Nutri-cereal tissue-specific transcriptome atlas during development: functional integration of gene expression to identify mineral uptake pathways in little millet (Panicum sumatrense)

Pahari, Shankar; Vaid, Neha; Soolanayakanahally, Raju; Kagale, Sateesh; Pasha, Asher; Esteban, Eddi; Provart, Nicholas; Stobbs, Jarvis A.; Vu, Miranda; Meira, Debora; Karunakaran, Chithra; Boda, Praveen; Prasannakumar, Mothukapalli K.; Nagaraja, Alur; Jain, Ashwani Kumar

Publisher's version / Version de l'éditeur:
https://doi.org/10.1111/tpj.16749
The Plant Journal, 2024-04-04

the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2024) doi: 10.1111/tpj.16749

RESOURCE

Nutri-cereal tissue-specific transcriptome atlas during development: Functional integration of gene expression to identify mineral uptake pathways in little millet (*Panicum sumatrense*)

Shankar Pahari[1], Neha Vaid[2], Raju Soolanayakanahally[1,*], Sateesh Kagale[3], Asher Pasha[4], Eddi Esteban[4], Nicholas Provart[4], Jarvis A. Stobbs[5], Miranda Vu[5], Debora Meira[6], Chithra Karunakaran[5], Praveen Boda[7], Mothukapalli K. Prasannakumar[7], Alur Nagaraja[7] and Ashwani Kumar Jain[8]

[1]*Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada, Saskatoon, Saskatchewan, Canada,*
[2]*Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada,*
[3]*Aquatic and Crop Resource Development, National Research Council Canada, Saskatoon, Saskatchewan, Canada,*
[4]*Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada,*
[5]*Canadian Light Source Inc,Saskatoon, Saskatchewan, Canada,*
[6]*Advanced Photon Source, Argonne National Laboratory, Argonne,IL, United States,*
[7]*Department of Plant Pathology, University of Agricultural Sciences, Bangalore, India, and*
[8]*College of Agriculture, Rewa, Madhya Pradesh, India*

Received 2 September 2022; revised 8 March 2024; accepted 14 March 2024.
*For correspondence (e-mail raju.soolanayakanahally@agr.gc.ca).

SUMMARY

Little millet (*Panicum sumatrense* Roth ex Roem. & Schult.) is an essential minor millet of southeast Asia and Africa's temperate and subtropical regions. The plant is stress-tolerant, has a short life cycle, and has a mineral-rich nutritional profile associated with unique health benefits. We report the developmental gene expression atlas of little millet (genotype JK-8) from ten tissues representing different stages of its life cycle, starting from seed germination and vegetative growth to panicle maturation. The developmental transcriptome atlas led to the identification of 342 827 transcripts. The BUSCO analysis and comparison with the transcriptomes of related species confirm that this study presents high-quality, in-depth coverage of the little millet transcriptome. In addition, the eFP browser generated here has a user-friendly interface, allowing interactive visualizations of tissue-specific gene expression. Using these data, we identified transcripts, the orthologs of which in *Arabidopsis* and rice are involved in nutrient acquisition, transport, and response pathways. The comparative analysis of the expression levels of these transcripts holds great potential for enhancing the mineral content in crops, particularly zinc and iron, to address the issue of "hidden hunger" and to attain nutritional security, making it a valuable asset for translational research.

Keywords: little millet, developmental transcriptome, mineral ion transport, differential gene expression, zinc and iron.

INTRODUCTION

The global population of 7.6 billion depends on rice, wheat, and maize for 42.5% of their daily caloric requirements (Food and Agriculture Organization of the United Nations, 2018). While the green revolution-led productivity increase decreased global malnutrition from 37% to 12% (Food and Agriculture Organization of the United Nations, 2018), yield enhancement has correlated with a decline in the nutritional quality of cereals (Fan et al., 2008). Additionally, high-yielding crops, such as rice and wheat, have replaced nutrient-rich but less profitable minor

cereals. At present, 25% of the world's population suffers from micronutrient, vitamin, and mineral deficiencies, referred to as "*hidden hunger*" (Food and Agriculture Organization of the United Nations, 2018). Vitamins and minerals serve as essential co-factors for several enzymatic reactions that regulate average growth and body functions, and their deficiencies are known to affect women and children disproportionately. Supplementation of diet with nutritionally superior mineral-rich minor cereals from the *Poaceae* family, such as rye, millet, oats, and sorghum, offers a viable solution to mitigate hidden hunger (Vetriventhan et al., 2020).

Minor millets are annual grasses with tiny seeds grown mainly in Asia and Africa. They include pearl millet (*Pennisetum glaucum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine coracana*), kodo millet (*Paspalum scrobiculatum*), barnyard millet (*Echinochloa crusgalli*), and little millet (*Panicum sumatrense*). The millets are characterized by seeds that are individually attached to the main stem by stalk instead of being enclosed in ears like the major cereals (Taylor & Kruger, 2016). Minor millets are gluten-free grain and rich sources of calories, fiber, proteins, antioxidants, micronutrients, phenolic phytochemicals, essential vitamins, and minerals (Goron & Raizada, 2015). Due to their unique nutritional profile, minor millets have been categorized as "*smart-food crops*" and "*nutri-cereals*" (Vetriventhan et al., 2020). These cereals also require low agricultural inputs, such as water and fertilizers, and are resistant to crop diseases and abiotic stresses (Amadou et al., 2013; Habiyaremye et al., 2017). These traits make minor millets most suitable for marginal agricultural lands or under environmentally harsh conditions, where major cereals rice, wheat, and maize show poor productivity. Identification of their nutritional and agricultural value along with a recent drive to preserve ancient crop germplasms and genetic diversity has fueled an immense research interest in these species.

Little millet is among the least studied crop (Johnson et al., 2019). Little millet was domesticated in India and is grown in temperate and tropical regions of India, China, East Asia, and Malaysia (Kalaisekar et al., 2017). Genetically, little millet is majorly a tetraploid species ($2n = 36$) (Hamoud et al., 1994; Saha et al., 2016), although a hexaploid variety has also been reported ($2n = 54$) (Chen & Renvoize, 2006). The crop shows a high degree of adaptability in diverse growth conditions (Nirmalakumari et al., 2010) and is resistant to pests, drought, waterlogging, and salinity stress (Ajithkumar & Panneerselvam, 2014; Bhaskaran & Panneerselvam, 2013; Sivakumar et al., 2006). Nutritionally, little millet is a rich source of fats, protein, iron, zinc, flavonoids, and phenolic acids (Pradeep & Guha, 2011; Selvi et al., 2015; Vetriventhan et al., 2020). The seed has a low glycemic index with the highest dietary fiber content in cereals, making it an ideal grain for the diabetic population (Kumar et al., 2018). The crop, however, suffers from lower productivity than significant cereal crops (Plaza-Wüthrich & Tadele, 2012).

Multi-omic resources are essential for the fundamental understanding of crops' biology and for devising crop improvement strategies. For example, transcriptomic analyses enabled the identification of drought tolerance-associated genes and subsequent development of breeding strategies for pearl millet (Dudhate et al., 2018; Serba & Yadav, 2016). For little millet, apart from chloroplast genome sequence (Sebastin et al., 2018), and genomically uncharacterized collection of over 450 accessions (Upadhyaya et al., 2014), phenotypic, genetic, and molecular data are unavailable (Johnson et al., 2019). The lack of little millet transcriptome has severely hindered crop improvement strategies. It has restricted the identification and incorporation of its unique genetics for highly sought-after agronomic traits, such as stress tolerance, low fertilizer and irrigation needs, and high seed mineral content in the ongoing breeding programs. Building a spatio-temporal gene expression atlas of little millet is the first and essential step in uncovering the genetic networks that provide this crop with its unique nutritional and stress-resilient features.

Here, we have generated the transcriptome of little millet across ten distinct tissue types encompassing the crop's entire life cycle. We have identified tissue- and life cycle-specific gene expression patterns that can be utilized for future molecular studies. Additionally, by harnessing transcriptome data, we have highlighted potential mineral transporter genes that might play a role in shaping the crop's impressive mineral profile. Moreover, we delved into the evolutionary connections between little millet and other minor and major cereals.

RESULTS

Despite a great potential for a hardy cereal crop, little millet is severely understudied. This study was undertaken to generate a growth stage-specific transcriptome atlas of little millet as an information resource to enable gene discovery and breed improved varieties.

Transcriptome sequencing and quality assessment

The little millet (genotype JK-8) developmental transcriptome was built with ten tissue types, representing three growth phases in its life cycle; emergence phase [germinating seeds (GS), radicle (RD), plumule (PU)], vegetative phase [young leaf (YL), young root (YR), crown meristem (CM), vegetative stem (VS)] and reproductive phase [early panicle (PE), mid panicle (PM), late panicle (PL)] (Figure 1, Table S1). The experimental workflow deployed for the little millet growth stage-specific transcriptome analysis is illustrated in Figure S1. The Illumina sequencing platform generated 325.4 million paired-end raw reads from 28

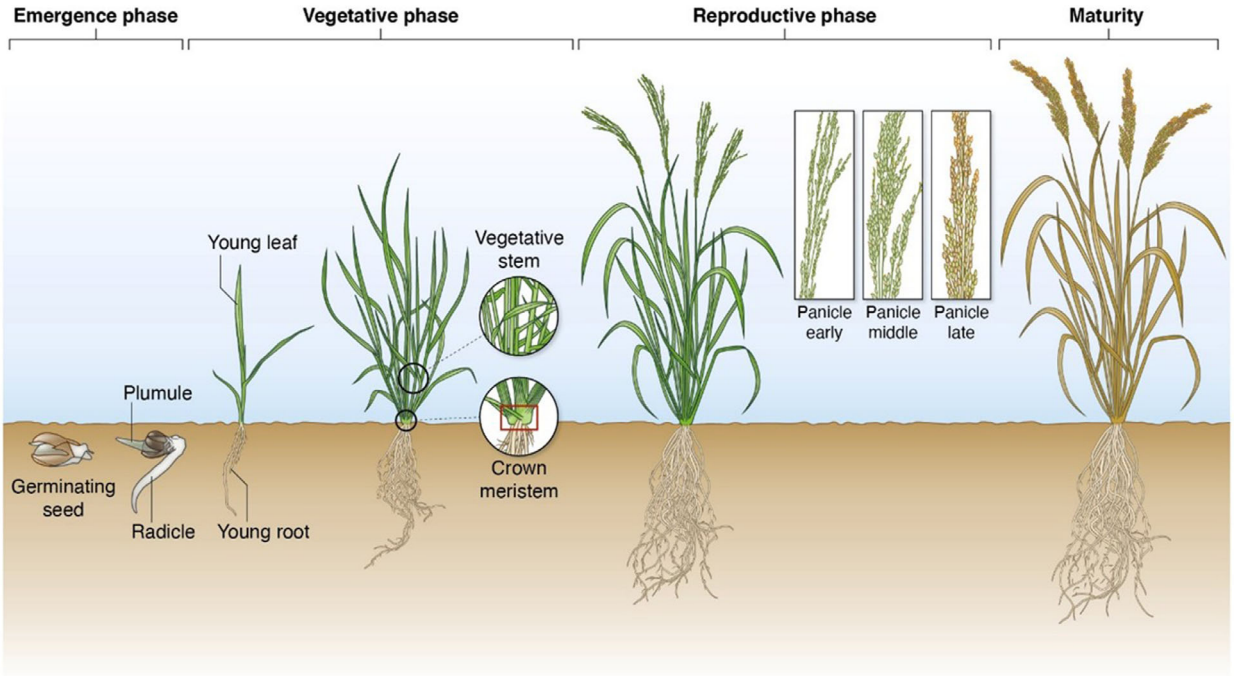

Figure 1. The ten tissue samples used to represent three growth phases of little millet (genotype JK-8).
Germinating seed (GS), radicle (RD), and plumule (PU) represent the emergence phase; young leaf (YL), young root (YR), crown meristem (CM), and vegetative stem (VS) represent the vegetative phase; and panicle early (PE), panicle mid (PM), and panicle late (PL) represent the reproductive phase. The life cycle of the crop is completed in approximately 90 days (Artwork by Debbie Maizels, Zoobotanica Scientific Illustration).

libraries, accounting for ~82 GB of sequencing data (Table S1). The sequences were filtered to remove adaptor sequences and low-quality and ambiguous reads to provide 258.8 million paired reads (79.5% of total reads) suitable for downstream analysis. Due to the lack of a reference genome for little millet, a reference transcriptome was built using *de novo* Trinity V 2.9.1 assembler (Grabherr et al., 2011). Trinity assembly yielded a total of 342 827 transcripts (Data S1). The coverage assessment of the assembled transcriptome showed a representation of approximately 86% of the input RNA-seq reads, suggesting that the majority of transcriptionally active genes have been captured in our assembly. Further, a survey of 3236 orthologs in the Benchmarking Universal Single-Copy Orthologs (BUSCO) set of Liliopsida (odb10 database) to assess the quality of the assembled transcriptome and annotation coverage showed that 92.3% of transcripts were recovered completely. This comprised 23.4% and 68.9% of complete single and duplicate copies of BUSCOs, respectively. Of the remaining 7.7%, our search identified 5.3% fragmented and 2.4% missing BUSCOs.

