

DATA NOTE

Open Access



# Transcriptomic analysis of developing sorghum grains to detect genes related to cell wall biosynthesis and remodelling

Camille Costes<sup>1</sup>, Sergi Navarro Sanz<sup>2,3,4</sup>, Caroline Calatayud<sup>2,3</sup>, Alexandre Soriano<sup>2,3</sup>, Hamza Mameri<sup>5</sup>, Nancy Terrier<sup>3\*</sup> and Mathilde Francin-Allami<sup>1</sup>

## Abstract

**Objective** Sorghum (*Sorghum bicolor* (L.) Moench) is the fifth most important grain produced in the world. Interest for cultivating sorghum is increasing all over the world in the context of climate change, due to its low input and water requirements. Like other cultivated cereals, sorghum has significant nutritional value thanks to its protein, carbohydrate and dietary fiber content, these latter mainly consisting of cell wall polysaccharides. This work describes for the first time a transcriptomic analysis dedicated to identify the genes involved in the biosynthesis and remodelling of cell walls both in the endosperm and outer layers of sorghum grain during its development. Further analysis of these transcriptomic data will improve our understanding of cell wall assembly, which is a key component of grain quality.

**Data description** This research delineates the steps of our analysis, starting with the cultivation conditions and the grain harvest at different stages of development, followed by the laser microdissection applied to separate the endosperm from the outer layers. It also describes the procedures implemented to generate RNA libraries and to obtain a normalized and filtered table of transcript counts, and finally determine the number of putative cell wall-related genes already listed in literature.

**Keywords** Sorghum, Grain development, Laser microdissection, Transcriptome, Cell wall polysaccharides

## Objective

Sorghum (*Sorghum bicolor* [L.] Moench) is a "C4" cereal whose interest in agricultural practice lies mainly in its resistance to drought and high temperatures [1].

In addition, the attraction of sorghum grain for human consumption is growing, due to its gluten free and its richness in health promoting macro and micro-nutrients [2]. Cell walls are an important variable in nutritional quality of grains as a dietary fiber intake. They are mainly composed of polysaccharides, whose composition may differ greatly between endosperm and outer layers. Our objective is to gain a better understanding of cell wall polysaccharide deposition during grain filling. For this, we monitored whole gene expression both in endosperm and outer layers during grain development to identify those who could be involved in cell wall biosynthesis and remodelling. While several works focused on the composition of the cell walls of dry mature grains [3–6], little research has

\*Correspondence:

Nancy Terrier  
nancy.terrier@inrae.fr

<sup>1</sup> INRAE, UR1268 BIA Nantes, France

<sup>2</sup> CIRAD, UMR AGAP Institut, 34398 Montpellier, France

<sup>3</sup> CIRAD, INRAE, UMR AGAP Institut, Univ Montpellier, Institut Agro, Montpellier, France

<sup>4</sup> Montpellier Ressources Imagerie, BioCampus, University of Montpellier, CNRS, INSERM, Montpellier, France

<sup>5</sup> UMR IATE, Univ Montpellier, INRAE, Institut-Agro Montpellier, 34060 Montpellier, France



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

concentrated on the development of sorghum grains [7, 8]. The transcriptome of developing grains focusing on cell wall genes has been established for cereals such as barley [9], wheat [10] or rice [11]. In developing sorghum grains, transcriptomic analyses have focused mainly on starch, dhurrin and tannins [1, 12–14]. Transcriptomic analyses specifically targeting the cell wall gene set have been carried out mostly on sorghum stem [15, 16]. Our transcriptomic results will be combined with analyses of polysaccharide contents to reach a comprehensive view of cell wall assembly during sorghum grain development.

