

Seed Shattering in a North American Oryzae grain: Developmental and Genomic Signatures of Early Domestication

Reneth Millas

University of Minnesota

Lillian McGilp

University of Minnesota

Alan Mickelson

University of Minnesota

Maybell Banting

University of Minnesota

Claudia Castell-Miller

University of Minnesota

Jennifer Kimball

jkimball@umn.edu

University of Minnesota

Research Article

Keywords: *Zizania palustris*, seed retention, abscission layer, *Oryza sativa*, gene expression, multiple sequence alignment, domestication

Posted Date: July 14th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7032638/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Northern Wild Rice (NWR; *Zizania palustris* L.) is an aquatic grain endemic to North America and a member of the Oryzaeae tribe. As an outcrossing crop with a short breeding history, domestication progress in cultivated NWR (cNWR) is ongoing and seed shattering remains a major barrier to yield stability. In this study, we investigated the developmental and genetic mechanisms underlying seed retention by integrating phenotypic, anatomical, and molecular analyses across wild and cultivated populations. Time-course phenotyping using four methods revealed a ~90% reduction and two-week delay in shattering in cNWR relative to wild populations. Histological analysis indicated anatomical reorganization of the abscission layer in cNWR, consistent with selection for seed retention. Comparative genomic analyses identified multiple NWR orthologs of key *Oryza sativa* shattering genes, revealing lineage-specific gene duplication, pseudogenization, and divergence from *Z. latifolia*. Expression profiling of candidate genes via RT-qPCR across floret developmental stages suggested a regulatory role for *ZpSh5c*, a putative *OsSh5* ortholog, in modulating seed shattering. In cNWR, delayed and reduced expression of *ZpSh5c* mirrored observed differences in shattering timing, highlighting its potential involvement in abscission zone development. Together, these findings provide new insights into the anatomical and molecular basis of seed shattering in NWR and demonstrate the utility of comparative frameworks for accelerating trait improvement in emerging, non-model crops.

PLAIN LANGUAGE SUMMARY

Northern Wild Rice (NWR, *Zizania palustris*) has been grown in irrigated paddies in the United States since the 1950s. Seed shattering, that is, when ripe seed separates from the plant pedicel, reduces grain yields in cultivated NWR (cNWR) production. However, little is known about this process. This study compared the seed shattering of two cNWR populations to a wild, unselected population. We found that cNWR seed retention increased ~90% compared to the time when the wild type shatters, and cNWR began shattering about two weeks later than the wild type. Imaging of a layer of cells that lead to shattering at the pedicel, where the seed and stem meet, also revealed differences between cNWR and the wild type. Being closely related to white rice, we leveraged available white rice genomic resources to identify potential NWR genes that may be associated with seed shattering. We found that the gene, *ZpSh5c*, is expressed in the same pattern as shattering occurs in both wild and cultivated plants, suggesting its possible role in the regulation of shattering in NWR.

INTRODUCTION

Seed shattering is a natural seed dispersal mechanism widespread among wild plant species, including the wild ancestors of many modern crops (Konishi et al., 2006; Li & Gill, 2006). As species were domesticated, traits that favored seed retention were selected, leading to the replacement of wild type alleles with mutations that reduced or eliminated shattering (Konishi et al., 2006; Lin et al., 2012). A key target of selection was a developmentally regulated abscission event involving the formation and breakdown of an abscission layer, a specialized cell layer, at the seed–pedicel junction (Oba et al., 1995; Fuller & Allaby, 2009; Watanabe et al., 2003). Disruption or loss of this abscission layer increases seed retention, significantly improving yields and harvest efficiency, which are critical milestones in crop domestication.

Numerous genes and quantitative trait loci (QTL) underlying seed shattering have been identified in crop species, particularly in *Oryza sativa* L. and related species (Oba et al., 1995; Cai and Morishima, 2000; Thomson et al., 2003; Lin et al., 2007; Konishi et al., 2006; Ishikawa et al., 2010; Lee et al., 2016; Wu et al., 2017). Several of these have been cloned or fine-mapped, revealing their roles in regulating the development and function of the abscission layer. Notable examples include the *sh4* gene, which encodes a MYB3 transcription factor, (Li et al., 2006; Purugganan & Fuller, 2009; Zhou et al., 2012); *qSH1*, which encodes a BEL1-type homeobox gene (Konishi et al., 2006); *Sh5*, a highly homologous gene of *qSH1* (Yoon et al., 2014); *SHAT1*, an *APETALA2* (AP2) transcription factor (Zhou et al., 2012); and *SH1*, a YABBY transcription factor (Lin et al., 2012). Together, these genes can modulate the development and degradation of the abscission zone and have been repeatedly targeted during the independent domestication of cereal crops.

Across crops, studies have revealed that seed shattering pathways are generally conserved, with many genes either orthologous or functionally analogous across divergent lineages (Dong & Wang., 2015; Ishikawa et al., 2022). This pattern of evolutionary convergence reflects similar selection pressures during domestication that were used to increase seed retention (Doebley, 2006; Lenser and Theißen, 2013; Olsen & Wendel, 2013; Purugganan & Fuller, 2009; Tranbarger et al., 2017). Comparative genomic studies have helped to shape our understanding of these convergent evolutionary patterns and are increasingly applied to identify candidate domestication genes in wild and understudied crop species (Kahler et al., 2014; Fu et al., 2019; Liu et al., 2022). For example, recent studies, have reported that *SH1* likely plays a role in seed shattering across a wide range of grass species, including *Lolium perenne* (Perennial ryegrass), *Setaria italica* (Foxtail millet), and *Zizania latifolia* (Manchurian wild rice) (Fu et al., 2019; Liu et al., 2022; Odonkor et al., 2018; Yu et al., 2023). By leveraging such comparative frameworks, researchers can identify conserved regulatory genes and causal variants for key traits, ultimately accelerating the improvement of emerging, non-model crops.

One emerging crop with strong potential for comparative genomics is Northern Wild Rice (NWR, *Zizania palustris* L.), an annual aquatic grass and Crop Wild Relative (CWR) of *O. sativa* within the Oryzaeae tribe. Endemic to the Great Lake Region of North America, NWR production in irrigated, diked paddies began in the region in the 1950s. In 1968, the first cultivated NWR (cNWR) variety was released, featuring improved seed retention compared to its wild counterparts (Oelke, 1982). Like *Oryza* species (Jin et al., 1993; Jin et al., 1982, 1990, 1995; Jin and Inouye, 1982a, b, 1985), seed shattering in NWR is associated with development of an abscission layer at the seed-pedicel junction composed of parenchyma and sclerenchyma cells, where variation in cell thickness and number contributes to genotype-specific differences (Hanten et al., 1980). Previous genetic studies have suggested that seed shattering in NWR is a recessive, quantitative trait regulated by at least two or three genes (Woods & Clark, 1976; Elliott & Perlinger, 1977; Kennard et al., 2002). Despite ongoing efforts, modern cNWR varieties and breeding germplasm still exhibit some degree of seed shattering, primarily due to the species' outcrossing nature and limited breeding history (Elliott, 1980). The trait's persistence contributes to substantial yield losses, particularly after late-summer storms or high

winds (Imle, 2001). As such, NWR offers a valuable system for exploring the genetic basis of seed shattering and applying insights from domesticated relatives to accelerate trait improvement in a semi-domesticated species.

Improving seed retention in cNWR requires a better understanding of the developmental, anatomical, and genetic factors that regulate seed shattering. Despite its agronomic importance, key aspects of this trait remain underexplored. Given the close phylogenetic relationship between *O. sativa* and NWR, comparative genomics offers a powerful framework for candidate gene discovery (Haas et al., 2021). Therefore, this study aimed to: (1) Assess the phenotypic variation in seed shattering using four quantitative phenotyping methods; (2) Examine the abscission layer anatomy of two cNWR populations and one wild NWR population using histological analysis; (3) Identify and compare the protein sequences of putative *O. sativa* orthologs in NWR and *Z. latifolia*; and (4) Evaluate the expression profiles of putative NWR orthologs of *O. sativa* seed shattering-related genes across stages of panicle development and seed maturation.

MATERIALS AND METHODS

Plant materials. Three NWR populations were used in this study, including a wild type population exhibiting a profuse shattering habit, collected from Sullivan Lake, Minnesota (MN), USA in 2020 (46.149519 N; - 93.936795 W); a cNWR breeding line, 'FY-C20' developed for seed length; and a cNWR variety, 'Itasca-C12', released in 2007 (Porter et al., 2008), which served as the industry's standard for seed retention at the time of this study. Seeds were removed from cool storage (3°C) (McGill et al., 2023), rinsed three times, and placed in new plastic bags containing fresh water on a lab benchtop. After 10 days, germinated seedlings were transplanted in the greenhouse.

Plant growth conditions. Experiments were conducted at the University of Minnesota (UMN) Plant Growth Facilities in Saint Paul, MN, in the spring of 2021. Greenhouse conditions were maintained at 22°C ($\pm$ 2°C) with a 16-hour photoperiod. Cone-tainers (112.98 cm²) were filled with a steamed soil mix supplemented with 0.14 g urea (20 lb/acre) and 0.14 g of iron chelate (1 lb/ 1000 ft²), then placed in aluminum aquaponic tanks (66 cm x 183 cm x 70 cm). Tanks were filled with water (13°C) up to 2–4 cm above the surface of the cone-tainers one week before seedlings were planted. To limit algae growth, water was circulated under 5 psi using a 1056GPH 276W submersible pump (Simple Deluxe, Duarte, California, USA) and treated twice weekly with Algaefix (1 ml/10 gallons of water; API, Chalfont, Pennsylvania, USA). After germination, one seedling was transplanted per cone-tainer. Plants were top-dressed with urea (20lb/acre) after tillering every 2–3 weeks.

