REVISTA
DESAFIOS
V.12, n.7, dezembro/2025 – DOI: 10.20873/saberesemcirculação3

ADVANCES IN REFERENCE GENE VALIDATION FOR GENE EXPRESSION STUDIES IN CASSAVA (*Manihot esculenta* Crantz): AN OVERVIEW OF THE MAIN GENES AND STUDIED CONDITIONS

AVANÇOS NA VALIDAÇÃO DE GENES DE REFERÊNCIA PARA ESTUDOS DE EXPRESSÃO GÊNICA EM MANDIOCA (Manihot esculenta CRANTZ): UM PANORAMA DOS PRINCIPAIS GENES E CONDIÇÕES ESTUDADAS

AVANCES EN LA VALIDACIÓN DE GENES DE REFERENCIA PARA ESTUDIOS DE EXPRESIÓN GÉNICA EN YUCA (Manihot esculenta CRANTZ): UNA VISIÓN GENERAL DE LOS PRINCIPALES GENES Y CONDICIONES ESTUDIADAS

Jaderson Roney Gomes de Oliveira:
Doutorando em Biodiversidade e Biotecnologia pelo Programa da Rede BIONORTE. Universidade Federal do Tocantins (UFT).
E-mail: jaderson.roney@uft.edu.br | https://orcid.org/0000-0001-8374-992X

Ivo Pontes Araújo:
Mestrando em Agroenergia Digital. Universidade Federal do Tocantins (UFT).
E-mail: ivo.pontes@uft.edu.br | https://orcid.org/0000-0002-7216-3656

Solange Aparecida Ságio:
Docente do Programa de Pós-Graduação em Agroenergia Digital, Universidade Federal do Tocantins (UFT).
E-mail: solangesagio@uft.edu.br | https://orcid.org/0000-0002-9329-4516

André Almeida Lima:
Doutor em Fisiologia Vegetal. Universidade Federal de Lavras (UFLA).
E-mail: andrelima4@yahoo.com.br | https://orcid.org/0000-0002-3262-7858

Horllys Gomes Barreto:
Docente do Programa de Pós-Graduação em Biodiversidade e Biotecnologia, Rede BIONORTE, Universidade Federal do Tocantins (UFT).
E-mail: horllys@uft.edu.br | https://orcid.org/0000-0001-7236-6296

Como citar este artigo:

OLIVEIRA, J. R. G. et al. Advances in reference gene validation for gene expression studies in cassava (*Manihot esculenta Crantz*): an overview of the main genes and studied conditions. **Desafios. Revista Interdisciplinar da Universidade Federal do Tocantins**. Palmas, v. 12, n. 7, p. 247-265, 2025. DOI: https://doi.org/10.20873/saberesemcirculação3

REVISTA DESAFIOS

ABSTRACT:

Reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) is widely used in gene expression studies, with the proper selection of reference genes being essential to ensure result reliability. This study analyzed the available literature to identify the most commonly studied, selected, and validated reference genes in gene expression research in cassava (*Manihot esculenta* Crantz). Through a bibliographic survey, five publications on the selection and validation of reference genes in cassava were analyzed. A total of 55 gene evaluations were conducted across the analyzed studies, considering that some genes were investigated more than once. The results indicated that *PP2A*, *UBQ10*, *GTPb*, *EF1*, *TUB*, *TBP*, *NAC*, *HisD*, *26S*, and *ZFP* are among the most stable genes in the selected studies. Among them, *PP2A* and *GTPb* stood out as the most stable genes of three and two studies, respectively. In contrast, widely used genes such as *ACT* and *GAPDH* were not validated in any of the analyzed studies when considering all samples from each experiment. The geNorm and NormFinder were the most frequently used algorithms to assess the stability of reference genes, while RefFinder was employed in only one study. This review highlights the importance of experimental validation specific to cassava and suggests that adopting multiple analysis tools may enhance the accuracy and reliability of gene expression studies.

KEYWORDS:RT-qPCR; Normalization; Gene Expression.

RESUMO:

A técnica de reação em cadeia da polimerase em tempo real quantitativa, com transcrição reversa (RT-qPCR), é amplamente utilizada em estudos de expressão gênica, sendo essencial a seleção adequada de genes de referência para garantir a confiabilidade dos resultados. Este estudo analisou a literatura disponível para identificar os genes de referência mais comumente estudados, selecionados e validados em pesquisas de expressão gênica em mandioca (Manihot esculenta Crantz). Por meio de um levantamento bibliográfico, foram analisadas cinco publicações sobre seleção e validação de genes de referência em mandioca. Um total de 55 avaliações de genes foram realizadas nos estudos analisados, considerando que alguns genes foram investigados mais de uma vez. Os resultados indicaram que os genes PP2A, UBQ10, GTPb, EF1, TUB, TBP, NAC, HisD, 26S e ZFP estão entre os mais estáveis nos estudos realizados. Dentre eles, destacam-se os genes PP2A e GTPb, que estão entre os mais estáveis em três e dois estudos, respectivamente. Em contraste, genes amplamente utilizados, como ACT e GAPDH, não foram validados em nenhum dos estudos analisados, ao analisar todas as amostras de cada experimento. Os algoritmos geNorm e NormFinder foram os mais utilizados para avaliar a estabilidade dos genes de referência, enquanto o RefFinder foi empregado em apenas um dos estudos. Esta revisão

destaca a importância da validação experimental específica para a cultura da mandioca, além de sugerir que a adoção de múltiplas ferramentas de análise pode aumentar a precisão e a confiabilidade dos estudos de expressão gênica.

***Palavras-chave**:*

PALAVRAS CHAVE: RT-qPCR; Normalização; Expressão Gênica.

