COMPARATIVE TRANSCRIPTOME ANALYSIS OF RUBBER TREE (*Hevea brasiliensis*) BETWEEN HIGH AND LOW YIELD RUBBER CLONES RELATED TO NATURAL RUBBER BIOSYNTHESIS

MOHD FAHMI BIN ABU BAKAR

UNIVERSITI SAINS MALAYSIA

2020

COMPARATIVE TRANSCRIPTOME ANALYSIS OF RUBBER TREE (*Hevea brasiliensis*) BETWEEN HIGH AND LOW YIELD RUBBER CLONES RELATED TO NATURAL RUBBER BIOSYNTHESIS

by

MOHD FAHMI BIN ABU BAKAR

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

February 2020

ACKNOWLEDGEMENT

In the name of Allah s.w.t., the most gracious, the most merciful, praise is to Allah s.w.t., the peace and blessing is unto Muhammad s.a.w. I wish to express my sincere thanks to my supervisor, Professor Ahmad Sofiman Othman who constantly gave his intellectual guidance, endless support and encouragement at each obstacle faced during the duration of my research.

Many thanks also to Malaysian Rubber Board and Sime Darby Plantation Berhad for providing me the information and samples required for this study to be done. I would like to record my gratefulness to lab mates (Khairul Lukman, Dr. Farah Alia, Siti Nurfazilah and Nur Afieqah Kaz), Dr. Hafsah Jaafar and Dr. Mohd Noor Abdul Ghani (former officer of Malaysian Rubber Board and Sime Darby Plantation Berhad) for their assistance and always lend me a helping hand in giving idea and support throughout the study.

Besides, the blessing from both of my parents (Abu Bakar Yusoff and Rokiah Hashim) and my parent-in-law (Abd Kadir Osman and Smadiyah Ahmad), I would like to extend my sincere thanks and appreciation to my dearest wife Siti Nurkhalida Abd Kadir whose love, encouragement and support, have given me the peace of mind to complete this research with success. I would also like to express my thankfulness and appreciation to my family (Mohd Faizal, Nor Azliza, Nor Syuhada, Muhammad Firdaus and Ummi Sakinah) for their support, care and concern during the project.

My candidature as PhD student in USM was supported by SLAB (Skim Latihan Akademik Bumiputera) from Ministry of Education of Malaysia (MOE) while this project was supported by grant provided from Rubber Industry Smallholder Development Authority (RISDA).

Alhamdulillah

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

<	Less than
%	Percentage
°C	Degree Celcius
μg	Microgram
μL	Microliter
mL	Mililiters
mM	Milimolar
rpm	Rotation per minute
bp	Basepair
DNA	Deoxyribonucleic acid
BLAST	Basic Local Alignment Search Tool
CTAB	Hexadecyl-trimethylammonium bromide
EDTA	ethylenediaminetetraacetic acid
MRB	Malaysian Rubber Board
IPP	Isopentenyl Diphosophate
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
MVA	Mevalonate
MEP	2-C-methyl-D-erythritol 4-phosphate
NR	Natural Rubber
IRRDB	International Rubber Research and Development Board
RRIM	Malaysian Rubber Research Institute
RM	Ringgit Malaysia

PB	Prang Besar
SST	Small Scale Trials
LST	Large Scale Trials
OFT	On Farm Trials
AACT	Acetyl-CoA-Acetyltransferase
HMGS	Hydroxymethylglutaryl-CoA Synthase
HMGR	Hydroxymethylglutaryl-CoA Reductase
MVK	Mevalonate Kinase
PMK	Phosphomevalonate Kinase
MVD	Diphosphomevalonate Decarboxylase
DXS	1-Deoxy-D-Xylulose-5-Phosphate Synthase
DXR	1-Deoxy-D-Xylulose-5-Phosphate Reductoisomerase
MCT	2-C-Methyl-D-Erythritol 4-Phosphate Cytidylyltransferase
CMK	4-(Cytidine 5'-Diphospho)-2-C-Methyl-D-Erythritol Kinase
MDS	2-C-Methyl-D-Erythritol 2,4-Cyclodiphosphate Synthase
HDS	4-Hydroxy-3-Methylbut-2-en-1-yl Diphosphate Synthase
HDR	4-Hydroxy-3-Methylbut-2-enyl Diphosphate Reductase
DPMD	diphosphomevalonate decarboxylase
MPD	mevalonate phosphate decarboxylase
M3K	mevalonate-3-kinase
PMVK	phosphomevalonate kinase
IPK	isopentenyl phosphate kinase
M3PK	mevalonate-3-phosphate-5-kinase
MBD	mevalonate-3,5-bisphosphate decarboxylase

DXP	deoxyxylulose-5-phosphate
RB	rubber biosynthesis
CPT	*cis*-prenyltranferase
SRPP	small rubber particle protein
REF	rubber-elongation factor
TF	transcription factors
SUT	sucrose transporter
SuSy	Sucrose synthase
RNA-seq	RNA sequencing
NGS	Next Generation Sequencing
NaCl	sodium chloride
DEPC	Diethylpyrocarbonate
CIA	chloroform-isoamyl alcohol
PCIA	phenol-chloroform-isoamyl alcohol
RISDA	Rubber Industry Smallholders Development Authority
NCBI	National Center for Biotechnology Information
FPKM	Fragments Per Kilobase of exon per Million

ANALISIS PERBANDINGAN TRANSKRIPTOM ANTARA KLON POKOK GETAH (*Hevea brasiliensis*) BERPRESTASI TINGGI DAN RENDAH YANG BERKAITAN DENGAN BIOSINTESIS GETAH ASLI

ABSTRAK

Hevea brasiliensis atau pokok getah merupakan spesies yang diketahui ramai dan tergolong dalam keluarga Euphorbiaceae adalah berasal dari hutan hujan Amazon. Spesies ini telah menyumbang sebahagian besar saham kepada ekonomi pertanian di Malaysia semenjak 1950an kerana kebolehan pokok ini untuk menghasilkan getah asli yang mempunyai nilai ekonomi yang tinggi. Getah asli mempunyai sifat yang istimewa seperti fleksibel, keanjalan yang tinggi dan mempunyai penyebaran haba yang cekap kerana kehadiran *cis*-1,4-polisoprene. *Cis*-1,4-polisoprene ini disintesis melalui dua laluan biosintesis; laluan biosintesis isoprenoid and laluan biosintesis getah. Penghasilan *cis*-1,4-polisoprene yang tinggi telah dijumpai berkaitan dengan penghasilan lateks yang tinggi. Sebagai tambahan, beberapa faktor-faktor lain termasuk penglibatan faktor transkripsi dan pengangkutan sukrosa ke dalam sel latisifer juga mungkin memainkan peranan dalam penghasilan lateks yang tinggi. Memandangkan produksi lateks yang tinggi adalah berkaitan dengan biosintesis getah pada tahap genetik, faktor-faktor pembatas utama dalam penghasilan getah asli masih dalam perbincangan. Walau bagaimanapun, masih terdapat kekurangan maklumat mengenai pengekspresan gen-gen yang berkaitan dengan biosintesis getah yang boleh membantu untuk memahami bagaimana penghasilan lateks yang tinggi boleh diperolehi pada tahap genetik. Jadi, penyelidikan ini menumpukan untuk mendapatkan pengetahuan dan beberapa maklumat penting berkaitan dengan gen-gen yang mengawal penghasilan getah asli.

