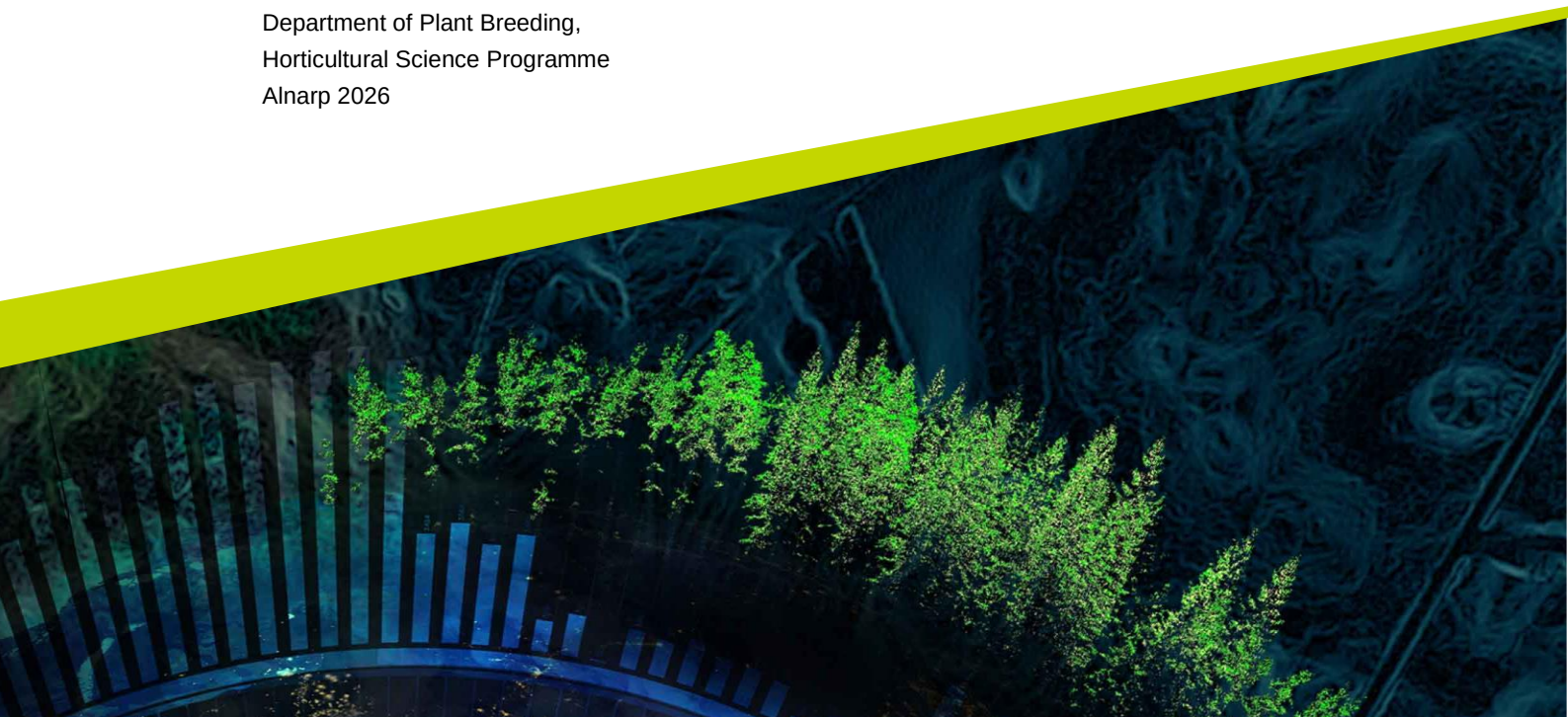
Genetic dissection of tiller traits in timothy grass (*Phleum pratense L.)* using genome-wide association analysis

Chigoziri Ekhuemelo

Degree project/Independent project • 30 credits
Swedish University of Agricultural Sciences, SLU
Faculty of Landscape Architecture, Horticulture and Crop Production Science,
Department of Plant Breeding,
Horticultural Science Programme
Alnarp 2026



Genetic dissection of tiller traits in timothy grass (*Phleum pratense* L.) using genome - wide analysis

Genetisk kartläggning av skottegenskaper hos timotej (*Phleum pratense* L.) med hjälp av genomomfattande analys

Chigoziri Ekhuemelo

Supervisor:	Yousef Rahimi, SLU, Department of Plant Biology, Plant Genomics and Plant Breeding**.**
Assistant supervisor:	Girma Bedada,,SLU, Department of Plant Biology, Plant Genomics and Plant Breeding
Examiner:	Mulatu Geleta Dida, SLU, Department of Plant Breeding.
Credits:	30 credits
Level:	Second cycle, A2E, Master's thesis
Course title:	Independent Project in Horticultural Science, A2E
Course code:	Ex 0948
Programme/education:	Horticultural Science Programme
Course coordinating dept:	Department of Plant Breeding
Place of publication:	Alnarp
Year of publication:	2026
Keywords:	*Phleum pratense*, timothy grass, tiller traits, GWAS, functional annotation, association mapping

Swedish University of Agricultural Sciences

Faculty of Landscape Architecture, Horticulture and Crop Production Science

Department of Plant Breeding

Abstract

Timothy grass (*Phleum pratense* L.) is a cool-season perennial forage valued for its high herbage yield, nutritional quality, and good digestibility. Despite its agricultural importance, the genetic architecture underlying tiller-related traits remains poorly understood. This study aimed to identify genomic regions and candidate genes associated with tiller composition and biomass-related traits using genome-wide association analysis (GWAS). Previously generated phenotypic data from a diverse panel of domesticated (cultivars, landraces, and breeding lines) and non-domesticated (wild and semi-wild) timothy accessions were utilized. Brief assessment confirmed substantial variation among traits, with vegetative tillers accounted for the majority of total tillers (86.61%), while elongated and generative tillers represented smaller proportions (9.51% and 3.88%, respectively). The lowest variability was recorded for vegetative tillers (16.83%), whereas generative tillers exhibited extremely high variability (191.01%). Several traits, including the proportion of generative tillers ($H^2 = 0.623$), vegetative tillers ($H^2 = 0.592$), and generative tiller length ($H^2 = 0.546$), showed moderate heritability, indicating partial genetic control. A total of approximately 29,000 single nucleotide polymorphisms (SNPs) were used for the GWAS. No significant marker–trait associations (MTAs) were detected for total tiller number (TTN) or total dry weight (TDW) at the Bonferroni threshold, whereas six MTAs were identified at a lower suggestive threshold ($-\log10(P) \geq 4$), mainly on chromosome 2PpB.This suggests strong environmental influence on TTN and TDW. However, multiple MTAs were identified for tiller composition (the percentage of vegetative, elongating, and generative tillers) and their dry matter across different significance thresholds. Of these, 10 MTAs exceeded the Bonferroni threshold and were mainly located on chromosomes 1PpC and 2PpB. Functional annotation of genes located near significant SNPs revealed candidate genes involved in protein binding, transport, transcriptional regulation, photosynthesis, and protein turnover. Additional annotations, including ABC transporters and UDP-glycosyltransferases, indicate involvement in metabolite transport, detoxification, and cellular homeostasis. For dry matter traits, candidate genes included those related to pectinesterase activity, DNA recombination, transmembrane transport, and F-box proteins involved in protein degradation and stress responses. A candidate gene associated with elongated tiller dry weight was identified near a significant SNP on chromosome 7PpB, suggesting a potential locus influencing biomass accumulation. Overall, this study provides new insights into the genetic basis of tiller-related traits in timothy grass and highlights genomic regions that may support future research and breeding efforts aimed at improving biomass production and adaptation.

Keywords: Phleum pratense, timothy grass, tiller traits, GWAS, functional annotation, association mapping.

Table of Contents

List of tables

.

List of Figures

Abbreviations

Abbreviation	Description
BAM	Binary alignment map
BED	Browser Extension Data
BLUES	Best linear unbiased estimators
BWA-MEM	Burrows wheeler aligner-maximal exact matches
CV	Coefficient of variation
DNA	Deoxyribonucleic acid.
Eggnog	Evolutionary genealogy of genes: Non-supervised orthologous groups
ELONG_L	Length of elongated tillers.
ELONG_DW	Dry weight of elongated tillers.
ELONG_TN_P	Percentage of elongated tillers.
FarmCPU	Fixed and Random Model Circulating Probability Unification
GATK	Genome Analysis Toolkit
GBS	Genotype by Sequencing
GEN_L	Length of generative tillers.
GEN_DW	Dry weight of generative tillers.
GEN_TN_P	Percentage of generative tillers.
gVCF	Genomic variant call format
GWAS	Genome wide association studies
H^2	Broad sense heritability.
NORDGEN	Nordic genetic resource center
MTA	Marker trait associations
PFAMs	Protein families
REMAP	Retrotransposon-microsatellite amplified polymorphism
REML	Restricted maximum likelihood
SD	Standard deviation
SLU	Swedish University of Agricultural Sciences
SNPs	Single nucleotide polymorphisms
TDW	Total dry weight.
VEG_L	Length of vegetative tillers.
VEG_DW	Dry weight of vegetative tillers.
VEG_TN_P	Percentage of vegetative tillers.

1. Background

Timothy (*Phleum pratense* L.) is a temperate grass of the Gramineae family grown for use as livestock feed. It is one of the most important perennial forage grasses in Northern Europe and in Sweden (Ivehag, 2018; Issa 2022). Timothy grass has high nutritive value making it a desirable forage crop in the temperate regions where it is popularly grown as a perennial grass with the possibility of several harvests of the tillers within its lifespan (Issa 2022; Rahimi *et al*. 2023). Its popularity is based on its winter hardiness, ability to adapt to different photoperiods and palatability for livestock

Variations in forage quality among timothy grass genotypes are primarily determined by growth habit, tillering capacity and developmental stages. These characteristics are closely linked to the genetic origin of the plant. Like other forage and cereal crops productivity in timothy grass is related to the tillering traits. Increase in biomass production and yield in timothy is dependent on the growth stage and the relative contribution of vegetative (VEG), elongated (ELONG), and generative (GEN) tillers (Rahimi, 2024).

Traditional selection based on tiller phenotypes is labour-intensive, slow, and ineffective while conventional breeding efforts have primarily emphasized visible traits like yield and quality, overlooking the genetic mechanisms driving tiller behaviour that are essential for long-term productivity (Alqudah et al., 2021; Bajgain et al., 2019; Rahimi *et al*., 2023).
Previous studies have shown substantial phenotypic diversity among domesticated and non-domesticated timothy accessions as well as closely related species of timothy grass such as *P. nodosum* and *P. alpinum.* These species differ in key traits like growth, biomass, and developmental timing, with timothy generally exhibiting higher growth and flowering earlier than wild accessions, indicating selection for development timing rather than just yield (Rahimi et al., 2023; Rahimi, 2024).

Genome wide association studies (GWAS) are genomic analysis approaches used to identify genetic markers associated with traits of interest in germplasm collections for use in breeding programmes. Genetic characterization of timothy grass remains largely untapped and there are currently no genomic studies identifying the genomic regions in control of tiller traits and no mapping of tiller characteristics across diverse populations of timothy in the Nordic regions and beyond. This absence of genetic insights hinders marker-assisted selection and genomic prediction for enhancing tiller output and total forage productivity (Bajgain et al., 2019; Rahimi *et al*., 2023).

The availability of genome-wide sequence data for multiple genotypes per accession provides an opportunity to link phenotypic variation in tiller traits to underlying genetic variation using GWAS approaches.

1.2 Objectives

The aim of the study was

(i) to characterize phenotypic variation in tiller-related traits between domesticated (cultivars, landraces, breeding lines) and non-domesticated (wild and semi-wild) and among groups of timothy accessions.

(ii) To identify genomic regions associated with tillering traits using GWAS.

(iii) To Identify potential candidate genes underlying significant marker trait associations (MTAs).

The study results will contribute to better understanding of tiller traits and facilitate their improvement through marker assisted selection ultimately enhancing timothy grass productivity.

