UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Botucatu

FCA
FACULDADE DE CIÊNCIAS AGRONÔMICAS
UNESP | BOTUCATU

ANA CAROLINA PICININI PETRONILIO

INNOVATIVE TOOLS TO PREDICT AND IMPROVE SEED PHYSIOLOGICAL QUALITY: CASE STUDIES IN TOMATO AND SOYBEAN

Botucatu

2024

ANA CAROLINA PICININI PETRONILIO

INNOVATIVE TOOLS TO PREDICT AND IMPROVE SEED PHYSIOLOGICAL QUALITY: CASE STUDIES IN TOMATO AND SOYBEAN

Thesis presented to the São Paulo State University (UNESP), School of Agricultural Sciences, Botucatu to obtain the title of Doctor of Agronomy – Agriculture.

Advisor: Edvaldo Aparecido Amaral da Silva

Botucatu

2024

P497i

Petronilio, Ana Carolina Picinini

Innovative tools to predict and improve seed physiological quality : case studies in tomato and soybean / Ana Carolina Picinini Petronilio. -- Botucatu, 2024

79 p. : il.

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agronômicas, Botucatu

Orientador: Edvaldo Aparecido Amaral da Silva

1. Seed physiology. 2. Seed technology. 3. Environmental stresses. 4. Chlorophyll. 5. Seed priming. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agronômicas, Botucatu. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Botucatu

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: INNOVATIVE TOOLS TO PREDICT AND IMPROVE SEED PHYSIOLOGICAL QUALITY: CASE STUDIES IN TOMATO AND SOYBEAN

AUTORA: ANA CAROLINA PICININI PETRONILIO
ORIENTADOR: EDVALDO APARECIDO AMARAL DA SILVA

Aprovada como parte das exigências para obtenção do Título de Doutor em Agronomia (Agricultura), pela Comissão Examinadora:

Prof. Dr. EDVALDO APARECIDO AMARAL DA SILVA (Participaçao Presencial)
Producao Vegetal / Faculdade de Ciencias Agronomicas de Botucatu

Pesquisadora Dr.ª CLÍSSIA BARBOZA MASTRANGELO (Participaçao Virtual)
Centro de Energia Nuclear na Agricultura / Universidade de São Paulo

Prof.ª Dr.ª JULIANA JOICE PEREIRA LIMA (Participaçao Virtual)
Centro de Ciências Agrárias / Universidade Federal de São Carlos

Prof.ª Dr.ª MARIA PAULA BARION ALVES NUNES (Participaçao Presencial)
Produção Vegetal / Faculdade de Ciencias Agronomicas de Botucatu

Voluntário Livre-Docente JOÃO NAKAGAWA (Participaçao Presencial)
Produção Vegetal / Faculdade de Ciências Agronômicas de Botucatu

Botucatu, 01 de março de 2024

Aos meus amados pais,

Sandra e Rosival,

dedico.

ACKNOWLEDGMENTS

To God for being my daily support during this journey. To my parents for all their support and for believing in my dreams even though I need to be hundreds or thousands of kilometers away to make them come true. To my family, especially my uncles and great encouragers, Ronaldo Picinini, Ricardo Picinini, José Naief Tayer, and Márcio Jampani.

During these five years of graduate school, I was lucky enough to meet many incredible people, have previously unimaginable experiences, and have dreams come true. Therefore, I thank each person who contributed to my personal and professional development and the completion of this work.

To Prof. Dr. Edvaldo Ap. Amaral da Silva, for his guidance, encouragement, and teachings. For teaching me the importance of always seeking the practical applicability of our research and for introducing me to molecular biology. Besides being a mentor, he was also a friend many times.

To Prof. Dr. João Nakagawa for his teachings and for being a professional inspiration who contributed so much to the seed sector.

To Thiago Batista and Gustavo Fonseca for their friendship and teachings. For encouraging me always to strive to be better, and for the work developed in partnership that generated the results of this thesis.

To Dennis Souza for all the friendship and fun hours of work shared in the laboratory. More than a friend, Dennis became a brother.

Isabela Lopes for all her help in conducting the experiments.

To my friends and colleagues at the Seed Physiology and Molecular Biology Laboratory: Samara Perissato, Carolina Cardoso, Larissa Chamma, Renata Arruda, Brunna Rithielly, Taíse Rodrigues, Vitor Gonçalves, Yago Triboni, Rômulo Lima, Mónica Córdoba, Michelane Lima, Felipe Carvalho and Tamires Freitas, for scientific learning, help with experiments and emotional support throughout these years. Having you made the exhausting graduation routine lighter and more enjoyable.

To Ruan Marchi for his friendship, all the laughter in the department corridors, and all his teachings.

To Lennis Afraire, Maria Gabriela Andrade and Aline Oliveira for their friendship.

To Dr. Clíssia Mastrangelo and Cena/USP for the partnership that culminated in the results of this thesis.

To Dr. Fernando Henning and Dr. Francisco Krzyzanowski for their guidance during my undergraduate internship, teaching me a passion for research, and encouraging me to pursue a graduate degree.

To Dr. Monica Schmidt for welcoming me as a visiting student, and for all the teachings given.

To my colleagues at Schmidt's lab: Mariam Marand, Cyrus Mou, Alexis Grace, Huang Yicheng, and Greg Clark, for welcoming me. Thank you for all your help and friendship throughout these 12 months in Tucson.

To my Brazilian friends and colleagues whom I met in Arizona, Andriele Muller, Bianca Ferreira, Ramon Domingues, Felipe Dorigão, Diego Costa, Sara Paim, Caíque Nonato, George Barbosa and Vanessa Mendonça.

To Prof. Dr. Carlos Alexandre Costa Crusciol and for all the encouragement throughout this journey.

To FCA employees Eliane, Iara, Adelina, and Talita. In particular, to Valéria Giandoni, you became more than just a laboratory technician; you became a friend. Thank you for all the care and affection. I will never forget the endless help with experiments and all the delicious cakes and lunches that I will never forget.

To the examining board members for their availability and for contributing to this important step in my life.

To the Plant Production Department FCA UNESP Botucatu for the structure and for being a pleasant working environment.

To The University of Arizona for providing the structure where I carried out part of my thesis.

To the Coordination for the Improvement of Higher Education Personnel – Brazil – CAPES for granting a scholarship. Processes: DS 88887.670559/2022-00; CAPES/PRINT 88887.685215/2022-00; DS 88887.924553/2023-00.

“[1]Filho meu, se aceitares as minhas palavras e esconderes contigo os meus mandamentos, [2]para fazeres atento à sabedoria o teu ouvido e para inclinares o coração ao entendimento, [3]e, se clamares por inteligência, e por entendimento alçares a voz, [4]se buscares a sabedoria como a prata e como a tesouros escondidos a procurares, [5]então, entenderás o temor do Senhor e acharás o conhecimento de Deus. [6]Porque o Senhor dá a sabedoria, e da sua boca vem a inteligência e o entendimento." (Provérbios 2:1-6).

BÍBLIA. Bíblia Sagrada. Traduzida por João Ferreira de Almeida. Barueri, SP: Sociedade Bíblica do Brasil, 1999.

ABSTRACT

The seed quality is made up of physical, sanitary, genetic, and physiological attributes. Seed physiological quality refers to factors that allow rapid and uniform germination and emergence under broad climatic conditions, which ensures vigorous initial plant establishment and produces an increased yield. This research sought to elucidate how the physiological quality of seeds can be predicted and improved. To this end, this research used soybean and tomato species as models. 1 – the physiological quality of soybean seeds produced under water and heat stress during the maturation phase was determined, and seed multispectral images were acquired. Machine learning models associated with multispectral imaging technology were developed to verify the possibility of separating seed groups autonomously. The quality of stressed seeds was lower compared to non-stressed ones. The multispectral images produced markers that made it possible the segmentation between seed groups. The ML models showed high performance in recognizing stressed seeds. This work opens up the possibility of using this technology as a results manager, as it provides information about the environment of seed production and its consequences on the physiological quality of the seeds. 2 – we performed soybean genetic transformations for further study of the molecular factors involved in seed chlorophyll degrading/retention. For this, the GmABI5 gene was overexpressed and silenced (RNAi) using the glycinin promoter in a hygromycin-resistant expression cassette. Somatic embryos were induced from the cotyledons of immature seeds and when they reached the globular stage they were transformed via bioballistics. Transformed embryos were selected for resistance to hygromycin. The presence of the expression cassette in the resistant embryos was confirmed via PCR and they were regenerated into seedlings. This work will allow the functional study of the ABI5 gene and also the investigation into the possibility of reproducing a genotype tolerant to chlorophyll retention in soybean seeds. 3 – we performed gene expression studies in tomato seeds during osmo-priming to investigate mechanisms involved in the reduction of longevity. For this, we collected a total of seven samples: four samples during the 60 hours of priming, one control group (without priming), one sample after post-priming drying, and one after post-priming thermal shock and drying. We investigated transcripts associated with stress response. The seed vigor and longevity after treatment were also determined. Primed seeds had their longevity impaired compared to unprimed ones. Genes from the heat-shock

protein family were down-regulated during the priming process. This research brings new insights into the mechanisms involved in the reduction of longevity of primed seeds and allows the use of these transcripts to monitor longevity in primed tomato seeds.

Keywords: seed quality; multispectral imaging; environmental stress; chlorophyll fluorescence; primed seeds.

RESUMO

A qualidade de sementes é composta por atributos físicos, sanitários, genéticos e fisiológicos. A qualidade fisiológica de sementes refere-se a fatores que permitem que a germinação e a emergência ocorram de forma rápida e uniforme sob amplas condições climáticas, o que garante um estabelecimento inicial vigoroso das plantas e resulta em aumento de produtividade. Esta pesquisa buscou investigar como a qualidade fisiológica das sementes pode ser estimada e aprimorada. Para tanto, esta pesquisa teve como modelo as espécies de soja e tomate. 1 – foi determinada a qualidade fisiológica de sementes de soja produzidas sob estresse hídrico e térmico durante a fase de maturação e adquiridas imagens multiespectrais das sementes. Modelos de aprendizado de máquina associados à tecnologia de imagem multiespectral foram desenvolvidos para verificar a possibilidade de separar os grupos de sementes de forma autônoma. A qualidade das sementes estressadas foi inferior em comparação às não estressadas. As imagens multiespectrais produziram marcadores que possibilitaram a segmentação dos grupos de sementes. Os modelos de aprendizado de máquina apresentaram alto desempenho no reconhecimento de sementes estressadas. Este trabalho possibilita a utilização desta tecnologia como gestora de resultados, pois fornece informações sobre o ambiente de produção de sementes e suas consequências na qualidade fisiológica das sementes. 2 – Foram realizadas transformações genéticas em soja para futuramente investigar os fatores moleculares envolvidos na degradação/retenção da clorofila nas sementes. Para isso, o gene *Gm*ABI5 foi superexpresso e silenciado (RNAi) utilizando o promotor glicinina em um cassete de expressão resistente à higromicina. Embriões somáticos foram induzidos a partir de cotilédones de sementes imaturas e quando os mesmos atingiram o estágio globular foram transformados via biobalística. Os embriões transformados foram selecionados quanto à resistência à higromicina. A presença do cassete de expressão nos embriões resistentes foi confirmada via PCR e eles foram regenerados em plântulas. Este trabalho permitirá o estudo funcional do gene ABI5 e a investigação da possibilidade de reproduzir um genótipo tolerante à retenção de clorofila em sementes de soja. 3 – foram realizados estudos de expressão gênica em sementes de tomate durante o condicionamento osmótico (*osmo-priming*) para investigar mecanismos envolvidos na redução da longevidade. Para isso, foram coletados um total de sete amostras: quatro amostras durante as 60 horas de priming,

uma amostra do grupo controle (sem priming), uma amostra após secagem pós-priming e uma após choque térmico pós-priming e secagem. Foi investigada a expressão gênica de moléculas associadas à resposta ao estresse. O vigor e a longevidade das sementes após o tratamento também foram determinados. As sementes condicionadas tiveram sua longevidade reduzida em comparação às não condicionadas. Os genes da família das proteínas de choque térmico foram regulados negativamente durante o processo de *priming*. Esta pesquisa traz novas percepções sobre os mecanismos envolvidos na redução da longevidade de sementes condicionadas e permite o uso destes genes para monitorar a longevidade em sementes de tomate submetidas ao *priming*.

Palavras-chave: qualidade de sementes; imagem multiespectral; estresse ambiental; fluorescência da clorofila; sementes condicionadas.

SUMMARY

GENERAL INTRODUCTION

In agriculture, seeds are one of the main inputs (Carvalho; Nakagawa, 2012). A new generation of plants begins through seeds, assuring plant life and food security worldwide. Much attention is given to producing high-quality seeds because it supports crop production. Seed quality consists of genetic, physical, sanitary, and physiological attributes. Physiological quality refers to viability, desiccation tolerance, vigor, and longevity, which ensure proper seedling establishment (Bewley et al., 2013).

Seed physiological quality can impact crop yield by affecting the speed and uniformity of germination and emergence and plant density (Tekrony et al., 1991). Low-quality seed lots result in a higher rate of abnormal seedlings, reducing population density (considering density recommended limits) and impacting productivity. On the other hand, high-quality seeds produce seedlings that emerge earlier and uniformly during a suitable sowing time. Seedling emergence during the suitable sowing time ensures that environmental conditions will be more favorable for growth than if the seedling emerges later, when the climate conditions might be less advantageous. Fast and uniform emergence also avoids competition between crop plants and weeds (Finch-Savage, 1995). Seed physiology studies have contributed to technology development and crop yield increase. High-quality seeds improved crop yield by more than 4% in corn production (Graven; Carter, 1991). According to Rajala et al. (2011), barley seeds with superior quality improved grain yield. High-quality seeds improved rice yield up to 19% in the Philippines (Diaz et al., 1998). The progressive increase in the level of vigor in soybean seeds led to an increase of up to 28 $kg.ha^{-1}$ in grain productivity (Bagateli et al., 2019). In this regard, molecular-level seed physiology research, seed priming, and imaging techniques are ways to predict and improve seed quality and will be addressed in this research.

In recent years, imaging analysis has been proposed as an alternative to traditional seed tests to assess seed quality. The traditional tests depend mainly on a visual inspection by the analyst, which can lead to a subjective interpretation of the results. Furthermore, these methods are destructive and mostly time-consuming. Therefore, there is a need to advance the use of more assertive tools, which promote time savings and eliminate subjectivity, and which additionally allow the integration of machine learning and process automation. Several imaging technologies are available nowadays, such as X-ray, X-ray microtomography, X-ray fluorescence, near-infrared

spectroscopy, and magnetic resonance imaging that can assess seeds (Bianchini et al., 2021; Cotrim et al., 2019; Gomes-Junior et al., 2019) and multispectral imaging that can assess both seeds and seedlings (Elmasry et al., 2019; Galletti et al., 2020). X-rays were used to detect mechanical damage in several species, such as beans (Gomes-Junior et al., 2019; Mondo et al., 2009), and were associated with seed physiological quality. Autofluorescence multispectral images associated with artificial intelligence tools have been used to recognize soybean maturation stages and predict seed quality (Barboza Da Silva et al., 2021; Batista et al., 2022). Multispectral imaging and machine learning methods can also be used to provide information regarding pigments. Chl fluorescence and anthocyanin in peanut seedlings were reported as efficient in segregating seeds with superior quality (Fonseca De Oliveira et al., 2022).

Seed physiological quality is a multigenic trait and is largely influenced by the environment (Tripathi; Khare, 2016). Over the years, in plant breeding programs, much attention has been given to improving grain quality traits such as oil and protein quality and yield (Rajcan; Hou; Weir, 2005), grain yield, and insect and disease resistance (Krzyzanowski, 1998). However, little attention is given to associating the selection of the aforementioned traits with seed physiological quality traits (Kalaji; Pietkiewicz, 2004). There is a need to further research seed physiological quality since it is essential to enable the expression of the productive potential of cultivars. Some researchers have studied the genetic gains regarding seed physiological quality in breeding research. In soybean, Monteiro et al. (2021) showed that the heritability of physiological quality traits was moderately high, which confirms the possibility of breeding to improve seed physiological quality traits. Mello Filho et al. (2004) highlighted the possibility of breeding to improve protein quality while keeping high-quality seeds and satisfactory yield. Hatzig et al. (2018) suggest the use of molecular markers of mean germination time for selection of germination capacity in oilseed rape cultivars.

Molecular-level seed physiology studies aim to investigate the molecular factors involved in the seed physiological quality attributes. This approach has been applied in the soybean “greenish-seed problem” research. Soybean greenish seeds are mature seeds with chlorophyll (chl) retention. This usually happens in response to abiotic stresses during the soybean maturation phase. Chl retention has been a significant problem in these last decades in Brazil because it impacts seed and oil quality (Pádua et al., 2007; Luccas, 2018). Besides the environmental circumstances,

genetic factors are also involved in the susceptibility of chl retention in soybean seeds (Ajala-Luccas et al., 2023; Pádua et al., 2009; Teixeira et al., 2016). Teixeira et al. (2016) suggested that in susceptible genotypes, the impaired expression of genes related to photosynthesis (D1, D2. PPH2, NYC1) might be involved in chlorophyll retention. Batista (2022) showed that the acquisition of longevity is impacted in stressed seeds with higher chl content. In the same study, the expression of genes related to chl degradation (D2, NYC1) is impacted in immature stressed seeds. Moreover, seeds that suffered stresses during the maturation phase had a higher expression of the HSP21 gene, which is a protection mechanism against stresses, compared to non-stressed seeds (Batista, 2022). Another approach for molecular-level seed physiology studies was presented by Ducatti et al. (2022) who reported transcripts associated with vigor in soybean seeds.

Thus, the above-mentioned transcripts can be used to monitor the quality attributes and provide knowledge that can be applied in plant breeding programs aiming at marker-assisted selection of superior genetic materials for seed quality. In addition, the genes discovered through molecular-level seed physiology studies and plant breeding might also be applied in genetic transformation and gene editing, which are additional tools that might improve seed quality. Genetic transformation and gene editing can be applied in gene functional studies to understand the behavior of a candidate gene, and furthermore, it can also provide genetically modified varieties for high-quality seed production.

Seed priming is a technology used to increase the speed and uniformity of germination and improve seedling emergence and stress tolerance through controlled seed hydration. The germination process starts during the treatment, but radicle protrusion is impeded. Thus, seeds are dried after a determined priming time. Therefore, germination and seedling emergence are faster when the seeds are sown because the metabolism was previously activated during priming treatment. Priming is particularly helpful in seeds with slow and uneven germination (Marcos-Filho, 2015). Besides the increase in seed physiological quality by priming (Brocklehurst; Dearman, 1983; Forti et al., 2020; Silveira et al., 2023), studies show crop yield improvement. Salicylic acid priming promoted a fruit yield increase of 33% and improved plant resilience to water stress in tomatoes and also improved yield and high-temperature germination tolerance in carrots (Chakma et al., 2021; Mahmood-Ur-Rehman et al., 2020). In wheat seeds, osmo-priming increased grain productivity by 1.10 tons per

hectare (Farooq et al., 2020). Hussain and collaborators (2006) obtained increases in the productivity of sunflower achenes through NaCl and KNO_3 osmo-priming.

Thus, studies of imaging analysis, molecular-physiology level studies, and seed priming, are tools that can predict and improve seed quality (Paparella et al., 2015; Fonseca De Oliveira et al., 2022; Ducatti, et al., 2022). Therefore, this research aims to study how seed physiological quality can be predicted and improved through innovative tools using two economically important crops as models, tomato, and soybean.

CHAPTER 1

MAPPING FEATURES TO SEGMENT SOYBEAN STRESSED SEEDS USING MACHINE LEARNING MODELS BASED ON MULTISPECTRAL IMAGING: AN ACCURATE TOOL TO SOLVE CLASSIC PROBLEMS

ABSTRACT

Extreme environmental conditions have been recurrent during the last few years and have impacted crop seed quality, mainly but not limited to, soybeans (Glycine max (L) Merrill). To overcome this, seed companies often demand innovative tools to address seed quality factors. Machine learning models based on multispectral imaging are a novel seed quality analysis approach. Thus, we hypothesize that segmenting stressed and non-stressed soybean seeds would be possible with this technology, opening a new opportunity for seed quality management and elucidating quality factors. Soybean seeds (cultivar BR/MG 46-Conquista) were produced under water deficit and heat during maturation (from R5.5 onwards). Multispectral images were acquired from stressed and non-stressed seeds, and the reflectance, autofluorescence, physical properties, and chlorophyll parameters were extracted from the images. In parallel, we determined seed vigor. We designed machine learning models using multispectral imaging data based on three algorithms: neural network, support vector machine, and random forest. Our results demonstrated that the stressed seeds have spectral markers that enable their recognition. Concomitantly, these markers had a direct relationship with seed vigor. The machine learning models developed based on neural network algorithm showed the highest performance in segmenting stressed seeds (90% accuracy). Here, we report a new approach for multispectral imaging with the potential to identify soybean seeds of lower vigor as a result of unfavorable environmental conditions during seed maturation.

