MONASH University

Genetic determinants of rice bran oil quantity and quality: genomic, proteomic and transcriptomic approaches

By

Gopal Ji Tiwari

THESIS

Submitted to the School of Science

Monash University Malaysia

In Partial Fulfilment of the Requirement for the Degree of

DOCTOR OF PHILOSOPHY (PhD)

MONASH UNIVERSITY MALAYSIA
October, 2016

Copyright notice

Dedicated to my

Late Dadi-Baba

&

Loving parents

Table of contents

Acknowledgements

Completion of this doctoral dissertation was possible with the support of several people. Foremost, I would like to express my special appreciation and thanks to my supervisor Prof. Sadequr Rahman for his patience, motivation, enthusiasm and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I feel blessed and could not have imagined having a better advisor and mentor for my PhD study.

I also thank my co-supervisors Dr. Song Beng Kah and Dr. Juan Joon Ching for their constant support, constructive comments and encouragement throughout my PhD journey.

I thank all my current and previous lab mates for maintaining the friendly atmosphere in the lab and also for their suggestions, critical comments during weekly lab meetings which helped me to enhance my presentation skills. Also my sincere thanks to all the academic and administrative staff members in School of Science for being so cooperative and nice to me, making me feel very comfortable throughout my PhD tenure.

I would like to convey my gratitude to Monash University Malaysia, for the scholarship, research funds and selecting my application for the platform mobility scheme. I would also like to thank Dr. David Steer for his guidance while I carried out part of my research work in Biomedical Proteomics Facility lab, Monash University, Australia.

A good support system is important to stay joyful and have lots of fun while studying. I was lucky to have Alviya Sultana, Snigdha Tiash and Uday Kundap from Monash University for being my stress busters and sharing lots of memorable moments together. Also, very heartfelt thanks to my friends Madhur, Sainath Nair and Naveen with whom I explored Malaysia and for making me feel a home away from home.

Finally I would like to thank Lord Shiv-shiva and my entire family specially my parents for believing in me and supporting me in my tough times with their love, care and emotional support.

PART A: General Declaration

Monash University

Declaration for thesis based or partially based on conjointly published or unpublished work

General Declaration

In accordance with Monash University Doctorate Regulation 17.2 Doctor of Philosophy and Research Master's regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 3 original paper published in peer reviewed journal. The core theme of the thesis is to identify new genetic targets for further improvement of rice bran oil. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself the candidate, working within the School of Science under the supervision of Prof. Sadequr Rahman.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of **chapters 1, 2 and 4** my contribution to the work involved the following:

Thesis chapter	Publication title	Publication status*	Nature and extent of candidate's contribution
1	Improving the economic value of rice bran oil.	Published	Review of literature, data collection and manuscript preparation
2	Lipase genes expressed in rice bran: LOC_Os11g43510 encodes a novel rice lipase.	Published	Experimental design and conduct, samples collection and process, data collection, result acquisition,

			statistical analysis, manuscript preparation
4	RNAi-mediated down-regulation of the expression of OsFAD2-1: effect on lipid accumulation and expression of lipid biosynthetic genes in the rice grain.	Published	Experimental design and conduct, samples collection and process, data collection, result acquisition, statistical analysis, manuscript preparation
3.1	Identification of Os01g0817700 as a new rice lipase gene by fluorescence based proteomics approach using methylumbelliferyl-derivative substrate.	Advanced draft ready	Experimental design and conduct, samples collection and process, data collection, result acquisition, statistical analysis, manuscript preparation

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed: ……  ……

Date: **……27/10/2016……**

Abstract

Any increase in the economic value of the rice crop could have a large impact on the economies of Asia. Rice bran can be used to produce rice bran oil (RBO) which is emerging as a major vegetable oil. In this research, new genetic targets to improve RBO quality have been investigated using genomic, proteomic and transcriptomic approaches. The first chapter gives a detailed background about rice grain structure, effect of lipases on lipids in rice and strategies to increase the quantity and quality of oil. In the second chapter, using the genomic approach 125 putative lipase gene sequences derived from *Oryza sativa ssp*. japonica genome were systematically analyzed using bioinformatics tools. Based on the analysis, LOC_Os11g43510 was selected and it was experimentally demonstrated to be highly expressed in rice bran. The transcript from LOC_Os11g43510 was cloned and expressed as a secretory protein in *P. pastoris* X-33 which retained demonstrated lipase activity.

The third chapter describes the use of a proteomics approach based on fluorescence based lipase activity probes and liquid chromatography-tandem mass spectrometry (LC/MS/MS) detection analysis. The proteomics approach used identified Os01g0817700 as a novel putative lipase in rice bran. Bioinformatics and expression studies also suggested that Os01g0817700 has two lipase motifs and is also expressed in the bran and embryo of the rice grain. These data suggest that Os01g0817700 is a novel putative lipase expressed in rice.

In the fourth chapter, a transcriptomic approach was used to study the subset of lipid-metabolism genes being affected when *OsFAD2-1* is down-regulated by RNAi to produce high oleic (HO) rice. The transcriptomic analysis in the HO rice suggests that a suite of key genes (FatA, LACS, SAD) involved in fatty acid biosynthesis are concurrently down-regulated. Moreover, a decrease in the expression of oil body proteins (caleosin and steroleosin) was also observed in the HO lines. All of these genes are suitable targets for gene manipulation in order to further increase the oleic acid content.

Very few plant lipases have been studied and reported so far. This study demonstrates an alternative route of identifying lipases in rice bran and also helps to identify additional genes involved in determining rice oil composition. Overall, this study has used different approaches to identify new genetic targets for further improvement of RBO.

Chapter 1

Review of Literature

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter 1

Declaration by candidate

In the case of **Chapter 1**, the nature and extent of my contribution to the work was the following:

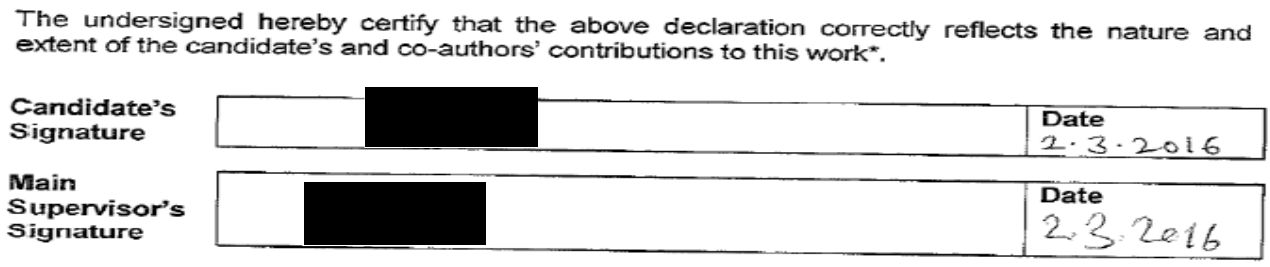
Nature of contribution	Extent of contribution (%)
Review of literature, data collection and manuscript writing	80

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Ayesha Aumeeruddy	Editing of Manuscript	N/A (not a student registered under Monash University)
Sadequr Rahman	Manuscript preparation and submission	N/A (not a student registered under Monash University)

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.



*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

1.1 Structure of rice grain

Rice is one of the most important crops for mankind. It feeds nearly half the world's population and accounts for more than 50% of their daily calorie intake (Maclean and Dawe, 2002). The rice grain is a seed of the monocot grass *Oryza sativa*, having a genome size of 430MB (Arumuganathan and Earle, 1991). The mature rice kernel obtained from the plant (also known as paddy), consists of an outer protective layer known as the hull made up of two parts, the lemma (outer) covering the dorsal portion and the palea (inner) covering the ventral part of the seed (Juliano and Bechtel, 1985).On de-husking paddy, the hull is removed and brown rice is obtained. Brown rice consists of the pericarp, the seed-coat and nucellus, the germ or embryo, and the endosperm (Fig.1). Any pigment is usually confined to the pericarp (Juliano and Bechtel, 1985) although rice varieties with black endosperm do exist (Abdel-Aal et al., 2006). The hull contributes (20 %), bran and germ (10 %), and starchy endosperm (70 %) to a mature rice grain or kernel mass (Orthoefer, 1996).

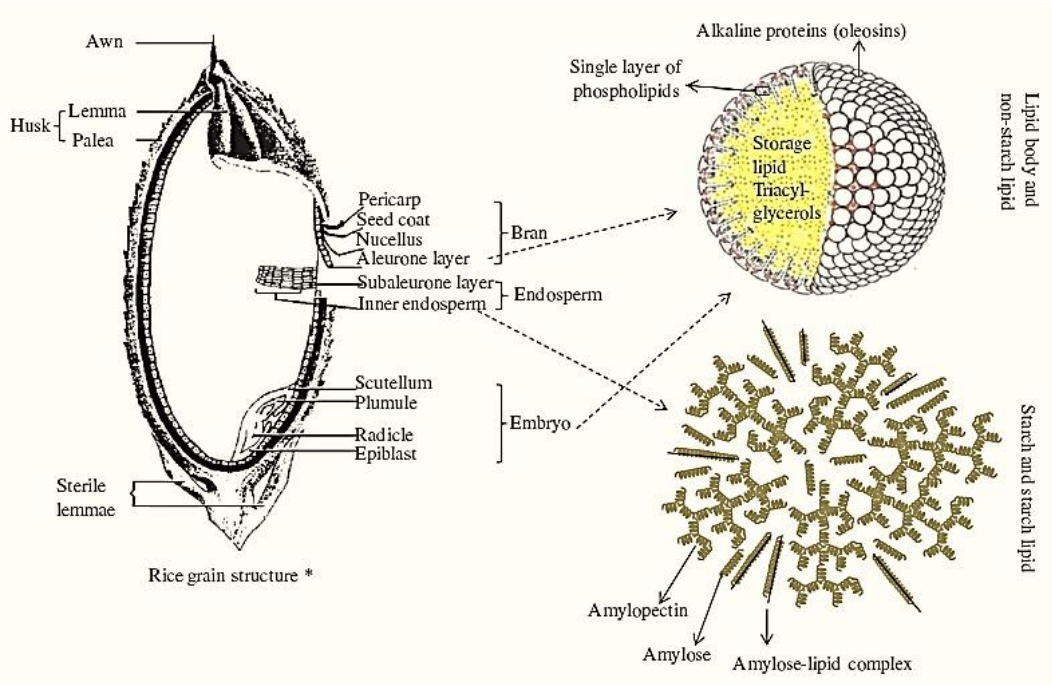


Fig.1. Structure of rice kernel (Adapted from Liu et al., 2013)

1.2 Rice milling products

Milling is an important step as it transforms paddy to white rice, giving it a good appearance for marketing and consumption. The main objective of milling is to separate white rice from hull and bran. White rice is generally used for eating and cooks faster than brown rice and it also can be stored for a longer time. Hulls are mostly used as animal feed and underutilized but food industries are showing interest in them as a dietary supplement. The bran is the most nutritious part of grain having a higher concentration of proteins, fats, vitamins and minerals than in the starchy endosperm. The removal of bran leads to loss of 90% of the nutrition and fat stored in germ (embryo) of rice grain. Table 1 shows the comparison of constituents among white rice, bran and hull.

Table 1. Nutrients among white rice, bran and hull of rice (Rice: Chemistry and Technology 3rd edition).

% in each fraction	White	Bran	Hull
Protein	4.5-10.5	11.3-14.9	2-2.8
Crude Fat	0.3-0.5	15-19.7	0.3-0.8
Available Carbs	77-89	34-62	22-34
Crude Fiber	0.2-0.5	7-11.4	34.5-45.9
Ash	0.3-0.8	6.9-9.9	13.2-21

The by-products in rice milling are rice hull, rice germ and bran layers, and fine broken grains. Broken grains are mixed with milled rice. Finally, the finished rice product is packed for marketing (Champagne, 2004).

1.3 Lipids in rice grain

Apart from starch, rice grains contain a much smaller proportion of lipids which contributes to processing and nutritional properties (Moazzami et al., 2011). Almost all of the lipids in rice are located in the outer layers of the grain, and are found in the bran which is a valuable by product of milling and contains a high concentration of other nutritional compounds as well. The bran

fraction also includes the germ or embryo. Rice bran oil is in steady demand as a so-called "healthy oil" and is a popular cooking oil in several in Asian countries (Sugano and Tsuji, 1997, Ghosh, 2007). From a nutritional point of view, the interest in rice bran oil has been growing, mainly because of its health benefits which include a reduction in both serum and LDL cholesterols in those who consume it (Wilson et al., 2000).

1.3.1 Distribution and accumulation of lipid in rice grain

Oil comprises 60% by weight in oil seeds like rapeseed, mustard, soybean whereas cereal grain lipid content is very low (Ohlrogge and Jaworski, 1997). The endosperm mostly contains free fatty acids (Fujino, 1978). Milled rice from which the bran layer is polished has 0.2 to 2% lipid when compared to brown rice having about 1-4% lipid (Juliano and Bechtel, 1985).

Oleic acid, linoleic acid and palmitic acid are major fatty acids present in rice bran oil. The content of palmitic acid (15.20%- 19.56%), oleic (37.9%-47.5%) and linoleic (38.2% - 30.4%) acid depends on the cultivator (Taira et al., 1988). A variation in the proportion of specific fatty acids is seen during lipid accumulation in rice grain which occurs up to 12 days after flowering. Oleic and linoleic acid increases up to 16 days after flowering. Palmitic acid remains constant and the portion of α-linoleic acid decreases (Choudhury and Juliano, 1980).

1.3.2 Comparison of oil content in colored rice grains

All unprocessed rice is one of three different colours brown, red and purple or black rice. The colour of the grain is due to the colour of the pericarp. Polishing grain to remove the pericarp and other outer layers produces white rice and the bran. However, sometimes in black rice the anthocyanins producing the colour can diffuse throughout the endosperm and in that case, even polished rice will appear black. Most of the rice consumed is white but the most nutritious part of the grain has been already removed by polishing.

The oil content of different coloured rice has been found to vary. Lipid content is found to be highest in purple or black rice (12-13%) while brown rice is ranging from 2.76-3.84 % on a dry weight basis (Frei and Becker, 2005). Phytochemicals like tocopherol, tocotrienol, and γ-oryzanol in the rice bran have been thoroughly studied (Dykes and Rooney, 2007, Liu, 2007,

Zhang et al., 2010). Like any other phytochemicals, the content of phenolics and flavonoids also depends on quality traits (grain color, size and weight) of rice grain. Black rice has been reported to have the highest content of phenolics, flavonoids also with maximum and antioxidant activity followed by red rice and lowest in white rice (Zhang et al., 2010).

1.4 Rice bran oil

1.4.1 Rice bran oil composition

In comparison with other vegetable oils, crude rice bran oil tends to contain higher levels of non-triglyceride components, most of which are removed during further refining processes (www.Ricebranoil.info). Table 3 compares the fatty-acid composition of rice bran oil with that of peanut, soybean, and cottonseed oils.

Table 3. Comparison of fatty-acid composition among selected oils. (Data for rice bran oil are from Riceland Foods, Inc.)

Fatty acid (%)	% Rice bran	% Peanut	% Soybean	% Cottonseed
Myristic (14:0)	0.2	0	0.2	0.8
Palmitic (16:0)	15.0	8.1	10.7	27.3
Stearic (18:0)	1.9	1.5	3.9	2.0
Oleic (18:1)	42.5	49.9	22.8	18.3
Linoleic (18:2)	39.1	35.4	50.8	50.5
Linolenic (18:3)	1.1	0	6.8	0
Arachidic (20:0)	0.5	1.1	0.2	0.3
Behenic (22:0)	0.2	2.1	0.1	0

The lipid components and fatty acid (FA) distribution of different acyl lipids from five different rice bran cultivars were compared. Triacylglycerides (TAG) make up about 85% of the total lipid, followed by phospholipids (~ 6.5%) and free fatty acids (~4.5%) (Yoshida et al., 2011).

1.4.2 Limitations on the use of rice oil

In spite of being a potential raw material for nutraceuticals or functional food, rice bran is underutilized due to the presence of certain enzymes, particularly lipase, but also lipoxygenase and peroxidase which affect the quality and shelf life of oil in the rice bran (Gong et al., 2013). Thus most of the rice bran is discarded or used as animal feed (Hu et al., 2009).

Milling, which produces the bran, ruptures the cells, mixes the cellular contents, and initiates hydrolytic, autolytic, and oxidative degradation of the oil is often accompanied by chemical degradative effects (Becker, 2007). Hydrolytic degradation of lipids usually takes by the action of lipases in breaking down TAG into FFA and glycerol. Free fatty acids produced by the hydrolysis reaction make rice bran unsuitable for edible use (Ju and Vali, 2005). On the other hand, lipid oxidation is a free radical chain reaction between unsaturated fats and oxygen. Oxidation can be non-enzymatic or autolytic. In autocatalytic oxidation incorporation of the oxygen molecule in lipids is catalyzed by free radicals whereas, in enzymatic oxidation, when enzymes like lipoxygenases oxidize lipids leading to volatile lipids and rancidity (Gutteridge, 1995, Chaiyasit et al., 2007). Crude rice bran oil is rapidly degraded due to high levels of very active lipases and the activity of these needs to be reduced in order to better utilize the oil. Detailed studies about the identification and characterization of lipases in rice are presented in chapters two and three of this thesis and provides new genomic and proteomic approaches to study lipases in plants.

1.4.3 Rice bran stabilization

As indicated above, the shelf life of freshly milled rice bran is very short as lipases endogenous to the bran cause decomposition of lipids (TAG) into free fatty acids (FFA), making it unsuitable for human consumption and decreasing its economic value for oil extraction. Dry heating, wet heating, and extrusion are a few methods which are used for the stabilization of rice bran (Sayre et al., 1982). Refrigeration, lowering pH and chemicals such as sodium metabisulfite have also been introduced to decrease lipase activity and promote stabilization of bran (Cheruvanky et al.,

2003).Although these methods are helpful in increasing shelf life of rice bran, there are still quality problems, and longer term solutions to the problems are required. Therefore new methods to reduce lipase activity in rice bran or to develop new lipase deficient transgenic rice lines are warranted. Thus the first step would be to identify and study lipases present in rice bran.

1.5 Lipases in rice bran

Most of the characterization of lipases in rice have been at the protein level. Aizono et al. (1976) and Fujiki et al. (1978), identified two soluble rice bran lipases designated Lipase I and Lipase II. Lipase I was found to be a 40kDa protein which cleaves fatty acids from positions sn-1 and sn-3 of the triacylglycerol. Lipase II has a molecular mass of 34kDa with an optimum pH of 7.5. In 2001, the first thermally stable rice bran lipase was identified and characterized; it was found to be a 9.4kDa polypeptide that showed maximum catalytic activity at 80°C, pH 11.0 and showed phospholipase A2 properties (Bhardwaj et al., 2001).

Few rice bran lipases have been studied at the gene level, and a limited number of genes have been functionally identified as lipases through the activity of the encoded protein. Lipase II (also called RBL) has been transformed into *E.coli* strain Rosetta (DE3) pLysS and *Pichia pastoris*. It was found that after purification the crude recombinant lipase showed maximum lipase activity on triacetin or other short chain triacylglycerol (Vijayakumar and Gowda, 2013).

1.6 Strategies to increase the quantity and quality of oil

To further increase the commercial value and acceptance of RBO among consumers the quality and quantity of RBO must be enhanced. Addressing the increase in demand of plant oil is therefore one of the main challenges for oilseed biotechnology. One way to achieve this would be to increase the oil content in seed crops via biotechnological approaches (Drexler et al., 2003, Lu et al., 2011). There is a demand for improving the fatty acid profile of vegetable oils, especially in the food industry. Recent advances in metabolic engineering have enabled the production of fish-type omega-3 fatty acids by oilseeds (Ruiz-Lopez et al., 2014, Adarme-Vega et al., 2014, Ruiz-Lopez et al., 2015). There are therefore the twin requirements of producing vegetable oils in greater quantity and of improved quality for specific purposes in the future.

1.6.1 Synthesis and degradation of TAG in seeds

As oilseeds contain far higher amounts of oils than cereals, lipid biosynthesis and accumulation in plants has been mostly studied in oilseeds.

Triacylglycerides (TAG) are the major storage lipids in oilseeds which are utilized during seed germination for plant growth and development (Li et al., 2012). In order to modify oilseed lipids knowledge of seed TAG biosynthesis and metabolism is required (Fig. 2). There are several biochemical factors which affect the fatty acid composition in TAG (Lu et al., 2011). In the classical pathway fatty acids (mostly oleic acid (18:1) and less palmitic acid (16:0) and, stearic acid (18:0))are transported from plastids and are further desaturated or modified by phosphatidylcholine (PC) before entering TAG (Bates et al., 2009). In contrast, the *de novo* pathway involves two sequential acylations of glycerol-3-phosphate to produce phosphatidic acid (PA). Later diacylglycerol (DAG) acts on PA and converts it into PC involving PA phosphatases and CDP-choline:DAG choline phosphotransferase (CPT). Later,FAD2 and FAD3 acts on PC and produce polyunsaturated molecular species by desaturation of 18:1 to linoleic acid (18:2) and finally into linolenic acid (18:3) (Bates et al., 2007, Bates et al., 2009, Williams et al., 2000). During germination, triacylglycerols stored in “oil bodies” or “oleosomes” are quickly used in the production of energy for the synthesis of the sugars, amino acids (mainly asparagine, aspartate, glutamine and glutamate) and carbon chains required for embryonic growth (Huang et al., 1988, Borek et al., 2006, Ejedegba et al., 2007, Quettier and Eastmond, 2009).

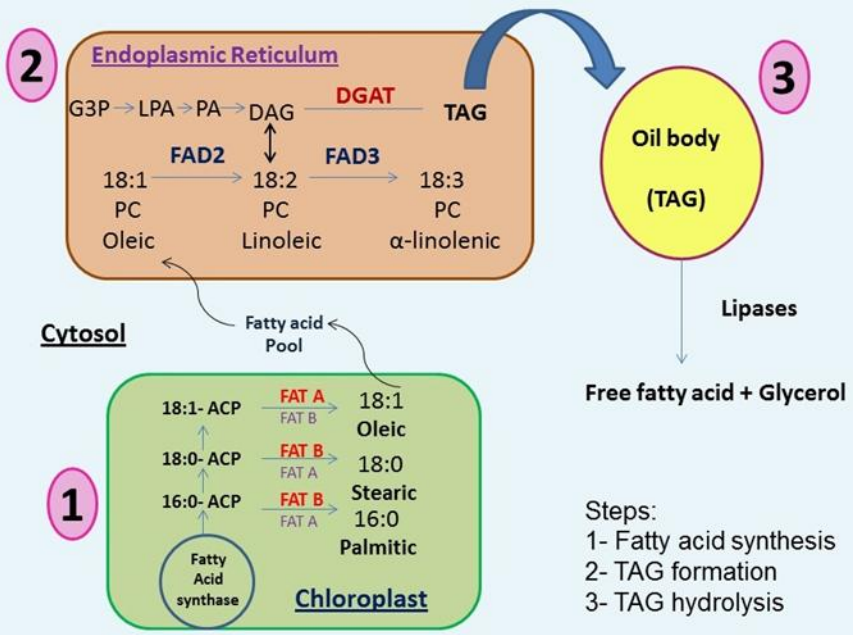


Fig. 2. Overview of lipid biosynthesis and its breakdown. ACP, acyl carrier protein, FAT A, Acyl ACP thioesterase A; FAT B, Acyl-ACP thioesterase B; FAD , fatty acyl desaturases, PC phosphatidylcholine, DGAT, acyl-CoA: DAG acyltransferase, G3P, glycerol 3-phosphate, LPA, lyso-phosphatidic acid, PA phosphatidic acid, DAG, diacylglycerol, TAG triacylglycerol.

1.6.2 Genes involved in improving oil quantity and quality: oilseeds

Diacylglycerolacyltransferase (DGAT; EC 2.3.1.20), is the key enzyme related to triglyceride synthesis (Chen and Farese, 2000) and is associated with seed germination and oil accumulation. DGAT is a membrane bound protein located in the endoplasmic reticulum (Kennedy, 1961). DGAT1 and DGAT 2 are the most common isoforms present in animals (Cases et al., 1998, Cases et al., 2001) and plants (Shockey et al., 2006, Lardizabal et al., 2001). Plants and animals which lack DGAT expression are found to have lower TAG accumulation (Zou et al., 1999, Stone et al., 2004, Smith et al., 2000). DGAT acts in the final and committed step of TAG synthesis so is therefore considered as a rate-limiting enzyme in plant storage lipid accumulation (Norton and Harris, 1975, Ichihara et al., 1988, Turchetto-Zolet et al., 2011).There are two types of DGATs in oilseeds and grasses with different roles. The DGAT1 gene has been cloned in Arabidopsis (Lu et al., 2003) and has been shown to be a key enzyme in TAG accumulation in developing seeds (Cahoon et al., 2007, Zheng et al., 2008, Zhang et al., 2009). DGAT2 was determined to regulate the incorporation of unusual FAs such as epoxy or hydroxyl FAs in TAGs

in Stokesia and Vernonia (Cahoon et al., 2007) to limit their accumulation (which might cause membrane dysfunction) (Beaudoin and Napier, 2004, Coleman and Lee, 2004).Phospholipid:diacylglycerolacyltransferase (PDAT; E.C.2.3.1.158) is another class of enzyme responsible for acyl-CoA independent conversion of di-acyl glycerol (DAG) to tri-acyl glycerol (TAG) (Abbadi et al., 2004, Beaudoin and Napier, 2004). The PDAT route could be a mechanism of incorporation of unusual fatty acids (acyl editing) in plants such as *Crepispalaestina*and *Ricinuscommunis* (Xu et al., 2012). DGAT and PDAT have overlapping functions as double mutation of *dgat1* and *pdat1* resulted in sterile pollen and embryo; suggesting that DGAT1 and PDAT1 are the primary enzymes for oil accumulation in Arabidopsis seeds (Zhang et al., 2009).

1.6.3 Enhancing the quality of oil: oilseeds

Polyunsaturated fatty acids (PUFAs) which are composed of 18- carbons are precursors required for synthesizing 20- and 22-carbon PUFAs (e.g. arachidonic, eicosapentaenoic, and docosahexaenoic acids) which act as substrates for the synthesis of prostaglandins, leukotrienes, and other important signaling molecules (Lu et al., 2009). In addition to this, saturated fatty acids are regarded as unhealthy, because of their cholesterol raising effects. Mono and polyunsaturated fatty acids are much better from this point of view. However, polyunsaturated fatty acids are unstable and lead to oxidative instability and rancidity. The most desirable oils are therefore those that are high in monosaturated fatty acids (Shahidi and Wanasundara, 2002, Kamal-Eldin, 2006).

A key gene involved in the production of high oleic acid lines is the fatty aciddesaturase 2 (FAD2) gene. The function of fatty acid desaturases is catalyzing the conversion of linoleic acid (18:2) to linolenic acid (18:3) in the pathway shown in Fig.2. FAD2, is a delta 12 desaturase which converts the monounsaturated fatty acid oleic acid (18:1Δ9) to linoleic acid (18:2Δ9, Δ12), which can be subsequently desaturated to α-linolenic acid (18:3Δ9, Δ12, Δ15) by FAD3 (Voelker and Kinney, 2001).

Soybean lines carrying both homozygous insertion/deletion mutant (indel) *FAD2-1A* alleles offer a simple route for the development of high oleic acid commercial soybean varieties (Pham et al.,

2011). Two high oleic acid soybean (*Glycine max.* (L.) Merr.) mutant lines, 'M23' and 'KK21'with >80% oleic acid content were developed by combining *GmFAD2-1a* and *GmFAD2-1b* mutant alleles. Both carry a unique non-functional mutant allele of GmFAD2-1a which is a member of GmFAD2 gene family (*GmFAD2 1a, GmFAD2 1b, GmFAD2 2a, GmFAD2 2b*) (Takagi and Rahman, 1996, Anai et al., 2008).

The monounsaturated FFA play an important role in many physiological processes, including cell membrane function and the development and functioning of the nervous system. An increase in monounsaturated fatty acids would also increase the utility of rice bran oil as a feedstock biodiesel (Fallen et al., 2011).

Elevation in oleic acid level and a reduction of saturated fatty acids level in soybean can also be achieved by combining ribozyme-mediated suppression of FAD2-1 and palmitoyl-thioesterase encoding *FatB* gene (Buhr et al., 2002). The *FatB* gene family could also be engineered to achieve over 80% oleic acid content in soybean (Hoshino et al., 2010). The function of the palmitoylthioesterase is to remove palmitic acid from further extension in the pathway.

The low oxidative stability in soybean oil is due to the presence of high levels of polyunsaturated FA, which leads to rancidity, reduces the shelf-life of food products as well as increases the viscosity of soy-based biodiesels (Canakci et al., 1999). High oleic acid content in soybean lowers saturates, increases monounsaturated fatty acid content above that present in commodity soybean oil, making it attractive from a nutritional standpoint (Bilyeu et al., 2003).

Gene technology and plant breeding are combining to provide alterations in the proportions of individual fatty acids of oilseeds to improve their nutritional value. A number of high-oleic oils have been developed in order to provide high stability cooking oils (Liu et al., 2002). Fatty acid desaturases (FAD-s) play a prominent role in plant lipid metabolism.

1.6.4 Enhancing oil quality in cereals

In maize, silencing of the FAD2 gene which catalyzes the conversion of oleic acid (18:1) to linoleic acid (18:2) resulted in oil with an increased oleic acid to linoleic acid ratio which helped in enhancing oxidative stability of the oil (Jaworski and Cahoon, 2003). Down regulation of the

FAD2 gene in rice seed led to an overall increase in oleic acid with a decrease in palmitic acid and linoleic acid (Zaplin et al. 2013). Such altered oleic acid to palmitic acid ratio in rice seeds should lead to healthier and more stable rice bran oil.

1.7 Role of oleosins in lipid accumulation

Oleosins are proteins ranging in size from 15 to 26 kDa (Murphy and Ross, 1998, Hsieh and Huang, 2004) embedded in outer phospholipid monolayer of spherical oleosomes (0.6–2 µm in diameter) which help in storing TAG in plants (Huang, 1992, Napier et al., 1996). Apart from oil bodies oleosins have also been found in the pollen tube of Arabidopsis (Kim et al., 2002). The main role of oleosin is maintaining stability of oil bodies (Tzen and Huang, 1992, Huang, 1992), long-term storage and mobilization of oil bodies (Murphy and Vance, 1999). Overexpression of the oleosin 3 (OLE3) gene in *Saccharomyces cerevisiae* suggested that oleosins have dual roles, both acting as monoacylglycerolacyltransferases (MGAT) and phospholipases (Parthibane et al., 2012). Oleosins are usually present in two isoforms (Tzen et al., 1990); two oleosins from inbred maize line (16 kDa and 18 kDa) (Vance and Huang, 1987), one oleosin in Brassica (19 kDa) (Lee et al., 1991), two oleosins from soybean (both 24 kDa) (Kalinski et al., 1991) and two oleosin from rice (16 and 18 kDa) (Wu et al., 2010) and barley (Aalen, 1995).

Transgenic rice lines have been developed using RNA interference (RNAi) where two oleosins isoforms (OLE16 and OLE18) were down regulated and showed a high impact on the structure of the oil bodies; the overall TAG content of rice seed was also decreased (Wu et al., 2010). This suggests oleosins are crucial for the stability of oil bodies and for maintaining their proper structure inside cereal grains. In contrast, overexpression of two embryo–specific soybean oleosins in rice led to 36.93 and 46.06 % increase of lipid content in the transgenic rice seeds compared to the non-transgenic control (Whitelaw et al., 2013). In addition overexpression of oleosins in Arabidopsis enhanced the accumulation of hydroxyl fatty acids in transgenic Arabidopsis seeds (Lu et al., 2006). In a recent study co-expression of PDAT1 and oleosin led to the overall enhancement of TAG synthesis in Arabidopsis leaves without affecting plant growth and membrane lipid composition. This suggests that PDAT1 can play a crucial role in engineering high levels of TAG accumulation in green biomass (Fan et al., 2013). In maize the

transcription factor LEAFY COTYLEDON1 (LEC1) under the control of the oleosin promoter elevated levels of seed oil (increased by48%) due to increased storage lipids in the maize embryo (Shen et al., 2010).

1.8 Transcriptomic approaches to identifying other genes involved in regulating oleic acid content

With the development of functional genomics, large-scale transcriptome analysis-based studies can be done using expressed sequence tag (EST) library sequencing, microarray hybridization and serial analysis of gene expression (SAGE). In the past, these techniques had been used in analysis of the transcriptome from Arabidopsis, maize and rice (Prioul et al., 2008, Teoh et al., 2013, Shankar et al., 2014, Wickramasuriya and Dunwell, 2015). However, in recent years, more advanced techniques like next generation high-throughput RNA sequencing technology (RNA-Seq) have been developed to determine expression profiles, and quantify RNA transcripts with higher sensitivity (Nagalakshmi et al., 2008, Ozsolak and Milos, 2011). Understanding the cellular and molecular mechanisms of storage lipid production in plants could help develop transgenic plants with tailor-made fatty acid composition and amount (Cahoon et al. 2007) by circumventing any bottlenecks involved in TAG biosynthesis.

1.9 Objectives of the thesis

This thesis reports the investigation of the genetic determinants of rice oil quality. Rice oil quality can be defined in different ways. One important aspect is the proportion of oil that is recovered as triglycerides. This is because if the triglycerides are degraded into constituent glycerol and fatty acids then the value of the oil is sharply reduced. So this thesis reports on research on the lipases that are present in rice bran using two different approaches. A second important aspect of rice oil quality is the composition of the fatty acids that make up the triglycerides. Here it is generally important to maximize the proportion of oleic acids. This research has taken advantage of an existing high oleic acid rice line in order to identify additional genes that may be critical for further increasing the oleic acid content.

The specific questions that the thesis has investigated are:

1) Is it possible to identify novel lipases that are expressed in rice bran by analyzing the genome of rice?

2) Is it possible to use flurogenic lipase activity probes and proteomics for the identification of active lipases in the rice bran?

3) Does a transcriptomic analysis of a *FAD2-1* RNAi high oleic rice line allow one to identify additional genes involved in determining rice oil composition?

1.10 References

Aalen, R. B. 1995. The transcripts encoding two oleosin isoforms are both present in the aleurone and in the embryo of barley (Hordeum vulgare L.) seeds. *Plant molecular biology,* 28**,** 583-588.

Abbadi, A., Domergue, F., Bauer, J., Napier, J. A., Welti, R., Zähringer, U., Cirpus, P. & Heinz, E. 2004. Biosynthesis of very-long-chain polyunsaturated fatty acids in transgenic oilseeds: constraints on their accumulation. *The Plant Cell,* 16**,** 2734-2748.

Abdel-Aal, E.-S. M., Young, J. C. & Rabalski, I. 2006. Anthocyanin composition in black, blue, pink, purple, and red cereal grains. *Journal of Agricultural and Food Chemistry,* 54**,** 4696-4704.

Adarme-Vega, T. C., Thomas-Hall, S. R. & Schenk, P. M. 2014. Towards sustainable sources for omega-3 fatty acids production. Current opinion in biotechnology, 26, 14-18.

Aizono, Y., Funatsu, M., Fujiki, Y. & Watanabe, M. 1976. Purification and Characterization, of Rice Bran Lipase II. *Agricultural and Biological Chemistry,* 40**,** 317-324.

Anai, T., Yamada, T., Hideshima, R., Kinoshita, T., Rahman, S. M. & Takagi, Y. 2008. Two high-oleic-acid soybean mutants, M23 and KK21, have disrupted microsomal omega-6 fatty acid desaturase, encoded by GmFAD2-1a. *Breeding Science,* 58**,** 447-452.

Arumuganathan, K. & Earle, E. 1991. Nuclear DNA content of some important plant species. *Plant molecular biology reporter,* 9**,** 208-218.

Bates, P. D., Durrett, T. P., Ohlrogge, J. B. & Pollard, M. 2009. Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. *Plant physiology,* 150**,** 55-72.

Bates, P. D., Ohlrogge, J. B. & Pollard, M. 2007. Incorporation of newly synthesized fatty acids into cytosolic glycerolipids in pea leaves occurs via acyl editing. *Journal of Biological Chemistry,* 282**,** 31206-31216.

Beaudoin, F. & Napier, J. A. 2004. 8 Biosynthesis and compartmentation of triacylglycerol in higher plants. *Lipid Metabolism and Membrane Biogenesis.* Springer.

Becker, R. 2007. Fatty acids in food cereal grains and grain products. *Fatty Acids in Foods and their Health Implications***,** 303-316.

Bhardwaj, K., Raju, A. & Rajasekharan, R. 2001. Identification, purification, and characterization of a thermally stable lipase from rice bran. A new member of the (phospho) lipase family. *Plant Physiology,* 127**,** 1728-1738.

Bilyeu, K., Palavalli, L., Sleper, D. & Beuselinck, P. 2003. Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. *Crop Science,* 43**,** 1833-1838.

Borek, S., Ratajczak, W. & Ratajczak, L. 2006. Ultrastructural and enzymatic research on the role of sucrose in mobilization of storage lipids in germinating yellow lupine seeds. *Plant Science,* 170**,** 441-452.

Buhr, T., Sato, S., Ebrahim, F., Xing, A., Zhou, Y., Mathiesen, M., Schweiger, B., Kinney, A. & Staswick, P. 2002. Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. *The Plant Journal,* 30**,** 155-163.

Cahoon, E. B., Shockey, J. M., Dietrich, C. R., Gidda, S. K., Mullen, R. T. & Dyer, J. M. 2007. Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Current opinion in plant biology,* 10**,** 236-244.

Canakci, M., Monyem, A. & Van Gerpen, J. 1999. Accelerated oxidation processes in biodiesel. *TRANSACTIONS-AMERICAN SOCIETY OF AGRICULTURAL ENGINEERS,* 42**,** 1565-1572.

Cases, S., Smith, S. J., Zheng, Y.-W., Myers, H. M., Lear, S. R., Sande, E., Novak, S., Collins, C., Welch, C. B. & Lusis, A. J. 1998. Identification of a gene encoding an acyl CoA: diacylglycerol acyltransferase,

a key enzyme in triacylglycerol synthesis. *Proceedings of the National Academy of Sciences,* 95**,** 13018-13023.

