PROTEOMIC PROFILING OF FATTY ACID BIOSYNTHETIC ENZYMES FROM OIL PALM CHROMOPLAST

A thesis

submitted in partial fulfilment

of the requirement for the Degree of

Doctor of Philosophy

at

Lincoln University

by

Benjamin Lau Yii Chung

Lincoln University

2015

Declaration

Parts of this thesis are to be submitted for publication and have been presented in advance of submission of the thesis.

Publications

Benjamin Lau Yii Chung, Santanu Deb-Choudhury, James D. Morton, Jolon M. Dyer, Umi Salamah Ramli, Stefan Clerens. (201x) Optimization of oil palm chromoplast isolation and protein extraction for proteomic analyses, *in preparation*.

Benjamin Lau Yii Chung, James D. Morton, Santanu Deb-Choudhury, Jolon M. Dyer, Umi Salamah Ramli, Stefan Clerens. (201x). Development of gel-based and non-gel proteomic approaches to characterise protein expression of oil palm fatty acid biosynthetic enzymes, *in preparation*.

Benjamin Lau Yii Chung, James D. Morton, Santanu Deb-Choudhury, Jolon M. Dyer, Umi Salamah Ramli, Stefan Clerens. (201x). New insights into the phosphorylation role in the regulation of oil palm fatty acid biosynthetic enzymes, *in preparation.*

Presentations

The chromoplast proteome of oil palm fruit: A poster presentation at the 19th Lorne Proteomics Symposium, Lorne, Victoria, Australia – 6-0th February 2014.

Abstract

Abstract of a thesis submitted in partial fulfilment of the requirements for the Degree of Ph.D.

PROTEOMIC PROFILING OF FATTY ACID BIOSYNTHETIC ENZYMES FROM OIL PALM CHROMOPLAST

by

Benjamin Lau Yii Chung

Plant fatty acid metabolism has proven to be amenable to manipulation by conventional breeding, genetic and metabolic engineering to enhance the fatty acid profile. This can be done by engineering palm fruit to synthesise more oleic acid at the expense of palmitic acid. This would produce an oil with greater perceived nutritional quality and higher market value. Although the biochemistry of fatty acid biosynthesis in plants is well described, crosstalk between transcriptional and metabolic controls in regulating fatty acid composition remains poorly understood. Our hypothesis is that phosphorylation is one of the main regulators of acetyl-CoA carboxylase, fatty acid synthase complex and stearoyl-ACP-desaturase in increasing the oleic acid level between oil palm (*Elaeis guineensis* Jacq. var. Tenera) low and high oleic acid varieties. This study utilised advanced proteomic techniques to isolate, detect and identify chromoplast-based phosphorylated proteins associated with the fatty acid biosynthesis pathway. Sub-organelle isolation using differential centrifugation enriched the chromoplast fraction that contained the fatty acid biosynthetic enzymes before their protein extraction. Gel-based and non-gel based mass spectrometry techniques were then employed to separate and improve the identification of key fatty acid biosynthetic enzymes. Protein expression was analysed using isobaric labelling strategy. Five key enzymes, namely the β-ketoacyl-ACP reductase (EC 1.1.1.100), β-hydroxyacyl-ACP dehydrogenase (EC 4.2.1.58 and 4.2.1.59), 3-enoyl-ACP reductase (EC 1.3.19), β-ketoacyl-ACP synthase (EC 2.3.1.41) and stearoyl-ACP desaturase (EC 1.14.99.6) were identified using GeLC-MS/MS strategy. An additional two subunits of acetyl-CoA carboxylase (EC 6.4.1.2) were identified from the 2DLC-

MS/MS strategy. The expression of β-hydroxyacyl-ACP dehydrogenase and β-ketoacyl-ACP synthase was up-regulated in the high oleic acid variety. In contrast, 3-enoyl-ACP reductase was down-regulated in the high oleic acid variety. The existence of other differentially regulated metabolic enzymes associated with fatty acid biosynthesis suggested that the control of fatty acid production, particularly the synthesis of oleic acid, involves more than just the main fatty acid biosynthetic enzymes. Subsequently, the role of phosphorylation in regulating these fatty acid biosynthetic enzymes was investigated using a novel combination of neutral loss-triggered MS^3 and Selected Reaction Monitoring. Acetyl-CoA carboxylase and 3-enoyl-ACP reductase were postulated to be phosphorylated in both low oleic acid and high oleic acid-producing oil palms during the fruit maturation stage of 20^{th} week after anthesis. However, other fatty acid biosynthetic enzymes from these oil palm varieties did not show any indication of phosphorylation despite the prediction of phosphoserine-containing peptides. The location of the phosphorylated serine residues in the protein domains of acetyl-CoA carboxylase and 3-enoyl-ACP reductase suggested that phosphorylation could have regulated their enzyme activities. This study has produced a robust method to capture and identify chromoplast-based enzymes that are related to plant fatty acid biosynthesis. The differences in their protein expression levels suggested that fatty acid biosynthetic enzymes were differentially regulated and phosphorylation might be involved in this regulation, at least in the enzyme activity. The outcomes reported in this thesis have significantly improved the knowledge of the possible regulation mechanisms in plant fatty acid biosynthesis. The logical extension of this work in future efforts will be to determine the biological significance of this differential protein expression and to understand the exact role of phosphorylation in the regulation of these enzymes.

Keywords

Oil palm, chromoplast, fatty acid biosynthesis, oleic acid, proteomic, phosphorylation, post-translational modification, electrophoresis, liquid chromatography, iTRAQ, neutral loss-triggered MS^3, Selected Reaction Monitoring

Acknowledgments

Undertaking a PhD is a huge personal challenge. In truth, it would have been even harder for me to accomplish this endeavour without help from these following people. To begin with, this research project would not have materialised without the scholarship provided by the Malaysian Palm Oil Board (MPOB), and I am therefore very much indebted. I am grateful to Dr. Mohd Din Amiruddin and Fadila Ahmad Malike, for the provision of the oil palm fruit materials as well as my colleagues back in the MPOB Metabolic Laboratory, for assisting in sample collection and processing, notably Hasliza Hassan and Abdul Rahim Mohamad. My gratitude also goes to Dr. Umi Salamah Ramli for permitting me to take leave from my position in MPOB to undertake this PhD programme and supporting me throughout the programme.

A special thanks to the Protein and Biomaterials Team at AgResearch Lincoln for giving me the opportunity to carry out my PhD work here for the past three years. I acknowledge the efforts of my supervisors; Associate Professor James Morton, Dr. Stefan Clerens, Dr. Santanu Deb-Choudhury and Dr. Jolon Dyer for critical reading of the thesis chapters and manuscripts; and to Karl Gately of Lincoln University for the instrumentation and chemical purchases. My appreciation also goes to Dr. Mark Condina and Dr. Jaran Jai-nhuknan from Bruker Daltonics for their technical assistance.

I thank the fellow lab team mates in the Proteins and Biomaterials Team and the Information Services Technology team; in particularly Ancy Thomas, Dr. Stephen Haines and Dr. Jeff Plowman for their assistances during my time in AgResearch. My gratitude is also dedicated to Cheng Vern Jou and Dr. Jessica Gathercole for their friendship and support. I am certain that all the entertaining social events and lunch parties that we had throughout these years will not be out of my mind for a long while.

Most importantly, to my loving partner, Christina, no words can describe the gratitude I have for your endless love, sacrifices and supports that has allowed me to complete such a great endeavour. It would be near impossible to endure the solitude without you.

"The secret of success is making the right decisions and to make the right decisions, you will need experience. To get the experience, you will have to make the wrong decisions"

Table of contents

List of tables

List of figures

Abbreviations used

1D SDS-PAGE	one-dimensional sodium dodecyl sulphate-polyacrylamide gel electrophoresis
2DE	two-dimensional gel electrophoresis
2DLC	two-dimensional liquid chromatography
ACCase	acetyl-CoA carboxylase
ACP	acyl carrier protein
CDE	cell wall digestive enzymes
CID	collision-induced dissociation
CoA	coenzyme A
GeLC-MS/MS	gel electrophoresis-liquid chromatography-tandem mass spectrometry
HO	high oleic acid variety
iTRAQ	isobaric tags for relative and absolute quantitation
IPG	immobilised pH gradient
KAS	ketoacyl-ACP synthase
kDA	kiloDalton
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LO	low oleic acid variety
MudPIT	multidimensional protein identification technology
m/z	mass-to-charge ratio
NCBI	National Centre for Biotechnology Information
$NLMS^3$	neutral loss-triggered MS/MS/MS
nm	nanometres

Abbreviations used

PTM	post-translational modification
SAD	stearoyl-ACP desaturase
SCX	strong cation exchange
SRM	Selected Reaction Monitoring
TCA	trichloroacetic acid
WAA	week after anthesis

1.0 Introduction

1.1 Research rationale

For many years, oil from African oil palm or the *Elaeis guineensis* Jacques has been sourced for food and energy. This tree of life grows in the wild in West Africa and Equatorial Africa and was first exploited by the ancient Egyptians about 5000 years ago (Obahiagbon, 2012). Through the centuries, the people of Africa have used palm oil in their sustenance and economy. The oil is regarded as a versatile and nutritious edible plant oil; free of trans fat and rich in vitamins and antioxidants (Hayes & Pronczuk, 2010; Obahiagbon, 2012; Rice & Burns, 2010; Sen, Rink, & Khanna, 2010; Sundram, 2000; Sundram, Sambanthamurthi, & Tan, 2003). Oil palm is commercially planted throughout Malaysia, covering 5.23 million hectares as of December 2013, as well as in Indonesia and other tropical countries (Division, n.d.). Currently, this oil crop is the world highest oil bearing crop with about 4 tonnes being produced annually per hectare (Basiron, 2012; Kirkland, 2011; Singh et al., 2013). By November 2014, Malaysia had contributed 40.4% of the global oils and fats export trade, supplying around 45 million metric tonnes for the global consumption of vegetable oils (http://www.fas.usda.gov/commodities/oilseeds/other-oilseeds). Palm oil plays a crucial role in supporting Malaysia's economy. It dominates the local edible oil market, and is the fundamental raw material for the oleochemical and food industries (Soyatech, n.d.). Overall, the palm oil industry contributed about 13.7% of Malaysia's Gross Domestic Product in 2011 (Sun, 2012; Szulczyk, 2013). With an expected exportable surplus of 20 million tonnes of palm oil in 2035 (Yean & ZhiDong, 2012), Malaysia has the potential to remain a formidable competitor in the world vegetable oil and biofuel market.

Strategically designed and exhaustive research and development programmes in diverse areas including crop physiology, tissue culture and biotechnology, have been applied by the Malaysian Palm Oil Board to ensure that oil palm stays competitive and sustainable in the vegetable oil and fat global market. Diversification into innovative high-value products derived from oil palm, spurred by advances in genetic engineering techniques, offers bright prospects for raising the economic value of the oil palm. Industry prognoses suggest that the palm oil industry's huge contributions toward Malaysia's Gross Domestic Product can be increased by

7.6% by the year 2020 if the value-added returns from efficiency and innovation materialise. Palm oil exports are also expected to increase by 7% per annum to approximately Malaysian Ringgits (MYR) 84.6 billion (USD 270.72 billion) by 2020 (National Economic Advisory Council, 2010). The export figure will rise even further if novel palm oil-based products and services can be effectively commercialised and marketed.

Advances in genetic modifications are being propelled by significant improvements in DNA sequencing efficiency as well as the increase in rate of synthesis and larger *de novo* assemblies of DNA. Such progress in genome engineering and design has significantly enhanced the tools available for extensive and intentional manipulation of plant genes to address specific environmental problems such as drought and economic goals such as increasing the value of palm oil (Carr & Church, 2009; Lawlor, 2012; Mittler & Blumwald, 2010). Unique genetic variations associated with particular phenotypic traits, for example shell thickness, that have been discovered in oil palm genome sequencing projects offer the possibility to improve the efficiency and productivity of the oil palm industry. The oil palm genome sequencing projects have the potential to drive the industry towards a more sustainable economic model essentially based on renewable resources, also known as the 'bio-based' economy (Langeveld, Dixon, & Jaworski, 2010). They are also support the Malaysia government's goals to achieve higher yields of fresh fruit bunches at 35 tonnes per hectare (from the current 20.18 tonnes per hectare) and oil extraction rate of 25% per hectare (from 20% currently) annually by 2020. Production of unsaturated high oleic acid palms for the feedstock and oil markets is the first direct contribution of genetic engineering in this field (Yunus & Parveez, 2008). Crude palm oil has the potential to escalate to about MYR 500 (USD 1,500) per hectare per year provided the content of oleic acid exceeds 65% (Wahid, Abdullah, & Henson, 2004). Existing palm oil from mesocarp contains 53% saturated fatty acids of which 44% is palmitic acid. Less than 33% of the total palm oil content is the unsaturated oleic acid (Sundram, 2000; Sundram et al., 2003). Reducing the saturated fatty acid content in palm oil would significantly improve nutritional value of palm oil.

Genetic manipulation of fatty acid biosynthesis genes in mesocarp played a vital role in producing high oleate palms, in addition to the current selective breeding programme to select high oleic acid producing palms (as used in this thesis) (Murphy, 2005, 2006). Alteration of the fatty acid yields is also achievable by manipulating substrates and intermediates from the biosynthesis of fatty acids. However, it is necessary to comprehend the fundamentals of

fatty acid biosynthesis before embarking on efforts to raise the oleic acid content through metabolic manipulation of these substrates and intermediates. Although findings on the pathway of fatty acid biosynthesis in plants are well documented, it has not been possible to decipher the exact regulators of the fatty acid biosynthesis. Previous research has shown that the regulation of fatty acid biosynthetic pathway involves many proteins. Acetyl-CoA carboxylase, the catalyst of the first reaction in the fatty acid biosynthesis, has long been considered to determine the flux of fatty acid biosynthesis in both plants and animals (Baud & Lepiniec, 2009; Harwood, 1996; Joyard et al., 2010). In addition, levels of specific fatty acids are controlled by fatty acid synthase and stearoyl-ACP-desaturase at both transcriptional and post-translational levels (Kachroo et al., 2007; Ramli, Baker, Quant, & Harwood, 2002; Sambanthamurthi, Abrizah, & Ramli, 1999). Therefore, a thorough insight into the fatty acid biosynthetic pathways involving these three key enzymes has the potential to enhance the efforts of the genetic engineering programme or even to circumvent genetic manipulation in producing high oleate palm as well as palm oil with targeted compositions.

The hypothesis underlying this research is that phosphorylation is one of the key regulators not only for acetyl-CoA carboxylase and fatty acid synthase in controlling overall fatty acid synthesis, but also for stearoyl-ACP desaturase in determining the oleic acid composition between different oil palm varieties. The thesis focusses on investigating the role of phosphorylation as the main post-translational modification of enzymes involved in oleic acid biosynthesis in order to gain a better insight into its mechanism. Mass spectrometry-based proteomics approaches were utilised to characterise the protein expression and phosphorylation that may regulate the rate of oleic acid production through the actions of the fatty acid biosynthetic enzymes. Robust protocols had been developed to isolate the fatty acid factory, that is, the chromoplast and improve the extraction of its proteins. A combination of gel-based and non-gel based proteomic techniques were then used to separate, detect and quantify the differentially expressed fatty acid biosynthetic enzymes. Characterisation of phosphorylated enzymes related to fatty acid biosynthesis was carried out using a novel combination of neutral loss-triggered MS^3 and Selected Reaction Monitoring. Possible sites of phosphorylation on serine residue in each of the fatty acid biosynthetic enzymes were also predicted and were found to have different phosphorylation states. It is anticipated that the procedures developed in this study will also be hugely advantageous for future undertakings in oil palm proteomic research.

1.2 Purpose and contribution of findings

The main research aims were;

(i) The development of protocols for chromoplast isolation and subsequent protein extraction from low and high oleic acid palm varieties

(ii) The development of proteomic-based strategies for detection, enrichment and quantification of differentially regulated enzymes associated with fatty acid biosynthesis

(ii) Characterisation of the phosphorylation in fatty acid biosynthetic enzymes with mass spectrometry-based techniques

The work done in this thesis provides a substantial contribution to the understanding of the regulation of oleic acid production through phosphorylation of some key enzymes in the fatty acid biosynthetic pathway. While conventional breeding programmes have significantly contributed in producing high yielding oil palm planting materials, the progress has been slow due to limited genetic diversity and the need for tedious and time-consuming procedures. Thus, information currently available from genetic engineering efforts and knowledge generated in this study are of both fundamental interest and essential for optimal commercial production of novel planting materials with desired fatty acid compositions, such as the high oleate palm. Furthermore, the robust techniques acquired have significant potential to be employed in investigations on various oil palm organelles using proteomic approaches. As part of the long term gains, mining of protein based biomarkers to indicate type and amount of fatty acids, for instance, will almost certainly enhance the oil crop improvement programme by aiding the plant breeders to screen oil palm germplasms for specific fatty acid compositions. Additionally, the differences in the abundances of key fatty acid biosynthetic enzymes will also help to determine the fatty acid compositions in oil palm.

2.0 Literature Review

2.1 The African oil palm tree (*Elaeis guineensis* Jacques)

The oil palm comprises two species of the Arecaceae (formerly known as the Palmae), or palm family. The order of this palm family is not well ascertained. For instance, Cronquist (Cronquist, 1981) consigned the Arecaceae in the order Arecales while Bentham and Hooker's *Genera plantarum* put the palms with the Flagellariaceae and Juncaceae in the series Calycinae. Another system put palm tree under the order Principes (Engler and Prantl's system). The family Arecaceae, to which the genus *Elaeis* belongs, is considered to be the first monocotyledon to have branched out from the primitive dicotyledon stock, and therefore, the progenitor of all monocotyledons. The classification of oil palm in the Arecaceae family was based on palms introduced into Martinique and the oil palm receiving its botanical name from Jacquin (Jacquin, 1763). The name of the genus *Elaeis*, is derived from a Greek word, *elaion,* which means 'oil'. The specific name *guineensis* shows that the first specimen described was collected in Guinea Coast, West Africa.

The Portuguese introduced oil palm to Brazil and other tropical countries in the 16th century. Three hundred years later, the Dutch brought the oil palm seeds from West Africa to be planted in Bogor, Indonesia in 1848. The seeds were from the *dura* variety and the progeny from the four seedlings were then planted in Deli as ornamentals. Thus, they are known as the Deli *dura.* In 1875, Botanical Garden in Singapore received its first oil palm and three years afterwards, the oil palm arrived in Malaysia. The oil palm was initially planted as an ornamental plant. Eventually in 1917 commercial oil palm plantations were established (Sundram, 2000). The *pisifera* palms were introduced to Malaysia by Algemene Vereniging van Rubberplanters ter Oostkust van Sumatra (AVROS) in 1957. Oil palm planting materials after 1961 were largely of the cross between Deli *dura* and *pisifera*, known as the *E. guineensis* var. tenera. *Tenera* is the only commercial variety in Malaysia today (Seng, 1987) despite the existence of another species of oil palm in Malaysia, the *E. oleifera*. *E. oleifera* from South America is only valued for its potential in transferring useful characteristics to interspecific hybrids with *E. guineensis*, even though it contains more than 69% oleic acid in comparison to the commercially cultivated *E. guineensis* var. tenera which has about 30% of oleic acid content.

Dura (D), *pisifera* (P) and *tenera* (DxP) are the three varieties of *E. guineensis* (Figure 2.1). The thick shelled homozygous dominant (*Sh+Sh+*) *dura* crossed to the shell-less homozygous recessive (*Sh-Sh-*) *pisifera*, resulted in the thin shelled heterozygous (*Sh+Sh-*) *tenera*. The DxP hybrid is used as planting material due to it having a higher proportion of oil-bearing mesocarp and consequent higher oil yield than *dura* (Hai, 2002). The shell-less *pisifera* has the highest mesocarp content but is not suitable as a planting material for commercialisation as it is female sterile. Thus, *pisifera* bunches would never develop to maturity. There are different types of *pisifera* and *dura* (Lim, Teo, Rao, & Chia, 2003; Teo, Rao, Chia, & Lim, 2004). These varieties, known as the parent planting materials, provide the varieties of *tenera* planting materials due to their different origins, genetic content and selection criteria (for example, the presence or absence of endocarp) (Breure, 2006; Esnan, Zakaria, & Wahid, 2004). In addition to these varieties, the Malaysian Palm Oil Board (MPOB) had also developed thirteen novel variations of planting material, designated as the PORIM Series (PS), through its breeding and selection programme. Each of these planting materials has one or more unique traits (Table 2.1). The *E. guineensis* varieties used in this study were obtained from the PS12 breeding population, which consists of 15 oil palm trees from the *tenera* variety with oleic acid content exceeding 48%. Normal oleic acid content in the standard DxP hybrids is around 30% of oleic acid. The standard DxP hybrid palms were obtained from the oil palm germplasm currently managed by MPOB. Due to the narrow genetic base of oil palm breeding populations, consisting of merely four Bogor *dura* and a handful of *pisifera* palms and also the continuing threats of genetic erosion, MPOB has actively searched for oil palm germplasm centred on its origin in Africa and Latin America (Rajahnaidu, 1994). *E. guineensis* has been sourced from Nigeria, Cameroon, Zaire (Congo DR), Tanzania, Madagascar, Angola, Senegal, Gambia, Sierra Leone, Guinea and Ghana.

The palm fruit (Figure 2.2) is a central hard-shelled nut (kernel) surrounded by an outer shell (endocarp). The soft oily mesocarp which contains the commercial palm oil lies in between the endocarp and exocarp. The nut or oil palm seed contains one to three kernels, from which a different type of oil, the palm kernel oil, is extracted (Figure 2.3) (Corley & Tinker, 2003). Development of oil palm fruit begins around two weeks after anthesis (WAA). Anthesis in this context means the period when the first functional flower from oil palm female inflorescence is apparent. The endosperm starts its oil accumulation about 12^{th} WAA and nears completion

Dura **Pisifera** **Tenera**

Figure 2.1. The different varieties of oil palm *E. guineensis.*

Table 2.1. PORIM Series with their properties (Swierszcz, Skurski, & Simons, 2012).

PORIM Series	Purpose or characteristics
1	Development of high yielding and dwarf palm
2	Breeding for high unsaturated oil (indicated by high iodine value)
3	Breeding for high lauric oil
4	Breeding for high carotenoid content, *E. oleifera*
5	*Tenera* with thin shell and high vitamin E
6	*Dura* with bigger fruit size and high bunch index
7	Very high bunch index
8	Very high vitamin E content
9	*Bactris gasipaes*
10	Long bunch
11	*E. guineensis* with high carotene content
12	**High content of oleic acid**
13	Low levels of lipase activity

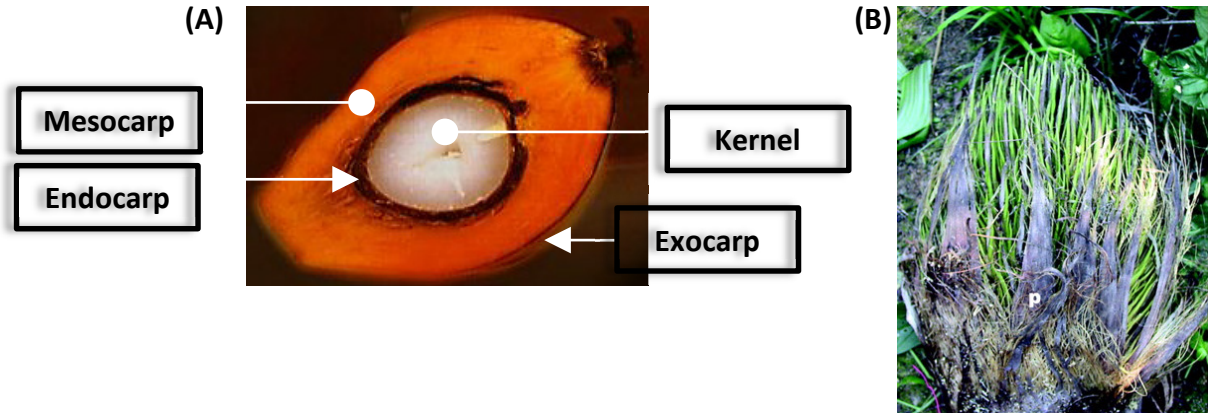


Figure 2.2. (A) Fresh oil palm fruit bunch and (B) the female inflorescence (Adam et al., 2005).

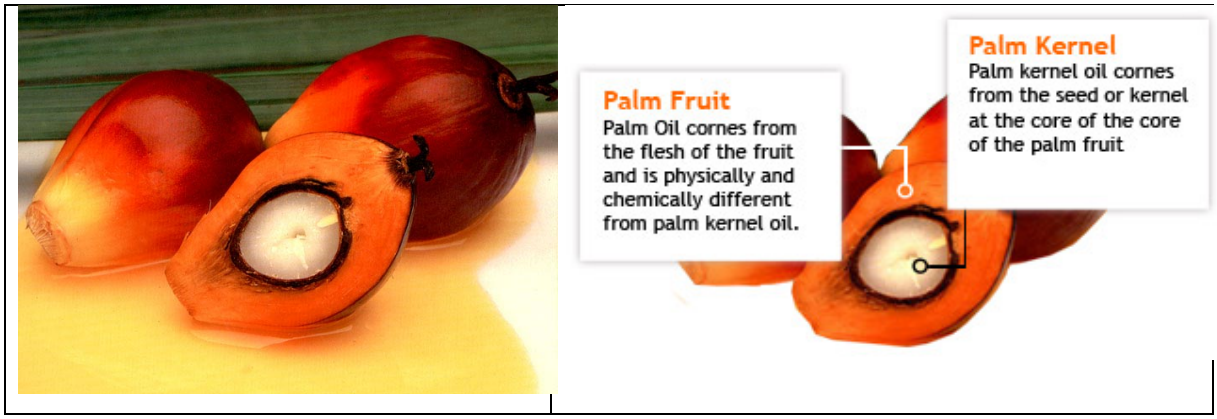


Figure 2.3. Descriptions of oil palm fruit (http://www.palmoilprices.net/main.php?mode=tplmain&c=palmoil).

by 16th WAA. Physiological changes that can be observed at this stage include the gradual hardening of the endosperm and endocarp. Subsequently, the endocarp become the hard shell and encloses the rigid white endosperm, also known as the kernel. Concurrently, mesocarp starts its oil deposition around the time when the oil deposition in endosperm nears its end, which is 15th WAA. The deposition process continues until the ripening of oil palm fruit occurs around 20th WAA. Due to slight differences in pollination time of the flower, fruits throughout the bunch do not mature concurrently. Primarily, fruits at the end of each spikelet ripen followed by those at the base of the bunch. The size and colour of the fruits within the bunch also vary, with those on the outside usually larger and deep orange in colour when ripe. The inner fruits are smaller and paler.

Oil palm has economic value up to 20-25 years with an average annual oil production of 3.5-5 tonnes per hectare. A tree bears about 8-12 fruit bunches weighing 15-25 kg per bunch each year. Each of these bunches can produce up to 3000 fruits (American Palm Oil Council, n.d.).

The oil palm fruit is the only commercial oil-producing plant that yields two different types of oil; palm oil from the mesocarps and palm kernel oil from the kernel. Through the refining, bleaching and deodorisation processes, crude palm oil is obtained from the mesocarp. The palm oil is generally orange or red in colour or the more commonly recognized bright golden oil. In nature, palm oil is partially solid. The fractionation of palm oil into liquid olein and solid stearin increased its usefulness in food applications. Olein is usually used as cooking or frying oil while stearin finds its many applications in solid fat formulations and food processing. Meanwhile, the palm kernel oil, extracted from the kernel, is used in specialty fats production for various food products and is also an important raw material for the oleochemicals industry.

2.2 Palm oil value and sustainability

Vegetable oils had a global production of 186.4 million tonnes in 2012 (Figure 2.4), thus cementing their position as a major agricultural commodity. Palm oil is amongst the 17 major vegetable oils and fats produced by oil palm, soybean, rapeseed, sunflower and other crops with a global production of 52 million tonnes (32%) in 2012. Palm oil was the most commonly consumed vegetable oils in 2012. In that year, the palm oil was consumed by around three billion people in 150 countries. India is by far the largest consumer of vegetable oils and fats, followed by China, European Union, the United States of America and Malaysia. The worldwide consumption of palm oil was anticipated to be at 52 million tonnes by the first quarter of 2013 (http://www.fas.usda.gov/).

Malaysia produces 45% of the world's palm oil, and is the second largest palm oil producer after Indonesia (American Palm Oil Council, n.d.). In 2013, Malaysia produced 19 million tonnes of crude palm oil, from slightly over 5 million hectares of plantations. Total exports of oil palm products, consisting of palm oil, palm kernel oil, palm kernel cake, oleochemicals, biodiesel and refined products, increased to 25.7 million tonnes in 2013, an increment of around 1.1 million tonnes from 2012. This has contributed to approximately MYR 61.36 billion (USD 17.18 billion) of Malaysia's total export earnings in 2013. The oil palm industry contributed MYR 71.45 billion (USD 20.01 billion) or 14.1% more in total export revenues in 2012 than 2013 from the export of palm oil and derived products. This was due to the lower export prices of all palm products (Division, n.d.). Accordingly, the oil palm industry became the largest contributor to external trade after electrical and electronic products as well as the most important foreign exchange earner for Malaysia (Malaysia, 2011, February 2). As the

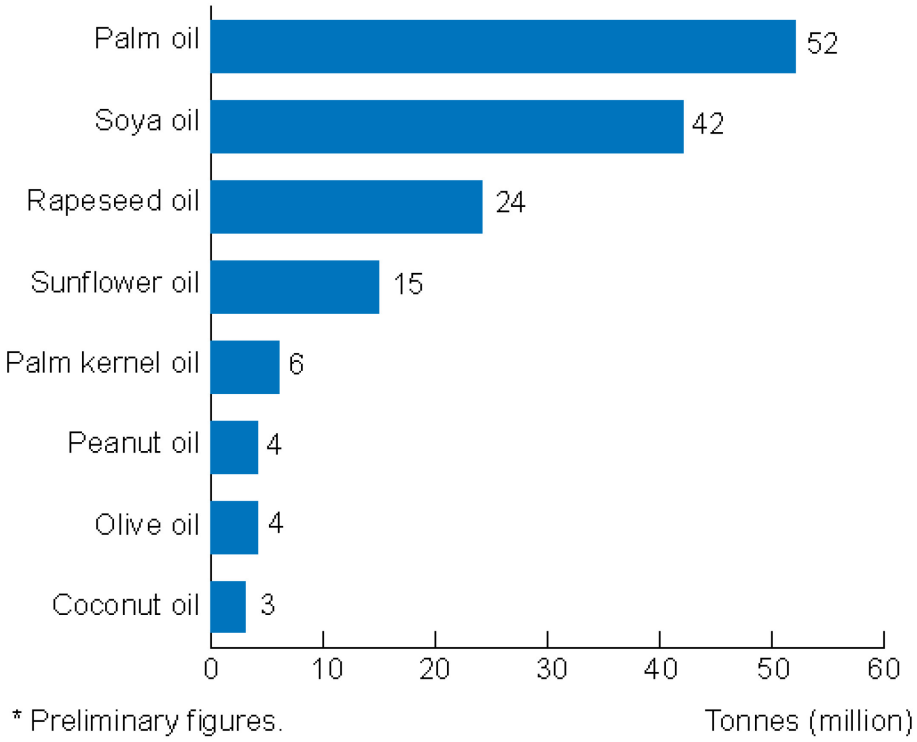


Figure 2.4. Global production of vegetable oils and fats in 2012 (AAK Report 2012).

second major producer and exporter of palm oil and its products, Malaysia has a crucial role in delivering the world's oils and fats.

The fatty acid composition of vegetable oils influences both their physical properties and nutritional characteristics. This fact and the huge market size for oil palm has generated substantial interest in altering plant fatty acid production for both food and non-food uses, such as a renewable industrial feedstock (Burton, Miller, Vick, Scarth, & Holbrook, 2004; Ohlrogge, 1994; Ohlrogge, Browse, & Somerville, 1991; Weselake et al., 2008). Recent initiatives to manipulate the quality and quantity of plant fatty acids have escalated due to growing concerns about generating sustainable agricultural output and novel strategies to make use of plants as a source of raw materials for industrial products, such as biodiesel (Murphy, 2005). Palm oil, which is predominantly oleic, palmitic and lauric acids, is considered a high quality raw material for biodiesel and cooking oil (Bamgboye & Hansen, 2008). The cetane index (used to substitute the cetane number of diesel fuel) is influenced by the fatty acid composition. A very strong positive relationship exists between the calculated cetane index of crude oil and oleic acid concentration (Duffield, Shapouri, Graboski, McCormick, & WIlson, 1998).

Previous studies have reported various approaches in investigating the regulation of lipid synthesis in oil crops. Quantitative trait loci associated with seed oil and fatty acid composition have been examined in important crops like *Brassica napus* (Burns et al., 2003) and the model plant *Arabidopsis* (Hobbs, Flintham, & Hills, 2004). In addition, comparison of wild-type with the low-lipid wrinkled 1 mutation of *Arabidopsis* using cDNA microarrays studied more than 3500 genes to reveal the gene that may exert significant control (Ruuska, Girke, Benning, & Ohlrogge, 2002). A more comprehensive assessment of fatty acid biosynthesis regulation using metabolic control analysis has also been reported (Fell, 1992). Two oil crops, the olive and oil palm, (Ramli et al., 2002) had employed one of the approaches known as the top-down control analysis (Brand, 1996). Such experiments allow identification of a metabolic pathway that is more amenable to genetic modification and also generates in depth knowledge of steps that will exert significant control and may be worth genetically manipulated.

Plant lipid metabolism has proven particularly amenable to modifications both by classical breeding approaches and more recently by genetic manipulations such as hairpin RNA-mediated gene silencing (Liu, Singh, & Green, 2002a; Schultz & Ohlrogge, 2001; Thelen & Ohlrogge, 2002a; Voelker & Kinney, 2001). Selective breeding exploiting natural variants or induced mutations had been utilised to develop a range of improved oils in the major temperate oilseed crops. This has led to high-stearic soybean (*Glycine max*) (Graef, Miller, Fehr, & Hammond, 1985), high-stearic sunflowers (*Helianthus annuus* L.) (Osorio, Fernandez-Martinez, Mancha, & Garces, 1995), high-oleic rapeseed (*Brassica napus*) (Auld, Heikkinen, Erickson, Sernyk, & Romero, 1992) and high-oleic peanut (*Arachis hypogaea)* (Norden, Gorbet, Knauft, & Young, 1987). The Malaysian Palm Oil Board has developed new breeding materials with high content of oleic acid through traditional breeding. The genetic engineering programme in the Malaysian Palm Oil Board has also focussed on engineering higher oleic acid content in oil palm at the expense of palmitic acid.

Important developments have been made in the area of oil palm genetic engineering within the last 10 years. Nearly all the genes and promoters essential in fatty acid biosynthetic pathway modifications have been isolated and studied. A part of the gene encoding heteromeric acetyl-CoA carboxylase subunit biotin carboxylase, the enzyme that catalysed the initial committed step in lipid biosynthesis and is thought to be an important fatty acids flux-controlling enzyme, has been cloned from oil palm fruit mesocarp (Budiani, Santoso, Aswidinnoor, & Suwanto, 2008; Wan Omar et al., 2008). β-ketoacyl ACP synthase II and acyl-

ACP thioesterase are the other key candidates for genetic manipulation to generate high oleate and stearate palms. That strategy was to overexpress β-ketoacyl ACP synthase II, the enzyme that catalysed the addition of two carbons to palmitoyl-ACP (C16:0) to produce an 18-carbons fatty acid. At the same time, an antisense strategy had been employed to inhibit the synthesis of palmitoyl-ACP thioesterase. Therefore, the production of fatty acids was channelled in favour of oleic acid at the expense of palmitic acid. Gene constructs containing mesocarp-specific promoter for oil palm transformation had been generated using full-length cDNA clones. It is also crucial to down-regulate oleoyl-CoA desaturase to prevent overproduction of oleic acid (C18:1) from stearic acid (C18:0) fatty acid. For that reason, the partial cDNA clone encoding the gene had been isolated (Shahwan, 2006). A full length cDNA clone for another key enzyme in fatty acid biosynthesis, the stearoyl-ACP desaturase had been isolated as well and its antisense construct was introduced into oil palm to increase its stearic acid composition (Akmar et al.; Shah, Rashid, & San, 2000). Although there have been other notable technical engineering successes (Liu et al., 2002a), food safety concerns, such as the antibiotic resistance markers used in the production of the original transgenic crops, has held up efforts to release transgenic oils to the market (Ohlrogge, Mhaske, Beisson, & Ruuska, 2004).

Industrially, oleic acid will be beneficial in generating chemical derivatives that can act as alternatives to petrochemical feedstock. The high melting point of the oleic acid will also enable palm oil to enter the liquid or salad oil market. The programme thus has the potential to change palm oil from a commodity oil to a high value oil with industrial applications in addition to food applications. Oleic acid profiles can also function as markers to assist in selective breeding programmes. Although traditional breeding offers one alternative, the more direct approach of manipulating the fatty acid composition through genetic manipulation, has been limited due to insufficient knowledge regarding the lipid metabolism regulation as a whole (Ohlrogge & Jaworski, 1997). An understanding of the biochemical factors and metabolic pathways regulating oleic acid composition is therefore necessary to exploit these opportunities.

2.3 The origin of plant plastids

Plastids are found in almost all plant cells, algal cells, several taxons of marine molluscs and at least one phylum of parasitic protists (Wise, 2006). Plastid family members (Figure 2.5) play

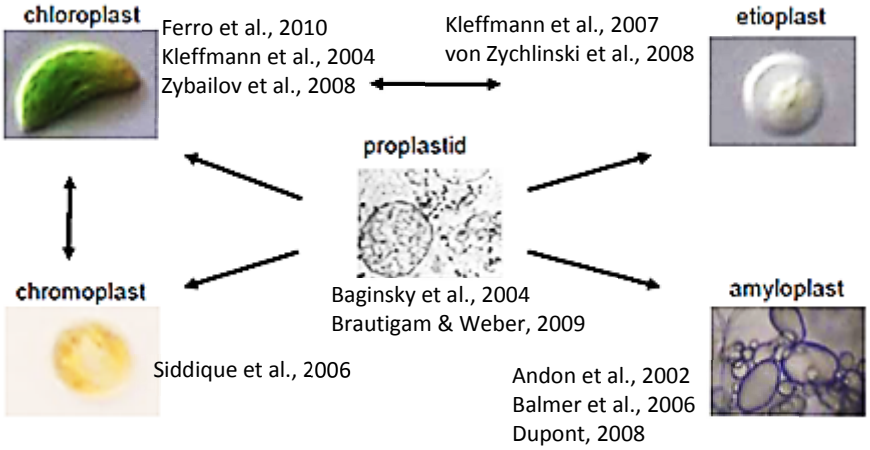


Figure 2.5. Types of plastid in plant and related studies (Andon et al., 2002; Baginsky, Siddique, & Gruissem, 2004; Balmer, Vensel, DuPont, Buchanan, & Hurkman, 2006; Bräutigam & Weber, 2009; Dupont, 2008; Ferro et al., 2010; Kleffmann et al., 2004; Kleffmann et al., 2007; Siddique, Grossmann, Gruissem, & Baginsky, 2006; von Zychlinski et al., 2005; Zybailov et al., 2008).

critical roles in plant metabolism through processes such as photosynthesis, lipid and amino acid synthesis, starch and oil storage, fruit and flower coloration, and nitrogen and sulphur assimilation, as well as synthesis of secondary metabolites (van Wijk & Baginsky, 2011). Three evolutionary lines of plastids are known to exist in the glaucophytes (small group of freshwater microscopic algae), the red lineage and the green lineage (Keeling, 2004). The plastids of the Kingdom Plantae form the green lineage.

Proplastids occur in meristematic tissue, in addition to embryonic tissues and cells in tissue culture and are the developmental precursors to all types of plastid (Pyke, 1999; Thomson & Whatley, 1980). In plants, differentiation of proplastids to their final forms depends on their environment and location. For instance, if the development of the meristematic or embryonic tissues continue in the dark or in extremely low light environment, plastid development will be arrested at etioplast stage (Domanskii et al., 2003; Harsanyi, Boddi, Boka, & Gaborjanyi, 2005). Etioplasts, or chloroplasts that have not been subjected to light, are normally found in the shoot tissues. However, etioplasts will proliferate, expand, become green and develop into chloroplasts if these shoot tissues are exposed to light during their development (Krishna, Joshi, Vani, & Mohanty, 1999; Ljubičić, Wrischer, & Ljubešić, 1998). Meanwhile, leucoplasts are non-pigmented plastids, of which amyloplasts (starch storing), elaioplasts (oil and lipid storing) and proteinoplasts (protein storing) are generally known (Negm, Cornel, & Plaxton,

1995). The structure and function of amyloplasts and elaioplasts have been extensively reported but there is limited literature on proteinoplasts. Proplastids in root tissues typically develop into colourless starch-containing amyloplast (James, Denyer, & Myers, 2003; Yu, He Mu, Mu-Forster, & Wasserman, 1998). Oil containing leucoplasts are known as elaioplasts and numerous oil droplets that contain proteins such as plastoglobule-associated proteins (PAP) dominate the inner structure of elaioplasts (Hernández-Pinzón, Ross, Barnes, Damant, & Murphy, 1999; Wu et al., 1997). Oleosomes are discrete non-plastid organelles derived from the rough endoplasmic reticulum that store oil (Yatsu, Jacks, & Hensarling, 1971). Light and electron microscope observations by Thomson and Whatley (Thomson & Whatley, 1980) indicated that proteinoplasts are plastids that contain mainly large and visible protein inclusions. Nonetheless, there is no conclusive evidence that all proteinoplasts function only as protein storage reservoirs. Chromoplasts are the brightly coloured plastids in fruits, flowers, leaves and even several roots. They contain a substantial amount of red, yellow or orange carotenoids (Juneau, Le Lay, Böddi, Samson, & Popovic, 2002). In general, chromoplasts are assumed to be involved in attracting biotic agents (vectors) including the fruit dispensing animals. Chromoplasts usually develop from chloroplasts in ripened fruit (Bouvier, Backhaus, & Camara, 1998) but they can also develop directly from proplastids in other tissues (Ljubesic, 1972).

This study is interested in the chromoplasts of the oil palm fruit as fatty acids are synthesised in the chromoplast stroma before they are exported to the endoplasmic reticulum as acyl-CoA esters (Bouvier et al., 1998). The chloroplasts and chromoplasts are defined collectively as plastids in this thesis.

2.4 Plant fatty acid biosynthesis

The biochemical reactions involved in fatty acid biosynthesis by chloroplasts have been extensively studied. The pathway of fatty acid biosynthesis (Figure 2.6) is essentially identical in all plants, and the first reaction of fatty acid biosynthesis occurs in the chloroplast stroma (Awai et al., 2001; Benning, 2008; Benning, 2009a; Harwood, 1996; Heinz & Roughan, 1983; Joyard & Douce, 1977; Joyard et al., 2010; Ohlrogge & Browse, 1995; Ohlrogge & Jaworski, 1997; Roughan & Slack, 1982; Slabas et al., 1994; Slabas & Fawcett, 1992; Stumpf, 1969). All carbon atoms in fatty acids synthesised in chloroplast originate from acetyl-CoA. Pyruvate is

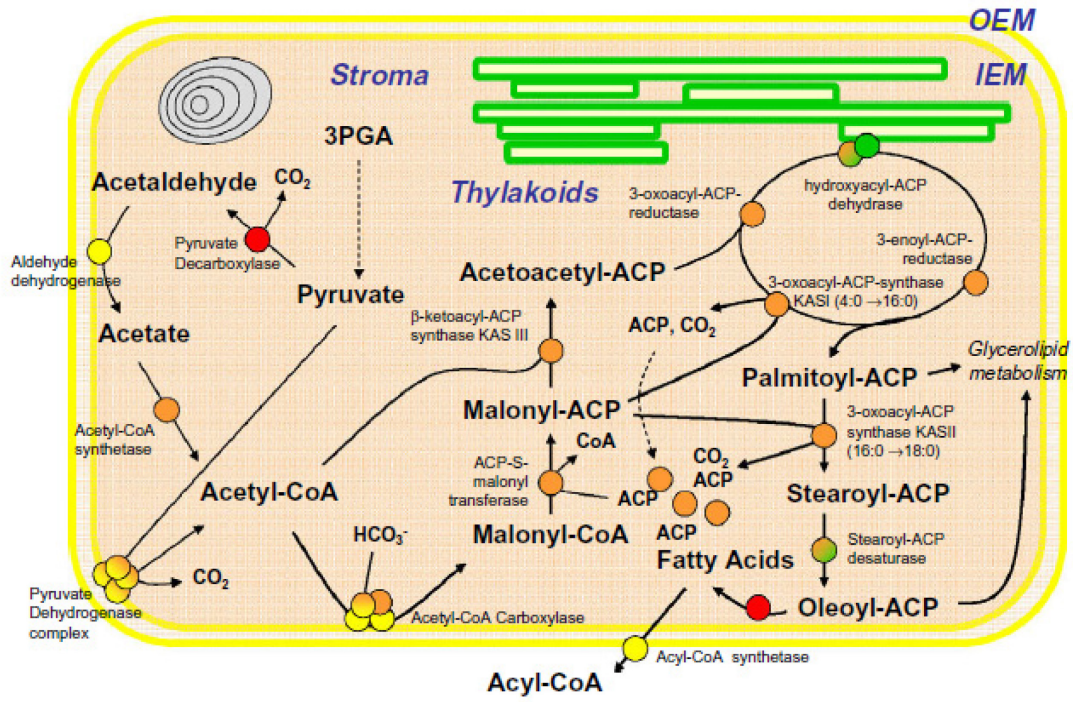


Figure 2.6. Fatty acid biosynthesis in plant. ACP, acyl-carrier protein; KAS, ketoacyl ACP synthase; 3PGA, 3-phosphoglycerate; IEM and OEM, inner and outer envelope membrane (Joyard et al., 2010).

generally considered to be the precursor for acetyl-CoA via the $C_2 \rightarrow C_3$ pathway (Bao, Focke, Pollard, & Ohlrogge, 2000; Givan, 1983; Joyard et al., 2010). 16- and 18-carbon fatty acids are formed by addition of two carbon units derived from malonyl-CoA, which is synthesised from acetyl-coA and carbon dioxide by acetyl-coA carboxylase (ACCase), in an ATP-dependent manner. Subsequently, a complex of discrete enzymes, known as fatty acid synthase (FAS) catalyses the essential carbon chain extension processes.

The production of fatty acids needs the presence of acyl-carrier protein for all catalytic activities. ACCase activity and protein levels have been shown to correlate with the amount of oil synthesised in developing seeds (Kang, Ridout, Morgan, & Rawsthorne, 1994; Roesler, Savage, Shintani, Shorrosh, & Ohlrogge, 1996; Turnham & Northcote, 1983). Two different forms of ACCase are present in all plants: the heteromeric/prokaryotic form (650-700 kDa) and a homodimeric/multifunctional form (250 kDa) (Cronan Jr & Waldrop, 2002; Harwood, 1996; Sasaki & Nagano, 2004b). However, only the heteromeric/prokaryotic form has been reported to exist in the plastids (Harwood, 1996) and catalyse fatty acid biosynthesis. The plastid ACCase consists of three distinct nuclear-encoded subunits: biotin carboxyl-carrier protein (BCCP), biotin carboxylase (BC) and the alpha (α) subunit of carboxyltransferase and one chloroplast-encoded subunit: β subunit of carboxyltransferase.

The primary condensation steps involve two enzymes. Firstly, a malonyl-CoA:ACP transacylase (32.4 kDa) catalyses the transacylation of malonate from malonyl-CoA to activated holo-ACP, resulting in generation of malonyl-ACP. ACP enters the fatty acid biosynthetic pathway at this point. Then β-ketoacyl-ACP synthase (KAS) III (46 kDa) (Jones et al., 2003) starts the chain lengthening through condensation of acetyl-CoA with acetoacetyl-ACP. Malonyl-CoA:ACP transacylase and KAS III are restricted to the stroma (Peltier et al., 2006; Sambanthamurthi et al., 1999; Zybailov et al., 2008).

Acetoacetyl-ACP undergoes a sequential reduction-dehydration-reduction process to yield 4:0-ACP, the primary substrate for KAS I (45 kDa) (Mekhedov, Cahoon, & Ohlrogge, 2001). The system then undergoes seven or eight elongation cycles to produce 16- or 18-carbon fatty acids, respectively. Beisson and co-workers (Beisson et al., 2003) proposed that there are five isoforms of KAS. However, only one isoform of KAS I was identified in chloroplasts (Froehlich et al., 2003; Zybailov et al., 2008). KAS I is accountable for the condensation in each elongation cycle from 4:0-ACP up to 16:0-ACP (palmitoyl-ACP). KAS II (57.6 kDa) (Carlsson, LaBrie, Kinney, Von Wettstein-Knowles, & Browse, 2002) is involved in the last elongation steps from 14:0-ACP and 16:0-ACP to 18:0-ACP (stearoyl-ACP). Both KAS I and KAS II have been identified in chloroplasts (Peltier et al., 2006); (Zybailov et al., 2008).

Plastid fatty acid biosynthesis produces saturated acyl-ACPs with carbon length up to 18 carbon atoms (16:0, 18:0). In spite of this, in most plant tissues, more than 75% of the fatty acids are unsaturated (Ohlrogge & Browse, 1995). A stearoyl-ACP desaturase (SAD) (45.4 kDa) (Shanklin & Somerville, 1991) introduces Δ9-desaturation in plants to form the unsaturated oleic acid (18:1) (Shanklin & Cahoon, 1998). SAD is required for unsaturated fatty acid synthesis and is thought to regulate the synthesis of unsaturated fatty acids. However, Sambanthamurthi and colleagues (Sambanthamurthi et al., 1999) have described in their studies that a more active KAS II was able to convert palmitoly-ACP to stearoyl-ACP and subsequently generate oleoyl-ACP, an unsaturated fatty acid. They also found that unsaturated fatty acids levels were not elevated by increasing the activity of SAD. The *Arabidopsis* genome encodes seven plastidial SAD (Beisson et al., 2003; Kachroo et al., 2007) but only two proteins, namely the SS12 (or FAB2) and DES5, were identified in chloroplasts (Peltier et al., 2006; Zybailov et al., 2008).

After synthesis in the chloroplast via the fatty acid biosynthetic pathway, acyl-ACPs could either form glycerolipids in the plastid through the Kornberg-Pricer route or be translocated from the plastid as fatty acids in the acyl-ACP form. These fatty acids supply the extraplastidial compartments for membrane synthesis (phospholipids), storage (triacylglycerol) lipids and wax (Joyard et al., 2010; Rawsthorne, 2002). Acyl-ACP thioesterase hydrolyses acyl-ACPs to release free fatty acids and regenerate ACP (Ohlrogge & Browse, 1995). This step signifies the conclusion of a fatty acid biosynthesis cycle.

2.5 Mass spectrometry and proteomics

The term proteomics was introduced by Wilkins and colleagues (Wilkins, Pasquali, et al., 1996). They described it as a requirement in the repertoire of 'omics technologies' for identifying and understanding proteins and their functions in a cell, tissue, organ and organism. They had also introduced the term 'proteome' as the 'total protein complement of a 'genome' (Wilkins, Pasquali, et al., 1996). In other words, the term 'proteome' was coined in analogy to the term 'genome,' the entire entity of genes of a particular organism (Dreger, 2003). Proteome analyses have become more prevalent than they were a decade ago. Previously, most of the work reported depends on a fundamental technique known as two-dimensional gel electrophoresis (O'Farrell, 1975) to separate proteins. Nonetheless, in the past few years, numerous techniques have been developed and enhanced for proteomic work. These techniques have been used to investigate proteins in various organisms as well as proteomes expressed in cells, tissues, organelles or sub-cellular fractions. These techniques complement the use of two-dimensional gel electrophoresis rather than to substitute the technique (Bachi & Bonaldi, 2008; Lilley & Dupree, 2006; Mann, 2009; Oeljeklaus, Meyer, & Warscheid, 2009; Schulze & Usadel, 2010; Thelen & Peck, 2007; Wilm, 2009).

The proteome is much more dynamic than the genome and is strictly regulated by an array of mechanisms. It constantly changes depending on cell types, cycles, stimuli and other tissue or organ specific factors. There are a multitude of regulatory processes taking place simultaneously in a cell proteome. These include synthesis and degradation of proteins, protein transport, post-translation modification, exocytosis and endocytosis. The cell requires exquisite regulation of proteins to successfully orchestrate all these processes. Prominent mechanisms of protein regulation include post-translational modification by other proteins or functional groups, protein cleavage and thus activation of additional function, localisation of proteins in different cell compartments and translocation between them. All of these

contribute to the final concentration, location and activity of each protein. Altered behaviour of proteins due to regulation, or the absence of such regulation can result in a myriad of pathological conditions (Jensen, 2004; Kwon, Choi, Choi, Ahn, & Park, 2006; Larsen, Trelle, Thingholm, & Jensen, 2006; Mann & Jensen, 2003; Seo & Lee, 2004; Ytterberg & Jensen, 2010).

Characterisation of the proteome of several organisms has revealed massive amounts of information pertaining to cell development processes (de Godoy et al., 2008). Subsequent elucidation of the protein complexes involved in those processes has revealed that various protein-protein interaction mechanisms were involved (de Godoy et al., 2008; Gingras, Aebersold, & Raught, 2005; Selbach & Mann, 2006; Sharma et al., 2009; Wessels et al., 2009). These studies revealed that cell development processes are much more distinct, dynamic and intricate than had been initially thought. This hugely dynamic regulation presents technological challenges for the proteomic researcher, and proteomic experiments need to be interpreted in the context of the exact cell conditions.

Up until 25 years ago, Edman degradation (Edman, 1964) was the only recognised technique to reliably determine the amino acid sequence of a protein. However, one of the drawbacks is that this method required a free N-terminal end while the majority of proteins had their N-terminus blocked by an acetyl or acyl groups. Mass spectrometry is an analytical chemistry technique to measure the mass-to-charge ratio (*m/z*) of a charged peptide. The introduction of mass spectrometry-based proteomics empowered the shift from analysis of single proteins to analysing the entire population of expressed proteins in a cell, or the proteome (Wilkins, Sanchez, et al., 1996).

Mass spectrometry-based proteomics is normally segregated into 'top-down' and 'bottom-up' or shotgun proteomics (Kelleher, Lin, et al., 1999; Yates, Ruse, & Nakorchevsky, 2009) (Figure 2.7). Some ambiguities about the usage of these two terms exist, but the most common usage describes top-down as methods that use masses of intact proteins and their peptide fragments for successful identification. Complementary fragmentation reactions, such as electron capture dissociation (Zubarev, Kelleher, & McLafferty, 1998) and electron transfer dissociation (Coon et al., 2005; Syka, Coon, Schroeder, Shabanowitz, & Hunt, 2004) are ideal for top-down proteomics approach since these methods give a more thorough protein backbone sequencing and retain labile post-translational modifications (Kelleher, Zubarev, et al., 1999). Top-down data are generally analysed utilising expressed sequence tags (Cargile, McLuckey, & Stephenson, 2001; Mortz et al., 1996) or the *de novo* sequencing technique

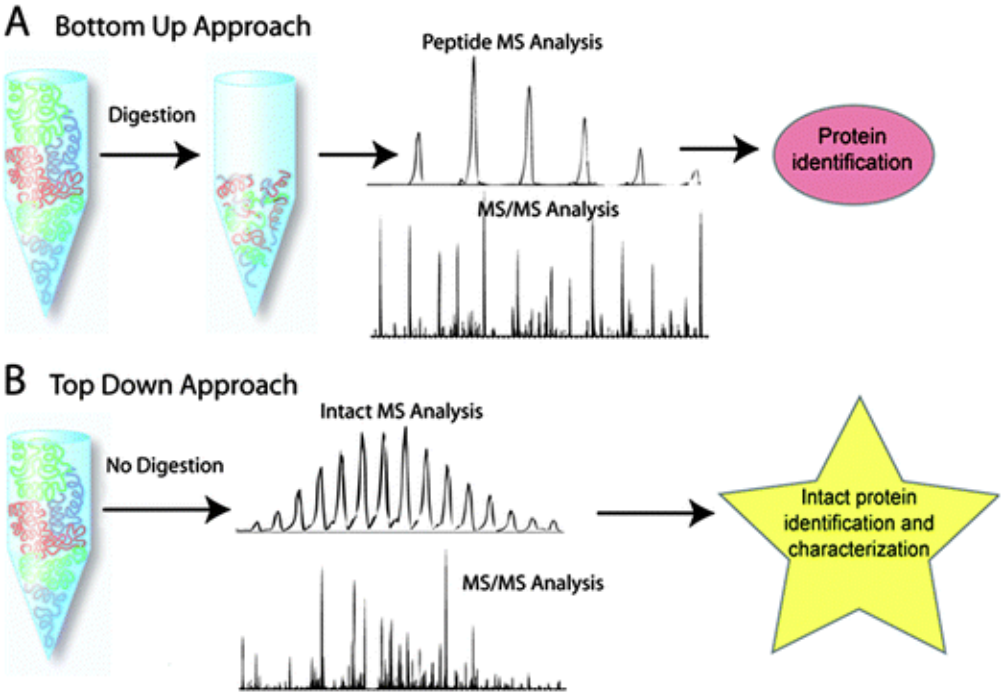


Figure 2.7. The different approaches used in mass spectrometric analysis (Kellie et al., 2010).

(Horn, Zubarev, & McLafferty, 2000). While the advantages of top-down proteomics such as higher sequence coverage of target proteins (Kelleher, Lin, et al., 1999), better characterization of post-translational modifications (Siuti & Kelleher, 2007; Zabrouskov et al., 2005) and improved reliability of protein quantification are irrefutable (Du, Parks, Sohn, Kwast, & Kelleher, 2006; Pesavento, Mizzen, & Kelleher, 2006; Waanders, Hanke, & Mann, 2007); there is no efficient method to fragment large proteins available. Therefore, the capacity of the top-down approach has been restricted to the analysis of single proteins or simple protein mixtures (Yates et al., 2009). However, in the last decade, there have been some attempts to use this approach for studying complex mixtures (Bunger, Cargile, Ngunjiri, Bundy, & Stephenson, 2008; Sharma et al., 2007).

Bottom-up proteomics is comparable to shotgun sequencing where intact deoxyribonucleic acid (DNA) is fragmented into smaller nucleotide fragments and sequenced. The full DNA sequence is then determined *in silico* by combining the sequencing information. Likewise, a routine bottom-up proteomic procedure involves digestion of a single to a few proteins into peptides and analysis of these peptides with mass spectrometry, routinely for characterisation of their amino acid sequences. Shotgun proteomics utilises the bottom-up technique to analyse complex mixture of peptides with mass spectrometry after separation using high performance liquid chromatography (HPLC). Separated fractions are eluted directly into a mass spectrometer, for example a quadruple-time-of-flight or an ion trap system that is

capable of tandem mass spectrometric analysis in an automated mode. Most bottom-up applications require tandem data acquisitions in which fragmentations of peptides are induced using collision-activated dissociation (Lee, Tan, & Chung, 2010) or also known as collision-induced dissociation (Yates et al., 2009). The complexity of the peptide mixtures largely dictates the ionisation techniques used to produce peptide ions. Matrix-assisted laser desorption ionisation (MALDI) alone is suitable for less complicated peptide mixtures; whereas HPLC coupled with MALDI or electrospray ionisation-based mass spectrometry system are often ideal for complex peptide mixtures. Eventually, several steps of peptide ion fragmentations generate the tandem mass spectrometric data in the form of mass to charge ratios that can be deconvoluted by robust bioinformatics software. The most commonly used approach for tandem protein identifications is by matching the tandem mass spectrometric data to the available sequence databases (Eng, McCormack, & Yates III, 1994; Perkins, Pappin, Creasy, & Cottrell, 1999). The peptide fragment ions generated can be used to match to a particular database of predicted fragmented masses for the peptide sequences of the same or similar organisms.

Large scale bottom-up or shotgun proteomics essentially started with the introduction of multidimensional protein identification technology (MudPIT) (Link et al., 1999; Washburn, Wolters, & Yates III, 2001; Wolters, Washburn, & Yates III, 2001). In MudPIT, a complex peptide mixture is loaded into a microcapillary or nano column with packing materials such as reversed phase and strong cation exchange. This microcapillary column is positioned in line with a high pressure liquid chromatography and a tandem mass spectrometry system (Fournier, Gilmore, Martin-Brown, & Washburn, 2007). Two areas in which the MudPIT had led to significant improvements over two-dimensional gel electrophoresis approaches are the detection and identification of membrane proteins and of low abundance proteins (Washburn et al., 2001). The success of the MudPIT approach for the large scale analysis of the yeast (Washburn et al., 2001) and rice proteomes (Koller et al., 2002) inspired the development of numerous coupled two-dimensional separations and mass spectrometry systems (Fournier et al., 2007).

2.5.1 Proteomics: The technologies and complications with plant samples

A typical proteomics workflow (Figure 2.8) involves protein extraction, separation and quantitation of peptides, and protein identification with mass spectrometry. The proteome of a cell or tissue at any given time is extremely complicated and diverse. Various approaches

have been applied to investigate a subset of proteins but never the whole proteome due to the complexity of the proteins and its wide dynamic range. Protein and peptide separation essentially encompasses two complementary approaches; gel-based and non-gel based (or gel-free). They vary in terms of peptide extraction, separation and detection. At the end, each of these approaches only covers specific protein subgroups but not the entire proteome. The gel-based approach is the cornerstone of proteomic analysis as mentioned earlier. Gel electrophoresis is a powerful technique to separate complex protein mixtures to yield qualitative and quantitative high resolution snapshots of intact proteins (two-dimensional) and polypeptides (one-dimensional), resulting in a good overview of protein isoforms variety and post-translation modifications.

The limitations of the gel-based approach, however, include the inability of the gel technique to resolve hydrophobic proteins or proteins with extreme sizes and isoelectric points. Thus, mass spectrometric analysis of these proteins is often not achievable unless specific techniques are employed. Non-gel based approaches have been developed mainly because of the need to reduce technical variation for high-throughput workflow that cannot be achieved using the gel electrophoresis techniques. In the gel-free approach, liquid chromatography is coupled to electrospray ionization (tandem mass spectrometry was used to separate peptide mixtures either using only reversed-phase column or with combination of different columns (as in MudPIT). Datasets of acquired tandem mass spectra were then used for protein database searches using a search engine such as Mascot.

One of the typical challenges in plant proteomics is the low protein concentration specially in circumstances where the amount of tissues are limited, as cell wall and vacuole make up most of the cell mass. Only 1-2% of the total cell volume make up the cytosol (Carpentier et al., 2008). Specialised procedures are essential to induce plant cell wall disruption to release the proteins. In addition, plant extracts also contain numerous non-proteinaceous compounds such as polyphenols, pigments, polysaccharides, nucleic acids and lipid. Several major crops had been reported to contain high amounts of these interfering compounds. For instance, oxidative enzymes (polyphenol oxidase), phenolic compounds, latex and carbohydrates are abundant in banana, *Musa* spp. and stalk tissues (Amalraj et al., 2010; Gooding, Bird, & Robinson, 2001; Wuyts, De Waele, & Swennen, 2006). These contaminants are co-purified with the precipitated proteins, rendering them difficult to solubilise. Solubilisation of proteins

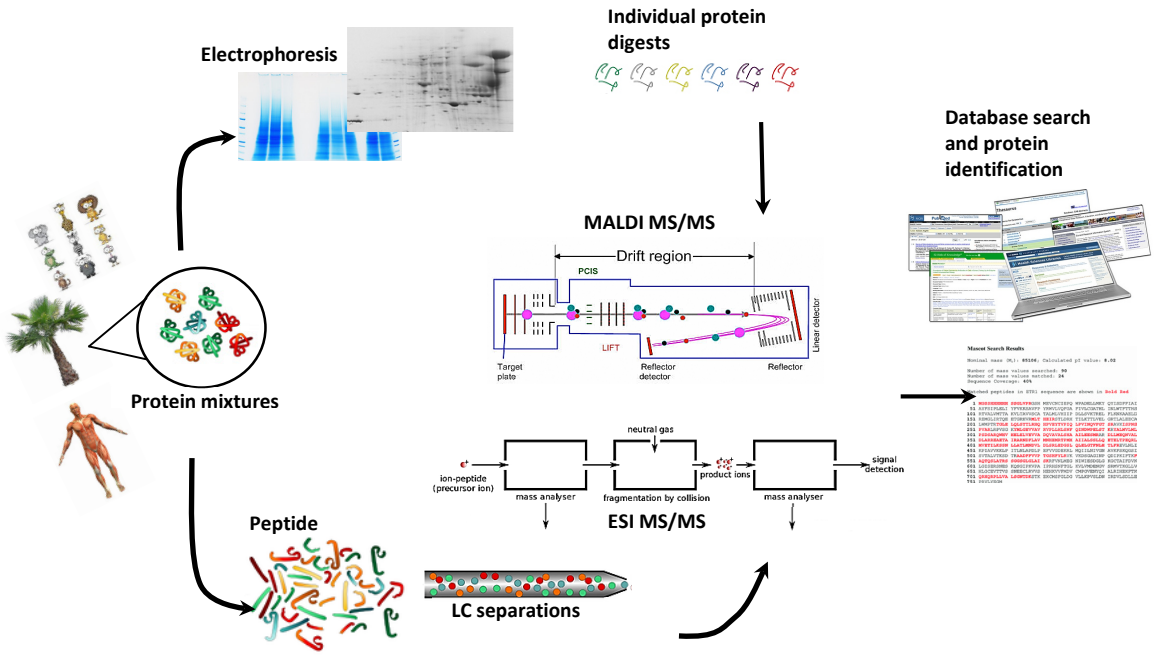


Figure 2.8. Typical proteomics analysis strategies.

is decisive in order to resolve them for further downstream analysis using techniques such as Western blot and mass spectrometry. Exhaustive protocols to remove contaminants during sample preparation have been described (Chatterjee, Gupta, Bhar, & Das, 2012; Gómez-Vidal, Salinas, Tena, & Lopez-Llorca, 2009; Saravanan & Rose, 2004).

Plant tissues also have distinctive subcellular components. This has led to the emergence of plant organelle proteomics in the modern proteomic workflow (Agrawal et al., 2011; Lilley & Dupree, 2007). Pre-isolation of various subcellular organelles by differential centrifugation is a robust approach to investigate cell or tissue sub-proteomes. This approach had been employed in our study to enrich the chromoplasts from other plant organelles such as mitochondria.

The existence of a high dynamic range of protein abundances in plant tissues confers additional complications to the protein analyses. For example, 40% of the total protein content of green tissues consists of ribulose-1,5-bisphosphate carboxylase oxygenase or RuBisCo (McCabe et al., 2001) while storage proteins are the most abundant proteins in seeds (Carpentier et al., 2008). The presence of those highly abundant proteins complicates the detection of low abundance proteins by means of protein electrophoresis and mass spectrometry. Normally, these low abundance proteins, such as the regulatory proteins are the proteins that we are interested in (Jensen, 2004). Various fractionation techniques have

been developed to deal with this wide dynamic range, which can be up to 12 orders of magnitude (Corthals, Wasinger, Hochstrasser, & Sanchez, 2000). They are generally divided into electrophoresis- or chromatography-based fractionations to separate a subset of proteins (van Wijk & Baginsky, 2011). For instance, isoelectric focussing that exploits charge differences of proteins has been utilised to fractionate proteins to capture the less abundant proteins (Hey, Posch, Cohen, Liu, & Harbers, 2008; Horth, Miller, Preckel, & Wenz, 2006). Other approaches are based on the principle of affinity chromatography such as ATP and metal affinity (Bayer, Stael, Csaszar, & Teige, 2011), hydroxypapatite affinity- (Mamone, Picariello, Ferranti, & Addeo, 2010) and immobilised metal affinity- chromatography (Aryal, Krochko, & Ross, 2011). The latter technique is extremely useful in the enrichment of phosphorylated proteins in phosphoproteomics studies.

2.5.2 Quantitation in proteomics and its application in plants

Early developments in quantitative proteomics were propelled by studies on yeast and mammalian cell lines (Schulze & Usadel, 2010). Quantitative information is needed to elucidate changes in protein expression. Staining of proteins on gels with specific stains such as Coomassie is routine to determine their intensities but this approach is tedious and error-prone. The intensities of the liquid chromatography peaks detected using an ultraviolet-visible spectrophotometric detector are usually not proportionate to the amount of proteins in a given sample. The reason is that different types of protein absorb different ultraviolet wavelengths which give the chromatograms. For examples, aromatic amino acids (tryptophan, tyrosine and cysteine) absorb ultraviolet wavelength at 280 nm (less accurate) while peptide backbone absorbs ultraviolet wavelength at 215-235 nm (more accurate, relatively). Thus, stable isotope approaches were introduced into mass spectrometry-based proteomics to allow a more accurate and reliable determination of relative variation in abundance of peptides. There are several strategies used today in quantitative proteomics and all of these methods have their advantages and disadvantages. Investigations using a quantitative proteome analysis approach rely either on a non-mass spectrometry-based quantitation technique (Carpentier et al., 2008) such as Difference Gel Electrophoresis (DIGE) or a mass spectrometry-based quantitation technique (Remmerie et al., 2011).

Determination of protein ratios using gel-based methods has the potential for error because of gel-to-gel inconsistencies of the separated protein profiles. DIGE addresses these difficulties and that explains why this technique is the most commonly employed in non-mass

spectrometry-based proteomic quantitation, as well as quantitative analysis in plant proteomes. In DIGE, up to two different protein samples and a reference standard (containing equal amounts of both protein samples) are labelled with fluorescent dyes such as CyDyes (Cy3, Cy5, Cy2). The two protein samples are then pooled prior to separation with gel electrophoresis (Timms & Cramer, 2008; Unlu, Morgan, & Minden, 1997). Protein ratios between the two different samples are calculated by measuring the fluorescence for each protein spot and thus revealing the quantitative data for protein isoforms or differentially regulated proteins (Bindschedler & Cramer, 2011). DIGE has been commonly used in plant proteomics studies. For example, in investigations of elicitation effects in plant symbiotes and plant pathogen interactions (Amey et al., 2008; Chivasa et al., 2006; Gerber, Laukens, De Vijlder, Witters, & Dubery, 2008; Schenkluhn, Hohnjec, Niehaus, Schmitz, & Colditz, 2010) as well as studies on environmental stresses (Casati, Zhang, Burlingame, & Walbot, 2005; Renaut, Hausman, & Wisniewski, 2006; Zhou, Sauvé, & Thannhauser, 2009). Gomez and co-workers (Gómez, López, Pintos, Camafeita, & Bueno, 2009) also demonstrated that DIGE coupled to MALDI-TOF analysis could be used to identify differentially expressed proteins in organisms lacking assembled genomes. Pro-Q Diamond, a fluorescent dye that binds to the phosphate moiety of phosphorylated proteins has also been successfully employed to specifically label and quantify phosphorylated protein isoforms in plant (Agrawal & Thelen, 2009; Agrawal & Thelen, 2006; Barsan et al., 2010b; Chitteti & Peng, 2007; Gerber, Laukens, Witters, & Dubery, 2006; Steinberg et al., 2003).

The mass spectrometry-based quantitative methods include both label free quantitation (Niittylä, Fuglsang, Palmgren, Frommer, & Schulze, 2007; Stulemeijer, Joosten, & Jensen, 2009) and chemical isotope labelling. Mass spectrometry signals from different liquid chromatography runs are known to be inconsistent due to technical variations for instance, and therefore generate significant error in quantitative proteomic studies. Despite that fact, label-free methods involving liquid chromatography is becoming increasingly prevalent as it circumvents the need for costly protein labelling and is generally suitable for all types of organisms as well as most workflows (Bindschedler & Cramer, 2011; Kieffer, Dommes, Hoffmann, Hausman, & Renaut, 2008). Label-free quantitation compares the chromatographic peak areas of extracted ions. Extracted ion chromatograms exploit the additional separation dimensions for higher confidence in the quantitative signals instead of simply comparing the mass spectrometry signals between different analytical liquid

chromatography runs. In principal, peptide areas are aligned according to their mass to charge ratios (*m*/*z*) and elution time tags in several liquid chromatography runs. The chromatographic peaks are then integrated with peak integration software such as Xalign (Zhang, Asara, Adamec, Ouzzani, & Elmagarmid, 2005) and Msalign (Palmblad, Mills, Bindschedler, & Cramer, 2007). In order to be able to do that, the liquid chromatography runs must be reproducible, which sometimes can be a challenging task. Reiland and co-workers used this approach to determine the dynamic regulation of protein phosphorylation in *Arabidopsis* (Reiland et al., 2009).

Spectral counting is an alternative approach that is practical, label-free and measures protein abundance in a semi-quantitative manner (Lilley & Dupree, 2007). Conversely, this method does not integrate chromatographic peaks nor align the retention time of peptides (Abdallah, Dumas-Gaudot, Renaut, & Sergeant, 2012) although it agrees with Extracted Ion Chromatogram peak area measurements (Old et al., 2005). Instead, statistical tools such as G-test and *t*-test are used to count the total number of tandem mass spectra identified for all the peptides from a particular protein to generate the quantitation data (Bindschedler & Cramer, 2011). While this method is reproducible, it requires many biological and technical replicates for each sample analysed. This can be difficult when several experimental conditions and/or time points are analysed. This approach had been successfully employed to quantify proteins in several studies in plant systems (Friso, Majeran, Huang, Sun, & van Wijk, 2010; Gammulla, Pascovici, Atwell, & Haynes, 2010; Lee et al., 2009; Zybailov et al., 2009).

Usage of differential labelling techniques could circumvent these limitations in label-free quantitation engaging liquid chromatography. These approaches rely on the assumption that both labelled and unlabelled peptides exhibit the same chromatographic and ionisation properties but are distinguishable by a mass-shift signature (Abdallah et al., 2012). Specific isotope labelled amino acids (^{13}C or ^{15}N) (Benschop et al., 2007) in the metabolic protein labelling technique known as Stable Isotope Labelling with Amino Acids in Cell Cultures (SILAC), and chemical labels such as in Isobaric tags for Relative and Absolute Quantitation (iTRAQ) have been used to quantitate changes in plant proteomes (Jones, Bennett, Mansfield, & Grant, 2006; Nuhse, Bottrill, Jones, & Peck, 2007). Labelling methods used in relative quantitation proteomics studies are classed into two categories depending on whether the labels are tagged directly to the peptides or not. Considering that SILAC uses labelling of plant

cells or whole plants, this approach is not suitable for quantitative proteomic studies using the current sample materials, which are oil palm mesocarps.

