frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 12 December 2024
DOI 10.3389/fpls.2024.1485819

Check for updates

OPEN ACCESS

EDITED BY
Shri Mohan Jain,
University of Helsinki, Finland

REVIEWED BY
Sebastian John Adams,
University of Mississippi, United States
Kurella Brs Visarada,
ICAR-Indian Institute of Millets Research (IIMR), India

*CORRESPONDENCE
Ayalew Ligaba-Osena
alosena@uncg.edu

†PRESENT ADDRESS
Eric Whisnant,
Department of Biology, Duke University,
Durham, NC, United States

RECEIVED 24 August 2024
ACCEPTED 18 November 2024
PUBLISHED 12 December 2024

CITATION
Whisnant ED, Keith C, Smieska L, Chia J-C, Bekele-Alemu A, Vatamaniuk OK, VanBuren R and Ligaba-Osena A (2024) Biggest of tinies: natural variation in seed size and mineral distribution in the ancient crop tef [*Eragrostis tef* (Zucc.) Trotter]. *Front. Plant Sci.* 15:1485819.
doi: 10.3389/fpls.2024.1485819

COPYRIGHT
© 2024 Whisnant, Keith, Smieska, Chia, Bekele-Alemu, Vatamaniuk, VanBuren and Ligaba-Osena. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Biggest of tinies: natural variation in seed size and mineral distribution in the ancient crop tef [*Eragrostis tef* (Zucc.) Trotter]

Eric D. Whisnant[1†], Christian Keith[1], Louisa Smieska[2], Ju-Chen Chia[3], Abreham Bekele-Alemu[1], Olena K. Vatamaniuk[3], Robert VanBuren[4] and Ayalew Ligaba-Osena[1*]

[1]Laboratory of Plant Molecular Biology and Biotechnology, Department of Biology, The University of North Carolina at Greensboro, Greensboro, NC, United States, [2]Cornell High Energy Synchrotron Source, Cornell University, Ithaca, NY, United States, [3]Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, United States, [4]Department of Horticulture, Michigan State University, East Lansing, MI, United States

Tef [*Eragrostis tef* (Zucc.) Trotter] is the major staple crop for millions of people in Ethiopia and Eritrea and is believed to have been domesticated several thousand years ago. Tef has the smallest grains of all the cereals, which directly impacts its productivity and presents numerous challenges to its cultivation. In this study, we assessed the natural variation in seed size of 189 tef and 11 accessions of its wild progenitor Indian lovegrass (*Eragrostis pilosa* (L.) P. Beauv.) and explored the mineral distribution of representative accessions. Our findings revealed significant natural variation in seed size and mineral concentration among both the tef and *E. pilosa* accessions. We observed significant variation in seed length, seed width, and seed area among the accessions of both *Eragrostis* spp. we analyzed. Using representative accessions of both species, we also found significant variation in 1000-grain weight. The observed variation in seed size attributes prompted us to use comparative genomics to identify seed size regulating genes based on the well-studied and closely related monocot cereal rice [*Oryza sativa* (L.)]. Using this approach, we identified putative orthologous genes in the tef genome that belong to a number of key pathways known to regulate seed size in rice. Phylogenetic analysis of putative tef orthologs of ubiquitin-proteasome, G-protein, MAPK, and brassinosteroid (BR)-family genes indicate significant similarity to seed size regulating genes in rice and other cereals. Because tef is known to be more nutrient-dense than other more common cereals such as rice, wheat, and maize, we also studied the mineral concentration of selected accessions using ICP-OES and explored their distribution within the seeds using synchrotron-based X-ray fluorescence (SXRF) microscopy. The findings showed significant variation in seed mineral concentration and mineral distribution among the selected accessions of both *Eragrostis* spp. This study highlights the natural variation in seed size attributes, mineral concentration, and distribution, while establishing the basis for

understanding the genetic mechanisms regulating these traits. We hope our findings will lead to a better understanding of the evolution of tef at the genetic level and for the development of elite tef cultivars to improve seed size, yield, and quality of the grains.

KEYWORDS

***Eragrostis tef*, *Eragrostis pilosa*, seed size, mineral distribution, mineral concentration, seed size regulating genes**

1 Introduction

Agricultural research has historically overlooked crops such as tef [*Eragrostis tef* (Zucc.) Trotter] which are often called "orphan crops" or "underutilized crops", defined as crops that are regionally limited, indigenous, grown for subsistence farming, and often neglected scientifically (Talabi et al., 2022). Orphan crops typically lack the scale and size to compete with more globally popular crops. This means most agricultural research focuses on a small set of major crops. Orphan crops present a rich source of genetic diversity for highly nutritional foods, which are often limited in production to small areas of the world (Naylor et al., 2004; Talabi et al., 2022). Tef is an ancient cereal grain primarily grown in the Horn of Africa. Its exact period of domestication remains unclear but is believed to have occurred several thousand years ago, with estimates ranging from ~2000 to 8000 years ago (D'Andrea, 2008; Bultosa, 2016; Vavilov, 1951). Ethiopia is the center for its domestication, a hotspot for tef biodiversity, and the primary location of its production. Tef is Ethiopia's major food crop, grown on over three million hectares of land by six million farmers (Tadele and Hibistu, 2021).

Tef is a self-pollinating allotetraploid C_4 cereal (Assefa et al., 2015; Cheng et al., 2017) in the Chloridoideae subfamily of the Poaceae (grass family) (Cannarozzi et al., 2014; VanBuren et al., 2020). Unlike wheat, barley, and rice, tef grows efficiently in hot and arid climates and has natural resistance to many biotic and abiotic stresses (Bekele-Alemu and Ligaba-Osena, 2023; Girija et al., 2022). Tef is the only *Eragrostis* species that is actively cultivated of the 350 species in the *Eragrostis* genus (Bultosa and Taylor, 2004).

The evolution of tef from its wild progenitors has been under speculation for many years. Based on morphological, cytological, biochemical, and genomic data, it is believed that *Eragrostis pilosa* (L.) P. Beauv. was an original progenitor of tef (Jones et al., 1978; Bekele and Lester, 1981; Ingram and Doyle, 2003). Additionally, *E. pilosa* is the only species with which tef successfully crosses among the *Eragrostis* species and is known to have several useful traits including wider environmental adaptation, lodging tolerance, and early maturity (Ayele et al., 1999; Tefera et al., 2003; Ingram and Doyle, 2003). Their ancestral relationship has been reasserted by genotyping by sequencing (Kebede et al., 2018). This makes *E. pilosa* an interesting species for the introgression of agronomically desirable traits into tef (Talabi et al., 2022). Reintroducing traits from wild progenitors back into domesticated species has been shown to improve desirable traits in maize, lentils, and wheat (Zhang et al., 2023; Rajpal et al., 2023; Keilwagen et al., 2022).

Tef is often touted for its nutritional superiority compared to other more popular cereals such as wheat, rice, and maize. It is gluten-free, becoming an important source of food for people with gluten intolerances and Celiac disease (Spaenij-Dekking et al., 2005). Tef has a low glycemic index, which can be valuable for those with diabetes (Habte et al., 2022). Tef grains also contain 9.4–13.3% protein with an excellent balance of essential amino acids present (Bultosa and Taylor, 2004). Additionally, studies have shown tef to have antioxidant activity *in vitro* (Kotásková et al., 2016; Shumoy et al., 2017) and through mammalian cell-based studies (Cotter et al., 2023), adding to the list of characteristics that make tef a nutritionally important crop. Tef grains have been shown to have higher mineral content than other cereals crops. More specifically, tef has a greater iron concentration than other cereals including wheat, barley, rice, and sorghum (Mohammed et al., 2009; Gebru et al., 2020). Dietary iron is integral as a micronutrient for preventing and treating iron-deficiency anemia. Ligaba-Osena et al. (2021) showed that tef grains have more bioavailable iron in cell-based assays versus rice and wheat. This same study showed that tef outranked other sampled cereals in essential minerals including Fe, Ca, S, K, Mg, P, Mn, and Zn concentrations. Others have shown similar findings when analyzing the mineral content in tef grains (Nyachoti et al., 2021; Habte et al., 2022), highlighting the value of the grains for dietary supplementation of key macro- and micronutrients. Moreover, using elite cultivars grown in Ethiopia, Ereful et al. (2022) reported exceptionally prominent levels of the key micronutrients Fe and Zn and began to establish the genetic basis associated with these traits. Therefore, it is important to keep the nutritional quality of the grain into consideration for those interested in developing elite tef varieties.

Despite the nutritional advantages of tef consumption, its global consumption has been thwarted due to challenges associated with its cultivation and low yield. Attempts to breed high yielding varieties of tef through modern techniques have been hindered due to the paucity of research on tef globally. Moreover, molecular breeding techniques are less developed in tef due to a lack of

transformation and regeneration methods. Recently, the use of morphogenic regulator genes has shown promise in tef trait improvement (Beyene et al., 2022), yet this technology remains to be optimized for wide application. There remain many limitations to tef cultivation. For example, tef is prone to shattering and lodging, which directly diminishes the yield. Lodging occurs when the stalk prematurely breaks or bends and is estimated to decrease the yield by 30-35%. This has prompted molecular breeding research to develop lodging-resistant varieties (Ben-Zeev et al., 2020). Jöst et al. (2015) identified semi-dwarf varieties of tef via a mutation in the α – *tubulin 1* gene, which improved tolerance to lodging. Beyene et al. (2022) have recently developed lodging tolerant semi-dwarf tef lines via CRISPR/Cas9 genome editing. Ligaba-Osena et al. (2020) previously found that supplementation with silicon (Si) improved overall plant performance and grain yield. Tef research has been lagging due to a lack of awareness from the scientific community and funding agencies. However, there is an increasing interest in tef research, due in part to its increase in global popularity for its nutritious and gluten-free grains, and the quality of its straw for animal feed (Miller, 2009; Davison et al., 2011; Anderson and Volesky, 2012).

Furthermore, tef grains are the smallest of all cereals, estimated to be only about 1/150th the size of a wheat grain (Bultosa and Taylor, 2004). It is believed that tef's seed size directly limits the harvest yield. Additionally, the small seed size presents challenges during seed sowing in the field, which can lead to poor population control and uneven distribution (Mengie et al., 2021). Overcrowding of crops leads to competition for light, water, and nutrients. This has been shown to diminish biomass and grain yield in tef (Mengie et al., 2021). Due to their small size, tef seeds are easily lost during harvest, which results in diminished yield. Unlike cereals with larger grains, tef cannot be mechanically harvested using standard equipment. Tef farmers must either harvest the grains by hand, or build/purchase specialty equipment, which is often expensive and adds an additional layer of difficulty to tef cultivation. Traditional breeding methods have been unsuccessful for breeding larger grain size in tef, which has necessitated the improvement of seed size via molecular breeding.

Seed development is an extraordinarily complex and multi-pathway process which is under the control of many transcription factors and hormones (Su et al., 2021; Alam et al., 2022). Thus, developing methodologies for the manipulation of seed development processes is difficult and complex. In many cereals, such as rice, large seed size is an economically important attribute which is used as a measure for yield and quality (Li et al., 2022). For example, large seed size has been associated with improved yield and germination in durum wheat (Akinci et al., 2008). Understanding the genetic mechanisms regulating seed size is a major area of research, in the hopes of developing higher yield varieties of crops. However, a comprehensive understanding of the mechanisms regulating seed size is not totally understood. Abiotic factors likely also affect seed development, leading to alterations in seed size and weight (Ma et al., 2023). The advancement of high-throughput sequencing technology has paved the way for robust genomic analysis. Such analyses interested in seed size regulation have shown potential for identifying genes involved in seed size regulation in *Arabidopsis*, rice, wheat, maize, and soybean (Alam et al., 2022).

In rice, seed size is almost entirely determined by the size of the hull (Alam et al., 2022). However, final seed size is regulated by many different regulatory pathways. The current models of seed size development suggest six essential pathways that regulate seed size: 1) ubiquitin-proteasome signaling, 2) mitogen activated protein kinase (MAPK) signaling, 3) haiku (IKU) signaling 4) guanine nucleotide-binding proteins (G-protein) signaling, 5) transcription factors, and 6) phytohormone signaling (Alam et al., 2022; Li et al., 2022; Li and Li, 2016). Broadly, each of these pathways have been shown to influence final grain length, width, and mass. Simplifying further, there are two overarching phases in seed development. First, seed morphogenesis is marked by cell proliferation, embryo development, endosperm development, and formation of the cotyledon. Second, the seed undergoes maturation of the embryo and enters into a dehydrated dormant state (Alam et al., 2022; Badoni et al., 2023; Mohapatra and Sahu, 2022). Seed development at the genetic level is extraordinarily complex, but better understood only in a handful of species.

Here we report the seed size phenotypes of 189 genetically unique tef and 11 *E. pilosa* accessions. Even among the smallest grains in the world, we have identified significant phenotypic variation in seed size attributes among most of the tef population. Additionally, we have begun to establish the possible genetic mechanisms regulating seed size in tef via comparative genomics, using sequences from rice (*Oryza sativa*) to generate putative orthologous coding sequences from the tef genome. Overall, we explored the natural variation in seed size in tef, as well as seed mineral concentration and localization in the grains, and attempted to establish an understanding of the genetic mechanisms potentially regulating seed size in tef. To our knowledge, this is the first study to quantify natural variation in seed size, map and quantify mineral distribution, and identify putative seed size regulating genes in an *Eragrostis* species.

2 Methodology

2.1 Seed images, measurements, and analysis

Bulked seeds of the tef diversity panel grown under field conditions were obtained from the USDA-ARS National Plant Germplasm System, Plant Germplasm Introduction and Testing Research Unit (Pullman, Washington, USA). The seeds were placed on a 76mm x 25mm concavity slide and imaged using a KEYENCE BZ-X710 All-in-one Fluorescence Microscope (Keyence Corporation of America, Ithica, Il) using the setting in Supplementary Table S1. ImageJ software v1.54g (Schneider et al., 2012) was used to determine seed size attributes including seed length, seed width, and seed area of 189 tef and 11 *E. pilosa* accessions. The length measurements were taken from the longest portion of the seed and the width measurements were taken from the widest portion of the seed, while the seed area corresponds to the region within the perimeter of the image (Figure 1). For each of the 189 tef and 11 *E. pilosa* accessions, ten seeds (replicates) were measured for length, width, and area. Each seed image included a 1000 µm scale bar. Ten replicate measurements of the scale bar were taken and averaged to generate a calibration factor from pixels to µm (seed length and width; Equations 1 and 2) or pixels^2 to

FIGURE 1
An example of how seed size attributes were measured in ImageJ. Images from left to right show measurement of seed length, seed width, and seed area. These measurements were converted to μm for seed length and seed width, and μm^2 for seed area.

μm^2 (area; Equation 3); the calibration factor is assigned $pixels_c$ in Equations 1–3. According to the user guide (https://imagej.net/imaging/spatial-calibration), ImageJ takes area measurements as a count of $pixels^2$ and straight-line measurements in pixels. We maintained the same microscope settings to acquire all the images. We found that after converting the seed measurement to a standardized unit (μm or μm^2), the initial parameters were not as important, as long as the calibration was properly set.

$$Width\ (\mu m) = pixels_w \times \frac{1000\,\mu m}{pixels_c} \quad (1)$$

$$Length\ (\mu m) = pixels_l \times \frac{1000\,\mu m}{pixels_c} \quad (2)$$

$$Area\ (\mu m^2) = pixels_a^2 \times \left(\frac{1000\,\mu m}{pixels_c}\right)^2 \quad (3)$$

Additionally, we also determined 1000-grain weight of representative tef accessions selected based on seed area analysis, and those accessions that are considered reference cultivars in tef research. For each accession, one thousand seeds were counted manually, and the mass was measured. This was conducted in quadruplicate and then averaged. Some common tef accessions used in this study include Dabi, a brown cultivar with medium-sized seeds and a reference cultivar; Magna, an ivory (white) variety with high consumer preference and relatively large seeds; and Dessie, another common, brown-colored cultivar widely grown in the U.S. Dabi, Magna, and Dessie have retained their common name, whereas other accessions are labeled with their Plant Introduction (PI) number (Supplementary File 1).

2.2 Plant growth conditions tef seed mineral analysis

Seeds of selected *E. tef* accessions varying in seed size (small, medium, and large) along with the reference cultivar (Dabi) and the cultivar commonly grown in U.S. (Dessie), were grown under greenhouse conditions with supplemental light at constant temperature (28°C). Twenty seeds of each accession was planted in 2 L pots containing Sun Gro professional Mix (Sun Gro Horticulture, Agawam, MA). Plants were supplied with Osmocote® Smart-release® controlled-release fertilizer (Scotts Company, LLC, Marysville, OH) following the manufacturer's recommended application rates. Seedlings were thinned down to five. Plants were supplemented with 1 L of 4.74 g/L Miracle-grow fertilizer (Scotts Miracle-Gro Co., Marysville, OH) solution (24–8–16) per flat containing two pots, and 1 L of 250 mg/L M.O.S.T. soluble trace elements (JR Peters, Allentown, PA) per flat every four weeks. Selected *E. pilosa* accessions were grown under the same conditions. All accessions were replicated four times. Seeds of plants grown in each pot were bulked at harvest, dried and processed for mineral analysis.

2.3 Seed mineral analysis

A total of 500 mg of ground seed was digested using concentrated HNO_3 for 30 min in a microwave at 200°C, and mineral content was analyzed using Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES).