Pengetahuan ini adalah berkaitan dengan laluan-laluan biosintesis terlibat, hubungan antara tisu-tisu dan pengekspresan gen-gen yang berkaitan dengan penghasilan getah asli yang sangat penting dalam menghasilkan klon getah berhasil tinggi melalui program pembiakbakaan getah. Sebanyak 12 spesimen tisu-tisu daripada batang, daun and lateks dari RRIM 3001 and RRIM 712 telah diasingkan untuk mendapatkan keseluruhan RNA. Setiap spesimen kemudian dihantar untuk penjujukan RNA melalui penjujukan berpasangan menggunakan Illumina NextSeq 500 dan menghasilkan min 75,742,501 bacaan awal. Penilaian kualiti pada bacaan awal telah dijalankan dan telah menjana min sebanyak 73,816,244 bacaan bersih untuk digunakan dalam analisis seterusnya. Bacaan bersih kemudian telah dihimpunkan menggunakan program Cufflink untuk menjana 101,269 transkrip. Kesemua 101,269 transkrip telah dianotasi menggunakan beberapa pengkalan data termasuk NCBI (Pusat Maklumat Bioteknologi Kebangsaan), UniProt-SwissProt (Sumber Protain Universal), GO (Gen Ontologi), KEGG (Ensiklopedia Gen dan Genom Kyoto) dan Pfam (Domain Protein) menggunakan program Blast2GO untuk mencari nama gen dan maklumat lain. Analisis pembezaan dampingan antara tisu-tisu dari RRIM 3001 dan RRIM 712 telah dijalankan untuk mencari pengekspresan gen-gen yang terlibat dalam biosinthesis isoprenoid melalui laluan Mevalonate (MVA) and laluan 2-C-methyl-D-erythritol 4-phosphate (MEP) serta laluan biosintesis getah. Analisis telah menunjukkan gen-gen Asetil-CoA-Asetiltransferase (AACT) dan Difosfomevalonate Dekarbosilase (MVD) melalui laluan MVA manakala Protein Partikel Kecil Getah (SRPP) dan Faktor Pemanjangan getah (REF) melalui laluan biosintesis getah mempunyai nilai pengekspresan sangat tinggi dalam tisu-tisu batang, daun dan lateks dari RRIM 3001 berbanding RRIM 712. Faktor Tindakbalas Etilena (ERF), iaitu faktor transkripsi juga telah dijumpai berekspres sangat tinggi dalam klon berprestasi

tinggi (RRIM 3001) berbanding klon berprestasi rendah (RRIM 712). Analisis amplifikasi kuantitatif telah dijalankan untuk pengesahan lima gen yang dipilih menggunakan klon getah yang mempunyai ciri-ciri berprestasi tinggi, sederhana dan rendah. Pengekspresan bagi lima gen adalah berkadar langsung dengan ciri penghasilan getah. Oleh itu, lima gen ini boleh digunakan untuk penentuan klon getah baru yang berprestasi tinggi.

COMPARATIVE TRANSCRIPTOME ANALYSIS OF RUBBER TREE (*Hevea brasiliensis*) BETWEEN HIGH AND LOW YIELD RUBBER CLONES RELATED TO NATURAL RUBBER BIOSYNTHESIS

ABSTRACT

Hevea brasiliensis or rubber tree is a well-known species belonging to the Euphorbiaceae family are native to Amazon rainforest. The species which has contributed a major share to the agricultural economy of Malaysia since 1950s because of the ability to produce natural rubber with high economic value. Natural rubber has special properties such as flexible, high elasticity and efficient heat dispersion due to presence of *cis*-1,4-polyisoprene. *Cis*-1,4-polyisoprene is synthesized via two biosynthesis pathways; isoprenoid and rubber biosynthesis pathways. The high production of *cis*-1,4-polyisoprene within rubber tree are found to be related to high latex production. In addition, several other factors including transcription factors involvement and sucrose transport to the laticifer cell may played some role in the high latex production. Since the high latex production is related to the rubber biosynthesis at genetic level, the main limiting factors in the natural rubber production are still discussed. However, there are still lack of information about expression of genes related to rubber biosynthesis which can help to understand how high latex production can be achieved at genetic level. Therefore, this study is focussing on gaining some knowledge and important information related to genes that control natural rubber production. The knowledges are related to biosynthesis pathways involved, relationships of the tissues and expression of genes related to natural rubber production which is very important for development of high yield rubber clone through rubber breeding program. A total of twelve specimens

from bark, leaf and latex tissues of RRIM 3001 and RRIM 712 were isolated for total RNA. Each specimen was then sent for RNA sequencing through paired-end sequencing on the Illumina NextSeq 500 and generated a mean of 75,742,501 raw reads. The quality assessment of raw reads was performed and a mean of 73,816,244 of clean reads used for further analysis. The clean reads were then assembled using Cufflink program to generate 101,269 transcripts. The 101,269 transcripts were annotated using several databases including NCBI (National Center of Biotechnology Information), UniProt-SwissProt (Universal Protein Resources), GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes) and Pfam (Protein Domain) using Blast2GO program to find genes name and other information. The differential expression analysis between tissues of RRIM 3001 and RRIM 712 was then performed to find expression of genes involved in isoprenoid biosynthesis through Mevalonate (MVA) and 2-C-methyl-D-erythritol 4-phosphate (MEP) pathways as well as rubber biosynthesis pathways. The analysis showed that Acetyl-CoA-Acetyltransferase (AACT) and Diphosphomevalonate Decarboxylase (MVD) genes via MVA pathway while Small Rubber Particle Protein (SRPP) and Rubber Elongation Factor (REF) genes through rubber biosynthesis pathway have the highly expressed within bark, leaf and latex tissues of RRIM 3001 compared to RRIM 712. Ethylene Response Factor (ERF), a transcription factor was also found expressed highly within high yield clone (RRIM 3001) than low yield clone (RRIM 712). The quantitative amplification analysis was done to validate the five selected genes using rubber clones with high, moderate and low yield characteristics. The expression of five genes were directly proportional to the yield characteristics. Therefore, the five selected genes could be used for determination of new rubber clone with high latex yield characteristics.

