MUTATION PROFILING OF *RBCL* AND *MATK* GENES AND PHYLOGENETIC RELATIONSHIPS OF SELECTED MEMBERS OF THE FABACEAE FAMILY

OBADIRE, TOBI TEJUMADE
(22PCO02402)
B. Tech, Agricultural Technology, Federal University of Technology Akure,
Ondo State, Nigeria

AUGUST, 2024

MUTATION PROFILING OF *RBCL* AND *MATK* GENES AND PHYLOGENETIC RELATIONSHIPS OF SELECTED MEMBERS OF THE FABACEAE FAMILY

BY

OBADIRE, TOBI TEJUMADE
(22PCO02402)
B. Tech, Agricultural Technology, Federal University of Technology Akure,
Ondo State, Nigeria

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 BIOLOGY IN THE DEPARTMENT OF BIOLOGICAL SCIENCES, COLLEGE OF SCIENCE AND TECHNOLOGY, COVENANT UNIVERSITY, OTA, OGUN STATE, NIGERIA

AUGUST, 2024

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 Biology in the Department of Biological Sciences, College of Science and Technology, Covenant University, Ota, Nigeria.

Miss Adefunke F. Oyinloye
(Secretary, School of Postgraduate Studies) **Signature and Date**

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

DECLARATION

I, **OBADIRE, TOBI TEJUMADE (22PCO02402),** declare that I carried out this research 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 wholly or partially for the award of any degree elsewhere. All sources of data and scholarly information used in this dissertation are duly acknowledged.

OBADIRE, TOBI TEJUMADE

Signature and Date

CERTIFICATION

We certify that this dissertation titled "**MUTATION PROFILING OF RBCL AND MATK GENES AND PHYLOGENETIC ANALYSIS OF SELECTED MEMBERS OF THE FABACEAE FAMILY"** is an original research work carried out by **OBADIRE, TOBI TEJUMADE (22PCO02024)** 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 Biology (BIOTECHNOLOGY AND BIOINFORMATICS).

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

Dr. Opeyemi I. Ayanda
(Head of Department) **Signature and Date**

Prof. Adegoke E. Adegbite
(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, Mrs Tobi-David and the entire management and staff of Covenant University for providing me with 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 your diverse training and programs organized towards capacity building for successful postgraduate research. Thank you to the Head of the Department of Biological Sciences, Dr Opeyemi I. Ayanda, and the department staff for their constant help.

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

I want to acknowledge Dr. Adetutu Bello and Dr. Rotimi Aworunse for their guidance, constant help, and support during this period.

I appreciate my parents, Mr and Mrs. Obadire, for their unrelenting support, prayers and constant encouragement. I also appreciate my uncles, the Oladoyins, my aunty, Mrs Elufisan, and my grandma, Mrs Dorcas Oladoyin, for their financial support in attaining this degree. I want to thank all my friends and loved ones whose love, help, input and support during this time were unwavering. I am grateful to you all; thank you.

Finally, I thank 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

CONTENTS **PAGES**

LIST OF APPENDIXES

APPENDIXES **LIST OFAPPENDIXES** **PAGES**

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

ABSTRACT

Gene mutation profiling involves detecting the type and effect of mutation on a gene. The rbcL and matK genes are chloroplast genes involved in photosynthesis; mutations on these genes can be deleterious or beneficial to the plant. Integrating underutilized legumes into genetic studies alongside major legumes using barcoding methodologies is essential for their genetic improvement. This research aimed to provide detailed mutation information on the rbcL and matK genes of the selected legumes and gain insight into their classification based on DNA barcodes. The sequences of sixteen legumes were used in the study. Five amino acid sequence entries from these species, reference sequences and complete chloroplast genomes were retrieved from the National Center for Biotechnology Information database. The gene sequences were aligned with their reference sequences using MULTALIN to identify amino acid substitutions in the genes. The type and effect of these mutations on protein function and stability were predicted using the SIFT, PROVEAN, and I-MUTANT v.2.0 online servers. The rbcL and MatK genes of *Macrotyloma geocarpum* were isolated, amplified, and sequenced. Consensus sequences of the sequence entries from 15 species were generated and aligned with those sequenced using the Muscle algorithm. Phylogenetic analysis was conducted using the neighbor-joining algorithm in MEGA 11 software. Results revealed 32 and 22 amino acid substitutions in the matK and rbcL genes, respectively. Five substitutions in the matK gene across three species were predicted to be deleterious by all three predictive software, and one substitution in the rbcL gene was predicted to be beneficial. The double barcodes (matK+rbcL) and single barcode (rbcL) classified the selected legumes randomly, however, matK gene classified the legumes down to the genus and species levels. It also revealed that *Cajanus cajan*, an underutilized legume, is the closest to *Vigna unguiculata*, implying the possibility of genetically improving *C. cajan* using the *V. unguiculata* genome. The matK gene also gave similar result to that of chloroplast genome. Consequently, matK gene is recommended as a core barcode sequence in the Fabaceae family.

Keywords: Mutation, rbcL, matK, Underutilized legumes, Phylogenetic Analysis