2.4 Synchrotron x-ray fluorescence imaging

The procedures for synchrotron x-ray fluorescence (SXRF) imaging were described previously with slight modifications (Chia et al., 2023). Briefly, the seeds were placed between a layer of metal-free Kapton film and Kapton tape before being mounted onto 35-mm slide frames. The spatial distribution of elements was imaged via SXRF microscopy at the Functional Materials Beamline (FMB) of the Cornell High Energy Synchrotron Source (CHESS). FMB employs an undulator source and an energy of 9.7 keV was selected using a side-

bounce monochromator (Smieska et al., 2023; Stoupin et al., 2021). The beam was focused to 3 μm tall x 15 μm wide containing approximately 3e10 ph/s using a set of compound refractive lenses (RXoptics, Monschau, Germany). SXRF signal was collected in flyscan mode along the vertical axis with a pixel size of 20 μm and a dwell time of 80 ms, using a Vortex ME4 detector (Hitachi, Japan) and an Xspress3 signal processor (Quantum Detectors, UK). SXRF peak areas were fit in Praxes software (https://github.com/praxes/praxes) which is based on PyMCA (http://dx.doi.org/10.1016/j.sab.2006.12.002). Thin foil calibration standards (Micromatter, Canada) were used to obtain a flux to concentration conversion and provide area densities in units of $\mu g/cm^2$.

2.5 Comparative genomics

A list of seed development genes in rice was compiled in a review by Li et al. (2022) and Li and Li (2016). We used these lists as a reference to identify genes in rice that are specifically implicated in controlling grain size. Rice gene nucleotide coding sequences (CDS) were generated by searching the Gene ID number in NCBI (National Center for Biotechnology Information) database (https://www.ncbi.nlm.nih.gov/); Search criteria: All Databases). Using the rice CDS sequences as query, we searched the most similar CDS from wheat (*Triticum aestivum*; taxid: 4565), maize (*Zea mays*; taxid: 4577), barley (*Hordeum vulgare* L.; taxid: 4513), sorghum (*Sorghum bicolor* (L.) Moench; taxid: 4558) and *Arabidopsis* (*Arabidopsis thaliana* (L.) Heynh.; taxid: 3702). We also included rice (*Oryza sativa* (japonica cultivar group); taxid: 39947) in this search as a control comparison to ensure the correctness of the query sequence. Then, the tef CDSs were gathered using CoGeBlast, using the rice gene as the reference sequence (https://genomevolution.org/coge/CoGeBlast.pl). Rice CDS nucleotide sequences were BLAST against the tef genome (Selected Genome: *Eragrostis tef* (tef) id 50954 PacBio unmasked vV3; Maker, PacBio: (id 50954) vV3 unmasked 577,738,711nt). Rice sequences were used as the reference in all but two instances, where the *Sorghum bicolor* sequence was used to derive the *RGG1* and *AGO17* orthologs from tef. To simplify the phylogenetic analysis, we selected a single allele with the highest predicted similarity. However, since tef contains two sub-genomes, we have included both alleles of each putative seed size regulating gene in Supplementary File 2. The BLAST parameters were set at an E-Value cutoff of 1e-30, with other parameters left standard. This cutoff was set to ensure the tef sequences we acquired could be classified as putative homologs or orthologs using the standard cutoff for determining homology using BLAST ($E< 1e\text{-}5$) (Choudhuri, 2014), or the cutoff others have indicated for added level of scrutiny for nucleotide comparisons ($E< 1e\text{-}10$) (Pearson, 2013).

Gene phylogenies were grouped by their common regulatory pathway or functional category. Sequences were aligned using MUSCLE and the gene phylogenies constructed using a Maximum Likelihood tree building algorithm in MEGA v11.0.13 (Tamura et al., 2021). The trees were converted to Newick file format and uploaded to Interactive Tree of Life (ITOL) for final formatting and annotation (https://itol.embl.de/) (Letunic and Bork, 2024). The Newick files are included as Supplementary Data.

Functional analysis of rice genes was conducted using ShinyGO v0.80 (http://bioinformatics.sdstate.edu/go/), which was used to identify key biological processes involved in seed size regulation in rice (Ge et al., 2020). Our list consisted of reported seed size regulating genes that have been cloned and characterized in rice, compared against the rice genome background (Oryza sativa Japonica Group genes IRGSP-1.0; TaxID: 39947). Genes were clustered by GO Biological Processes (Supplementary File 3), and extracted using FDR (false discovery rate; $FDR< 0.05$) and enrichment score. ShinyGO defines Fold Enrichment as the percentage of genes in the list belonging to a pathway, divided by the corresponding percentage in the background (Ge et al., 2020; http://bioinformatics.sdstate.edu/go74/#:~:text=Fold%20Enrichment%20is%20defined%20as,a%20certain%20pathway%20is%20overrepresented). Quoted from the ShinyGO v.80 website, "FDR refers to how likely the enrichment is by chance; Fold Enrichment indicates how drastically genes of a certain pathway is overrepresented." The purpose of this analysis was to highlight the pathways which are important to the regulation of seed size in rice. For those focused on tef breeding, genes of those pathways may be a key starting place for gene manipulation in tef, bearing the mechanisms regulating seed size in both cereals are similar.

2.6 Data analysis

Analysis of seed size attributes and ICP-OES results was performed in SAS Enterprise Guide v8.3.8.206 (SAS Institute Inc., Cary, NC USA). We conducted one-way analysis of variance (ANOVA) to determine if there was significant difference in mean seed width, seed length, and seed area among the 189 tef and 11 *E. pilosa* accessions. For the tef accessions, we binned accessions from each seed size attribute as small, medium, or large, and selected 10 accessions from each group for simplicity. These represent the 10 largest, medium, and smallest accessions for seed area, length, or width. For seed length, seed width, and seed area of the 30 representative accessions, we performed another one-way ANOVA, followed by a Student-Newman-Keuls (SNK) *post-hoc* test for means comparison. Lastly, ICP-OES results were averaged and analyzed by one-way ANOVA, followed by a SNK *post-hoc* test. Figures were generated using GraphPad Prism v10.0.3 (GraphPad software, Boston, Massachusetts, USA).

3 Results

3.1 Natural variation in seed size of *E. tef* and *E. pilosa* accessions

In this study, we analyzed the natural variation in seed size of 189 tef and 11 *E. pilosa* genotypes. Seed images were acquired by microscopy and seeds size attributes, including seed length, width, and area was measured. Additionally, 1000-grain weight was generated for selected tef and *E. pilosa* accessions.

The distribution of seed areas of sampled tef accessions is normal ($\bar{x}$ = 525,565 μm^2; sd = 49,102 μm^2) but slightly left skewed (Supplementary Table S2; Figure 2A). The center half (interquartile range, IQR) of the tef accessions fall within a range of 498,457 μm^2 and 555,9739 μm^2, with the largest and smallest accessions being nearly 25% larger and 27% smaller than the mean, respectively. One-way ANOVA revealed a statistically significant difference ($p< 0.0001$) between the 189 tef varieties in mean seed area. Seed areas of accessions PI494370 and PI494453 were identified as the largest and smallest, respectively (Figure 3A). Seeds of PI494370 are about 51% larger than the smallest accession PI494453. To better represent our data, we selected ten accessions with the largest, medium, and smallest areas (30 in total). One-way ANOVA revealed that the difference in seed area among the 30 selected accessions is statistically significant (Figure 3A). A *post-hoc* test indicated that the ten largest accessions were significantly larger than the ten smallest. However, no significant difference was found between the medium ten accessions and six of the larger and one of the smaller tef accessions.

The distribution of tef accessions by seed length is normal ($\bar{x}$ = 1142 μm; sd = 65.75 μm), and slightly left skewed, with the IQR of the accessions falling between 1095μm and 1185μm (Figure 2B). The 189 tef accessions were significantly different ($p< 0.0001$) in seed length. The longest accession was identified as PI494370, which also has the largest seed area. The seed length of PI494370 is 31% larger than PI193511, the shortest accession (Figure 3B), and were found to be significantly different in the 30-accession comparison ($p< 0.05$; Figure 3B).

Like seed area and length, the distribution of accessions by seed width is normal ($\bar{x}$ = 582.4 μm; sd = 33.98 μm) and slightly left skewed. The IQR falls within 559.1 μm and 605.7 μm. One-way ANOVA showed a statistically significant difference in seed width for the 189 tef accessions ($p< 0.0001$; Figure 2C). The seed width of accessions PI524446 and PI494307 were the widest and most

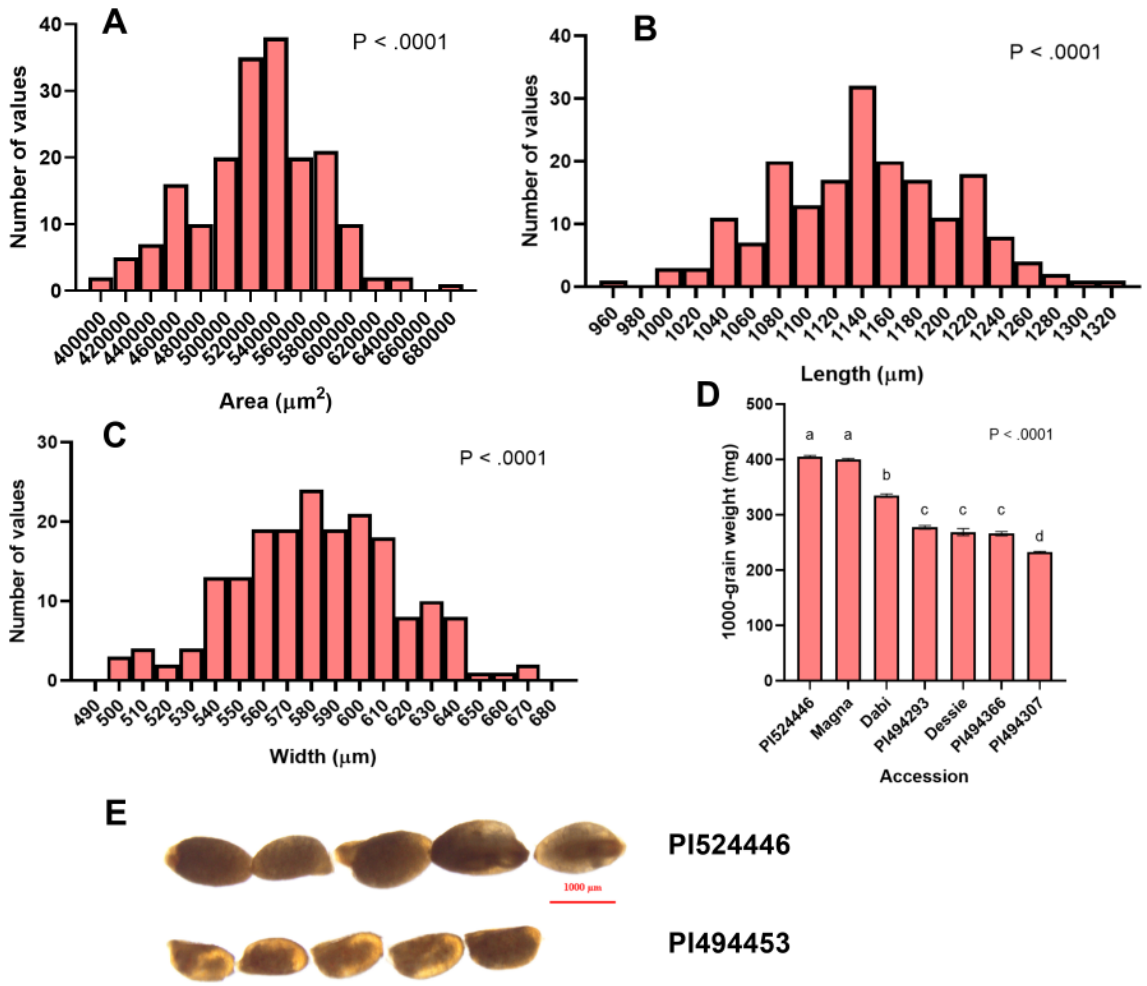

FIGURE 2
Measurements of tef seed area, length, width, and 1000-grain weight. **(A)** Frequency distribution of tef mean seed area (μm^2), **(B)** Seed length (μm), **(C)** Seed width (μm), and **(D)** 1000-grain weight for representative tef varieties. Ten replicate measurements (n = 10) of seed length, seed width, and seed area were taken and averaged for each accession. Four replicates (n = 4) were weighed and averaged to generate the 1000-grain weight. Data was analyzed by one-way ANOVA and mean comparison using the Student-Newman-Keuls (SNK) *post-hoc* test. One-way ANOVA found that differences between accessions for all seed size attributes was statistically significant ($p< 0.0001$). Bars bearing the same letter are not statistically significant ($p< 0.05$). **(E)** Visual comparison of accessions with large (PI524446) and small (PI494453) seeds. Image was acquired under 1x magnification using a Nikon SMZ18 Stereoscope (Nikon, Kanagawa, Japan).

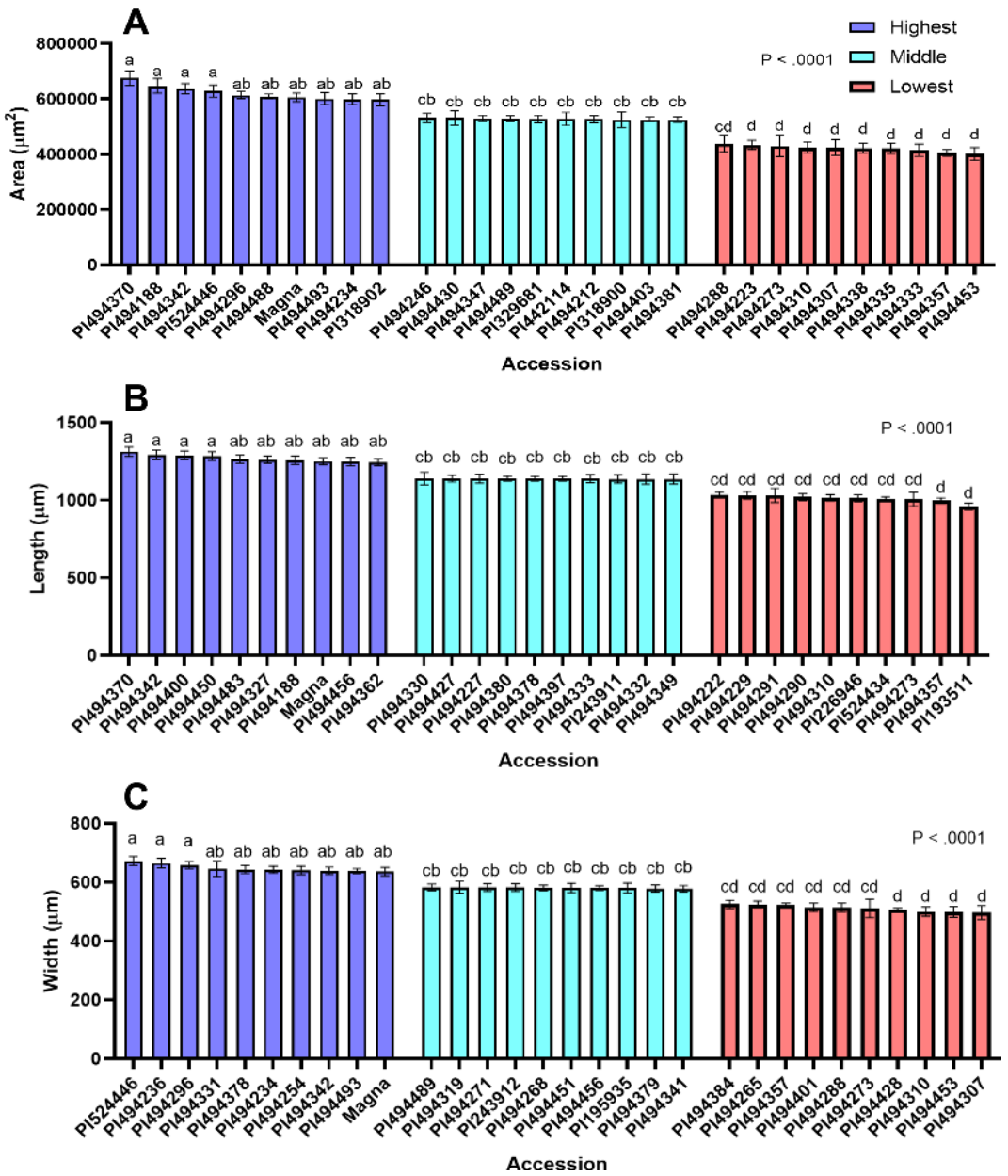

FIGURE 3
Seed size attributes of 30 representative tef accessions. Accessions were binned as high, medium, or low for seed area, length, and width. The data of 10 accessions representing the highest, medium, and lowest categories (30 total) are presented for mean seed area **(A)**, length **(B)** and width **(C)**. Ten replicate measurements (n = 10) of seed length, seed width, and seed area were taken and averaged for each accession. The 30 accessions were analyzed by one-way ANOVA. One-way ANOVA found that differences between accessions for all seed size attributes was statistically significant (*p<0.0001*). Statistical mean comparison used the SNK *post-hoc* test. Bars bearing the same letter are not significantly different (*p< 0.05*).

narrow, respectively, among the 189 tef accessions (Figure 3C). Seed of PI524446 are 30% wider than that of PI494307. One-way ANOVA of the 30-accession comparison of seed width revealed significant differences among the accessions. A *post-hoc* test indicated significant differences between the ten widest and ten most narrow accessions (Figure 3C). No significant difference was found between the ten medium width accessions and seven widest and six most narrow accessions. A summary of these findings are in Supplementary Table S2.

To determine the 1000-grain weight we counted 1000 seeds of seven tef accessions representing large, medium, and small seeds, and took the weights of the seeds. PI524446, which also had the widest seeds, had the highest 1000-grain weight, while PI494307 was found to have the lowest 1000-grain weight. The latter also has the smallest seed width. The 1000-grain weight of PI524446 was 1.54 times higher than that of PI494307, whose difference was statistically significant (Figure 2D).

