H3K27me3 peaks than with H3K4me3 peaks (Fig. 2B). The results are in agreement with LMRs in mice

[23]. H3K27me3 marks proximal and distal regulatory elements [24], whereas H3K4me3 mainly appears in transcript start sites [25].

It is difficult to identify the target genes of enhancers, because enhancers can work from a distance and in either



orientation [26]. In order to identify target genes regulated by distal regulatory elements, we used expression data (RNA-seq) for ten upstream and downstream genes from each distal regulatory element respectively. These 20 genes were filtered according to the correlation coefficient between the LMR DNA methylation level and gene expression level. Genes that are positively regulated by the enhancers should show a significant negative correlation. Using this method, we identified 97,909 enhancer-gene pairs, with 56,867 enhancers and 23,367 target genes (Table S1). There were 746,845 TFBS in the 56,867 enhancers, 13 TFBS in an enhancer on average. Each enhancer was associated with an average of 1.7 genes, and each gene was associated with an average of 4.2 enhancers. Of these enhancer-gene pairs, the target genes of 4468 enhancer-gene pairs were transcription factors, regarding to 1116 transcription factors. To investigate the relationship between putative enhancers and linked target genes, we determined the specific distances separating the identified enhancer-gene pairs and their frequencies using window sizes of 50 kb (Fig. 2C). We found that approximately 24% of enhancer-gene pairs were within the 150 kb range, and approximately 76% spanned 200 kb or larger genomic distances, with a median distance of 394 kb. We then selected the enhancer-gene pairs where a single enhancer was linked to a single gene

and determined how often the linked gene corresponded to the nearest TSS. Approximately 59% of enhancers had only one putative target gene. Only in 5% of these putative enhancer-gene pairs was the nearest TSS.

**Analysis of transcription factors under abiotic stress**

A total of 2329 transcription factors were identified in hulless barley genome using iTAK tools, that are mainly B3,NAC,bHLH family (Table S2, Fig. 3A). The LMRs were associated with enhancers containing TF-binding sites that recruit TFs to regulate the expression of nearby genes. To identify specific TFs that may play important roles in establishing and maintaining cell fates and adapting to extreme environments, we first identified the sequence motifs that appeared to be overrepresented in the specific hypomethylated LMRs of four tolerant varieties (Fig. 3B). This identification was achieved by performing HOMER analysis on repeat-masked sequences within the LMRs. Under drought stress, 52 enriched motifs and associated transcription factors were identified, including NAC, bHLH, MYB, TCP, and bZIP. Under cold stress, 67 enriched motifs and associated transcription factors were identified, including EBF, MYB-related, NAC, AP2, bHLH, TCP, and bZIP. Under salt stress, 136 enriched motifs and associated transcription factors were identified, including TCP, MYB, NAC, and bZIP. Under

**Table 2** The sequencing results from ChIP-seq

Sample	RawReads	CleanReads	CleanBases
In_0119	130,331,912	120,699,622	17,133,390,386
In_0226	150,631,522	135,910,486	19,191,709,878
In_BQ3	110,756,256	96,622,958	14,027,721,777
In_DQ	84,770,450	80,465,832	11,423,743,392
In_GND7	99,378,278	85,323,580	12,349,042,608
In_NC3	85,717,146	76,974,628	11,144,732,567
In_NC6	95,183,346	85,590,522	12,406,825,955
In_XL	98,977,998	93,442,800	13,212,397,069
IP_0119_H3K27ME3	120,120,702	112,937,754	15,975,609,893
IP_0119_H3K4ME3	103,402,144	95,106,898	13,445,058,497
IP_0226_H3K27ME3	138,540,060	126,498,530	17,784,286,347
IP_0226_H3K4ME3	144,379,848	131,660,326	18,556,727,187
IP_BQ3_H3K27ME3	99,171,422	87,892,364	12,712,914,671
IP_BQ3_H3K4ME3	87,163,154	75,060,996	10,758,236,629
IP_DQ_H3K27ME3	109,696,894	103,489,644	14,706,537,230
IP_DQ_H3K4ME3	88,993,448	84,378,304	11,940,270,399
IP_GND7_H3K27ME3	125,591,094	110,606,552	15,736,874,897
IP_GND7_H3K4ME3	98,072,560	85,838,950	12,342,566,744
IP_NC3_H3K27ME3	84,343,530	75,903,352	10,976,121,973
IP_NC3_H3K4ME3	80,751,230	72,863,388	10,543,634,758
IP_NC6_H3K27ME3	83,565,116	75,700,530	10,969,572,587
IP_NC6_H3K4ME3	84,149,000	75,787,454	10,950,362,320
IP_XL_H3K27ME3	91,925,772	87,490,906	12,416,299,085
IP_XL_H3K4ME3	88,750,186	84,198,968	11,895,135,311
Total	2,484,363,068	2,260,445,344	322,599,772,160