Isotope-Coded Affinity Tags (ICAT) (Figure 2.9) was one of the first differential isotope labelling containing three elements: a specific chemical reactive group which bound specifically to cysteinyl residues, an isotope mass tag with light or heavy isotopes and a biotin tag for affinity purification (Gygi et al., 1999). Peptide pairs with 8 kDa mass-shifts are detected in mass spectrometry scans and their ion intensities are compared for relative quantitation. As tagged cysteine-containing peptides are purified by affinity chromatography, the sample complexity is reduced. However, the obvious disadvantage is that only cysteine-containing peptides are captured by the affinity column. Thus, this impaired the identification and quantitation of proteins with more than one significant peptide as about one in seven proteins do not contain cysteine (Abdallah et al., 2012). A study by Majeran and co-workers had revealed that non-MS-based (2-DE), ICAT and label-free quantitative techniques are complementary (Majeran, Cai, Sun, & van Wijk, 2005). ICAT had been utilised to determine the localisation of *Arabidopsis thaliana* organelle proteins (Dunkley, Dupree, Watson, & Lilley, 2004). In addition, since ICAT labels specifically to thiol groups, this method has been widely used to study the redox-status of proteins in plants (Hagglund et al., 2010; Miles et al., 2009).

Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) was developed at first for peptide level labelling (Abdallah et al., 2012). The different between iTRAQ and ICAT is that in ICAT, tagged proteins from different samples are pooled before trypsinisation to eliminate vial-to-vial variations. In iTRAQ, the chemical tags label the peptides instead. The iTRAQ isobaric tags have slight differences in their molecular structures (Figure 2.10) and thus generate various fragment ions (also known as reporter ions) in tandem mass spectrometry scans. The overall molecule mass is kept constant at 145 Da (iTRAQ-4plex) and 304 Da (iTRAQ-8plex) by the presence of a mass balance group (carbonyl).

iTRAQ reagents label the peptide N-terminals and the amino groups of lysine side chains. The advantage of the iTRAQ approach is that it allows comparison of four (iTRAQ-4plex) to eight samples (iTRAQ-8plex) in a single experiment. Relative quantitation is ascertained after peptide fragmentations in MS/MS scans by measuring the intensity of the reporter ions for the peptide in the mass region of *m*/*z* 114-118 and *m*/*z* 114-121, for 4plex and 8plex, respectively (Abdallah et al., 2012; Bindschedler & Cramer, 2011). The iTRAQ method is able

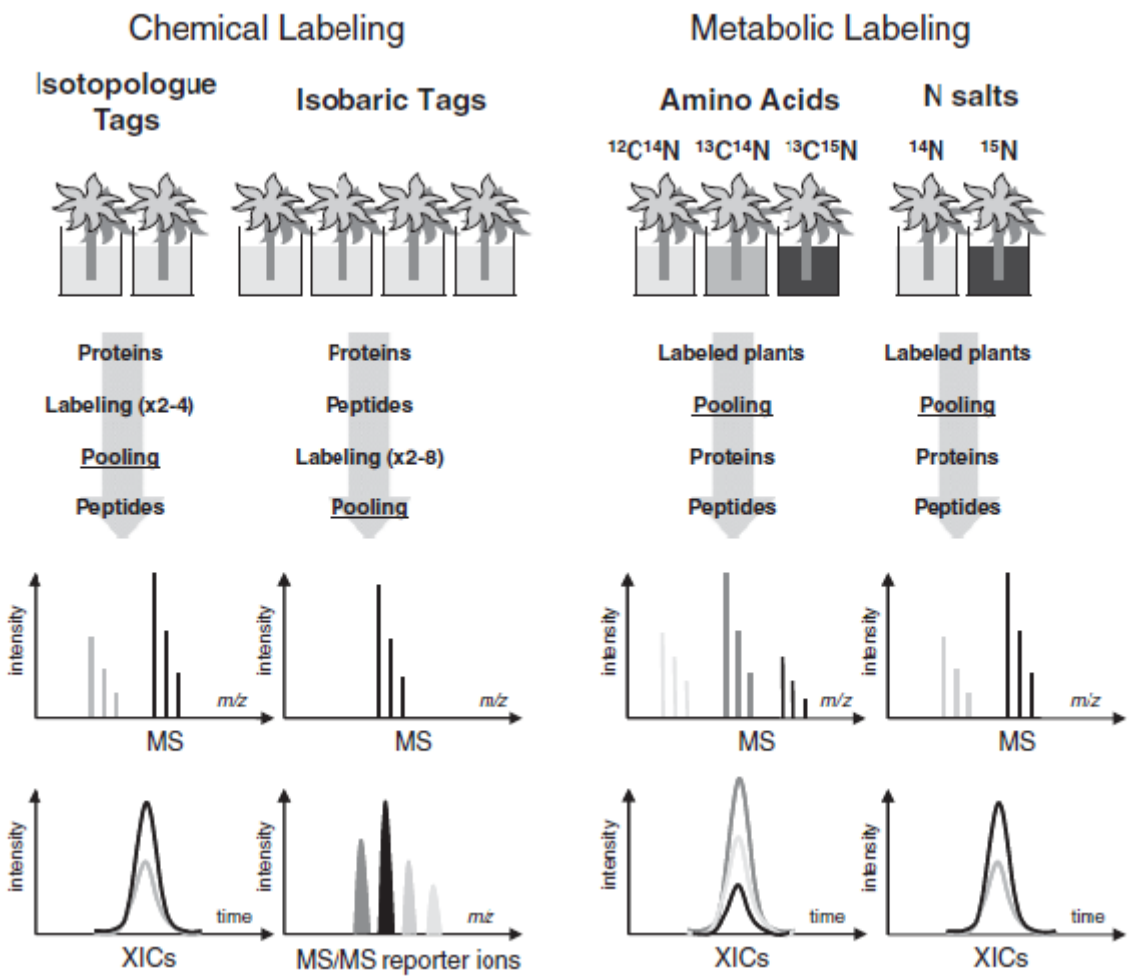


Figure 2.9. General strategies for isotope labelling methods in quantitative proteomics (Bindschedler & Cramer, 2011).

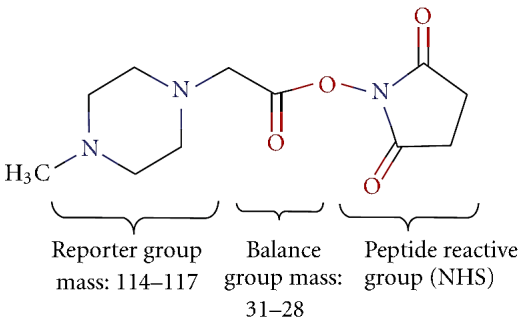


Figure 2.10. Structure of iTRAQ tag (Abdallah et al., 2012)

to give accurate quantitation spanning two orders of magnitude for low-complexity samples. Peptide cofragmentation happens when two or more closely spaced peptides in MS/MS are selected instead of the single peptide (Ow et al., 2009; Perkel, 2009). Peptide cofragmentation effect is reduced with a high accuracy mass spectrometer.

iTRAQ reagents have been successfully employed in several quantitative plant proteomics studies. Plant responses towards pathogens have been investigated using this approach (Fan et al., 2011; Marsh et al., 2010; Melo-Braga et al., 2012; Mohammadi, Anoop, Gleddie, & Harris, 2011) as well as the signalling role played by trimeric G proteins in plants (Kaffarnik, Jones, Rathjen, & Peck, 2009; Zhao, Stanley, Zhang, & Assmann, 2010). Other studies utilised iTRAQ to investigate the proteomes of grape berries (Lucker, Laszczak, Smith, & Lund, 2009) and oil palm mesocarp at different stages of ripening (Loei et al., 2013). Quantitative shotgun proteomics using iTRAQ was also employed to characterise differential phosphorylation of Arabidopsis in response to microbial elicitation (Jones et al., 2006).

2.6 Pitfalls of plant proteomics without a reference plant

Model plants are commonly used in investigating the physiological properties of cells, tissues, organelles or whole organisms. Simplicity of study design, biological relevance and economics have an important impact on the plant models employed (Carpentier et al., 2008). The green plants or *Viridiplantae* have only a diminutive number of completed and publicly available plant genomes (Figure 2.11). Presently, there are over 300,000 known species of land plants alone but the model plants only represent a handful of species and families. The existing plant genomes are the classical *Arabidopsis thaliana* (thale cress), economically important crops such as *Glycine max* (soybean), *Medicago truncatula* (barrel medic), *Populus trichocarpa* (poplar), *Vitis vinifera* (wine grape), *Oryza sativa* (rice), *Sorghum bicolor* (sorghum) and *Zea mays* (maize) (www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html), as well as other plants like the *Brachypodium distachyon* (purple false brome) (Vogel et al., 2010). However, none of these plant genomes are completely annotated (Remmerie et al., 2011). Moreover, only *Oryza sativa*, *Sorghum bicolor*, *Zea mays* and *Brachypodium distachyon* are monocotyledons while the rest are dicotyledons. Given that oil palm species used in this study does not have a standard reference plant and is a monocotyledon, from the technical outlook, those plants mentioned earlier are unlikely to be suitable as a model organism in the study. Encouragingly, a comprehensive genome sequencing project led by the Malaysian Palm Oil

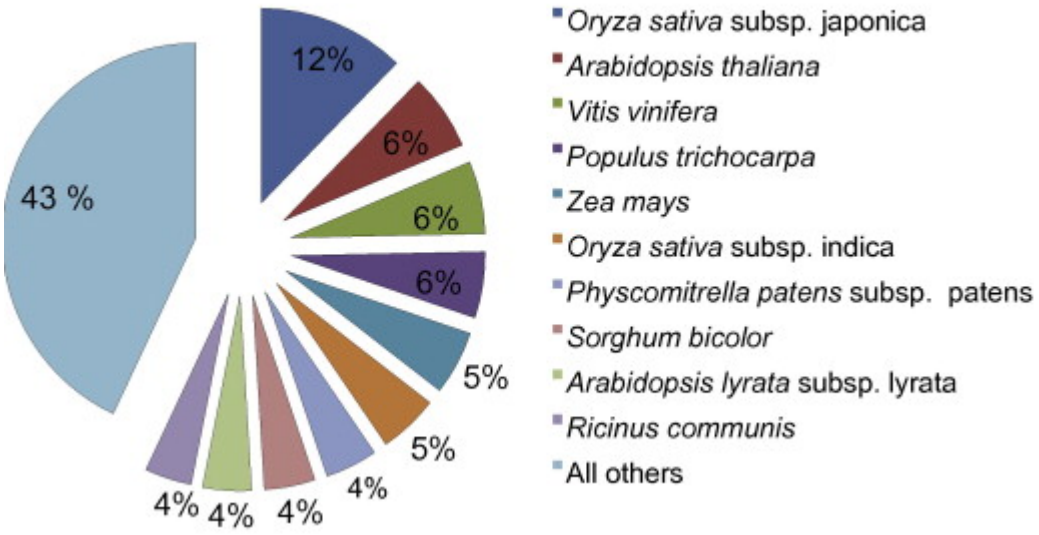


Figure 2.11. Accessible plant genomic databases in UNIPROT. Contributions of different plant species to the entries in UNIPROT database are shown in percentage (Remmerie et al., 2011).

Board and St. Louis based Orion Genomics, USA for the two key oil palm species, *E. oleifera and E. guineensis* was completed in 2010. However, to apply these species to investigate various basic biological processes including the intended post-translational modifications study in fatty acid biosynthesis is still not feasible. The annotations of the oil palm genome data is on-going (Low, personal communication, 1 March, 2012). Nevertheless, Uthaipaisanwong and co-workers successfully characterised the oil palm chloroplast genome sequence (Uthaipaisanwong et al., 2012) and there are 822 partial and full sequences of *E. guineensis* proteins currently available in NCBI protein database (as of 4th March, 2013). This information can support our oil palm organelle proteomics study in the chromoplasts.

Without the genomic sequences, efforts to identify and characterise protein with mass spectrometry efforts are significantly ameliorated with the availability of expressed sequence tag (EST) sequences (Yang et al., 2010). A large collection of 19,243 *Elaeis guineensis* ESTs had been assembled to give 10,258 unique sequences (Ting et al., 2010) while there are 40,809 ESTs deposited in the NCBI database. The EST databases are indispensable as those sequence tags can be translated into the six reading frames to identify proteins (homology based) using appropriate software. Nonetheless, the size and quality of EST databases have profound effects on the outcome of the protein identification. The usual limitations of EST databases are that more often than not, most proteins are either not or poorly denoted by short EST sequences that only partly cover the whole protein sequence. Bases misread, insertion or deletion errors during sequencing of ESTs can lead to high error rates (about 0.3%) in EST

sequences, thus reducing the accuracy of peptide matching (Hoff, 2009; Pedretti et al., 2001). Successful protein identifications with only peptide mass fingerprints employing EST databases are not feasible due to the limitations with EST databases. In addition, EST sequences are rarely sufficient in providing significant protein coverage and a satisfactory number of matching peptides (Rowley et al., 2000).

Reliance on complete plant genomes can be lessened as annotating the biological function of proteins can be facilitated by a homology based approach. According to Carpentier, Remmerie and their respective colleagues (Carpentier et al., 2008); (Remmerie et al., 2011), there is a requirement for cross-species analysis. When using other species for protein identification with mass spectrometric data, orthologue sequences are preferable as they are more likely to share similar functions (Gabaldón, 2007). Homologous sequences originate from a sequence in a common ancestor. The sequences are considered different or orthologues when they diverged by a speciation (inter-species) event. Paralogous sequences are sequences that came from a common ancestor and are present in the same genome. However, duplication event (intra-species) transpired in the sequences to produce paralogous sequences which may or may not share similar functions (Fitch, 1970).

Database dependent- or independent strategies can be used to execute confident cross-species protein identification. In the former approach, search engines such as Mascot are used to search peptide sequence data that contains precursor ion mass and a list of product ion masses against a taxonomically confined database (Perkins et al., 1999). Only a massive amount of peptide masses generated can guarantee its success as this increases the probability of matching several peptides to the homologous protein. This has been demonstrated for pea chloroplast proteins (Peltier et al., 2000) and maize proteins (Chang et al., 2000). In a database independent strategy, fragmentation spectra are utilized to obtain *de novo* peptide sequences. MSBLAST, which uses a combination of BLAST search and peptide *de novo* sequences, had been adapted for tandem mass spectrometry data to increase the accuracy and hit rate of protein identification (Shevchenko et al., 2001). Analysing modified (in peptide sequences) proteins without completed and annotated genomes has proved to be a daunting effort. The identification of modified peptides and their modification sites is essentially based on single amino acid identifications and becomes impossible when the peptide sequence is not available in any plant database. Application of a *de novo* sequencing strategy is able to facilitate the identification of modified peptides and may even help to locate

the modification site, albeit with requirements for high quality tandem mass spectrometry spectra and certain preferred fragmentation techniques such as electron capture dissociation or electron transfer dissociation (Carpentier et al., 2008; Remmerie et al., 2011).

Our work is focussed on proteins located in the chromoplast as this organelle is reported to have a major role in fatty acid biosynthesis (Awai et al., 2001; Benning, 2008; Benning, 2009a; Harwood, 1996; Heinz & Roughan, 1983; Joyard & Douce, 1977; Joyard et al., 2010; Ohlrogge & Browse, 1995; Ohlrogge & Jaworski, 1997; Roughan & Slack, 1982; Slabas et al., 1994; Slabas & Fawcett, 1992; Stumpf, 1969). While the annotated oil palm genome is yet to materialise, which hampers efforts to identify these proteins, there are several databases in which the identities of the proteins could be ascertained. PPDB, launched in 2004, is a Plant Proteome DataBase for *Arabidopsis thaliana* and maize (*Zea mays*) (Sun et al., 2009). PPDB was developed to accommodate plant plastids, but over time, the database expanded to cover the entire proteomes of those two plants. As a result, in November 2007, the database was renamed to Plant PDB. The database consists of cell type-specific proteomes (maize) or specific sub-organelle proteomes such as chloroplasts, thylakoids and nucleoids as well as whole leaf proteome (maize and *A. thaliana*). More than 1,500 *A. thaliana* proteins, prominently from the plastids, have been assigned with subcellular locations.

Information on protein localisation also helps in understanding the function of proteins and their biological inter-relationships. The SubCellular Proteomic Database (SUBA) provides the hypothetical localisation of many proteins that were identified in various sub-plastidial compartments in *Arabidopsis thaliana* (Heazlewood, Verboom, Tonti-Filippini, Small, & Millar, 2007). AT_CHLORO is another comprehensive chloroplast proteome database that focusses on the localisation of proteins in the stroma, thylakoids and envelope membrane (Ferro et al., 2010). The database contains a comprehensive repertoire of 1323 proteins that were identified by 10,654 unique peptide sequences. These proteins were present in highly purified chloroplasts and their sub-fractions from *Arabidopsis thaliana* leaves. These databases are valuable as resources for protein identifications even though the current study focusses on the mesocarp plastids and not the leaf. Combined with the limited genome sequences of other plants, they help in understanding the regulation of chromoplastic proteins and their respective roles in fatty acid biosynthesis involving post-translational modifications.

2.7 Post-translational modifications (PTMs)

Almost all proteins are modified in some way after synthesis. Many physiological responses result from differential protein modifications rather than changes in protein expression levels. These modifications (Figure 2.12) do not create novel proteins but rather a new 'protein species' since the translated protein sequence remains unaltered (Jungblut et al., 1996; Schluter, Apweiler, Holzhutter, & Jungblut, 2009). The modifications occur through covalent binding of functional groups such as phosphates, sulphates, carbohydrates and lipids (Bond, Row, & Dudley, 2011). This event, which is known as PTM, is one of the key mechanisms that changes the properties of a protein in cells and greatly enhances the structural diversity and functionality of proteins. This is feasible because PTMs provide a larger repertoire of chemical properties than is possible using the 20 amino acid specified by the genetic code. Protein PTMs could result in alterations in activity, localisation, production, interactions with other proteins and half-life (Bienvenut et al., 2011; Endler & Baginsky, 2011; Mann & Jensen, 2003). Modifications are often permanent, but some modifications, such as phosphorylation, are reversible and can be used to switch protein activity 'on' and 'off' in response to intracellular and extracellular signals. For example, in a signal transduction process, kinase cascades are activated or inactivated through reversible addition and removal of phosphate groups. The esterification of an amino acid side chain through the addition of a phosphate group introduces a strong negative charge, which can subsequently modify the conformation of the protein and alter its stability, activity and potential to interact with other molecules. Genomic sequencing has revealed that protein kinases are probably coded by 2-3% of all eukaryotic genes (Venter et al., 2001). PTM is therefore a dynamic phenomenon with a central role in many biological processes. Generally, in regulatory pathways, the status of serine, threonine and tyrosine is regulated by protein kinases and phosphatases (Adams, 2001; Johnson & Hunter, 2005). Interference with the activities of the kinases and phosphatases indirectly disrupts these regulatory pathways and may cause disease (Cohen, 2001; Lim, 2005).

The complexity of the proteome is increased significantly by PTMs, particularly in eukaryotes where many proteins exist as a heterogeneous mixture of alternative modified forms. Ideally, it would be possible to catalogue the proteome systematically and quantitatively in terms of the types of PTMs that are present, and specify the modified sites in each case. However, such

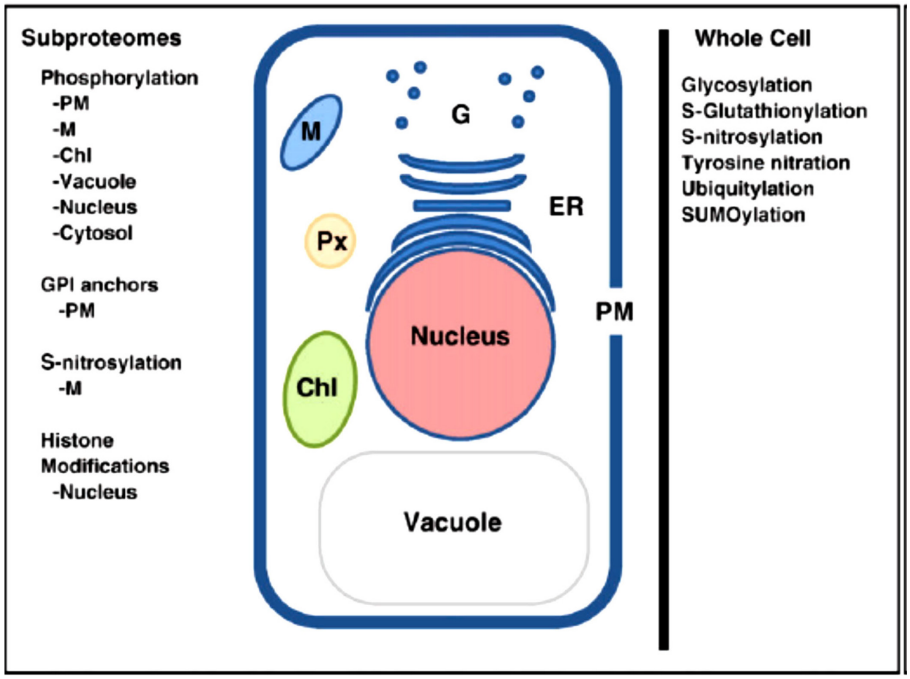


Figure 2.12. Modification-specific proteomes and their subcellular localizations. The sub-cellular compartments are the mitochondria (M), chloroplast (Chl), peroxisomes, (Px), plasma membrane (PM), endoplasmic reticulum (ER), Golgi apparatus, (G), vacuole, and nucleus (Ytterberg & Jensen, 2010).

attempts are complicated by the sheer diversity involved and the transient nature of certain modifications. Every protein could potentially be modified in hundreds of different ways, and might contain multiple modification target sites allowing different forms of modification to take place either singly or in combinations. Thus, it remains the case that most PTMs are discovered unintentionally when individual proteins, complexes, or pathways are studied. It is impossible to predict modifications accurately from the genome sequence. Even when a definitive modification motif is present; it is not necessarily the case that such or any modification will happen.

Until recent years, the analysis of PTMs at the proteomic level has received limited consideration due to the lack of appropriate techniques (Bond et al., 2011). However, improved separation methods can resolve different post-translational variants, and gels can be stained with reagents that recognize particular types of modified proteins. Mass spectrometry is at present the method of choice to characterise chemical additions and substitutions. Mass spectrometry analysis can be used to identify peptides carrying chemical adducts and can deduce their positions in the protein sequences.

Signalling proteins and regulatory molecules, which are the least abundant proteins in the cell, are often regulated by phosphorylation. Since the stoichiometry of phosphorylation is usually low, the modified target protein may be present in limiting amounts and may be difficult to detect and quantify. Ultimately, even if adequate amounts of a particular variant are available, a large quantity of the sample is required for the full characterization of modifications compared to the relatively simple matter of protein identification. Currently, affinity-based techniques are employed to improve the chances of detecting their targets by isolating sub-proteomes with particular types of modification (Reinders & Sickmann, 2005).

Investigations into PTMs and differentially expressed proteins are essential to comprehend cellular responses towards changes in environmental conditions (Vissers et al., 2009). It is clear that plants induce a complex array of pathways and protein phosphorylation cascades during biotic and abiotic stresses (Emes, 2009). There are over 200 possible PTMs (Jensen, 2006; O'Donovan, Apweiler, & Bairoch, 2001; Prabakaran, Lippens, Steen, & Gunawardena, 2012). A recent survey of the SwissProt database showed that a total of 87,308 experimentally detected amino acid residues modifications have been reported (Figure 2.13) (Khoury, Baliban, & Floudas, 2011). Phosphorylation is the most common and extensively studied PTM using mass spectrometry approaches (Asara, Christofk, Freimark, & Cantley, 2008; Beausoleil et al., 2004; Bond et al., 2011; Jensen, 2004; Kwon et al., 2006; Larsen et al., 2006; Lu, Ruse, Xu, Park, & Yates III, 2007; Mann & Jensen, 2003; Seo & Lee, 2004; Ytterberg & Jensen, 2010). The justification is that phosphorylation is one of the primary mechanisms in cellular process regulation (Graves & Krebs, 1999).

A mass spectrometry-based approach facilitates both absolute and relative quantitation of peptides and their PTMs. Internal standard peptides are employed for absolute quantitation for certain proteins and their defined PTMs (Gerber, Rush, Stemman, Kirschner, & Gygi, 2003). Relative quantitation is performed with either the peptide intensity profiling (PIP) or stable isotope labelling (SIL) using stable isotope-encoded chemical precursor molecules or alkylating reagents (Aebersold & Mann, 2003; Mann & Jensen, 2003).

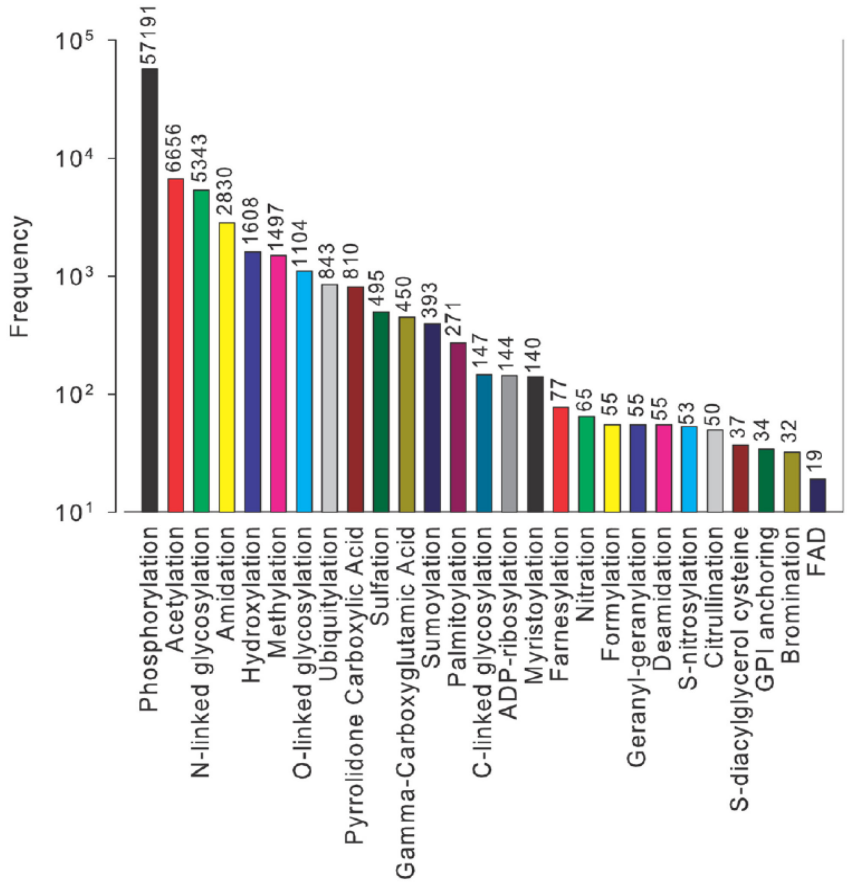


Figure 2.13. Experimental post-translational modifications curated from SwissProt database (Khoury et al., 2011).

2.7.1 Tracking the phosphopeptides

A wide array of approaches can be used to scrutinise phosphorylation changes in cell or tissues. Radiolabelling is a classical technique that uses radiolabeled ^{32}P-orthophosphate to detect phosphoproteins. Radioactivity can be very inconvenient, harmful and detrimental in the long term, both to the users and samples (Bendt et al., 2003; Su, Hutchison, & Giddings, 2007). Alternatively, after separation by two-dimensional gel electrophoresis (O'Farrell, 1975), phosphoproteins can be directly visualised on the gel using phosphospecific fluorescent stains and phosphospecific antibodies, which are non-radioactive (Barsan et al., 2010b; Boudsocq, Droillard, Barbier-Brygoo, & Laurière, 2007; Chitteti & Peng, 2007; Gerber et al., 2006; Patton, 2002). Immuno- or Western blot is the most common method used to assess the phosphorylation state of a protein using phosphospecific antibodies (for phosphorylated tyrosine, serine and threonine) transferred from a one-dimensional or two-dimensional gel electrophoresis (Bockus & Scofield, 2009; Kaufmann, Bailey, & Fussenegger, 2001). In direct staining, phosphospecific stains such as a fluorescent phosphosensor dye, Pro-Q Diamond (Invitrogen) bind directly to the phosphate moiety of phosphoproteins (Chitteti & Peng, 2007;

Schulenberg, Aggeler, Beechem, Capaldi, & Patton, 2003; Steinberg et al., 2003). The advantages of this stain are in its compatibility with other staining methods and the ensuing mass spectrometry analysis. This is particularly crucial when trypsinisations are performed directly on the gel pieces. A similar phospho-specific staining kit called Phos-tag had been used previously in which a Zn^{2+} ion chelator with high selectivity was coupled to a fluorophore (Nakanishi et al., 2007). The suitability of these stains in phosphoproteome analysis had been described in previous reports. Agrawal and Thelen (Agrawal & Thelen, 2006) identified 70 non-redundant phosphoproteins which belonged to the major functional classes from a Pro-Q Diamond stained two-dimensional gel containing rapeseed (*Brassica napus*) proteins. However, while phosphoproteins could be detected, the stains would not indicate the phospho-sites, which is vital in the characterisation of phosphorylation events. Special techniques are used to investigate membrane phosphoproteins due to the limitations in two-dimensional gel electrophoresis technique. Integral membrane proteins tend to aggregate during the isoelectric focusing migration and thus, it is not possible to separate them in the second dimension of two-dimensional gel electrophoresis.

The low abundance of phosphorylated proteins in cellular extracts and their relatively low degree of phosphorylation pose major challenges (Aebersold & Goodlett, 2001; Mann et al., 2002; Simpson, 2003). In mass spectrometric analysis, non-phosphorylated peptides often compete with the phosphorylated peptides for ionisation. As a result, many phosphoprotein peaks are difficult to detect, either because they have low signal to noise ratio or they are not ionised at all. Therefore, to tackle this obstacle, enrichment techniques, which are commonly applied prior to separation using liquid chromatography, have been used. Immobilized metal affinity chromatography (IMAC) is one of the methods that are used to enrich phosphopeptides from complex mixtures based on affinity of positively charged metal ions (Fe^{3+}, Al^{3+}, Ga^{3+} or Co^{2+}) towards phosphate moieties. Iminodiacetate and nitrilotriacetate are the prototypical metal-binding ligands used in IMAC stationary phases (Dunn, Reid, & Bruening, 2010; Grimsrud et al., 2010; Sun, Chiu, & He, 2005; Thingholm, Jensen, & Larsen, 2009). The Fe(III)-NTA complex is perhaps the most frequently utilised to enrich phosphopeptides although the use of other metal-ligand complexes had also been reported (Nuhse, Stensballe, Jensen, & Peck, 2003; Posewitz & Tempst, 1999). Most recently, Zr(IV)-phosphonate immobilized on various stationary phases had also been employed for phosphopeptide enrichment by several groups (Dong, Zhou, Wu, Ye, & Zou, 2007; Feng et al.,

2007; Wei et al., 2008; Yu et al., 2007; Zhou et al., 2006). The phosphopeptides can be eluted by different salt- and/or pH gradients prior to mass spectrometry analysis. Nonetheless, several challenges arise when using IMAC. Leaching of ions from the column during enrichment steps, non-specific binding of peptides that contain the acidic amino acids glutamic and aspartic acid and higher specificity for multiply phosphorylated peptides are amongst those complications (Dunn et al., 2010).

Metal oxide affinity chromatography (MOAC) is another valuable technique to isolate phosphopeptides from complex mixtures with high selectivity and recoveries (Dunn et al., 2010; Heintz et al., 2004; Hsu et al., 2009; Schmidt, Csaszar, Ammerer, & Mechtler, 2008). The metal oxides are often more stable at high temperatures and broad pH range (Nawrocki, Dunlap, McCormick, & Carr, 2004). Titanium oxide (TiO_2) is the most popular metal oxide resin used as a selective affinity support to capture phosphorylated peptides (Ikeguchi & Nakamura, 1997; Ikeguchi & Nakamura, 2000; Larsen, Thingholm, Jensen, Roepstorff, & Jørgensen, 2005; Pinkse, Uitto, Hilhorst, Ooms, & Heck, 2004; Sano & Nakamura, 2004a, 2004b). At acidic pH, TiO_2 has a positively charged surface (Kosmulski, 2002) that permits very selective enrichment of phosphopeptides from complex samples by their affinities (phosphate groups) toward porous TiO_2 particles (Titansphere) (Thingholm, Larsen, Ingrell, Kassem, & Jensen, 2008). Water-soluble phosphates are desorbed under alkaline conditions. Strong cation exchange and titanium dioxide-type columns have both been used in phosphopeptide enrichment and stable-isotope labelling by amino acids in cell culture (SILAC) for quantitation to study phosphorylation changes (Olsen et al., 2006).

Technical variations and bias in quantitative analyses are often reported to occur after phosphopeptide enrichment (Bindschedler & Cramer, 2011). Nonetheless, successful identification and quantitation has been reported using a combination of enrichment strategies and label-free quantitation as in the case of *Arabidopsis* phosphopeptides from a plasma membrane fraction following sucrose treatment (Niittylä et al., 2007) and a hypersensitive response study in tomato plants (Stulemeijer et al., 2009). In addition, iTRAQ labelling has been successfully used to quantify phosphorylated peptides in *Arabidopsis* cells as their defence response to *Pseudomonas syringae* induction (elicidators) (Jones et al., 2006). As a rule of thumb, it is more effective to perform chemical labelling prior to any enrichment strategies due to the fact that enrichment steps confer technical bias in quantitative analyses (Bindschedler & Cramer, 2011).

Hydrophilic interaction chromatography can also be used in the pre-separation stage of peptides prior to phosphopeptide enrichment such as IMAC or TiO_2 affinity purifications, in addition to MudPIT LC. HILIC separates polar biomolecules by the binding of the polar biomolecules to the neutral, hydrophilic stationary phase in hydrophilic interaction chromatography through hydrogen bonds. These bonds can be broken by reducing the organic composition in the mobile phase and the peptides eluted based on their polarities (McNulty & Annan, 2008).

2.7.2 Bioinformatics in post-translational modification analysis

Automated prediction of post-translational modification (PTM) sites is one of the main interest areas for bioinformatics investigations. *In vivo* and *in vitro* determinations of modified proteins and their PTM sites are not only time-consuming and tedious, but often restricted to the availability and optimisation of enzymatic reactions in order to determine the type of modifications and sites (Blom, Sicheritz-Pontén, Gupta, Gammeltoft, & Brunak, 2004; Xue et al., 2005; Zanzoni et al., 2011). Tandem mass spectrometry spectra offer the most informative fingerprints of modified peptides. The spectra encode not only peptide sequences, but also the masses and sequence positions of modifications. For these reasons, computational techniques have been employed to manage the massive amounts of fragmentation spectra, modified protein determination and individual PTM site identification with high accuracy as well as efficiency (Yoo, Ho, Zhou, & Zomaya, 2008). The current PTM prediction tools basically are classed into four major groups based on their types of classification schemes (Basu & Plewczynski, 2010).

The first group comprises general PTM related resources such as *PROSITE* (Sigrist et al., 2002) which predicts types of PTMs based on their sequence pattern consensus. Several signature recognition methods are combined to probe a query protein sequence against observed protein signatures. The *Scansite* tool predicts kinase-specific and signal transduction relevant motifs (Yaffe et al., 2001). Conserved sequence motifs represent imprints of important biochemical properties or biological functions of those proteins.

The second group consists of various neural network prediction tools. These tools covers phosphorylation related prediction servers such as *NetPhos* (Blom, Gammeltoft, & Brunak, 1999) and *NetPhosK* (Blom et al., 2004; Hjerrild et al., 2004). *NetPhosK* is the most popular since the server allows a preferred 'threshold' value to be indicated during prediction.

The third group of the prediction tools encompasses different support vector machine based prediction techniques. These methods are constructed on the basis that adjacent residues to the phospho-sites represent the main determinant for kinase specificity (Zanzoni et al., 2011; Zhou et al., 1993). For instant, *PredPhospho* (Kim, Lee, Oh, Kimm, & Koh, 2004) aims to predict phosphorylation sites and the type of kinase that acts at each site. *AutoMotifServer* (Plewczynski, Tkacz, Wyrwicz, & Rychlewski, 2005) also predicts PTM sites in protein sequences using support vector machine classifier with both linear and polynomial kernels. *KinasePhos 2.0* is the web server to identify protein kinase-specific phosphorylation sites based on amino acid residues sequences and coupling patterns (Wong et al., 2007). *PHOSIDA* is also capable of predicting phosphosites (Gnad, Gunawardena, & Mann, 2011).

The final group consists of remaining types of machine learning based PTM prediction tools. *PPSP* adopted Bayesian Decision Theory to predict kinase-specific phosphorylation sites and has been reported to produced precise prediction of the probable phosphorylation sites for about 70 protein kinase groups (Basu & Plewczynski, 2010). *Ascore*, a probability-based score that was developed by Beausoleil and his colleagues, used the presence and intensities of site-determining ions in tandem mass spectrometry spectra to calculate the probability of exact phosphorylation site localisation (Beausoleil, Villen, Gerber, Rush, & Gygi, 2006). *Ascore* re-evaluates the results from search engine on phosphopeptide and designates a confidence value to each of the phosphorylated site. *PhosphoScore* is another algorithm which acts similarly to *Ascore*. *PhosphoScore* considers both the match quality and the normalized intensity of observed spectra peaks compared to a theoretical spectrum. *PhosphoScore* was employed successfully in the studies done by Ruttenberg and colleagues (Ruttenberg, Pisitkun, Knepper, & Hoffert, 2008). *PTMap* is a sequence alignment software used to identify protein PTMs and polymorphisms (Chen, Chen, Cobb, & Zhao, 2009). The selection of peak, adjustment of inaccurate mass shifts and precise localisation of PTM sites are the features that improved searching speed and accuracy of *PTMap*. This software is the first algorithm that contains a scoring system which concentrates on unmatched peaks to eliminate false positives, thus increasing the accuracy and sensitivity of the PTM identifications.

There are numerous database which provide information on PTMs. *dbPTM* database (Lee et al., 2006) gathers various information such as the catalytic sites, protein domains and protein variations, in addition to these software or tools. These databases include a majority of experimentally validated PTM sites from SwissProt and *Phospho.ELM*. *Phospho.ELM*

comprises over 40,000 amino acid serine, threonine and tyrosine non-redundant phosphorylation sites from vertebrates, *Drosophila melanogaster* and *Caenorhabditis elegans* (Dinkel et al., 2011). Similarly, *PHOSIDA* (www.phosida.com) comprises more than 80,000 phosphorylated, N-glycosylated or acetylated sites from nine different species (Gnad et al., 2011). For each of the phosphosites, *PHOSIDA* lists matching kinase motifs, predicts secondary structures, conservation pattern, and its dynamic regulation upon stimulus. Unfortunately, none of these species are plants. *PhosphoSitePlus* (www.phosphosite.org) has 130,000 non-redundant modification sites, primarily on phosphorylation, ubiquitinylation and acetylation (Hornbeck et al., 2012).

2.8 Research outline

This thesis aims to develop robust methodologies specifically to extract chromoplast proteins from the oil palm fruit mesocarp and isolate the enzymes catalysing fatty acid biosynthesis from those extracted proteins. As mentioned in the literature review, in addition to interfering compounds, plant proteins are usually present in low abundance, especially when they are regulatory enzymes. These challenges required the enhancement of general organelle protein extraction approaches, which will be described in Chapter 3.

Subsequently, different gel-based and non-gel based proteomic techniques will be used to characterise and quantitate these fatty acid biosynthetic enzymes in terms of their expression (Chapter 4) and phosphorylation states (Chapter 5). Since regulatory enzymes are generally present in low abundance, complementary bottom-up proteomic techniques will be applied to capture the trypsin-digested proteins prior to mass spectrometric analyses. Quantitative analysis of protein expression level based on two different varieties of oil palm (low and high oleic acid varieties) will help to determine the role of the individual fatty acid biosynthetic enzymes as well as other metabolic enzymes in regulating the different level of oleic acid in these oil palm varieties.

Post-translational modified proteins are likely to present in low abundance, thus conferring another complication. Therefore, sensitive mass spectrometry-based techniques such as neutral loss-triggered MS^3 and Selected Reaction Monitoring will be used to reliably identify phosphorylated peptides through detecting the loss of the phosphate group. Determination of the phosphorylation state of fatty acid biosynthetic enzymes will provide better insight into the role of phosphorylation in the regulation mechanism of plant fatty acid biosynthesis.

3.0 Optimisation of oil palm chromoplast isolation and protein extraction procedures for proteomic analyses

3.1 Introduction

Plant fatty acid biosynthesis involves a series of pathways that require tight regulation orchestrated by different proteins. Lipid biosynthesis has been shown to occur in plastids (Awai et al., 2001; Benning, 2008, 2009b; Harwood, 1996; Heinz & Roughan, 1983; Joyard & Douce, 1977; Joyard et al., 2010; Ohlrogge & Browse, 1995; Ohlrogge & Jaworski, 1997; Roughan & Slack, 1982; Slabas et al., 1994; Slabas & Fawcett, 1992; Stumpf, 1969), both in photosynthetic plastids or leaf chloroplasts and non-photosynthetic plastids such as flower and fruit chromoplasts. The chromoplast has an important role in flower and fruit colouration as well as in promoting plant reproduction through pollination (Juneau et al., 2002). In this study, the oil palm mesocarp chromoplast plays an important role in harbouring the fatty acid biosynthetic factory apart from other essential metabolic pathways (Lopez-Juez & Pyke, 2005; Neuhaus & Emes, 2000).

Proteomics has been an effective way to study protein composition and cellular functions of subcellular organelles such as chromoplast. Several proteomic studies on fruit chromoplast have been published (Barsan et al., 2010a; Hansen & Chiu, 2005; Siddique et al., 2006; Tetlow, Bowsher, & Emes, 2003; Zeng et al., 2011a). Superior quality and effective recovery of proteins is crucial for a thorough proteomic analysis. However, compared to other organisms, plants impart unique problems when it comes to protein extraction for proteome studies. Plant tissues possess lower amounts of proteins and extraction of these proteins is frequently hampered by the presence of interfering compounds, such as cell walls, polysaccharides, lipids and phenolics. A vast proportion of plant tissue consists of water which provides the turgor pressure for cell integrity (Saravanan & Rose, 2004) (Figure 3.1). Unfortunately, these characteristics lead to the lower protein amount per cell mass. The cell wall is an important plant tissue component (Figure 3.2) (Xie, Pan, Liu, Ye, & Huo, 2007) which is difficult to disrupt by simple mechanical force. It also hampers protein recovery as membrane proteins tend to bind non-specifically to its polysaccharide matrices. Secondary metabolites, such as phenolic compounds, lipids and polysaccharides, are produced in higher amounts in plant tissues.

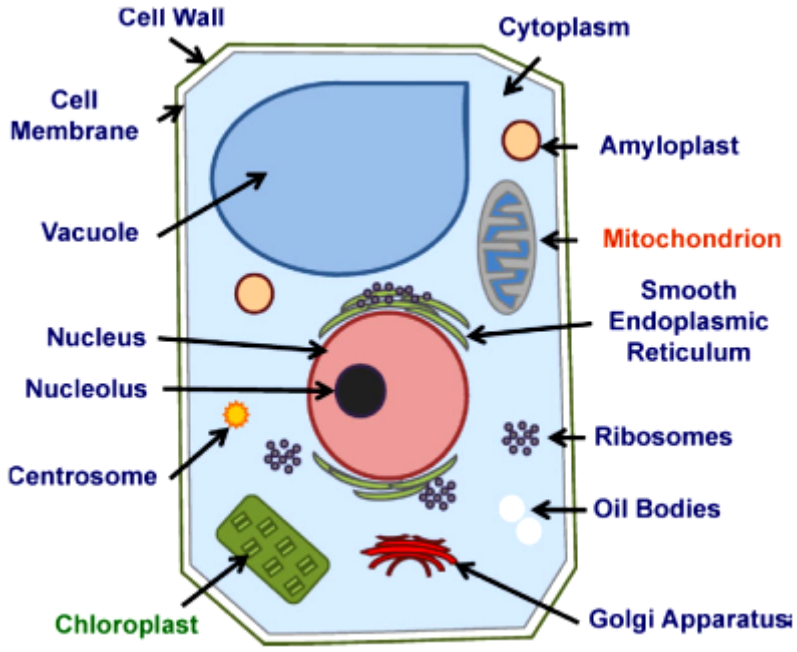


Figure 3.1. The plant cell (Agrawal et al., 2011).

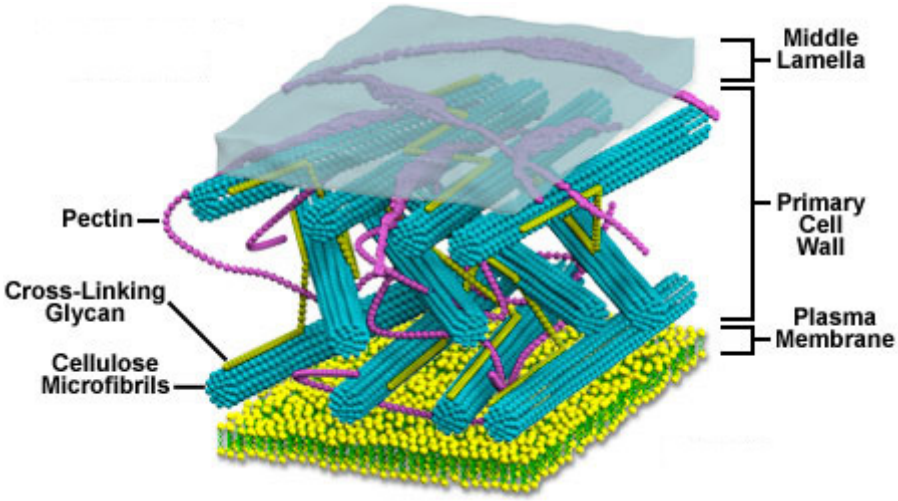


Figure 3.2. Components of plant cell wall (Source: http://micro.magnet.fsu.edu/cells/plants/cellwall.html).

These metabolites play vital roles in the makeup of plant structure and defence systems. Unfortunately, they interfere with the protein extraction, resulting in 25% less protein extracted (Carpentier et al., 2005). In addition, these secondary metabolites also affect protein separation and focussing (based on isoelectric point of protein) using gel electrophoresis and isoelectric focusing techniques. For instance, carbohydrate caused the protein sample to be highly viscous and block the gel pores. Carbohydrate also prolonged the protein focusing time in isoelectric focusing. Resolution of extracted proteins with gel electrophoresis can be very

poor due to interference of lipid contaminant. The *Elaeis guineensis* Jacques var. Tenera used in this study, is known for its high oil content (approximately 85% of the dry mass) (Dussert et al., 2013). Thus, the protein extraction strategy is a decisive issue to address when embarking on a plant-based proteomic analysis.

Working on a specific organelle reduces the complexity of the protein compositions compared to the proteome of whole cells or tissues. This approach allows analysis of biological and functional aspects of the organelle, in this case, the fatty acid biosynthesis regulation. In this study, two approaches were used to reduce sample complexity. Initial differential centrifugation (low speed and short centrifugation time) enriched the chromoplast by selectively removing other organelles (Agrawal et al., 2011). In the second stage, density gradient centrifugation was employed. In principle, organelle concentrates at a specific gradient generated where the organelle density is equal to the density of the surrounding medium. This phase is also known as the isopycnic point during centrifugation. Nonetheless, the same organelle would have different rates of sedimentation in the density-gradient centrifugation as the sedimentation rate depends very much on the organelle content, lipid to protein ratio, size and shape (Araùjo, Hube, & Stasyk, 2008). Numerous protocols to isolate other organelles or cell compartments have been reported and summarised are in Figure 3.3.

'*Every species of plant is a law unto itself*' (Gleason, 1926) means there is no general methodology for protein extraction that can be applied to all plant tissues. To our knowledge, this is the first report of organelle protein extraction from *E. guineensis* var. Tenera. Several studies involving plant oil seeds and olive leaf showed that additional steps are often needed in preparing these samples for protein extraction (Wang et al., 2003; Wang, Vignani, Scali, & Cresti, 2006; Wang et al., 2004). Earlier studies indicated that oil palm mesocarps were recalcitrant to typical protein extraction approaches (Lau, personal communication, 1 July 2012). Protein extraction from the chromoplast required a more labour intensive strategy. Direct recovery of proteins after organelle disruption by means of sonication in extraction buffer proved to be inadequate for oil palm mesocarp. A phenol based approach was employed to isolate the proteins from other surrounding materials such as lipid and nucleic acids. This method has been reported by Hurkman and Tanaka for proteomic analysis (Hurkman & Tanaka, 1986). Phenol extraction was not developed to extract protein initially as the approach is more commonly used to deproteinate carbohydrates and nucleic acid for

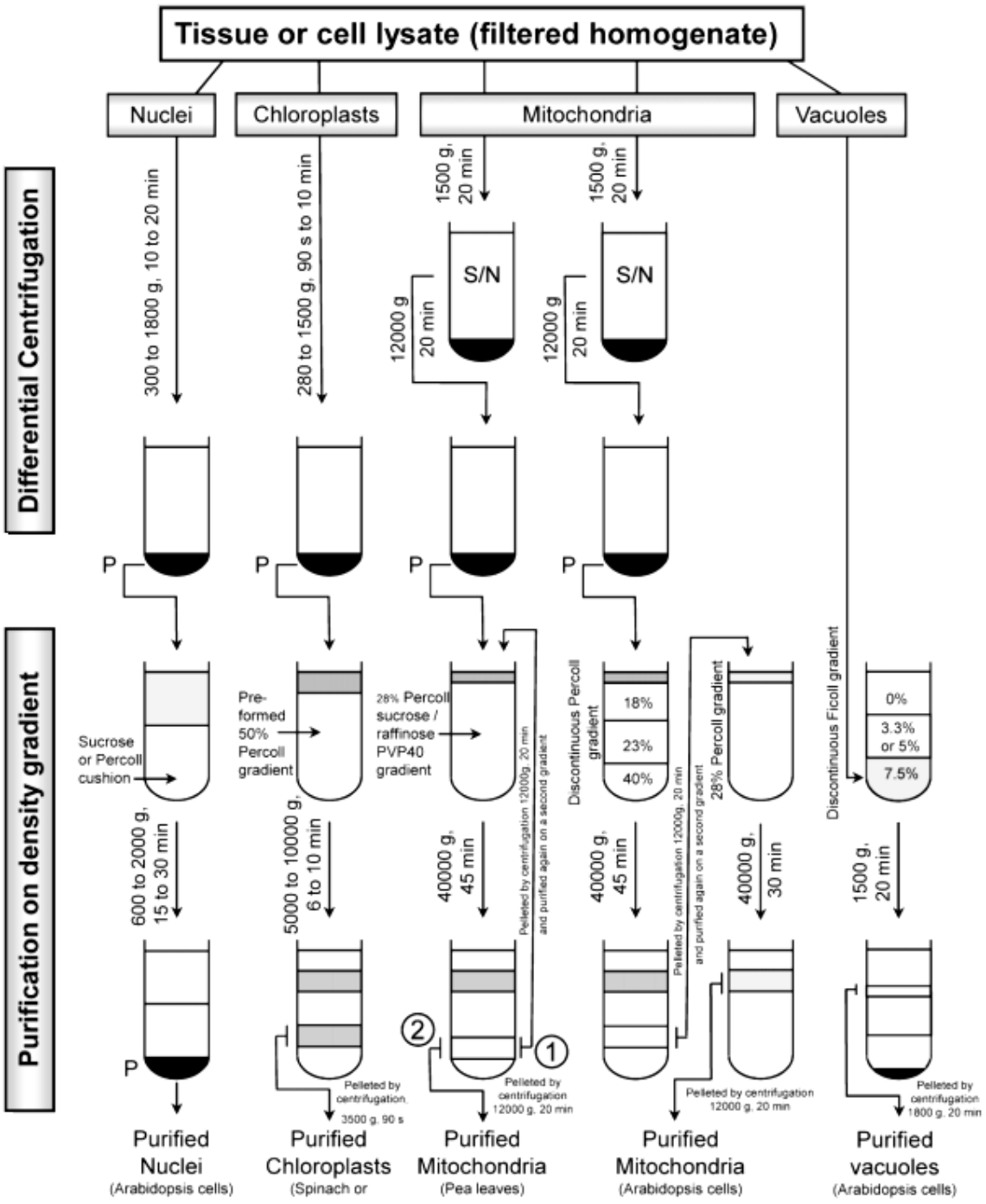


Figure 3.3. Reported protocols to isolate plant organelles (Agrawal et al., 2011).

molecular studies. Now, it is widely utilised to extract proteins from recalcitrant plant tissues such as root (Xie et al., 2007) and leaf (Fan, Wang, Kuang, & Li, 2009; Gómez-Vidal et al., 2009; He & Wang, 2008; Wang et al., 2003). Protein extraction with phenol followed by precipitation using ammonium acetate in methanol or trichloroacetic acid in acetone renders proteins suitable for subsequent mass spectrometric analysis. Therefore, to prepare high quality proteins from oil palm mesocarps, a methodology developed by Wang and co-workers (Wang

et al., 2003; Wang et al., 2006; Wang et al., 2004) was modified to effectively circumvent the obstacle posed by the high oil content of oil palm mesocarps.

Mass spectrometry is a robust technique for measuring the peptide ion masses for protein identification. This approach was preferred to the conventional Western Blot that employed polyclonal antibodies to detect specify proteins in subcellular compartments for organelle purity assessment (Barsan et al., 2010a). The preference was due to the fact that Western Blot depends greatly on the specificity and availability of primary antibodies for the targeted proteins while mass spectrometry techniques do not have such limitations. The results derived from these tandem mass spectrometry analyses were used to assess the purity of isolated chromoplasts by predicting the subcellular localisation of the identified chromoplast proteins (Emanuelsson, Brunak, von Heijne, & Nielsen, 2007; Emanuelsson, Nielsen, Brunak, & von Heijne, 2000). The protein identities are also crucial to enable the functional prediction of the oil palm chromoplast proteins based on rice (*Oryza sativa*), *Arabidopsis thaliana* and maize (*Zea mays*) protein homologues. In general, there are three main approaches in protein identification using tandem mass spectrometry. These are the database-driven approach, *de novo* sequencing and peptide sequence tag approach (Nesvizhskii, Vitek, & Aebersold, 2007). The database-driven approach was applied to determine the identity of proteins from oil palm chromoplasts. Sequence homology searches of these proteins with plastidial proteins from rice, *Arabidopsis* and maize in the Plant Protein Database (PPDB) (Sun et al., 2009) were performed.

This chapter describes the method developments to isolate oil palm chromoplasts and extract their total proteins for mass spectrometric analysis. It is anticipated that the methodologies developed here will facilitate future work on the organelle proteome in oil palm and aid in comprehensive understanding of associated biosynthetic and signalling pathways. This chapter also represents the first report on oil palm plastids utilising advanced proteomics techniques.

3.2 Materials

3.2.1 Plant materials

Fruits (Figure 3.4) of the standard DxP oil palm crosses (*Elaeis guineensis* var. Tenera) (denoted as low oleic acid, LO) and from the PORIM Series 12 (PS12) breeding populations (denoted as high oleic acid, HO) were grown and harvested from the MPOB research stations at Bangi,

Selangor and Hulu Paka, Terengganu, Malaysia. The fruit mesocarps were sliced (Figure 3.4), snap-frozen in liquid nitrogen and stored at –80°C until use. The seeds were discarded. Oil palm bunches (progenies) with different amount of oleic acids (in percentage) at 20th week after anthesis were collected from three LO and HO palms, respectively (Table 3.1). Fruit mesocarps obtained from the bunch were divided into four technical replicates of 100 g each.

3.2.2 General chemicals

Most of the general chemicals were purchased from Sigma (St. Louis, MO, USA), BDH (BDH Laboratory Supplies, Poole, England) and Merck (Merck KGaA, Darmstadt, Germany). Precast gels were from Bio-Rad (Bio-Rad Laboratories, Inc., Hercules, CA), organic solvents from Fisher (Fisher Scientific, Loughborough, UK) and Protease Inhibitor Cocktail Tablets were procured from Roche (Roche Diagnostic, IN, USA). Double-distilled water was used in preparation for all general solutions.

3.3 Methodology

3.3.1 Removal of lipid by organic solvents

Prior to isolation of chromoplasts, it is essential to remove lipids from the sliced mesocarps to avoid any interference in the subsequent protein extraction phase. A method was developed to remove lipids from the oil palm mesocarps based on the method of Wang and co-workers (Wang et al., 2006). Sliced mesocarps were homogenised in liquid nitrogen with a cold Waring blender (Blender B011, Dynamics Corporation, Greenwich, USA) at low speed for 10 s, followed by grinding with a standard ceramic mortar and pestle. The powdered mesocarps were then mixed with cold acetone containing 10% trichloroacetic acid and 1 mM dithiothreitol. The slurry was then centrifuged at 13,000 g for 10 min at 4°C (RA-300 rotor, Kubota 7820, Kubota Corporation, Tokyo, Japan). The supernatant was discarded and the washing step was repeated once. Subsequently, cold 80% methanol containing 0.1 M ammonium acetate was added to the precipitate; mixed and centrifuged as before. After discarding the supernatant, the precipitated mesocarp pellet was washed with cold 80% acetone. The mixture was mixed well and centrifuged again at 13,000 g for 10 min at 4°C. The resulting pellet was air-dried.

Table 3.1. Biological replicates for both the oil palm varieties used in the study.

High oleic acid (HO)			Low oleic acid (LO)		
Palm no.	**Progeny**	**Oleic acid percentage (%)**	**Palm no.**	**Progeny**	**Oleic acid percentage (%)**
0.306/319	PK 540	48.85	0.306/79	PK 540	39.03
0.306/319	PK 540	48.85	0.306/76	PK 540	39.03
0.337/249	PK 1254	48.90	0.337/214	PK 1201	37.90

3

Standard DxP cross (LO) PS12 breeding population (HO) Sliced mesocarps

Figure 3.4. The two oil palm varieties used in the study.

3.3.2 Isolation and subsequent purification of oil palm chromoplasts

The isolation method of chromoplasts was adapted from Jain, Fan and their co-workers (Fan et al., 2009; Jain, Katavic, Agrawal, Guzov, & Thelen, 2008). Washed mesocarp pellet was transferred into a beaker containing cell wall digestive enzymes (2% w/v cellulase, 0.1% w/v pectinase, 0.6 M sorbitol, 0.1 M DTT, 5 mM 2-(4-morpholino)-ethane sulfonic acid (MES)-KOH, pH 5.5). The suspension was then incubated in an incubator shaker (Labnet 211DS, Labnet Instrument, Inc., NJ, USA) for 6 h at 37°C. After cell wall digestion, the mixture was sieved through two layers of Miracloth (Calbiochem, EMB Millipore Corporation, Billerica, MA) into a fresh beaker on ice to separate non-macerated plant materials from protoplasts. After that, the filtrate was centrifuged at 1750 g for 5 min at 4°C (RA-300 rotor, Kubota 7820) to collect intact chromoplasts. The chromoplast pellet was gently re-suspended in the extraction buffer (0.7 M sucrose, 1 M Tris-HCl, pH 8.3, 5 M NaCl, 50 mM DTT, 1 mM EDTA containing a tablet of Roche protease inhibitors for every 50 mL of buffer). The mixture was agitated gently for 5 min to avoid the disruption of the chromoplasts before proceeding to its purification according

to Barsan and co-workers (Barsan et al., 2010a) with some modifications to the sucrose gradient concentrations used. Following re-suspension, the pellet was loaded onto a gradient of 1.2 M, 1.7 M and 2.2 M sucrose (in 5 mL of extraction buffer each) in a centrifuge tube. After that, they were centrifuged (JA-30.50 rotor, Avanti J-301, Beckman Coulter, Brea, CA) at 62,000 g for 45 min at 4°C to yield three interphases. The interphases were carefully retrieved and retained for protein extraction.

3.3.3 Organelle protein extraction and precipitation

We used both the chromoplast pellets (before purification) and interphases for subsequent proteomic analysis. The re-suspended chromoplast pellets in extraction buffer (0.7 M sucrose, 1 M Tris-HCl, pH 8.3, 5 M NaCl, 50 mM DTT and protease inhibitor cocktail) or the interphases were sonicated for 15 min at 4°C to release the proteins from the plastids. An equal volume of fresh 50 mM, pH 8.0 Tris-saturated phenol was added to the mixture thereafter. The mixture was further agitated for 10 min just before centrifugation at 15,000 g for 15 min at 4°C (RA-300 rotor, Kubota 7820) for phase separation. Following that, the upper phase was transferred to a new centrifuge tube while the bottom phase was discarded. Chromoplast proteins were precipitated by adding five volumes of cold ammonium acetate-saturated methanol and incubated at -20°C overnight before being centrifuged at 15,000 g for 15 min at 4°C (RA-300 rotor, Kubota 7820) to obtain the protein pellet. The protein pellets were then rinsed with cold ammonium acetate-saturated methanol until a whitish colour was obtained for the pellet. This was followed by washing the pellet with cold 80% acetone for three times. Proteins were precipitated after each wash by centrifuging at 15,000 g for 5 min at 4°C (RA-300 rotor, Kubota 7820). At the end of the washing phase, the supernatant was carefully decanted before the chromoplast protein pellet was air-dried.

3.3.4 Protein quantitation

The presence of interfering compounds often renders protein estimation using conventional protein assays such as Bradford unusable. Therefore, commercially available 2D Quant Kit (GE Healthcare Life Sciences, Uppsala, Sweden) was used to determine protein content in the samples. Bovine serum albumin provided with the kit was used as the protein standard. Each quantitation was performed in duplicate.

3.3.5 One dimensional (1D) sodium dodecyl sulphate-polyacylamide gel electrophoresis (SDS-PAGE)

1D SDS-PAGE was used to evaluate the overall purity and quality of protein extracts. The approach was adapted to fractionate the extracts in order to enhance the prospects of detecting less abundant proteins and consequently increasing the efficiency of the proteomic analysis. For this purpose, the precipitated chromoplast proteins (from the first stage of fractionation) were dissolved in Laemlli buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 25% glycerol, 0.01% bromophenol blue, 0.005% β-mercaptoethanol) and denatured by heating at 95°C for 4 min. 45 µg protein per lane was loaded onto a 1.0 mm precast Mini-PROTEAN® TGX™, 4-20% polyacrylamide gel (Bio-Rad Laboratories Inc., Hercules, CA). Electrophoresis was conducted in a Bio-Rad mini-PROTEAN® Tetra Cell apparatus (Bio-Rad Laboratories) at 150 V for 1 h. Following electrophoresis, the separated proteins were fixed for 30 min in a fixing solution (50% ethanol, 10% acetic acid) and stained with an in-house prepared Colloidal Coomassie G-250. The gel was destained with Milli-Q water (H_2O) until the background was clear before acquiring the gel image with a DLSR camera (Nikon D100, Nikon Corporation, Japan). The settings were ISO 200, f16 for the aperture and shutter speed was $1/40^{th}$ of a second. Image processing was done with the Analysis software (Soft Imaging System, Olympus Corporation).

3.3.6 Two dimensional gel electrophoresis (2DE)

2DE was used in this study to evaluate the effectiveness of different extraction approaches in terms of number of spots obtained and spot intensities. In addition to that, this technique was also used to assess the effects of solvent washes and CDE on the mesocarp protein yield. Dried protein pellet acquired after overnight precipitation was solubilised in rehydration buffer (RB) (7 M urea, 2 M thiourea, 4% CHAPS, 0.25% Pharmalyte and 0.4% DTT). The volume of RB needed was very much dependent on the amount of proteins acquired from the different protocols evaluated (in the case of method development for protein extraction). Sonication for about 15 min was required to assist the resolubilisation of the protein pellet. Any non-solubilised proteins after the sonication were removed by centrifugation. Protein quantitation was carried out prior to isoelectric focussing (IEF). For IEF, Bio-Rad ReadyStrip™ IPG strip, 7 cm, pH 3-10 (Bio-Rad Laboratories) were rehydrated with 125 µL of RB containing approximately 100 µg proteins overnight. For comparison between the LO and HO varieties, Bio-Rad ReadyStrip™ IPG strip, 11 cm, pH 3-11 were rehydrated with 200 µL of RB containing about 400 µg proteins overnight. IEF was performed according to the following focusing

Table 3.2. Focusing parameters used for isoelectric focusing.

Phase	Volt (V)	mA (Max)	Time (hours)	Volt hours
1	100	20	3	300
2	500	20	2	1000
3	1000	5	2	2000
4	8000	5	14	84000

parameters (Table 3.2) set on the Bio-Rad Protean IEF Cell (Bio-Rad Laboratories) at 20°C.

For the second dimension, a precast 4-20% gradient SDS-PAGE gel (Bio-Rad Laboratories Inc.) was used. After the rehydration, proteins in the IPG strip were reduced with 1% DTT (in RB) for 15 min followed by alkylation with 4% iodoacetamide for another 15 min before transferring to the gel (second dimension) for separation.

3.3.7 Protein digestion assisted by sodium deoxycholate

In-gel digestion – Each lane from the 1D SDS-PAGE separation was cut into seven uniform slices denoted as S1, S2, S3, until S7 and washed in 200 mM ammonium bicarbonate in 50% acetonitrile for 1 h at 37°C. Proteins were reduced and alkylated with thiol-free 50 mM tris(2-carboxyethyl)phosphine (TCEP) and 55 mM iodoacetamide, respectively. An ionic detergent, sodium deoxycholate (1%, 4 µL) was added to the dehydrated gel plugs prior to protein digestion with modified sequencing grade trypsin (50:1 µg) (Promega, Madison, WI, USA) in 50 mM NH_4HCO_3 for 16 h at 37°C. The solubility of the proteins was increased in the presence of sodium deoxycholate and this would enhance trypsinisation (Koehn et al., 2011; Lin et al., 2008; Proc et al., 2010). The resulting peptides were extracted by increasing the concentrations of acetonitrile gradually from 10% to 80% and pooled. The pooled peptide extract was acidified with 0.5% formic acid prior to centrifugation at 10 000 g for 15 min to precipitate out sodium deoxycholate (SDC). The digests were then dried in a centrifugal evaporator (CentriVap Concentrator, Labconco, MO, USA). *In-solution digestion* – About 50 µg of precipitated proteins were re-suspended in 0.1 M ammonium bicarbonate before reduction and alkylation using 50 mM TCEP and 55 mM iodoacetamide, respectively. A 4 µL of SDC was incorporated and removed accordingly as previously described for in-gel protein digestion. After digestion with 1 µg of trypsin at 37°C for 16 h, the peptide solution was dried in a centrifugal evaporator (CentriVap Concentrator).

3.3.8 LC-MS/MS for organelle purity assessment

Separation and identification of the protein digests was conducted with a nano-Advance Splitless nano-liquid chromatography (nanoLC) system coupled to an amaZon speed ETD ion trap mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). Trypsin digests were reconstituted in 30 µL of 0.1% formic acid (FA) and 5% acetonitrile (ACN) and 5 µL injected into the nanoLC system (Bruker Daltonik) for analysis. The peptides were separated with an in house-packed C18 Phenomenex Aeris XB trap column (3 µm, 0.1 x 100 mm) (Phenomenex, Inc., Torrance, CA) and a prepacked Magic C18 AQ analytical column (3 µm, 0.1 x 150 mm) (Bruker-Michrom Bioresources, Inc., Auburn, CA). Equilibration of the columns were performed with 95% solvent A (2% ACN, 0.1% FA) and 5% solvent B (98% ACN, 0.1% FA). Gradient from 0-45% solvent B was employed to elute the bound peptides over 45 min at a flow rate of 800 nL min^{-1}. The eluted peptides were electrosprayed into the ion trap mass spectrometer with a spray voltage of 1500 V and a capillary temperature of 150°C. Mass spectrometric survey scan was acquired in the ion trap with mass range from *m/z* 310-1400. The resolution for the survey scan was set to 'Enhanced Resolution' and the scanning speed was 8,100 u per sec. MS/MS conditions consisted of 'Xtreme Resolution' scan range from *m/z* 100-3000. Up to three of the most intense multiple charged ions (1+, 2+, and 3+) per scan were fragmented via collision induced dissociation in the linear ion trap to obtain tandem mass spectra. All tandem mass spectra were collected using a 1.00 V of fragmentation amplitude, an isolation window of 4.0 u and a scanning speed of 52,000 u per sec.

3.3.9 Data acquisition and analysis

Data acquisitions were performed with ESI Compass 1.3 for amaZon (trapControl Version 7.0). The mass spectrometer was operated in data-dependent mode and calibrated using an electrospray Tuning Mix (Agilent Technologies, Inc., Santa Clara, CA). Data analysis was performed using Bruker Compass DataAnalysis (Version 4.0, SP5). Peak lists were generated using the DataAnalysis Script Editor. These peaks list were then sent to Bruker ProteinScape data management (Version 3.1) for protein identification. The protein identifications were made with tandem mass spectra in the peak list using Mascot software (Matrix Science, London, UK). The peptide sequences were searched against the NCBI non-redundant (nr) protein database. For gel segments, individual peak list was created for each gel segment and individual Mascot searches carried out. For the interphases, peak list from each interphase analysis, was merged into a single MGF file using in-house "Concatenate" software and

searched against the Magnoliophyta database (308,693 sequences) in NCBInr. Mass tolerances for peptide and fragment ions were set to 0.3 Da and 0.6 Da, respectively; and the instrument setting was specified as 'ESI trap'. Semi-trypsin was designated as the protease and two missed cleavage were allowed. Carbamidomethylation on cysteine residue was set as the fixed modification while oxidation of methionine and deamidation of asparagine and glutamine residues were searched as variable modifications. Proteins were accepted if they had at least one top ranking peptide (Rank 1) with a Mascot score of more than 50.0 ($p < 0.01$). All database searches were also performed on the decoy database with a 2% the false discovery rate (FDR). Decoy database contained randomised sequence of Magnoliophyta taxonomy. A unique non-redundant protein list was compiled by homology-based matching with the Magnoliophyta taxonomy.

3.3.10 Protein description

Protein descriptions were performed either using annotations related to each protein entry or through homology-based comparisons with established plant protein databases using ClustalW multiple alignment software. Identified proteins were functionally classified employing MapMan Bins (http://mapman.mpimp-golm.mpg.de/pageman/index/shtml) using annotations retrieved from rice, Arabidopsis and maize databases. Homologues of the identified proteins were searched against Plant Protein Database with a cut-off expectation value of $1E^{-100}$ of having significant homology. Predictions of subcellular localisation were carried out using TargetP 1.1 software (http://www.cbs.dtu.dk/services/TargetP). Predictions were made based on the predicted presence of any of the N-terminal presequences: chloroplast transit peptide, mitochondria targeting peptide or secretory pathway signal peptide. These sequences essentially determine the final organelle localisation after their synthesis in the cell cytosol. Otherwise, predictions were made using rice, *Arabidopsis* and maize homologues. Those proteins whose close homologues were experimentally demonstrated to localise exclusively outside plastids were considered non-chromoplast contaminants. A protein is considered a 'close homologue' if the protein has at least 70% pair-wise identity with the query sequence and/or a cut-off expectation value of $1e^{-4}$ following Basic Local Alignment Search Tool for proteins searches against the NCBInr protein database (www.ncbi.nlm.nih.gov.BLAST/).

3.4 Results and discussion

3.4.1 Delipidating the mesocarps

The mesocarp powder was delipidated by washing with two different organic solvents since different solvents have diverse capabilities to extract various classes of lipids. Figure 3.5 shows the effect of delipidation on the protein yield. It is evident that the presence of excessive lipid decreased the effectiveness of the protein extraction as shown by the 1.6- fold reduction in protein yield. Even though another report contradicted that delipidation enhanced protein yield (Rodrigues, Torres, da Silva Batista, Huergo, & Hungria, 2012), the plant materials used in that study were roots and not oily mescocarp. Delipidation not only generated higher protein yield compared to the undelipidated mesocarps, but delipidated mesocarp proteins were also better resolved in SDS-PAGE (Figure 3.6). This could be due to the presence of excessive lipids, polyphenol and other contaminants in the unwashed samples that can reduce protein resolution. Comparisons of 2DE gels produced by solvent-washed and without solvent-washed mesocarp proteins (Figure 3.7) clearly demonstrated the importance of eliminating the lipids prior to incubating the mesocarps with cell wall digestive enzymes (to be discussed in Section 3.4.2) and subsequent protein extraction (to be discussed in Section 3.4.4). We did not perform spot to spot comparisons for these gels as it was apparent that many protein spots were missing in the without solvent-washed gel and the protein spot intensities were also noticeably lower.

A combination of trichloroacetic acid (TCA) and acetone was used in the first wash, instead of aqueous TCA or acetone alone as the resulting pellet was finer (Figure 3.8) (Damerval, De Vienne, Zivy, & Thiellement, 1986; Görg et al., 1997). The effectiveness of lipid removal from the ground mesocarps depended greatly on the fineness of the sample powder. The surface area of the plant materials is also crucial for an effective cell wall digestion with cell wall digestive enzymes in the later stage (Section 3.4.2). Usage of TCA and acetone substantially increased the fineness of the powder and thereby providing a larger surface area for the enzymes. Other methods have been reported to increase the surface area such as mincing with blades, grinding in a mortar and pestle and shearing in a Waring blender (Jain et al., 2008). In this study, the mesocarps were first ground with a Waring blender utilising low to high grinding speed within 15 seconds before the solvent washes. The grinding time is decisive as Hansen and co-workers found that longer blending time damaged the chromoplast

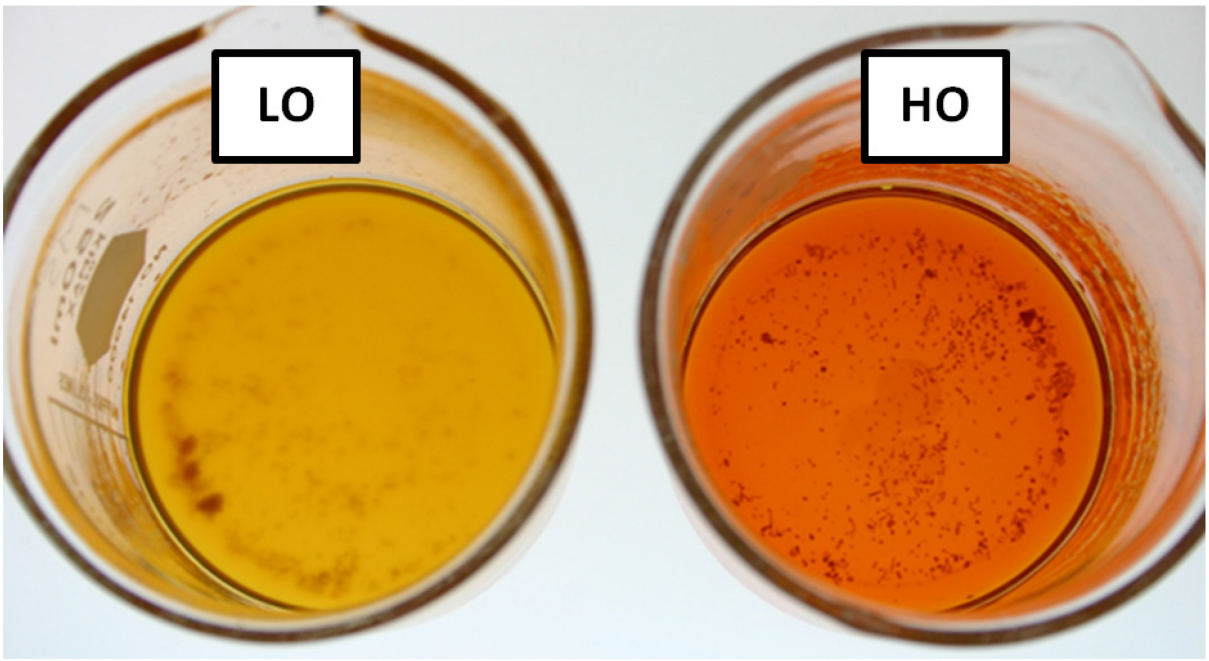


Figure 3.5. Lipid removed from the ground mesocarps.