### Data description

In this study, MACIA, a sorghum cultivar with white grain without tannin, was used. The plants were sown and grown in greenhouses at CIRAD Montpellier in 2022 according to the conditions described in Table 1 [17]. The grains were harvested in triplicate at 7, 13 and 30 days after flowering (DAF), covering the beginning of development until near maturity, frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  before microdissection step. All the following experiments were carried out under RNase-free conditions. Grains were embedded as blocks in O.C.T (VWR® Q Path® O.C.T. Compound Mounting Medium, PA, USA) using liquid nitrogen to harden the medium [18]. From these blocks, sections of  $18\ \mu\text{m}$  were cut using a Cryostat microtome (MICROM HM 520, Leica, IL, USA) before to be adhered to frame slides (steel frame, PET-membrane  $1.4\ \mu\text{m}$ ; RNase-free.  $4\ \mu\text{m}$ ; RNase-free, Leica, Germany). The sections were fixed on the slide by immersion in a 50 mL tube of iced pure ethanol for 30s. Then, they were immersed for 4 min in a 50 mL tube of iced 70% ethanol to remove O.C.T. Finally, the sections were dried at  $40^{\circ}\text{C}$  for 10 min [19]. Once the slides had dried, outer layers and endosperms were separated by laser microdissection using a Leica LMD7000 laser (Laser parameters: Power 50, Speed 6, Aperture 14, Pulse Frequency 1260, Specimen Balance 12) as previously described [20]. Samples were collected in 50  $\mu\text{L}$  tubes (RNase-free) containing cell lysis buffer (Nucleospin RNA, mini kit for RNA

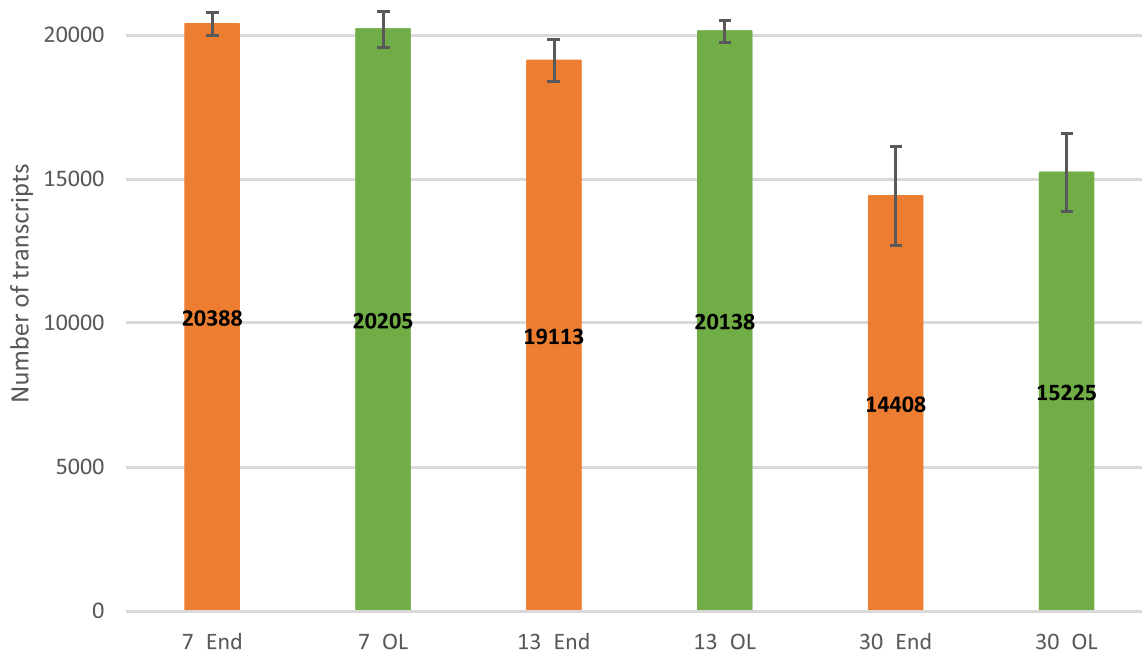
purification, Macherey–Nagel, ref 740,955.50, Germany). The tubes were then frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  for storage prior to RNA extraction. Total RNA was extracted following instructions of the NucleoSpin RNA extraction kit (Macherey–Nagel, Germany). RNA quantity and quality were evaluated by automated electrophoresis (4200 TapeStation system, Agilent Technologies, CA, USA). RNA libraries were prepared according to the Illumina protocol with the TruSeq RNA Library Prep Kit (Illumina, CA, USA). The indexed libraries were pooled in 18-plex and subjected to pair-end  $2\times 150$  bp sequencing on an Illumina HiSeq2500 (GENEWIZ, By Azenta Life Sciences, Germany). The raw sequence files are available on the short-read archive under the bioproject PRJNA1029354 [21]. The software Cutadapt (version 3.5) [22] was used to eliminate low-quality sequences with a cut-off of 30, remaining adapters, as well as resulting reads shorter than 35 bases. Reads were then mapped across the entire sorghum Sbicolor\_313\_v3\_2 genome using HISAT2 [23, 24]. Read count was then performed using featureCounts (from subread 2.0.1 [25]) with the parameters  $-M$   $-fraction$  to count multimapping reads. In our study, reads were aligned to 34129 transcripts among 34496 genes of the sorghum genome (Sbicolor\_453\_v3.2).