Seed shattering phenotypic data collection. The degree of seed shatter was measured on the main stem panicles of three individual plants per population per time point over the course of NWR seed ripening at the Principal Phenological Stage (PPS) 8 (Duquette & Kimball, 2020) using four different methods. Individual plants were collected at their respective days after anthesis (DAA), including 21, 28, 35, and 42 DAA. The first method was a visual estimation of the percentage of shattered seed at the time of data collection, ranging from 0 to 100%, where 0 = no shatter and 100% = complete shatter. The second method was a drop test, whereby a panicle was dropped three times from a height of 60.96 cm, and the proportion of shattered seeds was calculated relative to the total number of seeds per panicle (Altendorf et al., 2021). The third and fourth methods measured breaking tensile strength (BTS) using a Chatillon digital force gauge (Ametek, Inc, Berwyn, Pennsylvania), either by pulling the seed from the pedicel (BTS-P) or bending (BTS-B). For BTS-P, the seed was gently gripped with forceps attached to the gauge, and force was applied in a straight, downward direction until detachment occurred. The maximum force required to remove the seed was recorded in grams-force (g), a unit of force where 1g $\approx$ 0.0098 Newtons. For BTS-B, force was applied laterally to the seed at the point of attachment and gently bent until detachment occurred. The force required to break the seed from the pedicel was also recorded in g.

Histology of the seed-pedicel junction. Seed-pedicel junctions of five plants per population were sampled before reaching the hard dough stage of seed ripening (PPS 85), which corresponded to 21 DAA for the wild type population's individuals and 28 DAA for the cultivated line's individuals. Seeds were bisected, placed in a vacuum chamber, and fixed in a mixture of formalin-acetic acid-alcohol (FAA; 53:5:10:32) solution containing 0.1% Triton X-100. Fixed samples were stored in 70% ethanol at 4°C before being processed at the Histology and Research Laboratory, UMN. Upon arrival, the samples were fixed in 2% paraformaldehyde in a 0.1 M phosphate buffer (pH 7.0), dehydrated in an incremental ethanol and ethanol-xylene series (3:1, 1:1, 1:3, 100% xylene), and embedded in Paraplast Plus (Leica Biosystems, Wetzlar, Germany). Sections (10 μ m) were cut using a rotary microtome, stained with 0.05% toluidine blue O in 0.1 M phosphate buffer (pH 6.8), and imaged using an Axio Scan.Z1 slide scanner (Zeiss Group, Oberkochen, Baden-Württemberg, Germany).

Plant tissue collection. Three plant tissues were collected from three individual plants per population in liquid nitrogen as independent biological replicates. The sampled tissues included flag leaf at the heading stage (PPS 51); male florets at 7 DAA at PPS 66 when all flowering was complete and before pollen shed (PPS 69); and female florets at 1, 7, 18, 28, and 35 DAA throughout seed development (PPS 7) and maturation (PPS 8). In the case of the wild type population, female florets were collected only up to 18 DAA, as most seeds had shattered by the next time point. Tissue samples were stored at -80°C until further use.

RNA isolation and cDNA synthesis. Total RNA was extracted using a RNeasy® Plant Mini Kit (Qiagen, Valencia, CA) following the manufacturer's instructions. RNA concentration and purity were measured using a NanoDrop™ 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). To eliminate DNA contamination, RNA samples were treated with the DNA-free™ DNA Removal Kit (Thermo Fisher Scientific, Waltham, MA). Complementary DNA (cDNA) was synthesized from total RNA using the SuperScript™ III First-Strand Synthesis SuperMix (Invitrogen, Thermo Fisher Scientific, Waltham, MA), following the manufacturer's protocol. Synthesized cDNA was stored at -20°C until further use.

Reverse Transcription quantitative Polymerase Chain Reaction (RT-qPCR) assays. Gene expression quantification was performed using an Applied Biosystems 7500 Fast Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA) on 96-well plates. Each 20 μ L reaction contained 50 ng of cDNA template, 10 μ L of iTaq™ Universal SYBR® Green Supermix (Bio-Rad, Hercules, CA), and 0.2 μ L (10 mM) each of forward and reverse primers (Supplemental Table 1) using standard thermal cycling conditions (95°C for 20 s, followed by 40 cycles of 95°C for 3 s and 60°C for 30 s). Each gene was evaluated using

three biological replicates and two technical replicates per biological replicate, tissue type, and time point. Negative controls of no-template and water as template were used in all runs.

Selection and evaluation of candidate reference genes for RT-qPCR assays. To identify suitable internal reference genes for RT-qPCR normalization, five NWR genes were selected as candidates based on their stable expression in other plant species (Kozera & Rapacz, 2013; Joseph et al., 2018) and across eight NWR tissue types (Haas et al., 2021). These included putative homologs of actin 1 (ACT1), actin 4 (ACT4), eukaryotic translation initiation factor 5-like (EIF-5), glyceraldehyde 3-phosphate dehydrogenase (GAPDH), and ubiquitin (UBI). Primers were designed using PrimerQuest (Integrated DNA Technologies; Coralville, Iowa, USA) based on exon sequences extracted from the NWR reference genome v1.0 (Haas et al., 2021) using in-house R script. The amplification efficiency and expression stability of each primer set were assessed in leaf, male floret, and female floret tissues across all time points. Standard curves were generated using a 5-fold serial dilution of cDNA (200, 40, 8, 1.6, and 0.32 ng/μL). Sterile DNase- and RNase-free water was used as a control in all reactions.

Evaluation and selection of putative seed shattering orthologs. Protein sequences of well-characterized *O. sativa* seed shattering genes, *SH1*, *qSH1*, *sh4*, *Sh5*, and *SHAT1*, were retrieved from the *O. sativa* ssp. *japonica* reference genome (GenBank accession: GCF_034140825.1). Putative orthologs identified in *Z. latifolia* in Yan et al. 2022 were obtained from the Chinese wild rice 'Jaiobai' genome (GCA_0483537475.1; Guo et al., 2015). These *O. sativa* and *Z. latifolia* protein sequences were used as queries in BLASTp searches against the *Z. palustris* 'Itasca-C12' reference genome assembly v1.0 (GCA_019279435.1; Haas et al., 2021) using the NCBI BLASTp tool (www.ncbi.nlm.nih.gov/). Candidate orthologs in NWR were identified based on percent protein identity with a cut-off rate of 70%. Multiple sequence alignments (MSA) were performed using CLUSTALW v2.0 (Larkin et al., 2007) and a maximum likelihood (ML) phylogenetic tree with 100 bootstraps was calculated with the randomized accelerated ML (RAxML v8.2.11; Stamatakis, 2014) software. The resulting alignments and phylogram were visualized with ETE3 3.1.2 (Huerta-Cepas et al., 2016) via GenomeNET (<https://genome.ip/tools/ete/>), using default parameters. Protein sequence alignments for *qSh1* and *Sh5* orthologs were also conducted using NCBI's COBALT (Constraint-based Multiple Alignment Tool) with default parameters to visualize sequence divergence among homologs. Orthogroups were inferred using OrthoFinder version 2.5.4 (Emms & Kelly, 2015) and protein domain annotation was performed using Pfam version 37.4 (Finn et al., 2014) via the InterProScan web server (<https://www.ebi.ac.uk/interpro/>) to identify conserved domains in the predicted protein sequences. Based on these analyses, gene-specific primers (Supplemental Table 2) were designed for selected orthologs of *SH1* and *Sh5* for downstream RT-qPCR validation of gene expression, following the experimental protocol described above.

Data analysis. To assess variation in seed shattering, an analysis of variance (ANOVA) was conducted in R version 4.3.1 (R Core Team, 2021) for each of the four phenotyping methods, across four developmental timepoint. Genotype and timepoint were both considered fixed effects. A statistical significance threshold of 0.05 was used in all analyses. Tukey's Honestly Significant Difference (HSD) tests were performed using the *Agricolae* package version 1.3.7 (de Mendiburu, 2021) to identify pairwise differences among groups. To compare phenotyping methods, Pearson's correlation coefficients were calculated with the *Hmisc* package version 5.2.2 (Harrell, 2023).

For reference gene selection, primer efficiency was assessed for each candidate reference gene primer pair by constructing linear regression curves based on the logarithm of the cDNA and the threshold cycle (C_T). Primer efficiency (E) was calculated from the slope (S) of the standard curves generated from serial dilutions, using the equation $E = [10^{(1/-S)} - 1] \times 100\%$ (Rutledge & Cote, 2003). The coefficients of determination (R^2) were also computed. Expression stability across tissue types and timepoints were evaluated using RefFinder (Xie et al., 2023), which also utilizes geNorm (Vandesompele et al., 2002; Hellemans et al., 2008) and NormFinder (Andersen et al., 2004). The best candidate reference genes with amplification efficiencies within the optimal range ($100\% \pm 10\%$) and minimal variation in expression were selected for assessing gene expression of candidate shattering genes.

Gene expression levels of seed shattering-related genes in female and male tissues were quantified using the $2^{-\Delta\Delta C_T}$ method (Livak & Schmittgen, 2001). For each sample, threshold cycle (C_T) values of the target gene were first normalized to the selected reference gene to obtain CT values. Relative expression was then calculated by comparing normalized CT values of each tissue sample to those of the same genes expressed in flag leaf tissue, which was used as the baseline calibrator. The data was transformed using $\log_2$ and presented as relative gene expression. This approach enabled quantification of fold changes in gene expression relative to a consistent non-reproductive tissue control. To evaluate differential expression between the wild type population and the cNWR populations, ANOVA and Tukey's HSD were also performed.

RESULTS

Seed shattering variability in Northern Wild Rice populations

To evaluate genetic and phenotypic variation in seed shattering, we quantified this trait in three *Z. palustris* populations using four phenotyping methods at 21, 28, 35, and 42 days after anthesis (DAA). These methods included visual ratings, a drop test, and two biomechanical tension force methods (BTS-P and BTS-B). Statistical analyses (ANOVA, Tukey's HSD; Table 1–2) revealed visual ratings and the drop test results were significantly influenced by both genotypes and DAA, while BTS-P had significant variation by genotype and BTS-B had no statistical differences (Table 1). In general, visual and drop test scores showed increasing seed shattering over time, while BTS-B and BTS-P values, reflecting resistance to detachment, decreased (Table 2). Importantly, differences in the shattering phenotype were not due to variation in developmental timing, as all populations reached comparable phenological stages within one to three days of each other (Figure S1).

Table 1
Analysis of Variance (ANOVA) of four phenotyping methods to characterize seed shattering in Northern Wild Rice (NWR; *Zizania palustris*) post-anthesis.

