HAYATI
Journal of Biosciences

Vol. 33 No. 6, November 2026 1605-1616
DOI:10.4308/hjb.33.6.1605-1616
ISSN: 1978-3019
EISSN: 2086-4094

Research Article

Check for updates

OPEN ACCESS

Identification and Characterization of a Panicle Weight-Associated SCoT Marker in Foxtail Millet

Muhammad Anthony Rizwandi[1], Marisa Vidya Luthfiani[1], Andrean Heskiel Wokspakrik[1,2], Darda Efendi[3], Willy Bayuardi Suwarno[3], Sintho Wahyuning Ardie[3*]

[1]Plant Breeding and Biotechnology Study Program, Faculty of Agriculture, IPB University, Jl. Meranti, Dramaga Campus, Bogor 16680, West Java, Indonesia
[2]Papua University, Crop Production Study Program, Faculty of Agriculture, Manokwari, 98314, West Papua, Indonesia
[3]Department of Agronomy and Horticulture, Faculty of Agriculture, IPB University, Jl. Meranti, Dramaga Campus, Bogor 16680, West Java, Indonesia

ARTICLE INFO

Article history:
Received May 6, 2026
Received in revised form July 9, 2026
Accepted September 16, 2026
Available Online October 29, 2026

KEYWORDS:
bioinformatics analysis,
climate-resilient crop,
functional RNA,
TASSEL 5.0

Copyright (c) 2026 @author(s).

ABSTRACT

Foxtail millet (*Setaria italica* L.) is a climate-resilient cereal crop increasingly recognized for its role in food security and sustainable agriculture. This study aimed to identify SCoT loci associated with main panicle weight (MPW), characterize their functional relevance, and validate their potential for selecting high-productivity genotypes. Seven Indonesian foxtail millet local genotypes were analyzed using 12 SCoT primers, yielding 109 loci. Association analysis identified seven loci associated with MPW. Among them, the SCoT-32 (700 bp) showed a significant association ($R^2 = 0.6841$, $p = 0.0217$), with the band occurring only in high-MPW genotypes and absent in low-MPW genotypes. Sequence characterization of the SCoT32_700 locus positioned it near an NB ARC domain gene on chromosome 7, suggesting a regulatory role in panicle development. Bioinformatics analyses classified the locus as a putative long non-coding RNA (lncRNA), supported by stable secondary structure predictions and RNAcentral homology to known *Setaria italica* lncRNAs. Validation in 20 F_5 lines partially confirmed the association, with the 700 bp fragment explaining 17.2% of MPW variation. These findings highlight the potential of SCoT markers, particularly SCoT32_700, as functional markers for yield-related traits.

1. Introduction

The accelerating impacts of climate change, including rising temperatures, unpredictable rainfall, and increasing soil degradation, necessitate the adoption of climate-resilient crops (Ardie *et al.* 2025). Foxtail millet (*Setaria italica* L.) has emerged as a climate-resilient crop to sustain food production under these challenging conditions due to its remarkable tolerance to various abiotic stresses (Ejigu *et al.* 2025). Beyond its resilience, foxtail millet offers substantial nutritional benefits, including high levels of dietary fiber, essential amino acids, and micronutrients (Gupta *et al.* 2025; Shukla *et al.* 2026). However, despite these advantages, foxtail millet's overall productivity remains considerably low. The typical productivity of foxtail millet ranges only between 2-3 t ha^{-1} (Harish *et al.* 2024).

To address this challenge, precision breeding of foxtail millet has become a priority. The development of functional markers (FMs) is essential for accelerating breeding programs by enabling efficient selection of

*Corresponding Author
E-mail Address: sintho_wa@apps.ipb.ac.id

genotypes carrying desirable traits (Salgotra & Stewart 2020). One promising approach is the use of Start Codon Targeted (SCoT) markers. These markers are designed on conserved sequences surrounding the start codon, thereby generating amplicons that are likely to represent coding DNA sequences (CDS) of genes (Collard & Mackill 2009; Rai 2023). This property makes SCoT markers particularly suitable for identifying loci with direct functional relevance to agronomic traits and for subsequent FM development.

Among the traits that determine foxtail millet productivity, panicle-related traits are central. Panicle weight, in particular, integrates multiple yield components and has been reported as a key determinant of overall productivity. Several panicle traits that influence yield have been documented (Liu *et al.* 2022; Luthfiani *et al.* 2026), and the association of SCoT amplicons with these traits offers a pathway to develop FMs for breeding high-yielding foxtail millet varieties. Thus, linking SCoT loci to panicle weight provides a strategic opportunity to enhance productivity through marker-assisted selection.

The present study aims (1) to identify SCoT loci associated with main panicle weight per plant in foxtail millet, and (2) to characterize the associated loci to facilitate their future application in functional marker development. By integrating marker–trait association with locus characterization, this work aims to contribute to precision breeding strategies to enhance foxtail millet productivity across diverse agro-climatic conditions.

2. Materials and Methods

2.1. Genetic Materials

Seven Indonesian foxtail millet local genotypes used in this study were Hambapraing, Mauliru-2, NTB-1, Buru, ICERI-5, ICERI-6, and ICERI-7 (Luthfiani *et al.* 2026). Seeds of each genotype were germinated in planting trays filled with a mixture of soil and rice husk charcoal (1:1, v/v), with three seeds sown per hole. The genomic DNA was extracted from young leaves (approximately 10 days after sowing) using a modified CTAB method (Aboul-Maaty & Oraby 2019). DNA concentration and purity were measured with a spectrophotometer (MN-917 MaestroNano, MaestroGen, Taiwan).

2.2. Trait-Marker Association Analysis

A total of 109 SCoT loci were obtained from the amplification of 12 SCoT primers (Supplementary Table 1) across seven genotypes (Rizwandi *et al.* 2026). Phenotypic data for main panicle weight (MPW) were obtained from Luthfiani *et al.* (2026) (Supplementary Table 2). Genotypes were clustered into high and low MPW groups using Euclidean distance in RStudio (Supplementary Figure 1). Association analysis was performed using TASSEL 5.0 (Bradbury *et al.* 2007) with a general linear model (GLM) to identify loci associated with MPW based on p-values ($p<0.05$). Statistical significance was further evaluated by 10,000 phenotype permutations to calibrate the null distribution and reduce false-positive signals.

2.3. Re-amplification of Genomic DNA and PCR Conditions

SCoT primers producing loci associated with MPW were used for re-amplification of genomic DNA from the seven foxtail millet genotypes. PCR reactions were prepared in 50 µL volumes for sequencing purposes, containing 12.5 µL DNA template (12 ng μL^{-1}), 2.5 µL primer (10 µM), 25 µL DNA polymerase (MyTaq™ HS Red Mix, 2×), and 10 µL nuclease-free water. Amplification was performed using the following program, starting with initial denaturation at 94°C for 5 minutes, 35 cycles of denaturation at 94°C for 5 seconds, annealing at 50°C for 1 minute, and extension at 72°C for 30 seconds, followed by a final extension at 72°C for 10 minutes. The PCR products (8 µL) were separated by electrophoresis in a 2% (w/v) agarose gel with 1× TAE buffer at 50 V for 90 minutes. Gels were stained with ethidium bromide (0.5 µg mL^{-1}) for 10 min and visualized under UV light (Gel Doc EZ™, Bio-Rad, USA).