Phylogenetic analysis

For phylogenetic comparison, the orthologs of the transcript set were searched in small-grained dicot (quinoa) and monocots (ragi, foxtail millet, broom corn, pearl millet, and little millet), major cereals (rice and maize), and outgroup dicot species (*Arabidopsis* and alfalfa) due to their proximity to the little millet and their potential to provide insights into specific aspects of mineral nutrition (Figure 2a). Comparison of the gene coding sequences from closely related species produced a concatenated alignment of 3 475 812 base pairs from a supermatrix of 2752 orthologous genes. This alignment was used to define phylogenetic relationships of these species with little millet. The tree suggested that proso millet (*Panicum miliaceum*) is the closest relative to little millet, followed by pearl millet (*Pennisetum glaucum*) and foxtail millet (*Setaria italica*) (Figure 2a). The substitution per site rate (Ks) was calculated by comparing pairs of homologous genes within the little millet genome (Figure 2b). The Ks analysis revealed the peak at 0.049 corresponding to the latest polyploidy event which occurred at ~3.53 million years ago.

Global view of the transcriptome

Prior to assessing little millet transcriptome dynamics, the reproducibility of generated data was assessed using MA plots and scatter plots (Figure S2). MA plots showed that most log-ratios (representing individual genes) were clustered close to zero on the y-axis for replicates of all tissue samples (Figure S2). Similarly, scatter plots showed that most transcripts in biological replicates fall on the $x = y$ line (shown as black dots in Figure S2), indicating high

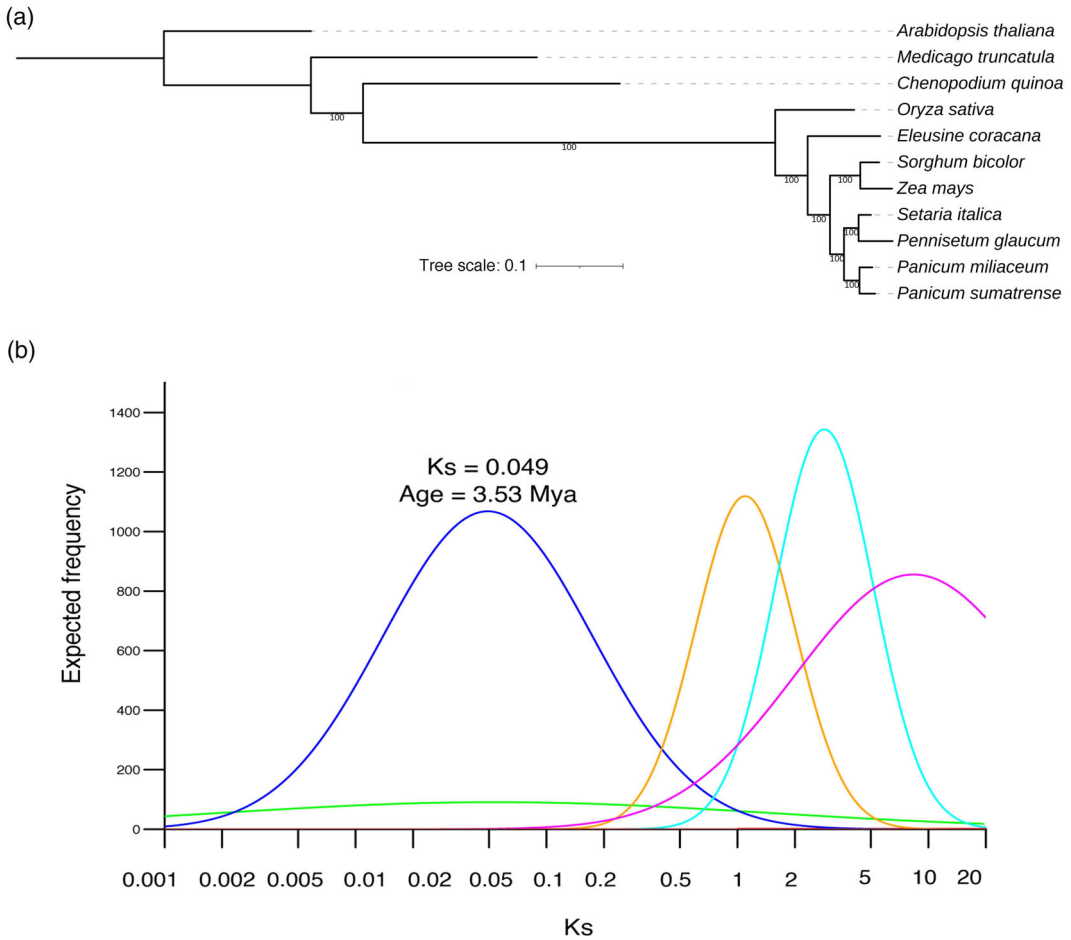

Figure 2. Phylogenetics and polyploidy of little millet.
(a) A maximum likelihood tree depicting evolutionary relationships among little millet and its closely related species was generated based on sequences of 2752 orthologous genes resulting in a concatenated alignment of 3 475 812 base pairs. The tree was visualized using the Interactive Tree of Life (iTOL) Web server. Clade support values (100) positioned next to nodes indicate complete support, and the branch lengths on the tree represent the estimated nucleotide substitutions per site.
(b) Mixture models fitted to Gaussian components in the histograms of frequency distributions of Ks values obtained by comparing pairs of homologous genes within the little millet genome.

reproducibility in the biological replicates. The variability between the replicates and tissue samples was further visualized with distribution plot of log-transformed transcript expression values of all samples (Figure S3). The distribution of expression values showed subtle difference in transcript counts while being consistent with respect to replicates within a tissue.

The relationship between the tissue samples and the replicates was visualized using correlation heatmap. The tissues were found to predominantly cluster according to their function (Figure S4a). For example, RD and YR clustered together and showed a low correlation with most of the aboveground tissues. The tissues of the reproductive phase (PE, PM, PL) showed a higher correlation with each other (Figure S4a). The transcriptional relationship between the tissue samples and their replicates was further ascertained by reducing the TPM expression data into a graphical two-dimensional principal component analysis (PCA) plot. The PCA plot showed a clear separation of the belowground (GS, RD, YR) and the aboveground samples along the first principal component (PC1) with 45% variance (Figure S4b). Similarly, PM tissues were scattered between PE and PL on the second principal component axis, indicating that this tissue captures the transition between the two stages. The samples VS and YL were grouped with PE, PU, and CM and separated from the underground tissues and late-stage reproductive tissues (PM and PL) along PC1 and PC2, respectively (Figure S4b).

Dynamics of gene expression

To streamline and understand gene expression dynamics in different tissues of little millet, the transcripts with TPM

>0.50 in at least two replicates of each sample were filtered as being expressed. The expressed transcripts were further categorized as expressed (TPM 0.5–10) and highly expressed (TPM >10) transcripts (Figure 3a). In our dataset, 214 796 transcripts (Data S2), corresponding to 63% of the assembled transcripts, were expressed in at least one tissue type (Figure 3b). The total number of expressed transcripts in different tissues ranged from 105 050 in CM to 30 216 in YR. Overall, the aboveground tissues (PU, CM, PE, PM, and PL) showed higher number of expressed transcripts compared to belowground tissues, RD and YR (Figure 3b).

Transcripts specific to an individual tissue can help understand specialized tissue specific processes. In our analysis, CM showed the highest number of tissue specific transcripts (i.e., 11.9%; 12 470 genes) followed by PU with 11%. The fewest number of uniquely expressed transcripts was found in YR and PL, accounting for 6.4% and 7.3% of the expressed transcripts, respectively (Figure 3b). Heatmap constructed using the $\log_2$-transformed TPM values to depict transcript abundances across different tissues indicating their tissue specificities is presented in Figure S5a. Further, 12 922 transcripts, representing 6% of the total expressed transcripts, were expressed in all tissues suggesting their housekeeping function (Figure 3b). A list of constitutively and uniquely expressed transcripts identified in this study has been provided in Data S3. Of the constitutively expressed transcripts, 26 transcripts had a coefficient of variance ≤6% across all tissues, indicating their relatively stable expression (Figure S5b). This list presents candidates that can be used as reference transcripts for gene expression analysis in little millet (Data S4). Of the 26 transcripts, *Arabidopsis* and rice orthologs were annotated for 15 transcripts, several of which encode component of essential cell functions such as *ubiquitin-conjugating enzyme, dehydrogenases, and kinases* (Figure S5c).

Among the expressed transcripts, 45 359 unique transcripts were annotated based on their orthologs in *Arabidopsis* and assigned GO terms for the biological process, molecular function, and cellular component categories (Figure S6). Biological process category had highest number of genes represented under metabolic processes (9343) and biosynthetic processes (4833) followed by anatomical structure development (3723) and response to stress (3717). The molecular function category showed the highest representation of protein binding (4024), catalytic

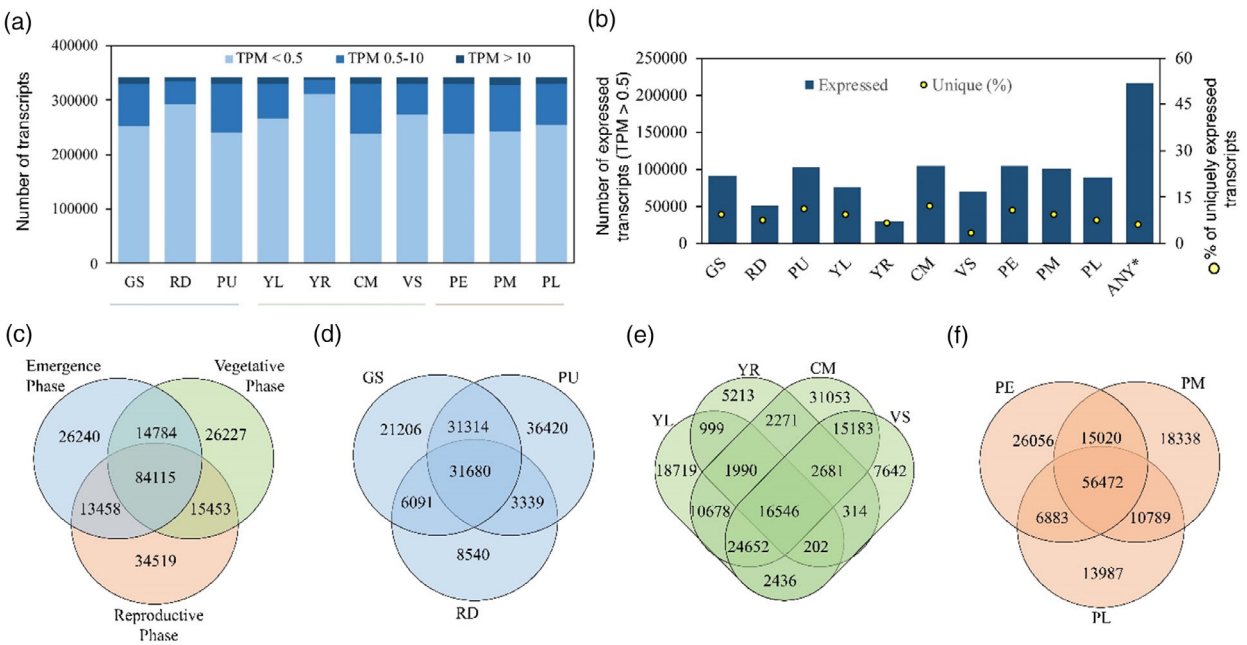

Figure 3. Dynamics of transcript expression in little millet.
(a) The number of transcripts expressed at different levels based on normalized TPM values in 10 tissues during the three growth stages. The bars indicate the number of transcripts expressed in each sample. Transcripts were categorized based on their TPM values in at least 2 replicates: (i) no expression (TPM <0.5), (ii) expressed (0.5 < TPM < 10), and (iii) highly expressed (TPM > 10).
(b) The number of transcripts categorized as expressed or highly expressed in 3a for each tissue is presented as bars, and % of transcripts specifically expressed in the tissue is presented as an yellow dot on the secondary y-axis. The bar "Any*" represents the total number of transcripts expressed in at least one of the ten tissues used in the study. The yellow dot for Any* represents the % of expressed transcripts present in all ten tissues.
(c-f) Venn diagram of the number of transcripts common and unique across three growth phases (c), across three tissues of emergence phase (d), four tissues of vegetative phase (e), and three tissues of reproductive phase (f). Transcripts with a normalized expression level TPM >0.5 in at least one of the 10 tissues analyzed were log-transformed before analysis. Tissue samples in (a) and (b) are color underlined representing their growth phases as in (c).

activity (3619), transferase activity (2577), and hydrolase activity (1972). The cellular component included nucleus (6709), cytoplasm (3775), chloroplast (3495), and mitochondrion (2436) as top represented categories. Transcription factors (TF) are the key regulators of all cellular functions. From the *Arabidopsis* orthologs of annotated little millet transcripts, 1103 transcription factors belonging to 49 families were identified (Figure S6d). These included C2H2 (119 genes), bHLH (116 genes), C3H (94 genes), and MYB (90 genes). Analysis of tissue-specific expression of a TF showed that 122 genes were expressed in GS belonging to 25 families, followed by PU with 89 genes in 23 families (Data S5).