**Table 3** The results of calling peaks for histone modification

Samples	Peak Number	
	H3K27ME3	H3K4ME3
Z0119	54,849	43,334
Z0226	75,442	60,175
BQ3	49,582	36,347
DQ	26,127	22,963
GND7	43,166	44,047
NC3	33,355	34,742
NC6	39,553	35,151
XL	18,999	17,859

low nitrogen stress, 151 enriched motifs and associated transcription factors were identified, including TCP, bZIP, HB, bHLH, MYB, NR, and AP2EREBP.

For the stress-related transcription factors, the regulatory network was constructed by Fimo, the results show that there are complex regulatory relationships among transcription factors (Fig. 3C). For the potential regulatory network under drought stress NAC transcription

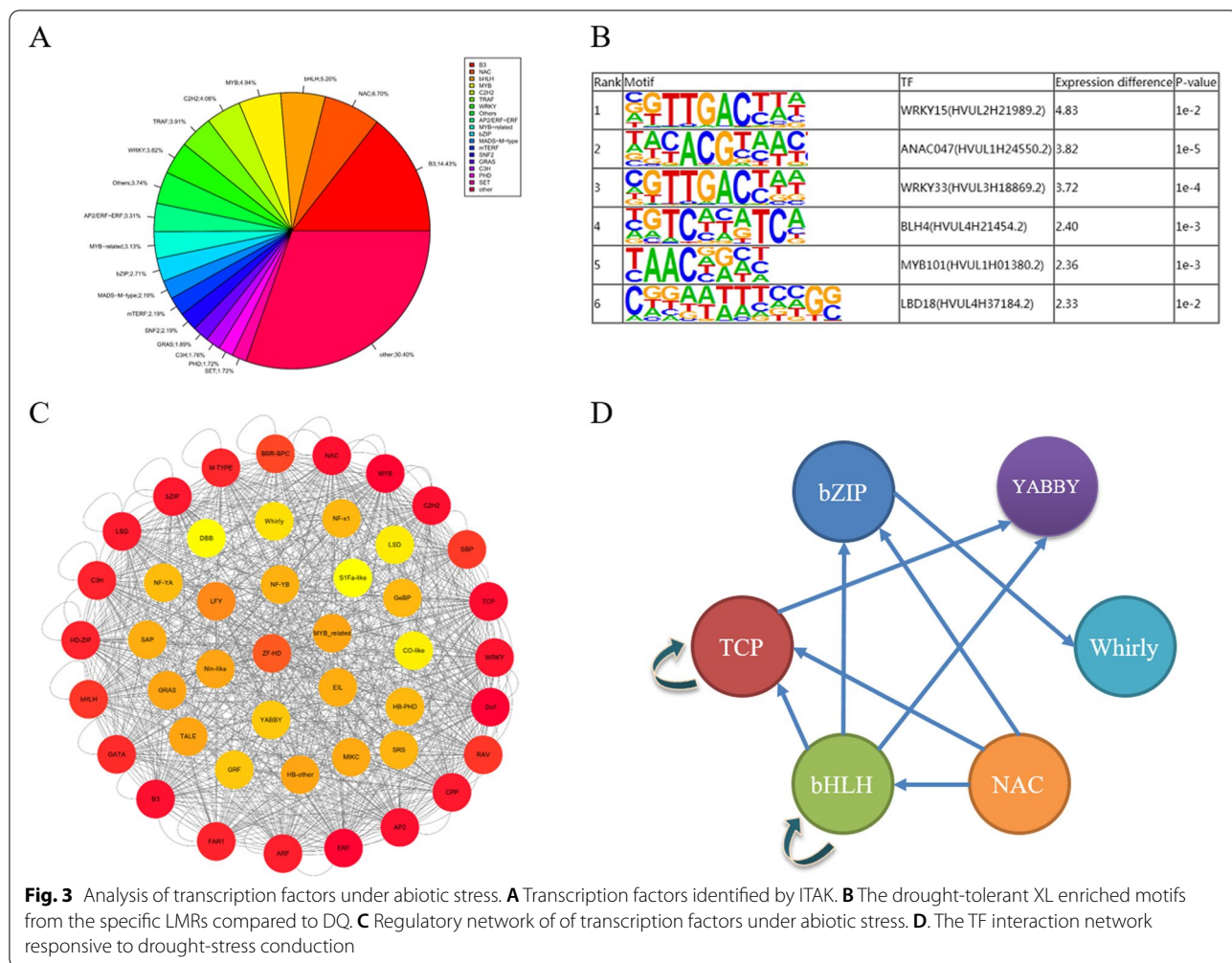
factor have been identified to be associated with drought stress [27]. There were 3156 significantly up-regulated genes in the XL16 sample compared to DQ, containing 181 transcription factors. There were 639 hypo-DMRs in XL16 compared to DQ, with 215 hypo-DMRs overlapped with the LMRs. Based the enhancer-gene pairs identified in the study, The target genes comprised of the transcription factors bHLH, TCP, bZIP, YABBY. Furthermore, the bHLH transcription factor could target the TCP, bZIP, YABBY transcription factors. The bHLH, TCP, bZIP transcription factors could target themselves too (Fig. 3D).