2.4. Sequencing and Bioinformatics Analysis

Only the associated fragment from the genotype with the highest MPW (Hambapraing) was sequenced. The PCR products (42 µL) obtained were purified using the ZymoClean Gel DNA Recovery Kit (Zymo Research, D4001). The purified fragments were ligated into the pTA2 vector using the TOYOBO Target Clone™ Kit (Toyobo, TAK 201) with a reaction mixture consisting of ddH_2O, 2× ligation buffer, pTA2 vector (50 ng μL^{-1}), purified PCR product, and T4 DNA ligase according to the manufacturer's instructions. The ligation mixture was transformed into DH10B competent cells using the Thermo Scientific Transformation Kit (EC0113), and colonies grown on selective media were screened. Colony PCR reactions analyzed selected colonies, with 10 µL volumes containing 5 µL MyTaq HS Red Mix (2×), 0.4 µL T3 (10 µM, 5'-GCAATTAACCCTCACTAAAGG-3') and 0.4 µL

T7 (10 µM, 5'-TAATACGACTCACTATAGGG-3') primers, and 4.2 µL ddH_2O. PCR protocol followed these conditions: initial denaturation at 95°C for 3 minutes, 35 cycles of denaturation at 95°C for 15 seconds, annealing at 45°C for 30 seconds, and extension at 72°C for 45 seconds, followed by a final extension at 72°C for 2 minutes. Positive colonies were cultured, and plasmids were isolated using the ZR Plasmid Miniprep-Classic Kit (Zymo Research, D4015). The isolated plasmids were subjected to bidirectional sequencing using T3 and T7 primers. PT Genetika Science Indonesia carried out sequencing procedures.

Bioinformatics analyses were performed to determine the identity and potential function of the fragment. As an initial step, nucleotide sequences were aligned against the foxtail millet reference genome (cultivar Yugu1, GenBank accession AGNK02000000) using BLAST (Basic Local Alignment Search Tool) at Phytozome (https://phytozome-next.jgi.doe.gov/). Then, nucleotide BLAST (blastn) was performed against the NCBI database (core_nt) to identify potential matches with annotated genes. If no significant matches were detected, further analyses were conducted to evaluate whether the sequence encodes a protein: coding potential was assessed using CPC2 (Kang *et al.* 2017) via its web server (http://cpc2.cbi.pku.edu.cn), and open reading frames (ORFs) were identified using ORF Finder (Rombel *et al.* 2002) at NCBI. If the sequence is unlikely to encode proteins, its potential as functional RNA was examined. Long non-coding RNAs (lncRNAs) were predicted using PLEK (Li *et al.* 2014), which applies an improved k-mer scheme. Secondary structure prediction was conducted using RNAfold from the ViennaRNA Package 2.0 (Lorenz *et al.* 2011). Non-coding RNA sequences were further validated using RNAcentral (The RNAcentral Consortium 2019), which integrates data from 28 expert databases and provides stable URS identifiers, genomic annotations, and functional information. Genomic diagrams were reconstructed to identify candidate genes located within the LD decay region (Jaiswal *et al.* 2019) surrounding the associated locus, based on BLAST reference results from the Phytozome database.

2.5. Validation of Associated SCoT Loci in 20 F_5 Foxtail Millet Lines

Associated SCoT loci were validated in 20 F_5 lines derived from the ICERI-6 × Botok-4 cross, which exhibited contrasting MPW. Five lines with the lowest average MPW and five with the highest average MPW, with two individuals per line, were selected for validation (Supplementary Table 3). Genomic DNA was isolated using the modified CTAB method (Aboul-Maaty & Oraby 2019). The PCR amplification was performed using the associated SCoT primers in a 10 µL reaction containing 5 µL DNA polymerase (MyTaq HS Red Mix, 2×), 2.5 µL DNA template (12 ng μL^{-1}), and 2.5 µL primer (10 µM). The PCR products were separated by electrophoresis in a 2% (w/v) agarose gel with 1× TAE buffer at 50 V for 90 minutes. Gels were stained with ethidium bromide (0.5 µg mL^{-1}) for 10 min and visualized under UV light (Gel Doc EZ™, Bio-Rad, USA). Molecular data analysis was conducted by scoring DNA bands from the visualized DNA. Banding patterns were scored as binary data (1 = presence, 0 = absence). The contribution of associated loci and MPW in the 20 lines was analyzed using regression analysis in RStudio.

3. Results

3.1. Association Analysis of SCoT Loci with Main Panicle Weight in Foxtail Millet

The Manhattan plot (Figure 1A) illustrates the distribution of $-\log_{10}(p)$ values across all tested markers, with several loci exceeding the significance threshold ($p \leq 0.05$), indicating strong associations with MPW. The QQ plot (Figure 1B) shows that most of the observed *p*-values lie close to the diagonal line, which represents the expected distribution when no true association exists. This indicates that the statistical model was well-calibrated and that the significant loci identified are more likely to represent true biological signals.

A total of seven loci, produced by six SCoT primers, were significantly associated with MPW (Table 1). Among them, locus SCoT16_475 exhibited the strongest association, with the lowest *p*-value (0.0010) and the highest coefficient of determination ($R^2 = 0.9033$), indicating a major contribution to MPW variation. Other loci also showed significant associations ($p = 0.0217$; $R^2 \approx 0.6841$), while SCoT15_700 displayed a weaker association ($p = 0.0499$; $R^2 = 0.5696$). The consistency between the tabulated results and the patterns observed in the Manhattan and QQ plots strengthens the evidence that these loci are genuinely associated with MPW variation and may serve as potential molecular markers for yield-related traits in foxtail millet.

To confirm that the identified loci were consistently associated with MPW, re-amplification and visualization were performed using the six SCoT primers (Figure