Method	Source of Variation	Df	Sums of Squares	Mean Squares	F-value	p-value
Visual	DAA	3	635.42	211.81	5.45	< 0.01
	Genotype	2	69,705.56	34,852.78	896.21	< 0.0001
	DAA:Genotype	6	266.67	44.44	1.14	0.37
	Residuals	24	933.33	38.89		
Drop Test	DAA	3	2,527.16	842.39	16.01	< 0.0001
	Genotype	2	36,198.54	18,099.27	343.97	< 0.0001
	DAA:Genotype	4	189.28	47.32	0.90	0.48
	Residuals	20	1,052.39	52.62		
BTS-P	DAA	2	18,193.62	9,096.81	3.50	0.06
	Genotype	2	33,073.60	16,536.80	6.36	< 0.01
	DAA:Genotype	2	5,465.79	2,732.90	1.05	0.38
	Residuals	14	36,419.62	2,601.40		
BTS-B	DAA	2	0.00	0.00	0.70	0.51
	Genotype	2	0.00	0.00	2.05	0.17
	DAA:Genotype	2	0.00	0.00	1.10	0.36
	Residuals	14	0.00	0.00		

Legend. BTS-P: Breaking tensile strength by pulling; BTS-B: breaking tensile strength by bending; DAA: days after anthesis; Df: degrees of freedom

Table 2
Tukey's mean separations analysis for four phenotyping methods to characterize seed shattering in Northern Wild Rice (NWR; *Zizania palustris*) over time and populations.

	Visual (%)				Drop Test (%)				BTS_Pulling (g)				BTS_Bending (g)			
Genotype / DAA	21	28	35	42	21	28	35	42	21	28	35	42	21	28	35	42
Itasca-C12	0 b	0 b	0 b	8.33 b	0 d	0 d	6.25 cd	23.23 bc	NA	206.78 a	168.67 ab	142.93 ab	NA	0.014 a	0.023 a	0.01 a
FY-C20	0 b	0 b	1.67 b	13.33 b	0 d	0 d	10.42 cd	35.27 b	NA	185.67 ab	139.67 ab	44.20 b	NA	0.027 a	0.014 a	0.005 a
Wild type	86.67 a	98.33 a	100 a	100 a	91.67 a	98.33 a	100 a	100 a	NA	91.33 ab	NA	NA	NA	0 a	NA	NA

Legend: Days after anthesis, DAA; not applicable, NA; grams, g; Same letters indicate no statistically significant difference (P = 0.05)

Methods: A visual rating, 0-100% shattering; a drop test, 0-100% shattering; breaking tensile strength by pulling, BTS-P, and breaking tensile strength by bending (BTS_B) were measured in force (g)

For visual ratings, the wild type population consistently exhibited greater seed shattering than both cNWR populations, with no significant differences found between Itasca-C12 and FY-C20. Shattering in the wild type increased from 86.67% at 21 DAA to 100% by 35 DAA (Table 2). FY-C20 began shattering at 35 DAA, reaching 13.3% by 42 DAA, while Itasca-C12 showed delayed onset, with 8.3% shattering by 42 DAA (Table 2). The drop test also effectively differentiated the wild type from both cultivated populations across all time points (Table 1; Table 2). The visual and drop test methods were highly correlated with one another ($r = 0.99$, $p > 0.001$) across all time points (Supplementary Table 3).

For BTS-P, only one time point could be assessed for the wild type, which yielded 91.3 g of force at 28 DAA, compared to 185.7 g for FY-C20 and 206.8 g for Itasca-C12 (Table 2). By 42 DAA, Itasca-C12 required greater force to detach seeds than FY-C20 (144 g vs. 44 g; Table 2). Although BTS-P differences among populations were not statistically significant, Itasca-C12 exhibited higher BTS-P values compared to FY-C20 across all time points. BTS-P was negatively and significantly correlated with both visual and drop test results across time points, except with the visual method at 35 DAA (Supplementary Table 3). BTS-B displayed non-significant negative correlations with the visual and drop test methods across timepoints. BTS-B was also non-significantly positively correlated with BTS-P results across timepoints, with one significant correlation between BTS-B and BTS-P at 35 DAA (Supplementary Table 3).

Histological analysis of the seed-pedicel abscission layer. We examined the anatomical structure of the abscission layer in a wild type population, characterized by high levels of seed shattering, and two cultivated populations, Itasca-C12 at 28 DAA and FY-C20 at 21 DAA, which exhibit reduced seed shattering. All three populations developed a distinct abscission layer at the seed-pedicel junction (Fig. 1). In the wild type population (Fig. 1b), the

abscission layer was well-defined, consisting of densely packed and structurally organized parenchyma cells. In contrast, the cultivated populations appear to have more loosely arranged cells and larger intercellular spaces disrupting the continuity of the layer. Among the cultivated populations, FY-C20 showed a more compact and organized abscission layer than Itasca-C12, which had a more diffuse structure and well-defined vascular bundle (Fig. 1c and 1d, respectively).

Putative relationships of seed shattering-related genes in *Oryzae*. Protein sequences of five *O. sativa* genes involved in seed abscission layer development, *SH1* (*OsSH1*), *qSH1* (*OsqSH1*), *Sh5* (*OsSh5*), *sh4* (*Ossh4*), and *SHAT1* (*OsSHAT1*), were queried against the *Z. palustris* genome in NCBI GenBank using the BlastP tool (Supplementary Table 4). Multiple orthologs were identified in NWR, consistent with a whole-genome duplication (WGD) event in the *Zizania* lineage following its divergence from *Oryza* approximately 26 to 30 million years ago (Guo et al., 2015; Haas et al., 2021; Yan et al., 2022). Previously reported candidate orthologs (Haas et al., 2021) were confirmed as top hits based on percent sequence similarity and conserved domain structure, except for *OsSHAT1*, where ZPchr0004g38673 (*ZpSHAT1a*) and ZPchr0458g2256 (*ZpSHAT1b*) were stronger candidates than the previously reported ZPchr0013g34051 (Haas et al., 2021). Additionally, ZPchr0458g22823 (*Zpsh4c*) was identified as a novel candidate of *Ossh4*, based on high protein sequence similarity to *ZpSh4a*, *ZpSh4b*, and *Ossh4* (Table 3). A percent protein identity (PI) matrix was generated to evaluate the sequence similarity among *O. sativa*, *Z. palustris*, and *Z. latifolia* seed shattering potential orthologs (Table 3). The highest PI between *O. sativa* and *Z. palustris* was observed for *SH1* (89.34%), while the lowest PI was for *Sh5a* (71.23%; Table 3). The PI comparisons between NWR and *Z. latifolia* ranged from 70.00% for *Sh5a* to 98.84% for *SH1*.

Table 3

Percent protein identity (PI) of seed shattering-related genes in *Oryza sativa* (*Os*) and putative *Zizania palustris* (*Zp*) and *Zizania latifolia* (*Zl*) orthologs.

<i>O. sativa</i> (<i>Os</i>)		<i>Z. palustris</i> (<i>Zp</i>)		<i>Z. latifolia</i> (<i>Zl</i>)		Percent Identify Comparisons between Species			Gene description
Gene	GenBank ID	Candidate Gene	GenBank ID	Candidate genes	GenBank IDs (a and b, respectively)	Zp vs. Os	Zp vs. Zl a/b	Os vs. Zl a/b	
<i>OsSH1</i>	NP_001389121.1	<i>ZpSH1</i>	ZPchr0003g18426	<i>ZlSH1a</i> / <i>ZlSH1b</i>	ABZP36_005173 / ABZP36_018673	89.84	91.98 / 98.84	91.94 / 91.91	YABBY transcription factor (Lin et al., 2012)
<i>OsSh5</i>	XP_015637721.1	<i>ZpSh5a</i>	ZPchr0001g31104	<i>ZlSh5a</i>	ABZP36_013876	71.23	78.85 / 70.00	89.56 / 83.92	Bel1-type homeobox (Yoon et al., 2014)
		<i>ZpSh5b</i>	ZPchr0005g15825			84.07	90.05 / 95.11		
		<i>ZpSh5c</i>	ZPchr0010g10516	<i>ZlSh5b</i>	ABZP36_030161	87.5	95.64 / 89.25		
		<i>ZpSh5d</i>	ZPchr0010g7757			85.51	95.64 / 89.25		
<i>OsqSH1</i>	XP_015641948.1	<i>ZpqSH1a</i>	ZPchr0001g31318	<i>ZlqSH1a</i>	ABZP36_025045	83.42	85.44 / 91.43	85.66	Bel1-type homeobox (Konishi et al., 2006)
		<i>ZpqSH1b</i>	ZPchr0007g5872	<i>ZlqSH1b</i>	ABZP36_016879	83.74	90.86 / 83.28	85.07	
<i>OsSHAT1</i>	XP_015636848.1	<i>ZpSHAT1a</i>	ZPchr0004g38673	<i>ZlSHAT1a</i>	ABZP36_003055	83.65	90.32 / 88.45	85.74	AP2 transcription factor (Zhou et al., 2012)
		<i>ZpSHAT1b</i>	ZPchr0458g22566	<i>ZlSHAT1b</i>	ABZP36_007933	84.69	90.30 / 96.67	84.02	
<i>Ossh4</i>	XP_015635337.1	<i>Zpsh4a</i>	ZPchr0004g38781	<i>Zlsh4a</i> / <i>Zlsh4b</i>	ABZP36_008339 / ABZP36_004261	80	78.05 / 80.18	76.47 / 80.12	Trihelix transcription factor (Li et al., 2006)
		<i>Zpsh4b</i>	ZPchr0458g22499			79.17	87.41 / 92.78		
		<i>Zpsh4c</i>	ZPchr0458g22823			80.17	86.36 / 84.21		

Legend: Zl a and b genes identified in Yan et al., 2022; Zl a and b genes GenBank ID in the 'Jaiobai' genome (GCA_0483537475.1; Guo et al., 2015); See Supplementary Table 4 for protein sequences