**More natural variation distributed in LMRs**

To determine natural variation patterns, we used variation information from previously published data sets [28]. After filtering with VCFtools, 22 million SNPs were used for subsequent analysis. The LMRs, corresponding to 75.8 Mb of sequence, were found to contain 911,187 SNPs, with 12 SNPs per kb of sequence, whereas the frequency for the whole genome is only 6 SNPs per kb on average.

**Discussion**

In our study, we characterized cis-regulatory elements related to LMRs from the whole genome of hulless barley. In terms of the nature of the regulatory elements, these LMRs are thought to function as enhancers. Thus far, many species have used LMRs for identifying enhancers [29–31]. It has been shown that lowly methylated and CpG-poor regions (LMRs, ~30% methylation) have enhancer characteristics, such as in serving as active histone marks, promoting DNase hypersensitivity, and increasing target gene expression [32]. The links from enhancers to target genes were deciphered via correlation analysis of the relationship between methylation level and gene expression. Using this method, a total of 56,867 enhancers could be linked to 23,367 putative target genes. In the majority of cases, enhancers (59%) were linked with the expression of only one gene. We found that the putative target gene was typically not the nearest gene. In fact, the identified gene was the nearest gene in only approximately 5% of enhancer–gene pairs. We detected links under the assumption that anti-correlation between an enhancer and the expression level of a nearby gene indicates functional regulation. Further experimental studies are needed to determine, with certainty, that the enhancers regulate their putative target genes. Other assays, such as Hi-C and ChIA-PET technology can also be used to validate the indicated enhancer–gene pairs [33, 34]. Many genes could not be linked for any enhancers, the reason for this may be that the sample size was too small to allow identification of links at the required level of significance.



**Fig. 3** Analysis of transcription factors under abiotic stress. **A** Transcription factors identified by ITAK. **B** The drought-tolerant XL enriched motifs from the specific LMRs compared to DQ. **C** Regulatory network of of transcription factors under abiotic stress. **D**. The TF interaction network responsive to drought-stress conduction

In 17% of cases, the LMRs overlapped with the H3K27me3 peaks, which may function as inactive enhancers [35]. Changes in the histone modification status of an enhancer region may be due to the gain or loss of site-specific transcription factors. To obtain insight into which site-specific TFs participate in response to different stresses, we examined the correspondence between stress-specific lowly H3K27-methylated enhancers and known regulatory factor recognition sequence motifs. We found associated TFs that recognize stress-specific little-H3K27-methylated enhancers, such as the NAC transcription factor [27, 36].