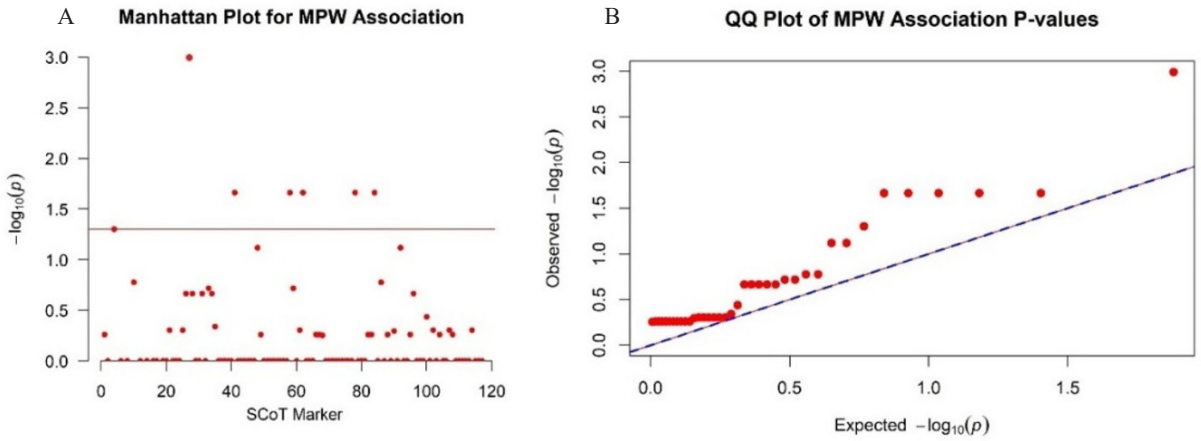


Figure 1. (A). Manhattan plots; red lines are the threshold $p \leq 0.05$, and (B) Quantile–quantile (Q-Q) plots of marker–trait associations for main panicle weight traits analyzed using the general linear model (GLM) with SCoT markers using 10,000 permutations

Table 1. Association analysis results between SCoT markers and main panicle weight traits based on the GLM ($p<0.05$)

Primer name	Locus size (bp)	Marker name	$p<0.05$	R^2
SCoT-16	475	SCoT16_475	0.0010	0.9033
SCoT-21	750	SCoT21_750	0.0217	0.6841
SCoT-23	950	SCoT23_950	0.0217	0.6841
SCoT-23	450	SCoT23_450	0.0217	0.6841
SCoT-31	625	SCoT31_625	0.0217	0.6841
SCoT-32	700	SCoT32_700	0.0217	0.6841
SCoT-15	700	SCoT15_700	0.0499	0.5696

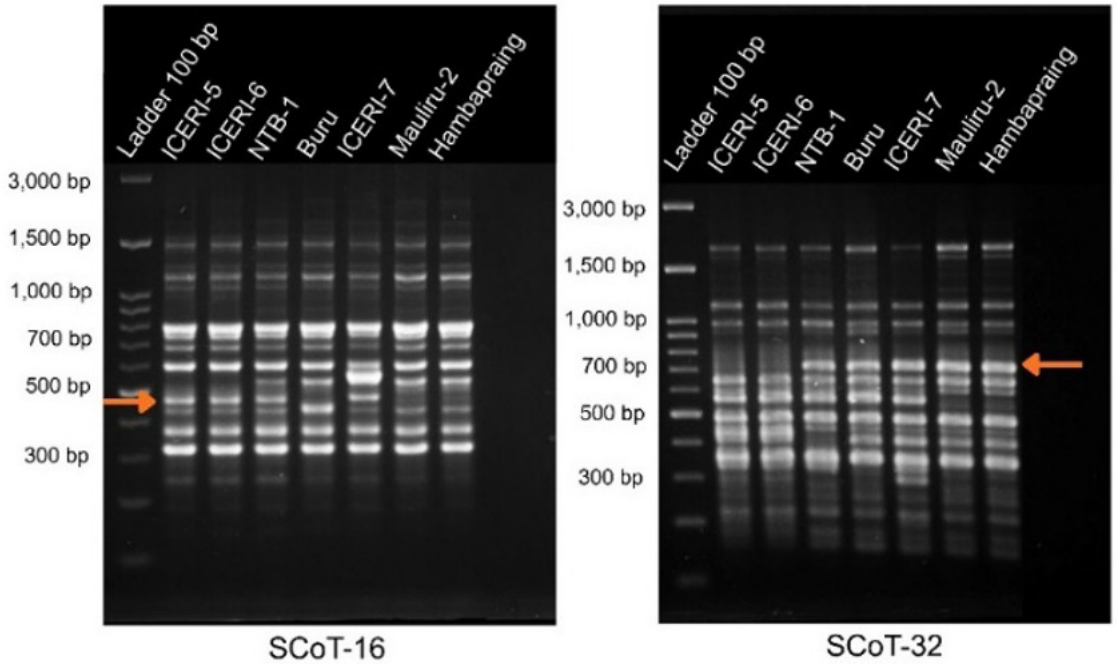


Figure 2. Agarose gel profiles of seven foxtail millet genotypes amplified with SCoT 16 and SCoT 32 primers associated with MPW; bp = base pairs

2, Supplementary Figure 2). Among the seven SCoT markers associated with MPW, only SCoT16_475 and SCoT32_700 displayed consistent, group-specific banding patterns (Figure 2), while SCoT15_700, SCoT21_750, SCoT23_450, SCoT23_950, and SCoT31_625 showed at least one inconsistency (Supplementary Figure 2). SCoT16_475 was consistently detected in low-MPW genotypes (ICERI-5, NTB-1, and ICERI-6), whereas SCoT32_700 appeared exclusively in high-MPW genotypes (Buru, ICERI-7, Mauliru-2, and Hambapraing). Given its specific presence in high MPW genotypes, SCoT32_700 represents a more promising marker candidate for the development of molecular tools targeting high panicle weight in foxtail millet. Accordingly, subsequent analyses focused solely on SCoT32_700 to further evaluate its potential as a yield-associated marker.

3.2. Sequence Characterization and Bioinformatics Analysis of the SCoT32_700 Locus

The sequence of SCoT32_700 has been deposited in GenBank under accession number PZ309115. Nucleotide BLAST against the *Setaria italica* reference genome (*Setaria italica* v2.2) revealed that SCoT32_700 aligns to a region on chromosome VII (Table 2). A subsequent BLAST search against the NCBI database yielded no significant hits, indicating no similarity to any annotated genes. The coding potential assessment using ORF Finder detected ten short ORFs (10-97 amino acids), with only ORF5 showing limited coding capacity. This result was corroborated by CPC2, which assigned a moderate probability (0.5562) of coding potential exclusively to ORF5, whereas the remaining ORFs were predicted as non-protein-coding. PLEK further classified the sequence as 100% non coding and strongly predicted its identity as a long non coding RNA (lncRNA).

Secondary structure prediction with RNAfold yielded a minimum free energy (MFE) of -285.00, indicating a stable RNA structure characteristic of lncRNAs. RNAcentral further identified 33 lncRNA hits, with the best alignment showing 56% identity to known *Setaria* lncRNAs. Of these, nine lncRNAs met the criteria of query coverage above 44% and sequence identity greater than 55% (Table 3), while the remaining hits are listed in Supplementary Table 4.