CHAPTER 1

INTRODUCTION

1.1 Overview

Hevea brasiliensis, commonly known as rubber tree, is a perennial cross-pollinating and monoecious plant that belongs to the Euphorbiaceae family (Saha & Priyadarshan, 2012). It is native to Amazon rainforest which can be found in the northern part of South America, from Brazil to Venezuela, and Colombia to Peru and Bolivia (Priyadarshan, 2003). According to Malaysia Rubber Board (2009), rubber tree requires a deep soil of the loamy texture with free drainage and can be planted to a maximum elevation of 500 m above sea level. Currently, rubber tree grown widely in Asia including Malaysia, Indonesia, Thailand, Vietnam, Sri Lanka, China, India and Papua New Guinea while in Africa including Nigeria, Ivory Coast, Cameroon, Liberia and Gabon (Saha & Priyadarshan, 2012).

Within genus *Hevea*, only *H. brasiliensis* has been exploited as a natural rubber producer with high economic value and has contributed a major share to the agricultural economy of Malaysia (Rozhan, 2015). The rubber industry has been a pillar of the Malaysian economy since 1950's and continues to be a major contributor until present day (Malaysia Rubber Board, 2009). Presently, Malaysia is the world's fourth biggest producer of natural rubber after Thailand, Indonesia and Vietnam with 7.1% of total natural rubber exports and contribute to USD 936.5 million for Malaysian economy in 2018 (www.worldstopexports.com). According to data

provided by Department of Statistics Malaysia, natural rubber production in June 2017 was 50,614 tons, an increase of 7.3% compared to 47,188 tons in the previous year at the same month with smallholding sector being a major contributor exceeding 90% from the total production. However, exports of natural rubber declined by 12.9% from 51,328 in June 2016 tons to 44,683 tons in June 2017.

Natural rubber produced by *H. brasiliensis* is one of the most important polymers which provide major industrial raw material. Rubber in the form of latex is composed of 94% cis-1,4-polyisoprene and 6% proteins in addition of fatty acids (Sakdapipanich, 2007). They have several important properties such as flexibility, impermeability to liquids and resistance to abrasion (Mantello *et al.,* 2014). These unique properties make natural rubber widely use in various applications (Cornish, 2001). Natural rubber has been used in more than 40,000 products, including over 400 medical devices (Cornish, 2001). According to Department of Statistics Malaysia, rubber gloves industry remains as the main domestic consumer with consumption of more than 70% followed by rubber thread (9.9%), tyres and tubes (7.3%) and others (8.9%).

Natural rubber is made up of *cis*-isoprene units, a derived product from isopentenyl diphosophate (IPP) (Lau *et al.,* 2016). There are about 5,000 to 10,000 isoprene units organized in an unbranched chain required to form *cis*-isopropene (Kush, 1994). The synthesis of *cis*-isopropene usually takes place in the cytoplasm of highly specialized cells, known as laticifers (Lau *et al.,* 2016). Laticifers are differentiated from the cambium and arranged in the inner bark of *H. brasiliensis* (De

Fay & Jacob, 1989). The cytoplasm of laticifers tubules which is known as latex expelled when the bark is wounded or cut by tapped (Gomez & Moir, 1979). Previous study has shown that the biosynthesis of IPP proceeds via two distinct routes: the cytoplasmic mevalonate pathway (MVA) and the plastidic 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway. IPP are produced from acetyl-coenzyme A via MVA pathway and pyruvate or glyceraldehyde 3-phosphate via MEP pathway (Tritsch *et al.*, 2010). The synthesis of IPP via MVA pathway involves six genes including acetyl-coenzyme A (Acetyl-CoA), acetoacetyl-coenzyme A (Acetoacetyl-CoA), hydroxymethylglutaryl coenzyme-A (HMG-CoA), mevalonate, mevalonate monophosphate (Mevalonate OP) and mevalonate diphosphate (Mevalonate OPP) to become IPP. Moreover, the synthesis of IPP through MEP pathway begins with pyruvate or glyceraldehyde 3-phosphate followed by 1-deoxy-D-xylulose 5-phosphate (DXP), 2-C-methyl-D-erythritol 4-phosphate (MEP), 2-C-methylD-erythritol-4-phosphate (CDP-ME), 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol (CDP-ME2P), 2-C-methyl-D-erythritol 2,4-cyclodiphosphate (MECDP), 4-hydroxy-3-methyl-but-2-en-1-yl diphosphate (HMBDP).

Recent surge in high-throughput sequencing efforts has enhanced the genetic resources available for *H. brasiliensis* (Rahman *et al.*, 2013*;* Lau *et al.*, 2016). The genomic data of *H. brasiliensis* is crucial in obtaining information at chromosome and genome level and obtaining novel insights into latex production trait. In this context, RNA-seq is one of the potential tools for understanding the biological process in *H. brasiliensis* governing agronomic traits for latex biosynthesis.

RNA-seq is a transcriptome profiling approach, which uses deep-sequencing technologies. Previously, the hybridization-based or sequence tag-based approaches have been used to address many biological issues. However, RNA-seq as a single platform that enables studying expression levels of entire transcriptome/ specific genes, differential splicing, allele-specific expression, RNA editing, fusion transcripts and SNP analysis (Wilhelm & Landry, 2009). RNA-seq also provide accurate transcriptome profiling that most closely resembles the cell biological processes (Wang *et al.,* 2009). There are multiple RNA-seq studies have been done using rubber tissues as the specimens as reported by Li *et al.* (2015), Li *et al.* (2016), Montoro *et al.* (2018), Abdul Rahman *et al.,* (2019) and Roy *et al.* (2019). On the other hand, RNA-seq also have been reported to be used for transcriptome profiling studies in maize, rice and soybeans (Eveland *et al.,* 2010; Severin *et al.,* 2010; Zhang *et al.,* 2010; Xu *et al.,* 2012).

1.2 Problem Statement

In Malaysia, natural rubber production coordinated by Malaysian Rubber Board and Sime Darby Plantation have supplied approximately 50,000 tons every month to meet the demands. The rubber industries in Malaysia have contribution nearly USD 1 million every year for Malaysian economy. However, the natural rubber production still insufficient to meets the global demands. Although rubber breeders produced RRIM 3001 clone which can produce approximately 3,000 kg/ha/yr of natural rubber, the reducing rubber planting area has caused Malaysian export declined over the year. However, most smallholders in Malaysia do not grow RRIM 3001 clone due to their supplies. Moreover, conventional breeding method

could take more than 20 years to produce new rubber clones with expected natural rubber production. The main challenge is how to increase the natural rubber production to meet the demands. The problem could be solved via molecular assisted breeding technique through rubber breeding program which can help rubber breeders to produce new rubber clones with higher latex yield at a faster rate. However, lack of information at genetic level related to the natural rubber biosynthesis in terms of genes directly related to high latex yield and expression of genes within rubber tree have slower the process. Although several genes have been suggested as limiting factors for natural rubber biosynthesis, there are no concrete dataset which could determine the high yield latex production characteristics. The hypothesis of this research is the high expression of genes related to isoprenoid biosynthesis and rubber biosynthesis pathways could determine high natural rubber production using transcriptome analysis.