A growth stage-wise comparison identified 84 115 transcripts expressed in all growth stages, while 26 240, 26 227, and 34 519 transcripts were specifically expressed in the emergence, vegetative, and reproductive phases, respectively (Figure 3c). Within the emergence phase, 31 680 transcripts were shared between GS, RD, and PU (Figure 3d). The four tissues of vegetative phase (YL, YR, CM, and VS) shared 16 546 transcripts (Figure 3e). Similarly, 56 472 transcripts were shared among 3 tissues of reproductive phase (Figure 3f).

Analysis of differentially expressed genes

The pairwise comparison of transcript expression between the studied tissue samples was statistically tested using ANOVA and adjusted for FDR (Data S2). Altogether, 75 887 transcripts with P_{adj} <0.01 were identified as differentially expressed transcripts and were used for the determination of the optimal number of clusters based on similarity in their expression patterns (Figure S7). Due to the high correlation among the biological replicates within all tissues (R^2 > 0.86, Table S2), the transcript expression values [$\log_2$ (TPM + 1)] of replicates were averaged for cluster analyses. Further, Z-score was calculated as a measure of standard deviations for each transcript's expression in a specific tissue from its mean expression across all tissues and plotted as a heatmap to visualize the expression patterns (Figure 4a). Owing to the maximum number of transcripts allowed by complex heatmap algorithm, 60 000 transcripts with lowest P_{adj} values (Data S2) were used. The analysis showed that cluster 1 grouped transcripts with high expression in PU, CM, and VS. Cluster 2 comprised of transcripts highly expressed in reproductive phase (PE, PM, PL), cluster 3 categorized highly expressed transcripts in RD and YR, and cluster 4 represented transcripts with relatively high expression in YL (Figure 4a). The boxplot depicting the average Z-score of all genes in each tissue is shown in the lower panel of Figure 4a, and the overall expression pattern of transcript in each tissue within each cluster is shown as boxplot in Figure 4b.

The functional categorization of genes within each cluster was conducted using GO term overrepresentation test. Of the 20 563, 16 488, 10 605, and 12 343 transcripts in the four clusters, *Arabidopsis* orthologs were annotated for 8658, 4596, 2642, and 4335 transcripts. The list of overrepresented GO terms for biological process, molecular function, and cellular component categories for each cluster is highlighted in Figure 4c, and the complete list of corresponding data with fold enrichment and *P*-values is provided in Data S6. In brief, GO terms overrepresented in cluster 1 include regulation of metabolic and developmental processes, embryo development, zinc ion binding, and ATP hydrolysis activity. Cluster 2 showed overrepresentation of GO terms such as floral organ development, pollen tube growth, vegetative to reproductive transition of meristem, and calcium ion transport, while cluster 3 represented the belowground tissues (RD, YR) and showed overrepresentation of GO terms including response to oxidative stress, response to water deprivation, root morphogenesis, and heme binding. Cluster 4 comprised the most photosynthetically active tissue (YL) and correspondingly showed overrepresentation of GO terms associated with photosynthesis, metal ion transport, and phosphatase activity (Figure 4c, Data S6).

Pairwise differential gene expression

To simplify pairwise comparison of the transcripts between each tissues, the tissue samples were clustered into four major tissue groups based on hierarchical clustering dendrogram (Figure S8a). Group 1 consisted of GS, PU, VS, and CM, group 2 with PE, PM, and PL; group 3 with RD and YR; and group 4 with YL. Pairwise comparisons of gene expression were performed between all 4 groups from the transcript abundance counts data (Data S2) that had P_{adj} from the ANOVA test <0.05 (Data S2) to identify differentially expressed genes using DESeq2. Genes were considered to be differentially expressed when P_{adj} <0.01 and $\log_2$FC >2. The overall distribution of the differentially expressed genes (DEGs) for each group comparison is depicted with enhanced volcano maps in Figure S8b.

Briefly, comparison of group 3 (constituting the below-ground tissues) with all other groups showed the highest numbers of DEGs, with group 3 vs group 4 (below-ground tissues vs. photosynthetically active tissues) showing the highest number of DEGs (25 042; 10 316 up-regulated and 14 726 downregulated) (Data S7). The comparison of group 1 vs group 2 (actively growing vs. reproductive tissues) showed the least number of DEGs 8842 (3010 up-regulated and 5832 downregulated) (Figure S8b). The genes showed a wide range of differential expression from a 35-fold increase to 38-fold decrease in expression. We filtered a set of genes with fold changes >8 and < −8 and annotated the top DEGs from this set. The detailed list of annotated genes from orthologs in *Arabidopsis* and rice is presented in Data S7b, and some of these genes are boxed in Figure S8 as well.

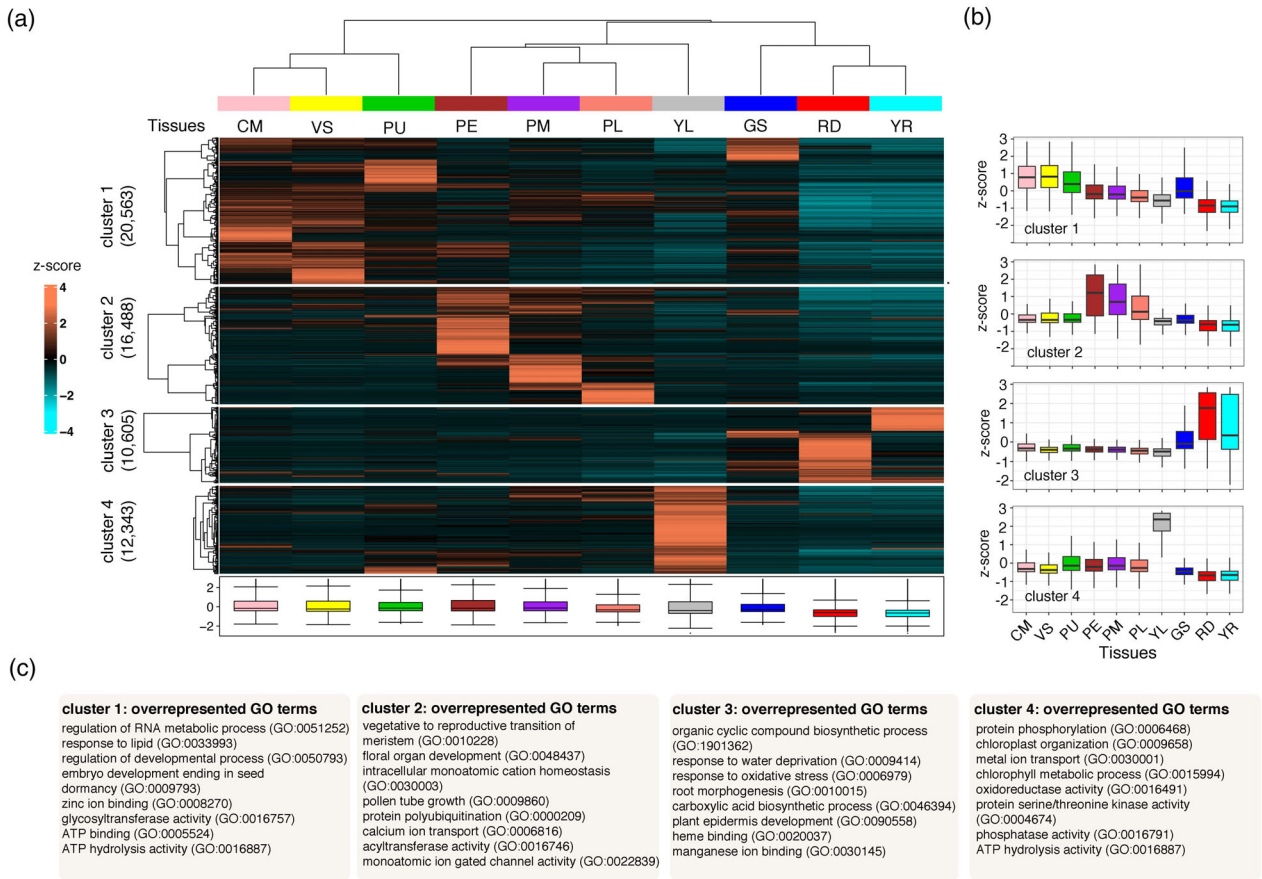

Figure 4. Clustering and heatmap of differentially expressed transcripts (ANOVA, top ranking P_{adj} <0.01) between ten tissue types.
Unsupervised hierarchical clustering demonstrated the clustering of genes into four clusters.
(a) Numbers within parentheses represent the number of transcripts in that cluster. The Z-score of the normalized log-transformed TPM values is visualized by the color key. Coral indicates high expression, black represents intermediate expression, and turquoise is indicative of low expression in the heatmap. Expression in each tissue is summarized in box plot in lower panel.
(b) Distribution of gene expression values as Z-score in each cluster.
(c) Top eight GO terms for each cluster based on overrepresentation test.

Characterization of the mineral-rich profile of little millet

Millets are exceptionally rich in mineral nutrients, such as calcium (Ca), iron (Fe), and zinc (Zn), as compared to major cereals. We quantified a subset of nutrients and mineral ions and found iron to be the most abundant mineral (0.32 ppm/g), followed by zinc (0.28 ppm/g), similar to the previous reports (Chandel et al., 2014) (Figure 5a, left panel). Protein contributed 15.05% of the total seed weight, followed by crude fiber content (8.06%) (Figure 5a, right panel). Figure 5b shows a virtual cross-section of a seed from the micro-computed tomography (SR-μCT) imaging. Further, using X-ray CT technique we were able to explore the details of seed coat, seed coat porosity, and inclusions of germ and endosperm (Figure 5c). These platforms allowed us to examine the anatomical details of seeds at cellular level non-destructively. The brighter regions in the germ show high-density compositions/inclusions (minerals) compared to the low-density starchy endosperm. Further, the relative quantity of several minerals' localization was observed using the micro-X-ray fluorescence (μXRF) imaging, whereby the seed coat is enriched in manganese (Mn), while the germ is rich in potassium (K), calcium (Ca), copper (Cu), iron (Fe), and zinc (Zn) (Figure 5d).

To delve further into the unique genes and pathways that might contribute to the superior mineral profile of little millet seeds, we cataloged a total of 76 unique little millet transcripts whose orthologs in *Arabidopsis* and rice are known to have role in mineral ion transport pathways (Whitt et al., 2020; Data S8). The gene ontology (GO) enrichment analysis identified the highest fold enrichment for zinc ion transmembrane transporter activity, followed by cadmium and iron transporters (Figure S9a). Several other unspecified metal ion transporters were also identified, which would be attractive candidates for future characterization studies. We looked at the tissue-specific expression patterns of these genes across all samples.

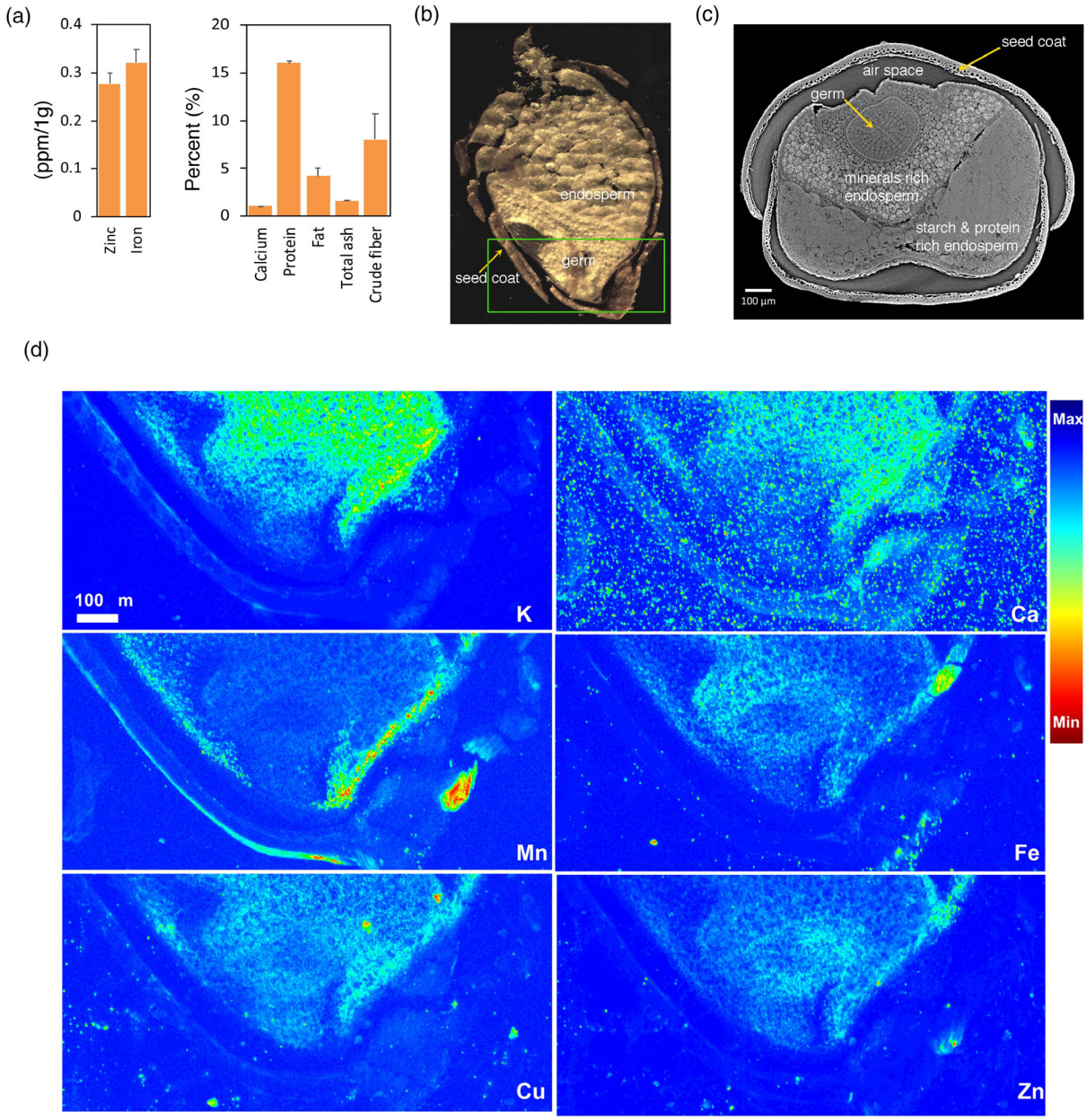

Figure 5. Nutrient elemental composition and distribution in little millet seed.
(a) Nutrient content in little millet seed.
(b) Virtual slice from the X-ray micro-computed tomography datasets of little millet seed.
(c) A microscope image of the cross-section of the seed imaged at the CLS@APS (20-ID) beamline at the Advanced Photon Source showing seed coat, and inclusions of germ and endosperm.
(d) The micro-X-ray fluorescence elemental maps of the little millet sample sections (green box in b) using the microprobe setup at the CLS@APS (20-ID) beamline at the Advanced Photon Source. Elemental densities are visualized as jet colors with lower image densities mapped with "cool" colors and higher densities with "hot" colors for a given element.

