BIOINFORMATIC ANALYSIS OF DEGRADOME DATA FROM AFRICAN OIL PALM (*ELAEIS GUINEENSIS JACQ*.) INFLORESCENCE

SRI DEVI A/P RAMARAO

FACULTY OF SCIENCE

UNIVERSITY OF MALAYA

KUALA LUMPUR

2012

BIOINFORMATIC ANALYSIS OF DEGRADOME DATA FROM AFRICAN OIL PALM (*ELAEIS GUINEENSIS JACQ*.) INFLORESCENCE

SRI DEVI A/P RAMARAO

(SGJ 100014)

SUBMITTED TO

INSTITUTE OF BIOLOGICAL SCIENCES

FACULTY OF SCIENCE

UNIVERSITY OF MALAYA

IN PARTIAL FULFILMENT

OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF BIOINFORMATICS

2012

ABSTRACT

African Oil Palm (*Elaeis guineensis Jacq.*) is an economically important plant since it is the highest yielding oil-producing crop in the world. Long reproductive cycle and slow seed maturation result in lengthy breeding cycles and a slow rate of improvement through selection. Demand to increase oil palm yield can be supported by analysis of oil palm sequence data towards a functional genomic approach to support marker assisted selection.

In this study degradome data from inflorescence of the African Oil Palm was analysed to identify miRNA/siRNA cleavage products and SNPs in miRNA targets. Degradome analysis tools CleaveLand, SeqTar and PARESnip were used to analyse degradome data towards the objective of identifying miRNA/target pairs that may be important in oil palm flowering and hence yield of this major crop.

Total miRNA targets predicted by CleaveLand, SeqTar, PARESnip were 1304, 663318 and 12532 respectively. SeqTar is suitable to be used to predict a wide range of targets without missing any possibilities of the predicted targets in oil palm degradome data. SeqTar is also suitable to predict precursor novel miRNA targets in oil palm degradome data. CleaveLand can only be used to predict known miRNA targets and it is not suitable to predict novel miRNA targets in oil palm degradome data.

Of the three currently available tools for degradome analysis, PARESnip was found to be the fastest and most convenient tool. PARESnip is user friendly and user configurable tool. PARESnip also shows the highest number of prediction of highly significant miRNA targets compared to CleaveLand and SeqTar. Thus, PARESnip can be considered as a more reliable tool to analyse degradome data from African Oil Palm inflorescence.

Computational prediction methods have been successfully employed to find candidate miRNA targets in African Oil Palm inflorescence data.

ABSTRAK

Kelapa Sawit Afrika (*Elaeis guineensis Jacq.)* adalah tanaman yang penting dari segi ekonomi kerana ia merupakan pengeluar minyak yang tertinggi di dunia. Kitar reproduktif yang lama dan maturasi benih yang perlahan mengakibatkan kitar pembiakan yang panjang dan kadar perlahan peningkatan melalui pilihan. Permintaan untuk meningkatkan hasil kelapa sawit boleh disokong oleh analisis data turutan minyak kelapa sawit ke arah pendekatan genomik berfungsi untuk menyokong seleksi dibantu marker.

Dalam kajian ini, data degradome dari pembungaan kelapa sawit Afrika telah dianalisa untuk mengenal pasti produk hasil belahan miRNA/siRNA dan SNPs dalam target miRNA. Program-program penganalisis degradome seperti CleaveLand, SeqTar dan PARESnip telah digunakan untuk menganalisa data degradome ke arah objektif mengenal pasti pasangan miRNA target yang mungkin penting dalam pembungaan kelapa sawit dan seterusnya meningkatkan hasil tanaman utama ini.

Jumlah pasangan miRNA target yang diramalkan oleh CleaveLand, SeqTar dan PARESnip adalah 1304, 663 318 dan 12 532 masing-masing. SeqTar sesuai diguna untuk meramal miRNA target dalam pelbagai jenis tanpa sebarang kemungkinan ketinggalan pasangan target. CleaveLand hanya boleh digunakan untuk meramalkan miRNA target orthologus tetapi tidak sesuai untuk miRNA target novel dalam data kelapa sawit.

Antara ketiga-tiga program ini yang diguna untuk menganalisa degradome, PAREsnip adalah paling pantas dan mudah untuk diguna. PARESnip merupakan mesra pengguna dan boleh dikonfigurasi. PARESnip juga menunjukkan bilangan tertinggi ramalan miRNA target yang amat ketara berbanding CleaveLand dan SeqTar. Oleh itu, PAREsnip boleh dianggap sebagai program yang lebih sesuai untuk menganalisa data degradome dari pembungaan kelapa sawit Afrika.

Kaedah ramalan komputer telah berjaya digunakan untuk mencari calon-calon miRNA target dalam data pembungaan kelapa sawit Afrika.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank the God to have given me the strength to pursue my ambition and fulfilling my promise to myself. I would like to take this opportunity to thank all who have given unconditional support to fulfil my ambition.

First of all thank you to my supervisor Professor Dr. Jennifer Ann Harikrishna for giving me the opportunity to do my study under her kind supervision, advice and guidance whenever I need it. Thank you for the support and encouragements to complete my thesis. I would like to thank my co-supervisor Dr. Khang Tsung Fei for his supervision and support to complete this thesis.

Not forgetting PhD students Ranga for his guidance on Linux has been a great help to complete this study. Thanks to Hui Li for her help in providing needed files to this study. Not forgetting all my classmates for encouragement, guidance and supports.

Last but not least, I greatly appreciate my parents, husband and family members for their unconditional love and freedom to give me free reigns in my studies.

TABLE OF CONTENTS

CHAPTER 2: LITERATURE REVIEW

CHAPTER 3: METHODOLOGY

CHAPTER 4: RESULTS

CHAPTER 5: DISCUSSION

LIST OF FIGURES

LIST OF TABLES

APPENDIX

APPENDIX A

APPENDIX B

APPENDIX C

APPENDIX D

APPENDIX E

APPENDIX F

APPENDIX G