Inter-individual genetic variation is a major cause of phenotypic diversity. Most (88%) genome-wide association study (GWAS) loci occurred in noncoding DNA, suggesting regulatory functions [37]. In our study, we found that more SNPs were distributed in the LMRs. These SNPs are valuable resources for GWAS analysis and further functional studies.

**Conclusions**

Our study primarily identified cis-regulatory elements in hulless barley via whole-genome bisulfite sequencing. A total of 231,440 lowly methylated regions were identified in our study. Through correlation with expression data, 97,909 enhancer–gene pairs were discovered for all the LMRs. The transcription factor regulatory network was inferred based on the characterization of TFBS in the CREs. The histone marks corresponding to H3K27me3 modification indicated that many LMRs were suppressed under specific abiotic stresses. The cis-regulatory elements identified in our study contained numerous single nucleotide polymorphisms in the natural population.

**Methods**

**Plant materials and treatments**

Eight Tibetan hulless barley (*H. vulgare* subsp. *vulgare*) varieties were used for methylome analysis in this study, which includes the drought-tolerant and susceptible varieties XL16 vs. DQ, salt-tolerant and susceptible varieties

Z0119 vs. Z0226, cold-tolerant and susceptible varieties NC3 vs. NC6, and low-nitrogen -tolerant and susceptible varieties BQ3 vs. GND7. Eight hullless barley varieties were preserved in the Agricultural Research Institute, Tibet Academy of Agricultural and Animal Husbandry Sciences, Lhasa, China for many years, all the methods involving the plant and its material complied with relevant institutional, national, and international guidelines and legislation.

Drought stress was induced by 21% polyethylene glycol (PEG) 6000 solutions for 48 h. For the assessment of salinity tolerance, seedlings were stressed with 200 mM NaHCO<sub>3</sub> and Na<sub>2</sub>CO<sub>3</sub> at the two-leaf and one needle stages for 72 h, respectively.

For the assessment of cold tolerance, seedlings were grown in a growth chamber at a temperature of 4 °C for 72 h. For the assessment of low-nitrogen tolerance, sterilized seeds were germinated on moistened filter paper deficient in nitrogen in a growth chamber for 128 h. Healthy seeds were sterilized after soaking in 2% H<sub>2</sub>O<sub>2</sub> for 40 min, rinsed in sterile water, and then germinated on moistened filter paper at 16–18 °C with 14 h light/10 h dark photoperiods and relative humidity of 80% in a growth chamber. After germination, the seedlings were grown in a growth chamber at 16 °C, relative humidity of 80%, and a photoperiod of 14 h. The roots and leaves were harvested from 10-day-old seedlings belonging to two genotypes under normal and stressful conditions. Two plants from each pot were considered biological replicates. The aforementioned tissue samples were snap-frozen in nitrogen and immediately stored at –80 °C [38].

#### **DNA extraction and whole-genome bisulfite sequencing (BS-seq)**

DNA was isolated using the cetyltrimethylammonium bromide (CTAB) method from hull-less barley leaves. The integrity of DNA was checked by agarose gel electrophoresis and the concentration was measured via a non-ultraviolet method [39]. Wuhan IGENEBOOK Biotechnology Co., Ltd conducted the bisulfite treatment, library construction, and sequencing. Unmethylated lambda DNA was spiked in to evaluate the conversion rate for all libraries. Finally, paired-end bisulfite-treated sample libraries were constructed and sequenced.

#### **Chromatin immunoprecipitation (ChIP) assay**

ChIP assays were performed according to methods previously described [40] by Wuhan IGENEBOOK Biotechnology Co., Ltd (<http://www.igenebook.com>). Briefly, 3 g of barley leaves were washed twice in cold PBS buffer and cross-linked with 1% formaldehyde for 10 min at room temperature. They were then quenched by adding glycine (final concentration, 125 mmol/L).