Cases, S., Stone, S. J., Zhou, P., Yen, E., Tow, B., Lardizabal, K. D., Voelker, T. & Farese, R. V. 2001. Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. *Journal of Biological Chemistry,* 276**,** 38870-38876.

Chaiyasit, W., Elias, R. J., Mcclements, D. J. & Decker, E. A. 2007. Role of physical structures in bulk oils on lipid oxidation. *Critical Reviews in Food Science and Nutrition,* 47**,** 299-317.

Champagne, E. T. 2004. *Rice: chemistry and technology*, American Association of Cereal Chemists.

Chen, H. C. & Farese, R. V. 2000. DGAT and triglyceride synthesis: a new target for obesity treatment? *Trends in cardiovascular medicine,* 10**,** 188-192.

Cheruvanky, R., Johnson, I. & Williamson, G. 2003. Phytochemical products: rice bran. *Phytochemical functional foods***,** 347-376.

Choudhury, N. H. & Juliano, B. O. 1980. Effect of amylose content on the lipids of mature rice grain. *Phytochemistry,* 19**,** 1385-1389.

Coleman, R. A. & Lee, D. P. 2004. Enzymes of triacylglycerol synthesis and their regulation. *Progress in lipid research,* 43**,** 134-176.

Drexler, H., Spiekermann, P., Meyer, A., Domergue, F., Zank, T., Sperling, P., Abbadi, A. & Heinz, E. 2003. Metabolic engineering of fatty acids for breeding of new oilseed crops: strategies, problems and first results. *Journal of plant physiology,* 160**,** 779-802.

Dykes, L. & Rooney, L. 2007. Phenolic compounds in cereal grains and their health benefits. *Cereal Foods World,* 52**,** 105-111.

Ejedegba, B., Onyeneke, E. & Oviasogie, P. 2007. Characteristics of lipase isolated from coconut (Cocos nucifera linn) seed under different nutrient treatments. *African Journal of Biotechnology,* 6.

Fallen, B. D., Pantalone, V. R., Sams, C. E., Kopsell, D. A., Vaughn, S. F. & Moser, B. R. 2011. Effect of soybean oil fatty acid composition and selenium application on biodiesel properties. *Journal of the American Oil Chemists' Society,* 88**,** 1019-1028.

Fan, J., Yan, C., Zhang, X. & Xu, C. 2013. Dual role for phospholipid: diacylglycerol acyltransferase: enhancing fatty acid synthesis and diverting fatty acids from membrane lipids to triacylglycerol in Arabidopsis leaves. *The Plant Cell,* 25**,** 3506-3518.

Frei, M. & Becker, K. 2005. A greenhouse experiment on growth and yield effects in integrated rice–fish culture. *Aquaculture,* 244**,** 119-128.

Fujiki, Y., Aizono, Y. & Funatsu, M. 1978. Characterization of minor subunit of rice bran lipase. *Agricultural and Biological Chemistry,* 42**,** 2401-2402.

Fujino, Y. 1978. Rice lipids. *Cereal Chem,* 55**,** 559-571.

Ghosh, M. 2007. Review on recent trends in rice bran oil processing. *Journal of the American Oil Chemists' Society,* 84**,** 315-324.

Gong, Z., Yu, G. P., Dou, C. R. & Xu, M. X. Effect of Heating Treatment of Fresh Rice Bran on Stabilization. Advanced Materials Research, 2013. Trans Tech Publ, 1200-1205.

Gutteridge, J. 1995. Lipid peroxidation and antioxidants as biomarkers of tissue damage. *Clinical chemistry,* 41**,** 1819-1828.

Hoshino, T., Takagi, Y. & Anai, T. 2010. Novel GmFAD2-1b mutant alleles created by reverse genetics induce marked elevation of oleic acid content in soybean seeds in combination with GmFAD2-1a mutant alleles. *Breeding science,* 60**,** 419-425.

Hsieh, K. & Huang, A. H. 2004. Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. *Plant Physiology,* 136**,** 3427-3434.

Hu, G., Huang, S., Cao, S. & Ma, Z. 2009. Effect of enrichment with hemicellulose from rice bran on chemical and functional properties of bread. *Food Chemistry,* 115**,** 839-842.

Huang, A. H. 1992. Oil bodies and oleosins in seeds. *Annual review of plant biology,* 43**,** 177-200.

Huang, A. H., Lin, Y.-H. & Wang, S.-M. 1988. Characteristics and biosynthesis of seed lipases in maize and other plant species. *Journal of the American Oil Chemists' Society,* 65**,** 897-899.

Ichihara, K. I., Takahashi, T. & Fujii, S. 1988. Diacylglycerol acyltransferase in maturing safflower seeds: its influences on the fatty acid composition of triacylglycerol and on the rate of triacylglycerol synthesis. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism,* 958**,** 125-129.

Jaworski, J. & Cahoon, E. B. 2003. Industrial oils from transgenic plants. *Current opinion in plant biology,* 6**,** 178-184.

Ju, Y.-L. & Vali, S. R. 2005. Rice bran oil as a potential resource for biodiesel: a review. *Journal of Scientific and Industrial Research,* 64**,** 866.

Juliano, B. & Bechtel, D. B. 1985. The rice grain and its gross composition.

Kalinski, A., Loer, D. S., Weisemann, J. M., Matthews, B. F. & Herman, E. M. 1991. Isoforms of soybean seed oil body membrane protein 24 kDa oleosin are encoded by closely related cDNAs. *Plant molecular biology,* 17**,** 1095-1098.

Kamal-eldin, A. 2006. Effect of fatty acids and tocopherols on the oxidative stability of vegetable oils. *European Journal of Lipid Science and Technology,* 108**,** 1051-1061.

Kennedy, E. P. Biosynthesis of complex lipids. Federation proceedings, 1961. 934.

Kim, H. U., Hsieh, K., Ratnayake, C. & Huang, A. H. 2002. A novel group of Oleosins is present inside the pollen ofArabidopsis. *Journal of Biological Chemistry,* 277**,** 22677-22684.

Lardizabal, K. D., Mai, J. T., Wagner, N. W., Wyrick, A., Voelker, T. & Hawkins, D. J. 2001. DGAT2 is a new diacylglycerol acyltransferase gene family Purification, cloning, and expression in insect cells of two polypeptides from Mortierella ramanniana with diacylglycerol acyltransferase activity. *Journal of Biological Chemistry,* 276**,** 38862-38869.

Lee, W. S., Tzen, J., Kridl, J. C., Radke, S. E. & Huang, A. 1991. Maize oleosin is correctly targeted to seed oil bodies in Brassica napus transformed with the maize oleosin gene. *Proceedings of the National Academy of Sciences,* 88**,** 6181-6185.

Li, R., Yu, K. & Hildebrand, D. F. 2010. DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. *Lipids,* 45**,** 145-157.

Li, W., Ling, H., Zhang, F., Yao, H., Sun, X. & Tang, K. 2012. Analysis of Arabidopsis genes encoding putative class III lipases. *Journal of plant biochemistry and biotechnology,* 21**,** 261-267.

Liu, Q., Singh, S. & Green, A. 2002. High-oleic and high-stearic cottonseed oils: nutritionally improved cooking oils developed using gene silencing. *Journal of the American College of Nutrition,* 21**,** 205S-211S.

Liu, R. H. 2007. Whole grain phytochemicals and health. *Journal of Cereal Science,* 46**,** 207-219.

Lu, C., Fulda, M., Wallis, J. G. & Browse, J. 2006. A high-throughput screen for genes from castor that boost hydroxy fatty acid accumulation in seed oils of transgenic Arabidopsis. *The Plant Journal,* 45**,** 847-856.

Lu, C., Napier, J. A., Clemente, T. E. & Cahoon, E. B. 2011. New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Current opinion in biotechnology,* 22**,** 252-259.

Lu, C., Xin, Z., Ren, Z. & Miquel, M. 2009. An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of Arabidopsis. *Proceedings of the National Academy of Sciences,* 106**,** 18837-18842.

Lu, C. L., De Noyer, S. B., Hobbs, D. H., Kang, J., Wen, Y., Krachtus, D. & Hills, M. J. 2003. Expression pattern of diacylglycerol acyltransferase-1, an enzyme involved in triacylglycerol biosynthesis, in Arabidopsis thaliana. *Plant molecular biology,* 52**,** 31-41.

Maclean, J. L. & Dawe, D. C. 2002. *Rice almanac: Source book for the most important economic activity on earth*, Int. Rice Res. Inst.

Moazzami, A., Lampi, A.-M. & Kamal-Eldin, A. 2011. Bioactive lipids in cereal and cereal products. *Fruit and Cereal Bioactives: Sources, Chemistry, and Applications***,** 229.

Murphy, D. J. & Ross, J. H. 1998. Biosynthesis, targeting and processing of oleosin-like proteins, which are major pollen coat components in Brassica napus. *The Plant Journal,* 13**,** 1-16.

Murphy, D. J. & Vance, J. 1999. Mechanisms of lipid-body formation. *Trends in biochemical sciences,* 24**,** 109-115.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. & Snyder, M. 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science,* 320**,** 1344-1349.

Napier, J. A., Stobart, A. K. & Shewry, P. R. 1996. The structure and biogenesis of plant oil bodies: the role of the ER membrane and the oleosin class of proteins. *Plant molecular biology,* 31**,** 945-956.

Norton, G. & Harris, J. 1975. Compositional changes in developing rape seed (Brassica napus L.). *Planta,* 123**,** 163-174.

Ohlrogge, J. B. & Jaworski, J. G. 1997. Regulation of fatty acid synthesis. *Annual review of plant biology,* 48**,** 109-136.

Orthoefer, F. T. 1996. Rice bran oil: healthy lipid source. *Food Technology,* 50**,** 62-64.

Ozsolak, F. & Milos, P. M. 2011. RNA sequencing: advances, challenges and opportunities. *Nature reviews genetics,* 12**,** 87-98.

Parthibane, V., Rajakumari, S., Venkateshwari, V., Iyappan, R. & Rajasekharan, R. 2012. Oleosin is bifunctional enzyme that has both monoacylglycerol acyltransferase and phospholipase activities. *Journal of Biological Chemistry,* 287**,** 1946-1954.

Pham, A.-T., Lee, J.-D., Shannon, J. G. & Bilyeu, K. D. 2011. A novel FAD2-1 A allele in a soybean plant introduction offers an alternate means to produce soybean seed oil with 85% oleic acid content. *Theoretical and applied genetics,* 123**,** 793-802.

Prioul, J. L., Méchin, V., Lessard, P., Thévenot, C., Grimmer, M., Chateau-Joubert, S., Coates, S., Hartings, H., Kloiber-Maitz, M. & Murigneux, A. 2008. A joint transcriptomic, proteomic and metabolic analysis of maize endosperm development and starch filling. *Plant biotechnology journal,* 6**,** 855-869.

Quettier, A.-L. & Eastmond, P. J. 2009. Storage oil hydrolysis during early seedling growth. *Plant Physiology and Biochemistry,* 47**,** 485-490.

Routaboul, J.-M., Benning, C., Bechtold, N., Caboche, M. & Lepiniec, L. 1999. The TAG1 locus of Arabidopsis encodes for a diacylglycerol acyltransferase. *Plant Physiology and Biochemistry,* 37**,** 831-840.

Ruiz-Lopez, N., Haslam, R. P., Napier, J. A. & Sayanova, O. 2014. Successful high-level accumulation of fish oil omega-3 long-chain polyunsaturated fatty acids in a transgenic oilseed crop. The Plant Journal, 77, 198-208.

Ruiz-Lopez, N., Haslam, R. P., Usher, S., Napier, J. A. & Sayanova, O. 2015. An alternative pathway for the effective production of the omega-3 long-chain polyunsaturates EPA and ETA in transgenic oilseeds. Plant biotechnology journal, 13, 1264-1275.

Shahidi, f. & wanasundara, U. N. 2002. Methods for measuring oxidative rancidity in fats and oils. *Food lipids: chemistry, nutrition and biotechnology***,** 387-403.

Sayre, R., Saunders, R., Enochian, R., Schultz, W. & Beagle, E. 1982. Review of rice bran stabilization systems with emphasis on extrusion cooking [to prevent oil breakdown also helps to control microorganism growth and insect populations]. *Cereal foods world*.

Shankar, A., Srivastava, A. K., Yadav, A. K., Sharma, M., Pandey, A., Raut, V. V., Das, M. K., Suprasanna, P. & Pandey, G. K. 2014. Whole genome transcriptome analysis of rice seedling reveals alterations in Ca 2+ ion signaling and homeostasis in response to Ca 2+ deficiency. *Cell calcium,* 55**,** 155-165.

Shen, B., Allen, W. B., Zheng, P., Li, C., Glassman, K., Ranch, J., Nubel, D. & Tarczynski, M. C. 2010. Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize. *Plant physiology,* 153**,** 980-987.

Shockey, J. M., Gidda, S. K., Chapital, D. C., Kuan, J.-C., Dhanoa, P. K., Bland, J. M., Rothstein, S. J., Mullen, R. T. & Dyer, J. M. 2006. Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *The Plant Cell,* 18**,** 2294-2313.

Singh, A. K., Fu, D.-Q., El-Habbak, M., Navarre, D., Ghabrial, S. & Kachroo, A. 2011. Silencing genes encoding omega-3 fatty acid desaturase alters seed size and accumulation of Bean pod mottle virus in soybean. *Molecular Plant-Microbe Interactions,* 24**,** 506-515.

Smith, S. J., Cases, S., Jensen, D. R., Chen, H. C., Sande, E., Tow, B., Sanan, D. A., Raber, J., Eckel, R. H. & Farese, R. V. 2000. Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. *Nature genetics,* 25**,** 87-90.

Stone, S. J., Myers, H. M., Watkins, S. M., Brown, B. E., Feingold, K. R., Elias, P. M. & Farese, R. V. 2004. Lipopenia and skin barrier abnormalities in DGAT2-deficient mice. *Journal of Biological chemistry,* 279**,** 11767-11776.

Sugano, M. & Tsuji, E. 1997. Rice bran oil and cholesterol metabolism. *The Journal of nutrition,* 127**,** 521S-524S.

Taira, H., Nakagahra, M. & Nagamine, T. 1988. Fatty acid composition of Indica, Sinica, Javanica, Japonica groups of nonglutinous brown rice. *Journal of agricultural and food chemistry,* 36**,** 45-47.

Takagi, Y. & Rahman, S. 1996. Inheritance of high oleic acid content in the seed oil of soybean mutant M23. *Theoretical and Applied Genetics,* 92**,** 179-182.

Teoh, K. T., Requesens, D. V., Devaiah, S. P., Johnson, D., Huang, X., Howard, J. A. & Hood, E. E. 2013. Transcriptome analysis of embryo maturation in maize. *BMC plant biology,* 13**,** 19.

Turchetto-Zolet, A. C., Maraschin, F. S., De Morais, G. L., Cagliari, A., Andrade, C. M., Margis-Pinheiro, M. & Margis, R. 2011. Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. *BMC evolutionary biology,* 11**,** 263.

Tzen, J. & Huang, A. 1992. Surface structure and properties of plant seed oil bodies. *The Journal of cell biology,* 117**,** 327-335.

Tzen, J. T., Lai, Y.-K., Chan, K.-L. & Huang, A. H. 1990. Oleosin isoforms of high and low molecular weights are present in the oil bodies of diverse seed species. *Plant physiology,* 94**,** 1282-1289.

Vance, V. B. & Huang, A. 1987. The major protein from lipid bodies of maize. Characterization and structure based on cDNA cloning. *Journal of Biological chemistry,* 262**,** 11275-11279.

Vijayakumar, K. & Gowda, L. R. 2013. Rice (Oryza sativa) lipase: Molecular cloning, functional expression and substrate specificity. *Protein expression and purification,* 88**,** 67-79.

Voelker, T. & Kinney, A. J. 2001. Variations in the biosynthesis of seed-storage lipids. *Annual review of plant biology,* 52**,** 335-361.

Whitelaw, E., Rahman, S., Li, Z., Liu, Q., Singh, S. P. & De Feyter, R. C. 2013. Altering the fatty acid composition of rice. Google Patents.

Wickramasuriya, A. M. & Dunwell, J. M. 2015. Global scale transcriptome analysis of Arabidopsis embryogenesis in vitro. *BMC genomics,* 16**,** 301.

Williams, J., Imperial, V., Khan, M. & Hodson, J. 2000. The role of phosphatidylcholine in fatty acid exchange and desaturation in Brassica napus L. leaves. *Biochem. J,* 349**,** 127-133.

Wilson, T. A., Ausman, L. M., Lawton, C. W., Hegsted, D. M. & Nicolosi, R. J. 2000. Comparative cholesterol lowering properties of vegetable oils: beyond fatty acids. *Journal of the american college of nutrition,* 19**,** 601-607.

Wu, Y.-Y., Chou, Y.-R., Wang, C.-S., Tseng, T.-H., Chen, L.-J. & Tzen, J. T. 2010. Different effects on triacylglycerol packaging to oil bodies in transgenic rice seeds by specifically eliminating one of their two oleosin isoforms. *Plant Physiology and Biochemistry,* 48**,** 81-89.

Xu, J., Carlsson, A. S., Francis, T., Zhang, M., Hoffman, T., Giblin, M. E. & Taylor, D. C. 2012. Triacylglycerol synthesis by PDAT1 in the absence of DGAT1 activity is dependent on re-acylation of LPC by LPCAT2. *BMC plant biology,* 12**,** 4.

Yoshida, H., Tanigawa, T., Kuriyama, I., Yoshida, N., Tomiyama, Y. & Mizushina, Y. 2011. Variation in fatty acid distribution of different acyl lipids in rice (Oryza sativa L.) brans. *Nutrients,* 3**,** 505-514.

Zhang, J., Martin, J. M., Beecher, B., Lu, C., Hannah, L. C., Wall, M. L., Altosaar, I. & Giroux, M. J. 2010. The ectopic expression of the wheat Puroindoline genes increase germ size and seed oil content in transgenic corn. *Plant molecular biology,* 74**,** 353-365.

Zhang, M., Fan, J., Taylor, D. C. & Ohlrogge, J. B. 2009. DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. *The Plant Cell,* 21**,** 3885-3901.

Zheng, P., Allen, W. B., Roesler, K., Williams, M. E., Zhang, S., Li, J., Glassman, K., Ranch, J., Nubel, D. & Solawetz, W. 2008. A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nature genetics,* 40**,** 367-372.

Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G. & Taylor, D. C. 1999. The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. *The Plant Journal,* 19**,** 645-653.

Malays. Appl. Biol. (2014) 43(2): 1–8

IMPROVING THE ECONOMIC VALUE OF RICE BRAN OIL

TIWARI, G.J., AUMEERUDDY, A. and RAHMAN, S.*

School of Science, Monash University, Jalan Lagoon Selatan, Bandar Sunway, 46150, Selangor, Malaysia

ABSTRACT

Cereal grains are usually considered to be sources of carbohydrate but they can also contain oil which is of considerable economic value. Rice oil is obtained from the bran produced by polishing brown rice. Key genes which control rice bran oil composition have been identified but oil quality is also affected by the rapid breakdown of the triglycerides following polishing. Several lipases in the bran have been identified by proteomic approaches. However, DNA sequence identity searches indicate that over a hundred putative lipase genes can be identified in the rice genome although it is not clear how many of these putative genes are expressed in the bran. A concerted research effort therefore is required to identify the lipases that are responsible for oil breakdown in the bran.

Key words: lipase, embryo, aleurone, cereal, grasses

INTRODUCTION

Cereals provide most of the world's calories through the starch that is present in the grain. The grain also contains lipids and these are of increasing importance, as ways to maximize the value of crops, are investigated. Furthermore grasses, which include cereals, are being considered as a source of biodiesel.

Oils from plants are primarily composed of triacylglycerols (TAGs), which have three fatty acid chains attached to the glycerol backbone through ester links. TAGs are stored in oil bodies which are surrounded by a monolayer comprising phospholipids and a diverse array of proteins, predominant among which are proteins termed oleosins. However, the main constituents of lipid bodies are TAGs (90 to 95%) (Ohlrogge and Browse, 1995).

Plant oils are an important component of our diet, as they serve 20-25% of the daily nutritional calorie intake in humans (Katan *et al.,* 1995). Plant oils are also used in chemical industries (e.g. in detergents, paints, inks, and plasticizers), food industries (e.g. in margarines, salad oils, fried foods) and also bio-based industrial formulations, like lubricants and drying oil (Lu *et al.,* 2011). Biodiesel production from plant oils is one of the main non-food applications (Emiliani and Pistocchi, 2006; Ramadhas *et al.,* 2005). Plant oil production has increased by nearly 50% overall over last decade to meet the increasing demand (Table 1). Naturally there is great demand for improvement and enhancement of plant oils.

* To whom correspondence should be addressed.

Plant oil types

Plant oils are grouped into two major classes namely, essential and fixed oils. Most of the fixed oils are derived from fruit or seeds of plants (e.g. soybean, rapeseed, cottonseed, sunflower seed, groundnut, palm, copra, sesame, linseed & castor seed, maize and coconut oils) (http://www.fas.usda.gov/cots/oilseed). Apart from seeds and fruits, plant oils are also extracted from the flowers, leaves, stems, bark and roots of herbs, bushes, shrubs and trees through distillation. The remainder of the review will concentrate on oils from seeds.

Among major oilseed crops, soybean is biggest source of edible oils followed by canola, and sunflower (Wilcox, 2004). However, palm oil is largest source overall. Maize is the most widely used cereal for the production of oils. Rice oil production is much lower (Table 1). The production of oil from the other cereals is very minor.

Accumulation of seed oils

In the seeds, lipids accumulate in the embryo or endosperm, depending on plant species (Baud and Lepiniec, 2010).

For eudicots, which include important oilseeds like soybean (Glycine max), sunflower (Helianthus annuus), linseed (Linum usitatissimum), safflower (Carthamus tinctorius) and the Brassicaceae,

Table 1. Production of the major plant oils

Oil	2002	2007	2012
Palm oil	26.1	39.7	50.1
Soybean	29	37.4	40
Rapeseed	13.3	17.9	23.5
Sunflower	7.8	11.1	15
Palm kernel	3.3	4.9	5
Maize oil	1.9	2.2	2.3
Estimated rice bran oil production	.7	.9	1.05

Quantities in millions of tonnes.
Source: FAOSTAT (http://faostat3.fao.org/faostat-gateway/go/to/browse/Q/*/E). Rice bran oil production is estimated based on rice production and proportion converted to oil from the bran currently.

embryonic tissues present between the integuments of the seed are major site of oil deposition. However, in castor bean, another important eudicot oilseed, the large endosperm tissue is the main oil storage tissue.

In cereals, which are monocots, both the embryo and the endosperm can be major sources of seed oil (Table 1).

Edible oils from monocot grasses

In cereals, which are monocot grasses like maize (*Zea mays*) and rice (*O. sativa*) the main product of the grain is starch. However, maize and oats are also good sources of plant oil as over 10% of the grain weight is oil (Morrison, 1977; Leonoval *et al.,* 2010). In other cereals such as rice, wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*) oils constitutes only 2-3% of total dry weight of grain (Fincher, 1989) but can be obtained in much higher proportions (approximately 20%) in the bran fraction that results from processing of the grain. In most of the cereals, oil is mainly stored in embryo and the aleurone which are minor parts of the seed but in oats (*Avena sativa*) the majority of lipids are stored in the endosperm (Leonova *et al.,* 2010). In terms of oil production from cereals, maize and rice are the only cereals so used to any significant degree. Although oats is rich in oil, the production of oats is very low compared to these two cereals, amounting to only about 23 million tonnes in 2013 compared to about 720 million tonnes for rice and over 1 billion tonnes for maize (http://faostat3.fao.org/faostat-gateway/go/to/browse/Q/QC/E). Oil from rice and maize are discussed in greater detail below.

RICE

Rice grains contain a much smaller proportion of lipids than starch (Tokuşoğlu and Hall III, 2011). Milled rice from which bran layer is removed has about 0.2% of lipid content when compared to brown rice having about 1-4% of lipid content (Juliano and Bechtel, 1985). Lipids in rice grain are stored as spherosomes of <1.5μm in the aleurone layer, <1μm in the sub-aleurone layer and <0.7μm in the embryo. The distribution of lipids is quite uneven as the outer layers of rice kernels (that is in the bran including the germ) have larger amounts of lipid than the inner parts (that is in the core or inner endosperm) (Bechtel and Pomeranz, 1977; Bechtel and Pomeranz, 1978; Juliano, 1983). Moreover three dimensional distribution analyses of lipids in rice grain showed more lipid on the dorsal than the ventral side of rice grain (Ogawa *et al.,* 2002).

Rice bran oil

Rice bran oil (RBO) is in increasing demand a popular cooking oil in several Asian countries (Sugano and Tsuji, 1997; Ghosh, 2007). The proportion and composition of rice bran oil can vary somewhat depending on the type of rice that is milled. Lipid content is higher in purple/black rice (12-13%) as compared to brown rice where it ranges from about 3-4% on dry weight basis (Frei and Becker, 2005). There appears to have been no effort to breed rice lines with increased oil content because of the focus on increasing grain yield through increasing starch content.

RBO is mostly triglyceride but also contains compounds such as oryzanol and tocotrienes having antioxidant and hypocholesterolemic properties (Carroll, 1990; Orthoefer, 1996; McCaskill and Zhang, 1999) and it has the potential to reduce both total serum and low density lipoprotein cholesterol levels in those who consume it (Wilson *et al.,* 2000). The triglyceride component of rice bran oil consists largely of esterified palmitic, oleic and linoleic acids. Table 2 shows the composition of rice bran oil during development.

The content of phytochemicals like tocopherol, tocotrienol, and γ-oryzanol in rice bran has been thoroughly studied (Dykes and Rooney, 2007; Zhang *et al.,* 2010; Liu, 2007). Like any other phytochemical, the phenolic and flavonoid content also depends on quality traits (grain color, size and weight) of rice grain. Table 3 indicates that black rice has the highest content of phenolics and flavonoids and the greatest maximum antioxidant property followed by red rice and white rice (Zhang *et al.,* 2010).

The biosynthesis of the oil in the grain during development has not been studied recently. Early studies indicated that oleic and linoleic acids increase up to 16 days after flowering whereas palmitic acid remains constant (Choudhury and Juliano, 1980).

Table 2. Lipid accumulation in rice grain

Days after flowering	Total lipids (µg/grain)	Fatty acid accumulation of non-starch lipids (% of 20-day grain)	Fatty acid composition of non-starch lipids (wt% of total)			
			Palmitic (16:0)	Oleic (18:1)	Linoleic (18.2)	α-Linoleic (18:3)
4	74	10	19	14	39	26
8	272	50	20	22	44	12
12	460					
16	470					
20	472	100	21	28	43	4
24	464					
28	468					

(Adapted from Choudhry and Juliano, 1980)

Table 3. Comparison of phenolics, flavonoids contents and antioxidant capacity among white, red and black rice genotypes

	Phenolics[a]	Flavonoids[a]	Antioxidant capacity[a]
Total rice			
Mean±SD	197.5±144.8	134.7±19.8	0.413±0.696
CV %	77.3	14.7	168.63
Range	108.1-1244.9	88.6-286.3	0.012-5.533
White rice			
Mean±SD	151.8± 19.5	131.6±14.2	0.196±0.073
CV %	12.9	10.8	37.33
Range	108.1-251.4	88.6-170.7	0.012-0.413
Red rice			
Mean±SD	470.1±107.2	147.2±18.0	1.705±0.600
CV %	22.8	12.3	35.22
Range	165.8-731.8	108.7-190.3	0.291-2.963
Black rice			
Mean±SD	1055.7±176.2	240.6±38.1	4.484±1.095
CV %	16.7	15.8	24.41
Range	841.0-1244.9	187.6-286.3	2.527-5.533

[a]Phenolics content was expressed as mg GAE/100 g, flavonoids content was expressed as mg RE/100 g, and antioxidant capacity was expressed as mMTAEC. (Adapted from Shen *et al.*, 2009).

There are few recent and systematic studies to screen for differences in rice bran oil composition. Earlier studies showed an oleic acid content around 40% (Taira *et al.,* 1988). However, as the critical genes involved in determining the composition are known in other plants (e.g. in cotton seed, Liu *et al.,* 2002) transgenic approaches have been undertaken in rice to increase the proportion of oleic acid at the expense of both palmitic acid (a saturated fatty acid) and linoleic acid (two double bonds that can go to the trans arrangement upon heating) as these latter two can have negative health implications. By knocking out the FAD2 gene using an RNAi approach, rice grains with almost the double the proportion of oleic acid were produced recently (Zaplin *et al.,* 2013). However, whether this translates into double the proportion of oleic acid in RBO has not yet been tested.

Little work seems have been done in trying to increase the oryzanol content, largely because the genes involved are not well delineated. Clearly a survey of the variation in oryzanol and tocopherol content among different rices would be of use in this regard.

LIMITATIONS ON THE USE OF RICE OIL

The production of RBO involves two broad steps. The first step is the production of the bran and stabilization of the oil within the bran. The second step is the extraction of the oil from within the bran.

To improve the value of RBO it is important to focus on the action of lipases and other enzymes that lead to degradation of the oil bodies. The oil bodies consist of a triglyceride core contained within phospholipid layer that is it itself interrupted and protected by a selection of proteins, predominant among which are the oleosins (Frandsen *et al.,* 2001). The oleosins and the related proteins caleosins are 15-30 kDa in mass. The breakdown of the triglyceride at the core requires disruption of these protective layers. The process has not been clearly delineated although it has been suggested that the oleosins contain binding motifs for lipases as the plant requires regulated breakdown during germination. However, if breakdown is initiated after bran production, either by premature triggering of the germinative breakdown cascade or by adventitious lipases the release of fatty acids leads to poor oil quality and lowers the value of the product. The free fatty acids produced that can be further acted on by lipoxygenases to produce rancid flavor.

A large number of different types of lipases have been characterized and some of these have TAG lipase activity (Matos and Pham-Thi, 2009). It is not known what proportion of them would be present in the bran. The complementary approach, isolating lipases from the bran has been attempted by a number of researchers (Bhardwaj *et al.,* 2001, Funatsu *et al.,* 1971; Aizono *et al.,* 1976; Fujiki *et al.,* 1978), but the genes corresponding to these activities have only been identified in a few cases (Kim 2004; Vijaykumar and Gowda 2013). Vijaykumar and Gowda (2013), purified a lipase activity that was identical to that reported by Aizino *et al.* (1976) and could relate it to a cDNA for lipase available at NCBI. They expressed the cDNA as protein and demonstrated lipase activity. The cDNA sequence contained the canonical GxSxG motif of lipases. Furthermore, they followed the accumulation of transcripts for this sequence by real-time PCR. However, the initial purification was based on hydrolysis of tributyrin which may limit the potential lipases assayed to a subset of those available (Vijaykumar and Gowda 2012).

The genome sequence of rice has been available for some time and more than a hundred lipases have been annotated in the rice genome (http://mpss.udel.edu/rice/mpss_index.php?). Although some of these may turn out not to be functional or of limited importance in terms of rice bran oil preparation, this approach may provide insight into the number of the different lipase and protease activities present in the rice bran. An example of a putative rice lipase gene recently isolated is *This1* (Liu *et al.,* 2013) and it is ubiquitously expressed, including in the panicle. However, its presence in the bran has not yet been demonstrated.

Once the relative importance of the various lipases in the bran have been evaluated, steps can be taken to eliminate their expression in the bran. This could take the form of establishing markers for the targeted lipases and then breeding to exclude such markers. Alternatively, if these lipases have critical roles otherwise, an RNAi or micro RNA approach using promoters that drive expression during late grain development could perhaps be used. The seed-specific promoter used by Zaplin *et al.* (2013) to reduce the proportion of linoleic acid in the rice bran is an example of the type of promoter that could be used. No lines lacking specific lipases appear to have been reported yet.

Rice bran stabilization and extraction

Physical and chemical methods used for inactivating rice bran lipase activity for stabilizing rice bran include dry heating, wet heating, and extrusion (Sayre *et al.,* 1982). Refrigeration and addition of chemicals additions such as sodium metabisulfite have also been used to decrease lipase activity and promote stabilization of bran (Tao, 2001; Cheruvanky *et al.,* 2003). Rice bran enzymes have also been deactivated by altering pH, which helps to increase the shelf life of rice bran for three months (Escamilla Castillo *et al.,* 2005).

However, these methods which have had relatively little success to date (Raghavendra *et al.,* 2007; Tao, 2001) are not promising long-term solutions to increase the shelf -life of rice bran. Rice lines lacking some of the lipoxygenases have demonstrated significantly better storability (Zhang *et al.,* 2007; Suzuki *et al.,* 1999) but as indicated earlier lines lacking lipases do not appear to have been produced.

RBO is always a by-product of the production of polished rice grains. Oil yield from rice bran extraction can be increased if rice bran is enzymatically treated with cellulase and pectinase prior to oil extraction by hexane (Sengupta and Bhattacharyya, 1996). An overview of rice bran oil extraction is shown in Fig. 1. RBO yield is highest when extracted with hexane (20.21%) followed by CO2–ethanol (18.23%) and supercritical CO2 (17.98%) respectively (Orthoefer, 2005). Hexane is commercially used for oil extraction from oilseeds although it is considered to be an air pollutant (Rosenthal *et al.,* 1996).

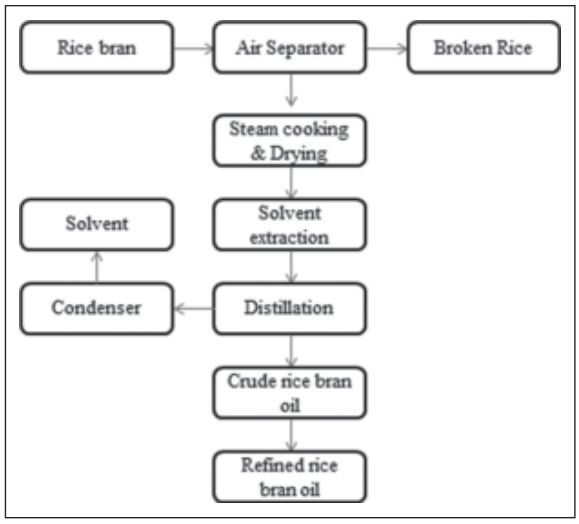


Fig. 1. Schematic diagram of rice bran oil extraction

OTHER CEREALS

Maize

Maize is often grown expressly for the oil and in such cases the corn is collected and processed by removing the germ (which contains about 85% of the oil) and oil content depends on its concentration and the area occupied by the embryo in the seed. However, corn oil can also be extracted from ground corn kernels (Hojilla-Evangelista *et al.,* 1992; Kwiatkowski and Cheryan, 2002) and corn fiber (Moreau *et al.,* 1996).

Maize oil is rich in polyunsaturated fatty acid (PUFA) content (65 to 85%) and thus fulfills the requirement of essential fatty acids in human nutrition (Goffman and Böhme, 2001). In maize oil, linoleic acid (18:2) alone comprises about 60% and monounsaturated fatty acid (MUFA) (oleic acid; 18:1) is about 24% of the total percentage of PUFA in maize oil. Among saturated fatty acids (SFA), palmitic acid (16:0) is almost 13% and stearic acid (18:0) is 1%. Thus maize thus has a high percentage of linoleic acid which is comparable to that of rice screw pressing and solvent extraction are two major methods of oil extraction; however these methods cannot be used to extract lipids present in the endosperm (MacRitchie and Gras, 1973). Commercially, corn oil is extracted either by hexane (Reiners, 1982; Stolp and Stute, 1982) or ethanol (Chien *et al.,* 1988; Chien *et al.,* 1990).

An overview of maize oil extraction is shown in Fig. 2.

Oats, wheat and barley

Oats apart from being a rich source of dietary soluble fiber beta-glucan (Glore *et al.,* 1994) also has a higher lipid content than other cereals as indicated previously, (see also Liu, 2011). The major fatty acids in oats oil are linoleic (18:2), oleic (18:1) and palmitic (16:0) (Welch, 1995). It also contains Vitamin E and antioxidant compounds which give oat oil cholesterol-reducing properties. (Youngs and Webster, 1986). The oleic content in

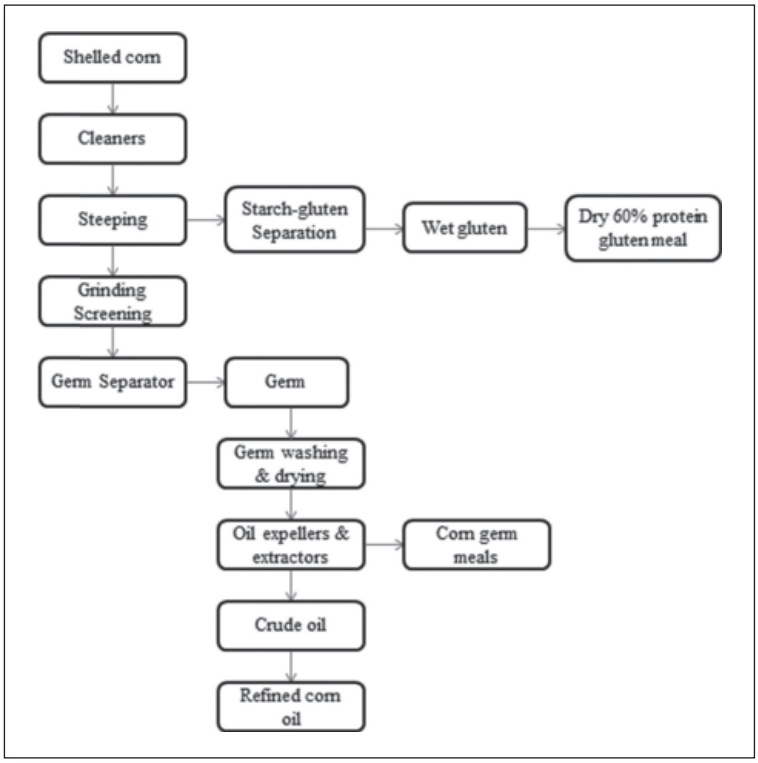


Fig. 2. Schematic diagram of corn oil extraction

oat oil is higher compared to commonly used soybean or sunflower oil but less than in canola and olive oils. The processing of oat oil is not economical and therefore, oat oil is not widely consumed or considered as edible oil.