The DIANE interactive workflow was used to normalize the data with the TMM method (Trimmed Mean of M values) and to process them [26]. Only genes with a sum of reads under all conditions equal or greater to 100 were conserved, thus resulting to a final count of 20,710 transcripts. The average number of transcripts detected is higher at the early stages of development, at 7 DAF and 13 DAF, compared with the older development stage at 30 DAF (Fig. 1, [27]).

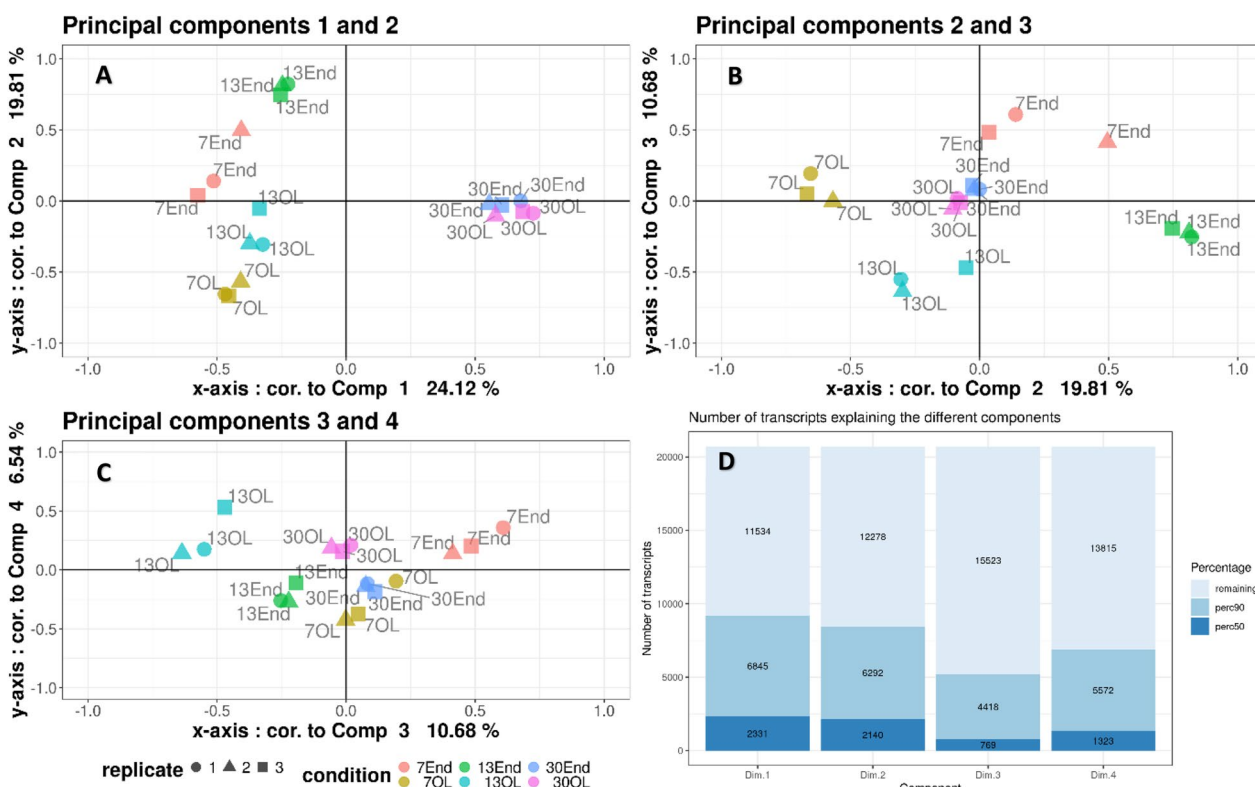
The Principal Component Analysis (PCA) plot of the filtered and normalized RNA-seq data indicated that the biological replicates are well grouped. Components 1 and 2 account for 43.9% of the variance. The component 1 separates the early stages of development (7 and 13 DAF) from the near-mature stage (30 DAF). Tissues are separated from each other by component 2, although this is less obvious for the samples collected at 30DAF (Fig. 2, [28]). A list of 655 sorghum cell wall genes was previously constructed [15, 29]. Among them, 384 were detected in our filtered data from our different samples (Additional file 1, [30]). The quantifications of these cell wall genes, obtained for each biological stage, tissue and replicate, have been detailed in Additional file 1 [30]. These data will be thoroughly analysed and linked to future biochemical composition analyses.