A cluster analysis of *O. sativa* shattering genes and putative *Z. palustris* and *Z. latifolia* orthologs identified two main clusters in the cladogram, each with moderate bootstrap support of 75%. The first main cluster consisted of two subclusters separating *SH1* (BS = 98%) from *qSH1* and *SH5* (BS = 75%). The second cluster (BS = 75%) consisted of *sh4* and *SHAT1* subclusters, supported by BS values of 92% and 93%, respectively (Fig. 2). Within the *SH1* subcluster, *OsSH1* formed a distinct branch from *Zizania* sequences (BS = 98%), while the *ZISH1b* sequence was the closest ortholog to *ZpSH1* (BS = 65%; 98.84% PI) (Fig. 2; Table 3). In the *qSH1* subcluster (BS = 75%), *ZpqSH1a* and *ZlqSH1b* grouped together (BS = 98%; 91.43% PI). *OsqSH1*, albeit with a weak bootstrap (52%), grouped more closely with *ZpqSH1b* and *ZlqSH1a*, who were 90.86% identical and in a clade with a BS of 98%. Within the *sh5* subcluster (BS = 66%), *OsSh5* separated from *Zizania* genes, which grouped with a BS of 82%. *ZpSH5b* clustered closely with *ZISH5b* (BS = 100%; 95.11% PI), while *ZpSh5a*, *ZpSh5c*, and *ZpSh5d* clustered with *ZISH5a* (BS = 91%; and PI ranging from 78.85 to 95.64%), although *ZISH5a* formed an independent branch separated from the *ZpSH5* orthologs (BS = 71%). The variation in PI among the *ZpSh5* orthologs and across species was largely attributed to *ZpSh5a*. Within the second main cluster, *SHAT1* subcluster was well supported (BS = 93%), integrated by low supported groups (BS = 39%) containing *OsSHAT1*, *ZISHATb*, and *ZpSHAT1b*, and another (BS = 58%; 90.32% PI) of *ZISHAT1a* and *ZpSHAT1a* (Fig. 2; Table 3). The *sh4* subcluster (92% BS) exhibited a stepwise branching pattern, starting with *Zlsh4b*, followed by *Zpsh4a* (BS = 13%), *OsSh4* (BS = 37%), and *Zlsh4a* (BS = 92%), with *Zpsh4b* and *Zpsh4c* forming a well-supported terminal cluster (BS = 99% BS). Among the five genes evaluated, *Zizania* orthologs of *OsSh4* showed some of the most divergent protein sequences, with PI percentages ranging from 76.47–80.17% (Table 3).

Orthogroup and Pfam analyses provided additional support for several conserved shattering-related orthologs (Supplementary Table 5 and Supplementary Table 6). *ZpSH1* was identified as an ortholog of *OsSH1* and assigned to orthogroup OG0010120, with similar sequences in *Z. latifolia* (Supplementary Table 5). All orthologs contained a YABBY protein domain (Supplementary Table 6). For *Sh5*, OrthoFinder analysis indicated that *ZpSh5b* was the strongest orthogonal candidate for *OsSh5* within orthogroup OG0013611, which also contained *ZISH5a* and *ZISH5b* (Supplementary Table 5). Two closely related sequences *ZpSh5c* and *ZpSh5d* grouped within orthogroup OG0026538, with no clear orthologs in *O. sativa* or *Z. latifolia*. *ZpSh5a* could not be assigned to an orthogroup. Pfam analysis found that all *Sh5* orthologs contain a homeobox domain and a POX domain, except for *ZpSh5a* (Supplementary Table 6). MSA revealed that *ZpSh5a* exhibited a truncated protein sequence compared to those of other orthologs (Supplementary Fig. 2). The alignment of sequences also revealed that *OsSh5* and all *ZpSh5* genes had 100% protein identities for their homeobox domains but differed in their POX domains, within which *OsSh5* differed from *ZpSh5b* and *ZpSh5c* by four and eight amino acids, respectively (Supplementary Table 7). Comparison of *OsqSH1* identified *ZpqSH1b* as a likely ortholog and it was placed in orthogroup OG0008858, along with *ZlqSH1a* and *ZlqSH1b* (Supplementary Table 5), while *ZpqSH1a* was not assigned an orthogroup. However, all *Z. palustris* orthologs in this group, including *ZpqSH1a*, have homeobox and POX domains (Supplementary Table 6). For *OsSHAT1*, *ZpSHAT1b* was the strongest orthogonal candidate and grouped in orthogroup OG0003844 with *ZISHAT1a* and *ZISHAT1b* (Supplementary Table 5), while *ZpSHAT1a* was not assigned to an orthogroup. Pfam searches indicated that all *SHAT1* ortholog sequences contained two tandem AP2/ERF domains (Supplementary Table 6). Lastly, *OsSh4* along with *Zpsh4a*, *Zpsh4b*, and *Zpsh4c* genes were grouped within the orthogroup OG0004851 (Supplementary Table 5). All *Z. palustris* ortholog sequences for *sh4* contained a Myb/SANT-like DNA-binding domain 4 (Supplementary Table 6).

Reference gene selection for RT-qPCR experiments. Five commonly used reference genes in plant gene expression studies, ACT1, ACT4, EIF-5, UBI, and GAPDH, were evaluated across leaf, female floret, and male floret tissues to identify those with the most stable expression for normalizing RT-qPCR expression data (Supplementary Table 1). Primer efficiencies ranged from 106–113% with R^2 values between 0.997 and 0.998 (Table 4). Stability rankings using RefFinder consistently identified *ACT1* as the most stable and *ACT4* as the least stable expressed gene (Table 4). Based on these results, ACT1 was selected as the internal reference gene for downstream expression analyses.

Table 4
Primer efficiency and stability of reference gene candidates for gene expression analyses in male and female floret tissues of Northern Wild Rice (NWR; *Zizania palustris*) based on RefFinder, which includes geNorm and NormFinder analyses.

Gene	Z. palustris gene	Chr.	Position (bp)	Description	R2	%E	RefFinder	geNorm	NormFinder
ACT1	ZPchr0012g19383	11	23,232,269	Actin 1	0.998	110	1.19	0.89	0.94
ACT4	ZPchr0010g7718	10	55,029,365	actin-related protein 4	0.998	113	5	4.9	8.9
EIF-5	ZPchr0006g44672	6	52,111,321	eukaryotic translation initiation factor 5-like	0.997	107	1.41	0.89	1.47
GADPH	ZPchr0008g11968	8	9,259,023	glyceraldehyde-3-phosphate dehydrogenase	0.997	106	3	1.31	1.74
UBIQ	ZPchr0007g6160	7	6,416,073	ubiquitin-conjugating enzyme E2-17 kDa isoform X1	0.998	106	4	2.14	2.13
Legend: primer efficiency, %E; RefFinder (Xie et al., 2023); geNorm (Vandesompele et al., 2002; Hellemans et al., 2008); NormFinder (Andersen et al., 2004); lower values indicate higher stability and rank									

Expression of SH1 and Sh5 orthologs in Z. palustris female florets. We analyzed the relative expression of NWR candidate orthologs of *OsSH1* and *OsSh5*, *ZpSH1* and *ZpSh5a*, *ZpSh5b*, and *ZpSh5c* in NWR female floret tissues (including the seed-pedicel junction) across a time course following anthesis (Supplementary Table 2). *ZpSH1* was selected due to its high expression in female panicle tissue (Haas et al., 2021) and its conserved role in seed shattering across grass species (Maity et al., 2021). Among four *ZpSh5* putative paralogs, *ZpSh5a*, *ZpSh5b*, and *ZpSh5c* were included based on their relatively high expression levels (Haas, et al., 2021) and to explore their potential roles in the regulatory pathways of seed shattering, considering their putative diversification following the WGD. *ZpSh5d* was omitted from RT-qPCR assays due to its low expression in 8 NWR tissues (Haas et al., 2021).

To characterize candidate gene expression, female florets were collected at 1, 7, 18, 28, and 35 DAA for RT-qPCR analysis. Data from the wild type population were unavailable at 28 and 35 DAA due to near-complete seed shattering (98%) by 28 DAA (Table 2). The RT-qPCR gene expression time course profiles of female florets are summarized across genes in Fig. 3 and by population in Supplementary Fig. 3. *ZpSH1* was upregulated compared to leaf tissue, except for FY-C20 at 35 DAA (Fig. 3a). Across DAA, the expression of *ZpSH1* in the wild type was high on 1 DAA, peaked at 7 DAA, and declined at the last timepoint of the population's collection, 18 DAA. FY-C20's expression was lower than the wild type from 1 to 18 DAA but with peaks at 7 and at 28 DAA and was slightly downregulated by 35 DAA. Itasca-C12 had higher expression than the wild type from 1 to 18 DAA, peaking at 1 DAA and then declining through the rest of the timepoints. For *ZpSh5a*, expression in female floret tissues was downregulated at 1 and 7 DAA for all populations compared to leaf tissue (Fig. 3b). Expression of the wild type was slightly upregulated at 18 DAA, its final timepoint. The expression of *ZpSh5a* in FY-C20 after 7 DAA was upregulated with a peak at 28 DAA, and undetectable at 35 DAA. In Itasca-C12, the gene was downregulated from 1 DAA through 18 DAA, undetectable at 28 DAA, and upregulated at 35 DAA. The expression of *ZpSh5b* in female floret tissue was downregulated compared with leaf tissue for all populations and timepoints except for the wild type at 1 DAA (Fig. 3c). The expression of *ZpSh5c* in female floret tissue was upregulated for all populations and timepoints compared with leaf tissue (Fig. 3d). The expression in the wild type was highest at 1 DAA, then declined through 18 DAA. Expression in FY-C20 and Itasca-C12 was higher at 1 DAA than 7 and 18 DAA, after which expression peaked at 28 DAA for both populations and then declined at 35 DAA (Fig. 3d). However, the expression of FY-C20 throughout the time course was higher relative to Itasca-C12.

ANOVA conducted on a per-gene basis revealed that genotype was not significant for any of the genes (Table 5). Expression of *ZpSH1* and *ZpSh5a* was influenced by DAA; *ZpSH1* expression at 7 DAA was different from 35 DAA and *ZpSh5c* expression at 1 and 7 DAA was different from 28 DAA (Supplementary Table 8). No significant differences between populations or time points were observed for candidate genes *ZpSh5b* and *ZpSh5c*. ANOVA conducted on a per population basis revealed significant differences between gene expression for all populations, where *ZpSH1* and *ZpSh5c* had higher expression than *ZpSh5a* and *ZpSh5b* (Table 5; Supplementary Table 8). For FY-C20, there were also significant differences between timepoints; 28 DAA was significantly different from 1 DAA and 35 DAA. For Itasca-C12, the interaction between gene and DAA was significant. Overall, *ZpSH1* had higher expression, with 1 and 7 DAA being significantly higher than *ZpSh5b* at 7, 18, and 28 DAA and also, *ZpSh5a* at 1 and 7 DAA. The expression for all timepoints of *ZpSh5c* was intermediate and did not differ significantly from any other genes (Supplementary Table 8).