During the course of the seed size measurement and analysis, we identified a unique accession PI442115, which was substantially smaller than the tef accessions in our sample. PI442115 was a total outlier in seed size and color compared to the other tef accessions. PI442115 more closely resembled *E. pilosa* in color and in size than

it did to tef (Figure 4). It was significantly smaller and much darker in color (Figure 5). This accession was mischaracterized in the USDA-ARS Germplasm Resource Information Network (GRIN) database as tef. This information prompted us to expand the study to include all the *E. pilosa* accessions available at the U.S. National Plant Germplasm Center.

We also measured seed length, width, and area for 11 *E. pilosa* accessions. Our analysis revealed significant variation among the seed size attributes of *E. pilosa*. Seed area of accession PI442487 (350,174 μm^2) was 1.79 times that of PI442115 (151,744μm^2). PI378082 has the longest seeds, and PI442115 has the shortest seeds among the *E. pilosa* accessions. The widest *E. pilosa* accession, PI263510, was 42% wider than PI442115. The 1000-grain weight of PI378082 is twice that of PI442115. Seed size attributes (area, length, width, and 1000-grain weight) of PI442115 are the smallest of all the tef and *E. pilosa* accessions we analyzed. A summary of these findings is in Supplementary Table S3. The analysis of *E. pilosa* grain length, width, area, and 1000-grain weight revealed significant variation in seed size among all of the available genotypes. We suspect there is a genetic basis behind this

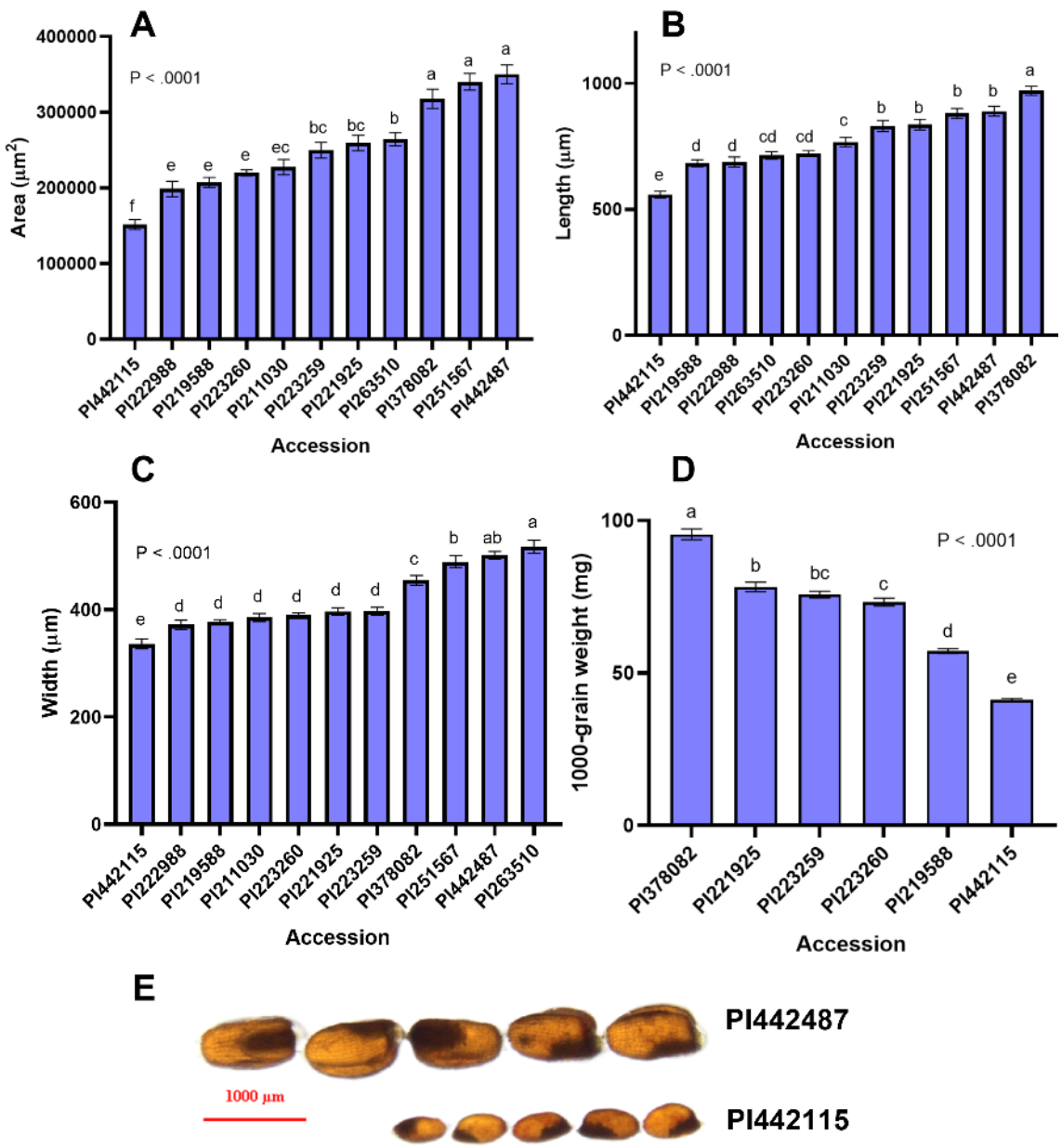

FIGURE 4
Seed size attributes of *E. pilosa* **(A)** Seed area (μm^2), **(B)** Seed length (µm) **(C)** Seed width (µm), and **(D)** 1000-grain weight for representative *E. pilosa* varieties. Ten replicate measurements (n = 10) of seed length, width, and area were taken and averaged for each accession. Four replicates (n = 4) were taken and averaged to generate the 1000-grain weight. Data were analyzed by one-way ANOVA. One-way ANOVA found that differences between accessions for all seed size attributes was statistically significant ($p<0.0001$). The SNK *post-hoc* test was used for means comparison. Bars bearing different letters are significantly different ($p< 0.05$). **(E)** Visual comparison of accessions with the largest (PI442487) and smallest (PI442115) seeds by area. Seeds images were acquired under 1x magnification, taken on a Nikon SMZ18 Stereoscope (Nikon, Kanagawa, Japan).

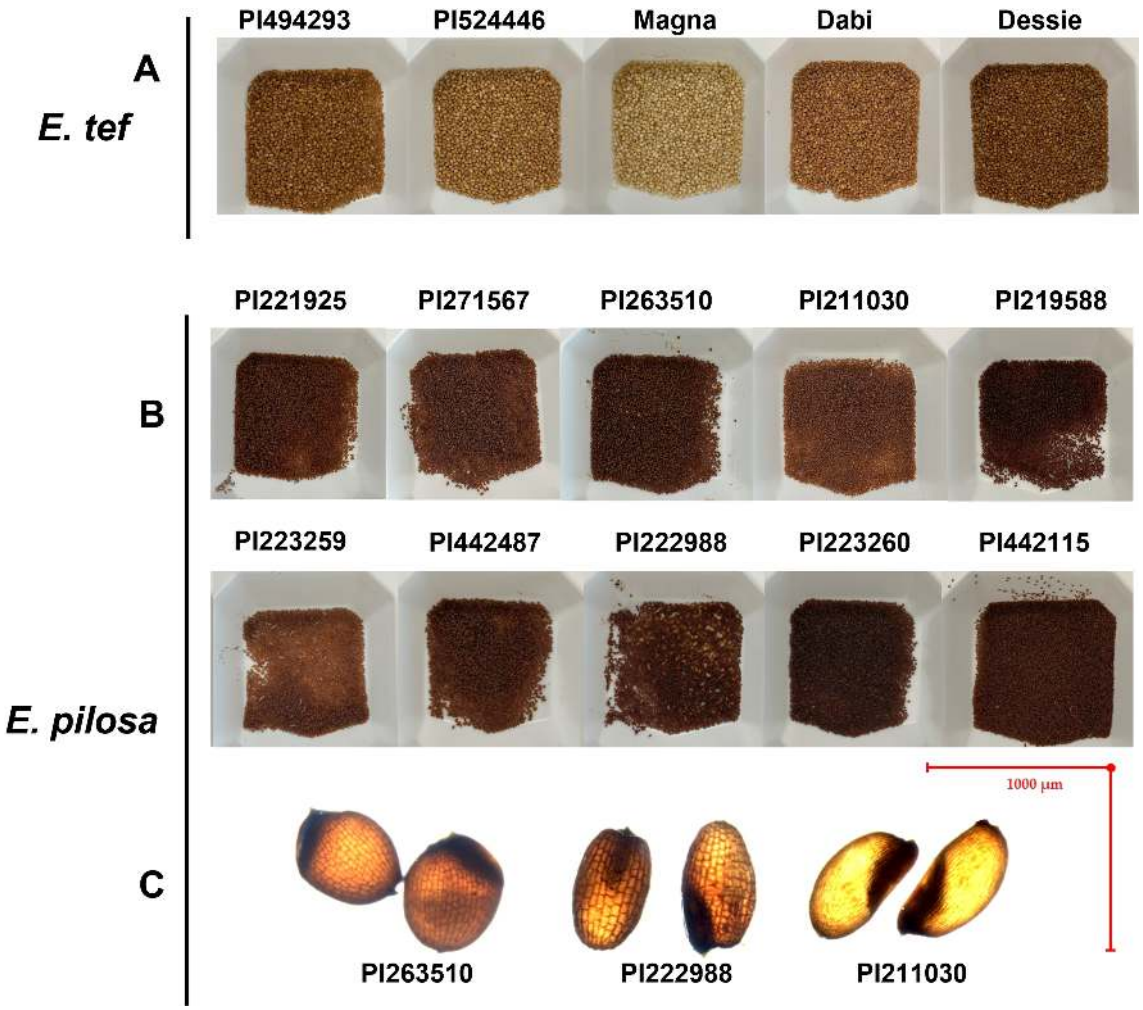

FIGURE 5
Comparison of selected **(A)** tef and **(B)** *E. pilosa* genotypes. **(C)** Images highlight phenotypes observed from *E. pilosa*, including the reticulate-patterned seed coat observed in a few accessions (e.g., PI222988) compared to the smooth seed coat in others (e.g., PI211030). A diminished length-to-width ratio was observed in PI263510. We observed differences in seed shape in some *E. pilosa* accessions, including more oblong seeds with emarginated apices (e.g., PI222988) in comparison to the more ovular seed shape of PI211030.

natural variation, but this remains to be validated. Furthermore, there is the potential for identifying alleles from tef's progenitor species that can shed light onto the evolution of tef domestication or can be utilized as a source of genes for introgression of agronomically desirable traits into tef. In rice, a handful of alleles regulating seed size have been identified in a similar way, using wild rice species such as *Oryza rufipogon*, the *O. sativa* progenitor, and *O. barthii*, the progenitor of African rice *O. glaberrima*, to identify alleles that were lost or selected for during domestication (Jiang et al., 2019; Luo et al., 2013; Wu et al., 2017).

Additionally, we identified morphological differences in both the tef and *E. pilosa* seeds. Most *E. pilosa* grains appear distinctly darker than tef grains. In comparison, tef grains have a lighter colored seed coat, with a smooth outer surface. Some of the *E. pilosa* grains showed clear darkened reticulate patterned seed coat (For an example: Figure 5C; PI222988). This was also reported by Kreichstitz et al. (2009). None of the tef accessions we sampled expressed this type of patterning. Interestingly, not all of the *E. pilosa* accessions showed reticulate seed coat patterning but retained the darker seed coat coloration. Specifically, PI211030, PI221925, and PI223259 have fingerprint-like patterning and more closely resemble tef grains in this aspect (For an example: Figure 5C; PI211030). Additionally, the non-reticulate *E. pilosa* accessions are more ovular (Figure 5C; PI211030), whereas the reticulate patterned seeds are oblong, with more pronounced emarginated apices (Figure 5C; PI222988). The seeds from PI263510 (*E. pilosa*) appeared to be more spherical, with a highly diminished length-to- width ratio (Figure 5C) (data not shown).

3.2 Mineral analysis

To understand the association between seed size and mineral content, we analyzed the mineral concentration of tef and *E. pilosa* accessions contrasting in seed size, along with a few reference varieties of tef using ICP-OES. As shown in Figure 6, the

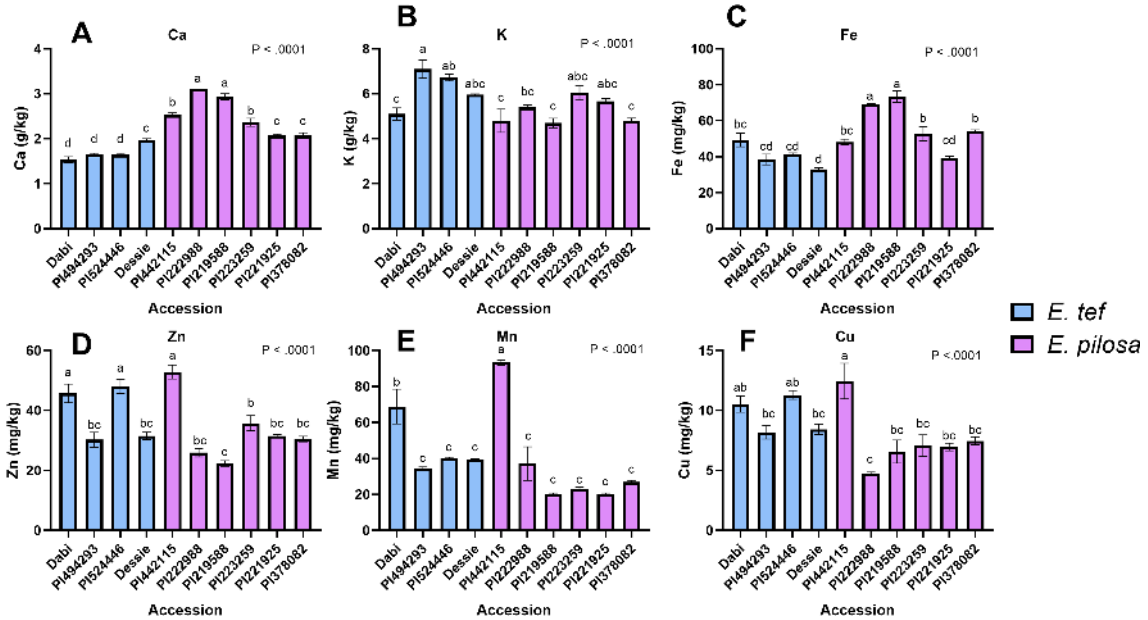

FIGURE 6
Elemental analysis using ICP-OES. **(A)** Ca, **(B)** K, **(C)** Fe, **(D)** Zn, **(E)** Mn, and **(F)** Cu. Blue bars indicate tef accessions and purple bars indicate *E. pilosa* accessions. Data were analyzed by one-way ANOVA. One-way ANOVA found that differences in mineral concentration for Ca, K, Fe, Cu, Mn, and Zn among the tef and *E. pilosa* accessions were statistically significant ($p< 0.0001$).The SNK *post-hoc* test was used for mean comparison. Bars bearing different letters are significantly different ($p< 0.05$).

differences in all mineral concentrations were statistically significant ($p< 0.0001$). The Ca concentrations in most *E. pilosa* accessions were higher than the tef accessions. The highest Ca concentration (3 g/kg) was detected for PI222988 and PI219588 (Figure 6A). The lowest Ca level (1.5 g/kg) was detected in the reference tef cultivar Dabi, along with the tef accessions with small seeds (PI494293), and large seeds (PI524446). The K concentration varied from 6.5-7.5g/kg for all accessions, and the highest concentration was detected for tef accession PI494293 (Figure 6B). The Fe concentration of two *E. pilosa* accessions PI219588 (73.47 mg/kg) and PI222988 (69.03 mg/kg) was significantly higher than tef and other *E. pilosa* accessions (Figure 6C). Seed Mn concentration was highest for PI442115 (93.37 mg/kg) followed by Dabi (68.91 mg/kg) (Figure 6E) while the concentration of Zn and Cu was higher for PI442115, P524446, and Dabi compared to the other accessions (Figures 6E, F). Our findings show that accession PI442115 tends to accumulate more micronutrients (Mn, Zn and Cu) than the other *E. pilosa* accession, but only some of the tef accessions. On the other hand, there was no marked difference in P and Mg concentration among the accessions (Supplementary Figure 1) while the concentration of S and B was higher in PI442115 (Supplementary Figure 1).

Next, using the ICP-OES data discussed above, we analyzed the relationship between seed size and mineral concentration. The seed area was plotted against mineral concentration, and a simple linear regression analysis was conducted to determine if there was a correlation between seed size and mineral content in tef or *E. pilosa*. We found no significant relationship between tef seed area and P ($p = 0.131; R = 0.87$), K ($p = 0.57; R = 0.43$); Ca ($p = 0.84; R = 0.17$), and Mg ($p = 0.64; R = 0.36$) concentration (Supplementary File 4; Supplementary Figure 2). We detected a significant positive correlation between tef seed area and S concentration ($p< 0.01; R = 0.99$) (Supplementary Figure 2). Among the micronutrients, we detected no significant relationship between tef seed area and Fe ($p = 0.59; R = 0.41$), Mn ($p = 0.57; R = 0.43$), Zn ($p = 0.12; R = 0.88$), Cu ($p = 0.10; R = 0.89$), or B ($p = 0.32; R = 0.68$) concentration. We found no significant association between *E. pilosa* seed area and P ($p = 0.25; R = 0.56$), K ($p = 0.71; R = 0.20$), Ca ($p = 0.14; R = 0.67$), and S ($p = 0.052; R = 0.89$) concentration. We detected a significant but weak negative correlation between *E. pilosa* seed area and Mg concentration ($p< 0.05; R = 0.89$). Lastly, we found no significant relationship between *E. pilosa* seed area and Fe ($p = 0.65; R = R = 0.24$), Mn ($p = 0.126; R = 0.69$), Zn ($p = 0.40; R = 0.42$), and Cu ($p = 0.42; R = 0.41$) concentration.