Nine *Setaria italica* lncRNAs showed 56-60% identity with the locus SCoT32_700, most of which were annotated with GO:0030533 (triplet codon-amino acid adaptor activity). Among these, URS000239627B_4555 exhibited the strongest statistical significance with an E-value of 2E-25 and high query coverage (71.91%), but its relatively large gap percentage (11.74%) reduced confidence in alignment quality. In contrast, URS00023AECFC_4555, together with the cluster of URS00023512BD_4555, URS00023836A4_4555, and URS00023C3035_4555, demonstrated higher identity (≈58-60%), moderate coverage (43-46%), and minimal gaps (~3%), making them the most reliable homologs. Overall, integration of identity, coverage, gap, and E-value metrics highlights URS00023AECFC_4555 and the cluster including URS00023512BD_4555,

Table 2. Summary of sequence-based bioinformatics analysis of SCoT32_700 loci

Analysis	Output value	Result	Interpretation
Nucleotide BLAST core_nt; blastn	N/A	No significant similarity found	No similarity with any annotated genes in the NCBI database
Nucleotide BLAST Genome (*Setaria italica* v2.2)	Max score: 1,136 Total score: 2,251 E-value: 0.0 Percent identity: 96%	*Setaria italica* strain Yugu1 chromosome VII, region 17,772,614 to 17,773,312	SCoT32_700 matches a region in chromosome VII.
ORF Finder	10 ORFs detected	ORF lengths 10–97 amino acids	Only ORF5 has coding potential
CPC2	ORF5 probability 0.5562	ORF5 = coding; others = noncoding	Only ORF5 shows limited coding capacity
PLEK	Score = -1.647560	Sequence 100% non-coding	Sequence predicted as putative lncRNA
RNAfold	MFE = -285.00	Stable RNA structure	Strong secondary structure typical of lncRNA
RNAcentral	33 lncRNA hits	Best hit 56% identity	Sequence resembles known *Setaria* lncRNAs

URS00023836A4_4555, and URS00023C3035_4555 as the strongest candidate homologs to SCoT32_700. At the same time, URS000239627B_4555 remains notable for its statistical significance and GO annotation despite alignment limitations.

While the sequence-based analyses consistently supported the non-coding nature of the locus and predicted it as an lncRNA, its potential relevance as a linked marker warranted further investigation by examining its genomic position and nearby functional genes within the LD-decay region. The BLAST analysis revealed that the fragment SCoT32_700 is located at position 17,772,614 bp on chromosome 7 (Figure 3). Upstream of this locus, transcript *Seita.7G073900.1* was identified at position 17,762,136 bp, approximately 9,898 bp away. Downstream, alternative transcripts *Seita.7G074000.1/2* were detected, spanning from 17,778,663 bp to 17,787,960 bp, with the start site located about 5,351 bp from the SCoT32_700 locus. The downstream gene (*Seita.7G074000.1/2*) is annotated with domains including protein kinase, leucine-rich repeat (LRR_1), and NB-ARC.

3.3. Validation of SCoT Locus Associated with Main Panicle Weight in F_5 Foxtail Millet Lines (ICERI 6 × Botok 4)

Validation of SCoT32_700 was conducted in 20 F_5 lines categorized into low and high MPW groups (Figure 4). Electrophoresis profiles showed that low MPW lines consistently lacked the 700 bp band, whereas several high MPW lines (S25-5, S25-10, S8-7, S8-9, and S16-6) displayed a clear band at this position. The presence of the 700 bp fragment explained ~17.2% of MPW variation (regression value = 0.172, p = 0.0695). These results suggest that SCoT32_700 showed potential as a supportive marker for identifying high-yielding genotypes has potential accuracy in distinguishing yield-related groups, although further validation in segregating populations is required.

Table 3. *Setaria italica* lncRNAs showing sequence similarity to the SCoT32_700 locus, extracted from RNAcentral

ID	E-value	Identity (%)	Query coverage (%)	Target coverage (%)	Gaps (%)	GO term qualifier
URS000239627B_4555	2E-25	56.41	71.91	0.14	11.74	GO:0030533
URS00023512BD_4555	6E-14	57.86	43.67	1.26	3.46	enables triplet
URS0002414B71_4555	6E-14	57.86	43.67	0.81	3.46	codon-amino
URS00023836A4_4555	6E-14	57.86	43.67	1.26	3.46	acid adaptor
URS00023C3035_4555	6E-14	57.86	43.67	1.26	3.46	activity
URS000240CD99_4555	6E-14	57.86	43.67	0.77	3.46	
URS00023AECFC_4555	4.3E-13	59.89	46.73	0.10	7.74	
URS000239F0AB_4555	5.7	57.92	29.35	22.97	10.86	Not available
URS000242DD2D_4555	5.8	56.59	41.59	0.22	7.40	

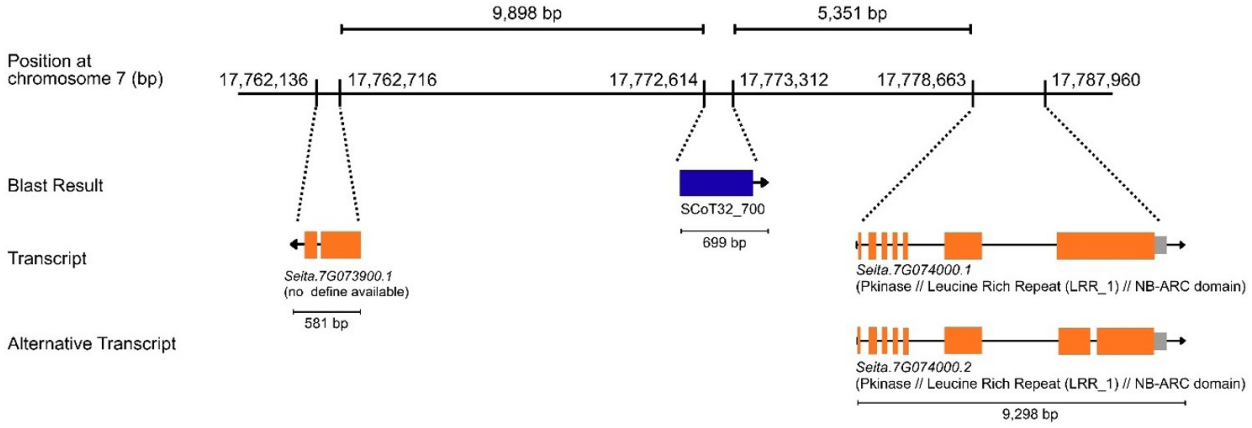


Figure 3. Genomic diagram showing structural positions of sequences and candidate genes within the LD decay region surrounding the SCoT32_700 locus on chromosome 7. The locus is positioned 9,898 bp downstream and 5,351bp upstream of *Seita.7G074000*; bp = base pair