A $\log_2$ (TPM + 1) expression level analysis of transcripts with their *Arabidopsis* and rice orthologs shows that little millet tissues have abundance of highly expressed genes involved in acquisition, transport, and response to mineral ions such as iron, zinc, copper, sulfur, calcium, manganese, sodium, and potassium (Figure 6a). Mineral ions in soil are taken up by the roots, transported to leaves and reproductive tissues, and finally accumulated in developing grains. Z-score values of expression levels allowed us to make tissue-wise comparison of those genes (Figure S9). Owing to their importance in biofortification, we focused our analysis to iron- and zinc ion-related gene

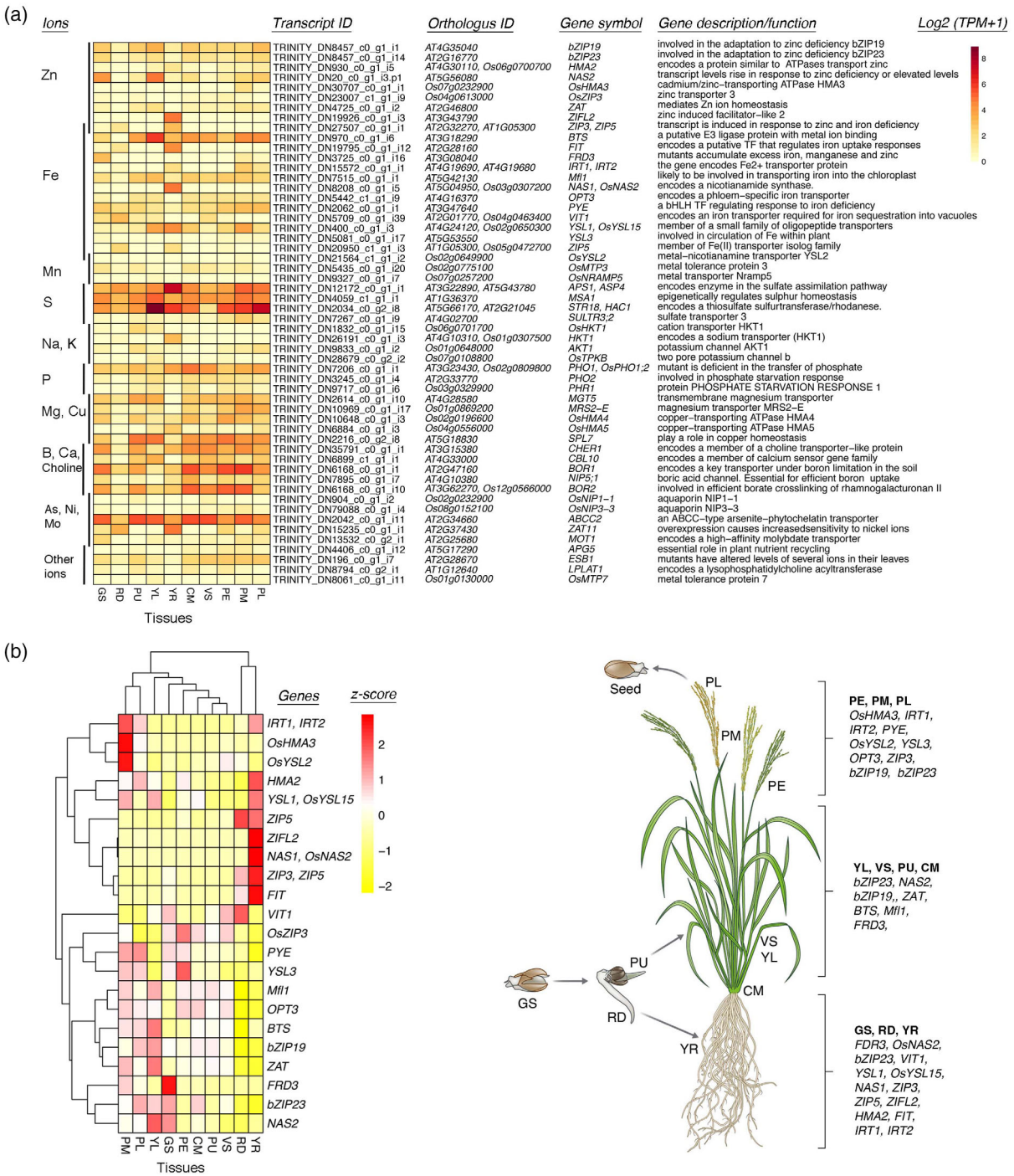

Figure 6. Expression of mineral ion transport genes in little millet.

(a) Heatmap of the differently expressed transcripts involved in transport, response, and regulation on mineral ions. Normalized $\log_2$ (TPM + 1) expression values are indicated by the color key.

(b) Clustering of genes involved in transport, response, and regulation of zinc and iron ions represented as a heatmap across the various tissue types. The Z-scores of normalized $\log_2$ (TPM + 1) expression values are indicated by the color key. Red indicates high expression, white represents intermediate expression, and yellow is indicative of low expression in the heatmaps. Schematic illustration of little millet plant on the right panel depicts groups of genes that are highly expressed in a specific set of tissues.

expression in specific tissues (Figure 6b). YR showed high expression of several iron and zinc transporters and regulators. This includes a major iron transporter *Iron Regulated Transporter 2* (*IRT2*) and the TF *FER-like Iron deficiency-induced TF (FIT)* both of which have been reported to be induced by iron deficiency (Schwarz

et al., 2020; Vert et al., 2009; Vert et al., 2001). Other iron-related genes highly expressed in roots include *YSL1* and *NAS1*. Similarly, Zn transporters *ZIFL2, ZIP3, and ZIP5* are highly expressed in roots (Figure 6b). Together with YR, the germinating seed (GS) and radicle (RD) also showed higher expression of some of these genes such as *FDR3*, *ZIP5*, and *VIT1*. Young leaves (YR) and other vegetative tissues displayed higher expression of genes involved in iron and zinc ion transport and homeostasis such as *bZIP23, bZIP19, ZAT, and BTS*. The reproductive tissues (PE, PM, and PL) which are the sites of grain development also contained an abundance of iron and zinc ion transport-related genes such as *HMA3, PYE, YSL3*, and *OPT3*.

The little millet eFP browser

The Bio-Analytic Resource for plant biology (BAR)/ eFP browser is a user-friendly online tool for visualizing gene expression levels in a tissue- or condition-specific manner. Using the expression of assembled transcripts across the 10 tissues used in this study, the eFP browser hosts a tissue-specific little millet transcript abundance visualization resource at http://bar.utoronto.ca/~asher/efp_little_millet/cgi-bin/efpWeb.cgi. As shown in Figure 7, transcript abundances (expressed as TPM values) in 10 tissues can be visualized at different stages of the little millet life cycle as a spectrum of yellow to red color (in ascending order of expression strength). Here, we demonstrate the dynamic expression of two little millet genes putatively involved in ion transport in different tissues. Little millet transcripts TRINITY_DN15235_c0_g1_i1 and TRINITY_DN970_c0_g1_i6, corresponding to characterized *Arabidopsis* genes *AT2G37430* (*ZAT11, Zinc finger of Arabidopsis thaliana*) and *AT3G18390* (*BTS, Brutus*) involved in nickel and iron ion transport, respectively (Hindt et al., 2017; Long et al., 2010), showed comparatively higher expression in YR (Figure 7a) and YL (Figure 7b).

DISCUSSION

Modern agricultural practices promote resource-intensive monoculture of selective crops that are bred to maximize productivity, while traits such as climate resilience or micronutrient content in seeds are often overlooked. These shortcomings have prompted a focus on resurrecting neglected ancient crops, including small millets, as nutrient-dense and sustainable food sources. (Food and Agriculture Organization of the United Nations, 2018). Little millet has been identified as the richest source of dietary fiber among major and minor cereals (Vetriventhan et al., 2020). It has approximately twice the zinc and four times the iron and crude fiber per gram of grain compared to major cereals, rice, and wheat (Chandel et al., 2014; Vetriventhan et al., 2020). It is a fast-maturing crop, requires minimal fertilizer and irrigation, and is resistant to heat and drought conditions (Goron & Raizada, 2015). These features make little millet an attractive model for gene mining and understanding the molecular pathways underlying the highly sought-after agronomic traits. However, the lack of foundational genetic and transcriptomic resources for little millet has severely restricted the utilization of its rich gene pool in nutrition enhancement and crop improvement strategies. This study aimed to build an expression atlas for little millet from progressive life cycle stages. To the best of our knowledge, the present study is

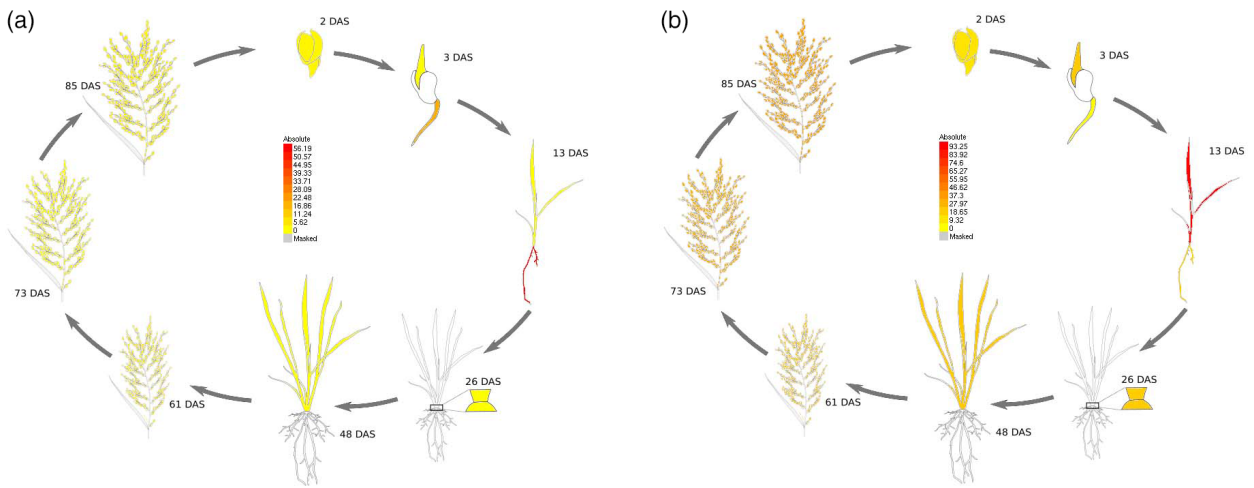

Figure 7. Little millet life cycle eFP browser at bar.utoronto.ca showing changes in gene expression.
(a) Expression of TRINITY_DN15235_c0_g1_i1, the ortholog in *Arabidopsis* (AT2G37430) encodes a member of the zinc finger family of transcriptional regulators (*ZAT11*). It is expressed in root tips, primary roots, cotyledons, and hypocotyl and is involved in nickel ion transport.
(b) Expression of TRINITY_DN970_c0_g1_i6, the ortholog of which in *Arabidopsis* (AT3G18390) encodes BRUTUS (*BTS*), a putative E3 ligase protein with metal ion binding and DNA-binding domains, which negatively regulates the response to iron deficiency.

the first developmental stage-specific transcriptomic dataset for an otherwise understudied crop, little millet.

