IN SILICO MUTATION PROFILING OF *RBCL* GENE AND CLASSIFICATION OF SELECTED EDIBLE MEMBERS OF THE CUCURBITACEAE FAMILY

OLORUNSOLA, HONEY AANU
(20PCO02179)

JULY, 2022

IN SILICO MUTATION PROFILING OF *RBCL* GENE AND CLASSIFICATION OF SELECTED EDIBLE MEMBERS OF THE CUCURBITACEAE FAMILY

BY

OLORUNSOLA, HONEY AANU
(20PCO02179)
B.Sc. Plant Science and Biotechnology, Ekiti State University, Ado-Ekiti

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE (M.Sc.) DEGREE IN APPLIED BIOLOGY AND BIOTECHNOLOGY IN THE DEPARTMENT OF BIOLOGICAL SCIENCES, COLLEGE OF SCIENCE AND TECHNOLOGY, COVENANT UNIVERSITY

JULY, 2022

ACCEPTANCE

This is to attest that this dissertation is accepted in partial fulfilment of the requirements for the award of the degree of Master of Science in Applied Biology and Biotechnology in the Department of Biological Sciences, College of Science and Technology, Covenant University, Ota, Nigeria.

Mr. Taiwo B. Erewunmi
(Secretary, School of Postgraduate Studies) **Signature and Date**

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies) **Signature and Date**

DECLARATION

I, **OLORUNSOLA, HONEY AANU (20PCO02179)** declare that this research was carried out by me under the supervision of Prof. Olawole O. Obembe of the Department of Biological Sciences, College of Science and Technology, Covenant University, Ota, Nigeria. I attest that the dissertation has not been presented either wholly or partially for the award of any degree elsewhere. All sources of data and scholarly information used in this dissertation are duly acknowledged.

OLORUNSOLA, HONEY AANU **Signature and Date**

CERTIFICATION

We certify that this dissertation titled "**IN SILICO MUTATION PROFILING OF *RBCL* GENE AND CLASSIFICATION OF SELECTED EDIBLE MEMBERS OF THE CUCURBITACEAE FAMILY"** is an original research work carried out by **HONEY, AANU OLORUNSOLA (20PCO02179)** in the Department of Biological Sciences, College of Science and Technology, Covenant University, Ota, Ogun State, Nigeria under the supervision of Prof. Olawole O. Obembe. We have examined and found this work acceptable as part of the requirements for the award of Master of Science in Applied Biology and Biotechnology.

Prof. Olawole O. Obembe
(Supervisor) **Signature and Date**

Prof. Solomon U. Oranusi
(Head of Department) **Signature and Date**

Prof. Abiodun E. Ayodele
(External Examiner) **Signature and Date**

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies) **Signature and Date**

DEDICATION

I dedicate this work to the Almighty God; my help in ages past and to my parents.

ACKNOWLEDGMENTS

All thanks and praises be to God Almighty; the giver of life with whom I made it this far, His name alone is highly exalted.

My appreciation goes to the Chancellor Bishop David Oyedepo, the current Vice-Chancellor, Professor Abiodun H. Adebayo, the Acting Registrar, Mr Emmanuel Igban, and the entire management and staff of Covenant University for providing me the facilities for the successful completion of my program.

To the current Dean of the School of Postgraduate Studies, Professor Akan Williams, and the Postgraduate School staff, I appreciate you for the diverse trainings and programmes organised towards capacity building for successful postgraduate research. And, a very big thank you goes to the Head of the Department of Biological Sciences, Professor Solomon U. Oranusi and the staff for the constant help.

My most profound gratitude and heartfelt appreciation goes to my supervisor, Prof. Olawole Obembe for his guidance and patience throughout the course of this research. I am forever grateful to you, Sir.

I want to acknowledge Dr. Opeyemi Soremekun, Dr. Isewon Itunu, Dr. Oduselu Gbolahan, Dr. Jacob Popoola and Mr. Rotimi Aworunse for their guidance, constant help and support during this period.

I appreciate my parents, Mr. Gbenga Olorunsola and Mrs. Modupe Olorunsola for their unrelenting support, prayers and encouragement at all times.

I want to appreciate all my friends and loved ones whose love, help, input and support during this time was unwavering. I am grateful to you all, thank you.

Finally, I say a big thank you to all members of the Department of Biological Sciences for their support and tutelage and also special thanks go to my colleagues and senior colleagues, God bless you all.

TABLE OF CONTENTS

LIST OF APPENDIX

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The Cucurbitaceae, with about 965 genera in 95 species is ranked next to the Solaneceae; the largest family of fruits and vegetables. Its members are among the world's most significant plant foods. As at 2022, the whole genome sequence of 18 species of this family have been published. Ribulose biphosphate carboxylase large (*rbcL)* is central to the global carbon cycle and is one of the single barcode genes of individual plants that have phylogenetic utility. However, there has not been a documented detailed investigation into the *rbcL* gene mutation of this family. This research aimed at providing information on Single amino acid substitution (SAAS) of the *rbcL* gene of selected cucurbits and gaining insight into their classification. The studied populations are entries of *Citrullus lanatus*, *Citrullus colocynthis*, *Cucurbita pepo*, *Cucurbita maxima, Cucurbita moschata,* *Cucumis sativus*, *Cucumis melo, Telfaria occidentalis*, *Momordica charantia, Lagenaria siceraria*, and *Coccinia grandis.* Five amino acid sequence entries of these members (including their reference sequence) were retrieved from the RefSeq database of the National Centre for Biotechnology Information. The sequences of each species were compared with their reference sequence using MUSCLE algorithm of the MEGA11 to check for the presence of SAASs in the gene. The type and effect of mutation on protein stability were predicted using the PPVED v1.0 and I-MUTANT v.3.0 online server, respectively. The *rbcL* genomic DNA isolation, amplification, purification, and sequencing of *Telfairia occidentalis* were carried out. Multiple sequence alignment of the translated amino acid sequence alongside the references of the other plant species was carried out using MUSCLE algorithm of MEGA11. A phylogenetic tree was constructed using Maximum-likelihood method. The gene sequences of six species were conserved while four species had the presence of SAASs. F108A was recorded in *Cucumis sativus,* D10X and M228I were recorded in *Citrullus lanatus*, T36- and A343X were recorded in *Lagenaria siceraria*, and I44X and Q375X were recorded in *Cucumis melo*. A343X and F108A were predicted to be functional mutations with a decrease in the stability of the protein binding site. T36- is well established to be a deleterious mutation. Only *Citrullus lanatus* had non- functional mutations. In total, 45 entries of 10 species were screened and only five entries of four species had mutations (10.9%). The phylogenetic tree revealed that *T. occidentalis* is positioned within the Cucurbitaceae family, although, it had the lowest bootstrap confidence compared to other members. To authenticate this positioning, *rbcL* sequences of ten species belonging to four closely related families within the Order Cucurbitales were added to those of the Cucurbitaceae, and a new tree was constructed. *Telfairia occidentalis* was centrally positioned between the Cucurbitaceae and the close relatives. The newly constructed tree also revealed that *T. occidentalis* was not only the closest to the root of the tree, but it was also closer to the relatives than to Cucurbitaceae. Conclusively, this study validated the conservation of *rbcL* gene sequence amongst the populations screened. It also demonstrated that *T. occidentalis* classification within the Cucurbitaceae cannot be ascertained but should be further investigated using other molecular markers.

Keywords: RbcL, Phylogeny, Cucurbitaceae, Mutation, Telfairia occidentalis.