2. Materials and Methods

2.1 Plant material and source of dataset

The study was conducted using previously generated phenotypic and genotypic data from timothy accessions obtained from the Nordic Genetic Resource Center (NordGen) gene bank and evaluated in an earlier greenhouse study (Rahimi, 2024). The dataset comprised of 212 timothy grass accessions representing the Northern European gene pool (Denmark, Finland, Germany, Iceland, the Netherlands, Norway, Russia, Sweden and United Kingdom). The plant material included cultivars, breeding lines, landraces, semi-wild, and wild accessions. In the original experiment, eight plants per accession were grown in a climate chamber for eight weeks at the Phytotron facility, SLU Uppsala. Four plants per accession were subsequently cloned and grown in a greenhouse for 14 days, after which two clones per plant were vernalized at 4 ℃ for 42 days. After vernalization, four genotypes per accession were evaluated, each with two clonal replicates, under controlled greenhouse conditions at SLU, Uppsala. Phenotypic data recorded included total dry weight (TDW), total tiller number (TTN), proportions of vegetative (VEG), elongated (ELONG), and generative (GEN) tillers, as well as dry biomass of each tiller type and their length.

2.1. 2 Sample collection for DNA extraction

Young leaf tissues were collected from four genotypes per accession from 832 greenhouse-grown plants obtained after removing low quality samples and immediately frozen in liquid nitrogen and stored at −80 °C. Genomic DNA was extracted using the DNeasy Plant Pro Kit with bead-beating, following the manufacturer's protocol with minor adjustments. Briefly, 100 mg of frozen tissue was homogenized using a TissueLyser II, followed by lysis with CD1 buffer. Subsequent steps followed the standard protocol. DNA concentration and purity were assessed using a NanoDrop 2000, and a subset of samples was verified on 1.5% agarose gels prior to library preparation.

2.2. Library preparation and genotyping-by-sequencing (GBS)

Library preparation was performed by first immobilizing Tn5 transposase on streptavidin-coated magnetic beads at a ratio of 0.3 µL Tn5 per 5 µL beads. The beads were washed with streptavidin-binding and dialysis buffers and subsequently stored at 4 ℃ according to the protocol described by Kucka (2020). Tagmentation was carried out by combining genomic DNA with tagmentation buffer and Tn5-loaded beads, followed by incubation at 55 ℃ to simultaneously fragment the DNA

and integrate sequencing adapters. Following tagmentation, DNA fragments were released from the beads using a wash buffer containing 0.6% SDS. Library amplification was then performed using Q5 High-Fidelity DNA Polymerase and Nextera primer sets to extend adapter sequences and enrich the libraries. PCR cycling conditions were optimized to ensure efficient amplification and proper fragment processing. The resulting libraries were assessed by agarose gel electrophoresis to verify fragment size distribution. Subsequently, libraries were purified and size-selected using AMPure XP beads, and DNA was eluted in nuclease-free water. Library concentrations were quantified using the AccuBlue® High Sensitivity dsDNA Quantitation Kit. Prepared libraries were pooled into four batches and submitted to Novogene (Germany) for 150 bp paired-end sequencing (PE150). To increase sequencing depth and improve the accuracy of downstream variant calling, all samples were sequenced twice.

2.3 Phenotypic variation and broad sense heritability

The phenotypic diversity of the domesticated and non-domesticated timothy accessions was determined through analysis of variance, mean comparisons and Pearson correlation analysis in RStudio Version 4.5.2 (Posit 2026). Best linear unbiased estimators (BLUEs) were calculated for all measured traits using linear mixed models implemented in R. For each trait, the following model was fitted:

$$y=\mu+Taxa+Block+\varepsilon$$

where Taxa (genotype) was treated as a fixed effect and Block as a random effect. Models were fitted using the lmer function from the lme4 package with restricted maximum likelihood (REML). BLUEs for each genotype were extracted as estimated marginal means using the emmeans package. These adjusted means were subsequently used for downstream analyses, including GWAS.

Broad-sense heritability (H^2) was estimated for each trait using a separate linear mixed model fitted with restricted maximum likelihood (REML) in R using the lmer function from the lme4 package. Genotype (Taxa) and replicate were treated as random effects. Variance components for genotypic and residual effects were extracted from the fitted models, and broad-sense heritability was calculated as:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_e^2}{r}}$$

where σ^2G represents the genotypic variance, σ^2e the residual variance, and r the number of replicates per genotype. In addition, descriptive statistics including mean, standard deviation, minimum, maximum, median, and coefficient of variation were calculated for each trait.

2.4 Read pre- processing and alignment

Raw paired-end sequencing reads were processed using fastp (Chen, 2025) to remove adapter sequences and low-quality bases. Quality control was performed both before and after trimming by evaluating read quality scores, GC content, sequence length distribution, and the presence of adapter contamination. Trimmed reads from each sequencing run were then aligned independently to the reference genome. Prior to alignment, the reference genome was indexed using BWA v0.7.18 (Li, 2013) and samtools v1.20, and a sequence dictionary was generated according to the Genome Analysis Toolkit (GATK) Best Practices workflow. Sequence alignment was performed using the BWA-MEM algorithm (Li, 2013), with read group information incorporated into each alignment to preserve metadata related to sample identity and sequencing run. The resulting BAM files were sorted and indexed using samtools. For each sample, BAM files generated from the two sequencing runs were merged, and BAM headers were standardized to ensure consistency among read groups. The merged BAM file was subsequently used for all downstream analyses. Alignment quality was assessed using Qualimap, which provided metrics including mapping quality, genome coverage, and duplication rate.

2.5 Variant discovery

Single nucleotide variant (SNV) calling was performed using the Genome Analysis Toolkit (GATK) v4.3.0.0 (Poplin et al., 2017; Van der Auwera and O'Connor, 2020) following the GATK Best Practices workflow. Variants were initially called for each sample individually using HaplotypeCaller, generating genomic VCF (gVCF) files. The resulting gVCF files were imported into a GenomicsDB database using GenomicsDBImport and subsequently jointly genotyped with GenotypeGVCFs to produce a multisample VCF file.

To obtain a set of high-quality SNVs, hard filtering was applied using GATK VariantFiltration with the following criteria: QD < 2.0, FS > 60.0, QUAL < 30.0, MQ < 40.0, ReadPosRankSum < −8.0, SOR > 3.0, and MQRankSum < −12.5. Insertions and deletions (indels) were excluded, and only SNPs were retained for downstream analyses. Following hard filtering, variants were further evaluated based on mean sequencing depth per site and per individual, genotype quality, and levels of missing data. SNPs with more than 20% missing data across individuals were excluded, while samples with more than 10% missing genotypes were removed from the dataset.

Additional soft filtering was then applied to retain only bi-allelic markers with a minor allele frequency (MAF) greater than 0.01. A MAF threshold of 0.01 was chosen to retain low-frequency variants that may contribute to population genetic structure and genetic diversity while still excluding extremely rare variants that are more likely to represent sequencing or genotyping errors. After quality control and filtering, a final dataset comprising 29,783 SNPs across 655 samples was retained for downstream analyses. Of the original 832 samples, 177 were excluded due to high levels of missing data or failure to meet quality control criteria.

2.6 Genome-wide association study

Genome-wide association studies (GWAS) were performed using the Fixed and Random Model Circulating Probability Unification (FarmCPU) method implemented in the rMVP package in R (Yin et al., 2021). BLUEs obtained from mixed linear models were used as phenotypic values for the GWAS analysis. To account for population structure and genetic relatedness among genotypes, principal components (PCs) and a kinship matrix were calculated using the rMVP package and incorporated into the model. Associations between SNP markers and phenotypic traits were visualized using Manhattan plots, where SNPs were plotted according to their chromosomal positions on the x-axis and the negative logarithm of the association P-values on the y-axis. Horizontal reference lines indicating significance thresholds were drawn at $-\log10(P) = 3$ ($P = 1 \times 10^{-3}$), $-\log10(P) = 4$ ($P = 1 \times 10^{-4}$), and the Bonferroni-corrected significance threshold.

2.7 Candidate gene identification and functional annotation

Significant SNPs identified by GWAS using FarmCPU ($-\log10P \geq 4$) were selected for candidate gene analysis. SNP genomic coordinates were converted to BED format and intersected with gene models from the reference genome annotation. Candidate genes were defined as genes located within a ±2 kb interval surrounding each significant SNP. This window was selected based on the estimated linkage disequilibrium (LD) decay in another subproject (Rahimi et al. 2026, Unpublished), where LD declined to $r^2 = 0.1$ within approximately 450–789 bp across the B, C, and D sub-genomes, with a genome-wide average of approximately 540 bp. The ±2 kb interval therefore exceeded the observed LD decay distance and was considered sufficient to capture genes likely to be in linkage disequilibrium with associated markers while minimizing the inclusion of unrelated loci.

Gene identifiers and associated annotation information were extracted and curated to generate a non-redundant candidate gene set. Protein sequences corresponding to candidate genes were retrieved and functionally annotated using eggNOG to infer putative gene functions, orthologous groups, functional categories, and Gene Ontology (GO) annotations.

3. Results

3.1 Phenotypic variation and broad-sense heritability of tiller traits

Summary statistics and broad-sense heritability estimates for the studied tiller traits are presented in Table 1. Substantial phenotypic variation was observed across all traits among the timothy accessions, indicating considerable diversity in tiller production. In terms of tiller composition, vegetative tillers (VEG_TN_P) dominated, accounting for an average of 86.61% of total tillers, whereas elongated (ELONG_TN_P) and generative tillers (GEN_TN_P) represented smaller proportions, with means of 9.51% and 3.88%, respectively.

Differences were also observed in biomass-related traits. Vegetative tillers exhibited the highest average dry weight (VEG_DW = 21.17), compared to elongated (ELONG_DW) and generative tillers (GEN_DW). In contrast, generative tillers were generally taller, as reflected by their higher mean length (GEN_L), compared to both vegetative (VEG_L) and elongated (ELONG_L) tillers.

The standard deviation values across all traits further showed the large phenotypic variability within the studied population. This is supported by the coefficients of variation (CV), which ranged from relatively low variability in VEG_TN_P (16.83%) to extremely high variability in GEN_TN_P (191.01%), suggesting heterogeneous distribution particularly for generative tiller proportion.

Broad-sense heritability (H^2) estimates indicated low to moderate genetic control of the studied traits. Heritability values ranged from 0.280 for elongated tiller length (ELONG_L) to 0.623 for the proportion of generative tillers (GEN_TN_P). Traits such as VEG_TN_P ($H^2 = 0.592$) and GEN_L ($H^2 = 0.546$) also exhibited relatively higher heritability, suggesting a stronger genetic control, whereas other traits may show more pronounced environmental effects.

Table 1. Summary statistics and broad-sense heritability estimate for the studied tiller traits among timothy accessions

Trait	N	Mean	SD	Min	Max	Median	CV(%)	H^2
TTN	1231	39.21	19.27	2.00	125.00	37.00	49.15	0.36
VEG_TN_P	1231	86.61	14.57	5.00	100.00	90.48	16.83	0.59
ELONG_TN_P	1231	9.51	11.25	0.00	95.00	6.45	118.22	0.47
GEN_TN_P	1231	3.88	7.41	0.00	51.06	0.00	191.01	0.62
VEG_DW	1231	21.17	14.93	0..23	122.00	18.00	70.52	0.40
ELONG_DW	778	16..41	15.48	0.24	111.97	11.06	94.33	0.29
GEN_DW	437	12.45	11.36	0.22	80.15	9.01	91.24	0.36
TDW	1231	35.96	26.95	0..23	195.15	30.51	74.94	0.46
VEG_L	1231	63.17	15.93	11.67	118.33	63.33	25.21	0.30
ELO_L	780	106.64	26.89	53.33	213.33	101.67	25.22	0.28
GEN_L	438	145.74	33.05	75.0	242.50	145.00	22.68	0.55

(N) the number of observations, (SD) standard deviation, (Min) minimum, (Max) maximum, (CV%) coefficient of variation, and (H^2) broad-sense heritability

3.2 Variation in tiller formation and biomass production across accession groups

The phenotypic variation of total tiller number (TTN) and total dry weight (TDW) across different groups of timothy (*Phleum pratense* L.) accessions is presented in Figure 1(a) and (b). The results in Figure 1(a) show significant differences in the total tiller numbers produced by the different groups of timothy grass accessions. The highest number of tillers were observed in the breeding lines and wild accessions while the lowest number of tillers were observed in landraces.