The wheat embryo – also known as the germ contains approximately 11% of oil (Sonntag, 1979) comprising a large proportion of polyunsaturated fatty acids and vitamin E. It is one of the richest natural sources of γ-tocopherol, a compound known to have high vitamin E activity (Kahlon, 1989). Most of the fatty acids (57%) are present are as triglycerides (Kahlon, 1989). The most abundant is linoleic acid (18:2) (42–59% of the total triglycerides), followed by palmitic acid (16:0) and oleic acid (16:1) (Kahlon, 1989; Hidalgo and Brandolini, 2008). The characteristics of barley oil are similar to that of wheat (Liu, 2011). As with oats oil production from wheat bran does not yet appear to be economically attractive.

CONCLUSIONS

All the important cereals contain appreciable amounts of lipids in the grain. Maize oil has been most widely exploited but rice bran oil is a promising second. Much work needs to be done on investigating differences in the oil composition among different types of rice and on reducing the activity of the lipases. Such reductions in lipase activity should help in increasing the economic value of the oil. Clearly a combination of genomic and proteomic approaches are needed to investigate the lipase genes in rice grain, in order to understand their roles in lipid metabolism. This will help in devising approaches to improve the quality as well as quantity of rice bran oil. In addition, work needs to be initiated to increase the overall quantity of rice bran oil in the rice grain without affecting grain production.

ACKNOWLEDGMENTS

We are grateful to Monash University Malaysia for supporting this research through the provision of financial assistance and facilities.

REFERENCES

Aizono, Y., Funatsu, M., Fujiki, Y. & Watanabe, M. 1976. Purification and characterization of rice bran lipase II. *Agricultural and Biological Chemistry*, **40:** 317-324.

Baud, S. & Lepiniec, L. 2010. Physiological and developmental regulation of seed oil production. *Progress In Lipid Research*, **49:** 235-249.

Bechtel, D.B. & Pomeranz, Y. 1977. Ultrastructure of the mature ungerminated rice (*Oryza sativa*) caryopsis. The caryopsis coat and the aleurone cells. *American Journal of Botany*, 966-973.

Bechtel, D.B. & Pomeranz, Y. 1978. Ultrastructure of the mature ungerminated rice (*Oryza sativa*) caryopsis. The starchy endosperm. *American Journal of Botany*, **65(6):** 684-691.

Bhardwaj, K., Raju, A. & Rajasekharan, R. 2001. Identification, purification, and characterization of a thermally stable lipase from rice bran. A new member of the (phospho) lipase family. *Plant Physiology*, **127:** 1728-1738.

Carroll, L. 1990. Functional properties and applications of stabilized rice bran in bakery products. *Food Technology*, **44:** 74-76.

Cheruvanky, R., Johnson, I. & Williamson, G. 2003. Phytochemical products: rice bran. *Phyto-Chemical Functional Foods*, 347-376.

Chien, J., Hoff, J. & Chen, L. 1988. Simultaneous dehydration of 95% ethanol and extraction of crude oil from dried ground corn. *Cereal Chem,* **65:** 484-486.

Chien, J., Hoff, J., Lee, M., Lin, H., Chen, Y. & Chen, L. 1990. Oil extraction of dried ground corn with ethanol. *The Chemical Engineering Journal*, **43:** B103-B113.

Choudhury, N.H. & Juliano, B.O. 1980. Effect of amylose content on the lipids of mature rice grain. *Phytochemistry*, **19:** 1385-1389.

Dykes, L. & Rooney, L. 2007. Phenolic compounds in cereal grains and their health benefits. *Cereal Foods World*, **52:** 105-111.

Emiliani, D. & Pistocchi, M. 2006. Biodiesel production from vegetable oils. *Industria Saccarifera Italiana*, **99:** 91-98.

EscamillaCastillo, B., VarelaMontellano, R., SánchezTovar, S.A., SolísFuentes, J.A. & DuránDeBazúa, C. 2005. Extrusion deactivation of rice bran enzymes by pH modification. *European Journal of Lipid Science and Technology*, **107:** 871-876.

Fincher, G.B. 1989. Molecular and cellular biology associated with endosperm mobilization in germinating cereal grains. *Annual Review of Plant Biology*, **40:** 305-346.

Frandsen, G, Mundy, J. & Tzen, J.T.C. 2001. Oil bodies and their associated proteins, oleosin and caleosin. *Physiologia Plantarum* **112:** 301-307.

Frei, M. & Becker, K. 2005. A greenhouse experiment on growth and yield effects in integrated rice-fish culture. *Aquaculture*, **244:** 119-128.

Fujiki, Y., Aizono, Y. & Funatsu, M. 1978. Characterization of minor subunit of rice bran lipase. *Agricultural and Biological Chemistry*, **42(12):** 2401-2402.

Funatsu, M., Aizono, Y., Hayashi, K., Watanabe, M. & Eto, M. 1971. Biochemical studies on rice bran lipase. I. Purification and physical properties. *Agricultural and Biological Chemistry*, **35(5):** 734-742.

Ghosh, M. 2007. Review on recent trends in rice bran oil processing. *Journal of the American Oil Chemists' Society*, **84:** 315-324.

Glore, S.R., Van Treeck, D., Knehans, A.W. & Guild, M. 1994. Soluble fiber and serum lipids: a literature review. *Journal of the American Dietetic Association*, **94:** 425-436.

Goffman, F.D. & Böhme, T. 2001. Relationship between fatty acid profile and vitamin E content in maize hybrids (Zea mays L.). *Journal of Agricultural and Food Chemistry*, **49:** 4990-4994.

Hidalgo, A. & Brandolini, A. 2008. Kinetics of carotenoids degradation during the storage of einkorn (Triticum monococcum L. ssp. monococcum) and bread wheat (Triticum aestivum L. ssp. aestivum) flours. *Journal of Agricultural and Food Chemistry*, **56:** 11300-11305.

Hojilla-Evangelista, M., Johnson, L. & Myers, D. 1992. Sequential extraction processing of flaked whole corn: Alternative corn fractionation technology for ethanol production. *Cereal Chemistry*, **69:** 643-643.

Juliano, B.C. 1983. Lipids in rice and rice processing. In *Lipids in Cereal Technology;* Barenes, P.J; Ed., Academic Press; New York, USA, 305-330.

Juliano, B.C. & Bechtel, D.B. 1985. The rice grain and its gross composition. In *Rice Chemistry and Technology*; 2nd Ed. Juliano, B.O; Ed. American Association of Cereal Chemists; St. Paul, Minnesota, 17-57pp.

Kahlon, T. 1989. Nutritional implications and uses of wheat and out kernel oil. *Cereal Foods World*, **34:** 872-875.

Katan, M.B., Zock, P.L. & Mensink, R.P. 1995. Trans fatty acids and their effects on lipoproteins in humans. *Annual Review of Nutrition*, **15:** 473-493.

Kim, Y. 2004. Cloning and expression of a lipase gene from rice (Oryza sativa cv. Dongjin). *Molecules and Cells*, **18:** 40-45.

Kwiatkowski, J.R. & Cheryan, M. 2002. Extraction of oil from ground corn using ethanol. *Journal of the American Oil Chemists' Society*, **79:** 825-830.

Leonova, S., Grimberg, Å., Marttila, S., Stymne, S. & Carlsson, A.S. 2010. Mobilization of lipid reserves during germination of oat (*Avena sativa* L.), a cereal rich in endosperm oil. *Journal of Experimental Botany*, **61:** 3089-3099.

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

Liu, KS. 2011. Comparison of lipid content and fatty acid composition and their distribution within seeds of 5 small grain species. *Journal of Food Science*, **76:** C334–C342.

Liu, W., Zhang, D., Tang, M., Li, D., Zhu, Y., Zhu, L. & Chen, C. 2013. THIS1 is a putative lipase that regulates tillering, plant height and spikelet fertility in rice. *Journal of Experimental Botany*, Vol. 64, No. **14:** 4389–4402.

Liu, L., Zhang, Y., Chen, N., Shi, X., Tsang, B. & Yu, Y.-H. 2007. Upregulation of myocellular DGAT1 augments triglyceride synthesis in skeletal muscle and protects against fat-induced insulin resistance. *Journal of Clinical Investigation*, **117:** 1679-1689.

Lu, C., Napier, J.A., Clemente, T.E. & Cahoon, E.B. 2011. New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Current Opinion in Biotechnology*, **22:** 252-259.

Macritchie, F. & Gras, P. 1973. The role of flour lipids in baking. *Cereal Chemistry*, **50:** 292-302.

Matos, A.R., Pham-Thi, A.T. 2009. Lipid deacylating enzymes in plants: Old activities, new genes. *Plant Physiology and Biochemistry* **47:** 491-503.

Mccaskill, D.R. & Zhang, F. 1999. Use of rice bran oil in foods: Developing nutraceuticals for the new millenium. *Food Technology*, **53:** 50-53.

Moreau, R.A., Powell, M.J. & Hicks, K.B. 1996. Extraction and quantitative analysis of oil from commercial corn fiber. *Journal of Agricultural and Food Chemistry*, **44:** 2149-2154.

Morrison, W. 1978. Cereal lipids. *Advances in Cereal Science and Technology*, **2:** 221-348.

Ohlrogge, J. and Browse, J. 1995. Lipid Biosynthesis. *The Plant Cell*, **7:** 957-970.

Ogawa, Y., Kuensting, H., Sugiyama, J., Ohtani, T., Liu, X., Kokubo, M., Kudoh, K. & Higuchi, T. 2002. Structure of a rice grain represented by a new three-dimensional visualisation technique. *Journal of Cereal Science*, **36:** 1-7.

Orthoefer, F.T. 1996. Rice bran oil: healthy lipid source. Food Technology.

Orthoefer, F.T. 2005. Rice bran oil. Bailey's industrial oil and fat products. **2:** 10.

Raghavendra, M., Kumar, P.R. & Prakash, V. 2007. Mechanism of inhibition of rice bran lipase by polyphenols: a case study with chlorogenic acid and caffeic acid. *Journal of Food Science*, **72:** E412-E419.

Ramadhas, A.S., Jayaraj, S. & Muraleedharan, C. 2005. Biodiesel production from high FFA rubber seed oil. *Fuel*, **84:** 335-340.

Reiners, R.A. 1982. Extraction of oil from vegetable materials. Google Patents.

Rosenthal, A., Pyle, D. & Niranjan, K. 1996. Aqueous and enzymatic processes for edible oil extraction. *Enzyme and Microbial Technology*, **19:** 402-420.

Sayre, R., Saunders, R., Enochian, R., Schultz, W. & Beagle, E. 1982. Review of rice bran stabilization systems with emphasis on extrusion cooking [to prevent oil breakdown also helps to control microorganism growth and insect populations]. *Cereal Foods World*, 27.

Sengupta, R. & Bhattacharyya, D. 1996. Enzymatic extraction of mustard seed and rice bran. *Journal of the American Oil Chemists' Society*, **73:** 687-692.

Shen, Y., Jin, L., Xiao, P., Lu, Y. & Bao, J. 2009. Total phenolics, flavonoids, antioxidant capacity in rice grain and their relations to grain color, size and weight. *Journal of Cereal Science*, **49:** 106-111.

Sonntag, N.O. 1979. Composition and characteristics of individual fats and oils. *Bailey's Industrial Oil and Fat Products*, **1:** 289-477.

Stolp, K.-D. & Stute, R.W. 1982. Process for obtaining corn oil from corn germ. Google Patents.

Sugano, M. & Tsuji, E. 1997. Rice bran oil and cholesterol metabolism. *The Journal of Nutrition*, **127:** 521S-524S.

Suzuki, Y., Ise, K., Li, C., Honda, I., Iwai, Y. & Matsukura, U. 1999. Volatile components in stored rice [*Oryza sativa* (L.)] of varieties with and without lipoxygenase-3 in seeds. *Journal of Agricultural and Food Chemistry*, **47:** 1119-1124.

Taira, H., Nakagahra, M. & Nagamine, T. 1988. Fatty acid composition of Indica, Sinica, Javanica, Japonica groups of nonglutinous brown rice. *Journal of Agricultural and Food Chemistry*, **36:** 45-47.

Tao, J. 2001. Method of stabilization of rice bran by acid treatment and composition of the same. Google Patents.

Tokuşoğlu, Ö. & Hall Iii, C.A. 2011. Fruit and Cereal Bioactives: Sources, Chemistry, and Applications, CRC Press.

Vijayakumar, K. & Gowda, L.R. 2012. Temporal expression profiling of lipase during germination and rice caryopsis development. *Plant Physiology and Biochemistry* **57** (2012) 245-253.

Vijayakumar, K. & Gowda, L.R. 2013. Rice (*Oryza sativa* lipase: Molecular cloning, functional expression and substrate specificity. *Protein Expression and Purification*, **88:** 67-79.

Weiss, E.A. 1997. Essential oil crops, Cab International.

Welch, R.W. 1995. The chemical composition of oats. The Oat Crop. Springer.

Wilcox, J.R. 2004. World distribution and trade of soybean. Soybeans: Improvement, Production, and Uses, 1-14.

Wilson, T.A., Ausman, L.M., Lawton, C.W., Hegsted, D.M. & Nicolosi, R.J. 2000. Comparative cholesterol lowering properties of vegetable oils: beyond fatty acids. *Journal of the American College of Nutrition*, **19:** 601-607.

Youngs, V.L. & Webster, F. 1986. Oat lipids and lipid-related enzymes. *Oats: Chemistry and Technology*, 205-226.

Zaplin Ella Simone, Liu Qing, Li Zhongyi, Butardo Vito, M., Blanchard Christopher, L. & Rahman Sadequr. 2013. Production of high oleic rice grains by suppressing the expression of the OsFAD2-1 gene. *Functional Plant Biology* **40:** 996-1004.

Zhang, Y., Yu, Z., Lu, Y., Wang, Y., She, D., Song, M. & Wu, Y. 2007. Effect of the absence of lipoxygenase isoenzymes on the storage characteristics of rice grains. *Journal of Stored Products Research*, **43:** 87-91.

Zhang, M.W., Zhang, R.F., Zhang, F.X. & Liu, R.H. 2010. Phenolic profiles and antioxidant activity of black rice bran of different commercially available varieties. *Journal of Agricultural and Food Chemistry*, **58:** 7580-7587.

Chapter 2

Genomic approach for the identification, cloning and expression studies of lipase genes in rice

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter 2

Declaration by candidate

In the case of **Chapter 2**, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Experimental design and conduct, samples collection and process, data collection, result acquisition, statistical analysis, manuscript preparation	80

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Chiang May Ying	Editing of manuscript	N/A (not a student registered under Monash University)
Jeremy Ryan De Silva	Editing of manuscript	N/A (not a student registered under Monash University)
Song Beng Kah	Experimental design and manuscript preparation	N/A (not a student registered under Monash University)
Lau Yee Ling	Experimental design	N/A (not a student registered under Monash University)
Sadequr Rahman	Experimental design, result discussion, manuscript preparation and submission	N/A (not a student registered under Monash University)

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

Candidate's Signature	[redacted]	**Date** 2·3·2016
Main Supervisor's Signature	[redacted]	**Date** 2.3.2016

*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

CHAPTER SUMMARY

Rice (*O.sativa*) is the first cereal crop whose entire genome is fully sequenced and available in online databases. This chapter aims to use the available genomic information and study lipase genes across the rice genome. Based on the online search in rice Massively Parallel Signature Sequencing (MPSS) database (http://mpss.udel.edu) a total number of 125 putative lipase genes were identified across the rice genome. Based on Fragments Per Kilobase of transcript per Million (FPKM) expression values obtained from the Michigan State University (MSU) Rice Genome Annotation Project Database (http://rice.plantbiology.msu.edu/) 69 lipase genes expressing in seed and leaf tissues of rice were selected for further bioinformatics study.

A bioinformatics study including homolog search in *A.thaliana*, lipase motif search and subcellular localization prediction analysis was performed on all the selected 69 putative lipase genes. Putative lipase genes carrying the lipase motif (GXSXG) and with ≥ 50% protein sequence identity with *A.thaliana* homologs were selected and phylogenetic analysis was performed with known triacylglycerol lipase (TAGL) genes in *A.thaliana*.

Based on bioinformatics analysis, nine *O.sativa* genes showing clear sequence identity by alignment with *A.thaliana* TAGL genes were selected for endpoint PCR expression analysis in the rice cultivar MR219, a popular Malaysian cultivar. Tissues analysed included leaf, root, embryo and endosperm of germinating seed, milky stage seed, developing seed, husk and bran. Gene specific primers were used to amplify the selected putative lipase genes using cDNA as template, synthesized from RNA extracted from the tissues stated above. Reverse transcription PCR analysis suggested that LOC_Os01g71010, LOC_Os011g43510, LOC_Os09g01590 and LOC_Os01g55650 were expressed in rice bran tissue and hence selected for real time PCR analysis. Real time PCR analysis showed that LOC_Os01g71010 and LOC_Os011g43510 were highly expressed in rice bran. Interestingly, phylogenetic analysis showed that Lipase II (MSU ID: LOC_Os07g47250), a previously identified TAGL in rice bran (Aizono et al., 1976, Vijayakumar and Gowda, 2013) in the same group and had >50% protein identity with LOC_Os01g71010. But LOC_Os11g43510 was found to be highly expressed far more in both

embryo and developing seed than in the roots compared to LOC_Os01g71010 suggesting it might be a seed-specific lipase and thus selected for further studies.

Oleic acid and linoleic acid are major component of rice bran oil (RBO). Therefore, in order to predict the substrate (fatty acid) specificity of putative lipase encoded by LOC_Os11g43510 molecular modeling and protein docking studies were performed. In absence of three dimensional structure for the protein encoded by LOC_Os11g43510, a three dimensional model was constructed using SWISS MODEL server. The validated model was used for protein docking studies with different triglycerides (as substrates). Molecular modeling suggested that the putative lipase encoded by LOC_Os11g43510 has high specificity for oleic and linolenic acids. Thus, the presence of LOC_Os11g43510 encoded lipase in rice bran might degrade the triglycerides in RBO leading to low oil yield and more free fatty acids (FFA) which reduce the overall quality of oil stored in the bran.

In order to demonstrate that the encoded lipase is active, LOC_Os11g43510 was cloned in pPICZαA, yeast expression vector and transformed into a methylotropic yeast *P.pastoris* for protein expression studies. After the protein induction, the supernatant with the induced protein was tested by in gel lipase assays (Zymograms) performed on the SDS PAGE. Zymogram analysis showed that the protein encoded by LOC_Os11g43510 is active and detected only after 24h of induction and is at a maximum at 72h of induction.

The supernatant from 72h of induction was collected and the expressed protein was partially purified using a Ni-NTA column. The eluted fraction was subjected to deglycosylation using PNGase F and SDS PAGE revealed the shift in mobility of protein bands from 65 and 50 kDa to the expected size of around 42 and 35 kDa approximately.

To further confirm activity of lipase encoded by LOC_Os11g43510, western blotting was conducted with the supernatant at 72h from the LOC_Os11g43510 construct and empty pPICZαA vector as control. A band was detected only in the supernatant from LOC_Os11g43510 construct and not in the empty vector. This suggests LOC_Os11g43510 is an active lipase in rice bran which might affect RBO quality and quantity.

This chapter concludes that LOC_Os11g43510 is the second identified TAGL in rice bran whose activity has been demonstrated experimentally. Also, this study provides a new approach to identifying and demonstrating the activity of lipases.

References:

Aizono, Y., Funatsu, M., Fujiki, Y. & Watanabe, M. 1976. Purification and Characterization, of Rice Bran Lipase II. *Agricultural and Biological Chemistry,* 40**,** 317-324.

Vijayakumar, K. & Gowda, L. R. 2013. Rice (Oryza sativa) lipase: Molecular cloning, functional expression and substrate specificity. *Protein expression and purification,* 88**,** 67-79.

Journal of Cereal Science 71 (2016) 43–52

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

Lipase genes expressed in rice bran: LOC_Os11g43510 encodes a novel rice lipase

Gopal Ji Tiwari [a, b], May Ying Chiang [a, b], Jeremy Ryan De Silva [c], Beng Kah Song [a, b], Yee Ling Lau [c], Sadequr Rahman [a, b, *]

[a] School of Science, Monash University, 46150, Bandar Sunway, Selangor, Malaysia
[b] Monash University Malaysia Genomics Facility, 46150, Bandar Sunway, Selangor, Malaysia
[c] Department of Parasitology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form
1 July 2016
Accepted 9 July 2016
Available online 11 July 2016

Chemical compounds studied in this paper:
p-nitrophenyl acetate (PubChem CID: 13243)
Triacetin (PubChem CID: 5541)
Triolein (PubChem CID: 5497163)
Trilinolein (PubChem CID: 5322095)
Triarachidin (PubChem CID: 522017)
Tricaprin (PubChem CID: 69310)
Tripalmitin (PubChem CID: 11147)
Tristrearin (PubChem CID: 11146)
Tribehnin (PubChem CID: 62726)

Keywords:
Triacylglycerol lipase
MUF-butyrate
GXSXG motif
Germination
A. thaliana

ABSTRACT

Rice is a highly consumed staple food all over the world. The economic value of the rice crop can be further increased by producing rice bran oil (RBO) from rice bran which is a by-product of rice milling. However, high utilization of RBO is difficult to achieve as lipases present in rice bran cause decomposition of lipids present in the form of triacyl glycerol (TAG) into free fatty acids (FFA). In this work, we selected and systematically analyzed 125 putative lipase gene sequences derived from *Oryza sativa* ssp. japonica genome using bioinformatic tools. LOC_Os11g43510 was experimentally demonstrated to be highly expressed in rice bran. Further, molecular modeling and protein docking studies suggested that the protein encoded by LOC_Os11g43510, has high affinity for oleic acid and linoleic acid, common components of TAG in rice. The transcript from LOC_Os11g43510 was cloned and expressed as a secretory protein in *Pichia pastoris* X-33. SDS-PAGE and zymography showed that expressed protein had lipase activity and was glycosylated. These results are consistent with the hypothesis that LOC_Os11g43510 encodes an active lipase that could be involved in TAG breakdown in rice. This study demonstrates an alternative route to identifying lipases in rice bran.

1. Introduction

Lipases (EC 3.1.1.3.) catalyze the hydrolysis of the ester bonds of tri-, di-, and monoglycerides into fatty acids and glycerol (Casas-Godoy et al., 2012). They are versatile enzymes, widely distributed in plants, animals and microorganisms (Gupta et al., 2004; Aravindan et al., 2007; Patil et al., 2011). Under low water activity lipases can also catalyze synthesis reactions involving esterification and transesterification (Reis et al., 2009). Most of the triacylglycerol (TAG) lipases possess a GXSXG motif, with serine serving as the nucleophile in the active site, which is characteristic of hydrolytic enzymes (Schrag and Cygler, 1997). Lipases are thus triacyl hydrolases.

Plant lipases are often present in the reserve tissues of germinating seedlings or in tissues with large amount of triacylglycerols where they play important roles in biological reactions such as lipolysis, esterification and transesterification, thus helping in plant growth and development (Li et al., 2012). Plant lipid biosynthesis has been well studied (see for example, Slabas and Fawcett, 1992). However, most plant lipases which degrade the synthesized lipids are still uncharacterized and their regulation is not very clear. In particular there is little information on lipases in the grass family.

* Corresponding author. School of Science, Monash University, 46150, Bandar Sunway, Selangor, Malaysia.
E-mail address: [redacted] (S. Rahman).

http://dx.doi.org/10.1016/j.jcs.2016.07.008
0733-5210/

Rice bran oil (RBO) is emerging as a popular oil as it is typically high in oleic-linoleic- type fatty acids and contains naturally occurring antioxidant compounds (Rogers et al., 1993; Goffman et al., 2003). The TAG oil bodies are protected by oleosins, caleosins and stereoleosins (Napier et al., 1996). However, the full utilization of bran oil is hampered by the presence of lipase activity, which leads to decomposition of TAG and accumulation of free fatty acids (FFA) in RBO (Funatsu et al., 1971). This shortens the shelf life and leads to rancidity in milled rice bran and makes the oil produced unsuitable for human consumption (Ramezanzadeh et al., 1999). One way of reducing rancidity could be by reducing the expression of lipase genes (Da Silva et al., 2006). However, despite its importance in rice only a few genes have been associated with proteins that demonstrate lipase activity (Brick et al., 1995; Morohoshi et al., 2011; Vijayakumar and Gowda, 2013; Seth et al., 2014). This is due to difficulties in both purification of proteins and expression of candidate genes. Plants are eukaryotes therefore codon bias problems occur when trying to express plant genes in prokaryotic hosts like *Escherichia coli*. Furthermore, the expressed proteins can form inclusion bodies which are difficult to work with, making *E. coli* not a very suitable host for plant lipase expression (Vijayakumar and Gowda, 2013, Seth et al., 2014). The eukaryotic methylotrophic yeast, *Pichia pastoris* may serve as a better host for plant lipase expression as it overcomes issues such as codon bias and improper protein folding and often the expressed protein is directly secreted into the supernatant for easy recovery (Macauley-Patrick et al., 2005; Seth et al., 2014).

In this study, we describe the discovery of putative lipase gene sequences in rice using bioinformatics tools. Phylogenetic analysis was performed and expression profiling of the identified genes was conducted. The gene LOC_Os11g43510 was found to be highly expressed in rice bran and was selected for molecular modeling and docking analysis. The cDNA derived from the LOC_Os11g43510 gene was cloned and expressed in methylotrophic yeast *P. pastoris* X33 and its lipase activity was demonstrated. Our findings provide an alternative route for identification of lipases in specific plant tissues.

2. Materials and methods

2.1. Plant material

Oryza sativa cv MR219 seeds were imbibed with water in a submerged condition for a period of 24 h and germinated at 25 ± 2 °C on moistened layers of filter paper in sterile petri dishes. At each sampling point, the whole seedling, roots and shoots were harvested. Zero to 4 days after imbibition (DAI) and 5–10 DAI were considered as germination and post-germination respectively. The germinated seeds were transferred to soil in pots and grown until the mature plants developed.

Milky seed (white starchy milky fluid) and developing seed (hard dough) are seed ripening stages. Samples of milky seed were collected about one week after flowering and developing seed about two weeks after flowering from growing rice plants under our conditions. Rice hull was separated from the grown mature rice seeds.

Rice bran was collected from the Faiza Sdn Bhd (Subang Jaya, Selangor, Malaysia) rice mill, stored at 4 °C and used for further studies.

2.2. Sequence retrieval and analysis

A basic query of predicted protein function for lipase genes was performed in the rice Massively Parallel Signature Sequencing database (MPPS) (http://mpss.danforthcenter.org/rice/mpss_index.php) and the list of all 125 lipase genes was retrieved. However, clearly other databases (e.g. RAP database, http://rapdb.dna.affrc.go.jp/)) and schemes could have been used as the start point for this search. Further, all the lipase genes were grouped based on their fragments per kilobase of exon per million fragments mapped (FPKM) expression value derived from Rice Genome Annotation Project Database http://rice.plantbiology.msu.edu/. The Arabidopsis Acyl-Lipid Metabolism database (ARALIP) (http://aralip.plantbiology.msu.edu/) was also searched to identify the lipase genes present in *A. thaliana* and homologs were searched in the list of *O. sativa* lipases using the Protein BLAST program found at the NCBI site (http://blast.ncbi.nlm.nih.gov/).

Motifs or domains of lipase genes were analyzed using the Scan Prosite tool found at ExPASy – PROSITE (http://prosite.expasy.org/). Only genes with encoding protein sequences with the GXSXG lipase motif were selected for further analysis. Selected sequences were aligned using ClustalW implemented in MEGA version 6 (Tamura et al., 2013). Deduced protein sequences were aligned using GONNET as the protein weight matrix (Gonnet et al., 1992). Values of 10 and 0.1 were used for gap opening and gap extension penalties, respectively. The multiple alignments were inspected visually, and columns with ≥50% gaps were excluded. Phylogenetic analysis was conducted using neighbor-joining (NJ) method in MEGA version 6 under the Jones-Taylor-Thornton amino acid matrix-based model of molecular evolution without rate variation among sites or homogeneous pattern among lineages. The phylogenetic tree was generated by implementing the NJ method. Subcellular Localization prediction for lipase genes was carried out using TargetP 1.1 (http://www.cbs.dtu.dk/services/TargetP/).

2.3. Molecular modeling and docking

The template for building 3D models was selected using the template identification tool SWISS MODEL (www.swissmodel.expasy.org). Consequently, based on the template, 3D structures of lipase proteins were modeled using SWISS-MODEL server, which is a fully automated protein structure homology-modeling server, accessible via the ExPASy web server, or from the program DeepView (Swiss Pdb-Viewer).

Once the 3D structures of lipases were modeled the geometrical aspects of modeled protein structures were evaluated using Qualitative Model Energy Analysis (QMEAN) server (http://swissmodel.expasy.org/qmean/cgi/index.cgi). Also Ramachandran plots for predicted 3D structures were generated using RAMPAGE server (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) for structure validation.

Three-dimensional structures of triglycerides and esters used for protein docking were generated using online SMILES Translator and Structure File Generator server (http://cactus.nci.nih.gov/translate/) Fatty acids to be used as ligands were optimized using CLC drug discovery workbench software version 2.5. Modeled 3D structures of lipases and fatty acid structures were imported into CLC drug discovery workbench software version 2.5 and molecular docking was performed (http://www.clcbio.com/products/clc-drug-discovery-workbench/) with optimized ligands (triglycerides).

2.4. Expression analysis of TAGL genes by RT-PCR

Semi-quantitative RT-PCR was carried out to analyze the relative expression of nine representative putative lipase genes. All PCR primers were designed with Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) according to product size (500–800 bp) specification. Primers were based on exon sequences that spanned introns in the genomic DNA, allowing one to easily

detect DNA contamination as product sizes from the cDNA and genomic DNA template would differ. Total RNA from *O. sativa* was extracted using RNeasy Plant Mini Kit (Qiagen). RNA concentration and purity factor (A260/A280 ratio) was determined by spectrometry.

Reverse transcription was performed using Reverse Transcription System kit (Promega) according to manufacturer's protocol. RT-PCR was carried out from three biological replicates using MyTaq DNA Polymerase system (Bioline) in a total volume of 50 μL. Equal volumes (2 μL) of each sample was analyzed in 1% agarose gel. Quantitative real time PCR reaction was also carried out in triplicates using iTaq universal SYBR Green supermix (BioRad) and 0.5 μM of each gene specific primer in a final volume of 10 μl reaction were used. The following standard thermal profile was used for all qRT-PCRs: 95 °C for 30 s; 40 cycles of 95 °C for 5 s and 60 °C for 20 s with no extension as amplicon size were small ranging from 100 to 120 bp. Data were analyzed using the Rotor-Gene Q series software (Qiagen) using comparative quantification method which compares the relative expression of target genes to the expression of a reference or "housekeeping" gene. In this case the ubiquitin gene (Genbank: AK061988) was used as the control. Equal volumes (2 μL) of each sample were analyzed in 1.5% agarose gels. Each experiment was repeated at least three times to ensure statistical significance.

2.5. Construction of the pPICZαA-L2 expression vector

The cDNA clone AK099612 selected from Knowledge-based Oryza Molecular biological Encyclopedia (KOME) database and obtained from the Rice Genome Resource Centre (http://www.rgrc.dna.affrc.go.jp/index.html.en), was used as the template for PCR. EcoRI and XbaI sites were introduced at the 5′ and 3′ ends using primer pairs L2F1 and L2R1 (Appendix A, Table A.1). Sub cloning in pPICZαA was done by ligating double digested insert and vector plasmid and later transformed into *E.coli* DH5α competent cells. Low salt LB plates containing zeocin (25 μg/mL) were used for selecting positive clones. Clones were further confirmed by double digestion and also by PCR using L2_F1 and L2_R1 primers (Appendix A, Table A.1). Sanger sequencing of positive clones was also done to check the correct reading frame of inserted gene with the α-factor secretory signal.

2.6. Expression of pPICZαA-L2 in P. pastoris

pPICZαA-L2 plasmid was linearized using *Sac* I and used for electroporation in electro-competent *P. pastoris* X33 cells using Micro Pulser™ (Bio-rad) with the preset protocol for yeast. After electroporation, transformants were selected on Yeast Extract Peptone Dextrose with Sorbitol (YPDS) agar, containing 100 μg/mL zeocin at 30 °C until colonies appeared. Positive clones were confirmed by PCR using L2F1 gene specific primer and AOX 3' primers (Appendix A, Table A.1).

Selected clones were cultured overnight in 200 mL of buffered glycerol-complex medium (BMGY) at 30 °C with shaking (250 rpm) until OD_{600} reached 2–6. The cells were harvested by centrifugation at 1300 × g for 5 min at room temperature and the pellet was resuspended in 100 mL of buffered methanol-complex medium (BMMY). To induce expression incubation was continued at 30 °C with shaking (250 rpm) for 72 h. Sterilized methanol (5 g/L) was added every 24 h to maintain and continue induction of pPICZαA-L2. At certain time points (0 h, 24 h, 48 h, 72 h) 1 mL of induced culture was centrifuged at 1300 × g for 3 min at room temperature. Supernatant and pellets were stored at −20 °C in different tubes for further analysis.

His_6-tag purification and deglycosylation of pPICZαA-L2 was carried out from the induced supernatant from the 72 h time point which was lyophilized. Approximately 2.8 g of lyophilized supernatant was mixed with 1 mL of Ni-NTA agarose (Qiagen). The column was packed and protein was partially purified as per manufacturer's protocol. Later partially purified protein was deglycosylated using PNGase F (Biolabs) as per manufacturer's protocol and analyzed by SDS PAGE.

Following PAGE, western blot analysis was conducted by the wet-transfer of protein bands from gel to nitrocellulose membranes (Whatman Inc., ME, US) using Tris-glycine pH 8.8 (0.025 M Tris and 0.192 M glycine) containing 20% methanol (v/v) and 0.1% SDS (W/V). HisProbe-HRP conjugate antibody (1:5000 dilution) independent of tag position was used to detect the expressed protein on the membrane. Western blotting was performed according to manufacturer's protocol using SuperSignal™ West HisProbe Kit (Thermo Scientific).

3. Results and discussion

3.1. Sequence analysis

Data available at the rice Massively Parallel Signature Sequencing (MPPS) database (http://mpss.danforthcenter.org/rice/mpss_index.php) was mined by using lipase as a keyword to search for all the lipase genes present in the annotated *O. sativa* genome and a list of 125 lipase genes was obtained (Appendix B). All of these lipase genes were grouped based on Fragments Per Kilobase of transcript per Million (FPKM) expression values obtained from the MSU Database (http://rice.plantbiology.msu.edu/). Among the 125 lipase genes initially retrieved, 69 lipase genes were reported to be expressed in seed and leaf tissue of *O. sativa*. The remaining 56 lipase genes were expressed in tissues other than the leaf or seeds and not selected for further studies (Appendix C).

The amino acid motif, GXSXG, is commonly found in lipases (Akoh et al., 2004). A lipase motif search analysis was performed for the proteins encoded by the 69 lipase genes. The lipase motif was not encoded in 15 lipase genes and these therefore not included for further study. The remaining 54 genes encoded proteins with the lipase motif in their deduced amino acid sequences. Of these, 7 of the encoded proteins contained the motif at the N-terminal end, 14 contained the motif at the C terminal end and 33 contained the lipase motif in the middle.

3.2. Phylogenetic inference of lipase genes

TAG is the major constituent in oil. Therefore our study specifically aimed to find TAG lipase genes in rice using Arabidopsis as a guide. To do this the 54 selected rice genes were compared against the annotated Arabidopsis genes for sequence identity using BLAST. The Arabidopsis genome encodes only 15 triacylglycerol lipase (TAGL) genes (http://aralip.plantbiology.msu.edu/).

Of these the proteins encoded by loci AT2G15230, AT3G57140, AT5G04040, and AT5G14930 are annotated as LIP1, SDP1-Like, SDP1 and SAG101 in the Arabidopsis genome database (www.arabidopsis.org) and have been well characterized (El-Kouhen et al., 2005, Eastmond, 2006; He et al., 2001; Kelly et al., 2013). Table 1 indicates homology percent identity/similarity for putative rice lipase proteins against experimentally demonstrated Arabidopsis lipase proteins. The TAGL genes from Arabidopsis are indicated by brackets. Nine lipase genes among all the 54 putative rice lipase genes encoded proteins with high levels of identity (≥50%) with the products of experimentally proven TAGL lipase genes from Arabidopsis (Table 1).

To further confirm the relationships a phylogenetic tree was constructed using proteins sequences containing GXSXG motifs that were encoded by the 54 lipase genes from rice and all TAGL

Table 1
Percent identity between rice lipases and Arabidopsis counterparts, GXSXG lipase motif position and subcellular localization prediction of putative rice lipases. The selected rice lipase genes are in bold.