RESUMEN:

La reacción en cadena de la polimerasa cuantitativa en tiempo real con transcripción inversa (RT-qPCR) se utiliza ampliamente en estudios de expresión génica, siendo esencial la selección adecuada de genes de referencia para garantizar la fiabilidad de los resultados. Este estudio analizó la literatura disponible con el fin de identificar los genes de referencia más comúnmente estudiados, seleccionados y validados en investigaciones de expresión génica en yuca (Manihot esculenta Crantz). A través de una revisión bibliográfica, se analizaron cinco publicaciones centradas en la selección y validación de genes de referencia en yuca. En total, se realizaron 55 evaluaciones de genes en los estudios analizados, considerando que algunos genes fueron investigados más de una vez. Los resultados indicaron que los genes PP2A, UBQ10, GTPb, EF1, TUB, TBP, NAC, HisD, 26S y ZFP se encuentran entre los más estables en los estudios seleccionados. Entre ellos, los genes PP2A y GTPb se destacaron como los más estables en tres y dos estudios, respectivamente. En contraste, genes ampliamente utilizados como ACT y GAPDH no fueron validados en ninguno de los estudios analizados al considerar todas las muestras de cada experimento. Los algoritmos geNorm y NormFinder fueron los más empleados para evaluar la estabilidad de los genes de referencia, mientras que RefFinder fue empleado en un solo estudio. Esta revisión resalta la importancia de la validación experimental específica para la yuca y sugiere que la adopción de múltiples herramientas de análisis puede mejorar la precisión y fiabilidad de los estudios de expresión génica.

PALABRAS CLAVE: *RT-qPCR, Normalización, Expresión Génica.*

INTRODUCTION

Quantitative real-time polymerase chain reaction with reverse transcription (RT-qPCR) is widely used to study gene expression in various organisms (ROCHA et al., 2015). Due to its high sensitivity, specificity, and reproducibility, this technique has been extensively employed in the field of molecular biology (WANG et al., 2024a; ZHANG et al., 2024). The accuracy and reliability of this method can be affected by several factors, including RNA quality, amplification efficiency, and sample-to-sample variation (SHI et al., 2024; ZHANG et al., 2024). In addition, data analysis can influence result interpretation, making normalization essential to compensate for uncontrollable experimental variations (COELHO et al., 2022).

Effective normalization relies on the selection of appropriate reference genes (DAUDE et al., 2024). These genes should exhibit moderate and stable expression levels across different developmental stages, experimental treatments, and organs, ensuring the precision and reliability of RT-qPCR results (GUO et al., 2024). Several reference genes, such as *Clathrin adaptor protein medium subunit* (*AP47*), *Ubiquitin* (*UBQ*) (FERNANDES-BRUM et al., 2017), *Ribosomal protein 24S* (*24S*), *Protein phosphatase 2A* (*PP2A*) (FREITAS et al., 2017), and *Tubulin* (*TUB*) (ZHANG et al., 2024), have been recommended for gene expression analysis in plants. However, even the most commonly used reference genes may vary depending on species and specific conditions.

In gene expression studies involving cassava (*Manihot esculenta* Crantz), RT-qPCR has proven essential for understanding gene regulation in response to various environmental and stress conditions, such as drought (HU et al., 2016; MORGANTE et al., 2020), cold (LI et al., 2020), salinity (SANTA BRÍGIDA et al., 2014), and pathogen infection (CALLEGARI et al., 2021; HERRERA et al., 2018). Several studies have investigated and validated reference genes in this crop, aiming to identify genes with stable expression and thus ensure the accuracy of RT-qPCR experiments (HU et al., 2016; MORENO et al., 2011; SALCEDO et al., 2014; SANTOS-SILVA et al., 2021; WHANKAEW et al., 2015). These studies typically analyze genes such as *Actin* (*ACT*), Serine-threonine phosphatase (*PP2A*), Elongation factor 1-alpha (*EF1α*), and Tubulin (*TUB*), among others, across different tissues and conditions. The validation of reference genes is crucial to ensure that the observed variations in target genes are associated with experimental treatments, rather than fluctuations in reference gene expression levels (COELHO et al., 2022).

In the present study, a literature review was performed to identify the most commonly used and validated reference genes in gene expression research on cassava using the RT-qPCR technique. We compiled a list of genes that can serve as a starting point for selecting reference genes in future gene expression studies for this crop.

MATERIAL AND METHODS

LITERATURE REVIEW

A systematic review was conducted using the Web of Science, Scopus and Google Scholar databases to identify studies comprizing the selection and validation of reference genes in *Manihot esculenta* for gene expression analyses via RT-qPCR. The search included all available studies in these databases, with no time restrictions, aiming to provide a comprehensive and up-to-date overview of the application of specific cassava reference genes, following the best methodological practices in the field. The keyword search were structured using Boolean operators to optimize data retrieval, and the following terms were applied: (i) reference genes; (ii) housekeeping genes; (iii) endogenous genes; (iv) constitutive genes; (v) Real-time PCR; (vi) RT-qPCR; (vii) *Manihot esculenta*.

EXPERIMENTAL CONDITIONS AND VALIDATED REFERENCE GENES

The most significant results were compiled to form a final list of tested and validated cassava reference genes, based on the selected studies. From each study, information regarding the selected genes, including gene name, description, identifier, experimental conditions, and the algorithms used for reference gene validation were gathered.

For the assessment of experimental conditions, factors such as the tissue type, plant developmental stage and stress conditions, were considered. These details were meticulously extracted to ensure that the context in which each gene was validated could be compared and properly categorized, providing a comprehensive overview of the adopted experimental conditions.