**Table 1** Greenhouse growing conditions for sorghum

Greenhouse conditions	Day	Night
Time	12 h	12 h
Temperature	$28^{\circ}\text{C}$	$22^{\circ}\text{C}$
Sunlight	Natural light complemented by artificial led light ( $700\ \mu\text{mol m}^{-2}\ \text{s}^{-1}$ )	
Moisture	55% average, up to 75% maximum	



**Fig. 1** Number of transcripts expressed in each sample. 7\_End: endosperm at 7 DAF (Days After Flowering); 13\_End: endosperm at 13 DAF; 30\_End: endosperm at 30 DAF; 7\_OL: outer layers at 7 DAF, 13\_OL: outer layers at 13 DAF, 30\_OL: outer layers at 30 DAF. Data are the mean ± Standard deviation of three biological replicates



**Fig. 2** Principal component analysis (PCA) of transcriptomic data from endosperm (End) and outer layers (OL) of sorghum grains at different developmental stages (7, 13 and 30 Days After Flowering (DAF)). A, B, C: representation of samples as functions of components 1 to 4. D: Number of transcripts explaining 50% and 90% of the different components. 7\_End: endosperm at 7 DAF; 13\_End: endosperm at 13 DAF; 30\_End: endosperm at 30 DAF; 7\_OL: outer layers at 7 DAF, 13\_OL: outer layers at 13 DAF, 30\_OL: outer layers at 30 DAF

## Limitations

During laser microdissection, the aleurone layer, made up of differentiated cells located between the endosperm and the outer layers, was mainly recovered with the outer layers. However, endosperm fractions could be contaminated by some aleurone cells during dissection.

## Abbreviations

CIRAD	Centre de coopération internationale en recherche agronomique
DAF	Days After Flowering
RNA	RiboNucleic Acid
PCA	Principal Component Analysis

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12863-024-01198-x>.

**Additional file 1.** List of the cell wall genes sorted from our transcriptomic analysis and their quantification.

## Acknowledgements

We acknowledge the imaging facility MRI, member of the France-BioImaging national infrastructure supported by the French National Research Agency (ANR-10-INBS-04, «Investments for the future»).

## Authors' contributions

MFA, HM and NT conceived the project, obtained funding and supervised the project. NT harvested the seeds. SNS and CCo developed the microdissection protocol and CCo carried out all the microdissection manipulations and performed the RNA extraction. CCo prepared the RNA library. AS, NT and CCo analyzed the final data. CCo wrote the original draft. CCo, MFA, HM and NT reviewed and edited the manuscript. All authors read and approved the final manuscript.

## Funding

This research was supported by TRANSFORM division (AIC SoWallDigest) from the French national institute for agriculture food and environment (INRAE).

## Availability of data and materials

The raw sequence files are available on the short-read archive under the bioproject PRJNA1029354.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 4 December 2023 Accepted: 24 January 2024