1.1 INTRODUCTION

Soybean is a significant source of protein and oil and is utilized in several industrial products (Contini et al., 2013). Due to the many applications of this crop, a strong market has developed worldwide (USDA, 2023). Many inputs are required to boost the soybean chain, and one of these inputs is seeds. Seeds are responsible for establishing a new set of plants in the field. High-quality seeds are primordial to providing uniform plant development that favors maximum crop productivity

(Krzyzanowski et al., 2018). To ensure high-quality seeds, seed companies strive and constantly seek innovations to overcome seed quality challenges.

One of the potential new approaches to be used in seed quality management is multispectral imaging associated with artificial intelligence. Multispectral imaging consists of reflectance or autofluorescence images obtained through the emission of several light wavelengths (Galletti et al., 2020). This technique can provide optical markers regarding seed chemical composition, pigments, and physical traits such as texture, area, length, width, color, and coat brightness (Bianchini et al., 2021; França-Silva et al., 2023). Intact seeds in the samples provide precise data since the image pixels are transformed into numbers through an algorithm, and consequently eliminate the subjectivity of human interpretation, typical of traditional seed quality tests (germination, accelerated aging, tetrazolium, etc.) (Batista et al., 2022; Galletti et al., 2020). The data generated combined with machine learning (ML) models allow the use of this technique in seed processing to evaluate their quality autonomously (Fonseca De Oliveira et al., 2022).

Multispectral imaging has recently been the object of study of several crop seed quality management programs. For example, Barboza da Silva et al. (2021b) showed that autofluorescence signals from the 365/400 nm wavelength associated with ML models could identify artificially aged soybean seeds. Chlorophyll (Chl) fluorescence, anthocyanin, 660, 690, and 780 nm light reflectance, and physical attributes associated with ML models were used to select high-quality peanut seeds (Fonseca de Oliveira et al., 2022). Autofluorescence-spectral imaging related to Chl was used to segment soybean seeds in different maturation stages (Batista et al., 2022). This technique was also used to identify viability, vigor, and hard seeds from Medicago sativa (Wang et al., 2023; Zhang et al., 2022). Texture, color, and 490 nm reflectance associated with artificial intelligence tools were used to discriminate pathogenic fungi in peanut seeds (Sudki et al., 2023).

The soybean seed market deals with several aspects affecting seed quality. In recent decades, reports show that soybean seeds have germination, vigor, and longevity impacted due to abiotic stress (Ajala-Luccas et al., 2023; Batista et al., 2022; Teixeira et al., 2016). One of the reported stresses that impact seeds is a combination of water deficit and heat (Pádua et al., 2009; Teixeira et al., 2016), and the occurrence of this stress during the plant maturation phase was outlined as being more critical than during the vegetative or embryogenic phases (Cohen et al., 2021). Environmental

stress negatively impacts seed physiological quality (Ajala-Luccas et al., 2023; Zorato et al., 2007) and is one of the challenges seed companies must overcome to ensure a high-quality product. Therefore, a tool that enables seed companies to know the causes of seed quality issues quickly and non-destructively would be helpful for seed quality management. While stress affects seed quality through physical and physiological changes, it could be possible to map this event beyond what the human eye can see. Thus, it would be possible to promptly identify reductions in seed quality associated with environmental stress.

Thus, we hypothesize that ML models based on multispectral imaging can be applied in the segmentation of stressed and non-stressed seeds as a tool for companies to keep track of the seed quality factors. To confirm this, we investigated multispectral imaging features and ML models that could effectively recognize abiotic stressed soybean seeds.

1.2 MATERIAL AND METHODS

Plant material

Soybean seeds from a cultivar susceptible to abiotic stresses (BR/MG 46-Conquista) (Pádua et al., 2009) were propagated in a greenhouse at the School of Agriculture, Department of Crop Sciences, UNESP/Botucatu/São Paulo/Brazil (Altitude 810m). Supplementary lighting was provided through LED full spectrum lamps (28W), with photosynthetically active radiation emission of 150 μmols m−2 s-1. The plants were cultivated in 180 pots with 11 L of soil with texture of 778 g kg-1 of sand. The soil acidity and fertilization were adjusted according to the crop needs.

The mean temperature during crop development was 28 ± 1 °C during the day and 21 °C ± 2 at night. The humidity of the pots was kept close to the field capacity (12 kPa), which had been previously determined. The humidity was controlled twice a day (10:00 am and 4:00 pm). It was measured with a humidity sensor positioned at a depth of 10 cm. If there was a difference between the reading and the field capacity humidity, the corresponding water volume was replaced to return the soil to the desired water potential (12 kPa).

The sample groups were composed of different environmental conditions during the seed maturation phase: non-stressful (control) and stressful conditions. From the R5.5 stage (Ritchie, Hanway, Thompson, 1982) onwards, 50% of the pots were randomly assigned to stressful environmental conditions (stressed group), and the

other 50% were kept in the conditions above and constituted the control group (non-stressed group). The environmental stress condition constituted a daytime average of 33 ± 1 ºC and nighttime 25 ± 0.4 ºC. Water supplementation was kept to a minimum so as to maintain a water potential of 27 kPa (water deficit) until the R9 stage (full maturation). The R5.5 stage approximately coincides with the end of the seed maturation phase (end of grain filling and disconnection from the mother plant) and with the beginning of the late maturation, an important phase for the seed physiological quality acquisition, mainly in terms of desiccation tolerance, vigor, and longevity (Lima et al., 2017).

On the seventh day, after the plants were subjected to stressful conditions, the third wholly expanded leaf from the top of the plants was collected to characterize their oxidative stress. The oxidative stress was determined with four replicates per treatment (stressed and non-stressed), constituting one leaf per replicate. The samples were immediately frozen in liquid nitrogen and stored in an ultra-freezer (-80 ºC) until the evaluation.

For hydrogen peroxide quantification (H2O2), 250 mg of leaf tissue was ground and homogenized in 3 mL of 0.1% (m/v) trichloroacetic acid (TCA) solution with 20% PVPP (m/m). The solution was centrifuged at 14,000 g at 4 ºC for 20 minutes (Hettich, Universal 320R, Tuttlingen, Germany). A 0.2 mL aliquot of the supernatant was added to the reaction of 0.2 mL of 100 mM potassium phosphate buffer solution (pH 7.5) and 0.8 mL of potassium iodide (KI) solution 1 M. The solution was incubated for one hour on ice in the dark. Then, the absorbance reading was performed on a spectrophotometer (Shimadzu, UV-2700, Kyoto, Japan) at 390 nm. The H2O2 concentration was calculated using the H2O2 standard curve at 1,000 µmol mL-1, and the results were expressed in mmol g-1 MF (Alexieva et al., 2001).

For lipid peroxidation (malondialdehyde – MDA), 250 mg of leaf tissue were ground in liquid nitrogen, homogenized in 3 mL of 0.1% (m/v) trichloroacetic acid (TCA) solution with 20% PVPP (m/m), and centrifuged at 14,000 g at 4 ºC, for 20 minutes (Hettich, Universal 320R, Tuttlingen, Germany). A 0.5 ml aliquot of the supernatant was added to 1.5 ml of 0.5% (w/v) 2-thiobarbituric acid (TBA) solution prepared in 20% TCA (w/v) and incubated at 90ºC for 20 minutes. The reaction was stopped by putting the tube in an ice bath for 10 minutes. Afterward, the samples were centrifuged at 10,000 g at 4 ºC for 5 minutes. Absorbances were determined using a spectrophotometer (Shimadzu, UV-2700, Kyoto, Japan) at 532 and 660 nm. The

malondialdehyde acid (MDA) concentration was calculated using the molar extinction coefficient of 155 mM-1 cm-1 and expressed as MDA nmol g-1 MF (Cakmak et al., 1991).

The seeds were harvested at the R9 stage (Ritchie et al., 1982), and unformed seeds were discarded. The seeds were stored in a cold chamber at 12 ºC/55% RH for one week, and then the tests were carried out.

Seed physiological quality assays

Before carrying out the seed quality tests, the water content was determined by the oven method (ISTA, 2020). The water content of all treatments was balanced to 0.15 g H2O/g DW-1 by pre-conditioning the seeds on a stainless-steel screen suspended inside a plastic box (11.0 x 11.0 x 3.5 cm) containing 40 mL of deionized water at 25 ºC for 16 hours.

Four replicates of 100 seeds were weighed on an analytical scale with a precision of 0.001 g to determine the weight of 100 seeds. To evaluate the seedling emergence and uniformity index, four replicates of 25 seeds were sown at a depth of 5 cm in 200-cell trays with a commercial substrate composed of sphagnum peat, expanded perlite, expanded vermiculite, and roasted rice hulls. The substrate was moistened with 60% of its water-holding capacity. The seedlings were cultivated under an average temperature of 25 ºC ± 0.8. Emerged seedlings were counted daily at the same time until the number of emerged seedlings stabilized (Krzyzanowski et al., 2020). For seedling emergence, data were expressed as a percentage of emerged seedlings. The uniformity index was expressed through the sum of the percentage of seedlings that emerged on the emergence peak day plus the day before and after the emergence peak, according to Ebone et al. (2020). For the accelerated aging test (AA), four replicates of 25 seeds were accommodated on a stainless-steel screen inside a plastic box (11.0 x 11.0 x 3.5 cm) containing 40 mL of deionized water. The boxes were sealed to avoid evaporation and put in an aging chamber at 41 ºC for 72 h. Afterward, a germination test was performed on paper towels moistened with a water quantity equivalent to 2.5 times the paper's weight. The paper rolls were kept at 25 ºC for five days in the dark when the normal seedlings were counted.

Multispectral imaging data acquisition

Multispectral images of 250 seeds per treatment were acquired. The seeds were placed on acetate sheets (5 cm x 8.5 cm) with double-sided tape. On each sheet, 50 seeds were accommodated in the same position (hilum facing left), and the images were captured using a VideometerLab4TM equipment (Videometer A/S, Herlev, Denmark), according to Batista et al. (2022).

Reflectance images were acquired at 19 wavelengths – 365 (UV), 405 (violet), 430 (indigo), 450 (blue), 470 (blue), 490 (cyan), 515 (green), 540 (green), 570 (yellow), 590 (amber), 630 (red), 645 (red), 660 (red), 690 (dark red), 780 (dark red), 850, 880, 940, and 970 nm (the last four wavelengths are in the near-infrared region) (Fonseca De Oliveira et al., 2022). Autofluorescence images were acquired using 28 different excitation/emission combinations at wavelengths 365/400, 365/500, 405/500, 430/500, 450/500, 470/500, 365/600, 405/600, 430/600, 450/600, 470/600, 490/600, 515/600, 540/600, 570/600, 365/700, 405/700, 430/700, 450/700, 470/700, 490/700, 515/700, 540/700, 570/700, 590/700, 630/700, 645/ 700, and 660/700 nm (Gomes et al., 2022). Before image acquisition, the light setting was adjusted to optimize the light intensity at each wavelength and allow all captured images to be compared. Representative samples were used for this calibration, which was later applied to all other samples.

After image acquisition, data were extracted using VideometerLabTM software (version 3.14.9). We built a segmentation mask based on the first image to separate the seeds from the background. Thus, the pixels in the background had values equal to zero, while each pixel different from zero was considered part of an object (Barboza Da Silva et al., 2021a). In the segmented images, a normalized canonical discriminant analysis algorithm (nCDA) was applied. The nCDA algorithm reduces the outliers and was used to detect the signals per pixel in each combination of wavelength employed and thus obtain numerical resolution values. In addition to the reflectance and autofluorescence variables, data were determined for the seed physical characteristics: area, length, and width, and for color descriptors: CIELab L*, CIELab a*, CIELab b* (Oliveira et al., 2021). The color descriptors are a way to describe color features in a objective way, thus, CIELab L* corresponds to brightness features, CIELab a* indicates how close an object is to red or green and CIELab b* means how close an object is to yellow or blue (Melgosa, 2000).

With the same sample preparation, we obtained multispectral images (2448 x 2448 pixels) using a SeedReporterTM (PhenoVation B.V., Wageningen, The Netherlands) instrument. The Chl *a* index images were acquired in the 710 and 770 nm reflectance (Gitelson et al., 2003) and the Chl *a* fluorescence was obtained by excitation at 620 nm wavelength and detection at 730 nm (Galletti et al., 2020). The image acquisition and data extraction details were obtained using the SeedReporterTM software version 5.5.1.

Seed segmentation using machine learning models

We designed ML models using the 20 most important variables of the multispectral data from each single seed (n=500 seeds). The models were based on neural network (NN) (solver: Stochastic Gradient Descent; hidden layer sizes: two layers with 25 neurons in each; Activation function: Tanh; Learning rate: adaptive; the maximum number of interactions: 5000) (Barboza Da Silva et al., 2021b), support vector machine (SVM) and random forest (RF) algorithms (Batista et al., 2022), in which 70% of data (n = 350) were automatically separated to train the algorithms (K-fold = 5), and 30% (n = 150) were used for the external validation test.

Statistical design

The physiological (*n*=8) and multispectral (*n*=500) data were compared by t-test ($P<0.05$). Additionally, in the multispectral data, we performed (i) Gini coefficient to rank variables regarding their importance in seed group segmentation, (ii) principal component analysis (PCA) to visualize the spatial seed group segregation (stressed and non-stressed) concerning the variables using “ggbiplot” R package. Accuracy, F1, precision, recall, and specificity were used to test the performance of the ML models generated. All metrics were calculated using a confusion matrix. The Python version 3.11 environment was used to generate the models. Pearson’s correlation coefficient was measured to investigate the relationship between the multispectral markers and the seed physiological quality traits using the “corrplot” R package.

1.3 RESULTS

Seed quality characterization

The hydrogen peroxide (H_2O_2) and lipid peroxidation (MDA) content (SUPPLEMENTARY FIGURE 1A, B) increased due to the stressful conditions in soybean plants. Moreover, seed weight was significantly affected by environmental

stress. Stressed seeds had less mass than the non-stressed seeds, which was 19.62 g and 23.63 g, respectively (FIGURE 1A). Therefore, these results confirmed the effect of plant stress in our research. In parallel, we performed seed physiological quality tests to show seed performance. Stressed seeds had their quality impaired in terms of seedling emergence and emergence uniformity. Stressed seeds had low seedling emergence performance in relation to non-stressed seeds (FIGURE 1B, C), with a substantial difference of 20% in the emergence uniformity, a robust index to measure disturbances in the seed-seedling transition directly related to seed vigor. However, there was no significant difference in the AA test between stressed and non-stressed seeds (FIGURE 1D).

Supplementary figure 1 – Oxidative stress characterization in soybean plants submitted to stressful conditions during maturation. Hydrogen peroxide (H_2O_2) in non-stressed and stressed plants (A). Malondialdehyde acid (MDA) in non-stressed and stressed plants (B). MDA is a product of lipid peroxidation. Different letters indicate a significant difference ($P \leq 0.01$) by t-test. Error bars show the standard deviation from four samples

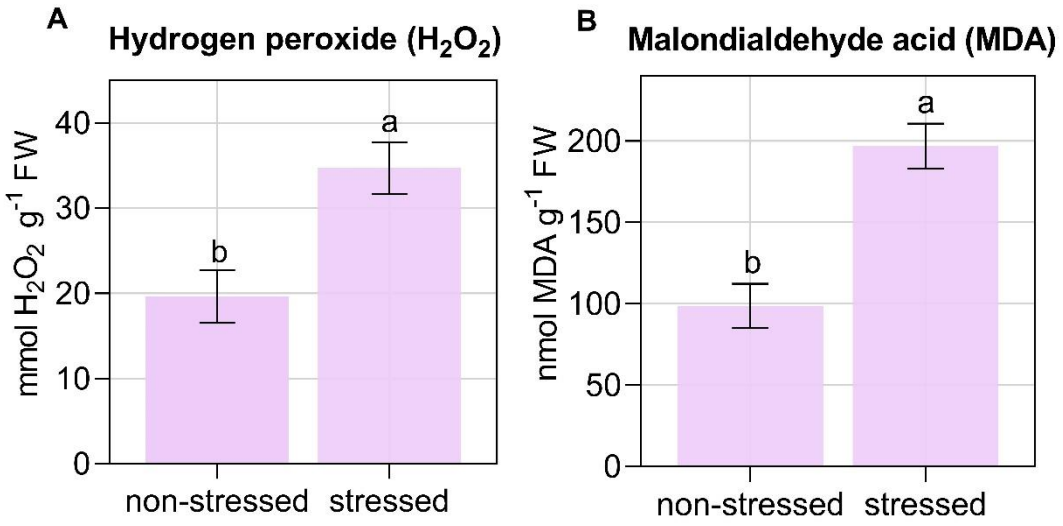
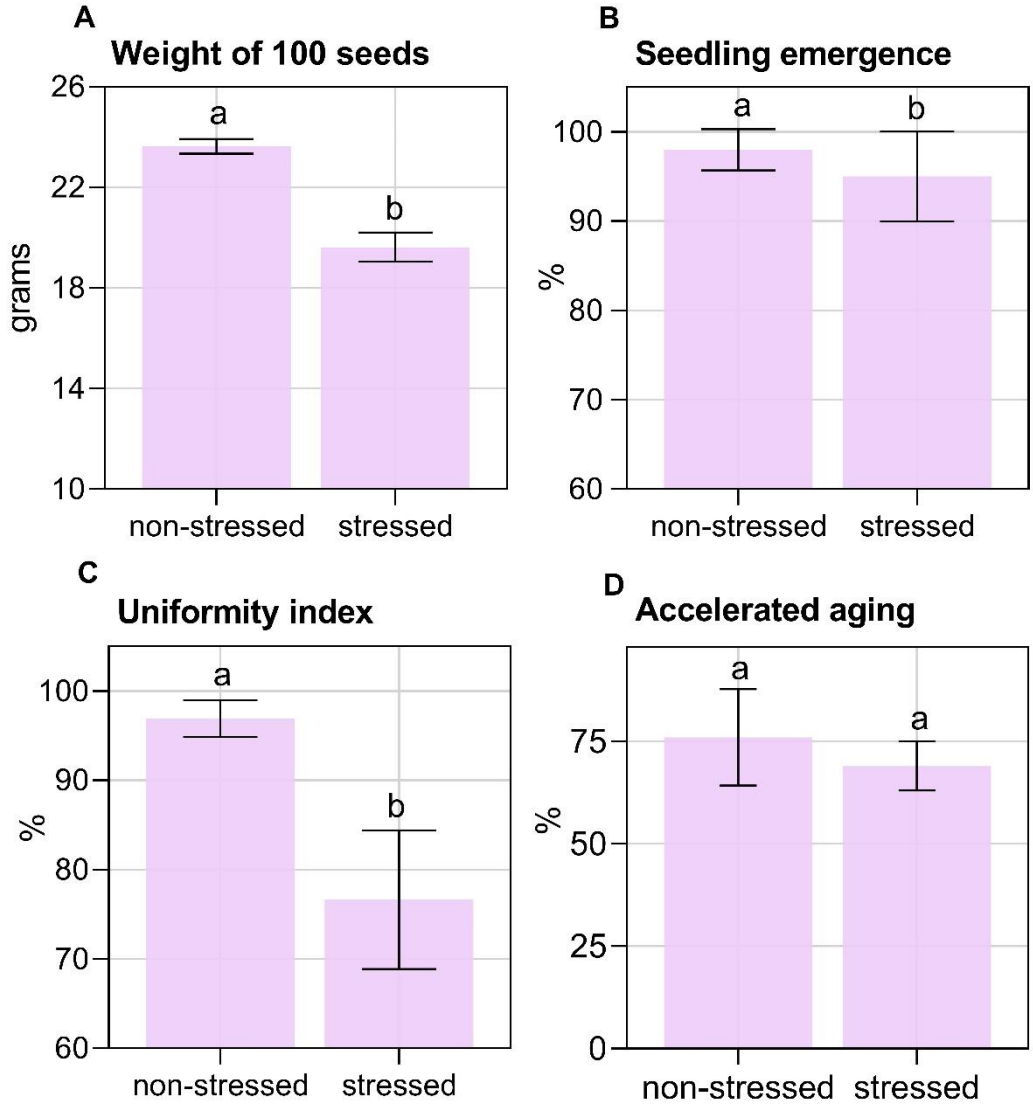


Figure 1 – Effect of environmental stress during maturation of soybean on seed weight and vigor. Seed weight (A), seedling emergence (B), emergence uniformity index (C), and accelerated aging (D) of non-stressed and stressed soybean seeds. Different letters indicate a significant difference ($P \leq 0.05$) by t-test. Error bars show the standard deviation from four samples