Afterward, samples were lysed, and chromatin was obtained following incubation on ice. The chromatin samples were sonicated to obtain soluble sheared chromatin with an average DNA length of 200–500 bp. A total of 20 µL of chromatin was stored at –20 °C for input DNA, and 100 µL of chromatin was used for immunoprecipitation using H3K4ME3 and H3K27ME3 antibodies (9751S, CST; ab6002, Abcam). A total of 10 µg of antibody was used in the immunoprecipitation reactions at 4 °C overnight. The next day, 30 µL of protein beads were added, and the samples were further incubated for 3 h. Afterward, the beads were washed once with 20 mM Tris/HCL (pH 8.1), 50 mM NaCl, 2 mM EDTA, and 1% Triton X-100, 0.1% SDS; twice with 10 mM Tris/HCL (pH 8.1), 250 mM LiCl, 1 mM EDTA, 1% NP-40, 1% deoxycholic acid, and twice with 1 × TE buffer (10 mM Tris–Cl at pH 7.5, 1 mM EDTA). Bound material was then eluted from the beads in 300 µL of elution buffer (100 mM NaHCO<sub>3</sub>, 1% SDS) and treated first with RNase A (final concentration, 8 µg/mL) for 6 h at 65 °C and then with proteinase K (final concentration, 345 µg/mL) overnight at 45 °C. Immunoprecipitated DNA was used to construct sequencing libraries following the protocol provided by the I NEXTFLEX® ChIP-Seq Library Prep Kit for Illumina® Sequencing (NOVA-514120, Bioo Scientific) and sequenced on Illumina Xten using the PE 150 method.

#### **BS-seq data analysis**

After removing low-quality reads, clean data were mapped to the reference genome using the software BitMapperBS (version 2.9), which permits an 8% mismatch per read [41]. Then, the methylation levels were calculated based on the methyl-cytosine percentage by MethGO [42]. Differentially methylated regions (DMRs) were identified using the software cgmaptools dmr tool [43]. LMR identification was conducted using the R package "MethylSeekR" (with the methylation level threshold under 20%) [44]. This package allows for identifying active regulatory regions from high-resolution WGBS methylomes and relies on the idea of transcription factor binding, which leads to a defined reduction in DNA methylation.

#### **ChIP-seq data analysis**

The raw reads were filtered using Trimmomatic (version 0.38) [45]. Then, the clean reads were mapped to the hullless barley reference genome [3] by BWA (version 0.7.15) [46], allowing up to two mismatches. SAMtools (version 1.3.1) [47] was used to remove potential PCR duplicates, and MACS2 software

(version 2.1.1.20160309) [48, 49] was used to call peaks of histone enrichment by default parameters (bandwidth, 300 bp; model fold, 5, 50; q-value, 0.05). The tags per kilobase of gene length per million mapped reads (TPM) were used to analyze the gene modification level (histone mark abundance).

### Linking the enhancer and target gene with methylation and expression changes

Each of the putative enhancer-gene relations (the closest 10 upstream genes and the closest 10 downstream genes) was tested for correlation between the methylation level of the enhancer and gene expression. The gene expression data were derived from published data [50, 51]. To select these genes, the enhancer–gene distance was defined as the distance from the enhancer to the gene transcriptional start site. The enhancer was considered to regulate the target gene when the correlation was significantly negative.

### The identification of transcription factors and TFBS

Transcription factors were detected using iTAK tools [52]. A total of 2329 transcription factors were identified in hulless barley genome (Table S2). The orthologous protein pairs between proteins of hulless barley and proteins with motifs were identified in JASPAR [53], resulting in the identification of 196 orthologous protein pairs. Furthermore, we examined TFBS in the enhancer for each putative transcription factor using the software FIMO with the parameter  $P < 0.0001$  [54], which allowed the association of enhancers with transcription factors in the hulless barley genome.

### Abbreviations

WGBS: Whole-genome bisulfite sequencing; LMRs: Lowly methylated regions; TF: Transcription factor; TFBS: Transcription factor binding site.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-09070-x>.

**Additional file 1: Table S1.** Identified enhancer–gene pairs.

**Additional file 2: Table S2.** Genome wide transcription factor identification.

### Acknowledgements

We thank Wuhan IGENEBOOK Biotechnology Co., Ltd for its help in data analysis.