The total dry weight which corresponds to dry biomass was significantly different among the various timothy groups with the highest total dry weight observed in cultivars and wild accessions (Figure 1b).

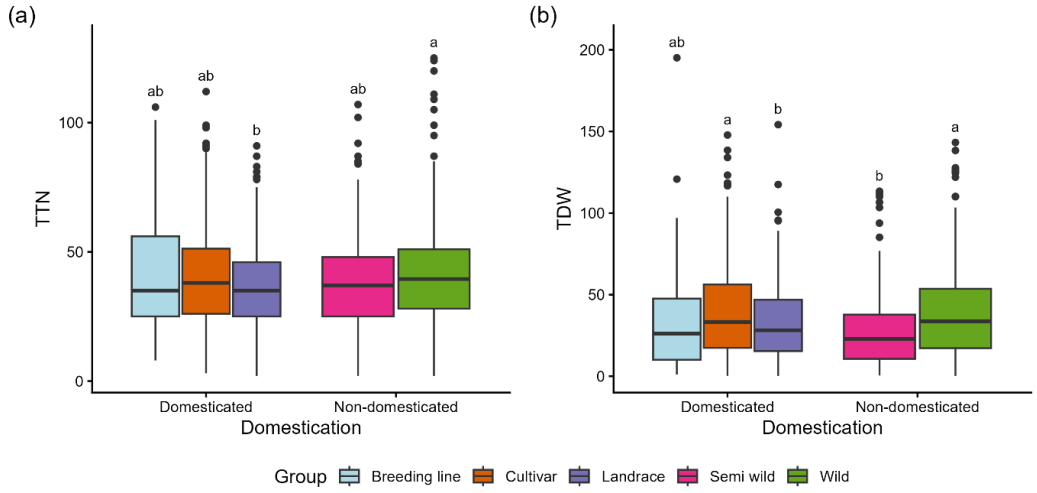


Figure 1. Phenotypic variation of total tiller number (TTN) and total dry weight (TDW) across different groups of timothy (*Phleum pratense* L.) accessions. Boxes represent the interquartile range (IQR), and horizontal lines within boxes indicate medians. Different letters above boxplots indicate significant differences among groups based on Tukey's honestly significant difference (HSD) test at $P < 0.05$.

3.3 Variation in the proportion of tiller types across accession groups

The phenotypic variation in the proportions of the three tiller types (vegetative, elongated, and generative tillers) across different groups of timothy accessions is shown in Figure 2a. The vegetative tillers were generally higher than all other tillers hence the inverted boxplot. Significant differences were observed in the proportion of vegetative tillers between domesticated and non-domesticated accessions. Semi-wild accessions exhibited the highest proportion of vegetative tillers, whereas cultivars showed the lowest proportion.

The highest elongated tiller and generative tiller percentages were observed in the cultivars and the wild accessions (Figure 2b, 2c).

Considering the vegetative dry weight, the result shows variation between the non-domesticated and the domesticated accessions. The wild accessions produced significantly more vegetative dry weight (biomass) compared with the breeding lines and cultivars. Although the vegetative dry weight produced by the semi wild accessions and the landraces did not differ significantly from each other, they produced vegetative dry weight that were statistically similar to both the wild and the breeding lines (Figure 2d).

There were no significant differences observed in the elongated and generative dry weight among and within various groups of timothy accessions studied. However, the domesticated timothy grass accessions had higher elongated and generative dry weight. Furthermore, the different groups of accessions did not show significant differences in the length of the three tiller types. Nevertheless, the landraces had the tallest vegetative tillers, the breeding lines produced the tallest elongated tillers and the cultivars produced the tallest generative tillers. (Figure 2e).

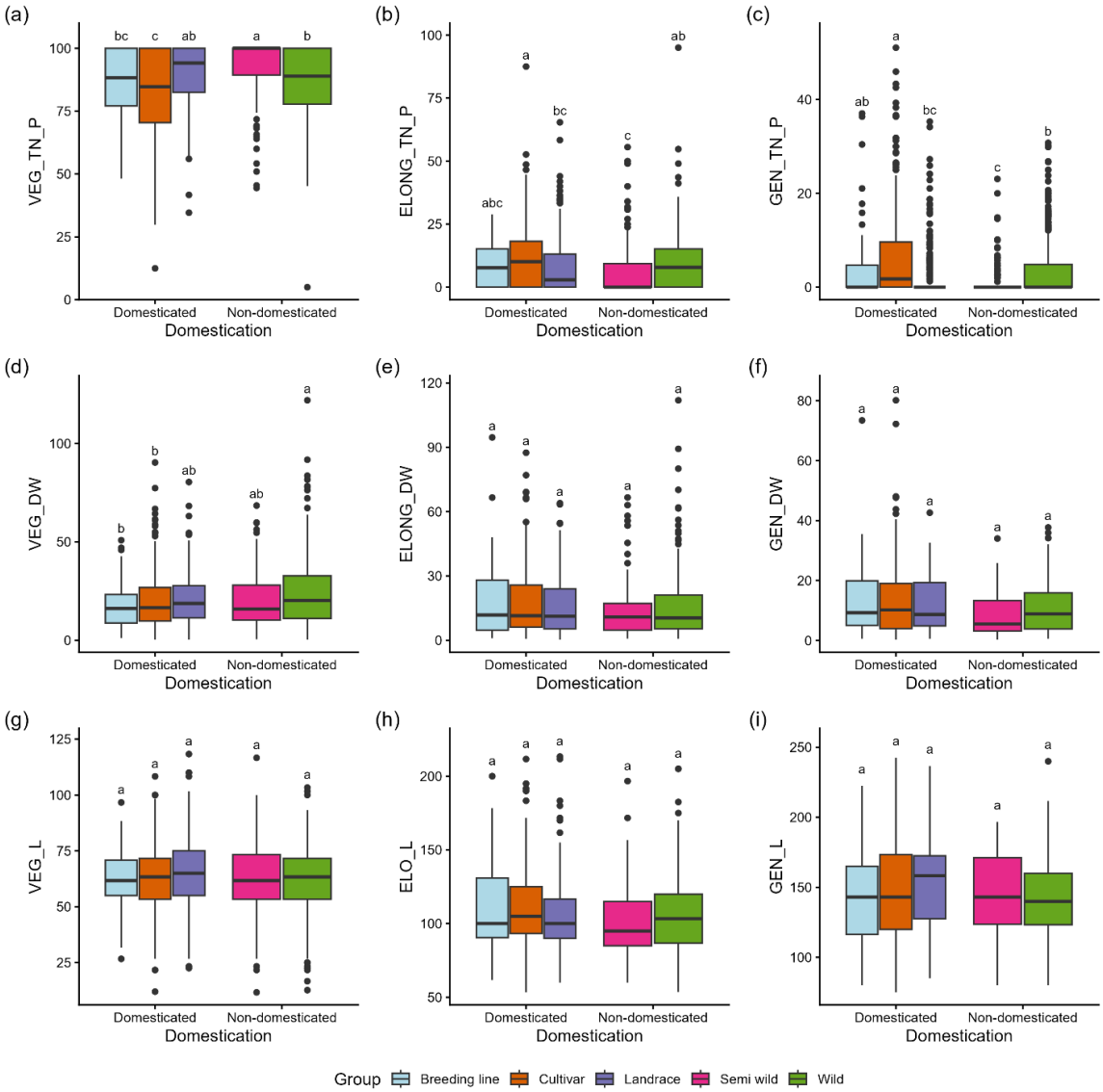


Figure 2.Phenotypic variation of vegetative (VEG), elongated (ELONG), and generative (GEN) tillers across different groups of timothy (*Phleum pratense* L.) accessions. Boxes represent the interquartile range (IQR), and horizontal lines within boxes indicate medians. Different letters above boxplots indicate significant differences among groups based on Tukey's honestly significant difference (HSD) test at $P < 0.05$.

3.4 Pairwise phenotypic correlations among evaluated traits

Pearson correlation analysis shows the pairwise relationship among the phenotypic traits across all accessions (Figure 3a), domesticated (Figure 3b) and non-domesticated (Figure 3c) group of accessions.

TTN in all accessions irrespective of the group (domesticated or non-domesticated) exhibited a strong positive correlation with vegetative dry weight (all accessions:

r=0.64***, domesticated accessions: r=0.60***, and non-domesticated accessions: r=0.68***. TTN also exhibited strong positive correlation with TDW of all group (all accessions: r=0.61***, domesticated accessions: r=0.63*** and non-domesticated accessions r=0.59***). Similarly, TDW showed strong positive relationships with elongated tiller dry weight (ELONG_DW: r=0.78*** and vegetative tiller dry weight (VEG_DW: r=0.72***) in all accessions suggesting that elongated and vegetative tillers contributed to the total yield of timothy accessions and groups. In contrast, the percentage vegetative tillers (VEG_TN_P) exhibited a weak negative correlation with most of the traits tested across all accessions. However, TDW had moderate positive correlation with proportions of elongated tillers (ELONG_TN_P: r= 0.49***) and generative tillers (GEN_TN_P: r=0.31***).

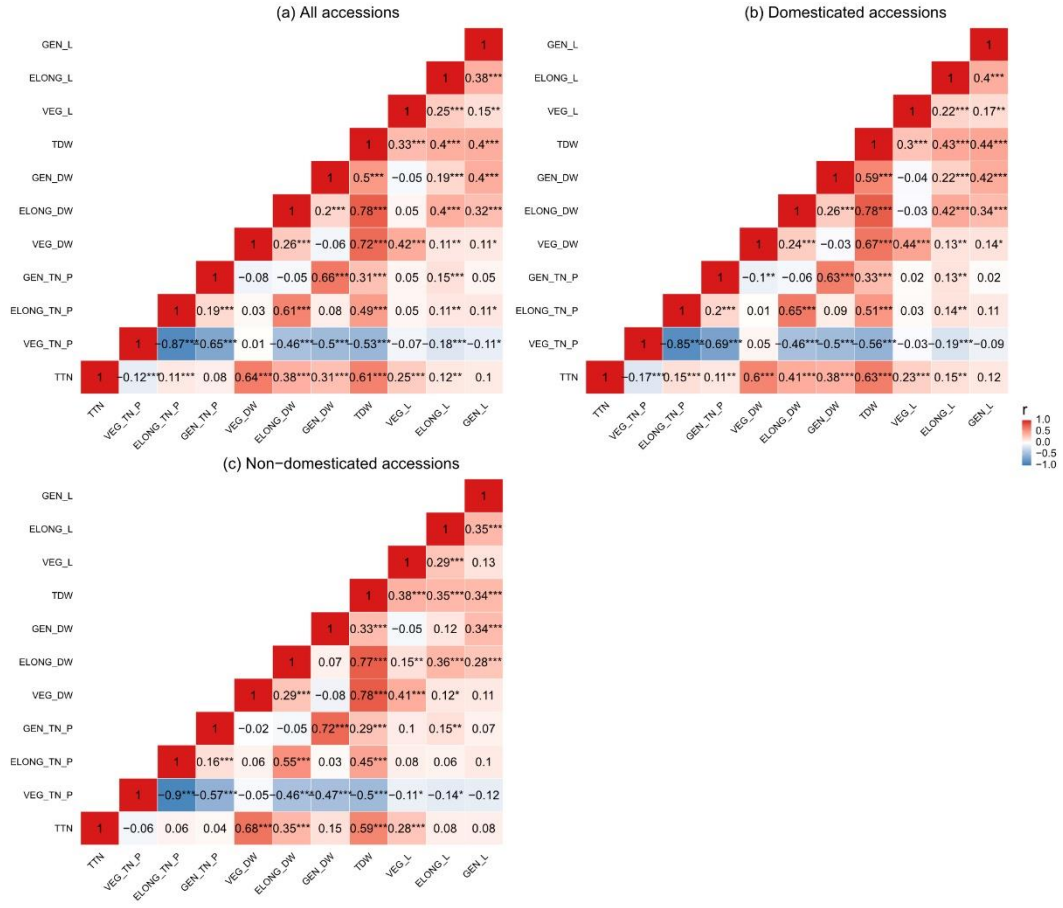


Figure 3. Pearson correlation heatmap showing pairwise relationships among phenotypic traits evaluated across timothy accessions. The lower triangle displays correlation coefficients (r), with colour intensity representing the magnitude and direction of association (blue = negative correlation; red = positive correlation), Correlations were considered significant at $p < 0.05$, with significance levels denoted as *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

3.5 Genome wide association analysis and distribution of significant SNPs

A total of 276 MTAs were identified for the studied traits at the relaxed significance threshold of $P< 1 \times 10^{-3}$. The results presented in Figure 4 show the distribution of associated SNPs across chromosomes under different significance threshold. The number and distribution of trait associated SNPs varied notably depending on the threshold applied.