Regarding validation algorithms, specific tools and methodologies employed for reference gene stability analysis were evaluated, including the geNorm, NormFinder, BestKeeper, Delta-CT, and RefFinder algorithms. Each study was evaluated concerning the algorithm used and the stability criteria adopted, enabling an accurate comparison among the studies.

CITATION ANALYSIS OF STUDIES

In order to quantify the relevance of each study related to reference genes in cassava, citation data were collected for each identified article. Citations were retrieved from Google Scholar and Web of Science, which provide comprehensive coverage of scientific journals and literature.

The data collection was conducted on October 12, 2024, ensuring temporal uniformity across all analyzed citations. In Google Scholar, each article was searched using its full title to avoid ambiguity in the results and ensure accurate citation counts for the intended study. The same search strategy was applied in the Web of Science database to verify consistency in citation numbers between the two platforms.

Discrepancies in citation counts were recorded and compared to assess variations between sources and evaluate the stability of citation metrics, thereby providing an overview of the relevance and impact of the studies included in the review.

TRENDS IN GENE EXPRESSION STUDIES

To identify the most frequently used reference genes in gene expression studies, a new literature search of scientific articles was conducted. The adopted methodology included a detailed search on the Web of Science platform to locate recent studies addressing gene expression in cassava.

To maximize the comprehensiveness and accuracy of the results, the keywords “PCR”, “gene expression”, “Cassava”, and "*Manihot esculenta*" were used, combined with Boolean operators as follows: “PCR” AND “gene expression” AND (“Cassava” OR “*Manihot esculenta*”). The search was temporally restricted to include only publications from the years of 2023 and 2024, ensuring the analysis of the most recent studies on the subject.

The inclusion criteria focused on studies addressing the gene expression of target genes, aiming to identify which reference genes were used for RT-qPCR data normalization in these studies. This step was essential for understanding recent trends in the selection of reference genes for gene expression studies in this crop.

The selected articles were organized and documented for subsequent comparative analysis, providing a solid basis for discussing the evolution of normalization practices in gene expression experiments involving cassava.

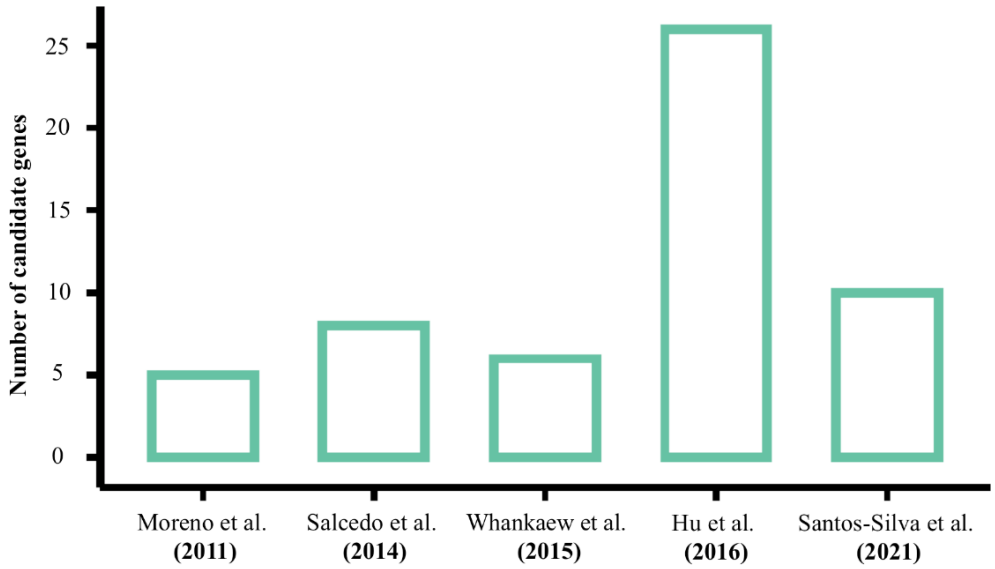
RESULTS AND DISCUSSION

LITERATURE SURVEY

A total of five publications related to the selection and validation of reference genes in cassava were identified after screening the articles retrieved from the three databases used (*Web of Science*, *Scopus*, and *Google Scholar*). The first study, published by Moreno et al. (2011), marked an initial milestone in the identification of reference genes for this crop.

In comparison with other agriculturally important species, such as coffee, there were already three publications on reference gene selection by 2011 (FERNANDES-BRUM et al., 2017), highlighting that gene expression studies in cassava are relatively recent. The data presented in Figure 1 show the number of candidate reference genes investigated in each publication, revealing considerable variation in the scope of the studies. The study developed by Hu et al. (2016) stands out for evaluating 26 genes, demonstrating a broader and more systematic approach for identifying suitable reference genes. In contrast, the pioneering study by Moreno et al. (2011) assessed only five genes, reflecting the initial limitations of research in this species.

Figure 1. Candidate reference genes evaluated per publication in cassava studies.



Other studies evaluated from 6 to 10 genes, adopting an intermediate approach (SALCEDO et al., 2014; SANTOS-SILVA et al., 2021; WHANKAEW et al., 2015). These findings suggest that, as knowledge regarding reference genes in cassava

advanced, the number of candidate genes tested increased, likely contributing to a more accurate and reliable selection of internal controls for RT-qPCR gene expression studies.

The variation in the number of genes investigated reflects the evolution of research and the increasing sophistication of the methodologies over time. The study by Hu et al. (2016), which tested a larger number of genes, highlights a significant effort to ensure accuracy in reference gene selection.

Differences in the number of genes analyzed can also be attributed to the adaptation of studies to specific experimental demands, such as environmental conditions, tissue types analyzed and physiological responses of interest.