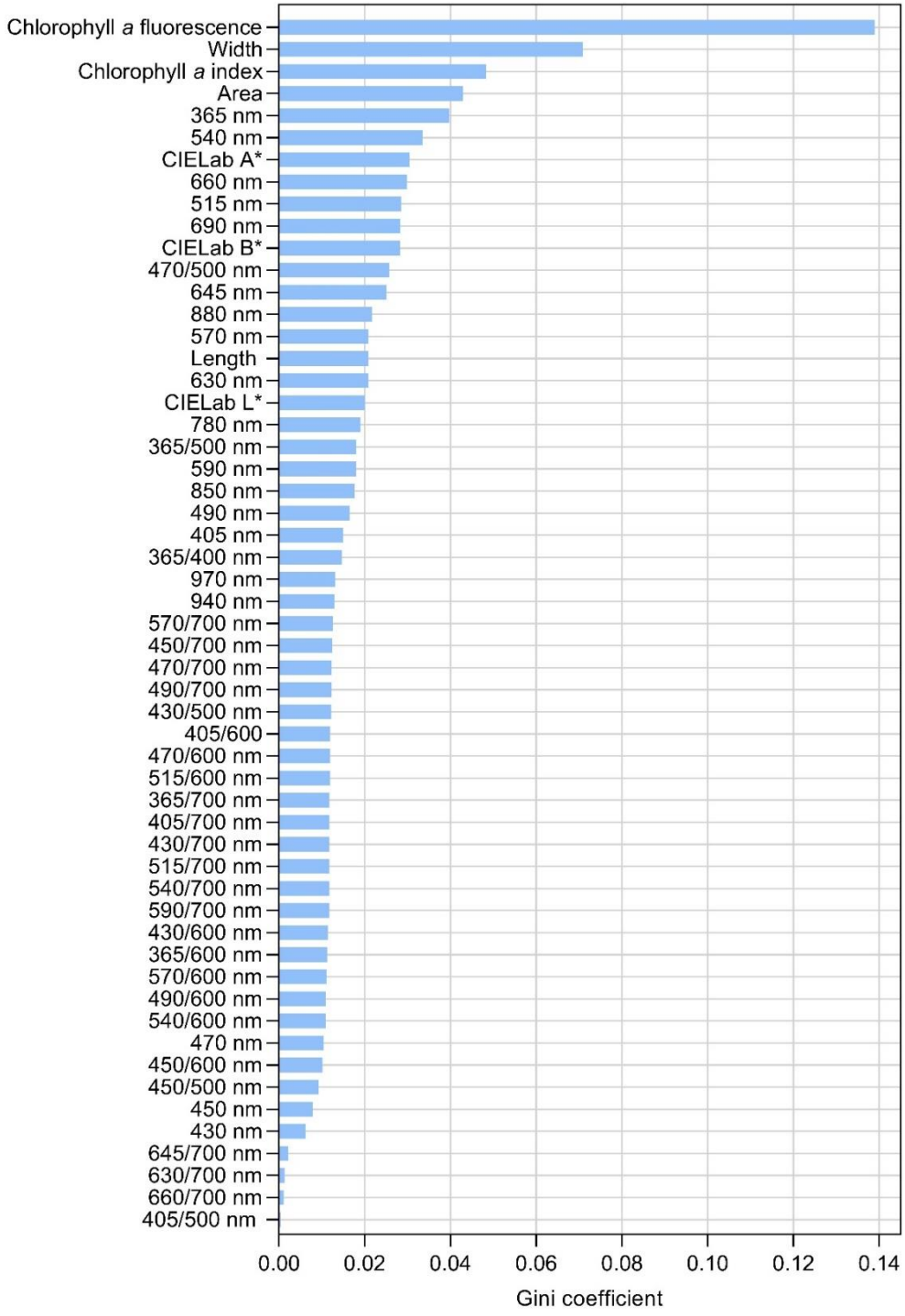


Multispectral parameters can segment stressed seeds

We applied the Gini coefficient to the 55 multispectral variables (reflectance, autofluorescence, physical properties, Chl *a* fluorescence, and Chl *a* index) obtained from the stressed and non-stressed seeds to rank the 20 most important. We identified that they are, in order: Chl *a* fluorescence, seed width, Chl *a* index, seed area, 365 nm, 540 nm, CIELab a*, 660 nm, 515 nm, 690 nm, CIELab b*, 470/500 nm, 645 nm, 880

nm, 570 nm, seed length, 630 nm, CIELab L*, 780 nm, 365/500 nm. The entire Gini-coefficient report is shown in FIGURE 2.

Figure 2 – Gini coefficient for the contribution of the multispectral variables in the segmentation of stressed and non-stressed soybean seeds (*n*=500 seeds). Chlorophyll *a* fluorescence and the excitation/emission combination of 405/500 nm were the variables that respectively contributed the most and the least to seed group segmentation



Using the 20 most important variables based on the Gini coefficient, we applied a PCA to explore seed segmentation (FIGURE 3). Component 1 (PC1) and component 2 (PC2) explained 70.9% and 14.7% of the significant variation, respectively. PCA efficiently separated the two groups of seeds, in which stressed seeds were associated with Chl *a* index, Chl *a* fluorescence, CIELab a*, and 365 nm reflectance. Meanwhile, the other 16 most important variables were associated with non-stressed seeds.

Figure 3 – Biplots of principal component analysis (PCA) for chlorophyll *a* index, chlorophyll *a* fluorescence, physical properties, reflectance, and autofluorescence spectral markers of stressed and non-stressed soybean seeds. The vectors indicate the correlation between the seed groups (stressed and non-stressed) and the dimensions PC1 and PC2

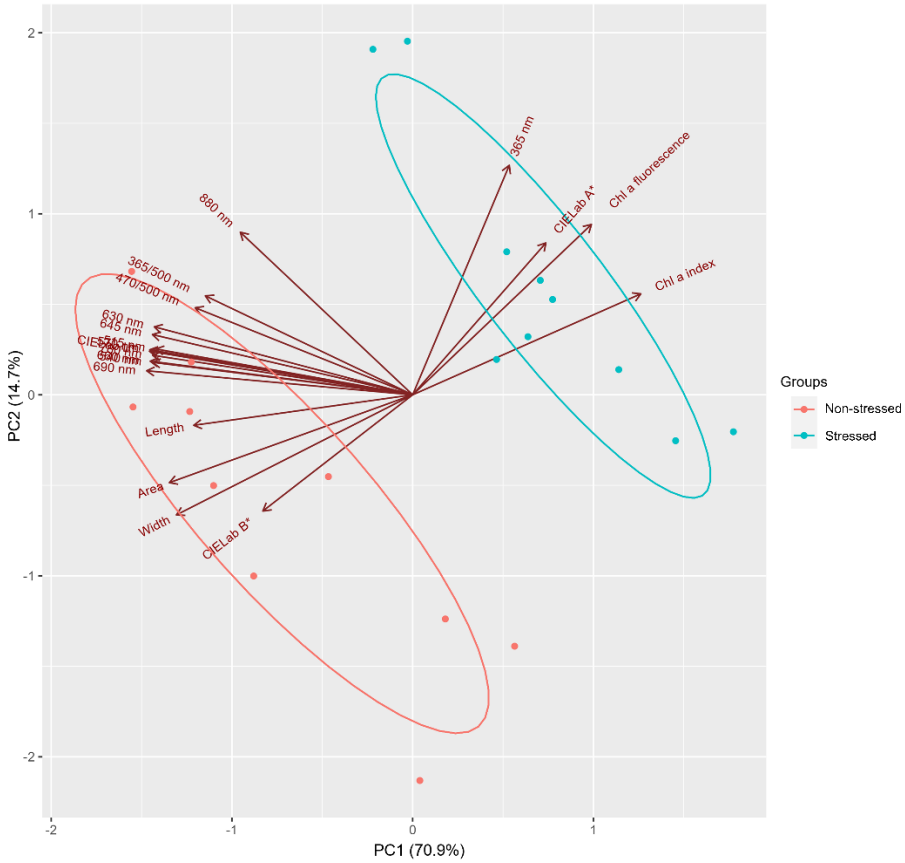


Figure 4 shows the means of the five most important markers for segmenting stressed and non-stressed seeds based on the Gini coefficient. Stressed seeds had higher values of Chl *a* fluorescence (FIGURE 4A), Chl *a* index (FIGURE 4B), and reflectance in the 365 nm wavelength (FIGURE 4E) ($P \leq 0.05$), as also shown in the images (Figure 5). However, they had smaller seed width (FIGURE 4C) and seed area (FIGURE 4D) compared to non-stressed seeds.

Figure 4 – The five most important multispectral variables to separate non-stressed seeds from stressed soybean seeds based on the Gini coefficient. Different letters indicate a significant difference ($P \leq 0.05$) by t-test. Error bars show the standard deviation from 250 samples

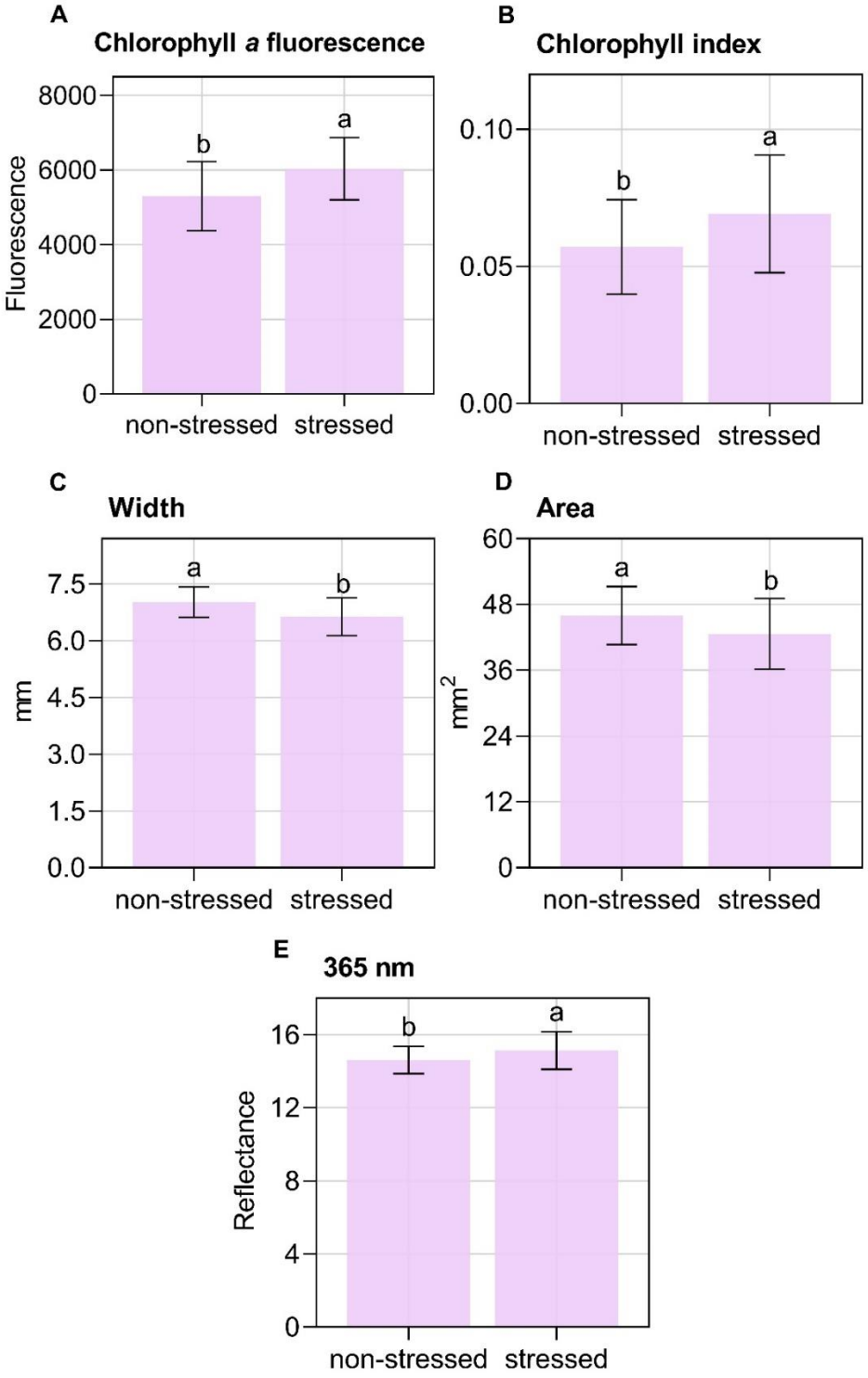
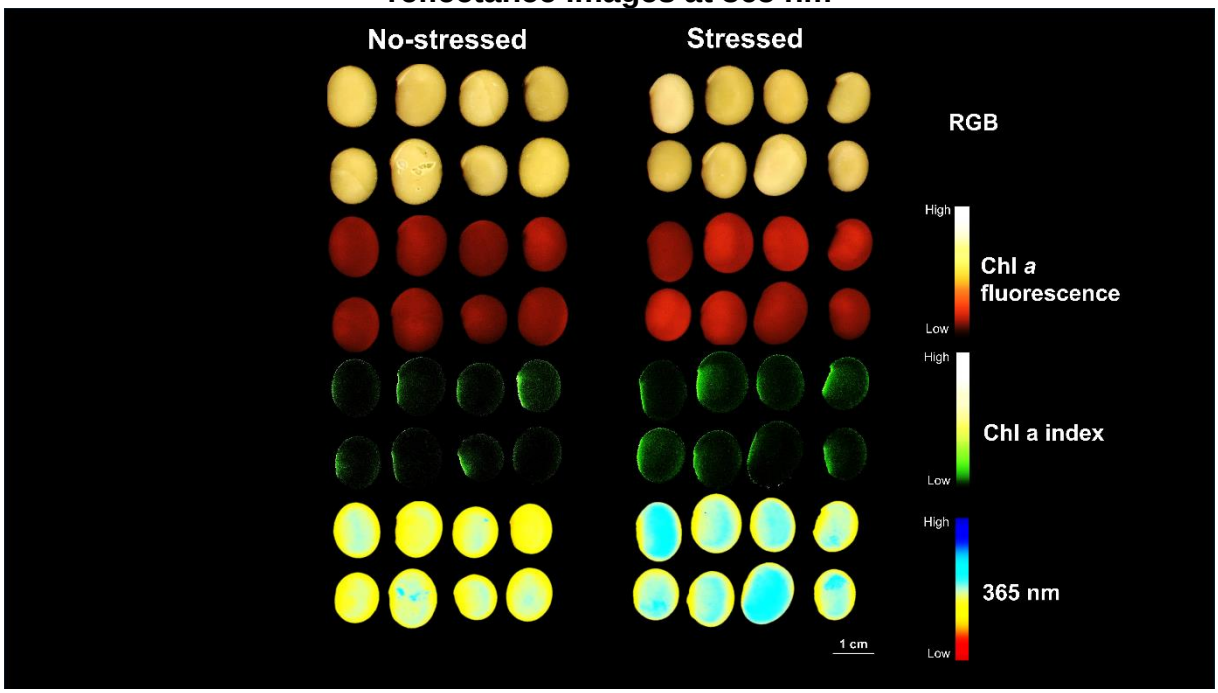


Figure 5 – Score pixels of non-stressed and stressed soybean seed images. RGB images (red, green and blue channels) and corresponding chlorophyll *a* fluorescence (Chl *a* fluorescence), chlorophyll *a* index (Chl *a* index), and reflectance images at 365 nm



Seed segmentation based on different machine learning models

Using NN as an algorithm, we obtained accuracy, F1, precision, recall, and specificity ≥ 0.91, while RF and SVM presented ≤ 0.89 in the training of the algorithms. In the test, the NN algorithm showed metrics ≥ 0.89; on the other hand, RF and SVM showed metrics ≤ 0.82 (FIGURE 6). In addition, the confusion matrix showed that using NN algorithm in the multispectral data, the probability of success in the prediction is greater than 90%. The probability of success for stressed seeds identification was 94%, demonstrating a high capacity for prediction through the parameters used to identify this class.

Figure 6 – Machine learning models and confusion matrices to discriminate stressed and non-stressed soybean seeds using multispectral data based on Random Forest (RF) (A), Support Vector Machine (SVM) (B), and Neural Network (NN) (C) algorithms. Models performance based on accuracy, F1, precision, recall, and specificity. Confusion matrices were created using the test dataset (*n*= 150 seeds)

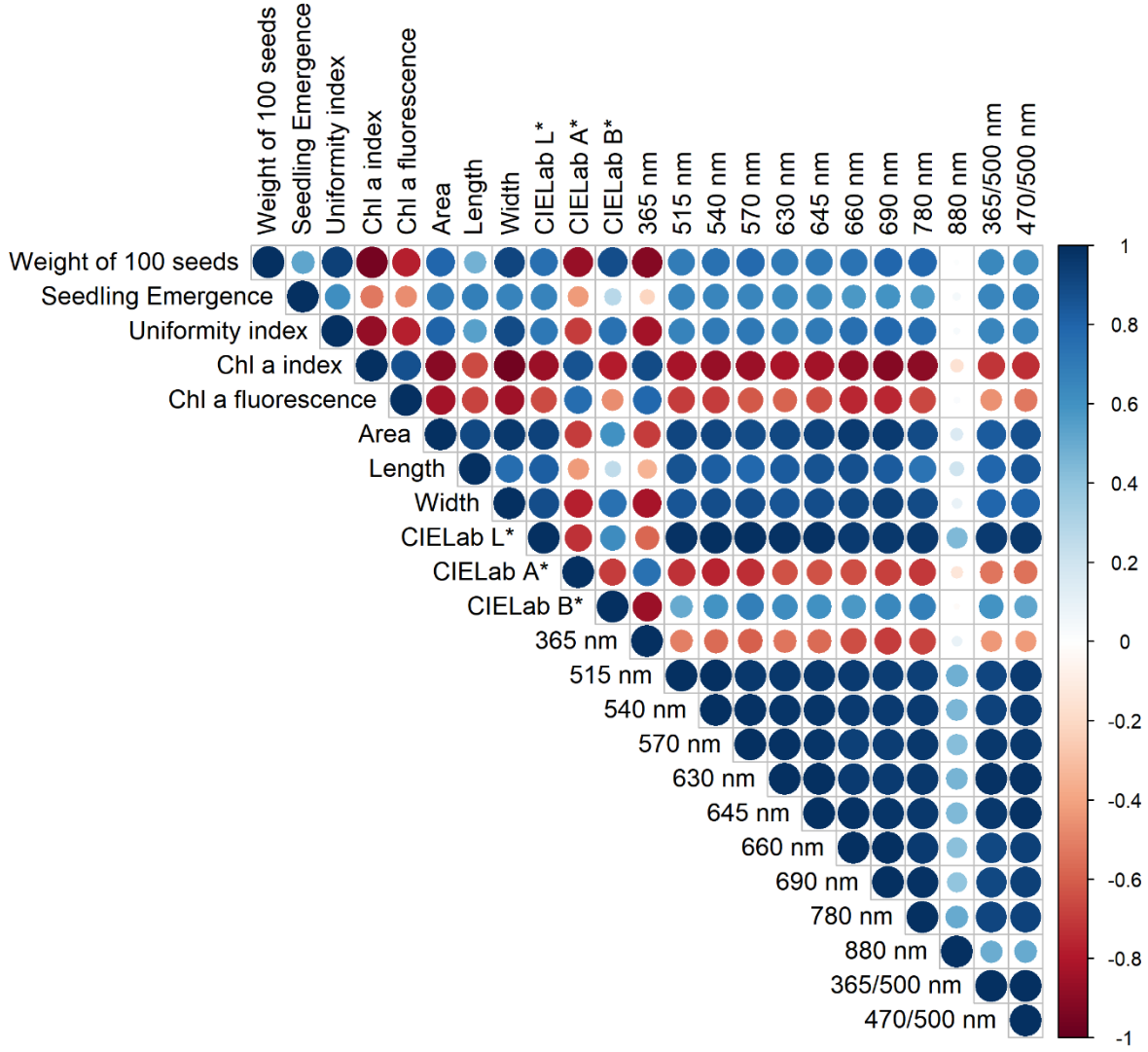
Correlation of the seed physiological traits with Chl a parameters, physical markers, reflectance, and autofluorescence

The 20 most important multispectral markers were inserted in a correlation matrix with the physiological data to check the relationship between the variables (FIGURE 7). The weight of 100 seeds was strongly correlated with Chl *a* index (-0.95). There was also a strong correlation coefficient between the weight of 100 seeds and physical properties, especially seed width (0.92). Furthermore, the weight of 100 seeds and

reflectance obtained the strongest correlation coefficient with 365 nm (-0.92) wavelength.

There was a strong correlation coefficient between the seedling emergence and seed area (0.70). The uniformity index was negatively correlated with Chl *a* parameters: Chl *a* index (-0.88) and Chl *a* fluorescence (-0.77). The uniformity index was also strongly correlated with physical markers, in which the highest correlation coefficient was obtained with seed width (0.89). Finally, the uniformity index was mainly correlated with the wavelength of 365 nm (-0.83).

Figure 7 – Pearson's correlation coefficients between seed physiological quality, physical properties, reflectance, and autofluorescence-spectral markers (nm) in stressed and non-stressed soybean seeds



1.4 DISCUSSION

Seed companies require seed quality management tools. Multispectral imaging associated with ML algorithms is a prospective innovation in this field. We asked

ourselves if this imaging technology could identify seeds that went through a stressful production environment and then if ML algorithms could automatically segregate the stressed seeds. For this, stressed soybean seeds (seeds produced under stressful conditions of water deficit and heat) and non-stressed soybean seeds – our control group - were analyzed for physiological quality, and also using multispectral imaging tools. Here, we demonstrate for the first time that by using emerging technologies, it is possible to classify the soybean seeds' origin environment and, thus, infer the consequences on seed physiological quality.