1.3 Importance of Study

This study is important for the determination of genes related to the high natural rubber production since there are lack of information due to the limiting data from previous studies. The development of high-throughput sequencing has increased the ability to gain more knowledge at molecular level through RNA-seq technique. The RNA-seq would provide the transcriptome dataset which could be manipulated for this study. The analysis using transcriptome dataset is the best option to uncover the complexity of biological functions in *H. brasiliensis* by studying gene expression, differential expression, relationships between tissues and identification of new transcripts leading to new function discovery. This information is very

important since there is incomplete information on genes that control agronomic traits especially on natural rubber production. This research findings can provide new information of transcriptome to address the biological and molecular issues focused on latex production as well as identifying the key regulatory genes and how the expression of genes can help in determination of high natural rubber production.

1.4 Objectives of the Study

The main objective of this study is to gather important information about genes related to natural rubber biosynthesis and searching for the genes which influence high natural rubber production through several biosynthesis pathways and other related factors. Based on the study purpose, there are three subprojects to be carried out which are:

Subproject 1: **Transcriptome analysis of the selected clones of rubber tree (*H. brasiliensis*)**

1. To determine total transcriptome activity between selected clones.
2. To identify the functional annotations from transcriptome dataset.
3. To determine qualitative and quantitative genes information from transcriptome dataset of selected rubber clones.

Subproject 2: **Comparative Transcriptome of High and Low Yield Rubber Clones based on Isoprenoid Biosynthesis Pathways, Rubber Biosynthesis Pathways, Transcription Factor and Sucrose Transport**

1. To identify the isoprenoid biosynthesis, rubber biosynthesis, transcription factor and sucrose transport related genes for rubber clones.
2. To determine the genes expression of each transcripts related to isoprenoid biosynthesis, rubber biosynthesis, transcription factor and sucrose transport.
3. To determine the differential expression of genes involved in isoprenoid biosynthesis, rubber biosynthesis, transcription factor and sucrose transport between high and low yield rubber clones.

Subproject 3: **Quantification of Expression from Selected Genes Related to the Natural Rubber Biosynthesis**

1. To determine the rubber content from latex tissue of selected rubber clones.
2. To determine the expression value of potential genes through quantitative amplification.
3. To compare the expression value from RNA-seq and relative transcript abundance quantification datasets.

CHAPTER 2

LITERATURE REVIEW

2.1 The Genus *Hevea*

Hevea brasiliensis (Willd. ex Adr. de Juss.) Muell. -Arg., which also known as rubber or Para rubber tree is a monoecious, tall, perennial and cross-pollinated tree species (Priyadarshan & Clement-Demange, 2004; Mantello *et al.*, 2012). Rubber tree can grow up to 30 m high with over 3 m girth diameter and can live up to 100 years in the wild (Priyadarshan & Clement-Demange, 2004; Lau *et al.*, 2016). It belongs to the genus *Hevea* within the Euphorbiaceae family and native to Amazon rainforest (Mantello *et al.*, 2012). Within the genus *Hevea*, there are 11 species including *H. benthamiana* Muell. -Arg., *H. brasiliensis, H. carmagoana* Pires, *H. camporum* Ducke, *H. guianensis* Aubl, *H. microphylla* Ule, *H. nitida* Mart. ex-Muel. -Arg., *H. pauciflora* (Spruce ex-Benth.) Muell. -Arg., *H. rigidifolia* (Spruce ex-Benth.) Muell. -Arg., *H. spruceana* (Benth.) Muell. -Arg. and *H. paludosa* Ule (Schultes, 1987; Priyadarshan & Goncalves, 2003). Of these 11 species, only *H. brasiliensis* has been commercially planted because of their latex or natural rubber (NR) yield which gave an economic value to rubber producing countries (Lau *et al.*, 2016). In addition, *H. brasiliensis* produces high quality rubber. According to Tan *et al.* (2017), NR quality is usually determined by their molecular weight where the higher the molecular weight, the better quality of NR. There are limited number of plants which can produce large amounts of high-quality NR where molecular weight is more than 1 million Daltons including *H. brasiliensis* (rubber), *Parthenium*

argentatum (guayule) and *Taraxacum koksaghyz* (Russian dandelion) (Mantello *et al.*, 2014). The higher production of high-quality NR has attracted more countries to invest in the rubber industries.

Historically, *Hevea* has been found documented in ancient religious books from Mexico about 600 B.C. while Christopher Columbus was known as the one of earliest person to describe the rubber morphology and discovered the crude product in the fifteenth century (Dijkman, 1951). An astronomer known as De La Condamine was the first person to send samples from Peru to France in 1736 with full details about the habitat and processing procedures for rubber tree which he named as 'caoutchouc' in French language for rubber (Dijkman, 1951). Since then, raw rubber material has been taken from the Amazonian rainforest to Europe and several discoveries were made from rubber materials (Malaysia Rubber Board, 2009). Furthermore, the name 'rubber' was given by Priestley while Fuse 'e Aublet, a botanist who described the genus *Hevea* in 1775 (Malaysia Rubber Board, 2009; Priyadarshan & Clement-Demange, 2004). Therefore, Clement-Demange *et al.* (2000) has suggested that *Hevea* species might have evolved in the Amazon rainforest more than 100,000 years.

On the other hand, *Hevea* species around the world came from centre of origin or their natural habitat which consists of Bolivia, Brazil, Colombia, Guyana, Peru, Surinam, and Venezuela (Priyadarshan & Clement-Demange 2004). There are four species can be found in Colombia, three in Venezuela, two in Bolivia and one in Guyana, Peru and Surinam (Priyadarshan & Clement-Demange, 2004). Moreover, all species can be found in Brazil except *H. microphylla* Ule while *H. guianensis* Aubl

has been recognised as the most widely adapted species within the genus. Currently, more than 10 countries within Asian, African and American continents including Thailand, Malaysia, Indonesia, China, India, Vietnam, Ivory Coast, Liberia, Sri Lanka, Philippines, Brazil, Cameroon, Nigeria Cambodia, Myanmar, Ghana, Gabon and Papua New Guinea have planted rubber tree and become rubber producers in the world. In addition, Southeast Asian countries have become leading rubber producers with more than 80% of total rubber in 2017 (www.worldstopexports.com).