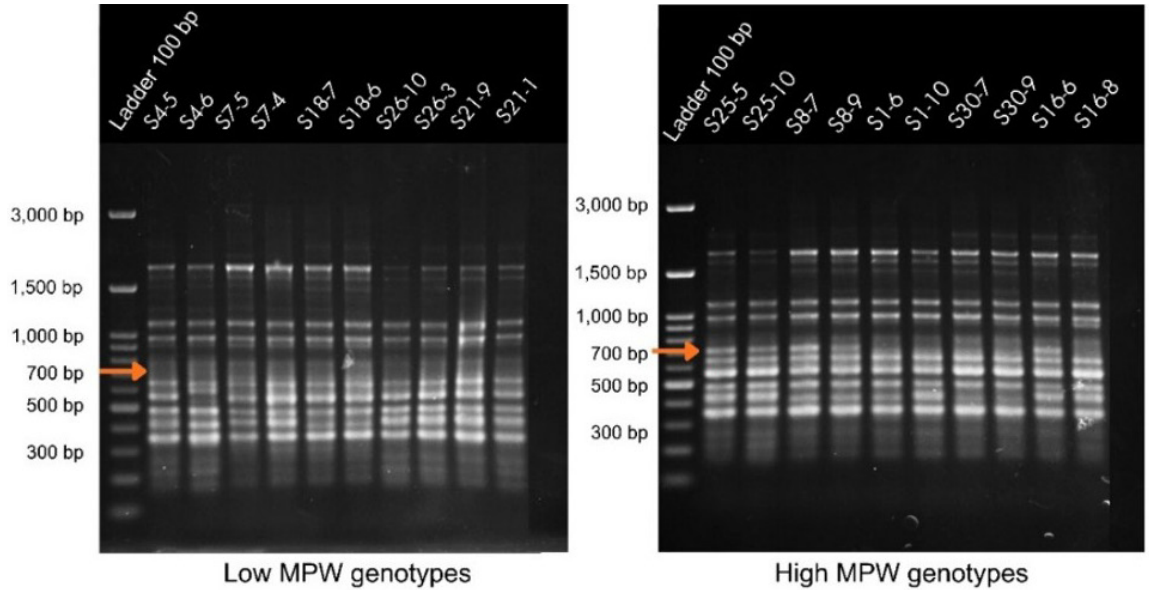


Figure 4. Agarose gel profile generated with SCoT-32 primer for ten low- and ten high-MPW F_5 genotypes; bp = base pairs

4. Discussion

The development of functional markers for yield-related traits is crucial to accelerate molecular breeding in foxtail millet. Main panicle weight (MPW) is a key yield component in cereals, including foxtail millet (Luthfiani *et al.* 2026), rice (Mushtaq *et al.* 2021), and sorghum (Wirnas *et al.* 2021). A genome-wide association study (GWAS) across 407 foxtail millet accessions identified 87 QTL regions, including 15 for panicle weight, with major QTL explaining up to 25.45% of phenotypic variance and narrowing candidate genes such as *Si3g11200*, *Si1g27910*, and *Si7g27560* (Liu *et al.* 2022). The limited variance explained reflects the polygenic nature of panicle traits and the need to uncover additional genes or regions for effective marker development. Identification of SCoT-based loci associated with MPW further demonstrates the potential of functional markers to support yield improvement. In this study, a 700 bp fragment of SCoT-32 was significantly associated with MPW, providing a promising candidate marker for selecting high-yielding foxtail millet genotypes.

Characterization of SCoT32_700 is important to gain deeper insights into its functional relevance. Although *in silico* analyses indicated that SCoT32_700 is non-protein-coding, its lncRNA-like nature is intriguing given emerging evidence that lncRNAs play critical roles in panicle development and yield regulation (Wang *et al.* 2018; Zhou *et al.* 2021; Wang *et al.* 2025). LncRNAs modulate gene expression through mechanisms such as epigenetic recruitment, R-loop formation, alternative splicing, and miRNA sequestration (Zicola 2023; Chanwala *et al.* 2026), and they interact with transcription factors to reshape chromatin architecture (Wang *et al.* 2022). Sequence similarity searches revealed that SCoT32_700 shares homology with 33 lncRNAs from *Setaria*, among which URS00023AECFC_4555 and the cluster including URS00023512BD_4555, URS00023836A4_4555, and URS00023C3035_4555 emerged as the strongest candidates based on identity, coverage, gap, and E-value metrics. Their annotation related to triplet codon–amino acid adaptor activity suggests a potential role in translational modulation, reflecting the classical function of tRNA in decoding mRNA codons and delivering amino acids. Specific tRNA modification patterns are crucial for maintaining translational accuracy (Masuda & Hou 2024).

Beyond functional relevance, SCoT32_700 also has potential as a linked marker. This locus was initially identified through trait–marker association analysis in a random population, which by itself does not guarantee linkage. However, if its genomic positioning falls within the linkage disequilibrium (LD) decay region of genes controlling panicle traits, it gains strong potential as a linkage marker. SCoT32_700 is located 6 kb upstream of *Seita.7G074000.1/2* on chromosome 7, a gene annotated with NB-ARC domains. LD in foxtail millet extends up to 136 kb (Jaiswal *et al.* 2019), supporting possible linkage with *Seita.7G074000.1/2*. While most NB-ARC and LRR domains form NBS-LRR proteins primarily involved in plant defense (Wang *et al.* 2023), a study in rice showed that NB-ARC genes such as *GNP12*

also regulate panicle traits (Pan *et al.* 2022). Since SCoT32_700 shows similarity with lncRNAs annotated with the GO term 'triplet codon–amino acid adaptor activity', it may represent a candidate regulatory locus requiring functional validation of *Seita.7G074000.1/2*. Its proximity to this NB-ARC/LRR gene highlights the need for further investigation into *Seita.7G074000.1/2* and its role in controlling panicle traits in foxtail millet.

To translate this potential into breeding utility, validation in segregating populations is essential. In this study, the marker was evaluated in F_5 lines derived from ICERI 6 × Botok-4, where ICERI 6 has low MPW and Botok-4 has high MPW (Ratnawati *et al.* 2024; Luthfiani *et al.* 2026). Importantly, Botok-4 was not included in the original trait–marker association panel, providing an independent genetic background for validation. Linear regression yielded a *p*-value of 0.0695, approaching the threshold of significance, and an R^2 of 0.172 (17.2%). This regression value suggests that MPW is controlled by multiple loci, with a single marker explaining only a small proportion of the variation .These results suggest that SCoT32_700 is associated with MPW but has not yet been validated as a linked marker. Given the polygenic nature of panicle traits (Liu *et al.* 2022; Li *et al.* 2025), additional markers are needed to capture the full genetic basis of MPW. This aligns with findings in rice, where QTL hotspots for grain traits involve several candidate genes (Satrio *et al.* 2025). While these findings strengthen the case for SCoT32_700, the small number of genotypes used in the trait–marker association analysis remains a limitation. This limitation indicates the necessity of evaluating SCoT32_700 across broader foxtail millet germplasm. Nevertheless, despite the restricted sample size, this study identified a potential lncRNA-like candidate associated with MPW, representing an important contribution to the development of yield-related functional markers in foxtail millet.

Acknowledgements

The authors gratefully acknowledge the dedicated managerial and technical staff of the Plant Molecular Biotechnology-2 Laboratory at IPB University for their invaluable support and contributions throughout this research.