Table 3.3. Effect of solvent wash on protein yield.

Sample	Protein yield (µg/µL)
Delipidated mesocarps	1.16 ± 0.02
Undelipidated mesocarps	0.71 ± 0.07

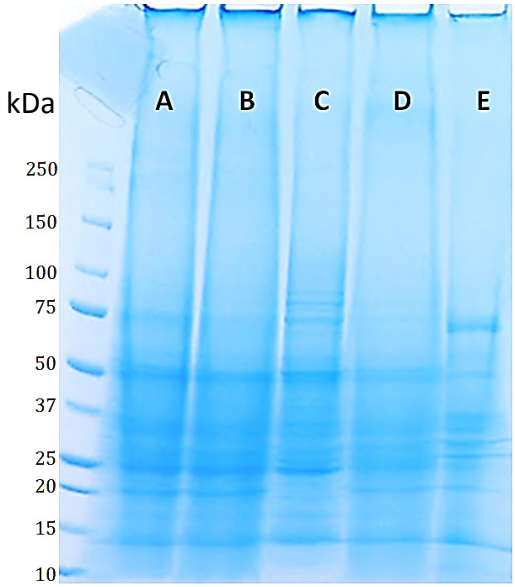
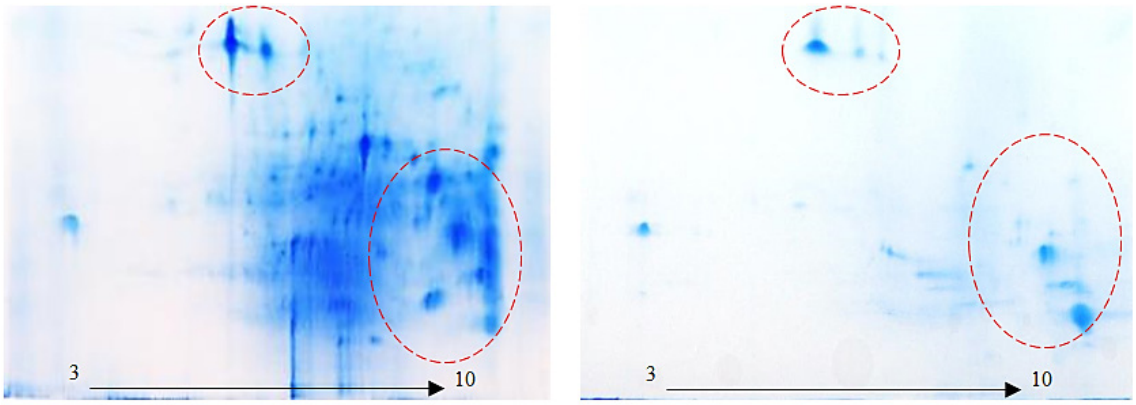


Figure 3.6. Effect of cell wall digestive enzymes. 50 µg of proteins (each lane) was separated with a pre-cast 4-20% gradient SDS-PAGE. From left: 3 hours (A); 6 hours without CDE (B); 6 hours (C); 6 hours without delipidation (D); 10 hours (E). Unless specified, all incubations were done in the presence of cell wall digestive enzyme.



With solvent washes Without solvent washes

Figure 3.7. The effects of washing homogenised mesocarps with solvents before protein extraction. 100 µg of proteins was rehydrated with 7 cm IPG strip, pH 3-10. Proteins were separated with 12% self-casted SDS-PAGE gel.

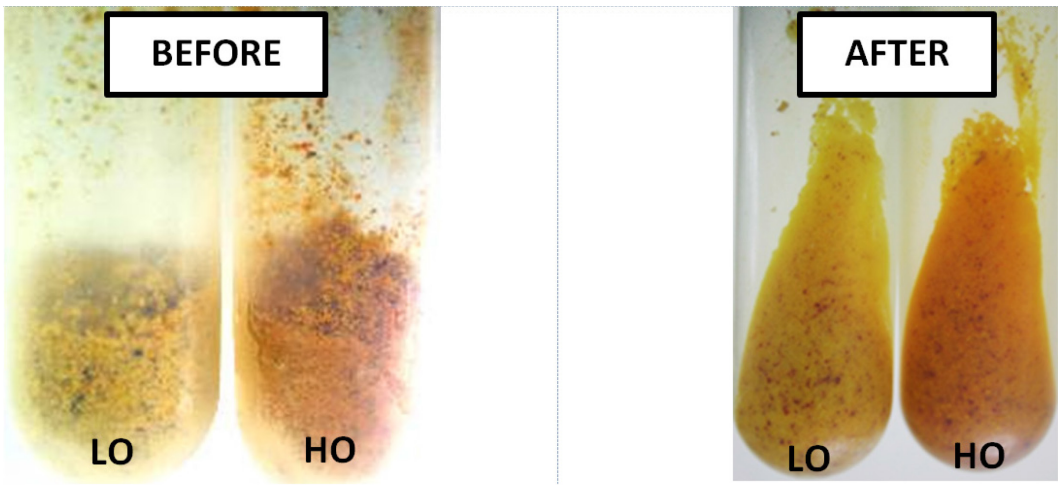


Figure 3.8. Finer ground mesocarps after solvent wash with TCA and acetone.

membrane during isolation (Hansen & Chiu, 2005). Fine powder was also reported to minimise proteolysis and other sources of protein degradation (Wang et al., 2006). The TCA and acetone washing step was repeated for the HO variety due to its significantly higher content of lipid than the LO variety. The first washing was repeated only once as prolonged exposure to low pH caused by TCA could possibly modify or degrade proteins.

Methanol was used in the second wash to further remove the lipid. Polar solvents such as methanol solubilise polar lipids like phospholipids and glycosphingolipids (Christie, 1993). Acetone on the other hand, could only dissolve simple lipids and glycolipids (Rastegari, Ahmad, Spencer, & Ismai, 2011). Methanol has a high dielectric constant and polarity (Wang et al., 2006), which is advantageous for recalcitrant plant tissues due to their high phenolic content. Since TCA was used in the first wash, the presence of ammonium salt in methanol was able to wash off residual TCA in addition to raise the pH to above 7 (Wang et al., 2006). Increase in pH could facilitate protein extractions by phenol, which will be discussed in Section 4.4.4. The last wash involved only acetone to remove the lipid residual. Although the complete removal of lipid could not be achieved, the reduction in lipid amounts was sufficient to increase the protein yield.

3.4.2 Incubation with cell wall digestive enzymes (CDE)

The delipidated mesocarps were incubated in a buffered solution of cellulase and pectinase. Protein yields were increased approximately two-fold from 0.64 ± 0.02 µg/µL to 1.16 ± 0.02 µg/µL when CDE were employed to break the cell wall. Precipitated protein pellets from mesocarps that were not treated with CDE were also smaller in size (Figure 3.9). As our starting materials are limited, it is critical to extract as much protein as possible. The amount of proteins extracted would be reduced even more since we only targeted the chromoplast for further analysis.

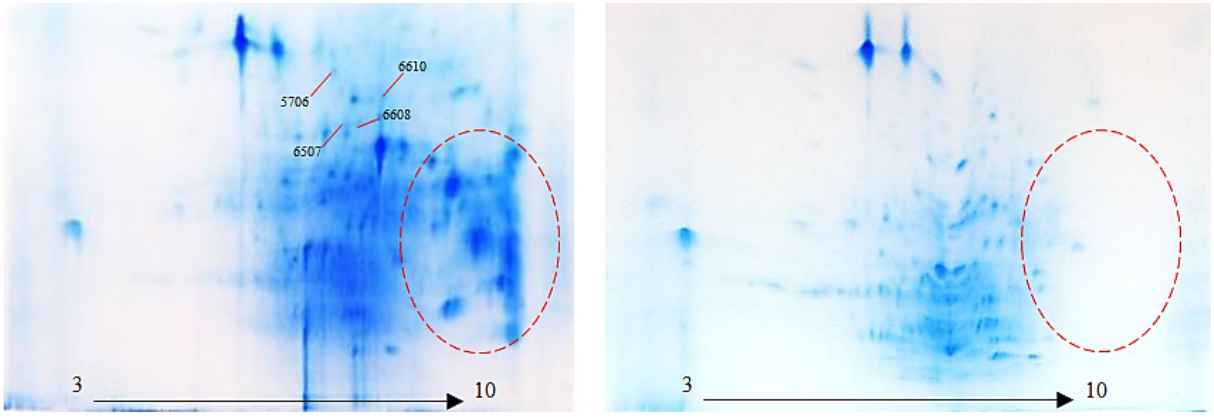
The 1D SDS-PAGE gel profiles (Figure 3.6) showed that several high molecular weight protein bands were missing from the separated mesocarp proteins without usage of CDE. These 1D SDS-PAGE gel profiles agreed with the 2DE gel profiles for both with and without CDE treatments. The 2DE gel profile for proteins extracted from mesocarps treated with and without CDE (Figure 3.10) demonstrated that CDE treatment generated more protein spots in the basic region (mostly membrane proteins) than those obtained without the enzymes (Görg, Weiss, & Dunn, 2004; Vincent, Wheatley, & Cramer, 2006). Lysing the cell wall with CDE would

allow phenol used in subsequent protein extraction, to disassociate the membrane proteins by breaking the lipid-protein interactions. Without CDE, phenol could not reach the membrane proteins to exert its effect. Protein spot to spot comparisons of the same set of 2DE gels also revealed that some protein spots were either absent or present in low intensities in gel produced from separated mesocarp proteins without CDE treatment (Figure 3.11). Several protein spots in the higher molecular weight regions were also not being extracted from the mesocarps without the CDE treatment, as indicated by the 1D SDS-PAGE profiles.

Essentially, cellulase breaks down the 1,4-β-D-glucosidic linkages (Figure 3.12.) in cellulose (Bayer, Chanzy, Lamed, & Shoham, 1998; Hayashi, Yoshida, Woo Park, Konishi, & Baba, 2005) while pectinase hydrolyases the 1,4-α-D-galactosiduronic linkages in pectin (Baldwin & Biggs, 1988; Harholt, Suttangkakul, & Vibe Scheller, 2010). Previous studies have reported the use of mechanical force to disrupt the plant cell wall to yield protoplasts. However, a study by Jain and co-workers revealed that their initial attempts to acquire rapeseed embryoplasts using mechanical disruption had resulted in high levels of contamination by protein storage vacuoles and mitochondrial proteins (Jain et al., 2008). As a result, they opted for cell wall digestive enzymes to lyse the cell for plastid preparations. Protoplast isolation after cell lysis has also been employed in studies involving maize endosperm (amyloplasts) (Echeverria, Boyer, Liu, & Shannon, 1985), mung beans (protein bodies) (Van der Wilden, Herman, & Chrispeels, 1980) and castor bean endosperm (plastids) (Nishimura & Beevers, 1978). Through the results obtained, the digestive enzymes had been demonstrated to be necessary to obtain the protoplasts for efficient protein extraction.

Temperature at 37°C and pH 5.5 were used as these are the optimal catalyst conditions for the cellulose and pectinase. Optimisations on the incubation time were also carried out. According to Jain and co-workers (Jain et al., 2008), 0.5 g of minced embryos from the seed coat was incubated in their enzyme digestion solution (containing digestive enzymes) for 1.5 hours. Therefore, to scale up, using similar materials to hour ratio, 5 g of materials would have to be incubated for 15 hours, which may significantly degrade heat-sensitive proteins (as demonstrated in this study). Our observations indicated that 6 hours incubation yielded about 1.16 ± 0.02 μg/μL protein and 3 hours or 10 hours incubation did not increase the protein yield substantially (Table 3.4). The 1D SDS-PAGE protein profiles for all the different incubation time (and without CDE) also indicated that 6 hours incubation was the most effective in increasing

Figure 3.9. Re-suspension of mesocarps in buffered CDE after 6 hours incubation. A; without CDE, B; with CDE. A larger lipid layer is visible in A. Different sizes of chromoplast protein pellets are shown on the respective sides.



6 hours incubation with CDE

6 hours incubation without CDE

Figure 3.10. The effects of CDE incubation on 2DE gel profile. 100 µg of proteins was rehydrated with 7 cm IPG strip, pH 3-10. Proteins were separated with 12% self-casted SDS-PAGE gel. Missing protein spots in the basic region is highlighted.

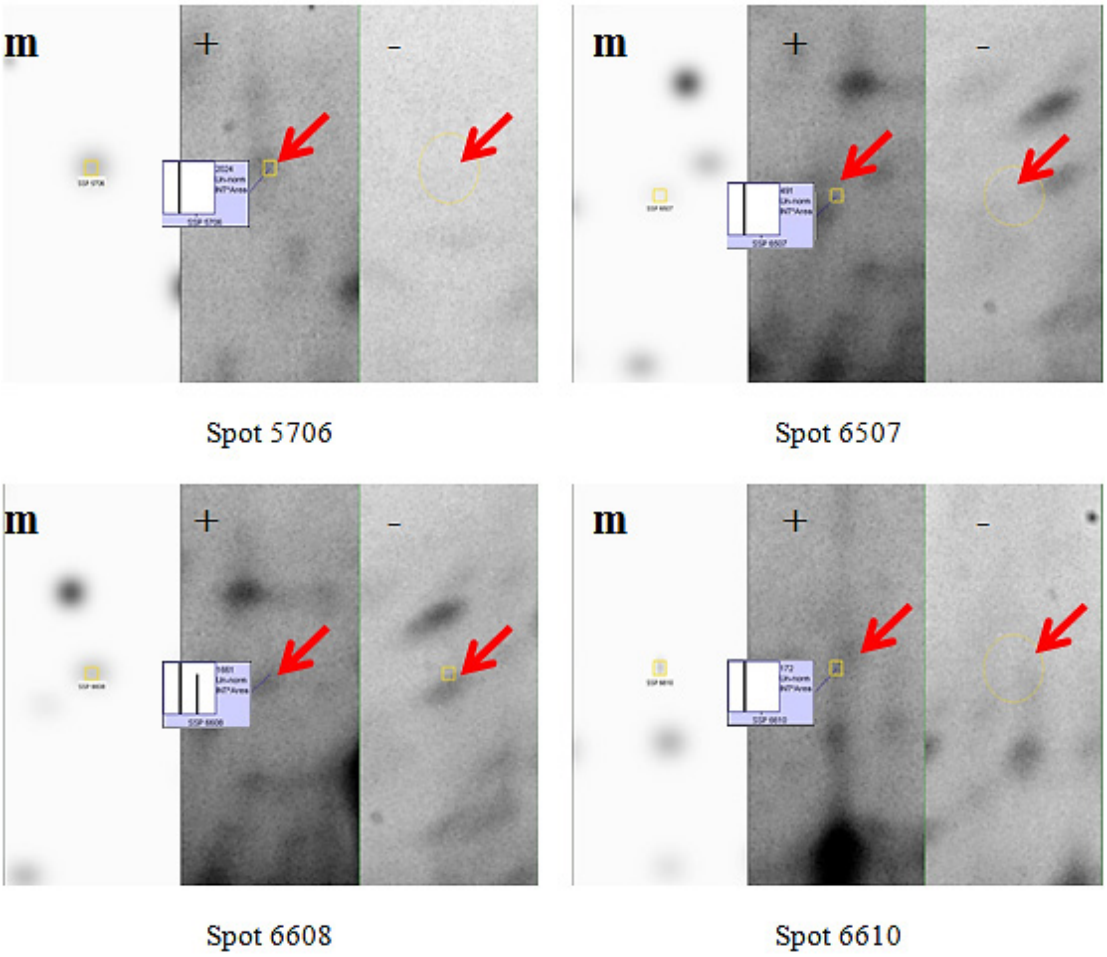


Figure 3.11. Spots analysed by PDQuest showing differences in spot intensities (shown in the box). From left: Master gel (m); with CDE (+); without CDE (-).

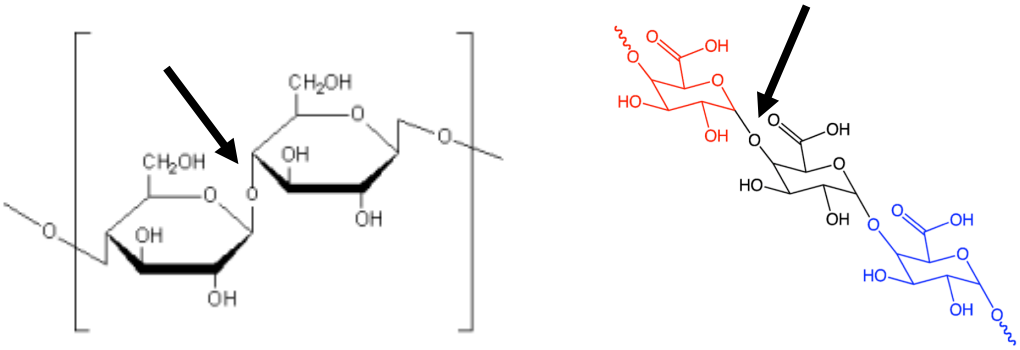


Figure 3.12. Structures of cellulose and pectin. The bond acted by the enzyme is shown by the arrow (Source: http://sciencegeist.net/to-ma-to-to-mah-to/).

Table 3.4. Protein yields from different incubation times. Washed mesocarps were incubated with the cell wall digestive enzymes at 3, 6 or 10 hours to evaluate the protein yields.

Incubation time (hours)	Protein yield (µg/µL)
3	1.01 ± 0.03
6	1.16 ± 0.02
10	1.04 ± 0.01

the yield of chromoplast protein (Figure 3.6).

Figure 3.13 shows the comparisons of the effects of different incubation times on the delipidated mesocarps with CDE using 2DE analysis. We found that after 6 hours incubation, the protein spots were better resolved. The analysis showed that 6 hours incubation produced more protein spots compared to 3 hours and 10 hours incubation. This result clearly implied that 6 hours was the most suitable duration to incubate the mesocarps with CDE to generate the protoplast. The conclusion also supported by the comparative analysis of selected protein spot intensities amongst each set of the three gels (Figure 3.14).

Although 3 hours incubation gave slightly lower yield as compared to 6 hours incubation, any additional proteins are crucial due to the limitations in oil palm materials used. 10 hours incubation with CDE should theoretically increase the protein yields but the results indicated otherwise. The probable explanation was that the cell wall digestive enzymes activity had reached a plateau between 6-10 hours due to limitation factor in the amount of mesocarps used. The long incubation time at 37°C could also lead to degradation of temperature-sensitive proteins, therefore, resulting in the decrease in protein yield after 10 hours incubation. Based on these results, it was decided to incubate the mesocarps with CDE for 6 hours to break the cell walls.

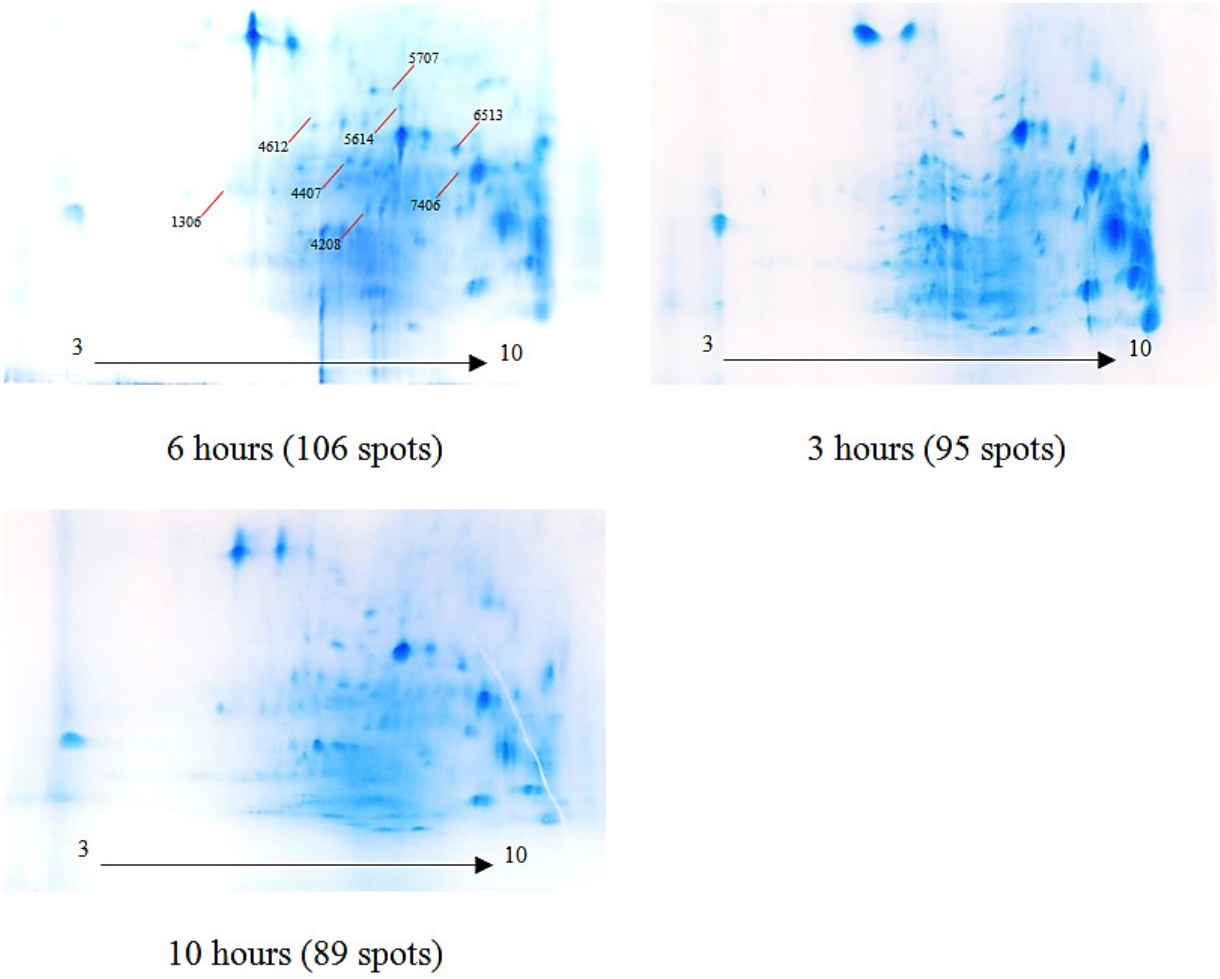


Figure 3.13. The effects of different incubation time of the homogenised mesocarps with CDE. 100 µg of proteins was rehydrated with 7 cm IPG strip, pH 3-10. Proteins were separated with 12% self-casted SDS-PAGE gel.

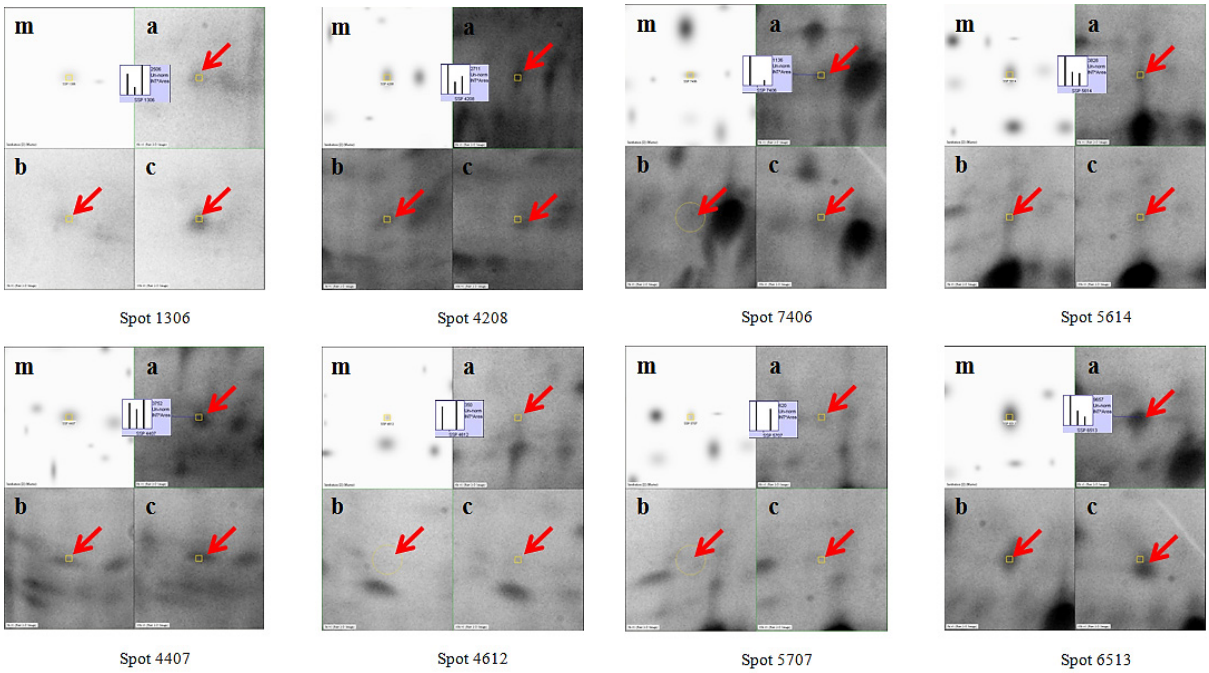


Figure 3.14. Spots analysed by PDQuest that showed differences in spot intensities after incubation at different times (shown in the box). Clockwise from left: Master gel (m); 6 hours (a); 3 hours (b); 10 hours (c).

3.4.3 Chromoplasts isolation

Following cell wall digestion, the chromoplasts were acquired through differential centrifugation before further purification with density gradient centrifugation. Three interphases were clearly visible (Figure 3.15) and each of these interphases was retrieved and their protein yield were determined (Table 3.5).

Differential centrifugation at low speed was used to selectively precipitate and enrich the larger and denser chromoplast from other cell organelles. The resulting fraction was further purified with density gradients made from 1.2 M, 1.7 M and 2.2 M sucrose based on a method by Barsan and co-workers (Barsan et al., 2010a). Despite the fact that most of other cell compartments were removed in the initial differential centrifugation, different purity of chromoplasts was still recorded in each interphase. As shown by the mass spectrometric analysis (Figure 3.16), a small percentage of mitochondria, ribosomes and nuclei was nevertheless co-precipitated in some of the interphases since the centrifugal force used in this study would still precipitate organelles with similar density to chromoplast. Therefore, instead of aiming to get the purest intact chromoplasts at the interface of 1.7 M and 2.2 M, as was achieved by Barsan and co-workers (Barsan et al., 2010a), this approach was used as a fractionation step to reduce the complexity of the chromoplasts mixture. Thus, the possibility

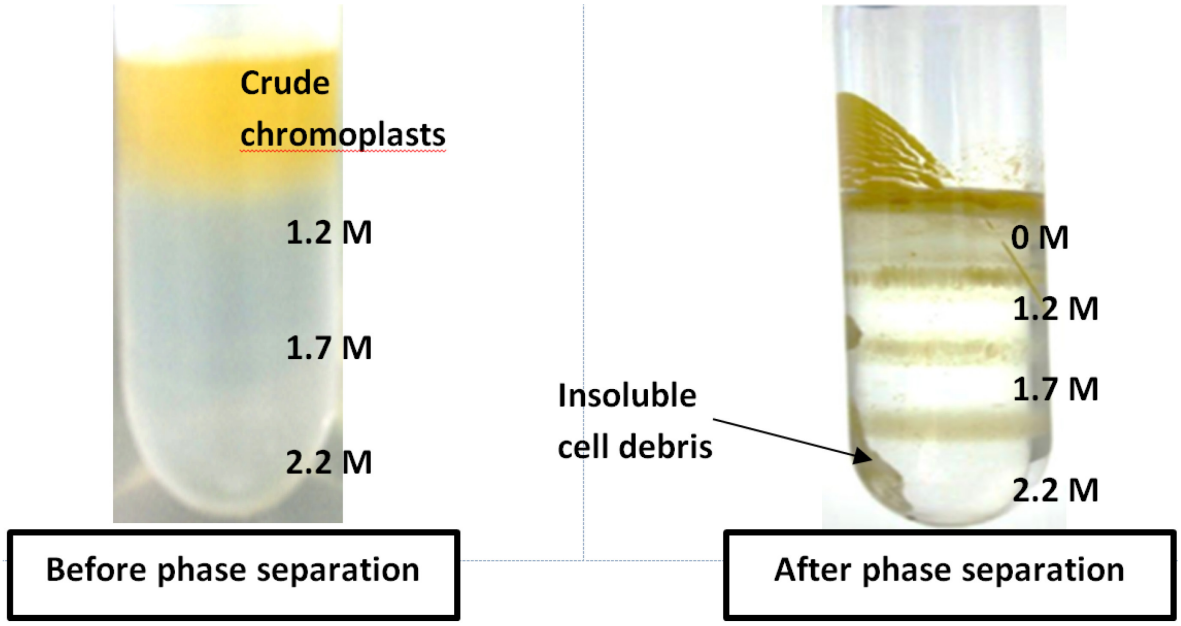


Figure 3.15. Interphases obtained after sucrose density gradient centrifugation.

Table 3.5. Protein yield for the interphases.

Interphase (molarity, M)	Protein yield (µg/µL)
1.2	0.12 ± 0.01
1.7	0.27 ± 0.02
2.2	0.48 ± 0.04

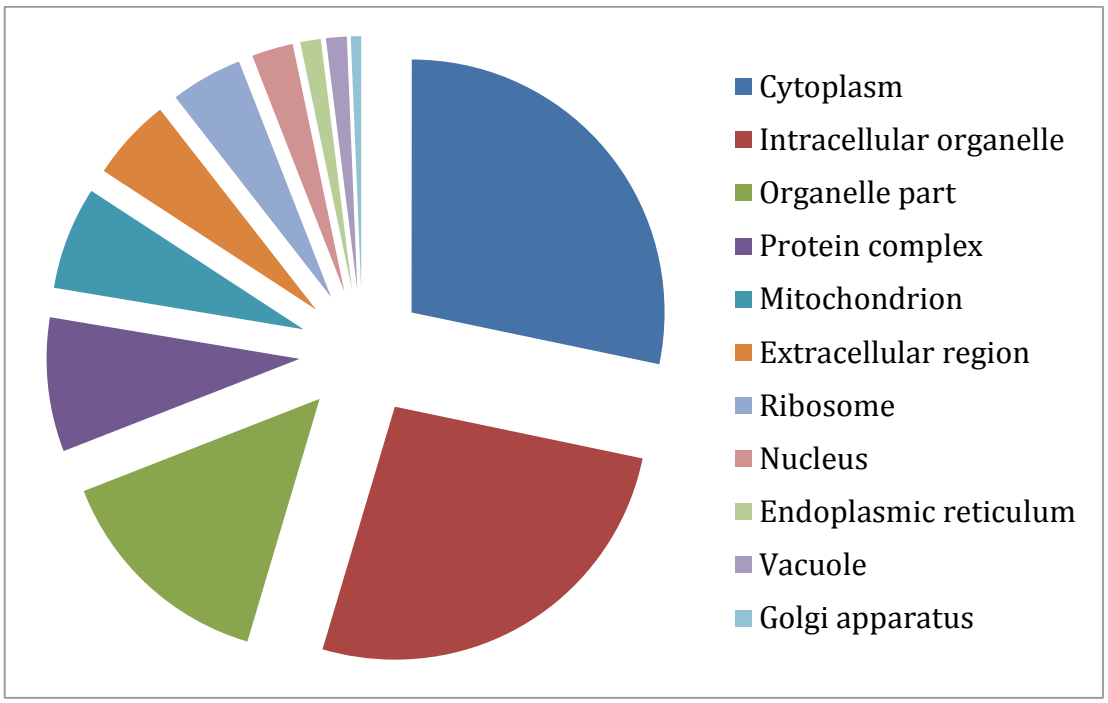


Figure 3.16. Purity assessment of organelle proteins. The quantity of other organelle-based proteins is indicated.

of getting more protein identified using mass spectrometry analysis in the later stage would be increased (Figure 3.17).

Differential centrifugation has been employed to isolate chromoplasts by Hansen and co-workers (Hansen & Chiu, 2005) in their works on tomato. Nycodenz (Axis-Shield PoC AS, Oslo, Norway) had been used to generate the density gradient in chromoplast isolation from sweet orange fruits (Zeng et al., 2011a), Percoll in bell pepper (Siddique et al., 2006); and sucrose in chromoplast isolation from tomato (Barsan et al., 2010a) and wild buttercup (Tetlow et al., 2003). The density gradients could either be continuous or discontinuous. In this study, a discontinuous density gradient consisted of three consecutive concentrations (1.2 M, 1.7 M and 2.2 M) was generated. Conversely in continuous gradient, the density increases linearly along the tube. Typically different organelles get enriched at different interphases of the medium during the centrifugation process. Although continuous gradient provides better resolution in general, the continuous density gradient was unable to be generated reliably.

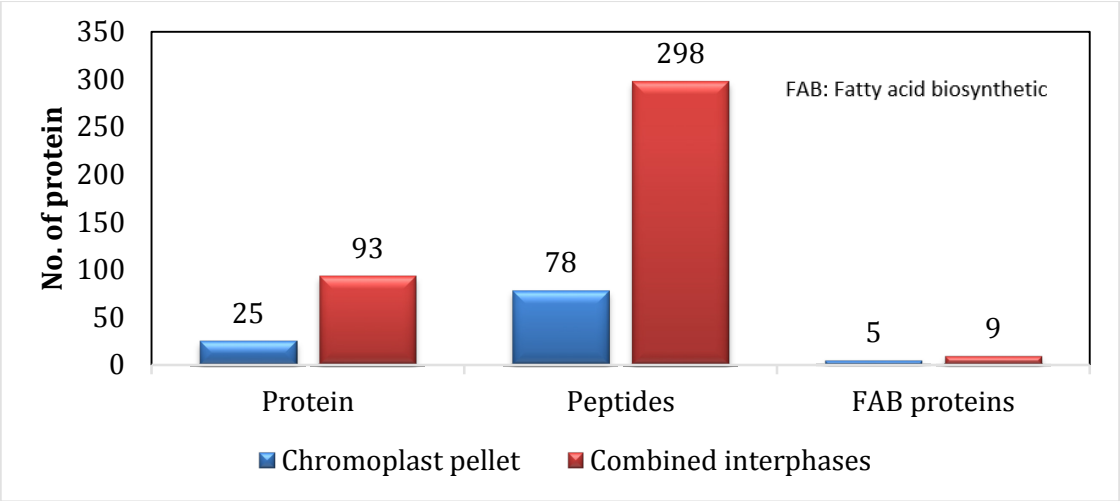


Figure 3.17. Effect of two different centrifugations on identified proteins.

3.4.4 Protein extraction

Proteins were extracted from the chromoplasts following isolation with differential and density gradient centrifugations. At present, there is no known report of proteomic studies on oil palm chromoplast. Since there are no established extraction methods for oil palm chromoplast proteins, different published protein extraction approaches to maximise protein yield for proteomic analysis were attempted (Chatterjee et al., 2012; Rastegari et al., 2011; Rodrigues et al., 2012) (Méchin, Damerval, & Zivy, 2006). Effectiveness and suitability of the approach was assessed based on protein yield, number of resolved protein spots and also intensities of selected spots. In this study, 100% phenol, phenol with sodium dodecyl sulfate (phenol/SDS), SDS alone and combination of TCA and acetone (TCA/acetone) extractions were trialled. Subsequent precipitation of the proteins was performed with either TCA/acetone or ammonium acetate/methanol.

Table 3.6 shows the protein yields generated from evaluation of different protein extraction and precipitation approaches. Protein concentrations for proteins from SDS alone and TCA/acetone extraction were unable to be determined. This was due to insolubility of precipitated proteins obtained with precipitant and co-precipitant provided with the 2D Quant kit (Figure 3.18). The results clearly show that phenol/SDS extraction followed by precipitation with ammonium acetate/methanol outperformed the 100% phenol extraction, regardless of the precipitation technique used. The effectiveness of the extraction protocols were further evaluated by comparing protein patterns produced with 1D SDS-PAGE (Figure 3.19). The protein profiles acquired for each of the extraction methods were consistent with the respective protein yields.

Table 3.6. Protein yield using different extraction approaches.

Extraction approach	Precipitation technique	Protein yield (µg/µL)
100% phenol	Ammonium acetate/methanol	1.84 ± 0.09
Phenol/SDS	Ammonium acetate/methanol	2.22 ± 0.05*
SDS alone	Ammonium acetate/methanol	Not determined
100% phenol	TCA/acetone	2.10 ± 0.03
TCA/acetone	-	Not determined

* Highest yield

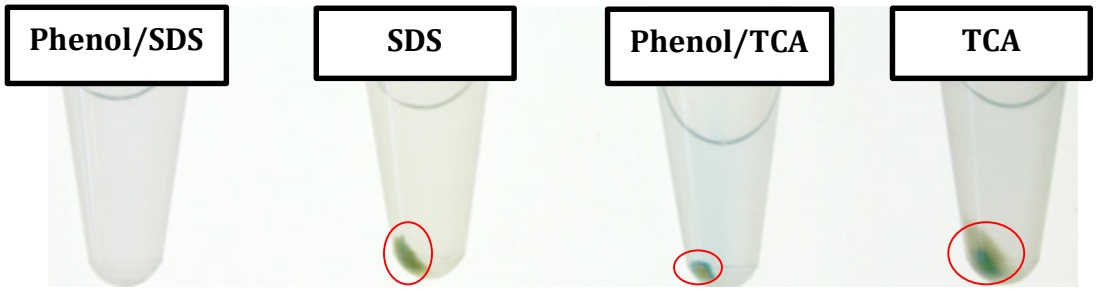


Figure 3.18. Protein pellet after precipitation with the precipitant and co-precipitants (2D Quant kit) were unable to be re-solubilised for protein content determination (indicated by red circles).

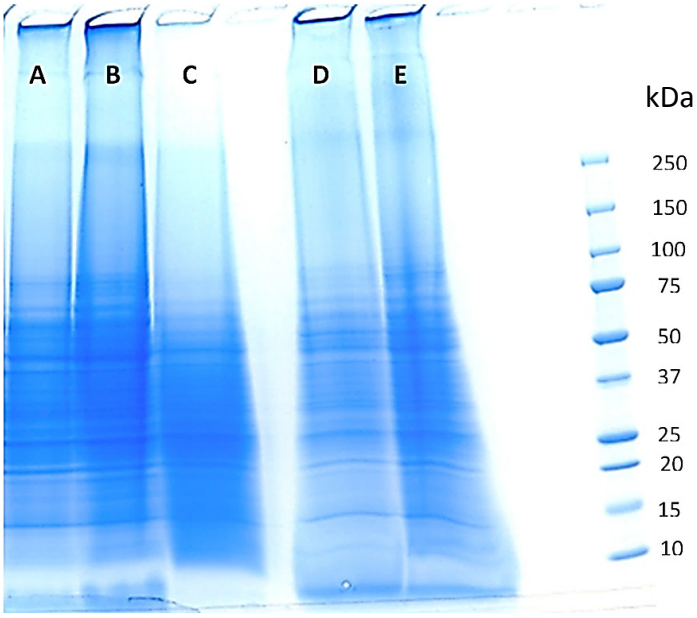


Figure 3.19. Effectiveness of different protein extraction approaches. Approximately 50 µg of proteins was loaded in lane A and B. As protein contents for lane C and D were not determined, the loaded sample volumes were standardised to 45 µL. Proteins were separated with 4-20% gradient SDS-PAGE gel. From left: 100% phenol + ammonium acetate/methanol (A); phenol/SDS + ammonium acetate/methanol (B); 100% phenol + TCA/acetone (C); SDS alone + ammonium acetate/methanol (D); TCA/acetone (E).

Assessment of the extraction approaches were also carried out using 2DE (Figure 3.20). The number of protein spots obtained correlated with the protein yield for each extraction method tested. As expected (based on protein yield), phenol/SDS generated more protein spots (150) compared to either 100% phenol (114 spots) or SDS alone (134 spots). All these proteins were precipitated with ammonium acetate/methanol. Proteins extracted with 100% phenol but precipitated with TCA/acetone produced 73 protein spots. Unfortunately attempts to focus TCA/acetone extracted proteins failed. Contaminants such as salts caused high conductivity in the IPG strips, resulted in burnt strips. Unlike phenol, TCA/acetone was unable to remove the contaminants from these proteins. High protein spot numbers for phenol/SDS and SDS alone extraction compared to 100% phenol indicated that the presence of SDS increased the solubility of some proteins. High protein yields for phenol/SDS (2.24 ±0.05 µg/µL) and SDS alone (not determined) might be attributed to effective extraction of proteins as shown in the 2DE gels. Comparison of 2DE gel profiles of 100% phenol, phenol/SDS and SDS alone, was similar. However, protein spots in 100% phenol extraction gel were better resolved and less streaking was observed. SDS could be detrimental to isoelectric focusing resulting in poor protein spot resolution for both phenol/SDS and SDS alone extracted proteins. However, several good resolution spots in the acidic and lower molecular weight regions were apparent, which could be an advantage. As the targeted fatty acid biosynthesis enzymes are mainly basic proteins, the advantage offered by SDS in solubilising more proteins, hence higher protein yield may not be as great.

Among all of the approaches tested in this study, 100% phenol extraction followed by TCA/acetone precipitation delivered the worst performance in term of number of protein spots and resolution. Only 73 protein spots were produced and there were protein aggregations on the basic region of the gel which caused severe vertical streaking. As stated before, proteins precipitated with TCA/acetone were difficult to dissolve and tended to produce vertical streaking in the gels. However, observations made by Gómez-Vidal and co-workers in their studies on *Phoenix dactylifera* L. leaves (Gómez-Vidal et al., 2009) contradicted this 2DE gel profile. They remarked that TCA probably helped in protein solubilisation. In the 2DE gels for all the methods evaluated, horizontal streaking was still detected, indicating insufficient focusing for the proteins. Spot to spot comparisons for selected protein in all these gels were also performed (Figure 3.21). Overall, comparable protein spot intensities for 100% phenol, phenol/SDS and SDS alone extraction were observed.

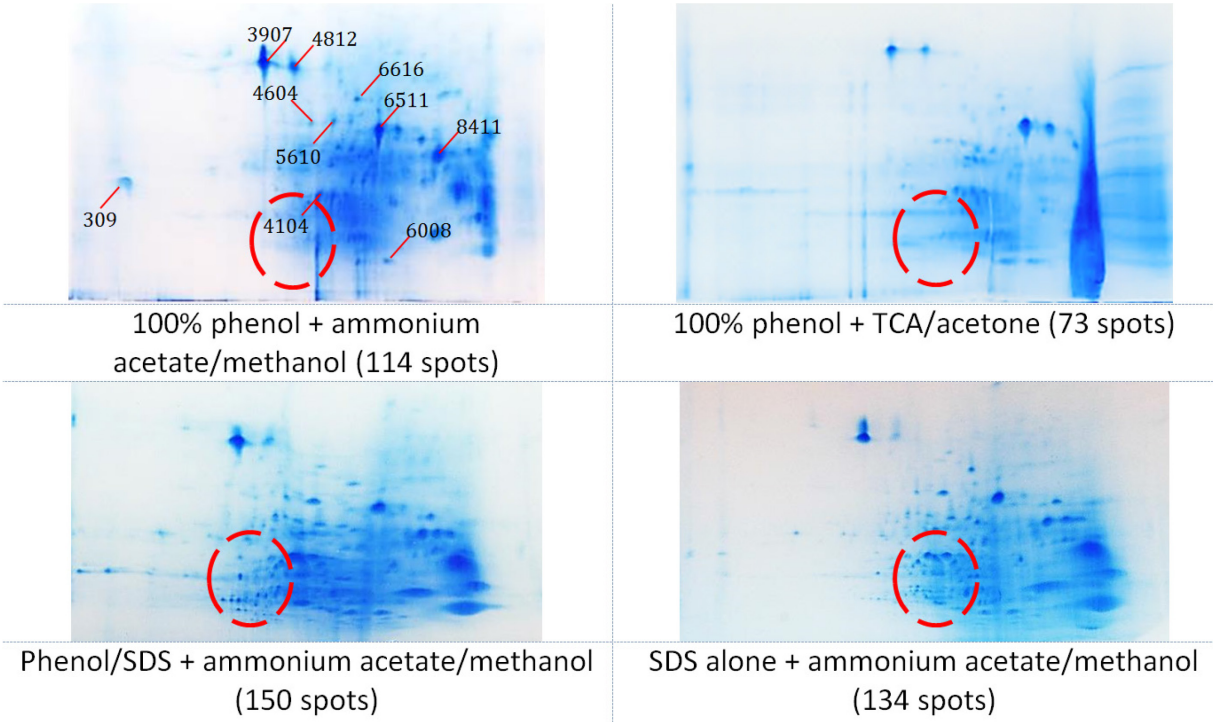


Figure 3.20. The effects of different extraction approaches on the 2DE profiles. 100 µg of proteins was rehydrated with 7 cm IPG strip, pH 3-10. Proteins were separated with 12% self-casted SDS-PAGE gel.

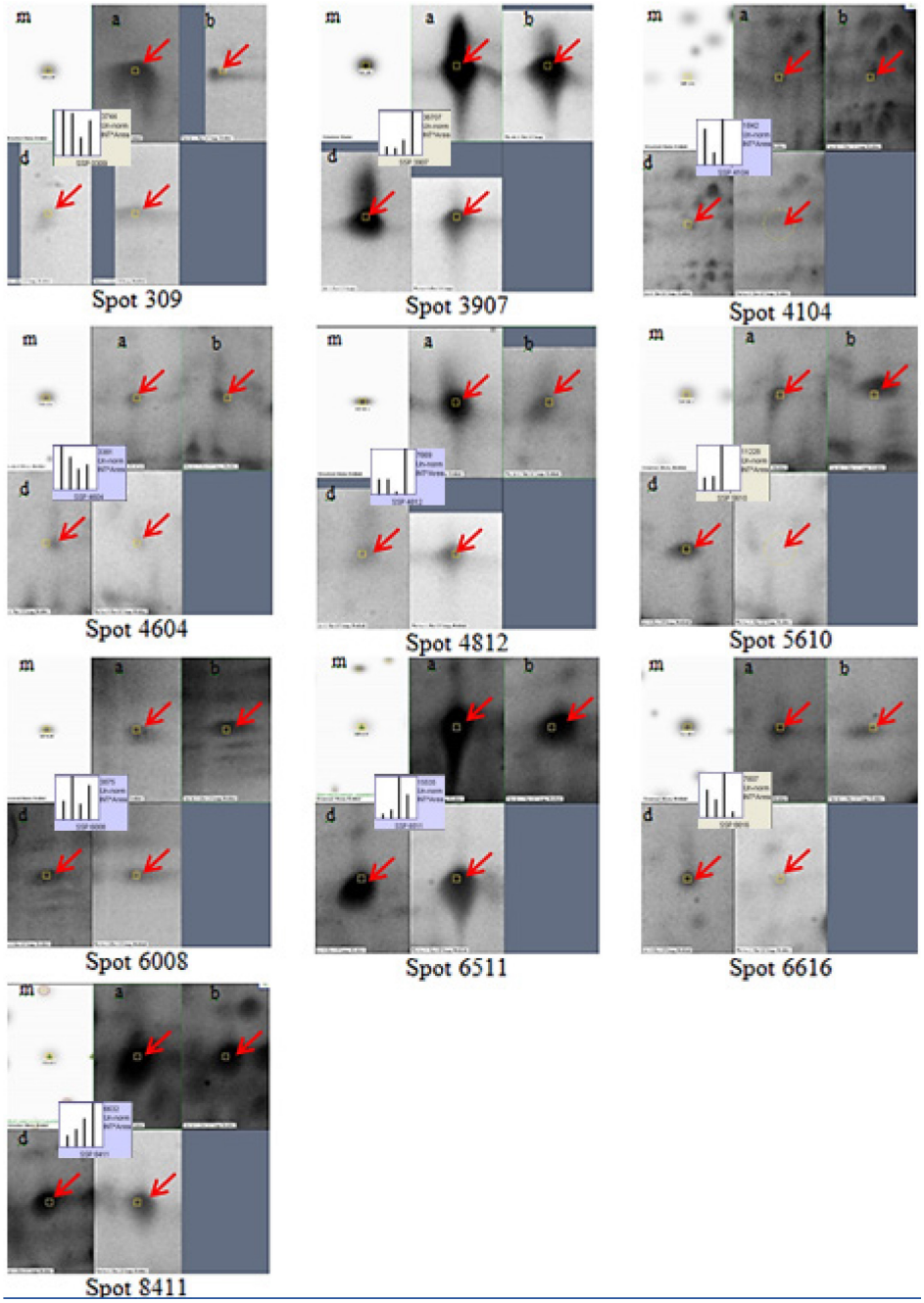


Figure 3.21. Spots analysed by PDQuest that showed differences in spot intensities. Clockwise from left: Master gel (m); 100% phenol (a); phenol/SDS (b); SDS alone (c); phenol/TCA (d).

However, the spot intensities for 100% phenol extraction and TCA/acetone precipitated proteins were clearly lower, indicating that it was crucial to use ammonium acetate/methanol precipitation after extracting the proteins with phenol. As reported by Wang and co-workers, prolonged exposure to TCA resulted in protein degradation or modifications (Wang et al., 2006).

Different buffers for re-suspending the precipitated pellet for downstream analyses were also examined. Based on the results in Table 3.7, rehydration buffer for isoelectric focusing was concluded to be superior for re-solubilising precipitated proteins. Rehydration buffer contained components such as detergents, chaotropes and reducing agents that could greatly enhance protein solubilisation. Unlike ammonium bicarbonate or Laemlli buffers, rehydration buffer contains urea and thiourea as chaotropes that disrupt non-covalent bonds resulting in disaggregation and denaturation of proteins (England & Haran, 2011). Protein solubilisation was further enhanced in the presence of 3-[(3-cholamidopropyl)dimethylammonio] propane sulfonate (CHAPS) and dithiothreitol (DTT). CHAPS is a detergent which is capable of disrupting hydrophobic interactions in proteins while DTT, a reducing agent, disrupts the disulfide bonds. As a result, proteins tend not to aggregate together, which helps in protein solubilisation.

Based on the results from the protein yield, 1D SDS-PAGE and 2DE, 100% phenol was ultimately employed to extract the mesocarp proteins despite the fact that phenol/SDS extraction was the most effective approach by far. The reason was that incompatibility of SDS with subsequent mass spectrometric analysis has been reported (Yeung, Nieves, Angeletti, & Stanley, 2008; Yu, Gilar, Lee, Bouvier, & Gebler, 2003). SDS is undoubtedly a powerful solubilising agent and used primarily in recovering membrane-bound proteins (Zheng, Song, Doncaster, Rowland, & Byers, 2007). Yet, the incomplete removal of SDS after protein extraction could also cause interference in isoelectric focusing (Yeung et al., 2008; Yu et al., 2003). Sheoran and co-workers stated in their work that both phenol and TCA/acetone extraction approaches were compatible with mass spectrometric analysis (Sheoran, Ross, Olson, & Sawhney, 2009). Phenol extraction followed by precipitation with ammonium acetate/methanol was preferred to TCA/acetone precipitation. TCA/acetone generated protein pellets that were difficult to solubilise in subsequent Laemlli or even urea buffer (Rastegari et al., 2011).

Table 3.7. Protein yield with different re-solubilisation buffers.

Buffer	Protein yield (µg/µL)
Urea-based (rehydration buffer)	1.53 ± 0.02
Ammonium bicarbonate	1.06 ± 0.02
Laemlli	0.93 ± 0.01

Phenol has been used to purify carbohydrates and subsequently nucleic acid (Hughes & Sinex, 1954; Kirby, 1956). This method is now routinely applied for deproteinisation of nucleic acids preparations. Interaction of phenol through hydrogen bonding with proteins results in the solubilisation of proteins, lipids and pigments in phenol phase while contaminants such as nucleic acid, polyphenol and carbohydrate remain in the Tris buffer phase (Carpentier et al., 2005) (Figure 3.22). The high pH of the buffer (pH 8) inhibits some common proteases (Hochstrasser, Harrington, Hochstrasser, Miller, & Merril, 1988), thus minimising protein degradation during the agitation phase prior to phase separation. In addition, basic environment have been reported to ionise phenolic contaminants, resulting in its inability to form hydrogen bonds with the proteins (Loomis & Battaile, 1966). Schuster and co-workers also mentioned that phenol reduced the chance of protein degradation caused by endogenous proteolytic activity (Schuster & Davies, 1983).

TCA/acetone is a highly effective precipitation method to eliminate proteolytic and DNA modifying enzymes such as nucleases (Rastegari et al., 2011) but this approach is not suitable for eliminating contaminants particularly polyphenols and lipids from plant. These metabolites can form hydrogen bonds and irreversible complexes with proteins through oxidation and covalent condensation (Loomis & Battaile, 1966). Several studies mentioned the unsuitability of TCA/acetone in extracting proteins from complex tissues (Carpentier et al., 2005; Saravanan & Rose, 2004; Wang et al., 2003). As shown in Figure 3.19, these interactions between metabolites and proteins caused vertical streaking in 1D SDS-PAGE gel as a result of charge heterogeneity for TCA/acetone precipitated proteins. Figure 3.23 shows that protein pellet precipitated with TCA/acetone was yellow to brownish which was due to polyphenolic oxidation as mentioned earlier (Wang et al., 2003). Proteins precipitated either with ammonium acetate/methanol or TCA/acetone after phenol extraction produced whitish pellets. This suggested that polyphenols have been effectively removed by phenol. The failure to solubilise the proteins for quantitation using the 2D Quant kit was almost certainly due to the presence of these interfering compounds in TCA/acetone extracted proteins. These results

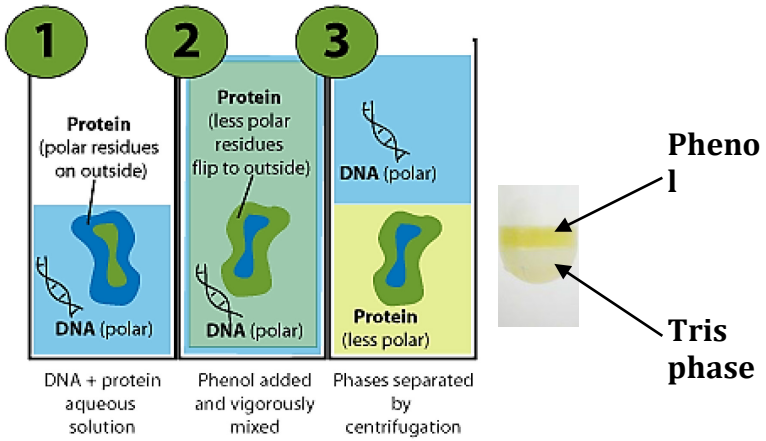


Figure 3.22. Principle of phenol extraction.

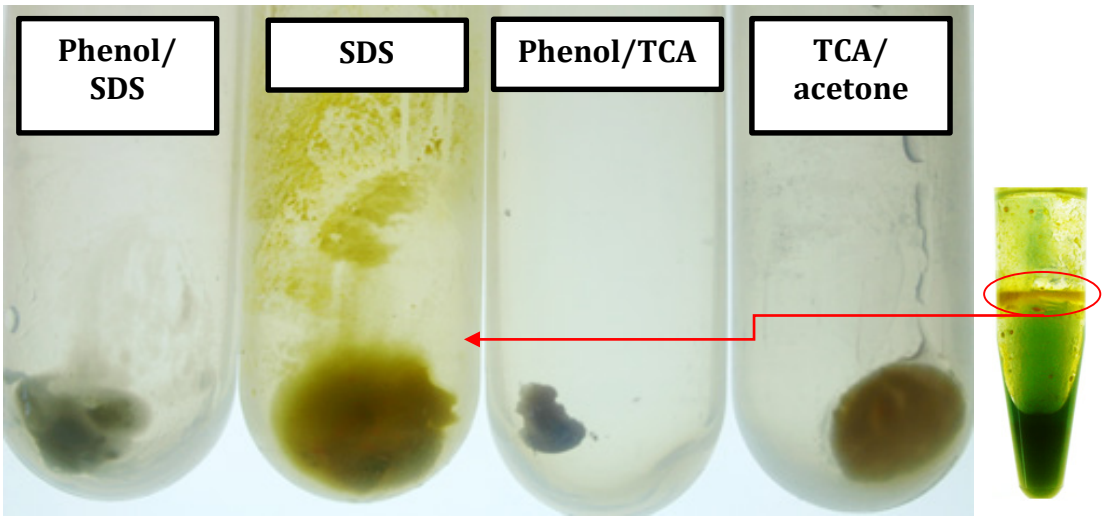


Figure 3.23. Acetone-washed chromoplast protein pellet. Picture on the far right depicted the re-suspended protein pellet extracted with SDS alone (notice the lipid layer).

emphasised the importance of eliminating contaminants prior to protein extraction with phenol.

3.4.5 1D SDS-PAGE and 2DE for oil palm varieties

The optimised techniques were applied to isolate, purify and extract chromoplast proteins from the low oleic acid (LO) and high oleic acid (HO) varieties. They were then resolved on 1D SDS-PAGE and 2DE gels (Figure 3.24 and Figure 3.25). These overview gel maps provide a useful reference set of proteins present in a cell, or tissue at any given time. In addition, the capability of this technique to detect differentially expressed proteins would almost allow the entire dynamic range of proteins approachable (Agrawal et al., 2011; De La Fuente et al., 2011). The gel-based techniques were used not only to investigate proteins that were

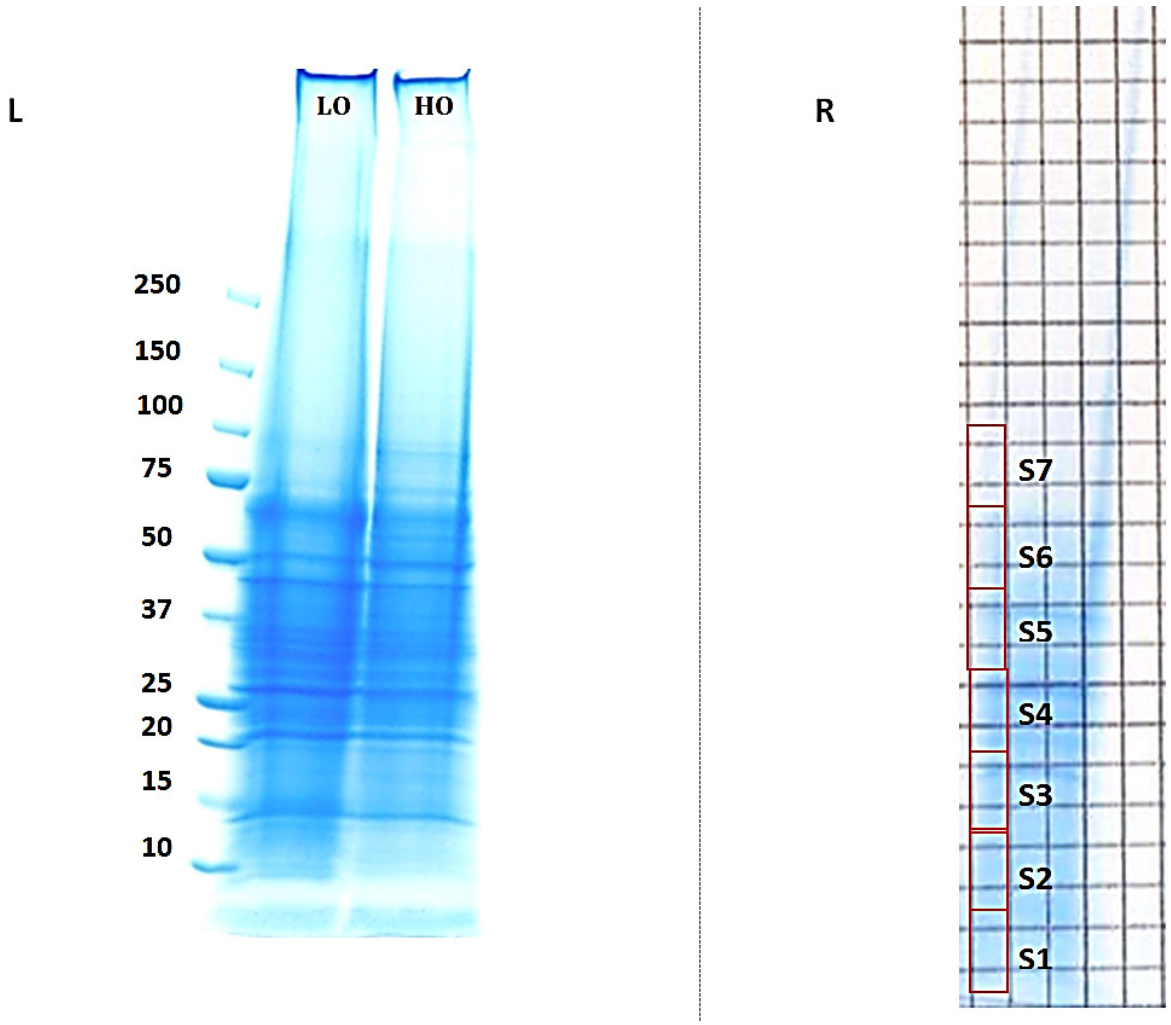


Figure 3.24. 1D SDS-PAGE profile of chromoplast proteins (L) from low oleic acid (LO) and high oleic acid (HO) varieties. 80 µg and 60 µg proteins from LO and HO varieties was loaded onto a 4-20% SDS-PAGE gel. Image on the right (R) depicts the segments excised for GeLC-MS/MS.

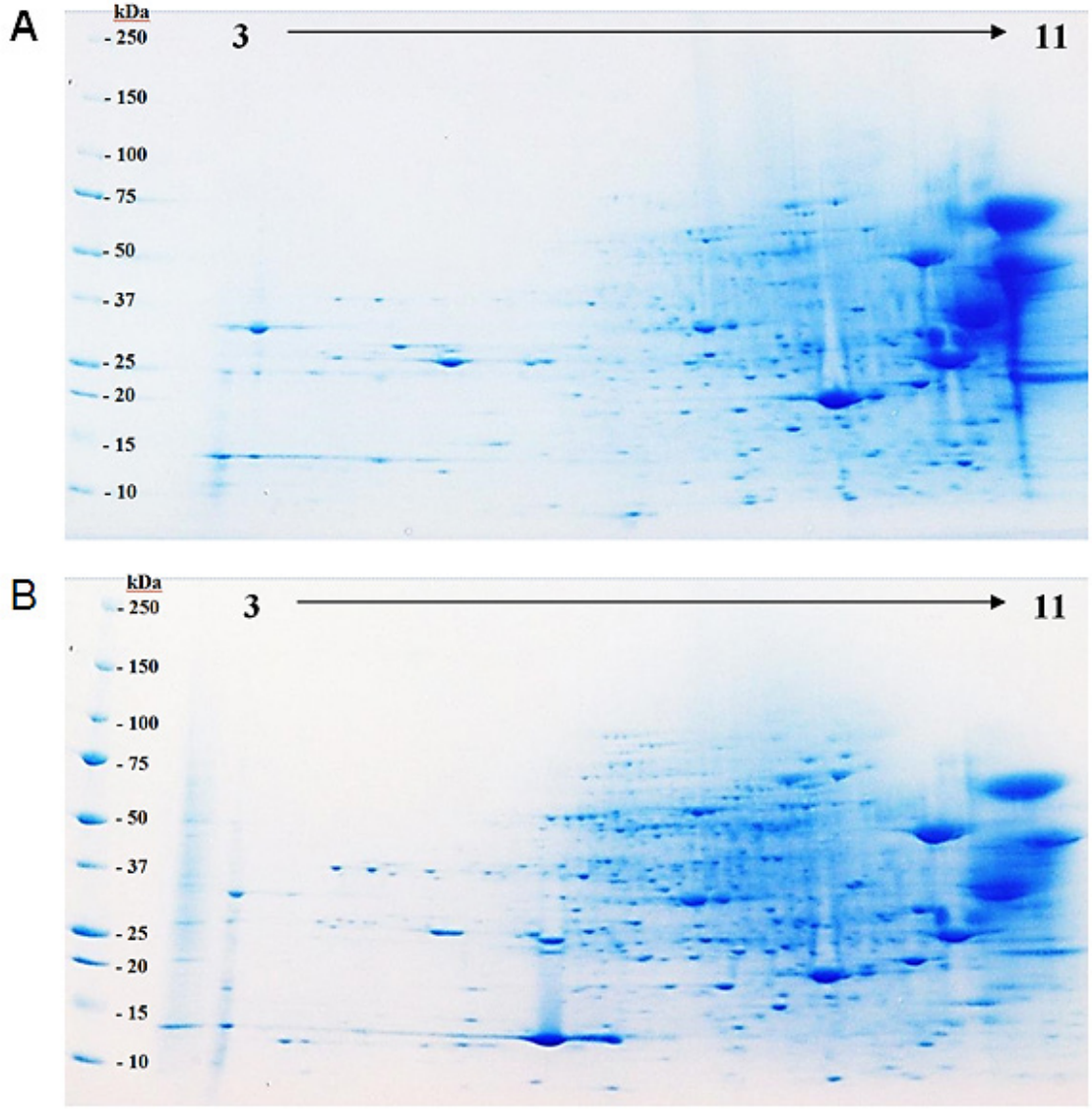


Figure 3.25. 2DE gel maps for (A) low oleic acid and (B) high oleic acid varieties. About 400 µg of proteins from low oleic acid and high oleic acid varieties was loaded.

differentially regulated in two different oil palm varieties, but also to determine the degree of the biological replicates variations (to be discussed in Chapter 4). Moreover, gel-based techniques are also capable of fractioning the extracted chromoplast proteins for organelle purity assessment with GeLC-MS/MS approach. Seven segments were excised from the 1D SDS-PAG (denoted as S1-S7) prior to mass spectrometric analysis (Figure 3.24). This approach increased the number of identified proteins in complex samples such as fruit mesocarps chromoplasts (to be discussed in Chapter 4). There are several reports where these techniques have been used to study chloroplast envelopes (Ferro et al., 2010), thylakoid and vacuole membranes (Agrawal et al., 2011; Peltier, Ytterberg, Sun, & van Wijk, 2004).

Gel-based approaches might not be the preferred tool in proteomics studies as non-gel based approaches such as liquid chromatography-mass spectrometry are less time-consuming and more consistent. However, gel-based techniques are still the choice for monitoring large protein sets as in specific studies done in root hair cell, leaf apoplasts and seeds, to name a few (Brechenmacher et al., 2009; De La Fuente et al., 2011; Goulet, Goulet, & Michaud, 2010; Hajduch et al., 2006).

3.4.6 LCMS for purity assessment

The aim of LCMS was to determine the purity of isolated organelle based on the subcellular localisation of identified proteins. Figure 3.26, Figure 3.27 and Figure 3.28 show the base peak chromatograms for the in-solution digested interphases (1.2 M, 1.7 M and 2.2 M) and in-gel digested chromoplast proteins (S1-7) from 1D SDS-PAGE. A base peak chromatogram (BPC) plots the intensity of the most intense peak in every mass spectrum throughout a liquid chromatography-mass spectrometry (LCMS) run (Bylund, Danielsson, & Markides, 2001; Niessen, 2006). A BPC is more informative compared to total ion chromatogram (TIC) as it corresponds to a single separated ion (instead of multiple ions in TIC) with the highest intensity. The results indicated that about 90% of these proteins were plastid-based while 10% were keratin contaminants, mitochondrion- and nucleus-based proteins (data not shown).

3.5 Conclusion

Apart from the fact that the oil palm fruit mesocarps are highly recalcitrant to general protein extraction due to high lipid content, it is also a plant-based material that contains high levels of interfering compounds such as phenolics. The method development described in this chapter was specifically targeted towards oil palm fruit mesocarps to obtain enriched

chromoplasts and maximum protein yield for proteomic characterisations. Removal of lipid from the mesocarps and incorporation of two cell wall digestive enzymes proved to be crucial and beneficial in increasing the effectiveness of the protein extraction. Subsequently, phenol extraction followed by protein precipitation with ammonium acetate in methanol was demonstrated to be the most effective way to acquire chromoplast proteins in terms of protein yield and good (in term of resolution and intensity) 2DE protein spots. Although SDS in phenol/SDS extraction approach gave the best protein yield and increased the number of 2DE protein spots, the SDS interfered with gel-based techniques and suppressed peptide ionisation in mass spectrometry approach. Even though the whole process was lengthy (Figure 3.29), good quality proteins for mass spectrometry analysis were able to be acquired. Most importantly, the first optimised methodology to obtain oil palm chromoplast proteins that are compatible with mass spectrometry analysis was reported here. Identified proteins revealed that most of the extracted proteins were plastid-based. This accomplishment was crucial for evaluating chromoplast proteins from two varieties in an attempt to reveal the key regulatory enzymes in fatty acid biosynthesis.

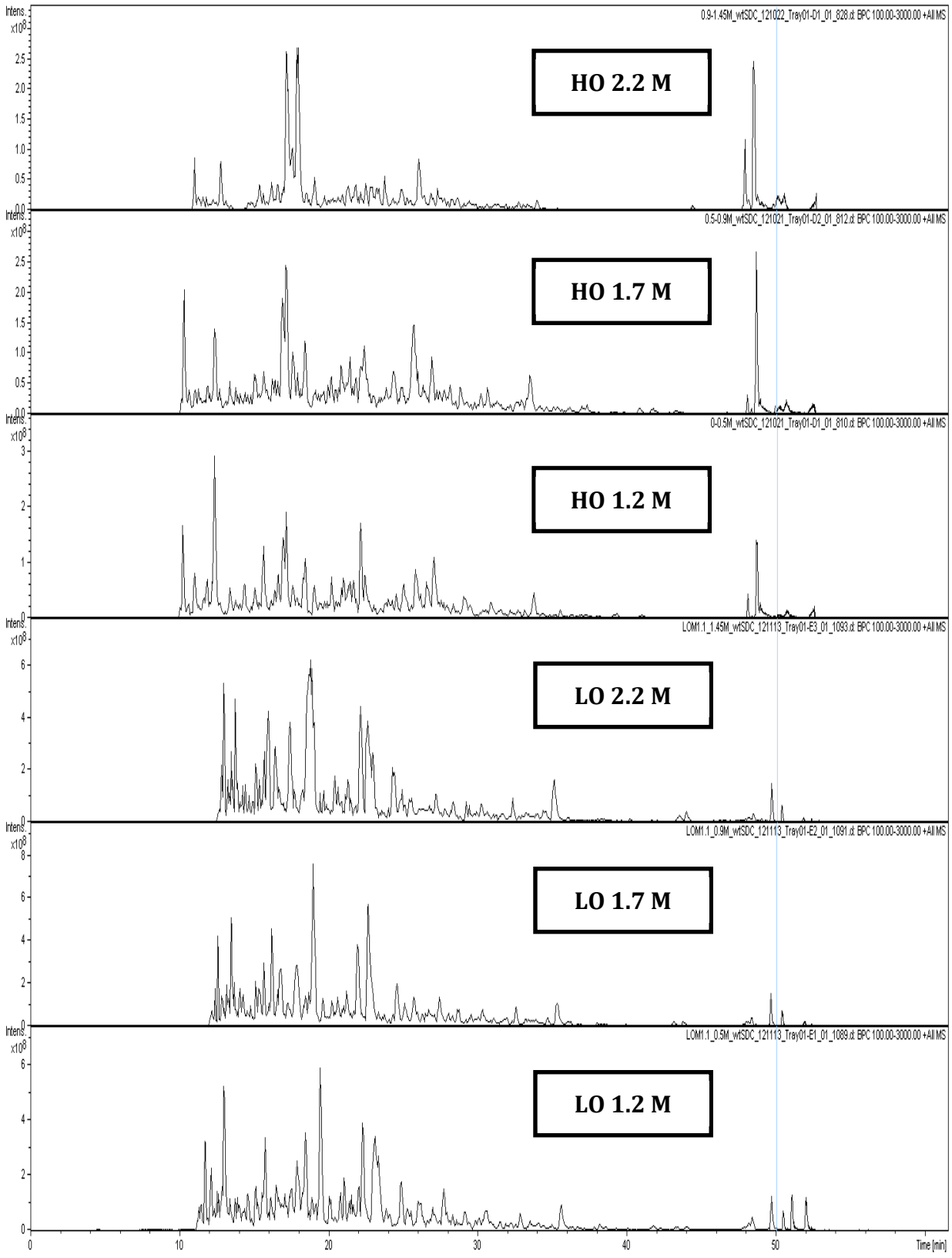


Figure 3.26. Interphases base peak chromatograms of low oleic acid (LO) and high oleic acid (HO) varieties.