In croplands worldwide, sowing seeds with high physiological quality is essential, because when establishing in the field, it provides fast and uniform germination and emergence. Our data showed a reduction in the seed vigor for stressed seeds, as is also reported in the literature (Bagateli et al., 2019; Finch-Savage; Bassel, 2016), which causes a disturbance in the seed-seedling transition, reducing emergence uniformity (FIGURE 1C). In the fields, uneven seedling emergence produces dominated plants in the plant community, which although being able to recover their height in the late growth stages due to phenotypic plasticity, are less productive because of the reduction in the number of nodes and pods per plant (EBONE et al., 2020). Reduced seedling emergence rate (FIGURE 1B) causes missing plants in the plant stand, directly reducing yield (Krzyzanowski et al., 2018).

To test and determine seed vigor, seed companies often apply AA tests in their lab work routine (França-Neto; Krzyzanowski, 2022). In our study, the AA test (72 h 41 ºC) could not detect seed vigor difference between the two groups (stressed and non-stressed) (FIGURE 1D). Thus, seeds with reduced vigor due to stressful conditions during the maturation phase might not be detected by traditional lab tests, such as AA. This problem is typical because the reduction in vigor is not as apparent as other factors that affect quality with a latent or immediate effect (such as mechanical damage).

We achieved an innovative methodology to diagnose reduction in seed vigor caused by stressful environments using Chl *a* fluorescence, Chl *a* index, seed width, seed area, reflectance variables, etc. since the changes in these properties promoted by stress during seed maturation were detected using multispectral imaging (FIGURE 2, 4, and 5). Stressed seeds presented high Chl *a* levels, as shown by a fluorescence and Chl *a* index (FIGURE 4A, 4B and 5). Stressful conditions during maturation promote Chl retention in soybean seeds (Pádua et al., 2007; Teixeira et al., 2016) that

impact the seed-seedling transition (Ajala-Luccas et al., 2023). Thus, Chl levels in stressed seeds were higher, which became a strong marker detected by multispectral images. This technique has high sensitivity and enables distinguishing groups based on the level of chlorophyll that does not degrade due to stressful conditions.

In our study, the greenish color (high chlorophyll content) of the seeds was not visible to the human eye as shown in the RGB images (Figure 5). High chlorophyll levels impact legume seeds' lifespan (Nakajima et al., 2012; Zinsmeister et al., 2023), but the reduction in the physiological quality of soybean seeds is generally detected only at the time of the effective use of the seeds, i.e., after six months of storage waiting to be commercialized. Thus, our method based on multispectral imaging, once incorporated data related to chlorophyll fluorescence, can be used as a tool for predicting severe problems that seeds may face after sewing (i.e., loss of viability and consequent commercial standard loss).

In seeds, several fluorescent chemical components can be detected by multispectral images at short wavelengths (Donaldson, 2020; García-Plazaola et al., 2015). Fluorescent compounds, such as chlorophyll, can alter the color properties of seeds, especially those related to brightness (Fonseca De Oliveira et al., 2022; Oliveira et al., 2021; Sudki et al., 2023. parallel, these changes potentially influence the reflectance properties of seeds, as shown in our study at 365 nm (FIGURE 4 and 5). This also occurred in peanut (Fonseca De Oliveira et al., 2022), tomato and carrot (Galletti et al., 2020). Therefore, the environmental stress conditions imposed on soybean plants generated seeds with a detectable post-harvest spectral pattern. This spectral pattern opens up a technological opportunity for assisted seed lot selection based on reflectance, physical properties, and color descriptors. This also have the potencial for a future seed testing standardization.

Regarding the shape markers, width and area were the most important variables for segmenting stressed seeds (FIGURE 2). Probably, stressed seeds had smaller width, area (FIGURE 4C, D), and weight (FIGURE 1A) because the plant needed to use its photoassimilates to deal with the oxidative stress (FIGURE SUPPLEMENTARY 1), and the accumulation of reserves in the seed was reduced (Sionit and Kramer, 1977; Staniak et al., 2023). Thus, the shape of the seeds was modified, which is a common effect in seeds produced over stressful conditions. Integrating shape properties with the chlorophyll fluorescence and reflectance reported here allows us to

generate a robust model to classify seeds that go through stress during their development (FIGURE 6).

Additionally, we demonstrated that our data could be interpreted in an autonomous way, especially using NN algorithms (FIGURE 6). This opens the possibility to turn our model into a “results manager,” i.e., identifying why there is a reduction in physiological quality in seeds when there is no noticeable damage, the so-called "silent enemies", as residual chlorophylls that impact the seed-seedling transition as reported by Ajala-luccas et al. (2023) and observed in our data (FIGURE 1, and 4).

In another way, our results allow more precise and sensitive diagnoses since the technology is highly associated with the uniformity index (FIGURE 7), an effective vigor test diagnosis. In this context, Barboza da Silva et al. (2021b) developed markers based on autofluorescence-spectral combined with ML to classify soybean seeds regarding their vigor. Therefore, our results additionally demonstrate the potential of these tools in this matter.

The technology proposed here can also be used in plant breeding studies. With global climate change, the tool can help to identify stress-tolerant materials by screening different genotypes submitted to stress. It can also be used in stress memory tests. Like phenotypic plasticity, stress memory is a plant survival strategy (Auler et al., 2021). The so-called plant stress memory is produced depending on the stress type and level since plants can carry out epigenetic modifications that will result in changes in their metabolism to ensure adaptation to the new environment (Lukić et al., 2020). This change can be passed on to the following generations and produce adaptations to new environmental scenarios (Liu et al., 2022). Stress memory assessments have been used in plant breeding programs to select stress-resistant materials (Villagómez-Aranda et al., 2022). The great advantage of adapting the tool in memory stress assays to seed bulking is that it allows quantification on a single seed.

Our study brings more knowledge about soybean seed quality evaluation through multispectral imaging associated with ML algorithms and presents a new approach to data management that enables detecting vigor differences caused by stress occurrence during seed maturation in a much more precise way than the traditional seed physiological quality tests.

1.5 CONCLUSION

Machine learning models based on multispectral imaging are highly effective for recognizing soybean seeds produced under stressful conditions, especially using features based on chlorophyll *a* fluorescence, seed width, chlorophyll *a* index, seed area, and reflectance at 365 nm. This approach can detect stressed soybean seeds with 90% accuracy using neural network algorithm.

REFERENCES

AJALA-LUCCAS, D. et al. The Seed–Seedling Transition in Commercial Soybean Cultivars with the Presence of Greenish Seeds in the Sample: A Perspective from Classical Genetic Parameters. **Agronomy**, Basel, Switzerland, Vol. 13, Page 1966, v. 13, n. 8, p. 1966, 26 jul. 2023.

ALEXIEVA, V. et al. The effect of drought and ultraviolet radiation on growth and stress markers in pea and wheat. **Plant, Cell & Environment**, Chichester, United Kingdom, 24, n. 12, p. 1337–1344, 1 dez. 2001.

BAGATELI, J. R. et al. Productive performance of soybean plants originated from seed lots with increasing vigor levels. **Journal of Seed Science**, Brasília, DF, v. 41, n. 2, p. 151–159, 2019.

BARBOZA DA SILVA, C. et al. A novel approach for Jatropha curcas seed health analysis based on multispectral and resonance imaging techniques. **Industrial Crops and Products**, Amsterdam, Netherlands, v. 161, p. 113186, 1 mar. 2021a.

BARBOZA DA SILVA, C. et al. Autofluorescence-spectral imaging as an innovative method for rapid, non-destructive and reliable assessing of soybean seed quality. **Scientific Reports**, London, United Kingdom, v. 11, n. 1, 1 dez. 2021b.

BATISTA, T. B. et al. A reliable method to recognize soybean seed maturation stages based on autofluorescence-spectral imaging combined with machine learning algorithms. **Frontiers in Plant Science**, Lausanne, Switzerland, v. 13, 2022.

BIANCHINI, V. DE J. M. et al. Multispectral and X-ray images for characterization of Jatropha curcas L. seed quality. **Plant Methods**, London, United Kingdom, v. 17, n. 1, 2021.

BRASIL, M. DA A. P. E A. **Regras para análise de sementes**. (Assessoria de Comunicação Social, Ed.) Brasília, DF: 2009.

CAKMAK, I.; CAKMAK, W. J. H.; HORST, W. J. Effect of aluminium on lipid peroxidation, superoxide dismutase, catalase, and peroxidase activities in root tips of soybean (Glycine max). **Physiologia Plantarum**, Hoboken, United States, 1991.

COHEN, I. et al. Meta-analysis of drought and heat stress combination impact on crop yield and yield components. **Physiologia Plantarum**, Hoboken, United States, v. 171, n. 1, p. 66–76, 2021.

CONTINI, E.; TALAMINI, D. J. D.; VIEIRA JUNIOR, P. A. **Cenário mundial de commodities: frango, soja e milho. - Portal Embrapa**, Brasília, DF. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/968636/cenario-mundial-de-commodities-frango-soja-e-milho>. Acesso em: 21 nov. 2023.

DONALDSON, L. Autofluorescence in plants. **Molecules,** Basel, Switzerland, v. 25, 2020.

EBONE, L. A. et al. Soybean Seed Vigor: Uniformity and Growth as Key Factors to Improve Yield. **Agronomy**, Basel, Switzerland, v. 10, n. 4, p. 545, 2020.

FINCH-SAVAGE, W. E.; BASSEL, G. W. Seed vigour and crop establishment: Extending performance beyond adaptation. **Journal of Experimental Botany**, Oxford, United Kingdom, v. 67, n. 3, p. 567–591, 2016.

FONSECA DE OLIVEIRA, G. R. et al. An approach using emerging optical technologies and artificial intelligence brings new markers to evaluate peanut seed quality. **Frontiers in Plant Science**, Lausanne, Switzerland, v. 13, 2022.

FRANÇA-NETO, J. B.; KRZYZANOWSKI, F. C. **Metodologia do teste de tetrazólio em sementes de soja**. Londrina, PR: Embrapa Soja, 2022.

FRANÇA-SILVA, F. et al. Advances in imaging technologies for soybean seed analysis. **Journal of Seed Science**, Brasília, DF, v. 45, p. 1–16, 2023.

GALLETTI, P. A. et al. Integrating Optical Imaging Tools for Rapid and Non-invasive Characterization of Seed Quality: Tomato (Solanum lycopersicum L.) and Carrot (Daucus carota L.) as Study Cases. **Frontiers in Plant Science**, Lausanne, Switzerland, v. 11, 21 dez. 2020.

GARCÍA-PLAZAOLA, J. I. et al. Autofluorescence: Biological functions and technical applications. **Plant Science**, Shannon, Ireland, v. 236, p. 136–145, 2015.

GITELSON, A. A.; GRITZ, Y.; MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. **Journal of plant physiology**, Muenchen, Germany, v. 160, n. 3, p. 271–282, 2003.

GOMES, W. P. C. et al. Application of multispectral imaging combined with machine learning models to discriminate special and traditional green coffee. **Computers and Electronics in Agriculture**, Amsterdam, Netherlands, v. 198, p. 107097, 2022.

KRZYZANOWSKI, F. C. et al. **Vigor de sementes: conceitos e testes**. Londrina, PR: Abrates, 2020.

KRZYZANOWSKI, F. C.; FRANÇA-NETO, F. B.; HENNING, A. A. **A alta qualidade da semente de soja: fator importante para a produção da cultura**. CT136. EMBRAPA, Londrina, PR, n. 136, p. 1–24, 2018.

LIMA, J. J. P. et al. Molecular characterization of the acquisition of longevity during seed maturation in soybean. **PLOS ONE**, San Francisco, United States, v. 12, n. 7, p. e0180282, 2017.

LIU, H.; ABLE, A. J.; ABLE, J. A. Priming crops for the future: rewiring stress memory. **Trends in Plant Science**, Cambridge, United States, v. 27, n. 7, p. 699–716, 1 jul. 2022.

MELGOSA, M. Testing CIELAB-Based Color-Difference Formulas. **Color Research and Application**, Hoboken, United States, v.25, n.1, 2000.

NAKAJIMA, S. et al. Chlorophyll b Reductase Plays an Essential Role in Maturation and Storability of Arabidopsis Seeds. **Plant Physiology**, Rockville, United States , v. 160, n. 1, p. 261–273, 3 set. 2012.

OLIVEIRA, N. M. et al. Hormetic effects of low-dose gamma rays in soybean seeds and seedlings: A detection technique using optical sensors. **Computers and Electronics in Agriculture**, Amsterdam, Netherlands, v. 187, p. 106251, 1 ago. 2021.

PÁDUA, G. P. et al. Incidence of green soybean seeds as a function of environmental stresses during seed maturation. **Revista Brasileira de Sementes**, Brasília, DF, v. 31, n. 3, p. 150–159, 2009.

PÁDUA, G. P. et al. Tolerance level of green seed in soybean seed lots after storage. **Revista Brasileira de Sementes**, Brasília, DF, v. 29, n. 3, p. 112–120, 2007.

RITCHIE, S. W.; HANWAY, J. J.; THOMPSON, H. E. **How a Soybean Plant Develops**. Des Moines, United States, p.1-24, 1982.

SIONIT, N.; KRAMER, P. J. Effect of Water Stress During Different Stages of Growth of Soybean. **Agronomy Journal**, Madison, United States, v. 26, p. 274-278, 1977.

STANIAK, M.; SZPUNAR-KROK, E.; KOCIRA, A. Responses of soybean to selected abiotic stresses—photoperiod, temperature, and water. **Agriculture**, Basel, Switzerland, v. 13, Page 146, v. 13, n. 1, p. 146, 2023.

SUDKI, J. M. et al. Fungal identification in peanuts seeds through multispectral images: Technological advances to enhance sanitary quality. **Frontiers in Plant Science**, Lausanne, Switzerland, v. 14, p. 1112916, 2023.

TEIXEIRA, R. N. et al. Gene expression profiling of the green seed problem in Soybean. **BMC Plant Biology**, London, United Kingdom, v. 16, n. 1, p. 1–15, 2016.

USDA. United States Department Of Agriculture. **Oilseeds: World Markets and Trade**. Washington, United States, 2023. Disponível em: <https://public.govdelivery.com/accounts/USDAFAS/subscriber/new>.

VILLAGÓMEZ-ARANDA, A. L. et al. Activating stress memory: eustressors as potential tools for plant breeding. **Plant Cell Reports**, Heidelberg, Germany, v. 41, p. 1481–1498, 2022.

WANG, X. et al. Multiple omics datasets reveal significant physical and physiological dormancy in alfalfa hard seeds identified by multispectral imaging analysis. **The Crop Journal**, Beijing, China, v. 11, n. 5, p. 1458–1468, 2023.

ZHANG, S. et al. Non-Destructive Testing of Alfalfa Seed Vigor Based on Multispectral Imaging Technology. **Sensors**, Basel, Switzerland, v. 22, Page 2760, v. 22, n. 7, p. 2760, 2022.

ZINSMEISTER, J. et al. ABSCISIC ACID INSENSITIVE 4 coordinates eoplast formation to ensure acquisition of seed longevity during maturation in Medicago

truncatula. **The Plant Journal**, Chichester, United Kingdom, v. 113, n. 5, p. 934–953, 1 mar. 2023.

ZORATO, M. DE F. et al. Presença de sementes esverdeadas em soja e seus efeitos sobre seu potencial fisiológico. **Revista Brasileira de Sementes**, Brasília, DF, v. 29, n. 1, p. 11–19, 2007.

CHAPTER 2

SILENCING AND OVEREXPRESSION OF GmABI5 GENE IN SOYBEAN GREENISH SEED STUDIES

ABSTRACT

Soybeans with chlorophyll retention have been frequent over the years, bringing issues for the seed and grain market. Greenish soybean seeds have lower physiological quality compared to seeds without chlorophyll retention. Besides the occurrence of adverse conditions during the seed maturation phase, genetic factors also play an important role as a determining factor for the occurrence of greenish seeds. Thus, this work aimed to perform silencing (RNAi) and overexpression of the candidate gene *Gm*ABI5 for future functional investigations. For the overexpression cassette, a fragment of 1359 base pairs was inserted between the regulatory elements of the glycinin protein. Two 422 bp fragments were inserted at the ends of an intron to form a hairpin and avoid gene expression for the silencing cassette. The transformation was carried out via a gene gun. Both cassettes had a hygromycin resistance gene used for the transformed embryos' selection phase. The presence of the expression cassette in transformed and hygromycin-resistant embryos was confirmed via PCR. The first embryos obtained were regenerated into seedlings and continue to be cultivated. Two hygromycin-resistant lines were confirmed for the presence of the cassette. The remaining embryos will be tested for the presence of the cassette; they are being cultivated, and the positive ones will be regenerated into seedlings and will be cultivated until homozygous lines are obtained.

2.1 INTRODUCTION

Soybean is an essential source of protein and oil and is widely used in animal feed, biodiesel production, and the pharmaceutical, cosmetic, and food industries (APROSOJA, 2023; Contini et al., 2018). Brazil has a prominent position in the world ranking of soy-producing countries. In 2021, the country was the production and export leader, with approximately 134 million tons and 86 million tons, respectively (FAO, 2023). Seeds are one of the most critical inputs in agriculture since they are the onset of a new field; they carry the genome for the new plant and allow the establishment of vigorous plants.

High quality seed production is critical for the seed technology market. In recent years, soybean greenish seeds have been a problem for Brazilian soybean growers due to the impairment of seed and oil quality. Soybean greenish seeds are more susceptible to degradation, leading to reduced physiological quality in terms of viability, vigor, and longevity (Pádua et al., 2007; Zinsmeister et al., 2016). These pigments also affect the grain industry since they lead to oxidation and changes in flavor, and demand additional bleaching steps (Diosady, 2005), causing financial losses.

Soybean greenish seeds are mature seeds with chlorophyll (Chl) retention. Chl is progressively degraded during maturation, and when maturity is reached, soybean seeds become naturally yellow (Lima et al., 2017; Ritchie, Hanway; Thompson, 1982). Soybean greenish seeds have their chlorophyll degradation process impaired in response to stresses such as high temperatures and drought during the maturation phase. Besides, the susceptibility to Chl retention may vary according to the genotype (Pádua et al., 2009; Teixeira et al., 2016).

There are few studies about the molecular mechanisms involved in chlorophyll retention/degradation in soybean seeds. Teixeira et al. (2016) reported altered expression of the STAY-GREEN 1 and STAY-GREEN 2 (D1, D2), PHEOPHORBIDASE 2 (PPH2), and NON-YELLOW COLORING 1 (NYC1_1) genes in a susceptible cultivar. *M. truncatula* abi5 mutant and Mtabi5 Mtabi4 double mutant seeds presented more chlorophyll and less seed quality than wild-type ones (Zinsmeister et al., 2016, 2023). However, the molecular mechanism associated with the retention and degradation of Chl is still an object of study.

ABI5 is a transcription factor involved in the control of late seed maturation in legumes. ABI5 plays a role in seed physiological traits such as dormancy, desiccation tolerance, and longevity through the regulation of raffinose family oligosaccharides (RFO) and late embryogenesis protein (LEA) content (Lynch et al., 2022). ABI5 is also involved in expressing photosynthetic genes (PhANGs – photosynthesis-associated nuclear genes) and dismantling the photosynthetic apparatus during seed maturation; therefore, it might also be involved with seed degreening (Zinsmeister et al., 2016). Thus, this study aimed to perform soybean genetic transformations with *Gm*ABI5 as a candidate gene for further functional investigations and possibly reproduce a genotype tolerant to environmental conditions that may cause chlorophyll retention.

2.2 MATERIALS AND METHODS

Candidate gene

The first steps of this research was perfomed at the School of Agriculture, Department of Crop Sciences, UNESP/Botucatu/São Paulo/Brazil. We performed gene expression preliminary studies to verify *Gm*ABI5 (Primers: 5' - GTGGAGAAGGTGGTGGAG 3' - CCAGTTCAACAGTGTATGCC) expression in samples of commercial soybean cultivars. The RNA was extracted using the NucleoSpin® RNA plant commercial Kit (Macherey-Nagel, Düren, Germany) following the manufacturer's instructions. The purity and quantity of total RNA were checked in a Nanodrop-2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). RNA integrity was checked in agarose gel at 1%. cDNA was synthesized by using a High-Capacity cDNA Reverse Transcription commercial kit (Applied Biosystems, Victoria, Australia) following the manufacturer's instructions. For a 20 µl reaction, we used 2.0 µl 10× RT Buffer, 0.8 µl 25× dNTP, 2.0 µl 10× RT Primer, 1 µl Reverse Transcriptase, 10 µl of the extracted RNA and 4.2 µl Nuclease-free water. The reaction was incubated in a thermocycler following these steps: 10 min at 25°C, 2 h at 37°C, followed by 5 min at 85°C and the synthesis ended at a constant temperature of 4°C.

The primers used were efficient above 1.8 and r^2 of approximately 1.0, according to the parameters of the LinRegPCR program (Ruijter et al., 2013). Gene expression analysis was performed in a thermocycler Eco Real-Time (Illumina, San Diego, USA) with SYBR Green qPCR ReadyMix (Sigma-Aldrich, St. Louis, USA), using two technical samples for each biological sample (three biological samples for each seed sample). For a reaction of 10 µl, 5 µl of SYBR Green, 1 µl of cDNA, and 0.25 µl of each primer was used, and the volume was adjusted with Nuclease-free water. The amplification consisted of 2 min at 50°C, 2 min at 95°C, then 45 cycles of 10 s at 95°C and 1 min at 60°C. At the end of the process, the melting curve was performed following these steps: 15 s at 95, 65, and 95°C, respectively. The normalized expression (NE) was assessed using the Proteasome gene as the reference gene, calculated according to NE = 2ΔCt, where ΔCt is the Cq of the reference gene – Cq target gene.