Using the Illumina sequencing platform, 28 samples belonging to 10 tissues produced a combined total of 258.41 million paired-end reads. This accounts to an average of 27.60 million per tissue (GS, RD, PU, YL, CM, PE, PM, PL) and 19 million per tissue (YR, VS). Although it is preferable to have a greater sequencing depth than our overall average of 9.24 million reads per sample, increasing depth beyond 10 million reads results in diminishing returns in terms of the ability to detect differentially expressed genes (Liu et al., 2013) or leads to an increase in unannotated single-exon transcripts, with a majority of these sequences originating from intronic regions (Patterson et al., 2019). *De novo* transcriptome assembly of the transcripts from all the tissues used in this study identified 342 827 transcripts. This number is substantially higher than the 55 527 and 37 908 protein-coding genes identified in related genomes of broomcorn millet (Zou et al., 2019) and foxtail millet (Thielen et al., 2020). We have also assessed the quality of the generated little millet transcriptome using the gold standard BUSCO analysis (Simão et al., 2015). Of the 3236 conserved plant proteins in BUSCO sets, our dataset had 92.3% full-length and 5.3% partial proteins indicating a thorough transcriptome coverage. This is comparable to 98% BUSCO coverage achieved by the broomcorn millet genome (Zou et al., 2019). Transcriptomic datasets have previously been used as valuable resources for understanding evolutionary and biological processes.

Defining trait-associated gene regulatory network

Little millet is a rich source of minerals, as shown in this and previous works (Figure 5) (Chandel et al., 2014, Vetriventhan et al., 2020). With the little millet transcriptome data, we sought to identify transcriptional patterns that might contribute to the mineral-dense profile of the crop. Recently, a total of 21 protein families were recognized as having functions related to mineral uptake in pearl millet (Satyavathi et al., 2022). In our study, we identified a total of 76 differentially expressed transcripts whose orthologs in *Arabidopsis* and rice are known to be involved in mineral ion-related pathways (Whitt et al., 2020). As little millet is exceptionally rich in iron and zinc, we focused on genes and molecular pathways associated with the uptake and accumulation of these mineral ions.

Graminaceae plants uptake Fe^{3+} as chelated complexes with phytosiderophores (Marschner & Römheld, 1994). Nicotianamine synthases (NAS) catalyze phytosiderophore synthesis (Kobayashi & Nishizawa, 2012), and overexpression of *OsNAS2* resulted in increased accumulation of iron and zinc in rice (Johnson et al., 2011; Lee et al., 2012; Singh et al., 2017). Our dataset showed very high expression of *NAS1/OsNAS2* in YR. Other member of NAS family, *NAS2*, was highly expressed in YL and GS as well as in PM and PL. The chelated Fe^{3+} is circulated within the plant by Yellow Striped-1 Like (YSL) protein family (Curie et al., 2009). We identified members of YSL family expressed in little millet tissues. While *YSL3* showed moderate expression in reproductive tissues (PE, PM, and PL), *YSL1/OsYSL15* showed much higher expression in the young roots. In the poly-metal hyperaccumulator *Thlaspi caerulescens*, of several YSL protein family members, only *YSL3* could rescue iron uptake defective yeast strain, indicating its role in the import and/or circulation of iron within plants (Gendre et al., 2007). We further identified homologs of three central regulators of iron sensing and signaling pathway among the mineral ion pathway genes in our dataset. These include two transcription factors, *Fe-deficiency Induced Transcription factor 1* (*FIT1*) and *POPEYE* (*PYE*), and their negative regulator *BRUTUS* (*BTS*). *FIT1*, which showed extremely high expression in young roots, can regulate the expression of over 40% Fe-accumulation-related genes in *Arabidopsis*, including *IRT1*, *IRT2*, and *NRAMP1* (Colangelo & Guerinot, 2004). At the same time, *IRT1* and *NRAMP1* act synergistically for iron uptake and transport in plants, while *IRT2* regulates *IRT1* function (Castaings et al., 2016; Curie et al., 2000; Vert et al., 2001, 2002, 2009). In our dataset, we observed high expression of *IRT1* and *IRT2* in YR and PM and moderate expression *OsNRAMP5* in most of the tissues. In addition, we also identified *Vacuolar Iron Transporter 1* (*VIT1*), which mediates the detoxification of cytosolic Fe^{3+} by its sequestration in the vacuole (Kim et al., 2006). Tissue-specific modulation of *VIT1* expression in wheat and rice was reported to cause iron biofortification in seeds (Connorton et al., 2017; Zhang et al., 2012).

The iron and zinc signaling pathways share several common genes. These include NAS, YSL, IRT, and ZIP protein families. Similar to Fe^{3+}, the association of NAS2 with Zn^{2+} facilitates its movement within the plant (Deinlein et al., 2012). Interestingly, similar to the high expression of *NAS2* in little millet, two hyperaccumulators of zinc, *Arabidopsis halleri* and *T. caerulescens*, show elevated expression of *NAS2*, and targeted reduction of the *NAS2* transcript resulted in reduced zinc accumulation in *A. halleri* (Deinlein et al., 2012). ZIPs are the primary zinc uptake and transport proteins in plants (Grotz et al., 1998). Overexpression of *Arabidopsis ZIP1* in cassava led to up to 10 times increase in zinc accumulation in tubers (Gaitán-Solís et al., 2015). Further, *ZIP1* expression was positively correlated with zinc hyper-accumulation in *T. caerulescens* and *A. halleri* (Becher et al., 2004; Pence et al., 2000). Our dataset found high expression of *ZIP3* and *ZIP5* in radicle and young roots. Further, comparative transcriptome analysis of zinc hyper-accumulator and non-accumulator species identified differential expression of *Metal Tolerance Protein 1* (*MTP1*) and *Heavy Metal-Associated* (*HMA*) gene families

between the two groups (Becher et al., 2004; Broadley et al., 2007). *MTP1*, which sequesters zinc from the cytosol to the vacuole (Desbrosses-Fonrouge et al., 2005; Gustin et al., 2009), showed high expression in all tissues of little millet, except RD. *HMA*s are ATP-dependent pumps that, similar to *MTP1*, are involved in cytoplasmic detoxification and are associated with zinc hyper-accumulation tolerance of *A. halleri* (Hanikenne et al., 2008; Kim et al., 2009; Morel et al., 2009). *OsHMA3* in our dataset showed relatively higher expression in PM, and *OsHMA4* and *OsHMA5* showed high expression in most of the vegetative and reproductive phase tissues. In our study, *YSL3* and *OsZIP3* showed higher expression in PE and *IRT1*, and *OsHMA3* and *OsYSL2* showed higher expression in PM compared to late stage of panicle development. Parallel to our study, Satyavathi et al. (2022) showed mineral transport process to be more active during panicle initiation stages in pearl millet. Although the genes identified in this study represent excellent candidates for Fe and Zn biofortification in crop plants, additional validation and in-depth functional characterization are necessary to uncover their specific roles in little millet.

Understanding the evolutionary relationship between little millet and other related species

Translational research for crop improvement is more feasible in genetically closely related species. Previously, Huang et al. (2016) found a close phylogenetic relationship among little millet, foxtail millet, pearl millet, and finger millet, while broom corn was only distantly related. Similarly, Kumari et al. (2013) found little millet closely related to foxtail millet and pearl millet, but distantly related to broom corn. It must be noted that the studies mentioned above utilized selected nuclear or chloroplast markers that represent a relatively more minor portion of the genome (Huang et al., 2016 used data from Grass Phylogeny Working Group II) as compared to the composite little millet transcriptome dataset used in this study.

Little millet has been domesticated from its wild weedy relative *Panicum psilopodium* (de Wet et al., 1983; Goron & Raizada, 2015); however, its progenitor species are unknown. The current study provides an overview of the evolutionary relatedness of the small millet species. With the available transcriptomes from other members of *Panicum* species and other small millets, the phylogenetic relationships and ploidy level of little millet have been generated in this study.

Generating publicly available datasets and visual tissue-specific expression atlas of little millet

The massive scientific advances in the last few decades have, in part, been attributed to the generation and availability of large-scale datasets (Kagale et al., 2016; Varmus, 2002). With this study, we contribute the first little millet transcriptome to the eFP browser, which allows for the visual inspection of the gene expression data. In addition, we have also provided a list of housekeeping genes that were found to be stably expressed in all the tissues. This gene list could be used in future transcript quantification studies in little millet, as reference for housekeeping genes.

Previous studies have reported a vast genetic diversity in little millet landraces due to several independent domestication events in different environmental conditions (de Wet et al., 1983). In addition, two independent studies (Nirmalakumari et al., 2010; Upadhyaya et al., 2014) have identified a wide range of phenotypic features in little millet germplasm collection. Therefore, a detailed transcriptional exploration of this genetic diversity, coupled with phenotypic assessment, is warranted. As the first step in this direction, our study presents a tissue-specific little millet transcriptome and lists the attractive candidates likely to be involved in mineral ion uptake, transport, and accumulation. This, together with tissue specific transcription factors identified in our study, could be supplemented with the recently published little millet transcriptome for abiotic stress responses (Das et al., 2020). Future studies aimed at improving little millet productivity could present the crop as an attractive supplement to the major cereals as a source of food and fodder.

MATERIALS AND METHODS

Plant material, assessment of seed nutritional composition, growth conditions, and RNA extraction

Seeds of little millet genotype JK-8 were obtained from the All India Coordinated Research Project (AICRP) on Small Millets located at the University of Agricultural Sciences, Bengaluru, India. Nutritional compositions of seeds including moisture, fat, calcium, protein, total ash, crude fiber, iron, and Zn contents were quantified as previously described (Heau et al., 1965; Horwitz, 1980; Raguramulu et al., 2003).

Ten tissue samples from various developmental stages of little millet were used to analyze its global transcriptome. The tissue samples represent three growth phases in the little millet life cycle: emergence phase (GS, RD, PU), vegetative phase (YL, YR, CM, VS), and reproductive phase (PE, PM, PL) (Figure 1).

The seeds were germinated on Whatman paper sheets moistened with Hoagland solution in Petri dishes (Hoagland & Snyder, 1933). Whole GS was harvested 2 days after sowing (DAS), and developing RD and PU tissues were separately harvested 3 DAS (Table S1). For vegetative and reproductive phases, the seeds were stratified in water-moistened germinating Whatman paper sheets for 72 h at 4°C, before sowing in a potting mix of Sunshine-2 (Sun Gro Horticulture, Vancouver, Canada; growing mix (60%), peat (30%), and vermiculite (10%)). The pots were maintained in a growth chamber at 25°C, 12 h/20°C, and 12 h (day/night) until sample harvest. The vegetative stage samples were harvested 13 DAS (YL and YR), 26 DAS (CM), and 48 DAS (VS). For the reproductive phase, panicles at three stages of growth PE, PM, and PL were harvested 12 days apart from each other at 61 DAS, 73 DAS, and 85 DAS, respectively.

RNA extraction, cDNA library construction, Illumina sequencing, and *de novo* assembly

The samples were harvested in three biological replicates and flash-frozen in liquid nitrogen. Total RNA was extracted from the harvested tissues using RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) for all samples except for GS, which was extracted using TRIzol (Sigma-Aldrich, Saint Louis, USA) following the manufacturer's protocol. RNA was purified using the RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany) and quantified using a NanoDrop ND-100 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). RNA integrity was evaluated using Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA).

A total of 28 paired-end cDNA libraries were prepared using a TruSeq RNA sample preparation kit (Illumina, San Diego, USA) following the manufacturer's protocol. The 28 libraries constituted two replicates of YR and VS and three replicates of the remaining 8 tissues. Of the three biological replicates, one replicate of YR failed the RNA integrity test, and one replicate of VS failed quality assessment after library preparation and, therefore, was not processed further. Paired-end sequencing was performed on the cDNA libraries multiplexed in two lanes of a flow cell (125 cycles) using the Illumina HiSeq 2500 (Illumina, San Diego, USA) platform at the National Research Council Canada, Saskatoon, SK, Canada. The quality of the raw reads was assessed by FastQC (https://github.com/s-andrews/FastQC), and low-quality reads and contaminants were filtered using Trimmomatic v0.38 (Bolger et al., 2014) by (i) removing adapter sequences, (ii) trimming low-quality reads, and (iii) removing sequences with a shorter length than 75 bp. The *de novo* assembly of the filtered reads was performed using Trinity assembler ver. 2.9.1 with the default settings. Assembled Trinity contigs were clustered at 100% identity using CD-HIT-EST (Huang et al., 2010), and duplicate transcripts were removed. The quality of transcriptome assembly was evaluated by (i) examining the RNA-seq read representation in the assembly and (ii) assessing the completeness of the assembly using BUSCO (Benchmarking Universal Single-Copy Orthologs).

Construction of phylogeny with related species and Ks analysis

The reciprocal best BLAST hit method was used to identify potential orthologs between little millet and various species, including small-grained dicots (quinoa) and monocots (ragi, foxtail millet, broom corn, pearl millet) and major cereals (rice and maize), as well as outgroup dicot species (*Arabidopsis* and alfalfa). Utilizing this approach, a phylogenomic data matrix was constructed, consisting of 2752 distinct sets of orthologous genes. The sequences within each orthologous gene set underwent local alignment using ClustalW (Larkin et al., 2007), and gaps and missing data were eliminated from each alignment through an automated alignment trimming tool, trimAL (Capella-Gutiérrez et al., 2009), with a gap threshold value (−gt) set at 1.