The increase in the number of genes tested in more recent studies indicates a more rigorous and robust selection of reference genes, leading to greater accuracy in RT-qPCR studies. This is consistent with the MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) guidelines (BUSTIN et al., 2009, 2025), which emphasize the importance of proper validation of reference genes to ensure the accuracy and reproducibility of data.

EXPERIMENTAL CONDITIONS AND RECOMMENDED REFERENCE GENES

The results presented in Table 1 highlight the diversity of tissues analyzed, experimental conditions applied and reference genes validated in cassava. The study conducted by Moreno et al. (2011) investigated cassava plants infected and non-infected by Potyvirus, providing initial insights into the stability of reference genes under biotic stress, which is a relevant condition in plant research. In contrast, Salcedo et al. (2014) and Whankaew et al. (2015) examined gene expression throughout the plant growth cycle in the absence of adverse conditions, thereby exploring gene expression patterns across different physiological stages. This approach contributes to the standardization of developmental studies.

The inclusion of abiotic (HU et al., 2016), biotic and mechanical stresses (SANTOS-SILVA et al., 2021) represents progress in understanding gene stability under multiple environmental and physical challenges. This is crucial for the proper validation of reference genes under conditions that simulate field-related stresses.

The variability in the conditions present among the different studies highlights the necessity of selecting reference genes with demonstrated stability across diverse

scenarios, ensuring that RT-qPCR results remain accurate and relevant to the wide range of conditions encountered in cassava research.

Table 1. Tissues, experimental conditions and validated reference genes in cassava studies, considering all samples from each study.

Reference	Tissues	Conditions/Treatments	Validated genes
MORENO *et al.* (2011)	Leaves, stems, and roots	Plants infected and non-infected by Potyvirus from cassava varieties	*PP2A*, *UBQ10* and *GTPb*
SALCEDO *et al.* (2014)	Leaves, stems, and roots	150, 210, 270 and 330 days after planting	*EF1* and *TUB*
WHANKAEW *et al.* (2015)	Leaves and roots	Two varieties at different developmental stages. 6, 9 and 12 months after planting	*TBP*
HU *et al.* (2016)	Leaves, stems, and roots	Varieties subjected to different environmental conditions (drought stress)	*NAC* and *HisD* (geNorm*), *26S*, *PP2A*, and *ZFP* (NormFinder*)
SANTOS-SILVA *et al.* (2021)	Leaves, stems, and roots	Biotic (*Macrophomina pseudophaseolina*) and abiotic stress (mechanical injury)	*GTPb* and *PP2A*

* Algorithm used to select the most stable genes

The diversity of experimental conditions applied in reference gene validation studies in cassava highlights how gene stability responses can be highly context-dependent. In light of this, it is suggested that future studies focus on tissues not yet investigated, such as embryogenic and non-embryogenic calluses, since development and differentiation in these tissues may generate distinct gene expression patterns.

Analyzing the stability of reference genes in embryogenic tissues will also allow the identification of robust candidates genes for expression studies, especially in tissue culture experiments. This could significantly contribute to advances in *in vitro* culture research, including areas such as regeneration and genetic improvement through genome editing.

Regarding the validated genes (Table 1), the results indicated that the following genes are among the most stable genes in the selected studies: *Serine-threonine phosphatase* (*PP2A*), *Ubiquitin 10* (*UBQ10*), *GTP binding* (*GTPb*), *Elongation factor 1-alpha* (*EF1*), *Tubulin 1 β* (*TUB*), *TATA box binding protein* (*TBP*), *Nascent polypeptide-associated complex subunit beta* (*NAC*), *Histidinol dehydrogenase* (*HisD*), *26S proteasome regulatory complex, subunit N10* (*26S*) and *Zinc finger C-x8-C-x5-C-x3-H type family protein* (*ZFP*). Among them, *PP2A* and *GTPb* stood out, since *PP2A* is among the most stable genes in three different studies (HU et al., 2016; MORENO et al., 2011; SANTOS-SILVA et al., 2021) , while *GTPb* was found to be among the most stable genes in two of them (MORENO et al., 2011; SANTOS-SILVA et al., 2021). This suggests that these genes exhibit high stability levels across the different conditions and/or treatments studied.

However, genes widely used in other organisms, such as *Actin* (*ACT*) and *Glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*), were not among the most stable genes when analyzing all samples from each experiment. This reinforces the need for experimental validation for each species.

The frequent use of *GAPDH* and *ACT* as reference genes is linked to the fact that they were widely used as standards during the peak of the Northern blotting technique in the 1980s and early 1990s (BUSTIN; NOLAN, 2004; CHAPMAN; WALDENSTRÖM, 2015). These genes are known to be activated by various biological factors. Furthermore, both contain many pseudogenes in several species, which can introduce errors in RT-qPCR experiments, especially if primers are not carefully evaluated to prevent the amplification of these pseudogenes in species where they are present (SUN et al., 2012).