Published online: 07 February 2024

## References

- Kang X, Gao W, Cui B, Abd El-Aty AM. Structure and genetic regulation of starch formation in sorghum (*Sorghum bicolor* (L.) Moench) endosperm: a review. *Int J Biol Macromol*. 2023;239:124315–30. <https://doi.org/10.1016/j.ijbiomac.2023.124315>.
- De Morais Cardoso L, Silva Pinheiro S, Stampini Duarte Martino H, Maria Pinheiro-Sant H, De Cardoso Morais L, Duarte Martino Stampini H. Critical Reviews in Food Science and Nutrition Sorghum (*Sorghum bicolor* L.): Nutrients, bioactive compounds, and potential impact on human health Sorghum (*Sorghum bicolor* L.): Nutrients, bioactive compounds, and potential impact on human health. *Food Sci Nutr*. 2017;5(7):372–90. <https://doi.org/10.1080/10408398.2014.887057>.
- Jaworski NW, Lærke HN, Bach Knudsen KE, Stein HH. Carbohydrate composition and in vitro digestibility of dry matter and nonstarch polysaccharides in corn, sorghum, and wheat and coproducts from these grains 1. *J Anim Sci*. 2015;93(3):1103–13. <https://doi.org/10.2527/jas.2014-8147>.
- Takoudjou Miafo AP, Koubala BB, Muralikrishna G, Kansci G, Fokou E. Non-starch polysaccharides derived from sorghum grains, bran, spent grain and evaluation of their antioxidant properties with respect to their bound phenolic acids. *Bioact Carbohydrates Diet Fibre*. 2022;28:100314. <https://doi.org/10.1016/j.bcdf.2022.100314>.
- Ramesh HP, Tharanathan RN. Non-cellulosic mixed linkage  $\beta$ -D-glucan in sorghum, *Sorghum bicolor* (L.) Moench - Localization and biological activity studies. *Indian J Exp Biol*. 2000;38(2):155–9.
- Betts NS, Fox GP, Kelly AM, Cruickshank AW, Lahnstein J, Henderson M, et al. Non-cellulosic cell wall polysaccharides are subject to genotype  $\times$  environment effects in sorghum (*Sorghum bicolor*) grain. *J Cereal Sci*. 2015;63:64–71. <https://doi.org/10.1016/j.jcs.2015.02.007>.
- Kavitha R, Chandrashekar A. Content and composition of nonstarchy polysaccharides in endosperms of sorghum varying in hardness during four developmental stages. *Cereal Chem*. 1997;74(1):22–4. <https://doi.org/10.1094/CCHEM.1997.74.1.22>.
- Ermawar RA, Collins HM, Byrt CS, Henderson M, O'Donovan LA, Shirley NJ, et al. Genetics and physiology of cell wall polysaccharides in the model C4 grass *Setaria viridis* spp. *BMC Plant Biol*. 2015;15(1):1–18. <https://doi.org/10.1186/s12864-020-00421-4>.
- Zhang R, Tucker MR, Burton RA, Shirley NJ, Little A, Morris J, et al. The dynamics of transcript abundance during cellularization of developing Barley Endosperm. *Plant Physiol*. 2016;170:1549–65. <https://doi.org/10.1104/pp.15.01690>.
- Pellny TK, Lovegrove A, Freeman J, Tosi P, Love CG, Paul Knox J, et al. Cell walls of developing wheat starchy endosperm: comparison of composition and RNA-Seq transcriptome. *Plant Physiol*. 2012;158(2):612–27. <https://doi.org/10.3389/fpls.2022.999747>.
- Wu T-Y, Müller M, Gruissem W, Bhullar NK. Genome wide analysis of the transcriptional profiles in different regions of the developing rice grains. Springer, New-York. 2020;13:62–81. <https://doi.org/10.1186/s12284-020-00421-4>.
- Xiao Q, Huang T, Cao W, Ma K, Liu T, Xing F, et al. Profiling of transcriptional regulators associated with starch biosynthesis in sorghum (*Sorghum bicolor* L.). *Front Plant Sci*. 2022;13:999747–78. <https://doi.org/10.3389/fpls.2022.999747>.
- Nielsen LJ, Stuart P, Pičmanová M, Rasmussen S, Olsen CE, Harholt J, et al. Dhurrin metabolism in the developing grain of *Sorghum bicolor* (L.) Moench investigated by metabolite profiling and novel clustering analyses of time-resolved transcriptomic data. *BMC Genomics*. 2016;17(1):1–24. <https://doi.org/10.1186/s12864-016-3360-4>.
- Wang L, Ni D, Yang F, Lin L, Yang Y, Sun C, et al. Transcriptome profiling of *Sorghum bicolor* reveals cultivar-specific molecular signatures associated with starch and phenolic compounds biosyntheses and accumulation during sorghum grain development under a creative commons attribution-non commercial 4.0 I. *Czech J Genet Plant Breed*. 2023;59(4):235–52.
- Hennet L, Berger A, Trabanco N, Ricciuti E, Dufayard JF, Bocs S, et al. Transcriptional regulation of Sorghum stem composition: key players identified through co-expression gene network and comparative genomics analyses. *Front Plant Sci*. 2020;11:1–24. <https://doi.org/10.3389/fpls.2020.00224>.
- McKinley B, Rooney W, Wilkerson C, Mullet J. Dynamics of biomass partitioning, stem gene expression, cell wall biosynthesis, and sucrose accumulation during development of *Sorghum bicolor*. *Plant J*. 2016;88(4):662–80. <https://doi.org/10.1111/TPJ.13269>.
- Costes C, Navarro Sanz S, Calatayud C, Soriano A, Mameri H, Terrier N, Francin-Allami M. Greenhouse growing conditions for sorghum. *Zenodo*. 2024. <https://doi.org/10.5281/zenodo.10470032>
- Brandt R, Mascher M, Thiel J. Laser capture microdissection-based RNA-seq of barley grain tissues. *Methods Mol Biol*. 2018 1723:397–409. [https://doi.org/10.1007/978-1-4939-9747-7\\_17](https://doi.org/10.1007/978-1-4939-9747-7_17)