S.No.	Loci (os)[a]	GXSXG Motif position	Total no. Amino acid	Homolog (At)[b]	Identity (%)	Subcellular localization
1	LOC_Os01g14080	183–187	359	AT5G50890	54	Un
2	LOC_Os01g15000	436–440	707	AT1G02660	48	C
3	LOC_Os01g20840	256–260	635	AT4G16070	56	S
4	LOC_Os01g21560	98–102, 138–142	329	AT5G19290	63	Un
5	LOC_Os01g46290	238–242	420	–	–	S
6	LOC_Os01g49380	168–172, 209–213	387	AT1G77420	61	M
7	LOC_Os01g49510	132–136, 169–173	364	AT5G38220	62	M
8	**LOC_Os01g55650**	239–243, 265–269	1044	[AT5G04040]	66	Un
9	LOC_Os01g47610	192–196	391	AT5G50890	51	Un
10	LOC_Os01g10250	177–181, 216–220	395	AT5G11650	71	Un
11	LOC_Os01g51360	231–235	465	AT2G42690	49	Un
12	**LOC_Os01g71010**	173–177	346	[AT5G18640]	57	S
13	LOC_Os01g43140	409–413	652	AT3g61680	48.5	M
14	LOC_Os01g46240	188–192	358	AT4g18550	48.2	M
15	LOC_Os01g67420	446–450	773	AT1G05790	48	S
16	LOC_Os01g46290	238–242	420	–	–	S
17	LOC_Os02g18480	160–164	349	AT4G10955	55	Un
18	LOC_Os02g28040	142–146, 655–659, 888–892	912	AT3G07400	56	Un
19	**LOC_Os02g52830**	303–307	482	[AT5G42930]	49	Un
20	LOC_Os02g55330	74–78, 111–115	264	AT4G31020	77	Un
21	LOC_Os02g03720	117–121	307	AT5G17670	73	M
22	LOC_Os02g09770	125–129, 162–166	389	AT3G30380	59	S
23	LOC_Os02g43700	363–367	544	AT4G16820	47	C
24	LOC_Os02g54010	264–268	657	AT3G14075	59	C
25	**LOC_Os03g51010**	179–183, 218–222	392	[AT1G73480]	55	M
26	LOC_Os03g61540	271–275	594	AT4G16070	42	S
27	LOC_Os03g50410	158–162	382	–	–	C
28	LOC_Os03g02740	340–344	904	AT2g42010 AT4g00240 AT4g11830 AT4g11840 AT4g11850	64.7 65.8 62 62.8 62.7	Un
29	LOC_Os04g41200	501–505	773	AT1G05790	45	S
30	LOC_Os04g43030	691–695, 792–796	870	AT4G13550	53	C
31	**LOC_Os05g32380**	338–342	577	AT1G06800	56	C
32	**LOC_Os05g30900**	169–173	342	[AT5G18640]	60	M
33	LOC_Os05g33820	175–179	471	[AT1G10740]	70	S
34	LOC_Os05g49840	228–232	407	AT4G18550	52	Un
35	LOC_Os06g42730	117–121, 154–158	389	AT3G30380	60	C
36	LOC_Os06g44060	475–479	512	AT5G62810	34	S
37	LOC_Os06g40170	543–547	832	AT3G15730	61	Un
38	LOC_Os06g40180	553–557	842	AT3G15730	61	Un
39	LOC_Os07g37840	101–105, 142–146	334	AT1G11090	62	Un
40	LOC_Os07g33670	597–601	1226	AT1G61850	66	Un
41	**LOC_Os07g47250**	185–189	358	[AT5G18640]	53	M
42	LOC_Os07g39810	152–156	253	–	–	Un
43	LOC_Os08g06420	113–117, 150–154	389	AT1G32190	61	S
44	**LOC_Os08g41780**	205–209	438	[AT5G14180]	57	S
45	**LOC_Os09g01590**	180–184	410	[AT2G15230]	63	S
46	LOC_Os09g39790	289–293	518	AT2G42450	55	M
47	LOC_Os09g23150	177–181,	397	AT1G73480	63	C
48	LOC_Os10g38060	188–192	1046	AT2G42010	61	Un
49	LOC_Os11g01040	79–83, 117–121	332	AT2G39420	54	Un
50	LOC_Os11g43510	193–197	366	[AT5G18640]	61	M
51	LOC_Os12g01030		332	AT2G39420	54	Un

Table 1 (*continued*)

S.No.	Loci (os)[a]	GXSXG Motif position	Total no. Amino acid	Homolog (At)[b]	Identity (%)	Subcellular localization
		79–83, 117–121				
52	LOC_Os12g18860	108–112, 145–149	377	AT4G24760	67	S
53	LOC_Os12g36770	267–271	417	AT4G16070	50	S
54	LOC_Os12g37630	515–519	625	AT1G09280	58	Un

[a] *O.sativa.*
[b] *A.thaliana*(C –chloroplast, M- mitochondria, S- secretory pathway, Un - unknown, [] – TAGL gene).

lipases from Arabidopsis (Fig. 1). The phylogenetic relationship was analyzed by MEGA version 6 using the Neighbor-Joining (NJ) method. It further confirmed that the selected nine rice putative genes were the closest homologues of the reported Arabidopsis TAGL genes (Fig. 1). It was clear that LOC_Os09g01590 and LOC_Os01g55650 encode proteins that are homologs of Arabidopsis LIP1 (At2g15230) and SDP1 (At5g04040) respectively with 63% and 66% sequence identity (Table 1). Moreover, LOC_Os01g55650 was found to be homologous to SDP1-like lipase with 60% identity. However, no homolog of the Arabidopsis SAG101 lipase gene was found in rice.

3.3. Expression analysis of O. sativa TAGL genes homologous to A. thaliana

Based on sequence identities (Table 1) and phylogenetic analysis (Fig. 1) *O.sativa* genes encoding at least one lipase motif and also showing the greatest identity to the TAGL genes which have been experimentally studied in *A.thaliana* were selected for further analysis. Reverse transcription PCR was carried out for nine lipase genes (LOC_Os01g71010, LOC_Os11g43510, LOC_Os09g01590, LOC_Os07g47250, LOC_Os05g30900, LOC_Os02g52830, LOC_Os05g33820, LOC_Os08g41780, LOC_Os01g55650) from *O. sativa* (Table 1).

Gene expression was tested in leaf, root, embryo and endosperm of germinating seed, milky stage seed, developing seed, husk and bran, using ubiquitin (Genbank: AK061988) as an internal control. From reverse transcription PCR analysis four genes LOC_Os01g71010, LOC_Os011g43510, LOC_Os09g01590 and LOC_Os01g55650 were found to be expressed in the bran and therefore selected for quantitative real time PCR analysis (Appendix A, Fig.A.1).

In leaf tissue LOC_Os01g71010, LOC_Os011g43510 and LOC_Os1g55650 genes were all were expressed 1.5 fold higher compared to ubiquitin, while LOC_Os09g01590 showed similar levels of expression to ubiquitin (Fig. 2a).

In roots, only LOC_Os01g71010 showed about 1000 fold increase in the expression compared to ubiquitin (Fig. 2b).

In embryo of germinating seed all four genes (LOC_Os01g71010, LOC_Os011g43510, LOC_Os09g01590 and LOC_Os01g55650) showed low expression levels suggesting none of these genes were involved in germination. However in the endosperm of the germinating seed, LOC_Os01g71010 showed 3.5 fold higher expression compared to control indicating it might be involved in breaking down starchy lipids (Liu et al., 2013), thus helping in seed germination. LOC_Os011g43510 was expressed lowest in endosperm compared to the other genes tested (Fig. 2a).

The first stage of seed ripening is known as milky seed. At this stage the expression of LOC_Os01g71010 and LOC_Os011g43510 was found to be increased by approximately two fold compared to control (Fig. 2a). At a later stage of seed development (approximately mid-development) LOC_Os011g43510 showed approximately 15 fold and LOC_Os01g71010 about two fold higher expression compared to the ubiquitin control gene (Fig. 2a). These results indicate that both of these genes are strongly expressed during seed development. The other selected genes showed low levels of expression.

In the husk LOC_Os09g01590 and LOC_Os01g55650 showed two fold higher compared to the control whereas the expression of LOC_Os01g71010 and LOC_Os011g43510 genes was similar to the control gene (Fig. 2a).

Expression of LOC_Os01g71010 and LOC_Os11g43510 was found to be significantly higher in bran tissue compared to the control ($(p < 0.001)$, $n = 3$) whereas LOC_Os01g55650 and LOC_Os09g01590 were expressed at low levels (Fig. 2b). Phylogenetic analysis (see earlier) had suggested that both LOC_Os01g71010 and LOC_Os11g43510 genes are grouped together and both of them are homologous to the same TAGL gene (AT5G18640) in *A. thaliana* (Fig. 1) with sequence identity of 57 and 61% respectively (Table 1). Phlyogenetic analysis also shows that both LOC_Os01g71010 and LOC_Os11g43510 putative lipase genes are homologues of Lipase II (MSU ID: LOC_Os07g47250) (Fig. 1) whose activity has been demonstrated recently by Vijayakumar and Gowda (2013).

In comparison to LOC_Os01g71010, LOC_Os11g43510 was found to be highly expressed in both embryo and developing seed but not in the roots suggesting it is a more seed-specific lipase and may have an important role in lipid breakdown in the grain. Moreover, it also has higher protein sequence identity (61%) with AT5G18640. Therefore, based on quantitative PCR and phylogenetic analysis LOC_Os11g43510 was selected for further studies.

3.4. Molecular modeling of LOC_Os11g43510

As indicated above quantitative expression data suggested that LOC_Os11g43510 encoded a lipase that was highly expressed in rice bran. Molecular modeling and docking studies were conducted to investigate whether the selected lipase genes could be expected to interact with oleic or linoleic acids which are the most abundant fatty acids in rice triacyl glycerols (TAG). TAGs with long and short fatty acids were thus used as substrates to investigate the substrate specificity. *Rhizomucormeihei* lipase (PDB id: 3TGL) having 25.9% identity (protein) with LOC_Os11g43510 encoded protein, was used as a template to generate the three dimensional model required for this analysis. The modelled structure was superimposed with the template (PDB id: 3TGL), and it showed overall root mean square deviation (RMSD) of 1.66 suggesting close structural similarity among the modelled and template structure (Fig. 3a and b). Geometrical aspects of the modelled structure were evaluated using QMEAN server which showed that more than 80% of the modelled secondary structure was in a favorable region (Appendix A, Fig. A.2). Also, the modelled structure was validated by predicting Ramachandran plots using the RAMPAGE program which indicates only 5% of residues fall in the disallowed region (Appendix A, Fig. A.2). Thus the predicted modelled structure is of good quality and it was used for protein docking studies with different

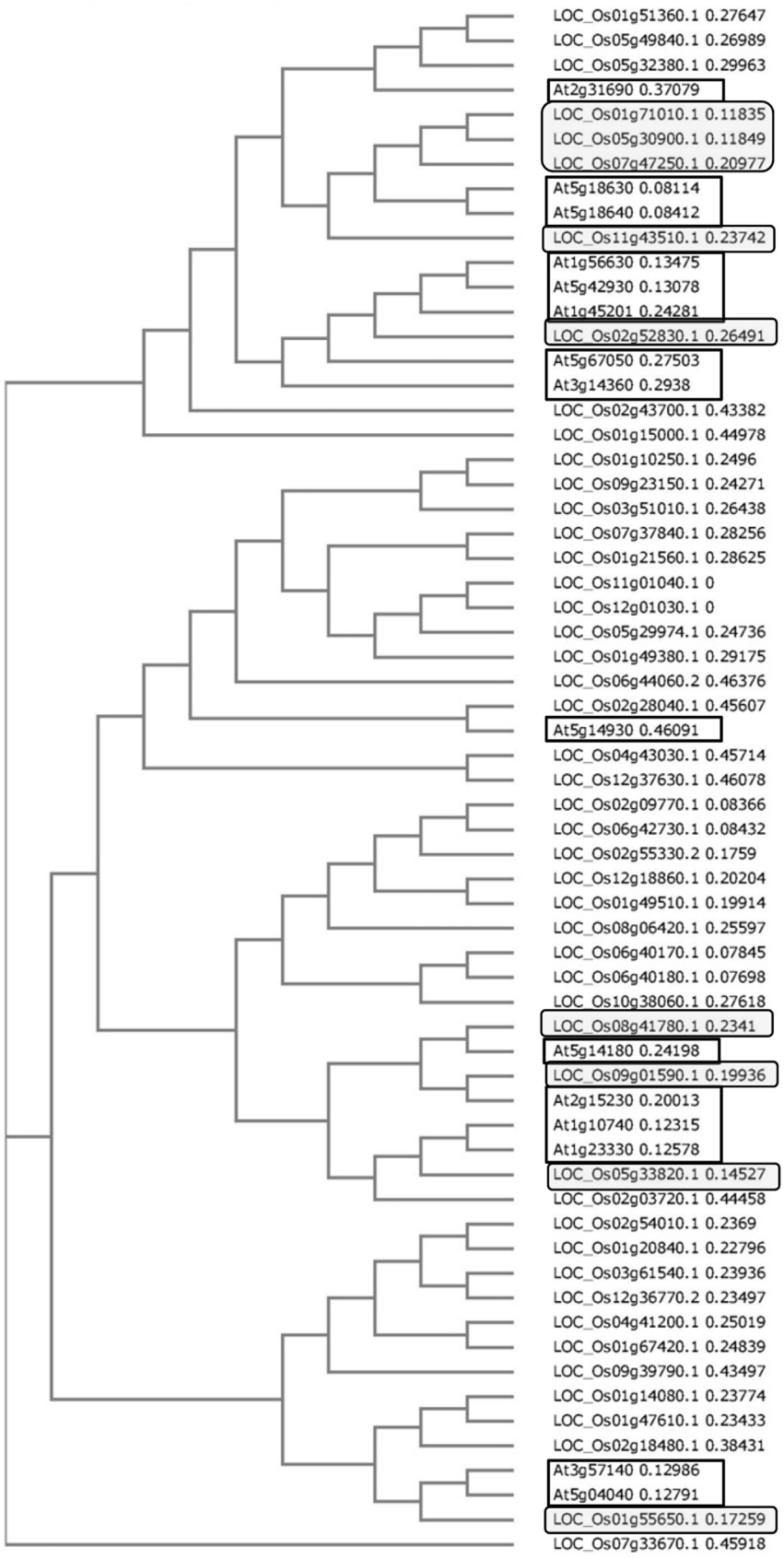


Fig. 1. Phylogenetic analysis of lipase proteins having GXSXG lipase motif and expressing higher in rice along with known TAGL (shown in rectangular box) in *A. thaliana* and selected putative rice TAGL (in shaded box) for further analysis. Distance displayed after each accession number.

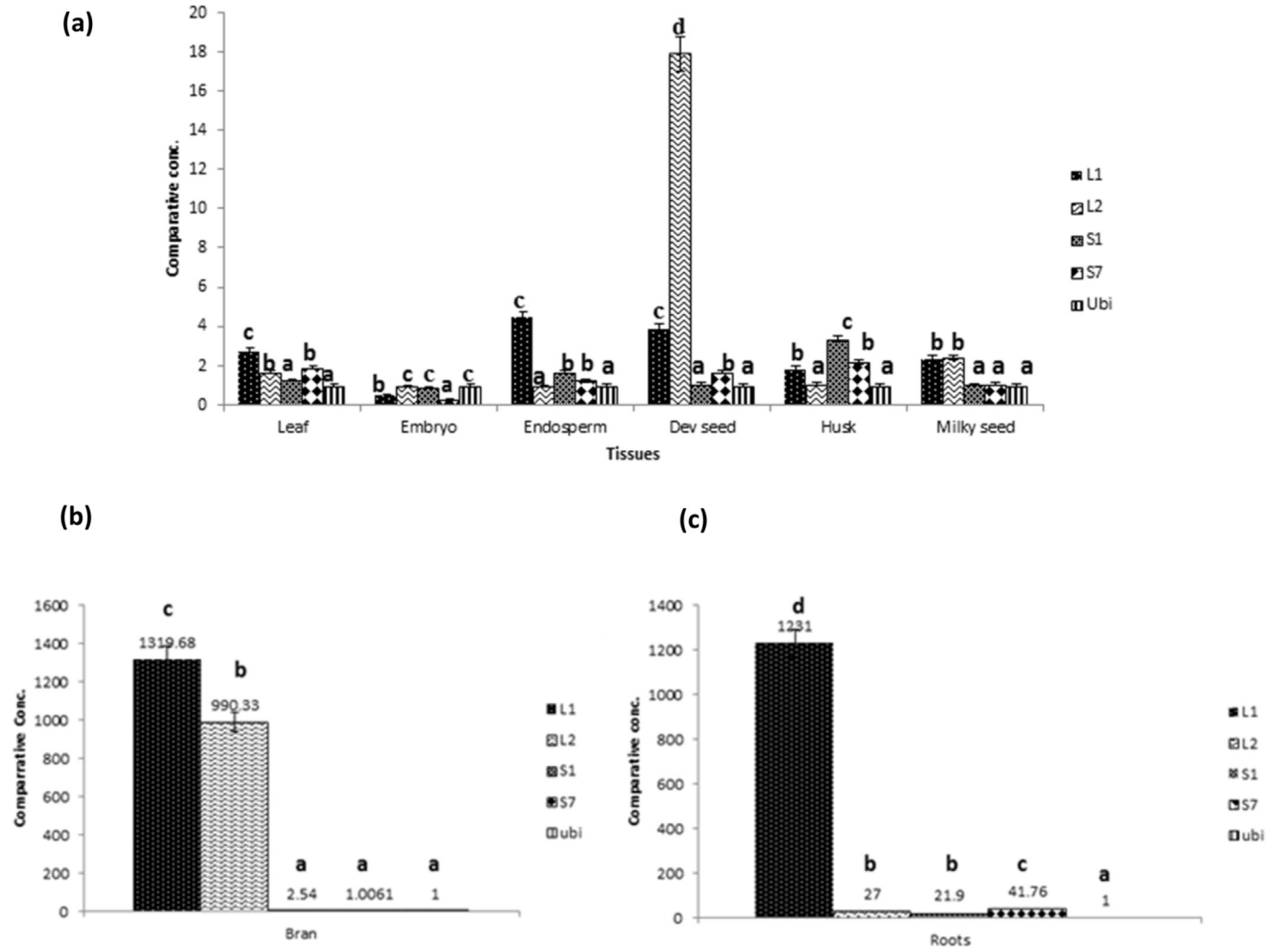


Fig. 2. Graphs showing transcript accumulation of lipase genes among different tissues of rice plant. (a) Comparison of relative expression of lipase genes among leaf, embryo, endosperm, developing seed, husk and milky seed tissues. (b) Comparison of relative expression of lipase genes in bran (c) Comparison of relative expression of lipase genes in roots. The genes analyzed are LOC_Os01g71010 (L1), LOC_Os011g43510 (L2), LOC_Os09g01590 (S1) and LOC_Os01g55650 (S7). Results are presented as mean ± SD where n = 3; different lowercase letter indicates significant difference between different types of genes within each tissue type (Data analyzed with one-way ANOVA).

triglycerides as ligands (substrate) (Appendix A, Table A.2).

Protein-ligand docking was performed using the CLC drug discovery workbench software that uses the automated docking software Molegro Dock algorithm and this was used to study the interaction of the modelled structure of the polypeptide encoded by LOC_Os11g43510 with various triglycerides (Appendix A, Table A.2). The protein docking results of various ligands (triglycerides) with the putative lipase is shown in Fig. 3(c–k). LOC_Os11g43510 encoded lipase was found to have very high affinity/specificity for triglycerides containing unsaturated (oleic and linoleic) fatty acid residues (Table 2). Table 2 also suggests that LOC_Os11g43510 has high affinity for 16–18 carbon chain fatty acids in a triglyceride and binding affinity of LOC_Os11g43510 declines with increase in fatty acid chain length (20–22 carbons) (Table 2). Thus the results presented in Table 2 further suggest that the protein encoded by LOC_Os11g43510 encodes a lipase that might degrade oleic and linoleic fatty acids. These fatty acids are known to be major components of RBO (Choudhury and Juliano, 1980; Zaplin et al., 2013).

3.5. Expression of LOC_Os11g43510 in P. pastoris

Based on the results above it is important to demonstrate that the protein encoded by LOC_Os11g43510 is an active lipase. A full length clone (Genbank: AK099612) encoded by LOC_Os11g43510 was obtained (see Materials and Methods). This clone is referred to as L2 hereafter. Despite repeated attempts, expression of L2 sequence as protein could not be achieved in *E.coli*. Therefore, *P. pastoris*, which is a methylotropic yeast, was used for expression of L2. L2 was cloned in the yeast expression vector pPICZαA using EcoRI and XbaI restriction sites. The vector contains the α-factor secretion signal which helps in secretion of expressed protein in the supernatant. Transformants were screened by PCR (Appendix A, Fig. A. 4) using L2F1 gene specific and 3′AOX1 primers and double digestion (Appendix A, Fig. A. 5).

Transformed and induced *P. pastoris* supernatants were analyzed for expression of the L2 sequence as protein. As induced bands could not be detected by Coomassie staining, zymography or in-gel lipase assay was carried out to detect successful lipase induction. The assay is based on the principle that proteins with lipase activity should cleave MUF-butyrate (see Materials and Methods) to produce a fluorescent product, which can be detected.

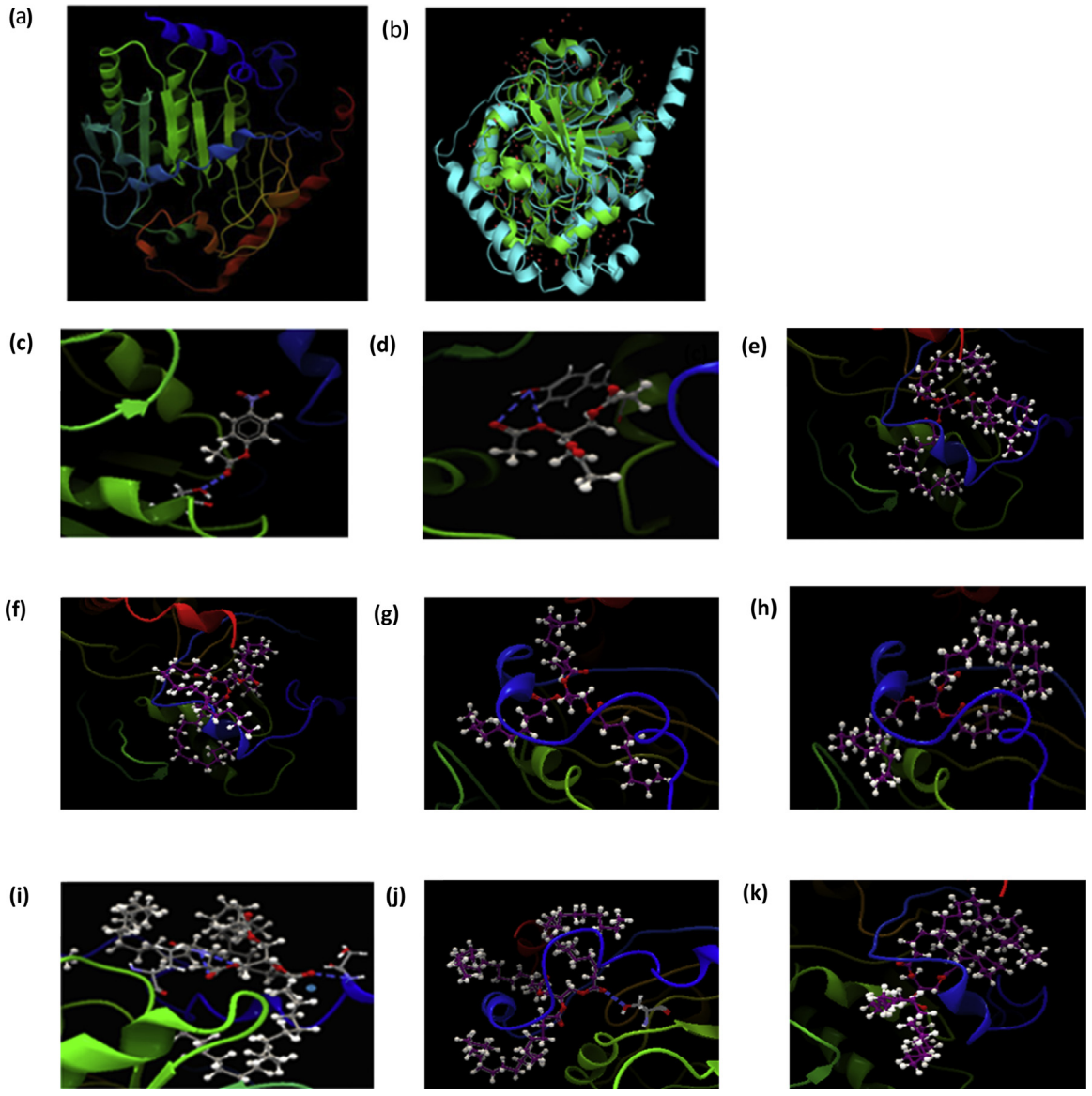


Fig. 3. (a) The three-dimensional modeled structure of proteins encoded by LOC_Os11g43510 (b) Superimposed structure of LOC_Os11g43510 (green) with its template PDB ID:3TGL (cyan) with overall root mean square deviation (RMSD) of 1.66. Prediction of substrate binding with proteins encoded by LOC_Os11g43510 (c) pNPA (d) Triacetin (e) Triolein (f) Trilinolein (g) Tricaprin (h) Tripalmitin (i) Tristearin (j) Triarachidin (k) Tribehnin. Ligands are shown in ball and stick model and dotted blue lines indicate the hydrogen bonding. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Docking score and hydrogen bonds observed between different docked substrates (triglycerides) with polypeptide encoded by LOC_Os11g43510. Lower the score higher is the affinity for that particular substrate.

S.No.	Substrate/ligand	Score	No. of H-bond	Amino acid residues involved in H-bonding
1	pNPA	−22.82	1	O → SER209
2	Triacetin	−23.36	2	O → TYR 235 O → TYR 235
3	Triolein	−86.36	0	
4	Trilinolein	−65.92	0	
5	Triarachidin	−52.13	1	O → SER272
6	Tricaprin	−70.87	0	
7	Tripalmitin	−78.12	0	
8	Tristrearin	−81.06	0	
9	Tribehnin	−55.94	0	

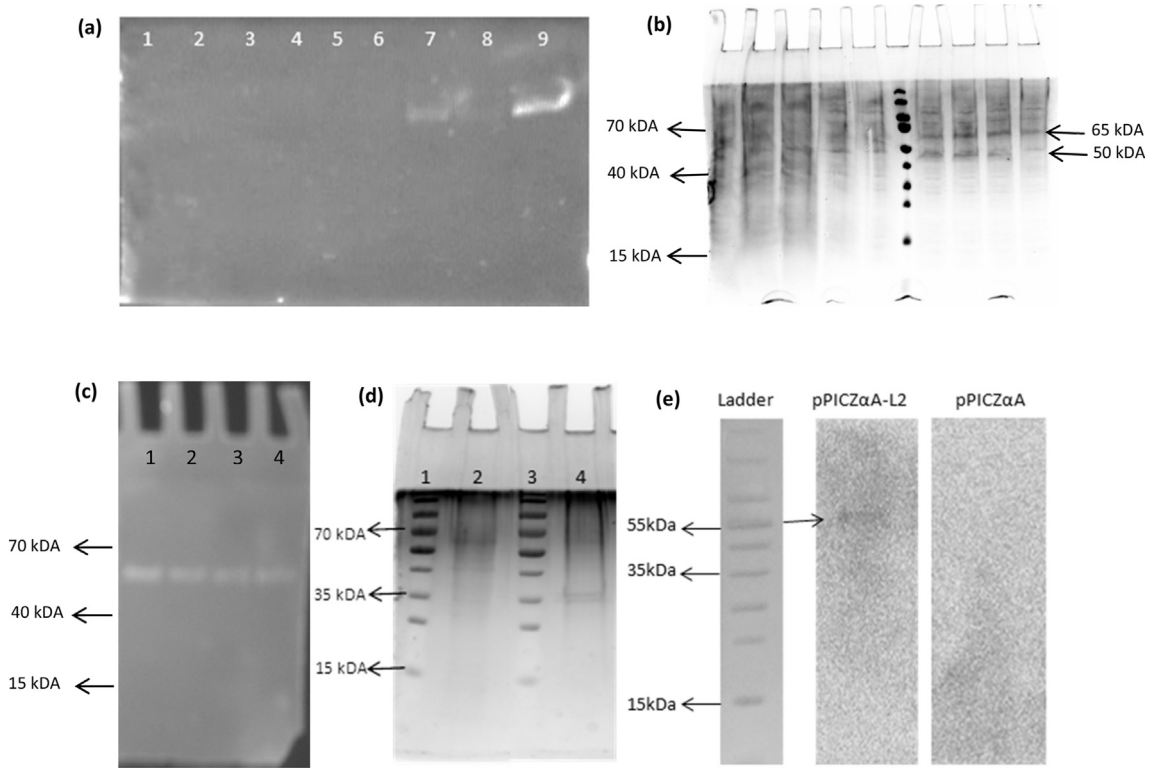


Fig. 4. Expression analysis of LOC_Os11g43510 (a) Detection of lipase activity at different time points by lipase-zymogram analysis using MUF butyrate. lane 1 Protein marker, lane 2 – Empty vector (0 h), lane 3 - clone (0 h), lane 4- Empty vector (24 h), lane 5- clone (24 h), lane 6- Empty vector (48 h), lane 7- clone (48 h), lane 8- Empty vector (72 h), lane 9- clone (72 h). (b) Coomassie stained 10% SDS-PAGEs of expression of recombinant pPICZαA-L2 in *P. pastoris* X-33 and His6-tag partial purification on a Ni-NTA agarose column. Lane 1- lyophilized culture supernatant from pPICZαA-L2 in *P. pastoris* X-33 after 72 h of induction, lane 2–3 – flow through, lane 4–5 – washing, lane 6 – 5 μl prestained protein marker, lane 7–10 – eluted protein. (c) Lipase activity staining of eluted fraction by MUF-butyrate lane 1–4 – eluted protein (same gel (b) used for activity staining). (d) Deglycosylation of partially purified expressed pPICZαA-L2 protein. Lane 1,3–5 μl prestained protein marker, lane 2 - partially purified protein, lane 4 – deglycosylated partially purified pPICZαA-L2 protein showing expected size of expressed protein ~35 kDa. (e) Western blotting analysis of pPICZαA-L2 clone transformed and expressed in *P. pastoris* and empty pPICZαA vector with HisProbe-HRP conjugate antibody confirms the expression of LOC_Os11g43510 (L2) in *P. pastoris* as no band was observed with empty pPICZαA vector.

Zymography results suggested that lipase activity increased with time. No activity was found at any time point in supernatants from cultures of *P. pastoris* containing a vector without any insert (Fig. 4a). From the gel image it is clearly seen that the fluorescent lipase activity signal increases with time and is continuing to increase at the 72 h time point (Fig. 4a). It was therefore decided to use the supernatant from this time point for further analysis as L2 is being highly expressed at the 72 h time point.

3.5.1. Partial purification and deglycosylation of expressed pPICZαA-L2 in P. pastoris

The expressed protein is tagged with six Histidines. His_6 tagged expressed protein was separated from supernatant using Ni-NTA affinity column. Fig. 4b shows the image of SDS PAGE gel where all the fractions (flow through, washing and elution) are stained with Coomassie blue stain. Eluted fractions show reduction of other secreted proteins compared to flow though and washing step fractions. In the eluted fractions two distinct bands corresponding to approximately 65 and 50 kDa were observed (Fig. 4b).

The same SDS PAGE gel with eluted fractions of partially purified L2 His_6 tagged expressed protein was tested for the presence of lipase activity using MUF-butyrate substrate. The lower band corresponding to 50 kDa showed lipase activity using MUF-butyrate (Fig. 4c). This is consistent with the expectation that the protein encoded by LOC_Os11g43510 encoded an active lipase that was being expressed in *P. pastoris*. However, the apparent molecular mass is higher than expected. This could be due to post translational glycosylation which is known to commonly occur in yeasts (eg Macauley-Patrick et al., 2005).

The eluted fraction was subjected to deglycosylation using PNGase F. SDS PAGE revealed the shift in mobility of protein bands from 65 to 50 kDa (Lane 2, Fig. 4d) to the around 42 and 35 kDa approximately (Lane 4, Fig. 4d). The expected mass of the L2 encoded protein is 35 kDa and the experimental result is consistent with the hypothesis that the L2 protein is glycosylated when expressed in *P. pastoris*. However, the relationships of the 35 kDa and 42 kDa polypeptides to the 50 kDa or 65 kDa polypeptides cannot be deduced with certainty.

3.5.2. Western blotting

The supernatant from induced pPICZαA-L2 and empty vector (pPICZαA) clone at 72 h was concentrated and subjected to western blotting analysis. The product from pPICZαA-L2 showed cross reactivity with HisProbe-HRP Conjugate antibody and a single band of approximately 50 kDa was detected and no band was seen in the lane with induced supernatant from empty vector (Fig. 4e). This demonstrates that the product in *P.pastoris* with lipase activity is being expressed from the introduced vector and is consistent with the hypothesis that the cDNA product of LOC_Os11g43510 (L2) was successfully cloned and expressed in *P. pastoris* and encodes a lipolytic activity. It is not clear why the 65 kDa polypeptide was not recognised.

4. Conclusions

Reduction of lipase activity is important for maintenance of the

rice bran oil quality. However, progress in identifying the lipases involved in rice oil breakdown has been slow, largely because of the difficulties of purifying active lipases. We have used publicly available databases to identify likely lipase genes and assessed the expression of a selection of these experimentally in different rice tissues. Based on the results obtained, the cDNA derived from one gene (LOC_Os11g43510) was expressed in *P. pastoris* and a lipase activity was demonstrated to be produced. The route followed in this paper indicates an alternative approach to purifying lipases from specific plant tissues and identifying them. In the future, the expression of the identified lipase gene LOC_Os11g43510 could be down regulated using different techniques such as gene editing or RNAi and the effect on stability of lipids in the bran could be assessed. In addition, a survey of germplasm could lead to the isolation of rice accessions not expressing this gene in the bran. Such material would allow direct evaluation of the importance of lipase in the breakdown of lipids in rice bran.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgement

We would like to thank Dr. Günther Daum, Graz University of Technology, Austria, for providing *P. pastoris* X33 strain for protein induction experiments. We also thank Monash University, Malaysia for research funding.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jcs.2016.07.008.

References

Akoh, C.C., Lee, G.-C., Liaw, Y.-C., Huang, T.-H., Shaw, J.-F., 2004. GDSL family of serine esterases/lipases. Prog. Lipid Res. 43, 534–552.
Aravindan, R., Anbumathi, P., Viruthagiri, T., 2007. Lipase applications in food industry. Indian J. Biotechnol. 6, 141.
Brick, D.J., Brumlik, M.J., Buckley, J.T., Cao, J.-X., Davies, P.C., Misra, S., Tranbarger, T.J., Upton, C., 1995. A new family of lipolytic plant enzymes with members in rice, arabidopsis and maize. FEBS Lett. 377, 475–480.
Casas-Godoy, L., Duquesne, S., Bordes, F., Sandoval, G., Marty, A., 2012. In: Sandoval, G. (Ed.), Lipases: an Overview. Lipases and Phospholipases: Methods and Protocols. Springer.
Choudhury, N.H., Juliano, B.O., 1980. Effect of amylose content on the lipids of mature rice grain. Phytochemistry 19, 1385–1389.
Eastmond, P.J., 2006. Sugar-dependent1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating Arabidopsis seeds. Plant Cell 18, 665–675.
El-Kouhen, K., Blangy, S., Ortiz, E., Gardies, A.-M., Ferté, N., Arondel, V., 2005. Identification and characterization of a triacylglycerol lipase in Arabidopsis homologous to mammalian acid lipases. FEBS Lett. 579, 6067–6073.
Funatsu, M., Aizono, Y., Hayashi, K., Watanabe, M., Eto, M., 1971. Biochemical studies on rice bran lipase: part I. Purification and physical properties. Agric. Biol. Chem. 35, 734–742.
Goffman, F.D., Pinson, S., Bergman, C., 2003. Genetic diversity for lipid content and fatty acid profile in rice bran. J. Am. Oil Chemists' Soc. 80, 485–490.
Gonnet, G.H., Cohen, M.A., Benner, S.A., 1992. Exhaustive matching of the entire protein sequence database. Science 256, 1443–1445.
Gupta, R., Gupta, N., Rathi, P., 2004. Bacterial lipases: an overview of production, purification and biochemical properties. Appl. Microbiol. Biotechnol. 64, 763–781.
He, Y., Tang, W., Swain, J.D., Green, A.L., Jack, T.P., Gan, S., 2001. Networking senescence-regulating pathways by using Arabidopsis enhancer trap lines. Plant Physiol. 126, 707–716.
Kelly, A.A., Shaw, E., Powers, S.J., Kurup, S., Eastmond, P.J., 2013. Suppression of the SUGAR-DEPENDENT1 triacylglycerol lipase family during seed development enhances oil yield in oilseed rape (Brassica napus L.). Plant Biotechnol. J. 11, 355–361.
Li, W., Ling, H., Zhang, F., Yao, H., Sun, X., Tang, K., 2012. Analysis of Arabidopsis genes encoding putative class III lipases. J. Plant Biochem. Biotechnol. 21, 261–267.
Liu, L., Waters, D.L., Rose, T.J., Bao, J., King, G.J., 2013. Phospholipids in rice: significance in grain quality and health benefits: a review. Food Chem. 139, 1133–1145.
Macauley-Patrick, S., Fazenda, M.L., Mcneil, B., Harvey, L.M., 2005. Heterologous protein production using the Pichia pastoris expression system. Yeast 22, 249–270.
Morohoshi, T., Oikawa, M., Sato, S., Kikuchi, N., Kato, N., Ikeda, T., 2011. Isolation and characterization of novel lipases from a metagenomic library of the microbial community in the pitcher fluid of the carnivorous plant Nepenthes hybrida. J. Biosci. Bioeng. 112, 315–320.
Napier, J.A., Stobart, A.K., Shewry, P.R., 1996. The structure and biogenesis of plant oil bodies: the role of the ER membrane and the oleosin class of proteins. Plant Mol. Biol. 31, 945–956.
Patil, K.J., Chopda, M.Z., Mahajan, R.T., 2011. Lipase biodiversity. Indian J. Sci. Technol. 4, 971–982.
Ramezanzadeh, F.M., Rao, R.M., Windhauser, M., Prinyawiwatkul, W., Tulley, R., Marshall, W.E., 1999. Prevention of hydrolytic rancidity in rice bran during storage. J. Agric. Food Chem. 47, 3050–3052.
Reis, P., Holmberg, K., Watzke, H., Leser, M., Miller, R., 2009. Lipases at interfaces: a review. Adv. Colloid Interface Sci. 147, 237–250.
Rogers, E., Rice, S., Nicolosi, R., Carpenter, D., Mcclelland, C., Romanczyk Jr., L., 1993. Identification and quantitation of γ-oryzanol components and simultaneous assessment of tocols in rice bran oil. J. Am. Oil Chemists' Soc. 70, 301–307.
Schrag, J.D., Cygler, M., 1997. Lipases and alpha/beta hydrolase fold. Methods Enzym. 284, 85.
Seth, S., Chakravorty, D., Dubey, V.K., Patra, S., 2014. An insight into plant lipase research–challenges encountered. Protein Expr. Purif. 95, 13–21.
Da Silva, M.A., Sanches, C., Amante, E.R., 2006. Prevention of hydrolytic rancidity in rice bran. J. Food Eng. 75, 487–491.
Slabas, A.R., Fawcett, T., 1992. The biochemistry and molecular biology of plant lipid biosynthesis. Plant Mol. Biol. 19, 169–191.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Vijayakumar, K., Gowda, L.R., 2013. Rice (Oryza sativa) lipase: molecular cloning, functional expression and substrate specificity. Protein Expr. Purif. 88, 67–79.
Zaplin, E.S., Liu, Q., Li, Z., Butardo, V.M., Blanchard, C.L., Rahman, S., 2013. Production of high oleic rice grains by suppressing the expression of the OsFAD2-1 gene. Funct. Plant Biol. 40, 996–1004.