In the first gene expression experiment, we studied the expression of the ABI5 gene in two different maturation stages (R7.2 and R9) in yellow seeds of the cultivar MG/BR 46 Conquista which were produced according to Batista (2022). In parallel, the

Chl autofluorescence was determined by using multispectral imaging (Batista, 2022). In the second study, we performed gene expression of the ABI5 gene in two commercial soybean cultivars, NA5909IPRO and 97Y07. The production environment of the commercial cultivars is described by Luccas (2018). The seed samples comprised greenish and yellow seeds obtained through visual segregation. Seeds were considered greenish when they showed any sign of green pigmentation and were considered yellow when they did not show any sign of green pigmentation (Luccas, 2018).

Expression cassette

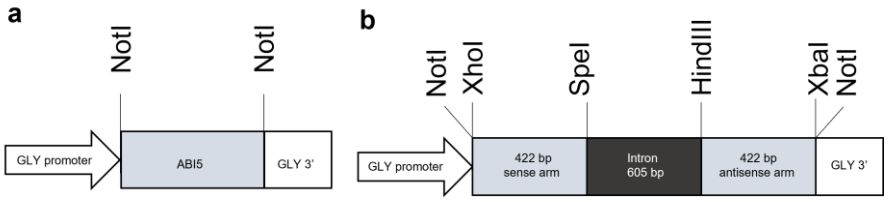
From this stage forward, the research was carried out at Dr. Schmidt's lab, College of Agriculture and Life Sciences, Bio5 Institute, University of Arizona/Tucson/AZ/EUA. We performed genetic transformation in soybean plants with expression alteration: increase (overexpression – ABI5OE) and silencing (RNAi – ABI5RNAi). The ABI5 cDNA nucleotide sequence was acquired in the GenBank database (Clark et al., 2016) under the accession number (XM_014763030.3).

For the overexpression cassette, a 1359 base pairs (bp) fragment containing the restriction sites for NotI was commercially synthesized in a plasmid with ampicillin (AMP) resistance and was cloned by *E. coli* transformation. The fragment was excised with NotI and placed in the open reading frame of an AMP-resistant vector containing the seed-specific storage protein glycinin (GLY) promoter. The plasmid also contained a hygromycin (HYG) resistance gene under the expression of the potato ubiquitin 3 regulatory elements (He et al., 2016) (pGLY/HYG). The resultant plasmid pGLYABI5OE/HYG (Figure 1a) was sequenced with a GLY promoter primer (5' - CCTCATTCACCTTCCTCTCTTC) to ensure the ABI5 open reading frame was placed correctly between the regulatory elements.

A 422 bp fragment containing NotI, XhoI, SpeI, Xba, and HindIII restriction sites was synthesized for the RNAi cassette in a plasmid with AMP resistance (pABI5). The cassette construction was according to Schmidt and Herman (2008). An intron was placed between two ABI5 arms forming a hairpin secondary structure in which the protein translation is impaired, and then the gene expression is silenced. The pABI5 and intron-containing plasmid were cloned by *E. coli* transformation. Two 422 bp fragments were restricted from the pABI5 vector with two separate restriction enzyme reactions: Xho/Spe and Xba/HindIII. The pABI5–Xho/Spe fragment was first inserted

into the new restriction sites at the 5’ end of the intron-containing vector. This plasmid was at this moment named pIntron-ABI5-HALF. The second arm pABI5-Xba/HindIII was inserted into pIntron-ABI5-HALF, generating the cassette containing a 422-bp ABI5 hairpin flanking the modified intron-containing plasmid; this vector was named pABI5-RNAi. Afterward, the pABI5RNAi open reading frame was NotI digested to move the entire ABI5 hairpin between the glycinin regulatory elements of the pGLY/HYG vector as described above (Figure 1b). Correct orientation of the hairpin between the regulatory elements was confirmed by sequencing using a glycinin promoter primer as previously described. The resulting vector was named pGLYRNAiABI5/HYG.

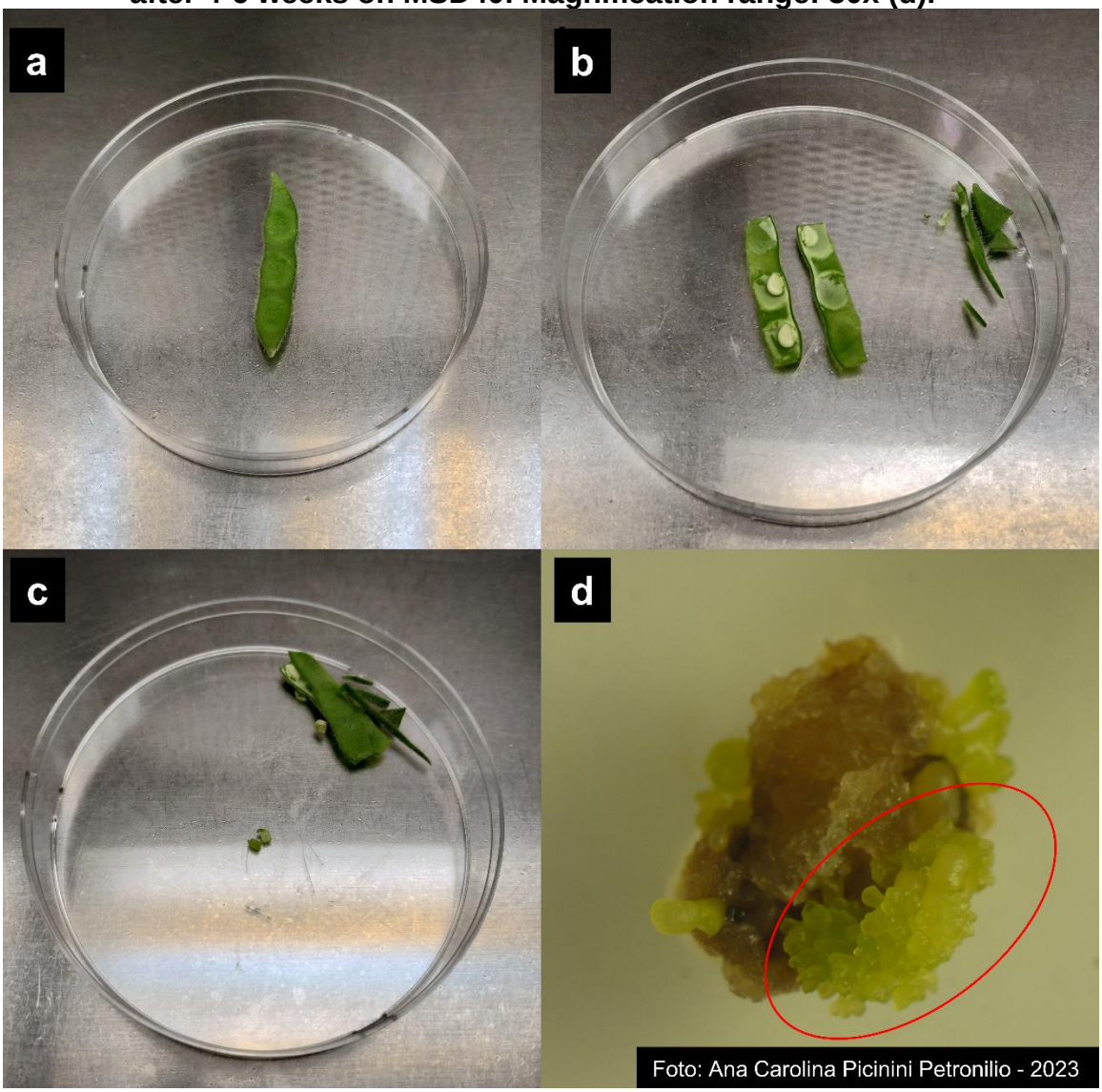
Figure 1 – Graphical representation of expression cassettes. The ABI5OE cassette contains the ABI5 cDNA sequence and the GLY promoter regulatory elements (a). ABI5RNAi cassette containing sense and antisense 422bp ABI5 fragments inserted between an intron and GLY promoter regulatory elements (b).



Tissue Culture and genetic transformation

Soybean somatic embryos (cultivar Jack) were induced from 5 mm long immature cotyledons (Figure 2a, b). The cotyledons were excised at their narrowest portion, and the portion with the embryonic axis was removed from the seed coat. The cotyledons were placed on MSD40 media plates with the flat side up (Figure 2c). The MSD40 media, according to Finer and Nagasawa (1988), had the pH modified according to Komatsuda and Ohyama (1988), and sucrose according to Lazzeri, Hildebrand, and Collins (1987). The plates were kept in the tissue culture room under 22 °C, low light intensity (10 µmol photons m^{-2} s^{-1}), eight hours photoperiod.

Figure 2 – Soybean somatic embryo induction. Soybean pod containing 5 mm long immature embryos (a, b). Excised embryos with the flat side up before being placed on MSD40 plates (c). Raspberry-appearance embryos after 4-6 weeks on MSD40. Magnification range: 30x (d).



After four-six weeks, raspberry-appearance embryos (Figure 2d) were transferred to FN-Lite proliferation media (Samoylov; Tucker; Parrott, 1998) and kept on a shaker at 125 rpm. The embryos were subcultured weekly for four weeks to new media, when they got a cluster appearance and became appropriate for gene gun shooting (Figure 3a, b). Four days before shooting, compact masses of globular-stage embryos were transferred to MSD20 (Wright et al., 1991) plates in a 30mm diameter area in the center of the plate (FIGURE 4a,b).

Figure 3 – Embryo clusters after four weeks of growth in FNL media

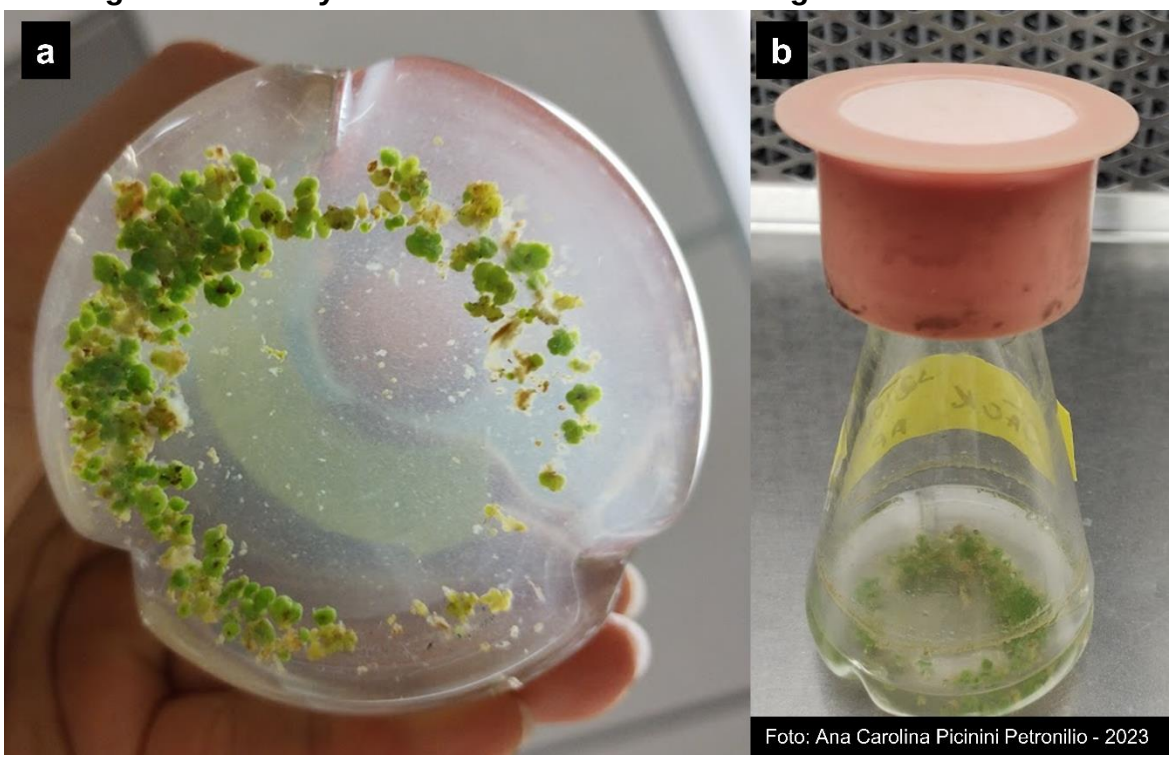
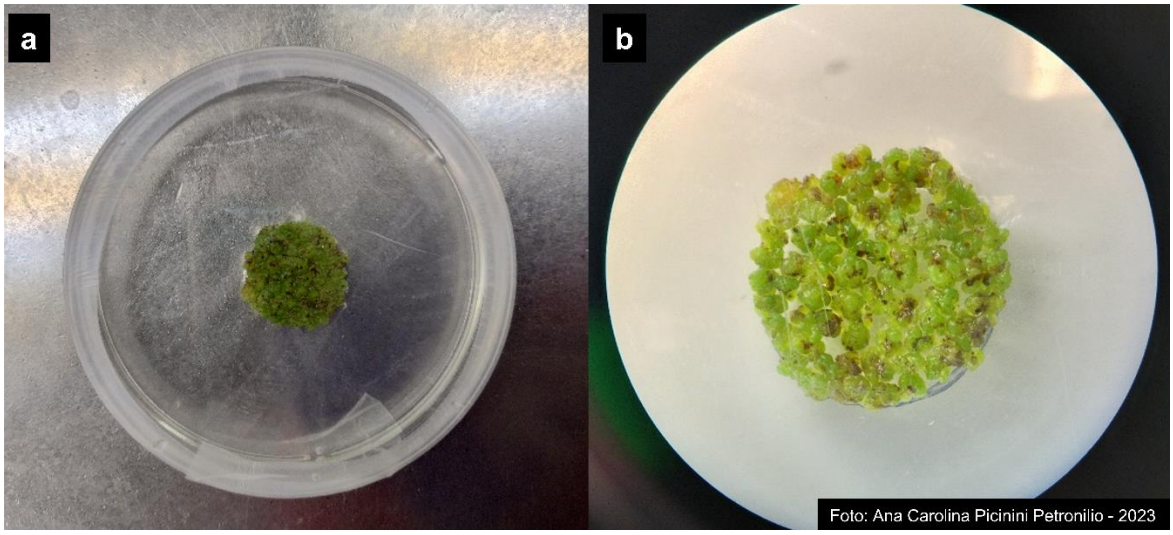


Figure 4 – Embryos on MSD20 ready for shooting (a, b)

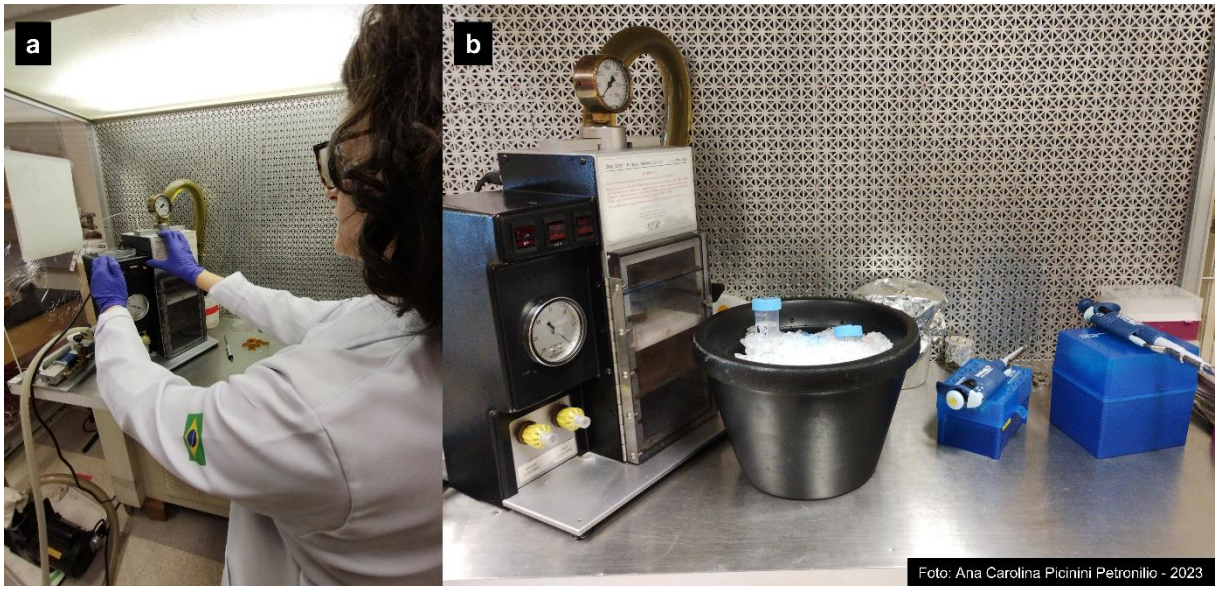


The somatic embryos were transformed using a Dupont PDS-1000 Biolistic Particle Delivery System with gold particles as a DNA microcarrier, accordingly to Dinkins et al. (2003), with some modifications. For 12 shots (0.4 mg of gold and 6.25 µg of DNA per shot), 10 mg of gold were rinsed in 1mL of 100%EtOH and centrifuged for 5 min at 7000rpm. The gold was resuspended in 169 µL of 100% EtOH, and three aliquots of 35µL were withdrawn. The aliquots were briefly centrifuged for 10 seconds

at 10000 rpm, the supernatant was removed, and 1 mL of sterile H_2O was added. Another centrifugation was performed for 5 minutes at 2000 rpm, and the supernatant was removed. 25µL of DNA was added to the metal pellet for each aliquot, and the microcarriers were resuspended. Subsequently, 220 µL of sterile H_2O were added, followed by 250 µL of 2.5M $CaCl_2$ and 50µL of 0.1M spermidine stock. The aliquots were placed on ice for 2-3 min, vortexed for 10 min, and centrifuged for 5 min at 500-1000 rpm. The supernatant was removed, the DNA/microcarrier was resuspended in 600 µL of 100% EtOH, centrifuged for 5 min at 500-1000 rpm, the EtOH was removed, and the aliquot was resuspended in 36µL of 100% EtOH by flicking. The aliquots were kept on ice for 1h.

The gene gun was prepared before shooting, and all parts were cleaned in 100% EtOH. Twenty minutes after the shooting, the embryos' plates were opened in the laminar flow to let any drop of water dry. The aliquots were sonicated for 10-15 sec, and 10 µL of the resuspended DNA/microcarriers were withdrawn and dispensed on the surface of a microcarrier disk. The shooting was performed using an 1100 psi rupture disk at an 8 cm distance from DNA to target tissue. After shooting, embryos were placed in FN-Lite media for one week. In this project, we performed a total of four gene gun protocols. The first two were the ABI5OE cassette, and the others were the ABI5RNAi cassette. There were 6 MSD20 plates for each protocol, and each plate received two shots. Figure 5 shows gene gun during shooting experiments.

Figure 5. Dupont PDS-1000 Biolistic Particle Delivery System during soybean somatic embrios transformation.



Selection of Cis/Intragenic plants

The embryos were submitted to hygromycin B (HYG) selection one week after transformation. Hygromycin B (PhytoTech Labs, Lenexa, Kansas – USA, 50mg/mL) was added to FN-Lite media in a final 20mg HYG/L concentration. The FN-Lite media + HYG solution was replaced weekly for 6-8 weeks. The embryos were screened using a stereomicroscope to select the ones that were green/alive. Each green embryo was considered as a different line (Figure 6) and was grown in FN-Lite media until it became a large enough tissue for subsequent analysis. The media was replaced every two weeks.

Figure 6 – HYG-resistant embryos were screened after six weeks of selection. Picture of the dissection microscope view. Magnification range: 30x.

Resistant embryos obtained from ABI5OE transformation, after nine weeks of growth, were screened by PCR to confirm the presence of the inserted expression cassette with HYG gene primers (5’ - CTCACTATTCCTTTGCCCTC 3’-CTGACCTATTGCATCTCCCG). Genomic DNA was extracted by the cetyl trimethyl ammonium bromide (CTAB) method. 150 ng genomic DNA were used in 25 µL total reaction containing 200 nM primers and 3 U Taq polymerase (NEB), and the following cycling parameters were used: one initial cycle at 95°C 4 min, then 30 cycles at 95°C 30 s, 55°C 30 s and 72°C 30s, and then a final extension cycle at 72°C for 7 min. GAPDH (glyceraldehyde-3-phosphate dehydrogenase) primers (5’ – TAACGAGCACGAGTACAAGCC 3’ - AATTCCCTTCAACTTGCCCTC) were used as an absolute control for PCR reactions. Soybean cultivar Jack was used as a negative control, and we used a HYG positive sample as a positive control.

Tissue culture embryos were histodifferentiatiated through SHAM media during approximately two weeks, according to Schmidt et al. (2005). When they reached the cotyledonary stage, they were submitted to desiccation (Figure 7), in which a determined number of embryos were placed in a sealed petri dish for about seven days. Afterward, embryos were germinated in MSO media (Figure 8a) (Murashige; Skoog, 1962; Gamborg et al., 1968), and when there was a visible shoot and roots

(Figure 8b), they were regenerated into plants (Figure 9). Tissue culture took, in total, approximately 9 months, from induction until conversion into plants.