- link.springer.com/protocol/[https://doi.org/10.1007/978-1-4939-7558-7\\_23](https://doi.org/10.1007/978-1-4939-7558-7_23)
19. Florez-rueda AM, Waser L, Grossniklaus U. Laser-assisted microdissection of plant embryos for transcriptional profiling. *Methods Mol Biol.* 2020;2122:127–39. [https://doi.org/10.1007/978-1-0716-0342-0\\_10](https://doi.org/10.1007/978-1-0716-0342-0_10).
  20. Maurer HC, Olive KP. Laser capture microdissection on frozen sections for extraction of high-quality nucleic acids. *Methods Mol Biol.* 2019;1882:253–9. [https://doi.org/10.1007/978-1-4939-8879-2\\_23](https://doi.org/10.1007/978-1-4939-8879-2_23). Cited 2023 Oct 27.
  21. 'NCBI Sequence Read Archive (<https://identifiers.org/ncbi/insdc.sra:SRP466948>). 2023
  22. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.* 2011;17:10–2. <https://doi.org/10.14806/ej.17.1.200>. Cited 2023 Oct 27.
  23. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements HHS public access. *Nat Methods.* 2015;12(4):357–60. <https://doi.org/10.1038/nmeth.3317>.
  24. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37:907–15. <https://doi.org/10.1038/s41587-019-0201-4>.
  25. Liao Y, Smyth GK, Shi W. Sequence analysis featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923–30. <https://doi.org/10.1093/bioinformatics/btt656>.
  26. Cassan O, Lèbre S, Martin A. SOFTWARE Open Access Inferring and analyzing gene regulatory networks from multi-factorial expression data: a complete and interactive suite. *BMC Genomics.* 2021;22:387–97. <https://doi.org/10.1186/s12864-021-07659-2>.
  27. Costes C, Navarro Sanz S, Calatayud C, Soriano A, Mameri H, Terrier N, Francin-Allami M. Number of transcripts expressed in each sample. Zenodo. 2024. <https://doi.org/10.5281/zenodo.10512775>
  28. Costes C, Navarro Sanz S, Calatayud C, Soriano A, Mameri H, Terrier N, Francin-Allami M. Principal component analysis (PCA) of transcriptomic data from endosperm (End) and outer layers (OL) of sorghum grains at different developmental stages (7, 13 and 30 Days After Flowering (DAF)). Zenodo. 2024. <https://doi.org/10.5281/zenodo.10517823>
  29. Rai KM, Thu SW, Balasubramanian VK, Cobos CJ, Disasa T, Mendu V. Identification, characterization, and expression analysis of cell wall related genes in *Sorghum bicolor* (L.) moench, a food, fodder, and biofuel crop. *Front Plant Sci.* 2016;7:1–19. <https://doi.org/10.3389/fpls.2016.01287>.
  30. Costes C, Navarro Sanz S, Calatayud C, Soriano A, Mameri H, Terrier N, Francin-Allami M. List of the cell wall genes sorted from our transcriptomic analysis and their quantification. 2024. Zenodo. <https://doi.org/10.5281/zenodo.10517869>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.