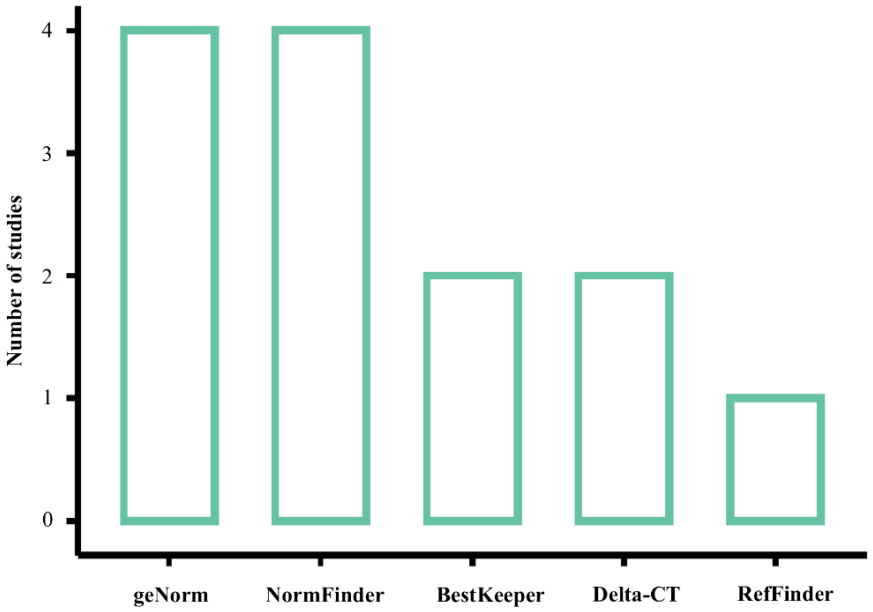
These results highlight the importance of carefully selecting reference genes specific to each experimental condition and indicate that genes commonly used in other plants may not be suitable for studies in cassava.

ALGORITHMS USED

The appropriate choice of tools for evaluating the stability of reference genes in RT-qPCR studies is not only a matter of preference, but also depends on the experimental context and the specific characteristics of the genes being studied. Among the main algorithms used for this purpose are geNorm, which calculates the geometric mean of gene expression to identify the most stable genes through an M value that decreases as stability increases (VANDESOMPELE et al., 2002); NormFinder, which uses a variance model to estimate stability more robustly, considering intragroup and intergroup variations (ANDERSEN et al., 2004); BestKeeper, which is based on correlations between Cq values to determine the most consistent genes (PFAFFL et al., 2004); and Delta-CT, which directly compares the Cq values of gene pairs across different samples to assess stability (SILVER et al., 2006). These algorithms, when used in combination, as observed in the RefFinder tool (XIE et al., 2012; XIE et al., 2023), allow for a more comprehensive and reliable analysis by integrating complementary approaches to ensure the selection of the most stable reference genes. Additionally, the combined use of these tools, as proposed by RefFinder, can provide a more robust and reliable analysis, integrating distinct methodologies to minimize biases and optimize the selection of the most stable genes. Therefore, the adoption of multiple analysis tools not only increases the reliability of the results but also provides greater flexibility in adapting the analyses to different experimental and biological conditions.

Figure 2 highlights the usage frequency of the main reference gene stability evaluation algorithms in the selected studies, highlighting the great use of geNorm and NormFinder, both mentioned in four articles. On the other hand, RefFinder was mentioned in only one of the five selected studies here. These data reflect the preference and reliability of these algorithms in the context of stability evaluation, as well as the common practice of combining approaches to ensure greater robustness in selecting the most stable genes.

Figure 2: Algorithms used in different studies to analyze the stability of candidate reference gene expression.

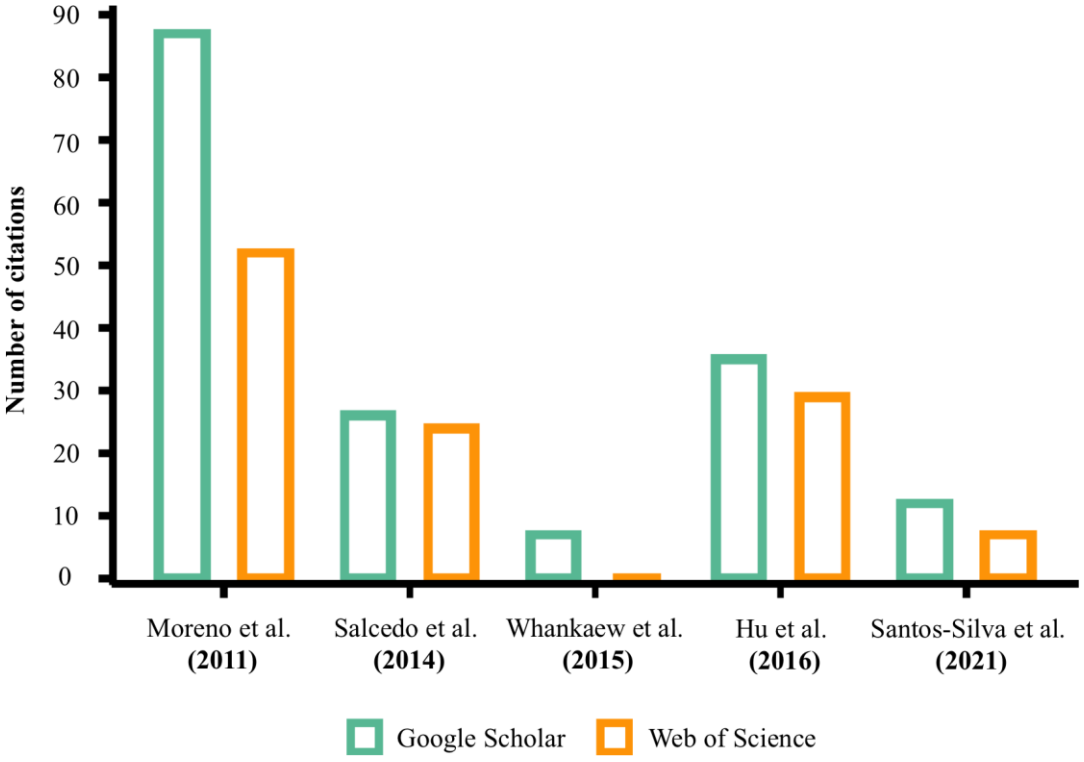


Currently, there is a database called RGeasy (Reference Genes Easy), which allows for both the deposition of reference gene experiment data, that is, Cq values, and the analysis of this data using any of the five algorithms mentioned through RefFinder. This analysis can be performed on combinations of conditions and/or treatments that have not been explored in publications, facilitating researchers' access to the best reference genes for their conditions and/or treatments of interest (DE SOUZA et al., 2024).