Figure 7 – Soybean somatic cotyledonary stage embryos during desiccation.

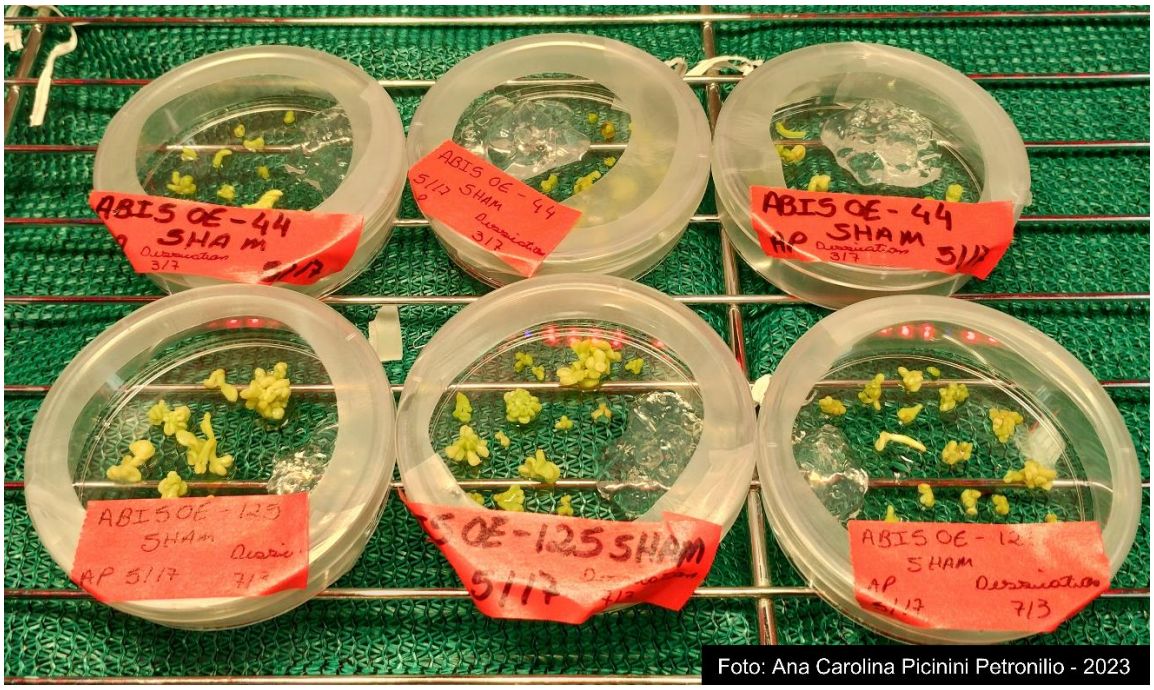


Figure 8 – Soybean somatic cotyledonary stage embryos during germination. Beginning of germination process (a). Germinated embryos with visible shoot and roots (b).

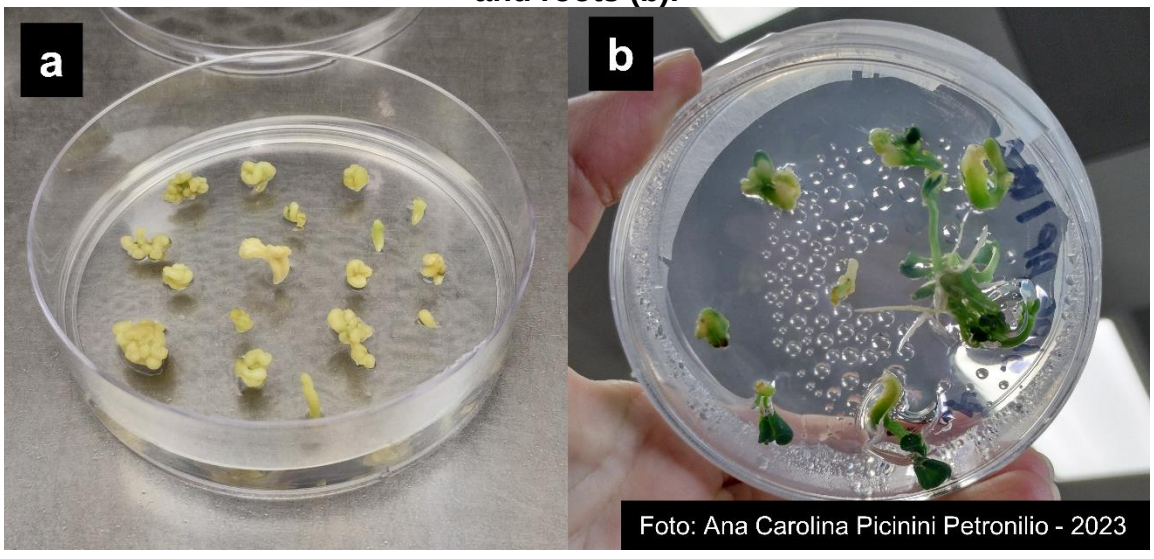


Figure 9 – Regenerated seedling from somatic embryogenesis.

2.3 RESULTS AND DISCUSSION

Figure 10a shows an increase in the expression of ABI5 transcripts from immature (R7.2) to mature (R9) soybean seeds. In parallel, chlorophylls *a* and *b* autofluorescence decreased in mature seeds (Figure 10b, c). In the first experiment, the BR/MG 46-Conquista was used because it is susceptible to abiotic stresses and chlorophyll retention (Pádua et al., 2009), therefore we supposed it would provide on-target results for ABI5 gene expression.

In addition, we used two soybean cultivars produced in commercial field environments according to Luccas (2018). There was a decreased expression of ABI5 transcripts in greenish seeds compared to yellow seeds in both commercial soybean cultivars (Figure 11). Thus, when ABI5 expression is increased, the chlorophyll content in the seeds tended to decrease.

Figure 10 – ABI5 gene expression and Chl fluorescence from MG/BR 46 Conquista cultivar. (A) Normalized gene expression from two different maturation stages. Each bar represents the mean of three biological samples ± standard deviation. 3.83-fold represents the fold change between R7.2 and R9 sample. (B) and (C) represent, respectively, Chl *a* and *b* fluorescence. Each bar represents the mean of four samples. Different letters indicate the difference between samples by t-test 95% of the confidence interval. The non-overlapping of the confidence interval constitutes the difference between the groups evaluated

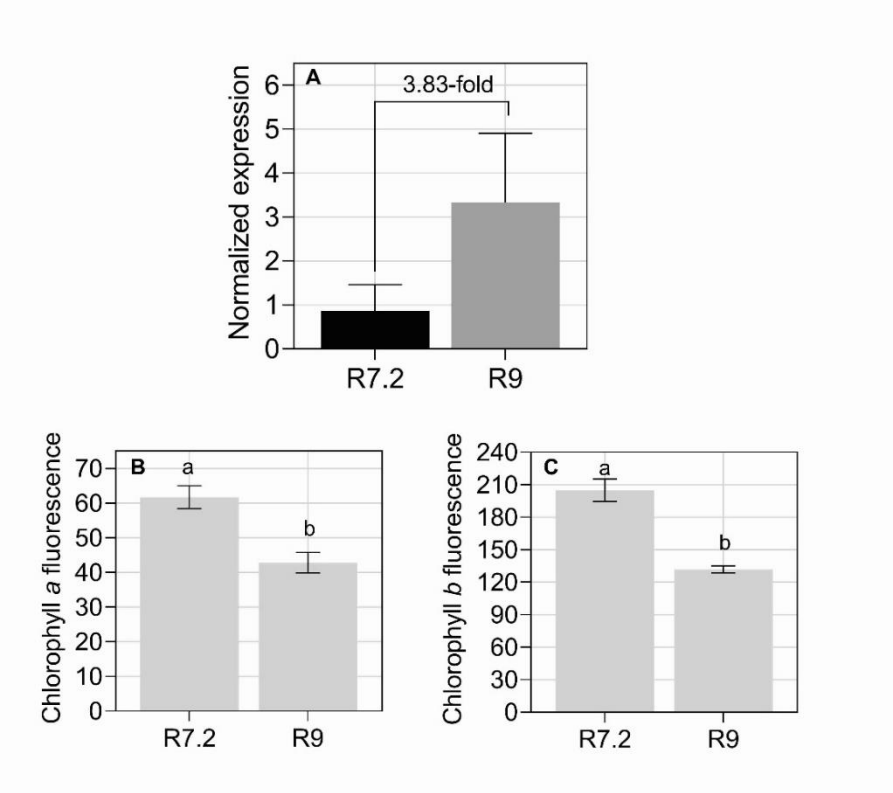
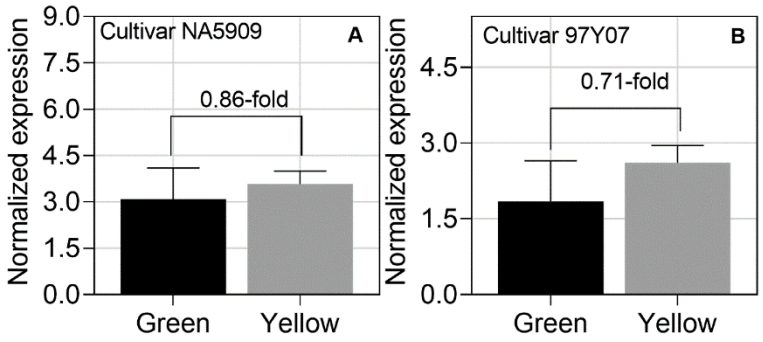


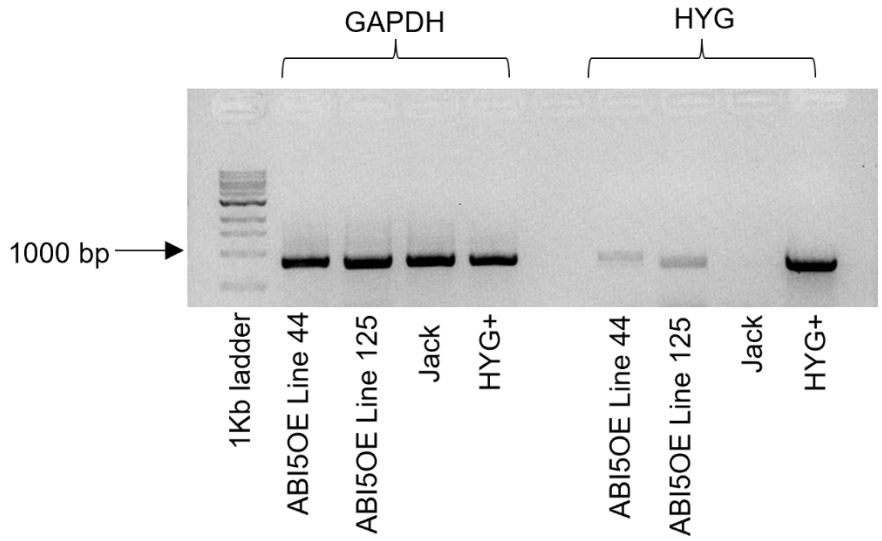
Figure 11– ABI5 gene expression from two different commercial soybean cultivars. (A) Normalized gene expression from greenish and yellow seeds from NA5909 cultivar. Each bar represents the mean of three biological samples ± standard deviation. 0.86-fold represents the fold change between greenish and yellow seeds. (B) Normalized gene expression from greenish and yellow seeds from 97Y09 cultivar. Each bar represents the mean of three biological samples ± standard deviation. 0.71-fold represents the fold change between greenish and yellow seeds. Data are not significantly different between samples by t-test 95% of the confidence interval



Therefore, this data indicated that ABI5 is a candidate gene for controlling chlorophyll degradation in soybean seeds and may be an essential gene in the greenish seed problem and environmental stress tolerance studies. In parallel, Zinsmeister et al (2016) suggested that chlorophyll degradation is under the control of ABI5, once there was remaining chlorophyll in abi5 mutant mature seeds of Pisum *sativum* and *M. truncatula*. Therefore, with these results and also based on literature, we decided to move forward with *Gm*ABI5 transformations.

After genetic transformation, the presence of the ABI5 overexpression cassette was confirmed in the first two HYG-resistant lines (Figure 12). The remaining ABI5OE and ABI5RNAi lines were not confirmed until this thesis publication because there was not enough tissue for DNA extraction yet. This step will be performed and subsequent analysis will be reported in a scientific paper.

Figure 12 – Electrophoretic pattern of the PCR product of HYG gene detection in *Gm*ABI5 overexpression studies. GAPDH - glyceraldehyde-3-phosphate dehydrogenase primers were used as absolute control for the PCR. HYG - Hygromycin gene. Jack – soybean cultivar used as HYG negative control. HYG+ - HYG positive control. bp – base pairs.



2.4 CONCLUSION

Two genetically modified *Gm*ABI5 overexpression lines have been confirmed as positive for the cassette presence, which indicates the gene gun transformation had succeeded. The remaining lines will be screened and grown until homozygous plants are obtained through self-pollination to characterize the *Gm*ABI5 overexpression effect in plants.

REFERENCES

APROSOJA. Associação Brasileira dos Produtores de Soja. **A soja: A origem do grão**. [S. L.], 2023. Available in: <https://aprosojabrasil.com.br/a-soja/>. Accessed on: 27 dez. 2023.

BATISTA, T. B. **MATURAÇÃO EM SEMENTES DE SOJA: ESTÁDIOS REPRODUTIVOS, DEGRADAÇÃO DA CLOROFILA E AQUISIÇÃO DA QUALIDADE FISIOLÓGICA**. 2022. Tese (Doutorado em Agronomia/Agricultura). Faculdade de Ciências Agronômicas, UNESP, Botucatu, SP, 2022.

CLARK, K. et al. GenBank. **Nucleic Acids Research**, Oxford, United Kingdom v. 44, n. 1, p. D67–D72, 2016.

CONTINI, E. et al. Série desafios do agronegócio brasileiro: complexo soja. Caracterização e Desafios Tecnológicos. **Embrapa**, Brasília, DF, p. 1-35, 2018.

DINKINS, R. D. et al. Recent Advances in Soybean Transformation. *In*: Jaiwal, P. K.; Singh, R. P. **Applied Genetics of Leguminosae Biotechnology**, Dordrecht, Netherlands: Springer, p.3-21, 2003.

DIOSADY, L. L. Chlorophyll removal from edible oils. **International Journal of Applied Science and Engineering**, Taichung, Taiwan, v. 3, n. 2, p. 81–88, 2005.

FAOSTAT. Food and Agriculture Organization of the United Nations. **Countries by commodities**. Rome, Italy, 2023. Available in : <https://www.fao.org/faostat/en/#rankings/countries_by_commodity_exports>. Accessed on: 27 dez. 2023.

FINER, J. J.; NAGASAWA, A. Development of an embryogenic suspension culture of soybean [Glycine max (L.) Merrill]. **Plant Cell, Tissue and Organ Culture**, Dordrecht, Netherlands, v. 15, p. 125–136, 1988.

GAMBORG, O. L.; MILLER, R. A.; OJIMA, K. Nutrient requirements of suspension cultures of soybean root cells. **Experimental Cell Research**, Waltham, United States, v. 50, n. 1, p. 151–158, 1968.

HE, Y. et al. Transgenic soybean production of bioactive human epidermal growth factor (EGF). **PLOS ONE**, San Francisco, United States, v. 11, n. 6, p. e0157034, 2016.

KOMATSUDA, T.; OHYAMA, K. Genotypes of high competence for somatic embryogenesis and plant regeneration in soybean Glycine max. **Theoretical and Applied Genetics**, Heidelberg, Germany, v. 75, p. 695–700, 1988.

LAZZERI, P. A.; HILDEBRAND, D. F.; COLLINS, G. B. Soybean somatic embryogenesis: Effects of nutritional, physical and chemical factors. **Plant Cell, Tissue and Organ Culture**, Dordrecht, Netherlands, v. 10, n. 3, p. 209–220, 1987.

LIMA, J. J. P. et al. Molecular characterization of the acquisition of longevity during seed maturation in soybean. **PLOS ONE**, San Francisco, United States, v. 12, n. 7, p. e0180282, 2017.

LUCCAS, D. A. **Caracterização fisiológica, bioquímica e molecular em sementes de soja (Glycine max (L.) Merr.) com retenção de clorofila**. 2018. Tese (Doutorado em Agronomia/Agricultura). Faculdade de Ciências Agronômicas, UNESP, Botucatu, SP, 2018.

LYNCH, T. et al. ABI5 binding protein2 inhibits ABA responses during germination without ABA-INSENSITIVE5 degradation. **Plant physiology**, Rockville, United States, v. 189, n. 2, p. 666–678, 1 jun. 2022.

MURASHIGE, T.; SKOOG, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. **Physiologia Plantarum**, Hoboken, United States, v. 15, n. 3, p. 473–497, 1 jul. 1962.

PÁDUA, G. P. DE et al. Tolerance level of greenish seed in soybean seed lots after storage. **Revista Brasileira de Sementes**, Brasília, DF , v. 29, n. 3, p. 112–120, 2007.

PÁDUA, G. P. DE et al. Response of soybean genotypes to the expression of greenish seed under temperature and water stresses. **Revista Brasileira de Sementes**, Brasília, DF, v. 31, n. 3, p. 140–149, 2009.

RITCHIE, S. W.; HANWAY, J. J.; THOMPSON, H. E. **How a Soybean Plant Develops**. Des Moines, United States, p.1-24, 1982.

RUIJTER, J. M. et al. Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications. **Methods**, San Diego, United States, v. 59, n. 1, p. 32–46, 2013.

SAMOYLOV, V. M.; TUCKER, D. M.; PARROTT, W. A. Soybean [Glycine max (L.) Merrill] embryogenic cultures: The role of sucrose and total nitrogen content on proliferation. **In Vitro Cellular and Developmental Biology - Plant**, Heidelberg, Germany, v. 34, n. 1, p. 8–13, 1998.

SCHMIDT, M. A. et al. Towards normalization of soybean somatic embryo maturation. **Plant Cell Reports**, Heidelberg, Germany, v. 24, n. 7, p. 383–391, 2005.

SCHMIDT, M. A.; HERMAN, E. M. Suppression of soybean oleosin produces micro-oil bodies that aggregate into oil body/ER complexes. **Molecular Plant**, Cambridge, United States v. 1, n. 6, p. 910–924, 2008.

TEIXEIRA, R. N. et al. Gene expression profiling of the greenish seed problem in Soybean. **BMC Plant Biology**, London, United Kingdom, v. 16, n. 1, p. 1–15, 1 fev. 2016.

WRIGHT, M. S. et al. A simple method for the recovery of multiple fertile plants from individual somatic embryos of soybean [Glycine max (L.) merrill]. **In Vitro Cellular & Developmental Biology - Plant**, Heidelberg, Germany, v. 27, n. 3, p. 153–157, jul. 1991.

ZINSMEISTER, J. et al. ABI5 is a regulator of seed maturation and longevity in legumes. **The Plant Cell**, Cary, United States, v. 28, n. 11, p. 2735–2754, 12 dez. 2016.

ZINSMEISTER, J. et al. ABSCISIC ACID INSENSITIVE 4 coordinates eoplast formation to ensure acquisition of seed longevity during maturation in Medicago truncatula. **The Plant Journal**, Chichester, United Kingdom, v. 113, n. 5, p. 934–953, 2023.

CHAPTER 3

OSMO-PRIMING IN TOMATO SEEDS DOWN-REGULATES GENES ASSOCIATED WITH STRESS RESPONSE AND LEADS TO REDUCTION IN LONGEVITY1

ABSTRACT

Tomato seeds subjected to osmo-priming show fast and more uniform germination. However, osmo-priming reduces seed longevity, which is a complex seed physiological attribute influenced by several mechanisms, including response to stress. Thus, to have new insights as to why osmo-primed tomato seeds show a short life span, we performed a transcript analysis during their priming. For that, we performed gene expression studies of the heat-shock protein family genes that were previously reported to be associated with the enhancement of longevity in primed tomato seeds. Physiological assays of germination, vigour and longevity tests were used to support the data. The results show that the short life span of osmo-primed tomato seeds is related to the decrease in the expression of transcripts associated with response to stress during the priming treatment. These results are important because they add information regarding which seed longevity mechanisms are impacted by the priming treatment. In parallel, it will allow the use of these genes as markers to monitor longevity in osmo-primed tomato seeds.

3.1 INTRODUCTION

Seed priming is a broadly used treatment by the seed industry aiming to improve the physiological performance of seed lots. During priming, the hydration of seeds is controlled allowing phases I and II of the germination process to start, and then the treatment is interrupted so that phase III (radicle protrusion) is deterred (Bradford, 1986). Fast and uniform germination are the main benefits of the priming technique. Seeds, such as Solanum lycopersicum, may have slow and uneven seed germination due to the mechanical restraint to radicle protrusion imposed by the micropylar endosperm (Toorop et al., 2008; Nonogaki et al., 2000; Bewley et al., 2013). The weakening of micropylar endosperm cells is a prerequisite for radicle protrusion of tomato seeds (Toorop et al., 2008). Thus, priming has been used to increase the speed

[1] Chapter written in accordance with the *Seed Science Research* journal's rules

of germination in tomato seeds (Batista et al., 2020). Apparently, priming improves physiological performance by favouring changes in the embryo and in the micropylar endosperm. For instance, Anese et al. (2011) showed that in *Solanum lycocarpum* seeds, growth of the embryo occurs concomitantly with weakening of the micropylar endosperm during the priming treatment.