The trimmed sequences were then realigned, and the alignments from the 2752 gene sets were concatenated using the Phyutility program (Smith & Dunn, 2008) to generate the final data matrix, comprising a total alignment length of 3 475 812 base pairs. Phylogenetic analysis was executed using the maximum likelihood method implemented in RAxML (Stamatakis, 2006), assuming the GTR + GAMMA model of sequence evolution. The robustness of the phylogenetic inference was evaluated through 1000 bootstrap replicates using the GTR + CAT approximation. The resulting tree was visualized using the Interactive Tree of Life (Letunic & Bork, 2019) Web server.

For Ks analysis, paralogous genes within little millet were identified by performing an all-against-all protein sequence similarity (BLASTP with an E-value cutoff of 1E-20) search. The 185 217 proteins predicted by transdecoder were used for this analysis. For each pair of paralogs, protein sequences were aligned using ClustalW (Larkin et al., 2007). The resulting protein alignments were used to produce the corresponding nucleotide alignments using PAL2NAL (Suyama et al., 2006). Ks values for each sequence pair were calculated based on codon alignments using the maximum likelihood method implemented in codeml of the PAML package (Yang, 2007) under the F3x4 model (Goldman & Yang, 1994). Mixture model analysis of Ks distributions was performed as described previously (Kagale et al., 2014).

Validation and quality assessment of the assembled transcriptome

Scatterplots and MA plots were generated to assess the expression dynamics of detected genes for each replicate of the tissue samples using the transcript-level estimates of fragment counts and the PtR (Perl-to-R) script included in the Trinity toolkit. To ensure the normalization of transcriptome data across all the samples, transcripts per kilobase million (TPM) were calculated for each transcript using Kallisto (Bray et al., 2016). Normalization allows samples to be compared regardless of the variabilities in library sizes (sequencing depths) or gene lengths (Li & Dewey, 2011; Wagner et al., 2012; Zhao et al., 2020).

Principal component analysis (PCA) was performed using the Bioconductor package DEseq2 (version 1.28.1) in R software (version 4.0.4) based on variance-stabilized normalized read counts (Love et al., 2014). To reduce the range of the data, the TPM values were transformed by adding one and taking the natural logarithm. The transformed values were then used to generate an expression level distribution plot of replicates using the ggplot2 (version 3.3.3) function in R (Wickham, 2009). Hierarchical clustering of samples was performed using hclust function in *R* for complete linkage method (cran.r-project.org/package=hclust1d). All R scripts, related libraries, and datafiles used to generate R plots are available publicly https://github.com/shankar7321/Little-millet.

Gene clustering and differential gene expression

The expression strength of each transcript was evaluated prior to downstream analyses. A transcript was considered to be expressed if the TPM value of at least two replicates was greater than 0.50. The transcript expression was considered to be tissue-specific when the TPM values for all other tissues were less than 0.50. Next, the coefficient of variation (CV) was used to identify constitutively expressed transcripts, where CV was calculated as the ratio between standard deviation and mean of $\log_2$ (TPM + 1) values for each transcript across the samples. Genes with CV ≤6% across the samples were considered stably expressed genes.

Statistical ANOVA was performed on the expressed set of transcripts using their TPM expression to assess the transcripts' *P*-values across all samples. Hierarchical clustering was performed on genes with *P*-values <0.01 after applying Benjamini–Hochberg correction for FDR (Benjamini & Hochberg, 1995) using the Euclidean distance metric and Ward's linkage method and plotted using the Complex Heatmap package (version 1.14.0) in *R* (Gu et al., 2016). The optimal number of clusters was determined using the elbow method (Charrad et al., 2014).

Z-scores from the mean of TPM expression values of the replicates for each tissue were used to construct clustered heatmap

and boxplots. The Z-scores were calculated as follows: Z = (X-Xmean)/SD, where X is the expression of a given gene in a tissue and Xmean and SD are the mean expressions and standard deviation, respectively, of that gene across all the selected tissues.

The differential expression between groups of tissues was analyzed in R using the DESeq2 package (version 1.28.1) (Love et al., 2014). Using a model based on the negative binomial distribution, DESeq2 can provide statistics that determine differences in gene expression data. A gene was considered to be differentially expressed if $\log_2$ fold change was ≥2 and false discovery rate (FDR)-adjusted *P*-value (P_{adj}) was ≤0.01. Volcano plots were produced using the enhanced volcano *R* package (Blighe et al., 2019).

GO annotation, TF identification, gene enrichment, and data plotting

For each ortholog of transcripts in *Arabidopsis*, a GO term was assigned and categorized to molecular function, biological process, and cellular component using the *Arabidopsis* Information Resource, version 10 (TAIR10) (www.arabidopsis.org) (Berardini et al., 2004). Transcription factors were identified using the *Arabidopsis* Gene Regulatory Information Server (AGRIS; http://arabidopsis.med.ohio-state.edu/). GO term overrepresentation analysis was performed using the web-based GO Enrichment Analysis tool (http://geneontology.org) Panther v.16.0 (Mi et al., 2019). The GO terms with *P*-value ≤0.05 (Fisher's exact and Bonferroni correction for multiple testing) were considered to be significantly enriched. The scatterplot was constructed using the ggplot2 package in *R* (Wickham, 2009). Heatmap was constructed using the pheatmap *R* package (Kolde, 2019). For the set of genes identified to be involved in mineral ion transport in *Arabidopsis* and rice (Whitt et al., 2020), gene name/symbol, gene function, functional classification, and protein class were assigned using the PANTHER classification system (http://www.pantherdb.org/) Panther v.16.0 (Mi et al., 2019), TAIR10 (Berardini et al., 2004) and Rice Annotation Project Database (RAP-DB) (Sakai et al., 2013). The GO term enrichment analysis for those genes and lollipop plot was conducted using ShinyGO v0.741 (Ge et al., 2020) and a web-based online tool (http://bioinformatics.sdstate.edu/go/).

Synchrotron-based X-ray micro-computed tomography (SR-μCT)

X-ray computed tomography data of the Indian pearl millet sample were collected using the BMIT-BM (05B1-1) beamline at the Canadian Light Source. A filtered white beam setup was used with 0.8 mm aluminum, which was used to generate a broadband beam with a mean energy of ~20 keV. A PCO Edge 5.5 camera coupled with a 10X optic peter objective achieves a field of view of 1.85 × 1.56 mm and an effective pixel size of 0.72 μm. The sample was mounted with dental wax to an aluminum SEM stub which is secured to a Huber goniometer stage. A sample-to-detector distance of 4.5 cm was used, and 3000 projection images were collected over a 180° rotation along with 20 flat (with X-ray beam on and without sample) and 20 dark (with X-ray beam off and without sample) images. The flat and dark images were used to normalize the projection images. X-ray image normalization and reconstruction into a 3D dataset were accomplished using the UFO-KIT software (Vogelgesang et al., 2012, 2016). Similar to other experiments conducted at the same beamline (Chen et al., 2021; Willick et al., 2020), phase retrieval was conducted with a delta/beta ratio of 200. Images were filtered using a Sarepy ring-removal algorithm, and a Laplace edge detection was used to further enhance contrast in the reconstructed slices.

Micro-X-ray fluorescence (XRF) elemental maps

The little millet seeds' micro-X-ray fluorescence elemental maps were collected using the microprobe setup at the CLS@APS (20-ID) beamline at the Advanced Photon Source. The seed was first soaked in water for 24 h and flash-frozen in water using liquid nitrogen. The sample was mounted to a cryo-chuck with frozen sectioning media and sectioned to 40 μm using a cryostat (Leica CM 1950, Leica Biosystems, Nussloch, Germany). The sections were then mounted on a Kapton tape and air-dried by placing them over a Teflon cutout to keep them flat until data collection. An incident X-ray energy of 12.8 keV was used to excite the sample section, and the focused X-ray beam spot size was 2 × 2 μm. The sample was kept 45° to the incident beam and the detector (Vortex 4-element silicon drift detector). The sample was raster-scanned with a dwell time of 200 ms per pixel. The XRF data analysis was completed using the PyMCA software (version 5.3.1; Solé et al., 2007).

ACCESSION NUMBERS

Sequence data from this paper can be found in the NCBI data libraries under the following accession number GEO (GSE183311).

ACKNOWLEDGMENTS

This work was supported by Agriculture and Agri-Food Canada to R.S. and toward eFP Browser web support N.P. received funding from NSERC. The X-ray computed tomography work described in this paper was performed at the Canadian Light Source, a national research facility of the University of Saskatchewan, which is supported by the Canada Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council (NSERC), the National Research Council (NRC), the Canadian Institutes of Health Research (CIHR), the Government of Saskatchewan, and the University of Saskatchewan. The micro-XRF data were collected from the CLS@APS and used resources of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, and were supported by the U.S. DOE under Contract No. DE-AC02-06CH11357 and the Canadian Light Source and its funding partners. Open Access funding provided by the Gouvernement du Canada Agriculture et Agroalimentaire Canada library.

CONFLICT OF INTEREST

The authors declare no conflict of interest and agree to the presented work.

AUTHOR CONTRIBUTIONS

RS conceived the study and managed the project. SP and RS performed the growth chamber study and generated RNA-seq data. SP, SK, and RS analyzed RNA-seq data. NV and SP performed analysis for mineral ion-associated genes. NV, AP, EE, and NP developed the eFP Browser. RS and CK conceived the synchrotron work. JS collected and analyzed the X-ray microcomputed tomography data. DM and MV collected and analyzed the X-ray fluorescence data. PB and MKP performed seed elemental composition. AN and AKJ provided the JK-8 genotype for the study. SP and NV wrote the first draft of the paper. RS and SK edited

and finalized the manuscript. RS, MKP, and NP contributed the materials, reagents, and web tools. All authors read and approved the manuscript.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Data S1. Normalized TPM counts of transcripts across 28 different samples collected from 10 tissues representing 3 growth phases.

Data S2. (a) List of expressed set of transcripts. ANOVA test was performed to assess the differences in means of transcripts and adjusted for FDR. Significance was cutoff at P_{adj} <0.01. Normalized TPM values were used to compute ANOVA. Mean of log2 (TPM + 1) for each tissue is shown. (b) Raw counts data used to compute differentially expressed transcripts using DESeq2. Transcripts having P_{adj} <0.05 from ANOVA test (Data S2a) were chosen.

Data S3. List of transcripts expressed in all tissues or uniquely expressed in specific tissue.

Data S4. Coefficient of variation (CV, %) of the constitutively expressed transcripts calculated from mean log2 (TPM + 1) expression values of transcripts. Transcripts with CV <6% are in bold and annotated with Arabidopsis or rice orthologs.

Data S5. (a) Number of transcription factors specific to each tissues belonging to different transcription factor families. (b) List of Arabidopsis orthologs expressed in specific tissues belonging to various transcription families.

Data S6. GO terms overrepresented in four clusters.

Data S7. (a) List of differentially expressed transcripts from a pairwise comparison between four groups of tissues using DESeq2 analysis. (b) Annotated list of differentially expressed transcripts. Arabidopsis and rice orthologs of transcripts with log2 fold change >8 and P_{adj} values <0.01 were annotated.

Data S8. (a) Transcripts and their ortholog genes in Arabidopsis involved in mineral ion transport, response and its acquisition.

(b) Transcripts and their ortholog genes in rice involved in mineral ion transport, response and its acquisition.

Figure S1. RNA-Seq processing pipeline used to generate gene expression atlas of little millet.

Figure S2. Scatter plot and MA plot for each replicate of the tissue samples used to access the expression dynamics of the detected transcripts.

Figure S3. Analysis of global gene expression in all tissues. Distribution plot depicting transcript expression for all 28 samples. TPM normalized values were $\log_2$ transformed and were used to represent the distribution of expression values.

Figure S4. Sample relationship and its distribution. (a) Heatmap of hierarchical clustering of Pearson's pairwise correlations for 28 samples. The color scale indicates the degree of correlation. (b) Two-dimensional PCA plot of tissue included in the study based on their normalized TPM expression values.

Figure S5. Screening of the tissue specific and housekeeping genes. (a) Heatmap of tissue specific expression profile. The color scale at the top represents $\log_2$ (TPM + 1) values. (b) Scatterplot of transcripts consecutively expressed in all tissues with their mean and standard deviation (SD) values. Transcripts in orange are proposed to be stably expressed with housekeeping functions. (c) List of housekeeping genes annotated from corresponding orthologs in *Arabidopsis* and rice.

Figure S6. Functional categories and transcription factor families of expressed transcripts. Bar graph representing % of genes in each GO terms of three categories, (a) Biological processes, (b) Molecular functions, and (c) Cellular components. (d) Distribution pattern of genes belonging to various transcription factor families. N represents total number of genes in each functional category.

Figure S7. Elbow plot constructed using unsupervised hierarchical clustering to identify the optimal number of clusters. Blueline indicates optimum number of clusters to be 4.