CITATIONS OF THE STUDIES

The number of citations of the articles on the selection and validation of reference genes in cassava (Figure 3) reflects the impact and relevance of each study in this field. The five evaluated articles have, to date*, received 167 and 112 citations in the Google Scholar and Web of Science databases, respectively, highlighting the significance of the conducted research.

Figure 3: Number of citations for each article found in the bibliographic search, according to the Google Scholar and Web of Science databases.



* Search performed on April 7, 2025.

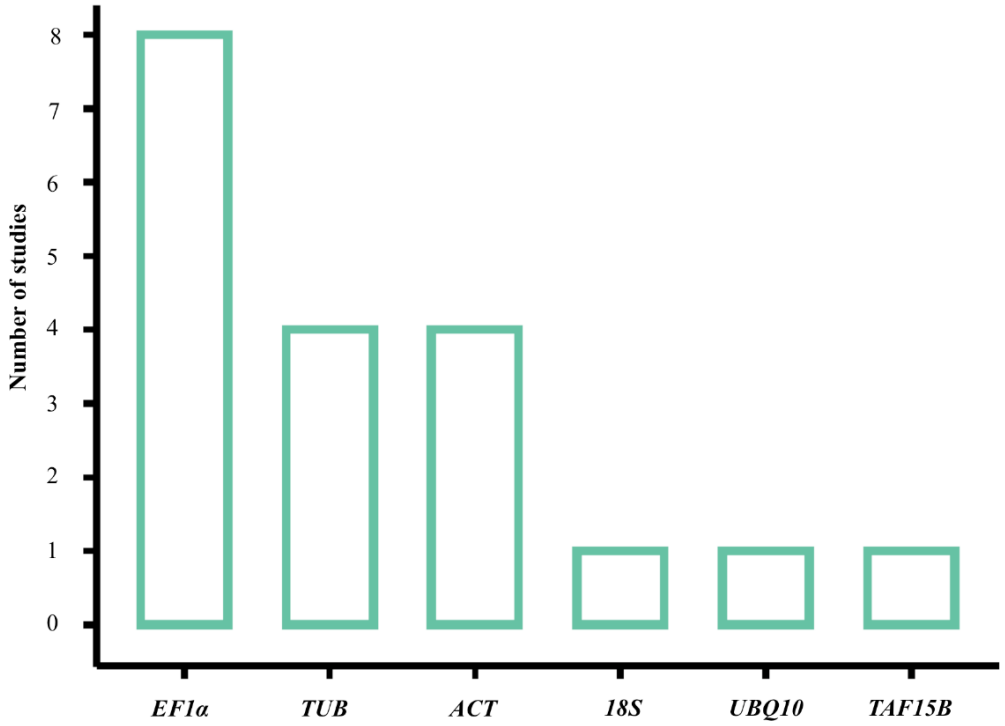
The study by Moreno et al. (2011), for instance, is the most cited one, with 87 and 52 citations in the Google Scholar and Web of Science databases, respectively. This study stands out as one of the pioneering investigations in the field. The article by Hu et al. (2016) is the second one in number of citations, with 35 and 29 citations in the same databases, highlighting its importance in the development of gene expression studies via RT-qPCR.

In contrast, the most recent study by Santos-Silva et al. (2021), with 12 (Google Scholar) and 7 (Web of Science) citations, is beginning to gain visibility in the field. These data suggest that interest in the validation of reference genes in cassava it is increasing, with these studies serving as a foundation for future works. In addition, the number of citations may also indicate the applicability of the analyzed reference genes in other crops and experimental contexts, broadening the perspectives for further investigations.

TRENDS IN GENE EXPRESSION STUDIES

Based on the bibliographic search conducted on the Web of Science platform, it was possible to identify the most frequently used reference genes in recent gene expression studies in cassava. Figure 4 presents the number of studies that employed each analyzed reference gene.

Figure 4: Reference genes used in gene expression studies in cassava in the years of 2023 and 2024. Genes: *Elongation Factor 1-alpha* (*EF1α*), *Tubulin* (*TUB*), *Actin* (*ACT*), *18S RNA ribosomal* (18S), *Ubiquitin 10* (*UBQ10*) and *TATA-binding protein-associated factor 15* (*TAF15b*).



The *EF1α* gene (*Elongation Factor 1-alpha*) stood out as the most used reference genes, being cited in 8 studies. This result reinforces the relevance of *EF1α* as one of the most robust reference gene in RT-qPCR studies due to its stable expression under different experimental conditions (SALCEDO et al., 2014). However, considering the results from the first bibliographic search, *EF1α* was tested in three studies (SALCEDO et al., 2014; SANTOS-SILVA et al., 2021; WHANKAEW et al., 2015), being validated in only one study (SALCEDO et al., 2014). This predominance reflects the search for genes that ensure the accuracy and reproducibility of gene

expression data, as recommended by the MIQE guidelines (BUSTIN et al., 2009, 2025). It is important to highlight that the MIQE guidelines also suggest the use of two or more reference genes for data normalization, in order to increase the reliability of the results (BUSTIN et al., 2009, 2025). This was observed in only 18.75% of the studies.