3.3 SXRF imaging

Inductively coupled plasma optical emission spectrometry (ICP-OES) is a useful strategy for quantifying minerals in different plant tissues including seeds. However, this method does not provide information on the spatial distribution of mineral elements in those tissues. Therefore, we used synchrotron-based X-ray fluorescence (SXRF) microscopy to visualize minerals in the mature seeds. This method provides semiquantitative elemental distribution at high sensitivity and high spatial resolution (Donner et al., 2012, 2013). Here we report the results of SXRF imaging of the reference tef variety Dabi, small (PI494293) and large (PI524446) seed tef accession, and the *E. pilosa* accession with the smallest seed (PI442115).

SXRF imaging detected calcium (Ca) apparently localized to the seed coat and was detected in the embryo (Figure 7A). It is barely detected in the endosperm. The Ca signal is lower in the reference

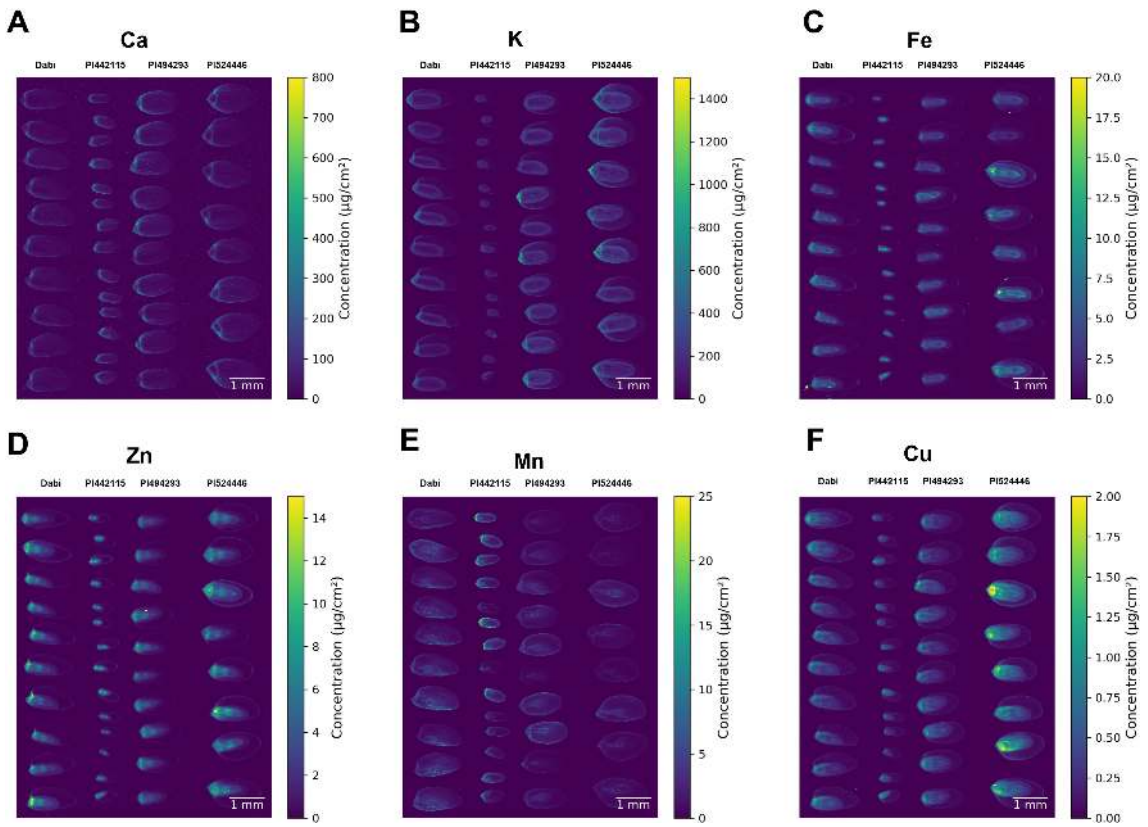

FIGURE 7
SXRF images of tef and *E. pilosa* seeds. Images show elemental localization of **(A)** Ca, **(B)** K, **(C)** Fe, **(D)** Zn, **(E)** Mn, and **(F)** Cu. Seeds of the same variety are aligned as a column. Accessions left-to-right: (1) Dabi; tef, (2) PI442115; *E. pilosa* (3) PI494293; tef (4) PI524446; tef.

cultivar Dabi, while there is no marked difference between PI494293, PI524446, and PI442115. When Ca concentration was quantified, the median concentration of Ca was lower in Dabi and slightly elevated in PI442115, with no clear difference between PI494293 and PI524446 (Supplementary File 5). Potassium (K) was detected in the outer embryo and the seed coat while it appears absent from the endosperm (Figure 7B). For K, we also observed a substantial difference among the accessions. The K signal was stronger in accession PI494293 and PI524446 while it was weakest in the small-seeded *E. pilosa* accession PI442115. Quantification of K concentration showed that PI494493 and PI524446 exhibit elevated median levels of K, in comparison to Dabi and PI442115 (Supplementary File 5). Furthermore, there is a greater proportion of signals that exhibit high levels of K in PI494293 and PI524446 in comparison to the other accessions. It should be noted that since the K and Ca SXRF lines have low energy, these signals are measured mainly from the surface of the seed, on the order of 10s of µm. For heavier elements, SXRF signals are measured from greater depths (100s to 1000 µm).

Iron (Fe) is detected in the embryo and the seed coat (Figure 7C). There is no marked difference in the signal intensity of Fe among the accessions, although it is not visible in the seed coat of PI442115, which could be due to its small seed size. Quantification of Fe shows that there is no difference in median levels of Fe among the accessions in our sample (Supplementary File 5). Like Fe, zinc (Zn) is also localized in the embryo, and the seed coat (Figure 7D). The signal of Zn looks slightly higher in PI524446, followed by Dabi, compared to accessions PI494293 and PI442115. Quantification of Zn shows that median levels of Zn are slightly higher in PI524446, but otherwise there is no difference detected in other accessions (Supplementary File 5). The manganese (Mn) signal was higher in PI442115, followed by Dabi, where it is detected in the embryo and the seed coat (Figure 7E). The Mn signal was lower in the other accessions. Quantification of Mn showed that the median levels of Mn are elevated in Dabi and PI442115, but that a larger proportion of detected signals exhibit high levels of Mn in PI442115 (Supplementary File 5). The signal of copper (Cu) was higher in PI524446 as compared to Dabi and accessions PI494293 and PI442115 (Figure 7F). Quantification of Cu indicated elevated levels of Cu in PI524446, and a greater proportion of signals from PI524446 show higher levels of Cu than the other accessions (Supplementary File 5). Cu is detected in the embryo and the seed coat. Overall, seed size does not appear to affect the general patterns of mineral localization in the seed.

3.4 Comparative genomics and functional annotation of seed size regulating genes in rice

This study established that there is natural intraspecific variation in seed size, which indicates that there is likely a genetic

basis for this trait. It is known that there are a number of factors that influence final grain size, including numerous genetic factors that are under scrutiny (Li et al., 2022). We hypothesized that the mechanisms regulating seed size in both tef and rice are potentially similar, with other studies indicating high conservation of seed size regulating mechanisms among the cereals (Tao et al., 2020a; Long et al., 2024; Zhang et al., 2015). We gathered the sequences of genes reported to influence grain size in rice and identified putative orthologs in tef using the rice genes as a reference. Although C_4 grasses such as the millets and sorghum are more closely related to tef, the genetic mechanisms regulating seed size are less established in these species. The mechanisms regulating these traits are better characterized in rice. As a result, we chose to use rice as the reference.

72 seed size regulating genes in rice were collected, representing numerous regulatory mechanisms and gene families, and identified putative orthologs in tef. Then, we conducted a phylogenetic analysis to determine the similarity of the tef sequences to highly similar sequences from other common cereals. The phylogenetic analysis was limited to four gene families, including G-protein pathway genes, mitogen-activated protein kinase (MAPK) signaling, the ubiquitin-proteasome pathway, and brassinosteroid (BR) signaling and biosynthesis, all of which have been shown to influence grain size in rice. Lastly, we conducted a functional analysis of reported seed size regulating genes in rice to establish which mechanisms were most represented and highly influential in rice seed size regulation. For the analysis we used a list of seed size regulating genes from rice, analyzed against the background rice genome. The genes were clustered by the gene ontology (GO) term Biological Processes. Only genes that have been functionally cloned and characterized from rice were included in the list. For those focused on tef breeding, genes of highly enriched pathways may be a key starting place for gene manipulation in tef, if the mechanisms regulating seed size in both cereals are similar.

The functional analysis of seed size regulating genes suggested that G-protein-coupled receptor signaling (*FDR = 1.9E-10*) was the most significantly enriched biological process, followed by brassinosteroid-mediated signaling (*FDR = 2.4E-18*) (Figures 8, 9). The other highly enriched biological processes involve steroid hormone signaling and response. This indicates that BR signaling, BR biosynthetic regulation, and G-protein mediated signaling are highly important for the regulation of seed size in rice.

Based on the phylogenetic analysis, most genes reported to control grain size in rice show high similarity to the loci we identified in the tef genome. For all trees, only the most similar tef sub-genome feature was displayed for each gene. In the phylogenetic analysis of genes grouped into ubiquitin-mediated pathways, we clustered a total of 57 sequences from seven plant species, including rice, wheat, maize, sorghum, barley, tef, and *Arabidopsis*. These genes clustered into eight groups (LG1, GRX8, TUD1, GW6a, bZIP47, GW2, CLG1, and WTG1) (Figure 10). For genes grouped into the MAPK pathway for the phylogenetic analysis, we gathered a total of 56 sequences from seven plant species. These genes clustered into seven groups (MKKK10, MKKK70, MKP1, MPK6, MAPKK4, MKKK62, MKK62 +MKKK70). For the phylogenetic analysis we display one ortholog identified from tef from each of the rice reference sequences except in the MKKK70 cluster, where BLAST for *MKKK62* and *MKKK70* returned identical sequences in tef and sorghum (Figure 11). From the G-protein mediated signaling genes, we gathered 47 sequences among the seven species and clustered them into nine groups (RGG2, RGG1, DEP1, RGA1, RGB2, GS3 +DEP1+GGC2, GS3, and DEP1+GGC2). Two additional groups formed from BLAST returning identical sequences from *GS3*, *DEP1*, and *GGC2* in *Arabidopsis*, and BLAST returning identical sequences from *DEP1* and *GGC2* in barley. In the G-protein phylogenetic analysis, we report one tef ortholog in seven of the clusters RGG2, RGG1, DEP1, RGA1, RGB2, and GS3 (Figure 12). In BR signaling or biosynthetic-family genes, we collected 154 sequences from seven species, which clustered into 22 groups. An additional group resulted from BLAST of *GSK2* and *GSK3* returning a single *Arabidopsis* sequence. For the purposes of the phylogenetic analysis, we report 21 tef putative orthologs. We identified other putative orthologs from tef which showed significant similarity (*E< 1e-30*) with rice genes from other biological processes, but they are not included in our phylogenetic analysis. We have also included the corresponding homologous sub-genome feature from tef. The complete list of putative orthologous sequences that were identified in tef is available in Supplementary File 2.

For simplicity, the phylogenetic analysis of ubiquitin-proteasome family genes includes a single tef sub-genome feature with the highest degree of similarity. From genes involved in the ubiquitin-proteasome pathway, we identified putative orthologous sequences from tef for the genes encoding a ubiquitin-specific protease (UBP15; LG1), a RING ubiquitin E3 ligase (HRZ2), an otubain-like protease (OTUB1; WTG1), a CC-type glutaredoxin protein (GRX8;WIDE-GRAIN 1; WG1), a histone acetyltransferase (OslIHAT1; GW6a), a U-box E3 ubiquitin ligase (TUD1), an E3 ubiquitin ligase (GW2), and a bZIP transcription factor (bZIP47) (Figure 10). The tef ortholog of *OsUBP15* forms a single cluster with close similarity to the maize and sorghum *UBP15* orthologs, both of which encode UBP15. Our analysis showed a single cluster of genes referenced from *OsGRX8*, with high confidence of similarity between the rice and tef sequences (94.1% bootstrap value). Rice and tef also share some similarities for *TUD1* but to a lesser extent than UBP15. Our analysis indicates that both *TUD1* and *lHAT1* found in rice are conserved among our sampled Poaceae species, with little similarity to the predicted *Arabidopsis* sequences. Tef sequences share the same level of similarity to *OslHAT1* and the predicted barley and wheat orthologs. Similarly, tef shows equal similarity to the rice *bZIP47* and other grass-family *bZIP47* orthologs. *GW2* referenced genes form a single cluster, with the tef ortholog sharing the closest similarity to the maize and sorghum sequences. *CLG1/HRZ2* from rice is highly similar to the wheat sequence and shares a sub-cluster with tef. Oddly, the predicted *CLG1/HRZ1* ortholog in barley is most similar to the sorghum *OTUB1* sequence. The tef ortholog of *OTUB1* shares a high bootstrap value with *OsOTUB1* and other grass-family *OTUB1* predicted orthologs. Our analysis suggests the tef genome contains two copies of each of the aforementioned genes, i.e. two homologous sub-genome features with high degrees of similarity. In tef, we identified an additional, highly similar copy of *GW6a* on

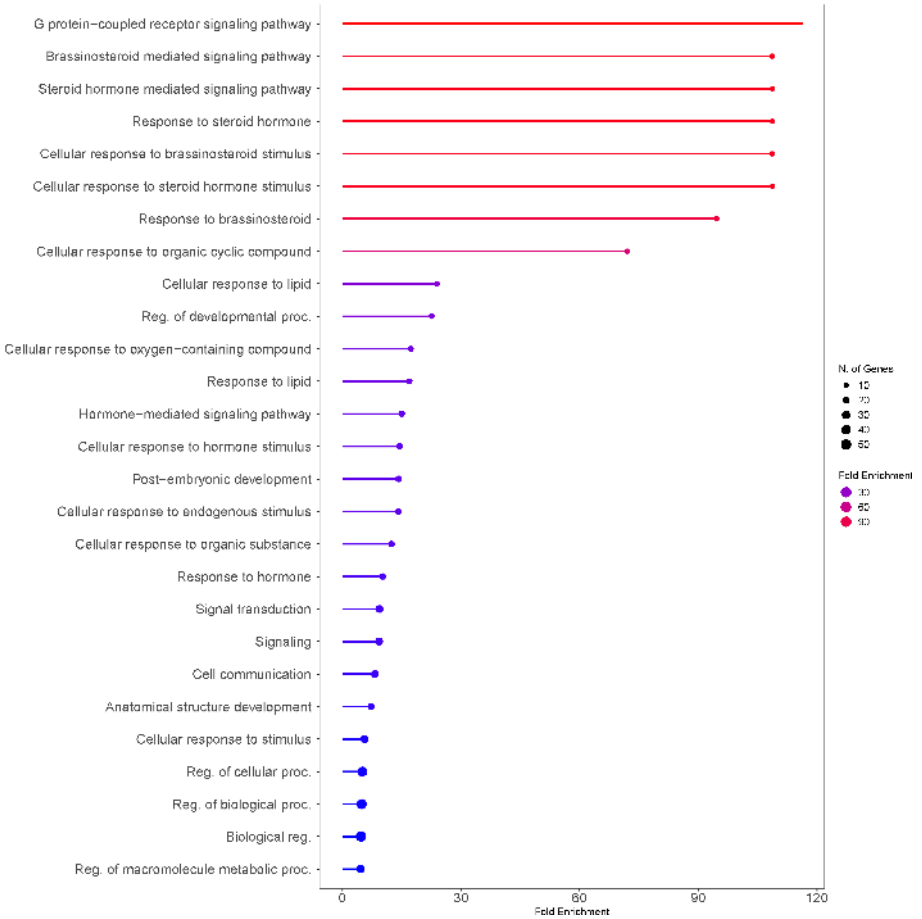

FIGURE 8
Number of genes that are strongly enriched in various categories of Biological Processes GO term using rice (*Oryza sativa* var. *japonica*) as a model. Reported seed size regulating genes from rice were listed and analyzed against the rice genome background (Oryza sativa Japonica Group genes IRGSP-1.0; Taxonomy ID: 39947). Genes were sorted by GO Biological Processes using ShinyGO v0.80. The size of the circle indicates the number of genes grouped into the biological function. Fold enrichment indicates the genes from our list that are overrepresented in the given pathway as a relationship to the background (Ge et al., 2020).

chromosome 1A, which was the second closest genomic feature returned from the BLAST results.

We report five orthologs in tef that resemble genes involved in MAPK-mediated regulation of seed size in rice. Tef has at least two homologous sub-genome features for each MAPK-family gene we analyzed. Notably, we found an additional copy of *MKK4* with significant similarity on chromosomes 1A and 1B. The putative tef ortholog of *OsMAPK6* shows low similarity to the rice gene, but closer similarity to the *T. aestivum* (wheat) *MAPK6* ortholog (XM_044568146.1), predicted to encode a MAP Kinase (Figure 11; Supplementary File 2). Both the tef and wheat orthologs of *MAPK6* form a cluster that shows higher similarity to the clusters for *MAPKKK70* and *MAPKKK62*. The tef ortholog of *MKK4* forms a cluster of MKK's in wheat, sorghum, rice, corn, and *Arabidopsis*. Our phylogenetic analysis indicates a high confidence of similarity (100% bootstrap value) between the predicted tef *MKKK10* sequence and the rice *MKKK10* sequence and form an obvious cluster with a high bootstrap value with the other five species we analyzed. Interestingly, the predicted tef ortholog of *MKP1* shows close similarity with the rice sequence for *MKKK70*, but next to the cluster of rice *MAPK6* and four other *MAPK6* orthologs. *OsMKKK62* shows high similarity to the sorghum and maize predicted orthologs of *MKKK70*, but little resemblance to the redundant kinase *OsMKKK70*. Both rice sequences *MKKK62* and *MKKK70* returned identical sequences in tef and sorghum. This indicates that tef and sorghum potentially lack the redundant version of *MKKK70*.