Figure S8. Dendrogram of tissue sample clustering and volcano plot of differentially expressed genes identified between tissue groups. (a) Cluster dendrogram obtained from complete linkage showing the global relationship among different tissue samples. 28 samples from 10 tissues formed 4 major groups. (b) Volcano plots to display differentially expressed genes. X-axis represents $\log_2$ fold-change between the two groups; and y-axis represents negative $\log_{10}$ *P*-value of the two groups. The red points indicate differentially expressed genes using the criteria P_{adj} <0.01 and FC >8 or FC <−8 as indicated by dashed lines. Number of DEGs upregulated and downregulated are specified in each plots with green and red arrows respectively. A subset of annotated genes under the criteria have been boxed.

Figure S9. Gene enrichment and expression of mineral ion transport genes. (a) Lollipop plot for gene ontology enrichment analysis of *Arabidopsis* ortholog genes involved in mineral ion transport. The horizontal axis indicates fold enrichment and the number of gene enriched in a GO term is represented by dot sizes. (b) Clustering of differently expressed transcripts involved in transport, response and regulation of mineral ions represented as a heatmap across the various tissue types. The Z-scores of normalized log TPM expression values are indicated by the color key. Red indicates high expression, white represents intermediate expression and yellow is indicative of low expression in the heatmaps.

Table S1. Details of tissues and a summary of the sequencing data generated for developing gene expression atlas.

Table S2. Pearson's correlation matrix between 28 samples calculated from log2 (TPM + 1) values of differentially expressed genes. Values in red font indicate the correlation between the same samples and those highlighted in yellow are between biological replicates.

OPEN RESEARCH BADGES

This article has earned Open Data and Open Materials badges. Data and materials are available at: The raw data publicly accessible at – https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE183311. The R scripts is available in Github – https://github.com/shankar7321/Little-millet. Developmental transcriptome of Little millet (Panicum sumatrense) (zenodo.org). Illumina HiSeq 2500 for little millet (Panicum sumatrense): https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE183311.

DATA AVAILABILITY STATEMENT

The developmental transcriptome of little millet (*Panicum sumatrense*): Developmental transcriptome of Little millet (Panicum sumatrense) (zenodo.org). Illumina HiSeq 2500 for little millet (*Panicum sumatrense*): https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE183311. R scripts and associated data files: https://github.com/shankar7321/Little-millet.

REFERENCES

Ajithkumar, I.P. & **Panneerselvam, R.** (2014) ROS scavenging system, osmotic maintenance, pigment and growth status of *Panicum sumatrense* Roth. under drought stress. *Cell Biochemistry and Biophysics*, **68**, 587–595.

Amadou, I., **Gounga, M.E.** & **Le, G.-W.** (2013) Millets: nutritional composition, some health benefits and processing – a review. *Emirates Journal of Food and Agriculture*, **25**, 501–508.

Becher, M., **Talke, I.N.**, **Krall, L.** & **Krämer, U.** (2004) Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *The Plant Journal*, **37**, 251–268.

Benjamini, Y. & **Hochberg, Y.** (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, **57**, 289–300.

Berardini, T.Z., **Mundodi, S.**, **Reiser, L.**, **Huala, E.**, **Garcia-Hernandez, M.**, **Zhang, P.** ***et al.*** (2004) Functional annotation of the *Arabidopsis* genome using controlled vocabularies. *Plant Physiology*, **135**, 745–755.

Bhaskaran, J. & **Panneerselvam, R.** (2013) Accelerated reactive oxygen scavenging system and membrane integrity of two *Panicum* species varying in salt tolerance. *Cell Biochemistry and Biophysics*, **67**, 885–892.

Blighe, K., **Rana, S.** & **Lewis, M.** (2019) EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling. R package version, 1.

Bolger, A.M., **Lohse, M.** & **Usadel, B.** (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Bray, N.L., **Pimentel, H.**, **Melsted, P.** & **Pachter, L.** (2016) Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, **34**, 525–527.

Broadley, M.R., **White, P.J.**, **Hammond, J.P.**, **Zelko, I.** & **Lux, A.** (2007) Zinc in plants. *New Phytologist*, **173**, 677–702.

Capella-Gutiérrez, S., **Silla-Martínez, J.M.** & **Gabaldón, T.** (2009) trimAl: a tool for automated alignment trimming in largescale phylogenetic analyses. *Bioinformatics*, **25**, 1972–1973.

Castaings, L., **Caquot, A.**, **Loubet, S.** & **Curie, C.** (2016) The high-affinity metal transporters NRAMP1 and IRT1 team up to take up iron under sufficient metal provision. *Scientific Reports*, **6**, 37222.

Chandel, G., **Meena, R.K.**, **Dubey, M.** & **Kumar, M.** (2014) Nutritional properties of minor millets: neglected cereals with potentials to combat malnutrition. *Current Science*, **107**, 1109–1111.

Charrad, M., **Ghazzali, N.**, **Boiteau, V.** & **Niknafs, A.** (2014) NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, **61**, 1–36.

Chen, J., **Ghazani, S.M.**, **Stobbs, J.A.** & **Marangoni, A.G.** (2021) Tempering of cocoa butter and chocolate using minor lipidic components. *Nature Communications*, **12**, 5018.

Chen, S. & **Renvoize, S.A.** (2006) *Panicum* Linnaeus, Sp. Pl. 1:55. 1753. *Flora of China*, **22**, 504–510.

Colangelo, E.P. & **Guerinot, M.L.** (2004) The essential basic helix-loop-helix protein FIT1 is required for the iron deficiency response. *The Plant Cell*, **16**, 3400–3412.

Connorton, J.M., **Jones, E.R.**, **Rodríguez-Ramiro, I.**, **Fairweather-Tait, S.**, **Uauy, C.** & **Balk, J.** (2017) Wheat vacuolar iron transporter TaVIT2 transports Fe and Mn and is effective for biofortification. *Plant Physiology*, **174**, 2434–2444.

Curie, C., **Alonso, J.M.**, **Jean, M.L.**, **Ecker, J.R.** & **Briat, J.-F.** (2000) Involvement of NRAMP1 from *Arabidopsis thaliana* in iron transport. *The Biochemical Journal*, **347**, 749–755.

Curie, C., **Cassin, G.**, **Couch, D.**, **Divol, F.**, **Higuchi, K.**, **Le Jean, M.** ***et al.*** (2009) Metal movement within the plant: contribution of nicotianamine and yellow stripe 1-like transporters. *Annals of Botany*, **103**, 1–11.

Das, R.R., **Pradhan, S.** & **Parida, A.** (2020) De-novo transcriptome analysis unveils differentially expressed genes regulating drought and salt stress response in *Panicum sumatrense*. *Scientific Reports*, **10**, 21251.

de Wet, J.M.J., **Prasada Rao, K.E.** & **Brink, D.E.** (1983) Systematics and domestication of *Panicum sumatrense* (Graminae). *Journal d'agriculture traditionnelle et de botanique appliquée*, **30**, 159–168.

Deinlein, U., **Weber, M.**, **Schmidt, H.**, **Rensch, S.**, **Trampczynska, A.**, **Hansen, T.H.** ***et al.*** (2012) Elevated Nicotianamine levels in *Arabidopsis halleri* roots play a key role in zinc hyperaccumulation. *The Plant Cell*, **24**, 708–723.

Desbrosses-Fonrouge, A.-G., **Voigt, K.**, **Schröder, A.**, **Arrivault, S.**, **Thomine, S.** & **Krämer, U.** (2005) *Arabidopsis thaliana* MTP1 is a Zn transporter in the vacuolar membrane which mediates Zn detoxification and drives leaf Zn accumulation. *FEBS Letters*, **579**, 4165–4174.

Dudhate, A., **Shinde, H.**, **Tsugama, D.**, **Liu, S.** & **Takano, T.** (2018) Transcriptomic analysis reveals the differentially expressed genes and pathways involved in drought tolerance in pearl millet [*Pennisetum glaucum* (L.) R. Br]. *PLoS One*, **13**, e0195908.

Fan, M.-S., **Zhao, F.-J.**, **Fairweather-Tait, S.J.**, **Poulton, P.R.**, **Dunham, S.J.** & **McGrath, S.P.** (2008) Evidence of decreasing mineral density in wheat grain over the last 160 years. *Journal of Trace Elements in Medicine and Biology*, **22**, 315–324.

Food and Agriculture Organization of the United Nations. (2018) Future Smart Food: rediscovering hidden treasures of neglected and underutilized species for zero hunger in Asia. UN https://doi.org/10.18356/23b5f7ab-en

Gaitán-Solís, E., **Taylor, N.J.**, **Siritunga, D.**, **Stevens, W.** & **Schachtman, D.P.** (2015) Overexpression of the transporters *AtZIP1* and *AtMTP1* in cassava changes zinc accumulation and partitioning. *Frontiers in Plant Science*, **6**, 492.

Ge, S.X., **Jung, D.** & **Yao, R.** (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, **36**, 2628–2629.

Gendre, D., **Czernic, P.**, **Conéjéro, G.**, **Pianelli, K.**, **Briat, J.-F.**, **Lebrun, M.** ***et al.*** (2007) TcYSL3, a member of the YSL gene family from the hyperaccumulator *Thlaspi caerulescens*, encodes a nicotianamine-Ni/Fe transporter. *The Plant Journal*, **49**, 1–15.

Goldman, N. & **Yang, Z.** (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular Biology and Evolution*, **11**, 725–736.

Goron, T.L. & **Raizada, M.N.** (2015) Genetic diversity and genomic resources available for the small millet crops to accelerate a new green revolution. *Frontiers in Plant Science*, **6**, 157.

Grabherr, M.G., **Haas, B.J.**, **Yassour, M.**, **Levin, J.Z.**, **Thompson, D.A.**, **Amit, I.** ***et al.*** (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644–652.

Grotz, N., **Fox, T.**, **Connolly, E.**, **Park, W.**, **Guerinot, M.L.** & **Eide, D.** (1998) Identification of a family of zinc transporter genes from *Arabidopsis* that respond to zinc deficiency. *PNAS*, **95**, 7220–7224.

Gu, Z., **Eils, R.** & **Schlesner, M.** (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847–2849.

Gustin, J.L., **Loureiro, M.E.**, **Kim, D.**, **Na, G.**, **Tikhonova, M.** & **Salt, D.E.** (2009) MTP1-dependent Zn sequestration into shoot vacuoles suggests dual roles in Zn tolerance and accumulation in Zn-hyperaccumulating plants. *The Plant Journal: For Cell and Molecular Biology*, **57**, 1116–1127.

Habiyaremye, C., **Matanguihan, J.B.**, **D'Alpoim Guedes, J.**, **Ganjyal, G.M.**, **Whiteman, M.R.**, **Kidwell, K.K.** ***et al.*** (2017) Proso millet (*Panicum miliaceum* L.) and its potential for cultivation in the Pacific northwest, U.S.: a review. *Frontiers in Plant Science*, **7**, 1961.

Hamoud, M.A., **Haroun, S.A.**, **MacLeod, R.D.** & **Richards, A.J.** (1994) Cytological relationships of selected species of *Panicum* L. *Biologia Plantarum*, **36**, 37.

Hanikenne, M., **Talke, I.N.**, **Haydon, M.J.**, **Lanz, C.**, **Nolte, A.**, **Motte, P.** ***et al.*** (2008) Evolution of metal hyperaccumulation required cis -regulatory changes and triplication of HMA4. *Nature*, **453**, 391–395.

Heau, W., **Menzel, R.**, **Roberts, H.** & **Freee, M.** (1965) *Methods of soil and plant analysis*. Department of Agriculture, USA: Agriculture research service.

Hindt, M.N., **Akmakjian, G.Z.**, **Pivarski, K.L.**, **Punshon, T.**, **Baxter, I.**, **Salt, D.E.** ***et al.*** (2017) Brutus and its paralogs, BTS LIKE1 and BTS LIKE2, encode important negative regulators of the iron deficiency response in Arabidopsis thaliana. *Metallomics*, **9**, 876–890.

Hoagland, D.R. & **Snyder, W.C.** (1933) Nutrition of strawberry plant under controlled conditions: (a) effects of deficiencies of boron and certain other elements: (b) susceptibility to injury from sodium salts. *Proceedings of the American Society For Horticultural Science*, **30**, 288–294.

Horwitz, W. (1980) *Official Method of Analysis, 13th (Edn)*. Washington, DC: Association of Analytical Chemists.

Huang, P., **Shyu, C.**, **Coelho, C.P.**, **Cao, Y.** & **Brutnell, T.P.** (2016) Setaria viridis as a model system to advance millet genetics and genomics. *Frontiers in Plant Science*, **7**, 1781.

Huang, Y., **Niu, B.**, **Gao, Y.**, **Fu, L.** & **Li, W.** (2010) CD-HIT suite: a web server for clustering and comparing biological sequences. *Bioinformatics*, **26**, 680–682.

Johnson, A.A.T., **Kyriacou, B.**, **Callahan, D.L.**, **Carruthers, L.**, **Stangoulis, J.**, **Lombi, E.** ***et al.*** (2011) Constitutive overexpression of the OsNAS gene family reveals single-gene strategies for effective iron- and zinc-biofortification of rice endosperm. *PLoS One*, **6**, e24476.