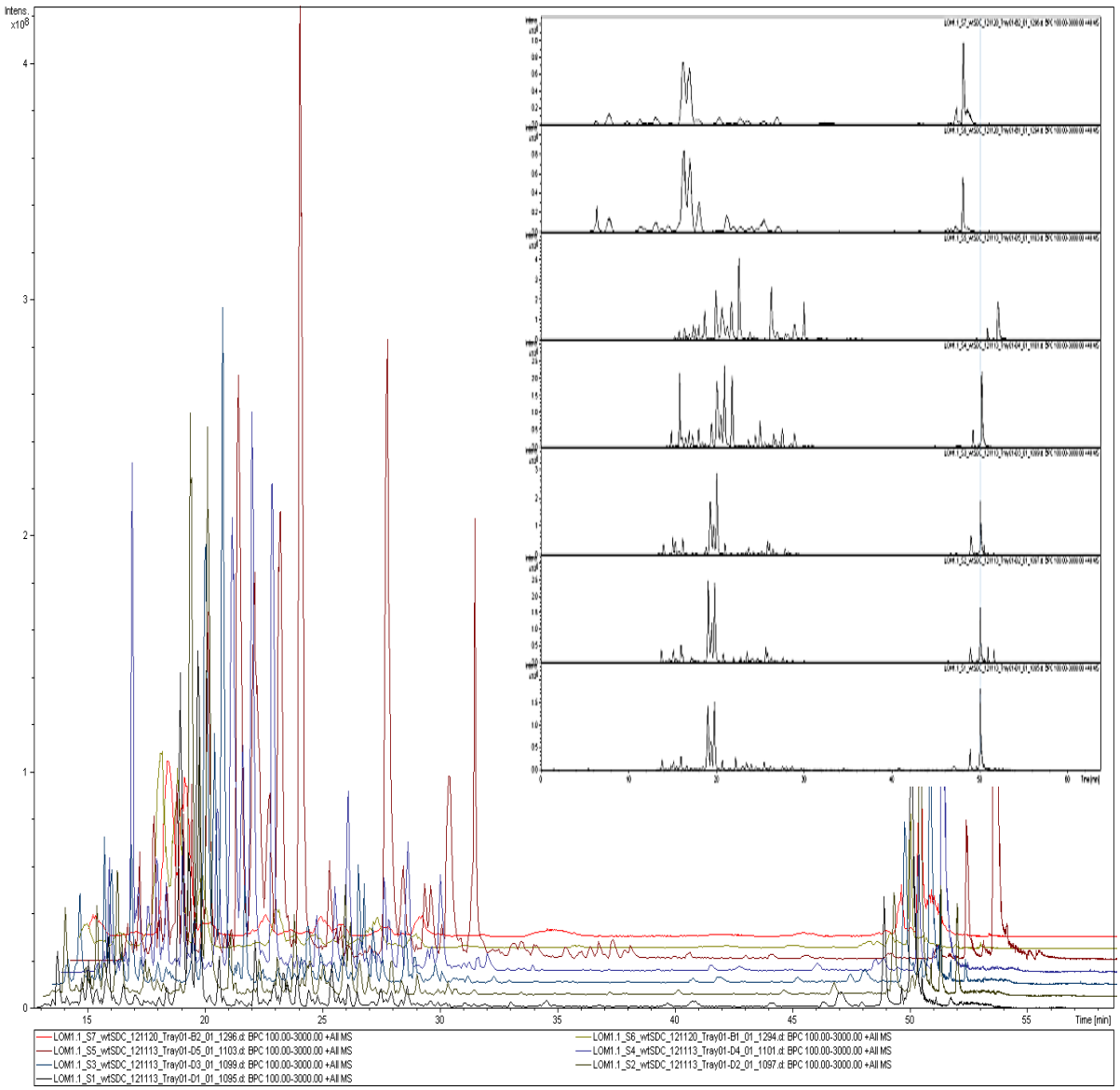


Figure 3.27. Base peak chromatograms for gel segments (LO S1-7). The inset shows the individual base peak chromatograms.

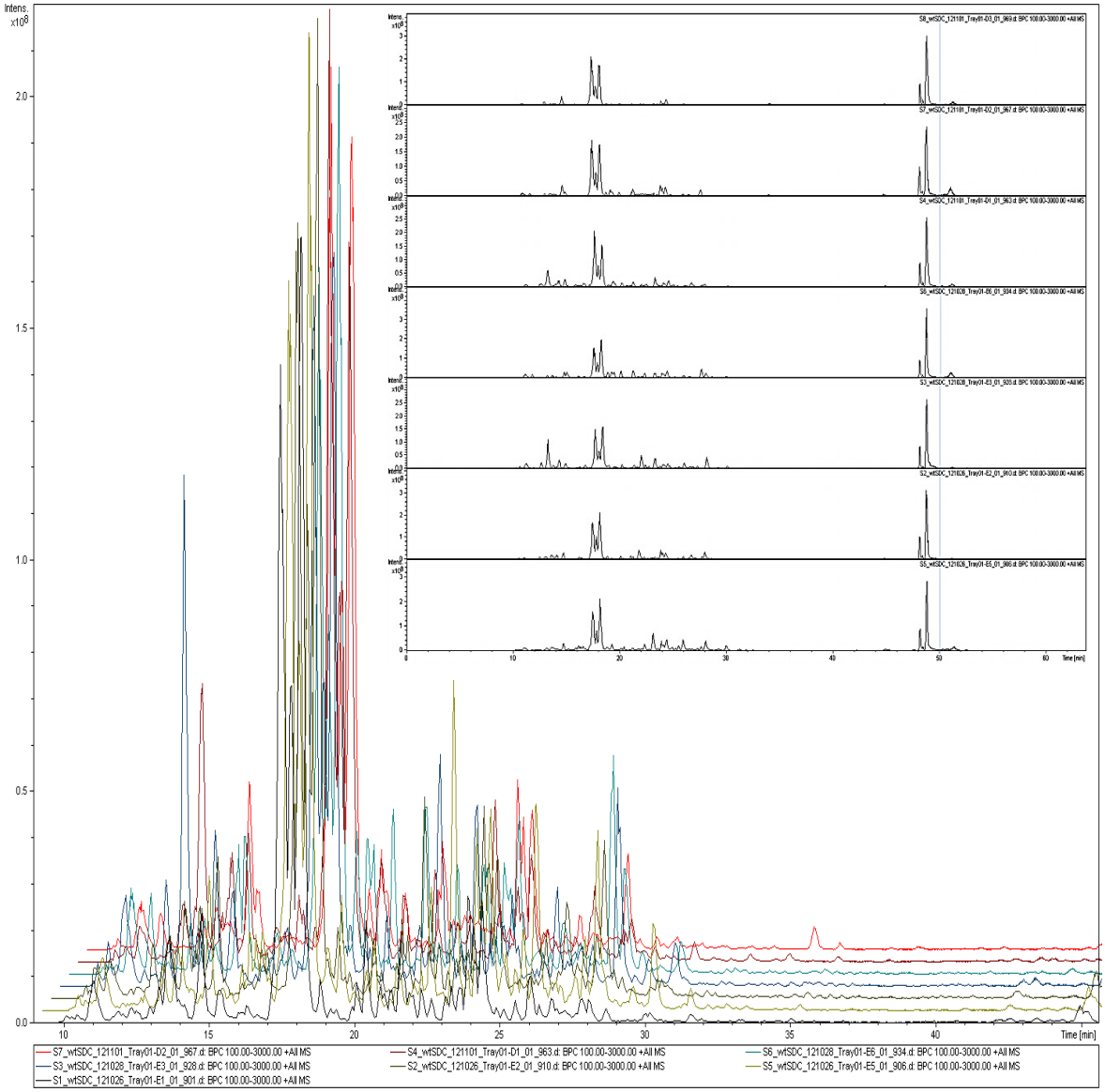


Figure 3.28. Base peak chromatograms for gel segments (HO S1-7).

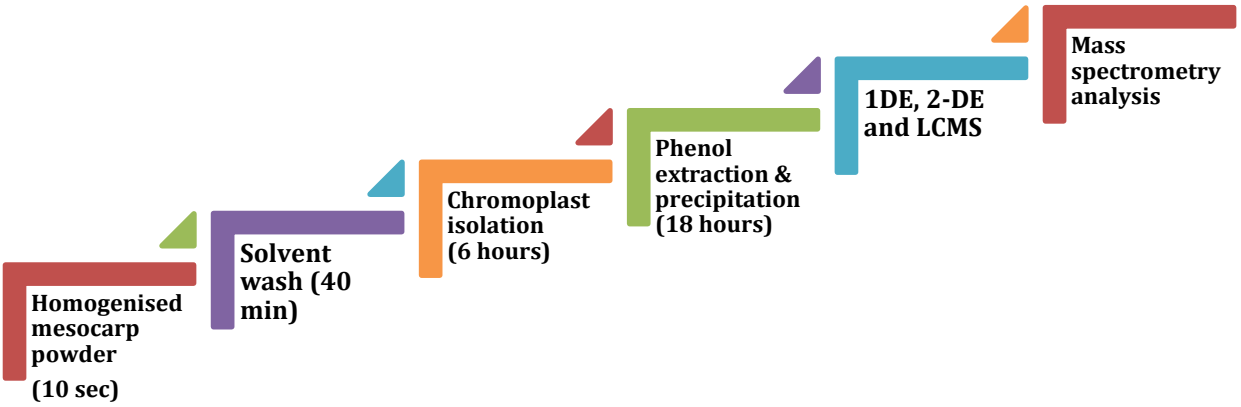


Figure 3.29. Overview of workflow. Approximate time required for each step (per sample) until protein extraction is indicated in bracket.

4.0 Development of gel-based and non-gel based proteomic approaches to characterise protein expression of oil palm fatty acid biosynthetic enzymes

4.1 Introduction

Classic gel electrophoresis techniques were applied in this study to observe the biological variation in three replicates from low and high oleic acid varieties of oil palm. This was followed by proteomic investigation of their protein expression using gel electrophoresis-based mass spectrometric analysis, termed as GeLC-MS/MS. GeLC-MS/MS essentially involves the digestion of excised gel-separated proteins for mass spectrometric analysis and has proved to be a key approach in protein identification (Shevchenko, Wilm, Vorm, & Mann, 1996). GeLC-MS/MS approach has been applied to profile differential protein expression, proteome analyses and biomarker discoveries (Beer, Tang, Barnhart, & Speicher, 2011; Daher et al., 2010; Graham et al., 2007; Paulo, Kadiyala, Banks, Steen, & Conwell, 2012; Petricka et al., 2012; Zeng et al., 2011b). Now this technique is typically used in tandem with a complementary liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Balbuena, He, Salvato, Gang, & Thelen, 2012; He et al., 2012; Salvato et al., 2014).

Non-gel based assessment of fatty acid biosynthetic enzyme expressions was performed via the application of LC-MS/MS. This approach has become the technique of choice for most large-scale proteomic studies such as in yeast (Picotti, Bodenmiller, Mueller, Domon, & Ruedi, 2009), *Arabidopsis* (Wienkoop & Weckwerth, 2006) and *Picea abies* (Zulak et al., 2009). A typical shotgun proteomics workflow consists of trypsin digested proteins followed by separation of the peptides by columns of different attributes such as a reverse-phase column. Separated peptides are then fragmented in a mass spectrometer to generate product ions for their identification or downstream characterisation. This peptide centric proteomics is capable of producing sequence-specific information and performing relative quantitation of the peptide. LC-MS/MS is undoubtedly less laborious and more repeatable in term of generated data than two-dimensional gel electrophoresis (Abdallah et al., 2012). Furthermore, this approach can be high throughput and sensitive (requires less protein for analysis) than a gel-based approach.

Complex mixtures of proteins potentially produce peptides that 'congest' a mass spectrometric analysis. In this study, the complexity of the oil palm derived proteins was resolved with the application of two-dimensional liquid chromatography (2DLC). In 2DLC, a strong cation exchange column is normally utilised to provide orthogonal peptide separation for shotgun proteome analyses. The combination of a strong cation exchange column with a reverse-phase column was done 'offline'. An online combination gives the advantages of full automation and less sample handling (thus reducing technical variation). On the other hand, an offline combination provides the flexibility of independently optimising the analytical performance of two columns with different attributes. Peptide separation had been shown to be more effective in an offline 2DLC approach (Nagele, Vollmer, Horth, & Vad, 2004). In the study, optimisation of the strong cation exchange fractionation were also carried out 'offline' prior to the separation in the second dimension with a reverse-phase column to improve the detection of low abundance peptides from fatty acid biosynthetic enzymes.

Novel comparative quantitation of fatty acid biosynthetic enzymes from the low oleic acid and high oleic acid varieties was performed using isobaric iTRAQ labelling. Digested protein samples were labelled with six iTRAQ isobaric tags at amino groups of lysine side chains and *N*-termini of all peptides (Abdallah et al., 2012). The multiplexed labelled peptides samples were then pooled and fractionated offline using a strong cation exchange column to improve the detection of the lower abundance peptide variants. Subsequently, the labelled peptides were separated by a reverse phase nanoHPLC and the resolved labelled peptides analysed with tandem mass spectrometry. Collision-induced dissociation caused the iTRAQ-tagged peptides to fragment and release the reporter ions (at *m*/*z* 113, 114, 115, 116, 117 and 118) as well as *b*- and *y*-ions among other fragment ions for protein identification. The generated peak ratios of the six tandem mass spectrometry reporter ions were used to quantify the relative peptide abundance of targeted fatty acid biosynthetic enzymes. A quadrupole-time-of-flight mass spectrometer was used in this study for its ability to detect the reporter ions at the low *m/z* range.

Chapter 4 describes the application of gel-based and non-gel based approaches to investigate the presence of differentially expressed enzymes that are related to the oil palm oleic acid biosynthesis. Proteomic techniques used to determine the extent of biological variation among replicates in the study are also explained. The reliability of the replicates is crucial to ensure that any differences in protein expression of the fatty acid biosynthetic enzymes

occurred primarily due to their differential regulation of oleic acid production in oil palm. Subsequently, isobaric tags were used to investigate those differences in enzyme expressions. Proteome changes of oil palm mesocarp from different development stages has previously been documented to understand the mechanism of palm oil production (Loei et al., 2013). However, this chapter outlines and reports the first targeted attempt to improve and quantitate the differences in fatty acid biosynthetic enzyme expression from two different oil palm varieties.

4.2 Materials

4.2.1 Plant materials

Fruit bunches (Figure 4.1) of the standard DxP oil palm crosses (*Elaeis guineensis* var. Tenera) (denoted as low oleic acid variety) and from the PORIM Series 12 (PS12) breeding populations (denoted as high oleic acid variety) were grown and harvested from the Malaysian Palm Oil Board research stations at Bangi, Selangor and Hulu Paka, Terengganu, Malaysia. The fruit mesocarps were sliced, snap frozen in liquid nitrogen and stored at – 80°C until use. The seeds were discarded. Oil palm bunches (progenies) with different levels of oleic acids at the 20th week after anthesis were collected from three low oleic acid and high oleic acid palms, respectively (Table 4.1). Fruit mesocarps obtained from the bunch were divided into four technical replicates of 100 g each.

Figure 4.1. The two oil palm varieties used in this study.

Table 4.1. Biological replicates for the oil palm varieties used in the study.

High oleic acid			Low oleic acid		
Palm no.	**Progeny**	**Oleic acid content (%)**	**Palm no.**	**Progeny**	**Oleic acid content (%)**
0.306/319	PK 540	48.85	0.306/79	PK 540	39.03
0.306/319	PK 540	48.85	0.306/76	PK 540	39.03
0.337/249	PK 1254	48.90	0.337/214	PK 1201	37.90

4.2.2 Chemicals

Unless otherwise stated, general chemicals used in this study were purchased from Sigma (St. Louis, MO, USA), BDH (BDH Laboratory Supplies, Poole, England) and Merck (Merck KGaA, Darmstadt, Germany). Organic solvents were acquired from Fisher (Fisher Scientific, Loughborough, UK). Specific chemicals such as precast gels were procured from Bio-Rad (Bio-Rad Laboratories, Inc., Hercules, CA), different ranges of Immobiline DryStrip from GE Healthcare (GE Healthcare Bio-Sciences AB, Uppsala, Sweden), Protease Inhibitor Cocktail Tablets from Roche (Roche Diagnostic, IN, USA), sequencing grade modified trypsin from Promega (Promega Corporation, WI, USA) and 8-plex iTRAQ reagents from Applied Biosystems (AB Sciex Pte. Ltd., Foster City, CA, USA). Double-distilled water was used in preparation for all general solutions.

4.3 Methodology

4.3.1 Chromoplast isolation and subsequent protein extraction

The delipidation method was based on the method developed by Wang and co-workers (Wang et al., 2006) and described in detail in Chapter 3. Sliced mesocarps were homogenised in liquid nitrogen with a cold Waring blender (Blender B011, Dynamics Corporation, Greenwich, USA) at low speed for 10 s, followed by grinding with a standard ceramic mortar and pestle. The powdered mesocarps were then mixed with cold acetone containing 10% trichloroacetic acid and 1 mM dithiothreitol. The slurry was then centrifuged at 13,000 g for 10 min at 4°C (RA-300 rotor, Kubota 7820, Kubota Corporation, Tokyo, Japan). The supernatant was discarded and the washing step was repeated once. Subsequently, cold 80% methanol containing 0.1 M ammonium acetate was added to the precipitate; mixed and centrifuged as before. After discarding the supernatant, the precipitated mesocarp pellet was washed with cold 80% acetone. The mixture was mixed well and centrifuged again at 13,000 g for 10 min at 4°C. The resulting pellet was air-dried.

The isolation method of chromoplasts was adapted from Jain, Fan and co-workers (Fan et al., 2009; Jain et al., 2008) and described in detail in Chapter 3. Washed mesocarp pellet was transferred into a beaker containing cell wall digestive enzymes (2% w/v cellulase, 0.1% w/v pectinase, 0.6 M sorbitol, 0.1 M DTT, 5 mM 2-(4-morpholino)-ethane sulfonic acid (MES)-KOH, pH 5.5). The suspension was then incubated in an incubator shaker (Labnet 211DS, Labnet Instrument, Inc., NJ, USA) for 6 h at 37°C. After cell wall digestion, the mixture was sieved through two layers of Miracloth (Calbiochem, EMB Millipore Corporation, Billerica, MA) into a fresh beaker on ice to separate non-macerated plant materials from protoplasts. After that, the filtrate was centrifuged at 1750 g for 5 min at 4°C (RA-300 rotor, Kubota 7820, Kubota Corporation, Tokyo, Japan) to collect intact chromoplasts. The chromoplast pellet was gently re-suspended in the extraction buffer (0.7 M sucrose, 1 M Tris-HCl, pH 8.3, 5 M NaCl, 50 mM DTT, 1 mM EDTA containing a tablet of Roche protease inhibitors for every 50 mL of buffer). The mixture was agitated for 5 min and then sonicated for 15 min at 4°C to release the proteins. . An equal volume of fresh 50 mM, pH 8.0 Tris-saturated phenol was added to the mixture thereafter. The mixture was further agitated for 10 min just before centrifugation at 15,000 g for 15 min at 4°C (RA-300 rotor, Kubota 7820) for phase separation. Following that, the upper phase was transferred to a new centrifuge tube while the bottom phase was discarded. Chromoplast proteins were precipitated by adding five volumes of cold ammonium acetate-saturated methanol and incubated at -20°C overnight before being centrifuged at 15,000 g for 15 min at 4°C (RA-300 rotor, Kubota 7820) to obtain the protein pellet. The protein pellets were then rinsed with cold ammonium acetate-saturated methanol until a whitish colour was obtained for the pellet. This was followed by washing the pellet with cold 80% acetone for three times. Proteins were precipitated after each wash by centrifuging at 15,000 g for 5 min at 4°C (RA-300 rotor, Kubota 7820). At the end of the washing phase, the supernatant was carefully decanted before the chromoplast protein pellet was air-dried. Commercially available 2D Quant Kit (GE Healthcare Life Sciences, Uppsala, Sweden) was utilised to determine protein content in the samples. Bovine serum albumin provided with the kit was used as the protein calibration standard and each of the quantitation was performed in duplicates.

4.3.2 One-dimensional (1D) gel electrophoresis

1D SDS-PAGE is a complementary technique to strong cation exchange used to fractionate the chromoplast proteins in order to improve the detection of low abundance proteins. For this

purpose, the precipitated chromoplast proteins were dissolved in Laemlli buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 25% glycerol, 0.01% bromophenol blue, 0.005% β-mercaptoethanol) and denatured by heating at 95°C for 4 min. 50-100 µg protein was loaded into each lane on a 1.0 mm precast Mini-PROTEAN® TGX™ 4-20% or 12% polyacrylamide gel (Bio-Rad Laboratories Inc., Hercules, CA). Electrophoresis was conducted in a Bio-Rad mini-PROTEAN® Tetra Cell apparatus (Bio-Rad Laboratories Inc., Hercules, CA) at 200 V for 1 h. Following electrophoresis, the separated proteins were fixed for 30 min in a fixing solution (50% ethanol, 10% acetic acid) and stained with an in-house prepared Colloidal Coomassie G-250. The gel was destained with Milli-Q water (H_2O) until the background was clear before acquiring the gel image with a DLSR camera (Nikon D100, Nikon Corporation, Japan). The settings were ISO 200, f16 for the aperture and shutter speed was $1/40^{th}$ of a second. Image processing was done with the Analysis software (Soft Imaging System, Olympus Corporation). Comparative analysis of the protein bands were performed using the Phoretix 1D gel analysis software (TotalLab Ltd, Newcastle, UK).

4.3.3 Two-dimensional gel electrophoresis (2DE)

2DE was used to qualitatively evaluate the biological variations of the replicates in terms of overall number, relative intensities and resolution of protein spots obtained in addition to screening differentially expressed proteins among the replicates from low oleic acid and high oleic acid varieties. Dried protein pellet acquired after overnight precipitation was solubilised in rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 0.25% Pharmalyte and 0.4% DTT). The volume of rehydration buffer needed depended on the amount of proteins acquired. Sonication for about 15 min was required to assist the resolubilisation of the protein pellet. Any non-solubilised proteins after the sonication were removed by centrifugation. Protein quantitation was carried out prior to isoelectric focussing. For the first dimension isoelectric focusing, Bio-Rad ReadyStrip™ IPG strip, 11 cm, pH 3-10 (Bio-Rad Laboratories Inc., Hercules, CA) were rehydrated with 200 µL of rehydration buffer containing 200 µg proteins overnight. For comparison between the low oleic acid and high oleic acid varieties, Bio-Rad ReadyStrip™ IPG strip, 11 cm, pH 3-11 were rehydrated with 200 µL of rehydration buffer containing about 200 µg proteins overnight. IEF was performed according to the following focusing parameters (Table 4.2) set on the Bio-Rad Protean IEF Cell (Bio-Rad Laboratories Inc., Hercules, CA) at 20°C.

Table 4.2. Parameters used for isoelectric focusing.

Phase	Volt (V)	mA (Max)	Time (hours)	Volt hours
1	100	20	3	300
2	500	20	2	1000
3	1000	5	2	2000
4	8000	5	14	84000

For the second dimension, a precast 4-20% gradient SDS-PAGE gel (Bio-Rad Laboratories Inc., Hercules, CA) was used. After the rehydration, proteins in the IPG strip were reduced with 1% DTT for 15 min followed by alkylation with 4% iodoacetamide for another 15 min before transferring to the gel (second dimension) for separation.

4.3.4 Protein digestion (in-gel and in-solution) and fractionation

In-gel digestion – Each lane representing different biological replicates from the 1D SDS-PAGE separation was excised into eight uniform segments (5.0 x 7.5 mm per segment) denoted as S1, S2, S3, until S8. The gel segments were sliced further into even size plugs before destaining in 200 mM ammonium bicarbonate in 50% acetonitrile for 1 h at 37°C. Proteins were reduced and alkylated with thiol-free 50 mM tris(2-carboxyethyl)phosphine and 150 mM iodoacetamide, respectively. An ionic detergent, sodium deoxycholate (1% w/v) was added to the dehydrated gel plugs prior to protein digestion with modified sequencing grade trypsin (Promega, Madison, WI, USA) in 50 mM NH_4HCO_3 for 16 h at 37°C. The solubility of the proteins was increased in the presence of 4 µL of sodium deoxycholate to enhance trypsinisation (50:1 µg) (Koehn et al., 2011; Lin et al., 2008; Proc et al., 2010). The resulting peptides were extracted with increasing concentrations of acetonitrile from 10% to 80% and pooled. 0.5% formic acid was added to the peptide extract before the precipitated sodium deoxycholate was removed by centrifugation at 14, 000 g for 15 min at ambient temperature. The pooled peptide extract was then dried in a centrifugal evaporator (CentriVap Concentrator, Labconco, MO, USA). *In-solution digestion* – About 50 µg of TCA/acetone-precipitated chromoplast proteins were re-suspended in 0.1 M ammonium bicarbonate before reduction and alkylation using 50 mM tris(2-carboxyethyl)phosphine and 150 mM iodoacetamide, respectively. As with the in-gel protein digestion, 4 µL of sodium deoxycholate (1% w/v) was added. After digestion with 1 µg of trypsin at 37°C for 16 h and removal of sodium deoxycholate through acidification with 0.5% formic acid, the peptide solution was dried in a centrifugal evaporator (CentriVap Concentrator). *Peptides clean-up* – The dried labelled peptide pellet was resuspended in 100 µL of 0.1% formic acid. Acetonitrile, methanol

and 0.1% formic acid-conditioned Empore solid phase extraction disks (3M Company, MN, CA) were added to the peptide solution and incubated at ambient temperature with slight agitation for 4 h. The bound peptides on the C18 membrane disks were eluted with 50% ACN in 0.1% FA for 2.5 h. *SCX fractionation* – The peptide eluent was dried with a centrifugal evaporator (CentriVap Concentrator) and resuspended in 50 µL of 0.1% FA. The resuspended peptides were then fractionated using a 5 µm strong cation exchange column packed with Spherisorb (BioX-SCX, 500 µm ID, 15 mm). Elution was done with 1%, 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 80% and 100% of 2 M ammonium formate buffer containing 2% ACN, pH 3.5. Each of the collected fractions was then dried with a centrifugal evaporator (CentriVap Concentrator).

4.3.5 Quantitation using iTRAQ isobaric tags

Labelling of the peptides with six different isobaric tags was performed using commercial 8-plex iTRAQ reagents (AB Sciex Pte. Ltd., CA, USA) (Table 4.3, Figure 4.2). Precipitated proteins from six different biological replicates of low oleic acid and high oleic acid varieties containing 200 µg proteins each were resuspended in 300 µL 0.5 M triethylammonium bicarbonate, pH 8 buffer prior to reduction and alkylation with 50 mM tris(2-carboxyethyl)phosphine and 150 mM iodoacetamide, respectively. After 16 h of digestion with modified sequencing grade trypsin (Promega, Madison, WI, USA) in a 1:40 µg trypsin:protein ratio, the peptides solution was dried using a centrifugal evaporator (CentriVap Concentrator) and resuspended in 50 µL of 0.5 M triethylammonium bicarbonate, pH 8. Each of the 113, 114, 115, 116, 117 and 118 iTRAQ reagents were mixed with 50 µL of isopropanol and added to each of the digests. The mixtures were incubated at room temperature for 2 h for the labelling process. The labelling reaction was terminated with the addition of 100 µL of mass spectrometry grade water and incubated further for another 30 min at ambient temperature. All of the individually labelled protein digests were then pooled together and dried. The dried labelled peptides were then resuspended in 100 µL of mass spectrometry grade water and washed for another three times before drying with a centrifugal evaporator (CentriVap Concentrator). The dried labelled peptide pellet was resuspended in 200 µL of 0.1% formic acid. Acetonitrile, methanol and 0.1% formic acid-conditioned Empore solid phase extraction disks (3M Company, MN, CA) were added to the peptide solution and incubated at ambient temperature with slight agitation for 4 h. The bound peptides on the C18 membrane disks were eluted with 50% ACN in 0.1% FA for 2.5 h.

Table 4.3. Experimental design for the iTRAQ labelling. [a] and [b] represent the low oleic acid and high oleic acid biological replicates, respectively.

Replicate	iTRAQ reagents	Replicate	iTRAQ reagents	Replicate	iTRAQ reagents
LO1[a] + HO1[b]	113 + 116	LO1[a] + HO2[b]	113 + 117	LO1[a] + HO3[b]	113 + 118
LO2 [a]+ HO1[b]	114 + 116	LO2[a] + HO2[b]	114 + 117	LO2[a] + HO3[b]	114 + 118
LO3 [a]+ HO1[b]	115 + 116	LO3[a] + HO2[b]	115 + 117	LO3[a] + HO3[b]	115 + 118

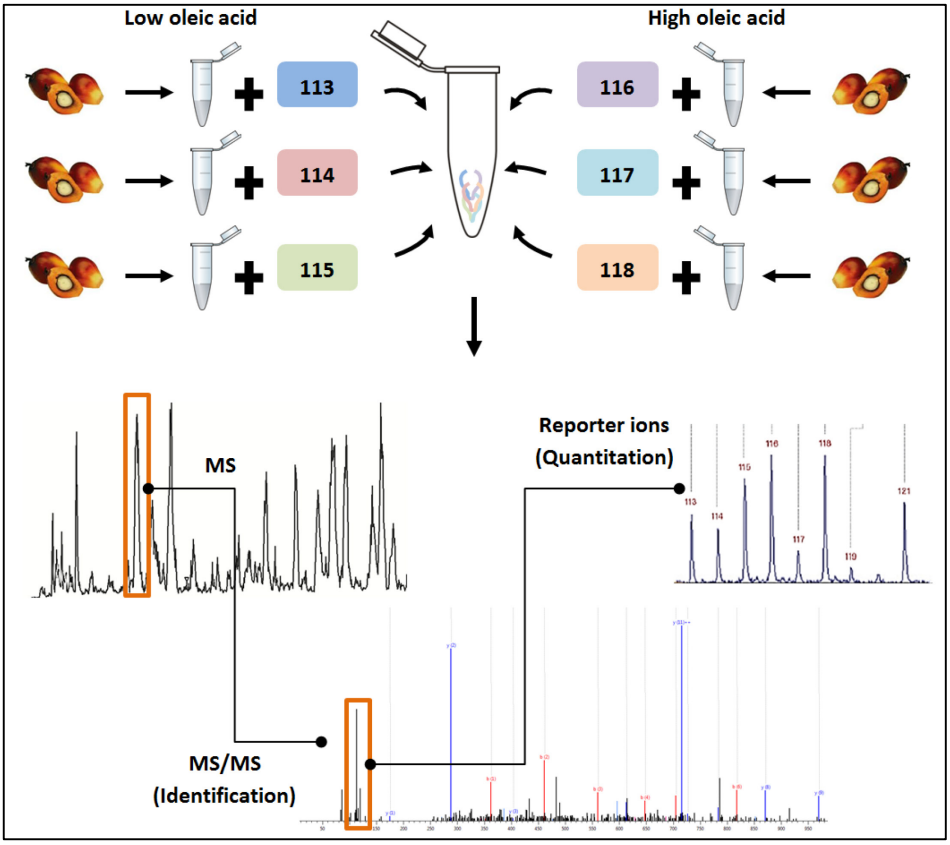


Figure 4.2. Workflow of peptides labelling with 8-plex iTRAQ tags

The eluent was dried with a centrifugal evaporator (CentriVap Concentrator) and resuspended in 30 µL of 0.1% FA. The resuspended labelled peptides were then fractionated using a strong cation exchange column (Phenomenex SCX cartridge, 4 x 2 mm). Elution was performed with 3%, 6%, 10%, 15%, 20%, 25%, 40%, 80% and 100% of 2 M ammonium formate buffer containing 2% ACN, pH 3.5. Each of the collected fractions was then dried with centrifugal evaporator (CentriVap Concentrator).

4.3.6 Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

Separation and identification of the protein digests was conducted with a nano-Advance Splitless nano-liquid chromatography (nanoLC) system (Bruker Daltonik GmbH, Bremen, Germany) coupled to an amaZon speed ETD ion trap mass spectrometer (Bruker Daltonik).

Before loading onto the nanoLC system, the tryptic digests were reconstituted in 30 µL of 0.1% FA and 5% ACN. 5 µL was injected into the nanoLC system for peptide separation with an Intensity C18P (1.8µ, 0.1 x 300 mm) (Bruker-Michrom Bioresources, Inc., Auburn, CA, USA) reverse-phase columns equilibrated with 98% solvent A (2% ACN, 0.1% FA) and 2% solvent B (98% ACN, 0.1% FA). Gradient of 2-45% solvent B was employed to elute the bound peptides during 45 min at a flow rate of 400 nL min^{-1}. The eluted peptides injected by electrospray into the ion trap mass spectrometer with a spray voltage of 1300 V and a capillary temperature of 150°C. Precursor survey scan was acquired with mass ranged from *m/z* 310-1400, the resolution was set to 'Enhanced Resolution' and the scanning speed was 8100 u per sec. Tandem MS condition consisted of 'Xtreme Resolution' scan ranged from *m/z* 100-3000. Up to three of the most intense multiple charged ions (1+, 2+, and 3+) per scan were fragmented (collision-induced dissociation) in the linear ion trap to obtain tandem mass spectra. All tandem mass spectra were collected using 1.00 V of fragmentation amplitude, an isolation window of 4.0 u and scanning speed of 52,000 u per sec.

For quantitation analysis, iTRAQ labelled strong cation exchange-fractionated peptides were eluted from an Intensity C18P column (Bruker-Michrom Bioresources) using a gradient of 2-65% solvent B in 60 min at a flow rate of 800 nL min^{-1}. The peptide ions were electrosprayed into the Impact HD quadrupole time-of-flight mass spectrometer (Bruker Daltonik) with a spray voltage of 1300 V and a capillary temperature of 150°C. Precursor survey scan was acquired with a mass range of from *m/z* 350-1200. MS/MS settings consisted of scan ranged from *m/z* 50-2200. Up to 10 of the most intense multiple charged ions (1+, 2+, and 3+) per scan were fragmented via collision-induced dissociation to generate tandem mass spectra.

4.3.7 Data acquisition and analysis

Data acquisition (in positive mode) was executed with ESI Compass 1.4 for amaZon (trapControl Version 7.1) (Bruker Daltonik). The mass spectrometer was externally calibrated using Electrospray Tuning Mix Positive (Agilent Technologies, Inc., Santa Clara, CA, USA). Analyses of the data were performed using the Compass DataAnalysis (Version 4.1 SR1) (Bruker Daltonik). Peak lists were generated using the DataAnalysis Script Editor. These peak lists in eXtensible Markup Language (XML) format were then sent to Bruker ProteinScape data management software (Version 3.1) (Bruker Daltonik) for protein identification. The protein identifications were executed with tandem mass spectrometry peak lists using Mascot software (Matrix Science, London, UK). The peptide sequences were searched against

Arabidopsis thaliana, *Elaeis* sp.*, Zea mays, Brassica oryza* plant taxonomies (308,973 sequences as on 24th December 2013) in the Uniprot protein database and the protein list was compiled using the Protein Extractor module in ProteinScape. Mass tolerances for peptide and fragment ion were set to 100 ppm and 0.6 Da, respectively; and the instrument setting was specified as 'ESI trap'. Semi-trypsin was designated as the protease with two missing cleavages allowed. Carbamidomethylation on cysteine was set as the fixed modification while oxidation of methionine and deamidation of asparagine and glutamine were searched as variable modifications. Proteins were accepted if they had at least one top ranking peptide (Rank 1) with a Mascot ion score of more than 50.0, indicating identity or extensive homology ($p < 0.05$). All database searches were also performed on the 'decoy' database with 2% false discovery rate (FDR). The decoy database consisted of randomised sequences of the searched taxonomies. Individual identified protein lists were compiled using the Protein Extractor module in ProteinScape.

4.3.8 iTRAQ quantitation analysis

Data acquisitions were performed using Compass for otofSeries 1.5 (for otofControl Version 3.2) (Bruker Daltonik). The mass spectrometer was externally calibrated using Electrospray Tuning Mix Positive (Agilent Technologies, Inc., Santa Clara, CA, USA). Analyses of the data were performed using the Compass DataAnalysis (Version 4.1 SR1) (Bruker Daltonik). Peak lists were generated using the DataAnalysis Script Editor. These XML format peak lists were then sent to ProteinScape data management software (Version 3.1) (Bruker Daltonik). For protein identification, the peptide sequences were searched against *Arabidopsis thaliana*, *Elaeis* sp.*, Zea mays, Brassica oryza* plant taxonomies (308,973 sequence entries as on 24th December 2013) in the Uniprot protein database. Peptide and fragment ion mass tolerances were set to 20 ppm and 0.1 Da, respectively; and the instrument setting was specified as 'ESI-QUAD-TOF'. Semi-trypsin was designated as the protease with two missed cleavage allowed. Carbamidomethylation on cysteine and iTRAQ8plex (on lysine and *N*-terminal) were set as the fixed modifications while oxidation of methionine, deamidation (of asparagine and glutamine) and iTRAQ8plex of tyrosine were set as variable modifications. Proteins were accepted if they had at least one top ranking peptide (Rank 1) with a Mascot ion score of more than 50.0, indicating identity or extensive homology ($p < 0.05$). All database searches were also performed on the 'decoy' database. Individual identified protein list of each strong cation exchange fractions and two iTRAQ experiments were compiled using the Protein Extractor

module in ProteinScape. For relative protein quantitation analysis of identified iTRAQ labelled peptides with in-house Mascot server, the XML format peak lists of all the strong cation exchange fractions were compiled into a single Mascot Generic Format file using the “Export MGF” functionality in ProteinScape. The quantitation analysis was performed using Mascot algorithm based on the Reporter Protocol (Matrix Science, London, UK). The iTRAQ quantitation method contained the mass of the reporter ions used, their ratios and the specific modifications for iTRAQ (as described above). Rosner’s method were used to identify outliers and critical values for a significance level (p-value) 0.05 were used. At least two peptides were used for the quantitation of their abundance.

4.4 Results

4.4.1 The extent of biological variations between replicates

A two-dimensional gel electrophoresis (2DE) approach was used to analyse the protein profile of three biological replicates from the low oleic acid and high oleic acid varieties. The effects of 2-hydroxyethyl disulfide as a destreaking reagent, non-linear IPG strip and different pH ranges of IPG strip were first evaluated to enhance the effectiveness of this technique in term of resolution and reproducibility. Chromoplast proteins for both varieties were separated on a linear pH 3-10 IPG strip without the presence of 2-hydroxyethyl disulfide (Figure 4.3). The appearance of smeared protein spots and streaking was apparent in those gels, especially in the basic region.

Figure 4.4 shows the effects of a non-linear IPG strip and addition of 2-hydroxyethyl disulfide on the protein spot resolution and quality of the gel. Overall, there was better resolution of proteins in the basic region of the gel. Presence of 2-hydroxyethyl disulfide also reduced the horizontal streaking in the gel. Relatively, the number of proteins spots visualised in the 2DE gel of chromoplast proteins using the combination of a 2-hydroxyethyl disulfide and non-linear IPG strip was higher.

Most of the key fatty acid biosynthetic enzymes have basic isoelectric points, causing them to be focussed in the basic region of a 2DE gel. Increased protein load, length of the IPG strip and even a narrower pH gradient could aggravate the streaking effect of basic proteins during the focussing step. Migration of reducing agent, dithiothreitol from the basic part of the IPG strip during focussing and non-specific oxidation of basic proteins poses other common problems in 2DE technique, resulting in poor resolution and reproducibility of protein spots. 2-hydroxyethyl disulfide reduces the formation of inter- and intramolecular disulfide bonds between proteins that primarily cause the streaking by maintaining the protein thiol groups in an oxidised state (Acin, Rayo, Guerrero, & Quero, 2009; Depagne & Chevalier, 2012; Lamberti et al., 2007). Therefore, it acted as a “destreaking” reagent to eliminate most of the streaking from the gel as observed in Figure 4.4. In contrast, resolution of proteins in the acidic region of all the gels did not change much compared to those in the basic region, even with the addition of 2-hydroxyehyl disulfide. Since the sulfhydryl groups of cysteines in proteins are stabilised as mixed disulfides at low pH, the reoxidation of thiol groups is minimised.

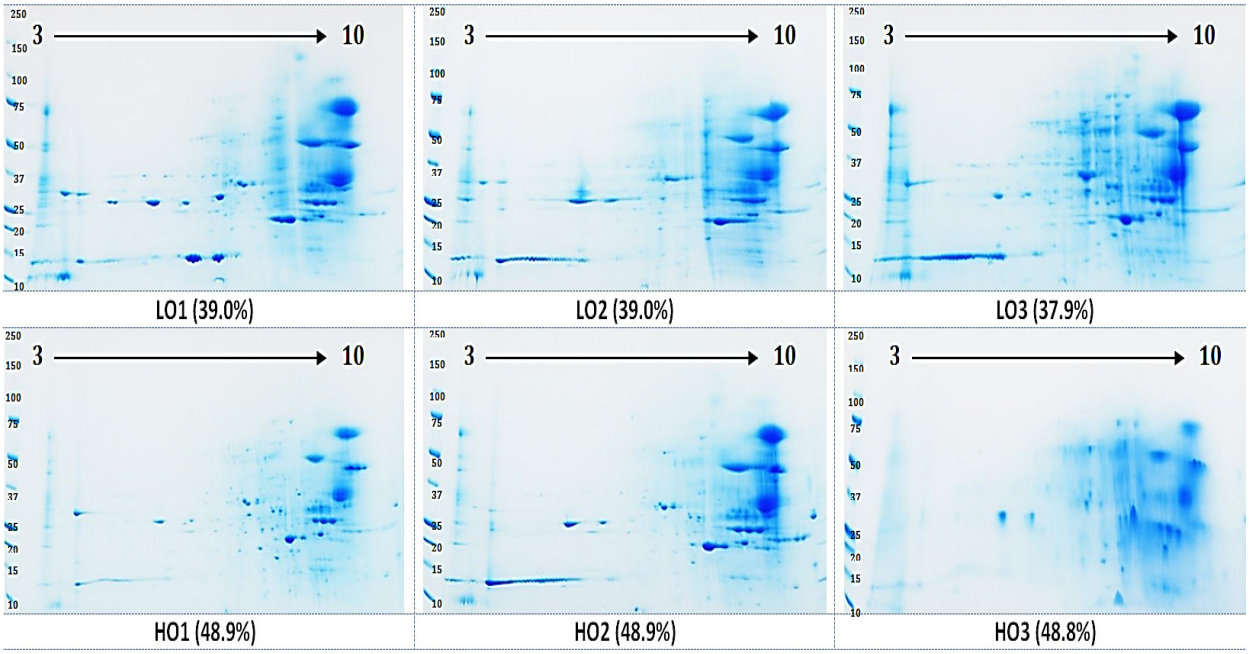


Figure 4.3. The biological replicates of low oleic acid (LO1-3) and high oleic acid (HO1-3) varieties. The 11 cm linear IPG strip, pH 3-10 was rehydrated with 200 µg of proteins. Proteins were separated with a precast 4-20% gradient SDS-PAGE gel. Percentage (%) indicates the amount of oleic acid present in each replicate.

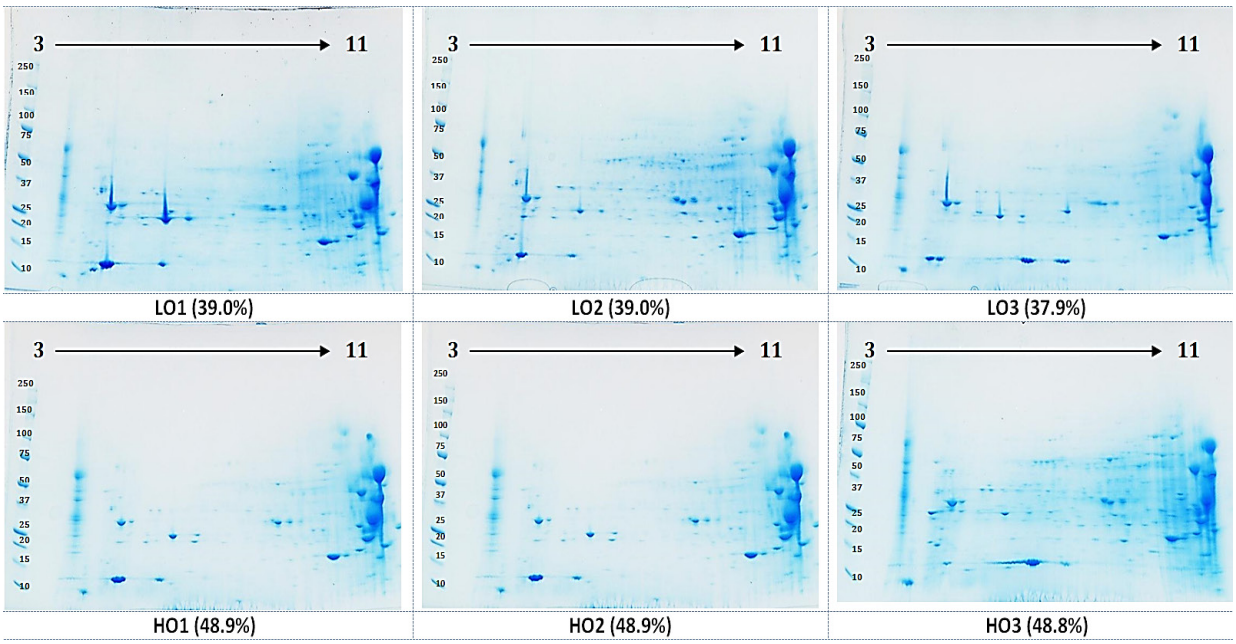


Figure 4.4. The biological replicates of low oleic acid (LO1-3) and high oleic acid (HO1-3) varieties. 11 cm non-linear IPG strip, pH 3-11 was rehydrated with 200 µg of proteins. Proteins were separated with a precast 4-20% gradient SDS-PAGE gel. 2-hydroxyethyl disulfide was added to all samples prior to the first dimension separation. Percentage (%) indicates the amount of oleic acid present in each replicate.

Non-linear IPG strips were used to enable the proteins to be focussed more evenly along the length of the IPG. Hence the resolution of the proteins in the region of pH 5-7 (in a pH 3-11 IPG strip) was improved (Figure 4.4) compared to a linear IPG strip (Figure 4.3). With the execption of acetyl-CoA carboxylase beta subunit (isoelectric point, pI of 5.6) and β-ketoacyl-ACP reductase (pI of 5.1), most of the fatty acid biosynthetic enzymes have a pI between 6.4 and 10.0. Therefore, a non-linear IPG strip proved to be ideal for these proteins.

In addition to using an additional reducing agent (2-hydroxyethyl disulphide) and non-linear IPG strip, a narrower basic range of pH (for the IPG strip) was employed to 'amplify' the low-abundance fatty acid biosynthetic enzymes of the oil palm chromoplast. Figure 4.5 shows the protein profiles of the chromoplast proteins from low oleic acid and high oleic acid varieties focussed using pH 7-11 non-linear IPG strips. Sixty nine protein spots from the gel for the low oleic acid variety (Figure 4.6) were digested and analysed mass spectrometrically to determine their protein identities (data not shown), in a search for the key enzymes involved in fatty acid biosynthesis. However, none of the proteins identified were associated with fatty acid biosynthesis. These results indicated that there was likely to be insufficient detection of low-abundance fatty acid biosynthetic enzymes using these approaches. The narrow range of the IPG strip also caused extremely alkaline proteins to streak as demonstrated in Figure 4.5. The 2DE gel profiles (Figure 4.5) suggested that the fatty acid biosynthetic proteins might be present in very low-abundance and could not be stained with colloidal Coomassie Brilliant Blue. Hence the same gel was subsequently stained with silver nitrate to enhance the detection of low-abundance proteins (around nanogram range). The silver stain revealed (Figure 4.7) many proteins undetectable using Coomassie Brilliant Blue stain and that possibly may be the reason why none of the sixty nine protein spots (as detected with Coomassie Brilliant Blue) analysed contained fatty acid biosynthetic proteins.

A comparative profile analysis was performed on eleven reference protein bands separated with one-dimensional SDS-PAGE using Phoretix 1D analysis software (Figure 4.8). The analysis revealed that the intensities of the reference protein bands were similar across the low oleic acid and high oleic acid replicates (Figure 4.9 and Figure 4.10). These findings showed that expression profiles for major proteins in bands 1 to 11 from the replicates were similar. Thus, this indicated that the variation in proteins between the two palm varieties was minimal. A more in-depth evaluation of the replicates of low oleic acid and high oleic acid varieties was

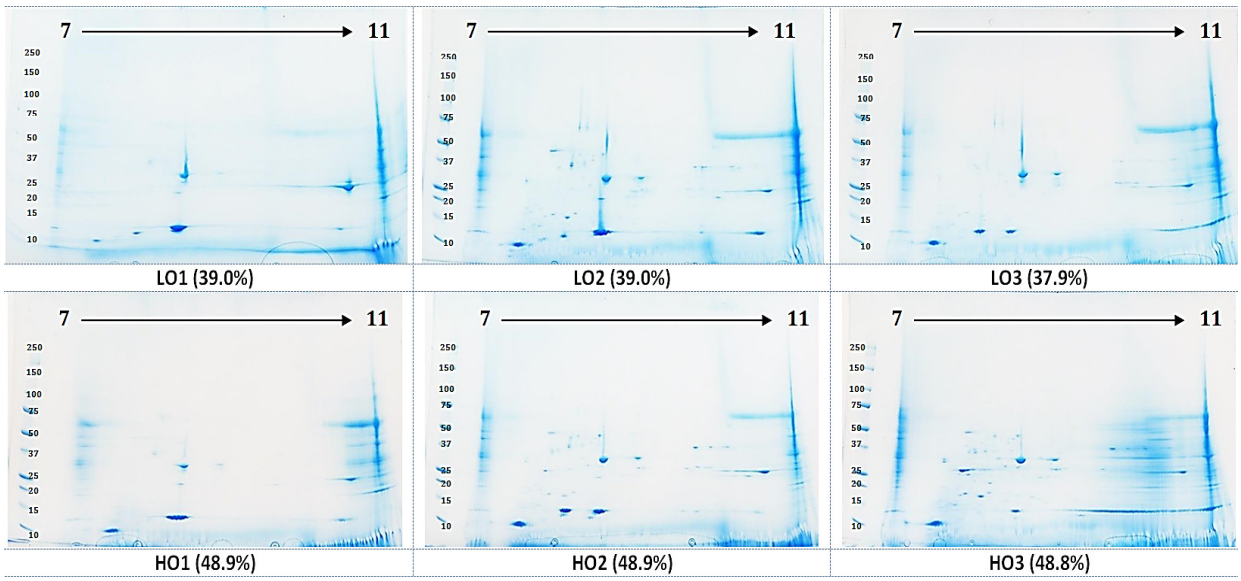


Figure 4.5. The biological replicates of low oleic acid (LO1-3) and high oleic acid (HO1-3) varieties. The 11 cm non-linear IPG strip, pH 7-11 was rehydrated with 200 µg of proteins. Proteins were separated with a precast 4-20% gradient SDS-PAGE gel. 2-hdryoxyethyl disulfide was added to all samples prior to the first dimension separation. Percentage (%) indicated the amount of oleic acid present in each replicate.

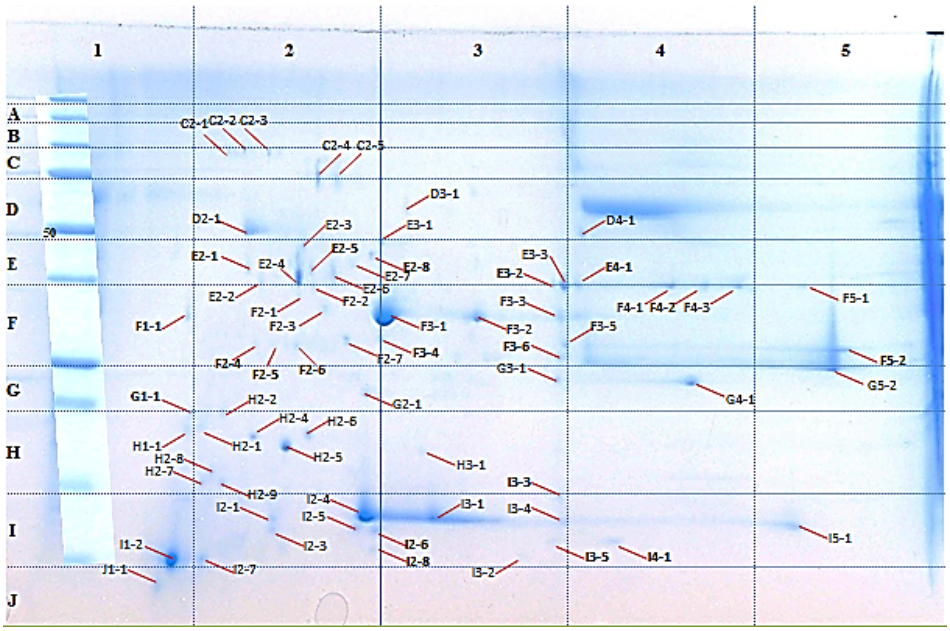


Figure 4.6. Spots excised from 2DE gel for low oleic acid variety. The 11 cm non-linear IPG strip, pH 7-11 was rehydrated with 200 µg of proteins. Proteins were separated with a precast 4-20% gradient SDS-PAGE gel.

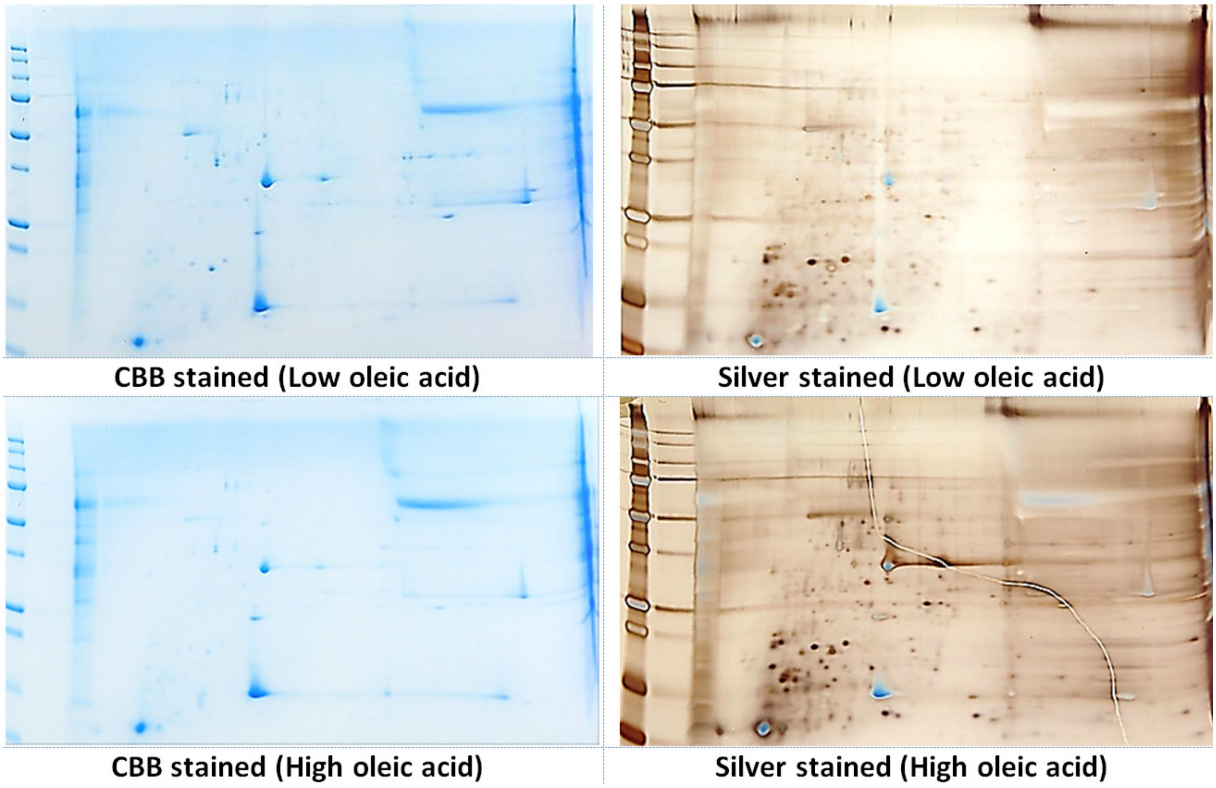


Figure 4.7. Utilisation of silver stain (right) revealed more proteins previously undetectable with Coomassie Brilliant Blue (CBB). The 11 cm non-linear IPG strip, pH 7-11 was rehydrated with 200 µg of proteins. Proteins were separated with a precast 4-20% gradient SDS-PAGE gel.

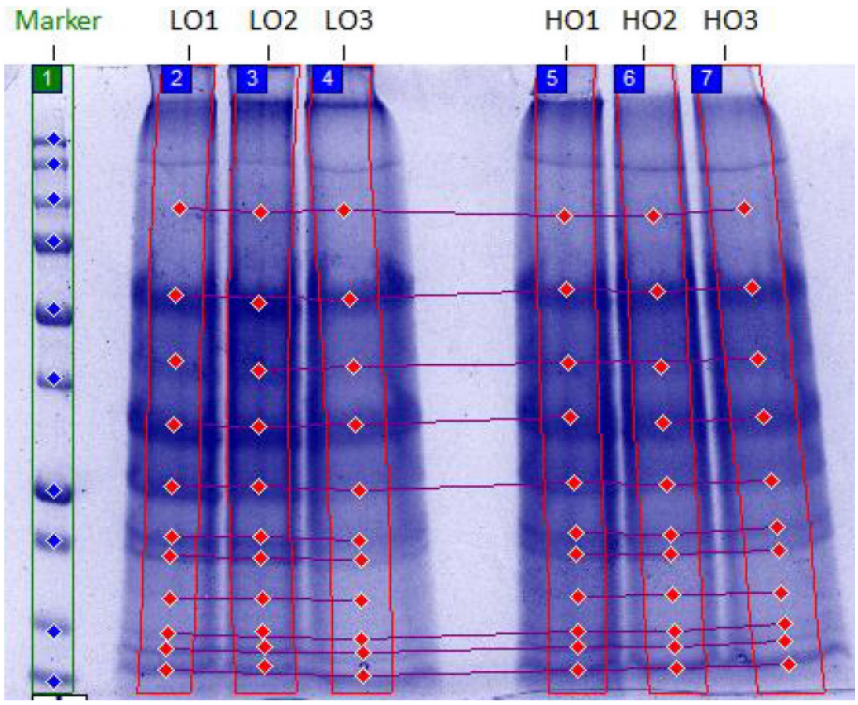


Figure 4.8. The reference protein bands used for comparative profile analysis (♦). LO1, LO2 and LO3 represented low oleic acid replicates. HO1, HO2 and HO3 represented high oleic acid replicates. The gel bands were normalised using the Retardation factor (Rf) of the protein bands with Phoretix 1D. The normalisation is to minimise the effect of non-uniform migration distance in protein band profile comparisons.

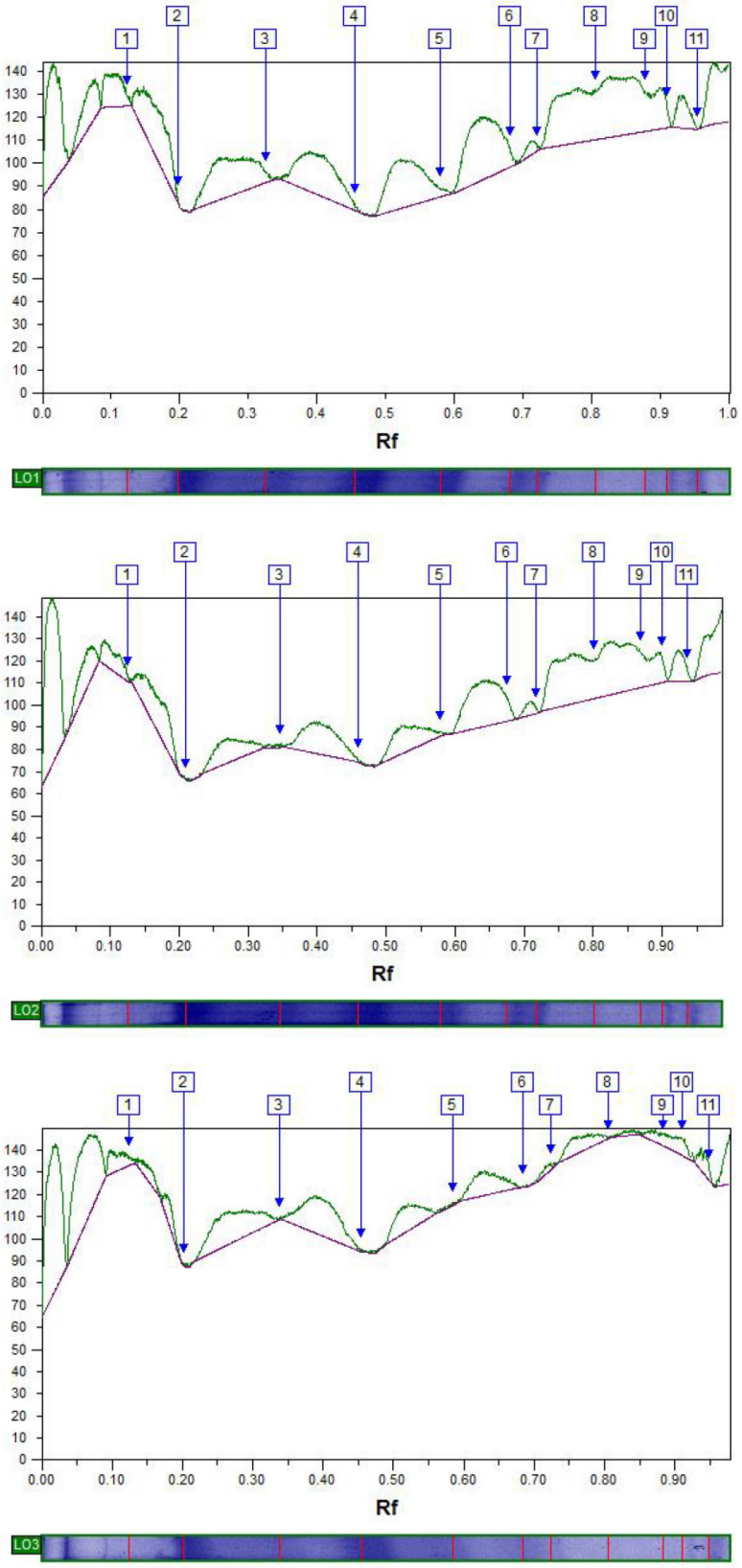


Figure 4.9. Qualitative intensity comparison of the reference protein bands. 11 individual protein band patterns low oleic acid (LO) replicates were compared using Phoretix 1D.

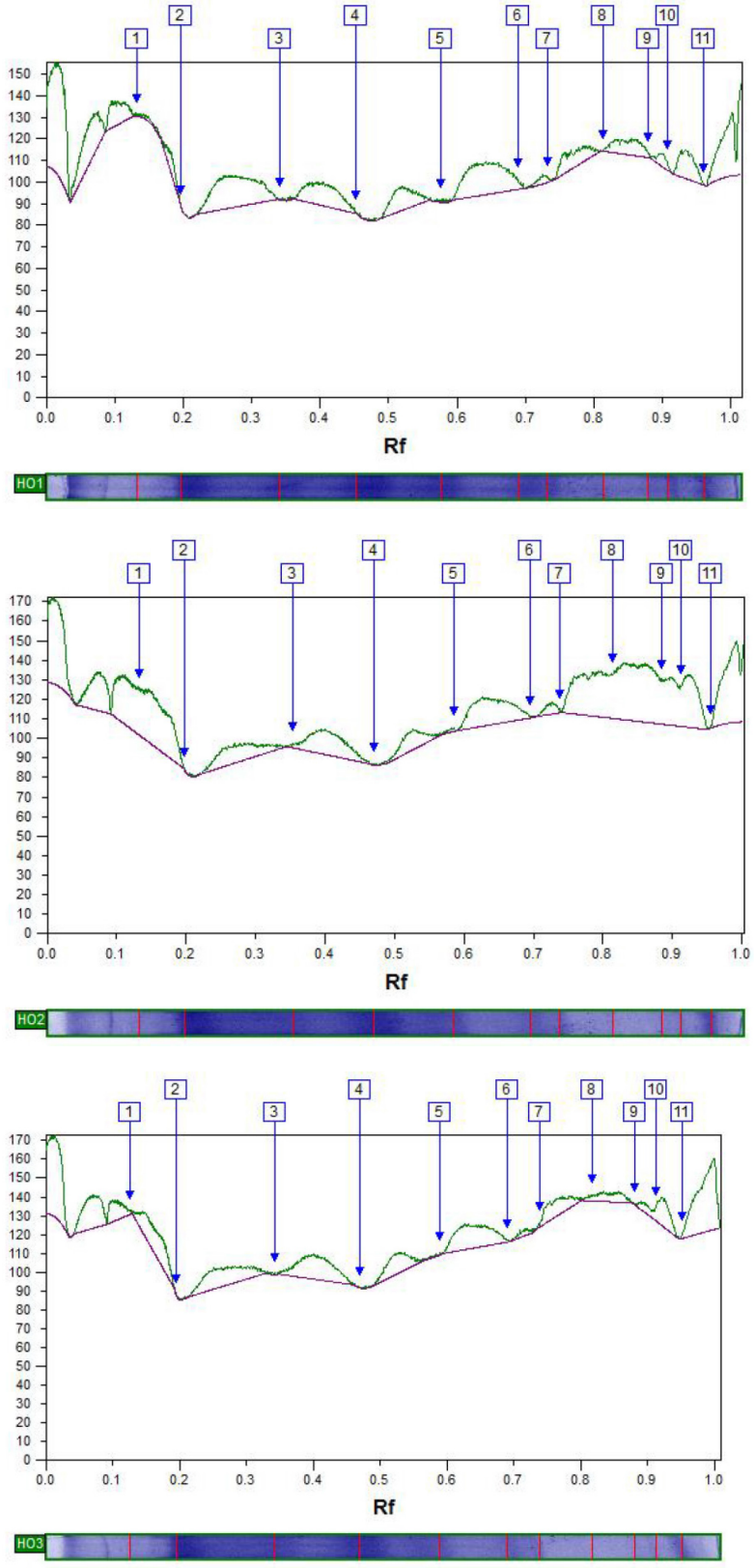


Figure 4.10. Qualitative intensity comparison of the reference protein bands. 11 individual protein band patterns from each high oleic acid (HO) replicates were compared using Phoretix 1D.

Table 4.4. Comparison of protein and peptide yields for low oleic acid (LO) and high oleic acid (HO) replicates. [a] and [b] indicated the two different plantation areas in which the sampling was made.

Low oleic acid variety			High oleic acid variety		
Replicate	**Proteins**	**Peptides**	**Replicate**	**Proteins**	**Peptides**
LO1	72[a]	245	**HO1**	66[a]	216
LO2	59[a]	173	**HO2**	50[b]	143
LO3	55[b]	176	**HO3**	62[b]	175

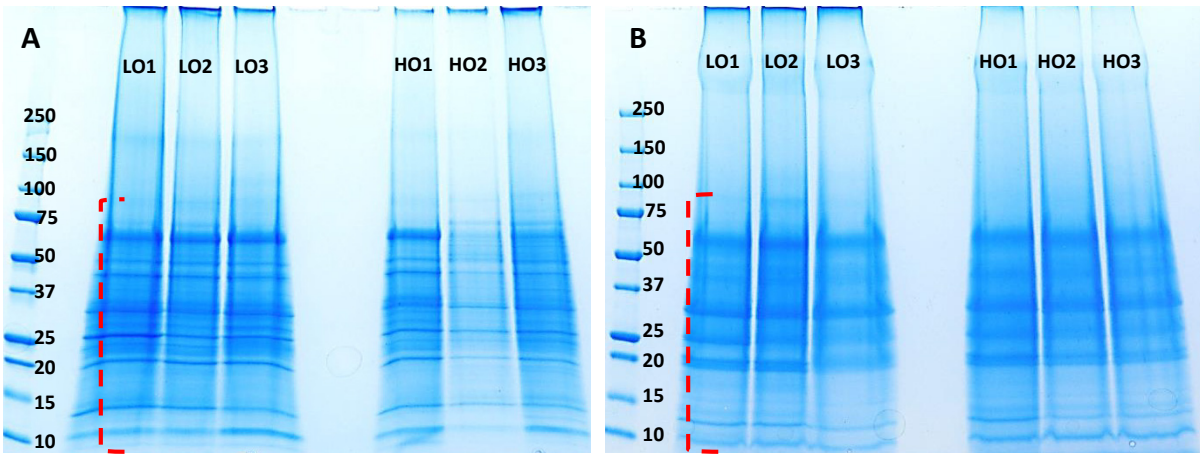
performed using a complementary liquid-chromatography tandem mass spectrometry (LC-MS/MS) approach to further assess if there were any differences between those replicates. Table 4.4 compares the number of proteins and peptides detected for each of the replicates from two different palm varieties. The results demonstrate that the number of identified proteins for LO1 (low oleic acid replicate 1) was slightly higher compared to the other two replicates. Replicate 1 for high oleic acid variety (HO1) had the highest number of proteins identified (66). These slight biological variations in the number of proteins identified could be caused by the different sampling areas, harvesting time of the oil palm fruits or random chance. Figure 4.4 demonstrated that there were some obvious differences in the expression of proteins from the two different varieties of oil palm tree and the same conclusion was made based on results obtained from mass spectrometric analysis. Only 48 proteins (193 peptides) were identified in both the low oleic acid and high oleic acid varieties (compiled from three biological replicates). The low oleic acid variety had 58 unique proteins (148 peptides) while the high oleic acid variety had 50 unique proteins (110 peptides). There were clear protein presence differences between the low oleic acid and high oleic acid varieties. These results revealed that there were definitely biological differences to be discovered, and that the hypothesis of differentially expressed fatty acid biosynthetic enzymes was worthy of further exploration.

4.4.2 GeLC-MS/MS to detect the fatty acid biosynthetic enzymes

SDS-PAGE followed by in-gel digestion, termed GeLC-MS/MS, is currently a widely used approach in proteomics. GeLC-MS/MS also offers the benefit of fractionation for complex protein mixtures as different gel segments are digested prior to LC-MS/MS. Prior to utilising GeLC-MS/MS, the percentage of polyacrylamide gel used to separate the chromoplast

proteins from low oleic acid and high oleic acid varieties was optimised. As presented in Figure 4.11, protein bands were well resolved on the 4-20% gradient SDS-PAGE gel and qualitative observation indicated that the overall protein profiles for the biological replicates from both varieties were similar. Protein loads for the replicates were increased 2-fold to 100 µg but major differences in terms of protein profile were still not apparent (Figure 4.11). Instead, the resolution of the protein bands decreased with the increment in protein load. Subsequently, the gel lane ranging from 10 kDa to about 100 kDa (indicated with red dotted line) was excised into eight similar segments and was in-gel digested. Sixty eight non-redundant proteins were identified from both the low oleic acid and high oleic acid replicates. Neither acetyl-CoA carboxylase nor β-ketoacyl-ACP synthase were detected at this stage. Other fatty acid biosynthetic enzymes identified were β-ketoacyl-ACP reductase, β-hydroxyacyl-ACP dehydrogenase, 3-enoyl-ACP reductase, β-ketoacyl-ACP synthase and stearoyl-ACP desaturase.

A single percentage SDS-PAGE gel of 12% was used to improve the number of identified proteins associated with fatty acid biosynthesis. Figure 4.12 showed that the overall protein profiles of the biological replicates from both varieties were similar, just as visualised on the 4-20% gradient SDS-PAGE gel. The resolution of the protein bands was inferior because the 12% SDS-PAGE gel was not a gradient gel. The same GeLC-MS/MS strategy was applied on the gel. Database search results revealed that the number of identified proteins with Mascot score of more than 60 from both low oleic acid and high oleic acid replicates was elevated to almost 2-fold or 132 non-redundant proteins (Figure 4.13). The utilisation of a single-percentage 12% SDS-PAGE gel for the separation of proteins between 10 kDa to about 100 kDa also resulted in a significant increase (31%) in the number of the identified peptides. GeLC-MS/MS of 12% SDS-PAGE gel generated 228 more peptides with a Mascot score of more than 35 (Figure 4.13). Furthermore, in contrast to the 4-20% gradient SDS-PAGE gel, β-ketoacyl-ACP synthase was successfully detected from the excised gel segments (Figure 4.12) but acetyl-CoA carboxylase remained undetected. Table 4.5 reveals that chromoplast proteins separated on a single-percentage 12% SDS-PAGE yielded significantly better outcomes in terms of number of peptides and sequence coverage for the enzymes involved in fatty acid biosynthesis.



50 µg 100 µg

Figure 4.11. The SDS-PAGE gel of the low oleic acid and high oleic acid biological replicates. 50 µg (A) and 100 µg (B) of proteins were loaded in each lane. Proteins were separated with a precast 4-20% gradient SDS-PAGE gel. Red dotted lines indicated the region excised into eight segments for GeLC-MS/MS.

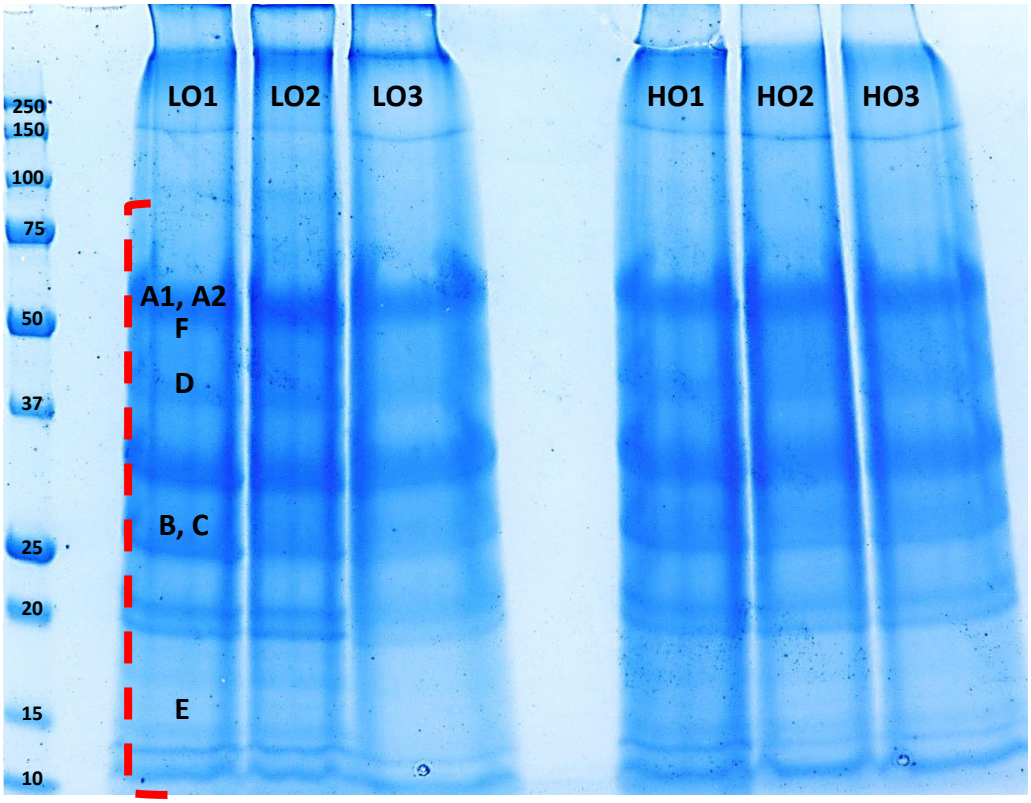


Figure 4.12. The 12% SDS-PAGE gel of low oleic acid and high oleic acid biological replicates. 100 µg of proteins was loaded in each lane. Proteins were separated with a precast 12% SDS-PAGE gel. The location of fatty acid biosynthetic enzymes in the gel (as determined experimentally) were indicated; A1. acetyl-CoA carboxylase; A2. acetyl-CoA carboxylase subunit beta (theoretical); B. β-ketoacyl-ACP reductase; C. β-hydroxyacyl-ACP dehydrogenase; D. 3-enoyl-ACP reductase; E. β-ketoacyl-ACP synthase; F. stearoyl-ACP desaturase.