Table 5
Analysis of Variance (ANOVA) analyses of the relative gene expression of four Northern Wild Rice (NWR; *Zizania palustris*) candidate orthologs of *Oryza sativa* genes related to seed shattering in three NWR populations.

Gene	Source of Variation	Df	Sum of Squares	Mean Squares	F-value	P-value
ZpSH1	DAA	4	83.44	20.86	2.87	0.044
	Genotype	2	12.93	6.46	0.89	0.424
	Replication	1	10.69	10.69	1.47	0.237
	DAA:Genotype	6	23.95	3.99	0.55	0.766
	Residuals	25	181.9	7.28		
ZpSh5a	DAA	4	94	23.5	6.66	0.001
	Genotype	2	23.35	11.68	3.31	0.054
	Replication	1	4.13	4.13	1.17	0.29
	DAA:Genotype	6	28.58	4.76	1.35	0.275
	Residuals	24	84.73	3.53		
ZpSh5b	DAA	4	6.26	1.56	0.44	0.777
	Genotype	2	6.58	3.29	0.93	0.415
	Replication	1	1.02	1.02	0.29	0.6
	DAA:Genotype	5	20.97	4.19	1.18	0.36
	Residuals	16	56.69	3.54		
ZpSh5c	DAA	4	8.35	2.09	0.4	0.806
	Genotype	2	16.89	8.44	1.62	0.221
	Replication	1	20.25	20.25	3.89	0.062
	DAA:Genotype	6	22.38	3.73	0.72	0.641
	Residuals	21	109.36	5.21		
Genotype	Source of Variation	Df	Sum of Squares	Mean Squares	F-value	P-value
Wild type	DAA	2	18.91	9.46	2.35	0.118
	Genes	3	201.23	67.08	16.66	< 0.001
	Rep	1	0.09	0.09	0.02	0.882
	DAA:Genes	6	32	5.33	1.32	0.286
	Residuals	23	92.61	4.03		
FY-C20	DAA	4	85.9	21.48	3.51	0.018
	Genes	3	121.08	40.36	6.59	0.001
	Rep	1	19.06	19.06	3.11	0.088
	DAA:Genes	11	36.53	3.32	0.54	0.859
	Residuals	31	189.89	6.13		
Itasca-C12	DAA	4	4.99	1.25	0.27	0.892
	Genes	3	245.14	81.71	18	< 0.001
	Rep	1	26.69	26.69	5.88	0.021
	DAA:Genes	12	117.75	9.81	2.16	0.04
	Residuals	33	149.82	4.54		
*Degrees of freedom, df; days after anthesis, DAA						

Expression of SH1 and Sh5 orthologs in male florets. To investigate potential overlap in the regulatory mechanisms between male and female floret shattering, we analyzed expression in male florets at 7 DAA. Similar to the expression patterns in female florets, *ZpSh1* and *ZpSh5c* were upregulated in all populations (Fig. 3). Expression was highest in Itasca-C12 for *ZpSH1* and in the wild type for *ZpSh5c*. At 7 DAA, the expression of *ZpSh5a* in male florets was downregulated for all populations, consistent with the early DAA in female florets. For *ZpSh5b*, wild type male florets were upregulated while the

cultivated populations were downregulated, consistent with the expression patterns of female florets at 1 DAA (Fig. 3). When averaged across genes and populations the male floret expression at 7 DAA was significantly correlated with 1 (0.82), 7 (0.80), and 18 (0.38) DAA in the female floret tissue (Supplementary Table 9).

DISCUSSION

Selection-driven changes in cultivated Northern Wild Rice seed retention

Seed shattering is a key adaptive trait in wild grasses, enabling seed dispersal, but it presents a major obstacle in grain domestication. In cNWR, strong selection pressure for seed retention has modified the timing and structure of seed abscission, despite a short domestication history. Our results show that cultivated populations exhibited a 90% reduction and 2-week delay in shattering relative to the wild type (Table 2), possibly due to alteration of the abscission zone. Histological observations support this, where cultivated populations displayed less organized and compact abscission layers, consistent with delayed or disrupted cell separation processes (Fig. 1). These changes parallel those seen in domesticated *Oryza* species (Jin and Inouye, 1981; Oba et al., 1995; Ji et al., 2006; Lin et al., 2007; Thurber et al., 2013; Li et al., 2024) and suggest that similar cellular mechanisms are being targeted in NWR. The observed changes in tensile strength and shattering progression, particularly in the higher seed-retaining Itasca-C12 (Table 2), suggest population-specific modifications to abscission layer timing or robustness, reflecting the ongoing selection for this trait. While this study identified a potential relationship between seed shattering and the development of cNWR's abscission layer, further investigation into the timing of abscission layer development and degradation is needed, as histological assessments were limited to a single time point per population.

Gene duplication and divergence underlie shattering variation

Gene duplication and divergence are central to trait evolution following WGD events. Our comparative genomic analyses revealed that NWR harbors multiple orthologs, similar to *Z. latifolia* (Yan et al., 2022), of known *O. sativa* seed shattering genes, including *sh4*, *SHAT1*, *qSH1*, and *Sh5*. This redundancy, arising from a WGD and subsequent gene duplications in NWR (Haas et al., 2021), suggests that regulatory divergence rather than gene presence/absence may drive variation in NWR seed shattering. For example, *SHAT1* orthologs showed evidence of divergence in NWR. *ZpSHAT1b* clustered with known functional orthologs, while *ZpSHAT1a* did not cluster with any known orthogroup, possibly reflecting neofunctionalization or regulatory repurposing (Wang et al., 2012). The exception was a single functional ortholog of *OsSH1*, *ZpSh1*, which retained a YABBY transcription factor and consistently clustered with orthologs from *O. sativa* and *Z. latifolia* (Supplementary Table 5), indicating its conserved role in abscission regulation.

In particular, *ZpSh5*, a Bel1-homeobox domain, exhibited paralog divergence among four strong *Sh5* candidates. *ZpSh5a*'s loss of key homeobox and POX domains, for example, suggests it may be transitioning toward pseudogenization (Panchy et al., 2016). The retention of multiple *Sh5*-like genes, with varying degrees of structural and functional integrity, reflects the history of WGD and subsequent segmental or transposon-mediated duplications (Fedoroff, 2012; Huang et al., 2022; Leister, 2004; Wang et al., 2012). The differential expression of *ZpSh5* paralogs also supports a role for subfunctionalization or regulatory partitioning in NWR genes following duplication. While some copy number differences between *Z. palustris* and *Z. latifolia* could be due to assembly artifacts, others likely reflect true genomic expansion, consistent with the larger genome size of *Z. palustris*. This complexity highlights the challenge of linking genotype to phenotype in recently duplicated genomes and the need for functional assays to dissect roles of individual paralogs.

Developmental timing of gene expression differentiates shattering phenotypes

Developmental regulation of gene expression is a key determinant of phenotypic variation in domestication traits such as seed shattering. In this study, we identified a general trend of higher expression of *ZpSH1* and *ZpSh5c* relative to the other tested genes. *ZpSH1* exhibited higher expression in the early DAA compared to later DAA (Fig. 3a), consistent with *SH1*'s known role in promoting abscission layer formation in grasses (Lin et al., 2012; Yu et al., 2023) and supports the hypothesis that *ZpSH1* functions as a conserved regulator of shattering in NWR. The expression of *ZpSh5c*, by contrast, was reduced in cultivated populations compared to the wild type (Fig. 3d). The delay in peak expression of FY-C20 (28 DAA) aligned with FY-C20's shattering time course, consistent with *Sh5*'s known role in regulating abscission zone differentiation and downstream enzymes (Yoon et al., 2014) and suggests that altered expression timing could contribute to extended seed retention in NWR. A similar pattern of delay has been observed in U.S. weedy rice (*O. sativa*), which has re-evolved its shattering ability, compared to its wild progenitor, *Oryza rufipogon* (Thurber et al., 2011). These findings support a model in which selection for seed retention in NWR has acted on regulatory changes, such as the timing and magnitude of gene expression, and that population-specific responses to artificial selection vary in cNWR.

While sample size was limited, strong correlations between male and female floret expression in the early DAA also suggest that these tissues may have similar shattering mechanisms. In NWR, male floret shattering occurs prior to female floret shattering and has been used as an early, indirect selection tool for improving seed retention (Kennard et al., 2002). These results reinforce the utility of this selection method in breeding programs.

Overall, the developmental and regulatory changes identified in this study resemble patterns observed in *Oryza*, where domestication involved sequential, quantitative shifts in expression across multiple shattering-related loci (Ishikawa et al., 2022). We cannot rule out the unknown interactions between these genes and others not included in this study that can modulate seed shattering in NWR and should be subjects of future research to understand mechanisms of seed shattering in this native North American species.

CONCLUSIONS

Despite the short breeding history of cNWR, the reduction of seed shattering is already evident at the anatomical, physiological, and gene expression levels. Delayed abscission layer development, increased tensile strength, and altered expression of key regulatory genes collectively support the hypothesis that

selection has reshaped the developmental and genetic control of shattering in cNWR. Comparative genomic analyses revealed duplicated orthologs of *O. sativa* shattering genes, consistent with the WGD history and genomic complexity of *Z. palustris*. Among these, *ZpSH1* and *ZpSh5c* showed biologically relevant expression variation across populations, suggesting roles in modulating abscission zone formation and seed retention. *ZpSh5c*'s delayed expression in one of the cultivated populations aligns with its delayed shattering phenotype and may reflect a target of selection. Together, these findings highlight the potential of cNWR as a model for investigating the early stages of domestication. Its phylogenetic proximity to *Oryza* and ongoing artificial selection offer a rare opportunity to study the molecular evolution of domestication traits in real time. Further integrative studies across broader germplasm, incorporating functional assays and regulatory network analysis, will be key to resolving the complex genetic basis of seed retention in this North American cereal.

Abbreviations

analysis of variance, ANOVA

breaking tensile strength, BTS

breaking tensile strength by pulling, BTS-P

breaking tensile strength by bending, BTS-B

chromosome, ch

complementary DNA, cDNA

cultivated Northern Wild Rice, cNWR

days after anthesis, DAA

Northern Wild Rice, NWR

Oryza sativa, *Os*

principal phenological stage, PPS

quantitative trait loci, QTL

reverse transcription-quantitative polymerase chain reaction, RT-qPCR

University of Minnesota, UMN

whole genome duplication, WGD

Zizania latifolia, *Zp*

Zizania palustris, *Zl*

Declarations

Ethics Approval and Consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Most data is provided within the manuscript or supplementary information files. Other raw data files associated with this project have also been submitted to the Data Repository for the University of Minnesota (DRUM) and can be accessed via doi XXXX.