Chapter 3

Proteomics approach for the identification of active lipases in rice bran using flurogenic lipase activity probes

3.1.1 Introduction

Rice (*Oryza sativa*) is the most popular staple food worldwide; apart from carbohydrates it also contains lipids present in outer layer of rice seeds known as bran. Rice bran oil (RBO) is considered to be a "healthy oil" with about 35% oleic acid content and a high smoke point (Tiwari et al., 2014). Due to the presence of lipases in rice bran, however, the stored oil easily gets hydrolyzed by lipase enzyme which affects both the quality and quantity of extracted (Funatsu et al., 1971, Ramezanzadeh et al., 1999).

Lipases (EC 3.1.1.3) are versatile enzymes capable of catalyzing both hydrolysis and synthesis reactions (Freire et al., 2008) thus giving them wide industrial use (Caro et al., 2000, Pinyaphong and Phutrakul, 2009, Mounguengui et al., 2013). Lipases are very diverse in nature and can be isolated from different biological sources (Seth et al., 2014). Plant lipases have major roles during seed germination by providing energy and nutrition for growth and development of the embryo by breaking down the stored lipids/triacyl glycerols (TAG) of seeds (Barros et al., 2010, Kelly et al., 2011). Till date, a small number of lipases from plants have been reported, among which only two are from rice (Seth et al., 2014). Much of the difficulty has been in purifying the enzymes (Seth et al., 2014). In contrast in gel zymography is an attractive technique for identifying enzymes. It is accurate in determining molecular mass of enzymes because of short time period of staining required preventing diffusion of proteins in the gel. Moreover, zymography is a very straightforward method which could be used to detect proteins from crude protein extracts and even if the protein is in very low concentration and not visible by Comassie staining (Prim et al., 2003).

There are various methods (eg. radiometric, titrimetric) used to measure lipolytic activity (Prim et al., 2003). One method uses a flurogenic substrate, 4-methylumbelliferone (MUF) and its derivatives to measure lipase activity. MUF substrates can also be used for in-gel assays on separated proteins (Prim et al., 2003).

This study has helped to analyze lipase activity directly from crude protein extracts (without any further purification) from different parts of the rice plant by using different MUF derivatives

(MUF- butyrate, MUF-heptonate, MUF-oleate). One of the activity bands was excised and analyzed by liquid chromatography-tandem mass spectrometry (LC- MS/MS) which identified as being encoded by Os01g081770. The encoded protein is an active lipase in rice bran, where most of the lipids accumulate in the form of TAG. Further the expression of Os01g0817700 was confirmed by endpoint PCR and transcriptomics analysis in different rice lines and in different tissues. These results demonstrate that the protein encoded by Os01g081770 represents an unreported lipase in rice bran.

3.1.2 Materials and Methods

3.1.2.1 Plant material

Oryza sativa L. ssp. indica var. MR219, BD192 and Indonesia black rice (IDB) seeds were imbibed with water for a period of 24 h and germinated at 25 ± 2°C on moistened layers of filter paper in sterile petridishes. At each sampling point, the whole seedling, roots and shoots were analyzed . 0-4 DAI and 5-10 DAI were considered as germination and post-germination phases respectively. The germinated seeds were transferred to soil in pots and grown until the mature plants developed. Rice bran from rice variety MR219 was collected from nearby Faiza rice mill, Subang, Malaysia stored at 4°C and used for further studies.

3.1.2.2 Extraction of proteins from rice tissues

A portion of rice tissue samples (leaf, roots, husk, bran, germinating seed, developing seed) was homogenized with liquid nitrogen to powder (300mg) then mixed with 800µL lysis buffer containing 62.5mM TrisHcl (pH 7.4), 10%glycerol, 0.1%SDS, 2mM EDTA, 5% β Mercaptoethanol. The mixture was vortexed vigorously for approximately 5min and then placed on ice for 10 min. The homogenate was centrifuged at 12,000 rpm for 10 min, and the protein concentration of supernatant (stored in -80°C) was assayed by Bradford method (BioRad, Hercules, CA, USA) and subjected to gel electrophoresis.

3.1.2.3 Gel Electrophoresis of Proteins and Overlay Activity Assay

Stock solutions of methylubeliferyl (MUF)-butyrate (0.5M), MUF-heptonate (0.5M) and MUF-oleate (0.5M; Sigma, USA) were made in Dimethyl sulfoxide (DMSO) and stored at −20 ºC. The final concentration of 200 µM was used for further assays. SDS-PAGE was performed in 10% (w/v) gels, essentially as described by Laemmli (1970). After the run, gels were soaked for 30 min in 2.5% Triton X-100® at room temperature, briefly washed in 50mM phosphate buffer, pH 7.0, and covered by a solution of 100µM MUF-butyrate or 200 µM MUF-oleate in the same buffer (Prim et al., 2003). Activity bands become visible in after UV illumination. Detection of lipolytic activity on MUF-butyrate and MUF- heptonate takes less than 1 min, while hydrolysis of MUF-oleate usually requires 15 min incubation at room temperature (Prim et al., 2003). Following zymogram analysis, SDS-PAGE gels were subsequently stained with Coomassie Brilliant Blue R®-250, and protein bands were visualized.

3.1.2.4 Identification of protein by LC/MS analysis

The lipase activity band from rice bran that appeared in the gel, after zymography with MUF-butyrate was excised under protective UV-light and gel slices were digested by trypsin using the method described by Sanders et al (2007). The products were analyzed by LC-MS/MS using the HCT ULTRA ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany) coupled online with nanoflow HPLC (Ultimate 3000, Thermo Scientific, Bremen, Germany). Data from LC-MS/MS run was exported in Mascot generic file format (*.mgf) and searched against an in house curated *Oryza sativa* database of proteins sequences obtained from uniprot using the MASCOT search engine (version 2.4). The following search parameters were used: missed cleavages, 1; peptide mass tolerance, ± 20ppm; peptide fragment tolerance, ± 0.04 Da; peptide charge, 2+, 3+ and 4+; fixed modifications, carbamidomethyl; Variable modification, oxidation (Met).

3.1.2.5 Sequence retrieval and analysis

Lipase gene sequence was retrieved using MSU, Version 7.0 rice database (rice.plantbiology.msu.edu). Motifs and domains of lipase genes were analyzed using Scan Prosite tool found at ExPASy – PROSITE (http://prosite.expasy.org/) and Interproscan 5 (http://www.ebi.ac.uk/interpro/search/sequence-search).

3.1.2.6 Molecular modeling and docking

The template for building 3D models was selected using the template identification tool SWISS MODEL (www.swissmodel.expasy.org). Subsequently, based on the template, 3D structures of lipase proteins were modeled using SWISS-MODEL server, which is a fully automated protein structure homology-modelling server, accessible via the ExPASy web server, or from the program DeepView (Swiss Pdb-Viewer).

Once the 3D structures of lipases were modeled the geometrical aspects of modeled protein structures were evaluated using Qualitative Model Energy Analysis (QMEAN) server (http://swissmodel.expasy.org/qmean/cgi/index.cgi). Also Ramachandran plots for predicted 3D structures were generated using RAMPAGE server (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) for structure validation.

Three-dimensional structures of triglycerides and ester used for protein docking were generated using online SMILES Translator and Structure File Generator server (http://cactus.nci.nih.gov/translate/) Fatty acids to be used as ligands were optimized using CLC drug discovery workbench software version 2.5. Modeled 3D structures of lipases and fatty acid structures were imported into CLC Drug Discovery Workbench software version 2.5 and molecular docking was performed (http://www.clcbio.com/products/clc-drug-discovery-workbench/) with optimized ligands (triglycerides).

3.1.2.7 Preparation of RNA and cDNA

Total RNA from leaf, germinating BD192 seeds and rice bran was extracted according to manufacturer's protocol using RNeasy Plant Mini Kit (Qiagen). RNA concentration and purity factor (A260/A280 ratio) was determined by spectrometry.

Reverse transcription or cDNA synthesis was performed using Reverse Transcription System kit (Promega) according to manufacturer's protocol. RT-PCR was carried out using MyTaq DNA Polymerase system (Bioline). Equal Volume (2µL) of each sample was analyzed in 1% agarose gel.

3.1.2.8 Semi-quantitative RT-PCR

All PCR primers were designed with Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) according to product size (100-300bp) specification. The rice α-tubulin gene (OsTuba1) was used as internal control. PCR was carried out using MyTaq DNA Polymerase system (Bioline). Primer sequence for *OsTuba1* was derived from Zaplin et al. (2013). Other Sequence of primers used is listed in the appendix (Table 2). Endpoint thermal cycling reaction was conducted with amplification program set as follow: 35 cycles of 2 mins at 95°C, 55°C, 72°C for 30s each and a final extension of 72°C for 7 min using Bio Rad MyCycler TM thermal cycler PCR machine. PCR products were verified agarose gel electrophoresis using 0.5 x TBE buffer.

3.1.2.9 Transcriptomics analysis

Purified total RNA was checked for quality using an RNA Nano kit on the 2100 Bioanalyzer instrument (Agilent Technologies) to ensure RNA integrity number was greater than 8 indicating minimal degradation and was normalized to 1µg starting amounts in 50µl. Sequencing libraries were prepared using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina (New England Biolabs Inc.) following manufacturer's instructions. Quantification and size estimation of libraries were performed on a Bioanalyzer 2100 High Sensitivity DNA chip (Agilent, USA). Libraries were finally normalised to 2nM and sequenced on the Miseq System (Illumina Inc.) generating 150bp length single end reads. Statistical analyses were performed using Qiagen CLC Genomics Workbench version 7.0.4. All statistical analysis was done using IBM SPSS Statistics version 20 and CLC Gaussian-based T-test.

3.1.3 Results and Discussion

3.1.3.1 Zymogram analysis

Zymography is simple, sensitive and functional assay for lipolytic activity. It is a convenient assay that can be used in conjunction with polyacrylamide gel electrophoresis. As zymogram analysis requires short assaying time and same gels can be stained with a conventional dye to

determine the molecular mass of the active proteins, it was decided to directly investigate lipase activity among different rice tissues using different MUF-derivative substrates.

The assay is based on the capability of lipase to catalyze the hydrolysis of fatty acid ester bond of the non-fluorescent compound like, 4-MUB to yield one molecule of the highly fluorescent compound, 4-MU, and one molecule of butyric acid. Each molecule of 4-MU produced, therefore, reflects the hydrolysis of one molecule of fatty acid by the lipase. Activity bands become visible in a short time after UV illumination.

MUF-butyrate and MUF- heptonate took short time (<1 min) for activity staining whereas, MUF-Oleate took around 15min for staining and producing fluorescent lipase activity bands under UV light. Fig.1 shows SDS PAGE results and zymograms of different tissues. MUF-butyrate stained gel showed bands corresponding to 32-35 kDa in all the tissues except roots. A similar pattern was also seen when gel stained with MUF-oleate; except in husk and roots, other bands corresponding to 25kDa were also observed. On the other hand, MUF-heptonate staining showed faint bands in leaf and germinating seed tissue, whereas husk and roots showed bands corresponding to size of 55-60 kDa approximately. A common band of 10-15 kDa size was seen in bran tissue, when stained with all three types of MUF-derivative substrates (Fig.1). Results suggest that different types of lipases are present among different rice tissues ranging from 25-60 KDa based on their substrate specificity. Lipase activity was very intense in developing grain tissue consistent with the knowledge that lipases help in providing nutrition for plant growth and development by breaking down the lipids present in the seed (Kelly et al., 2011). Lipase activity was also observed in rice bran, which also degrades or breakdown the oil present in rice bran adversely affecting RBO production.

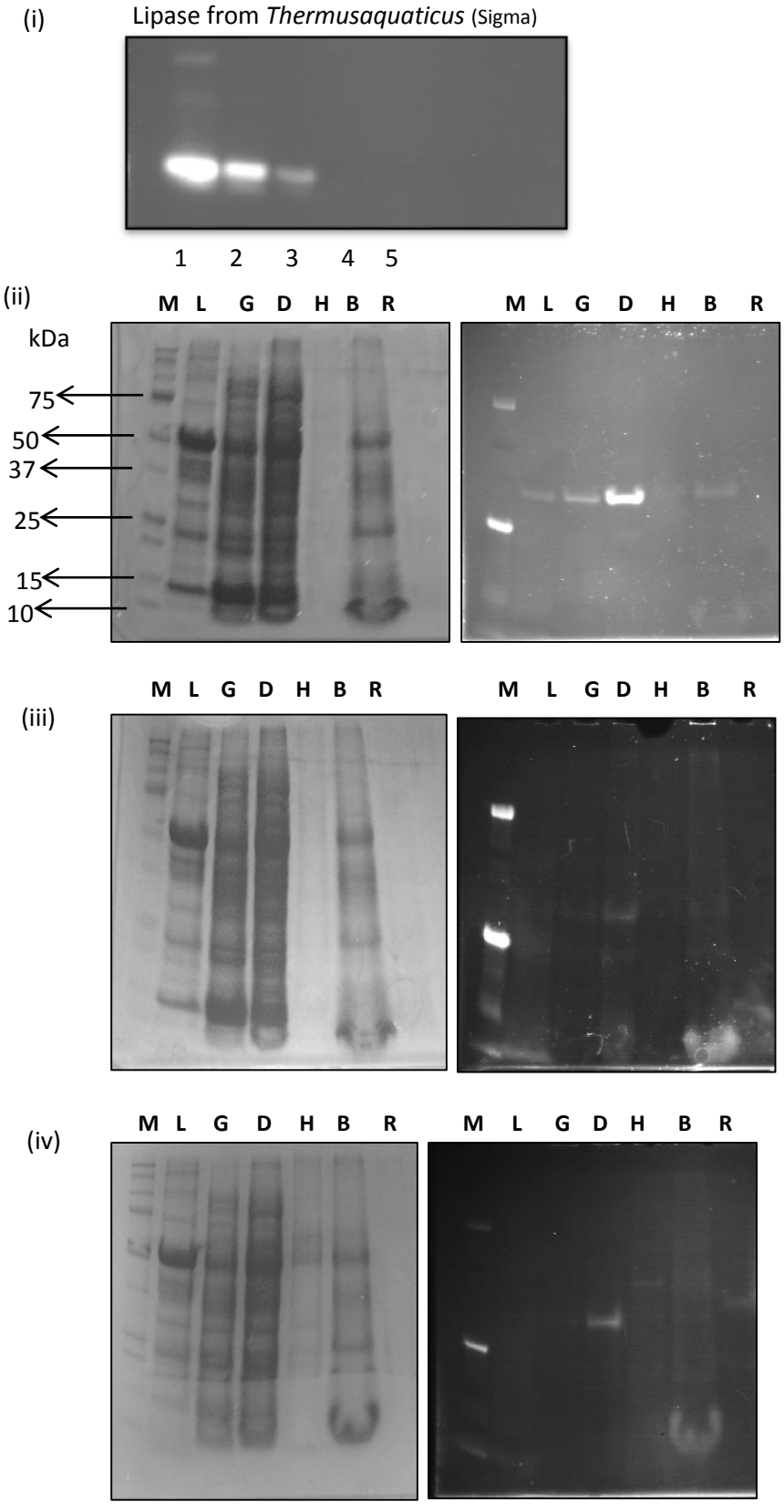


Fig. 1. Zymogram analysis performed on SDS-PAGE gels. The samples loaded corresponds to M- 250Kd protein marker, L- leaf, G – germinating seed, D- developing seed, H – husk, B – Bran, R – Roots. (i) Zymogram from an SDS-polyacrylamide gel of different dilutions of lipase from *T. aquaticus* analyzed with MUF-butyrate. Samples (1-5) correspond to stock of 1mg/100µl,1/3rd dilution, 1/9th dilution, 1/27th dilution , 1/81 dilution. (ii- iv) SDS-PAGE of different tissue fractions from rice plant, analyzed for lipolytic activity using (ii) MUF-butyrate, (iii) MUF- oleate and (iv) MUF-heptonate (right) and subsequently stained with Comassive Brilliant Blue ®-250 (left).

3.1.3.2 Protein identification by mass spectrometry

LC-MS/MS was used to analyze excised lipase activity bands from SDS PAGE gel (Fig.1, (ii)), from rice bran tissue stained with MUF-butyrate. The theoretical digested mass was matched with empirical peptide mass values and database was queried by Mascot search. The protein encoded by Os01g0817700 was the top most hit with a Mascot score of 287 and Exponentially Modified Protein Abundance Index (emPAI) score of 1.20, suggesting the hit is genuine (Appendix, Fig.1). Further, the protein sequence encoded by Os01g0817700 was found to encode two (GXSXG) lipase motifs in regions 175-179 and 268-272 respectively, further indicated it might be a lipase (Fig. 2a). These motifs were present in the middle region which is a typical characteristic of lipases and esterases (Akoh et al., 2004). In the database Os01g0817700 was identified and annotated as putative 2, 3-bisphosphoglycerate-independent phosphoglycerate mutase with MSU locus id LOC_Os01g60190 submitted in MSU database Version 7.0 (http://rice.plantbiology.msu.edu/).

3.1.3.3 Domains and lipase motif analysis of Os01g0817700

As indicated above the protein encoded by Os01g0817700 was found to have two lipase motifs positioned in the middle of the sequence (Fig. 2a). A domain search was also performed using Interproscan 5 which revealed the presence of Alkaline phosphatase like, alpha/beta/alpha domains (Fig. 2b). A gene ontology search for Alkaline phosphatase like, alpha/beta/alpha domains revealed that these domains are known to be involved in membrane lipid metabolic process (Appendix, Fig.2).

3.1.3.4 Expression analysis of Os01g0817700 in different rice lines

Semi-quantitative RT-PCR analysis of Os01g0817700 gene was performed with cDNA as template, synthesized from RNA extracted from leaf, germinating seed and rice bran tissues from different *Indica* rice lines namely MR219, BD192 and IDB respectively. Due to unavailability of rice bran from BD192 and IDB, only MR219 was used for expression analysis in bran. *OsTuba*1 gene from rice was used as internal control.

(b)

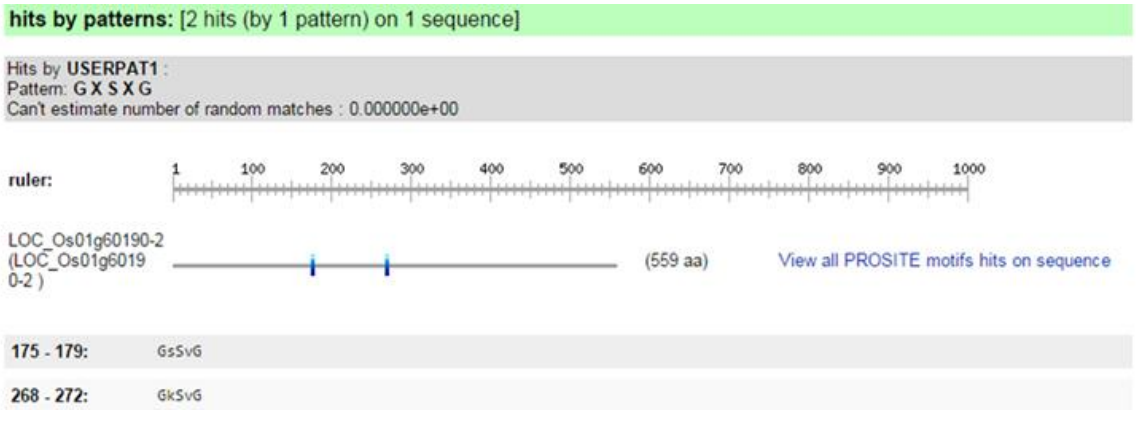


Fig. 2. Domain and motif search. (a) List of domains identified using Interproscan 5 (b) GXSXG lipase motif search using ExPasy ScanProsite tool.

Semi-quantitative RT-PCR analysis showed that Os01g0817700 was expressed in leaf and germinating seed in BD192 and IDB, as well as in bran from MR219 (Fig. 3). In MR219 rice line the expression of Os01g0817700 was similar in all three tissues studied including rice bran. The expression of Os01g0817700 appeared much higher in germinating seed compared to leaf in BD192 and IDB rice lines. This finding was in agreement with transcriptomics data (Fig. 6), supporting the fact that lipases are generally more active during seed germination (Barros et al., 2010); consistent with the suggestion that Os01g0817700 is a new lipase gene.

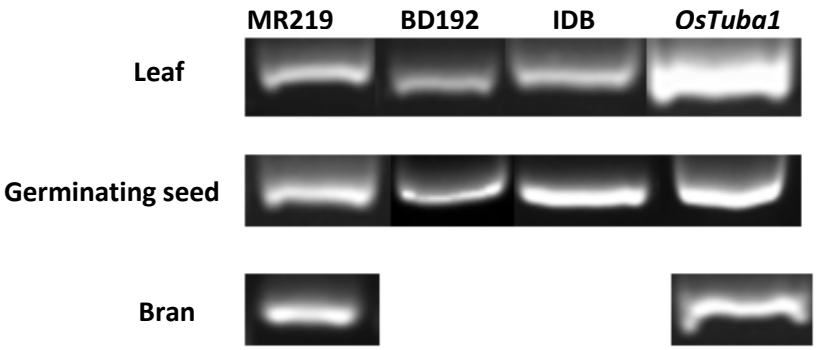


Fig. 3. Semi quantitative expression profiles of identified Os01g0817700 gene in various rice tissues in three *Indica* rice lines namely MR219, BD192 and IDB.

3.1.3.5 *Molecular modeling and protein docking studies of Os01g0817700*

To study the specificity of the putative lipase gene with short and long chain fatty acids in triacylglycerols (TAG), crystal structure of *Leishmania mexicana* phosphoglycerate mutase (PDB id 3IGZ) having 55% identity (protein) with the protein encoded by Os01g0817700 was selected as a template to generate a three dimensional model of the protein encoded by Os01g0817700 (Fig. 4a). Alignment of the modeled structure of Os01g0817700 with its template (PDB ID: 3IGZ) displayed an overall root mean square deviation (RMSD) value of 0.076 (Fig. 4b) suggesting high similarity of the modeled structure with the template. QMEAN Z-score was -1.001 suggesting geometrical aspects of the modeled structure are perfect (Appendix, Fig.3). The modeled structure was further validated using RAMPAGE server which suggested only 1.3% of amino acids residues are in outlier region (Appendix, Fig.4). After validating the quality

of modeled three dimensional structure, it was used for protein docking studies with different triglycerides substrates.

The CLC drug discovery workbench software was used to study the interaction of the modeled structure of the polypeptide encoded by Os01g0817700 with various triglycerides. The structures of the docking study for various ligands with the polypeptide encoded by Os01g0817700 modeled structures are shown in (Fig.5). The polypeptide encoded by Os01g0817700 showed very high affinity/specificity for triglycerides with 10-18 carbon chain fatty acids (Table 1). At 20 DAF stage, oleic and linoleic fatty acids have been reported to be the most abundant fatty acids in rice grain (Choudhury and Juliano, 1980).

(a)

(b)

Fig. 4. (a) The three-dimensional modeled structure of protein encoded by Os01g0817700 (b) Superimposed structure of Os01g0817700 (red) with their template 3IGZ (cyan) with overall root mea square deviation (RMSD) of 0.076.

Our docking results suggest that the Os01g0817700 encoded protein might affect the oleic acid content in rice oil, as it has high affinity for unsaturated fatty (oleic and linoleic) acids (Table 1). Binding affinity/specificity of the Os01g0817700 encoded protein declines with increase in fatty acid chain length (20-22 carbons) (Table 1). Thus Os01g0817700 could be a potential lipase degrading RBO as triglycerides in rice bran oil have a majority of fatty acids in the range of 16-18 carbon length and selectivity of seed lipases is towards the major TAG constituents of the

seed oil (Lin et al., 1986). Hence these results suggest that the protein encoded by Os01g0817700 might be a lipase that is expressed in the developing grain.

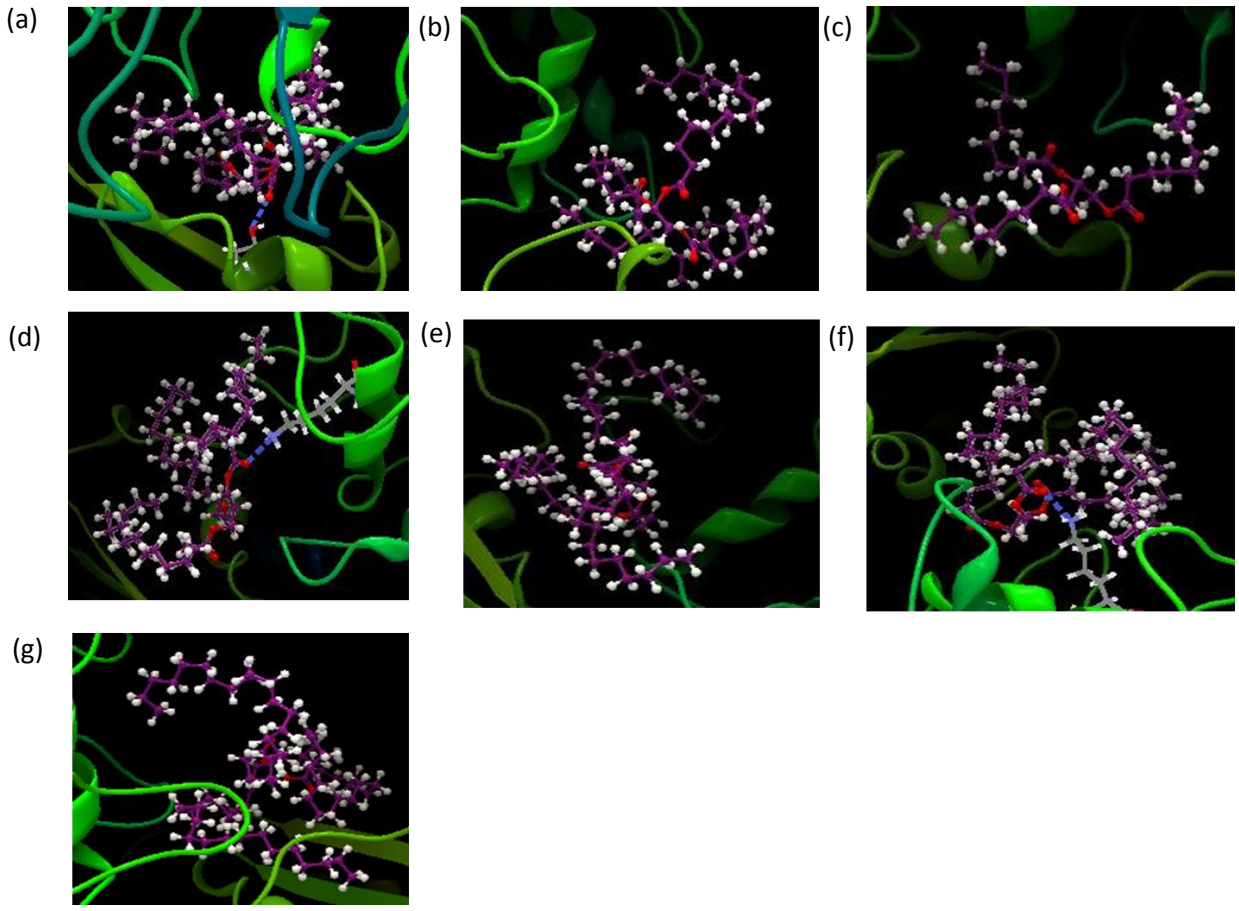


Fig. 5. Protein docking showing the prediction of substrate (triglyceride) binding with proteins encoded by Os01g0817700 (a) Triolein (b) Trilinolein (c) Tricaprin (d) Tripalmitin (e) Tristearin (f) Triarachidin (g) Tribehnin. Ligands are shown in ball and stick model and dotted blue lines indicate the hydrogen bonding.

Table 1. Docking score observed between different docked substrates (triglycerides) with Os01g0817700. Lower the score higher is the affinity for that particular substrate.

S.No.	Substrate/ligand	Score	No. of H-bond	Residues
1	Triolein	-56.74	1	O->THR366
2	Trilinolein	-52.37	0	
3	Triarachidin	-45.83	1	H-> LYS295
4	Tricaprin	-73.46	0	
5	Tripalmitin	-69.53	1	H->LYS295
6	Tristrearin	-64.81	0	
7	Tribehnin	-39.05	0	

3.1.3.6 Transcriptomics analysis of Os01g0817700 in BD192 and IDB rice lines

RNA extracted from embryo of germinating seed and leaf of BD192 and IDB rice lines were subjected to Miseq System (Illumina Inc.) and RNA-seq reads generated were mapped against reference rice genome (Nipponbare, MSU, Version 7.0, October 2011) (http://rice.plantbiology.msu.edu/index.shtml) and the transcript abundance of Os01g0817700 in BD192 and IDB rice lines were compared and analyzed (Fig.6).

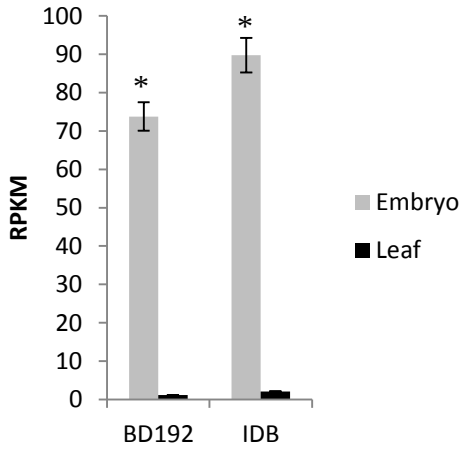


Fig. 6. Bar chart showing differential transcript expression of identified Os01g0817700 gene in embryo and leaf tissue from BD192 and IDB rice lines. Data analyzed using CLC-Bio Genomic Workbench. Baggerley's test was conducted for analyzing genes between the tissues. Symbol '*' indicates significant difference at $0.05 < p$ value ≤ 0.1, n=3, RPKM = reads per kilo base per million.

In both BD192 and IDB rice lines, RPKM value comparison between different tissues demonstrated that Os01g0817700 is expressed significantly higher in embryos than in leaves with p-values of 6.05E-03 in BD192 and 5.27E-04 in IDB. Also, when comparing between the rice lines, expression of Os01g0817700 is higher in IDB compared to BD192 rice line but no significant difference was observed. Detailed expression data can be seen in appendix table 1. The oil content of different coloured rice has been found to vary. Lipid content is found to be highest in purple/black rice (12-13%) as compared to present in brown rice ranging from 2.76-3.84 % on dry weight basis (Frei and Becker, 2005). Detailed mapped contig data of RNA-Seq reads from BD192 and IDB rice lines in leaf and embryo tissues is in appendix Table 3.

3.1.4 Conclusion

A fluorescence based lipase activity detection approach helped to detect several bands with lipase activity present in different protein extracts from rice plant tissues Through LC-MS/MS analysis, one of such lipase activity band from rice bran was identified as being encoded by Os01g0817700. The polypeptide encoded by Os01g0817700 was predicted to have two lipase motifs and found to be active and expressed in leaf, germinating seed and rice bran suggesting it might be a novel lipase. Modeling and protein docking studies predicted that the protein encoded by Os01g0817700 has higher affinity for oleic and linoleic fatty acids which are known to be a major component in RBO. RT-PCR data supports that Os01g0817700 is expressed in leaf, embryo tissue of all three tested rice lines as well as in the bran of MR219. Transcriptomics data support that the Os01g0817700 gene is highly expressed in seed embryo than in leaf. Therefore it is likely that Os01g0817700 encodes an active bran lipase which might reduce oleic and linoleic content resulting in RBO quality reduction. Further studies are required to confirm the biological activity of putative lipase, Os01g0817700 by gene cloning and over expression studies as demonstrated in chapter 2.

3.1.5 References

Akoh, C. C., Lee, G.-C., Liaw, Y.-C., Huang, T.-H. & Shaw, J.-F. 2004. GDSL family of serine esterases/lipases. *Progress in lipid research,* 43**,** 534-552.

Barros, M., Fleuri, L. & Macedo, G. 2010. Seed lipases: sources, applications and properties-a review. *Brazilian Journal of Chemical Engineering,* 27**,** 15-29.

Caro, Y., Villeneuve, P., Pina, M., Reynes, M. & Graille, J. 2000. Lipase activity and fatty acid typoselectivities of plant extracts in hydrolysis and interesterification. Journal of the American Oil Chemists' Society, 77, 349-354.

Choudhury, N. H. & Juliano, B. O. 1980. Lipids in developing and mature rice grain. *Phytochemistry,* 19**,** 1063-1069.

Freire, GDM and Castilho, FL, lipases in Biocatalysis. In Bon et al. (org). Enzymes in biotechnology: Production, Application and Market. Rio de Janeiro, Interscience (2008).

Frei, M. & Becker, K. 2005. Fatty acids and all-trans-β-carotene are correlated in differently colored rice landraces. *Journal of the Science of Food and Agriculture,* 85**,** 2380-2384.

Funatsu, M., Aizono, Y., Hayashi, K., Watanabe, M. & Eto, M. 1971. Biochemical Studies on Rice Bran Lipase: Part I. Purification and Physical Properties. *Agricultural and Biological Chemistry,* 35**,** 734-742.

Kelly, A. A., Quettier, A.-L., Shaw, E. & Eastmond, P. J. 2011. Seed storage oil mobilization is important but not essential for germination or seedling establishment in Arabidopsis. *Plant Physiology,* 157**,** 866-875.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *nature,* 227**,** 680-685.

Lin, Y., Yu, C. & Huang, A. H. 1986. Substrate specificities of lipases from corn and other seeds. *Archives of biochemistry and biophysics,* 244**,** 346-356.

Mounguengui, R. W. M., Brunschwig, C., Baréa, B., Villeneuve, P. & Blin, J. 2013. Are plant lipases a promising alternative to catalyze transesterification for biodiesel production? *Progress in Energy and Combustion Science,* 39**,** 441-456.

Pinyaphong, P. & Phutrakul, S. 2009. Synthesis of cocoa butter equivalent from palm oil by Carica papaya lipase-catalyzed interesterification.

Prim, N., Sánchez, M., Ruiz, C., Pastor, F. J. & Diaz, P. 2003. Use of methylumbeliferyl-derivative substrates for lipase activity characterization. *Journal of Molecular Catalysis B: Enzymatic,* 22**,** 339-346.

Ramezanzadeh, F. M., Rao, R. M., Windhauser, M., Prinyawiwatkul, W., Tulley, R. & Marshall, W. E. 1999. Prevention of hydrolytic rancidity in rice bran during storage. *Journal of agricultural and food chemistry,* 47**,** 3050-3052.

Sanders, W. S., Bridges, S. M., Mccarthy, F. M., Nanduri, B. & Burgess, S. C. 2007. Prediction of peptides observable by mass spectrometry applied at the experimental set level. *BMC bioinformatics,* 8**,** S23.

Seth, S., Chakravorty, D., Dubey, V. K. & Patra, S. 2014. An insight into plant lipase research–challenges encountered. *Protein expression and purification,* 95**,** 13-21.

Tiwari GJ, Aumeeruddy A, Rahman S. Improving the economic value of rice bran oil. Malaysian J Appl Biol. 2014;43:1–8.

Zaplin, E. S., Liu, Q., Li, Z., Butardo, V. M., Blanchard, C. L. & Rahman, S. 2013. Production of high oleic rice grains by suppressing the expression of the OsFAD2-1 gene. Functional Plant Biology, 40, 996-1004.

3.2.1 Introduction

Apart from the MUF-substrates discussed in section 3.1, fluorescence lipase activity probes could also be used for detecting lipases directly from plant tissue. Fluorescence lipase activity probes are more specific in identifying lipases as probes can be designed which can differentiate among different types of lipases (Adam et al., 2002, Birner-Gruenberger et al., 2005). Screening the lipolytic proteome from rice tissues will allow one to identify the lipases which are present in active form. Screening of the lipolytic proteome is a difficult task but if we can device a method for profiling and identification of lipases it will be very helpful for stabilizing and increasing the rice bran oil. Selected novel lipases can be overexpressed and studied. Thus knowledge about active lipases in rice bran will help to take further steps to prevent the breakdown of lipids by the action of lipases present in rice bran.

Fluorescence probes are generally used to study the biological membrane properties. Fluorescence activity probes consist of three main parts (i) a recognition site specific for certain enzyme species, (ii) a reactive site that forms a covalent bond with the substrate/target, and (iii) a fluorescence tag for visualization and/or purification of the covalently bound target (Cravatt and Sorensen, 2000, Adam et al., 2002, Campbell and Szardenings, 2003, Speers and Cravatt, 2004) (Fig. 7).

In the presence of lipase, the reactive group attached with 7-nitrobenz-2-oxa-1, 3-diazole (NBD) probes (mimicking as substrate for lipase) is cleaved and lipase forms a covalent bond with the NBD probe and remains attached to the probe. Thus, this provides a new method for selection and identification of lipase from the proteome pool (Birner-Gruenberger et al., 2005).

Birner-Gruenberger and co-workers (2005) designed several such NDB probes with different reactive groups for studying the lipolytic proteome in rats. The NBD-HE-HP probe with *sn-1* triacylglycerol as a reactive group could recognize wide range of lipolytic enzymes suggesting it to be a general tool for recognizing (Birner-Gruenberger et al., 2005). The advantage of using thesefluorescence activity probes is that they (probes) react only with the active form of an enzyme. Thus use of the probes provides a direct approach for the identification of lipases from

the proteome of any tissue. It was therefore decided to use the probe in rice. To the best of our knowledge, this is the first use of such probes have been used for studying the lipolytic proteome in the plant.