2.2 *Hevea brasiliensis* in Malaysia

In June 1876, about 70,000 seeds were taken from the Rio Tapajoz region of the upper Amazon Brazil and transported to Kew Botanic Gardens, England by Wickham (Baulkwill, 1989). In September 1877, a total of 1,911 germinated seeds were sent to Ceylon (Sri Lanka) while 22 seedlings have been sent to Singapore (13 seedlings) and Kuala Kangsar (9 seedlings) (Malaysia Rubber Board, 2009). Therefore, the source of rubber tree in Malaysia originated from the Wickham collection. However, several expeditions were conducted by MRB and International Rubber Research and Development Board (IRRDB) which involved breeders and researchers from Malaysia to collect new seeds from wild species of *H. brasiliensis* for rubber genetic resources expansion (Malaysia Rubber Board, 2009).

Through rubber breeding program, Malaysian breeders have produced more than 180 rubber clones since 1920s (Malaysia Rubber Board, 2009). According to Dijkman (1951), rubber breeding was initiated at the beginning of twentieth century with the introduction of very strict selection methodologies including bud grafting,

generative and vegetative. The first rubber estate was established in Melaka in 1903 and breeding program carried out by Department of Agriculture until 1925. After that, the rubber research program was continued by Rubber Research Institute of Malaya (previous name of MRB) (Malaysia Rubber Board, 2009). Moreover, the production of new rubber clones generally involves four stages including evaluation via nursery screening, Small Scale Trials (SST), Large Scale Trials (LST) and On Farm Trials (OFT). Each stage of cultivation involves a minimum of 10 years (7 years immaturity & 3 years initial yielding phase). Hence conventionally a period of 30-35 years is required for commercial release (Malaysia Rubber Board, 2009).

According to MRB (2009), seven series of clones had been developed and recommended to the industry produced by Malaysian Rubber Research Institute (RRIM) under the names RRIM 500 (1928-1931), RRIM 600 (1937-1941), RRIM 700 (1947-1958), RRIM 800 (1959-1965), RRIM 900 (1966-1973), RRIM 2000 (1974-2008) and RRIM 3000 (2009 until now) series clones. Moreover, Sime Darby Plantation also had developed rubber series under the names PB 100, PB 200 and PB 300 series clones. Interestingly, some of these clones are also widely planted in other rubber growing countries. In the early years, the development of new rubber clone was focussing on improving the NR yield. With the introduction of new genetic materials from Brazil in the 1950s in the development of the RRIM 900, RRIM 2000 and RRIM 3000 series clones have successfully increased the yield potential up to 3,000 kg/ha/year (Malaysia Rubber Board, 2009). More recently, the highest NR production has been reported was RRIM 3001 with approximately 3,000 kg/ha/year of yield.

In Malaysia, rubber trees have continued to be planted although the rubber planting area has been declining since 1982 (Malaysia Rubber Board, 2009). According to data provided by MRB and Department of Statistics Malaysia, in 2016 total rubber planting area was 10,764,000 hectares with Peninsular Malaysia covers 7,713,000 hectares, Sarawak with 1,797,000 hectares and Sabah with 1,254,000 hectares while in 2006, the total rubber planted area was 12,635,900 hectares with Peninsular Malaysia as the major contributor with 10,425,900 hectares followed by Sarawak with 1,556,100 hectares and Sabah with 652,800 hectares. Of the total 10,764,000 hectares, about 92.6% from the natural rubber production involved smallholder's sector while 7.4% involved estate. The decline in rubber planting area has caused Malaysian export to drop from third biggest natural rubber producer to fourth biggest natural rubber producer country in the world (Malaysia Rubber Board, 2009).

2.3 Economic Importance

The rubber industry has high contribution to the Malaysian economy since 1950s (Malaysia Rubber Board, 2009). Recently, Malaysia is the world's fourth contributor of natural rubber after Thailand, Indonesia and Vietnam with 7.1% of total natural rubber exports which has contributed about RM 936.5 million for Malaysian economy in 2018 (www.worldstopexports.com). In addition, Department of Statistics Malaysia has reported that natural rubber production in July 2018 was 56,397 tons, an increase of 2.1% compared to previous year production with smallholding sector as the biggest contributor with 90% from the total production. Moreover, natural rubber exports from Malaysia has showed an increase of 7.8%

with the yield about 56,490 tons where China as the highest importer with approximately 52.4% followed by Germany with 11.6%, Iran with 7.6%, Finland with 4.3% and USA with 2.5% of the total exports in July 2018.

Apart from that, Malaysia also has exported downstream products such as rubber products, rubberwood product and other rubber products. According to data provided by Department of Statistics Malaysia, Malaysia has gained approximately RM 33.30 billion from rubber industry including natural rubber exports which accounted about 4.27% of the total national exports in 2016. Of these RM 33.3 billion, rubber products have contributed RM 17.99 billion, rubberwood about RM 8.12 billion and other rubber products have contributed RM 0.94 billion. Rubber products including raw product, tyres, gloves, catheters and latex have contributed about 2.31% while rubberwood products including furniture, toys and kitchen accessories have contributed about 1.04% of total national exports. In addition, other rubber products which consists of synthetic rubber, reclaimed rubber, waste rubber, compound rubber and unvulcanised rubber have contributed about 0.12% of total national exports.

On the other hand, rubber industries have contributed to the socio-economy involving more than 200,000 smallholder families and over 64,000 workers in the downstream manufacturing sector (Malaysia Rubber Board, 2009). Although the natural rubber price has been floating, in July 2018 the Department of Statistics Malaysia have released the data which showed that the industry has contributed approximately RM 17.8 million to the workers and smallholders in terms of salaries and wages. Compared to July 2017, the salaries and wages have increased from RM

13.9 million. These statistics have shown the importance of the rubber industry as the contributor to the socio-economy of the people as well as Malaysian economy for more than 50 years.

2.4 Natural Rubber Biosynthesis

Natural Rubber (NR) production is an important biological feature of *H. brasiliensis* which makes them very special compared to other crops. According to Tang *et al.* (2016), NR makes up approximately one-third of the total latex volumes in the cytoplasm of the laticifers in *Hevea*. NR can be extracted by tapping the bark, a commercial method which can harvest the latex for continuous production (Tang *et al.*, 2016).

NR is synthesized on the surface of rubber particles suspended in the cytoplasm of the highly specialized plant cells known as laticifers or latex vessels (D'Auzac *et al.*, 1997). The morphological study has found that there are primary and secondary laticifers involved in the NR production, however only secondary laticifers in the bark of rubber trees are economically exploited. Primary laticifers which is present in the bark of young stems or shoots usually disappear in the bark of more than six year old rubber trees where it is distributed at random in the phloem (Tan *et al.*, 2017). Moreover, primary laticifers are differentiated from the meristematic cells. The initial laticifer tubes in the meristematic cells have characteristics such as thin, straight and un-branched, however the primary laticifer tubes would then grow and expand and forming a necklace-like structures towards

the neighbouring parenchyma cells (Tan *et al.*, 2017). The expand laticifer tubes were then connected by bridges to form a laticifer network with neighbouring tubes.