At the most relaxed threshold ($P< 1 \times 10^{-3}$), 276 SNPs were detected across the genome. These associations were unevenly distributed, with clear clustering on Chromosome 1 and 2. In particular, chromosome such as 1PpC and 2PpB consistently showed the highest number of associated SNPs, each with more than 30 SNPs (Figure 4).

Additional clusters of associations were observed on chromosomes 3PpB, 1PpB and 2PpC suggesting that multiple loci contributed to tiller traits variation. When a more stringent threshold ($P< 1 \times 10^{-4}$) was applied, the number of associated SNPs decreased substantially to 62 SNPs, representing a smaller set of more robust associations. Despite the reduction, Chromosomes/sub genome 1PpC and 2PpB remained prominent, indicating that these regions likely harbour loci with stronger genetic effects.

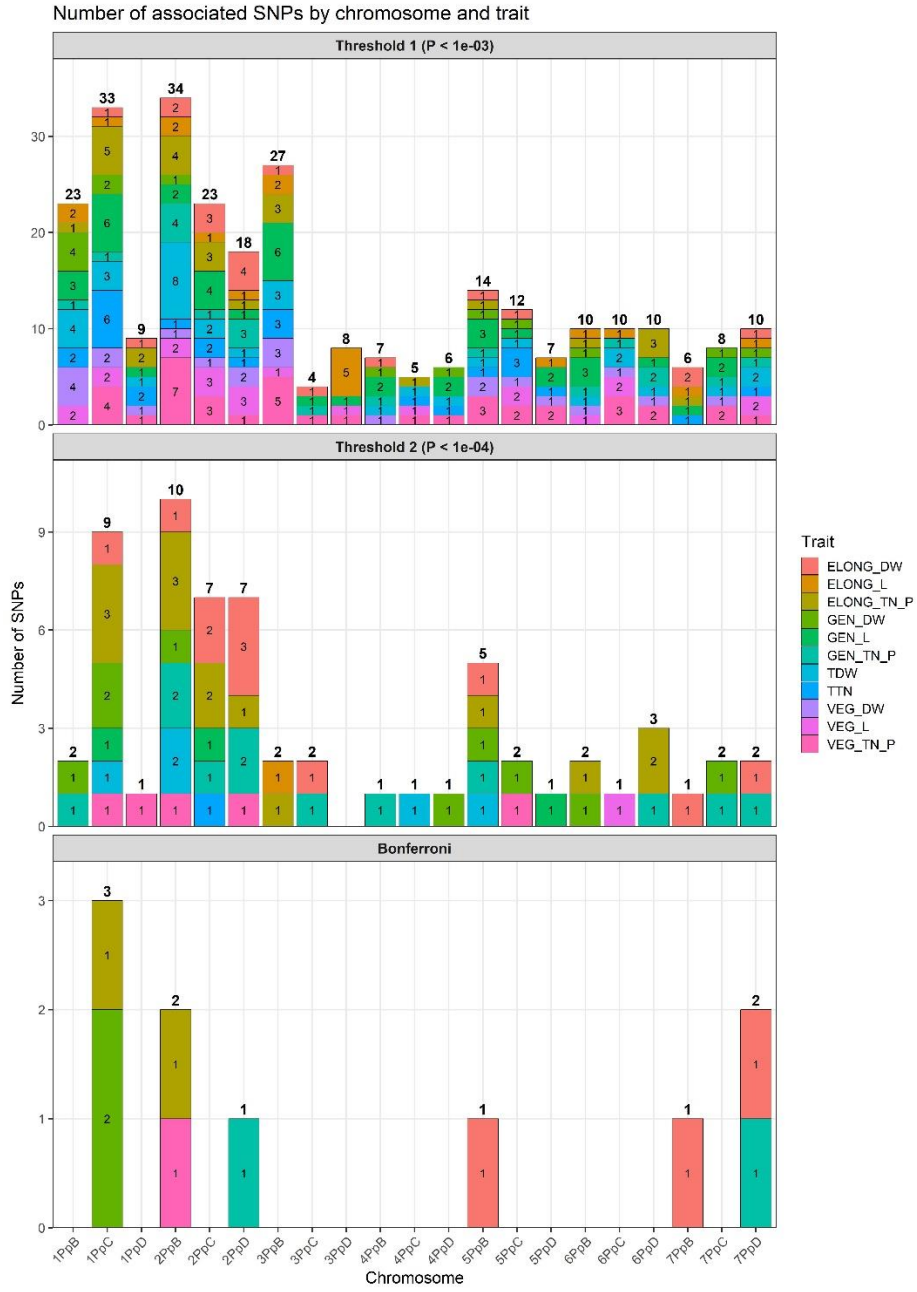


Figure 4.: Distribution of associated SNPs across chromosomes at different significance threshold. Stacked bar plots show the number of SNPs associated with each trait per chromosome at P < 1 × 10-3, P < 1 ×10 -4, and Bonferroni-corrected threshold respectively. Numbers within bars indicate SNP counts per trait and totals per chromosome are shown above each bar.

3.6. GWAS with tiller traits

3.6.1 Markers associated with tiller number and dry biomass

Genome-wide association analysis revealed a number of markers linked to total tiller numbers (TTN) and total dry weight (TDW) at the significance threshold (-

log 10(p) ≥ 3). Specifically, a total of 26 SNPs were associated with TTN while 34 SNPs were identified for TDW distributed across multiple chromosomes (Figure 5).

Increasing the stringency to -log 10(P) ≥ 4 reduced the number of significant associations with only one SNP remaining for TTN and 5 SNPs for TDW, indicating that most associations detected at the lower threshold were of moderate significance. At the Bonferroni-corrected threshold, no significant SNPs were detected for either TTN or TDW, suggesting that none of the associations surpassed the strict genome- wide significance level after correction for multiple testing. The Manhattan and Quartile-quartile (QQ) plots are shown in Figure 6.

Overall, the association analysis indicates that while several genomic regions maybe involved in the genetic control of TTN and TDW, their effect are likely small to moderate.

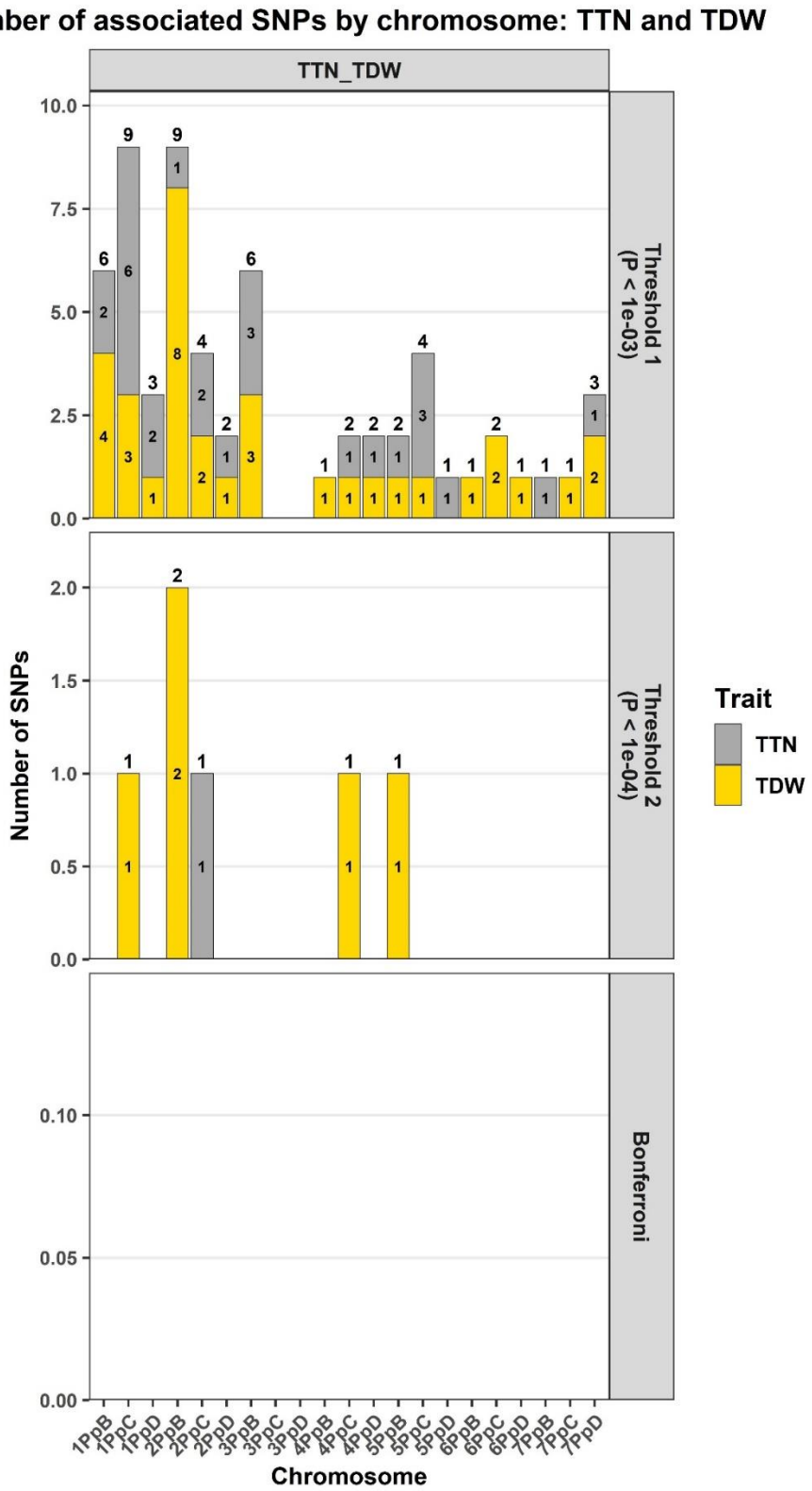


Figure 5. Number of SNPs associated with total tiller number (TTN) and total dry weight (TDW) by Chromosome.

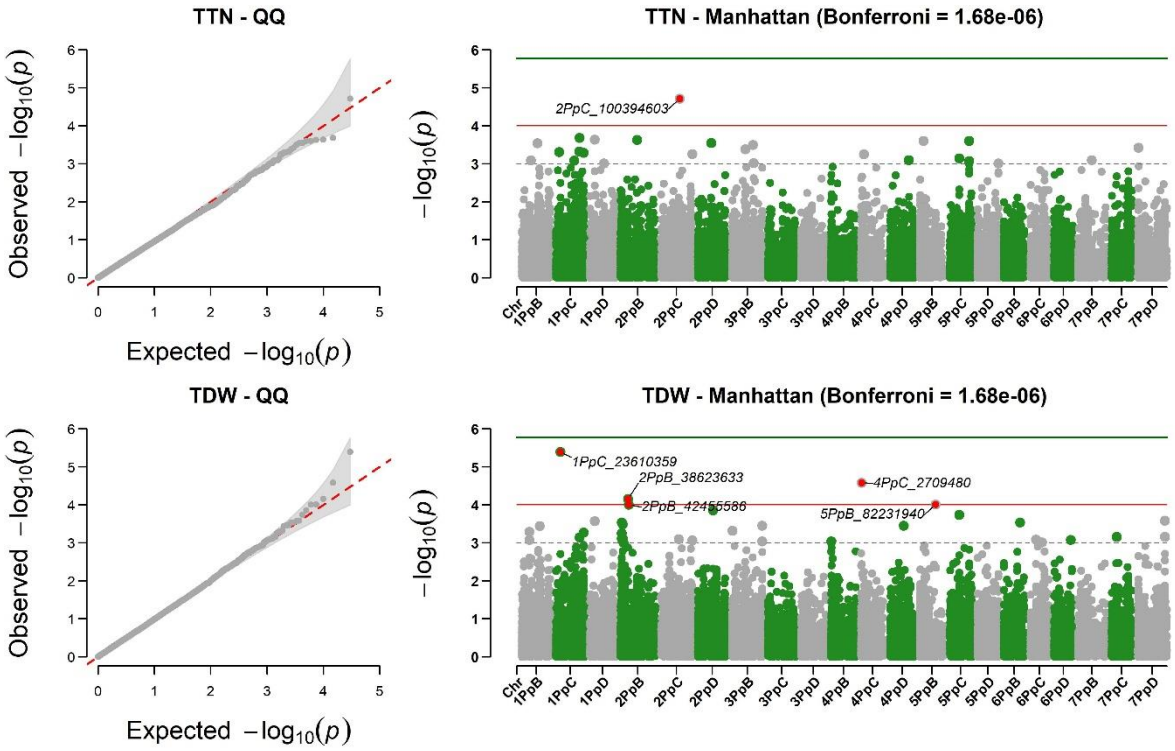


Figure 6. Genome-wide association results for total tiller number (TTN) and total dry weight (TDW). Quartile-quartile (QQ) plots (left) and Manhattan plots (right) are shown for each trait. Horizontal lines indicate the suggestive significance thresholds ($P < 1\times 10^{-3}$, $P < 1 \times 10^{-4}$) and the Bonferroni-corrected threshold.