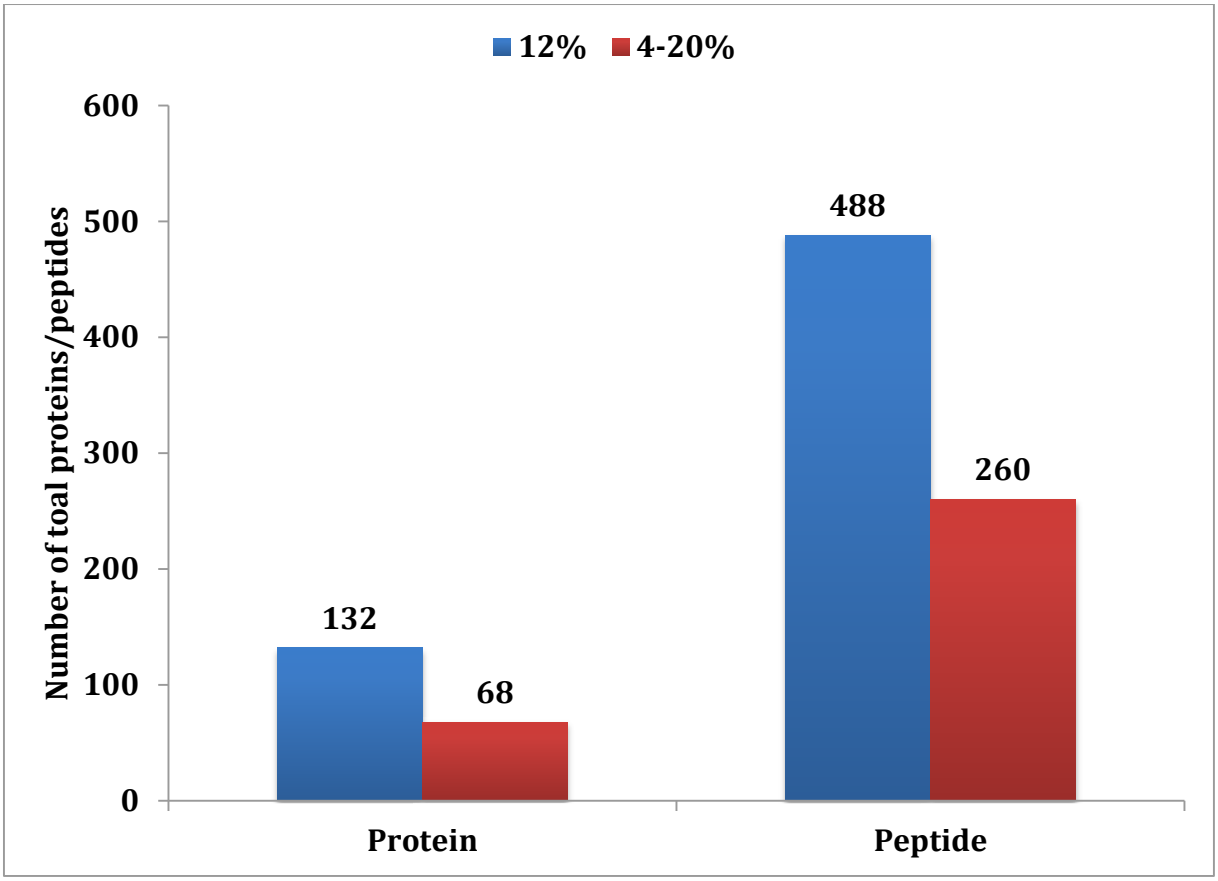


Figure 4.13. Total number of proteins and peptides observed with a Mascot score of more than 60 and 35, respectively. The results were obtained with GeLC-MS/MS on 12% SDS-PAGE and 4-20% gradient SDS-PAGE gels.

Table 4.5. Peptides and sequence coverage for each fatty acid biosynthetic enzymes, as compiled from low oleic acid and high oleic acid biological replicates. These results were obtained from GeLC-MS/MS on single-percentage 12% and 4-20% gradient SDS-PAGE gels.

Enzymes	Gel percentages			
	Continuous 12%		Gradient 4-20%	
	Number of peptides	Sequence coverage (%)	Number of peptides	Sequence coverage (%)
β-ketoacyl-ACP reductase	12	61.6	8	37.6
β-hydroxyacyl-ACP dehydrogenase	10	35.0	7	22.3
3-enoyl-ACP reductase	23	58.3	15	43.1
β-ketoacyl-ACP synthase	2	16.2	0	0
Stearoyl-ACP desaturase	2	22.3	1	7.6

Consequently, GeLC-MS/MS analyses on the biological replicates for low oleic acid and high oleic acid separated with a 12% SDS-PAGE gel were performed to observe the number of peptides detectable from each variety. Figure 4.14 presents the number of non-redundant unique peptides and sequence coverage corresponding to the fatty acid biosynthetic enzymes, compiled from three biological replicates for each variety. The mass spectrometric analysis revealed that stearoyl-ACP desaturase (a key enzyme in oleic acid biosynthesis) from high oleic acid variety had less number of detected peptides (2) compared to low oleic acid (4). The difference of sequence coverage of stearoyl-ACP desaturase between these two varieties was 2.3%, indicating longer peptides were detected for the high oleic acid variety. β-ketoacyl-ACP reductase from low oleic acid variety had additional coverage of 10.6% in its protein sequence. This corresponded to a higher number of detected peptides identified to β-ketoacyl-ACP reductase in the low oleic acid variety. Other fatty acid biosynthetic enzymes from the two varieties had similar number of detected peptides and percentages of the sequence coverages. Table 4.6 shows the fatty acid biosynthetic unique peptide sequences compiled from low oleic acid and high oleic acid biological replicates.

Enzymes involved in fatty acid biosynthesis have theoretical molecular weights ranging from 15 kDa to 60 kDa, with the exception of the carboxyl transferase subunit alpha of acetyl-CoA carboxylase (81 kDa). Different percentage SDS-PAGE gels were tested in the search for the fatty acid biosynthetic enzymes. Gradient gels are effective in separating proteins with a broad range of molecular weights, especially high molecular weight proteins. The reason is that the top part of the gel contains larger pore sizes, therefore allowing resolution of large proteins. However, the decreasing pore sizes towards the bottom of a 4-20% SDS-PAGE gel prevent small proteins from being separated effectively. Hence, broad range gradient gels such as the 4-20% are more suited for global evaluation of protein profile and typically, give better protein band resolution. β-ketoacyl-ACP synthase has a theoretical molecular weight of 15.7 kDa and is likely present in low-abundance. The same gel segment of 4-20% gradient SDS-PAGE contained other higher-abundance proteins, thus likely masking β-ketoacyl-ACP synthase so that it was not detected mass spectrometrically. Indeed, low-abundance proteins analysed with an electrospray ionisation-based mass spectrometer system have less chance to get ionised compared to high-abundance proteins if present in the same gel segment (Wang, Xie, Young, & Li, 2009).

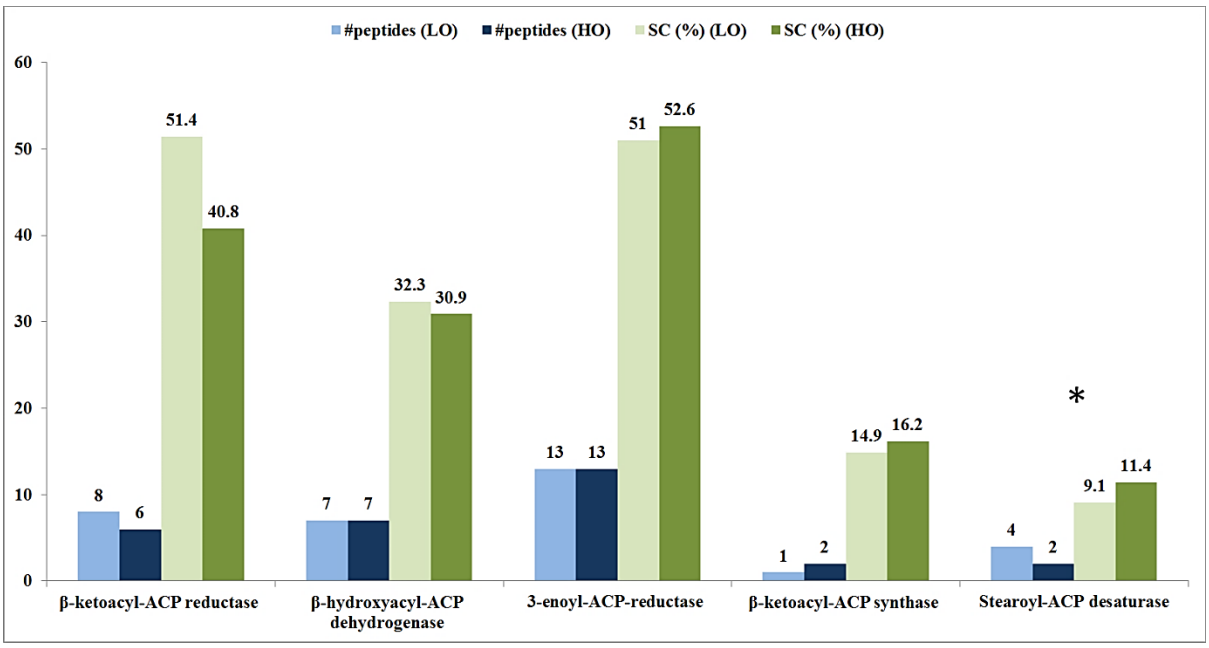


Figure 4.14. Number of unique peptides and sequence coverage (SC) per fatty acid biosynthetic enzyme obtained through GeLC-MS/MS. The results were compiled from three biological replicates for each low oleic acid (LO) and high oleic acid (HO) variety. * denotes the key enzyme in oleic acid biosynthesis.

Table 4.6. Unique peptides detected with GeLC-MS/MS from low oleic acid (LO) and high oleic acid (HO) varieties that matched to the respective fatty acid biosynthetic enzymes. Peptides not detected corresponding to LO or HO are denoted with (**x**).

Enzyme	Accession	LO	HO	Unique peptide sequence
β-ketoacyl-ACP reductase	G8FGI6		x	AGVIGFTK
		x		AIALALGK
			x	EIEASGGQAIIIFFGDISKEDDVESMIK
				IAADAWGTVDILVNNAGITR
				IINIASVVGLTGNAGQANYSASK
				ILQTIPLGR
				VLINYATSTEEAEEVSK
		x		NINVNAVAPGFIASPMTAQLGEDVEKK
β-hydroxyacyl-ACP dehydrogenase	G8FGF9	x		RIPPFPTVMDINQIR
				TEEEVPIEK
				ENFFFAGIDK
				FPFLLVDR
			x	FRKPVIAGDTLVMR
				IPPFPTVMDINQIR
				VIEYQPGVTAVGIK
				KPVIAGDTLVMR
3-enoyl-ACP reductase	H6TNL1			AALESDTMVLAFEAGR
				GVVSGLPIDLR
				GYLAAISASSYSFVSLLK
				HFLPIMNPGGASISLTYIASER
				MIEYSYANAPLQK
				NDFGSIDILVHSLANGPEVTKPLLETSR
				VYPLDAVYDTPEDVPDDVR
				AFVAGVADDNGYGWAIAK
			x	GKFDESR
				KGYLAAISASSYSFVSLLK
				RLPNGSLMEIIK
		x		TIPGYGGGMSSAK
				VNTISAGPLGSR
				YAGASNWTVK
β-ketoacyl-ACP synthase	G8FGI5			QLALSDDTPVTGESK
		x		VAKPETVQK
Stearoyl-ACP desaturase	C325I9			ATFISHGNTAR
				DYADILEFLINR
			x	HGDLLNK
			x	IPFSWIYGR

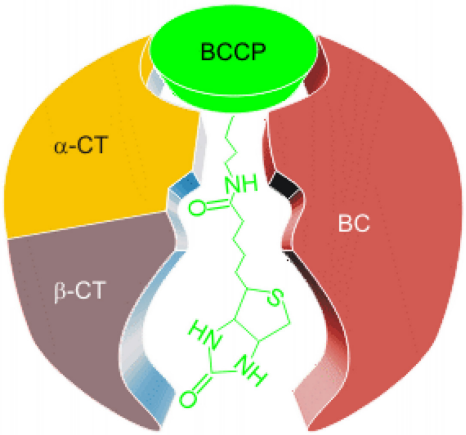


Figure 4.15. Heteromeric form of acetyl-CoA carboxylase in plant. Biotin-carboxyl carrier protein (BCCP), biotin carboxylase (BC), carboxyltransferase (CT) (Li-Beisson et al., 2013).

A single percentage 12% SDS-PAGE gel enabled β-ketoacyl-ACP synthase to be detected mass spectrometrically. Optimal protein separation in a single percentage gel occurs in the lower half of the gel. In addition, a continuous percentage gel is able to separate proteins that are close in molecular weights, as in the case of fatty acid biosynthetic enzymes. The advantage of a single-percentage gel in separating proteins with adjacent molecular weights was clearly demonstrated in this study as more peptides corresponding to fatty acid biosynthesis were detected, resulting in higher sequence coverage of the matched enzymes (Table 4.5). The GeLC-MS/MS analysis revealed that all these enzymes were detected in the gel segments corresponding to their theoretical molecular weights (Figure 4.12), except acetyl-CoA carboxylase. Acetyl-CoA carboxylase remained undetectable with the GeLC-MS/MS approach. The plant acetyl-CoA carboxylase is composed of a biotin carboxyl carrier protein, biotin carboxylase and carboxyltransferase (Figure 4.15). As mentioned earlier, only the carboxyltransferase subunit alpha has a high molecular weight (81 kDa), while the biotin carboxylase and biotin carboxyl carrier protein have molecular weights of 53 kDa and 21 kDa, respectively (Cronan Jr & Waldrop, 2002; Sasaki & Nagano, 2004a). Therefore, these two subunits of acetyl-CoA carboxylase were expected to be detected since the analysed gel segments ranged from 10 kDa to 100 kDa. This particular enzyme could have been expressed in a very low amount, in which case a complementary but much more sensitive approach to detect it was needed. Another possible reason this heteromeric enzyme was not detected could be that post-translational modifications had altered the molecular weight of the enzymes (Halligan et al., 2004; Podkowinski & Tworak, 2011), giving it a higher molecular weight, beyond the analysed molecular weight range.

4.4.3 Offline two-dimensional liquid chromatography (2DLC)-tandem mass spectrometry (MS/MS) to enhance the detection of fatty acid biosynthetic enzymes

A non-gel based approach in analysing the expression of fatty acid biosynthetic enzymes presented a potential alternative to a gel-based approach. Essentially, chromoplast peptides from different biological replicates of low oleic acid and high oleic acid varieties were fractionated using a strong cation exchange column before their subsequent separation with a reverse-phase column for mass spectrometric characterisation. Offline 2DLC-MS/MS has been proven to be a successful strategy to maximise the performance of a reverse-phase LC-MS/MS in other studies (Wang et al., 2009; Yun et al., 2011).

In order to determine the suitable fractionation range to capture all the fatty acid biosynthetic enzymes, different fractionation ranges were attempted (Table 4.7). Figure 4.16 shows an important trend on the effect of different fractionation ranges on the number of identified peptides and proteins. The number of peptides increased from 114 peptides (1st fractionation range) and 217 peptides (2nd fractionation range) to 497 peptides identified using the 3rd fractionation range (Figure 4.16). More peptides with a Mascot score above 35 were obtained using the 3rd fractionation range. A rise in the number of total proteins with a Mascot score of more than 60 with the 2nd and 3rd fractionation range was also observed compared to the 1st fractionation range (Figure 4.16).

Table 4.8 summarises the number of peptides related to fatty acid biosynthesis eluted with the tested fractionation ranges. Most of the fatty acid biosynthetic peptides eluted between 1% and 25%. The peptides eluted later (30-80%) were mostly the same peptides that had already eluted by 25%. Furthermore, the previously undetected β-ketoacyl-ACP synthase peptides (using the 1st and 2nd fractionation ranges) were successfully eluted using the 3rd fractionation range. Peptides for stearoyl-ACP desaturase were also eluted earlier in the 1% and 5% fractions (with the 3rd fractionation range). The successful detection of these two enzymes was very significant in terms of being able to evaluate the potential expression differences between low oleic acid and high oleic acid varieties. Stearoyl-ACP desaturase is the key enzyme in the conversion of saturated fatty acids (stearic acid) to unsaturated fatty acids (oleic acid).

Table 4.7. The three strong cation exchange fractionation ranges used in this study. The percentage (%) indicates the concentration of ammonium formate used.

Designation	Fractionation range tested
1st range	10%, 20%, 30%, 40%, 60%, 80%, 100%
2nd range	5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 100%
3rd range	1%, 5%, 10%, 20%, 30%, 40%, 60%, 80%, 100%

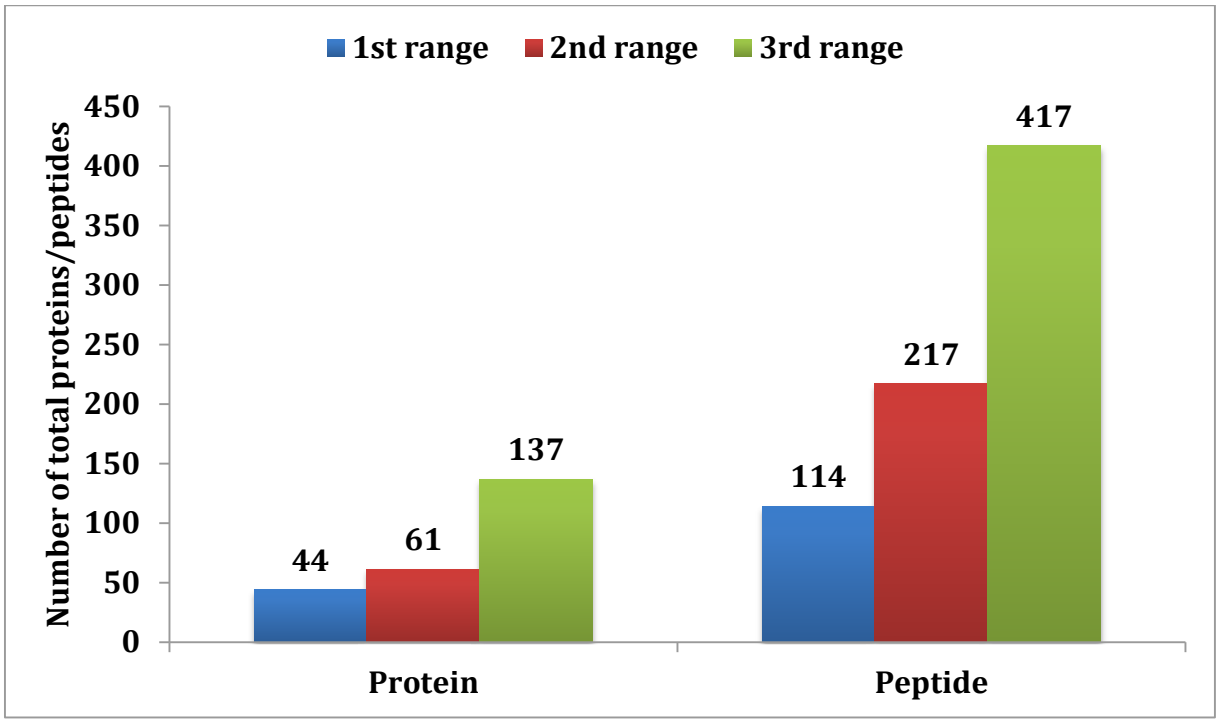


Figure 4.16. Total number of proteins and peptides observed with a Mascot score of more than 60 and 35, respectively. The results were obtained with different fractionation ranges using a strong cation exchange column.

Table 4.8. Distribution of peptides per fatty acid biosynthetic enzymes over the collected strong cation exchange fractions compiled from three biological replicates for low oleic acid and high oleic acid varieties. Coloured boxes indicate the different fractionation ranges. The numbers in the boxes indicate the number of peptides that matched to the respective fatty acid biosynthetic enzyme.

Enzymes	Fractions												
	Flow through	1	5	10	15	20	25	30	35	40	60	80	100
Acetyl-CoA carboxylase (biotin carboxylase)				2		4						1	1
			3										
				3									
Acetyl-CoA carboxylase (carboxyl transferase subunit beta)	2			1									
					1								
		2	1	2						1	1		
3-ketoacyl-ACP reductase				1				2		2			
	3		1	3	1								
	4	5	3	3				1			1		
3-hydroxyacyl-ACP dehydrogenase	2			2		3							2
	4		4	2				1	1				3
	3	4	3	2								1	
Enoyl-ACP reductase	3			5		1							
	3		3	1	3								
	6	5	3	2	1							1	1
β-ketoacyl-ACP synthase													
	1	2											
Stearoyl-ACP desaturase							1	2					
		1	1	1									

Flow through, 10%, 20%, 30%, 40%, 60%, 80%, 100%
Flow through, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 100%
Flow through, 1%, 5%, 10%, 20%, 30%, 40%, 60%, 80%, 100%

Table 4.8 also indicates that more peptides were eluted per fraction using the 3rd fractionation range while the 1st fractionation range generated the smallest number of eluted peptides per fraction. With the exception of acetyl-CoA carboxylase and stearoyl-ACP desaturase, all other peptides known to be related to fatty acid biosynthesis were eluted in the flow through, indicating that they were either present in higher-abundance relative to acetyl-CoA carboxylase and stearoyl-ACP desaturase or were more highly charged. Figure 4.17 and Figure 4.18 display the base peak chromatograms for all the biological replicates fractionated with the 3rd fractionation range. It was apparent that most peptides were able to be eluted using less than 25% (v/v) ammonium formate. The results on the chromatograms corresponded with the number of peptides eluted per fraction as presented in Table 4.8. Therefore, more fractionations could be done in an elution region that gave the most peptides (< 25%) while other regions (30-100%) could be collected less frequently.

Figure 4.19 presents the compiled number of non-redundant unique peptides and sequence coverage corresponding to the fatty acid biosynthetic enzymes. The overall outcome of the mass spectrometric analysis showed that the stearoyl-ACP desaturase from both the low oleic and high oleic acid varieties had almost similar number of peptides, covering 23-25% of the stearoyl-ACP desaturase full sequence. It was also evident that offline 2DLC-MS/MS approach identified more peptides that were matched to β-ketoacyl-ACP synthase and stearoyl-ACP desaturase than GeLC-MS/MS. In contrast, fewer peptides were detected for both β-hydroxyacyl-ACP dehydrogenase and 3-enoyl-ACP desaturase. Hence, their sequence coverage was also reduced. Only β-ketoacyl-ACP reductase showed a similar number of detected peptides and sequence coverage as obtained with GeLC-MS/MS. Table 4.9 shows the compiled fatty acid biosynthetic unique peptide sequences from low oleic acid and high oleic acid varieties. Comparison of the detected unique peptides from 2DLC-MS/MS with GeLC-MS/MS revealed that in general, more peptides matched to the fatty acid biosynthetic enzymes were identified with 2DLC-MS/MS (Table 4.10).

The GeLC-MS/MS analysis revealed that fatty acid biosynthetic enzymes were present in low-abundance. Hence, complementary 2DLC-MS/MS was used, which is able to overcome the obstacles in gel-based technique in detecting subtle quantitative differences of the fatty acid

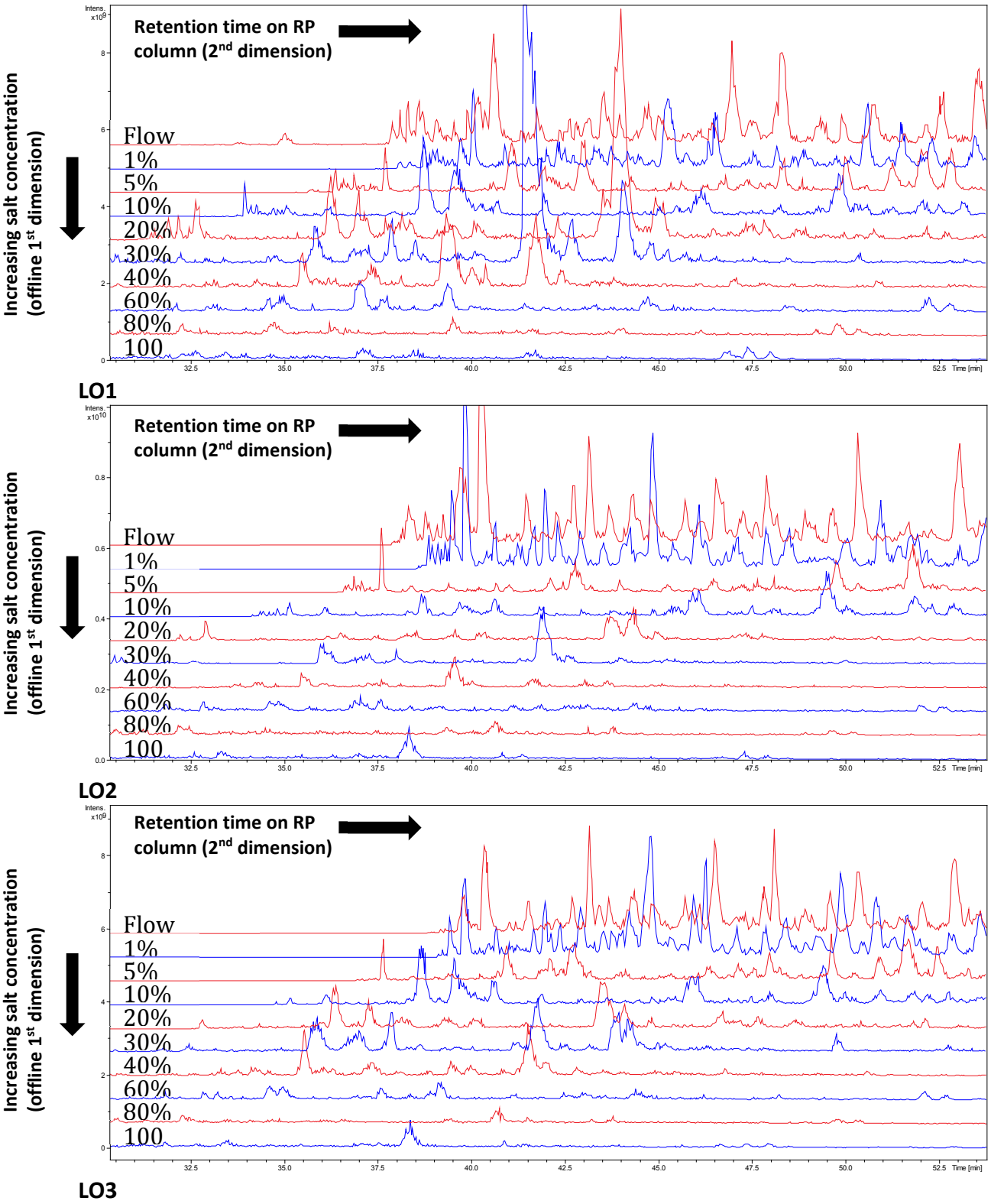


Figure 4.17. Stacked elution profiles (base peak chromatograms) of strong cation exchange eluates fractionated with the 3rd fractionation range and separated on a reverse-phase column. LO1-3 represent the biological replicates of low oleic acid variety.

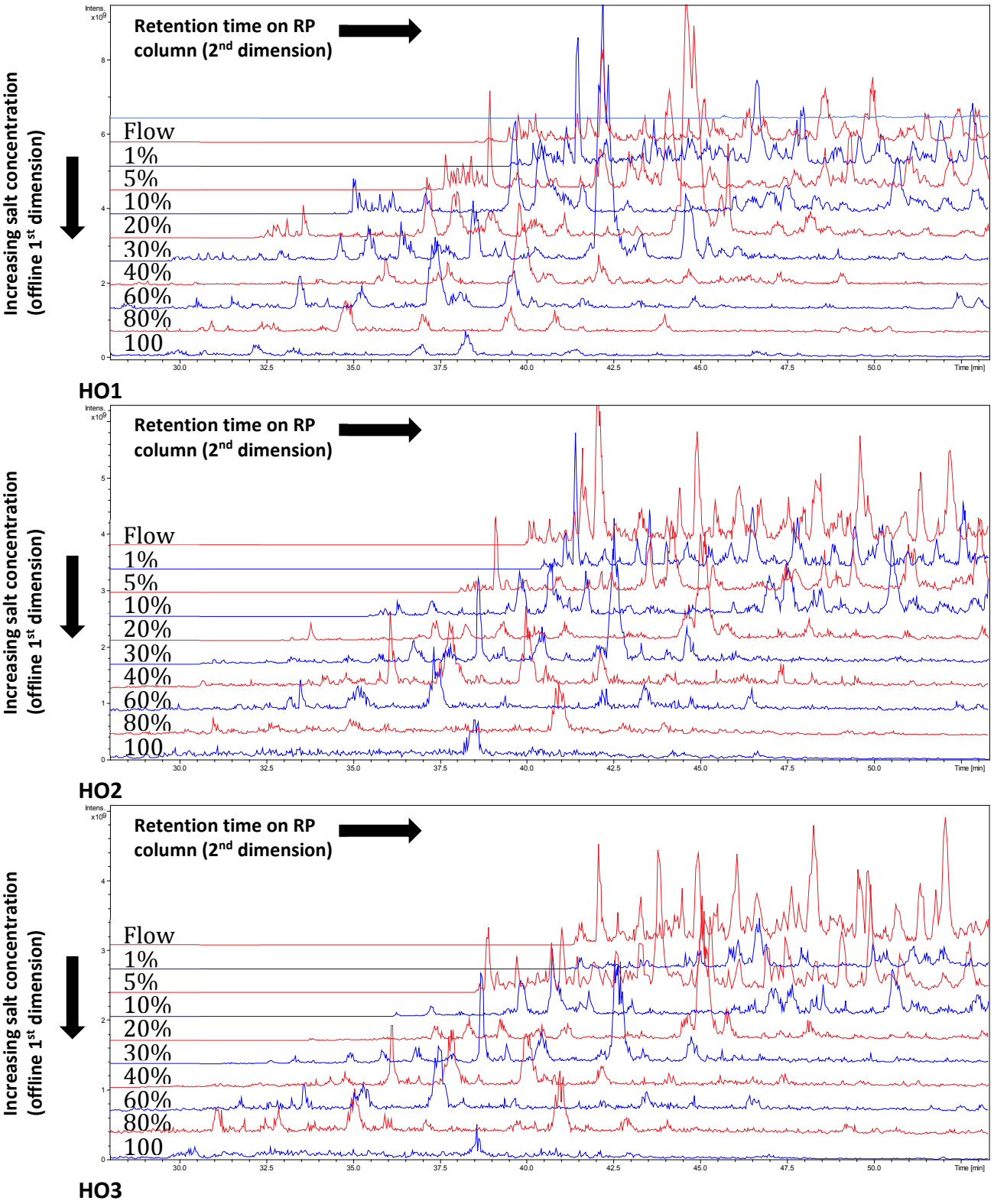


Figure 4.18. Stacked elution profiles (base peak chromatograms) of strong cation exchange eluates fractionated with the 3rd fractionation range and separated on a reverse phase column. HO1-3 represent the biological replicates of high oleic acid variety.

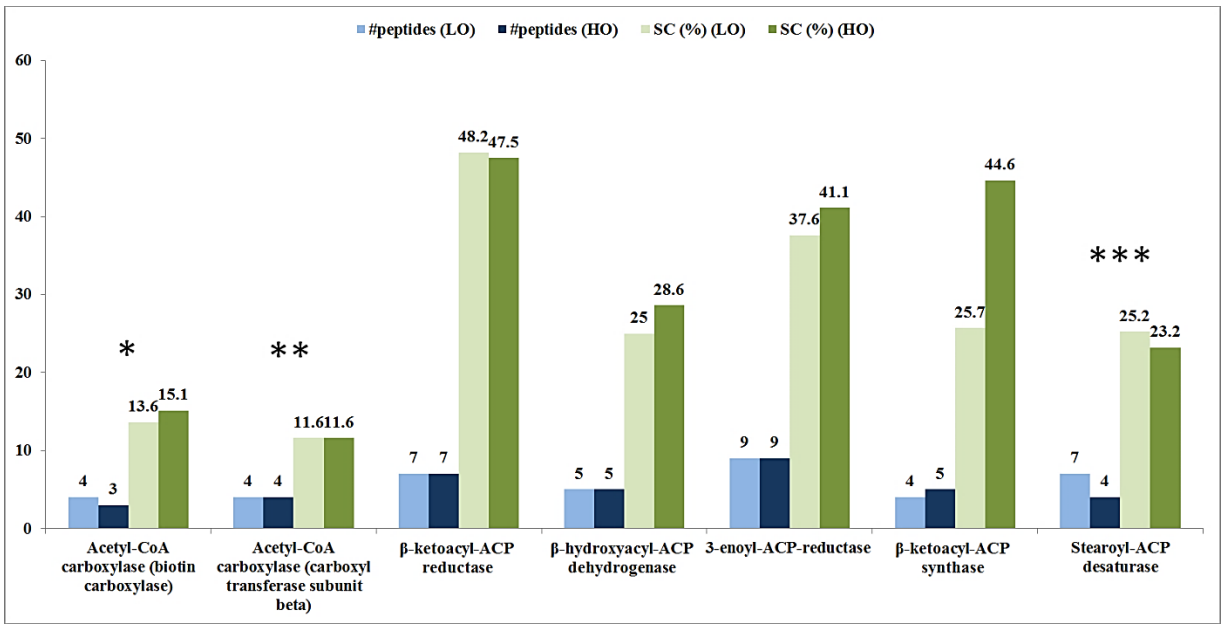


Figure 4.19. Number of unique peptides and sequence coverage (SC) per fatty acid biosynthetic enzyme obtained with an offline 2DLC-MS/MS. The results were compiled from three biological replicates for each low oleic acid (LO) and high oleic acid (HO) variety. * and ** denote the subunits of acetyl-CoA carboxylase. *** indicates the key enzyme in oleic acid biosynthesis.

Table 4.9. Unique peptides detected with 2DLC-MS/MS from low oleic acid (LO) and high oleic acid (HO) varieties that matched to the respective fatty acid biosynthetic enzymes. Peptides not detected corresponding to LO or HO are denoted with (x).

Enzyme	Accession	LO	HO	Unique peptide sequence
Acetyl-CoA carboxylase (biotin carboxylase)	D2CFM8			LLEEAPSPALTPELR
				LIVWAPTR
			x	ITSYLPSGGPFVR
		x		ALDDTIITGIPTTIEYHK
			x	HIEFQVLADK
Acetyl-CoA carboxylase (carboxyl transferase subunit beta)	H9LAY3			DIWSLISDDTFLVR
				SLPVIIVCASGGAR
				NVDYLFDIR
				SVLSELFQLHGFFPSNQNSK
β-ketoacyl-ACP reductase	G8FGI6			IAADAWGTVDILVNNAGITR
				VLINYATSTEEAEEVSK
				IINIASVVGLTGNAGQANYSASK
				IEQAQNVEAPVAIVTGGSR
				ILQTIPLGR
				NINVNAVAPGFIASPMTAQLGEDVEKK
				AIALALGK
β-hydroxyacyl-ACP dehydrogenase	G8FGF9			IPPFPTVMDINQIR
				VIEYQPGVTAVGIK
				FPFLLVDR
				TEEEVPIEK
				ENFFFAGIDK
3-enoyl-ACP reductase	H6TNL1			VYPLDAVYDTPEDVPDDVR
				VNTISAGPLGSR
				GVVSGLPIDLR
				NDFGSIDILVHSLANGPEVTKPLLETSR
				YAGASNWTVK
				AFVAGVADDNGYGWAIAK
				GYLAAISASSYSFVSLLK
				FTNVVTRALSGESEK
			x	GKFDESR
		x		GASQAKFASSSYISSVKPLR
β-ketoacyl-ACP synthase	G8FGI5			QLALSDDTPVTGESK
				VAKPETVQK
				VCDIVKKQLALSDDTPVTGESK
				GFSGLKSVSFSIQR
		x		LQAASR
Stearoyl-ACP desaturase	C325I9			DYADILEFLINR
				IPFSWIYGR
				TENSPYLGFIYTSFQER
			x	RRSGGSFVAVASMTSAAVSTR
			x	LGVYTARDYADILEFLVDR
			x	KPFMPPR
			x	LNDVALCLSPPLKAR
		x		SLEDWAENNILVHLKPVEK

Table 4.10. Unique peptides identified in low oleic acid (LO) and high oleic acid (HO) varieties corresponding to the respective fatty acid biosynthetic enzymes using GeLC- MS/MS and 2DLC-MS/MS. Peptides detected only with GeLC-MS/MS or 2DLC-MS/MS are indicated with the colourless box.

Enzyme	Accession	GeLC-MS/MS	2DLC-MS/MS	Unique peptide sequence
β-ketoacyl-ACP reductase	G8FGI6			AGVIGFTK
				EIEASGGQAIIIFFGDISKEDDVESMIK
				IEQAQNVEAPVAIVTGGSR
				AIALALGK
				IAADAWGTVDILVNNAGITR
				IINIASVVGLTGNAGQANYSASK
				ILQTIPLGR
				VLINYATSTEEAEEVSK
				NINVNAVAPGFIASPMTAQLGEDVEKK
β-hydroxyacyl-ACP dehydrogenase	G8FGF9			RIPPFPTVMDINQIR
				FRKPVIAGDTLVMR
				TEEEVPIEK
				ENFFFAGIDK
				FPFLLVDR
				IPPFPTVMDINQIR
				VIEYQPGVTAVGIK
3-enoyl-ACP reductase	H6TNL1			AALESDTMVLAFEAGR
				HFLPIMNPGGASISLTYIASER
				MIEYSYANAPLQK
				TIPGYGGGMSSAK
				GVVSGLPIDLR
				GYLAAISASSYSFVSLLK
				NDFGSIDILVHSLANGPEVTKPLLETSR
				VYPLDAVYDTPEDVPDDVR
				AFVAGVADDNGYGWAIAK
				GKFDESR
				RLPNGSLMEIIK
				VNTISAGPLGSR
				YAGASNWTVK
β-ketoacyl-ACP synthase	G8FGI5			QLALSDDTPVTGESK
				VAKPETVQK
				VCDIVKKQLALSDDTPVTGESK
				GFSGLKSVSFSIQR
				LQAASR
Stearoyl-ACP desaturase	C325I9			ATFISHGNTAR
				HGDLLNK
				TENSPYLGFIYTSFQER
				RRSGGSFVAVASMTSAAVSTR
				LGVYTARDYADILEFLVDR
				KPFMPPR
				LNDVALCLSPPLKAR
				SLEDWAENNILVHLKPVEK
				DYADILEFLINR
				IPFSWIYGR

biosynthesis-related enzymes. The most significant achievement using the non-gel based approach was that two main subunits of acetyl-CoA carboxylase (Figure 4.14) were detected in contrast to the GeLC-MS/MS approach. Commonly, a strong cation exchange column is used in series with a reverse-phase column (Nagele et al., 2004; Peng, Elias, Thoreen, Licklider, & Gygi, 2003; Slebos et al., 2008; Vollmer, Nagele, & Horth, 2003; Yun et al., 2011). This set up is termed as online 2DLC. In this study however, the fractionation of the two varieties' replicates was done 'offline' with a strong cation exchange column which was not in series with a reverse-phase column. The step gradients of ammonium formate in the first dimension (strong cation exchange column) without intermitting reverse-phase chromatography intervals (as in online 2DLC) enable better separation and peak capacity. The advantage of a superior separation of peptides with an independent strong cation exchange column is possibly the reason acetyl-CoA carboxylase was able to be detected with a 2DLC-MS/MS approach and not with the GeLC-MS/MS.

4.4.4 Isobaric tags to determine the expression of fatty acid biosynthetic enzymes

Changes in the protein expressions of fatty acid biosynthetic enzymes between low oleic and high oleic acid producing palms were quantitated using the isobaric labelling strategy. Three biological replicates from low oleic and high oleic varieties were individually labelled with 113, 114, 115, 116, 117 and 118-iTRAQ reagents and separated using offline 2DLC-MS/MS. To examine the dynamics of the proteome changes between low oleic and high oleic palms, the normalised protein ratios from all three replicates of the low oleic acid variety were compared to the normalised protein ratios from all three replicates of the high oleic acid variety (from two different iTRAQ experiments). In-house Mascot server results (Table 4.11) indicated that β-hydroxyacyl-ACP dehydrogenase and β-ketoacyl-ACP synthase were up-regulated significantly (ratios of 0.905 and 0.902, respectively) ($p < 0.05$) in the high oleic acid palm. 3-enoyl-ACP reductase exhibited a significant decrease (ratio of 1.088) ($p < 0.05$) in expression level in the high oleic acid palm while acetyl-CoA carboxylase was considered not to show any significant alteration in the expression level (ratio of 0.959) as determined by in-house Mascot server. The same unique peptide from both varieties is needed to make a relative estimation of protein abundance. Since this was not achieved for β-ketoacyl-ACP reductase and stearoyl-ACP desaturase, differences in their expression could not be determined with in-house Mascot server.

Table 4.11. The protein ratios for the fatty acid biosynthetic enzymes as determined from in-house Mascot server based on two iTRAQ experiments. Ratios that are significantly different from unity ($p < 0.05$) are bolded and indicated with *.

Enzyme	Score	# of peptides	Ratios (low oleic acid/high oleic acid)	Geometric standard deviation	Up/down regulated in high oleic acid variety
Acetyl-CoA carboxylase	157.0	4	0.959	1.144	Unchanged
β-hydroxyacyl-ACP dehydrogenase	81.0	6	**0.905***	1.033	Up
3-enoyl-ACP reductase	129.0	6	**1.088***	1.025	Down
β-ketoacyl-ACP synthase	240.0	11	**0.902***	1.127	Up

Two iTRAQ experiments identified a total of 114 proteins and 140 proteins, respectively, with at least 95% confidence using the in-house Mascot server. In addition to the fatty acid biosynthesis related enzymes, Table 4.12 shows the presence of other differentially expressed proteins between low oleic and high oleic varieties that were significant ($p < 0.05$). The results revealed that several metabolic enzymes; fructose-biphosphate aldolase, malate dehydrogenase and sucrose synthase were up-regulated in the high oleic acid variety. The expression levels of uncharacterised proteins involved in ATP binding were also increased as the oil palm fruit produced more oleic acid. The elevated expression of an uncharacterised protein which transfers acyl groups was also observed. Notably, lipase which is involved in lipid metabolism was expressed less in high oleic acid variety. From these results, it was clear that several proteins involved indirectly with fatty acid biosynthesis had been expressed differently between the low oleic acid and high oleic acid varieties.

Table 4.12. The protein ratios for other identified proteins as determined from in-house Mascot server based on two iTRAQ experiments. Proteins involved indirectly with fatty acid biosynthesis were bolded. All the listed proteins show significant differential expression ($p < 0.05$).

Enzyme	Accession	Score	# pep.	Ratios low oleic acid/ high oleic acid	Geometric standard deviation	Up/ down regulated in high oleic acid var.	GO (inferred from UniProtKB)
Lipase	**K4NZ15**	**372**	**18**	**1.29**	**1.138**	**Down**	**Lipid metabolism**
Fructose-bisphosphate aldolase	**B3TLY1**	**240**	**7**	**0.92**	**1.046**	**Up**	**Glycolytic process**
Malate dehydrogenase	**G8FGJ2**	**167**	**5**	**0.94**	**1.014**	**Up**	**Carbohydrate metabolic process**
Sucrose synthase	**K7V5Z8**	**56**	**5**	**1.16**	**1.051**	**Down**	**Sucrose synthase activity**
Ribulose 1,5-biphosphate carboxylase/oxygenase large subunit	**A9QBD0**	**107**	**5**	**1.11**	**1.085**	**Down**	**Magnesium ion binding (carbon fixation)**
Elongation factor 1-alpha	**B8APM5**	**108**	**8**	**1.05**	**1.024**	**Down**	**GTP binding**
Nucleoside diphosphate kinase 1	**P39207**	**56**	**4**	**1.14**	**1.056**	**Down**	**ATP binding**
Uncharacterised protein	**C0HHR4**	**226**	**6**	**0.94**	**1.050**	**Up**	**ATP binding**
Uncharacterised protein	**A2Z1N0**	**61**	**4**	**0.84**	**1.066**	**Up**	**ATP binding**
Uncharacterised protein	**C0P972**	**278**	**24**	**0.90**	**1.074**	**Up**	**Transferase activity, transferring acyl groups**
Histone H2B	M4C8F5	240	18	1.07	1.019	Down	Nucleosome assembly
Fibrillin-like protein	Q7XZF3	91	6	0.78	1.107	Up	Structural molecule activity
Disulfide isomerase	B3TM09	69	8	1.16	1.156	Down	Cell redox homeostasis
Uncharacterised protein	A3APM5	61	8	0.90	1.092	Up	Response to cadmium ion
Aldolase-like TIM barrel family protein	Q9LYR4	50	4	1.14	1.076	Down	Lignin biosynthetic process

4.5 Discussion

In higher plants, plastidial acetyl-CoA is used to generate malonyl-CoA for fatty acid biosynthesis. This reaction is catalysed by acetyl-CoA carboxylase (ACCase). In these studies, a 2DLC-MS/MS approach successfully detected and separated both the main subunits of the multisubunit heteromeric complex of ACCase (Figure 4.15). The biotin carboxylase carboxylates a biotin prosthetic group covalently bound to biotin carboxyl carrier protein (BCCP). The immediate transfer of a carboxyl group from the carboxylated biotin to acetyl-CoA to form malonyl-CoA (sole precursor for long-chain fatty acids) (Rawsthorne, 2002) is catalysed by carboxyltransferase (CT). Mass spectrometric analysis of iTRAQ labelled ACCase peptides showed no significant changes in its expression between low and high oleic acid varieties (Table 4.11). However, Roesler and co-workers demonstrated that transgenic plants with 10- to 20-fold increase in the enzymatic activity of ACCase resulted in the enrichment of oleic acid in the *Brasscia napus* seeds (13.7% increment in oleic acid content) (Roesler, Shintani, Savage, Boddupalli, & Ohlrogge, 1997). Increment in the expression of ACCase from 12th week after anthesis until fruit maturation was also observed by Loei and co-workers, although the expression was not compared between low-oil-yielding fruits and high-oil yielding fruits (Loei et al., 2013). Since the high oleic variety has 10.2% more oleic acid than the low oleic acid variety, the expression of ACCase was anticipated to be up-regulated but that was not the case in oil palm mesocarp. These results could possibly be explained by studies that reported the effect of enhanced ACCase activity varied in different plant organelles (seed and leaf) (Madoka et al., 2002). Another likely explanation was that the heteromeric form of ACCase is too labile to be isolated (Sasaki & Nagano, 2004a). Therefore, the number of peptides detected for both subunits of ACCase might not reflect the total peptides of the enzyme in mesocarp. Previous studies of fatty acid biosynthesis in plants also implicated the importance of ACCase for regulation. However, these studies were focussed on the overall production of lipids and not specifically on oleic acid (Chen, Mooney, et al., 2009; Klaus, Ohlrogge, Neuhaus, & Dormann, 2004; Page, Okada, & Harwood, 1994; Post-Beittenmiller, Roughan, & Ohlrogge, 1992; Thelen & Ohlrogge, 2002b).

The elongation of the carbon chain to 16- or 18-carbon (C16:0, C18:0) fatty acid is carried out by Type II fatty acid synthase (FAS) (Type I FAS present in animals and fungi) (Joyard et al., 2010). This easily dissociable multisubunit complex consists of β-ketoacyl-acyl carrier protein (ACP) reductase, β-hydroxyacyl-ACP dehydrogenase, 3-enoyl-ACP reductase and β-ketoacyl-

ACP synthase (KAS). Based on GeLC-MS/MS and 2DLC-MS/MS results (Figure 4.14 and Figure 4.19), these enzymes that involved in the elongation of fatty acid chain (except for β-ketoacyl-ACP synthase), were found to have more detected peptides in both varieties compared to acetyl-CoA carboxylase, β-ketoacyl-ACP synthase and stearoyl-ACP desaturase. This might suggest that those enzymes were present in higher abundance relative to acetyl-CoA carboxylase, β-ketoacyl-ACP synthase and stearoyl-ACP desaturase in fruit mesocarp during maturation. Quantitation results indicated that expression of β-hydroxyacyl-ACP dehydrogenase increased while 3-enoyl-ACP reductase showed a lower expression level in high oleic acid variety (Table 4.11). β-hydroxyacyl-ACP dehydrogenase produces 16-carbon palmitoyl-ACP (C16:0-ACP), which is the final product of fatty acid elongation. The relatively higher expression of β-hydroxyacyl-ACP dehydrogenase in high oleic acid variety could functionally increase the production of palmitoyl-ACP, which is the precursor for 18-carbon stearoyl-ACP. 3-enoyl-ACP reductase regenerates acyl-ACP for the next cycle of fatty acid biosynthesis. Unexpectedly, the expression of 3-enoyl-ACP reductase was lower in the high oleic acid variety. However, a number of the core enzymes of fatty acid biosynthesis that were genetically overexpressed or underexpressed in transgenic soybean or *Brassica napus* seeds did not show any significant increase or only a marginal increase in seed oil yield (Ohlrogge & Jaworski, 1997; Roesler et al., 1997). In a different study, the overexpression of spinach KAS III in tobacco had even decreased the fatty acid content (Dehesh, Tai, Edwards, Byrne, & Jaworski, 2001). The expression levels of β-hydroxyacyl-ACP dehydrogenase was also reported to increase throughout the developing stages in low-oil-yielding and high-oil-yielding fruits (Loei et al., 2013). As for β-ketoacyl-ACP reductase and 3-enoyl-ACP reductase, the expression level of β-ketoacyl-ACP reductase decreased in high-oil-yielding fruits while the expression level of 3-enoyl-ACP reductase did not change much. In general, the results above showed that β-ketoacyl-ACP reductase, β-hydroxyacyl-ACP dehydrogenase and 3-enoyl-ACP reductase were expressed more relative to acetyl-CoA carboxylase, KAS and stearoyl-ACP desaturase, independent of the level of oleic acid in the oil palm mesocarp.

The findings from the quantitation analysis revealed the significant increase in the expression of KAS (Table 4.11). KAS has three isoforms and they were not able to be distinguished mass spectrometrically (due to lack of studies on these isoforms). KAS III essentially initiates the condensation reaction between acetyl-CoA and malonyl-ACP and the subsequent condensations between these two building blocks are performed by KAS I. KAS II is implicated

in the final addition of 2-carbon to 16-carbon palmitoyl-ACP to produce the 18-carbon stearoyl-ACP (C18:0-ACP). Therefore, it was anticipated that expression of KAS II would be up-regulated to supply the stearoyl-ACP in high oleic acid variety, at the expense of palmitoyl-ACP. Since the results indicated that the expression of KAS was elevated in high oleic acid variety, the identified KAS in this study could well be KAS II. Increasing the expression of KAS II had raised the production of C18:1-ACP, as demonstrated in previous study (Sambanthamurthi et al., 1999). Loei and co-workers also described an increase in the expression of KAS in mature high-oil-yielding fruits but as observed in this study, their isoforms were not able to be differentiated (Loei et al., 2013).

Stearoyl-ACP desaturase (SAD) is the key fatty acid biosynthetic enzyme responsible for introduction of the double bond into the saturated stearoyl-ACP (C18:0-ACP) to produce mono-unsaturated oleoyl-ACP (C18:1-ACP). As SAD was not detected in the quantitative evaluation, the relative abundance of this enzyme between the two varieties could not be determined. However, results from both GeLC-MS/MS and 2DLC-MS/MS showed no substantial difference in the detected number of peptides between low oleic acid and high oleic acid varieties. This observation contrasted with the findings from previous studies as the expression of SAD was up-regulated in oleic acid accumulating tissues in earlier published work. Loei and co-workers described that the level of SAD in oil-accumulating fruit had been elevated although the result could not be validated (Loei et al., 2013). Plants defective in SAD were also reported to accumulate high levels of C18:0-ACP and low level of C18:1-ACP fatty acids (Kachroo et al., 2003; Kachroo et al., 2007). A potential explanation of the lack of up-regulation of SAD in the high oleic acid variety is a feedback inhibition mechanism by the end product; oleic acid. The accumulation of oleic acid could result in the inhibition or even down-regulation of SAD, which controls its synthesis. Purified monomeric ACCases from maize, diatom and *Brassica napus* were revealed to be inhibited by palmitoyl-CoA (Nikolau & Hawke, 1984; Roessler, 1990) and stearoyl-ACP (Andre, Haslam, & Shanklin, 2012). Decanoyl-ACP was also reported to inhibit KAS activity in canola, spinach and *Cuphea* (Bruck, Brummel, Schuch, & Spener, 1996; Schuch, Winter, Brück, Brummel, & Spener, 1997). The exact mechanism regulating these feedbacks in fatty acid biosynthesis is unknown (Ramli et al., 2002; Shintani & Ohlrogge, 1995). These feedback mechanisms could occur through biochemical or post-translational modification of fatty acid biosynthetic enzymes (Andre et al., 2012). The

relevance of feedback inhibition in this study is unclear because the changes in the pools of stearoyl-ACP and oleoyl-ACP were not measured.

The findings stated here underlined the prospect of other metabolic enzymes in regulating oleic acid level in the high oleic variety. In this study, several enzymes involved indirectly with fatty acid biosynthesis have been observed to be regulated differently between low oleic acid and high oleic acid palms (Table 4.12). Notably, the increased levels of fructose-biphosphate aldolase and malate dehydrogenase would increase the carbon flux towards fatty acid biosynthesis. Fructose-biphosphate aldolase generates triose phosphates, dihydroxyacetone phosphate and glyceraldehyde 3-phosphate to produce pyruvate via the glycolysis pathway. Meanwhile, malate dehydrogenase is involved in the production of glucose via gluconeogenesis through the oxidation of malate to oxaloacetate. The synthesis of glucose is also enhanced in the high oleic acid variety through the reduction in the expression of sucrose synthase, which produces sucrose at the expense of glucose. Hence, the increased supplies of glucose and pyruvate in the high oleic acid variety may elevate the supply of acetyl-CoA, the essential fatty acid biosynthesis precursor. A decrease in the expression of lipase, which catalyses the metabolism of oleic acid esters, could also raise the level of oleic acid in the high oleic acid variety, albeit indirectly. The results agreed with a study on sunflower seeds that reported the lowered oil content as the effect of lowered activities of glycolytic enzymes. Likewise, the level of seed oil in *Arabidopsis* was found to be affected by pyruvate kinase (Andre, Froehlich, Moll, & Benning, 2007; Troncoso-Ponce, Garces, & Martinez-Force, 2010).

The expression levels of elongation factor 1-alpha and ribulose 1,5-biphosphate carboxylase/oxygenase large subunit were also found to be differentially regulated. Fatty acid biosynthesis is an anabolic process that demands an abundant supply of ATP and NADPH. Therefore, the reduced expression of these proteins in high oleic acid palms might drive metabolic resources such as ATP and NADPH toward metabolic processes related to lipid biosynthesis, for example, glycolysis and fatty acid biosynthesis. In addition to the reduction of other metabolic processes that utilise ATP, the decrease in the expression of nucleoside diphosphate kinase 1 and elevated expressions of two uncharacterised proteins involved in ATP binding may have enhanced the supply of ATP for lipid biosynthesis. Loei and co-workers had found a reduction in the level of expression of proteins that compete for metabolic resources such as ATP and NADPH with the whole lipid biosynthesis process (Loei et al., 2013).

Another protein that was found to be up-regulated in the high oleic acid variety was an uncharacterised protein from *Zea mays*, which acts to transport acyl groups, probably during fatty acid biosynthesis. This elevated expression coincided with the increased fatty acid biosynthesis activity. As mentioned earlier, ACCase produced the malonyl-CoA. Before the malonyl-CoA can be used for fatty acid biosynthesis, the malonyl group needs to be transferred to ACP, catalysed by a malonyl-CoA:ACP malonyltransferase (MCMT). Up-regulating the expression of this enzyme in the high oleic acid variety could have increased the malonyl-ACP precursor pool for FAS as malonyl-CoA levels in plastids are normally very low (Thelen & Ohlrogge, 2002a). Therefore, an enhanced supply of malonyl-ACP precursor could result in more oleic acid being produced, even if the expressions of FAS and SAD remain unaltered. Since the identified protein was an uncharacterised protein and the fact that MCMT has not been sequenced in plant, this protein could not be determined whether it was the MCMT.

Several studies that manipulated the gene expression of oleoyl desaturase and acyl-ACP thioesterase had managed to increase the level of desirable fatty acids (Cahoon & Shanklin, 2000; Dehesh et al., 2001; Dormann, Voelker, & Ohlrogge, 2000; Kinney, 1996; Lardizabal et al., 2000; Stoutjesdijk, Hurlestone, Singh, & Green, 2000; Verwoert, van der Linden, Nijkamp, & Stuitje, 1994; Voelker, Jones, Cranmer, Davies, & Knutzon, 1997). Therefore it is worth looking at how the expressions of these fatty acid biosynthetic enzymes (in theory) might explain the differences in oleic acid content between these two varieties. Oleoyl desaturase introduces a second double bond in C18:1-ACP to form linoleic acid (C18:2-ACP). Thus, the down-regulation of this enzyme potentially increased the oleic acid content at the expense of linoleic acid. Kinney, Liu and co-workers showed that the suppression of oleoyl desaturase in soybean did increase the C18:1-ACP (Kinney, 1996; Liu, Singh, & Green, 2002b). Termination of FAS reactions is mediated by an acyl-ACP thioesterase which hydrolyse acyl-ACPs into free fatty acids and ACP (to be recycled). High activity of acyl-ACP thioesterase had been shown to increase levels of oleic acid in oil palm mesocarp (Sambanthamurthi et al., 1999). This thioesterase has higher specific activity towards C16:0 ACP than C18:1-ACP; and that could explain how the up-regulation of its expression could raise the levels of oleic acid. None of these four enzymes were able to be detected mass spectrometrically in this study. Furthermore, their regulation at the protein level has never been comprehensively investigated.

4.6 Conclusions

The study described the isolation and identification of six main fatty acid biosynthetic enzymes from oil palm fruit mesocarp through the application of gel- and non-gel based proteomic techniques. Comparative analysis between low oleic acid and high oleic acid oil palm varieties revealed that three of those fatty acid biosynthetic enzyme (based on the iTRAQ experiments), were differentially regulated (Figure 4.20). The findings here indicate that other metabolic proteins and lower abundant enzymes associated with fatty acid production could also be correlated with the different level of oleic acid in the oil palm mesocarp. This targeted study provides a clearer understanding on the overall fatty acid biosynthesis regulation in oil palm. Further validation of the enzymes showing differential regulations is required to corroborate their biological significance on oleic acid production in oil palm fruit.

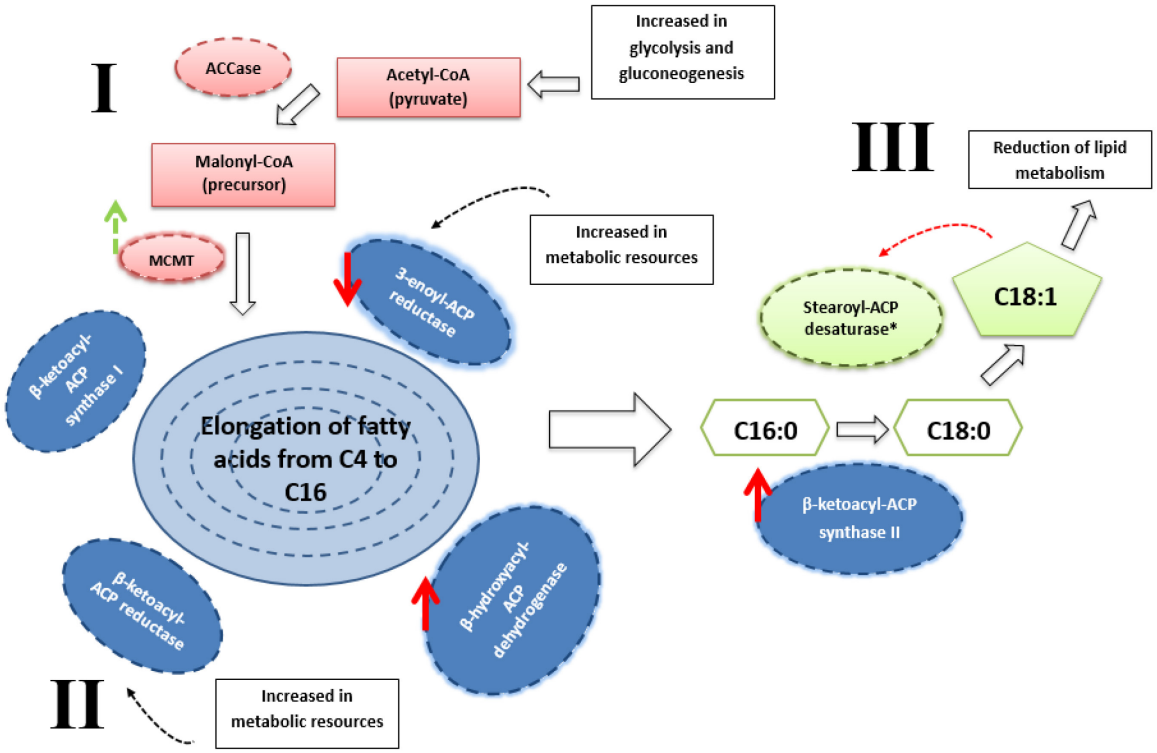


Figure 4.20. Overview of protein expression changes in fatty acid biosynthetic enzymes between low oleic acid and high oleic acid varieties. These changes (→) or plausible change (⇢) were characterised by iTRAQ experiments in high oleic acid variety. Circular boxes with dotted lines are the enzymes involved in various stages of fatty acid biosynthesis; **I**. Generation of malonyl-ACP; **II**. Seven condensation cycles to generate a 16-carbon fatty acid; **III**. Desaturation of stearic acid (C18:0-ACP) to form oleic acid (C18:1-ACP). Stearoyl-ACP desaturase (*) is the key enzyme in oleic acid production. C18:1-ACP was proposed as the signal molecule that inhibited the expression of SAD through feedback inhibition mechanism (⇢).

5.0 New insights into the possible role of phosphorylation in the regulation of oil palm fatty acid biosynthetic enzymes

5.1 Introduction

Regulation of proteins in a wide range of cellular processes is largely coordinated by a series of post-translational modifications especially reversible phosphorylation (Molina, Horn, Tang, Mathivanan, & Pandey, 2007; Nakagami, Sugiyama, Ishihama, & Shirasu, 2012). Protein phosphorylation is not directly dictated by the genome and thus, the nature of the phosphorylation events is unpredictable and transient. Three enzymes catalysing fatty acid biosynthesis in the oil palm (*Elaeis guineensis* Jacq. variant of tenera) fruit mesocarp were demonstrated to be differentially expressed between low and high oleic acid palms, as reported in Chapter 4. Therefore, the next key step was to understand the potential role of phosphorylation in the regulation of fatty acid biosynthetic enzymes. Despite a detailed understanding of many aspects of plant lipid metabolism and the genes coding for the key enzymes involved, how their activity is regulated is still far from clear (Cahoon, Shanklin, & Ohlrogge, 1992; Harwood, 1996; Kinney, 1994; Knutzon et al., 1992; Loei et al., 2013; Mekhedov, de Ilarduya, & Ohlrogge, 2000; Sambanthamurthi et al., 1999). Past research had reported that acetyl-CoA carboxylase, the enzyme that generates malonyl-CoA for subsequent fatty acid elongation processes, was deactivated by phosphorylation (Brownsey, Boone, Elliott, Kulpa, & Lee, 2006; Hardie, 1992; Hardie & Pan, 2002). However, the role of phosphorylation in controlling other enzymes associated with fatty acid biosynthesis is still ambiguous.

Comprehensive modification analysis or modificomics, as introduced by Jensen and his co-workers (Marx, 2013), can be performed with either a gel-based mass spectrometry approach or shotgun proteomic approaches (Du, Liang, Pei, & Ma, 2011; Fukao, Ferjani, Fujiwara, Nishimori, & Ohtsu, 2009; Tada & Kashimura, 2009). In a modification-specific gel stains approach, specific fluorescent stains such as Pro-Q Diamond used in this study, are applied directly to a polyacrylamide gel to detect the phosphate group attached to serine, threonine or tyrosine, and subsequently permit the fluorescent visualisation of the modified proteins (Di Domenico et al., 2011; Steinberg et al., 2003). The limitation is that a sufficient amount of

protein is the prerequisite to study the modifications of a single protein. At least several tens of picomoles or about a microgram of protein would be needed to increase the chance of detecting and characterising the modifications in proteins. This limitation means the utilisation of mass spectrometry-based approaches in studying protein modifications is becoming more common.

The phosphate group on a phosphopeptide is relatively labile and tends to break away in the form of phosphoric acid (HPO_3 or H_3PO_4) during collision-induced dissociation (CID). Thus, the fragmentation of phosphoamino residue-containing (serine, threonine and tyrosine) precursor ions generates neutral losses of 80 Da (HPO_3) or 98 Da (H_3PO_4) (Moon, Shin, & Kim, 2009). Usually, the mechanism for loss of H_3PO_4 from phosphoserine and phosphothreonine-containing peptide ions is the result of a β-elimination reaction (DeGnore & Qin, 1998). In a β-elimination reaction, the hydrogen atom on the α-carbon of the phosphorylated amino acid residue is transferred to the phosphate oxygen (Figure 5.1). As a result, dehydroalanine- (69 Da) or dehydroaminobutyric acid- (83 Da) containing product ions from phosphorylated serine or threonine residues, respectively, and H_3PO_4 are produced. Loss of H_3PO_4 is more dominant for serine phosphorylated peptides and in a lesser extent in threonine phosphorylated peptides. This might be caused by the steric hindrance of the β-methyl group in the side chain of threonine (Boersema, Mohammed, & Heck, 2009). Tyrosine phosphorylated peptides give a much lower extent of neutral loss and these are in the form of HPO_3.

Mass spectrometry-based strategies such as the neutral loss-triggered MS^3 ($NLMS^3$) and Selected Reaction Monitoring (SRM) are essentially built on the detection of the characteristic neutral loss generated during the CID fragmentation of phosphoamino-containing peptides. In the $NLMS^3$ mode of operation, the diagnostic neutral loss of H_3PO_4 (98 Da) from the precursor ion in a tandem MS scan automatically triggers the MS^3 fragmentation of the neutral loss precursor ion. The aim of the MS^3 is to compensate for the lack of sequence-specific information in the MS^2 spectra of phosphorylation-modified peptides (Boersema et al., 2009) although a study by Villen and co-workers indicated that the collection of MS^3 scans did not improve the informative spectra of the peptides identified (Villen, Beausoleil, & Gygi, 2008).

MS^3 operates in a data-dependent manner, in which the MS^3 is triggered by the presence of an intense product ion peak with the mass of a neutral loss (Figure 5.2). This strategy has been

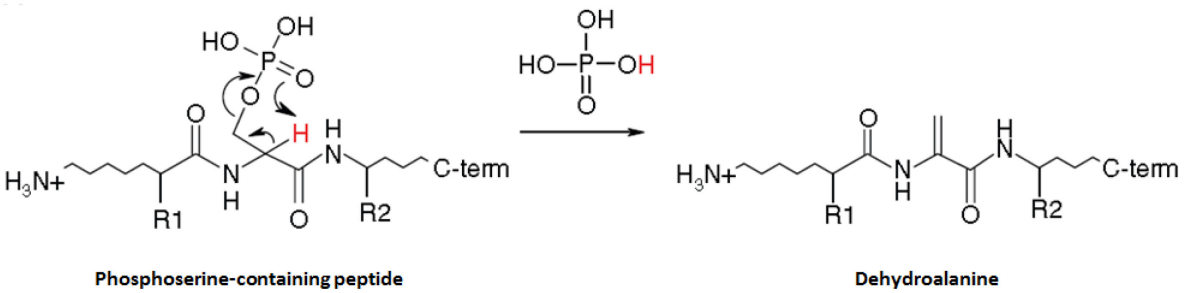


Figure 5.1. The loss of phosphoric acid (H3PO4) as a result of β-elimination reaction from the phosphoserine-containing peptide (Palumbo et al., 2011).

extended to detect phosphopeptides in this study using the Bruker amaZon ETD speed ion trap mass spectrometer. These neutral loss species from the product ions were calculated on the basis of the product ion mass and charge state, resulting in the neutral product ions of *m*/*z* 48.99 or *m*/*z* 32.66, relative to the doubly or triply charged phosphorylated product ion. A common problem with a neutral loss scan to detect phosphopeptides is that unassigned peptides may generate ions with a mass similar to the neutral losses as well (Villen et al., 2008; Yocum & Chinnaiyan, 2009). There are also instances when a phosphopeptide fails to generate the specified neutral loss and therefore are not detected (Cox et al., 2005).

SRM is an ideal complementary technique to reliably target and quantitate low abundance phosphopeptides of interest (Domanski, Murphy, & Borchers, 2010; Fan, Mohareb, Jones, & Bessant, 2012; Lange, Picotti, Domon, & Aebersold, 2008; Martinez-Marquez et al., 2013; Wolf-Yadlin, Hautaniemi, Lauffenburger, & White, 2007). SRM is predominantly performed on a triple quadrupole mass spectrometer as the availability of additional mass filter (third quadruple, Q3) is exploited to isolate targeted fragment ion for MS^2 (Figure 5.3). However, in this study, the Bruker Impact HD quadrupole-TOF was used to scan the targeted precursor ion for any loss of neutral loss species (98 Da) instead. Absence of the Q3 mass filter in a quadrupole-TOF implies that only neutral loss species at 98 Da correspond to the loss of H_3PO_4 can be detected after CID fragmentation of the selected precursor ions in q2. The ideal prerequisites to targeted SRM experiments are the prior knowledge of the primary sequence, type of phosphorylation, sequence motif and predicted fragmentation pathways to identify the potential phosphopeptides (for example, a neutral loss of 98 Da from phosphoserine and phosphothreonine peptides but not phosphotyrosine peptide).

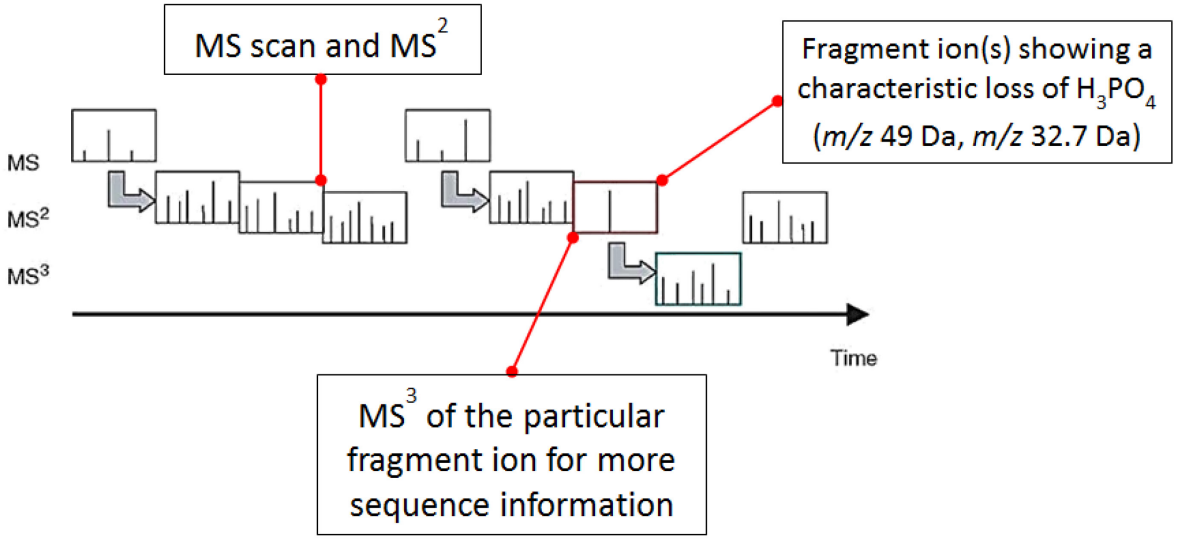


Figure 5.2. MS^3 of product ion with a neutral loss to yield more sequence-specific information.

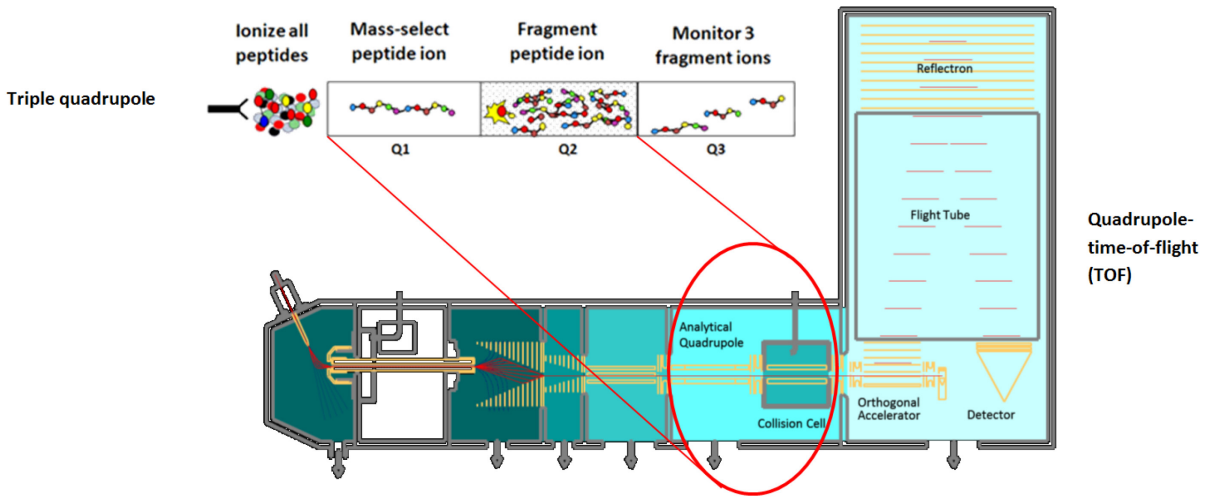


Figure 5.3. The equivalent of the Q1 and Q2 in a triple quadrupole and Bruker Impact HD quadrupole-TOF. Selected precursor ions are isolated in the analytical quadrupole mass filter or Q1 and CID fragmented in the collision cell or q2 to detect the neutral loss of 98 Da from phosphopeptides. The absence of another mass filter, Q3, means the selection of targeted product ions is not possible in a quadrupole-TOF.

Essentially, Chapter 5 describes the novel combination of $NLMS^3$ and SRM approaches to detect the presence of phosphopeptides associated with fatty acid biosynthesis in the oil palm. Until now, $NLMS^3$ has been mainly used to acquire more peptide sequence-specific information while most of the SRM-based applications have involved biomarker studies and quantitation phosphoproteomics. This study also compares the phosphorylation profiles of these fatty acid biosynthetic phosphopeptides from low oleic acid and high oleic acid-producing palms to determine if phosphorylation plays any role in the regulation of oleic acid levels in these palms. This chapter also presents plausible phosphorylation sites in the key fatty acid biosynthetic enzymes based on the identified phosphopeptides.