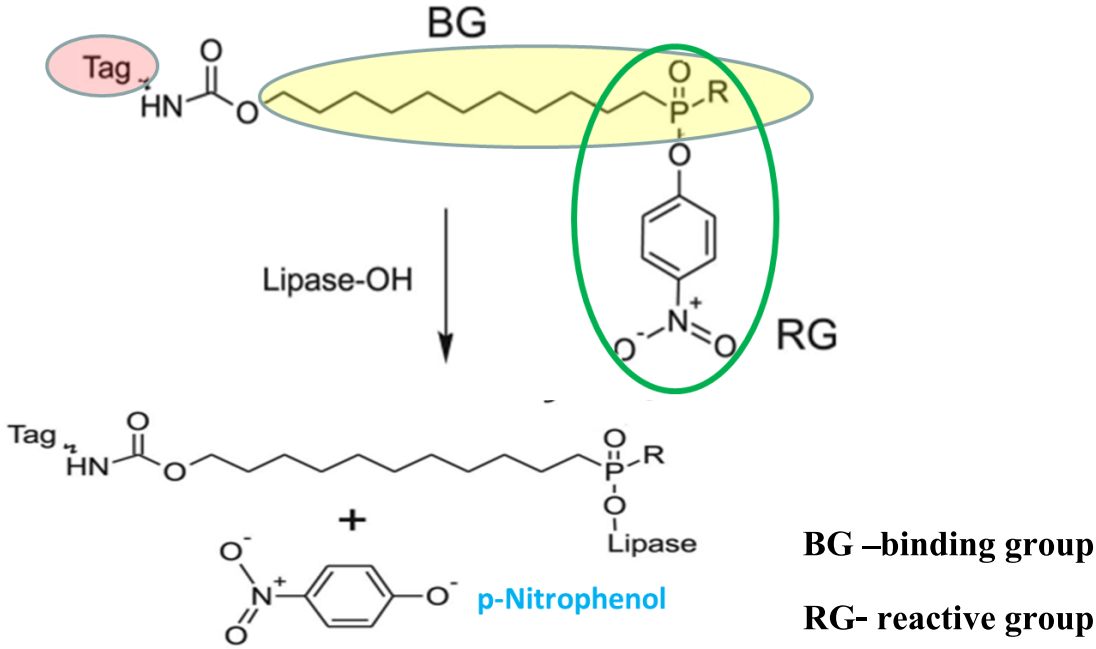


Fig. 7. Schematic representation of NBD-HE-HP activity probe and lipase interaction in an activity based proteomics approach for identifying lipases.

3.2.2 Materials and Methods

3.2.2.1 Plant materials

Rice bran from variety MR219 was collected from nearby Faiza rice mill, Subang, Malaysia, stored at 4°C and used for further studies.

3.2.2.2 Extraction of proteins from rice bran

100 mg of bran tissue was frozen in liquid nitrogen and ground to a fine powder. A total of 800µL lysis buffer (62.5mM Tris HCl (pH 7.4), 10%glycerol, 0.1%SDS, 2mM EDTA, 5% β Mercaptoethanol) was added to the tissue, vortexed vigorously for 5 min and placed on ice for 10

min. The homogenate was centrifuged at 12,000 rpm for 10 min. The supernatant was retained and stored at -80°C freezer prior to Bradford assay (BioRad, Hercules, CA, USA) and gel electrophoresis.

3.2.2.3 Activity tagging

Proteins extracted from rice bran were tagged with NBD-HE-HP tags. 10 µL of a 10 mM solution of Triton X-100 in CHCl3 (final concentration, 1 mM) and 20 µl of activity tag dissolved in CHCl3 (1 nmol/10 µLFµ; final concentration, 20µM) were mixed with 50 µg of protein sample dissolved in lysis buffer. The organic solvent was evaporated in speed vac. (MiVac duo concentrator, Genelac). 100 µL of protein sample (0.5 mg/ml) was added to the mixture and incubated under light protection for 2 hours at 37°C. Tagged protein was precipitated using 2-D clean-up kit (GE healthcare). The resultant pellet was dissolved in 2D sample buffer (7 M urea, 2 M thiourea, 4% CHAPS, 60 mM DTT, 2% Pharmalyte pH 3–10, 0.002% bromphenol blue).

3.2.2.4 2-D Gel Electrophoresis and Visualization

2D gel electrophoresis was performed as described by Gorg et al. (1985). In the first dimension, 50 or 500 µg of protein were isoelectrically focused in 7cm immobilized nonlinear pH 3–10 gradient at 6.5 kV-h (Ettan IPG Phor II, Amersham). In the second dimension, proteins were separated by 10% SDS-PAGE on 7cm gels in the second dimension, respectively. Experiment was performed in triplicate.

Lipase activity spots were visualized scanned at a resolution of 100 µm (Bio-Rad Molecular ImagerTM FX Pro Plus) and scanned at 605 nm and an excitation wavelength of 488 nm.

3.2.3 Results and Discussion

The activity tag probe forms a covalent bond with active site of detected lipases. Hence stable probe-protein complexes were formed after the activity tagging of extracted protein from rice bran. Probe-protein complexes were analyzed on the basis of their fluorescence after two-dimensional electrophoretic separation (Birner-Gruenberger et al., 2005). Scanned gel image Fig.

8shows the overall lipolytic profiling in rice bran after the separation of NBD-HE-HP tagged proteins by 2-D electrophoresis. Most of the fluorescent spots visible may be lipases and few spots might also appear due to auto-fluorescent proteins having fluorescent prosthetic group. Similar fluorescent spots were also observed with NBD-HE-HP tagged proteins from animal tissue.

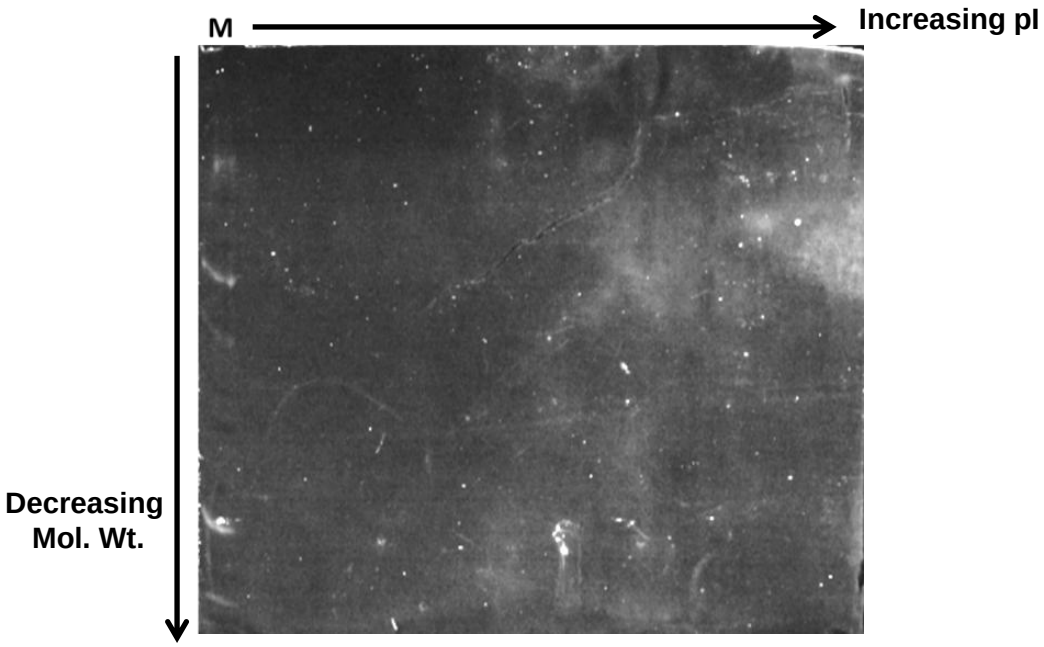


Fig. 8. Lipolytic proteome of rice bran tissue. Protein from rice bran tagged with NBD-HE-HP probe separated and fluorescent lipase activity spots detected in 2D gel electrophoresis. M-Protein marker.

3.2.4 Conclusion

This method of identifying active lipase directly from plant tissue is very straight forward and will help to study the lipolytic proteome from any plant tissue. So far very few active plant lipases have been identified and characterized (Pahoja and Sethar, 2002). Objective of this study has been achieved by screening the lipolytic proteome in rice bran. Spots could be analysed and the genes encoding these active enzymes can be identified. In the future, it is intended that several spots that consistently appeared could be excised from the gel and further analyzed by

liquid chromatography-tandem mass spectrometry (LC-MS/MS) prior to identification of their coding genes.

3.2.5 References

Adam, G. C., Sorensen, E. J. & Cravatt, B. F. 2002. Chemical strategies for functional proteomics. *Molecular & Cellular Proteomics,* 1, 781-790.

Birner-Gruenberger, R., Susani-Etzerodt, H., Waldhuber, M., Riesenhuber, G., Schmidinger, H., Rechberger, G., Kollroser, M., Strauss, J. G., Lass, A. & Zimmermann, R. 2005. The lipolytic proteome of mouse adipose tissue. *Molecular & Cellular Proteomics,* 4, 1710-1717.

Campbell, D. A. & Szardenings, A. K. 2003. Functional profiling of the proteome with affinity labels. *Current opinion in chemical biology,* 7, 296-303.

Cravatt, B. F. & Sorensen, E. J. 2000. Chemical strategies for the global analysis of protein function. *Current opinion in chemical biology,* 4, 663-668.

Görg, A., Postel, W., Günther, S. & Weser, J. 1985. Improved horizontal two-dimensional electrophoresis with hybrid isoelectric focusing in immobilized pH gradients in the first dimension and laying-on transfer to the second dimension. *Electrophoresis,* 6, 599-604.

Pahoja, V. M. & Sethar, M. A. 2002. A review of enzymatic properties of lipase in plants, animals and microorganisms. *Pakistan J. Appl. Sci,* 2, 474-484.

Speers, A. E. & Cravatt, B. F. 2004. Chemical strategies for activity-based proteomics. *ChemBioChem,* 5, 41-47.

Chapter 4

Transcriptomic analysis of *FAD2-1* RNAi high oleic rice lines to identify changes in the expression of genes involved in lipid metabolism

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter 4

Declaration by candidate

In the case of **Chapter 4**, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Experimental design, sample processing, data collection, result acquisition, statistical analysis, manuscript preparation	70

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Qing Liu	Sample preparationexperimental design, result discussion, manuscript preparation	N/A (not a student registered under Monash University)
Pushkar Shreshtha	Participated in lipid analysis experiment, result discussion, manuscript preparation,	N/A (not a student registered under Monash University)
Zhongyi Li	Sample preparation experimental design, result discussion, manuscript preparation	N/A (not a student registered under Monash University)
Sadequr Rahman	Sample preparation experimental design, result discussion, manuscript preparation and submission	N/A (not a student registered under Monash University)

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work*.

Candidate's Signature	[redacted]	**Date** 2·3·2016
Main Supervisor's Signature	[redacted]	**Date** 2.3.2016

*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

CHAPTER SUMMARY

This chapter provides an insight into the regulation of fatty acids composition of stored lipids in the rice grain. *De novo* biosynthesis pathways involved in lipid accumulation are very well known, but target genes responsible for accumulation of mono-unsaturated fatty acids (*MUFA*) and poly-unsaturated fatty acids (*PUFA*) in rice is unclear. Better understanding of genes responsible for fatty acid and lipid biosynthesis will also help in genetic engineering to increase the specific type of fatty acid in accumulated lipids (TAG) in seeds (in this case oleic acid accumulation). This may ultimately help in increasing quality of rice bran oil so that it becomes comparable to olive oil.

A rice line with high oleic acid accumulation due to the targeting of the *OsFAD2* gene is available (Zaplin et al. 2013). In this chapter, Illumina sequencing technology was used to compare and analyze the expression levels of genes involved in the entire lipid biosynthesis pathway. RNA was extracted from wild type and *OsFAD2* RNAi high oleic (HO) rice lines at three developmental stages (10, 15 and 20 DAA) and the expression of genes involved in the lipid biosynthesis pathway was analyzed using CLC genomics workbench software. High – throughput Illumina sequencing provided several million short reads of cDNA from the RNA library prepared from RNA extracted at different time points. The results help to identify key genes whose expression are affected in the high oleic acid line and thus are clearly involved in the accumulation of oleic acid in oil bodies. These results will also help to identify the genes involved in seed development and lipid storage in rice and other cereals.

Paper III discusses in detail the possible roles of genes identified which might be useful to further alter fatty acid composition in rice. This is the first time any transcriptomics analysis has been reported on any transgenic rice line affected in lipid biosynthesis. This study helps to compare and understand the effect of *OsFAD2* knockdown on the entire lipid biosynthesis pathway in rice. Moreover identification of new target genes might also help to further increase oleic acid content (to like that in olive oil) and overall oil content in rice seed. Results presented in paper III will also be helpful for future researchers to develop a rice line with enhanced lipid

accumulation with high percentage of oleic acid content. The results from rice, being a model plant will also help in understanding the transcriptome of lipid pathway in other cereal grains.

Reference:

Zaplin, E. S., Liu, Q., Li, Z., Butardo, V. M., Blanchard, C. L. & Rahman, S. 2013. Production of high oleic rice grains by suppressing the expression of the *OsFAD2-1* gene. *Functional Plant Biology,* 40**,** 996-1004.

Tiwari et al. BMC Plant Biology (2016) 16:189
DOI 10.1186/s12870-016-0881-6

BMC Plant Biology

RESEARCH ARTICLE Open Access

RNAi-mediated down-regulation of the expression of *OsFAD2-1*: effect on lipid accumulation and expression of lipid biosynthetic genes in the rice grain

Gopal Ji Tiwari[1,2], Qing Liu[3], Pushkar Shreshtha[3], Zhongyi Li[3] and Sadequr Rahman[1,2*]

Abstract

Background: The bran from polished rice grains can be used to produce rice bran oil (RBO). High oleic (HO) RBO has been generated previously through RNAi down-regulation of *OsFAD2-1*. HO-RBO has higher oxidative stability and could be directly used in the food industry without hydrogenation, and is hence free of *trans* fatty acids. However, relative to a classic oilseed, lipid metabolism in the rice grain is poorly studied and the genetic alteration in the novel HO genotype remains unexplored.

Results: Here, we have undertaken further analysis of role of *OsFAD2-1* in the developing rice grain. The use of Illumina-based NGS transcriptomics analysis of developing rice grain reveals that knockdown of *Os-FAD2-1* gene expression was accompanied by the down regulation of the expression of a number of key genes in the lipid biosynthesis pathway in the HO rice line. A slightly higher level of oil accumulation was also observed in the HO-RBO.

Conclusion: Prominent among the down regulated genes were those that coded for FatA, LACS, SAD2, SAD5, caleosin and steroleosin. It may be possible to further increase the oleic acid content in rice oil by altering the expression of the lipid biosynthetic genes that are affected in the HO line.

Keywords: Rice bran oil, Triacylglycerol, Oleic acid, FAD2, Transcriptome

Background

Rice is one of the most important crops for mankind as it provides nearly half of the world's population a source of dietary energy [1]. Apart from starch, rice grains contain a small proportion of lipids (1–4 % of the grain) located mostly in the bran. Rice bran oil (RBO) is extracted from rice bran as a by-product of milling and is commercially available as a food grade vegetable oil [2, 3]. Triacylglycerols (TAGs) make up about 85 % of the total lipids in RBO, followed by phospholipids (~6.5 %) and free fatty acids (~4.5 %) [4]. RBO is also rich in compounds such as oryzanol and tocotrienes having antioxidant and cholesterol–reducing activities [5–8]. TAGs inRBO are composed of three main fatty acids: palmitic acid, oleic acid and linoleic acid. The relative content of palmitic (15–20 %), oleic (36–48 %) and linoleic acids (30–38 %) depends on the cultivar and environment [9, 10].

Linoleic acid can undergo non-enzymatic oxidation because of the presence of the two reactive double bonds in the molecule [11, 12] which reduces the shelf-life of RBO and leads to wastage of 60–70 % of RBO [6, 13]. Therefore, partial hydrogenation has often been used to enhance the oxidative stability of RBO, resulting in nutritionally undesirable *trans* fatty acids as a by-product. *Trans* fatty acids have been found to increase the risk of cardiovascular diseases and have been prohibited in foods in an increasing number of countries in the world [14–17]. On the other hand, oleic acid is both oxidatively stable and nutritionally desirable, hence favored for direct food applications without partial hydrogenation.

* Correspondence: sadequr.rahman@monash.edu
[1]School of Science, Monash University Malaysia, 46150 Bandar Sunway, Selangor, Malaysia
[2]Monash University Malaysia Genomics Facility, 46150 Bandar Sunway, Selangor, Malaysia
Full list of author information is available at the end of the article

The microsomal enzyme Δ12 fatty acid desaturase (FAD2) converts oleic acid into linoleic acid while associated with phosphatidylcholine in the endoplasmic reticulum (ER). A total of 18 desaturase genes have been annotated in rice genome, among which are the four *FAD2* genes investigated by Zaplin et al. [18]. These were termed *OsFAD2–1*, *–2*, *–3* and *–4*. Among these four genes, the expression of *OsFAD2–1* was reduced by RNA interference (RNAi) suppression which resulted in an increase in the proportion of oleic acid and a reduction of the proportions of linoleic and palmitic acids in T_3 grains. Our previous results suggested that the *OsFAD2–1*gene was an effective target for raising oleic acid levels at the expense of the oxidatively unstable linoleic acid and the cholesterol-raising palmitic acid [18].

Most reports of genetic modification and characterisation of oil accumulation in plants have so far been carried out in Arabidopsis and classic dicot oilseed crops and focused mainly on trait development [19–24]. We have therefore decided to investigate further the role of the *OsFAD2-1* gene in the rice grain. The comparative analysis of lipid fractions in wild type (WT) and HO-RBO was carried out. We also describe the use of Illumina-based NGS transcriptomic analysis on the same selected HO rice line to study the effect of RNAi down-regulation of *OsFAD2-1* on the grain transcriptome, especially on other genes that are involved in lipid biosynthesis and turnover. Preliminary qPCR experiments confirmed the transcriptomic results for some of the selected genes. In this paper we also show that the down-regulation of *OsFAD2-1*with a seed-specific promoter to produce HO rice line was not associated with compromised oil accumulation in the grain, but rather a modest increase.

Results and discussion

Analysis of lipid composition in rice grains from HO rice line and its null segregant

Total lipids were analysed from the HO rice grains. These grains were from the homozygous transgenic line containing the *OsFAD2-1* RNAi construct that was used for transcriptomics analysis. The total lipids in the HO rice grain were composed of 55.0 % oleic acid, 19.8 % linoleic acid and 16.8 % palmitic acid, whereas the grains from a null segregant (a sister line derived from the same original transformation event that does not contain the *OsFAD2-1 RNAi* construct) comprised 32.3 % oleic acid, 40.7 % linoleic acid and 18.6 % palmitic acid (Table 1). The oleic acid content from HO rice line was significantly higher than that from its null segregant ($p < 0.05$). Similar changes were also observed in TAG and phosphatidylcholine (PC) pools, however, there were somewhat different fatty acid compositional profiles for polar lipids, such as the phosphatidylethanolamine (PE) and phosphatidylcholine (PC) pools. The overall results are in broad agreement with the results from Zaplin et al. [18] from an earlier generation of this material (Additional file 1).

Grains from *OsFAD2-1* RNAi line contained higher levels of total lipids (2.9 % by dry weight) compared to 2.6 % in its null segregant ($p < 0.05$), which was reflected by the significant increases in both TAG and polar lipids.

Transcriptome analysis of rice immature endosperms from HO rice line and its null segregant

RNAseq reads from three developmental stages of endosperm of both the HO rice line and its null segregant were mapped against the reference rice genome (cultivar *Nipponbare*) [25] to generate the mapped contigs as

Table 1 Fatty acid composition of rice grains of *OsFAD2-1 RNAi* line and its null segregant line

	Total lipids		Triacylglycerols		Polar lipid pool		Free fatty acids		PC		PE	
	Control	Fad2	Control	Fad2	Control	Fad2	Control	Fad2	Control	Fad2	Control	Fad2
Cl4:0	0.6 ± 0.0	0.3 ± 0.0	0.5 ± 0.0	0.3 ± 0.0	2.5 ± 0.1	1.2 ± 0.1	1.6 ± 0.0	0.9 ± 0.1	1.1	0.6	1.6	1.2
Cl6:0	18.6 ± 0.2	16.8 ± 0.4	18.4 ± 0.1	16.3 ± 0.3	26.7 ± 1.0	25.2 ± 0.8	18.7 ± 0.6	22.2 ± 0.6	21.0	17.8	25.8	24.1
Cl6:1	0.3 ± 0.0	0.3 ± 0.0	0.3 ± 0.0	0.3 ± 0.0	0.2 ± 0.0	0.3 ± 0.0	0.2 ± 0.0	0.3 ± 0.0	0.3	0.3	0.3	0.3
Cl8.0	2.4 ± 0.0	2.6 ± 0.1	2.4 ± 0.1	2.6 ± 0.1	1.8 ± 0.1	2.1 ± 0.1	3.0 ± 0.1	3.8 ± 0.2	1.5	1.4	2.2	19
08.1	32.3 ± 0.4	55.0 ± 0.7	33.8 ± 0.3	56.2 ± 0.7	24.1 ± 1.0	43.9 ± 14	11.8 ± 0.3	45.4 ± 0.9	38.8	55.4	25.4	41.8
C18:ld11	1.0 ± 0.0	1.1 ± 0.0	1.0 ± 0.0	1.1 ± 0.0	1.1 ± 0.0	1.2 ± 0.0	0.6 ± 00	0.8 ± 0.0	1.4	1.4	1.3	1.3
Cl8.2	40.7 ± 0.4	19.8 ± 0.7	40.2 ± 0.5	19.8 ± 0.6	36.6 ± 0.4	20.0 ± 0.6	58.9 ± 0.6	21.4 ± 1.3	33.2	20.4	40.3	26.2
Cl8:3n3	1.7 ± 0.1	1.5 ± 0.1	1.7 ± 0.1	1.4 ± 0.1	1.5 ± 0.2	1.5 ± 0.1	2.5 ± 0.0	2.4 ± 0.2	1.4	1.2	1.3	1.1
C20:0	0.7 ± 0.0	0.8 ± 0.0	0.7 ± 0.0	0.8 ± 0.0	0.3 ± 0.0	0.3 ± 0.0	0.3 ± 00	0.4 ± 0.0	0.2	0.2	0.2	0.3
C20:1d11	0.4 ± 0.0	0.6 0.0	0.4 ± 0.0	0.6 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.1 ± 0.1	02 ± 0.0	0.2	0.2	0.1	0.2
C22:0	0.4 ± 0.0	0.4 ± 0.0	0.2 ± 0.0	0.2 ± 0.0	1.3 ± 0.1	1.1 ± 0.1	0.5 ± 0.0	0.5 ± 0.0	0.3	0.3	0.4	0.4
C24:0	0.8 ± 0.0	0.8 ± 0.0	0.4 ± 0.0	0.4 ± 0.0	3.7 ± 0.3	3.2 ± 0.3	1.9 ± 0.1	1.8 ± 02	0.7	0.7	1.2	1.3
% oil/wt	2.6 ± 0.1	2.9 ± 0.1	1.8 ± 0.1	2.1 ± 0.1	0.21 ± 0.01	0.23 ± 0.00	0.07 ± 0.00	0.08 ± 0.01	0.06	0.08	0.02	0.02

Control: represents grains from null segregant; Fad2: represents grains from *OsFAD2-1*RNAi line; numbers represent mean ± SE in percentage (%); Mean Values are from three repeat analyses of lipid samples which were extracted separately from three independent grain samples

summarised in Table 2. In total, 1.5–9 million of contigs per sample were assembled which included approximately 80–94 % counted contigs for use in further analysis, and 6–20 % un-counted contigs, defined as the total number of fragments after sequencing which could not be mapped, either as intact or as broken pairs. Among the counted contigs, 75–86 % were unique, and 3–10%were non-specific contigs, defined as the reads which have multiple equally good alignments to the reference and therefore have to be excluded from the RNA-seq analysis.

The genes analysed could be grouped broadly into four categories: genes known to be involved in fatty acid biosynthesis and degradation, genes involved in TAG metabolism, transcriptional factors and other genes found to be affected (Additional file 2 and Additional file 3). A total of 55,801 different gene transcripts were detected in the overall analyses out of which 1,617 (2.9 %) genes at 10 days after anthesis (DAA), 1,175 (2.1 %) genes at 15 DAA and 626 (1.12 %) genes at 20 DAA showed significant differences in expression between the null segregant and the HO rice line.

Table 2 Mapped contig results of RNA-Seq reads from null segregant (NG) and *OsFAD2-1*RNAi rice lines at three grain developmental stages

Contigs	Null segregant			*Os-FAD2*-1 RNAi		
	Sample1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
10 DAA						
Counted contigs	1,474,350	2,451,049	912,841	2,305,750	7,974,195	3,179,294
Unique contigs	1,403,705	2,334,799	858,280	2,090,665	7,406,550	3,050,886
Non-S contigs	70,645	116,250	54,561	215,085	567,645	128,408
Un-C contigs	380,469	691,998	678,922	221,969	1,119,121	590,977
Total contigs	1,854,819	3,143,047	1,591,763	2,527,719	9,093,316	3,770,271
Counted contigs (%)	79.49	77.98	57.35	91.22	87.69	84.33
Unique contigs (%)	75.68	74.28	53.92	82.71	81.45	80.92
Non-S contigs (%)	3.81	3.70	3.43	8.51	6.24	3.41
Un-C contigs (%)	20.51	22.02	42.65	8.78	12.31	15.67
15 DAA						
Counted contigs	1,721,045	4,038,637	6,507,485	1,260,698	4,102,787	5,184,375
Unique contigs	1,580,034	3,759,033	5,944,747	1,211,621	3,877,771	4,889,384
Non-S contigs	141,011	279,604	562,738	49,077	225,016	294,991
Un-C contigs	210,716	347,568	496,403	385,614	1,123,069	436,145
Total contigs	1,931,761	4,386,205	7,003,888	1,646,312	5,225,856	5,620,520
Counted contigs (%)	89.09	92.08	92.91	76.58	78.51	92.24
Unique contigs (%)	81.79	85.70	84.88	73.60	74.20	86.99
Non-S contigs (%)	7.30	6.37	8.03	2.98	4.31	5.25
Un-C contigs (%)	10.91	7.92	7.09	23.42	21.49	7.76
20 DAA						
Counted contigs	2,945,375	1,943,916	1,348,074	3,914,475	791,645	3,627,328
Unique contigs	2,797,599	1,778,024	1,212,290	3,446,816	734,727	3,386,969
Non-S contigs	147,776	165,892	135,784	467,659	56,918	240,359
Un-C contigs	447,772	250,284	168,761	441,185	464,829	492,027
Total contigs	3,393,097	2,194,200	1,516,835	4,355,660	1,256,474	4,119,355
Counted contigs (%)	89.09	88.59	88.87	89.87	63.01	88.06
Unique contigs (%)	86.80	81.03	79.92	79.13	58.48	82.22
Non-S contigs (%)	82.45	7.56	8.95	10.74	4.53	5.83
Un-C contigs (%)	4.36	11.41	11.13	10.13	36.99	11.94

Non-S contigs- Non-specific contigs; *Un C contigs*-Un-counted contigs

Expression of genes involved in fatty acid biosynthesis and degradation

De novo fatty acid biosynthesis occurs primarily in plastids, although it also occurs in the mitochondrion to a much lesser extent [26, 27]. The first addition of a malonyl group to an acetyl group is catalysed by KASIII, while the subsequent acyl chain elongation up to C16 and the final two-carbon extension to form C18 fatty acid while associated with acyl carrier protein (ACP) are catalysed by KASI and KASII, respectively (Additional file 4: Table S1). None of the putative transcripts for *KAS* genes were affected by the RNAi down-regulation of *OsFAD2-1* gene (LOC_Os02g48560) (Additional file 5).

Termination of fatty acid elongation in plastids is catalysed by acyl-ACP thioesterase enzymes (Fat), 25 unigenes of which have been annotated in the Rice Genome project [25]. Among them *FatA* and *FatB* are represented by LOC_Os09g32760 and LOC_Os06g05130, respectively. FatA preferentially catalyses the cleavage of the thioester bond of oleyl-ACP, and is also regarded as one of the key enzymes responsible for oleic acid concentration in oil and FatB has substrate preference forC16 - C18 saturated fatty acids [28]. Expression of *FatA* was found significantly reduced at 15 DAA by -1.62 fold ($p = 0.04$) equivalent to -0.91 log 2 fold (Table 3). This is in contrast to the transcript abundance of *FatB* that was not affected in the RNAi-*OsFAD2-1*line, compared to the null segregant control, in all three developmental stages analysed (Fig. 1a). Significant differences in the expression levels of *FatA* and *FatB* were not observed at 10 and 20 DAA.

The first desaturation step of a saturated fatty acid occurs in the plastids, catalysed by stearoyl-ACP desaturase (SAD). SAD is a soluble plastidial enzyme that introduces the first double bond into stearic acid and to a lesser extent palmitic acid to form oleic acid and palmitoleic acid, respectively. LOC_Os01g69080annotated as *SAD2*gene was highly expressed in rice grains at 10 DAA. In comparison to the null segregant, the expression level of *SAD2* was reduced by −1.6 and −1.35 fold in the HO rice grains at 15 DAA ($p = 0.02$) and 20 DAA ($p = 0.01$) respectively, while no significant difference was observed at 10 DAA (Table 3, Fig. 1a). *SAD5* (LOC_Os04g31070) expression was also found to be down regulated at 15 DAA by -1.88 fold ($p = 2.17E-4$) and −1.12 log2 fold change (Table 3, Fig. 1a). No significant change in expression was found in other unigenes annotated for encoding SAD in the HO line compared to null segregant (Additional file 2).

The nucleotide sequence alignment match between either of *SAD2* or *SAD5* and *OsFAD2-1* is generally low and stretches of 20 nucleotide DNA sequences with significant identity were not found. It is therefore unlikely that the decrease in expression level of *SAD* genes in HO line was due to cross silencing. As SAD is an upstream fatty acid desaturase of FAD2, it is tempting to assume that the reduction in the expression of *OsFAD2-1* leading to the build-up of oleic acid may have a feedback effect that leads to the down regulation of SAD expression which is responsible for oleic acid production.

Oleic acid could be further modified by FAD2 in endoplasmic reticulum (ER) through the eukaryotic pathway or by FAD6 in plastids via the prokaryotic pathway. In the previous study [18], four genes in the rice genome were putatively identified as *FAD2* that are present in the eukaryotic pathway, LOC_Os02g48560 (*OsFAD2-1*), LOC_Os07g23430 (*OsFAD2-2*), LOC_Os07g23410 (*Os FAD2-3*) and LOC_Os07g23390 (*OsFAD2-4*). Transcriptome analysis showed that the expression patterns of all the four *OsFAD2* genes were consistent with the previous data of Zaplin et al. [18] and the analysis of publicly available transcriptome data (Additional file 6: Table S2). The analysis of transcriptome data described in this paper showed that, only *OsFAD2-1* transcripts were found in all three grain developmental stages (10, 15 and 20 DAA) (Table 4). The highest expression level of *OsFAD2-1* was found in the early developmental stage in the null segregant line and it declined as the grains developed. Such a finding is consistent with Wang et al. [29] who found that in sesame most of the genes related to lipid biosynthesis were highly expressed at early stage of seed development, which is at 10 DAA. This may suggest that the biosynthesis of polyunsaturated fatty acids is initiated at a rather early stage of grain development. Such a factor needs to be considered for the choice of promoter that drives the hairpin expression cassette of the *OsFAD2-1* sequence in RNAi construct. The HO rice line was generated by using a storage protein promoter, Bx17, which becomes most active from the mid-stage of endosperm development onwards [18]. It is tempting to assume that further enhancement of oleic acid accumulation above that observed in the current transgenic lines is possible when an alternative grain- or bran- specific promoter that is active from early grain development is employed.

The expression of *OsFAD2-1* in the HO rice lines was significantly down regulated in all the three developmental stages examined, with the most marked reduction by -2.05 fold ($p = 9.15E-6$) and -1.22 log2 fold at 15 DAA (Table 3, Fig. 1a). This is anticipated because *OsFAD2-1*was specifically targeted by RNAi mediated gene silencing. However, the down-regulation of *OsFAD2-1* expression did not result in detectable level of alteration in the already very low expression of *OsFAD2-2, -3, -4* genes at 10, 15 and 20 DAA stages.

Effect on long chain fatty acyl-CoA synthetases (LACS) genes

Long chain fatty acyl-CoA synthetases (LACS) are known to be involved in the breakdown of complex fatty acids. Among a total of five annotated *LACS* unigenes in rice, LOC_Os05g25310 was found to be significantly

Table 3 Differential expression of genes in the metabolism of Fatty acid and TAG biosynthesis

Feature ID	Gene abbreviation	DAA	Weighted proportions fold change	*P*-value	RNAi/WT mean fold change	RNAi/WT mean log2 fold change
LOC_Os09g32760	FATA	10	−1.1	0.54	0.93	−0.11
LOC_Os09g32760	FATA	15	−1.62	*0.04	0.53	−0.91
LOC_Os09g32760	FATA	20	−1.28	0.22	0.80	−0.32
LOC_Os01g69080	SAD2	10	−1.04	0.76	0.98	−0.03
LOC_Os01g69080	SAD2	15	−1.57	*0.02	0.55	−0.85
LOC_Os01g69080	SAD2	20	−1.35	*0.01	0.75	−0.41
LOC_Os04g31070	SAD5	10	−1.31	0.2	0.78	−0.37
LOC_Os04g31070	SAD5	15	−1.88	*2.17E-4	0.46	−1.12
LOC_Os04g31070	SAD5	20	−1.01	0.92	1.01	0.01
LOC_Os05g25310	LACS	10	−1.3	0.13	0.78	−0.36
LOC_Os05g25310	LACS	15	−1.45	*0.04	0.59	−0.76
LOC_Os05g25310	LACS	20	−1.37	0.23	0.74	−0.43
LOC_Os01g70090	ECH1	10	−1.11	0.56	0.91	−0.13
LOC_Os01g70090	ECH1	15	−1.64	[a]0.03	0.53	−0.93
LOC_Os01g70090	ECH1	20	1.06	0.79	1.11	0.15
LOC_Os02g48560	FAD2	10	−2.1	*2.02E-3	0.48	−1.05
LOC_Os02g48560	FAD2	15	−2.05	*9.15E-6	0.43	−1.22
LOC_Os02g48560	FAD2	20	−1.77	*0.04	0.59	−0.75
LOC_Os06g22080	DGAT2	10	−1.31	0.45	0.76	−0.39
LOC_Os06g22080	DGAT2	15	−1.71	*7.73E-3	0.51	−0.98
LOC_Os06g22080	DGAT2	20	−1.16	0.28	0.89	−0.17
LOC_Os02g50174	Caleosin	10	1.52	0.27	1.56	0.64
LOC_Os02g50174	Caleosin	15	−1.33	*0.04	0.65	−0.63
LOC_Os02g50174	Caleosin	20	−1.97	*5.02E-3	0.51	−0.97
LOC_Os03g12230	Caleosin	10	−1.14	0.42	0.88	−0.18
LOC_Os03g12230	Caleosin	15	−1.58	*6.60E-3	0.55	−0.86
LOC_Os03g12230	Caleosin	20	−1.27	0.5	0.81	−0.31
LOC_Os04g32080	STEROLEOSIN	10	−1.36	0.42	0.73	−0.45
LOC_Os04g32080	STEROLEOSIN	15	−1.36	*0.03	0.64	−0.65
LOC_Os06g22080	STEROLEOSIN	20	−1.16	0.28	0.89	−0.17
LOC_Os02g49410	LEC1	10	−1.23	0.42	0.81	−0.30
LOC_Os02g49410	LEC1	15	−1.66	*3.91E-3	0.53	−0.92
LOC_Os02g49410	LEC1	20	−1.44	0.12	0.71	−0.49

*represents significant *p*-values

down regulated by -1.45 fold ($p = 0.04$) and -0.76 log2 fold in the HO line at 15 DAA (Table 3, Fig. 1a) compared to the null segregant. Such reduction of LOC_Os05g25310 was also verified by real time quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) (Fig. 2) indicating the significant reduction of the expression at 15 DAA developmental stage. The significance of such a down-regulation remains unclear. There was no significant change in the expression of LOC_Os05g25310 at 10 and 20 DAA. Expression of LOC_Os05g25310 was the highest at 10 DAA with a gradual decrease as the rice grain development progressed.

Effects on TAG assembly

As the major storage lipid in oilseeds, TAG is utilized to fuel seed germination and early seedling establishment prior to autotrophy by photosynthesis [30, 31]. Given the potential importance of the HO trait in rice bran oil, it is pivotal to understand whether and how the TAG biosynthesis, turnover and catabolism are impacted upon in the HO grains.