In contrast, the secondary laticifers cells differ from primary laticifers by their distribution pattern and growth. The secondary laticifer cells which originate from cambium cells have thick, straight, and smooth cell walls. Interestingly, only the secondary laticifers can be found within mature tree trunk. They are present in rings parallel to the vascular cambium in the secondary phloem of the trunk. Moreover, the secondary laticifer rings number varies in diff erent rubber clones where the clones with greater number of secondary laticifer rings usually have a higher rubber yield. Therefore, the secondary laticifers have become the focus of researchers due to their great impact to the latex yield (Tan *et al.*, 2017).

NR is a latex polymer which consists of 94% *cis*-1,4-polyisoprene and another 6% including proteins and fatty acid (Sakdapipanich, 2007). NR biosynthesis has involved several processes including carbon fixation in the leaf, loading and transportation of the sucrose and specific metabolic pathway to generate biosynthesis precursors and storage of polyisoprenes in the laticifer cells. The key factors which makes NR has special properties such as flexible, high elasticity and efficient heat dispersion is due to the presence of *cis*-1,4-polyisoprene (Mantello *et al.*, 2014). *Cis*-1,4-polyisoprene is a biopolymer which is made up of repeating isopentenyl diphosphate (IPP) units in the *cis*-configuration (Makita *et al.*, 2017). There are two biosynthetic pathways which can generate the *cis*-1,4-polyisoprene; isoprenoid and rubber biosynthesis pathways. The production of IPP unit via isoprenoid biosynthesis involving the condensation of either acetyl CoA or glyceraldehyde-3-phosphate and

pyruvate as the precursor. The IPP unit was then polymerized to form *cis*-1,4-isoprene unit. The polymerization process is essential in the development of rubber particles.

2.5 Isoprenoid Biosynthesis

Isoprenoids are an important component of cell membranes in all organisms and one of the oldest known biomolecules where they have been found on 2.5 billion years sediments (Summons *et al.*, 1999; Hoshino & Gaucher, 2018). Moreover, isoprenoids are one of the largest family of organic compounds which consists over 65,000 compounds (Buckingham, 2007). Isoprenoids also have diverse roles in all organisms including serve as cell-signalling agents, lipids in cell membranes, quinones in electron transport chains, as photosynthetic pigments, as hormones and as plant defense compound (Lange *et al.*, 2000; Gershenzon & Dudareva, 2007; Crowell & Huizinga 2009; Nowicka & Kruk 2010; Jain *et al.*, 2014).

NR which is made up of *cis*-1,4-polyisoprene is derived from isopentenyl diphosphate (IPP). The isoprenoid biosynthesis is responsible in the production of IPP. Interestingly, IPP is synthesized through condensations process of the five-carbon compound within eubacteria, archaebacteria and eukaryotes via two distinct and independent biosynthetic routes. Previously, researchers have found that the IPP production in mammals and yeast starts from acetyl-CoA which then proceeds through the intermediate mevalonic acid (MVA) and concluded that the pathway to be same in all organisms (Qureshi & Porter, 1981). However, studies done by Rohmer *et al.* in 1993 has found that the IPP production in bacteria and green algae

could be generated from pyruvate and glyceraldehyde-3-phosphate. Therefore, Pérez-Gil & Rodríguez-Concepción (2013) concluded that IPP is derived via two distinctive pathways: the mevalonate (MVA) pathway and the methylerythritol 4-phosphate (MEP) pathway. Through the MVA pathway, IPP is derived from the condensation of acetyl-CoA while the MEP pathway involves the condensation of either pyruvate or glyceraldehyde-3-phosphate (Lange *et al.*, 2000).

In 1993, Rohmer *et al.* has suggested that IPP production of archaea and eukaryotes are through MVA pathway while eubacterial, algae and higher plants were found to produce IPP via MEP pathway. However, in 1999, Rohmer has found that all living organisms including archaebacteria, animal and fungi usually used either MVA or MEP pathway for IPP production. Interestingly, plants use both pathways for isoprenoid biosynthesis but only differ in the localization where MVA pathway are used within cytoplasm while MEP pathway usually used in the chloroplast. Studies by Lichtenthaler in 1997 has shown that although isoprenoid biosynthesis of some bacteria is via MVA pathway, most bacteria usually produce IPP via MEP pathway. In addition, photosynthetic eukaryotes have both the MVA and the MEP pathways where MEP pathway is originated from an ancestral MEP pathway within a symbiotic of cyanobacteria and eukaryotes earlier. Figure 2.1 shows the summary of the genes and enzymes involved in MVA and MEP pathways.