3.6.2 Markers associated with tiller types

The GWAS result for the percentage tiller types at the -log 10(P) ≥ 3) threshold showed a total of 37 SNPs associated with the proportion of VEG tillers while a total of 26 SNPs and 20 SNPs were associated with the proportion of ELONG and GEN tillers respectively (Figure 7).

At a stringent threshold -log 10(P) ≥ 4 the number of significant SNPs reduced to 5 SNPs for proportion of VEG, 14 SNPs for ELONG and 12 SNPs for GEN.

At the Bonferroni-corrected threshold only one significant SNP was associated with the proportion of VEG, tillers while the proportion of ELONG and GEN tillers were associated with 2 significant SNPs. The Manhattan and Quartile-quartile (QQ) plots for markers associated with the proportion of tillers are shown in Figure 8.

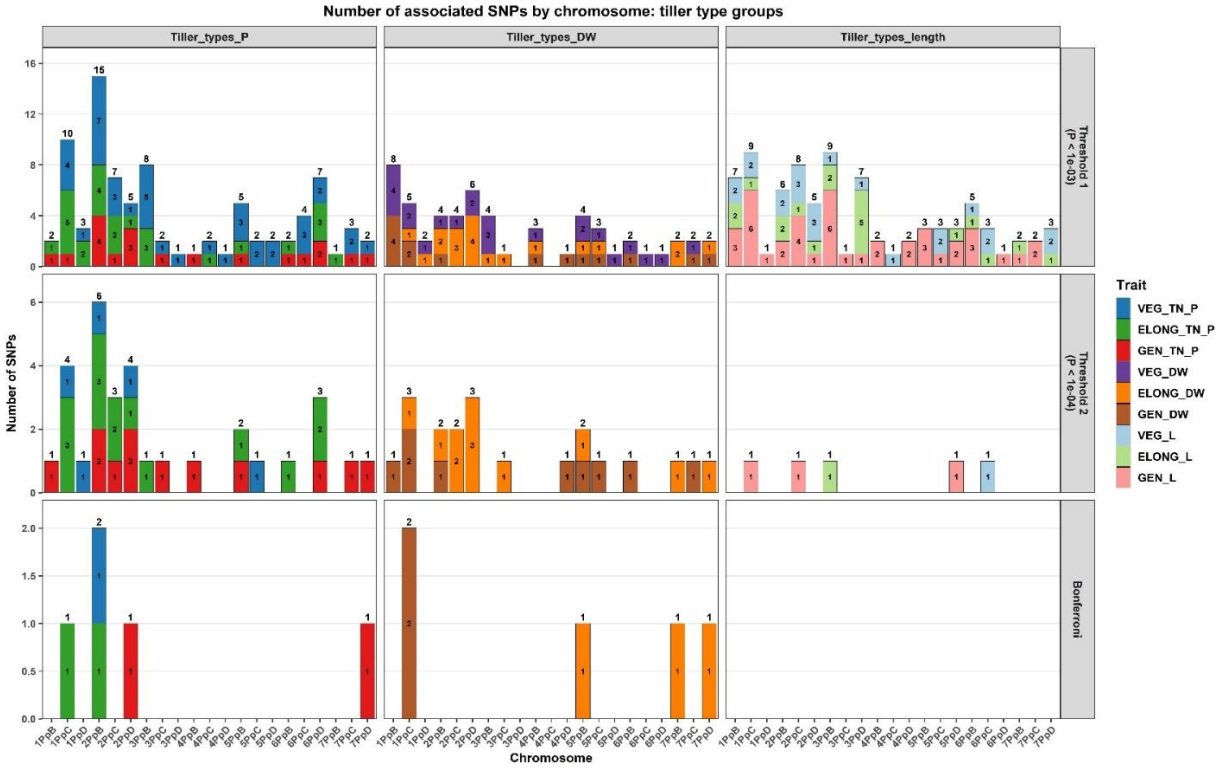


Figure 7: Number of SNPs associated with the proportions, dry matter and length of the three tiller types at different significance thresholds.

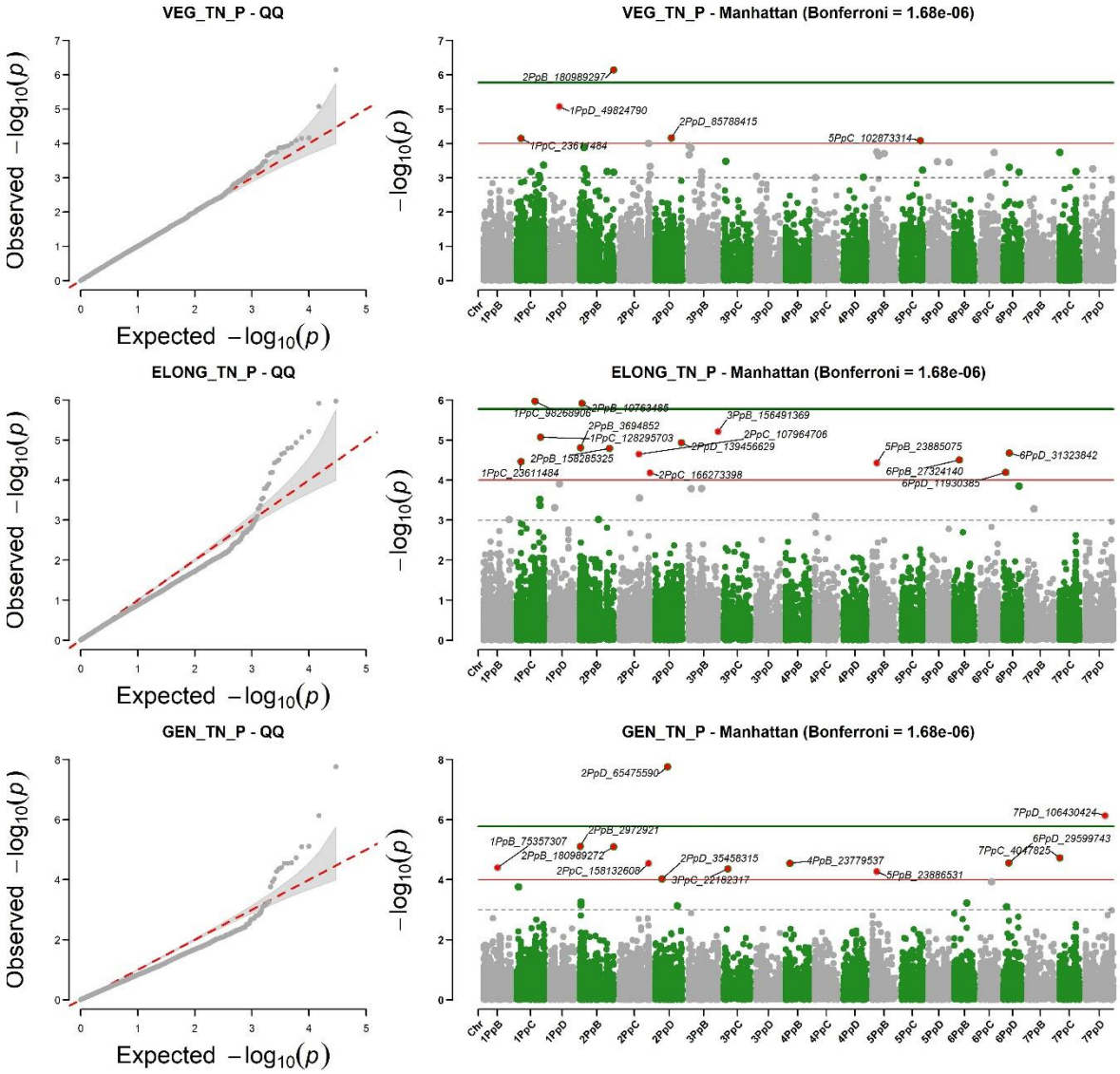


Figure 8. Genome-wide association results for the percentage of VEG (VEG_TN_P), ELONG (ELONG_TN_P), GEN (GEN_TN_P) tillers. Quartile-quartile (QQ) plots (left) and Manhattan plots (right) are shown for each trait. Horizontal lines indicate the suggestive signific thresholds ($P < 1 \times 10^{-3}$, $P < 1 \times 10^{-4}$) and the Bonferroni-corrected threshold.

3.6.3 Markers associated with dry biomass from various tiller types

Genome wide association result for the dry matter from three tiler types at -log 10 (P) ≥ 3 threshold revealed a total of 21 significant SNPs associated with biomass of VEG tillers, 30 significant SNPs associated with biomass of ELONG tillers and 12 significant SNPs associated with biomass of GEN tillers (Figure 9).

There was no significant SNP associated with biomass of VEG tillers. At the -log 10(P) ≥ 4 threshold, whereas 8 and 5 significant SNPs were associated with the dry matter of ELONG and GEN tillers respectively.

At the Bonferroni-corrected threshold, biomass from ELONG tillers had 3 significant SNPs located on 5PpB, 7PpB and 7PpD while biomass of GEN tillers had 2 significant SNPs located on 1PpC. The results suggests that these genomic regions have a moderate to high effect on dry matter accumulation of the three tiller types in timothy grass.