Johnson, M., **Deshpande, S.**, **Vetriventhan, M.**, **Upadhyaya, H.D.** & **Wallace, J.G.** (2019) Genome-wide population structure analyses of three minor millets: Kodo millet, little millet, and Proso millet. *The Plant Genome*, **12**, 190021.

Kagale, S., **Nixon, J.**, **Khedikar, Y.**, **Pasha, A.**, **Provart, N.J.**, **Clarke, W.E.** ***et al.*** (2016) The developmental transcriptome atlas of the biofuel crop *Camelina sativa*. *The Plant Journal*, **88**, 879–894.

Kagale, S., **Robinson, S.J.**, **Nixon, J.**, **Xiao, R.**, **Huebert, T.**, **Condie, J.** ***et al.*** (2014) Polyploid evolution of the Brassicaceae during the Cenozoic era. *The Plant Cell*, **26**, 2777–2791.

Kalaisekar, A., **Padmaja, P.G.**, **Bhagwat, V.R.** & **Patil, J.V.** (2017) *Insect pests of millets: systematics, bionomics and management*, 1st edition. New York: Elsevier, Academic Press.

Kim, S.A., **Punshon, T.**, **Lanzirotti, A.**, **Li, L.**, **Alonso, J.M.**, **Ecker, J.R.** ***et al.*** (2006) Localization of iron in *Arabidopsis* seed requires the vacuolar membrane transporter VIT1. *Science*, **314**, 1295–1298.

Kim, Y.-Y., **Choi, H.**, **Segami, S.**, **Cho, H.-T.**, **Martinoia, E.**, **Maeshima, M.** ***et al.*** (2009) AtHMA1 contributes to the detoxification of excess Zn(II) in *Arabidopsis*. *The Plant Journal*, **58**, 737–753.

Kobayashi, T. & **Nishizawa, N.K.** (2012) Iron uptake, translocation, and regulation in higher plants. *Annual Review of Plant Biology*, **63**, 131–152.

Kolde, R. (2019) Pheatmap: pretty heatmaps. *R package version 1.0. 12. 2019*.

Kumar, A., **Tomer, V.**, **Kaur, A.**, **Kumar, V.** & **Gupta, K.** (2018) Millets: a solution to agrarian and nutritional challenges. *Agriculture & Food Security*, **7**, 31.

Kumari, K., **Muthamilarasan, M.**, **Misra, G.**, **Gupta, S.**, **Subramanian, A.**, **Parida, S.K.** ***et al.*** (2013) Development of eSSR-markers in *Setaria italica* and their applicability in studying genetic diversity, cross-transferability and comparative mapping in millet and non-millet species. *PLoS One*, **8**, e67742.

Larkin, M.A., **Blackshields, G.**, **Brown, N.P.**, **Chenna, R.**, **McGettigan, P.A.**, **McWilliam, H.** ***et al.*** (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

Lee, S., **Kim, Y.-S.**, **Jeon, U.S.**, **Kim, Y.-K.**, **Schjoerring, J.K.** & **An, G.** (2012) Activation of rice nicotianamine synthase 2 (OsNAS2) enhances iron availability for biofortification. *Molecules and Cells*, **33**, 269–275.

Letunic, I. & **Bork, P.** (2019) Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, **47**, W256–W259.

Li, B. & **Dewey, C.N.** (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.

Liu, Y., **Zhou, J.** & **White, K.P.** (2013) RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics*, **30**, 301–304.

Long, T.A., **Tsukagoshi, H.**, **Busch, W.**, **Lahner, B.**, **Salt, D.E.** & **Benfey, P.N.** (2010) The bHLH transcription factor POPEYE regulates response to iron deficiency in Arabidopsis roots. *Plant Cell*, **22**, 2219–2236.

Love, M.I., **Huber, W.** & **Anders, S.** (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 1–21.

Marschner, H. & **Römheld, V.** (1994) Strategies of plants for acquisition of iron. *Plant and Soil*, **165**, 261–274.

Mi, H., **Muruganujan, A.**, **Ebert, D.**, **Huang, X.** & **Thomas, P.D.** (2019) PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, **47**, D419–D426.

Morel, M., **Crouzet, J.**, **Gravot, A.**, **Auroy, P.**, **Leonhardt, N.**, **Vavasseur, A.** ***et al.*** (2009) AtHMA3, a P1B-ATPase allowing Cd/Zn/Co/Pb vacuolar storage in *Arabidopsis*. *Plant Physiology*, **149**, 894–904.

Nirmalakumari, A., **Salini, K.** & **Veerabadhiran, P.** (2010) Morphological characterization and evaluation of little millet (*Panicum sumatrense* Roth. ex. Roem. and Schultz.). *Electronic Journal of Plant Breeding*, **1**, 148–155.

Patterson, J., **Carpenter, E.J.**, **Zhu, Z.**, **An, D.**, **Liang, X.**, **Geng, C.** ***et al.*** (2019) Impact of sequencing depth and technology on de novo RNA-Seq assembly. *BMC Genomics*, **20**, 604.

Pence, N.S., **Larsen, P.B.**, **Ebbs, S.D.**, **Letham, D.L.D.**, **Lasat, M.M.**, **Garvin, D.F.** ***et al.*** (2000) The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*. *PNAS*, **97**, 4956–4960.

Plaza-Wüthrich, S. & **Tadele, Z.** (2012) Millet improvement through regeneration and transformation. *Biotechnology and Molecular Biology Reviews*, **7**, 48–61.

Pradeep, S.R. & **Guha, M.** (2011) Effect of processing methods on the nutraceutical and antioxidant properties of little millet (*Panicum sumatrense*) extracts. *Food Chemistry*, **126**, 1643–1647.

Raguramulu, N., **Madhavan Nair, K.** & **Kalyana Sundaran, S.** (2003) *Laboratory techniques*. Hyderabad: Nin, Icmr Publications.

Saha, D., **Gowda, M.V.C.**, **Arya, L.**, **Verma, M.** & **Bansal, K.C.** (2016) Genetic and genomic resources of small millets. *Critical Reviews in Plant Sciences*, **35**, 56–79.

Sakai, H., **Lee, S.S.**, **Tanaka, T.**, **Numa, H.**, **Kim, J.**, **Kawahara, Y.** ***et al.*** (2013) Rice annotation project database (RAP-DB): an integrative and interactive database for rice genomics. *Plant and Cell Physiology*, **54**, e6.

Satyavathi, C.T., **Tomar, R.S.**, **Ambawat, S.**, **Kheni, J.**, **Padhiyar, S.M.**, **Desai, H.** ***et al.*** (2022) Stage specific comparative transcriptomic analysis to reveal gene networks regulating iron and zinc content in pearl millet [Pennisetum glaucum (L.) r. br.]. *Scientific Reports*, **12**, 276. Available from: https://doi.org/10.1038/s41598-021-04388-0

Schwarz, B., **Azodi, C.B.**, **Shiu, S.-H.** & **Bauer, P.** (2020) Putative cis-regulatory elements predict iron deficiency responses in *Arabidopsis* roots. *Plant Physiology*, **182**, 1420–1439.

Sebastin, R., **Lee, G.-A.**, **Lee, K.J.**, **Shin, M.-J.**, **Cho, G.-T.**, **Lee, J.-R.** ***et al.*** (2018) The complete chloroplast genome sequences of little millet (*Panicum sumatrense* Roth ex Roem. And Schult.) (Poaceae). *Mitochondrial DNA Part B Resources*, **3**, 719–720.

Selvi, V.M., **Nirmalakumari, A.** & **Senthil, N.** (2015) Genetic diversity for zinc, calcium and iron content of selected little millet genotypes. *Journal of Nutrition & Food Sciences*, **5**, 1–5.

Serba, D.D. & **Yadav, R.S.** (2016) Genomic tools in pearl millet breeding for drought tolerance: status and prospects. *Frontiers in Plant Science*, **7**, 1724.

Simão, F.A., **Waterhouse, R.M.**, **Ioannidis, P.**, **Kriventseva, E.V.** & **Zdobnov, E.M.** (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, **31**, 3210–3212.

Singh, S.P., **Keller, B.**, **Gruissem, W.** & **Bhullar, N.K.** (2017) Rice NICOTIANAMINE SYNTHASE 2 expression improves dietary iron and zinc levels in wheat. *Theoretical and Applied Genetics*, **130**, 283–292.

Sivakumar, S., **Mohan, M.**, **Franco, O.L.** & **Thayumanavan, B.** (2006) Inhibition of insect pest α-amylases by little and finger millet inhibitors. *Pesticide Biochemistry and Physiology*, **85**, 155–160.

Smith, S.A. & **Dunn, C.W.** (2008) Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics*, **24**, 715–716.

Solé, V.A., **Papillon, E.**, **Cotte, M.**, **Walter, P.** & **Susini, J.** (2007) A multiplatform code for the analysis of energy-dispersive X-ray fluorescence spectra. *Spectrochimica Acta Part B: Atomic Spectroscopy*, **62**, 63–68.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

Suyama, M., **Torrents, D.** & **Bork, P.** (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**, W609–W612.

Taylor, J.R.N. & **Kruger, J.** (2016) Millets. In: Caballero, B., Finglas, P.M. & Toldrá, F. (Eds.) *Encyclopedia of Food and Health*. Oxford: Academic Press, pp. 748–757.

Thielen, P.M., **Pendleton, A.L.**, **Player, R.A.**, **Bowden, K.V.**, **Lawton, T.J.** & **Wisecaver, J.H.** (2020) Reference genome for the highly Transformable-Setaria viridisME034V. *G3 (Bethesda)*, **10**, 3467–3478.

Upadhyaya, H.D., **Dwivedi, S.L.**, **Singh, S.K.**, **Singh, S.**, **Vetriventhan, M.** & **Sharma, S.** (2014) Forming core collections in barnyard, kodo, and little millets using morphoagronomic descriptors. *Plant Genetic Resources*, **54**, 2673–2682.

Varmus, H. (2002) Genomic empowerment: the importance of public databases. *Nature Genetics*, **32**, 3.

Vert, G., **Barberon, M.**, **Zelazny, E.**, **Séguéla, M.**, **Briat, J.-F.** & **Curie, C.** (2009) *Arabidopsis* IRT2 cooperates with the high-affinity iron uptake system to maintain iron homeostasis in root epidermal cells. *Planta*, **229**, 1171–1179.

Vert, G., **Briat, J.F.** & **Curie, C.** (2001) *Arabidopsis* IRT2 gene encodes a root-periphery iron transporter. *The Plant Journal: For Cell and Molecular Biology*, **26**, 181–189.

Vert, G., **Grotz, N.**, **Dédaldéchamp, F.**, **Gaymard, F.**, **Guerinot, M.L.**, **Briat, J.-F.** ***et al.*** (2002) IRT1, an *Arabidopsis* transporter essential for iron uptake from the soil and for plant growth. *The Plant Cell*, **14**, 1223–1233.

Vetriventhan, M., **Azevedo, V.C.**, **Upadhyaya, H.D.**, **Nirmalakumari, A.**, **Kane-Potaka, J.**, **Anitha, S.** ***et al.*** (2020) Genetic and genomic resources, and breeding for accelerating improvement of small millets: current status and future interventions. *The Nucleus*, **63**, 217–239.

Vogelgesang, M., **Chilingaryan, S.**, **dos_Santos Rolo, T.** & **Kopmann, A.** (2012) UFO: A scalable GPU-based image processing framework for online monitoring. In: *In 2012 IEEE 14th International Conference on High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems*. Liverpool, UK: IEEE, pp. 824–829.

Vogelgesang, M., **Farago, T.**, **Morgeneyer, T.F.**, **Helfen, L.**, **dos Santos Rolo, T.**, **Myagotin, A.** ***et al.*** (2016) Real-time image-content-based beamline control for smart 4D X-ray imaging. *Journal of Synchrotron Radiation*, **23**, 1254–1263.

Wagner, G.P., **Kin, K.** & **Lynch, V.J.** (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences*, **131**, 281–285.

Whitt, L., **Ricachenevsky, F.K.**, **Ziegler, G.**, **Clemens, S.**, **Walker, E.**, **Maathuis, F.** ***et al.*** (2020) A curated list of genes that affect the plant ionome. *Plant Direct*, **4**, 1–15.

Wickham, H. (2009) *Elegant graphics for data analysis*. 2nd Edition. New York: Springer.

Willick, I.R., **Stobbs, J.**, **Karunakaran, C.** & **Tanino, K.K.** (2020) Phenotyping plant cellular and tissue level responses to cold with synchrotron-based fourier-transform infrared spectroscopy and X-Ray computed tomography. In: *Plant cold acclimation*. New York, NY: Humana, pp. 141–159.

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.

Zhang, Y., **Xu, Y.-H.**, **Yi, H.-Y.** & **Gong, J.-M.** (2012) Vacuolar membrane transporters OsVIT1 and OsVIT2 modulate iron translocation between flag leaves and seeds in rice. *The Plant Journal*, **72**, 400–410.

Zhao, S., **Ye, Z.** & **Stanton, R.** (2020) Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols. *RNA*, **26**, 903–909.

Zou, C., **Li, L.**, **Miki, D.**, **Li, D.**, **Tang, Q.**, **Xiao, L.** ***et al.*** (2019) The genome of broomcorn millet. *Nature Communications*, **10**, 436.