References

Aboul-Maaty, N.A.F.A., Oraby, H.A.S., 2019. Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. *Bulletin of the National Research Centre*. 43, 1-10. https://doi.org/10.1186/s42269-019-0066-1

Ardie, S.W., Nugroho, R.B., Dirpan, A., Anshori, M.F., 2025. Foxtail millet research in supporting climate change resilience efforts: Bibliometric analysis and focused literature review. *Heliyon*. 11, e42348. https://doi.org/10.1016/j.heliyon.2025.e42348

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., Buckler, E.S., 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 23, 2633-2635. https://doi.org/10.1093/bioinformatics/btm308

Chanwala, J., Rosenkranz, I., Kindgren, P., 2026. The functions of long noncoding RNAs in plants. *Current Opinion in Plant Biology*. 89, 102830. https://doi.org/10.1016/j.pbi.2025.102830

Collard, B.C.Y., Mackill, D.J., 2009. Start codon targeted (SCoT) polymorphism: A simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Molecular Biology Reporter*. 27, 86-93. https://doi.org/10.1007/s11105-008-0060-5

Ejigu, D., Pushpalatha, R.S.S., Padil, V., Gangadharan, B., Thendiyath, R.K.S., Upadhyay, G., Harilal, S., 2025. Millets as a dual-purpose crop for sustainable nutritional and energy security: A comprehensive review. *Journal of Agrometeorology*. 27, 245-257. https://doi.org/10.54386/jam.v27i2.2892

Gupta, H., Choudhury, N., Nickhil, C., Deka, S.C., 2025. Exploring the nutritional and bioactive profile, health advantages, starch extraction, modification, and diverse applications of millets: A comprehensive overview. *Cereal Research Communications*. 53, 1949-1968. https://doi.org/10.1007/s42976-025-00660-5

Harish, M.S., Bhuker, A., Chauhan, B.S., 2024. Millet production, challenges, and opportunities in the Asia-Pacific region: a comprehensive review. *Frontiers in Sustainable Food Systems*. 8, 1386469. https://doi.org/10.3389/fsufs.2024.1386469

Jaiswal, V., Gupta, S., Gahlaut, V., Muthamilarasan, M., Bandyopadhyay, T., Ramchiary, N., Prasad, M., 2019. Genome-wide association study of major agronomic traits in foxtail millet (*Setaria italica* L.) using ddRAD sequencing. *Scientific Reports*. 9, 5020. https://doi.org/10.1038/s41598-019-41602-6

Kang, Y.J., Yang, D.C., Kong, L., Hou, M., Meng, Y.Q., Wei, L., Gao, G., 2017. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Research*. 45, W12-W16. https://doi.org/10.1093/nar/gkx428

Li, A., Zhang, J., Zhou, Z., 2014. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinformatics*. 15, 311. https://doi.org/10.1186/1471-2105-15-311

Li, H., Zhu, Q., Wang, T., Zhao, Y., Fan, J., Yang, L., Wang, S., Ma, Y., Chen, L., Li, X., Dong, S., Chu, X., Yuan, X., Wang, J., Gao, L., Yang, G., 2025. Genetic linkage map construction and QTL mapping for panicle-related traits in foxtail millet (*Setaria italica* (L.) P. Beauv.). *Molecular Breeding*. 45, 72. https://doi.org/10.1007/s11032-025-01594-w

Liu, X., Yang, Y., Hou, S., Men, Y., Han, Y., 2022. The integration of genome-wide association study and homology analysis to explore the genomic regions and candidate genes for panicle-related traits in foxtail millet. *International Journal of Molecular Sciences*. 23, 14735. https://doi.org/10.3390/ijms232314735

Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., Hofacker, I.L., 2011. ViennaRNA Package 2.0. *Algorithms for Molecular Biology*. 6, 1-14. https://doi.org/10.1186/1748-7188-6-26

Luthfiani, M.V., Pahlevi, M.R., Purwoko, B.S., Ardie, S.W., 2026. Panicle trait diversity and its impact on the productivity of Indonesian foxtail millet (*Setaria italica* L. Beauv.) Genotypes. *HAYATI Journal of Biosciences*. 33, 566-574. https://doi.org/10.4308/hjb.33.3.566-574

Masuda, I., Hou, Y.M., 2024. A tRNA modification pattern that facilitates interpretation of the genetic code. *Frontiers in microbiology*. 15, 1415100. https://doi.org/10.3389/fmicb.2024.1415100

Mushtaq, D., Kumar, B., Singh, P., Sharma, M., Sheera, A., 2021. Exploring the relationship between yield and yield attributing traits in advanced breeding lines of rice (*Oryza sativa* L.). *International Journal of Environment and Climate Change*. 11, 350-358. https://doi.org/10.9734/ijecc/2021/v11i1230586

Pan, Y.H., Chen, L., Guo, H.F., Feng, R., Lou, Q.J., Rashid, M.A.R., Zhu, X.Y., Qing, D.J., Liang, H.F., Gao, L.J., Huang, C.C., Zhao, Y., Deng, G.F., 2022. Systematic analysis of NB-ARC gene family in rice and functional characterization of *GNP12*. *Frontiers in Genetics*. 13, 887217. https://doi.org/10.3389/fgene.2022.887217

Rai, M.K., 2023. Start codon targeted (SCoT) polymorphism marker in plant genome analysis: current status and prospects. *Planta*. 257, 34. https://doi.org/10.1007/s00425-023-04067-6

Ratnawati, S., Jannah, R.M., Dewi, Y.I., Rizqullah, R., Suwarno, W.B., Ardie, S.W., 2024. The genetic variability of Indonesian local foxtail millet accession based on agro-morphological traits and early salinity tolerance evaluation utilizing *SiDREB2*-based SNAP marker. *HAYATI Journal of Biosciences*. 31, 82-93. https://doi.org/10.4308/hjb.31.1.82-93

Rizwandi, M.A., Pahlevi, M.R., Efendi, D., Ardie, S.W., 2026. Combining marker informativeness and redundancy metrics to refine SCoT primer selection in foxtail millet diversity analysis. *Journal of Tropical Biodiversity and Biotechnology*. 11, jtbb28721. https://doi.org/10.22146/jtbb.28721

Rombel, I.T., Sykes, K.F., Rayner, S., Johnston, S.A., 2002. ORF-FINDER: a vector for high-throughput gene identification. *Gene*. 282, 33-41. https://doi.org/10.1016/S0378-1119(01)00819-8

Salgotra, R.K., Stewart, C.N., 2020. Review: functional markers for precision plant breeding. *Journal of Molecular Sciences*. 21, 4792. https://doi.org/10.3390/ijms21134792

Satrio, R.D., Fendiyanto, M.H., Nurrahmah, N., Anofri, N., Ikhsan, M., Nugroho, S., Hairmansis, A., Miftahudin, M., 2025. Rice QTL hotspots related with seed grain size, shape, weight, and color based on genome wide association study and linkage mapping. *Scientific Reports*. 15, 21470. https://doi.org/10.1038/s41598-025-05814-3

Shukla, P., Aggarwal, P.R., Choudhary, P., Muthamilarasan, M., 2026. Exploring the nutraceutical potential of foxtail millet: a minor millet with a major impact on nutrition security. *Proceedings of the Indian National Science Academy*. 92, 52-62. https://doi.org/10.1007/s43538-025-00402-5