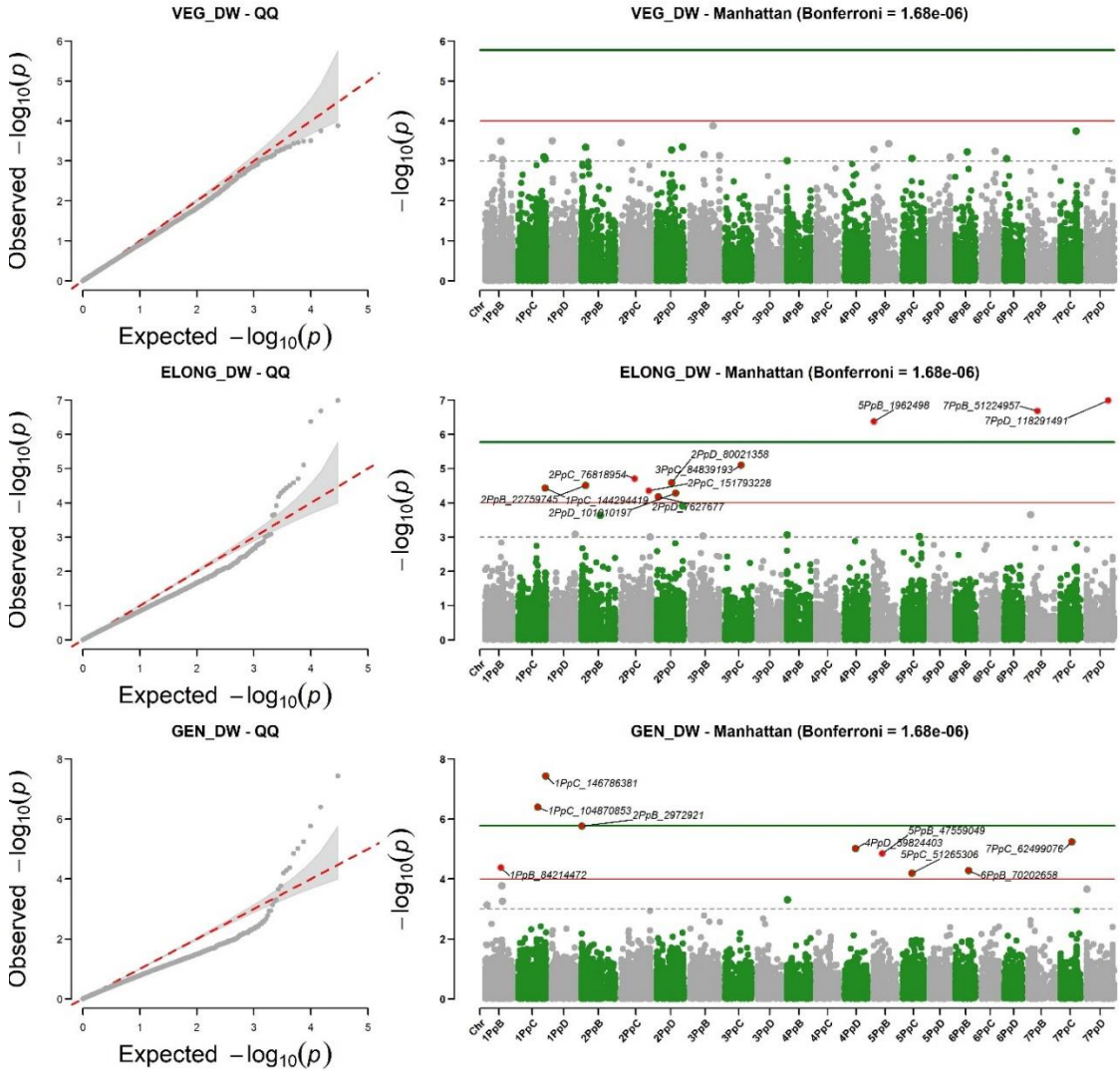


Figure 9. Genome-wide association results for the dry matter of VEG (VEG_DW), ELONG (ELONG_DW), and GEN (GEN_DW), tillers. Quartile-quartile (QQ) plots (left) and Manhattan plots (right) are shown for each trait. Horizontal lines indicate the nominal significance threshold ($P < 1 \times 10^{-3}$, $P < 1 \times 10^{-4}$) and the Bonferroni-corrected threshold.

3.6.4 Markers associated with length of tiller types

Genome-wide association analysis revealed multiple clusters of marker trait associations for the length of the three tillers at threshold of -log 10(P) ≥ 3. A total of 22 SNPs were associated with the length of VEG, tillers, 16 SNPs were identified for the length of ELONG tillers while 43 SNPs distributed across multiple chromosomes were associated with the length of GEN tillers (Figure 8). One significant SNP each was associated with the length of VEG and ELONG tillers while 3 significant SNPs were associated with the length of GEN tillers at the -log 10(p) ≥ 4 significant thresholds (Figure 10). No SNPs associated with the length of any of the tiller types at the Bonferroni-corrected threshold although several associations surpassed the suggestive threshold ($P < 1 \times 10^{-3}$ and $P < 1 \times 10^{-4}$). This suggests that the genetic control of VEG, ELONG and GEN tiller length is likely polygenic and may involve loci with small to moderate effects.

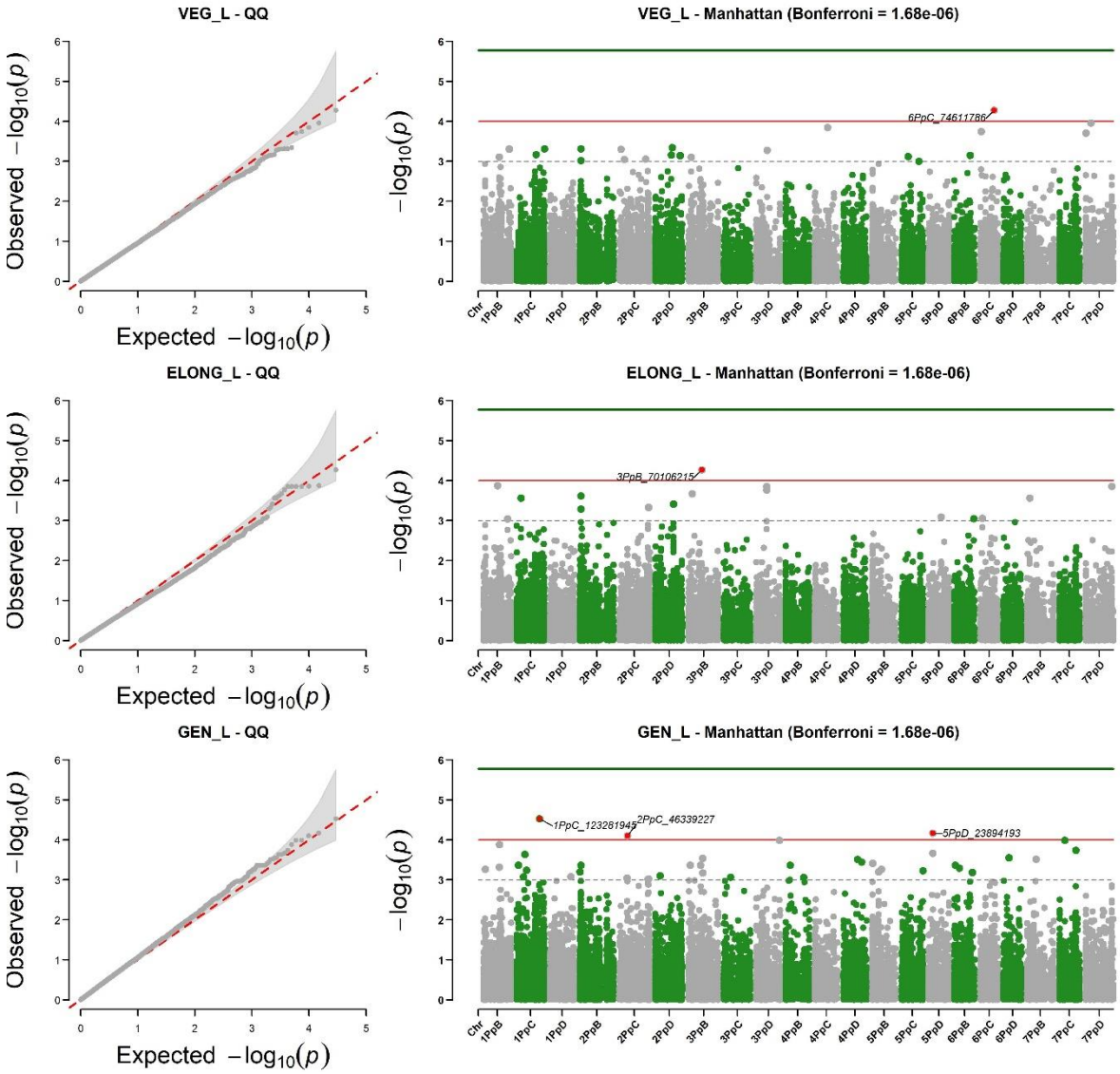


Figure 10. Genome-wide association results for the length of VEG (VEG_L), ELONG (ELONG_L), GEN (GEN_L), tillers. Quartile-quartile (QQ) plots (left) and Manhattan plots (right) are shown for each trait. Horizontal lines indicate the nominal significance threshold ($P < 1 \times 10^{-3}$, $P < 1 \times 10^{-4}$) and the Bonferroni-corrected threshold.

3.7 Functional annotation of candidate genes

Functional annotation of genes located near significant SNPs identified several candidate genes with predicted roles in protein binding, disease resistance, transport, transcriptional or post-transcriptional regulation, photosynthesis, and protein turnover (Table 2). Several candidate genes were annotated as proteins of unknown function, including genes such as DUF563, DUF1191, and DUF1668 domains. For traits related to tiller number, several candidate genes were associated with domains involved in stress response and signaling. For example, candidate genes near significant SNPs for ELONG_TN_P and GEN_TN_P included annotations such as NB-ARC, PPR, UDP-glycosyltransferase, ABC transporter transmembrane region, BTB/MATH, and ribosomal protein S11. The presence of NB-ARC and disease-resistance-related genes suggests that some loci may be involved in defense-related signaling pathways, which can indirectly influence growth and reproductive development. Linkage with ABC transporter and UDP-glycosyltransferase genes indicate possible roles in metabolite transport, detoxification, and cellular homeostasis.

Candidate genes linked with dry biomass produced from ELONG (ELONG_DW) and GEN (GEN_DW) tillers included genes related to disease resistance, DNA recombination, pectinesterase activity, transmembrane amino acid transport, transposase-related domains, and F-box proteins. The identification of a pectinesterase-associated gene suggests a possible role in cell wall modification, which may affect biomass accumulation and tissue development. F-box proteins are commonly involved in protein degradation through ubiquitin-mediated pathways and may regulate developmental or stress-response processes associated with dry matter production. One notable candidate gene associated with ELONG_DW was annotated as a photosystem II reaction center-related protein.

This gene was located near a significant SNP on chromosome 7PpB (7PpB_51224957) and may be involved in photosynthetic electron transport. Because photosynthesis directly contributes to biomass accumulation, this locus may represent a biologically relevant candidate for dry weight-related variation.

For tiller length-related traits, including length of VEG (VEG_L) and GEN (GEN_L) tillers, fewer significant loci were detected compared with tiller number and tiller dry biomass traits. Candidate genes identified for these traits included genes annotated as PPR repeat-containing proteins, genes with unknown function, and loci associated with general regulatory or organellar processes. PPR proteins are often involved in RNA processing and organelle gene expression, suggesting that mitochondrial or chloroplast-related regulation may contribute to variation in plant length traits.

Table 2. Functional annotation of candidate genes (within ± 2kb) underlying significant MTAs identified by GWAS

Trait	SNP	CHR	POS	GeneID	Gstart	Gend	Description	PFAMs
ELONG_TN_P	2PpB_158285325	2PpB	1,58E+08	*PPRT2PpB6G038518*	1,58E+08	1,58E+08	Protein of unknown function (DUF563)	DUF563
ELONG_TN_P	2PpB_158285325	2PpB	1,58E+08	*PPRT2PpB6G038519*	1,58E+08	1,58E+08	-	-
ELONG_TN_P	2PpB_158285325	2PpB	1,58E+08	*PPRT2PpB6G038520*	1,58E+08	1,58E+08	Protein of unknown function (DUF1191)	DUF1191
ELONG_TN_P	2PpC_166273398	2PpC	1,66E+08	*PPRT2PpC6G051434*	1,66E+08	1,66E+08	-	NA
ELONG_TN_P	2PpC_166273398	2PpC	1,66E+08	*PPRT2PpC6G051435*	1,66E+08	1,66E+08	-	-
ELONG_TN_P	2PpD_139456629	2PpD	1,39E+08	*PPRT2PpD6G061058*	1,39E+08	1,39E+08	ADP binding	NB-ARC
ELONG_TN_P	3PpB_156491369	3PpB	1,56E+08	*PPRT3PpB6G072602*	1,56E+08	1,56E+08	Belongs to the peptidase A1 family	TAXi_C,TAXi_N
ELONG_TN_P	5PpB_23885075	5PpB	23885075	*PPRT5PpB6G121102*	23884650	23884932	meprin and TRAF homology	BTB,MATH
ELONG_TN_P	5PpB_23885075	5PpB	23885075	*PPRT5PpB6G121103*	23885153	23886665	ABC transporter transmembrane region	ABC_membrane, ABC_tran
GEN_TN_P	1PpB_75357307	1PpB	75357307	*PPRT1PpB6G005547*	75358775	75360272	ATP synthase subunit beta	ATP-synt_ab,ATP-synt_ab_N
GEN_TN_P	2PpC_158132608	2PpC	1,58E+08	*PPRT2PpC6G050754*	1,58E+08	1,58E+08	Belongs to the UDP-glycosyltransferase family	UDPGT
GEN_TN_P	3PpC_22182317	3PpC	22182317	*PPRT3PpC6G075080*	22179655	22180424	-	-

Table 2. Continued. Functional annotation of candidate genes (within ± 2kb) underlying significant MTAs identified by GWAS.