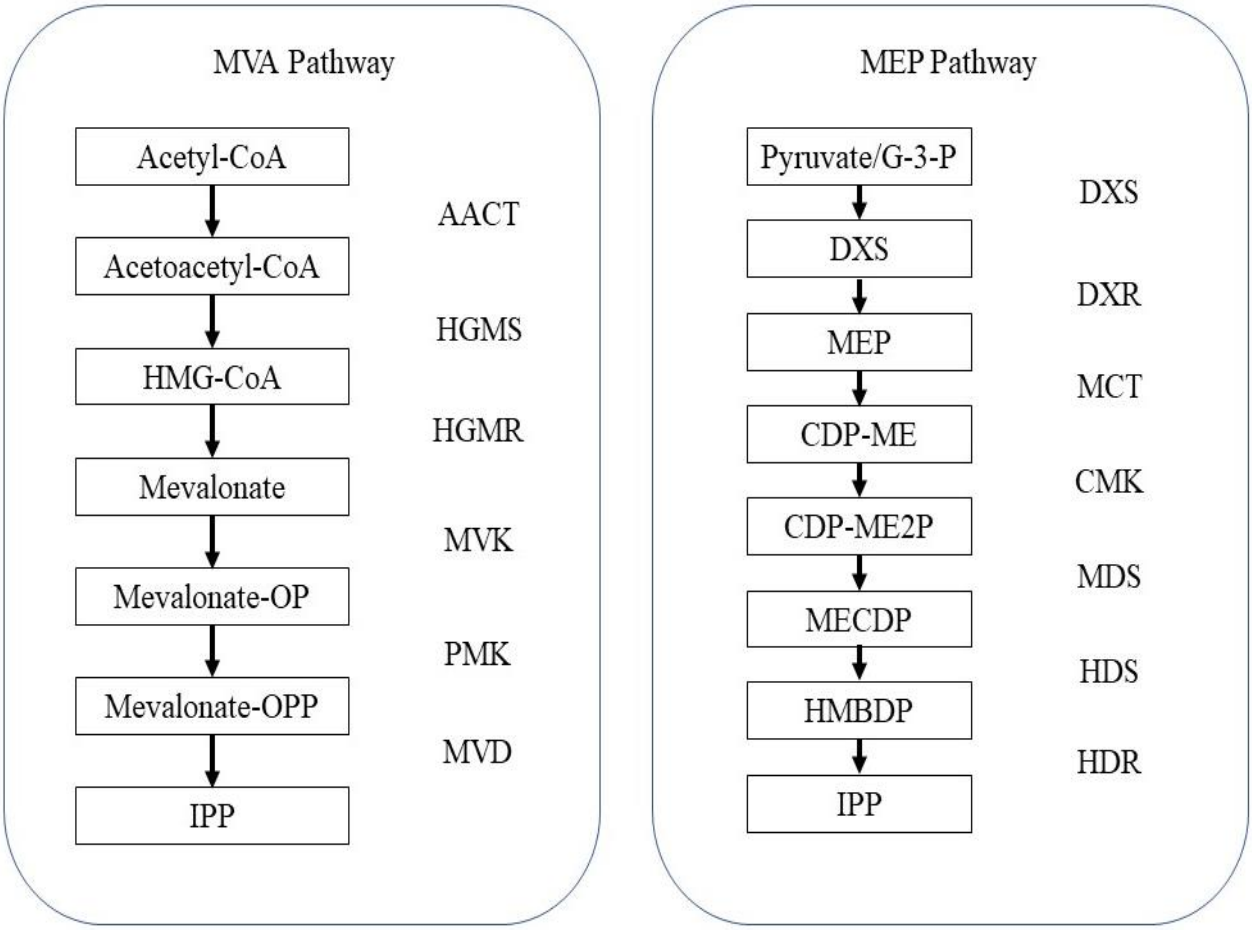

Figure 2.1: The summary of genes and enzymes involved in MVA and MEP pathways according to Lau *et al.*, (2016). MVA pathway; AACT (Acetyl-CoA-Acetyltransferase), HMGS (Hydroxymethylglutaryl-CoA Synthase), HMGR (Hydroxymethylglutaryl-CoA Reductase), MVK (Mevalonate Kinase), PMK (Phosphomevalonate Kinase), MVD (Diphosphomevalonate Decarboxylase). MEP pathway; DXS (1-Deoxy-D-Xylulose-5-Phosphate Synthase), DXR (1-Deoxy-D-Xylulose-5-Phosphate Reductoisomerase), MCT (2-C-Methyl-D-Erythritol 4-Phosphate Cytidylyltransferase), CMK (4-(Cytidine 5'-Diphospho)-2-C-Methyl-D-Erythritol Kinase), MDS (2-C-Methyl-D-Erythritol 2,4-Cyclodiphosphate Synthase), HDS (4-Hydroxy-3-Methylbut-2-en-1-yl Diphosphate Synthase), HDR (4-Hydroxy-3-Methylbut-2-enyl Diphosphate Reductase).

2.5.1 Mevalonate (MVA) Pathway

Mevalonic acid is a primary precursor to the production of IPP via mevalonate (MVA) pathway. MVA pathway which was first discovered in 1950s is an important metabolic pathway which responsible in synthesizing multiple compound such as cholesterol, isopentenyl tRNA and ubiquinone (Bloch, 1992; Buhaescu & Izzedine, 2007). MVA pathway which begin with acetyl-CoA and ends with IPP production are present in archaea, eukaryotes and some bacteria. According to Mantello *et al.* (2014), MVA pathway was proceed with the involvement of six enzymes including Acetyl-CoA-Acetyltransferase (AACT), Hydroxymethylglutaryl-CoA Synthase (HMGS), Hydroxymethylglutaryl-CoA Reductase (HMGR), Mevalonate Kinase (MVK), Phosphomevalonate Kinase (PMK) and Diphosphomevalonate Decarboxylase (MVD) to produce single isoprene unit.

The IPP production via MVA pathway is dependent on enzymes present which produced earlier during transcription and translation process and each enzyme has different role. The first enzyme, *AACT* would catalyse the conversion of acetyl-Coenzyme A to acetoacetyl-Coenzyme A while HMGS would catalyse the conversion of acetoacetyl-Coenzyme A to hydroxymethylglutaryl-Coenzyme A and HMGR would catalyse the conversion of hydroxymethylglutaryl-Coenzyme A to mevalonate. Moreover, MVK catalyse the conversion of mevalonate to mevalonate monophosphate, PMK catalyse the conversion of mevalonate monophosphate to mevalonate diphosphate while MVD catalyse the conversion of mevalonate diphosphate to isopentenyl diphosphate or IPP (Makita *et al.*, 2017).

Although the first enzyme (AACT) is not specifically involved in the isoprenoid biosynthesis, the other five enzymes are uniquely dedicated to the MVA pathway (Lange *et al.*, 2000). Each gene involved in MVA pathway are important in the biosynthesis of isoprenoid. According to Miziorko (2011), AACT which belongs to thiolase group which has function in transferase activity enzyme has been widely observed in metabolism of prokaryotes and eukaryotes. AACT (EC 2.3.1.9) enzyme catalysed the biosynthetic reaction via MVA pathway through condensation process which formed acetoacetyl-CoA from two acetyl-CoA molecules. Moreover, HMG-CoA synthase or HMGS enzyme is responsible in the biosynthesis of 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). The biosynthetic reaction which catalysed by HMGS (EC 2.3.3.10) enzyme is irreversible and the process is done through condensation process of acetoacetyl-CoA. Furthermore, HMG-CoA reductase or HMGR (EC 1.1.1.34) enzyme has been reported presence in eukaryotes, archaebacteria, and some eubacteria. The HMGR enzyme react by two steps reduction process of HMG-CoA to form mevalonate.

In addition, mevalonate kinase which is also known as MVK (EC 2.7.1.36) enzyme can be found in eukaryotes, archaea, and certain eubacteria. MVK enzyme is responsible in the transfer of ATP's γ-phosphoryl to the C5 hydroxyl oxygen of mevalonic acid to form mevalonate-5-phosphate and ADP in MVA pathway. It can be inhibited by geranyl diphoshophate and farnesyl diphosphate in the isoprenoid biosynthesis process. Phosphomevalonate kinase or PMK (EC: 2.7.4.2) enzyme is present in eukaryotes and some eubacteria. However, the amino acid sequences for animal was not orthologous to those for PMK in plants, fungi, and bacteria where it catalysed different enzymatic reaction depending on their source. PMK enzyme also

catalysed the reversible reaction of mevalonate-5-phosphate and ATP to form mevalonate-5-diphosphate and ADP. Mevalonate diphosphate decarboxylase which is known as MVD (EC: 4.1.1.33) enzyme is responsible in the catalysation of ATP dependent decarboxylation of mevalonate-5-diphosphate to form isopentenyl 5-diphosphate. The reaction catalysed by MVK enzyme has been measured in animals, plants and yeast and the activity is important to the MVA pathway of isoprenoid biosynthesis.