Competing interests

The authors declare no conflicts of interest

Funding

This work was supported by the State of Minnesota, Agricultural Research, Education, Extension and Technology Transfer program.

Author contributions

R. M.: data generation, expression data generation, figure generation, writing; A. M.: histology data generation; M. B.: data generation, writing, editing; L. M.: figure generation, writing, editing; C. C. M.: data analysis and interpretation; editing; J. K.: original writing, editing, figure generation

Acknowledgements

The authors would like to thank the University of Minnesota's Clinical and Translational Science Institute (<https://ctsi.umn.edu/services/specimen/histology-digital-imaging>) for their histological imaging of the seed-pedicel abscission layer.

References

1. Altendorf, K. R., DeHaan, L. R., Larson, S. R., & Anderson, J. A. (2021). QTL for seed shattering and threshability in intermediate wheatgrass align closely with well-studied orthologs from wheat, barley, and rice. *The Plant Genome*, 14(3), e20145. <https://doi.org/10.1002/tpg2.20145>
2. Andersen, C. L., Jensen, J. L., & Ørntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*, 64(15), 5245–5250. <https://doi.org/10.1158/0008-5472.CAN-04-0496>
3. Cai, H. W., & Morishima, H. (2000). Genomic regions affecting seed shattering and seed dormancy in rice. *Theoretical and Applied Genetics*, 100(6), 840–846. <https://doi.org/10.1007/s001220051360>
4. de Mendiburu, F. (2021). *agricolae tutorial* (Version 1.3-5). Universidad Nacional Agraria, La Molina, Peru.
5. Doebley, J. F., Gaut, B. S., & Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell*, 127(7), 1309–1321. <https://doi.org/10.1016/j.cell.2006.12.006>
6. Dong, Y., & Wang, Y. Z. (2015). Seed shattering: From models to crops. *Frontiers in Plant Science*, 6, Article 476. <https://doi.org/10.3389/fpls.2015.00476>
7. Duquette, J., & Kimball, J. A. (2020). Phenological stages of cultivated northern wild rice according to the BBCH scale. *Annals of Applied Biology*, 176(3), 350–356. <https://doi.org/10.1111/aab.12613>
8. Elliott, W. A., & Perlinger, G. J. (1977). Inheritance of shattering in wild rice. *Crop Science*, 17(6), 851–853. <https://doi.org/10.2135/cropsci1977.0011183x001700060008x>
9. Emms, D. M., & Kelly, S. (2015). OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, 16, Article 157. <https://doi.org/10.1186/s13059-015-0721-2>
10. Fedoroff, N. V. (2012). Transposable elements, epigenetics, and genome evolution. *Science*, 338(6108), 758–767. <https://doi.org/10.1126/science.338.6108.758>
11. Finn, R. D., Bateman, A., Clements, J., Coghill, P., Eberhardt, R. Y., Eddy, S. R., ... & Punta, M. (2014). Pfam: the protein families database. *Nucleic acids research*, 42(D1), D222–D230.
12. Fu, Z., Song, J., Zhao, J., & Jameson, P. E. (2019). Identification and expression of genes associated with the abscission layer controlling seed shattering in *Lolium perenne*. *AoB Plants*, 11, ply076. <https://doi.org/10.1093/aobpla/ply076>
13. Fuller, D. Q., & Allaby, R. (2009). Seed dispersal and crop domestication: Shattering, germination and seasonality in evolution under cultivation. In L. Østergaard (Ed.), *Annual Plant Reviews: Fruit Development and Seed Dispersal* (Vol. 38, pp. 238–295). Wiley-Blackwell.
14. Guo, L., Qiu, J., Han, Z., Ye, Z., Chen, C., Liu, C., ... & Fan, L. (2015). A host plant genome (*Zizania latifolia*) after a century-long endophyte infection. *The Plant Journal*, 83(4), 600–609. <https://doi.org/10.1111/tpj.12906>
15. Haas, M., Kono, T., Macchietto, M., Millas, R., McGilp, L., Shao, M., ... & Kimball, J. (2021). Whole genome assembly and annotation of northern wild rice, *Zizania palustris* L., supports a whole genome duplication in the *Zizania* genus. *The Plant Journal*, 107(6), 1806–1816. <https://doi.org/10.1111/tpj.15419>
16. Hanten, H. B., Ahlgren, G. E., & Carlson, J. B. (1980). The morphology of grain abscission in *Zizania aquatica*. *Canadian Journal of Botany*, 58(21), 2269–2273. <https://doi.org/10.1139/b80-261>
17. Harrell, F. E., Jr., & Harrell, M. F. E., Jr. (2023). *Package 'Hmisc'*. CRAN. <https://CRAN.R-project.org/package=Hmisc>
18. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., & Vandesompele, J. (2008). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology*, 8, R19. <https://doi.org/10.1186/gb-2007-8-2-r19>
19. Huang, Y. L., Zhang, L. K., Zhang, K., Chen, S. M., Hu, J. B., & Cheng, F. (2022). The impact of tandem duplication on gene evolution in Solanaceae species. *Journal of Integrative Agriculture*, 21(4), 1004–1014. [https://doi.org/10.1016/S2095-3119\(21\)63658-3](https://doi.org/10.1016/S2095-3119(21)63658-3)
20. Huerta-Cepas, J., Serra, F., & Bork, P. (2016). ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Molecular Biology and Evolution*, 33(6), 1635–1638. <https://doi.org/10.1093/molbev/msw046>
21. Imle, P. T. (2001). *QTL verification and testcross analysis of seed shattering in wild rice (Zizania palustris L.)* [Master's thesis, University of Minnesota]. University of Minnesota Digital Conservancy.
22. Ishikawa, R., Castillo, C. C., Htun, T. M., Numaguchi, K., Inoue, K., Oka, Y., ... Ishii, T. (2022). A stepwise route to domesticate rice by controlling seed shattering and panicle shape. *Proceedings of the National Academy of Sciences*, 119(26), e2121692119. <https://doi.org/10.1073/pnas.2121692119>
23. Ishikawa, R., Thanh, P. T., Nimura, N., Htun, T. M., Yamasaki, M., & Ishii, T. (2010). Allelic interaction at seed-shattering loci in the genetic backgrounds of wild and cultivated rice species. *Genes and Genetic Systems*, 85(4), 265–271. <https://doi.org/10.1266/ggs.85.265>

24. Ji, H. S., Chu, S. H., Jiang, W., Cho, Y. I., Hahn, J. H., Eun, M. Y., ... Koh, H. J. (2006). Characterization and mapping of a shattering mutant in rice that corresponds to a block of domestication genes. *Genetics*, 173(2), 995–1005. <https://doi.org/10.1534/genetics.105.054031>
25. Jin, I., & Inouye, J. (1981). On the degree of grain shedding of Japonica-Indica hybrid rice bred in Korea. *Japanese Journal of Crop Science*, 50, 181–185. <https://doi.org/10.1626/jcs.50.181>
26. Jin, I. D. & Inouye, J. (1982a). Relation between grain shedding and pedicel morphology near the abscission layer of japonica-indica hybrid rices bred in Korea. *Jpn. J. Crop Sci.* 51: 271–275.
27. Jin, I. D. & Inouye, J. (1982b). Relationship between grain shedding and abscission layer in pedicel of japonica-indica hybrid rices in Korea. *Jpn. J. Breed.* 51: 43–50.
28. Jin, I. D., & Inouye, J. (1985). On the degree of grain shedding, histological peculiarity of abscission region and esterase isozyme genotype of Bulu and Tjereh rice varieties originated in Indonesia. *Japanese Journal of Crop Science*, 54, 373–378.
29. Jin, I. D., Terao, H., & Inouye, J. (1982). On the cracking of abscission layer in Asian rice cultivar (*Oryza sativa* L.). *Japanese Journal of Crop Science*, 51, 542–545.
30. Jin, I. D., Inouye, J., & Quat, N. N. (1990). Histological peculiarities of the abscission layers of African rice, *Oryza glaberrima* Steud., and its relation with degree of grain shedding. *Japanese Journal of Crop Science*, 59, 475–480.
31. Jin, I. D., Bae, Y. H., & Inouye, J. (1995). Formation and development of abscission layer between pedicel and rachilla, and changes in grain shedding during ripening in African rice, *Oryza glaberrima* Steud. *Korean Journal of Crop Science*, 40, 103–112.
32. Jin, I. D., Sano, Y., & Inouye, J. (1992). Histological similarities of abscission layers in the pedicel of Asian and African rices and their relatives. *Japanese Journal of Crop Science*, 61(2), 257–263. <https://doi.org/10.1626/jcs.61.257>
33. Joseph, J. T., Poolakkalody, N. J., & Shah, J. M. (2018). Plant reference genes for development and stress response studies. *Journal of Biosciences*, 43, 173–187. <https://doi.org/10.1007/s12038-018-9730-3>
34. Kahler, A. L., Kern, A. J., Porter, R. A., & Phillips, R. L. (2014). Maintaining food value of wild rice (*Zizania palustris* L.) using comparative genomics. In R. Tuberosa, A. Graner, & E. Frison (Eds.), *Genomics of plant genetic resources: Volume 2. Crop productivity, food security and nutritional quality* (pp. 233–248). Springer Netherlands.
35. Kennard, W., Phillips, R., & Porter, R. (2002). Genetic dissection of seed shattering, agronomic, and color traits in American wildrice (*Zizania palustris* var. *interior* L.) with a comparative map. *Theoretical and Applied Genetics*, 105(6–7), 1075–1086. <https://doi.org/10.1007/s00122-002-0988-z>
36. Konishi, S., Izawa, T., Lin, S. Y., Ebana, K., Fukuta, Y., Sasaki, T., & Yano, M. (2006). An SNP caused loss of seed shattering during rice domestication. *Science*, 312(5778), 1392–1396. <https://doi.org/10.1126/science.1126410>
37. Kozera, B., & Rapacz, M. (2013). Reference genes in real-time PCR. *Journal of Applied Genetics*, 54, 391–406. <https://doi.org/10.1007/s13353-013-0173-x>
38. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ... Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>
39. Lee, G. H., Kang, I. K., & Kim, K. M. (2016). Mapping of novel QTL regulating grain shattering using doubled haploid population in rice (*Oryza sativa* L.). *International Journal of Genomics*, 2016, Article 2128010. <https://doi.org/10.1155/2016/2128010>
40. Leister, D. (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends in Genetics*, 20(3), 116–122. <https://doi.org/10.1016/j.tig.2004.01.004>
41. Lenser, T., & Theißen, G. (2013). Molecular mechanisms involved in convergent crop domestication. *Trends in Plant Science*, 18(12), 704–714. <https://doi.org/10.1016/j.tplants.2013.08.007>
42. Li, X., Lowey, D., Lessard, J., & Caicedo, A. L. (2024). Comparative histology of abscission zones reveals the extent of convergence and divergence in seed shattering in weedy and cultivated rice. *Journal of Experimental Botany*, 75(16), 4837–4850. <https://doi.org/10.1093/jxb/erae221>
43. Li, C., Zhou, A., & Sang, T. (2006). Rice domestication by reducing shattering. *Science*, 311(5769), 1936–1939. <https://doi.org/10.1126/science.1123604>
44. Li, F., Numa, H., Hara, N., Sentoku, N., Ishii, T., Fukuta, Y., ... Kato, H. (2019). Identification of a locus for seed shattering in rice (*Oryza sativa* L.) by combining bulked segregant analysis with whole-genome sequencing. *Molecular Breeding*, 39, Article 20. <https://doi.org/10.1007/s11032-019-0935-z>
45. Li, W., & Gill, B. S. (2006). Multiple genetic pathways for seed shattering in the grasses. *Functional & Integrative Genomics*, 6(4), 300–309. <https://doi.org/10.1007/s10142-005-0015-y>
46. Lin, Z., Griffith, M. E., Li, X., Zhu, Z., Tan, L., Fu, Y., ... Sun, C. (2007). Origin of seed shattering in rice (*Oryza sativa* L.). *Planta*, 226(1), 11–20. <https://doi.org/10.1007/s00425-006-0460-4>
47. Lin, Z., Li, X., Shannon, L. M., Yeh, C.-T., Wang, M. L., Bai, G., ... Yu, J. (2012). Parallel domestication of the Shattering1 genes in cereals. *Nature Genetics*, 44(6), 720–724. <https://doi.org/10.1038/ng.2281>
48. Liu, H., Fang, X., Zhou, L., Li, Y., Zhu, C., Liu, J., ... Lin, Z. (2022). Transposon insertion drove the loss of natural seed shattering during foxtail millet domestication. *Molecular Biology and Evolution*, 39(6), msac078. <https://doi.org/10.1093/molbev/msac078>
49. Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods*, 25(4), 402–408. <https://doi.org/10.1006/meth.2001.1262>
50. McGilp, L., Castell-Miller, C., Haas, M., Millas, R., & Kimball, J. (2023). Northern wild rice (*Zizania palustris* L.) breeding, genetics, and conservation. *Crop Science*, 63(4), 1904–1933. <https://doi.org/10.1002/csc2.21045>