Although the priming treatment improves physiological performance as mentioned earlier, it negatively affects seed longevity as was shown by Liu et al. (1996), Bruggink et al. (1999), Buitink et al. (2000) and Batista et al. (2020). Seed longevity is the ability of a seed to remain viable during storage in the dry state (Leprince et al., 2017) and is acquired over seed maturation. Several elements are involved in this process, including the non-reducing sugars, which play a role in membrane protection and form the cytoplasm glassy state, late embryogenesis abundant proteins (LEAs), the seed coat, which protects against oxygen absorption by the embryo, RNA-binding proteins that conserve seed mRNA in the dry state, an array of antioxidant molecules, and heat-shock proteins (HSPs), which are protective proteins that play a role as chaperones under stress conditions, according to the review by Zinsmeister et al. (2020).

According to Batista et al. (2020), some of the mechanisms associated with seed longevity are negatively impacted during the priming treatment. Sano and Seo (2019) proposed that the short life span of primed *Arabidopsis thaliana* seeds is associated with the progress of the cell cycle during the priming treatment. On the other hand, Wang et al. (2018) showed that primed and stored rice seeds presented events associated with the reduction in starch metabolism, the consumption of starch reserves from the endosperm, the accumulation of malondialdehyde and the decrease of antioxidant activities of the enzymes, which could explain the short longevity observed.

Recently, Batista et al. (2020) showed that a heat-shock treatment is able to enhance longevity in primed tomato seeds and preserve the vigour of the seeds during storage. The authors showed that transcripts associated with stress response are up-regulated after the heat-shock treatment in primed tomato seeds. They proposed that these transcripts may have a positive influence on the enhancement of longevity. Based on these results and considering the importance of stress response molecules to maintain seed longevity in primed tomato seeds, we hypothesize that the loss of storability observed in osmo-primed tomato seeds is due to the down-regulation of genes associated with stress response that apparently occurs during the priming treatment. To confirm this, we tested this hypothesis by performing gene expression

studies of transcripts related to stress response during the osmo-priming treatment in tomato seeds.

3.2 MATERIAL AND METHODS

Seed production

Tomato seeds from the LA1509 accession, donated by the Tomato Genetics Resource Center (https://tgrc.ucdavis.edu/), were produced as described in Batista et al. (2020). Red fruits without the presence of greenish color were harvested. The fruits were cut with the aid of a knife and the seeds were extracted by hand. The seeds were then treated with a sodium hypochlorite solution at an initial concentration of 9% at a ratio of 1:1 of seeds for a period of 30 min to remove the remnants of the fruit. Seeds were stored at 12°C and 60% ±2% relative humidity (RH) until the beginning of the experiment. The seed lot was mixed and empty seeds were discarded.

Priming, drying and heat-shock treatment

The priming treatment was performed according to Batista et al. (2020), which consisted of placing the seeds into tubes containing a polyethylene glycol (PEG) 6000 solution with an osmotic potential of −1.0 MPa at 20°C, for 60 h in the dark. The tubes were placed over a mixer to shake the solution. The PEG solutions were renewed three times to avoid changes in the osmotic potential. After the priming treatment, seeds were washed in running water, placed over paper towels, and kept for 24 h at 20°C and 60% ±2% RH for drying. In parallel, another group of seeds were placed over paper towels and exposed to 38°C and 32% RH, in an oven with air circulation for 2 h following the heat-shock treatment protocol as defined by Batista et al. (2020). After that, these seeds were maintained in the drying conditions described above. The unprimed seeds were used as a control group.

Physiological assays

After treatment, the seeds were stored in hermetic glass pots and placed in a cold chamber at 12°C and 60% RH ±2% which resulted in a moisture content of 0.08 ± 0.01 g H2O/g DW−1. After that, the following tests were performed:

Seed germination and vigour – Seven replications of 50 seeds each were germinated in plastic boxes (11 × 11 × 3.5 cm) with a substrate of paper towel moistened with distilled water equal to 2.5 times the weight of the substrate at 25°C, in an 8 h light and 16 h dark cycle. A radicle length of 2 mm or more was used as a

germination criterion. We determined seed vigour by first germination count on the 5th day and the calculation of t50 (time to 50% of germination) through the analysis of cumulative germination data. The curve fitting module of the Germinator software package was used to calculate t50 (Joosen et al., 2010). Total germination was determined on the 14th day.

Longevity – This test was performed by placing the seeds on a support over a saturated solution of NaCl (75% RH) at 35°C in hermetically sealed glass bottles. At different time intervals, seed viability was assessed by placing the seeds under the same conditions as described for the germination assay. The viability data were transformed into probits to determine the moment when the germination was reduced to half (p50), by using the equation: $v = Ki - p/\sigma$, according to Ellis and Roberts (1980) where v is viability in days, Ki is the initial germination in probit values, p is expected death over time and σ is the slope of the curve.

RNA extraction and quantitative real-time PCR

Three biological samples of 100 seeds each were collected from seeds that had been primed, unprimed and primed plus heat shock. At the same time, we collected three biological samples of 100 seeds each at 3, 12, 42 and 60 h during priming (in phase II, data not shown). All seeds were stored at −80°C. Total RNA was extracted using the NucleoSpin® RNA plant commercial Kit (Macherey-Nagel, Düren, Germany) following the manufacturer's instructions. The purity and quantity of total RNA was checked in a Nanodrop-2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The integrity of the RNA was checked in agarose gel at 1%.

cDNA was synthesized by using a High-Capacity cDNA Reverse Transcription commercial kit (Applied Biosystems, Victoria, Australia) following the manufacturer's instructions. For a 20 µl reaction, we used 2.0 µl 10× RT Buffer, 0.8 µl 25× dNTP, 2.0 µl 10× RT Primer, 1 µl Reverse Transcriptase, 10 µl of the extracted RNA and 4.2 µl Nuclease-free water. The reaction was incubated in a thermocycler following these steps: 5 min at 25°C, 2 h at 37°C, followed by 5 min at 85°C and the synthesis was ended at a constant temperature of 4°C.

The genes used for Real-Time PCR were selected based on the work done by Batista et al. (2020). In order to confirm the effect of priming, we selected expansin which is a gene related to germination (Chen and Bradford, 2000; Chen et al., 2001). Thus, the following genes were studied: *HEAT STRESS TRANSCRIPTION FACTOR*

B-2B (HSFB2b), DNAJ PROTEIN HOMOLOG, HEAT SHOCK PROTEIN 70 (HSP70), SMALL HEAT SHOCK PROTEIN PRECURSOR (er-HSP), 15.7 KDA HEAT SHOCK PROTEIN (HSP15.7) and EXPANSIN (EXP2). The primers used were efficient above 1.8 and r2 of approximately 1.0, according to the parameters of the LinRegPCR program (Ruijter et al., 2013). The primers used are listed in Table 1.

We performed the gene expression analysis in a thermocycler Eco Real-Time (Illumina, San Diego, USA) with SYBR Green qPCR ReadyMix (Sigma-Aldrich, St. Louis, USA), using two technical samples for each biological sample. For a reaction of 10 µl, 5 µl of SYBR Green, 1 µl of cDNA and 0.25 µl of each primer were used and the volume was adjusted with Nuclease-free water. The amplification consisted of 2 min at 50°C, 2 min at 95°C; then 45 cycles of 10 s at 95°C and 1 min at 60°C. At the end of the process, the melting curve was performed following these steps: 15 s at 95, 65 and 95°C, respectively. The normalized expression (NE) was assessed using a geometric means of three reference genes, *UBIQUITIN CONJUGATING ENZYME 21 (UBC21), METALLOTHIONEIN (MTP)* and *ARF-LIKE GTPASE FAMILY PROTEIN (ASAR1)*, calculated according to: $NE = 2\Delta Ct$, where ΔCt is the geometric mean of the Cq reference genes – Cq target gene.

Table 1. Primer sequences used as target and reference genes (mRNAs) in RT-qPCR reactions.

Gene name	Forward (5' – 3')	Reverse (5' – 3')	Bp
UBC21*	GGACGGCTCTTGTTAAAGG	TGGATACTGCTCTGGAACTG	86
MTP*	CTACACCGAAAGCAGCAC	CAGCCATTCTCAGCAACAG	110
ASAR1*	TAGCGACTGTTCCCTTCC	TTACCCTTGCCAGTAGTGAC	114
HSFB2b	ATGAAGATATGAGCCCACGG	GCGGTTGTACTTGATCCTG	102
dnaJ protein homolog	TCTCCAACAGAAAGATCACCC	CCAAGTGTGCCAATATGAACTG	116
HSP 70	AATCCCTCCAGCTCCCAG	GCCGTGACAGAAAGAATACC	81
HSP 15.7	GACGAATTCCACGGTAAAGAG	TGATCCACCTTCACATCTTCC	116
er-HSP	ACCAAATGATAAGCAGCAATCC	TTCGCCGTCTCTTTCCAG	121
EXP	ACTTGTGGTGCTTGTTATGAG	TGTTAGGTAGAGACGGGTTC	115

*reference genes; bp) base pairs

Statistical analysis

We checked the normality of the data by the Shapiro–Wilk test. The data fulfilled the assumption for normality. Thus, a one-way analysis of variance (ANOVA) was performed and the significant physiological characteristics or transcript levels were separated using Fisher's Least Significant Difference (LSD) test at 0.05-confidence level. The sigmoidal behaviour during storage was adjusted using the Boltzmann equation parameters.

3.3 RESULTS

Physiological characterization

Unprimed seeds started to germinate after 60 h of imbibition and reached 50% of germination (t50) in 104 h, while primed and primed plus heat shock seeds started to germinate after 36 h, and t50 was reached in 64 h (Fig. 1A). Therefore, primed and primed plus heat shock seeds showed an increase in the speed of germination with a significant reduction ($P < 0.001$) in t50 of 40 h in relation to unprimed seeds (Fig. 1B).

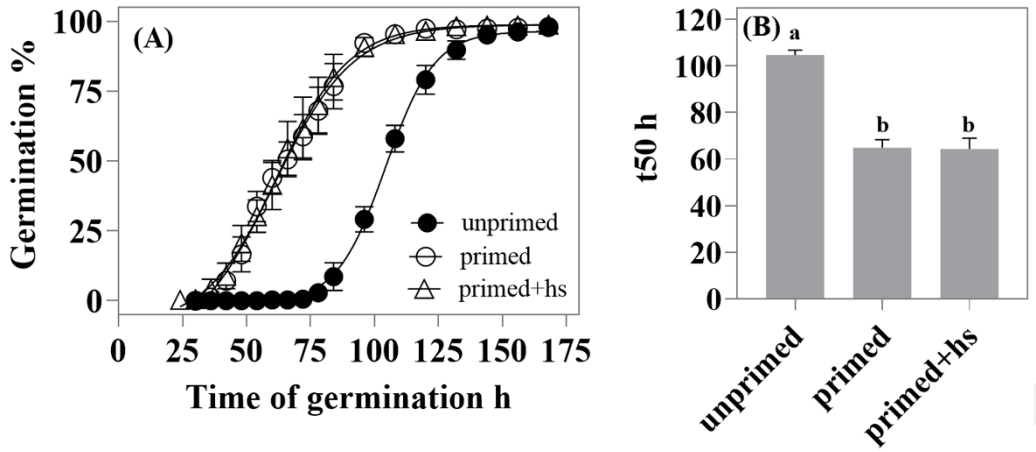


Fig. 1. Effect of the priming and heat-shock treatments on seed germination and vigor in *S. lycopersicum* seeds over time. Seed germination time curve (A). Time to 50% of germination – t50 (B) in unprimed, primed, and primed plus heat shock seeds. Different letters indicate a significant difference ($P \leq 0.05$) by Fisher's LSD test. n.s., not significant. Error bars show the standard deviation from seven samples.

Consequently, there was an increase ($P < 0.001$) greater than 17% in the first germination count in primed and primed plus heat shock seeds in relation to unprimed seeds (Fig. 2A). Unprimed, primed and primed plus heat shock seeds reached their maximum germination rate (≥98%) on the 7th day (Fig. 1A). Thus, the priming and heat-shock treatment did not affect the final germination (Fig. 2B).

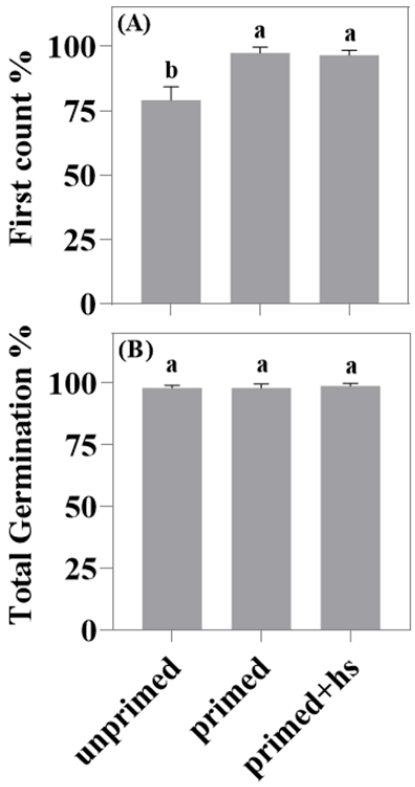


Fig. 2. Effect of the priming and heat-shock treatments on seed germination and vigour in *S. lycopersicum* seeds. First germination count on the 5th day (A) and total germination (B) in unprimed, primed and primed plus heat shock seeds. Different letters indicate a significant difference ($P \leq 0.05$) by Fisher's LSD test. n.s., not significant. Error bars show standard deviation from seven samples.

At the beginning of storage, unprimed, primed and primed plus heat shock seeds displayed 100% germination. Nevertheless, the germination of primed seeds was reduced to 50% on the 26th day, while primed plus heat shock seed germination was reduced to 50% on the 42nd day, and unprimed seeds reached p50 after 112 d (Fig. 3A). Thus, there was a significant reduction ($P < 0.001$) of 86 d in longevity (p50, days) of primed seeds in relation to unprimed seeds (Fig. 3B). The heat-shock treatment promoted an increase in longevity of 16 d compared to primed seeds ($P < 0.001$).

Through controlled storage, it was possible to verify the reduction in storability for primed and primed plus heat shock seeds, which lost germination capacity completely after 100 d, while more than 85% of unprimed seeds remained viable (Fig. 3A).

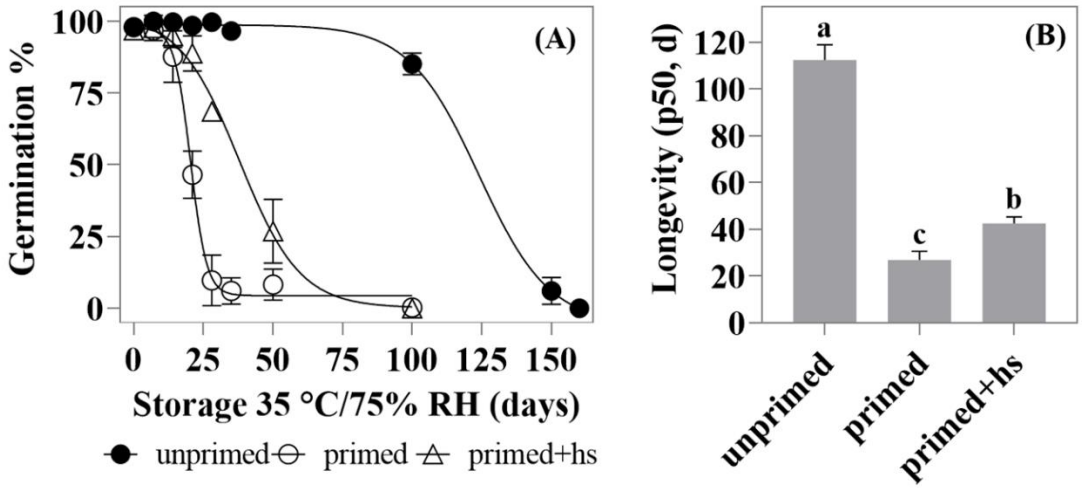


Fig. 3. Germination during storage and seed longevity in primed, primed plus heat shock and unprimed *S. lycopersicum* seeds. Germinability during storage (A) and longevity expressed in p50 (the number of days in which the seed lot has lost 50% of viability during storage) (B) from unprimed, primed and primed plus heat shock seeds. Germination during storage was fitted using Boltzmann sigmoid [y = top + (bottomtop)/ 1 + [v50 − (x/slope)]. Different letters indicate a significant difference (P ≤ 0.05) by Fisher's LSD test. Error bars show standard deviation from seven samples.

The pattern of Gene expression profile during Priming

We found that for primed seeds without the heat-shock treatment, there was a significant reduction ($P < 0.001$) in the transcript levels of HSFB2b; HSP 70; DNAJ PROTEIN HOMOLG; ER-HSP and HSP 15.7 over the process of priming, mainly from 12 h of the priming treatment, which was not re-established after drying (Fig. 4A–E). The expression of HSP 15.7 increased at 3 h of priming in relation to the unprimed seeds, however this decreased after 12 h, similar to the other genes after this period (Fig. 4E).

However, in primed plus heat shock seeds, there was an increase in the expression level of genes related to the response to stress in relation to seeds collected at 42 and 60, and primed seeds (Fig. 4A until E) when the reduction of transcript levels was drastic. Nonetheless, the heat-shock treatment re-established the transcript levels of HSFB2b, DNAJ PROTEIN HOMOLG and ER-HSP genes to the same level as found at 12 h (Fig. 4A, B, D).

To confirm the effect of the priming treatment at the molecular level, we studied the pattern of EXP expression levels. Thus, we verified that the expression increased after 42 h during the priming treatment, and in primed seeds, it was higher than in unprimed seeds (Fig. 4F).

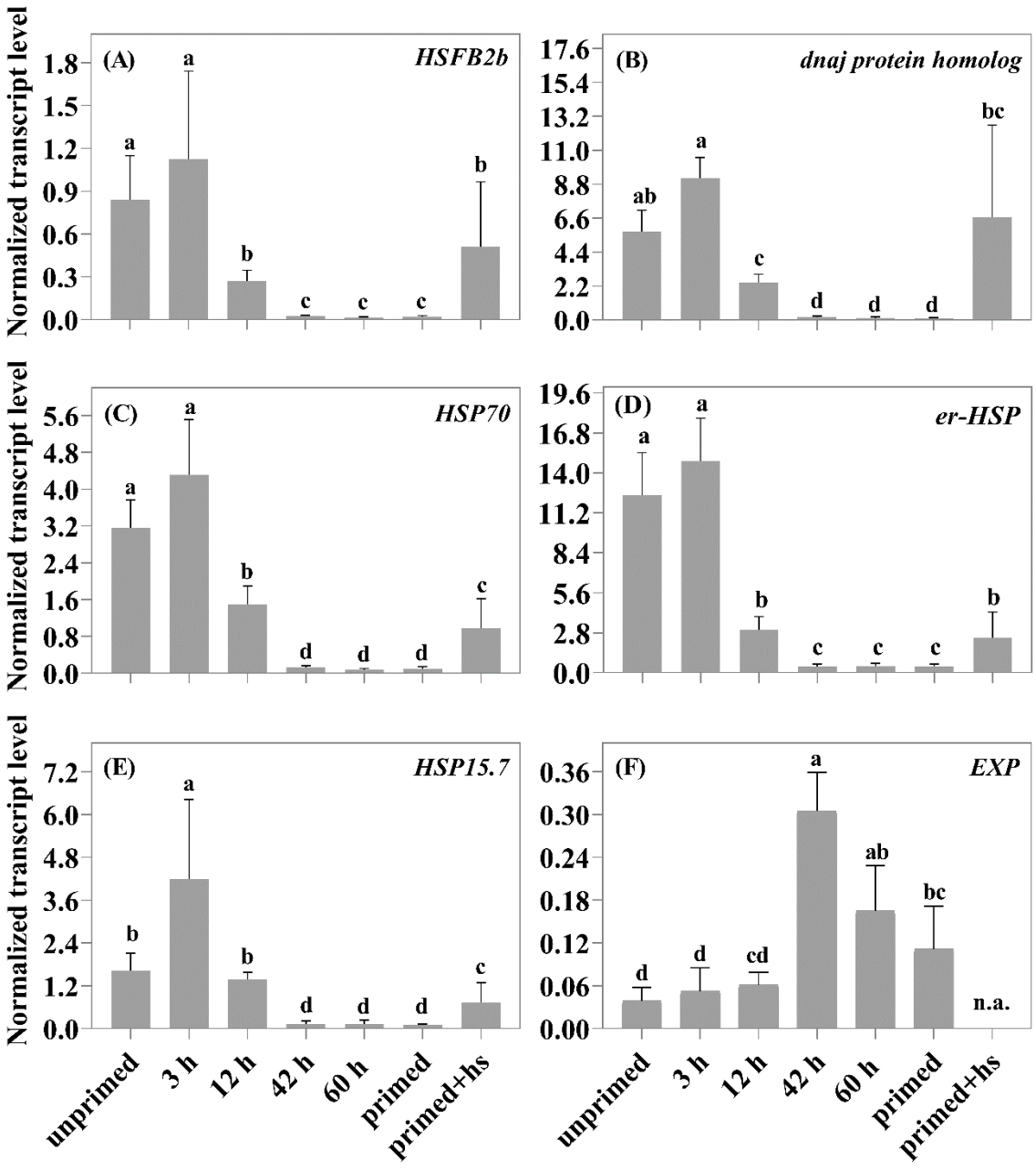


Fig. 4. Gene expression before (unprimed seeds), during and after the priming treatment and plus heat-shock treatment of *S. lycopersicum* seeds. Different letters indicate a significant difference (P ≤ 0.05) by Fisher' LSD test. n.a., not available. Error bars indicate standard deviation from six technique samples.