Following *EF1α*, *TUB* and *ACT* were the second most frequent used reference genes, being used in 4 studies each. *ACT*, in particular, is commonly used as a reference gene in gene expression studies. However, this study shows that although *ACT* was frequently tested as a candidate reference gene, it was not validated in any recent study when all the samples of each experiment were taken into account. This suggests that, despite being commonly used, the stability of this gene may be questioned depending on the experimental conditions, reinforcing the need for specific validation for each context.

CONCLUSION

The five cassava reference gene studies analyzed here enabled the observation that the genes *PP2A*, *UBQ10*, *GTPb*, *EF1*, *TUB*, *TBP*, *NAC*, *HisD*, *26S*, and *ZFP* are among the most stable ones when analyzing all samples from each study. Among them, *PP2A* and *GTPb* stood out, being among the most stable in three and two studies, respectively. In contrast, widely used reference genes, such as *ACT* and *GAPDH*, were not validated in any of the studies analyzed here when all the samples from each study were considered. The use of multiple algorithms enhances the reliability of the results by integrating different validation approaches. These methodological advances are essential for ensuring that future gene expression studies in cassava are based on more robust and reliable data.

Acknowledgments

This work was financially supported by the “Fundação de Amparo à Pesquisa do Estado do Tocantins” (FAPT), the “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior” (CAPES), the “Conselho Nacional de Desenvolvimento Científico e Tecnológico” (CNPq), the “Rede de Biodiversidade e Biotecnologia da Amazônia Legal” (Bionorte), the “Programa de Pós-Graduação em Agroenergia Digital” and the Universidade Federal do Tocantins (PROPESQ-UFT).

References

ANDERSEN, C. L.; JENSEN, J. L.; ØRNTOFT, T. F. Normalization of Real-Time Quantitative Reverse Transcription-PCR Data: A Model-Based Variance Estimation Approach to Identify Genes Suited for Normalization, Applied to Bladder and Colon Cancer Data Sets. **Cancer Research**, v. 64, n. 15, p. 5245–5250, 2004.

BUSTIN, S. A. *et al.* The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. **Clinical Chemistry**, v. 55, n. 4, p. 611–622, 2009.

BUSTIN, S. A. *et al.* MIQE 2.0: Revision of the Minimum Information for Publication of Quantitative Real-Time PCR Experiments Guidelines. **Clinical Chemistry**, 24 abr. 2025.

BUSTIN, S.; NOLAN, T. Pitfalls of Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction. **Journal of biomolecular techniques : JBT**, v. 15, p. 155–166, 2004.

CALLEGARI, D. M.; LIMA, A. M.; FERREIRA BARROS, N. L.; SIQUEIRA, A. S.; MOURA, E. F.; BATISTA DE SOUZA, C. R. Changes in transcript levels of cassava superoxide dismutase and catalase during interaction with *Phytopythium* sp. **Physiological and Molecular Plant Pathology**, v. 114, 2021.

CHAPMAN, J. R.; WALDENSTRÖM, J. With reference to reference genes: A systematic review of endogenous controls in gene expression studies. **PLoS ONE**, v. 10, n. 11, 2015.

COELHO, T. C.; CHALFUN-JUNIOR, A.; BARRETO, H. G.; SOUZA DUARTE, M. DE; OLIVEIRA GARCIA, B. DE; TEIXEIRA, P. D.; GIONBELLI, T. R. S.; LADEIRA, M. M. Reference gene selection for quantitative PCR in liver, skeletal muscle, and jejunum of *Bos indicus* cattle. **Revista Brasileira de Zootecnia**, v. 51, 2022.

DAUDE, M. M.; SÁGIO, S. A.; RODRIGUES, J. N.; LIMA, N. M. P.; LIMA, A. A.; SARMENTO, M. I.; SARMENTO, R. A.; BARRETO, H. G. Reference genes for Eucalyptus spp. under *Beauveria bassiana* inoculation and subsequently infestation by the galling wasp *Leptocybe invasa*. **Scientific Reports**, v. 14, n. 1, 2024.

FERNANDES-BRUM, C. N.; GARCIA, B. DE O.; MOREIRA, R. O.; SÁGIO, S. A.; BARRETO, H. G.; LIMA, A. A.; FREITAS, N. C.; LIMA, R. R. DE; CARVALHO, C. H. S. DE; CHALFUN-JÚNIOR, A. A panel of the most suitable reference genes for RT-qPCR expression studies of coffee: screening their stability under different conditions. **Tree Genetics and Genomes**, v. 13, n. 6, 2017.

FREITAS, N. C.; BARRETO, H. G.; FERNANDES-BRUM, C. N.; MOREIRA, R. O.; CHALFUN-JUNIOR, A.; PAIVA, L. V. Validation of reference genes for qPCR analysis of *Coffea arabica* L. somatic embryogenesis-related tissues. **Plant Cell, Tissue and Organ Culture**, v. 128, n. 3, p. 663–678, 2017.

GUO, W.; YANG, Y.; MA, B.; WANG, W.; HU, Z.; LENG, P. Selection and Validation of Reference Genes for Gene Expression Studies in *Euonymus japonicus* Based on RNA Sequencing. **Genes**, v. 15, n. 1, 2024.

HERRERA, M.; PORTILLO, D.; PULIDO, M. A.; DÍAZ TATIS, P. A.; LÓPEZ CARRASCAL, C. E. Estudio de la Expresión de Genes que Codifican para Putativas Proteínas PR en Yuca (*Manihot esculenta* Crantz). **Acta Biológica Colombiana**, v. 23, p. 242–252, 2018.