5.2 Materials

5.2.1 Plant materials

Independent fruit bunches (Figure 5.4) of the standard DxP oil palm crosses (*Elaeis guineensis* var. Tenera) (denoted as low oleic acid) and from the PORIM Series 12 (PS12) breeding populations (denoted as high oleic acid) were harvested from the Malaysian Palm Oil Board research stations at Bangi, Selangor and Hulu Paka, Terengganu, Malaysia. The fruit mesocarps were sliced, snap frozen in liquid nitrogen and stored at – 80°C until use. The seeds were discarded. Oil palm bunches (progenies) with different levels of oleic acids at 20th week after anthesis were collected from three low oleic acid and high oleic acid palms, respectively as listed in Table 5.1. Fruit mesocarps obtained from the bunch were divided into four technical replicates of 100 g each.

Standard DxP cross PS12 breeding population Sliced mesocarps

Figure 5.4. The two oil palm varieties used in the study.

Table 5.1. Biological replicates for both the oil palm varieties used in the study.

High oleic acid			Low oleic acid		
Palm no.	**Progeny**	**Oleic acid content (%)**	**Palm no.**	**Progeny**	**Oleic acid content (%)**
0.306/319	PK 540	48.85	0.306/79	PK 540	39.03
0.306/319	PK 540	48.85	0.306/76	PK 540	39.03
0.337/249	PK 1254	48.90	0.337/214	PK 1201	37.90

5.2.2 Chemicals

Most of the general chemicals were purchased from Sigma (St. Louis, MO, USA), BDH (BDH Laboratory Supplies, Poole, England) and Merck (Merck KGaA, Darmstadt, Germany). Organic solvents were acquired from Fisher (Fisher Scientific, Loughborough, UK). Specific chemicals

such as Protease Inhibitor Cocktail Tablet was from Roche (Roche Diagnostic, IN, USA), sequencing grade modified trypsin and mass spectrometry grade rLys-C from Promega (Promega Corporation, WI, USA), Pro-Q Diamond Phosphoprotein Gel Stain and PeppermintStick Phosphoprotein Molecular Weight Standards from Molecular Probes (Molecular Probes, Inc. OR, USA). Double-distilled water was used in preparation for all general solutions.

5.3 Methodology

5.3.1 Chromoplast isolation and subsequent protein extraction

Sliced mesocarps were homogenised in liquid nitrogen with a cold Waring blender (Blender B011, Dynamics Corporation, Greenwich, USA) at low speed for 10 s, followed by grinding with a standard ceramic mortar and pestle. The powdered mesocarps were then mixed with cold acetone containing 10% trichloroacetic acid and 1 mM dithiothreitol. The slurry was centrifuged at 13,000 g for 10 min at 4°C (RA-300 rotor, Kubota 7820, Kubota Corporation, Tokyo, Japan). The supernatant was discarded and the washing step was repeated once. Subsequently, cold 80% methanol containing 0.1 M ammonium acetate was added to the precipitate; mixed and centrifuged as before. After discarding the supernatant, the precipitated mesocarp pellet was washed by resuspending the pellet in cold 80% acetone. The mixture was mixed well and centrifuged again at 13,000 g for 10 min at 4°C. The resulting pellet was air-dried.

Washed mesocarp pellet was then transferred into a beaker containing cell wall digestive enzymes (2% w/v cellulase, 0.1% w/v pectinase, 0.6 M sorbitol, 0.1 M DTT, 5 mM 2-(4-morpholino)-ethane sulfonic acid (MES)-KOH, pH 5.5). The suspension was then incubated in an incubator shaker (Labnet 211DS, Labnet Instrument, Inc., NJ, USA) for 6 h at 37°C. After cell wall digestion, the mixture was sieved through two layers of Miracloth (Calbiochem, EMB Millipore Corporation, Billerica, MA) into a fresh beaker on ice to separate non-macerated plant materials from protoplasts. After that, the filtrate was centrifuged at 1750 g for 5 min at 4°C (RA-300 rotor, Kubota 7820) to collect intact chromoplasts. The chromoplast pellet was gently re-suspended in the extraction buffer (0.7 M sucrose, 1 M Tris-HCl, pH 8.3, 5 M NaCl, 50 mM DTT, 1 mM EDTA containing a tablet of Roche protease inhibitors in every 50 mL of buffer). The mixture was agitated for 5 min and then sonicated for 15 min at 4°C to enhance the release of the chromoplast proteins. An equal volume of fresh 50 mM, pH 8.0 Tris-

saturated phenol was added to the mixture thereafter. The mixture was further agitated for 10 min just before centrifugation at 15,000 g for 15 min at 4°C (RA-300 rotor, Kubota 7820) for phase separation. Following that, the upper phase was transferred to a new centrifuge tube while the bottom phase was discarded. Chromoplast proteins were precipitated by adding five volumes of cold ammonium acetate-saturated methanol and incubated at -20°C overnight before being centrifuged at 15,000 g for 15 min at 4°C (RA-300 rotor, Kubota 7820) to obtain the protein pellet. The protein pellets were then rinsed with cold ammonium acetate-saturated methanol until a whitish colour was obtained for the pellet. This was followed by washing the pellet with cold 80% acetone for three times. Proteins were precipitated after each wash by centrifuging at 15,000 g for 5 min at 4°C (RA-300 rotor, Kubota 7820). At the end of the washing phase, the supernatant was carefully decanted before the chromoplast protein pellet was air-dried. Commercially available 2D Quant Kit (GE Healthcare Life Sciences, Uppsala, Sweden) was utilised to determine protein content in the samples. Bovine serum albumin provided with the kit was used as the protein calibration standard and each of the quantitation was performed in duplicates.

5.3.2 Gel electrophoresis and staining

Chromoplast proteins were dissolved in Laemlli buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 25% glycerol, 0.01% bromophenol blue, 0.005% β-mercaptoethanol) and denatured by heating at 95°C for 4 min. 100 µg protein was loaded into each lane on a 1.0 mm precast Mini-PROTEAN® TGX™ 8-16% polyacrylamide gel (Bio-Rad Laboratories Inc., Hercules, CA). Electrophoresis was conducted in a Bio-Rad mini-PROTEAN® Tetra Cell apparatus (Bio-Rad Laboratories) at 200 V for 1 h. Following electrophoresis, the separated proteins were fixed for 30 min in a fixing solution (50% ethanol, 10% acetic acid) and stained with Pro-Q Diamond phosphoprotein stain according to the manufacturer's instructions. Stained gel was visualised and imaged using a Typhoon FLA 7000 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) with excitation/emission wavelength of 532/580 nm. The gel was subsequently stained with an in-house prepared Colloidal Coomassie G-250 to obtain the total protein expression profile. The gel was destained with Milli-Q water (H_2O) until the background was clear before acquiring the gel image with Typhoon FLA 9500 (GE Healthcare Bio-Sciences).

5.3.3 In-solution protein digestion and fractionation

200 µg of chromoplast proteins were re-suspended in 0.1 M ammonium bicarbonate and 1 M urea before reduction and alkylation using 50 mM tris(2-carboxyethyl)phosphine and 150 mM

iodoacetamide, respectively. An ionic detergent, sodium deoxycholate (1% w/v) was added to the protein solution prior to digestion with 4 µg of modified sequencing grade trypsin (Promega, Madison, WI, USA) in 50 mM NH_4HCO_3 for 16 h at 37°C. After digestion, the digests solution was acidified with 0.5% formic acid. Sodium deoxycholate precipitate was removed from the peptide digests by centrifugation at 14 000 g (RA-300, Kubota 7820) for 15 min at ambient temperature. The peptide solution was then dried in a centrifugal evaporator (CentriVap Concentrator, Labconco, MO, USA). *Peptides clean-up* – The dried peptide pellet was resuspended in 200 µL of 0.1% formic acid. Acetonitrile, methanol and 0.1% formic acid-conditioned Empore solid phase extraction disks (3M Purification, Inc., MN, USA) were added to the peptide solution and incubated at ambient temperature with slight agitation for 4 h. The bound peptides on the C18 membrane disks were sequentially eluted with 50% ACN in 0.1% FA for 2.5 h. *SCX fractionation* – The peptide eluent was dried with a centrifugal evaporator (CentriVap Concentrator) and resuspended in 250 µL of 0.1% FA. The resuspended peptides were then fractionated using a strong cation exchange cartridge (Phenomenex SCX cartridge, 4 x 2 mm). Elution was done with 0.06 M, 0.12 M, 0.2 M, 0.3 M and 0.4 M ammonium formate buffer containing 2% ACN, pH 3.5. Each of the collected fractions was then dried with a centrifugal evaporator (CentriVap Concentrator).

5.3.4 Liquid chromatography-tandem mass spectrometry

Neutral loss-triggered MS^3 – Separation and spectra acquisition of the protein digests was conducted with a nano-Advance Splitless nano-liquid chromatography (nanoLC) (Bruker Daltonik GmbH, Bremen, Germany) system coupled to an amaZon speed ETD ion trap mass spectrometer (Bruker Daltonik). The tryptic digests were reconstituted in 30 µL of 0.1% FA and 5% ACN and 10 µL was injected into the nanoLC system for peptide separation with a ProntoSIL C18AQ (3µ, 100 µ x 150 mm) (nanoLCMS Solutions, LLC., CA, USA) reverse-phase column equilibrated with 95% solvent A (2% ACN in 0.1% FA) and 5% solvent B (98% ACN in 0.1% FA). Gradient of 0-45% solvent B was employed to elute the bound peptides during 45 min at a flow rate of 800 nL min^{-1}. The neutral loss feature in the trapControl software (Version 7.1) was activated and the loss of *m/z* 49, *m/z* 32.7, *m/z* 58 and *m/z* 38.7 was included as the neutral loss masses. The eluted peptides were electrosprayed into the ion trap mass spectrometer using a spray voltage of 1400 V and a capillary temperature of 150°C. Precursor survey scan was acquired with a mass range from *m/z* 310-1400, the resolution was set to 'Enhanced Resolution' and the scanning speed was 8,100 u per sec. Tandem MS conditions

consisted of 'Enhanced Resolution' scan ranging from *m/z* 100-3000. Up to five of the most intense multiple charged ions (1+, 2+, and 3+) per scan were fragmented via collision-induced dissociation in the linear ion trap. The neutral loss scans were searched from the tandem MS derived from the second scan, or MS^2. All tandem mass spectra were collected using 1.00 V of fragmentation amplitude, an isolation window of 4.0 u and scanning speed of 52,000 u per sec. *Selected Reaction Monitoring* – Separation and spectra acquisition of the protein digests was conducted with a nano-Advance Splitless nano-liquid chromatography (nanoLC) (Bruker Daltonik) system coupled to an Impact HD quadrupole-time-of-flight mass spectrometer (Bruker Daltonik). The mass spectrometer was externally calibrated using Electrospray Tuning Mix Positive (Agilent Technologies, Inc., Santa Clara, CA, USA). The tryptic digests were reconstituted in 30 µL of 0.1% FA and 5% ACN and 5 µL was injected for peptide separation with a ProntoSIL C18AQ (3µ, 100 µ x 150 mm) (nanoLCMS Solutions) reverse-phase column equilibrated with 95% solvent A (2% ACN, 0.1% FA) and 5% solvent B (98% ACN, 0.1% FA). A gradient of 2-65% solvent B in 60 min at a flow rate of 800 nL min^{-1} was used. The MRM feature in the otofControl software was activated and the precursor isolation width and collision energy (eV) for the unique peptides was determined (Table 5.2). These unique peptides were identified by database searches using Mascot to the fatty acid biosynthetic enzymes in a different Q-TOF experiment. The peptides were electrosprayed with a spray voltage of 1300 V and a capillary temperature of 150°C. SRM was done in a data-dependent manner. Precursor survey scan was acquired with a mass range of from *m/z* 350-2900. MS/MS settings consisted of scan ranged from *m/z* 50-2200. Up to three of the most intense multiple charged ions (1+, 2+, and 3+) per scan were fragmented via collision-induced dissociation.

5.3.5 Data acquisition and analysis

Data acquisitions in positive mode were performed with ESI Compass 1.4 for amaZon (trapControl Version 7.1) and ESI Compass for otofSeries 1.7 for Impact HD (tofControl Version 3.4) (Bruker Daltonik). Analyses of the data were performed using the Compass DataAnalysis (Version 4.1 SR1) (Bruker Daltonik). Generated peak lists in eXtensible Markup Language (XML) format were then sent to ProteinScape (Version 3.1) (Bruker Daltonik) for protein identification using Mascot software (Matrix Science, London, UK). The peptide sequences were searched against *Arabidopsis thaliana*, *Elaeis* sp*., Zea mays, Brassica and Oryza* plant taxonomies (308,973 sequences as on 24th December 2013) in the Uniprot protein database and the protein list was compiled using the Protein Extractor module

Table 5.2. Parameters for Selected Monitoring Reaction experiments for the fatty acid biosynthetic associated phosphopeptides. Each selected precursor ion is unique to the target protein and is between seven to 24 amino acids in length.

Enzyme	Precursor (measured MH+)	Precursor sequence	Precursor isolation width (Da)	Collision energy (eV)
Acetyl-CoA carboxylase (biotin carboxylase subunit)	914.5678	KTLVSAPLK	± 4	20
	939.6208	LVLSTSEK	± 4	20
	1004.5964	KLVLSTSEK	± 4	30
	1393.8551	ITSYLPSGGPFVR	± 5	30
	1635.9405	LLEEAPSPALTPELR	± 5	30
	1641.7513	SEAAAAFGNDGVYLEK	± 5	30
Acetyl-CoA carboxylase (β-carboxyl transferase subunit)	1399.8551	SLPVIIVCASGGAR	± 5	30
	1662.9378	TVPDGSQAAEYLFHK	± 5	30
	1679.9249	DIWSLISDDTFLVR	± 5	30
β-ketoacyl-ACP reductase	1882.9150	VLINYATSTEEAEEVSK	± 5	40
	2248.3432	IINIASVVGLTGNAGQANYSASK	± 7.5	40
β-hydroxyacyl-ACP dehydrogenase	1759.8880	SVDSAVKTEEEVPIEK	± 5	30
	1910.0020	CSVDSAVKTEEEVPIEK	± 5	25
	1926.8520	CSVDSAVKTEEEVPIEK	± 5	30
3-enoyl-ACP reductase	838.4068	GKFDESR	± 4	30
	958.5292	TISAGPLGSR	± 4	35
	1171.6420	VNTISAGPLGSR	± 5	30
	1125.6630	GVVSGLPIDLR	± 5	30
	1282.6146	TIPGYGGGMSSAK	± 5	20
	1737.9233	AALESDTMVLAFEAGR	± 5	30
β-ketoacyl-ACP synthase	1546.759	QLALSDDTPVTGESK	± 5	30
	1560.7770	QLALSDDTPVTGESK	± 5	30
	1572.8090	QLALSDDTPVTGESK	± 5	30
	1688.8760	KQLALSDDTPVTGESK	± 5	30
Stearoyl-ACP desaturase	1160.618	ISMASTVGPSTK	± 5	20
	1507.8675	TIQYLIGSGMDPR	± 5	20
	2492.1890	RSGGSFVAVASMTSAAVSTRVENK	± 7.5	30

in ProteinScape. Mass tolerances for peptide and product ions were set to 100 ppm and 0.6 Da (ion trap) or 20 ppm and 0.1 Da (quadrupole-TOF), respectively; the instrument setting was specified as 'ESI trap' or 'ESI-QUAD-TOF'. Semi-trypsin was designated as the protease with two missing cleavages allowed. Carbamidomethylation on cysteine was set as the fixed modification while oxidation of methionine and deamidation of asparagine and glutamine were searched as variable modifications. In error tolerant searches, phosphorylation of serine (S), threonine (T) and tyrosine (Y) amino acids were included as variable modifications.

Proteins were accepted if they had at least one top ranking peptide (Rank 1) with a Mascot ion score of more than 40.0 ($p < 0.05$). All database searches were also performed against the decoy database to determine the false discovery rate. A decoy database contained randomised sequences of the searched taxonomies. Individual identified protein lists were compiled using the Protein Extractor module in ProteinScape. All phosphopeptides presented in this work were validated either through neutral loss-triggered MS^3 or Selected Reaction Monitoring approaches.

5.3.6 Bioinformatic tools

Theoretical digestion with trypsin on each of the protein sequences was performed with PeptideCutter available online at ExPASy website (http://web.expasy.org/peptide_cutter/). Trypsin was specified as the digestion enzyme and one missed cleavage was allowed. Phosphorylation sites within the protein sequences were predicted computationally using KinasePhos 2.0 and has 95% specificity (http://kinasephos2.mbc.nctu.edu.tw/) (Wong et al., 2007).

5.4 Results

5.4.1 Total phosphoprotein expression profile

The use of Pro-Q diamond staining allowed the detection of phosphorylated proteins separated on a 8-16% gradient SDS-PAGE gel. This enabled the investigation of phosphoproteome changes between two different palm varieties, namely, the low oleic acid and high oleic acid-producing palms as well as between biological replicates of each variety (Figure 5.5). Comparison between low oleic acid replicates (LO1-LO3) revealed the dissimilarity in their phosphorylation states (Figure 5.5A) although their total protein (Figure 5.5B) showed similar expressions. Phosphoproteins of high oleic acid replicates were more consistent with regard to phosphorylation and total protein expressions (Figure 5.5). The signal intensity of the phosphostained proteins also indicated that there appeared to be more phosphorylated residues in the high oleic variety compared to the low oleic acid variety (Figure 5.5A) since signal intensity correlates with the total number of phosphorylated residues in the peptide. Given that each of the protein bands consisted of more than one protein, a direct comparison of the phosphorylation state between fatty acid biosynthetic enzymes (positions as indicated in Figure 5.5) in low oleic acid and high oleic acid varieties was not possible using this approach. The protein expression levels for the total proteins from both varieties seemed to show greater consistency than the phosphorylation state (Figure 5.5B). Table 5.3 shows the phosphorylation states for each of the biological replicates. As presented, the phosphoproteome of low oleic acid replicates varied more while the variation was slightly less in high oleic acid replicates.

In the previous chapter (Chapter 4), the protein expression of some of the fatty acid biosynthetic enzymes was revealed to be different. Those differential expressions were not evident in this work, based on the gel-based approach (Figure 5.5B). Therefore, the phosphorylation states of the overall phosphoproteins from both varieties could not be used as the direct indication of the phosphorylation states for these low abundant fatty acid biosynthetic enzymes. Differences in the overall phosphorylation state for each of the biological replicates as displayed in Table 5.3 may have resulted from unanticipated variations in harvesting time, sampling location or the fruit post-harvesting process.

Amongst the low oleic acid replicates, fewer protein bands containing phosphorylated proteins were observed in LO3 compared to LO1 and LO2. The sampling areas (East and

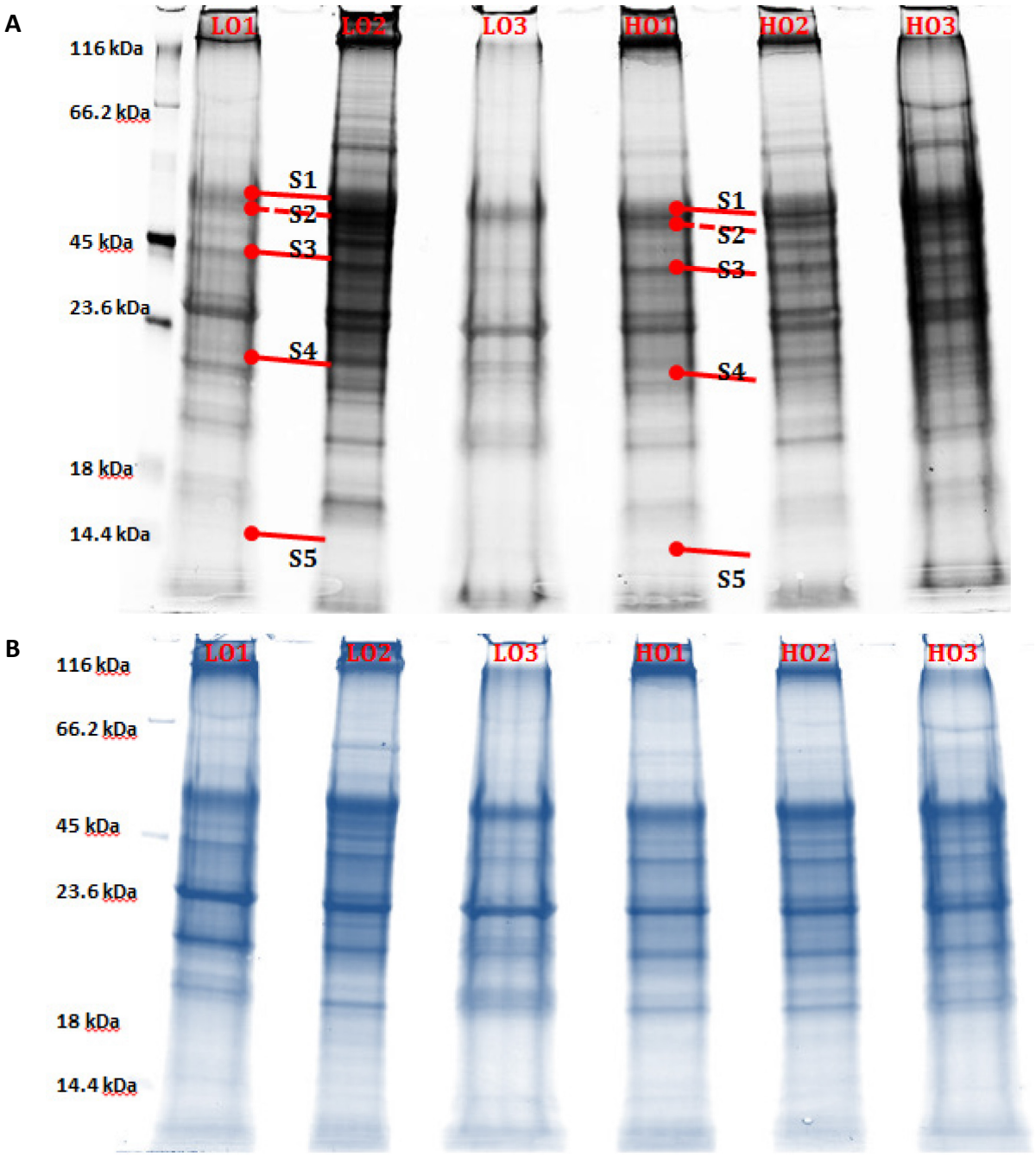
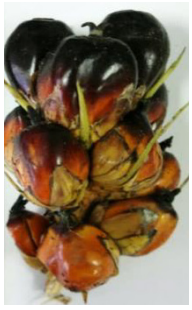


Figure 5.5. Screening of total phosphoproteins (A) and protein (B) expression patterns from low oleic acid (LO) and high oleic acid (HO) varieties separated on 8-16% gradient SDS-PAGE. Location of fatty acid biosynthetic enzymes as determined from previous works is indicated; S1: Acetyl-CoA carboxylase subunits (theoretical); S2. stearoyl-ACP desaturase; S3. 3-enoyl-ACP reductase; S4. β-ketoacyl-ACP reductase and β-hydroxyacyl-ACP dehydrogenase; S5. β-ketoacyl-ACP synthase.

Table 5.3. Phosphorylation state corresponding to the different biological replicates from low oleic acid (LO) and high oleic acid (HO) varieties.

Spikelet	Fruitlet	Description	Phosphorylation state
		**LO1** 39.03% oleic acid East Peninsular	
		**LO2** 39.03% oleic acid East Peninsular	
		**LO3** 37.90% oleic acid West Peninsular	
		**HO1** 48.85% oleic acid East Peninsular	
		**HO2** 48.85% oleic acid East Peninsular	
		**HO3** 48.90% oleic acid West Peninsular	

West Peninsular) and the content of oleic acid (different of 1.13% between LO3 and LO1/LO2) could possibly affect the phosphorylation pattern. In addition, the harvesting time of the fruit used in this study varied based on evidence of the epicarp colour between the replicates. This indicated a different fruit ripening time. The phosphorylation pattern of each of the high oleic acid replicates (Table 5.3) showed better similarity. Although all replicates had the same oleic acid content, the harvesting time (based on epicarp colour) for HO3 was different from HO1 and HO2. That probably explained the slightly higher signal intensities for several protein bands in HO3. Fruit ripening is controlled by the hormone ethylene. Different phosphorylation patterns of isoforms of 1-aminocyclopropane-1-carboxylic acid (ACC) synthase and ACC oxidase had been shown during the regulation of ethylene production in tomato (Alexander & Grierson, 2002). The phosphorylation pattern of different phosphoproteins was also found to be altered during different ripening stages in both tomato (Kamiyoshihara, Tieman, Huber, & Klee, 2012; Vescovi & Viale, 1993) and sweet orange (Zeng et al., 2011b).

5.4.2 Phosphopeptides prediction

Phosphorylation sites on serine (S), threonine (T) and tyrosine (Y) residues were predicted using the web-based KinasePhos 2.0. The aim was to acquire masses for phosphopeptides containing the predicted phosphorylated S, T or Y residues for the neutral loss-triggered MS^3 and SRM analyses. Table 5.4 summarises the number of predicted phosphosites and the probable number of phosphoserine-containing peptides obtained from *in silico* digest of the target enzymes (based on available protein sequences). The phosphosite predictions were performed on full amino acid sequences for the two subunits of acetyl-CoA carboxylase and fragment amino acid sequences for the other fatty acid biosynthetic enzymes. The prediction indicated that the subunit β-carboxyl transferase of acetyl-CoA carboxylase contains the most possible phosphorylation sites on serine residues. Meanwhile, β-ketoacyl-ACP reductase and β-hydroxyacyl-ACP dehydrogenase were predicted to have the least possible phosphorylation sites on serine residues. The subunit β-carboxyl transferase of acetyl-CoA carboxylase was predicted to generate 25 peptides containing the predicted phosphosites although the enzyme had 49 possible phosphosites on serine residues. This implied that the enzyme had more peptides containing multi-phosphosites compared to the other enzymes. None of the threonine or tyrosine residues was predicted by KinasePhos 2.0 to be phosphorylated.

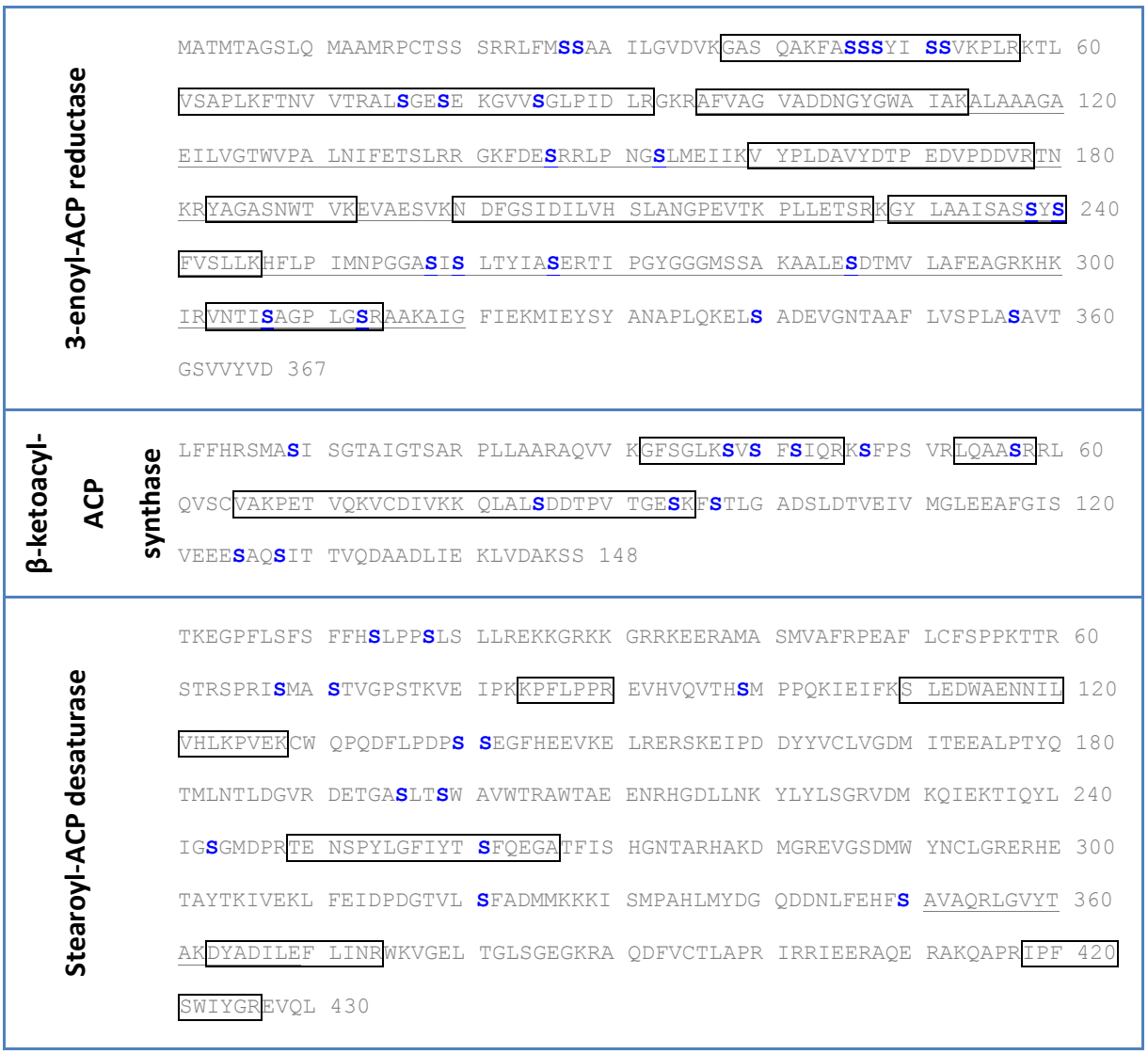
Table 5.4. Number of predicted phosphosites on serine residues and phosphoserine-containing peptides for fatty acid biosynthetic enzymes using KinasePhos 2.0. All entries are from *Elaeis guineensis* var. Tenera, unless specified.

Enzyme	Unitprot accession no.	NCBI accession no.	AA length	Sequence status	Predicted phosphosites	Predicted no. of phosphopeptides
Acetyl-CoA carboxylase BC	D2CFM8	209168853	536	Complete	**18**	**15**
Acetyl-CoA carboxylase β-CT	H9LAY3	383931199	492	Complete	**49**	**25**
β-ketoacyl-ACP reductase	G8FGI6	353441094	255	Partial	**7**	**5**
β-hydroxyacyl-ACP dehydrogenase	G8FGF9	353441060	220	Partial	**5**	**3**
3-enoyl-ACP reductase	H6TNL1	374255957	367	Partial	**22**	**19**
β-ketoacyl-ACP synthase	G8FGI5	353441032	148	Partial	**11**	**9**
Stearoyl-ACP desaturase	C3W5I9 (*E. oleifera*)	228478496	430	Partial	**13**	**19**

In addition to the predicted phosphosites, Table 5.5 also indicates the location of protein domains sourced from Pfam and the identified peptides associated with fatty acid biosynthesis acquired using both gel-based and non-gel based approaches carried out in previous chapter (Chapter 4). The acetyl-CoA carboxylase subunits, biotin carboxylase and β-carboxyl transferase, have 14 and 18 predicted phosphosites, respectively in their protein domains. However, only two peptides containing the predicted phosphosites were identified for subunit biotin carboxylase and one peptide for the subunit carboxyl transferase. As for β-ketoacyl-ACP reductase, seven predicted phosphosites are located in the protein domain and two of the detected peptides from previous work contained these phosphosites. 3-enoyl-ACP reductase has 10 phosphosites in the protein domain and two phosphosites are situated in the protein domain of β-ketoacyl-ACP synthase. Two peptides identified to these enzymes contained the predicted phosphosites. β-hydroxyacyl-ACP dehydrogenase has one possible phosphorylation site in the protein domain but the protein domain for stearoyl-ACP desaturase has no possible phosphorylation sites as predicted with KinasePhos 2.0. None of the detected peptides for these two enzymes contained the serine residues.

Table 5.5. Amino acid sequences from the different fatty acid biosynthetic enzymes. The predicted phosphorylated serine residue (s), the protein domains (underlined) and the identified peptides corresponding to the respective fatty acid biosynthetic enzymes from previous work (boxed) are indicated.

Enzyme	Sequence
Acetyl-CoA carboxylase (subunit biotin carboxylase)	MDSMIACKSG CLPPGLFIRP TRGIRSSQCS FMVGNGPKFC KISFPRKELL VRSGNPKKNG 60 GALNATCCDE KILVANRGEI AVRVIRAAHE MGIPCVAVHS TIDQDALHVR LADEAVCIGE 120 APSSQSYLFI PNVLSAAVSR GCTMLHPGYG FLAENAGFVD ICKEHGINFI GPNPDSIRVM 180 GDKSTARETM KKAGVPTVPG SDGLLQSTEE AVKLAHEVGF PVMIKATAGG GGRGMRLAYE 240 PEQFVKLLQQ AKSEAAAAFG NDGVYLEKYI QNPRHIEFQV LADKYGNVVH FGERDCSIQR 300 RNQKLLEEAP SPALTPELRK AMGDAAVAAA ASIGYIGGWN CGNFLLDERG SFYFMEMNTR 360 IQVEHPVTEM ISSTDLIEEQ IRVALGERLT YKQEDIVLRG HSIECRINAE DAFKGFRPGP 420 GKITSYLPSG GPFVRMDSHV YPGYVVPPSY DSLLGKLIVW APTREKAIER MKRALDDTII 480 TGIPTTIEYH KLILDIEDFR NGKVDTAFIP KHEKDLTAPH KLVLSTSEKE LAGVGA 536
Acetyl-CoA carboxylase (subunit β-carboxyl transferase)	MEKWWFNSML SNDKLEHRCG LNKSMDSLDA IGHTSGSEEP ILNGTEKNIP SRSDSGSYSF 60 RNVDYLFDIR DIWSLISDDT FLVRDSNGDS FSVCFDIENQ IFEIDNDSSF LSELESFFSS 120 YLNNGSKRNN HNYYHYIYDT QSSWNNHINS CIDNYLRFEV SINSSISGGT DNYSDSYIYS 180 FICTENVTGS ESGSSGIRTS KNGSDFNIRR RSNDFGRKKK YRHLWIQCEN CYGLNYKKFF 240 RSKMNICEQC GYHLKMSSSD RIELSIDPGT WDPMDEDMVS MDPIEFHSEG KPYRDRIDSY 300 QRKTGLTEAV QTGIGQLNGI PIAIGVMDFK FMGGSMGSVV GEKITRLIEY ATNRSLPVII 360 VCASGGARMQ EGSLSLMQMA KISSASHNYQ SNKKLFYVSI LTSPTTGGVT ASFGMLGDVI 420 VAEPNAYIAF AGKRVIEQTL NKTVPDGSQA AEYLFHKGLF DPIVPRNPLK SVLSELFQLH 480 GFFPSNQNSK NL 492
β-ketoacyl-ACP reductase	QASIEQAQNV EAPVAIVTGG SRGIGKAIAL ALGKAGCKVL INYATSTEEA EEVSKEIEAS 60 GGQAIIFFGD ISKEDDVESM IKIAADAWGT VDILVNNAGI TRDTLLMRMK KSQWQEVIDV 120 NLTGVFLCTQ AAAKLMMKKK KGKIINIASV VGLTGNAGQA NYSASKAGVI GFTKTVAREY 180 ASRNINVNAV APGFIASPMT AQLGEDVEKK ILQTIPLGRY GQPEEVAGLV EFLALNPAAD 240 YITGQVFTID GGMVM 255
β-hydroxyacyl-ACP dehydrogenase	SAPPLHSNPF LSCHGSNTDL LAFPSPKSQR RNALPPVLSL ANPKSPAWAL GLERRRSFLT 60 RCSVDSAVKT EEEVPIEKRI PPFPTVMDIN QIRDILPHRF PFLLVDRVIE YQPGVTAVGI 120 KNVTINDNFF PGHFPERPIM PGVLMVEAMA QVGGLVMLQP EVGGSRENFF FAGIDKVRFR 180 KPVIAGDTLV MRMTLIKLQK RFGIAKMDGK AYVGGDLVCE 220


3-enoyl-ACP reductase
MATMTAGSLQ MAAMRPCTSS SRRLFMSSAA ILGVDVKGAS QAKFASSSYI SSVKPLRKTL 60
VSAPLKFTNV VTRALSGESE KGVVSGLPID LRGKRAFVAG VADDNGYGWA IAKALAAAGA 120
EILVGTWVPA LNIFETSLRR GKFDESRRLP NGSLMEIIKV YPLDAVYDTP EDVPDDVRTN 180
KRYAGASNWT VKEVAESVKN DFGSIDILVH SLANGPEVTK PLLETSRKGY LAAISASSYS 240
FVSLLKHFLP IMNPGGASIS LTYIASERTI PGYGGGMSSA KAALESDTMV LAFEAGRKHK 300
IRVNTISAGP LGSRAAKAIG FIEKMIEYSY ANAPLQKELS ADEVGNTAAF LVSPLASAVT 360
GSVVYVD 367
β-ketoacyl-ACP synthase
LFFHRSMASI SGTAIGTSAR PLLAARAQVV KGFSGLKSVS FSIQRKSFPS VRLQAASRRL 60
QVSCVAKPET VQKVCDIVKK QLALSDDTPV TGESKFSTLG ADSLDTVEIV MGLEEAFGIS 120
VEEESAQSIT TVQDAADLIE KLVDAKSS 148
Stearoyl-ACP desaturase
TKEGPFLSFS FFHSLPPSLS LLREKKGRKK GRRKEERAMA SMVAFRPEAF LCFSPPKTTR 60
STRSPRISMA STVGPSTKVE IPKKPFLPPR EVHVQVTHSM PPQKIEIFKS LEDWAENNIL 120
VHLKPVEKCW QPQDFLPDPS SEGFHEEVKE LRERSKEIPD DYYVCLVGDM ITEEALPTYQ 180
TMLNTLDGVR DETGASLTSW AVWTRAWTAE ENRHGDLLNK YLYLSGRVDM KQIEKTIQYL 240
IGSGMDPRTE NSPYLGFIYT SFQEGATFIS HGNTARHAKD MGREVGSDMW YNCLGRERHE 300
TAYTKIVEKL FEIDPDGTVL SFADMMKKKI SMPAHLMYDG QDDNLFEHFS AVAQRLGVYT 360
AKDYADILEF LINRWKVGEL TGLSGEGKRA QDFVCTLAPR IRRIEERAQE RAKQAPRIPF 420
SWIYGREVQL 430

5.4.3 Neutral loss-triggered MS^3 ($NLMS^3$)

Tandem mass spectrometric analysis of phosphorylated peptides under collision-induced dissociation tends to generate neutral losses of *m/z* 48.99 or *m/z* 32.66, relative to the doubly or triply charged product ion, respectively. In addition, the loss of a water molecule with the phosphoric acid ($H_2O+H_3PO_4$) from the serine residue-containing peptide ion is also common. Hence, the Data Dependent experiment was created to also selectively trigger MS^3 scans on the MS/MS product ions for which *m/z* 57.99 and/or *m/z* 38.66 losses were detected. For the product ions that showed the neutral losses, it was then determined if they were related to the fatty acid biosynthetic enzymes, based on the precursor ion masses of their phosphopeptides. As only serine residues were predicted to be phosphorylated (by KinasePhos 2.0), this research focussed only on neutral losses from serine residues.

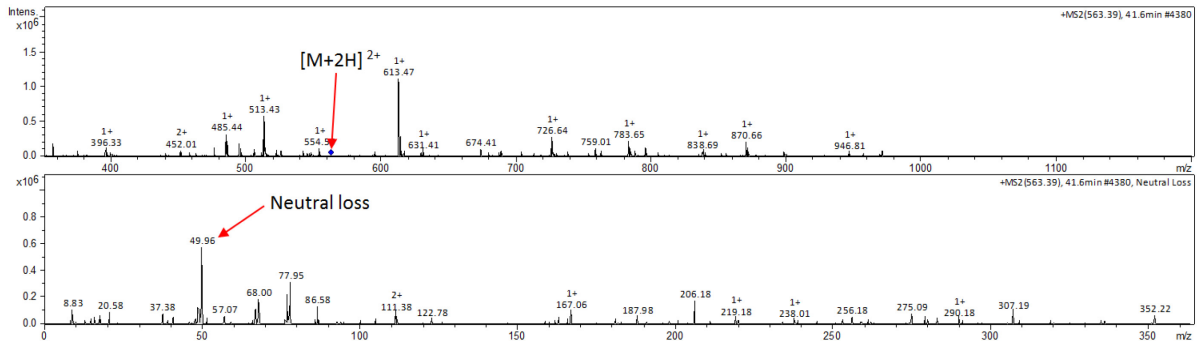
Table 5.6 summarises the number of detected phosphoserine-containing peptides and the peptide product ions showing the neutral losses from low and high oleic acid varieties. The results, based on three different $NLMS^3$ experiments using three biological replicates for each variety (in each experiment), revealed that only a single peptide was found to give a neutral loss of *m/z* 98 after fragmentation. Fragmentation of the peptide with mass of *m/z* 563.4 in doubly charged state produced a neutral loss of *m/z* 49 (Figure 5.6). This phosphopeptide corresponded to peptide GVVVSGLPIDLR based on their precursor ion masses. Although numerous phosphoserine-containing peptides were predicted (Table 5.4) for the other fatty acid biosynthetic enzymes, none of these phosphopeptides were detected and selected for $NLMS^3$. The subunit β-carboxyl transferase of acetyl-CoA carboxylase and β-hydroxyacyl-ACP dehydrogenase did not have any phosphopeptide with precursor ion masses matched to these two enzymes.

Peptide with the mass of *m/z* 563.4 from both varieties exhibited the same neutral losses (Figure 5.6). This suggested that the 3-enoyl-ACP reductase might be phosphorylated at 20^{th} week after anthesis since peptide GVVSGLPIDLR is one of its unique peptides. Since there was only one serine residue in peptide GVVSGLPIDLR, the phosphoric acid group would most probably derive from a phosphorylated serine residue. Further phosphosite mapping is required to validate the prediction. This phosphosite was predicted by KinasePhos 2.0 to be GVVpSGLPIDLR.

Table 5.6. Number of phosphoserine-containing peptides predicted and product ion displaying a neutral loss of phosphoric acid from NLMS^3 analyses. Three different batches of biological replicates (in three different NLMS^3 experiments) from low oleic acid/high oleic acid varieties (LO/HO) were used. Prediction of phosphosites and phosphopeptides were explained in Section 5.4.2.

Enzyme	Predicted phosphorylated sites	No. of phosphoserine -containing peptides	Detected unique peptides (LO/HO)			Product ion with neutral loss of *m/z* 98
			Exp. 1	Exp. 2	Exp. 3	
Acetyl-CoA carboxylase (biotin carboxylase)	18	15	-/-	5/3	-/2	-/-
Acetyl-CoA carboxylase (β-carboxyl transferase)	49	25	-/-	-/-	-/-	-/-
β-ketoacyl-ACP reductase	7	5	1/1	1/1	1/1	-/-
β-hydroxyacyl-ACP dehydrogenase	5	3	-/-	-/-	-/-	-/-
3-enoyl-ACP reductase	22	19	2/2	4/2	2/3	1/1
β-ketoacyl-ACP synthase	11	9	-/1	1/1	-/-	-/-
Stearoyl-ACP desaturase	13	19	-/-	-/1	-/1	-/-

A. Low oleic acid variety



B. High oleic acid variety

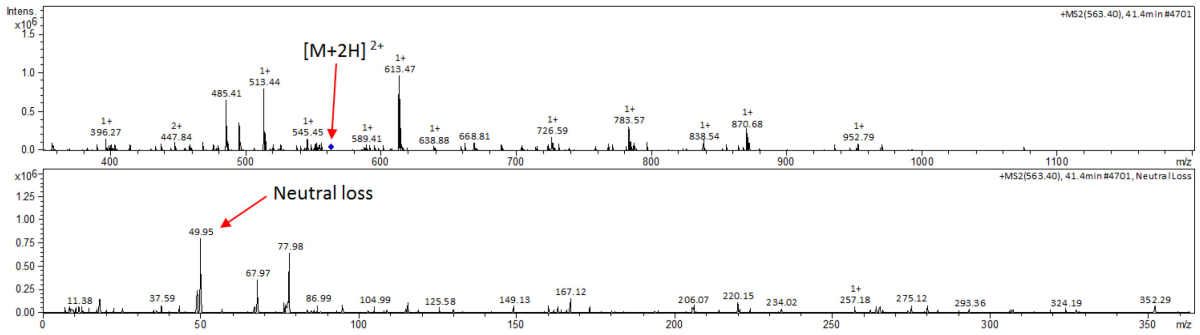


Figure 5.6. Neutral loss detected from the product ion with *m/z* 563.4 $[M+H]^{2+}$ using $NLMS^3$ approach. The neutral loss of *m/z* 49 (in the neutral loss spectra) was calculated based on the product ion spectrum.

It was noticeable that the number of detected peptides varied between experiments for subunit biotin carboxylase of acetyl-CoA carboxylase and 3-enoyl-ACP reductase. Reasons for these variations could be resulting from the dynamic range of protein concentrations in different experiments (Yocum & Chinnaiyan, 2009), the ionic suppression (Jessome & Volmer, 2006) by competing low and high abundant peptides to capture protons and the inconsistency of phosphorylated peptide ionisation.

5.4.4 Selected reaction monitoring (SRM)

In SRM method development, targeted peptides (Table 5.2) were pre-selected for fragmentation via collision-induced dissociation in a data-dependent manner rather than a MS^3 scan. This approach was used to determine if these predicted phosphoserine-containing peptides from different enzymes involved in the fatty acid biosynthesis are phosphorylated through the loss of the phosphoric acid group attached to the side chain of a serine residue. At least two unique peptides for each of the targeted proteins were selected for SRM experiments (Table 5.2). Three ionic forms were detected for peptide QLLALSDDTPVTGESK of β-ketoacyl-ACP reductase. Two different ionic forms for peptide CSVDSAVKTEEEVPIEK were detected for β-hydroxyacyl-ACP dehydrogenase. Although the presence of ionic forms (resulted from different charged peptides) was not a desirable attribute for these targeted peptides (James & Jorgensen, 2010), all the precursors were included for SRM experiments in order to enhance the detection of the targeted peptides since they are present in low abundances. These selected peptides were also consistently detected over the two different oil palm varieties investigated in this study.

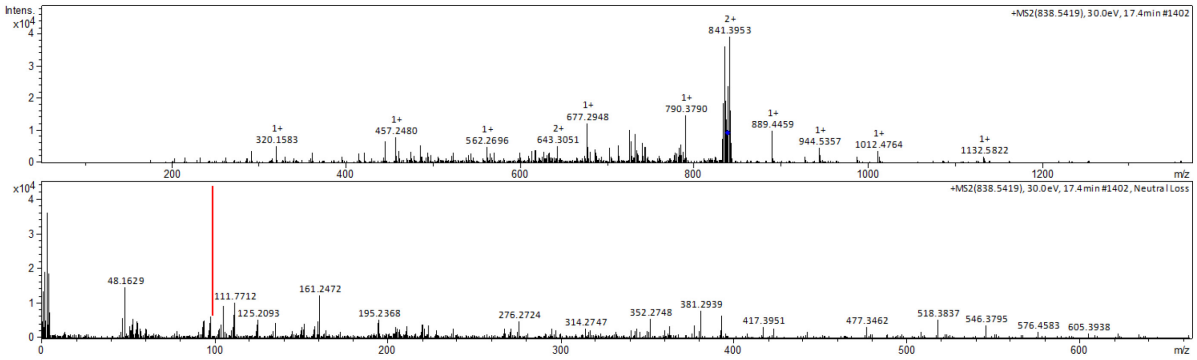
Table 5.7 summarises the number of detected precursor ions that showed a phosphoric acid (98 Da) loss from different SRM experiments for both low and high oleic acid varieties. SRM experiments had revealed that two peptides with masses of *m/z* 1004.59 and *m/z* 1393.74, produced neutral losses of 98 Da. These phosphopeptides corresponded to peptides KLVLSTSEK and ITSYLPSGGPFVR from the subunit biotin carboxylase of acetyl-CoA carboxylase, based on their precursor ion masses. Fragmentation of peptides with masses of *m/z* 838.47, *m/z* 1125.66 and *m/z* 1171.64 also yielded neutral losses of 98 Da. Their masses matched to the precursor ion masses for peptides GKFDESR, GVVSGLPIDL and VNTISAGPLGSR from 3-enoyl-ACP reductase. The same peptides from both low oleic acid and high oleic acid varieties exhibited these neutral losses.

Meanwhile, Figure 5.7 and Figure 5.8 depict the MS/MS spectra for the peptides with neutral loss of 98 Da from both varieties. As shown, the product ion peak for the neutral loss was not a dominant peak as the selected peptide ion scanning was performed in a positive ion mode, which was not tailored to detect the loss of neutral loss specifically. The abundance of the neutral losses between low oleic acid and high oleic acid varieties were not comparable as there was no internal standardisation in this study. Figure 5.9 shows the selected MS/MS spectra for several serine containing peptides from other fatty acid biosynthetic enzymes that did not produce any neutral loss after collision-induced fragmentation. The low abundance of several selected peptides also hampered the attempt to determine if these serine containing peptides were actually phosphorylated through the loss of phosphoric acid group. Signals for these peptides were probably reduced by ionic suppression during fragmentation (Jessome & Volmer, 2006).

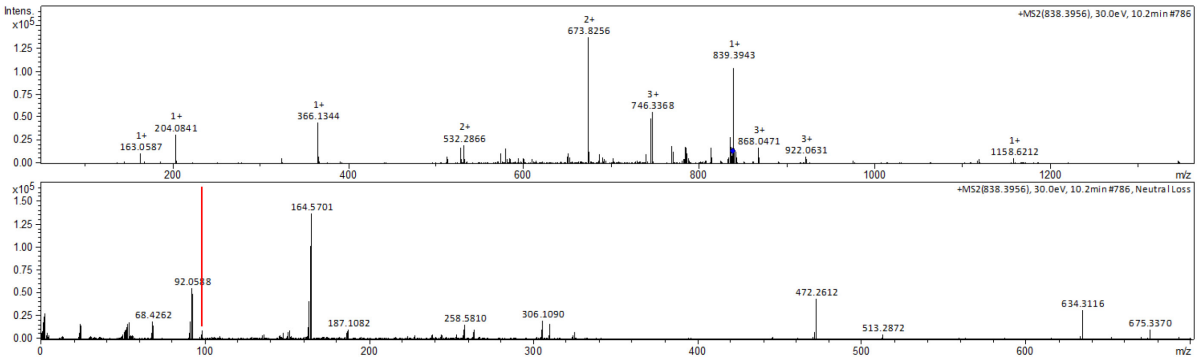
Table 5.7. Number of phosphoamino-containing peptides predicted and product ion with the neutral loss of phosphoric acid (98 Da) obtained from SRM analyses. Three different batches of biological replicates from low oleic acid/high oleic acid varieties (LO/HO) were used. Prediction of phosphosites and phosphopeptides is explained in Section 5.4.2.

Enzyme	Predicted phosphorylated sites	No. of phosphoserine-containing peptides	Detected unique peptides (LO/HO)			Product ion with neutral loss of 98 Da
			Exp. 1	Exp. 2	Exp. 3	
Acetyl-CoA carboxylase (biotin carboxylase)	18	15	2/2	4/3	3/4	2/2
Acetyl-CoA carboxylase (β-carboxyl transferase)	49	25	1/1	3/2	3/1	-/-
β-ketoacyl-ACP reductase	7	5	-/2	2/1	2/1	-/-
β-hydroxyacyl-ACP dehydrogenase	5	3	1/1	-/-	-/-	-/-
3-enoyl-ACP reductase	22	19	2/4	4/4	4/4	3/3
β-ketoacyl-ACP synthase	11	9	1/1	3/1	1/1	-/-
Stearoyl-ACP desaturase	13	19	-/1	1/1	1/2	-/-

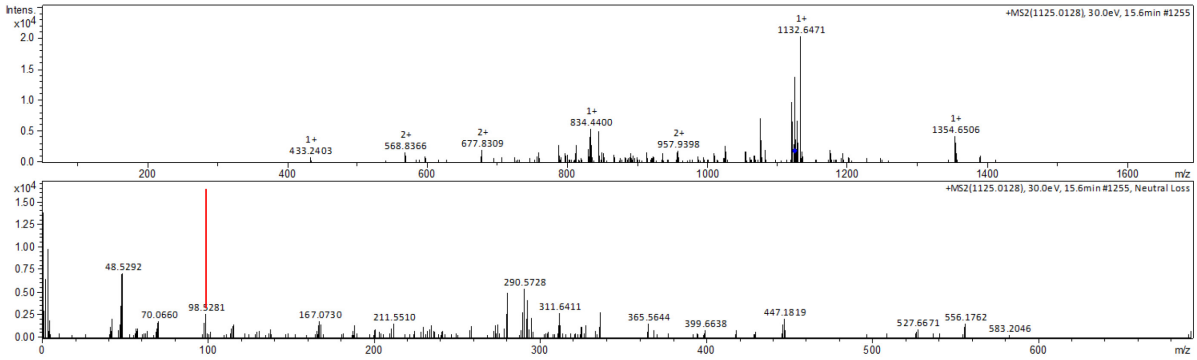
m/z 838.47 – GKFDEpSR in low oleic acid variety



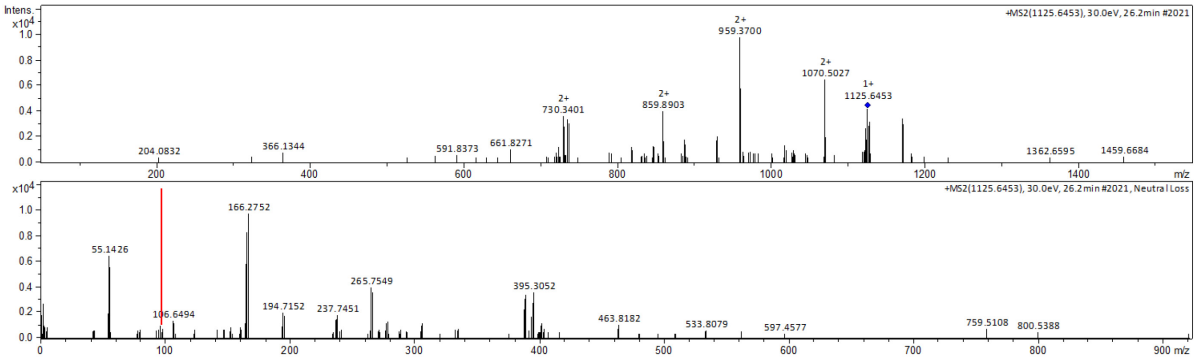
m/z 838.47 – GKFDEpSR in high oleic acid variety



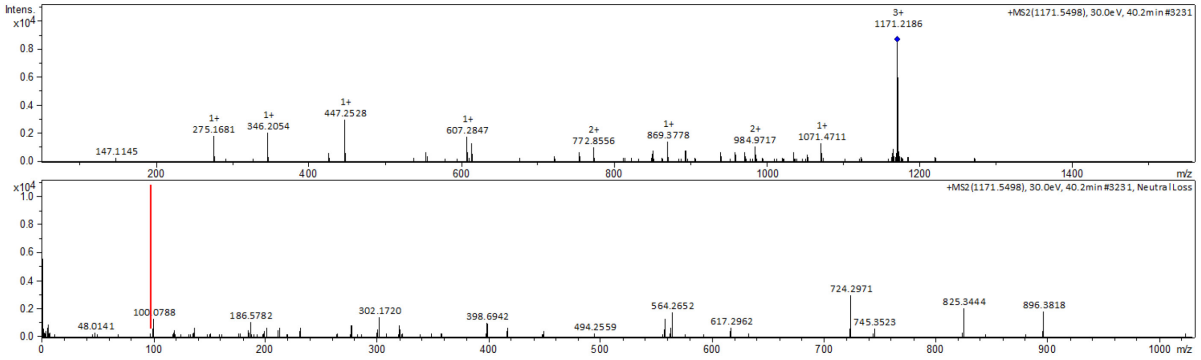
m/z 1125.6626 – GVVpSGLPIDLR in low oleic acid variety



m/z 1125.6626 – GVVpSGLPIDLR in high oleic acid variety



m/z 1171.642 – VNTIpSAGPLGpSR in low oleic acid variety



m/z 1171.642 – VNTIpSAGPLGpSR in high oleic acid variety

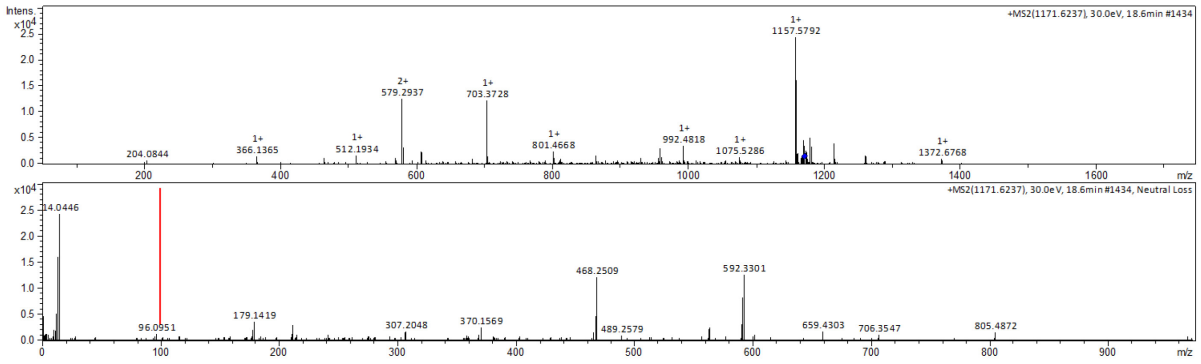
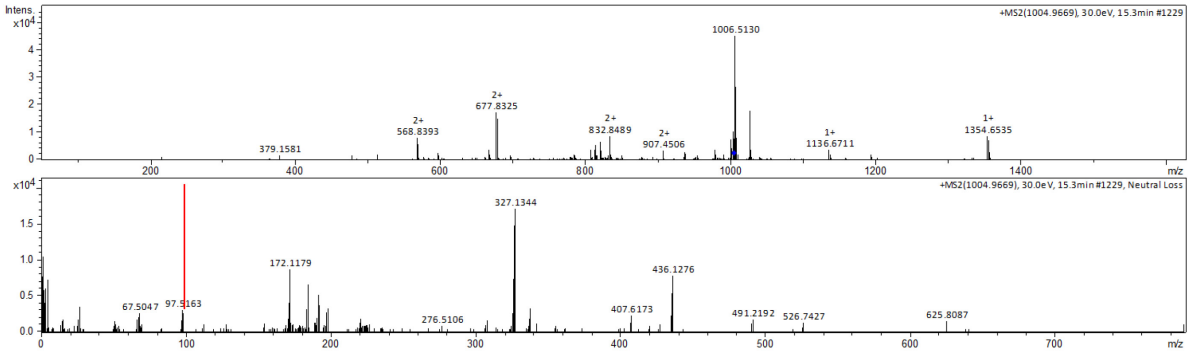
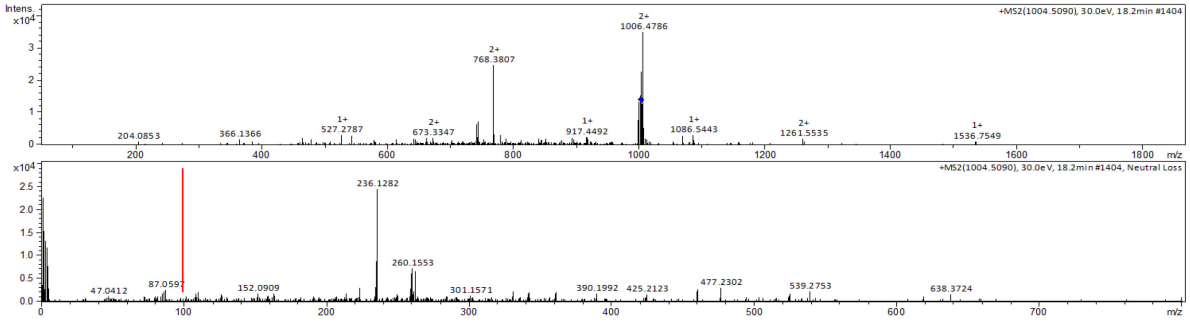


Figure 5.7. Precursor ions (indicated with a blue diamond) with a neutral loss detected (indicated with a red line), using SRM approach. The neutral loss of 98 Da (in the neutral loss spectra) was calculated based on the precursor ion spectrum. Predicted phosphorylated serine residues with KinasePhos 2.0 are indicated in red.

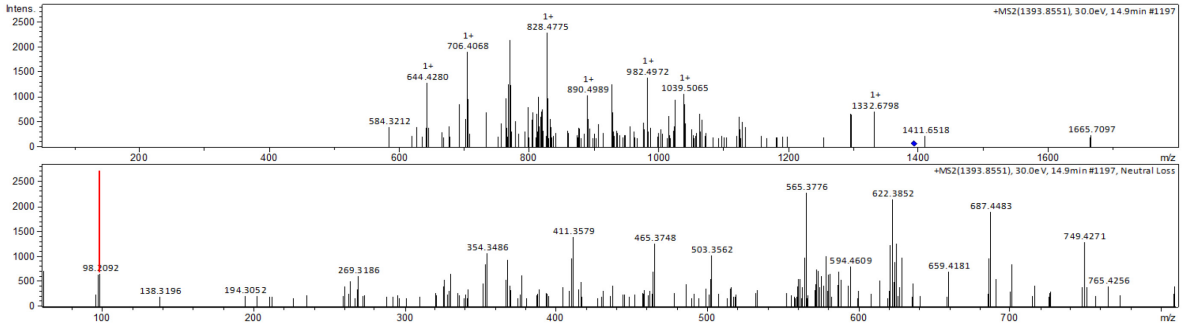
m/z 1004.5964 – KLVLSTpSEK in low oleic acid variety



m/z 1004.5964 – KLVLSTpSEK in high oleic acid variety



m/z 1393.745 – ITpSYLPpSGGPFVR in low oleic acid variety



m/z 1393.745 – ITpSYLPpSGGPFVR in high oleic acid variety

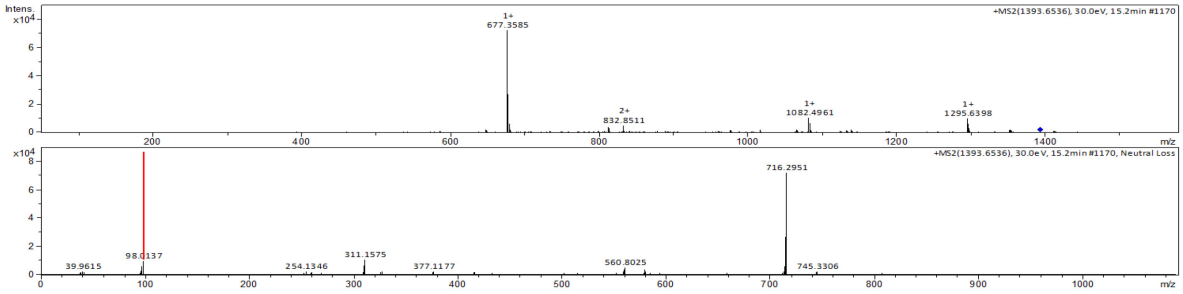
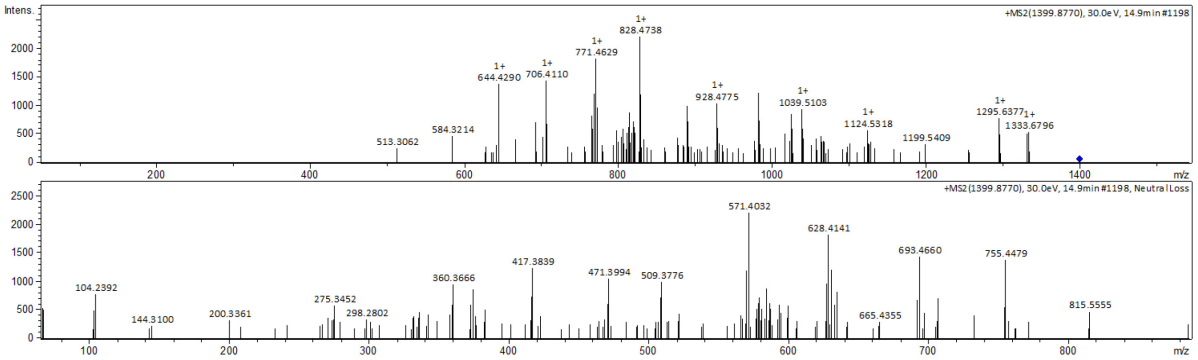
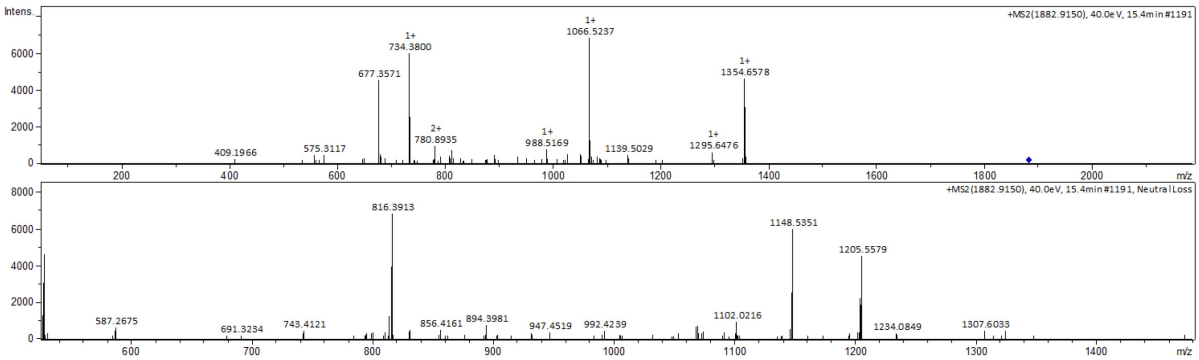


Figure 5.8. Precursor ions (indicated with a blue diamond) with a neutral loss (indicated with a red line), detected using SRM approach. The neutral loss of 98 Da (in the neutral loss spectra) was calculated based on the precursor ion spectrum. Predicted phosphorylated serine residues with KinasePhos 2.0 are indicated in red.

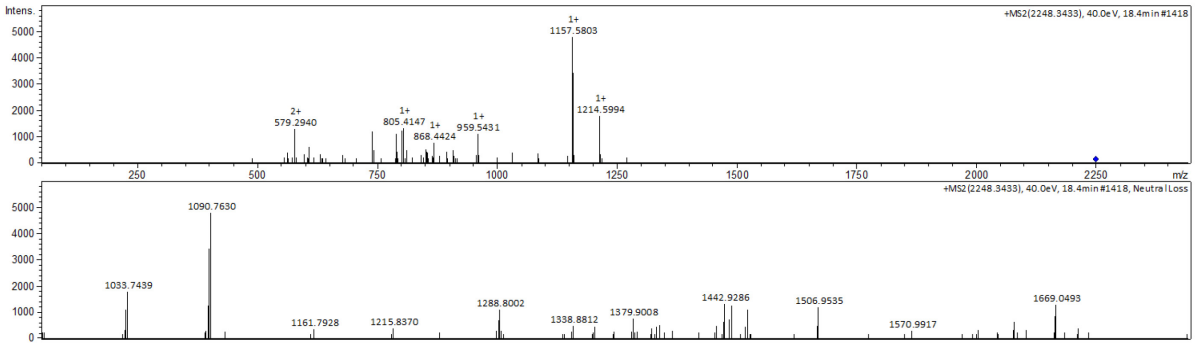
m/z 1679.9249 – DIWSLIpSDDTFLVR



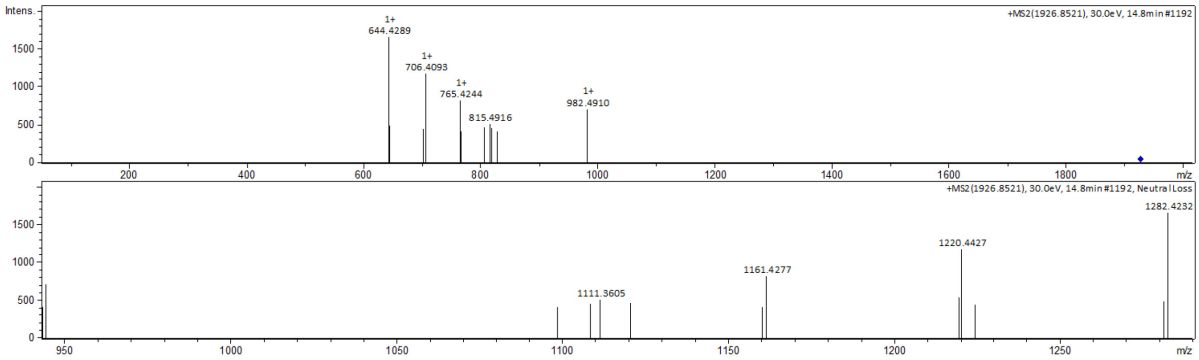
m/z 1882.9150 – VLINYATpSTEEAEEVSK



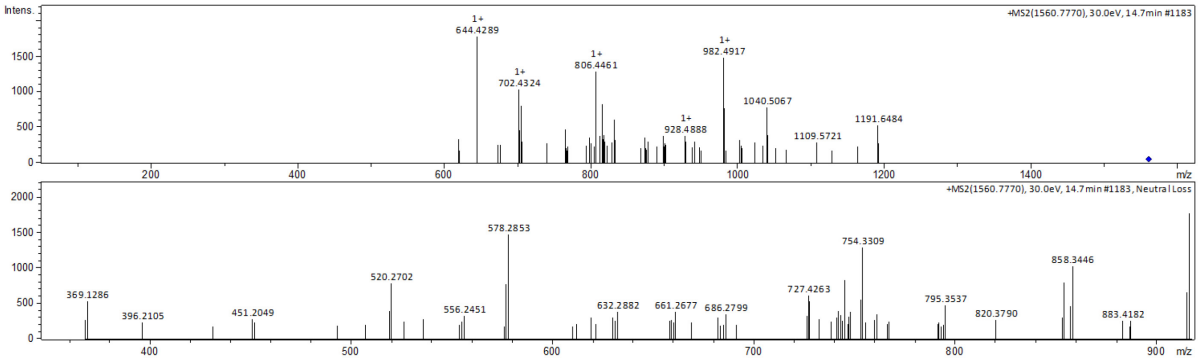
m/z 2248.3432 – IINIApSVVGLTGNAGQANYpSApSK



m/z 1926.8520 – CSVDpSAVKTEEEVPIEK



m/z 1560.7770 – QLALpSDDTPVTGEpSK



m/z 1507.8675 – TIQYLIGpSGMDPR

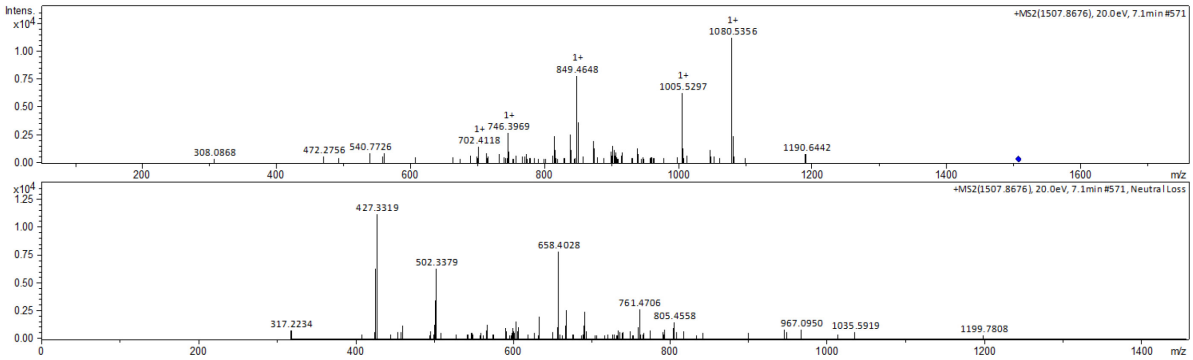


Figure 5.9. Precursor ions (indicated with a blue diamond) without any neutral loss, detected using SRM analysis for other fatty acid biosynthetic enzymes. Predicted phosphorylated serine residues with KinasePhos 2.0 are indicated in red. These peptides have precursor ion masses that matched to the precursor ions obtained in the SRM experiment.

5.5 Discussion

By employing complementary phosphopeptide detection approaches (NLMS^3 and SRM), several peptides showing the neutral losses that matched to the unique peptides identified to the subunit biotin carboxylase of acetyl-CoA carboxylase and 3-enoyl-ACP reductase were detected. The NLMS^3 approach did not detect any neutral loss from the fragmentation of peptides *m/z* 1004.59 and *m/z* 1393.74 that corresponded to peptides KLVLSTSEK and ITSYLPSGGPFVR from the subunit biotin carboxylase. However, the serine residues in these peptides were revealed to exhibit phosphoric acid losses in SRM experiments, indicating the importance of these complementary approaches. Both of the peptides were predicted to have phosphorylated serine residue at Ser-527 (KLVLSTpSEK); Ser-425 and Ser-429 (ITpSYLpSGGPFVR) (Table 5.5). Ser-527 is located in the carboxyl terminal while Ser-425 and Ser-429 are located in the biotin carboxylation domain of subunit biotin carboxylase. These phosphopeptide candidates have never been reported before and the predicted phosphosites in different protein domains of subunit biotin carboxylase suggest a crucial role of phosphorylation in the functioning of acetyl-CoA carboxylase. Phosphorylation of serine residues in the C-terminal domain of Arabidopsis RNA polymerase II has been reported to regulate the transcription and cotranscriptional RNA processing (Hajheidari, Farrona, Huettel, Koncz, & Koncz, 2012). AMP-activated protein kinase (AMPK) had been denoted to be the main kinase regulator of acetyl-CoA carboxylase (Brownsey et al., 2006; Hardie, 1992; Hardie & Pan, 2002). Their studies also demonstrated that the regulation of mammalian acetyl-CoA carboxylase was regulated by the phosphorylation of multiple serine residues.