51. Oba, S., Sumi, N., Fujimoto, F., & Yasue, T. (1995). Association between grain shattering habit and formation of abscission layer controlled by grain shattering gene *sh-2* in rice (*Oryza sativa* L.). *Japanese Journal of Crop Science*, 64(3), 607–615. <https://doi.org/10.1626/jcs.64.607>
52. Odonkor, S., Choi, S., Chakraborty, D., Martinez-Bello, L., Wang, X., Bahri, B. A., ... & Devos, K. M. (2018). QTL mapping combined with comparative analyses identified candidate genes for reduced shattering in *Setaria italica*. *Frontiers in Plant Science*, 9, 918. <https://doi.org/10.3389/fpls.2018.00918>
53. Oelke, E. A. (Ed.). (1982). *Wild rice production in Minnesota*. University of Minnesota, Agricultural Extension Service.
54. Olsen, K. M., & Wendel, J. F. (2013). A bountiful harvest: Genomic insights into crop domestication phenotypes. *Annual Review of Plant Biology*, 64, 47–70. <https://doi.org/10.1146/annurev-arplant-050312-120048>
55. Panchy, N., Lehti-Shiu, M., & Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant Physiology*, 171(4), 2294–2316. <https://doi.org/10.1104/pp.16.00523>
56. Porter, R. A., Kahler, A. L., & Phillips, R. L. (2008). Inheritance and characterization of gynoeious “pistillate” panicle in American wildrice. *ASA-CSSA-SSSA Meeting Abstracts*, Abstract No. 552-3.
57. Purugganan, M. D., & Fuller, D. Q. (2009). The nature of selection during plant domestication. *Nature*, 457, 843–848. <https://doi.org/10.1038/nature07895>
58. Rutledge, R. G., & Côté, C. (2003). Mathematics of quantitative kinetic PCR and the application of standard curves. *Nucleic Acids Research*, 31(16), e93. <https://doi.org/10.1093/nar/gng093>
59. Stamatakis, A. (2014). RAXML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
60. Thomson, M. J., Tai, T. H., McClung, A. M., Lai, X. H., Hinga, M. E., Lobos, K. B., ... & McCouch, S. R. (2003). Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theoretical and Applied Genetics*, 107(3), 479–493. <https://doi.org/10.1007/s00122-003-1270-8>
61. Thurber, C. S., Hepler, P. K., and Caicedo, A. L. (2011). Timing is everything: early degradation of abscission layer is associated with increased seed shattering in US weedy rice. *BMC plant biology*, 11, 1-10.
62. Thurber, C. S., Jia, M. H., Jia, Y., & Caicedo, A. L. (2013). Similar traits, different genes? Examining convergent evolution in related weedy rice populations. *Molecular Ecology*, 22(3), 685–698. <https://doi.org/10.1111/mec.12147>
63. Tranbarger, T. J., Tucker, M. L., Roberts, J. A., & Meier, S. (2017). Editorial: Plant organ abscission: From models to crops. *Frontiers in Plant Science*, 8, 196. <https://doi.org/10.3389/fpls.2017.00196>
64. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7), research0034.1. <https://doi.org/10.1186/gb-2002-3-7-research0034>
65. Wang, Y., Wang, X., & Paterson, A. H. (2012). Genome and gene duplications and gene expression divergence: a view from plants. *Annals of the New York Academy of Sciences*, 1256(1), 1-14.
66. Watanabe, K., Oba, S., & Horiuchi, T. (2003). Allelic test of rice shattering genes *sh1* and *sh2* in an F₂ population derived from the cross between Momigaredatsu and Dee-Geo-Woo-Gen (*Oryza sativa* L.). *SABRAO Journal of Breeding and Genetics*, 35, 57–64.
67. Woods, D. L., & Clark, K. W. (1976). Preliminary observations on the inheritance of non-shattering habit in wild rice. *Canadian Journal of Plant Science*, 56, 197–198.
68. Wu, W., Liu, X., Wang, M., Meyer, R. S., Luo, X., Ndjondjop, M. N., ... Zhu, Z. (2017). A single-nucleotide polymorphism causes smaller grain size and loss of seed shattering during African rice domestication. *Nature Plants*, 3, 17064. <https://doi.org/10.1038/nplants.2017.64>
69. Xie, F., Wang, J., & Zhang, B. (2023). RefFinder: A web-based tool for comprehensively analyzing and identifying reference genes. *Functional & Integrative Genomics*, 23(2), 125. <https://doi.org/10.1007/s10142-023-01055-7>
70. Yan, N., Yang, T., Yu, X. T., Yang, Y., Zhong, S., Liu, Y., ... Ma, X. (2022). Chromosome-level genome assembly of *Zizania latifolia* provides insights into its seed shattering and phytocassane biosynthesis. *Communications Biology*, 5, 36. <https://doi.org/10.1038/s42003-021-02993-3>
71. Yoon, J., Cho, L. H., Kim, S. L., Choi, H., Koh, H. J., & An, G. (2014). The BEL1-type homeobox gene *SH5* induces seed shattering by enhancing abscission-zone development and inhibiting lignin biosynthesis. *The Plant Journal*, 79(5), 717–728. <https://doi.org/10.1111/tpj.12581>
72. Yu, Y., Hu, H., Voytas, D. F., Doust, A. N., & Kellogg, E. A. (2023). The YABBY gene *SHATTERING1* controls activation rather than patterning of the abscission zone in *Setaria viridis*. *New Phytologist*, 240(2), 846–862. <https://doi.org/10.1111/nph.19157>
73. Zhou, Y., Lu, D., Li, C., Luo, J., Zhu, B. F., Zhu, J., & Han, B. (2012). Genetic control of seed shattering in rice by the APETALA2 transcription factor *Shattering Abortion1*. *The Plant Cell*, 24(3), 1034–1048. <https://doi.org/10.1105/tpc.111.094383>

Figures

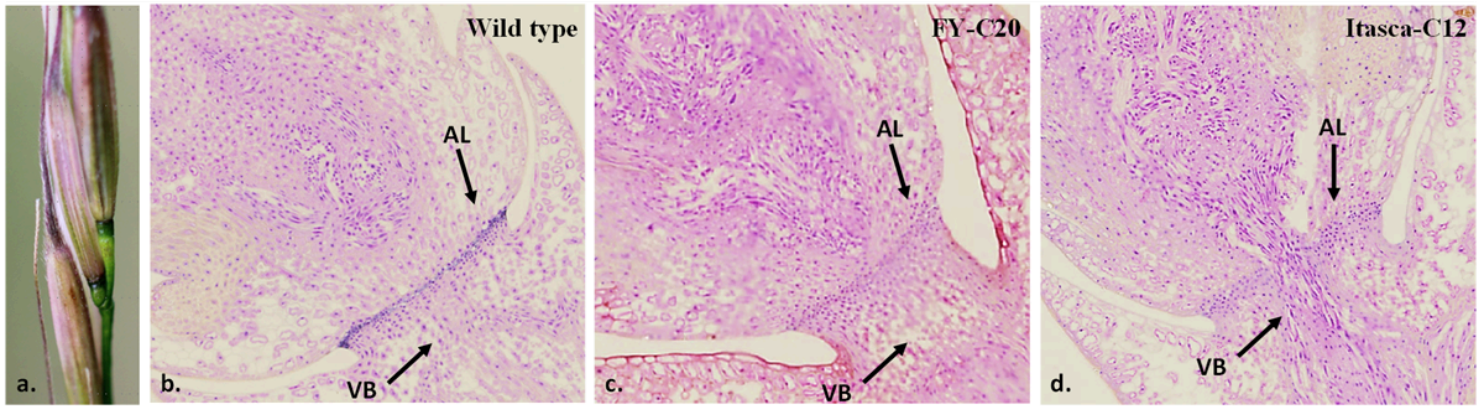


Figure 1

Morphology and abscission layer histology of the seed-pedicel junctions in Northern Wild Rice (NWR; *Zizania palustris*) populations at the hard dough seed phenological stage.