Using the sequences for the standard seed size regulating G-proteins in rice (Figure 12), we were able to identify all but one of tef putative orthologs using rice sequences. Each G-protein family gene we analyzed returned two homologous sub-genome features in tef. For *RGG1*, we used the sequence from *S. bicolor* (sorghum) (XM_002464159.2) to generate the CDS in tef. From the phylogenetic analysis, the *RGG1* tef ortholog shows no significant similarity to the rice *RGG1*, but high similarity to the *Z. mays* (maize) and sorghum sequences, which are annotated as Gγ proteins. We also predict coding sequences in tef that show high similarity to the rice genes *RGA1* and *RGB2*. In addition to *RGG1*, *DEP1* and *GGC2* orthologous tef sequences show closer similarity to the maize and sorghum sequences, than to the rice sequences. The

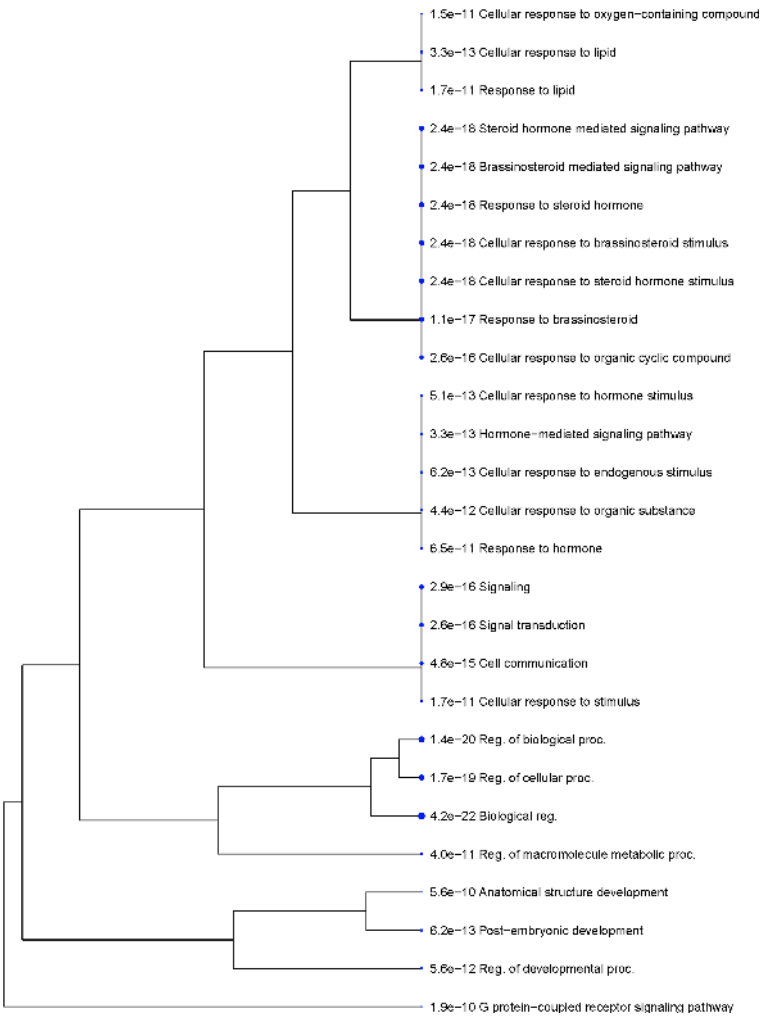

FIGURE 9
Functional clusters of genes are strongly enriched in distinct categories based on biological function. The size of the circle indicates the number of genes enriched in the group. GO terms of similar biological functions are clustered together. The numerical value in front of the GO ID and term indicates the enrichment False Discovery Rate (FDR; *FDR< 0.05*), which is an adjusted *p - value* to test the statistical significance of the enrichment. Reported seed size regulating genes from rice were listed and analyzed against the rice genome background (Oryza sativa Japonica Group genes IRGSP-1.0; Taxonomy ID: 39947).

tef ortholog of *GS3* shows lower similarity to the rice sequence but shares a high bootstrap value with a cluster that includes the *GS3* gene in rice, in addition to the barley, wheat, sorghum, and maize *GS3* putative orthologs.

The phylogenetic analysis of BR pathway associated genes were limited to 21 reference genes (Figure 13; Supplementary File 2). Fourteen of the reference genes returned two highly similar homologous sub-genome values in tef. We found a significantly similar third copy of *BRD2* on chromosome 6A. *GW5* and *GS5* returned two sub-genome values, but in both cases, other loci on different chromosomes were ahead of the second homolog in similarity. We also found several highly similar copies of *D11*/CYP724B1 along chromosomes 7A and 7B. *GW10* returned a third sub-genome feature on chromosome 5B (*E-value = ~ 0.0*). From the gene phylogeny, we identified a locus in tef that shows high similarity to the rice *GSK2* sequence, along with the sorghum and maize *GSK2* orthologs. The predicted tef *GSK3* ortholog shows high similarity to the maize and sorghum *GSK3* sequences, which both encode SHAGGY-related kinases. *BRD1*, *BRD2, GL3.1, GS5, BAK1* and *BZR1* show a high degree of conservation among the seven species we analyzed, with clear clustering of sequences with high bootstrap values. The cluster of genes generated from the *OsD11*/CYP724B1 sequence shows high similarity among the monocots, with *Arabidopsis* forming an outgroup of the cluster. The tef orthologous sequence of *OsD11*/CYP724B1 was most closely related to the maize gene, which also encodes a CYP724B1. The cluster forming *LAC*-related sequences formed a single cluster with a high bootstrap value. However, within the *LAC* cluster, the separation between the homologs is less clear. Regardless, the tef and sorghum *OsLAC* orthologs show the closest similarity. The *BRI, GW5, GS2/GRF4, OFP1, AGO17* and *OFP8* clusters indicate high conservation among the monocots, with low similarity to the *Arabidopsis* sequences. Interestingly, the tef orthologous sequences to *WRKY53* and *POW1* show low similarity to the rice

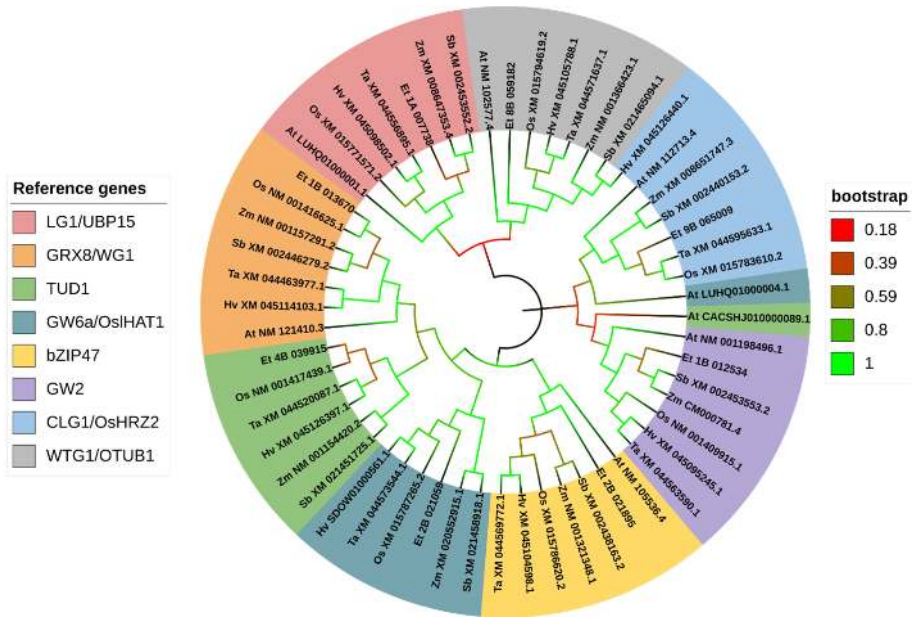

FIGURE 10
Gene phylogeny of ubiquitin-proteasome pathway genes regulating seed size. Genes were gathered from six monocot species (*Oryza sativa:* Os, *Eragrostis tef*; Et, *Hordeum vulgare*: Hv; *Triticum aestivum*: Ta; *Sorghum bicolor*: Sb; and *Zea mays*: Zm) and the model dicot species (*Arabidopsis thaliana*; At), using the rice sequence as the reference, unless otherwise noted. The first two letters in front of the gene ID represent the initials of the genus and species names. The *Arabidopsis UBP15* homolog LUHQ01000001.1 refers to CDS of the protein OAP15376.1. The *Arabidopsis GW6a/OslHAT* homolog LUHQ01000004.1 refers to CDS of the protein OAO96749.1. Genes are sorted and colored based on the rice gene they were referenced from. In the legend, gene names separated with a forward slash (/) indicate synonymous names. Gene names separated with a plus sign (+) indicate genes that were referenced from more than one rice sequence. Branches on the tree are colored to highlight the bootstrap value.

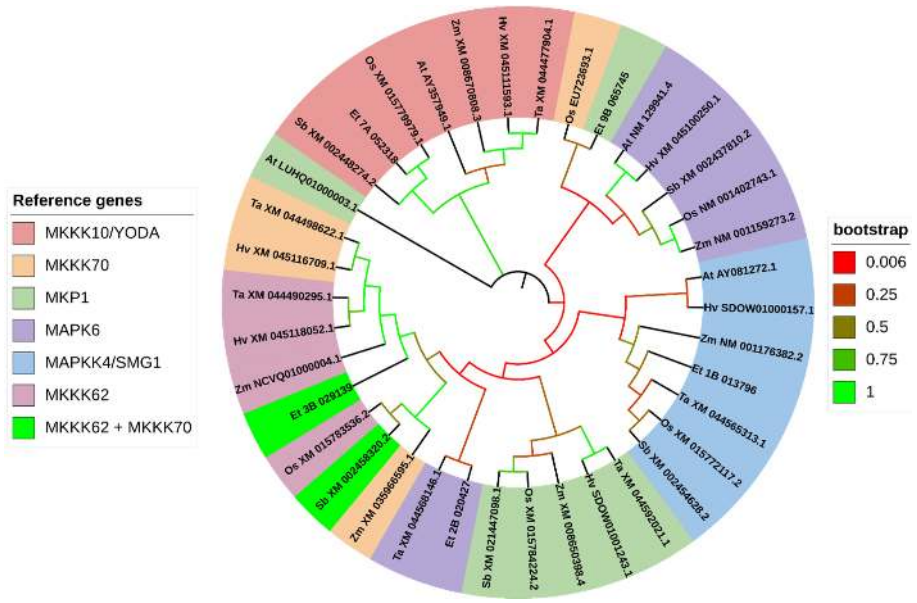

FIGURE 11
Gene phylogeny of genes associated with MAPK-mediated regulation of seed size. Genes were gathered from six monocot species (*Oryza sativa:* Os, *Eragrostis tef*; Et, *Hordeum vulgare*: Hv; *Triticum aestivum*: Ta; *Sorghum bicolor*: Sb; and *Zea mays*: Zm) and the model dicot species (*Arabidopsis thaliana*; At), using the rice sequence as the reference, unless otherwise noted. The first two letters in front of the gene ID represent the initials of the genus and species names. No significantly similar sequences were identified for *MKKK62* and *MKKK70* in *Arabidopsis thaliana*. *MKKK62* homolog Zm NCVQ01000004.1 represents the CDS associated with the protein PWZ31627.1 of *Z. mays* cultivar inbred line Mo17. The *Arabidopsis* homolog of *MKP1*, LUHQ01000003.1, encodes the protein OAP02192. Genes are sorted and colored based on the rice gene they were referenced from. In the legend, gene names separated with a forward slash (/) indicate synonymous names. Gene names separated with a plus sign (+) indicate genes that were referenced from more than one rice sequence. Branches on the tree are colored to highlight the bootstrap value.

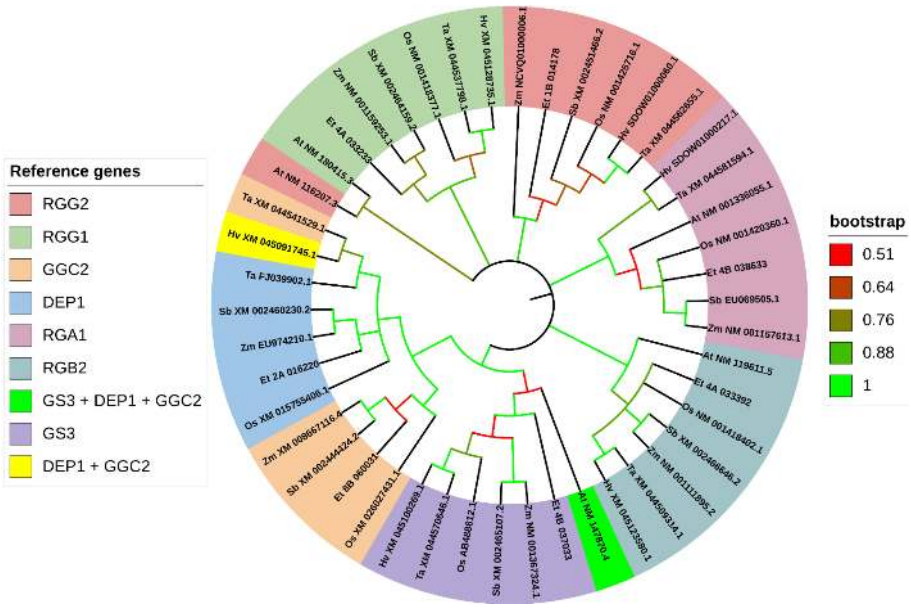

FIGURE 12
Gene phylogeny of G-protein mediated signaling pathway genes reported to control seed size in rice. Genes were gathered from six monocot species (*Oryza sativa:* Os, *Eragrostis tef*; Et, *Hordeum vulgare*: Hv; *Triticum aestivum*: Ta; *Sorghum bicolor*: Sb; and *Zea mays*: Zm) and the model dicot species (*Arabidopsis thaliana*; At), using the rice sequence as the reference. The first two letters in front of the gene ID represent the initials of the genus and species names. For *RGG1,* no significant ortholog was identified for tef using the rice reference. For *RGG1,* the *Sorghum bicolor* sequence was used to generate the predicted sequence in tef. The *Zea mays* ortholog of *RGG1* generated the same tef sequence. *RGG2* ortholog Zm NCVQ01000006.1 here refers to the CDS associated with the protein PWZ21711.1 of *Z. mays* cultivar inbred line Mo17. Genes are sorted and colored based on the rice gene they were referenced from. In the legend, gene names separated with a forward slash (/) indicate synonymous names. Gene names separated with a plus sign (+) indicate genes that were referenced from more than one rice sequence. Branches on the tree are colored to highlight the bootstrap value.

and other sequences we analyzed. The tef ortholog of *GS9* sequences appeared to be a partial sequence and shows low similarity to the main cluster with *OsGS9* and *GS9* orthologs in the other monocots. The tef *GS9* and *DLT* sequences form a two-leafed group in our analysis, both of which are not included in the main clusters. The cluster for *GW10* forms a clear group, including *Arabidopsis*, with a 99.2% bootstrap value. Oddly, the maize *GW10* ortholog is much more distant. The sorghum *GW10* sequence is predicted to encode CYP450-89A2, the same as *OsGW10* and exhibits the highest similarity to the putative tef ortholog.

4 Discussion

The purpose of this research was to explore the natural variation in seed size among 189 unique genotypes of tef. From our analysis, we have found significant variation in seed length, width, and area among most of the tef population. Using representative accessions, we also found significant variation in 1000-grain weight. Our research also led us to explore the seed size variation in the wild progenitor of tef, *E. pilosa*. Additionally, tef's nutritional qualities confer special attention, due to its superiority over more popular cereals such as rice, wheat, and maize (Ligaba-Osena et al., 2021). As such, we studied the overall mineral content in seeds of selected tef and *E. pilosa* grains and explored their distribution within the seeds using SXRF imaging. Lastly, we performed comparative genomics to identify putative seed size regulating genes in tef, conducted a phylogenetic analysis of seed size regulating genes across a variety of plant species, and conducted functional annotation of seed size regulating genes based on genes from the model monocot cereal *Oryza sativa* (rice).

4.1 Mineral concentration and spatial distribution in the grains

Studying the spatial distribution of minerals in seeds is important for understanding the fundamentals of seed development, nutrient bioavailability, bio-accessibility, and strategies for biofortification. Cereals provide a substantial portion of the mineral nutrients acquired through diet (Arafsha et al., 2023) and mineral malnutrition remains a global health concern and is widespread in both the developed and developing world (Zimmermann and Hurrell, 2007). To study the mineral content of biological samples such as grains, inductively coupled plasma mass spectrometry (ICP-MS) or ICP-OES has been widely used (Dame, 2020; Woldetsadik et al., 2024). However, it gives no insight into localization patterns and potential bioavailability.