The same serine residues that could have been phosphorylated for subunit biotin carboxylase were found in both low and high oleic acid varieties. Previous studies on regulation of fatty acid synthesis revealed that subunits of acetyl-CoA carboxylase were deactivated by phosphorylation (Brownsey et al., 2006; Hardie & Pan, 2002; Hwang et al., 2014). Phosphorylation inhibits the ATP-utilising production of malonyl-CoA, which is the mandatory carbon precursor for fatty acid biosynthesis. Therefore, the results from this study speculate that subunit biotin carboxylase of acetyl-CoA carboxylase had been deactivated in both varieties as the oil palm fruit reached maturation (20th week after anthesis), possibly as the result of phosphorylation. Although it had been proven that the regulation of plant lipid biosynthesis was regulated by acetyl-CoA carboxylase, the different level of oleic acid content between both varieties might not be regulated by the activation/deactivation of acetyl-CoA

carboxylase. NLMS^3 and SRM experiments did not show any presence of phosphorylated serine residue in peptides with their precursor ion masses matched to unique peptides that belong to the subunit β-carboxyl transferase. The only detected peptide for subunit β-carboxyl transferase was DIWSLISDDTFLVR (*m/z* 1679.92), which is not located in any catalytic domain.

NLMS^3 experiments identified peptide with *m/z* 1125.66 that corresponded to peptide GVVSGLPIDLR belonging to 3-enoyl-ACP reductase as exhibiting neutral losses in both low oleic acid and high oleic acid varieties. The same peptide was also found to have neutral losses in both varieties using the SRM approach. The phosphorylation was most likely to occur at Ser-85 as predicted with KinasePhos 2.0, which is located in the NAD(P)-binding domain of 3-enoyl-ACP reductase (Table 5.5). In addition to that, peptides with *m/z* 838.47 and *m/z* 1171.64, which were matched to peptides GKFDESR and VNTISAGPLGSR identified to 3-enoyl-ACP reductase, were also demonstrated to show neutral losses of 98 Da in both varieties (from SRM experiments). All of the potential phosphorylated Ser-146 (GKFDEpSR), Ser-307 and Ser-313 (VNTIpSAGPLGpSR) were located in the NAD(P)-binding domain (Table 5.5). Since 3-enoyl-ACP reductase required NADP as the electron donor to form a saturated fatty acid, therefore, the phosphorylation of these serine residues in the NAD(P)-binding domain may have involve in the enzyme regulation of the last step of fatty acid elongation at 20th week after anthesis. Given that all these peptides were found in both low oleic acid and high oleic acid varieties, these results indicate, similarly to acetyl-CoA carboxylase, that there is no direct evidence that phosphorylation regulates the generation of oleic acid but only the overall lipid biosynthesis through the possible deactivation of 3-enoyl-ACP reductase.

Although the predicted phosphoserine containing peptides (by KinasePhos 2.0) were identified in other fatty acid synthase enzymes, these peptides were not found to show any neutral loss in either NLMS^3 or SRM experiments. Peptide IINIASVVGLTGNAGQANYSASK (*m/z* 2248.34) for β-ketoacyl-ACP reductase has possible phosphosites located in the enzyme NAD(P)-binding domain while peptide QLALSDDTPVTGESK (*m/z* 1560.77) of β-ketoacyl-ACP synthase's phosphosites are located in the acyl carrier protein phosphopantetheine attachment site. Detected peptide CSVDSAVKTEEEVPIEK (*m/z* 1926.85) has predicted phosphosites that are not situated in FAbA-like domain in β-hydroxyacyl-ACP dehydrogenase. The functional significance of this predicted phosphopeptide was unclear. It was not apparent from these results whether phosphorylation plays any role in their enzyme regulation during the increased oleic acid production in high oleic acid variety. The key enzyme in the conversion

of saturated fatty acid to yield oleic acid in oil palm fruit is the stearoyl-ACP desaturase. However, the detected serine containing peptide from the study did not have any predicted phosphosite in the fatty acid desaturases domain of this enzyme. In addition, the peptide with *m/z* 1507.86 that matched to the precursor ion mass of peptide TIQYLIGSGMDPR did not show the neutral loss in either oil palm varieties.

Development and maturation of oil palm fruit happened at 20th week after anthesis (Dussert et al., 2013; Tranbarger et al., 2011). At this stage, Tranbarger and co-workers had reported a major increase in the lipid accumulation in the mesocarps, in the form of triacylglycerols (esterification of glycerol with three fatty acids) (Tranbarger et al., 2011). They also exhibited substantial changes in fatty acid composition (Dussert et al., 2013). Dussert and co-workers reported that palmitoyl-ACP (C16:0), the precursor for stearoyl-ACP (Figure 5.10) did not increase significantly after 20th week after anthesis. Based on their results, a possible explanation could be inferred based on the findings in this study. Decrease or inhibition of acetyl-CoA carboxylase activity could be resulted from phosphorylation on the serine residues in its domains. Thus, the availability of malonyl-CoA is diminished as the fruit reached maturation in both low oleic acid and high oleic acid varieties. The lack of malonyl-CoA subsequently caused the decrease in the amount of C16:0 as stated in previous works (Dussert et al., 2013). This study also suggested that neutral losses were detected from serine residues of peptide in the NAD(P)-binding domain of 3-enoyl-ACP reductase in both varieties. Since this enzyme generated a saturated fatty acid for the subsequent elongation phase, phosphorylation could have reduced its activity and halted the elongation steps. That explained only a small increment in C16:0 at 20th week after anthesis (0.3 mg [w/w]) and a rise of stearoyl-ACP (C18:O) by 0.4 mg (w/w) (Dussert et al., 2013).

Other fatty acid biosynthetic enzymes like β-ketoacyl-ACP reductase and β-hydroxyacyl-ACP dehydrogenase did not show any presence of phosphorylation, signified by the presence of a neutral loss. This probably indicated that dephosphorylation had reduced or inhibited their activities once the demand for C16:0 has been met on fruit maturation in both varieties. The fact that all these fatty acid biosynthetic enzymes showed similar phosphorylation state in low oleic acid and high oleic acid varieties, the main regulator of the oleic acid biosynthesis that gave a difference of 10% (in average) in the oleic acid content between the two varieties used in this study, could well be the stearoyl-ACP or one of the β-ketoacyl-ACP synthase isoenzymes (I and II). Unfortunately, β-ketoacyl-ACP synthase isoenzymes could not be distinguished in

this study although the detected serine-containing peptide (of β-ketoacyl-ACP synthase) from both varieties revealed no phosphorylation.

Dussert and co-workers indicated that the composition of oleic acid (C18:1) in mesocarp reduced from 30% to 11%, after 18th weeks after anthesis (Dussert et al., 2013). That implied that the activity of stearoyl-ACP desaturase, which is responsible to generate C18:1 through the unsaturation of C18:O had decreased (Figure 5.10). In this study, predicted phosphopeptide of stearoyl-ACP desaturase from both low oleic acid and high oleic acid varieties did not exhibit any phosphorylation. Therefore, it could be postulated that dephosphorylation had reduced or inhibited the enzyme activity. The difference in the oleic acid content of both varieties could be due to the expression and activity of stearoyl-ACP desaturase during development stage rather than when the fruit reached maturation at 20th week after anthesis.

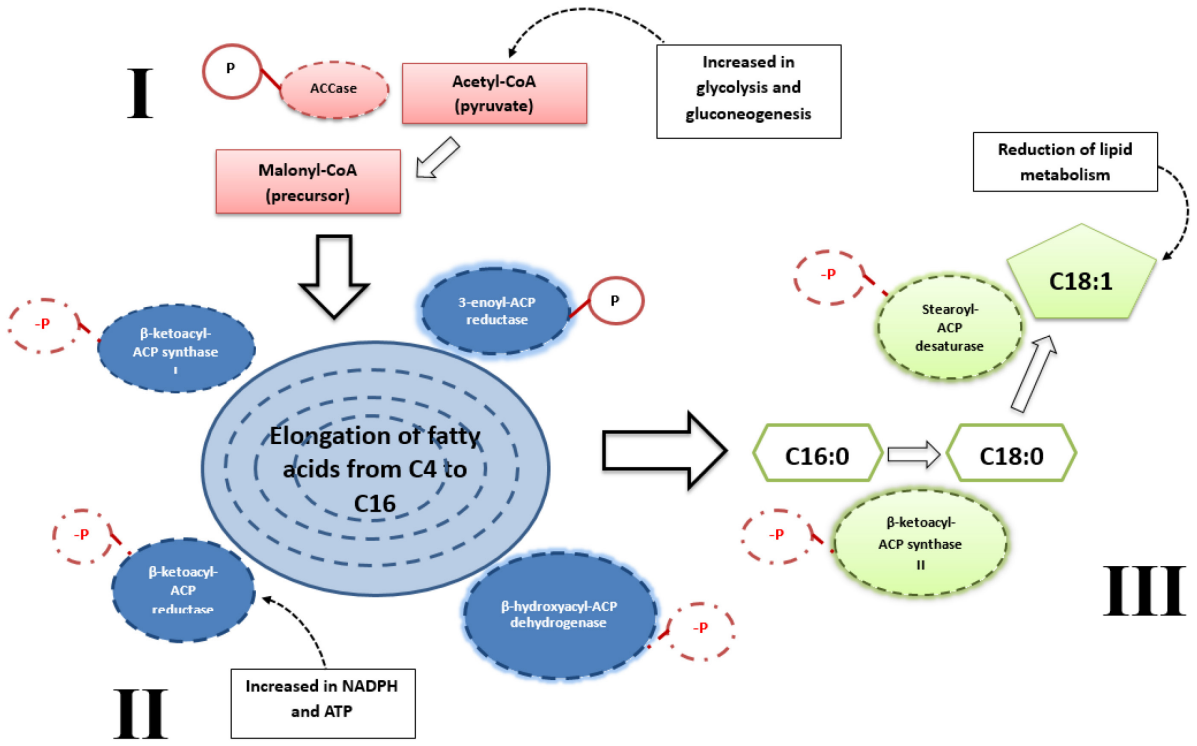


Figure 5.10. A schematic overview of oleic acid biosynthesis related enzyme phosphorylation from oil palm fruit. Circular boxes with dotted lines are the enzymes involved in various stages of fatty acid biosynthesis; I. Generation of malonyl-ACP; II. Seven condensation cycles to generate a 16-carbon fatty acid; III. Desaturation of stearic acid (C18:0-ACP) to form oleic acid (C18:1-ACP). Stearoyl-ACP desaturase is the key enzyme in oleic acid production. In this study, both acetyl-CoA carboxylase and 3-enoyl-ACP reductase were showed to be phosphorylated. Phosphorylation possibly deactivated 3-enoyl-ACP reductase as it had been demonstrated with acetyl-CoA carboxylase. It remains unclear whether phosphorylation plays any role in the regulation of other fatty acid biosynthetic enzymes, especially stearoyl-ACP desaturase.

5.6 Conclusions

The utilisation of non-targeted neutral loss-triggered MS^3 and Selected Reaction Monitoring was aimed at identifying phosphopeptide candidates through the loss of a phosphoric acid molecule (neutral loss). The results demonstrated that a combination of these approaches was valuable in detecting neutral losses from serine containing peptides. Figure 5.10 summarises the results obtained in this study using both approaches. The neutral losses suggest that phosphorylation could have occurred to these peptides. Therefore, it is possible to postulate that phosphorylation regulates the fatty acid biosynthesis through the deactivation of both acetyl-CoA carboxylase and 3-enoyl-ACP reductase, while dephosphorylation could have deactivated the other fatty acid biosynthetic enzymes from low oleic acid and high oleic acid varieties. The study also identified a number of potential phosphosites in the protein domains of fatty acid biosynthetic enzymes which could play a

regulatory role in their enzyme activities. Despite the high occurrence of predicted phosphosites in subunit β-carboxyl transferase of acetyl-CoA carboxylase, β-ketoacyl-ACP reductase, β-hydroxyacyl-ACP dehydrogenase and stearoyl-ACP desaturase, detection of their peptides containing these phosphorylation sites had been poor. Thus, the exact role of phosphorylation in their enzyme activities, especially in stearoyl-ACP desaturase, remained inconclusive. Further research is necessary to validate these phosphosites and increase the chance of detecting more phosphopeptides before the exact role of phosphorylation in the regulation of oleic acid biosynthesis can be established.

6.0 General discussion

The experimental work underpinning this thesis is described within three of the chapters (Chapters 3, 4 and 5). The first aim was to develop optimised methodologies to isolate and enhance the extraction of chromoplast proteins from the oil palm fruit mesocarps. Following the acquisition of mass spectrometry compatible proteins, a combination of different proteomic-based techniques was applied to investigate the expression of enzymes associated with plant fatty acid biosynthesis from low oleic acid and high oleic acid palms. Subsequently, advanced mass spectrometry-based approaches were exploited to characterise the role of phosphorylation in regulating the fatty acid biosynthetic enzyme activities between the two oil palm varieties.

6.1 Method development to provide chromoplast proteins for shotgun proteomic analysis

The method development described in Chapter 3 was specifically tailored for oil palm fruit mesocarps to isolate plastid-based (chromoplast) proteins and improve their extraction for proteomic analyses. Oily fruit-based proteins are known to be highly recalcitrant to common protein extraction approaches. In addition, these plant-based protein materials contain high levels of interfering compounds that are detrimental to proteomic techniques. Therefore, the mesocarps were first delipidated using a combination of acetone and methanol solvents. Their ability to extract different classes of lipid proved to be advantageous in removing lipids from the fruit mesocarps, especially for the high oleic acid variety. Utilisation of cellulase and pectinase enhanced the protein yields by breaking down the plant cell walls to release the plant organelles. Chromoplasts were separated from other plant organelles using a combination of differential and density gradient centrifugation approaches prior to protein extraction. These steps, especially the differential centrifugation, were critical to permit the extraction of mainly chromoplast proteins (up to 90% of total proteins) for proteomic investigation of chromoplast-based fatty acid biosynthetic enzymes. Phenol-based protein extraction and the precipitation of proteins with ammonium acetate were demonstrated to give the highest protein yield and quality. Phenol was used as its interaction with proteins through hydrogen bonding enabled the isolation of predominantly proteins from other interfering compounds. Ammonium acetate was used to precipitate and eliminate more

interfering compounds from the isolated proteins, something which the commonly used TCA/acetone protein precipitation approach was unable to do.

6.2 Utilisation of non-gel based proteomic technique to enhance the protein expression characterisation of fatty acid biosynthetic enzymes

Different proteomic techniques were first employed to investigate the biological variations within the three different replicates of the low oleic acid and high oleic acid varieties used in this study (Chapter 4). Strategies developed in Chapter 3 were used to provide the chromoplast proteins. A combination of gel- and non-gel based approaches provided 'deeper' findings with regard to their biological variations. The variations were not evident using the gel-based techniques (1D SDS-PAGE and 2DE) alone. However, the non-gel based technique (2DLC-MS/MS) gave a better overview of the degree of variation between these replicates. LO1 was found to give a higher number of identified proteins and peptides relative to their LO2 and LO3 counterparts. Meanwhile, HO1 and HO3 from the high oleic variety were found to be varied to a lesser extend compared to HO2, although these differences were small.

GeLC-MS/MS identified β-ketoacyl-ACP reductase, β-hydroxyacyl-ACP dehydrogenase, 3-enoyl-ACP reductase, β-ketoacyl-ACP synthase and stearoyl-ACP desaturase. However, non-gel based technique utilising offline 2DLC-MS/MS managed to identify an additional main fatty acid biosynthetic enzyme including its subunits. The complementary non-gel based technique was demonstrated to improve the number of detected unique peptides for β-ketoacyl-ACP synthase and stearoyl-ACP desaturase from both low oleic acid and high oleic acid varieties. In contrast, a lower number of peptides was detected for β-hydroxyacyl-ACP dehydrogenase and 3-enoyl-ACP reductase using this non-gel based technique. Both approaches gave a nearly similar number of detected peptides for β-ketoacyl-ACP reductase. Overall, enzymes involved in the elongation of the fatty acid chain, namely, the β-ketoacyl-ACP reductase, β-hydroxyacyl-ACP dehydrogenase and 3-enoyl-ACP reductase were found to have a higher number of detected peptides in both varieties compared to acetyl-CoA carboxylase, β-ketoacyl-ACP synthase and stearoyl-ACP desaturase. This might suggest that those enzymes were present in higher abundance relative to acetyl-CoA carboxylase, β-ketoacyl-ACP synthase and stearoyl-ACP desaturase in fruit mesocarp during this stage of maturation.

Chapter 4 also documented valuable information on the expression levels of different enzymes implicated with plant oleic acid biosynthesis, in addition to earlier proteomic studies (Loei et al., 2013; Sambanthamurthi et al., 1999). The information in this study was obtained through a combination of an offline 2DLC-MS/MS approach with an iTRAQ labelling strategy. Only peptides detected in both low oleic acid and high oleic acid varieties for acetyl-CoA carboxylase (subunit biotin carboxylase), β-hydroxyacyl-ACP dehydrogenase, 3-enoyl-ACP reductase and β-ketoacyl-ACP synthase were compared to determine their relative protein ratios. The outcomes indicated that both β-hydroxyacyl-ACP dehydrogenase and β-ketoacyl-ACP synthase expressions were up-regulated in the high oleic acid variety. The protein expression of 3-enoyl-ACP reductase was down-regulated in high oleic acid variety while the expression of acetyl-CoA carboxylase remained unchanged. In addition to examining the protein expression of fatty acid biosynthetic enzymes, the protein expressions of other metabolic enzymes associated with the fatty acid production were also investigated. Fructose-biphosphate aldolase, malate dehydrogenase and sucrose synthase were found to have higher levels of expression in the high oleic acid variety. Elevated expression of proteins involved in ATP binding and acyl group transfer were also noticed in high oleic acid variety. Notably, in contrast, the expression of lipase was higher in the low oleic acid variety.

Earlier studies (Dussert et al., 2013; Tranbarger et al., 2011) reported that the production of palmitoyl-ACP (C16:0) reached a plateau at the fruit maturation age of 20th week after anthesis. This coincided with the unaltered expression of acetyl-CoA carboxylase in both low oleic acid and high oleic acid varieties, since acetyl-CoA produced malonyl-CoA, the precursor for fatty acid elongation process to produce C16:0. The results reported in Chapter 4 showed that the expression of enzymes involved in elongation of fatty acids (β-hydroxyacyl-ACP dehydrogenase and β-ketoacyl-ACP synthase) to produce C16:0 were higher in the high oleic acid variety. Both enzymes are involved in the elongation of fatty acids. Therefore, the higher expression of these enzymes suggested that the C16:0 produced has been converted to stearoyl-ACP (C18:0), the precursor for oleic acid in the high oleic acid variety. Furthermore, the down-regulated 3-enoyl-ACP reductase in high oleic acid variety and the unaltered expression of acetyl-CoA carboxylase in both low oleic acid and high oleic acid varieties postulated that oleic acid production was regulated by other fatty acid biosynthetic enzymes in the downstream stages, rather than during the fatty acid elongation phase. Since the expression level of stearoyl-ACP desaturase (the main enzyme in oleic acid biosynthesis) could

not be determined, its role in regulating oleic acid biosynthesis could not be established. Influxes of acetyl-CoA (precursor for malonyl-ACP), NADPH and ATP could potentially increase the production of oleic acid in high oleic acid variety indirectly. Furthermore, the down-regulated expression of lipase in high oleic acid variety indicated that less oleic acid esters would be metabolised to generate free oleic acids. This may explain the higher level of oleic acids (in the form of esters) in the high oleic acid variety.

6.3 Plausibility of phosphorylation involvement in regulating fatty acid biosynthetic enzyme activities

Phosphorylation is an important mechanism to regulate protein activity. Chapter 5 described the utilisation of two advanced mass spectrometry-based techniques to detect the neutral loss of the phosphoric acid group from phosphorylated serine-containing peptides. Using the neutral-loss triggered MS^3 and Selected Reaction Monitoring approaches, peptides showing the neutral losses were found, which corresponded to the peptides identified to 3-enoyl-ACP reductase and subunit biotin carboxylase of acetyl-CoA carboxylase (based on precursor ion masses). These phosphopeptide candidates were detected in both low oleic acid and high oleic acid varieties, indicating that phosphorylation of these enzymes did not regulate the oleic acid biosynthesis. As the predicted phosphorylation sites of these peptides were found in the protein domains of 3-enoyl-ACP reductase and biotin carboxylase, the findings suggested a crucial role of phosphorylation in regulating the enzyme activities, perhaps during the palmitoyl-ACP production. Phosphorylation was known to deactivate acetyl-CoA carboxylase (Brownsey et al., 2006; Hardie & Pan, 2002; Hwang et al., 2014). Therefore, the phosphorylated peptide candidates for subunit biotin carboxylase could mean the activity of acetyl-CoA had reduced when the fruit reached maturation at 20^{th} week after anthesis. As stated in Chapter 4, the expression of 3-enoyl-ACP reductase decreased in high oleic acid variety. However, in this part of study (Chapter 5), phosphopeptide candidates were identified to 3-enoyl-ACP reductase from both varieties. The amount of protein does not necessarily equate to the protein activity (Cramer et al., 2013). Therefore, this could indicate that phosphorylation regulated the enzyme activities but not their protein expression. Its enzyme activity was reduced since palmitoyl-ACP did not show any substantial increment at fruit maturation stage.

There was no indication of possible phosphorylation in the serine-containing peptides for β-ketoacyl-ACP reductase, β-hydroxyacyl-ACP dehydrogenase and β-ketoacyl-ACP synthase.

However, the results described in Chapter 4 indicated that the protein expression of β-hydroxyacyl-ACP dehydrogenase and β-ketoacyl-ACP synthase was up-regulated in the high oleic acid variety. Similarly to 3-enoyl-ACP reductase, phosphorylation does not appear to play a role in determining the protein expression of β-hydroxyacyl-ACP dehydrogenase and β-ketoacyl-ACP synthase in the high oleic acid variety. A dephosphorylation event may have decreased the activities of all three enzymes instead, as the fruit reached the end of its development. Phosphorylation was not evident for both the subunit β-carboxyl transferase and the key enzyme in oleic acid biosynthesis, the stearoyl-ACP desaturase. Therefore, the role of phosphorylation in regulating stearoyl-ACP desaturase via increasing the oleic acid content remained inconclusive.

6.4 Concluding remarks

The robust methods to isolate and extract chromoplast proteins from oil palm mesocarp proved to be a significant improvement over existing but limited organelle proteomic studies in oily fruit. The obvious downside of the approach was the time needed to complete the whole process. The optimised chromoplast isolation and its protein extraction provided an excellent foundation for application to other plant organelle proteomic study in the future.

The relatively low stoichiometry of fatty acid biosynthetic enzymes, especially stearoyl-ACP desaturase, presents a very challenging analytical problem, even with the combination of different proteomic approaches. The situation becomes worse with very low expression of stearoyl-ACP desaturase, the key enzyme in oleic acid biosynthesis. Nonetheless, the combination of these proteomic techniques was highly valuable in obtaining the key enzymes involved in the fatty acid biosynthesis.

The outcomes reported throughout this thesis have significantly enhanced the knowledge of fatty acid biosynthesis in plant. The expression profiles for the key fatty acid biosynthetic enzymes provided new insights into their regulation in different oil palm varieties, with regard to oleic acid content. Novel insights have also been made on the possible phosphorylation states of each of the fatty acid biosynthetic enzymes. A plausible role for phosphorylation and dephosphorylation events in regulating the overall fatty acid production between low oleic acid and high oleic acid palms has also been described. In addition, the phosphorylated serine residues in the protein domains of the fatty acid biosynthetic enzymes could be exploited for further investigations into the exact role of phosphorylation. All of these are novel results.

6.5 Future perspectives

The low stoichiometry of phosphorylated peptides remains a huge obstacle in proteomic investigation, as in this study. Relatively lower abundance of these phosphopeptides makes it difficult to study them. Therefore, utilisation of affinity-based columns such as Immobilised Metal ion Affinity Chromatography (IMAC) and Titanium Dioxide (TiO_2) will potentially improve the detection of phosphorylated peptides. Although utilisation of TiO_2 column has been trialled in this study (not reported in this thesis), further optimisation would be essential to improve the binding efficiency of phosphopeptides to the column. To further reduce the suppression of phosphorylated peptides, these enrichment strategies could be used in combination with electron transfer dissociation (ETD). Unlike the collision-induced dissociation used in this study, peptide fragmentations with ETD retain the labile phosphate group in the peptide backbone fragments after dissociations. The apparent advantage of this type of fragmentation is that the site of modification can be identified more confidently. The non-existent neutral loss peak related to the phosphate group will enhance the identification of phosphopeptides caused by suppression of sequence diagnostic ions in the presence of neutral loss peak. ETD in combination with SCX-2DLC-MS/MS had also been attempted in this study but without success, at the time of this thesis. Further optimisation is considered to be required. Another approach to enhance the complete coverage of phosphoamino-containing peptides is to combine the use of a reverse-phase column (used in 2DLC-MS/MS approach in this study) with a Hydrophilic Interaction Chromatography or HILIC. This combination presents another prospective phosphopeptide enrichment strategy. The phosphate group is strongly hydrophilic and therefore, capable of altering the hydrophilicity of the attached peptide. Thus, the phosphopeptides would bind more tightly to the neutral, hydrophilic stationary phase in a HILIC column, relatively to non-phosphopeptides. This may have contributed to the inability to obtain more phosphoamino-containing peptides in this study as only a hydrophobic column was used (Chapter 5).

Once the problems of low stoichiometric phosphopeptides have been overcome, the next step is to investigate and determine the biological significance of the protein expression and phosphorylation states of the fatty acid biosynthetic enzymes during the oil palm development stages. The results from this study reflected that the role of phosphorylation in fatty acid biosynthesis will be better defined throughout the fruit development stages, rather than at the fruit maturation stage. iTRAQ labelling technique will be applied but in

combination with the Selected Reaction Monitoring (SRM) technique. SRM (Chapter 5) has the capability to selectively quantitate iTRAQ-labelled phosphopeptides of fatty acid biosynthetic enzymes since their unique peptides (with the corresponding *m/z*) had been identified (Chapter 4). The advantage of SRM should be exploited to detect the low abundance phosphopeptides more effectively. In addition, as more information on the oil palm genome is being uncovered, its non-redundant protein sequence database can be established. The database would include the descriptions of the function of a protein, its domain structure, reactions catalysed by the protein, similarities to protein sequences from other related species and perhaps, the post-translational modifications. Thus, the efforts to distinguish the isoforms of β-ketoacyl-ACP synthase to determine which isoform regulates the oleic acid level as well as enhancing the protein identification confidence can also be carried out.

Elucidating complex signalling pathway through phosphorylation events can be difficult and laborious. Combinations of various proteomic-based strategies have been used to identify the dynamic changes of phosphorylated enzymes in this study. In addition to that information, determination of the phosphorylation sites is equally important to understand the role the phosphorylation plays in the function of fatty acid biosynthetic enzymes. Several possible phosphosites had been reported in this study (Chapter 5). Validation of these sites can be carried out with site-directed mutagenesis on these enzymes. Serine, threonine or tyrosine residue deletion in the protein domains will be useful to confirm precise role of phosphorylation. In order to comprehend the entire signalling pathway of fatty acid biosynthesis, it is vital to determine the exact effect of phosphorylation on these fatty acid biosynthetic enzymes. Phosphorylation can activate or inhibit enzyme activity (as postulated in this study, Chapter 5) through conformational changes or even create recognition sites for other protein molecules. At an industrial level, identification of the protein kinase and phosphatase that interacts with these phosphosites can be beneficial. Protein kinases and phosphatases are usually the prime targets for the manipulation of the phosphorylation state of the target enzyme. The phosphorylation state may be able to be altered to increase the oleic acid and linoleic acid contents to yield a higher value-added palm oil for the global oil market.

7.0 References

Abdallah, C., Dumas-Gaudot, E., Renaut, J., & Sergeant, K. (2012). Gel-based and gel-free quantitative proteomics approaches at a glance. *International Journal of Plant Genomics, 2012*, 494572. doi:10.1155/2012/494572

Acin, P., Rayo, J., Guerrero, A., & Quero, C. (2009). Improved resolution in the acidic and basic region of 2-DE of insect antennae proteins using hydroxyethyl disulfide [Research Support, Non-U.S. Gov't]. *ELECTROPHORESIS, 30*(15), 2613-2616. doi:10.1002/elps.200900031

Adam, H., Jouannic, S., Escoute, J., Duval, Y., Verdeil, J.-L., & Tregear, J. W. (2005). Reproductive developmental complexity in the African oil palm (Elaeis guineensis, Arecaceae). *American Journal of Botany, 92*(11), 1836-1852. doi:10.3732/ajb.92.11.1836

Adams, J. A. (2001). Kinetic and catalytic mechanisms of protein kinases. *Chemical Reviews, 101*(8), 2271-2290. doi:10.1021/cr000230w

Aebersold, R., & Goodlett, D. R. (2001). Mass spectrometry in proteomics. *Chemical Reviews, 101*(2), 269-296. doi:10.1021/cr990076h

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature, 422*(6928), 198-207. doi:10.1038/nature01511

Agrawal, G., & Thelen, J. (2009). A high-resolution two dimensional gel- and Pro-Q DPS-based proteomics workflow for phosphoprotein identification and quantitative profiling. In M. Graauw (Ed.), *Phospho-Proteomics* (Vol. 527, pp. 3-19): Humana Press. doi:10.1007/978-1-60327-834-8_1

Agrawal, G. K., Bourguignon, J., Rolland, N., Ephritikhine, G., Ferro, M., Jaquinod, M., . . . Rakwal, R. (2011). Plant organelle proteomics: Collaborating for optimal cell function. *Mass Spectrometry Reviews, 30*(5), 772-853. doi:10.1002/mas.20301

Agrawal, G. K., & Thelen, J. J. (2006). Large scale identification and quantitative profiling of phosphoproteins expressed during seed filling in oilseed rape. *Molecular & Cellular Proteomics, 5*(11), 2044-2059. doi:10.1074/mcp.M600084-MCP200

Akmar, S. N., Cheah, A., Aminah, S. C., Leslie, C. L. O., Sambanthamurthi, R., & Murphy, D. J. Characterization and regulation of the oil palm (*Elaeis guineensis*) stearoyl-acp desaturase genes. *Journal of Oil Palm Research*(Special Issue), 1-17.

Alexander, L., & Grierson, D. (2002). Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening [Review]. *Journal of Experimental Botany, 53*(377), 2039-2055. doi:10.1093/jxb/erf072

Amalraj, R. S., Selvaraj, N., Veluswamy, G. K., Ramanujan, R. P., Muthurajan, R., Palaniyandi, M., . . . Viswanathan, R. (2010). Sugarcane proteomics: establishment of a protein extraction method for 2-DE in stalk tissues and initiation of sugarcane proteome reference map. *ELECTROPHORESIS, 31*(12), 1959-1974. doi:10.1002/elps.200900779

American Palm Oil Council. (n.d.). Retrieved from http://www.americanpalmoil.com/news.html.

Amey, R. C., Schleicher, T., Slinn, J., Lewis, M., Macdonald, H., Neill, S. J., & Spencer-Phillips, P. T. N. (2008). Proteomic analysis of a compatible interaction between Pisum sativum (pea) and the downy mildew pathogen Peronospora viciae. In A. Lebeda, P. N. Spencer-Phillips & B. M. Cooke (Eds.), *The Downy Mildews - Genetics, Molecular Biology and Control* (pp. 41-55): Springer Netherlands. Retrieved from http://dx.doi.org/10.1007/978-1-4020-8973-2_5. doi:10.1007/978-1-4020-8973-2_5

Andon, N. L., Hollingworth, S., Koller, A., Greenland, A. J., Yates III, J. R., & Haynes, P. A. (2002). Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. *PROTEOMICS, 2*(9), 1156-1168. doi:10.1002/1615-9861(200209)2:9<1156::aid-prot1156>3.0.co;2-4

Andre, C., Froehlich, J. E., Moll, M. R., & Benning, C. (2007). A heteromeric plastidic pyruvate kinase complex involved in seed oil biosynthesis in Arabidopsis [Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, Non-P.H.S.]. *Plant Cell, 19*(6), 2006-2022. doi:10.1105/tpc.106.048629

Andre, C., Haslam, R. P., & Shanklin, J. (2012). Feedback regulation of plastidic acetyl-CoA carboxylase by 18:1-acyl carrier protein in Brassica napus. *Proceedings of the National Academy of Sciences 109*(25), 10107-10112. doi:10.1073/pnas.1204604109

Araùjo, M., Hube, L., & Stasyk, T. (2008). Isolation of endocitic organelles by density gradient centrifugation. In A. Posch (Ed.), *2D PAGE: Sample Preparation and Fractionation* (Vol. 424, pp. 317-331): Humana Press. Retrieved from http://dx.doi.org/10.1007/978-1-60327-064-9_25. doi:10.1007/978-1-60327-064-9_25

Aryal, U. K., Krochko, J. E., & Ross, A. R. S. (2011). Identification of phosphoproteins in *Arabidopsis thaliana* leaves using polyethylene glycol fractionation, immobilized metal-ion affinity chromatography, two-dimensional gel electrophoresis and mass spectrometry. *Journal of Proteome Research, 11*(1), 425-437. doi:10.1021/pr200917t

Asara, J. M., Christofk, H. R., Freimark, L. M., & Cantley, L. C. (2008). A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *PROTEOMICS, 8*(5), 994-999. doi:10.1002/pmic.200700426

Auld, D. L., Heikkinen, M. K., Erickson, D. A., Sernyk, J. L., & Romero, J. E. (1992). Rapeseed mutants with reduced levels of polyunsaturated fatty acids and increased levels of oleic acid. *Crop Science, 32*, 657-662. doi:10.2135/cropsci1992.0011183X003200030016x

Awai, K., Maréchal, E., Block, M. A., Brun, D., Masuda, T., Shimada, H., . . . Joyard, J. (2001). Two types of MGDG synthase genes, found widely in both 16:3 and 18:3 plants, differentially mediate galactolipid syntheses in photosynthetic and nonphotosynthetic tissues in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences, 98*(19), 10960-10965. doi:10.1073/pnas.181331498

Bachi, A., & Bonaldi, T. (2008). Quantitative proteomics as a new piece of the systems biology puzzle. *Journal of Proteomics, 71*(3), 357-367. doi:10.1016/j.jprot.2008.07.001

Baginsky, S., Siddique, A., & Gruissem, W. (2004). Proteome analysis of tobacco bright yellow-2 (BY-2) cell culture plastids as a model for undifferentiated heterotrophic plastids. *Journal of Proteome Research, 3*(6), 1128-1137. doi:10.1021/pr0499186

Balbuena, T. S., He, R., Salvato, F., Gang, D. R., & Thelen, J. J. (2012). Large-scale proteome comparative analysis of developing rhizomes of the ancient vascular plant equisetum hyemale [Original Research]. *Frontiers in Plant Science, 3*, 131. doi:10.3389/fpls.2012.00131

Baldwin, E. A., & Biggs, R. H. (1988). Cell-wall lysing enzymes and products of cell-wall digestion elicit ethylene in citrus. *Physiologia Plantarum, 73*(1), 58-64. doi:10.1111/j.1399-3054.1988.tb09193.x

Balmer, Y., Vensel, W. H., DuPont, F. M., Buchanan, B. B., & Hurkman, W. J. (2006). Proteome of amyloplasts isolated from developing wheat endosperm presents evidence of broad metabolic capability. *Journal of Experimental Botany, 57*(7), 1591-1602. doi:10.1093/jxb/erj156

Bamgboye, A. I., & Hansen, A. C. (2008). Prediction of cetane number of biodiesel fuel from the fatty acid methyl ester (FAME) composition. *International Agrophysics, 22*(1), 21-29.

Bao, X., Focke, M., Pollard, M., & Ohlrogge, J. (2000). Understanding in vivo carbon precursor supply for fatty acid synthesis in leaf tissue. *The Plant Journal, 22*(1), 39-50. doi:10.1046/j.1365-313x.2000.00712.x

Barsan, C., Sanchez-Bel, P., Rombaldi, C., Egea, I., Rossignol, M., Kuntz, M., . . . Pech, J.-C. (2010a). Characteristics of the tomato chromoplast revealed by proteomic analysis. *Journal of Experimental Botany*. doi:10.1093/jxb/erq070

Barsan, C., Sanchez-Bel, P., Rombaldi, C., Egea, I., Rossignol, M., Kuntz, M., . . . Pech, J.-C. (2010b). Characteristics of the tomato chromoplast revealed by proteomic analysis. *Journal of Experimental Botany, 61*(9), 2413-2431. doi:10.1093/jxb/erq070

Basiron, Y. (2012). *Palm oil: The worlds's best edible oil?* Retrieved 4 June, 2013,from http://www.ceopalmoil.com/2012/05/palm-oil-the-world%E2%80%99s-best-edible-oil/.

Basu, S., & Plewczynski, D. (2010). AMS 3.0: prediction of post-translational modifications. *BMC Bioinformatics, 11*(1), 210.

Baud, S., & Lepiniec, L. (2009). Regulation of de novo fatty acid synthesis in maturing oilseeds of Arabidopsis. *Plant Physiology and Biochemistry, 47*(6), 448-455. doi:10.1016/j.plaphy.2008.12.006

Bayer, E. A., Chanzy, H., Lamed, R., & Shoham, Y. (1998). Cellulose, cellulases and cellulosomes. *Current Opinion in Structural Biology, 8*(5), 548-557. doi:10.1016/s0959-440x(98)80143-7

Bayer, R. G., Stael, S., Csaszar, E., & Teige, M. (2011). Mining the soluble chloroplast proteome by affinity chromatography. *PROTEOMICS, 11*(7), 1287-1299. doi:10.1002/pmic.201000495

Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villén, J., Li, J., . . . Gygi, S. P. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proceedings of the National Academy of Sciences of the United States of America, 101*(33), 12130-12135. doi:10.1073/pnas.0404720101

Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., & Gygi, S. P. (2006). A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature Biotechnology, 24*(10), 1285-1292. doi:10.1038/nbt1240

Beer, L., Tang, H.-Y., Barnhart, K., & Speicher, D. (2011). Plasma biomarker discovery using 3D protein profiling coupled with label-free quantitation. In R. J. Simpson & D. W. Greening (Eds.), *Serum/Plasma Proteomics* (Vol. 728, pp. 3-27): Humana Press. Retrieved from http://dx.doi.org/10.1007/978-1-61779-068-3_1. doi:10.1007/978-1-61779-068-3_1

Beisson, F., Koo, A. J. K., Ruuska, S., Schwender, J., Pollard, M., Thelen, J. J., . . . Ohlrogge, J. B. (2003). Arabidopsis genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. *Plant Physiology, 132*(2), 681-697. doi:10.1104/pp.103.022988

Bendt, A. K., Burkovski, A., Schaffer, S., Bott, M., Farwick, M., & Hermann, T. (2003). Towards a phosphoproteome map of Corynebacterium glutamicum. *PROTEOMICS, 3*(8), 1637-1646. doi:10.1002/pmic.200300494

Benning, C. (2008). A role for lipid trafficking in chloroplast biogenesis. *Progress in Lipid Research, 47*(5), 381-389. doi:10.1016/j.plipres.2008.04.001

Benning, C. (2009a). Mechanisms of lipid transport involved in organelle biogenesis in plant cells. *Annual Review of Cell and Developmental Biology, 25*(1), 71-91. doi:10.1146/annurev.cellbio.042308.113414

Benning, C. (2009b). Mechanisms of lipid transport involved in organelle biogenesis in plant cells. *Annu Rev Cell Dev Biol, 25*(1), 71-91. doi:10.1146/annurev.cellbio.042308.113414

Benschop, J. J., Mohammed, S., O'Flaherty, M., Heck, A. J. R., Slijper, M., & Menke, F. L. H. (2007). Quantitative phosphoproteomics of early elicitor signaling in *Arabidopsis*. *Molecular & Cellular Proteomics, 6*(7), 1198-1214. doi:10.1074/mcp.M600429-MCP200

Bienvenut, W. V., Espagne, C., Martinez, A., Majeran, W., Valot, B., Zivy, M., . . . Giglione, C. (2011). Dynamics of post-translational modifications and protein stability in the

stroma of Chlamydomonas reinhardtii chloroplasts. *PROTEOMICS, 11*(9), 1734-1750. doi:10.1002/pmic.201000634

Bindschedler, L. V., & Cramer, R. (2011). Quantitative plant proteomics. *PROTEOMICS, 11*(4), 756-775. doi:10.1002/pmic.201000426

Blom, N., Gammeltoft, S., & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology, 294*(5), 1351-1362. doi:10.1006/jmbi.1999.3310

Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., & Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *PROTEOMICS, 4*(6), 1633-1649. doi:10.1002/pmic.200300771

Bockus, L., & Scofield, R. H. (2009). Phosphoprotein detection on protein electroblot using a phosphate-specific fluorophore. In B. T. Kurien & R. H. Scofield (Eds.), *Protein Blotting and Detection* (Vol. 536, pp. 385-393): Humana Press. Retrieved from http://dx.doi.org/10.1007/978-1-59745-542-8_39. doi:10.1007/978-1-59745-542-8_39

Boersema, P. J., Mohammed, S., & Heck, A. J. (2009). Phosphopeptide fragmentation and analysis by mass spectrometry. *Journal of Mass Spectrometry, 44*(6), 861-878. doi:10.1002/jms.1599

Bond, A. E., Row, P. E., & Dudley, E. (2011). Post-translation modification of proteins; methodologies and applications in plant sciences. *Phytochemistry, 72*(10), 975-996. doi:10.1016/j.phytochem.2011.01.029

Boudsocq, M., Droillard, M.-J., Barbier-Brygoo, H., & Laurière, C. (2007). Different phosphorylation mechanisms are involved in the activation of sucrose non-fermenting 1 related protein kinases 2 by osmotic stresses and abscisic acid. *Plant Molecular Biology, 63*(4), 491-503. doi:10.1007/s11103-006-9103-1

Bouvier, F., Backhaus, R. A., & Camara, B. (1998). Induction and control of chromoplast-specific carotenoid genes by oxidative stress. *Journal of Biological Chemistry, 273*(46), 30651-30659. doi:10.1074/jbc.273.46.30651

Brand, M. D. (1996). Top down metabolic control analysis. *Journal of Theoretical Biology, 182*(3), 351-360. doi:10.1006/jtbi.1996.0174

Bräutigam, A., & Weber, A. P. M. (2009). Proteomic analysis of the proplastid envelope membrane provides novel insights into small molecule and protein transport across proplastid membranes. *Molecular Plant, 2*(6), 1247-1261. doi:10.1093/mp/ssp070

Brechenmacher, L., Lee, J., Sachdev, S., Song, Z., Nguyen, T. H. N., Joshi, T., . . . Stacey, G. (2009). Establishment of a protein reference map for soybean root hair cells. *Plant Physiology, 149*(2), 670-682. doi:10.1104/pp.108.131649

Breure, C. J. (2006). Performance of ASD's oil palm parent material in South Sumatra. *ASD Oil Palm Papers, 29*, 19-30. Retrieved from http://www.asd-cr.com/paginas/english/articulos/bol29-2en.html website:

Brownsey, R. W., Boone, A. N., Elliott, J. E., Kulpa, J. E., & Lee, W. M. (2006). Regulation of acetyl-CoA carboxylase [Review]. *Biochemical Society Transactions, 34*(Pt 2), 223-227. doi:10.1042/BST20060223

Bruck, F., Brummel, M., Schuch, R., & Spener, F. (1996). In-vitro evidence for feed-back regulation of β-ketoacyl-acyl carrier protein synthase III in medium-chain fatty acid biosynthesis. *PLANTA, 198*(2), 271-278. doi:10.1007/bf00206253

Budiani, A., Santoso, D., Aswidinnoor, H., & Suwanto, A. (2008). ACCase activity of oil palm mesocarp and cloning of gene fragment encoding biotin carboxylase subunit of ACCase. *Indonesian Journal of Agriculture, 1*(1), 44-50.

Bunger, M. K., Cargile, B. J., Ngunjiri, A., Bundy, J. L., & Stephenson, J. L. (2008). Automated proteomics of *E. coli* via top-down electron-transfer dissociation mass spectrometry. *Analytical Chemistry, 80*(5), 1459-1467. doi:10.1021/ac7018409

Burns, M. J., Barnes, S. R., Bowman, J. G., Clarke, M. H. E., Werner, C. P., & Kearsey, M. J. (2003). QTL analysis of an intervarietal set of substitution lines in *Brassica napus*: (i) Seed oil content and fatty acid composition. *Heredity, 90*(1), 39-48. doi:10.1038/sj.hdy.6800176

Burton, J. W., Miller, J. F., Vick, B. A., Scarth, R., & Holbrook, C. C. (2004). Altering fatty acid composition in oil seed crops. *Advances in Agronomy, 84*, 273-306. doi:10.1016/S0065-2113(04)84006-9

Bylund, D., Danielsson, R., & Markides, K. E. (2001). Peak purity assessment in liquid chromatography-mass spectrometry [Research Support, Non-U.S. Gov't]. *Journal of Chromatography A, 915*(1-2), 43-52.

Cahoon, E. B., & Shanklin, J. (2000). Substrate-dependent mutant complementation to select fatty acid desaturase variants for metabolic engineering of plant seed oils. *Proc Natl Acad Sci U S A, 97*(22), 12350-12355. doi:10.1073/pnas.210276297

Cahoon, E. B., Shanklin, J., & Ohlrogge, J. B. (1992). Expression of a coriander desaturase results in petroselinic acid production in transgenic tobacco. *Proceedings of the National Academy of Sciences of the United States of America, 89*(23), 11184-11188.

Cargile, B. J., McLuckey, S. A., & Stephenson, J. L. (2001). Identification of bacteriophage MS2 coat protein from E. coli lysates via ion trap collisional activation of intact protein ions. *Analytical Chemistry, 73*(6), 1277-1285. doi:10.1021/ac000725l

Carlsson, A. S., LaBrie, S. T., Kinney, A. J., Von Wettstein-Knowles, P., & Browse, J. (2002). A KAS2 cDNA complements the phenotypes of the Arabidopsis fab1 mutant that differs in a single residue bordering the substrate binding pocket. *The Plant Journal, 29*(6), 761-770. doi:10.1046/j.1365-313X.2002.01253.x

Carpentier, S. C., Panis, B., Vertommen, A., Swennen, R., Sergeant, K., Renaut, J., . . . Devreese, B. (2008). Proteome analysis of non-model plants: A challenging but powerful approach. *Mass Spectrometry Reviews, 27*(4), 354-377. doi:10.1002/mas.20170

Carpentier, S. C., Witters, E., Laukens, K., Deckers, P., Swennen, R., & Panis, B. (2005). Preparation of protein extracts from recalcitrant plant tissues: An evaluation of different methods for two-dimensional gel electrophoresis analysis. *PROTEOMICS, 5*(10), 2497-2507. doi:10.1002/pmic.200401222

Carr, P. A., & Church, G. M. (2009). Genome engineering. *Nature Biotechnology, 27*(12), 1151-1162. doi:10.1038/nbt.1590

Casati, P., Zhang, X., Burlingame, A. L., & Walbot, V. (2005). Analysis of leaf proteome after UV-B irradiation in maize lines giffering in sensitivity. *Molecular & Cellular Proteomics, 4*(11), 1673-1685. doi:10.1074/mcp.M500173-MCP200

Chang, W. W. P., Huang, L., Shen, M., Webster, C., Burlingame, A. L., & Roberts, J. K. M. (2000). Patterns of protein synthesis and tolerance of anoxia in root tips of maize seedlings acclimated to a low-oxygen environment, and identification of proteins by mass spectrometry. *Plant Physiology, 122*(2), 295-318. doi:10.1104/pp.122.2.295

Chatterjee, M., Gupta, S., Bhar, A., & Das, S. (2012). Optimization of an efficient protein extraction protocol compatible with two-dimensional electrophoresis and mass spectrometry from recalcitrant phenolic rich roots of Chickpea (Cicer arietinum L.). *International Journal of Proteomics, 2012*, 1-10. doi:10.1155/2012/536963

Chen, M., Mooney, B. P., Hajduch, M., Joshi, T., Zhou, M., Xu, D., & Thelen, J. J. (2009). System analysis of an Arabidopsis mutant altered in de novo fatty acid synthesis reveals diverse changes in seed composition and metabolism. *Plant Physiology, 150*(1), 27-41. doi:10.1104/pp.108.134882

Chen, Y., Chen, W., Cobb, M. H., & Zhao, Y. (2009). PTMap - A sequence alignment software for unrestricted, accurate, and full-spectrum identification of post-translational modification sites. *Proceedings of the National Academy of Sciences of the United States of America, 106*(3), 761-766. doi:10.1073/pnas.0811739106

Chitteti, B. R., & Peng, Z. (2007). Proteome and phosphoproteome dynamic change during cell dedifferentiation in Arabidopsis. *PROTEOMICS, 7*(9), 1473-1500. doi:10.1002/pmic.200600871

Chivasa, S., Hamilton, J. M., Pringle, R. S., Ndimba, B. K., Simon, W. J., Lindsey, K., & Slabas, A. R. (2006). Proteomic analysis of differentially expressed proteins in fungal elicitor-treated Arabidopsis cell cultures. *Journal of Experimental Botany, 57*(7), 1553-1562. doi:10.1093/jxb/erj149

Christie, W. W. (1993). Preparation of lipid extracts from tissues. In W. W. Christie (Ed.), *Advances in lipid methodology-Two* (pp. 195-213). Dundee: Oily Press.

Cohen, P. (2001). The role of protein phosphorylation in human health and disease. *European Journal of Biochemistry, 268*(19), 5001-5010. doi:10.1046/j.0014-2956.2001.02473.x

Coon, J. J., Ueberheide, B., Syka, J. E. P., Dryhurst, D. D., Ausio, J., Shabanowitz, J., & Hunt, D. F. (2005). Protein identification using sequential ion/ion reactions and tandem mass

spectrometry. *Proceedings of the National Academy of Sciences of the United States of America, 102*(27), 9463-9468. doi:10.1073/pnas.0503189102

Corley, R. H. V., & Tinker, P. B. (2003). The products of the oil palm and their extraction. In R. H. V. Corley & P. B. Tinker (Eds.), *The oil palm*. Oxford, United Kingdom: Blackwell Science Ltd.

Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., & Sanchez, J.-C. (2000). The dynamic range of protein expression: A challenge for proteomic research. *ELECTROPHORESIS, 21*(6), 1104-1115. doi:10.1002/(sici)1522-2683(20000401)21:6<1104::aid-elps1104>3.0.co;2-c

Cox, D. M., Zhong, F., Du, M., Duchoslav, E., Sakuma, T., & McDermott, J. C. (2005). Multiple reaction monitoring as a method for identifying protein posttranslational modifications. *Journal of Biomolecular Techniques, 16*(2), 83-90.

Cramer, G. R., Van Sluyter, S. C., Hopper, D. W., Pascovici, D., Keighley, T., & Haynes, P. A. (2013). Proteomic analysis indicates massive changes in metabolism prior to the inhibition of growth and photosynthesis of grapevine (Vitis vinifera L.) in response to water deficit. *BMC Plant Biol, 13*(1), 49. doi:10.1186/1471-2229-13-49

Cronan Jr, J. E., & Waldrop, G. L. (2002). Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research, 41*(5), 407-435. doi:10.1016/s0163-7827(02)00007-3

Cronquist, A. (1981). *An integrated system of classification of flowering plants*. Columbia: University Press.

Daher, Z., Recorbet, G., Valot, B., Robert, F., Balliau, T., Potin, S., . . . Dumas-Gaudot, E. (2010). Proteomic analysis of Medicago truncatula root plastids. *PROTEOMICS, 10*(11), 2123-2137. doi:10.1002/pmic.200900345

Damerval, C., De Vienne, D., Zivy, M., & Thiellement, H. (1986). Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *ELECTROPHORESIS, 7*(1), 52-54. doi:10.1002/elps.1150070108

de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Frohlich, F., . . . Mann, M. (2008). Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature, 455*(7217), 1251-1254. doi:10.1038/nature07341

De La Fuente, M., Borrajo, A., Bermudez, J., Lores, M., Alonso, J., Lopez, M., . . . Alvarez, G. (2011). 2-DE-based proteomic analysis of common bean (Phaseolus vulgaris L.) seeds [Research Support, Non-U.S. Gov't]. *Journal of Proteomics, 74*(2), 262-267. doi:10.1016/j.jprot.2010.10.004

DeGnore, J. P., & Qin, J. (1998). Fragmentation of phosphopeptides in an ion trap mass spectrometer. *American Society For Mass Spectrometry, 9*(11), 1175-1188. doi:10.1016/S1044-0305(98)00088-9

Dehesh, K., Tai, H., Edwards, P., Byrne, J., & Jaworski, J. G. (2001). Overexpression of 3-ketoacyl-acyl-carrier protein synthase IIIs in plants reduces the rate of lipid synthesis [Research Support, U.S. Gov't, Non-P.H.S.]. *Plant Physiology, 125*(2), 1103-1114. doi:10.1104/pp.125.2.1103

Depagne, J., & Chevalier, F. (2012). Technical updates to basic proteins focalization using IPG strips. *Proteome Science, 10*(1), 54. doi:10.1186/1477-5956-10-54

Di Domenico, F., Sultana, R., Barone, E., Perluigi, M., Cini, C., Mancuso, C., . . . Butterfield, D. A. (2011). Quantitative proteomics analysis of phosphorylated proteins in the hippocampus of Alzheimer's disease subjects. *Journal of Proteomics, 74*(7), 1091-1103. doi:10.1016/j.jprot.2011.03.033

Dinkel, H., Chica, C., Via, A., Gould, C. M., Jensen, L. J., Gibson, T. J., & Diella, F. (2011). Phospho.ELM: a database of phosphorylation sites—update 2011. *Nucleic Acids Research, 39*(suppl 1), D261-D267. doi:10.1093/nar/gkq1104

Division, E. a. I. D. (n.d.). *Overview of the Malaysian oil palm industry 2013*. Retrieved 4 June, 2013,from http://bepi.mpob.gov.my/images/overview/Overview_of_Industry_2012.pdfa.

Domanski, D., Murphy, L. C., & Borchers, C. H. (2010). Assay development for the determination of phosphorylation stoichiometry using multiple reaction monitoring methods with and without phosphatase treatment: application to breast cancer signaling pathways. *Anal Chem, 82*(13), 5610-5620. doi:10.1021/ac1005553

Domanskii, V., Rassadina, V., Gus-Mayer, S., Wanner, G., Schoch, S., & Rudiger, W. (2003). Characterization of two phases of chlorophyll formation during greening of etiolated barley leaves. *Planta, 216*(3), 475-483. doi:10.1007/s00425-002-0870-x

Dong, J., Zhou, H., Wu, R., Ye, M., & Zou, H. (2007). Specific capture of phosphopeptides by Zr4+-modified monolithic capillary column. *Journal of Separation Science, 30*(17), 2917-2923. doi:10.1002/jssc.200700350

Dormann, P., Voelker, T. A., & Ohlrogge, J. B. (2000). Accumulation of palmitate in Arabidopsis mediated by the acyl-acyl carrier protein thioesterase FATB1 [Research Support, Non-U.S. Gov't]. *Plant Physiology, 123*(2), 637-644.

Dreger, M. (2003). Proteome analysis at the level of subcellular structures. *European Journal of Biochemistry, 270*(4), 589-599. doi:10.1046/j.1432-1033.2003.03426.x

Du, H., Liang, Y., Pei, K., & Ma, K. (2011). UV radiation-responsive proteins in rice leaves: a proteomic analysis [Research Support, Non-U.S. Gov't]. *Plant and Cell Physiology, 52*(2), 306-316. doi:10.1093/pcp/pcq186

Du, Y., Parks, B. A., Sohn, S., Kwast, K. E., & Kelleher, N. L. (2006). Top-down approaches for measuring expression ratios of intact yeast proteins using Fourier transform mass spectrometry. *Analytical Chemistry, 78*(3), 686-694. doi:10.1021/ac050993p

Duffield, J., Shapouri, H., Graboski, M., McCormick, R., & WIlson, R. (1998). U.S. Biodiesel development: New markets for conventional and genetically modified agricultural products. In U. S. D. o. Agriculture (Ed.), *Economic Research Service*. Washington D.C.

Dunkley, T. P. J., Dupree, P., Watson, R. B., & Lilley, K. S. (2004). The use of isotope-coded affinity tags (ICAT) to study organelle proteomes in *Arabidopsis thaliana*. *Biochemical Society Transactions, 32*, 520-523.

Dunn, J. D., Reid, G. E., & Bruening, M. L. (2010). Techniques for phosphopeptide enrichment prior to analysis by mass spectrometry. *Mass Spectrometry Reviews, 29*(1), 29-54. doi:10.1002/mas.20219

Dupont, F. (2008). Metabolic pathways of the wheat (Triticum aestivum) endosperm amyloplast revealed by proteomics. *BMC Plant Biology, 8*(1), 39. doi:10.1186/1471-2229-8-39

Dussert, S., Guerin, C., Andersson, M., Joët, T., Tranbarger, T. J., Pizot, M., . . . Morcillo, F. (2013). Comparative transcriptome analysis of three oil palm fruit and seed tissues that differ in oil content and fatty acid composition. *Plant Physiology, 162*(3), 1337-1358. doi:10.1104/pp.113.220525

Echeverria, E., Boyer, C., Liu, K.-C., & Shannon, J. (1985). Isolation of amyloplasts from developing maize endosperm. *Plant Physiology, 77*(3), 513-519. doi:10.1104/pp.77.3.513

Edman, P. (1964). Determination of amino acid sequences in proteins. *Trombosis et Diathesis Haemorrhagica Supplementum, 13*, 17-20.

Emanuelsson, O., Brunak, S., von Heijne, G., & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools [10.1038/nprot.2007.131]. *Nature Protocols, 2*(4), 953-971.

Emanuelsson, O., Nielsen, H., Brunak, S., & von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology, 300*(4), 1005-1016. doi:10.1006/jmbi.2000.3903

Emes, M. J. (2009). Oxidation of methionine residues: the missing link between stress and signalling responses in plants. *Biochemical Journal, 422*, e1-e2. doi:10.1042/BJ20091063

Endler, A., & Baginsky, S. (2011). Use of phosphoproteomics to study posttranslational protein modifications in Arabidopsis chloroplasts. In R. P. Jarvis (Ed.), *Chloroplast Research in Arabidopsis* (Vol. 775, pp. 283-296): Humana Press. Retrieved from http://dx.doi.org/10.1007/978-1-61779-237-3_15. doi:10.1007/978-1-61779-237-3_15

Eng, J. K., McCormack, A. L., & Yates III, J. R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry, 5*(11), 976-989. doi:10.1016/1044-0305(94)80016-2

England, J. L., & Haran, G. (2011). Role of solvation efects in protein denaturation: From thermodynamics to single molecules and back. *Annual Review of Physical Chemistry, 62*(1), 257-277. doi:10.1146/annurev-physchem-032210-103531

Esnan, A. G., Zakaria, Z. Z., & Wahid, M. B. (2004). *A guide-oil palm industry*. Selangor: Malaysia: Millennium Edn.

Fan, J., Chen, C., Yu, Q., Brlansky, R. H., Li, Z.-G., & Gmitter, F. G. (2011). Comparative iTRAQ proteome and transcriptome analyses of sweet orange infected by "Candidatus Liberibacter asiaticus". *Physiologia Plantarum, 143*(3), 235-245. doi:10.1111/j.1399-3054.2011.01502.x

Fan, J., Mohareb, F., Jones, A. M., & Bessant, C. (2012). MRMaid: The SRM Assay Design Tool for Arabidopsis and Other Species [Methods]. *Frontiers in Plant Science, 3*, 164. doi:10.3389/fpls.2012.00164

Fan, P., Wang, X., Kuang, T., & Li, Y. (2009). An efficient method for the extraction of chloroplast proteins compatible for 2-DE and MS analysis. *ELECTROPHORESIS, 30*(17), 3024-3033. doi:10.1002/elps.200900172

Fell, D. A. (1992). Metabolic control analysis: a survey of its theoretical and experimental development [Review]. *Biochem Journal, 286 (Pt 2)*(Pt 2), 313-330. doi:Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1132899/?tool=pubmed

Feng, S., Ye, M., Zhou, H., Jiang, X., Jiang, X., Zou, H., & Gong, B. (2007). Immobilized zirconium ion affinity chromatography for specific enrichment of phosphopeptides in phosphoproteome analysis. *Molecular & Cellular Proteomics, 6*(9), 1656-1665. doi:10.1074/mcp.T600071-MCP200

Ferro, M., Brugière, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., . . . Rolland, N. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Molecular & Cellular Proteomics, 9*(6), 1063-1084. doi:10.1074/mcp.M900325-MCP200

Fitch, W. M. (1970). Distinguishing homologous from analogous proteins. *Systematic Zoology, 19*(2), 99-113. doi:10.2307/2412448

Fournier, M. L., Gilmore, J. M., Martin-Brown, S. A., & Washburn, M. P. (2007). Multidimensional separations-based shotgun proteomics. *Chemical reviews, 107*(8), 3654-3686. doi:10.1021/cr068279a

Friso, G., Majeran, W., Huang, M., Sun, Q., & van Wijk, K. J. (2010). Reconstruction of metabolic pathways, protein expression, and homeostasis machineries across maize bundle sheath and mesophyll chloroplasts: large-scale quantitative proteomics using the first maize genome assembly [Research Support, U.S. Gov't, Non-P.H.S.]. *Plant Physiology, 152*(3), 1219-1250. doi:10.1104/pp.109.152694

Froehlich, J. E., Wilkerson, C. G., Ray, W. K., McAndrew, R. S., Osteryoung, K. W., Gage, D. A., & Phinney, B. S. (2003). Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *Journal of Proteome Research, 2*(4), 413-425. doi:10.1021/pr034025j

Fukao, Y., Ferjani, A., Fujiwara, M., Nishimori, Y., & Ohtsu, I. (2009). Identification of Zinc-Responsive Proteins in the Roots of Arabidopsis thaliana Using a Highly Improved Method of Two-Dimensional Electrophoresis. *Plant and Cell Physiology, 50*(12), 2234-2239. doi:10.1093/pcp/pcp154

Gabaldón, T. (2007). Evolution of proteins and proteomes: a phylogenetics approach. *Evolutionary Bioinformatics 1*(1). doi:10.4137/ebo.s0

Gammulla, C. G., Pascovici, D., Atwell, B. J., & Haynes, P. A. (2010). Differential metabolic response of cultured rice (Oryza sativa) cells exposed to high- and low-temperature stress. *PROTEOMICS, 10*(16), 3001-3019. doi:10.1002/pmic.201000054

Gerber, I. B., Laukens, K., De Vijlder, T., Witters, E., & Dubery, I. A. (2008). Proteomic profiling of cellular targets of lipopolysaccharide-induced signalling in Nicotiana tabacum BY-2 cells [Research Support, Non-U.S. Gov't]. *Biochimica et Biophysica Acta, 1784*(11), 1750-1762. doi:10.1016/j.bbapap.2008.06.012

Gerber, I. B., Laukens, K., Witters, E., & Dubery, I. A. (2006). Lipopolysaccharide-responsive phosphoproteins in Nicotiana tabacum cells. *Plant Physiology and Biochemistry, 44*(5–6), 369-379. doi:10.1016/j.plaphy.2006.06.015

Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., & Gygi, S. P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences, 100*(12), 6940-6945. doi:10.1073/pnas.0832254100

Gingras, A.-C., Aebersold, R., & Raught, B. (2005). Advances in protein complex analysis using mass spectrometry. *The Journal of Physiology, 563*(1), 11-21. doi:10.1113/jphysiol.2004.080440

Givan, C. V. (1983). The source of acetyl coenzyme A in chloroplasts of higher plants. *Physiologia Plantarum, 57*(2), 311-316. doi:10.1111/j.1399-3054.1983.tb00917.x

Gleason, H. A. (1926). The individualistic concept of the plant association. *53*(1), 7-26.

Gnad, F., Gunawardena, J., & Mann, M. (2011). PHOSIDA 2011: the posttranslational modification database. *Nucleic Acids Research, 39*(suppl 1), D253-D260. doi:10.1093/nar/gkq1159

Gómez-Vidal, S., Salinas, J., Tena, M., & Lopez-Llorca, L. V. (2009). Proteomic analysis of date palm (Phoenix dactylifera L.) responses to endophytic colonization by entomopathogenic fungi. *ELECTROPHORESIS, 30*(17), 2996-3005. doi:10.1002/elps.200900192

Gómez, A., López, J. A., Pintos, B., Camafeita, E., & Bueno, M. Á. (2009). Proteomic analysis from haploid and diploid embryos of Quercus suber L. identifies qualitative and quantitative differential expression patterns. *PROTEOMICS, 9*(18), 4355-4367. doi:10.1002/pmic.200900179

Gooding, P. S., Bird, C., & Robinson, S. P. (2001). Molecular cloning and characterisation of banana fruit polyphenol oxidase. *PLANTA, 213*(5), 748-757. doi:10.1007/s004250100553

Görg, A., Obermaier, C., Boguth, G., Csordas, A., Diaz, J.-J., & Madjar, J.-J. (1997). Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins. *ELECTROPHORESIS, 18*(3-4), 328-337. doi:10.1002/elps.1150180306

Görg, A., Weiss, W., & Dunn, M. J. (2004). Current two-dimensional electrophoresis technology for proteomics. *PROTEOMICS, 4*(12), 3665-3685. doi:10.1002/pmic.200401031

Goulet, C., Goulet, M. C., & Michaud, D. (2010). 2-DE proteome maps for the leaf apoplast of Nicotiana benthamiana [Research Support, Non-U.S. Gov't]. *PROTEOMICS, 10*(13), 2536-2544. doi:10.1002/pmic.200900382

Graef, G. L., Miller, L. A., Fehr, W. R., & Hammond, E. G. (1985). Fatty acid development in a soybean mutant with high stearic acid. *Journal of the American Oil Chemists' Society, 62*(4), 773-775. doi:10.1007/bf03028752

Graham, R. L., Sharma, M. K., Ternan, N. G., Weatherly, D. B., Tarleton, R. L., & McMullan, G. (2007). A semi-quantitative GeLC-MS analysis of temporal proteome expression in the emerging nosocomial pathogen Ochrobactrum anthropi. *Genome Biology, 8*(6), R110. doi:10.1186/gb-2007-8-6-r110

Graves, J. D., & Krebs, E. G. (1999). Protein phosphorylation and signal transduction [Review]. *Pharmacology and Therapeutics, 82*(2-3), 111-121. doi:10.1016/s0163-7258(98)00056-4

Grimsrud, P. A., den Os, D., Wenger, C., D.,, Swaney, D. L., Schwartz, D., Sussman, M., R.,, . . . Coon, J. J. (2010). Large-scale phosphorylation analysis in *Medicago truncatula* roots provides insight into in vivo kinase activity in legumes. *Plant Physiology, 152*(1), 19-28. doi:10.1104/pp.109.149625

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., & Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology, 17*(10), 994-999. doi:10.1038/13690

Hagglund, P., Bunkenborg, J., Yang, F., Harder, L. M., Finnie, C., & Svensson, B. (2010). Identification of thioredoxin target disulfides in proteins released from barley aleurone layers [Research Support, Non-U.S. Gov't]. *Journal of Proteomics, 73*(6), 1133-1136. doi:10.1016/j.jprot.2010.01.007

Hai, T. C. (2002). *The palm oil industry in Malaysia: From seed to frying pan*. Selangor:Malaysia: WWF.

Hajduch, M., Casteel, J. E., Hurrelmeyer, K. E., Song, Z., Agrawal, G. K., & Thelen, J. J. (2006). Proteomic analysis of seed filling in *Brassica napus*. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. *Plant Physiology, 141*(1), 32-46. doi:10.1104/pp.105.075390

Hajheidari, M., Farrona, S., Huettel, B., Koncz, Z., & Koncz, C. (2012). CDKF;1 and CDKD protein kinases regulate phosphorylation of serine residues in the C-terminal domain of Arabidopsis RNA polymerase II. *Plant Cell, 24*(4), 1626-1642. doi:10.1105/tpc.112.096834

Halligan, B. D., Ruotti, V., Jin, W., Laffoon, S., Twigger, S. N., & Dratz, E. A. (2004). ProMoST (Protein Modification Screening Tool): a web-based tool for mapping protein modifications on two-dimensional gels. *Nucleic Acids Research, 32*(Web Server issue), W638-644. doi:10.1093/nar/gkh356

Hansen, L. U., & Chiu, M.-C. M. (2005). Isolation and composition of chromoplasts from tomatoes. *Journal of Agricultural and Food Chemistry, 53*(17), 6678-6682. doi:10.1021/jf058013p

Hardie, D. G. (1992). Regulation of fatty acid and cholesterol metabolism by the AMP-activated protein kinase. *Biochimica et Biophysica Acta, 1123*(3), 231-238. doi:10.1016/0005-2760(92)90001

Hardie, D. G., & Pan, D. A. (2002). Regulation of fatty acid synthesis and oxidation by the AMP-activated protein kinase. *Biochemical Society Transactions, 30*(Pt 6), 1064-1070. doi:10.1042/

Harholt, J., Suttangkakul, A., & Vibe Scheller, H. (2010). Biosynthesis of pectin. *Plant Physiology, 153*(2), 384-395. doi:10.1104/pp.110.156588

Harsanyi, A., Boddi, B., Boka, K., & Gaborjanyi, R. (2005). Pathogen affected greening process of barley seedlings infected with BSMV by seed transmission. *Cereal Research Communications, 33*(209-212). doi:Retrieved from http://nda.sztaki.hu/kereso/index.php?a=set&id=439

Harwood, J. L. (1996). Recent advances in the biosynthesis of plant fatty acids. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism, 1301*(1–2), 7-56. doi:10.1016/0005-2760(95)00242-1

Hayashi, T., Yoshida, K., Woo Park, Y., Konishi, T., & Baba, K. i. (2005). Cellulose metabolism in plants. In W. J. Kwang (Ed.), *International Review of Cytology* (Vol. Volume 247, pp. 1-34): Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/S0074769605470011. doi:10.1016/s0074-7696(05)47001-1

Hayes, K. C., & Pronczuk, A. (2010). Replacing trans fat: the argument for palm oil with a cautionary note on interesterification [Review]. *Journal of the Americal College of Nutrition, 29*(3 Suppl), 253S-284S.

He, C.-F., & Wang, Y.-M. (2008). Protein extraction from leaves of *Aloe vera* L., a succulent and recalcitrant plant, for proteomic analysis. *Plant Molecular Biology Reporter, 26*(4), 292-300. doi:10.1007/s11105-008-0040-9

He, R., Kim, M. J., Nelson, W., Balbuena, T. S., Kim, R., Kramer, R., . . . Gang, D. R. (2012). Next-generation sequencing-based transcriptomic and proteomic analysis of the common reed, Phragmites australis (Poaceae), reveals genes involved in invasiveness

and rhizome specificity. *Americal Journal of Botany, 99*(2), 232-247. doi:10.3732/ajb.1100429

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., & Millar, A. H. (2007). SUBA: the Arabidopsis subcellular database. *Nucleic Acids Research, 35*(suppl 1), D213-D218. doi:10.1093/nar/gkl863

Heintz, D., Wurtz, V., High, A. A., Van Dorsselaer, A., Reski, R., & Sarnighausen, E. (2004). An efficient protocol for the identification of protein phosphorylation in a seedless plant, sensitive enough to detect members of signalling cascades. *ELECTROPHORESIS, 25*(7-8), 1149-1159. doi:10.1002/elps.200305795

Heinz, E., & Roughan, P. G. (1983). Similarities and differences in lipid metabolism of chloroplasts isolated from 18:3 and 16:3 plants. *Plant Physiology, 72*(2), 273-279. doi:10.1104/pp.72.2.273

Hernández-Pinzón, I., Ross, J. H. E., Barnes, K. A., Damant, A. P., & Murphy, D. J. (1999). Composition and role of tapetal lipid bodies in the biogenesis of the pollen coat of Brassica napus. *Planta, 208*(4), 588-598. doi:10.1007/s004250050597

Hey, J., Posch, A., Cohen, A., Liu, N., & Harbers, A. (2008). Fractionation of complex protein mixtures by liquid-phase isoelectric focusing. In A. Posch (Ed.), *2D PAGE: Sample Preparation and Fractionation* (Vol. 424, pp. 225-239): Humana Press. Retrieved from http://dx.doi.org/10.1007/978-1-60327-064-9_19. doi:10.1007/978-1-60327-064-9_19

Hjerrild, M., Stensballe, A., Rasmussen, T. E., Kofoed, C. B., Blom, N., Sicheritz-Ponten, T., . . . Gammeltoft, S. (2004). Identification of phosphorylation sites in protein kinase A substrates using artificial neural networks and mass spectrometry. *Journal of Proteome Research, 3*(3), 426-433. doi:10.1021/pr0341033

Hobbs, D. H., Flintham, J. E., & Hills, M. J. (2004). Genetic control of storage oil synthesis in seeds of Arabidopsis. *Plant Physiology, 136*(2), 3341-3349. doi:10.1104/pp.104.049486

Hochstrasser, D. F., Harrington, M. G., Hochstrasser, A.-C., Miller, M. J., & Merril, C. R. (1988). Methods for increasing the resolution of two-dimensional protein electrophoresis. *Analytical Biochemistry, 173*(2), 424-435. doi:10.1016/0003-2697(88)90209-6

Hoff, K. J. (2009). The effect of sequencing errors on metagenomic gene prediction [Research Support, Non-U.S. Gov't]. *BMC Genomics, 10*(1), 520. doi:10.1186/1471-2164-10-520

Horn, D. M., Zubarev, R. A., & McLafferty, F. W. (2000). Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry. *Proceedings of the National Academy of Sciences, 97*(19), 10313-10317. doi:10.1073/pnas.97.19.10313

Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., . . . Sullivan, M. (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications

in man and mouse. *Nucleic Acids Research, 40*(Database issue), D261-270. doi:10.1093/nar/gkr1122

Horth, P., Miller, C. A., Preckel, T., & Wenz, C. (2006). Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. *Molecular & Cellular Proteomics, 5*(10), 1968-1974. doi:10.1074/mcp.T600037-MCP200

Hsu, J.-L., Wang, L.-Y., Wang, S.-Y., Lin, C.-H., Ho, K.-C., Shi, F.-K., & Chang, I.-F. (2009). Functional phosphoproteomic profiling of phosphorylation sites in membrane fractions of salt-stressed Arabidopsis thaliana. *Proteome Science, 7*(1), 42. doi:10.1186/1477-5956-7-42

Hughes, W. L., & Sinex, F. M. (1954). Chemistry of the proteins, peptides, and amino acids. *Annual Review of Biochemistry, 23*(1), 177-214. doi:10.1146/annurev.bi.23.070154.001141

Hurkman, W. J., & Tanaka, C. K. (1986). Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiology, 81*(3), 802-806. doi:10.1104/pp.81.3.802

Hwang, I. W., Makishima, Y., Kato, T., Park, S., Terzic, A., & Park, E. Y. (2014). Human acetyl-CoA carboxylase 2 expressed in silkworm Bombyx mori exhibits posttranslational biotinylation and phosphorylation. *Appl Microbiol Biotechnol, 98*(19), 8201-8209. doi:10.1007/s00253-014-5715-6

Ikeguchi, Y., & Nakamura, H. (1997). Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line preconcentration on titania. *Analytical Sciences, 13*(3), 479-483. doi:10.2116/analsci.13.479

Ikeguchi, Y., & Nakamura, H. (2000). Selective enrichment of phospholipids by titania. *Analytical Sciences, 16*(5), 541-543. doi:10.2116/analsci.16.541

Jacquin, N. J. (1763). Selectarum stirpium Americanarum historia. In *Ex Officina Krausiana*. Vindobonae:Vienna

Jain, R., Katavic, V., Agrawal, G. K., Guzov, V. M., & Thelen, J. J. (2008). Purification and proteomic characterization of plastids from Brassica napus developing embryos. *PROTEOMICS, 8*(16), 3397-3405. doi:10.1002/pmic.200700810

James, A., & Jorgensen, C. (2010). Basic Design of MRM Assays for Peptide Quantification. In P. R. Cutillas & J. F. Timms (Eds.), *LC-MS/MS in Proteomics* (Vol. 658, pp. 167-185): Humana Press. Retrieved from http://dx.doi.org/10.1007/978-1-60761-780-8_10. doi:10.1007/978-1-60761-780-8_10

James, M. G., Denyer, K., & Myers, A. M. (2003). Starch synthesis in the cereal endosperm. *Current Opinion in Plant Biology, 6*(3), 215-222. doi:10.1016/s1369-5266(03)00042-6

Jensen, O. N. (2004). Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol, 8*(1), 33-41. doi:10.1016/j.cbpa.2003.12.009

Jensen, O. N. (2006). Interpreting the protein language using proteomics. *Nature Reviews Molecular Cell Biology, 7*(6), 391-403. doi:10.1038/nrm1939

Jessome, L. L., & Volmer, D. A. (2006). Ion suppression: A major concern in mass spectrometry. *LCGC North Americca, 24*(5), 498-510.