Trait	SNP	CHR	POS	GeneID	gstart	Gend	Description	PFAMs
GEN_TN_P	3PpC_22182317	3PpC	22182317	*PPRT3PpC6G075082*	22180512	22180860	-	-
GEN_TN_P	5PpB_23886531	5PpB	23886531	*PPRT5PpB6G121102*	23884650	23884932	meprin and TRAF homology	BTB,MATH
GEN_TN_P	5PpB_23886531	5PpB	23886531	*PPRT5PpB6G121103*	23885153	23886665	ABC transporter transmembrane region	ABC_membrane,ABC_tran
GEN_TN_P	6PpD_29599743	6PpD	29599743	*PPRT6PpD6G179839*	29598541	29599567	30S ribosomal protein S11	RNA_pol_A_CTD,RNA_pol_A_bac,RNA_pol_L,Ribosomal_S11
GEN_TN_P	7PpC_40478225	7PpC	40478225	*PPRT7PpC6G197154*	4046991	4048181	-	-
ELONG_DW	1PpC_144294419	1PpC	1,44E+08	*PPRT1PpC6G019481*	1,44E+08	1,44E+08	disease resistance protein	NB-ARC
ELONG_DW	2PpB_22759745	2PpB	22759745	*PPRT2PpB6G030270*	22750886	22773730	Pectinesterase	PMEI,Pectinesterase
ELONG_DW	2PpB_22759745	2PpB	22759745	*PPRT2PpB6G030271*	22758472	22761685	DNA recombination	DUF223,REPA_OB_2,Rep_fac-A_C
ELONG_DW	3PpC_84839193	3PpC	84839193	*PPRT3PpC6G078060*	84838002	84838491	-	-
ELONG_DW	7PpB_51224957	7PpB	51224957	*PPRT7PpB6G191002*	51224618	51225029	Photosystem II (PSII) is a light-driven water plastoquinone oxidoreductase that uses light energy to abstract electrons from H(2)O, generating O(2) and a proton gradient subsequently used for ATP formation. It consists of a core antenna complex that captures photons, and an electron transfer chain that converts photonic excitation into a charge separation. The D1 D2 (PsbA PsbA) reaction center heterodimer binds P680, the primary electron donor of PSII as well as several subsequent electron acceptors	Photo_RC

Table 2. Continued. Functional annotation of candidate genes (within ± 2kb) underlying significant MTAs identified by GWAS.

Trait	SNP	CHR	POS	GeneID	Gstart	gend	Description	PFAMs
ELONG_DW	7PpB_51224957	7PpB	51224957	*PPRT7PpB6G191003*	51225283	51225652	-	-
ELONG_DW	7PpD_118291491	7PpD	1,18E+08	*PPRT7PpD6G214833*	1,18E+08	1,18E+08	Transmembrane amino acid transporter protein	Aa_trans
GEN_DW	1PpC_146786381	1PpC	1,47E+08	*PPRT1PpC6G019662*	1,47E+08	1,47E+08	source UniProtKB	DUF4216,DUF4218, Transpos_assoc, Transposase_21
GEN_DW	5PpB_47559049	5PpB	47559049	*PPRT5PpB6G122925*	47557343	47558447	Protein of unknown function (DUF1668)	-
GEN_DW	7PpC_62499076	7PpC	62499076	*PPRT7PpC6G200859*	62500878	62501466	A Receptor for Ubiquitination Targets	F-box,FBA_3
VEG_L	6PpC_74611786	6PpC	74611786	*PPRT6PpC6G172596*	74589756	74617734	-	-
GEN_L	1PpC_123281945	1PpC	1,23E+08	*PPRT1PpC6G017844*	1,23E+08	1,23E+08	PPR repeat	PPR,PPR_1,PPR_2,PPR_3
GEN_L	1PpC_123281945	1PpC	1,23E+08	*PPRT1PpC6G017845*	1,23E+08	1,23E+08	-	-
GEN_L	2PpC_46339227	2PpC	46339227	*PPRT2PpC6G043885*	46338548	46341405	-	-
GEN_L	5PpD_23894193	5PpD	23894193	*PPRT5PpD6G146763*	23892386	23892659	-	-

4. Discussion

Tillering is an essential developmental trait that determine productivity, survival and adaptability in grasses. Tillers function as photosynthetic units that intercept light, contributes to canopy structure, regrowth capacity and biomass accumulation. The production of tillers in grasses is dependent on both the genetic makeup and the environmental factors such as light, temperature and nutrient availability Productivity in timothy grass is related to the initiation and seasonal turnover of its tillers. Hence investigating the origin of tillering traits in timothy grass is important for improving stand persistence, optimizing forage yield and developing improved grass cultivars with the desired trait characteristics.

Although there are some previous studies on GWAS of tillering traits on other forage grasses such as Ryegrass and Switchgrass (Zhang *et al*., 2022), there is no GWAS study relating to tillering traits in timothy grass. By performing GWAS analysis we connect genetic variation with phenotypic traits observed in timothy grass and can identify genetic markers across the timothy grass genome related to its tillering capability. We also provide information on the plausible candidate genes, potential regulatory pathways and biological mechanisms suspected to be behind these tillering traits. In the present study, we investigated genetic control of tillering traits in 212 accessions of timothy grass using GWAS to provide insight into MTAs that could be useful in marker assisted breeding targeted towards higher biomass production in timothy grass improvement.

4.1 Phenotypic diversity, trait relationships and heritability

The correlation analysis revealed a strong positive correlation between total tiller number (TTN) and total dry weight (TDW) which suggests that tillering ability of timothy grass is a main determinant of biomass accumulation. The strong positive relationship exhibited between TTN and in all VEG _DW in all accessions suggests that timothy plants with more vegetative tillers tend to produce more biomass while Wild accessions producing more biomass compared with the breeding lines and cultivars which could be because of their adaptability to the environment and ability to withstand abiotic stresses. The trade-off is that stress resistance and resilience is diminished in cultivars and breeding lines through domestication and selection of agronomic traits (De Meyer et al., 2026).

Tillers are developed through shoot proliferation leading to high tiller density through the production of numerous meristematic growing points (Li et al. 2024). This indicates that biomass production in timothy grass may be dependent on shoot proliferation than on the size of tillers. While tiller numbers contribute in a positive

way to dry weight accumulation, there should be a healthy balance to sustain productivity (Chen et al. 2019)

High tillering genotypes may also contribute to weed management by the production of denser swards that limit weed establishment in forage systems (Li et al., 2024).

4.2 Genetic architecture of tiller traits based on GWAS

The GWAS analysis identified multiple MTAs across several tiller-related traits. However, no significant associations were detected for total tiller number or total dry weight at the Bonferroni significance threshold. This likely reflects the highly quantitative and environmentally influenced nature of these traits, which are probably controlled by numerous loci with individually small effects. Such polygenic inheritance patterns are common for complex agronomic traits in grasses and other crop species.

The distribution of significant SNPs across multiple chromosomes further supports the complex genetic architecture of tiller-related traits in timothy grass. Similar findings have been reported in other grass species, where tillering, biomass accumulation, and developmental traits are controlled by multiple genomic regions interacting with environmental factors (Alqudah et al., 2021; Zhang et al., 2022). The reduction in the number of significant SNPs under more stringent significance thresholds further suggests that many detected associations represent loci with moderate or small effects, which is typical for complex quantitative traits.

The identification of significant SNPs associated with tiller composition traits demonstrates the potential utility of genome-wide markers for future breeding applications in timothy grass. Although additional validation is required, these loci may eventually contribute to marker-assisted selection, genomic selection, and precision breeding approaches aimed at improving tillering characteristics, biomass production, and adaptation.

4.3 Functional annotation and candidate gene interpretation

Functional annotation of genes located near significant SNPs provided insights into the biological processes potentially underlying variation in tiller-related traits. The identified candidate genes were associated with diverse functional categories, including cell wall modification, transport, photosynthesis, transcriptional regulation, protein turnover, and disease resistance, all of which are processes known to influence plant growth and development.

A candidate gene associated with elongated tiller dry weight (ELONG_DW) was annotated as a pectinesterase-related gene, suggesting a potential role for cell wall remodeling in biomass accumulation. Pectinesterases regulate the degree of pectin methylesterification within the cell wall and thereby influence wall extensibility, tissue stiffness, and cell expansion (Micheli, 2001; Pelloux et al., 2007; Wolf et al., 2009). Variation in pectinesterase activity may therefore affect stem elongation and structural development in elongated tillers, ultimately influencing dry matter accumulation and biomass production.

Another functionally relevant candidate gene associated with ELONG_DW was annotated as a photosystem II (PSII)-related protein. Photosystem II is a central component of the photosynthetic electron transport chain and is essential for light-driven energy conversion and carbon assimilation (Nelson and Yocum, 2006; Aro et al., 1993). Variation in genes associated with PSII function may influence photosynthetic efficiency and assimilate availability, thereby affecting biomass accumulation and dry matter production. Previous studies have demonstrated that environmental and biotic stresses affecting PSII performance can substantially reduce biomass accumulation in grasses and crop species (Vinyard et al., 2013; Shevela et al., 2023), supporting the biological relevance of this association.

Several candidate genes identified within the genomic interval of detected loci for tiller composition traits were annotated with NB-ARC domains and other disease-resistance-related features. Although these genes are traditionally associated with plant defense and immune signaling, their association with this genomic region suggest the possibility of interaction between defense-related pathways, developmental regulation and growth processes. Their association with tiller traits in the present study may be an indication of possible interactions between stress-response signaling and developmental coordination.

Candidate genes annotated with pentatricopeptide repeat (PPR) domains were also identified near significant SNPs. PPR proteins are RNA-binding proteins involved in post-transcriptional regulation within mitochondria and chloroplasts, including RNA editing, transcript stabilization, and translational regulation (Barkan and Small, 2014). These proteins play important roles in organellar function, photosynthesis, respiration, and stress adaptation. Previous studies in rice and Arabidopsis have demonstrated associations between PPR proteins and tolerance to drought, oxidative stress, and pathogen responses through the maintenance of chloroplast function and reactive oxygen species homeostasis (Luo et al., 2022; Meng et al., 2024). The identification of PPR-associated loci in timothy grass therefore suggests potential roles in stress adaptation, growth stability, and environmental resilience.

Genes associated with ubiquitin-mediated protein turnover pathways, including BTB/MATH and F-box domain proteins, were also identified as gene of interest within the mapped interval. These proteins are components of the ubiquitin–proteasome system, which regulates protein degradation and cellular signaling processes involved in plant growth, development, and stress responses. Similar proteasome-related pathways have previously been reported in timothy grass and other plant species (Pashapu et al., 2024), supporting the importance of protein turnover mechanisms in developmental regulation.

In addition, candidate genes annotated as ABC transporters and UDP-glycosyltransferases were identified near significant SNPs. ABC transporters participate in diverse physiological processes, including metabolite transport, detoxification, and stress responses (Ishikawa et al., 2013), whereas UDP-glycosyltransferases are involved in secondary metabolism, hormone regulation, and defense responses (Ouyang, et al., 2023). The identification of these candidate genes suggests that tiller-related traits in timothy grass are influenced by interconnected pathways regulating growth, metabolism, environmental adaptation, and defense.

Conclusion

This study investigated the genetic architecture of tiller-related traits in timothy grass using genome-wide association studies (GWAS) in a diverse collection of domesticated and non-domesticated accessions. Considerable phenotypic variation was observed among accessions, particularly for tiller composition and biomass-related traits, indicating substantial genetic diversity within the studied material. The GWAS analysis identified multiple genomic regions associated with tiller-related traits, supporting the complex and polygenic nature of tiller development in timothy grass

Functional annotation of candidate genes revealed associations with biological processes related to cell wall modification, photosynthesis, stress responses, protein turnover, and transport. Candidate genes annotated with NB-ARC, PPR, BTB/MATH, ABC transporter, UDP-glycosyltransferase, and pectinesterase domains suggest that tiller development in timothy grass maybe regulated through interconnected pathways involving growth, developmental regulation, and environmental adaptation.