To study the spatial distribution of minerals, researchers have turned to synchrotron-based x-ray fluorescence (SXRF) imaging to visually study and quantify the minerals within biological tissues. XRF imaging is a non-destructive imaging technique that allows *in situ* 2D quantification of elements (Donner et al., 2012, 2013). Some studies have utilized XRF spectroscopy to study hyperaccumulators

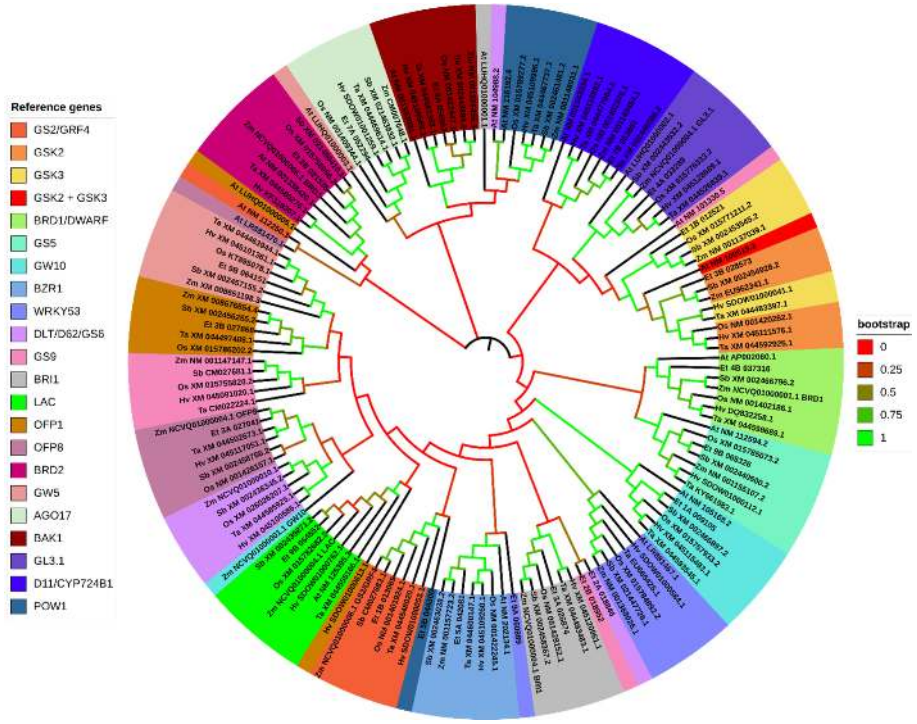

FIGURE 13
Gene phylogeny of genes involved in brassinosteroid (BR) signaling and biosynthesis reported to control grain size in rice. Genes were gathered from six monocot species (*Oryza sativa:* Os, *Eragrostis tef*; Et, *Hordeum vulgare*: Hv; *Triticum aestivum*: Ta; *Sorghum bicolor*: Sb; and *Zea mays*: Zm) and the model dicot species (*Arabidopsis thaliana*; At). The first two letters in front ot the gene ID represent the initials of the genus and species names. No significantly similar sequences were identified in *Arabidopsis* for the *AGO17* cluster. Some of the sequences in the tree above share a common label (gene ID). Here is the legend for those with common gene IDs: (Zm NCVQ01000006.1 (BRD1): PWZ55818.1) (Zm NCVQ01000004.1 (GL3.1): PWZ31897.1) (Zm NCVQ01000004.1 (BRI1): PWZ33145.1) (Zm NCVQ01000004.1 (LAC): PWZ31747.1) (Zm NCVQ01000004.1 (OFP8): PWZ31834.1) (Zm NCVQ01000006.1 (BRD2): PWZ21402.1) (Zm NCVQ01000006.1 (BRD2): PWZ22989.1) (*BRI1/DG1*: At LUHQ01000001.1: OAP19821.1) (*GL3.1*: At LUHQ01000002.1: OAP09351.1) (*GW5*: LUHQ01000003.1: OAP01951.1) (*OFP1*: LUHQ01000005.1: OAO94238.1). Genes are sorted and colored based on the rice gene they were referenced from. Genes names separated with a forward slash (/) indicate synonymous names. Gene names separated with a plus sign (+) indicate genes that were referenced from more than one rice sequence. Branches on the tree are colored to highlight the bootstrap value.

(Goudard et al., 2024). Others have utilized XRF imaging for studying micronutrient mobilization during seed germination *in vivo* (Takahashi et al., 2009). SXRF microscopy has proven to be useful in studying alterations in mineral accumulation in loss of function studies (Chia et al., 2023; Sheng et al., 2021). In this report, we utilized SXRF imaging to explore the distribution of minerals within the tef grains for the first time. Our findings show that Ca is associated with the seed coat and the embryo while K is detected in the outer embryo and seed coat, and the micronutrients are predominantly associated with the embryo and seed coat. The concentration of minerals was also variable among the genotypes we analyzed (Figure 5). A similar distribution pattern has been reported in rice, where minerals were shown to localize in the embryo and remain absent from the endosperm (Lu et al., 2013).

For processing and consumption, mineral localization is nutritionally important. In many cereals, the milling process separates the embryo and endosperm (Shewry et al., 2020). Because mineral localization is variable, this may inadvertently remove micronutrients from the final product. The whole seed is milled in some cereals, such as tef, leaving the endosperm and embryo components together. Milling the entire seed potentially increases the overall amount of nutrients in the flour. However, the localization of minerals in different seed tissues may alter bio-accessibility or dictate how easily minerals are accessed from seed tissues upon digestion (Arafsha et al., 2023). For example, the plant cell wall is resistant to digestion in the human digestive tract. The cells of the aleurone layer are composed of water-insoluble indigestible fibers (Arafsha et al., 2023; Brouns et al., 2012). The starchy endosperm is the most digestible part of the seed, but lacks minerals, which was evident in the tef accessions we analyzed (Gidley, 2024; Khalid et al., 2023) (Figure 5). Furthermore, the accessibility of minerals can be affected by how the mineral is stored in the seed chemically. For example, prominent levels of phytate can bind minerals in the seeds, making them less accessible during digestion (Mandha and Raes, 2023). Phytate tends to localize at the highest concentration in the aleurone layer and the embryo (Bohn et al., 2008). Baye et al. (2015) addressed this issue though the additions of phytate degrading enzymes in tef flour to improve iron bioavailability. Furthermore, tef is usually consumed after fermentation, which was reported to decrease the concentration of antinutrients such as phytate and increase mineral bioavailability (Fischer et al., 2014; Shumoy and Raes, 2017).

4.2 Evolution of seed size and mineral content of tef

Plant domestication has followed convergent, but highly similar processes globally among many cultivated crops (Fuller et al., 2014). In cereals, the major traits selected were shattering and lodging resistance, along with greater seed size to improve yield and quality. Many domestication traits were selected for purposely, whereas others may have been inadvertently acquired through genetic linkage, pleiotropy, or otherwise unconsciously (Purugganan, 2019). The increase in seed size among the cereals is contested, with some arguing that enlarged seed size was consciously selected for, which would bring higher yield and ease of handling. Others have argued that seed size was linked to characteristics such as height, suggesting that the enlargement of seed size was unconscious (Purugganan, 2019; Meyer and Purugganan, 2013). In our study, we show that tef seeds can be as much as 4.44 times larger by area and 2.35 times longer than its progenitor *E. pilosa*. The evolution of grain size in tef remains to be fully understood.

In some cases, individual genes that produce highly desirable traits and share commonality among closely related cultivars or between species, show signs of selection at the genetic level (Meyer and Purugganan, 2013). The capability to sequence genomes reliably and efficiently for genome wide association studies (GWAS) and quantitative trait loci (QTL) mapping has provided novel insights to the genomic evolution of domesticated cereals. In rice, this has revealed the genetic basis of seed size, a key determinant of yield in cereals (Tao et al., 2020b). We anticipate that future research will reveal such insights in tef.

Seed size is an extraordinarily complex trait that is influenced by both genetic and environmental conditions. Furthermore, there seems to be a tradeoff between offspring number and size, as reported by those studying natural variations and single gene mutants (Gnan et al., 2014; Zhang et al., 2017a, 2017b; Xiao et al., 2019). In natural populations, there is strong environmental pressure that determines seed size, and variation exists across species (Venable, 1992). Razzaque et al. (2023) showed that local adaptation to climate was a major predictor of seed size of the perennial grass *Panicum hallii*. They showed that annual temperature showcased the strongest predictability in determining seed size among *P. hallii* ecotypes. During the course of domestication, humans artificially selected against the natural plant dispersal mechanisms (shattering) and selected for increased seed size (Harlan et al., 1973). Harlan et al. (1973) revealed that the transition from wild grasses to their domesticated counterparts follow highly similar trajectories among species. Fuller et al. (2014) substantiated these ideas, highlighting extensive parallelism (convergence) in the evolution of non- shattering species and the enlargement of seeds in cereals. In tef, one estimate indicates a 72% difference in grain thickness in tef and *E. pilosa*, which is consistent with our estimates (D'Andrea, 2008). D'Andrea (2008) speculated that early tef cultivators encountered several obstacles during domestication and that tef's domestication history was highly unconventional in comparison to other cereals. In speculation about why tef grains have remained so small, one hypothesis suggests that natural introgression with *E. pilosa* may have helped maintain small seed sizes. Early cultivators more likely recognized tef's susceptibility to shattering and lodging and may have selected against large panicles and large, heavy grains (D'Andrea, 2008; Ketema, 1997). Lastly, water-conservation practices and attempts to reduce lodging resulted in minimal tilling practices in Ethiopia's semi-arid highlands, which is hypothesized to reduce selection pressure for large seeds (Harlan et al., 1973). Furthermore, D'Andrea (2008) notes that modern Ethiopian farmers do not select for large seeds, but rather for color, appearance, increased panicle branching, and more numerous grains, suggesting that any improvement to grain size in tef has been non-deliberate or unconscious (Purugganan, 2019; Meyer and Purugganan, 2013). For those interested in breeding large seed cultivars of tef, these limitations remain a major point of consideration.

The relationship between seed size and mineral content remains elusive in tef. Previously, there has been some speculation regarding the relationship between seed size and mineral content. It is suggested that higher mineral contents of the seed may confer a major evolutionary advantage. Additionally, smaller seeds are more easily dispersed and evolutionarily beneficial but have become agronomically undesirable in many seed-bearing crops. However, some have suggested that any benefit conferred from a larger seed diminishes if the external nutrient availability is low, when smaller, nutrient-dense seeds may be much more advantageous (Grubb and Burslem, 1998). Curiously, Grubb and Burslem (1998) found an inverse relationship between mineral content and seed size. They showed that larger seed size diminished N, P, K, Mg, and Ca concentrations in several different crop species. It is interesting to speculate that the ancestral species of tef, *E. pilosa*, may have benefited evolutionarily from this kind of selection. However, this remains to be validated. The accession with the smallest and lightest grains from either the tef or *E. pilosa* accessions was PI442115 (*E. pilosa*), which showed the highest concentration of Cu, Zn, and Mn from the ICP-OES, and had clearly higher levels of Mn in the SXRF (Figures 6, 7). However, in PI442115, the mineral concentration of other elements was relatively lower, potentially indicating some kind of trade off. The preliminary results of whether there is any relationship between seed size and mineral concentration in tef or *E. pilosa* seeds are inconclusive. We will continue to pursue this by analyzing the mineral content of more tef genotypes, where we can begin to unravel the relationship between seed size and mineral content further.

4.3 Overview of the genetic mechanisms regulating seed size

The genetic mechanisms regulating seed development and seed size are complex and poorly understood. Even in a model monocot such as rice, the underlying mechanisms regulating seed size remain fragmented. Regardless, developing novel rice cultivars with enhanced seed size is a major area of research, as larger seed size is often associated with improved yield and quality. In tef cultivation, its small seed size presents numerous challenges to its cultivation. Much of the general biology, let alone the genetic mechanisms regulating seed development, remains to be elucidated in tef. Additionally, the tef genome remains to be fully annotated.

In rice there are ~80 genes that have been implicated in seed size regulation (Li et al., 2022). Some of these have been selected for over

the course of domestication, whereas many others represent rare alleles identified from mutagenic studies or evaluation of rice progenitor species. Most of these genes regulate seed size via changes to cell proliferation, elongation, and enlargement in the seed hull (and related tissues), or in maternally derived tissues such as the embryo or endosperm (Li et al., 2022). Many of the genes encode receptors, proteins involved in various signaling mechanisms, hormone biosynthetic enzymes, transcriptional regulators, and transcription factors. For simplicity, they can be grouped by a shared pathway or function, but many have been shown to be multi-pathway regulating or multifunctional.

Our analysis suggests that G-protein mediated signaling is the most significantly (*FDR< 0.05*) enriched biological process regulating grain size in rice (Figures 8, 9). Several G-proteins have been implicated in seed size regulation in rice, many of which were identified as major quantitative loci. The rice genome encodes one Gα subunit (RGA1), one Gβ subunit (RGB1), and five Gγ subunits (RGG1, RGG2, GS3, DEP1, and GGC2), which compose the major G-proteins in rice (Mao et al., 2010; Fan et al., 2006; Sun et al., 2018; Tao et al., 2020b; Miao et al., 2019). Plants also exhibit a number of non-canonical extra-large GTP-binding proteins (XLGs), which are large atypical Gα subunits (Maruta et al., 2021). XLGs have been shown to regulate several complex traits including panicle architecture, stress tolerance, and some agronomic traits such as 1000-grain weight (Biswal et al., 2022; Zhao et al., 2022; Cantos et al., 2023). However, compared to the other G-proteins we used in our analysis, these XLGs are not as well characterized and their effects on seed size regulation are still under active investigation. As a result, we chose to focus our study on the better characterized G-protein subunits and omit XLGs from our analysis. Unlike animals, plants lack G-protein coupled receptors (GPCR's) and are self-activated (Urano and Jones, 2014). To transduce the signal, Gβ and Gγ subunits form a Gβγ heterodimer, which go on to interact with Gα (Maruta et al., 2021; Urano and Jones, 2014). Evidence suggests that various combinations of G-protein subunits exhibit unique alterations to key developmental processes and specific alterations to various grain size attributes, either positively or negatively (Sun et al., 2018; Pandey, 2020). We were able to identify putative orthologs in tef for all but one of the G-protein family genes using the rice CDS (Criterion from BLAST: *E< 1e-30*). We found no significantly similar sequences in tef using the *OsRGG1* sequence but found that the maize and sorghum *RGG1* orthologs produced the same tef sequence, which was validated by our phylogenetic analysis (Figure 12). Otherwise, the other tef sequences we identified pass the standard cutoff for the determination of homology using BLAST (Choudhuri, 2014; Pearson, 2013).

The ubiquitin-proteasome pathway regulates protein stability, activity, and degradation in eukaryotes. In plants, ubiquitin-mediated processes contribute to many biological processes, such as embryogenesis, hormone signaling, and senescence (Moon et al., 2004). Eight genes involved in ubiquitin-mediated regulation of seed size have been reported in rice. Many of these genes have been identified in other grasses and in *Arabidopsis*, where they have exhibited highly conserved functions (Du et al., 2014; Tang et al., 2023; Shi et al., 2019; Keren et al., 2020; Li and Li, 2014; Wang et al., 2020; Zombori et al., 2020; Brinton et al., 2018). Most of the genes grouped in ubiquitin-mediated pathways act as regulators of other systems, and exhibit crosstalk with the G-protein, BR, GA, and MAPK pathways, in addition to regulating the activity of some transcription factors (Bai et al., 2024; Yang et al., 2021; Hu et al., 2013; Hao et al., 2021; Tian et al., 2021). From tef, we report eight putative orthologs involved in ubiquitin-mediated regulation of grain size, along with their homologous sub-genome feature (Figure 10; Supplementary File 2) (Choudhuri, 2014; Pearson, 2013).

The mitogen activated protein kinase (MAPK) cascade is a central signaling mechanism in eukaryotes (Taj et al., 2010). MAPK cascades are comprised of closely related kinases that are sequentially phosphorylated, eventually leading to a final activated MAPK that phosphorylates transcription factors and other transcriptional regulators (Hardie, 1999). In rice, a three-gene MAPK module serves as a positive regulator of grain size (Liu et al., 2021; Xu et al., 2018; Liu et al., 2015; Duan et al., 2014), with a single phosphatase acting as the negative regulator of the module (Guo et al., 2018; Xu et al., 2018). Furthermore, two redundant MKKK's have been shown to regulate the activity of the transcription factor *OsWRKY53* and influence grain size via the MAPK module mentioned above (Liu et al., 2021; Tian et al., 2021). Of the six MAPK-related seed size regulating genes reported in rice, only five putative orthologs could be described in tef, along with their sub-genome homolog (Figure 11; Supplementary File 2). BLAST of the tef genome with *OsMKKK62* and *OsMKKK70* returned the same tef sequence, indicating that the tef genome likely lacks the paralog of this kinase.

Several phytohormones and their associated regulatory mechanisms have been linked to seed development, including brassinosteroids (BRs), auxin (IAA), gibberellins (GA), and cytokinin's (Li et al., 2022). The functional analysis concluded that brassinosteroid signaling and homeostasis was a significantly enriched biological process involved in seed size regulation in rice (Figures 8, 9; Supplementary File 3). We report 21 putative orthologs in tef from rice genes involved in BR biosynthesis and regulation, along with their sub-genome homolog (Figure 13). Furthermore, several transcription factors, functional proteins, microRNAs (miRNA's), and proteins involved in endosperm development are reported to determine final grain size and weight in rice. Of such genes, we report 29 putative orthologs in tef, along with their corresponding sub-genome homolog (Supplementary File 2).

4.4 Conclusions and limitations

This study characterized the seed size phenotype of 189 tef and 11 *E. pilosa* genotypes. Unfortunately, this is reflective of most, but not all of the tef population available at the USDA-ARS germplasm center. We feel our analysis likely reflects the true nature of the tef population but remains incomplete without more research. In our analysis we only utilized tef and *E. pilosa* accessions whose seeds exhibited relatively homogenous coloration and treated samples with different colored seeds as impure. To ensure that our seed size analysis was as accurate as possible, we excluded samples with multi-colored seeds. Additionally, the seeds used in this study were

from stock collections from the U.S. National Plant Germplasm Center to ensure they were sourced from a single location, grown under the same field conditions. Notably, replication of many of these accessions under greenhouse conditions indicated that seed size is true to type (data not shown).