HU, M.; HU, W.; XIA, Z.; ZHOU, X.; WANG, W. Validation of reference genes for relative quantitative gene expression studies in cassava (*Manihot esculenta* Crantz) by using quantitative real-time PCR. **Frontiers in Plant Science**, v. 7, n. MAY2016, 2016.

LI, S.; CHENG, Z.; PENG, M. Genome-wide identification of miRNAs targets involved in cold response in cassava. **Plant OMICS**, v. 13, n. 1, p. 57–64, 2020.

MORENO, I.; GRUISSEM, W.; VANDERSCHUREN, H. Reference genes for reliable potyvirus quantitation in cassava and analysis of Cassava brown streak virus load in host varieties. **Journal of Virological Methods**, v. 177, n. 1, p. 49–54, 2011.

MORGANTE, C. V.; NUNES, S. L. P.; CHAVES, A. R. DE M.; FERREIRA, C. F.; AIDAR, S. DE T.; VITOR, A. B.; OLIVEIRA, E. J. DE. Genetic and physiological analysis of early drought response in *Manihot esculenta* and its wild relative. **Acta Physiologiae Plantarum**, v. 42, n. 2, p. 22, 11 fev. 2020.

PFAFFL, M. W.; TICHOPAD, A.; PRGOMET, C.; NEUVIANS, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample

integrity: BestKeeper – Excel-based tool using pair-wise correlations. **Biotechnology Letters**, v. 26, n. 6, p. 509–515, 2004.

ROCHA, D. J. P.; SANTOS, C. S.; PACHECO, L. G. C. Bacterial reference genes for gene expression studies by RT-qPCR: survey and analysis. **Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology**, v. 108, n. 3, p. 685–693, 2015.

SALCEDO, A.; ZAMBRANA, C.; SIRITUNGA, D. Comparative Expression Analysis of Reference Genes in Field-Grown Cassava. **Tropical Plant Biology**, v. 7, n. 2, p. 53–64, 2014.

SANTA BRÍGIDA, A. B.; REIS, S. P. DOS; NAZARÉ MONTEIRO COSTA, C. DE; CARDOSO, C. M. Y.; LIMA, A. M.; SOUZA, C. R. B. DE. Molecular cloning and characterization of a cassava translationally controlled tumor protein gene potentially related to salt stress response. **Molecular Biology Reports**, v. 41, n. 3, p. 1787–1797, 12 mar. 2014.

SANTOS-SILVA, C. A. DOS; VILELA, L. M. B.; OLIVEIRA-SILVA, R. L. DE; SILVA, J. B. DA; MACHADO, A. R.; BEZERRA-NETO, J. P.; CROVELLA, S.; BENKO-ISEPPON, A. M. Cassava (*Manihot esculenta*) defensins: Prospection, structural analysis and tissue-specific expression under biotic/abiotic stresses. **Biochimie**, v. 186, p. 1–12, 2021.

SHI, L.; CAI, Y.; YAO, J.; ZHANG, Q.; HE, B.; LIN, S. Reference Genes Selection and Validation for *Cinnamomum burmanni* by Real-Time Quantitative Polymerase Chain Reaction. **International Journal of Molecular Sciences**, v. 25, n. 6, 2024.

SILVER, N.; BEST, S.; JIANG, J.; THEIN, S. L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. **BMC Molecular Biology**, v. 7, 2006.

SOUZA, M. R. DE; ARAÚJO, I. P.; COSTA ARRUDA, W. DA; LIMA, A. A.; SÁGIO, S. A.; CHALFUN-JUNIOR, A.; BARRETO, H. G. RGeasy: a reference gene analysis tool for gene expression studies via RT-qPCR. **BMC genomics**, v. 25, n. 1, p. 907, 2024.

SUN, Y.; LI, Y.; LUO, D.; LIAO, D. J. Pseudogenes as weaknesses of ACTB (Actb) and GAPDH (Gapdh) used as reference genes in reverse transcription and polymerase chain reactions. **PLoS ONE**, v. 7, n. 8, 2012.

VANDESOMPELE, J.; PRETER, K. DE; PATTYN, F.; POPPE, B.; ROY, N. VAN; PAEPE, A. DE; SPELEMAN, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome Biology**, v. 3, n. 7, 2002.

WANG, F.; LI, P.; LIU, Q.; NIE, G.; ZHU, Y.; ZHANG, X. Selection and Validation of Reference Genes in Sudan Grass (*Sorghum sudanense* (Piper) Stapf) under Various Abiotic Stresses by qRT-PCR. **Genes**, v. 15, n. 2, 2024.

WHANKAEW, S.; SRAPHET, S.; THAIKERT, R.; BOONSENG, O.; SMITH, D. R.; TRIWITAYAKORN, K. Validation of a reference gene for transcript analysis in cassava (*Manihot esculenta* Crantz) and its application in analysis of linamarase and α-hydroxynitrile lyase expression at different growth stages. **African Journal of Biotechnology**, v. 14, n. 9, p. 745–751, 2015.

XIE, F.; WANG, J.; ZHANG, B. RefFinder: a web-based tool for comprehensively analyzing and identifying reference genes. **Functional & Integrative Genomics**, v. 23, n. 2, p. 125, 2023.

XIE, F.; XIAO, P.; CHEN, D.; XU, L.; ZHANG, B. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. **Plant Molecular Biology**, v. 80, n. 1, p. 75–84, 2012.

ZHANG, P.; CHEN, SHUYING; CHEN, SIYU; ZHU, Y.; LIN, Y.; XU, X.; LIU, Z.; ZOU, S. Selection and Validation of qRT-PCR Internal Reference Genes to Study Flower Color Formation in *Camellia impressinervis*. **International Journal of Molecular Sciences**, v. 25, n. 5, 2024.