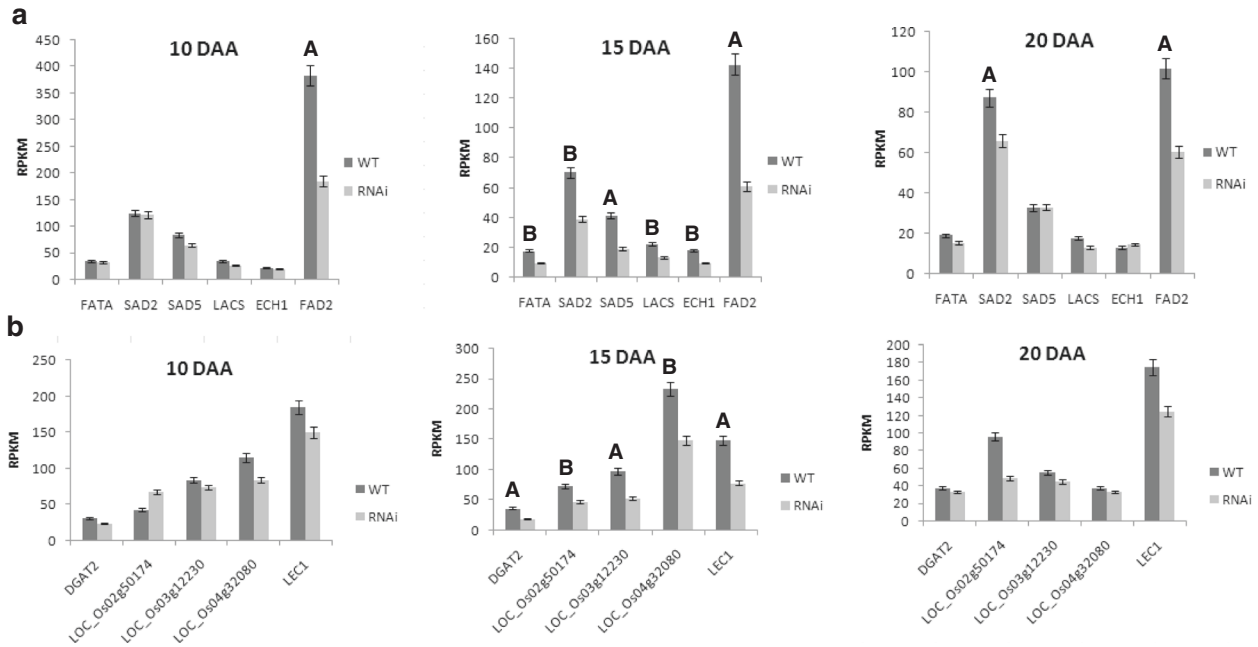


Fig. 1 Differential transcript expression of genes involved in rice (**a**) fatty acid biosynthesis and (**b**) lipid metabolism between the null segregates and *OsFAD2-1* RNAi lines. Developing stages of immature endosperm and gene types are indicated above each figure, the values on the y-axis represent RPKM, and gene locus and their names are labelled underneath. Data analysed using CLC-Bio Genomic Workbench. Baggerley's test was conducted for analysing genes between the null segregates (WT or NG) and *OsFAD2-1* RNAi (RNAi) lines. The letter **a**: indicates significant results at p value ≤ 0.01, the letter **b**: indicates significant results at 0.01 ≥ p value ≤ 0.05. FAT: Acyl-ACP thioesterase A, SAD: stearoyl-ACP desaturases, LACS: Long-chain acyl-CoA synthetase, ECH1: enoyl-CoA hydratase 1, FAD: fatty acid desaturases, DGAT:acyl-CoA:DAG acyltransferase, LEC1: Leafycotyledon1

TAG biosynthesis starts with glycerol-3-phosphate (G3P). Apart from glycolysis, G3P could also be produced by the action of glycerol kinase (GK). There are 14 unigenes encoding for GK as annotated in the rice genome database [25]. None of the *GK* genes was affected in their expression in any of the time points in the HO rice line. Also, there was no effect on the expression of the 18 annotated genes encoding for GPAT required to form lysophosphatidic acid (LPA) at the next stage of TAG assembly.

LPA is acylated by a lysophosphatidic acid acyltransferase (LPAAT) enzyme to form phosphatidic acid (PA). Again the expression of annotated *LPAAT* genes (http://rice.plantbiology.msu.edu/) was not affected in the HO rice. Diacylglycerol (DAG) is generated by removing the phosphate group from PA by phosphatidic acid phosphohydrolase (PAP). *PAP1* (LOC_Os01g63060), *PAP2* (LOC_Os05g21180) and *PAP3* (LOC_Os05g37910) have been annotated in the rice genome database [25]. TAG can be synthesised from DAG in two ways, the acyl-CoA dependent which is normally known as the Kennedy pathway or the acyl-CoA independent pathway. DGAT catalyses the last step of Kennedy pathway by transferring an acyl group from acyl-CoA to DAG to generate *de novo* TAG and has been implicated as the key enzyme in determining the oil content in seed oil [32, 33]. Expression of *DGAT2* (LOC_Os06g22080) was found to increase with the seed development in the null segregant. At 15 DAA, expression of *DGAT2* was significantly down regulated by -1.71 fold ($p = 7.73E\text{-}3$) and -0.98 log2 fold (Table 3) in the HO line. There was no significant difference in the expression level of *DGAT2* gene at other time points between HO and the null segregant line (Table 3, Fig. 1b). DGAT2 has been regarded as a key enzyme in incorporation of unusual fatty acids such as epoxy or hydroxyl fatty acids in TAG to prevent their accumulation in the form of free fatty acids which might cause membrane dysfunction [34, 35]. The other DGAT enzyme, DGAT1, has low expression in the endosperm and no effect was detected.

The acyl-CoA independent reactions are involved in the conversion of two DAGs into a monoacyl glycerol (MAG) and a TAG by DAG:DAG transacylase [36, 37] or the conversion of DAG to TAG by an acyl transfer from the sn-2 position of PC to DAG by Phospholipid:diacylglycerolacyltransferase (PDAT) using PC as acyl donor in TAG formation [34, 38]. In the null segregant, among the 8 annotated PDAT unigenes, the majority of them were found to express at high levels at 10 DAA and decrease in expression in mature grains. Such an expression pattern

Table 4 Expressionlevels of four *FAD2* genes in a null segregant (NG) and an *OsFAD2-1*RNAi line at 10, 15 and 20 DAA developmental stages

DAA	Gene	Rice Genome Annotation Project locus ID	NG 1 (RPKM)	NG2 (RPKM)	NG3 (RPKM)	NG (RPKM mean)	*OsFAD2-1*RNAi 1 (RPKM)	*OsFAD2-1*RNAi 2 (RPKM)	*OsFAD2-1 RNAi 3 (RPKM)*	*OsFAD2-1*RNAi (RPKM mean)
10	*OsFAD2–1*	LOC_Os02g48560	296.6	371.57	482.81	383.66	188.21	210.85	158.01	185.69
15	*OsFAD2–1*	LOC_Os02g48560	134.18	133.69	160.53	142.8	34.8	89.31	59.16	61.09
20	*OsFAD2–1*	LOC_Os02g48560	89.67	89.96	128.16	101.93	54.63	55.79	70.9	60.44
10	*OsFAD2–2*	LOC_Os07g23430	0	0	0	0	0	0	0	0
15	*OsFAD2–2*	LOC_Os07g23430	0	0	0	0	0	0	0	0
20	*OsFAD2–2*	LOC_Os07g23430	0	0	0	0	0.13	0	1.3	0.48
10	*OsFAD2–3*	LOC_Os07g23410	0	0	0.15	0.53	0.15	0.5	0	0.23
15	*OsFAD2–3*	LOC_Os07g23410	0.19	0.06	0	0	0	0	0	0
20	*OsFAD2–3*	LOC_Os07g23410	0.46	0.15	0.31	0	0.31	0	0	0.10
10	*OsFAD2–4*	LOC_Os07g23390	0	0	0	0	0	0	0	0
15	*OsFAD2–4*	LOC_Os07g23390	0	0.22	0	0.07	1.18	0.36	0	0.51
20	*OsFAD2–4*	LOC_Os07g23390	0	0	0	0	0	0.82	1.87	0.90

DAA-days after anthesis

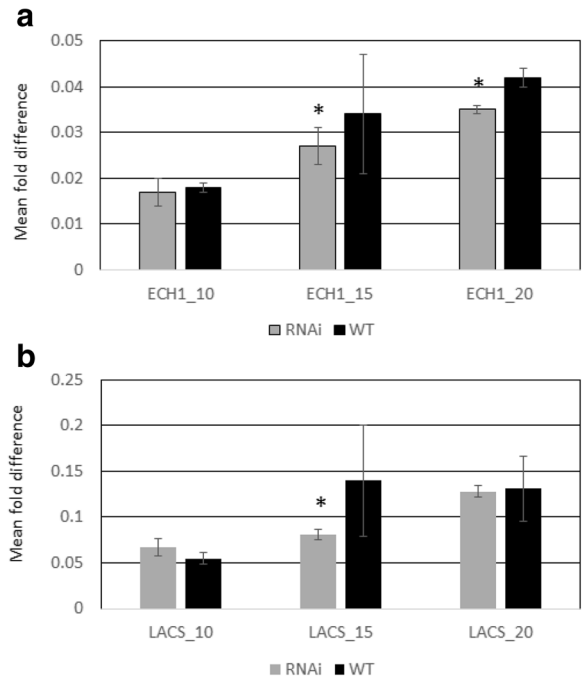


Fig. 2 Level of ECH1 and LACS transcripts of two *OsFAD2-1* RNAi lines (RNAi) and two null segregants (WT) at three different developmental stages. The ΔΔCT method [55] was used to determine the expression of the ECH1 (**a**) and LACS (**b**) gene transcripts normalised to the α-tubulin housekeeping gene from qRT-PCR data to produce a mean fold difference. Error bars are one standard error (s.e). ECH1_10: ECH1 gene at 10 DPA; ECH1_15: ECH1 gene at 15DAA; ECH1_20: ECH1 gene at 20 DAA; LACS_10: LACS gene at 10 DAA; LACS_15: LACS gene at 15 DAA; LACS_20: LACS gene at 20 DAA; OsFAD2–1RNAi Line 22–4 (4) and Line 22–4 (5) were used as transgenic lines, *OsFAD2–1* RNAi Line 22–4 (1) and Line 22–4 (2) were used as the null segregants. All four lines were derived from one *OsFAD2–1* RNAi 22–4 T_2 plant.* shows the significantly different at $P < 0.05$ levels

was not affected in the HO line. The PDAT route is a mechanism for incorporation of unusual fatty acids in *Ricinus communis* by their direct transfer from PC to DAG [39, 40]. As unusual fatty acids have not been reported in rice bran oil, the significance of PDAT in RBO biosynthesis remains unresolved. The consistent expression between WT and HO rice may indicate the PDAT is not a key enzyme determining the oleic acid accumulation in RBO.

Effect on genes involved in TAG packaging and oil body formation

TAG molecules synthesised are packaged and stored in oil bodies (OBs). OBs are maintained and protected by a single layer of PC and proteins which include oleosins, caleosins and steroleosins, with oleosin being the most abundant [41, 42]. Six oleosin genes, 9 caleosin genes and 1 steroleosin gene have been annotated in the rice genome database [25]. Our transcriptomics data showed that in the null segregant each of the three classes of oil body protein genes is expressed in all the three developmental stages examined, and increased as the grain developed. The expression of the oleosins was not found to be significantly affected in the HO line when compared to null segregant rice grain.

Caleosins are calcium- binding OB proteins. The expression of caleosins is reduced during germination to provide access to lipases for breakdown of TAG [53]. Among caleosins, the expression of LOC_Os02g50174 in the HO rice was significantly down regulated at both 15 and 20 DAA by -1.33 (and -0.63 log2 fold) and -1.97 fold ($p = 0.04$, 5.02E-3) (-0.97 log2 fold)respectively; (Table 3, Fig. 1b). Steroleosin has sterol-binding capacity and is mostly involved in signal transduction. The steroleosin unigene annotated as LOC_Os04g32080 was down regulated at 15 DAA by -1.36 fold ($p = 0.03$) and -0.65 log2 fold in the HO rice line (Table 3, Fig. 1b). It remains unclear how the down-regulation of *OsFAD2-1* in rice led to the down-regulation of OB protein gene expression. It is also of particular interest that such a change did not result in the reduction, but rather a modest increase of oil accumulation in HO rice.

Effects on genes involved in fatty acid and lipid catabolism

The key genes coding for the enzymes involved in β-oxidation or fatty acid catabolism were also analysed. In general, all enoyl-CoA hydratase (ECH), 3- hydroxyacyl-CoA dehydrogenase (HACDH), ketoacyl-CoA thiolase (KAT) and acyl-CoA thioesterase (ACT) genes were expressed at high levels at 10 DAA and their expression level gradually decreased as seed development progressed. In the HO line, at 15 DAA stage the expression of ECH1 (LOC_Os01g70090) was significantly reduced by -1.64 fold ($p = 0.03$) and -0.93 log2 fold, compared to the null segregant (Table 3, Fig. 1a). Such reduction of the expression was also supported by qRT-PCR analysis (Fig. 2).

In the HO line, the majority of lipases are found to be expressed at high levels in the early developmental stage at 10 DAA and gradually decreased at later stages. Down-regulation of lipase promotes TAG stabilisation in rice [43]. Among all four phospholipases (PLC1-4), *PLC2* was found to be highly expressed with maximum expression at 10 DAA in null segregant. There was no significant variation on the PLC gene expression between the HO and null segregant.

Expression of transcription factors that may be relevant to lipid accumulation

Apart from the genes that encode functional enzymes or proteins in the lipid biosynthesis or catabolism pathways, several transcription factors such as Leafy cotyledon1 (LEC1), LEC2 and FUSCA3 Like 1 (FL1), Wrinkled 1 (WRI1) and Abscisic acid-insensitive (ABI3) are also known to regulate fatty acid and TAG biosynthesis and

play an important role in lipid accumulation in seed, in addition to their roles in seed development and maturation [44–49]. At 15 DAA, the expression level of the unigene LOC_Os02g49410 annotated as *LEC1* was significantly reduced by -1.66 fold (p = 3.91E-3) and -0.92 log2 fold in the HO line compared to the null segregant (Table 3, Additional file 5).

Impact of *OsFAD2-1* RNAi down regulation on other genes

It was found that the expression of several genes not discussed above was also affected in the HO rice. These are not known to have a direct association with fatty acid and lipid biosynthesis (Additional file 7: Figure S1). For example, the expression of different storage protein genes were differentially regulated at all three stages in the HO rice grains (see Table 5). The expression patterns of additional selected genes being significantly affected in all the time points are also shown in Table 5. This data may facilitate the exploration of other potential molecular networks OsFAD2-1 might be involved, in addition to its key role in linoleic acid biosynthesis.

Conclusion

The transcriptomic analysis of the HO rice grains generated through RNAi down-regulation of *OsFAD2-1* suggests that a suite of key genes involved in fatty acid biosynthesis, TAG assembly and turnover have been differentially regulated in order to incorporate the increased level of oleic acid in TAG that is stored in the form of OBs. Further, the observation of a modest increase in TAG in the HO rice grains may also suggest that the availability of high level of oleic acid is likely favourable for TAG biosynthesis in rice. Overall, this study has delineated a subset of lipid-metabolism genes as being affected when *OsFAD2-1* is down-regulated and the proportion of oleic acid increases in TAG (Fig. 3). The impact on these genes is currently being verified by other techniques. It is envisaged that the genetic manipulation or co-expression of the genes clearly shown to be affected might lead to in further enhancement of the nutritionally desirable oleic acid and TAG accumulation in rice grains.

Methods

Plant materials

High oleic (HO) and null segregant rice (*O. sativa*cv. Nipponbare) seeds were harvested in CSIRO Agriculture, Australia where the HO rice line was previously developed [18]. One *OsFAD2-1*RNAi silencing line, *FAD2RNAi-22(4)* and a null segregant, *FAD2RNAi-22(8)* were used for this study. These were derived from the progeny from one single transformation event, *FAD2RNAi-22,* which had a dramatic reduction of the targeted gene expression and high level of oleic acid content [18]. Rice plants were grown in a containment glasshouse with a constant temperature regime of 27 °C (day and night) under natural light. Fifteen to twenty of immature seeds were collected at 10, 15 and 20 DAA respectively. The endosperms were isolated from the developing grains, frozen in liquid nitrogen and preserved at -80 °C freezer for RNA isolation. T_5 seeds from T_4 plants were analysed, whereas in Zaplin et al. [18], T_4 seeds from T_3 plants were analysed.

Rice grain lipid analysis

Mature brown rice grains were obtained by manual dehulling and ground with a CapMixTM capsule mixing device (3 M ESPE, Seefeld, Germany). Total lipids from ~300 mg above prepared rice flour samples were extracted with a mixture of chloroform/methanol/0.1 M KCl (at a ratio of 2/1/1, by volume). Fatty acid methyl esters (FAME) were prepared by incubating lipid samples in 1 N Methanolic-HCl (Supelco, Bellefonte, PA) at 80 °C for 2 h. TAG and polar membrane lipid pools were fractionated from total lipids in thin layer chromatography (TLC) (Silica gel 60, Merck, Darmstadt, Germany) using a solvent mixture of hexane/diethylether/acetic acid (at a ratio of 70/30/1, by volume) and individual membrane lipid classes were separated by TLC using a solvent mixture of chloroform/methanol/acetic acid/water (90/15/10/3, by volume). Authentic lipid standards were loaded and were run in separate lanes on the same plates for identification of lipid classes. Silica bands, containing individual class of lipid were used to prepare FAME as mentioned above and were analysed by gas chromatography GC-FID 7890A (Agilent Technologies, Palo Alto, CA) that was fitted with a 30 m BPX70 column (SGE, Austin, TX) for quantifying individual fatty acids on the basis of peak area of the known amount of heptadecanoin that was added in as an internal standard [50].

RNA isolation and transcriptomic analysis

Total RNA was isolated from endosperm at 10, 15 and 20 DAA following the method of Higgins et al. [51] with modifications. For each RNA preparation, three endosperms were first ground in liquid nitrogen, then further ground with 600 μL NTES buffer (containing 100 mMNaCl, 10 mMTris, pH8.0, 1 mM EDTA and 1 % SDS), 800 μL phenol/chloroform (Sigma-Aldrich, St. Louis, MO). Samples were transferred into Eppendorf tubes and centrifuged at 13,000x rpm for 5 min in a microcentrifuge. After transferring into new Eppendorf tubes, the supernatant was mixed with an equal volume of 4 M LiCl/10 mM EDTA solution and kept at -20 °C overnight for RNA precipitation. RNA samples were precipitated by centrifugation at 10,000x rpm for 15 min at room temperature (25 °C), rinsed with 70 % ethanol and air dried. RNA pellets were dissolved in 360 μL milliQ H_2O and 40 μL of 2 M NaOAc, pH5.8, which were then precipitated again with 1 mL 95 %

Table 5 Differential expression of non lipid genes between *OsFAD2-1*RNAi lines and their null segregant (NG)

Gene ID	Gene description	10 DAA (RPKM)				15 DAA (RPKM)				20 DAA (RPKM)			
		NG	RNAi	*p*-value	Fold change	NG	RNAi	*p*-value	Fold change	NG	RNAi	*p*-value	Fold change
LOC_Os05g26377	PROLM9 - precursor, expressed	10.42	34.97	3.33E-4	3.355	6.14	13.75	0.00	2.241	13.93	48.11	0.00	3.452
LOC_Os03g07226	Thioredoxin, putative, expressed	176.08	87.9	0.02	−2.00	234.84	134.73	2.16E-07	−1.743	431	74.28	0.00	−1.488
LOC_Os05g26770	PROLM18- precursor, expressed	144	391.26	5.56E-5	2.717	252.1	397.94	1.21E-05	1.578	783.1	2001.73	0.01	2.556
LOC_Os06g31070	PROLM24 precursor, expressed	7999.23	6629.43	0.03	−1.206	13109.13	8339.18	0.01	−1.571	21612.21	13605.77	0.01	−1.588
LOC_Os01g60410	Ubiquitinconjugating enzyme	392.22	271.47	0.02	−1.444	258.38	153.13	1.55E-05	−1.687	182.23	133.27	0.02	−1.367
LOC_Os03g55730	SSA2 - 2S albumin seed storage family protein precursor	7010.17	4731.88	4.97E-4	−1.481	7616.26	4233.57	0.01	−1.799	8507.59	5390.14	0.02	−1.578
LOC_Os05g33570	40S ribosomal protein S9-2	807.34	510.12	0.01	−1.582	402.52	183.09	5.65E-10	−2.198	99.06	61.26	0.04	−1.617

DAA- days after anthesis

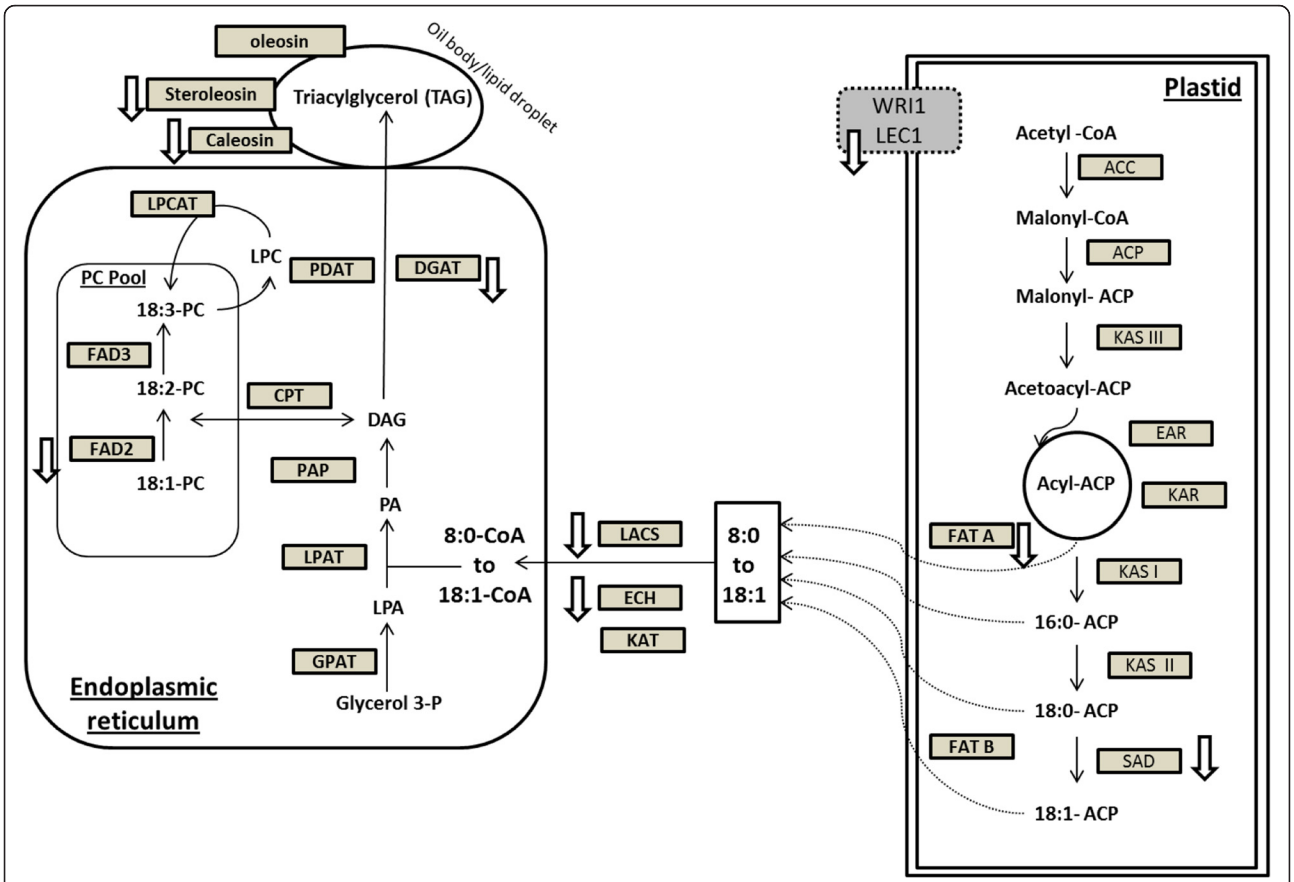


Fig. 3 Differential expression of mRNAs of *OsFAD2-1* RNAi lines on genes involved in fatty acid biosynthesis and lipid metabolism. Downward arrows (↓) indicate down regulated the expression of specific genes in *OsFAD2-1* RNAi lines. ACC: acetyl-CoA carboxylase, ACP: Acyl carrier protein, KASIII: Beta-ketoacyl-ACP synthase III, EAR: Enoyl-ACP reductase, KAR: ketoacyl-ACP reductase, KASI: Beta-ketoacyl-ACP synthase I, KASII: Beta-ketoacyl-ACP synthase II, SAD: Stearoyl-ACP desaturase, FATA: Acyl-ACP thioesterase A, FATB: Acyl-ACP thioesterase B, LACS: Long-chain acyl-CoA synthetase, ECH: Enoyl-CoA hydratase, KAT: Ketoacyl-CoA Thiolase, GPAT: acyl-CoA:G3P acyltransferase, LPAT: acyl-CoA:LPAacyltransferase, PAP: PA phosphatase, CPT: CDP-choline:DAGcholinephosphotransferase, FAD: fatty acid desaturases, LPCAT: acyl-CoA:LPCacyltransferase, PDAT: phospholipid:DAGacyltransferase, DGAT: acyl-CoA:DAGacyltransferase, LPA: lyso-phosphatidic acid, PA: phosphatidic acid, DAG: diacylglycerol, PC: phosphatidylcholine, LEC1: Leafy cotyledon1, WRI1: Wrinkled 1

ethanol and kept at -20 °C for 2 h. Samples were centrifuged, rinsed with 70 % ethanol and air dried as above. After drying, RNAs were dissolved in 20 μL DEPC water, and treated with RQ1 RNase-Free DNase (Promega, Madison, WI) following protocols. The quality of RNA samples were measured with Nanodrop 1000 Spectrophotometer for the ratios of OD 260 nm/280 nm (≥1.8) and OD260 nm/230 nm (≥1.8) and with Aligent Bioanalyser for RNA integrity number (RIN ≥ 6.5) score. RNA was normalised to 1 μg starting amounts in 50 μL. Sequencing libraries were prepared using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina (New England Biolabs Inc., Ipswich, MA) following manufacturer's instructions. Quantification and size estimation of libraries were performed on a Bioanalyser 2100 High Sensitivity DNA chip (Agilent Technologies, Waldbronn, Germany). Libraries were finally normalised to 2nM and sequenced on the Miseq System (Illumina Inc., San Diego, USA) generating 150 bp length single end reads.

Transcriptomic analysis of *OsFAD2* genes from published databases

Six rice RNAseq libraries were down-loaded from Rice Gnome Annotation Project [25] that contains RNAseq databases from different tissues of *Nipponbare* rice. The RNAseq libraries were named SRR352184, 352187, 352190, 342204, 352206 and 352207 and derived from 20 day leaves, post-emergence inflorescence, anthers, 25 DAA embryo, 25 DAA endosperm and 10 DAA grain respectively. The read lengths were 40 or 35 bp and each run produced about 25 million 'clean' reads.

Four rice *FAD2* genes, *OsFAD2-1* to *OsFAD2-4*, were used as reference sequences to conduct gene mapping search "Map to Reference" against the databases in Additional file 6: Table S2 using a bioinformatic analysis program, Geneious [52]. Parameters used were set as custom sensitivity (for sensitivity), and none (fast/read mapping) (for Fine Tuning). Advanced settings were used with 10 % gap, 25 bp

minimum overlap, 24 word length (words repeated more than 8 times were ignored), 2 % maximum mismatches per read, maximum gap 3, minimum overlap identity 80 %, index word length 14 and maximum ambiguity 4.

Statistics analysis

Analysis of variation was performed using Genstat version 16 for lipid content and oleic acid content. All transcriptomics data of HO rice line and its null segregant was analysed using Qiagen CLC Genomics Workbench version 7.0.4. All statistical analysis was done using IBM SPSS Statistics version 20 and CLC Baggerley's test (CLC Bio-Qiagen, Aarhus, Denmark). Details regarding the RNAseq analysis are available online athttp://www.clcbio.com/support/tutorials. For further verifying those differentially expressed genes determined by the method above, the read numbers for each cDNA were first converted to reads per kilo base per million (RPKM), then the ratios of RNAi and WT, and finally log2 value of the ratios.

RNA extraction and quantitative real-time PCR (qRT-PCR)

Total RNA from endosperms at 10, 15 and 20 DAA was extracted using NucleoSpin®RNA Plant Kit (MachereyNagel, Duren, Germany) and quantified using Nanodrop1000 (Thermo Fisher Scientific, Waltham, MA). A total of 0.5 μg of RNA templates was used for the cDNA synthesis in a 50 μL reaction with ramp at 50 °C using Super Script III reverse transcriptase (Thermo Fisher Scientific). The cDNA template (100 ng) was used in a 10 μL qRT-PCR reaction with the annealing temperature at 58 °C. The primers for ECH1 gene were ECH1F(5′ GATGCTGGCGTTGCAAAGAT3′) and ECH1R (5′TCCCTGCTTCTCAGCAAAAC A3′), for LACS gene were LACSF (5′TTGGCGAGGATGCACTGG 3') and LACSR (5'TGGAACTGATTGCAGGTAGCTT 3') which only amplified RT-PCR fragment in cDNAs. The primers for the *Tubulin* gene in rice were used as published [54]. The amplification was conducted in a Rotor-Gene 6000 (Corbett Life Sciences, Sydney, Australia) using Rotor Gene™ SYBR®Green PCR Kit (Qiagen, Hilden, Germany). Comparative quantification was analysed using *Tubulin* as a reference gene in the Real Time Rotary Analyzer Software (Corbett Life Sciences, Sydney, Australia). For each sample, triplicates of qRT-PCR reactions were performed.

Additional files

Additional file 1: Amounts and fatty acid profiles of various lipids in Fad2 silenced rice. (XLSX 31 kb)

Additional file 2: List of unigenes involved in Fatty acid biosynthesis and catabolism. (XLSX 12 kb)

Additional file 3: List of unigenes involved in TAG acid biosynthesis and catabolism. (XLSX 14 kb)

Additional file 4: Table S1. (DOCX 17 kb)

Additional file 5: Gene expression (RPKM) values of affected genes in Lipid biosynthesis pathway. (XLSX 44 kb)

Additional file 6: Table S2. (DOC 30 kb)

Additional file 7: Figure S1. (DOCX 122 kb)

Funding

We thank Monash University, Malaysia for research funding and the Monash University Malaysia Tropical Medicine and Biology Platform for partially funding this study.

Availability of data and materials

All the supporting data are included as additional files in this manuscript.

Authors' contribution

GJT, QL, ZL, SR contributed to the design of the research. GJT carried out the RNA analysis with the help of the Monash University Genomics Facility with RNA provided by ZL. PS carried out the lipid analysis, GJT, QL, PS, ZL, SR contributed to the analysis of the results and writing the manuscript. All authors have read and approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

No ethics approval was needed for this study.

Author details

[1]School of Science, Monash University Malaysia, 46150 Bandar Sunway, Selangor, Malaysia. [2]Monash University Malaysia Genomics Facility, 46150 Bandar Sunway, Selangor, Malaysia. [3]CSIRO Agriculture & Food, PO Box 1600, Canberra, ACT 2601, Australia.

Received: 9 July 2016 Accepted: 24 August 2016
Published online: 31 August 2016

References

1. Maclean JL. Rice almanac: Source book for the most important economic activity on earth. Int Rice Res Inst. 2002.
2. Juliano B. Lipids in rice and rice processing. Lipids Cereal Technol. 1983;440. http://www.sciencedirect.com/science/article/pii/B9780120790203500210.
3. Sugano M, Tsuji E. Rice bran oil and cholesterol metabolism. J Nutr. 1997;127(3):521S–4S.
4. Yoshida H, Tanigawa T, Kuriyama I, Yoshida N, Tomiyama Y, Mizushina Y. Variation in fatty acid distribution of different acyl lipids in rice (Oryza sativa L.) brans. Nutrients. 2011;3(4):505–14.
5. Carroll L. Functional properties and applications of stabilized rice bran in bakery products. Food Technol. 1990;44(4):74–6.
6. Orthoefer FT. Rice bran oil: healthy lipid source. Food Technol. 1996; 50(12):62–4.
7. McCASKILL DR, Zhang F. Use of rice bran oil in foods: Developing nutraceuticals for the new millenium. Food Technol. 1999;53(2):50–3.
8. Lai P, Li KY, Lu S, Chen HH. Phytochemicals and antioxidant properties of solvent extracts from Japonica rice bran. Food Chem. 2009;117(3):538–44.
9. Taira H, Nakagahra M, Nagamine T. Fatty acid composition of Indica, Sinica, Javanica, Japonica groups of nonglutinous brown rice. J Agric Food Chem. 1988;36(1):45–7.
10. Radcliffe JD, Imrhan VL, Hsueh AM. The use of soy protein isolate to reduce the severity of 13-cis retinoic acid-induced hypertriglyceridemia. Cancer Detect Prev. 1997;22(6):526–32.
11. Spiteller G. Linoleic acid peroxidation—the dominant lipid peroxidation process in low density lipoprotein—and its relationship to chronic diseases. Chem Phys Lipids. 1998;95(2):105–62.
12. Chaiyasit W, Elias RJ, McClements DJ, Decker EA. Role of physical structures in bulk oils on lipid oxidation. Crit Rev Food Sci Nutr. 2007;47(3):299–317.
13. Zhou Z, Robards K, Helliwell S, Blanchard C. Composition and functional properties of rice. Int J Food Sci Technol. 2002;37(8):849–68.

14. Mauger J-F, Lichtenstein AH, Ausman LM, Jalbert SM, Jauhiainen M, Ehnholm C, et al. Effect of different forms of dietary hydrogenated fats on LDL particle size. Am J Clin Nutr. 2003;78(3):370–5.
15. Esmaillzadeh A, Azadbakht L. Consumption of hydrogenated versus nonhydrogenated vegetable oils and risk of insulin resistance and the metabolic syndrome among Iranian adult women. Diabetes Care. 2008; 31(2):223–6.
16. Benatar JR, Gladding P, White HD, Zeng I, Stewart RA. Trans-fatty acids in New Zealand patients with coronary artery disease. Eur J Cardiovasc Prev Rehabil. 2011;18(4):615–20.
17. U.S. Food and Drug Administration. [http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm372915.htm].
18. Zaplin ES, Liu Q, Li Z, Butardo VM, Blanchard CL, Rahman S. Production of high oleic rice grains by suppressing the expression of the OsFAD2-1 gene. Funct Plant Biol. 2013;40(10):996–1004.
19. Weselake RJ, Shah S, Tang M, Quant PA, Snyder CL, Furukawa-Stoffer TL, et al. Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (Brassica napus) to increase seed oil content. J Exp Bot. 2008;59(13):3543–9.
20. Xu J, Francis T, Mietkiewska E, Giblin EM, Barton DL, Zhang Y, et al. Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (DGAT1) gene from Tropaeolum majus, and a study of the functional motifs of the DGAT protein using site-directed mutagenesis to modify enzyme activity and oil content. Plant Biotechnol J. 2008;6(8):799–818.
21. Jako C, Kumar A, Wei Y, Zou J, Barton DL, Giblin EM, et al. Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. Plant Physiol. 2001;126(2):861–74.
22. Clemente TE, Cahoon EB. Soybean oil: genetic approaches for modification of functionality and total content. Plant Physiol. 2009;151(3):1030–40.
23. Taylor DC, Zhang Y, Kumar A, Francis T, Giblin EM, Barton DL, et al. Molecular modification of triacylglycerol accumulation by over-expression of DGAT1 to produce canola with increased seed oil content under field conditions This paper is one of a selection of papers published in a Special Issue from the National Research Council of Canada-Plant Biotechnology Institute. Botany. 2009;87(6):533–43.
24. Van Erp H, Bates PD, Burgal J, Shockey J. Castor phospholipid: diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol. 2011;155(2):683–93.
25. MSU Rice Genome Annotation Project. [http://rice.plantbiology.msu.edu/index.shtml].
26. Hiltunen JK, Schonauer MS, Autio KJ, Mittelmeier TM, Kastaniotis AJ, Dieckmann CL. Mitochondrial fatty acid synthesis type II: more than just fatty acids. J Biol Chem. 2009;284(14):9011–5.
27. Hiltunen JK, Autio KJ, Schonauer MS, Kursu VS, Dieckmann CL, Kastaniotis AJ. Mitochondrial fatty acid synthesis and respiration. Biochim Biophys Acta Bioenergetics. 2010;1797(6):1195–202.
28. Salas JJ, Ohlrogge JB. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. Arch Biochem Biophys. 2002;403(1):25–34.
29. Wang L, Yu S, Tong C, Zhao Y, Liu Y, Song C, et al. Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis. Genome Biol. 2014;15(2):1.
30. Theodoulou FL, Eastmond PJ. Seed storage oil catabolism: a story of give and take. Curr Opin Plant Biol. 2012;15(3):322–8.
31. Vanhercke T, El Tahchy A, Liu Q, Zhou XR, Shrestha P, Divi UK, et al. Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. Plant Biotechnol J. 2014;12(2):231–9.
32. Lung S-C, Weselake RJ. Diacylglycerol acyltransferase: a key mediator of plant triacylglycerol synthesis. Lipids. 2006;41(12):1073–88.
33. Bates PD. The significance of different diacylgycerol synthesis pathways on plant oil composition and bioengineering. Frontiers Plant Sci. 2012;3:147.
34. Beaudoin F, Napier JA. 8 Biosynthesis and compartmentation of triacylglycerol in higher plants. Lipid Metabolism and Membrane Biogenesis: Springer; 2004. p. 267–87.
35. Coleman RA, Lee DP. Enzymes of triacylglycerol synthesis and their regulation. Prog Lipid Res. 2004;43(2):134–76.
36. Stobart K, Mancha M, Lenman M, Dahlqvist A, Stymne S. Triacylglycerols are synthesised and utilized by transacylation reactions in microsomal preparations of developing safflower (Carthamus tinctorius L.) seeds. Planta. 1997;203(1):58–66.
37. Kroon JT, Wei W, Simon WJ, Slabas AR. Identification and functional expression of a type 2 acyl-CoA: diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. Phytochemistry. 2006;67(23):2541–9.
38. Abbadi A, Domergue F, Bauer J, Napier JA, Welti R, Zähringer U, et al. Biosynthesis of very-long-chain polyunsaturated fatty acids in transgenic oilseeds: constraints on their accumulation. Plant Cell. 2004;16(10):2734–48.
39. Cahoon EB, Shockey JM, Dietrich CR, Gidda SK, Mullen RT, Dyer JM. Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. Curr Opin Plant Biol. 2007;10(3):236–44.
40. Cagliari A, Margis-Pinheiro M, Loss G, Mastroberti AA, de Araujo Mariath JE, Margis R. Identification and expression analysis of castor bean (Ricinus communis) genes encoding enzymes from the triacylglycerol biosynthesis pathway. Plant Sci. 2010;179(5):499–509.
41. Shimada TL, Shimada T, Takahashi H, Fukao Y, Hara-Nishimura I. A novel role for oleosins in freezing tolerance of oilseeds in Arabidopsis thaliana. Plant J. 2008;55(5):798–809.
42. Siloto RM, Findlay K, Lopez-Villalobos A, Yeung EC, Nykiforuk CL, Moloney MM. The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis. Plant Cell. 2006;18(8):1961–74. http://www.sciencedirect.com.ezproxy.lib.monash.edu.au/science/article/pii/S0168945205003602.
43. Chepyshko H, Lai C-P, Huang L-M, Liu J-H, Shaw J-F. Multifunctionality and diversity of GDSL esterase/lipase gene family in rice (Oryza sativa L. japonica) genome: new insights from bioinformatics analysis. BMC Genomics. 2012;13(1):1.
44. Parcy F, Valon C, Raynal M, Gaubier-Comella P, Delseny M, Giraudat J. Regulation of gene expression programs during Arabidopsis seed development: roles of the ABI3 locus and of endogenous abscisic acid. Plant Cell. 1994;6(11):1567–82.
45. Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, et al. LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. Proc Natl Acad Sci. 2001;98(20):11806–11.
46. Jakoby M, Weisshaar B, Dröge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, et al. bZIP transcription factors in Arabidopsis. Trends Plant Sci. 2002;7(3):106–11.
47. Cernac A, Benning C. WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. Plant J. 2004;40(4):575–85.
48. Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL, et al. Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. Proc Natl Acad Sci U S A. 2006;103(9):3468–73.
49. Dabbs P. Identification of Plant Transcription Factors that Play a Role in Triacylglycerol Biosynthesis. 2015.
50. Vanhercke T, Wood CC, Stymne S, Singh SP, Green AG. Metabolic engineering of plant oils and waxes for use as industrial feedstocks. Plant Biotechnol J. 2013;11(2):197–210.
51. Higgins T, Goodwin PB, Whitfeld PR. Occurrence of short particles in beans infected with the cowpea strain of TMV: II. Evidence that short particles contain the cistron for coat-protein. Virology. 1976;71(2):486–97.
52. Geneious. [http://www.geneious.com].
53. Chen JC, Tsai CC, Tzen JT. Cloning and secondary structure analysis of caleosin, a unique calcium-binding protein in oil bodies of plant seeds. Plant Cell Physiol. 1999;40(10):1079–86.
54. Toyota K, Tamura M, Ohdan T, Nakamura Y. Expression profiling of starch metabolism-related plastidic translocator genes in rice. Planta. 2006;223(2):248–57.
55. Livak KJ, Schmittgen TD, Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2– ΔΔCT method. Methods. 402-8;25(4).