The RNAcentral Consortium., 2019. RNAcentral: a hub of information for non-coding RNA sequences. *Nucleic Acids Research*. 47, D221-D229. https://doi.org/10.1093/nar/gky1034

Wang, X., Jiang, X., Yuan, X., Zhou, Z., Zhao, Z., Wang, L., Jiao, W., Ye, W., Song, Q., 2018. Overexpressing lncRNA LAIR increases grain yield and regulates neighbouring gene cluster expression in rice. *Nature Communications*. 9, 3516. https://doi.org/10.1038/s41467-018-05829-7

Wang, Y., Deng, X.W., Zhu, D., 2022. From molecular basics to agronomic benefits: Insights into noncoding RNA-mediated gene regulation in plants. *Journal of Integrative Plant Biology*. 64, 2290-2308. https://doi.org/10.1111/jipb.13420

Wang, X., Xu, Y., Fan, H., Cui, N., Meng, X., He, J., Ran, N., Yu, Y. 2023. Research progress of plant nucleotide-binding leucine-rich repeat protein. *Horticulturae*. 9, 122. https://doi.org/10.3390/horticulturae9010122

Wang, X., Jiang, X., Yuan, X., Zhou, Z., Zhao, Z., Wang, L., Jiao, W., Ye, W., Song, Q., 2025. Transcriptome-wide association uncovers lncRNAs controlling seed weight in soybean. *Advanced Science*. 12, e16794. https://doi.org/10.1002/advs.202416794

Wirnas, D., Oktanti, N., Rahmi, H.N., Andriani, D., Faturrahman, Rini, E.P., Marwiyah, S., Trikoesoemaningtyas, Sopandie, D., 2021. Genetic analysis for designing an ideotype of high-yielding sorghum based on existing lines performance. *Biodiversitas.* 22, 5286-5292. https://doi.org/10.13057/biodiv/d221208

Zhou, R., Sanz-Jimenez, P., Zhu, X.T., Feng, Z.W., Shao, L., Song, J.M., Chen, L.L., 2021. Analysis of rice transcriptome reveals the LncRNA/CircRNA regulation in tissue development. *Rice*. 14, 14. https://doi.org/10.1186/s12284-021-00455-2

Zicola, J., 2023. The long noncoding RNA family gets bigger: thousands of new loci identified in natural *Arabidopsis* accessions. *The Plant Cell*. 36, 6-7. https://doi.org/10.1093/plcell/koad242

Supplementary Materials

Supplementary Table 1. SCoT primers list used for amplification

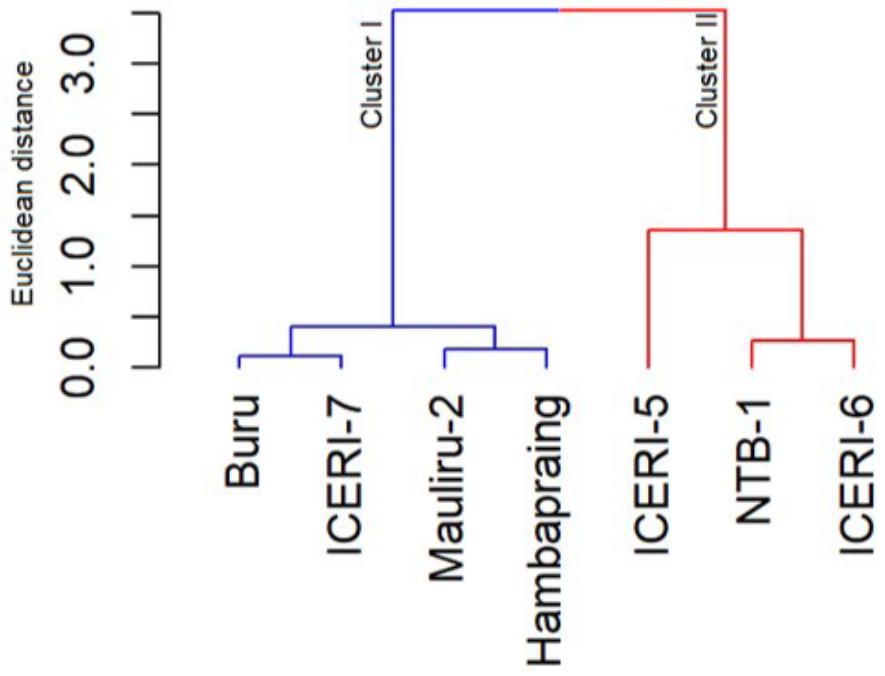
Primer name	Nucleotide base sequence (5'-3')	GC Content (%)	Tm (°C)
SCoT-15	ACGACATGGCGACCGCGA	66.7	62.3
SCoT-16	ACCATGGCTACCACCGAC	61.1	57.0
SCoT-18	ACCATGGCTACCACCGCC	66.7	60.7
SCoT-21	ACGACATGGCGACCCACA	61.1	59.7
SCoT-22	AACCATGGCTACCACCAC	55.6	54.7
SCoT-23	CACCATGGCTACCACCAG	61.1	55.7
SCoT-30	CCATGGCTACCACCGGCG	72.2	65.0
SCoT-31	CCATGGCTACCACCGCCT	66.7	63.8
SCoT-32	CCATGGCTACCACCGCAC	66.7	58.8
SCoT-33	CCATGGCTACCACCGCAG	66.7	58.5
SCoT-34	ACCATGGCTACCACCGCA	61.1	59.3
SCoT-35	CATGGCTACCACCGGCCC	72.2	61.3

Collard & Mackil (2009), Tm: melting temperature

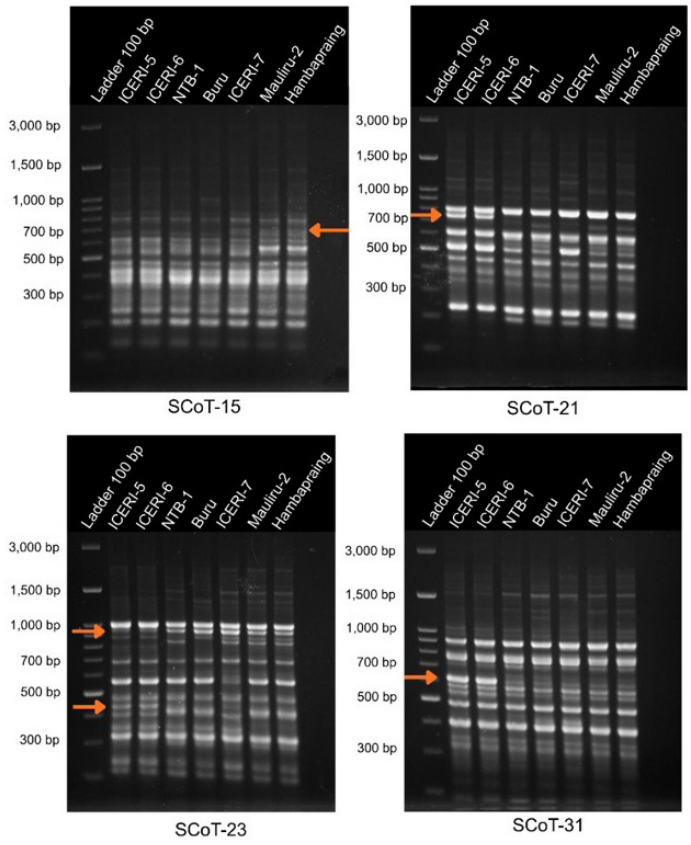
Supplementary Table 2. List of genotypes used for association analysis with main panicle weight (MPW) traits