### Authors' contributions

X.Z. and Y.W. conceived the idea of the work and designed the research; Q.X. and S.H. produced and analyzed the data; M.W., Y.W. and C.Y. managed the samples and the data; S.H. and Q.X. wrote the paper; X.Z. and G.G. revised the manuscript. All authors read and approved the final manuscript.

### Funding

The Financial Special Fund provided financial support for the field survey of materials (2015CZZX001 and XZNKY-2018-C-021). The financial support from China Agriculture Research System (CARS-05-01A-08) was provided to conduct the sequencing and data analysis.

### Availability of data and materials

The datasets generated during the current study are available in the NCBI SRA repository: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA708355>. The experimental materials, including NC3 (langkaziqingke), Z0119, and Z0226, were landraces collected by the Tibet Academy of Agricultural and Animal Husbandry Sciences. The materials NC6 (kunlun1), BQ3 (beijing3), GND7 (gan-nongda7), XL16 (ximala16), and DQ (diqingheiyuanguai) were cultivars bred by different groups in western China. The plant materials used in this study are available from the corresponding author upon request: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA708355>

### Declarations

#### Ethics approval and consent to participate

My manuscript does not report on or involve the use of any animal or human data or tissue, it is not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>State Key Laboratory of Hulless Barley and Yak Germplasm Resources and Genetic Improvement, Lhasa 850002, China. <sup>2</sup>Agricultural Research Institute, Tibet Academy of Agricultural and Animal Husbandry Sciences, Lhasa 850002, China. <sup>3</sup>College of Forestry, Henan Agricultural University, Zhengzhou 450002, People's Republic of China. <sup>4</sup>Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China.

Received: 7 April 2022 Accepted: 5 December 2022

Published online: 20 December 2022

### References

- Badr A, Muller K, Schafer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, et al. On the origin and domestication history of barley (*Hordeum vulgare*). *Mol Biol Evol.* 2000;17:499–510. <https://doi.org/10.1093/oxfordjournals.molbev.a026330>.
- Dequan M. The research on classification and origin of cultivated barley in Tibet Autonomous Region. *Sci Agri Sin.* 1988.
- Zeng X, Long H, Wang Z, Zhao S, Tang Y, Huang Z, et al. The draft genome of Tibetan hulless barley reveals adaptive patterns to the high stressful Tibetan Plateau. *Proc Natl Acad Sci.* 2015;112:1095–100. <https://doi.org/10.1073/pnas.1423628112>.
- Zeng X, Xu T, Ling Z, Wang Y, Li X, Xu S, et al. An improved high-quality genome assembly and annotation of Tibetan hulless barley. *Scientific Data.* 2020;7:139. <https://doi.org/10.1038/s41597-020-0480-0>.
- Zeng X, Guo Y, Xu Q, Mascher M, Guo G, Li S, et al. Origin and evolution of hulless barley in Tibet. *Nat Commun.* 2018;9:5433. <https://doi.org/10.1038/s41467-018-07920-5>.
- Consortium EP. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature.* 2007;447:799–816. <https://doi.org/10.1038/nature05874>.
- Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, et al. A comparative encyclopedia of DNA elements in the mouse genome. *Nature.* 2014;515:355–64. <https://doi.org/10.1038/nature13992>.
- Lane AK, Niederhuth CE, Ji L, Schmitz RJ. pENCODE: A Plant Encyclopedia of DNA Elements. *Annu Rev Genet.* 2014;48:49–70. <https://doi.org/10.1146/annurev-genet-120213-092443>.



9. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet.* 2013;14:204–20. <https://doi.org/10.1038/nrg3354>.
10. Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet.* 2014;15:394–408. <https://doi.org/10.1038/nrg3683>.
11. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet.* 2012;13:484–92. <https://doi.org/10.1038/nrg3230>.
12. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature.* 2011;480:490–5. <https://doi.org/10.1038/nature10716>.
13. Pfluger J, Wagner D. Histone modifications and dynamic regulation of genome accessibility in plants. *Curr Opin Plant Biol.* 2007;10:645–52. <https://doi.org/10.1016/j.pbi.2007.07.013>.
14. Berr A, Shafiq S, Shen WH. Histone modifications in transcriptional activation during plant development. *Biochim Biophys Acta.* 2011;1809:567–76. <https://doi.org/10.1016/j.bbagr.2011.07.001>.
15. Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet.* 2012;13:613–26. <https://doi.org/10.1038/nrg3207>.
16. Meng F, Zhao H, Zhu B, Zhang T, Yang M, Li Y, et al. Genomic editing of intronic enhancers unveils their role in fine-tuning tissue-specific gene expression in *Arabidopsis thaliana*. *Plant Cell.* 2021;33:1997–2014. <https://doi.org/10.1093/plcell/kob093>.
17. Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature.* 2006;442:86–90. <https://doi.org/10.1038/nature04815>.
18. van Dijk K, Ding Y, Malkaram S, Riethoven JJM, Liu R, Yang J, et al. Dynamic changes in genome-wide histone H3 lysine 4 methylation patterns in response to dehydration stress in *Arabidopsis thaliana*. *BMC Plant Biol.* 2010;10:238. <https://doi.org/10.1186/1471-2229-10-238>.
19. Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, et al. Histone H2A Monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol Cell.* 2008;29:69–80. <https://doi.org/10.1016/j.molcel.2007.11.002>.
20. Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, et al. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat Cell Biol.* 2007;9:1428–35. <https://doi.org/10.1038/ncb1663>.
21. Sani E, Herzyk P, Perrella G, Colot V, Amtmann A. Hyperosmotic priming of *Arabidopsis* seedlings establishes a long-term somatic memory accompanied by specific changes of the epigenome. *Genome Biol.* 2013;14:R59. <https://doi.org/10.1186/gb-2013-14-6-r59>.
22. Du Z, Li H, Wei Q, Zhao X, Wang C, Zhu Q, et al. Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in *Oryza sativa* L. *Japonica*. *Mol Plant.* 2013;6:1463–72. <https://doi.org/10.1093/mp/sst018>.
23. Van Nimwegen E. DNA-binding factors shape the mouse methylome at distal regulatory regions. 2012.
24. Attanasio C, Nord AS, Zhu Y, Blow MJ, Biddie SC, Mendenhall EM, et al. Tissue-specific SMARCA4 binding at active and repressed regulatory elements during embryogenesis. *Genome Res.* 2014;24:920–9. <https://doi.org/10.1101/gr.168930.113>.
25. Liu X, Wang C, Liu W, Li J, Li C, Kou X, et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature.* 2016;537:558–62. <https://doi.org/10.1038/nature19362>.
26. Mumbach MR, Satpathy AT, Boyle EA, Dai C, Gowen BG, Cho SW, et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet.* 2017;49:1602–12. <https://doi.org/10.1038/ng.3963>.
27. Hu H, Dai M, Yao J, Xiao B, Li X, Zhang Q, et al. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proc Natl Acad Sci.* 2006;103:12987–92. <https://doi.org/10.1073/pnas.0604882103>.
28. Zeng X, Yuan H, Dong X, Peng M, Jing X, Xu Q, et al. Genome-wide dissection of co-selected UV-B responsive pathways in the UV-B adaptation of Qingke. *Mol Plant.* 2020;13:112–27. <https://doi.org/10.1016/j.molp.2019.10.009>.
29. Oka R, Zicola J, Weber B, et al. Genome-wide mapping of transcriptional enhancer candidates using DNA and chromatin features in maize. *Genome Biol.* 2017;18:137. <https://doi.org/10.1186/s13059-017-1273-4>.
30. Yao L, Shen H, Laird PW, Farnham PJ, Berman BP. Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. *Genome Biol.* 2015;16:105. <https://doi.org/10.1186/s13059-015-0668-3>.
31. Lee HJ, Lowdon RF, Maricque B, Zhang B, Stevens M, Li D, et al. Developmental enhancers revealed by extensive DNA methylome maps of zebrafish early embryos. *Nat Commun.* 2015;6:6315. <https://doi.org/10.1038/ncomms7315>.