MVA pathway within domain eukaryotes is well-conserved while MVA pathways within domain archaea composed about three variations based on their routes in the IPP production; DPMD, MPD and M3K routes (Hoshino & Gaucher, 2018). Most bacteria have similar routes with the eukaryotic MVA pathway, while only Chloroflexi harbors is homologous to an archaeal-type MVA pathway (Nishimura *et al.*, 2013). The different routes for archaea was determined during the conversion process in the pathway. According to Hoshino & Gaucher (2018), DPMD routes occurs when mevalonate-5-phosphate is converted to mevalonate-5-diphosphate using phosphomevalonate kinase (PMVK) and finally to IPP using diphosphomevalonate decarboxylase (DPMD) enzyme. However, MPD route occur when mevalonate-5-phosphate is converted to Isopentenyl phosphate using mevalonate phosphate decarboxylase (MPD) enzyme before conversion to IPP using isopentenyl phosphate kinase (IPK) enzyme.

In addition, M3K route occur when mevalonate was converted to mevalonate-3-phosphate using mevalonate-3-kinase (M3K) enzyme instead of mevalonate-5-phosphate in DPMD and MPD routes and then converted again to mevalonate-3,5-

biphosphate using mevalonate-3-phosphate-5-kinase (M3PK) enzyme. The mevalonate-3,5-biphosphate was then converted to isopentenyl phosphate using mevalonate-3,5-bisphosphate decarboxylase (MBD) before conversion to isopentenyl phosphate and IPP which is same with MPD route. Figure 2.2 shows the genes and enzymes involved in MVA pathway via three routes based on description by Hoshino & Gaucher (2018).

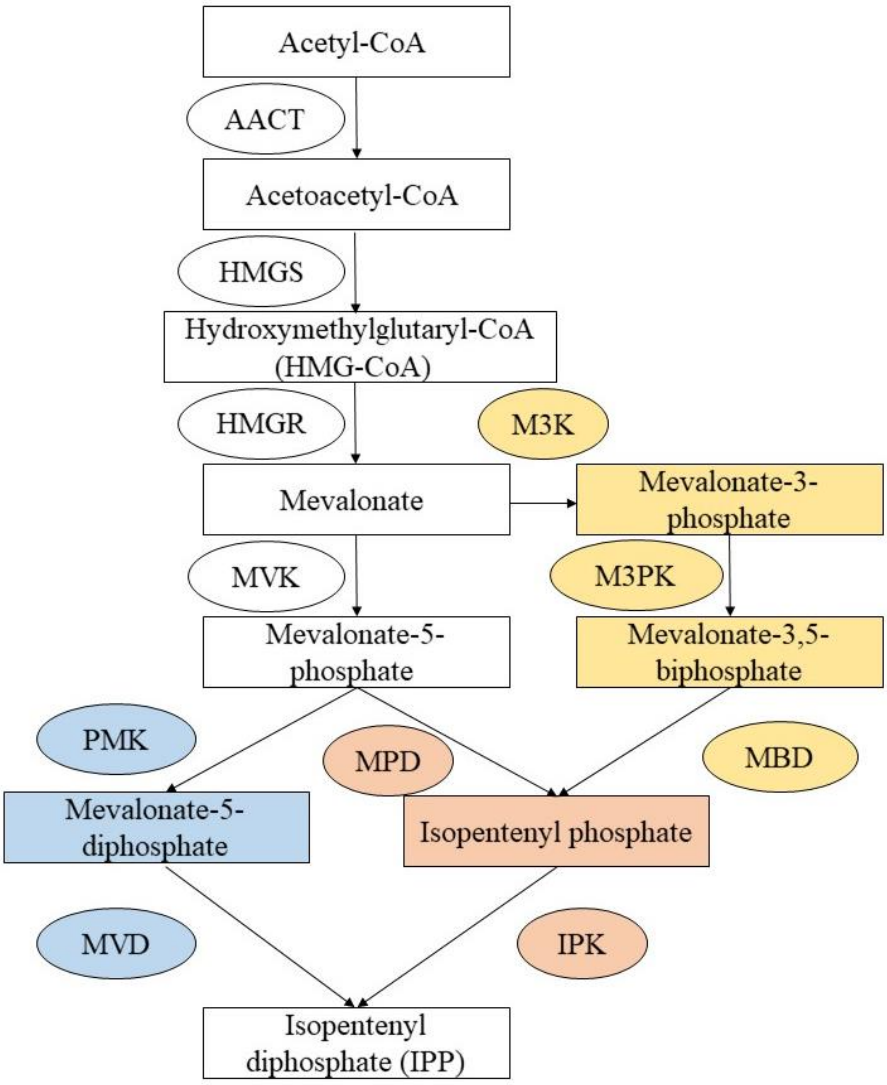

Figure 2.2: Three variations of the MVA pathway discovered in archaeal type based on Hoshino & Gaucher (2018). The coloured boxes indicate the route for genes and enzymes. The enzymes involved are AACT (Acetyl-CoA-acetyltransferase); HMGS (Hydroxymethylglutaryl-CoA synthase), HMGR (Hydroxymethylglutaryl-CoA reductase), MVK (Mevalonate kinase), PMK (Phosphomevalonate kinase), MVD (Diphosphomevalonate decarboxylase), MPD (Mevalonate phosphate decarboxylase), IPK (Isopentenyl phosphate kinase), M3K (Mevalonate-3-kinase), M3PK (Mevalonate-3-phosphate-5-kinase) and MBD (Mevalonate-3,5-bisphosphate decarboxylase).

2.5.2 Plastidic 2-C-Methyl-D-Erythritol-4-Phosphate (MEP) Pathway

Before 1980, the isoprenoid production was only known to be produced via MVA pathway. In the late 1980s, an alternative pathway named as MVA-independent pathway has been identified in bacteria which also responsible in the biosynthesis of the isoprene units (Rohmer *et al.*, 1993). The pathway has been initially found by several research groups, but Rohmer's group was the first to publish their work (Flesch & Rohmer, 1988; Rohmer *et al.*, 1993). The MVA-independent pathway was previously named as Rohmer pathway due to the founder of this pathway. However, its name was then changed after the first intermediate, deoxyxylulose-5-phosphate (DXP) pathway. Recently, the pathway is named after its first precursor, methylerythritol 4-phosphate (MEP) following the same rule applied to the MVA pathway and known as MEP pathway (Rodriguez-Concepsion & Boronat, 2002).

MEP pathway has been found present in bacteria, green algae and the malaria parasite (*Plasmodium falciparum*) (Boucher & Doolittle, 2000). However, it is absent from isoprenoid biosynthesis within archaebacteria, fungi and animals where these groups mainly produce isoprene unit via MVA pathway (Rohmer *et al.*, 1993). Interestingly, plant use both MVA and MEP pathways for isoprenoid biosynthesis in different location. In plant, isoprenoid biosynthesis proceeds via MEP pathway in plastids while through MVA pathway in cytoplasm and mitochondria (Lichtenthaler, 1997; Rohmer, 1999).