Furthermore, there is no literature on the anatomy of tef seeds, so the conclusions we make for the SXRF imaging are highly general. To better understand seed development in tef, the impact of genetic manipulation, and the localization of minerals, more research into the developmental anatomy of tef is required. Moreover, the tef genome annotation remains to be completed, so validation of the findings from our comparative genomic analysis is impossible without further research. Future research into tef should aim to address these limitations. Regardless, we hope the findings from this research can provide a starting point for those interested in breeding high quality, nutritious tef, and for those interested in the evolutionary genetics of tef domestication.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

EW: Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing – original draft. CK: Data curation, Formal analysis, Writing – original draft. LS: Data curation, Funding acquisition, Investigation, Resources, Writing – review & editing. J-CC: Investigation, Writing – review & editing. AB-A: Writing – review & editing, Data curation. OV: Writing – review & editing, Funding acquisition, Methodology, Resources. RV: Funding acquisition, Resources, Writing – review & editing, Conceptualization. AL-O: Conceptualization, Funding acquisition, Resources, Writing – review & editing, Methodology, Supervision.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This manuscript was financially supported by the National Institute of Health # 1SC2GM144193- 01 and National Science Foundation Plant Genome Research Program # 2327906 to AL-O. SXRF measurements were conducted at the Functional Materials Beamline at the Cornell High Energy Synchrotron Source (CHESS) which during the period of research was supported by AFRL under agreement number FA8650-22-2-5200. A portion of LS effort on this work was supported by XLEAP, NSF-BIO 2330043. Research in O.K.V. lab is funded by the Agriculture and Food Research Initiative Award 2021-67013-33798 from the USDA National Institute of Food and Agriculture.

Acknowledgments

The authors acknowledge graduate students Muhammed Numan for his help in the comparative genomics and functional analysis, Raza Ullah for help with seed imaging and contribution to generating seed size data, and undergraduates Daniel Araya Staples, Willy Pina, and Abdalkader Salaimah for their contribution to generating the seed size data used in this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2024.1485819/full#supplementary-material

References

Akinci, C., Yıldırım, M., and Bahar, B. (2008). The effects of seed size on emergence and yield of durum wheat. *J. Food Agric. Environ.* 6, 234–237.

Alam, I., Batool, K., Huang, Y., Liu, J., and Ge, L. (2022). Developing genetic engineering techniques for control of seed size and yield. *Int. J. Mol. Sci.* 23, 13256. doi: 10.3390/ijms232113256

Anderson, B. E., and Volesky, J. D. (2012). Summer annual forage grasses. University of Nebraska. *NebGuide G74-171*.

Arafsha, S. M., Aslam, M. F., Ellis, P. R., Latunde-Dada, G. O., and Sharp, P. A. (2023). Strategies to increase the bioaccessibility and bioavailability of iron and zinc from cereal products. *Proc. Nutr. Soc.*, 1–7. doi: 10.1017/S0029665123003543

Assefa, K., Cannarozzi, G., Girma, D., Kamies, R., Chanyalew, S., Plaza-Wüthrich, S., et al. (2015). Genetic diversity in tef [Eragrostis tef (Zucc.) Trotter. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00177

Ayele, M., Tefera, H., Assefa, K., and Nguyen, H. T. (1999). Genetic characterization of two Eragrostis species using AFLP and morphological traits. *Hereditas* 130, 33–40. doi: 10.1111/j.1601-5223.1999.00033.x

Badoni, S., Parween, S., Henry, R. J., and Sreenivasulu, N. (2023). Systems seed biology to understand and manipulate rice grain quality and nutrition. *Crit. Rev. Biotechnol.* 43, 716–733. doi: 10.1080/07388551.2022.2058460

Bai, C., Wang, G.-J., Feng, X.-H., Gao, Q., Wang, W.-Q., Xu, R., et al. (2024). OsMAPK6 phosphorylation and CLG1 ubiquitylation of GW6a non-additively enhance rice grain size through stabilization of the substrate. *Nat. Commun.* 15, 4300. doi: 10.1038/s41467-024-48786-0

Baye, K., Guyot, J.-P., Icard-Vernière, C., Rochette, I., and Mouquet-Rivier, C. (2015). Enzymatic degradation of phytate, polyphenols and dietary fibers in Ethiopian injera flours: effect on iron bioaccessibility. *Food Chem.* 174, 60–67. doi: 10.1016/j.foodchem.2014.11.012

Bekele, E., and Lester, R. N. (1981). Biochemical Assessment of the Relationships of Eragrostis tef (Zucc.) Trotter with some Wild Eragrostis Species (Gramineae). *Ann. Bot.* 48, 717–725. doi: 10.1093/oxfordjournals.aob.a086178

Bekele-Alemu, A., and Ligaba-Osena, A. (2023). Comprehensive in silico analysis of the underutilized crop tef (Eragrostis tef (Zucc.) Trotter) genome reveals drought tolerance signatures. *BMC Plant Biol.* 23, 506. doi: 10.1186/s12870-023-04515-1

Ben-Zeev, S., Rabinovitz, O., Orlov-Levin, V., Chen, A., Graff, N., Goldwasser, Y., et al. (2020). Less is more: lower sowing rate of irrigated tef (Eragrostis tef) alters plant morphology and reduces lodging. *Agronomy* 10, 570. doi: 10.3390/agronomy10040570

Beyene, G., Chauhan, R. D., Villmer, J., Husic, N., Wang, N., Gebre, E., et al. (2022). CRISPR/Cas9-mediated tetra-allelic mutation of the 'Green Revolution' SEMIDWARF-1 (SD-1) gene confers lodging resistance in tef (Eragrostis tef). *Plant Biotechnol. J.* 20, 1716–1729. doi: 10.1111/pbi.13842

Biswal, A. K., Wu, T.-Y., Urano, D., Pelissier, R., Morel, J.-B., Jones, A. M., et al. (2022). Novel mutant alleles reveal a role of the extra-large G protein in rice grain filling, panicle architecture, plant growth, and disease resistance. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.782960

Bohn, L., Meyer, A. S., and Rasmussen, S. K. (2008). Phytate: impact on environment and human nutrition. A challenge for molecular breeding. *J. Zhejiang Univ Sci. B* 9, 165–191. doi: 10.1631/jzus.B0710640

Brinton, J., Simmonds, J., and Uauy, C. (2018). Ubiquitin-related genes are differentially expressed in isogenic lines contrasting for pericarp cell size and grain weight in hexaploid wheat. *BMC Plant Biol.* 18, 22. doi: 10.1186/s12870-018-1241-5

Brouns, F., Hemery, Y., Price, R., and Anson, N. M. (2012). Wheat aleurone: separation, composition, health aspects, and potential food use. *Crit. Rev. Food Sci. Nutr.* 52, 553–568. doi: 10.1080/10408398.2011.589540

Bultosa, G. (2016). "Teff, overview," in *Reference module in food science* (Elsevier). doi: 10.1016/B978-0-08-100596-5.00018-4

Bultosa, G., and Taylor, J. R. N. (2004). ""TEFF,"," in *Encyclopedia of grain science*. Ed. C. Wrigley (Elsevier, Oxford), 281–290. doi: 10.1016/B0-12-765490-9/00172-5

Cannarozzi, G., Plaza-Wüthrich, S., Esfeld, K., Larti, S., Wilson, Y. S., Girma, D., et al. (2014). Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (Eragrostis tef). *BMC Genomics* 15, 581. doi: 10.1186/1471-2164-15-581

Cantos, C. F., dePamphilis, C. W., and Assmann, S. M. (2023). Extra-large G proteins have extra-large effects on agronomic traits and stress tolerance in maize and rice. *Trends Plant Sci.* 28, 1033–1044. doi: 10.1016/j.tplants.2023.04.005

Cheng, A., Mayes, S., Dalle, G., Demissew, S., and Massawe, F. (2017). Diversifying crops for food and nutrition security - a case of teff. *Biol. Rev. Camb Philos. Soc.* 92, 188–198. doi: 10.1111/brv.12225

Chia, J.-C., Woll, A. R., Smieska, L., and Vatamaniuk, O. K. (2023). ""Visualizing metal distribution in plants using synchrotron X-ray fluorescence microscopy techniques,"," in *Plant iron homeostasis: methods and protocols*. Ed. J. Jeong (Springer US, New York, NY), 177–189. doi: 10.1007/978-1-0716-3183-6_14

Choudhuri, S. (2014). ""Chapter 6 - Sequence Alignment and Similarity Searching in Genomic Databases: BLAST and FASTA**The opinions expressed in this chapter are the author's own and they do not necessarily reflect the opinions of the FDA, the DHHS, or the Federal Government," in *Bioinformatics for beginners*. Ed. S. Choudhuri (Academic Press, Oxford), 133–155. doi: 10.1016/B978-0-12-410471-6.00006-2

Cotter, C. J., Wright, A. J., Romanov, A. V., Graf, T. N., Whisnant, E. D., Flores-Bocanegra, L., et al. (2023). Evaluating the antioxidant properties of the ancient-crop tef (Eragrostis tef) grain extracts in THP-1 monocytes. *Antioxidants (Basel)* 12, 1561. doi: 10.3390/antiox12081561

D'Andrea, A. C. (2008). T'ef (Eragrostis tef) in ancient agricultural systems of highland Ethiopia. *Econ Bot.* 62, 547–566. doi: 10.1007/s12231-008-9053-4

Dame, Z. T. (2020). Analysis of major and trace elements in teff (*Eragrostis tef*). *J. King Saud Univ. - Sci.* 32, 145–148. doi: 10.1016/j.jksus.2018.03.020

Davison, J., Laca, M., and Creech, E. (2011). The potential for teff as an alternative forage crop for irrigated regions., 2011 Western Alfalfa & Forage Symposium (Las Vegas, NV), 11–13. Available online at: https://alfalfa.ucdavis.edu/sites/g/files/dgvnsk12586/files/media/documents/the_potential_for_tef_as_an_alternative_forage_crop_for_irrigated_regions_%282011%29_by_jay_davison_mike_laka_and_earl_creech_0.pdf.

Donner, E., de Jonge, M. D., Kopittke, P. M., and Lombi, E. (2013). ""Mapping element distributions in plant tissues using synchrotron X-ray fluorescence techniques,"," in *Plant mineral nutrients: methods and protocols*. Ed. F. J. M. Maathuis (Humana Press, Totowa, NJ), 143–159. doi: 10.1007/978-1-62703-152-3_9

Donner, E., Punshon, T., Guerinot, M. L., and Lombi, E. (2012). Functional characterisation of metal(loid) processes in planta through the integration of synchrotron techniques and plant molecular biology. *Anal. Bioanal Chem.* 402, 3287–3298. doi: 10.1007/s00216-011-5624-9

Du, L., Li, N., Chen, L., Xu, Y., Li, Y., Zhang, Y., et al. (2014). The ubiquitin receptor DA1 regulates seed and organ size by modulating the stability of the ubiquitin-specific protease UBP15/SOD2 in arabidopsis[W. *Plant Cell* 26, 665–677. doi: 10.1105/tpc.114.122663

Duan, P., Rao, Y., Zeng, D., Yang, Y., Xu, R., Zhang, B., et al. (2014). SMALL GRAIN 1, which encodes a mitogen-activated protein kinase kinase 4, influences grain size in rice. *Plant J.* 77, 547–557. doi: 10.1111/tpj.12405

Ereful, N. C., Jones, H., Fradgley, N., Boyd, L., Cherie, H. A., and Milner, M. J. (2022). Nutritional and genetic variation in a core set of Ethiopian Tef (Eragrostis tef) varieties. *BMC Plant Biol.* 22, 220. doi: 10.1186/s12870-022-03595-9

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., et al. (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* 112, 1164–1171. doi: 10.1007/s00122-006-0218-1

Fischer, M. M., Egli, I. M., Aeberli, I., Hurrell, R. F., and Meile, L. (2014). Phytic acid degrading lactic acid bacteria in tef-injera fermentation. *Int. J. Food Microbiol.* 190, 54–60. doi: 10.1016/j.ijfoodmicro.2014.08.018

Fuller, D. Q., Denham, T., Arroyo-Kalin, M., Lucas, L., Stevens, C. J., Qin, L., et al. (2014). Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record. *Proc. Natl. Acad. Sci.* 111, 6147–6152. doi: 10.1073/pnas.1308937110

Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36, 2628. doi: 10.1093/bioinformatics/btz931

Gebru, Y. A., Sbhatu, D. B., and Kim, K.-P. (2020). Nutritional composition and health benefits of teff (*Eragrostis tef* (Zucc.) trotter). *J. Food Qual.* 2020, e9595086. doi: 10.1155/2020/9595086

Gidley, M. J. (2024). Grain factors in food and feed that influence digestion and gut fermentation outcomes. *Cereal Chem.* 101, 288–298. doi: 10.1002/cche.10718

Girija, A., Han, J., Corke, F., Brook, J., Doonan, J., Yadav, R., et al. (2022). Elucidating drought responsive networks in tef (Eragrostis tef) using phenomic and metabolomic approaches. *Physiologia Plantarum* 174, e13597. doi: 10.1111/ppl.13597

Gnan, S., Priest, A., and Kover, P. X. (2014). The genetic basis of natural variation in seed size and seed number and their trade-off using arabidopsis thaliana MAGIC lines. *Genetics* 198, 1751–1758. doi: 10.1534/genetics.114.170746

Goudard, L., Blaudez, D., Sirguey, C., Purwadi, I., Invernon, V., Rouhan, G., et al. (2024). Prospecting for rare earth element (hyper)accumulators in the Paris Herbarium using X-ray fluorescence spectroscopy reveals new distributional and taxon discoveries. *Ann. Bot.* 133, 573–584. doi: 10.1093/aob/mcae011

Grubb, P. J., and Burslem, D. F. R. P. (1998). Mineral nutrient concentrations as a function of seed size within seed crops: implications for competition among seedlings and defence against herbivory. *J. Trop. Ecol.* 14, 177–185. doi: 10.1017/S0266467498000145

Guo, T., Chen, K., Dong, N.-Q., Shi, C.-L., Ye, W.-W., Gao, J.-P., et al. (2018). GRAIN SIZE AND NUMBER1 negatively regulates the osMKKK10-osMKK4-osMPK6 cascade to coordinate the trade-off between grain number per panicle and grain size in rice. *Plant Cell* 30, 871–888. doi: 10.1105/tpc.17.00959

Habte, M. L., Beyene, E. A., Feyisa, T. O., Admasu, F. T., Tilahun, A., and Diribsa, G. C. (2022). Nutritional values of teff (Eragrostis tef) in diabetic patients: narrative review. *Diabetes Metab. Syndr. Obes.* 15, 2599–2606. doi: 10.2147/DMSO.S366958

Hao, J., Wang, D., Wu, Y., Huang, K., Duan, P., Li, N., et al. (2021). The GW2-WG1-OsbZIP47 pathway controls grain size and weight in rice. *Mol. Plant* 14, 1266–1280. doi: 10.1016/j.molp.2021.04.011

Hardie, D. G. (1999). PLANT PROTEIN SERINE/THREONINE KINASES: classification and functions. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 97–131. doi: 10.1146/annurev.arplant.50.1.97

Harlan, J. R., de Wet, J. M. J., and Price, E. G. (1973). Comparative evolution of cereals. *Evolution* 27, 311–325. doi: 10.2307/2406971

Hu, X., Qian, Q., Xu, T., Zhang, Y., Dong, G., Gao, T., et al. (2013). The U-box E3 ubiquitin ligase TUD1 functions with a heterotrimeric G α Subunit to regulate brassinosteroid-mediated growth in rice. *PloS Genet.* 9, e1003391. doi: 10.1371/journal.pgen.1003391

Ingram, A. L., and Doyle, J. J. (2003). The origin and evolution of Eragrostis tef (Poaceae) and related polyploids: evidence from nuclear waxy and plastid rps16. *Am. J. Bot.* 90, 116–122. doi: 10.3732/ajb.90.1.116

Jiang, L., Ma, X., Zhao, S., Tang, Y., Liu, F., Gu, P., et al. (2019). The APETALA2-like transcription factor SUPERNUMERARY BRACT controls rice seed shattering and seed size. *Plant Cell* 31, 17–36. doi: 10.1105/tpc.18.00304

Jones, B. M. G., Ponti, J., Tavassoli, A., and And Dixon, P. A. (1978). Relationships of the Ethiopian cereal T'ef (Eragrostis tef (Zucc.) trotter): evidence from morphology and

chromosome number. *Ann. Bot.* 42, 1369–1373. doi: 10.1093/oxfordjournals.aob.a085583

Jöst, M., Esfeld, K., Burian, A., Cannarozzi, G., Chanyalew, S., Kuhlemeier, C., et al. (2015). Semi-dwarfism and lodging tolerance in tef (Eragrostis tef) is linked to a mutation in the α-Tubulin 1 gene. *J. Exp. Bot.* 66, 933–944. doi: 10.1093/jxb/eru452

Kebede, D., Cannarozzi, G., Weichert, A., and Tadele, Z. (2018). Genotyping by sequencing reasserts the close relationship between tef and its putative wild eragrostis progenitors. *Diversity* 10, 17. doi: 10.3390/d10020017

Keilwagen, J., Lehnert, H., Berner, T., Badaeva, E., Himmelbach, A., Börner, A., et al. (20221908). Detecting major introgressions in wheat and their putative origins using coverage analysis. *Sci. Rep.* 12, 1908. doi: 10.1038/s41598-022-05865-w

Keren, I., Lacroix, B., Kohrman, A., and Citovsky, V. (2020). Histone deubiquitinase OTU1 epigenetically regulates DA1 and DA2, which control arabidopsis seed and organ size. *iScience* 23. doi: 10.1016/j.isci.2020.100948

Ketema, S. (1997). Tef Eragrostis tef (Zucc.) Trotter. promoting the conservation and use of underutilized and neglected crops 12, 50.