32. Charlet J, Duymich CE, Lay FD, Mundbjerg K, Dalsgaard Sørensen K, Liang G, et al. Bivalent regions of cytosine methylation and H3K27 acetylation suggest an active role for DNA methylation at enhancers. *Mol Cell.* 2016;62:422–31. <https://doi.org/10.1016/j.molcel.2016.03.033>.
33. Feng S, Cokus SJ, Schubert V, Zhai J, Pellegrini M, Jacobsen SE. Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in *Arabidopsis*. *Mol Cell.* 2014;55:694–707. <https://doi.org/10.1016/j.molcel.2014.07.008>.
34. Zhao L, Wang S, Cao Z, Ouyang W, Zhang Q, Xie L, et al. Chromatin loops associated with active genes and heterochromatin shape rice genome architecture for transcriptional regulation. *Nat Commun.* 2019;10:3640. <https://doi.org/10.1038/s41467-019-11535-9>.
35. Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat Genet.* 2013;45:1198–206. <https://doi.org/10.1038/ng.2746>.
36. Fang Y, Liao K, Du H, Xu Y, Song H, Li X, et al. A stress-responsive NAC transcription factor SNAC3 confers heat and drought tolerance through modulation of reactive oxygen species in rice. *J Exp Bot.* 2015;66:6803–17. <https://doi.org/10.1093/jxb/erv386>.
37. Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci.* 2009;106:9362–7. <https://doi.org/10.1073/pnas.0903103106>.
38. Wang F, Wang X, Zhao C, et al. Alternative pathway is involved in the tolerance of highland barley to the low-nitrogen stress by maintaining the cellular redox homeostasis. *Plant Cell Rep.* 2016;35:317–28. <https://doi.org/10.1007/s00299-015-1886-6>.
39. Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc.* 2006;1:2320–5. <https://doi.org/10.1038/nprot.2006.384>.
40. Li N, Wei S, Chen J, Yang F, Kong L, Chen C, et al. OsASR2 regulates the expression of a defence-related gene, Os2H16, by targeting the GT-1 cis-element. *Plant Biotechnol J.* 2018;16:771–83. <https://doi.org/10.1111/pbi.12827>.
41. Cheng H, Xu Y. BitMapperBS: a fast and accurate read aligner for whole-genome bisulfite sequencing. *bioRxiv.* 2018:442798. <https://www.biorxiv.org/content/10.1101/442798v2.article-metrics>.
42. Liao W-W, Yen M-R, Ju E, Hsu F-M, Lam L, Chen P-Y. MethGo: a comprehensive tool for analyzing whole-genome bisulfite sequencing data. *BMC Genomics.* 2015;16:S11. <https://doi.org/10.1186/1471-2164-16-S12-S11>.
43. Guo W, Zhu P, Pellegrini M, Zhang MQ, Wang X, Ni Z. CGmapTools improves the precision of heterozygous SNV calls and supports allele-specific methylation detection and visualization in bisulfite-sequencing data. *Bioinformatics.* 2018;34:381–7. <https://doi.org/10.1093/bioinformatics/btx595>.
44. Burger L, Gaidatzis D, Schübeler D, Stadler MB. Identification of active regulatory regions from DNA methylation data. *Nucleic Acids Res.* 2013;41:e155–e155. <https://doi.org/10.1093/nar/gkt599>.
45. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
46. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25:1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078–9. <https://doi.org/10.1093/bioinformatics/btp352>.

48. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc.* 2012;7:1728–40. <https://doi.org/10.1038/nprot.2012.101>.
49. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9:R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
50. Zeng X, Bai L, Wei Z, Yuan H, Wang Y, Xu Q, et al. Transcriptome analysis revealed the drought-responsive genes in Tibetan hulless barley. *BMC Genomics.* 2016;17:386.
51. Yuan H, Zeng X, Ling Z, Wei Z, Wang Y, Zhuang Z, et al. Transcriptome profiles reveal cold acclimation and freezing tolerance of susceptible and tolerant hulless barley genotypes. *Acta Physiol Plant.* 2017;39:275. <https://doi.org/10.1007/s11738-017-2566-7>.
52. Zheng Y, Jiao C, Sun H, Rosli HG, Pombo MA, Zhang P, et al. iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol Plant.* 2016;9:1667–70.
53. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 2019;48:D87–92. <https://doi.org/10.1093/nar/gkz1001>.
54. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics.* 2011;27:1017–8. <https://doi.org/10.1093/bioinformatics/btr064>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