Although further validation is required, the identified MTAs and candidate genes provide a valuable foundation for future studies on the genetic regulation of tillering in timothy grass. Future research should focus on validating significant loci across multiple environments and developmental stages, as well as investigating the functional roles of candidate genes through transcriptomic and functional genomic

approaches. The integration of validated molecular markers into breeding programs may contribute to the development of timothy cultivars with improved biomass production, persistence, and adaptation to northern environments.

References

Alqudah AM, Sharma R, Börner A.(2021). Insight into the genetic contribution of maximum yield potential, spikelet development and abortion in barley. *Plants, People, Planet*. 3:721–736. https://doi.org/10.1002/ppp3.10203

Aro E.M, Virgin I, and Andersson B. (1993). Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. Biochim Biophys Acta. 5;1143(2):113-34. doi: 10.1016/0005-2728(93)90134-2. PMID: 8318516.

Bajgain, P., Zhang, X., & Anderson, J. A. (2019). Genome-wide association study of yield component traits in intermediate wheatgrass and implications in genomic selection and breeding. *Genes, Genomes, Genetics, 9*(8), 2429-2439 . https://doi.org/10.1534/g3.119.400073

Barkan, A and Small, I. (2014). Pentatricopeptide repeat proteins in plants. Annu Rev Plant Biol. 65:415-42. doi: 10.1146/annurev-arplant-050213-040159.

Chang D, Wu Y, Liu L, Lu-Thames S, Dong H, Goad C, Bai S, Makaju S, Fang T. (2016) Quantitative Trait Loci Mapping for Tillering-Related Traits in Two Switchgrass Populations. Plant Genome. Jul;9(2). doi: 10.3835/plantgenome2016.01.0010. PMID: 27898826.

Chen, XX., Zhang, W., Liang, X.Y, Liu,Y.M., Xu,S.J., Zhao,Q.Y., Du,Y.F., Zhang, L., Chen. X.P., and Zou, C. Q. (2019). Physiological and developmental traits associated with the grain yield of winter wheat as affected by phosphorus fertilizer management. Sci Reports 9, 16580 https://doi.org/10.1038/s41598-019-53000-

Chen, S. (2025). Fastp 1.0: An ultra-fast all-round tool for FASTQ data quality control and preprocessing. *Imeta*, *4*(5), e70078.

De Meyer, E., Van Cauter, F., Vandelook, F. (2026). Comparison of multi-stress resilience in wild and domesticated Cowpea. *Sci Rep* **16**, 5109 https://doi.org/10.1038/s41598-026-35860-4

Issa I. I. A. A. (2022). Multi-environment Screening of Timothy (*Phleum pratense*) Breeding Material for Better Forage Yield in Sweden, MSc. Thesis in Department of Plant Breeding Horticultural Science Programme, Swedish University of Agricultural Sciences, Alnarp , Sweden.

Ivehag, A., (2018). Spillage during timothy threshing: how much is spilled at different driving speeds?. First cycle, G1E. Alnarp: SLU, Dept. of Biosystems and Technology.

Ishikawa, K., Ito,K., Inoue,J and Semba, K.(2013). Cell growth by stable Rhg2/Gir2 complex formationunder amino acid starvation. Gene Cells 18(10): 859-872

Kucka, M. (2020). Tn5-on-beads DNA tagmentation protocol, Germany: Friedrich Miescher Laboratory. Available at: https://coilink.org/20.500.12592/j1gxn2

Li, N., Zhang, J., Zhang, L. Q., & Nie, P. (2010). Difference in genes between a high virulence strain G4 and a low virulence strain G18 of *Flavobacterium columnare* by using suppression subtractive hybridization. Journal of Fish Diseases, 33(5), 403–412. https://doi.org/10.1111/j.1365-2761.2009.01132.x.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li, J., Yao, X., Lai,H. Zhang, X and Zhong, J.(2024) The diversification of the shoot branching system: A quantitative and comparative perspective in meristem determinacy. Current Opinion in Plant Biology, Volume 81, 102574, https://doi.org/10.1016/j.

Luo Z, Xiong J, Xia H, Wang L, Hou G, Li Z, Li J, Zhou H., Li T. and Luo L. (2022). Pentatricopeptide Repeat Gene-Mediated Mitochondrial RNA Editing Impacts on Rice Drought Tolerance. Front Plant Sci. Jul 19;13:926285. doi: 10.3389/fpls.2022.926285.

Meng L, Du M, Zhu T, Li G, Ding Y, Zhang Q. (2024) PPR proteins in plants: roles, mechanisms, and prospects for rice research. Front Plant Sci. Jun 27;15:1416742. doi: 10.3389/fpls.2024.1416742

Micheli, F. (2001). Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *Trends in plant science*, *6*(9), 414-419.

Nelson N, and Yocum C.F. (2006) Structure and function of photosystems I and II. Annu Rev Plant Biol.;57:521-65. doi:10.1146/annurev.arplant.57.032905.105350. .

Ouyang, L., Liu, Y., Yao, R., He, D.,Yan, L., Chen, Y., Huai,D.,Wang,Z.,Yu,B.,Kang, Y.,Jiang, H., Lei, Y.,Liao, B. and Wang, X. (2023). Genome-wide analysis of UDP-glycosyltransferase gene family and identification of a flavonoid 7-*O*-UGT (*AhUGT75A*) enhancing abiotic stress in peanut (*Arachis hypogaea* L.). *BMC Plant Biol* **23**, 626https://doi.org/10.1186/s12870-023-04656-3

Pashapu, A. R., Dalmannsdottir, S., Jørgensen, M., Schubert, M., Rognli, O. A., & Kovi, M. R. (2024). Frost survival and gene expression in timothy (*Phleum pratense* L.) cultivars as affected by age and selection in diverse field environments. Physiologia Plantarum, 176(1), e14217. https://doi.org/10.1111/ppl.14217 .

Pelloux, J., Rusterucci, C., and Mellerowicz, E. J. (2007). New insights into pectin methylesterase structure and function. *Trends in plant science*, *12*(6), 267-277.

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A., Kling, D. E., Gauthier, L. D., Levy-Moonshine, A., & Roazen, D. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*, 201178.

Posit (2026) RStudio Desktop. RStudio: Integrated Development for R. Rstudio. PBC:Boston, MA, 02210 USA, Available online https://www.stats.bris.ac.uk/R/ (accessed on 26 January 2026).

Rahimi, Y.; Bedada, G.; Moreno, S.; Gustavsson, A.-M.; Ingvarsson, P.K.; Westerbergh, A. (2023) Phenotypic Diversity in Domesticated and Wild Timothy Grass, and Closely Related Species for Forage Breeding. *Plants 12*, 3494. https://doi.org/10.3390/plants12193494.

Rahimi, Y. (2024). Phenotypic and genetic diversity in wild and domesticated timothy and related *Phleum* species: implications for breeding. In Acta universitatis agriculturae Sueciae. Swedish University of Agricultural Sciences.

Rahimi, Y., Bedada, G.; Gustavsson, A.-M.; Öhlund, L., Westerbergh, A. and Ingvarsson, P.K.; (2026 unpublished manuscript). Dissecting the genetic architecture of forage yield and

developmental traits in timothy grass (*Phleum pratense* L.) using single and multi-trait GWAS.

Shevela D., Kern J.F, Govindjee G, and Messinger J. (2023) Solar energy conversion by photosystem II: principles and structures. Photosynth Res. Jun;156(3):279-307. doi: 10.1007/s11120-022-00991-y.

Shikanai T, and Fujii S. (2013). Function of PPR proteins in plastid gene expression. RNA Biol. 10(9):1446-56. doi: 10.4161/rna.25207.

Van der Auwera, G. A., and O'Connor, B. D. (2020). *Genomics in the cloud: using Docker, GATK, and WDL in Terra*. O'Reilly Media.

Vinyard, D.J., Ananyev, G.M., and Dismukes, C. (2013). Photosystem II: The Reaction Center of Oxygenic Photosynthesis*. *Annual Review Biochemistry*. 82:577-606. https://doi.org/10.1146/annurev-biochem-070511-100425

Wolf, S., Mouille, G., & Pelloux, J. (2009). Homogalacturonan methyl-esterification and plant development. *Molecular plant*, *2*(5), 851-860.

Xiaoheng X., Peng, L., Shunfeng, L., Guangyan, F., Miaoli, W., Zhongfu, Y., Gang, N., Linkai H. and Xinquan, Z. (2024). Genome-wide association analysis reveals novel candidate loci and a gene regulating tiller number in orchardgrass, Plant Physiology and Biochemistry, Volume 216, 109148.

Yang, B., & Li, Y. Z. (2022). Effects of Timothy Cladosporium Eyespot on Photosynthesis and Biomass. In Research Square. Research Square. https://doi.org/10.21203/rs.3.rs-1669625/v1

Yin, L, Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., Yuan, X., Zhu, M., Zhao, S., Li, X and Liu, X., (2021) rMVP: A Memory-Efficient, Visualization-Enhanced, and Parallel-Accelerated Tool for Genome-Wide Association Study, Genomics, Proteomics & Bioinformatics, Vol, 19 (4): 619-628.

Zhang L, MacQueen A, Weng X, Behrman KD, Bonnette J, Reilley JL, Rouquette FM Jr, Fay PA, Wu Y, Fritschi FB, Mitchell RB, Lowry DB, Boe AR, Juenger TE. (2022). The genetic basis for panicle trait variation in switchgrass (*Panicum virgatum*). Theor Appl Genet. 135(8):2577-2592. DOI: 10.1007/s00122-022-04096-x

Popular science summary

Timothy grass is one of the most important forage grasses used for livestock production in Sweden and other cold regions of the world. It is commonly grown for hay and silage and is valued for its high nutritional quality, winter hardiness, and ability to regrow after harvesting. Improving the productivity and persistence of timothy grass can contribute to more sustainable meat and dairy production by increasing the availability of high-quality forage for cattle and horses.

One of the most important characteristics of timothy grass is its ability to produce tillers, which are shoots that grow from the base of the plant. The number and type of tillers strongly influence biomass production, forage yield, and regrowth capacity. Plants with favourable tillering characteristics can therefore produce more forage and maintain productive stands for longer periods.

In this study, the genetic basis of tiller-related traits in timothy grass was investigated using genome-wide association studies (GWAS). A diverse collection of 212 timothy accessions was evaluated for traits related to tiller number, tiller composition, tiller length, and dry biomass production. By combining phenotypic measurements with genome-wide genetic marker data, genomic regions associated with tiller-related traits were identified.

Several candidate genes associated with plant growth, photosynthesis, stress responses, and developmental regulation were identified near significant genetic markers. These findings improve our understanding of how tillering and biomass production are regulated in timothy grass and provide valuable information for future breeding programs. In the long term, the identified genetic markers may support the development of improved timothy cultivars with higher forage yield, better persistence, and improved adaptation to northern growing conditions.
This study also contributes to a broader understanding of branching and growth regulation in perennial grasses and may support improvement efforts in related forage grass species.

Publishing and archiving

☐ YES, I/we have read and agree to the agreement for publication and the personal data processing that takes place in connection with this

☒ NO, I/we do not give my/our permission to publish the full text of this work. However, the work will be uploaded for archiving and the metadata and summary will be visible and searchable.

Acknowledgements

It is with a grateful heart I appreciate God Almighty for the grace to complete this study. I would like to express my deep gratitude to my Supervisor Dr Yousef Rahimi for his exceptional guidance and valuable contributions at each stage of the research. This helped me to complete the thesis within the timeline. I also appreciate the contributions and feedback from my Co-Supervisor Dr Girma Bedada whose experience was helpful in the research.

I extend my thanks to all Faculty and Staff whom I encountered during the course of my study, thank you for the training received.

I would like to appreciate my husband Dr David Ekhuemelo, my children Triumph Joseph and Treasure for their financial and moral support that enabled me to complete this study.

All colleagues and friends are also appreciated for the time spent learning together.