Legend. a. Representative image of the NWR seed and pedicel at early seed developmental stage; b. Unselected, highly shattering wild type genotype; c. Shattering resistant FY-C20 genotype and d. Higher shattering resistance Itasca-C20 genotype. AL: abscission layer; VB: vascular bundle.

Histological images were observed at a 20X magnification at the plant principal phenological stage 85 (hard dough seed) which corresponded to 21 days after anthesis (DAA) for the wild type, and 28 DAA for the cultivated populations.

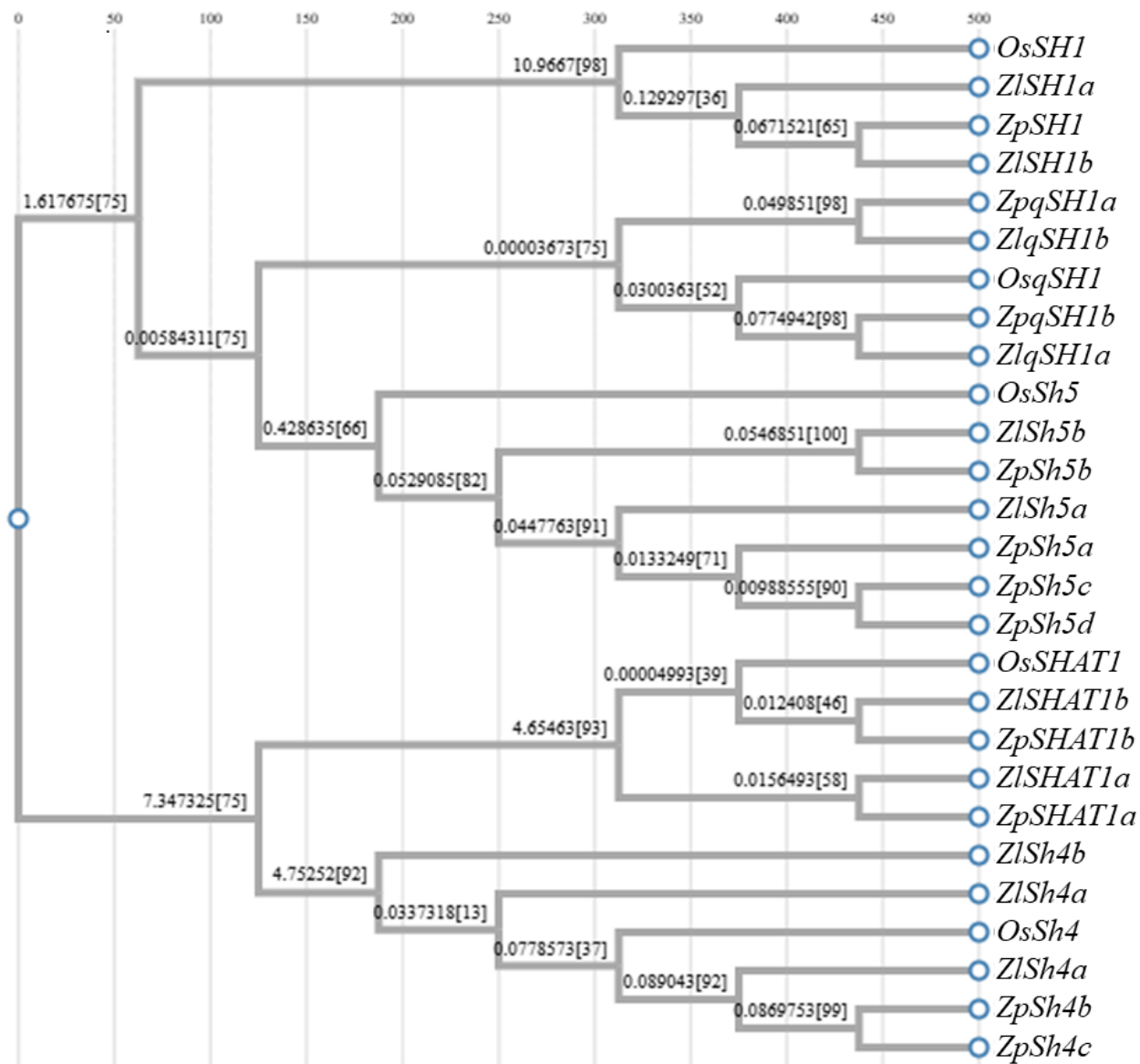


Figure 2

Phylogenetic tree of seed shattering *Oryza sativa* putative orthologs in *Zizania palustris* and *Zizania latifolia*.

Legend. *Os* = *Oryza sativa*, *Zp* = *Zizania palustris*, *Zl* = *Z. latifolia*; *SH1* = *Shattering1* (Lin et al., 2012), *qSH1* = QTL of seed shattering in chromosome 1 (Konishi et al., 2006), *SHAT1* = *Shattering Abortion1* (Zhou et al, 2012), *sh4* = (Li et al, 2006), *Sh5* (Yoon et al, 2014). Midpoint rooted cladogram constructed with aligned protein sequences of shattering genes using the ETE3 pipeline from Genome Net (www.genome.jp) and 100 bootstraps. Branch length and bootstrap values (in square brackets) are located at branches.

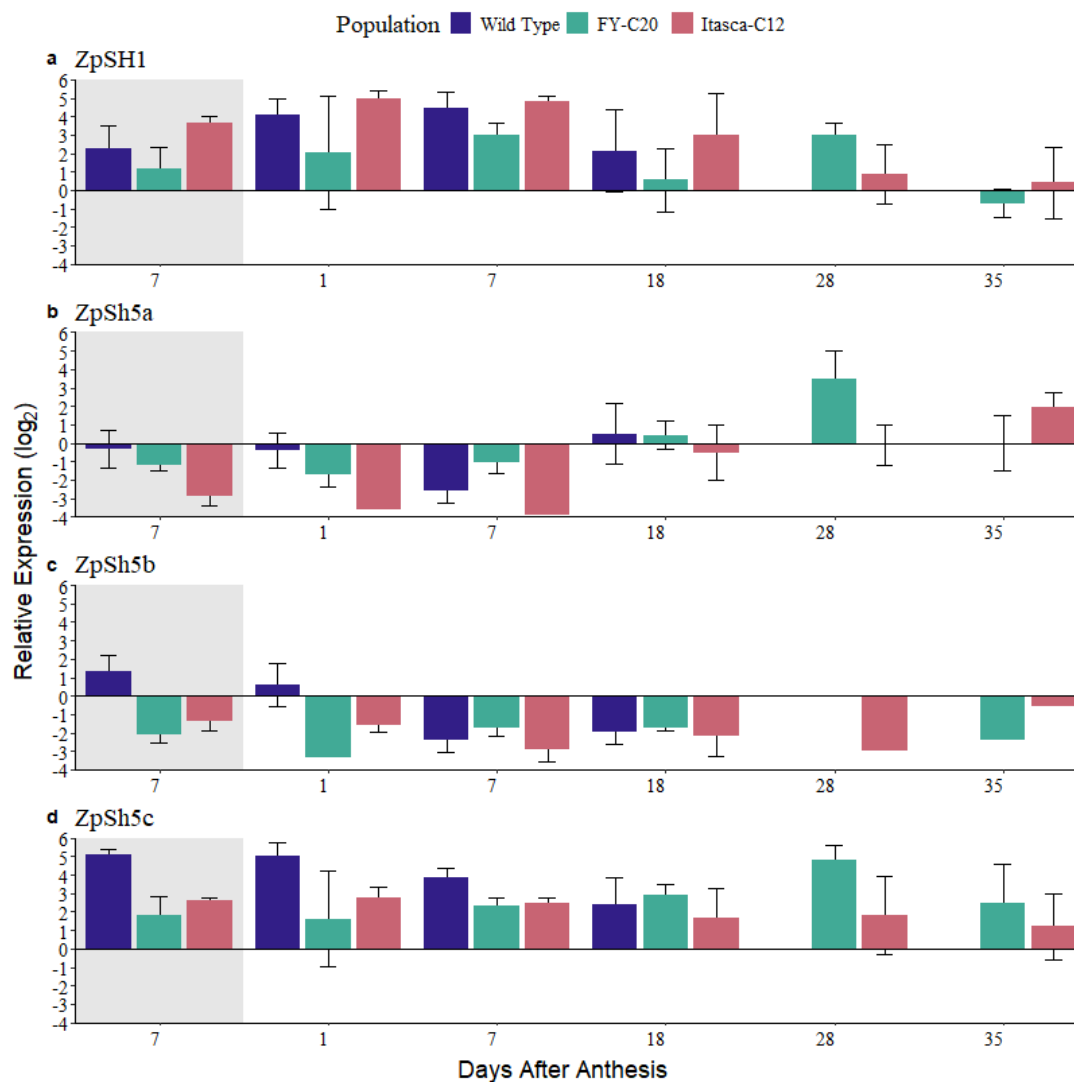


Figure 3

Relative expression of candidate seed-shattering genes in female and male floret tissues of Northern Wild Rice (NWR; *Zizania palustris*) post-anthesis across a wild type and two cultivated populations.

Legend: *Zp* = *Zizania palustris*, *qSH1* = QTL of seed shattering in chromosome 1 (Konishi et al., 2006), *Sh5* = (Yoon et al, 2014). Log₂ relative gene expression was measured in male florets collected at 7 days after anthesis (DAA) (indicated in gray) and in female florets collected at 1, 7, 18, 28, and 35 DAA across 4 genes

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SUPPLEMENTALTABLESANDFIGURES.docx](#)