3.4 DISCUSSION

Seed priming is an important technique used by the seed industry. Faster and more uniform seed germination are considered the main benefits promoted by the priming treatment. However, primed seeds have shorter longevity. Seed longevity maintains their life span during storage, which is important to ensure the propagation of the seeds over the time. Here, we demonstrated that during osmo-priming of tomato seeds, the

genes that code for proteins associated with longevity in primed seeds have their expression pattern affected.

In our study, the priming protocol increased the performance of *S. lycopersicum* seeds and the heat-shock treatment did not interfere with this, which was verified through the speed of germination and consequently germination rate in the first count (Figs 1 and 2A). However, there was a notable reduction in seed longevity in the primed seeds (Fig. 3), confirming previous studies by Gurusinghe et al. (2002) and Batista et al. (2020). This reduction in longevity is less pronounced when using the heat-shock treatment after priming (Fig. 3) as previously reported by Batista et al. (2020) and Gurusinghe et al. (2002).

Previous studies were performed to comprehend the mechanism associated with the loss of seed longevity in primed seeds. In *Impatiens* and pepper seeds, the reduced longevity after priming was unrelated to increased molecular mobility in the cytoplasm (Buitink et al., 2000). In primed Arabidopsis seeds, the short longevity was correlated with the advancement of the cell cycle (Sano and Seo, 2019).

Nevertheless, there is still a knowledge gap about how exactly this loss of longevity is controlled. Gurusinghe et al. (2002) found increased expression of BiP genes in extended longevity by heat treatment of primed tomato seeds. BiP family are heat shock protein family homologous. Recently, Batista et al. (2020) demonstrated that the transcripts associated with protection from stress are involved in the enhancement of longevity in primed tomato seeds. Thus, considering the role of the protection mechanism in seed longevity as reviewed by Zinsmeister et al. (2020), we investigated the expression pattern of transcripts associated with stress response during osmo-priming to confirm whether there is a significant change in the expression level of these genes that could be responsible for the short longevity of osmo-primed tomato seeds.

Our hypothesis proved correct since our results demonstrated that there is down-regulation of HSFB2b (Fig. 4A), which leads to a reduction in transcripts that code for HSPs (Fig. 4B until E) which are associated with stress protection (Guo et al., 2016). The heat-shock protein family is important to seeds due to their chaperone role in protecting against cellular damage under stress conditions. The HSPs are present during seed development and are always related to the acquisition of germinability, desiccation tolerance and consequently longevity (Kaur et al., 2016). It was demonstrated that heat shock factor A9 acts against deterioration in transgenic tobacco seeds (Prieto-Dapena et al., 2006), which implies its role in storability.

HSFA6B and sHSPs were correlated with the increase in seed longevity in soybean (Lima et al., 2017). Kaur et al. (2015) reported that OsHSP18.2 in *A. thaliana* is an ageing responsive protein that possibly protects and stabilizes the cellular proteins during maturation drying, desiccation and ageing in seeds by restricting reactive oxygen species accumulation and thereby improving seed vigour, longevity and seedling establishment. According to Zhang et al. (2018), a cytosolic class II small heat-shock protein, PfHSP17.2, confers resistance to heat stress in transgenic Arabidopsis, which may result in resistance to deterioration. The DnaJ proteins resemble HSPs and act as chaperones in response to heat stress and plant growth and development (Fan et al., 2017); the DnaJ protein homologous gene was associated with the enhanced longevity of primed tomato seeds, as shown by Batista et al. (2020).

The studies reported earlier showed an important role of the HSP family in seed quality, especially seed storability, and as demonstrated here, the expression of these genes was reduced in tomato seeds by the priming treatment (Fig. 4A until E). According to Batista et al. (2020), these genes are important to maintain viability in the dry state during storage in primed tomato seeds.

Our experiments in primed and primed plus heat-shock treatment seeds confirmed the importance of these transcripts to enhance longevity in primed tomato seeds as mentioned earlier, since it was demonstrated that there was an increase in the gene expression of transcripts related to seed longevity, and consequently, an enhancement of storability in the primed tomato seeds subjected to heat-shock treatment in comparison to seeds that were only primed (Figs 3 and 4A until E).

In addition, during seed germination, there is an activation of the protection system that protects the seeds against damage and reduces the effects of ageing (Bewley et al., 2013). In our study, the expression of HSP 15.7 increased at the beginning of the treatment while other genes maintained their initial level, and was gradually reduced with the advance of the treatment (Fig. 4A until E). As a result, after drying (primed seeds), the expression of genes related to stress response was lower than the initial level (Fig. 4A until E), which is associated with the shorter longevity found in osmo-primed tomato seeds (Fig. 3). Apparently, the deterioration process may begin already during the treatment.

Thus, transcripts that code for molecules associated with stress response are compromised during priming. It is true that several other mechanisms are associated

with longevity as mentioned earlier. However, the findings of Batista et al. (2020) have shown that the stress response genes are involved in enhancing the longevity in primed tomato seeds. This led us to explore these genes during priming and better understand how the reduction in the expression of these transcripts by priming caused the shorter life span in primed tomato seeds (Figs 3 and 4A until E). In addition, this research opens the possibility to use some of these transcripts as markers to monitor longevity of primed tomato seeds during storage.

REFERENCES

Anese S, da Silva EAA, Davide AC, Rocha Faria JM, Soares GCM, Matos ACB and Toorop PE (2011) Seed priming improves endosperm weakening, germination, and subsequent seedling development of Solanum lycocarpum St. Hil. Seed Science and Technology 39, 125–139.

Batista TB, Fernandez JG, da Silva TA, Maia J and da Silva EAA (2020) Transcriptome analysis in osmo-primed tomato seeds with enhanced longevity by heat shock treatment. AoB Plants 12, 1–10.

Bewley JD, Bradford KJ, Hilhorst HWM and Nonogaki H (2013) Seeds: physiology of development, germination and dormancy. New York,Springer.

Bradford KJ (1986) Manipulation of seed water relations via osmotic priming to improve germination under stress conditions. Horticulture Science 21, 1105–1112.

Bruggink GT, Ooms JJJ and van der Toorn P (1999) Induction of longevity in primed seeds. Seed Science Research 9, 49–53.

Buitink J, Hemminga MA and Hoekstra FA (2000) Is there a role for oligosaccharides in seed longevity? An assessment of intracellular glass stability. Plant Physiology 122, 1217–1224.

Chen F and Bradford KJ (2000) Expression of an expansin is associated with endosperm weakening during tomato seed germination. Plant Physiology 124, 1265–1274.

Chen F, Dahal P and Bradford KJ (2001) Two tomato expansin genes show divergent expression and localization in embryos during seed development and germination. Plant Physiology 127, 928–936.

Ellis RH and Roberts EH (1980) Improved equations for the prediction of seed longevity. Annals of Botany 45, 13–30.

Fan F, Yang X, Cheng Y, Kang Y and Chai X (2017) The DnaJ gene family in pepper (Capsicum annuum L.): comprehensive identification, characterization and expression profiles. Frontiers in Plant Science 8, 689.

Guo M, Liu J, Ma X, Luo D, Gong Z and Lu M (2016) The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses. Frontiers in Plant Science 7, 1–13.

Gurusinghe S, Powell ALT and Bradford KJ (2002) Enhanced expression of BiP is associated with treatment that extend storage longevity of primed tomato seeds. Journal of the American Society for Horticultural Science 127, 528–534.

Joosen RVL, Kodde J,Willems LAJ, Ligterink W, van der Plas LH Wand, Hilhorst HWM (2010) Germinator: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. Plant Journal 62, 148–159.

Kaur H, Petla BP, Kamble NU, Singh A, Rao V, Salvi P, Ghosh S and Majee M (2015) Differentially expressed seed aging responsive heat shock protein OsHSP18.2 implicates in seed vigor, longevity and improves germination and

seedling establishment under abiotic stress. Frontiers in Plant Science 6, 713.

Kaur H, Petla BP and Majee M (2016) Small heat shock proteins: roles in development, desiccation tolerance and seed longevity, pp. 3–18 in Asea A; Kaur P and Calderwood S (Eds) Heat shock proteins and plants. Cham, Springer International Publishing Switzerland.

Leprince O, Pellizzaro A, Berriri S and Buitink J (2017) Late seed maturation: drying without dying. Journal of Experimental Botany 68, 827–841.

Lima JJP, Buitink J, Lalanne D, Rossi RF, Pelletier S, da Silva EAA and Leprince O (2017) Molecular characterization of the acquisition of longevity during seed maturation in soybean. PLoS ONE 12, 1–25.

Liu Y, Bino RJ, van der Burg WJ, Groot SPC and Hilhorst HWM (1996) Effects of osmotic priming on dormancy and storability of tomato (Lycopersicon esculentum Mill.) seeds. Seed Science Research 6, 49–55.

Nonogaki H, Gee OH and Bradford KJ (2000) A germination-specific endo-β-mannanase gene is expressed in the micropylar endosperm cap of tomato seeds. Plant Physiology 123, 1235–1245.

Prieto-Dapena P, Castano R, Almoguera C and Jordano J (2006) Improved resistance to controlled deterioration in transgenic seeds. Plant Physiology 142, 1102–1112.

Ruijter JM, PfafflMW, Zhao S, Spiess AN, Boggy G, Blom J, Rutledge RG, Sisti D, Lievens A, De Preter K, Derveaux S, Hellemans J and Vandesompele J (2013) Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications. Methods 59, 32–46.

Sano N and Seo M (2019) Cell cycle inhibitors improve seed storability after priming treatments. Journal of Plant Research 132, 263–271.

Toorop PE, Van Aelst AC and Hilhorst HWM (2008) Endosperm cap weakening and endo-β-mannanase activity during priming of tomato (Lycopersicon esculentum cv. Moneymaker) seeds are initiated upon crossing a threshold water potential. Seed Science Research 8, 483–492.

Wang W, He A, Peng S, Huang J, Cui K and Nie L (2018) The effect of storage condition and duration on the deterioration of primed rice seeds. Frontiers in Plant Science 9, 1–17.

Zhang L, Hu W, Gao Y, Pan H and Zhang Q (2018) A cytosolic class II small heat shock protein, PfHSP17.2, confers resistance to heat, cold, and salt stresses in transgenic Arabidopsis. Genetics and Molecular Biology 41, 3, 649–660.

Zinsmeister J, Leprince O and Buitink J (2020) Molecular and environmental factors regulating seed longevity. Biochemical Journal 477, 305–323.

FINAL CONSIDERATIONS

This research contributed to the advancement of knowledge and technology in the area of seed physiology. The maternal environment directly impacts seed quality, and due to climate change creating challenging environments, reductions in seed quality have been noticed season after season. In this regard, the multispectral imaging technology associated with machine learning proposed here is a robust way of classifying seeds that have suffered stress during their formation. Besides this, since it is an autonomous system, it can be incorporated into the seed companies' daily routine. The greatest advantage however, is that this information can help seed companies better understand the causes of reduced quality in a seed lot. Alternatively, the overexpression and silencing of the ABI5 gene in soybean will enable, in further studies, a better understanding of the role this gene plays in the quality and retention of chlorophyll in soybean seeds. On the other hand, priming is a technique broadly used by companies for tomato seeds. Nevertheless, it impairs seed shelf life. Here we elucidated one of the mechanisms involved in reducing the longevity of primed tomato seeds and suggested the use of this knowledge as a marker to monitor the longevity of primed seeds. Thus, concepts generated here can be used at industrial and academic levels to solve classic problems regarding seed physiological quality.

REFERENCES

AJALA-LUCCAS, D. et al. The Seed–Seedling Transition in Commercial Soybean Cultivars with the Presence of Greenish Seeds in the Sample: A Perspective from Classical Genetic Parameters. **Agronomy**, Basel, Switzerland, Vol. 13, Page 1966, v. 13, n. 8, p. 1966, 26 jul. 2023.

BAGATELI, J. R. et al. Productive performance of soybean plants originated from seed lots with increasing vigor levels. **Journal of Seed Science**, Brasília, DF, v. 41, n. 2, p. 151–159, 2019.

BARBOZA DA SILVA, C. et al. Autofluorescence-spectral imaging as an innovative method for rapid, non-destructive and reliable assessing of soybean seed quality. **Scientific Reports**, London, United Kingdom, v. 11, n. 1, 1 dez. 2021b.

BATISTA, T. B. et al. A reliable method to recognize soybean seed maturation stages based on autofluorescence-spectral imaging combined with machine learning algorithms. **Frontiers in Plant Science**, Lausanne, Switzerland, v. 13, 2022.

BATISTA, T. B. **MATURAÇÃO EM SEMENTES DE SOJA: ESTÁDIOS REPRODUTIVOS, DEGRADAÇÃO DA CLOROFILA E AQUISIÇÃO DA QUALIDADE FISIOLÓGICA**. 2022. Tese (Doutorado em Agronomia/Agricultura). Faculdade de Ciências Agronômicas, UNESP, Botucatu, SP, 2022.

BEWLEY, J. D. et al. **Seeds: Physiology of Development, Germination and Dormancy**. 3rd Edition. New York, United States: Springer, 2013.

BIANCHINI, V. DE J. M. et al. Multispectral and X-ray images for characterization of Jatropha curcas L. seed quality. **Plant Methods**, London, United Kingdom, v. 17, n. 1, 2021.

BROCKLEHURST, P. A.; DEARMAN, J. Interactions between seed priming treatments and nine seed lots of carrot, celery and onion. **Annals of Applied Biology**, Chichester, United Kingdom, v.102, p.577-584, 1983.

CARVALHO, N. M. DE; NAKAGAWA, J. Importância da semente. *In*: CARVALHO, N. M. DE; NAKAGAWA, J. **Sementes: Ciência, tecnologia e produção**. 5. ed. Jaboticabal, SP: Funep, 2012, p. 6–11.

CHAKMA, R. et al. Foliar application and seed priming of salicylic acid affect growth, fruit yield, and quality of grape tomato under drought stress. **Scientia Horticulturae**, Amsterdam, Netherlands, v. 280, 5 abr. 2021.

COTRIM, M. F. et al. Studying the link between physiological performance of Crotalaria ochroleuca and the distribution of Ca, P, K and S in seeds with X-ray fluorescence. **PLoS ONE**, San Francisco, United States, v. 14, n. 9, 1 set. 2019.

DIAZ, C. et al. Seed quality and effect on rice yield: findings from farmer participatory experiments in Central Luzon, Philipines. **Philippine Journal of Crop Sciences**, Laguna, Phillipines, v. 23, n. 2, p. 111–119, 1998.

DUCATTI, K. R. et al. Transcripts Expressed during Germination Sensu Stricto Are Associated with Vigor in Soybean Seeds. **Plants**, Basel, Switzerland, v. 11, n. 10, 2022.

ELMASRY, G. et al. Recent applications of multispectral imaging in seed phenotyping and quality monitoring—An overview. **Sensors**, Basel, Switzerland, v.19, n. 5, 2019.

FAROOQ, M. et al. Influence of seed priming techniques on grain yield and economic returns of bread wheat planted at different spacings. **Crop and Pasture Science**, Clayton, Australia, v. 71, n. 8, p. 725–738, 1 ago. 2020.

FINCH-SAVAGE, W. E. Influence of seed quality on crop establishment, growth, and yield. *In*: BASRA, A. S. **Seed quality: basic mechanisms and agricultural implications**. New York, United States: Food Products Press, 1995. p. 361–385.

FONSECA DE OLIVEIRA, G. R. et al. An approach using emerging optical technologies and artificial intelligence brings new markers to evaluate peanut seed quality. **Frontiers in Plant Science**, Lausanne, Switzerland, v. 13, 2022.

FORTI, C. et al. Molecular dynamics of pre-germinative metabolism in primed eggplant (Solanum melongena L.) seeds. **Horticulture Research**, v. 7, n. 1, 1 dez. 2020.

GALLETTI, P. A. et al. Integrating Optical Imaging Tools for Rapid and Non-invasive Characterization of Seed Quality: Tomato (Solanum lycopersicum L.) and Carrot (Daucus carota L.) as Study Cases. **Frontiers in Plant Science**, Lausanne, Switzerland, v. 11, 21 dez. 2020.

GOMES-JUNIOR, F. G. et al. X-ray microtomography in comparison to radiographic analysis of mechanically damaged maize seeds and its effect on seed germination. **Acta Scientiarum - Agronomy**, Maringá, PR, v. 41, n. 1, 2019.

GRAVEN, L. M.; CARTER, P. R. Seed quality effect on corn performance under conventional and no-tillage systems. **Journal of Production Agriculture**, Madison, United States, 1991.

HATZIG, S. et al. Hidden effects of seed quality breeding on germination in oilseed rape (Brassica napus L.). **Frontiers in Plant Science**, Lausanne, Switzerland, v. 9, 2018.

HUSSAIN, M. et al. Influence of seed priming techniques on the seedling establishment, yield and quality of hybrid sunflower. **International journal of agriculture & biology,** Faisalabad, Pakistan, 2020.

KALAJI, M. H.; PIETKIEWICZ, S. Review: some physiological indices to be exploited as a crucial tool in plant breeding. **Plant breeding and seed science**, Warszawa, Poland, v. 49, p. 1–21, 2004.

KRZYZANOWSKI, F. C. Relationship between seed technology research and federal plant breeding programs. **Scientia Agrícola**, Piracicaba, SP, n. 55, 1998.

LUCCAS, D. A. **Caracterização fisiológica, bioquímica e molecular em sementes de soja (Glycine max (L.) Merr.) com retenção de clorofila**. 2018. Tese (Doutorado em Agronomia/Agricultura). Faculdade de Ciências Agronômicas, UNESP, Botucatu, SP, 2018.

MAHMOOD-UR-REHMAN, M. et al. Seed priming with salicylic acid improve seed germination and physiological responses of carrot seeds. **Pakistan Journal of Agricultural Sciences**, Faisalabad, Pakistan, v. 57, n. 2, p. 351–359, 1 mar. 2020.

MARCOS-FILHO, J. **Fisiologia de sementes de plantas cultivadas**. 2. ed. Londrina, PR: Abrates, 2015.

MELLO FILHO, O. L. DE et al. Grain yield and seed quality of soybean selected for high protein content. **Pesquisa Agropecuária Brasileira**, Brasília, DF, n. 5, p. 445–450, 2004.

MONDO, V. H. V. et al. Avaliação de danos mecânicos em sementes de feijão por meio da análise de imagens. **Revista Brasileira de Sementes**, Brasília, DF, v. 31, n. 2, p. 27–35, 2009.

MONTEIRO, F. F. et al. Breeding for yield and seed quality in soybean. **Euphytica**, Dordrecht, Netherlands, v. 217, n. 12, 1 dez. 2021.

PÁDUA, G. P. DE et al. Tolerance level of green seed in soybean seed lots after storage. **Rev Bras Sementes**, Brasília, DF, v. 29, n. 3, p. 112–120, 2007.

PÁDUA, G. P. DE et al. Response of soybean genotypes to the expression of green seed under temperature and water stresses. **Revista Brasileira de Sementes**, Brasília, DF, v. 31, n. 3, p. 140–149, 2009.

PAPARELLA, S. et al. Seed priming: state of the art and new perspectives. **Plant Cell Reports**, Heidelberg, Germany, Springer Verlag, , 24 ago. 2015.

RAJALA, A. et al. Seed quality effects on seedling emergence, plant stand establishment and grain yield in two-row barley. **Agricultural and food science**, Jokioinen, Finland,v. 10, p. 228-234, 2011.

RAJCAN, I.; HOU, G.; WEIR, A. D. **Advances in breeding of seed-quality traits in soybean**. **Journal of Crop Improvement**, New York, Unites States, 2005.

SILVEIRA, A. DE S. et al. Osmopriming with selenium: physical and physiological quality of tomato seeds in response to water deficit. **Journal of Seed Science**, Brasília, DF, v. 45, 2023.

TEIXEIRA, R. N. et al. Gene expression profiling of the green seed problem in Soybean. **BMC Plant Biology**, London, United Kingdom, v. 16, n. 1, p. 1–15, 1 fev. 2016.

TEKRONY, D. M. et al. Relationship of Seed Vigor to Crop Yield: A Review. **Crop Science**, Hoboken, United States, v. 31, p. 816–822, 1991.

TRIPATHI, N.; KHARE, D. Molecular approaches for genetic improvement of seed quality and characterization of genetic diversity in soybean: a critical review. **Biotechnology Letters,** Dordrecht, Netherlands, v. 38, n. 10, 2016.