Johnson, S. A., & Hunter, T. (2005). Kinomics: methods for deciphering the kinome. *Nature Methods, 2*(1), 17-25. doi:10.1038/nmeth731

Jones, A. L., Gane, A. M., Herbert, D., Willey, D. L., Rutter, A. J., Kille, P., . . . Harwood, J. L. (2003). Beta-ketoacyl-acyl carrier protein synthase III from pea (Pisum sativum L.): properties, inhibition by a novel thiolactomycin analogue and isolation of a cDNA clone encoding the enzyme [Research Support, Non-U.S. Gov't]. *PLANTA, 216*(5), 752-761. doi:10.1007/s00425-002-0925-z

Jones, A. M. E., Bennett, M. H., Mansfield, J. W., & Grant, M. (2006). Analysis of the defence phosphoproteome of Arabidopsis thaliana using differential mass tagging. *PROTEOMICS, 6*(14), 4155-4165. doi:10.1002/pmic.200500172

Joyard, J., & Douce, R. (1977). Site of synthesis of phosphatidic acid and diacyglycerol in spinach chloroplasts. *Biochimica et Biophysica Acta 486*(2), 273-285. doi:10.1016/0005-2760(77)90023-6

Joyard, J., Ferro, M., Masselon, C., Seigneurin-Berny, D., Salvi, D., Garin, J., & Rolland, N. (2010). Chloroplast proteomics highlights the subcellular compartmentation of lipid metabolism. *Progress in Lipid Research, 49*(2), 128-158. doi:10.1016/j.plipres.2009.10.003

Juneau, P., Le Lay, P., Böddi, B., Samson, G., & Popovic, R. (2002). Relationship between the structural and functional changes of the photosynthetic apparatus during chloroplast–chromoplast transition in flower bud of *Lilium longiflorum*. *Photochemistry and Photobiology, 75*(4), 377-381. doi:10.1562/0031-8655(2002)075<0377:rbtsaf>2.0.co;2

Jungblut, P., Thiede, B., Zimny-Arndt, U., Muller, E. C., Scheler, C., Wittmann-Liebold, B., & Otto, A. (1996). Resolution power of two-dimensional electrophoresis and identification of proteins from gels [Review]. *ELECTROPHORESIS, 17*(5), 839-847. doi:10.1002/elps.1150170505

Kachroo, A., Lapchyk, L., Fukushige, H., Hildebrand, D., Klessig, D., & Kachroo, P. (2003). Plastidial fatty acid signaling modulates salicylic acid- and jasmonic acid-mediated defense pathways in the Arabidopsis ssi2 mutant [R]. *Plant Cell, 15*(12), 2952-2965. doi:10.1105/tpc.017301

Kachroo, A., Shanklin, J., Whittle, E., Lapchyk, L., Hildebrand, D., & Kachroo, P. (2007). The *Arabidopsis* stearoyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. *Plant Molecular Biology, 63*(2), 257-271. doi:10.1007/s11103-006-9086-y

Kaffarnik, F. A. R., Jones, A. M. E., Rathjen, J. P., & Peck, S. C. (2009). Effector proteins of the bacterial pathogen *Pseudomonas syringae* alter the extracellular proteome of the host plant, *Arabidopsis thaliana*. *Molecular & Cellular Proteomics, 8*(1), 145-156. doi:10.1074/mcp.M800043-MCP200

Kamiyoshihara, Y., Tieman, D. M., Huber, D. J., & Klee, H. J. (2012). Ligand-induced alterations in the phosphorylation state of ethylene receptors in tomato fruit [Research Support, U.S. Gov't, Non-P.H.S.]. *Plant Physiology, 160*(1), 488-497. doi:10.1104/pp.112.202820

Kang, F., Ridout, C. J., Morgan, C. L., & Rawsthorne, S. (1994). The activity of acetyl-CoA carboxylase is not correlated with the rate of lipid synthesis during development of oilseed rape (*Brassica napus* L.) embryos. *Planta, 193*(3), 320-325. doi:10.1007/bf00201808

Kaufmann, H., Bailey, J. E., & Fussenegger, M. (2001). Use of antibodies for detection of phosphorylated proteins separated by two-dimensional gel electrophoresis. *Proteomics, 1*(2), 194-199. doi:Doi 10.1002/1615-9861(200102)1:2<194::Aid-Prot194>3.3.Co;2-B

Keeling, P. J. (2004). Diversity and evolutionary history of plastids and their hosts. *American Journal of Botany, 91*(10), 1481-1493. doi:10.3732/ajb.91.10.1481

Kelleher, N. L., Lin, H. Y., Valaskovic, G. A., Aaserud, D. J., Fridriksson, E. K., & McLafferty, F. W. (1999). Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *Journal of the American Chemical Society, 121*(4), 806-812. doi:10.1021/ja973655h

Kelleher, N. L., Zubarev, R. A., Bush, K., Furie, B., Furie, B. C., McLafferty, F. W., & Walsh, C. T. (1999). Localization of labile posttranslational modifications by electron capture dissociation: the case of gamma-carboxyglutamic acid [Comparative Study

Research Support, U.S. Gov't, P.H.S.]. *Analytical Biochemistry, 71*(19), 4250-4253. doi:10.1021/ac990684x

Kellie, J. F., Tran, J. C., Lee, J. E., Ahlf, D. R., Thomas, H. M., Ntai, I., . . . Kelleher, N. L. (2010). The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. *Molecular BioSystems, 6*(9), 1532-1539. doi:10.1039/c000896f

Khoury, G. A., Baliban, R. C., & Floudas, C. A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database [10.1038/srep00090]. *Scientific Reports, 1*. doi:10.1038/srep00090

Kieffer, P., Dommes, J., Hoffmann, L., Hausman, J.-F., & Renaut, J. (2008). Quantitative changes in protein expression of cadmium-exposed poplar plants. *PROTEOMICS, 8*(12), 2514-2530. doi:10.1002/pmic.200701110

Kim, J. H., Lee, J., Oh, B., Kimm, K., & Koh, I. (2004). Prediction of phosphorylation sites using SVMs. *Bioinformatics, 20*(17), 3179-3184. doi:10.1093/bioinformatics/bth382

Kinney, A. J. (1994). Genetic modification of the storage lipids of plants. *Current Opinion in Biotechnology, 5*(2), 144-151. doi:10.1016/s0958-1669(05)80027-8

Kinney, A. J. (1996). Development of genetically engineered soybean oils for food applications. *Journal of Food Lipids, 3*(4), 273-292. doi:10.1111/j.1745-4522.1996.tb00074.x

Kirby, K. S. (1956). A new method for the isolation of ribonucleic acids from mammalian tissues. *Biochemical Journal, 64*(3), 405-408.

Kirkland, R. (2011). *The role of palm oil in a sustainable dairy industry*. Retrieved 4 June, 2013,from http://www.volac.com/news/agriculture-news/news236/palm-oil-in-a-sustainable-world.

Klaus, D., Ohlrogge, J. B., Neuhaus, H. E., & Dormann, P. (2004). Increased fatty acid production in potato by engineering of acetyl-CoA carboxylase [Research Support, Non-U.S. Gov't]. *PLANTA, 219*(3), 389-396. doi:10.1007/s00425-004-1236-3

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., & Baginsky, S. (2004). The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. *Curr Biol, 14*(5), 354-362. doi:10.1016/j.cub.2004.02.039

Kleffmann, T., von Zychlinski, A., Russenberger, D., Hirsch-Hoffmann, M., Gehrig, P., Gruissem, W., & Baginsky, S. (2007). Proteome dynamics during plastid differentiation in rice. *Plant Physiology, 143*(2), 912-923. doi:10.1104/pp.106.090738

Knutzon, D. S., Thompson, G. A., Radke, S. E., Johnson, W. B., Knauf, V. C., & Kridl, J. C. (1992). Modification of Brassica seed oil by antisense expression of a stearoyl-acyl carrier protein desaturase gene. *Proceedings of the National Academy of Sciences of the United States of America, 89*(7), 2624-2628.

Koehn, H., Lau, B., Clerens, S., Plowman, J., Dyer, J., Ramli, U., & Deb-Choudhury, S. (2011). Combination of acid labile detergent and C18 Empore™ disks for improved identification and sequence coverage of in-gel digested proteins. *Analytical and Bioanalytical Chemistry, 400*(2), 415-421. doi:10.1007/s00216-011-4765-1

Koller, A., Washburn, M. P., Lange, B. M., Andon, N. L., Deciu, C., Haynes, P. A., . . . Yates III, J. R. (2002). Proteomic survey of metabolic pathways in rice. *Proceedings of the National Academy of Sciences of the United States of America, 99*(18), 11969-11974. doi:10.1073/pnas.172183199

Kosmulski, M. (2002). The significance of the difference in the point of zero charge between rutile and anatase. *Advances in colloid and interface science, 99*(3), 255-264. doi:10.1016/s0001-8686(02)00080-5

Krishna, K., Joshi, M. K., Vani, B., & Mohanty, P. (1999). Structure-function correlation during the etioplast-chloroplast transition in cucumber cotyledonary leaves. *Photosynthetica, 36*(1), 199-212. doi:10.1023/a:1007043409538

Kwon, S. J., Choi, E. Y., Choi, Y. J., Ahn, J. H., & Park, O. K. (2006). Proteomics studies of post-translational modifications in plants. *Journal of Experimental Botany, 57*(7), 1547-1551. doi:10.1093/jxb/erj137

Lamberti, C., Pessione, E., Giuffrida, M. G., Mazzoli, R., Barello, C., Conti, A., & Giunta, C. (2007). Combined cup loading, bis(2-hydroxyethyl) disulfide, and protein precipitation protocols to improve the alkaline proteome of Lactobacillus hilgardii. *ELECTROPHORESIS, 28*(10), 1633-1638. doi:10.1002/elps.200600496

Lange, V., Picotti, P., Domon, B., & Aebersold, R. (2008). *Selected reaction monitoring for quantitative proteomics: a tutorial* (Vol. 4) [Journal Article]. Retrieved from http://msb.embopress.org/msb/4/1/222.full.pdf. doi:10.1038/msb.2008.61

Langeveld, J. W. A., Dixon, J., & Jaworski, J. F. (2010). Development perspectives of the biobased economy: A review *Crop Sci., 50*(Supplement_1), S-142-S-151. doi:10.2135/cropsci2009.09.0529

Lardizabal, K. D., Metz, J. G., Sakamoto, T., Hutton, W. C., Pollard, M. R., & Lassner, M. W. (2000). Purification of a jojoba embryo wax synthase, cloning of its cDNA, and production of high levels of wax in seeds of transgenic arabidopsis. *Plant Physiology, 122*(3), 645-655.

Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., & Jørgensen, T. J. D. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular & Cellular Proteomics, 4*(7), 873-886. doi:10.1074/mcp.T500007-MCP200

Larsen, M. R., Trelle, M. B., Thingholm, T. E., & Jensen, O. N. (2006). Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *Biotechniques, 40*, 790-798. doi:10.2144/000112201

Lawlor, D. W. (2012). Genetic engineering to improve plant performance under drought: physiological evaluation of achievements, limitations, and possibilities. *Journal of Experimental Botany*. doi:10.1093/jxb/ers326

Lee, J., Feng, J., Campbell, K. B., Scheffler, B. E., Garrett, W. M., Thibivilliers, S., . . . Cooper, B. (2009). Quantitative proteomic analysis of bean plants infected by a virulent and avirulent obligate rust fungus. *Molecular & Cellular Proteomics, 8*(1), 19-31. doi:10.1074/mcp.M800156-MCP200

Lee, T. Y., Huang, H. D., Hung, J. H., Huang, H. Y., Yang, Y. S., & Wang, T. H. (2006). dbPTM: an information repository of protein post-translational modification [Research Support,

Non-U.S. Gov't]. *Nucleic Acids Research, 34*(Database issue), D622-627. doi:10.1093/nar/gkj083

Lee, Y. H., Tan, H. T., & Chung, M. C. M. (2010). Subcellular fractionation methods and strategies for proteomics. *PROTEOMICS, 10*(22), 3935-3956. doi:10.1002/pmic.201000289

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., . . . Ohlrogge, J. (2013). Acyl-lipid metabolism. *Arabidopsis Book, 11*, e0161. doi:10.1199/tab.0161

Lilley, K. S., & Dupree, P. (2006). Methods of quantitative proteomics and their application to plant organelle characterization. *Journal of Experimental Botany, 57*(7), 1493-1499. doi:10.1093/jxb/erj141

Lilley, K. S., & Dupree, P. (2007). Plant organelle proteomics. *Current Opinion in Plant Biology, 10*(6), 594-599. doi:10.1016/j.pbi.2007.08.006

Lim, C. C., Teo, K. W., Rao, V., & Chia, C. C. (2003). Performances of some pisiferas of Binga, Ekona, urt and angolan origins: Part 1 - breeding background and fruit bunch traits. *Journal of Oil Palm Research, 15*(1), 21-31.

Lim, Y. P. (2005). Mining the tumor phosphoproteome for cancer markers. *Clinical Cancer Research, 11*(9), 3163-3169. doi:10.1158/1078-0432.ccr-04-2243

Lin, Y., Zhou, J., Bi, D., Chen, P., Wang, X., & Liang, S. (2008). Sodium-deoxycholate-assisted tryptic digestion and identification of proteolytically resistant proteins. *Analytical Biochemistry, 377*(2), 259-266. doi:10.1016/j.ab.2008.03.009

Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., . . . Yates III, J. R. (1999). Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology, 17*(7), 676-682. doi:10.1038/10890

Liu, Q., Singh, S. P., & Green, A. G. (2002a). High-stearic and high-oleic cottonseed oils produced by hairpin RNA-mediated post-transcriptional gene silencing. *Plant Physiology, 129*(4), 1732-1743. doi:10.1104/pp.001933

Liu, Q., Singh, S. P., & Green, A. G. (2002b). High-stearic and high-oleic cottonseed oils produced by hairpin RNA-mediated post-transcriptional gene silencing [Research Support, Non-U.S. Gov't]. *Plant Physiology, 129*(4), 1732-1743. doi:10.1104/pp.001933

Ljubesic, N. (1972). Ultrastructural changes of plastids during the yellowing of the fruit of *Cucurbita pepo* var *pyriformis*. *Acta Botanical Croatica, 31*, 47-53. doi:Retrieved from http://www.abc.botanic.hr/index.php/abc/issue/archive

Ljubičić, J. M., Wrischer, M., & Ljubešić, N. (1998). Formation of the photosynthetic apparatus in plastids during greening of potato microtubers. *Plant Physiology and Biochemistry, 36*(10), 747-752. doi:10.1016/s0981-9428(98)80025-9

Loei, H., Lim, J., Tan, M., Lim, T. K., Lin, Q. S., Chew, F. T., . . . Chung, M. C. (2013). Proteomic analysis of the oil palm fruit mesocarp reveals elevated oxidative phosphorylation

activity is critical for increased storage oil production. *Journal of Proteome Research, 12*(11), 5096-5109. doi:10.1021/pr400606h

Loomis, W. D., & Battaile, J. (1966). Plant phenolic compounds and the isolation of plant enzymes. *Phytochemistry, 5*(3), 423-438.

Lopez-Juez, E., & Pyke, K. A. (2005). Plastid unleashed: their development and their integration in plant development. *The International Journal of Development Biology, 49*(5-6), 557-577. doi:10.1387/ijdb.051997el

Lu, B., Ruse, C., Xu, T., Park, S. K., & Yates III, J. (2007). Automatic validation of phosphopeptide identifications from tandem mass spectra. *Analytical Chemistry, 79*(4), 1301-1310. doi:10.1021/ac061334v

Lucker, J., Laszczak, M., Smith, D., & Lund, S. (2009). Generation of a predicted protein database from EST data and application to iTRAQ analyses in grape (Vitis vinifera cv. Cabernet Sauvignon) berries at ripening initiation. *BMC Genomics, 10*(1), 50.

Madoka, Y., Tomizawa, K., Mizoi, J., Nishida, I., Nagano, Y., & Sasaki, Y. (2002). Chloroplast transformation with modified accD operon increases acetyl-CoA carboxylase and causes extension of leaf longevity and increase in seed yield in tobacco. *Plant and Cell Physiology, 43*(12), 1518-1525. doi:10.1093/pcp/pcf172

Majeran, W., Cai, Y., Sun, Q., & van Wijk, K. J. (2005). Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. *The Plant Cell Online, 17*(11), 3111-3140. doi:10.1105/tpc.105.035519

Malaysia, D. o. S. (2011, February 2). Retrieved 2 February, from http://www.statistics.gov.my/portal/images/stories/files/LatestReleases/trade/bi/Dec10/Pre_External_Trade_DisBI.pdf.

Mamone, G., Picariello, G., Ferranti, P., & Addeo, F. (2010). Hydroxyapatite affinity chromatography for the highly selective enrichment of mono- and multi-phosphorylated peptides in phosphoproteome analysis. *PROTEOMICS, 10*(3), 380-393. doi:10.1002/pmic.200800710

Mann, M. (2009). Comparative analysis to guide quality improvements in proteomics. *Nature Methods, 6*(10), 717-719. doi:10.1038/nmeth1009-717

Mann, M., & Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nature Biotechnology, 21*(3), 255-261. doi:10.1038/nbt0303-255

Mann, M., Ong, S.-E., Grønborg, M., Steen, H., Jensen, O. N., & Pandey, A. (2002). Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends in biotechnology, 20*(6), 261-268. doi:10.1016/S0167-7799(02)01944-3

Marsh, E., Alvarez, S., Hicks, L. M., Barbazuk, W. B., Qiu, W., Kovacs, L., & Schachtman, D. (2010). Changes in protein abundance during powdery mildew infection of leaf tissues of Cabernet Sauvignon grapevine (Vitis vinifera L.). *PROTEOMICS, 10*(10), 2057-2064. doi:10.1002/pmic.200900712

Martinez-Marquez, A., Morante-Carriel, J., Selles-Marchart, S., Martinez-Esteso, M. J., Pineda-Lucas, J. L., Luque, I., & Bru-Martinez, R. (2013). Development and validation of MRM methods to quantify protein isoforms of polyphenol oxidase in loquat fruits [Research Support, Non-U.S. Gov't]. *Journal of Proteome Research, 12*(12), 5709-5722. doi:10.1021/pr4006712

Marx, V. (2013). Making sure PTMs are not lost after translation [10.1038/nmeth.2370]. *Nature Methods, 10*(3), 201-204.

McCabe, M. S., Garratt, L. C., Schepers, F., Jordi, W. J. R. M., Stoopen, G. M., Davelaar, E., . . . Davey, M. R. (2001). Effects of PSAG12-IPT gene expression on development and senescence in transgenic lettuce. *Plant Physiology, 127*(2), 505-516. doi:10.1104/pp.010244

McNulty, D. E., & Annan, R. S. (2008). Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Molecular & Cellular Proteomics, 7*(5), 971-980. doi:10.1074/mcp.M700543-MCP200

Méchin, V., Damerval, C., & Zivy, M. (2006). Total protein extraction with TCA-acetone. In (Vol. 355, pp. 1-8). Retrieved from http://www.springerprotocols.com/Abstract/doi/10.1385/1-59745-227-0:1

Mekhedov, S., Cahoon, E., & Ohlrogge, J. (2001). An unusual seed-specific 3-ketoacyl-ACP synthase associated with the biosynthesis of petroselinic acid in coriander. *Plant Molecular Biology, 47*(4), 507-518. doi:10.1023/a:1011832611885

Mekhedov, S., de Ilarduya, O. M., & Ohlrogge, J. (2000). Toward a functional catalog of the plant genome. A survey of genes for lipid biosynthesis. *Plant Physiology, 122*(2), 389-402. doi:10.1104/pp.122.2.389

Melo-Braga, M. N., Verano-Braga, T., Leon, I. R., Antonacci, D., Nogueira, F. C., Thelen, J. J., . . . Palmisano, G. (2012). Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (Vitis vinifera) mesocarp and exocarp owing to *Lobesia botrana* infection [Research Support, Non-U.S. Gov't]. *Molecular & Cellular Proteomics, 11*(10), 945-956. doi:10.1074/mcp.M112.020214

Miles, G. P., Samuel, M. A., Ranish, J. A., Donohoe, S. M., Sperrazzo, G. M., & Ellis, B. E. (2009). Quantitative proteomics identifies oxidant-induced, AtMPK6-dependent changes in *Arabidopsis thaliana* protein profiles. *Plant Signaling and Behavior, 4*(6), 497-505. doi:10.4161/psb.4.6.8538

Mittler, R., & Blumwald, E. (2010). Genetic engineering for modern agriculture: challenges and perspectives. *Annual Review of Plant Biology, 61*(1), 443-462. doi:10.1146/annurev-arplant-042809-112116

Mohammadi, M., Anoop, V., Gleddie, S., & Harris, L. J. (2011). Proteomic profiling of two maize inbreds during early gibberella ear rot infection. *PROTEOMICS, 11*(18), 3675-3684. doi:10.1002/pmic.201100177

Molina, H., Horn, D. M., Tang, N., Mathivanan, S., & Pandey, A. (2007). Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America, 104*(7), 2199-2204. doi:10.1073/pnas.0611217104

Moon, J. H., Shin, Y. S., & Kim, M. S. (2009). Utility of reaction intermediate monitoring with photodissociation multi-stage (MSn) time-of-flight mass spectrometry for mechanistic and structural studies: Phosphopeptides. *International Journal of Mass Spectrometry, 288*(1-3), 16-21. doi:DOI 10.1016/j.ijms.2009.07.008

Mortz, E., O'Connor, P. B., Roepstorff, P., Kelleher, N. L., Wood, T. D., McLafferty, F. W., & Mann, M. (1996). Sequence tag identification of intact proteins by matching tanden mass spectral data against sequence data bases. *Proceedings of the National Academy of Sciences of the United States of America, 93*(16), 8264-8267. doi:Retrieved from http://www.pnas.org/content/93/16/8264

Murphy, D. J. (2005). *Plant lipids. Biology, utilisation and biology*. Oxford:UK: Blackwell Publishing Ltd.

Murphy, D. J. (2006). Molecular breeding strategies for the modification of lipid composition. *In Vitro Cellular & Developmental Biology - Plant, 42*(2), 89-99. doi:10.1079/ivp2005734

Nagele, E., Vollmer, M., Horth, P., & Vad, C. (2004). 2D-LC/MS techniques for the identification of proteins in highly complex mixtures [Review]. *Expert Review of Proteomics, 1*(1), 37-46. doi:10.1586/14789450.1.1.37

Nakagami, H., Sugiyama, N., Ishihama, Y., & Shirasu, K. (2012). Shotguns in the front line: phosphoproteomics in plants. *Plant and Cell Physiology, 53*(1), 118-124. doi:10.1093/pcp/pcr148

Nakanishi, T., Ando, E., Furuta, M., Kinoshita, E., Kinoshita-Kikuta, E., Koike, T., . . . Nishimura, O. (2007). Identification on membrane and characterization of phosphoproteins using an Alkoxide-Bridged Dinuclear Metal Complex as a phosphate-binding tag molecule. *Journal of Biomolecular Techniques, 18*(5), 278-286. doi:Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2392992/?tool=pubmed

National Economic Advisory Council. (2010). *New economic model*. Putrajaya, Malaysia: Performance Management and Delivery Unit (PEMANDU) Retrieved from http://www.neac.gov.my/publications.php?subcategory=nem%20reports&ID=287&title=part%201%20-%20appendix%201%20-%20sustainability%20and%20the%20palm%20oil%20industry

Nawrocki, J., Dunlap, C., McCormick, A., & Carr, P. W. (2004). Part I. Chromatography using ultra-stable metal oxide-based stationary phases for HPLC. *Journal of Chromatography A, 1028*(1), 1-30. doi:10.1016/j.chroma.2003.11.052

Negm, F. B., Cornel, F. A., & Plaxton, W. C. (1995). Suborganellar localization and molecular characterization of nonproteolytic degraded leucoplast pyruvate kinase from developing castor oil seeds. *Plant Physiology, 109*(4), 1461-1469. doi:10.1104/pp.109.4.1461

Nesvizhskii, A. I., Vitek, O., & Aebersold, R. (2007). Analysis and validation of proteomic data generated by tandem mass spectrometry [Research Support, N.I.H., Extramural

Review]. *Nature Methods, 4*(10), 787-797. doi:10.1038/nmeth1088

Neuhaus, H. E., & Emes, M. J. (2000). Nonphotosynthetic metabolism in plastids. *Annual Review of Plant Physiology and Plant Molecular Biology, 51*(1), 111-140. doi:10.1146/annurev.arplant.51.1.111

Niessen, W. M. A. (2006). *The encyclopedia of mass spectrometry*: Elsevier.

Niittylä, T., Fuglsang, A. T., Palmgren, M. G., Frommer, W. B., & Schulze, W. X. (2007). Temporal analysis of sucrose-induced phosphorylation changes in plasma membrane proteins of *Arabidopsis*. *Molecular & Cellular Proteomics, 6*(10), 1711-1726. doi:10.1074/mcp.M700164-MCP200

Nikolau, B. J., & Hawke, J. C. (1984). Purification and characterization of maize leaf acetyl-coenzyme A carboxylase. *Archives of Biochemistry and Biophysics, 228*(1), 86-96. doi:10.1016/0003-9861(84)90049-3

Nishimura, M., & Beevers, H. (1978). Hydrolases in vacuoles from castor bean endosperm. *Plant Physiology, 62*(1), 44-48. doi:10.1104/pp.62.1.44

Norden, A. J., Gorbet, D. W., Knauft, D. A., & Young, C. T. (1987). Variability in oil quality among peanut genotypes on the Florida breeding program. *Peanut Science, 14*(1), 7-11. doi:10.3146/i0095-3679-14-1-3

Nuhse, T. S., Bottrill, A. R., Jones, A. M., & Peck, S. C. (2007). Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses [Research Support, Non-U.S. Gov't]. *The Plant Journal, 51*(5), 931-940. doi:10.1111/j.1365-313X.2007.03192.x

Nuhse, T. S., Stensballe, A., Jensen, O. N., & Peck, S. C. (2003). Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry [Research Support, Non-U.S. Gov't]. *Molecular & Cellular Proteomics, 2*(11), 1234-1243. doi:10.1074/mcp.T300006-MCP200

O'Donovan, C., Apweiler, R., & Bairoch, A. (2001). The human proteomics initiative (HPI). *Trends in biotechnology, 19*(5), 178-181. doi:10.1016/s0167-7799(01)01598-0

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins [Research Support, U.S. Gov't, Non-P.H.S.]. *The Journal of Biological Chemistry, 250*(10), 4007-4021. doi:Retrieved from http://www.jbc.org/content/250/10/4007.abstract

Obahiagbon, F. I. (2012). A review: aspects of the African oil palm (*Elaeis guineensis* jacq.) and the implications of its bioactives in human health. *American Journal of Biochemistry and Molecular Biology, 2*(3), 106-119. doi:10.3923/ajbmb.2012.106.119

Oeljeklaus, S., Meyer, H. E., & Warscheid, B. (2009). Advancements in plant proteomics using quantitative mass spectrometry. *Journal of Proteomics, 72*(3), 545-554. doi:10.1016/j.jprot.2008.11.008

Ohlrogge, J., & Browse, J. (1995). Lipid biosynthesis [Review]. *Plant Cell, 7*(7), 957-970. doi:10.1105/tpc.7.7.957

Ohlrogge, J. B. (1994). Design of new plant products: engineering of fatty acid metabolism. *Plant Physiology, 104*(3), 821-826. doi:10.1104/pp.104.3.821

Ohlrogge, J. B., Browse, J., & Somerville, C. R. (1991). The genetics of plant lipids. *Biochimica et Biophysica Acta 1082*(1), 1-26. doi:10.1016/0005-2760(91)90294-r

Ohlrogge, J. B., & Jaworski, J. G. (1997). Regulation of fatty acid synthesis. *Annual Review of Plant Physiology and Plant Molecular Biology, 48*(1), 109-136. doi:10.1146/annurev.arplant.48.1.109

Ohlrogge, J. B., Mhaske, V. B., Beisson, F., & Ruuska, S. (2004, 26 September - 1 October). Genomics approaches to lipid biosynthesis. New directions for a diverse plane Symposium conducted at the meeting of the Proceedings of the 4th International Crop Science Congress, Brisbane, Australia.

Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., . . . Ahn, N. G. (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Molecular & Cellular Proteomics, 4*(10), 1487-1502. doi:10.1074/mcp.M500084-MCP200

Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., & Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell, 127*(3), 635-648. doi:10.1016/j.cell.2006.09.026

Osorio, J., Fernandez-Martinez, J., Mancha, M., & Garces, R. (1995). Mutant sunflower with high concentration of saturated fatty acids in their oil. *Crop Science, 35*, 739-742. doi:10.2135/cropsci1995.0011183X003500030016x

Ow, S. Y., Salim, M., Noirel, J., Evans, C., Rehman, I., & Wright, P. C. (2009). iTRAQ underestimation in simple and complex mixtures: “The good, the bad and the ugly”. *Journal of Proteome Research, 8*(11), 5347-5355. doi:10.1021/pr900634c

Page, R. A., Okada, S., & Harwood, J. L. (1994). Acetyl-CoA carboxylase exerts strong flux control over lipid synthesis in plants [Research Support, Non-U.S. Gov't]. *Biochimica et Biophysica Acta 1210*(3), 369-372. doi:10.1016/0005-2760(94)90242-9

Palmblad, M., Mills, D. J., Bindschedler, L. V., & Cramer, R. (2007). Chromatographic alignment of LC-MS and LC-MS/MS datasets by genetic algorithm feature extraction. *American Society for Mass Spectrometry, 18*(10), 1835-1843. doi:10.1016/j.jasms.2007.07.018

Palumbo, A. M., Smith, S. A., Kalcic, C. L., Dantus, M., Stemmer, P. M., & Reid, G. E. (2011). Tandem mass spectrometry strategies for phosphoproteome analysis [Review]. *Mass Spectrometry Reviews, 30*(4), 600-625. doi:10.1002/mas.20310

Patton, W. F. (2002). Detection technologies in proteome analysis. *Journal of Chromatography B, 771*(1–2), 3-31. doi:10.1016/s1570-0232(02)00043-0

Paulo, J. A., Kadiyala, V., Banks, P. A., Steen, H., & Conwell, D. L. (2012). Mass spectrometry-based (GeLC-MS/MS) comparative proteomic analysis of endoscopically (ePFT) Collected Pancreatic and Gastroduodenal fluids [Pancreas and Biliary Tract]. *Clinical and Translational Gastroenterology, 3*, e14. doi:10.1038/ctg.2012.7

Pedretti, K., Scheetz, T., Braun, T., Roberts, C., Robinson, N., & Casavant, T. (2001). A parallel Expressed Sequence Tag (EST) clustering program. In V. Malyshkin (Ed.), *Parallel Computing Technologies* (Vol. 2127, pp. 490-497): Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/3-540-44743-1_51. doi:10.1007/3-540-44743-1_51

Peltier, J.-B., Cai, Y., Sun, Q., Zabrouskov, V., Giacomelli, L., Rudella, A., . . . van Wijk, K. J. (2006). The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Molecular & Cellular Proteomics, 5*(1), 114-133. doi:10.1074/mcp.M500180-MCP200

Peltier, J.-B., Friso, G., Kalume, D. E., Roepstorff, P., Nilsson, F., Adamska, I., & van Wijk, K. J. (2000). Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. *The Plant Cell Online, 12*(3), 319-342. doi:10.1105/tpc.12.3.319

Peltier, J.-B., Ytterberg, A. J., Sun, Q., & van Wijk, K. J. (2004). New functions of the thylakoid membrane proteome of *Arabidopsis thaliana* revealed by a simple, fast, and versatile fractionation strategy. *Journal of Biological Chemistry, 279*(47), 49367-49383. doi:10.1074/jbc.M406763200

Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., & Gygi, S. P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *Journal of Proteome Research, 2*(1), 43-50. doi:10.1021/pr025556v

Perkel, J. M. (2009). iTRAQ gets put to the test. *Journal of Proteome Research, 8*(11), 4885-4885. doi:10.1021/pr900897d

Perkins, D. N., Pappin, D. J. C., Creasy, D. M., & Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *ELECTROPHORESIS, 20*(18), 3551-3567. doi:10.1002/(sici)1522-2683(19991201)20:18<3551::aid-elps3551>3.0.co;2-2

Pesavento, J. J., Mizzen, C. A., & Kelleher, N. L. (2006). Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: Human histone H4. *Analytical Chemistry, 78*(13), 4271-4280. doi:10.1021/ac0600050

Petricka, J. J., Schauer, M. A., Megraw, M., Breakfield, N. W., Thompson, J. W., Georgiev, S., . . . Benfey, P. N. (2012). The protein expression landscape of the Arabidopsis root. *Proceedings of the National Academy of Sciences of the United States of America, 109*(18), 6811-6818. doi:10.1073/pnas.1202546109

Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B., & Ruedi, A. (2009). Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell, 138*(4), 795-806. doi:10.1016/j.cell.2009.05.051

Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B., & Heck, A. J. R. (2004). Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-nanoLC-ESI-MS/MS and titanium oxide precolumns. *Analytical Chemistry, 76*(14), 3935-3943. doi:10.1021/ac0498617

Plewczynski, D., Tkacz, A., Wyrwicz, L. S., & Rychlewski, L. (2005). AutoMotif server: prediction of single residue post-translational modifications in proteins. *Bioinformatics, 21*(10), 2525-2527. doi:10.1093/bioinformatics/bti333

Podkowinski, J., & Tworak, A. (2011). Acetyl-conenzyme A carboxylase - an attractive enzyme for biotechnology. *BioTechnologia, 92*(4), 321-335.

Posewitz, M. C., & Tempst, P. (1999). Immobilized gallium(III) affinity chromatography of phosphopeptides. *Analytical Chemistry, 71*(14), 2883-2892. doi:10.1021/ac981409y

Post-Beittenmiller, D., Roughan, G., & Ohlrogge, J. B. (1992). Regulation of plant Fatty Acid biosynthesis : analysis of acyl-coenzyme a and acyl-acyl carrier protein substrate pools in spinach and pea chloroplasts. *Plant Physiology, 100*(2), 923-930. doi:10.1104/pp.100.2.923

Prabakaran, S., Lippens, G., Steen, H., & Gunawardena, J. (2012). Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 4*(6), 565-583. doi:10.1002/wsbm.1185

Proc, J. L., Kuzyk, M. A., Hardie, D. B., Yang, J., Smith, D. S., Jackson, A. M., . . . Borchers, C. H. (2010). A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. *Journal of Proteome Research, 9*(10), 5422-5437. doi:10.1021/pr100656u

Pyke, K. A. (1999). Plastid division and development. *The Plant Cell Online, 11*(4), 549-556. doi:10.1105/tpc.11.4.549

Rajahnaidu, N. (1994). *PORIM oil palm genebank.* Palm Oil Research Institute, Bangi, Malaysia.

Ramli, U. S., Baker, D. S., Quant, P. A., & Harwood, J. L. (2002). Control analysis of lipid biosynthesis in tissue cultures from oil crops shows that flux control is shared between fatty acid synthesis and lipid assembly. *Biochemical Journal, 364*(2), 393-401. doi:10.1042/bj20010203

Rastegari, E., Ahmad, Z., Spencer, D. F., & Ismai, M. (2011). Two-dimensional profiling of proteins from *Curculigo latifolia* fruit by three different extraction protocols. *Journal of Medicinal Plants Research, 5*(16), 3719-3724.

Rawsthorne, S. (2002). Carbon flux and fatty acid synthesis in plants. *Progress in Lipid Research, 41*(2), 182-196. doi:10.1016/s0163-7827(01)00023-6

Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., . . . Baginsky, S. (2009). Large-scale Arabidopsis phosphoproteome profiling teveals novel vhloroplastkinase dubstrates and phosphorylation networks. *Plant Physiology, 150*(2), 889-903. doi:10.1104/pp.109.138677

Reinders, J., & Sickmann, A. (2005). State-of-the-art in phosphoproteomics. *PROTEOMICS, 5*(16), 4052-4061. doi:10.1002/pmic.200401289

Remmerie, N., De Vijlder, T., Laukens, K., Dang, T. H., Lemière, F., Mertens, I., . . . Witters, E. (2011). Next generation functional proteomics in non-model plants: A survey on techniques and applications for the analysis of protein complexes and post-translational modifications. *Phytochemistry, 72*(10), 1192-1218. doi:10.1016/j.phytochem.2011.01.003

Renaut, J., Hausman, J.-F., & Wisniewski, M. E. (2006). Proteomics and low-temperature studies: bridging the gap between gene expression and metabolism. *Physiologia Plantarum, 126*(1), 97-109. doi:10.1111/j.1399-3054.2006.00617.x

Rice, A. L., & Burns, J. B. (2010). Moving from efficacy to effectiveness: red palm oil's role in preventing vitamin A deficiency [Review]. *Journal of American College of Nutrition, 29*(3 Suppl), 302S-313S.

Rodrigues, E. P., Torres, A. R., da Silva Batista, J. S., Huergo, L., & Hungria, M. (2012). A simple, economical and reproducible protein extraction protocol for proteomics studies of soybean roots. *Genetics and Molecular Biology, 35*(1 (suppl)), 348-352. doi:10.1590/S1415-47572012000200016

Roesler, K., Shintani, D., Savage, L., Boddupalli, S., & Ohlrogge, J. (1997). Targeting of the Arabidopsis homomeric acetyl-coenzyme A carboxylase to plastids of rapeseeds. *Plant Physiology, 113*(1), 75-81. doi:10.1104/pp.113.1.75

Roesler, K. R., Savage, L. J., Shintani, D. K., Shorrosh, B. S., & Ohlrogge, J. B. (1996). Co-purification, co-imniunoprecipitation, and coordinate expression of acetyl-coenzyme A carboxylase activity, biotin carboxylase, and biotin carboxyl carrier protein of higher plants. *Planta, 198*(4), 517-525. doi:10.1007/bf00262637

Roessler, P. G. (1990). Purification and characterization of acetyl-CoA carboxylase from the Diatom *Cyclotella cryptica*. *Plant Physiology, 92*(1), 73-78. doi:10.1104/pp.92.1.73

Roughan, P. G., & Slack, C. R. (1982). Cellular Organization of Glycerolipid Metabolism. *Annual Review of Plant Physiology, 33*(1), 97-132. doi:10.1146/annurev.pp.33.060182.000525

Rowley, A., Choudhary, J. S., Marzioch, M., Ward, M. A., Weir, M., Solari, R. C. E., & Blackstock, W. P. (2000). Applications of protein mass spectrometry in cell biology. *Methods, 20*(4), 383-397. doi:10.1006/meth.2000.0951

Ruttenberg, B. E., Pisitkun, T., Knepper, M. A., & Hoffert, J. D. (2008). PhosphoScore: An open-source phosphorylation site assignment tool for MSn data. *Journal of Proteome Research, 7*(7), 3054-3059. doi:10.1021/pr800169k

Ruuska, S. A., Girke, T., Benning, C., & Ohlrogge, J. B. (2002). Contrapuntal networks of gene expression during Arabidopsis seed filling. *The Plant Cell Online, 14*(6), 1191-1206. doi:10.1105/tpc.000877

Salvato, F., Havelund, J. F., Chen, M., Rao, R. S., Rogowska-Wrzesinska, A., Jensen, O. N., . . . Moller, I. M. (2014). The potato tuber mitochondrial proteome. *Plant Physiology, 164*(2), 637-653. doi:10.1104/pp.113.229054

Sambanthamurthi, R., Abrizah, O., & Ramli, U. S. (1999). Biochemical factors that control oil composition in the oil palm. *Journal of Oil Palm Research*, 24-33.

Sano, A., & Nakamura, H. (2004a). Chemo-affinity of titania for the column-switching HPLC analysis of phosphopeptides. *Analytical Sciences, 20*(3), 565-566. doi:10.2116/analsci.20.565

Sano, A., & Nakamura, H. (2004b). Titania as a chemo-affinity support for the column-switching HPLC analysis of phosphopeptides: Application to the characterization of phosphorylation sites in proteins by combination with protease digestion and electrospray ionization mass spectrometry. *Analytical Sciences, 20*(5), 861-864. doi:10.2116/analsci.20.861

Saravanan, R. S., & Rose, J. K. C. (2004). A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *PROTEOMICS, 4*(9), 2522-2532. doi:10.1002/pmic.200300789

Sasaki, Y., & Nagano, Y. (2004a). Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding [Review]. *Bioscience, Biotechnology, and Biochemistry, 68*(6), 1175-1184.

Sasaki, Y., & Nagano, Y. (2004b). Plant acetyl-CoA carboxylase: Structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Bioscience, Biotechnology, and Biochemistry, 68*(6), 1175-1184. doi:10.1271/bbb.68.1175

Schenkluhn, L., Hohnjec, N., Niehaus, K., Schmitz, U., & Colditz, F. (2010). Differential gel electrophoresis (DIGE) to quantitatively monitor early symbiosis- and pathogenesis-induced changes of the Medicago truncatula root proteome. *Journal of Proteomics, 73*(4), 753-768. doi:10.1016/j.jprot.2009.10.009

Schluter, H., Apweiler, R., Holzhutter, H.-G., & Jungblut, P. (2009). Finding one's way in proteomics: a protein species nomenclature. *Chemistry Central Journal, 3*(1), 11. doi:10.1186/1752-153X-3-11

Schmidt, A., Csaszar, E., Ammerer, G., & Mechtler, K. (2008). Enhanced detection and identification of multiply phosphorylated peptides using TiO2 enrichment in combination with MALDI TOF/TOF MS. *PROTEOMICS, 8*(21), 4577-4592. doi:10.1002/pmic.200800279

Schuch, R., Winter, E., Brück, F. M., Brummel, M., & Spener, F. (1997). β-Ketoacyl-acyl carrier protein synthases in the regulation of fatty acid synthase activity in higher plants — An overview. *Fett/Lipid, 99*(8), 278-281. doi:10.1002/lipi.19970990804

Schulenberg, B., Aggeler, R., Beechem, J. M., Capaldi, R. A., & Patton, W. F. (2003). Analysis of steady-state protein phosphorylation in mitochondria using a novel fluorescent phosphosensor dye. *Journal of Biological Chemistry, 278*(29), 27251-27255. doi:10.1074/jbc.C300189200

Schultz, D., & Ohlrogge, J. B. (2001). Metabolic engineering of fatty acid biosynthesis. In T.-M. Kuo & Gardner (Eds.), *Lipid biotechnology*. New York:NY: Marcel Dekker, Inc.

Schulze, W. X., & Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics [Review]. *Annual Review of Plant Biology, 61*(1), 491-516. doi:10.1146/annurev-arplant-042809-112132

Schuster, A., M., & Davies, E. (1983). Ribonucleic acid and protein metabolism in Pea epicotyls. *Plant Physiology, 73*, 817-821.

Selbach, M., & Mann, M. (2006). Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK) [Research Support, Non-U.S. Gov't]. *Nature Methods, 3*(12), 981-983. doi:10.1038/nmeth972

Sen, C. K., Rink, C., & Khanna, S. (2010). Palm oil-derived natural vitamin E alpha-tocotrienol in brain health and disease. *J Am Coll Nutr, 29*(3 Suppl), 314S-323S.

Seng, T. K. (1987). *The botany of oil palm*. Paper presented at the meeting of the The Incorporated Society of Planters, Kuala Lumpur.

Seo, J., & Lee, K.-J. (2004). Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *Journal of Biochemistry and Molecular Biology, 37*(1), 35-44. doi:Retrieved from http://www.jbmb.or.kr/view_content.php3?cont=jbmb&number=162

Shah, F. H., Rashid, O., & San, C. T. (2000). Temporal regulation of two isoforms of cDNA clones encoding delta 9-stearoyl-ACP desaturase from oil palm (*Elaies guineensis*). *Plant Science, 152*, 27-33.

Shahwan, S. (2006). *Cloning and characterization of oleoyl-CoA desaturase gene from oil palm (Elaeis guineensis L.)*. Universiti Putra Malaysia, Selangor.

Shanklin, J., & Cahoon, E. B. (1998). Desaturation and related modifications of fatty acids. *Annual Review of Plant Physiology and Plant Molecular Biology, 49*(1), 611-641. doi:10.1146/annurev.arplant.49.1.611

Shanklin, J., & Somerville, C. (1991). Stearoyl-acyl-carrier-protein desaturase from higher-plants Is structurally unrelated to the animal and fungal homologs. *Proceedings of the National Academy of Sciences of the United States of America, 88*(6), 2510-2514. doi:DOI 10.1073/pnas.88.6.2510

Sharma, K., Weber, C., Bairlein, M., Greff, Z., Keri, G., Cox, J., . . . Daub, H. (2009). Proteomics strategy for quantitative protein interaction profiling in cell extracts [Research Support, Non-U.S. Gov't]. *Nature Methods, 6*(10), 741-744. doi:10.1038/nmeth.1373

Sharma, S., Simpson, D. C., Tolic, N., Jaitly, N., Mayampurath, A. M., Smith, R. D., & Pasa-Tolic, L. (2007). Proteomic profiling of intact proteins using WAX-RPLC 2-D

separations and FTICR mass spectrometry. *Journal of Proteome Research, 6*(2), 602-610. doi:10.1021/pr060354a

Sheoran, I. S., Ross, A. R. S., Olson, D. J. H., & Sawhney, V. K. (2009). Compatibility of plant protein extraction methods with mass spectrometry for proteome analysis. *Plant Science, 176*(1), 99-104. doi:10.1016/j.plantsci.2008.09.015

Shevchenko, A., Sunyaev, S., Loboda, A., Bork, P., Ens, W., & Standing, K. G. (2001). Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Analytical Chemistry, 73*(9), 1917-1926. doi:10.1021/ac0013709

Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels [Research Support, Non-U.S. Gov't]. *Analytical Chemistry, 68*(5), 850-858. doi:10.1021/ac950914h

Shintani, D. K., & Ohlrogge, J. B. (1995). Feedback inhibition of fatty acid synthesis in tobacco suspension cells. *The Plant Journal, 7*(4), 577-587. doi:10.1046/j.1365-313X.1995.7040577.x

Siddique, M. A., Grossmann, J., Gruissem, W., & Baginsky, S. (2006). Proteome analysis of Bell Pepper (*Capsicum annuum* L.) chromoplasts. *Plant and Cell Physiology, 47*(12), 1663-1673. doi:10.1093/pcp/pcl033

Sigrist, C. J. A., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., . . . Bucher, P. (2002). PROSITE: A documented database using patterns and profiles as motif descriptors. *Briefings in Bioinformatics, 3*(3), 265-274. doi:10.1093/bib/3.3.265

Simpson, R. J. (2003). Proteomic methods for phosphorylation site mapping. In R. J. Simpson (Ed.), *Protein and proteomics. A laboratory manual.* (pp. 597-668). New York, United States of America: Cold Spring Harbour.

Singh, R., Ong-Abdullah, M., Low, E.-T. L., Manaf, M. A. A., Rosli, R., Nookiah, R., . . . Sambanthamurthi, R. (2013). Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. *Nature, advance online publication*. doi:10.1038/nature12309

Siuti, N., & Kelleher, N. L. (2007). Decoding protein modifications using top-down mass spectrometry. *Nature Methods, 4*(10), 817-821. doi:10.1038/nmeth1097

Slabas, A. R., Brown, A., Sinden, B. S., Swinhoe, R., Simon, J. W., Ashton, A. R., . . . Elborough, K. M. (1994). Pivotal reactions in fatty acid synthesis. *Progress in Lipid Research, 33*(1-2), 39-46. doi:10.1016/0163-7827(94)90007-8

Slabas, A. R., & Fawcett, T. (1992). The biochemistry and molecular biology of plant lipid biosynthesis [Review]. *Plant Molecular Biology, 19*(1), 169-191. doi:10.1007/bf00015613

Slebos, R. J., Brock, J. W., Winters, N. F., Stuart, S. R., Martinez, M. A., Li, M., . . . Liebler, D. C. (2008). Evaluation of strong cation exchange versus isoelectric focusing of peptides for multidimensional liquid chromatography-tandem mass spectrometry [Research

Support, N.I.H., Extramural]. *Journal of Proteome Research, 7*(12), 5286-5294. doi:10.1021/pr8004666

Soyatech. (n.d.). *Palm oil facts*. Retrieved from http://www.soyatech.com/Palm_Oil_Facts.htm.

Steinberg, T. H., Agnew, B. J., Gee, K. R., Leung, W.-Y., Goodman, T., Schulenberg, B., . . . Patton, W. F. (2003). Global quantitative phosphoprotein analysis using Multiplexed Proteomics technology. *PROTEOMICS, 3*(7), 1128-1144. doi:10.1002/pmic.200300434

Stoutjesdijk, P. A., Hurlestone, C., Singh, S. P., & Green, A. G. (2000). High-oleic acid Australian Brassica napus and B. juncea varieties produced by co-suppression of endogenous Delta12-desaturases. *Biochemical Society Transactions, 28*(6), 938-940. doi:10.1042/bst0280938

Stulemeijer, I. J. E., Joosten, M. H. A. J., & Jensen, O. N. (2009). Quantitative phosphoproteomics of tomato mounting a hypersensitive response reveals a swift suppression of photosynthetic activity and a differential role for Hsp90 isoforms. *Journal of Proteome Research, 8*(3), 1168-1182. doi:10.1021/pr800619h

Stumpf, P. K. (1969). Metabolism of fatty acids. *Annual Review of Biochemistry, 38*(1), 159-212. doi:10.1146/annurev.bi.38.070169.001111

Su, H.-C., Hutchison, C., & Giddings, M. (2007). Mapping phosphoproteins in Mycoplasma genitalium and Mycoplasma pneumoniae. *BMC Microbiology, 7*(1), 63. doi:10.1186/1471-2180-7-63

Sun, M. (2012). *Palm oil poser*. Retrieved from http://www.asiaviews.org/business/38928-palm-oil-poser.

Sun, Q., Zybailov, B., Majeran, W., Friso, G., Olinares, P. D. B., & van Wijk, K. J. (2009). PPDB, the Plant Proteomics Database at Cornell. *Nucleic Acids Research, 37*(suppl 1), D969-D974. doi:10.1093/nar/gkn654

Sun, X., Chiu, J. F., & He, Q. Y. (2005). Application of immobilized metal affinity chromatography in proteomics. *Expert Review of Proteomics, 2*(5), 649-657. doi:10.1586/14789450.2.5.649

Sundram, K. (2000). Chemistry and biochemistry of palm oil. *Progress in Lipid Research*, 507-558.

Sundram, K., Sambanthamurthi, R., & Tan, Y. A. (2003). Palm fruit chemistry and nutrition [Review]. *Asia Pacific Journal of Clinical Nutrition, 12*(3), 355-362.

Swierszcz, I., Skurski, P., & Simons, J. (2012). Dipole and Coulomb forces in electron capture dissociation and electron transfer dissociation mass spectroscopy [Research Support, U.S. Gov't, Non-P.H.S.]. *The Journal of Physical Chemistry: A, 116*(7), 1828-1837. doi:10.1021/jp210915c

Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J., & Hunt, D. F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.

Proceedings of the National Academy of Sciences of the United States of America, 101(26), 9528-9533. doi:10.1073/pnas.0402700101

Szulczyk, K. R. (2013). The Economics of the Malaysian Palm Oil Industry and Its Biodiesel Potential. *SSRN Electronic Journal*. doi:10.2139/ssrn.2309130

Tada, Y., & Kashimura, T. (2009). Proteomic analysis of salt-responsive proteins in the mangrove plant, Bruguiera gymnorhiza. *Plant and Cell Physiology, 50*(3), 439-446. doi:10.1093/pcp/pcp002

Teo, K. W., Rao, V., Chia, C. C., & Lim, C. C. (2004). Performance of some pisiferas of binga, ekona, urt and angolan origins: Part 2 fruit bunch yields, vegetative growth and physiological traits. *Journal of Oil Palm Research, 16*(1), 22-38.

Tetlow, I. J., Bowsher, C. G., & Emes, M. J. (2003). Biochemical properties and enzymic capacities of chromoplasts isolated from wild buttercup (Ranunculus acris L.). *Plant Science, 165*(2), 383-394. doi:10.1016/s0168-9452(03)00199-7

Thelen, J. J., & Ohlrogge, J. B. (2002a). Metabolic engineering of fatty acid biosynthesis in plants. *Metabolic Engineering, 4*(1), 12-21. doi:10.1006/mben.2001.0204

Thelen, J. J., & Ohlrogge, J. B. (2002b). The multisubunit acetyl-CoA carboxylase is strongly associated with the chloroplast envelope through non-ionic interactions to the carboxyltransferase subunits. *Archives of Biochemistry and Biophysics, 400*(2), 245-257. doi:10.1016/s0003-9861(02)00025-5

Thelen, J. J., & Peck, S. C. (2007). Quantitative proteomics in plants: Choices in abundance. *The Plant Cell Online, 19*(11), 3339-3346. doi:10.1105/tpc.107.053991

Thingholm, T. E., Jensen, O. N., & Larsen, M. R. (2009). Analytical strategies for phosphoproteomics. *PROTEOMICS, 9*(6), 1451-1468. doi:10.1002/pmic.200800454

Thingholm, T. E., Larsen, M. R., Ingrell, C. R., Kassem, M., & Jensen, O. N. (2008). TiO2-based phosphoproteomic analysis of the plasma membrane and the effects of phosphatase inhibitor treatment. *Journal of Proteome Research, 7*(8), 3304-3313. doi:10.1021/pr800099y

Thomson, W. W., & Whatley, J. M. (1980). Development of Nongreen Plastids. *Annual Review of Plant Physiology, 31*(1), 375-394. doi:10.1146/annurev.pp.31.060180.002111

Timms, J. F., & Cramer, R. (2008). Difference gel electrophoresis. *PROTEOMICS, 8*(23-24), 4886-4897. doi:10.1002/pmic.200800298

Ting, N.-C., Zaki, N., Rosli, R., Low, E.-T., Ithnin, M., Cheah, S.-C., . . . Singh, R. (2010). SSR mining in oil palm EST database: application in oil palm germplasm diversity studies. *Journal of Genetics, 89*(2), 135-145. doi:10.1007/s12041-010-0053-7

Tranbarger, T. J., Dussert, S., Joet, T., Argout, X., Summo, M., Champion, A., . . . Morcillo, F. (2011). Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening, and functional specialization in lipid and carotenoid metabolism [Research

Support, Non-U.S. Gov't]. *Plant Physiology, 156*(2), 564-584. doi:10.1104/pp.111.175141

Troncoso-Ponce, M. A., Garces, R., & Martinez-Force, E. (2010). Glycolytic enzymatic activities in developing seeds involved in the differences between standard and low oil content sunflowers (Helianthus annuus L.) [Research Support, Non-U.S. Gov't]. *Plant Physiology and Biochemistry, 48*(12), 961-965. doi:10.1016/j.plaphy.2010.09.012

Turnham, E., & Northcote, D. H. (1983). Changes in the activity of acetyl-CoA carboxylase during rape-seed formation. *Biochemical Journal, 212*(1), 223-229. doi:Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1152033/?tool=pubmed

Unlu, M., Morgan, M. E., & Minden, J. S. (1997). Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *ELECTROPHORESIS, 18*(11), 2071-2077. doi:10.1002/elps.1150181133

Uthaipaisanwong, P., Chanprasert, J., Shearman, J. R., Sangsrakru, D., Yoocha, T., Jomchai, N., . . . Tangphatsornruang, S. (2012). Characterization of the chloroplast genome sequence of oil palm (Elaeis guineensis Jacq.). *Gene, 500*(2), 172-180. doi:10.1016/j.gene.2012.03.061

Van der Wilden, W., Herman, E. M., & Chrispeels, M., J. (1980). Protein bodies of mung bean cotyledons as autophagic organelles. *Proceedings of the National Academy of Sciences of the United States of America, 77*(1), 428-432.

van Wijk, K. J., & Baginsky, S. (2011). Plastid proteomics in higher plants: Current state and future goals. *Plant Physiology, 155*(4), 1578-1588. doi:10.1104/pp.111.172932

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., . . . Zhu, X. (2001). The sequence of the human genome. *Science, 291*(5507), 1304-1351. doi:10.1126/science.1058040

Verwoert, II, van der Linden, K. H., Nijkamp, H. J., & Stuitje, A. R. (1994). Developmental specific expression and organelle targeting of the Escherichia coli fabD gene, encoding malonyl coenzyme A-acyl carrier protein transacylase in transgenic rape and tobacco seeds [Research Support, Non-U.S. Gov't]. *Plant Molecular Biology, 26*(1), 189-202. doi:10.1007/bf00039531

Vescovi, E. G., & Viale, A. M. (1993). Protein Phosphorylation in Plastids from Ripening Tomato Fruits {Lycopersicon esculentum). *Plant and Cell Physiology, 34*(6), 905-911.

Villen, J., Beausoleil, S. A., & Gygi, S. P. (2008). Evaluation of the utility of neutral-loss-dependent MS3 strategies in large-scale phosphorylation analysis. *PROTEOMICS, 8*(21), 4444-4452. doi:10.1002/pmic.200800283

Vincent, D., Wheatley, M. D., & Cramer, G. R. (2006). Optimization of protein extraction and solubilization for mature grape berry clusters. *ELECTROPHORESIS, 27*(9), 1853-1865. doi:10.1002/elps.200500698

Vissers, J. P. C., Pons, S., Hulin, A., Tissier, R., Berdeaux, A., Connolly, J. B., . . . Ghaleh, B. (2009). The use of proteome similarity for the qualitative and quantitative profiling of reperfused myocardium. *Journal of Chromatography B, 877*(13), 1317-1326. doi:10.1016/j.jchromb.2008.10.024

Voelker, T., & Kinney, A. J. (2001). Variations in the biosynthesis of seed-storage lipids. *Annual Review of Plant Physiology and Plant Molecular Biology, 52*, 335-361. doi:10.1146/annurev.arplant.52.1.335

Voelker, T. A., Jones, A., Cranmer, A. M., Davies, H. M., & Knutzon, D. S. (1997). Broad-range and binary-range acyl-acyl-carrier protein thioesterases suggest an alternative mechanism for medium-chain production in seeds [Comparative Study]. *Plant Physiology, 114*(2), 669-677.

Vogel, J. P., Garvin, D. F., Mockler, T. C., Schmutz, J., Rokhsar, D., Bevan, M. W., . . . Int, B. I. (2010). Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature, 463*(7282), 763-768. doi:10.1038/nature08747

Vollmer, M., Nagele, E., & Horth, P. (2003). Differential proteome analysis: two-dimensional nano-LC/MS of E. coli proteome grown on different carbon sources. *Journal of Biomolecular Techniques, 14*(2), 128-135.

von Zychlinski, A., Kleffmann, T., Krishnamurthy, N., Sjolander, K., Baginsky, S., & Gruissem, W. (2005). Proteome analysis of the rice etioplast: metabolic and regulatory networks and novel protein functions. *Molecular & Cellular Proteomics, 4*(8), 1072-1084. doi:10.1074/mcp.M500018-MCP200

Waanders, L., Hanke, S., & Mann, M. (2007). Top-down quantitation and characterization of SILAC-labeled proteins. *Journal of the American Society for Mass Spectrometry, 18*(11), 2058-2064. doi:10.1016/j.jasms.2007.09.001

Wahid, M. B., Abdullah, S. N. A., & Henson, I. E. (2004, 26 September - 1 October). *Oil palm - achievements and potential*. Paper presented at the meeting of the Proceedings of the 4th International Crop Science Congress, Brisbane, Australia. Retrieved from http://regional.org.au/au/asa/2004/symposia/2/4/187_wahidmb.htm

Wan Omar, W. S., Willis, L. B., Rha, C., Sinskey, A. J., Ramli, U. S., Mat Yunus, A. M., . . . Sambanthamurthi, R. (2008). Isolation and utilization of acetyl-CoA carboxylase from oil palm (*Elaeis guineensis*) mesocarp. *Journal of Oil Palm Research, 2*(Special Issue on Malaysia-MIT Biotechnology Partnership Program), 97-107.

Wang, N., Xie, C., Young, J. B., & Li, L. (2009). Off-line two-dimensional liquid chromatography with maximized sample loading to reversed-phase liquid chromatography-electrospray ionization tandem mass spectrometry for shotgun proteome analysis. *Anal Chem, 81*(3), 1049-1060. doi:10.1021/ac802106z

Wang, W., Scali, M., Vignani, R., Spadafora, A., Sensi, E., Mazzuca, S., & Cresti, M. (2003). Protein extraction for two-dimensional electrophoresis from olive leaf, a plant tissue containing high levels of interfering compounds. *ELECTROPHORESIS, 24*(14), 2369-2375. doi:10.1002/elps.200305500

Wang, W., Vignani, R., Scali, M., & Cresti, M. (2006). A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *ELECTROPHORESIS, 27*(13), 2782-2786. doi:10.1002/elps.200500722

Wang, W., Vignani, R., Scali, M., Sensi, E., Tiberi, P., & Cresti, M. (2004). Removal of lipid contaminants by organic solvents from oilseed protein extract prior to electrophoresis. *Analytical Biochemistry, 329*(1), 139-141. doi:10.1016/j.ab.2004.02.044

Washburn, M. P., Wolters, D., & Yates III, J. R. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology, 19*(13), 242-247. doi:10.1038/85686

Wei, J., Zhang, Y., Wang, J., Tan, F., Liu, J., Cai, Y., & Qian, X. (2008). Highly efficient enrichment of phosphopeptides by magnetic nanoparticles coated with zirconium phosphonate for phosphoproteome analysis. *Rapid Communications in Mass Spectrometry, 22*(7), 1069-1080. doi:10.1002/rcm.3485

Weselake, R. J., Shah, S., Tang, M., Quant, P. A., Snyder, C. L., Furukawa-Stoffer, T. L., . . . Harwood, J. L. (2008). Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (Brassica napus) to increase seed oil content. *Journal of Experimental Botany, 59*(13), 3543-3549. doi:10.1093/jxb/ern206

Wessels, H. J. C. T., Vogel, R. O., van den Heuvel, L., Smeitink, J. A., Rodenburg, R. J., Nijtmans, L. G., & Farhoud, M. H. (2009). LC-MS/MS as an alternative for SDS-PAGE in blue native analysis of protein complexes. *PROTEOMICS, 9*(17), 4221-4228. doi:10.1002/pmic.200900157

Wienkoop, S., & Weckwerth, W. (2006). Relative and absolute quantitative shotgun proteomics: targeting low-abundance proteins in Arabidopsis thaliana. *Journal of Experimental Botany, 57*(7), 1529-1535. doi:10.1093/jxb/erj157

Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J.-C., . . . Hochstrasser, D. F. (1996). From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and arnino acid analysis. *Nature Biotechnology, 14*(1), 61-65. doi:10.1038/nbt0196-61

Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., & Williams, K. L. (1996). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology and Genetic Engineering Reviews, 13*, 19-50. doi:Retrieved from http://www.nottingham.ac.uk/ncmh/BGER/volume-13.html

Wilm, M. (2009). Quantitative proteomics in biological research. *PROTEOMICS, 9*(20), 4590-4605. doi:10.1002/pmic.200900299

Wise, R. R. (2006). The diversity of plastid form and function. In J. K. Hoober (Ed.), *The structure and function of plastids* (pp. 3-26). Dordrecht:Netherlands: Springer.

Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D. A., & White, F. M. (2007). Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling

networks. *Proceedings of the National Academy of Sciences of the United States of America, 104*(14), 5860-5865. doi:10.1073/pnas.0608638104

Wolters, D. A., Washburn, M. P., & Yates III, J. R. (2001). An automated multidimensional protein identification technology for shotgun proteomics. *Analytical Chemistry, 73*(23), 5683-5690. doi:10.1021/ac010617e

Wong, Y.-H., Lee, T.-Y., Liang, H.-K., Huang, C.-M., Wang, T.-Y., Yang, Y.-H., . . . Hwang, J.-K. (2007). KinasePhos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Research, 35*(suppl 2), W588-W594. doi:10.1093/nar/gkm322

Wu, S. S., Platt, K. A., Ratnayake, C., Wang, T. W., Ting, J. T., & Huang, A. H. (1997). Isolation and characterization of neutral-lipid-containing organelles and globuli-filled plastids from *Brassica napus* tapetum. *Proceedings of the National Academy of Sciences of the United States of America, 94*(23), 12711-12716. doi:Retrieved from http://www.pnas.org/content/94/23/12711.abstract

Wuyts, N., De Waele, D., & Swennen, R. (2006). Extraction and partial characterization of polyphenol oxidase from banana (Musa acuminata Grande naine) roots. *Plant Physiology and Biochemistry, 44*(5–6), 308-314. doi:10.1016/j.plaphy.2006.06.005

Xie, H., Pan, S., Liu, S., Ye, K., & Huo, K. (2007). A novel method of protein extraction from perennial *Bupleurum* root for 2-DE. *ELECTROPHORESIS, 28*, 871-875. doi:10.1002/elps.200600354

Xue, Y., Zhou, F., Zhu, M., Ahmed, K., Chen, G., & Yao, X. (2005). GPS: a comprehensive www server for phosphorylation sites prediction. *Nucleic Acids Research, 33*(suppl 2), W184-W187. doi:10.1093/nar/gki393

Yaffe, M. B., Leparc, G. G., Lai, J., Obata, T., Volinia, S., & Cantley, L. C. (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways [10.1038/86737]. *Nat Biotech, 19*(4), 348-353. doi:http://www.nature.com/nbt/journal/v19/n4/suppinfo/nbt0401_348_S1.html

Yang, L., Luo, Y., Wei, J., Ren, C., Zhou, X., & He, S. (2010). Methods for protein identification using expressed sequence tags and peptide mass fingerprinting for seed crops without complete genome sequences. *Seed Science Research, 20*(04), 257-262. doi:10.1017/S0960258510000243

Yates, J. R., Ruse, C. I., & Nakorchevsky, A. (2009). Proteomics by mass spectrometry: Approaches, advances, and applications. *Annual Review of Biomedical Engineering, 11*(1), 49-79. doi:10.1146/annurev-bioeng-061008-124934

Yatsu, L. Y., Jacks, T. J., & Hensarling, T. P. (1971). Isolation of spherosomes (oleosomes) from onion, cabbage, and cottonseed tissues. *Plant Physiology, 48*(6), 675-682. doi:10.1104/pp.48.6.675

Yean, G. P., & ZhiDong, L. (2012). A study on Malaysia's palm oil position in the world market to 2035. https://eneken.ieej.or.jp/data/4362.pdf

Yeung, Y.-G., Nieves, E., Angeletti, R. H., & Stanley, E. R. (2008). Removal of detergents from protein digests for mass spectrometry analysis. *Analytical Biochemistry, 382*(2), 135-137. doi:10.1016/j.ab.2008.07.034

Yocum, A. K., & Chinnaiyan, A. M. (2009). Current affairs in quantitative targeted proteomics: multiple reaction monitoring-mass spectrometry. *Briefings in Functional Genomics & Proteomics, 8*(2), 145-157. doi:10.1093/bfgp/eln056

Yoo, P., Ho, Y., Zhou, B., & Zomaya, A. (2008). SiteSeek: Post-translational modification analysis using adaptive locality-effective kernel methods and new profiles. *BMC Bioinformatics, 9*(1), 272.

Ytterberg, A. J., & Jensen, O. N. (2010). Modification-specific proteomics in plant biology. *Journal of Proteomics, 73*(11), 2249-2266. doi:10.1016/j.jprot.2010.06.002

Yu, L.-R., Zhu, Z., Chan, K. C., Issaq, H. J., Dimitrov, D. S., & Veenstra, T. D. (2007). Improved titanium dioxide enrichment of phosphopeptides from HeLa cells and high confident phosphopeptide identification by cross-validation of MS/MS and MS/MS/MS spectra. *Journal of Proteome Research, 6*(11), 4150-4162. doi:10.1021/pr070152u

Yu, Y.-Q., Gilar, M., Lee, P. J., Bouvier, E. S. P., & Gebler, J. C. (2003). Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. *Analytical Chemistry, 75*(21), 6023-6028. doi:10.1021/ac0346196

Yu, Y., He Mu, H., Mu-Forster, C., & Wasserman, B. P. (1998). Polypeptides of the maize amyloplast stroma. *Plant Physiology, 116*(4), 1451-1460. doi:10.1104/pp.116.4.1451

Yun, S. H., Park, G. W., Kim, J. Y., Kwon, S. O., Choi, C. W., Leem, S. H., . . . Kim, S. I. (2011). Proteomic characterization of the Pseudomonas putida KT2440 global response to a monocyclic aromatic compound by iTRAQ analysis and 1DE-MudPIT. *Journal of Proteomics, 74*(5), 620-628. doi:10.1016/j.jprot.2011.01.020

Yunus, A. M. M., & Parveez, G. K. A. (2008). Development of transformation vectors for the production of potentially high oleate transgenic oil palm. *Journal of Biotechnology, 11*(3). doi:10.2225/vol11-issue3-fulltext-7

Zabrouskov, V., Han, X., Welker, E., Zhai, H., Lin, C., van Wijk, K. J., . . . McLafferty, F. W. (2005). Stepwise deamidation of ribonuclease A at five sites determined by top down mass spectrometry†. *Biochemistry, 45*(3), 987-992. doi:10.1021/bi0517584

Zanzoni, A., Carbajo, D., Diella, F., Gherardini, P. F., Tramontano, A., Helmer-Citterich, M., & Via, A. (2011). Phospho3D 2.0: an enhanced database of three-dimensional structures of phosphorylation sites. *Nucleic Acids Research, 39*(suppl 1), D268-D271. doi:10.1093/nar/gkq936

Zeng, Y., Pan, Z., Ding, Y., Zhu, A., Cao, H., Xu, Q., & Deng, X. (2011a). A proteomic analysis of the chromoplasts isolated from sweet orange fruits [Citrus sinensis (L.) Osbeck]. *Journal of Experimental Botany, 62*(15), 5297-5309. doi:10.1093/jxb/err140

Zeng, Y., Pan, Z., Ding, Y., Zhu, A., Cao, H., Xu, Q., & Deng, X. (2011b). A proteomic analysis of the chromoplasts isolated from sweet orange fruits [Citrus sinensis (L.) Osbeck]. *Journal of Experimental Botany, 62*(15), 5297-5309. doi:10.1093/jxb/err140

Zhang, X., Asara, J. M., Adamec, J., Ouzzani, M., & Elmagarmid, A. K. (2005). Data pre-processing in liquid chromatography–mass spectrometry-based proteomics. *Bioinformatics, 21*(21), 4054-4059. doi:10.1093/bioinformatics/bti660

Zhao, Z., Stanley, B. A., Zhang, W., & Assmann, S. M. (2010). ABA-regulated G protein signaling in *Arabidopsis* guard cells: A proteomic perspective. *Journal of Proteome Research, 9*(4), 1637-1647. doi:10.1021/pr901011h

Zheng, Q., Song, J., Doncaster, K., Rowland, E., & Byers, D. M. (2007). Qualitative and quantitative evaluation of protein extraction protocols for apple and strawberry fruit suitable for two-dimensional electrophoresis and mass spectrometry analysis. *Journal of Agricultural and Food Chemistry, 55*(5), 1663-1673. doi:10.1021/jf062850p

Zhou, H., Xu, S., Ye, M., Feng, S., Pan, C., Jiang, X., . . . Zou, H. (2006). Zirconium phosphonate-modified porous silicon for highly specific capture of phosphopeptides and MALDI-TOF MS analysis. *Journal of Proteome Research, 5*(9), 2431-2437. doi:10.1021/pr060162f

Zhou, S., Sauvé, R., & Thannhauser, T. W. (2009). Proteome changes induced by aluminium stress in tomato roots. *Journal of Experimental Botany, 60*(6), 1849-1857. doi:10.1093/jxb/erp065

Zhou, S., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., . . . Cantley, L. C. (1993). SH2 domains recognize specific phosphopeptide sequences. *Cell, 72*(5), 767-778.

Zubarev, R. A., Kelleher, N. L., & McLafferty, F. W. (1998). Electron capture dissociation of multiply charged protein cations. A nonergodic process. *Journal of the American Chemical Society, 120*(13), 3265-3266. doi:10.1021/ja973478k

Zulak, K. G., Lippert, D. N., Kuzyk, M. A., Domanski, D., Chou, T., Borchers, C. H., & Bohlmann, J. (2009). Targeted proteomics using selected reaction monitoring reveals the induction of specific terpene synthases in a multi-level study of methyl jasmonate-treated Norway spruce (Picea abies). *The Plant Journal, 60*(6), 1015-1030. doi:10.1111/j.1365-313X.2009.04020.x

Zybailov, B., Friso, G., Kim, J., Rudella, A., Rodríguez, V. R., Asakura, Y., . . . van Wijk, K. J. (2009). Large scale comparative proteomics of a chloroplast Clp protease mutant reveals folding stress, altered protein homeostasis, and feedback regulation of metabolism. *Molecular & Cellular Proteomics, 8*(8), 1789-1810. doi:10.1074/mcp.M900104-MCP200

Zybailov, B., Rutschow, H., Frisco, G., Rudella, A., Emmanuelsson, O., Sun, Q., & Van Wijk, K. J. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE, 3*(4), 1-19. doi:10.1371/journal.pone.0001994

1.