Khalid, A., Hameed, A., and Tahir, M. F. (2023). Wheat quality: A review on chemical composition, nutritional attributes, grain anatomy, types, classification, and function of seed storage proteins in bread making quality. *Front. Nutr.* 10. doi: 10.3389/fnut.2023.1053196

Kotásková,, E., Sumczynski, D., Mlček, J., and Valášek, P. (2016). Determination of free and bound phenolics using HPLC-DAD, antioxidant activity and in vitro digestibility of Eragrostis tef. *J. Food Comp. Anal.* 46, 15–21. doi: 10.1016/j.jfca.2015.11.001

Kreitschitz, A., Tadele, Z., and Gola, E. M. (2009). Slime cells on the surface of Eragrostis seeds maintain a level of moisture around the grain to enhance germination. *Seed Sci. Res.* 19, 27–35. doi: 10.1017/S0960258508186287

Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.*, gkae268. doi: 10.1093/nar/gkae268

Li, P., Chen, Y.-H., Lu, J., Zhang, C.-Q., Liu, Q.-Q., and Li, Q.-F. (2022). Genes and their molecular functions determining seed structure, components, and quality of rice. *Rice* 15, 18. doi: 10.1186/s12284-022-00562-8

Li, N., and Li, Y. (2014). Ubiquitin-mediated control of seed size in plants. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00332

Li, N., and Li, Y. (2016). Signaling pathways of seed size control in plants. *Curr. Opin. Plant Biol.* 33, 23–32. doi: 10.1016/j.pbi.2016.05.008

Ligaba-Osena, A., Guo, W., Choi, S. C., Limmer, M. A., Seyfferth, A. L., and Hankoua, B. B. (2020). Silicon enhances biomass and grain yield in an ancient crop tef [Eragrostis tef (Zucc.) trotter. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.608503

Ligaba-Osena, A., Mengistu, M., Beyene, G., Cushman, J., Glahn, R., and Pineros, M. (2021). Grain mineral nutrient profiling and iron bioavailability of an ancient crop tef (Eragrostis tef). *Aust. J. Crop Sci.* 15, 1314–1324. doi: 10.3316/informit.288153232944381

Liu, S., Hua, L., Dong, S., Chen, H., Zhu, X., Jiang, J., et al. (2015). OsMAPK6, a mitogen-activated protein kinase, influences rice grain size and biomass production. *Plant J.* 84, 672–681. doi: 10.1111/tpj.13025

Liu, Z., Mei, E., Tian, X., He, M., Tang, J., Xu, M., et al. (2021). OsMKKK70 regulates grain size and leaf angle in rice through the OsMKK4-OsMAPK6-OsWRKY53 signaling pathway. *J. Integr. Plant Biol.* 63, 2043–2057. doi: 10.1111/jipb.13174

Long, Y., Wang, C., Liu, C., Li, H., Pu, A., Dong, Z., et al. (2024). Molecular mechanisms controlling grain size and weight and their biotechnological breeding applications in maize and other cereal crops. *J. Advanced Res.* 62, 27–46. doi: 10.1016/j.jare.2023.09.016

Lu, L., Tian, S., Liao, H., Zhang, J., Yang, X., Labavitch, J. M., et al. (2013). Analysis of Metal Element Distributions in Rice (Oryza sativa L.) Seeds and Relocation during Germination Based on X-Ray Fluorescence Imaging of Zn, Fe, K, Ca, and Mn. *PloS One* 8, e57360. doi: 10.1371/journal.pone.0057360

Luo, J., Liu, H., Zhou, T., Gu, B., Huang, X., Shangguan, Y., et al. (2013). An-1 encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. *Plant Cell* 25, 3360–3376. doi: 10.1105/tpc.113.113589

Ma, B., Zhang, L., and He, Z. (2023). Understanding the regulation of cereal grain filling: The way forward. *J. Integr. Plant Biol.* 65, 526–547. doi: 10.1111/jipb.13456

Mandha, J., and Raes, K. (2023). ""Chapter 10 - minerals and phytates,"," in *ICC handbook of 21st century cereal science and technology*. Eds. P. R. Shewry, H. Koksel and J. R. N. Taylor (Academic Press), 97–102. doi: 10.1016/B978-0-323-95295-8.00034-4

Mao, H., Sun, S., Yao, J., Wang, C., Yu, S., Xu, C., et al. (2010). Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl. Acad. Sci.* 107, 19579–19584. doi: 10.1073/pnas.1014419107

Maruta, N., Trusov, Y., Jones, A. M., and Botella, J. R. (2021). Heterotrimeric G proteins in plants: canonical and atypical Gα Subunits. *Int. J. Mol. Sci.* 22, 11841. doi: 10.3390/ijms222111841

Mengie, Y., Assefa, A., and Jenber, A. J. (2021). Sowing methods and seeding rates effects on yield and yield components of Tef(*Eragrostis tef* [Zucc.] Trotter) at Adet, North West Ethiopia. *Heliyon* 7, e06519. doi: 10.1016/j.heliyon.2021.e06519

Meyer, R. S., and Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* 14, 840–852. doi: 10.1038/nrg3605

Miao, J., Yang, Z., Zhang, D., Wang, Y., Xu, M., Zhou, L., et al. (2019). Mutation of RGG2, which encodes a type B heterotrimeric G protein γ subunit, increases grain size and yield production in rice. *Plant Biotechnol. J.* 17, 650–664. doi: 10.1111/pbi.13005

Miller, D. (2009).Teff Grass: A new alternative. In: *2009 western alfalfa & Forage symposium., in* california alfalfa & Forage symposium (Reno, NV: University of California Cooperative Ext. Davis, CA). Available online at: https://alfalfasymposium.ucdavis.edu/+symposium/2009/ (Accessed June 17, 2024).

Mohammed, M., Mustafa, A., and Osman, G. (2009). Evaluation of wheat breads supplemented with Teff (Eragrostis tef (ZUCC.) Trotter) Grain flour. *Aust. J. Crop Sci.* 3, 207–212.

Mohapatra, P. K., and Sahu, B. B. (2022). ""Fertilization and seed development in rice,"," in *Panicle architecture of rice and its relationship with grain filling*. Eds. P. K. Mohapatra and B. B. Sahu (Springer International Publishing), 63–86, Cham). doi: 10.1007/978-3-030-67897-5_4

Moon, J., Parry, G., and Estelle, M. (2004). The ubiquitin-proteasome pathway and plant development. *Plant Cell* 16, 3181–3195. doi: 10.1105/tpc.104.161220

Naylor, R. L., Falcon, W. P., Goodman, R. M., Jahn, M. M., Sengooba, T., Tefera, H., et al. (2004). Biotechnology in the developing world: a case for increased investments in orphan crops. *Food Policy* 29, 15–44. doi: 10.1016/j.foodpol.2004.01.002

Nyachoti, S., Adebayo, S., and Godebo, T. R. (2021). Elemental composition of teff (a gluten-free grain), maize and wheat: Staple crops in the Main Ethiopian Rift Valley. *J. Food Compost Anal.* 100, 103660. doi: 10.1016/j.jfca.2020.103660

Pandey, S. (2020). Plant receptor-like kinase signaling through heterotrimeric G-proteins. *J. Exp. Bot.* 71, 1742–1751. doi: 10.1093/jxb/eraa016

Pearson, W. R. (2013). An introduction to sequence similarity ("Homology") searching. *Curr. Protoc. Bioinf.* 42, 3.1.1–3.1.8. doi: 10.1002/0471250953.bi0301s42

Purugganan, M. D. (2019). Evolutionary insights into the nature of plant domestication. *Curr. Biol.* 29, R705–R714. doi: 10.1016/j.cub.2019.05.053

Rajpal, V. R., Singh, A., Kathpalia, R., Thakur, R., Mohd., K., Pandey, A., et al. (2023). The Prospects of gene introgression from crop wild relatives into cultivated lentil for climate change mitigation. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1127239

Razzaque, S., Heckman, R. W., and Juenger, T. E. (2023). Seed size variation impacts local adaptation and life-history strategies in a perennial grass. *Proc. Biol. Sci.* 290, 20222460. doi: 10.1098/rspb.2022.2460

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Sheng, H., Jiang, Y., Rahmati, M., Chia, J.-C., Dokuchayeva, T., Kavulych, Y., et al. (2021). YSL3-mediated copper distribution is required for fertility, seed size and protein accumulation in Brachypodium. *Plant Physiol.* 186, 655–676. doi: 10.1093/plphys/kiab054

Shewry, P. R., Wan, Y., Hawkesford, M. J., and Tosi, P. (2020). Spatial distribution of functional components in the starchy endosperm of wheat grains. *J. Cereal Sci.* 91, 102869. doi: 10.1016/j.jcs.2019.102869

Shi, C., Ren, Y., Liu, L., Wang, F., Zhang, H., Tian, P., et al. (2019). Ubiquitin specific protease 15 has an important role in regulating grain width and size in rice. *Plant Physiol.* 180, 381–391. doi: 10.1104/pp.19.00065

Shumoy, H., Gabaza, M., Vandevelde, J., and Raes, K. (2017). Soluble and bound phenolic contents and antioxidant capacity of tef *injera* as affected by traditional fermentation. *J. Food Composition Anal.* 58, 52–59. doi: 10.1016/j.jfca.2017.01.004

Shumoy, H., and Raes, K. (2017). Tef: The Rising Ancient Cereal: What do we know about its Nutritional and Health Benefits? *Plant Foods Hum. Nutr.* 72, 335–344. doi: 10.1007/s11130-017-0641-2

Smieska, L., Page, K. A., Ree, B., Zheng, B., Koerner, H., and Woll, A. R. (2023). The functional materials beamline at CHESS. *Synchrotron Radiat. News* 36, 4–11. doi: 10.1080/08940886.2023.2202578

Spaenij-Dekking, L., Yvonne, K.-W., and Frits, K. (2005). The Ethiopian cereal tef in celiac disease. *New Engl. J. Med.* 353, 1748–1749. doi: 10.1056/NEJMc051492

Stoupin, S., Krawczyk, T., Sagan, D., Temnykh, A., Smieska, L., Woll, A., et al. (2021). Side-bounce beamlines using single-reflection diamond monochromators at Cornell High Energy Synchrotron Source. *J. Synchrotron Radiat.* 28, 429–438. doi: 10.1107/S160057752100120X

Su, L., Wan, S., Zhou, J., Shao, Q. S., and Xing, B. (2021). Transcriptional regulation of plant seed development. *Physiologia Plantarum* 173, 2013–2025. doi: 10.1111/ppl.13548

Sun, S., Wang, L., Mao, H., Shao, L., Li, X., Xiao, J., et al. (2018). A G-protein pathway determines grain size in rice. *Nat. Commun.* 9, 851. doi: 10.1038/s41467-018-03141-y

Tadele, E., and Hibistu, T. (2021). Empirical review on the use dynamics and economics of teff in Ethiopia. *Agric. Food Secur.* 10, 40. doi: 10.1186/s40066-021-00329-2

Taj, G., Agarwal, P., Grant, M., and Kumar, A. (2010). MAPK machinery in plants. *Plant Signal Behav.* 5, 1370–1378. doi: 10.4161/psb.5.11.13020

Takahashi, M., Nozoye, T., Kitajima, N., Fukuda, N., Hokura, A., Terada, Y., et al. (2009). *In vivo* analysis of metal distribution and expression of metal transporters in rice seed during germination process by microarray and X-ray Fluorescence Imaging of Fe, Zn, Mn, and Cu. *Plant Soil* 325, 39–51. doi: 10.1007/s11104-009-0045-7

Talabi, A. O., Vikram, P., Thushar, S., Rahman, H., Ahmadzai, H., Nhamo, N., et al. (2022). Orphan crops: A best fit for dietary enrichment and diversification in highly deteriorated marginal environments. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.839704

Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120

Tang, S., Zhao, Z., Liu, X., Sui, Y., Zhang, D., Zhi, H., et al. (2023). An E2-E3 pair contributes to seed size control in grain crops. *Nat. Commun.* 14, 3091. doi: 10.1038/s41467-023-38812-y

Tao, Y., Miao, J., Wang, J., Li, W., Xu, Y., Wang, F., et al. (2020a). RGG1, involved in the cytokinin regulatory pathway, controls grain size in rice. *Rice* 13, 76. doi: 10.1186/s12284-020-00436-x

Tao, Y., Zhao, X., Wang, X., Hathorn, A., Hunt, C., Cruickshank, A. W., et al. (2020b). Large-scale GWAS in sorghum reveals common genetic control of grain size among cereals. *Plant Biotechnol. J.* 18, 1093–1105. doi: 10.1111/pbi.13284

Tefera, H., Assefa, K., and Belay, G. (2003). Evaluation of interspecific recombinant inbred lines of Eragrostis tef x E. pilosa. *J. Genet. Breed.* 57, 21–30.

Tian, X., He, M., Mei, E., Zhang, B., Tang, J., Xu, M., et al. (2021). WRKY53 integrates classic brassinosteroid signaling and the mitogen-activated protein kinase pathway to regulate rice architecture and seed size. *Plant Cell* 33, 2753–2775. doi: 10.1093/plcell/koab137

Urano, D., and Jones, A. M. (2014). Heterotrimeric G protein–coupled signaling in plants. *Annu. Rev. Plant Biol.* 65, 365–384. doi: 10.1146/annurev-arplant-050213-040133

VanBuren, R., Man Wai, C., Wang, X., Pardo, J., Yocca, A. E., Wang, H., et al. (2020). Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. *Nat. Commun.* 11, 884. doi: 10.1038/s41467-020-14724-z

Vavilov, N. I. (1951). The origin, variation, immunity and breeding of cultivated plants. *Soil Sci.* 72, 482. doi: 10.1097/00010694-195112000-00018

Venable, D. L. (1992). Size-number trade-offs and the variation of seed size with plant resource status. *Am. Nat.* 140, 287–304. doi: 10.1086/285413

Wang, J., Wang, R., Mao, X., Zhang, J., Liu, Y., Xie, Q., et al. (2020). RING finger ubiquitin E3 ligase gene TaSDIR1-4A contributes to determination of grain size in common wheat. *J. Exp. Bot.* 71, 5377–5388. doi: 10.1093/jxb/eraa271

Woldetsadik, D., Sims, D. B., Garner, M. C., and Hailu, H. (2024). Metal(loid)s profile of four traditional Ethiopian teff brands: geographic origin discrimination. *Biol. Trace Elem Res.* 202, 1305–1315. doi: 10.1007/s12011-023-03736-7

Wu, W., Liu, X., Wang, M., Meyer, R. S., Luo, X., Ndjiondjop, M.-N., et al. (2017). A single-nucleotide polymorphism causes smaller grain size and loss of seed shattering during African rice domestication. *Nat. Plants* 3, 1–7. doi: 10.1038/nplants.2017.64

Xiao, Y., Liu, D., Zhang, G., Gao, S., Liu, L., Xu, F., et al. (2019). Big Grain3, encoding a purine permease, regulates grain size via modulating cytokinin transport in rice. *J. Integr. Plant Biol.* 61, 581–597. doi: 10.1111/jipb.12727

Xu, R., Duan, P., Yu, H., Zhou, Z., Zhang, B., Wang, R., et al. (2018). Control of grain size and weight by the osMKKK10-osMKK4-osMAPK6 signaling pathway in rice. *Mol. Plant* 11, 860–873. doi: 10.1016/j.molp.2018.04.004

Yang, W., Wu, K., Wang, B., Liu, H., Guo, S., Guo, X., et al. (2021). The RING E3 ligase CLG1 targets GS3 for degradation via the endosome pathway to determine grain size in rice. *Mol. Plant* 14, 1699–1713. doi: 10.1016/j.molp.2021.06.027

Zhang, D., Li, J., Compton, R. O., Robertson, J., Goff, V. H., Epps, E., et al. (2015). Comparative genetics of seed size traits in divergent cereal lineages represented by sorghum (Panicoidae) and rice (Oryzoidae). *G3 Genes|Genomes|Genetics* 5, 1117–1128. doi: 10.1534/g3.115.017590

Zhang, M., Li, Y., Liang, X., Lu, M., Lai, J., Song, W., et al. (2023). A teosinte-derived allele of an HKT1 family sodium transporter improves salt tolerance in maize. *Plant Biotechnol. J.* 21, 97–108. doi: 10.1111/pbi.13927

Zhang, J.-P., Yu, Y., Feng, Y.-Z., Zhou, Y.-F., Zhang, F., Yang, Y.-W., et al. (2017b). MiR408 regulates grain yield and photosynthesis via a phytocyanin protein1. *Plant Physiol.* 175, 1175–1185. doi: 10.1104/pp.17.01169

Zhang, H., Zhang, J., Yan, J., Gou, F., Mao, Y., Tang, G., et al. (2017a). Short tandem target mimic rice lines uncover functions of miRNAs in regulating important agronomic traits. *Proc. Natl. Acad. Sci. U.S.A.* 114, 5277–5282. doi: 10.1073/pnas.1703752114

Zhao, Y., Shi, Y., Jiang, G., Wu, Y., Ma, M., Zhang, X., et al. (2022). Rice extra-large G proteins play pivotal roles in controlling disease resistance and yield-related traits. *New Phytol.* 234, 607–617. doi: 10.1111/nph.17997

Zimmermann, M. B., and Hurrell, R. F. (2007). Nutritional iron deficiency. *Lancet* 370, 511–520. doi: 10.1016/S0140-6736(07)61235-5

Zombori, Z., Nagy, B., Mihály, R., Pauk, J., Cseri, A., Sass, L., et al. (2020). RING-type E3 ubiquitin ligase barley genes (HvYrg1–2) control characteristics of both vegetative organs and seeds as yield components. *Plants (Basel)* 9, 1693. doi: 10.3390/plants9121693