Genotypes	Main panicle weight (g)
ICERI-5	3.81
ICERI-6	4.90
ICERI-7	7.03
Buru	6.92
Hambapraing	7.33
Mauliru-2	7.15
NTB-1	5.17

Luthfiani *et al.* (2026)



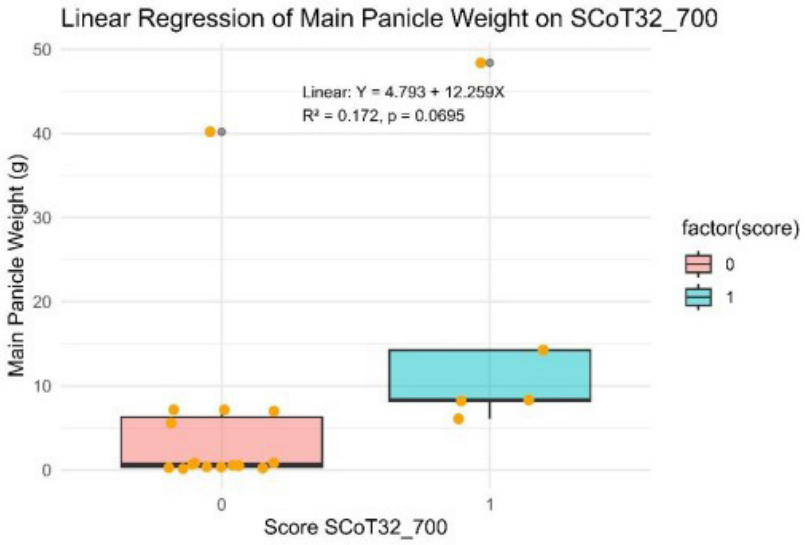
Supplementary Figure 1. Clustering dendrogram based on the Euclidean distance method for main panicle weight traits



Supplementary Figure 2. Agarose gel image obtained using SCoT primers associated with seven foxtail millet genotypes; bp = base pairs

Supplementary Table 3. F_5 foxtail millet genotypes derived from the ICERI-6 × Botok-4 cross used for the validation of SCoT loci associated with main panicle weight (MPW)

Genotype	Code	MPW (g)	MPW category
I6B4-7-281-273-3.12-3.5	S4-5	0.38	low
I6B4-7-281-273-3.12-3.6	S4-6	0.56	low
I6B4-7-370-53-3.7-3.5	S7-5	0.24	low
I6B4-7-370-53-3.7-3.4	S7-4	0.28	low
I6B4-7-271-155-3.16-3.7	S18-7	0.35	low
I6B4-7-271-155-3.16-3.6	S18-6	0.83	low
I6B4-7-240-326-3.10-3.10	S26-10	0.16	low
I6B4-7-240-326-3.10-3.3	S26-3	0.67	low
I6B4-7-209-132-3.10-3.9	S21-9	0.53	low
I6B4-7-209-132-3.10-3.1	S21-1	0.79	low
I6B4-7-240-325-3.3-3.5	S25-5	8.33	high
I6B4-7-240-325-3.3-3.10	S25-10	6.07	high
I6B4-7-360-201-3.10-3.7	S8-7	14.25	high
I6B4-7-360-201-3.10-3.9	S8-9	8.21	high
I6B4-7-292-249-3.3-3.6	S1-6	7.16	high
I6B4-7-292-249-3.3-3.10	S1-10	5.59	high
I6B4-7-224-3-3.13-3.7	S30-7	7.16	high
I6B4-7-224-3-3.13-3.9	S30-9	6.99	high
I6B4-7-375-158-3.11-3.6	S16-6	48.4	high
I6B4-7-375-158-3.11-3.8	S16-8	40.2	high



Supplementary Figure 3. Linear regression analysis of main panicle weight on SCoT32_700

Supplementary Table 4. lncRNA dataset of *Setaria italica* extracted from RNA central

ID	E-value	Identity	Query coverage (%)	Target coverage (%)	Gaps (%)
URS000239627B_4555	2E-25	56.41	71.91	0.14	11.74
URS00023512BD_4555	6E-14	57.86	43.67	1.26	3.46
URS0002414B71_4555	6E-14	57.86	43.67	0.81	3.46
URS00023836A4_4555	6E-14	57.86	43.67	1.26	3.46
URS00023C3035_4555	6E-14	57.86	43.67	1.26	3.46
URS000240CD99_4555	6E-14	57.86	43.67	0.77	3.46
URS00023AECFC_4555	4.3E-13	59.89	46.73	0.10	7.74
URS000239F0AB_4555	5.7	57.92	29.35	22.97	10.86
URS000242DD2D_4555	5.8	56.59	41.59	0.22	7.40
URS000238A75B_4555	15	54.58	34.77	4.38	13.36
URS00023DA38B_4555	16	54.41	34.63	1.40	13.41
URS000236F2FE_4555	23	65.38	14.33	0.07	5.77
URS000241AD24_4555	23	65.38	14.33	0.06	5.77
URS000242922F_4555	23	65.38	14.33	0.06	5.77
URS00023E0068_4555	23	65.38	14.33	0.04	5.77
URS00023619DC_4555	23	65.38	14.33	0.06	5.77
URS000237B821_4555	23	65.38	14.33	0.06	5.77
URS00023FCB8A_4555	23	65.38	14.33	0.04	5.77
URS00023CDE93_4555	23	65.38	14.33	0.04	5.77
URS00023E6B56_4555	23	65.38	14.33	0.07	5.77
URS000238FA19_4555	23	65.38	14.33	0.04	5.77
URS00023D5968_4555	1500	55.63	40.47	0.87	6.95
URS00023EF6B5_4555	1500	55.63	40.47	0.78	6.95
URS00023E14D7_4555	9600	59.04	11.40	20.91	1.20
URS00023D22FF_4555	36000	59.82	15.16	0.03	5.36
URS00023E45E4_4555	60000	63.86	11.13	0.11	4.82
URS00023770A3_4555	82000	62.00	13.49	0.47	6.00
URS00024220B9_4555	120000	60.40	13.77	0.13	4.95
URS000236AACD_4555	120000	60.40	13.77	0.15	4.95
URS00023636DE_4555	120000	60.40	13.77	0.13	4.95
URS00023CE997_4555	120000	60.40	13.77	0.58	4.95
URS00023E9C9C_4555	120000	60.40	13.77	0.12	4.95
URS0002371608_4555	170000	64.79	9.74	0.34	1.41