Chapter 5

Conclusions and Future Directions

Conclusions and Future Directions

This dissertation reports new genetic targets to improve rice bran oil quality by utilizing genomics, proteomics and transcriptomics approaches. Firstly, the genomic approach enabled the investigation of all the annotated 125 lipase genes in the rice genome (chapter 2). Putative homologs of known *A. thaliana* triacylglycerol lipase (TAGL) genes in rice were selected for further study. Two lipase genes LOC_Os01g71010 and LOC_Os11g43510 were found to be highly expressed in the rice bran. The molecular characterization of a putative lipase encoded by LOC_Os11g43510 was performed by cloning and expressing it in the methylotrophic yeast *P. pastoris* and its lipase activity was demonstrated. The route followed in this study could be an alternative approach to partially purifying lipases from specific plant tissues and identifying them. In the future, the expression of the identified lipase gene LOC_Os11g43510 could be down regulated through miRNA based gene silencing (Schwab et al., 2006, Ossowski et al., 2008, Huntzinger and Izaurralde, 2011) or rapidly evolving CRISPR technologies (Feng et al., 2013, Belhaj et al., 2013, Hsu et al., 2014) which might help to stabilize the lipids in the rice bran. Also, the investigation of other lipase genes found to be expressed in rice bran should be undertaken.

A proteomics approach was also used as a different way to identify active lipases directly from the protein samples extracted from rice bran (Chapter 3). Two different types of flurogenic lipase activity probes 4-methylumbelliferone (MUF) derivatives and N-(7-Nitrobenz-2-oxa-1, 3-diazol-4-yl) amine (NBD-HE-HP) (Susani-Etzerodt et al., 2006, Schmidinger et al., 2006) were used. These have different reaction mechanisms for identifying active lipases from the extracted proteome from rice bran. NBD-HE-HP is highly specific for lipase and forms irreversible and strong covalent bonds with lipase and such lipases can then be identified by fluorescence. In contrast, fatty acids present in MUF derivatives get cleaved in the presence of lipases and this results in a fluorescent signal in the presence of ultraviolet light. Using MUF-butyrate and LC/MS/MS analysis, one such lipase activity band from the rice bran proteome was identified as likely to be encoded by the gene Os01g0817700. Os01g0817700 was annotated as putative 2, 3-bisphosphoglycerate-independent phosphoglycerate mutase therefore could not be searched among the lipase genes searched in Chapter 2. The putative product would contain two lipase

motifs (GXSXG) and have an estimated molecular mass of ~35 kDa, in correspondence to expected molecular mass. Modeling and protein docking studies predicted the product of Os01g0817700 has affinity for oleic and linoleic fatty acids which are known to be major component in rice bran oil (RBO). Further studies are required to confirm the biological activity of identified putative lipase encoded by Os01g0817700 by gene cloning and over expression studies. Lipase activity bands that were detected by reacting with other MUF substrates (MUF-heptonate and MUF-oleate) could also be studied and investigated further by LC-MS/MS and cloning studies.

Activity tagging of rice bran extracted protein with the NBD-HE-HP lipase activity probe enabled the observation of the entire lipolytic proteome from rice bran (Chapter 3). Several lipase activity spots were observed after scanning the two-electrophoresis protein gel loaded with NBD-HE-HP tagged protein. This is the first time that the NBD-HE-HP lipase activity probe has been used on plant tissues to study the lipolytic proteome. This method provides a straight forward and quick approach for identifying tissue specific lipases from plants. To identify the detected lipases, spots need to be excised from the gel and analyzed by LC-MS/MS or protein fractionation techniques.

Better understanding of the lipid pathway and identification key genes affecting the lipid composition in rice oil would be beneficial to further increase the quality and quantity of RBO. The availability of a high oleic rice line (Zaplin et al., 2013) provided an excellent opportunity for this. Therefore, an Illumina-based NGS transcriptomic approach was performed to understand the expression of genes involved in lipid biosynthetic pathway and degradation. The high oleic rice line developed by Zaplin et al., (2013) was compared with negative segregants (wild-type) rice line at three different seed developmental stages, to study the changes in the expression of genes involved in lipid pathway (Chapter 4). Changes in the expression of genes suggest that these genes are directly or indirectly involved in oleic acid accumulation. This identification can be difficult as many steps in the proposed pathway include enzymes that could be encoded by any of a number of genes. Genes that were identified as being affected in expression coded for FatA (LOC_Os09g32760), SAD2 (LOC_Os01g69080) and LACS (LOC_Os05g25310) and these could be the key genes which regulate oleic acid accumulation in

RBO. Suitably altering the expression of the identified genes may help to further increase the oleic acid content comparable to that of olive oil. In the high oleic (HO) rice line, the majority of lipases were found to be expressed at high levels in the early developmental stage at 10 DAA and gradually decreased at later stages.

Overall, this study assists in understanding the regulation of lipases in rice and should help to devise an efficient method to overcome the problem of rice bran oil instability and degradation. It also should help in increasing the value of the oil produced. It is hoped that this research can be ultimately translated to provide better economic outcomes for rice farmers in the tropics.

References

Belhaj, K., Chaparro-Garcia, A., Kamoun, S. & Nekrasov, V. 2013. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant methods,* 9**,** 1.

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.-L., Wei, P., Cao, F., Zhu, S., Zhang, F. & Mao, Y. 2013. Efficient genome editing in plants using a CRISPR/Cas system. *Cell research,* 23**,** 1229.

Hsu, P. D., Lander, E. S. & Zhang, F. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell,* 157**,** 1262-1278.

Huntzinger, E. & Izaurralde, E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nature Reviews Genetics,* 12**,** 99-110.

Ossowski, S., Schwab, R. & Weigel, D. 2008. Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant Journal,* 53**,** 674-690.

Schmidinger, H., Susani-Etzerodt, H., Birner-Gruenberger, R. & Hermetter, A. 2006. Inhibitor and Protein Microarrays for Activity-Based Recognition of Lipolytic Enzymes. *ChemBioChem,* 7**,** 527-534.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N. & Weigel, D. 2006. Highly specific gene silencing by artificial microRNAs in Arabidopsis. *The Plant Cell,* 18**,** 1121-1133.

Susani-Etzerodt, H., Schmidinger, H., Riesenhuber, G., Birner-Gruenberger, R. & Hermetter, A. 2006. A versatile library of activity-based probes for fluorescence detection and/or affinity isolation of lipolytic enzymes. *Chemistry and physics of lipids,* 144**,** 60-68.

Zaplin, E. S., Liu, Q., Li, Z., Butardo, V. M., Blanchard, C. L. & Rahman, S. 2013. Production of high oleic rice grains by suppressing the expression of the OsFAD2-1 gene. *Functional Plant Biology,* 40**,** 996-1004.

6. Appendix

Chapter 2

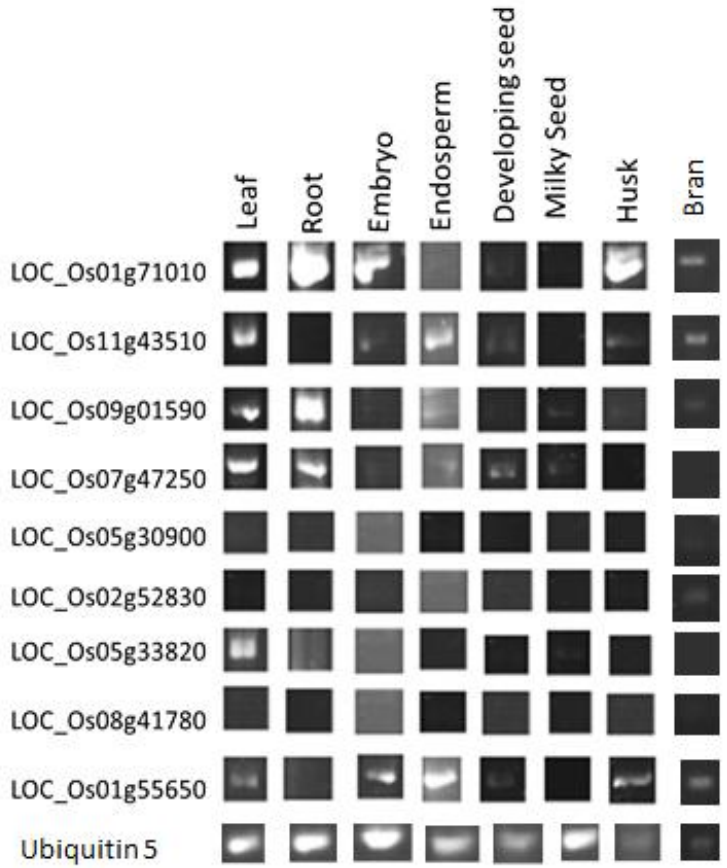


Fig.A.1 Expression profile of *O. sativa* genes encoding putative triacylglycerol lipases among different tissues

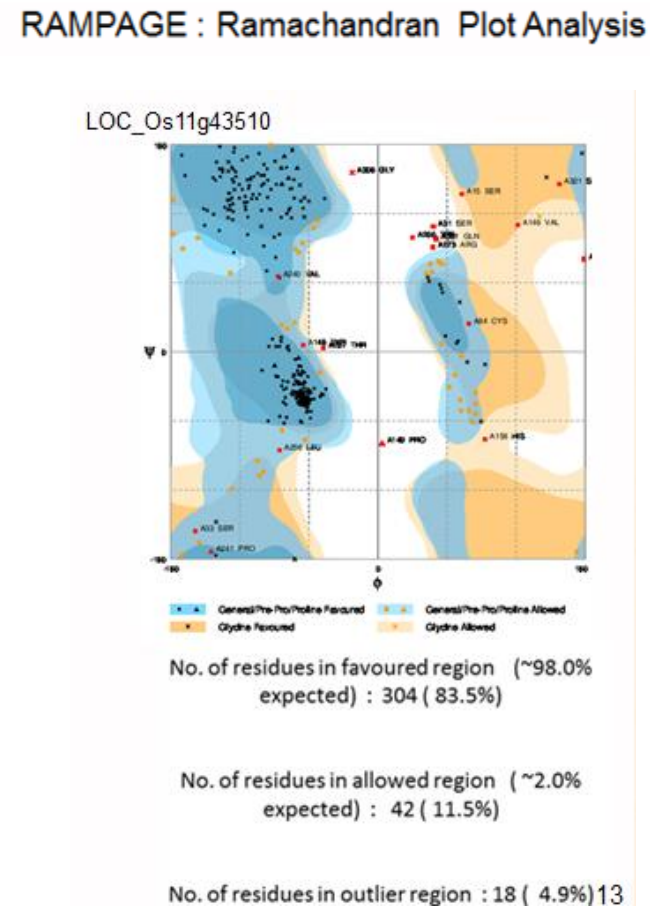


Fig. A.2 RAMPAGE analysis of modeled 3D structure of LOC_Os11g43510

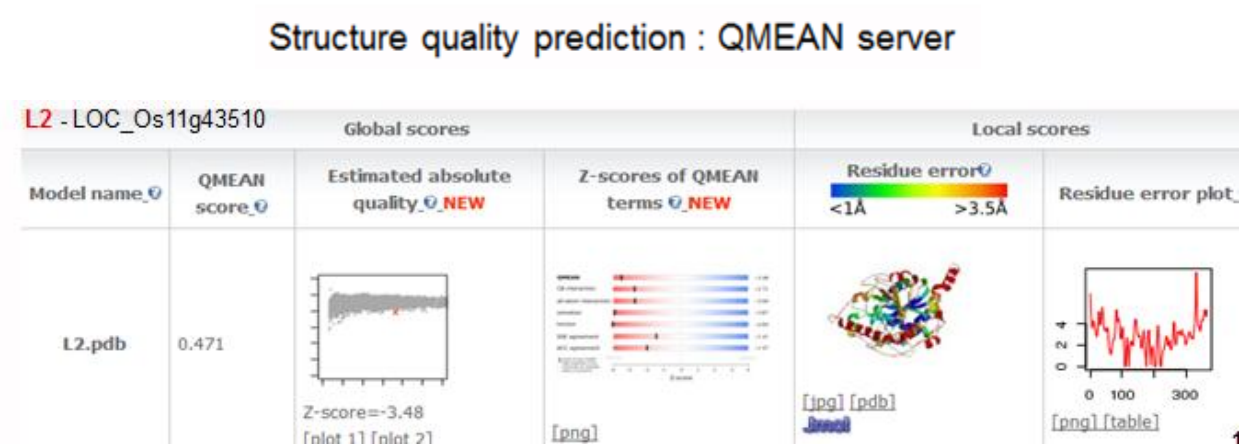


Fig. A.3 Structure quality prediction of modeled 3D structure of LOC_Os11g43510 using QMEAN server

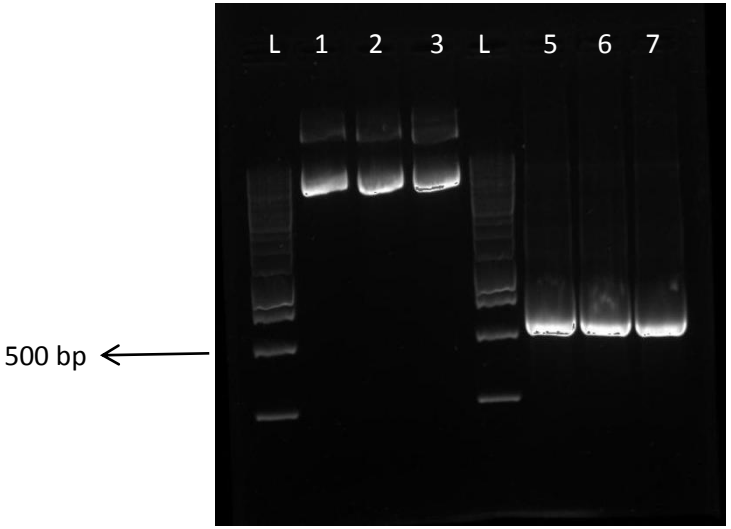


Fig. A.4 Plasmid and PCR analysis from selected positive pPICZαA-L2 clones in *E.coli*. L-1Kb ladder, lanes 1-3 –Plasmid from selected colonies, lanes 5-7 PCR products of expected size of 544bp using L2_F1 and L2_R1 primers.

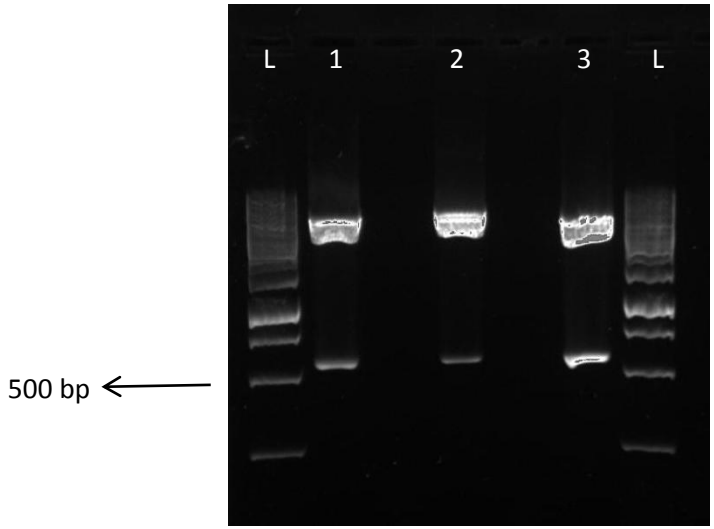


Fig. A. 5 Double digestion of plasmids from selected positive clone in *E.coli*with EcoRI and Xba I.

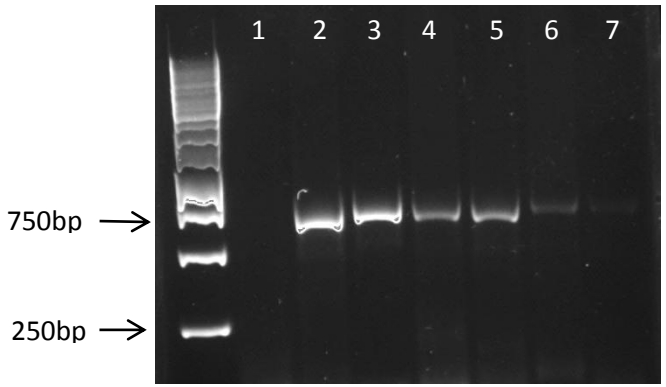
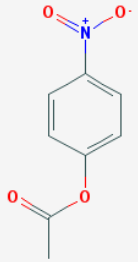
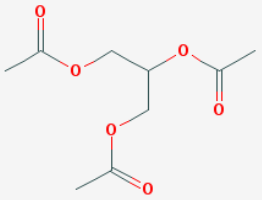
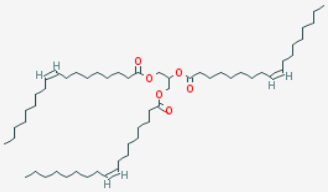


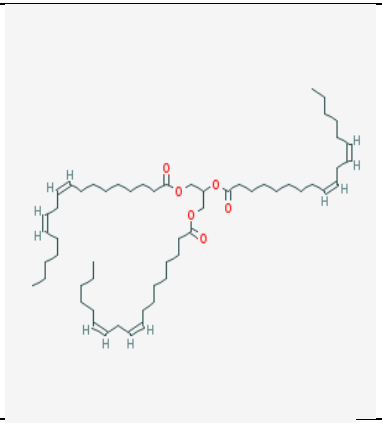
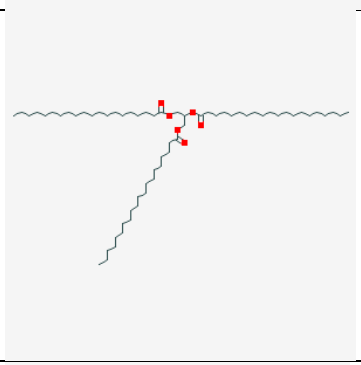
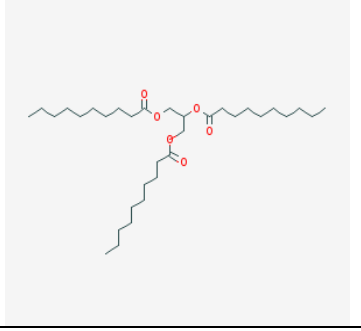
Fig. A.6 Agarose gel electrophoresis of PCR products confirming cloning and integration of LOC_Os11g43510 gene in *P.pastoris*. Lane 1 negative control, lane 2-7 PCR products from different selected colonies.

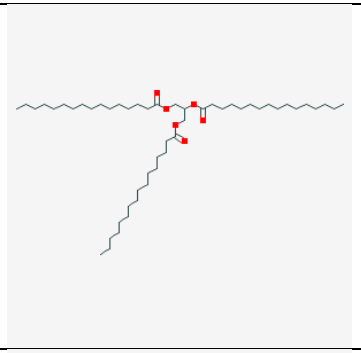
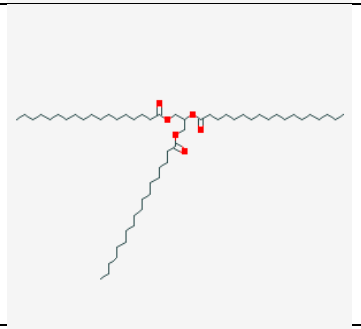
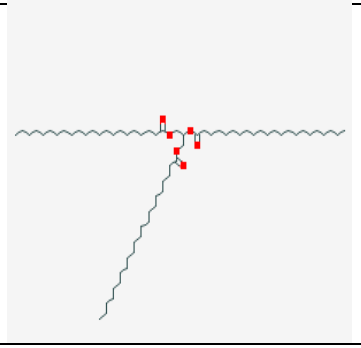
Table A.1 Primers used for cloning L2 in pICZα A

Target gene	Primer pair	Sequence
Lipase L2	L2F1 L2R1	5′-TAGATGAATTCATGGTGCAATATGCATCTGCTGTGT-3′ 5′-GCGTCTCTAGAACACGGATTGTTCTCGGGACACT-3′
PPICZαAL2	L2F1 3′AOX	5′-TAGATGAATTCATGGTGCAATATGCATCTGCTGTGT-3′ 5′-GCAAATGGCATTCTGACATCC-3′
Lipase L2	L2_F1 L2_R1	5′-ATGGTGCAATATGCATCTGCTGTGT-3′ 5′-ACACGGATTGTTCTCGGGACACT-3′

Table A.2 List of ligands used for Molecular docking studies.

S.No.	Ligand Name	Molecular formula	PubChem CID	2D structure
1.	p nitrophenyl acetate (pNPA)	$C_8H_7NO_4$	13243	
2.	Triacetin	$C_9H_{14}O_6$	5541	
3.	Triolein	$C_{57}H_{104}O_6$	5497163	

4.	Trilinolein	$C_{57}H_{98}O_6$	5322095	
5.	Triarachidin	$C_{63}H_{122}O_6$	522017	
6.	Tricaprin	$C_{33}H_{62}O_6$	69310	

7.	Tripalmitin	$C_{51}H_{98}O_6$	11147	
8.	Tristrearin	$C_{57}H_{110}O_6$	11146	
9.	Tribehnin	$C_{69}H_{134}O_6$	62726	

Supplemental Excel sheet 1

List of 125 lipase genes searched across rice genome using Massively Parallel Signature Sequencing (MPPS) database

Sno.	Gene	Title	Chr	Strand	5' End	3' End	BAC Name
1	LOC_Os01g07760	phospholipase D alpha 1 precursor, putative, expressed	1	c	3721117	3726086	P0583G08
2	LOC_Os01g07780	lipase/lipooxygenase, PLAT/LH2, putative, expressed	1	w	3732652	3733630	OSJNBa0089K24
3	LOC_Os01g10250	monoglyceride lipase, putative, expressed	1	c	5386541	5391216	B1046G12
4	LOC_Os01g11760	GDSL-like Lipase/Acylhydrolase family protein, expressed	1	w	6361830	6363959	P0515G01
5	LOC_Os01g11790	GDSL-like Lipase/Acylhydrolase family protein, expressed	1	w	6374063	6378839	P0515G01
6	LOC_Os01g14080	triacylglycerol lipase, putative, expressed	1	c	7876733	7879434	B1066G12
7	LOC_Os01g15000	triacylglycerol lipase, putative, expressed	1	w	8401049	8405176	P0705D01
8	LOC_Os01g20840	triacylglycerol lipase, putative, expressed	1	w	11600816	11605016	P0551A11
9	LOC_Os01g20860	phospholipase D. Active site motif family protein, expressed	1	c	11608139	11619557	P0551A11
10	LOC_Os01g21560	monoglyceride lipase, putative, expressed	1	w	12050165	12051594	B1153F04
11	LOC_Os01g33784	carboxylic ester hydrolase/ lipase, putative, expressed	1	w	18920136	18924936	P0516D04
12	LOC_Os01g43140	triacylglycerol lipase, putative, expressed	1	w	24938537	24942334	B1040D09
13	LOC_Os01g46090	lipase 1, putative	1	w	26530063	26533764	OJ1159_D09
14	LOC_Os01g46240	triacylglycerol lipase, putative, expressed	1	w	26623542	26625455	OJ1159_D09
15	LOC_Os01g46250	lipase-like protein, putative, expressed	1	w	26627042	26629913	OJ1159_D09
16	LOC_Os01g46290	triacylglycerol lipase, putative, expressed	1	w	26648231	26651658	OJ1159_D09
17	LOC_Os01g46370	triacylglycerol lipase, putative, expressed	1	c	26710443	26714042	OSJNBb0032K15
18	LOC_Os01g47610	triacylglycerol lipase, putative, expressed	1	c	27566966	27569358	OSJNBb0063G05
19	LOC_Os01g49380	monoglyceride lipase, putative, expressed	1	c	28710868	28714006	P0519D04
20	LOC_Os01g49510	esterase/lipase/thioesterase, putative, expressed	1	w	28806121	28810506	P0519D04

21	LOC_Os01g51360	triacylglycerol lipase, putative, expressed	1	c	29856574	29859694	P0456F08
22	LOC_Os01g52180	triacylglycerol lipase, putative, expressed	1	w	30340881	30341987	P0690B02
23	LOC_Os01g55650	patatin-like phospholipase family protein, expressed	1	c	32384248	32389677	P0512C01
24	LOC_Os01g62010	monoglyceride lipase, putative, expressed	1	w	36220130	36222027	P0506B12
25	LOC_Os01g67420	triacylglycerol lipase, putative, expressed	1	w	39485002	39492578	P0035F12
26	LOC_Os01g67430	triacylglycerol lipase, putative, expressed	1	w	39504018	39505525	P0035F12
27	LOC_Os01g67450	triacylglycerol lipase, putative	1	w	39522538	39523821	P0035F12
28	LOC_Os01g71010	lipase precursor, putative, expressed	1	c	41427345	41431905	P0492G09
29	LOC_Os01g73740	triacylglycerol lipase, putative	1	c	43040151	43040659	OJ1656_A11
30	LOC_Os02g03720	esterase/lipase/thioesterase family active site protein, putative, expressed	2	w	1560693	1563343	P0576F08
31	LOC_Os02g09770	esterase/lipase/thioesterase, putative, expressed	2	c	5038456	5043326	OSJNBb0031B09
32	LOC_Os02g18480	triacylglycerol lipase, putative, expressed	2	w	10758909	10762541	OJ1115_D03
33	LOC_Os02g18954	lipase/hydrolase, putative, expressed	2	w	11071693	11077519	OJ1124_E11
34	LOC_Os02g28040	triacylglycerol lipase, putative, expressed	2	c	16596799	16606368	OSJNBa0091C16
35	LOC_Os02g31200	esterase/lipase/thioesterase, putative	2	w	18685109	18685599	OSJNBa0004O05
36	LOC_Os02g42170	triacylglycerol lipase, putative, expressed	2	w	25358765	25359798	OJ1643_A10
37	LOC_Os02g43700	triacylglycerol lipase like protein, putative, expressed	2	w	26373590	26375810	P0491E01
38	LOC_Os02g44860	GSDL-motif lipase, putative, expressed	2	c	27152479	27156897	P0684A08
39	LOC_Os02g50000	GDSL-like Lipase/Acylhydrolase family protein, expressed	2	c	30548468	30550383	P0643A10
40	LOC_Os02g52830	triacylglycerol lipase, putative, expressed	2	c	32293905	32298113	OJ1004_A11
41	LOC_Os02g54010	triacylglycerol lipase, putative, expressed	2	w	33072355	33077687	OJ1369_G08
42	LOC_Os02g55330	esterase/lipase/thioesterase, putative, expressed	2	c	33870782	33874639	OJ1004_E04
43	LOC_Os02g58500	phospholipase A2, putative, expressed	2	c	35764745	35766255	OJ1149_C12
44	LOC_Os03g02740	phospholipase D beta 1, putative, expressed	3	c	1005922	1009991	OJ1263H11
45	LOC_Os03g18000	phosphoinositide-specific phospholipase C, putative	3	w	10003272	10003838	OSJNBb0027B12
46	LOC_Os03g18010	phosphoinositide-specific phospholipase C,	3	w	10004299	10006917	OSJNBb0027B12

		putative, expressed					
47	LOC_Os03g22670	triacylglycerol Lipase, putative, expressed	3	c	13066230	13072635	OSJNBa0006D11
48	LOC_Os03g25000	lipase 2, putative	3	w	14247529	14250741	OSJNBa0022F22
49	LOC_Os03g27370	phospholipase D alpha 1, putative, expressed	3	w	15652482	15657976	OSJNBa0065F09
50	LOC_Os03g30130	phospholipase C, putative, expressed	3	c	17162223	17165521	OSJNBb0059G13
51	LOC_Os03g50030	phospholipase A2, putative, expressed	3	w	28497524	28499203	OSJNBb0022E02
52	LOC_Os03g50410	lipase family protein	3	w	28715930	28717988	OSJNBb0033N16
53	LOC_Os03g51010	monoglyceride lipase, putative, expressed	3	c	29100693	29104260	B1377B10
54	LOC_Os03g61540	triacylglycerol lipase, putative, expressed	3	c	34835273	34839864	OSJNBa0078D06
55	LOC_Os03g62410	phospholipase D gamma 3, putative, expressed	3	c	35289129	35293830	OSJNBa0075M12
56	LOC_Os04g21160	gastric triacylglycerol lipase precursor, putative	4	c	11886348	11893334	OSJNBa0094P09
57	LOC_Os04g35100	phospholipase C, putative, expressed	4	c	21144839	21148155	OSJNBa0042L16
58	LOC_Os04g41200	lipase family protein, expressed	4	c	24212385	24221252	OSJNBa0084K20
59	LOC_Os04g43030	triacylglycerol lipase, putative, expressed	4	w	25241151	25249480	OSJNBb0065L13
60	LOC_Os04g56240	triacylglycerol lipase, putative, expressed	4	w	33311004	33314675	OSJNBa0071I13
61	LOC_Os05g03610	phospholipase C, putative, expressed	5	w	1513734	1519597	P0683F12
62	LOC_Os05g06140	lipase, putative, expressed	5	c	3047664	3052410	OSJNBa0072C16
63	LOC_Os05g07880	phospholipase D alpha 1 precursor, putative, expressed	5	c	4233782	4238352	P0685E10
64	LOC_Os05g07890	lipase/lipooxygenase, PLAT/LH2, putative, expressed	5	w	4240661	4241468	P0685E10
65	LOC_Os05g12330	esterase/lipase/thioesterase family protein, putative, expressed	5	c	7066396	7067821	OSJNBb0067H15
66	LOC_Os05g29050	phospholipase D p1, putative, expressed	5	c	16970428	16981729	OSJNBa0009L15
67	LOC_Os05g29974	monoglyceride lipase, putative, expressed	5	w	17255731	17263218	P0692D12
68	LOC_Os05g30900	lipase precursor, putative, expressed	5	c	17860814	17864526	OSJNBa0025P09
69	LOC_Os05g32380	triacylglycerol lipase, putative, expressed	5	c	18803258	18804753	OJ1562_H01
70	LOC_Os05g33820	lipase, putative, expressed	5	c	19851468	19855093	OSJNBb0014K18
71	LOC_Os05g43080	GDSL-motif lipase/hydrolase-like protein, putative	5	c	24938947	24939275	OSJNBb0013J02
72	LOC_Os05g43110	GDSL-like Lipase/Acylhydrolase family protein,	5	c	24946588	24949630	OSJNBb0013J02

		expressed					
73	LOC_Os05g49830	triacylglycerol lipase, putative, expressed	5	w	28514304	28515613	OJ1268_B08
74	LOC_Os05g49840	triacylglycerol lipase, putative, expressed	5	c	28515068	28519994	OJ1268_B08
75	LOC_Os06g10850	triacylglycerol lipase, putative, expressed	6	w	5656851	5660266	P0021C04
76	LOC_Os06g40170	phospholipase D alpha 2, putative, expressed	6	c	23907921	23912800	P0481H08
77	LOC_Os06g40180	phospholipase D alpha 2, putative, expressed	6	w	23920943	23923922	P0481H08
78	LOC_Os06g40190	phospholipase D alpha 2, putative, expressed	6	w	23927705	23931375	P0481H08
79	LOC_Os06g42730	esterase/lipase/thioesterase, putative, expressed	6	w	25697744	25701600	P0505A04
80	LOC_Os06g42860	gastric triacylglycerol lipase precursor, putative, expressed	6	c	25753841	25756446	OSJNBa0019I19
81	LOC_Os06g44060	phospholipase D. Active site motif family protein, expressed	6	c	26574128	26578687	P0453H04
82	LOC_Os06g46350	patatin-like phospholipase family protein, expressed	6	c	28108238	28109953	P0710B08
83	LOC_Os07g15680	phospholipase D beta 1, putative, expressed	7	c	9103800	9111224	P0046D03
84	LOC_Os07g28250	triacylglycerol lipase, putative	7	w	16501471	16503020	P0404G11
85	LOC_Os07g33670	patatin-like phospholipase family protein, expressed	7	c	20113836	20121270	OJ1657_A07
86	LOC_Os07g34400	triacylglycerol lipase, putative	7	c	20623397	20624993	OSJNBa0007H12
87	LOC_Os07g34420	triacylglycerol lipase, putative	7	w	20645291	20645911	OSJNBa0007H12
88	LOC_Os07g34440	triacylglycerol lipase, putative	7	w	20650553	20652352	OSJNBa0007H12
89	LOC_Os07g37840	monoglyceride lipase, putative, expressed	7	c	22691463	22692798	OJ1773_H01
90	LOC_Os07g39810	triacylglycerol lipase, putative, expressed	7	c	23849587	23850754	OJ1113_E01
91	LOC_Os07g47250	lipase precursor, putative, expressed	7	w	28238772	28243756	P0625E02
92	LOC_Os07g49330	phosphoinositide-specific phospholipase C, putative, expressed	7	c	29544132	29547288	P0627E10
93	LOC_Os08g04800	triacylglycerol lipase like protein, putative, expressed	8	w	2399689	2401410	P0025F03
94	LOC_Os08g06420	esterase/lipase/thioesterase, putative, expressed	8	c	3590010	3595354	P0577B11
95	LOC_Os08g31060	phospholipase D alpha 1, putative, expressed	8	c	19040682	19048021	OSJNBa0086F04
96	LOC_Os08g38092	esterase/lipase/thioesterase, putative, expressed	8	c	23997086	24001873	P0028A08

97	LOC_Os08g41780	gastric triacylglycerol lipase precursor, putative, expressed	8	c	26246575	26251798	OJ1789_C07
98	LOC_Os09g01590	gastric triacylglycerol lipase precursor, putative, expressed	9	w	409594	418218	P0414D03
99	LOC_Os09g23150	monoglyceride lipase, putative, expressed	9	c	13694146	13698817	B1040D06
100	LOC_Os09g25390	phospholipase D alpha 1, putative, expressed	9	c	15221651	15225443	OJ1740_D06
101	LOC_Os09g31050	phospholipase A2, group IVB isoform 6, putative, expressed	9	c	18672329	18673856	OSJNBa0046G16
102	LOC_Os09g33820	phospholipase A1, putative, expressed	9	w	19964721	19969670	P0450E05
103	LOC_Os09g37100	phospholipase D delta, putative, expressed	9	w	21391293	21397468	P0478E02
104	LOC_Os09g39790	triacylglycerol lipase, putative, expressed	9	c	22810710	22813586	B1331F11
105	LOC_Os10g38060	phospholipase D beta 1, putative, expressed	10	c	20047348	20056413	OSJNBa0096G08
106	LOC_Os10g41270	triacylglycerol lipase like protein, putative	10	w	21851099	21852703	OSJNBb0089A17
107	LOC_Os11g01040	monoglyceride lipase, putative, expressed	11	c	13244	14796	OSJNBa0029D01
108	LOC_Os11g03520	acyltransferase/ carboxylic ester hydrolase/ lipase, putative, expressed	11	c	1358189	1359545	OSJNBa0056E15
109	LOC_Os11g19290	triacylglycerol lipase, putative	11	c	11028408	11029781	OSJNBb0056F11
110	LOC_Os11g19340	triacylglycerol lipase, putative, expressed	11	c	11088166	11089823	OSJNBa0046F10
111	LOC_Os11g34440	phospholipase A2, putative, expressed	11	c	19673520	19675027	OSJNBb0005H02
112	LOC_Os11g43510	lipase precursor, putative, expressed	11	w	25756945	25760373	P0485F09
113	LOC_Os11g43760	triacylglycerol lipase, putative	11	w	25927218	25932871	P0682E05
114	LOC_Os12g01030	monoglyceride lipase, putative	12	c	11806	13099	OSJNBb0077A02
115	LOC_Os12g16180	phospholipase, putative, expressed	12	w	9244028	9247791	OJ1005_C11
116	LOC_Os12g17570	lipase/hydrolase, putative, expressed	12	c	10067403	10069844	OSJNBa0056I18
117	LOC_Os12g18860	esterase/lipase/thioesterase, putative, expressed	12	c	10932942	10939812	OSJNBb0034E16
118	LOC_Os12g36770	triacylglycerol lipase, putative, expressed	12	c	22496653	22501413	OSJNBa0027H05
119	LOC_Os12g37560	phospholipase C, putative, expressed	12	w	23015182	23019784	OSJNBb0076G11
120	LOC_Os12g37630	phospholipase/Carboxylesterase family protein, expressed	12	w	23068480	23071642	OSJNBb0076G11
121	LOC_Os12g37910	GDSL-motif lipase/hydrolase-like protein,	12	c	23261363	23265644	OSJNBa0011B18

		putative, expressed					
122	LOC_Os12g41720	patatin-like phospholipase family protein, expressed	12	c	25804958	25806873	OSJNBa0018C20
123	LOC_Os12g41970	triacylglycerol lipase, putative	12	c	25992801	25993931	OJ1327_A12
124	LOC_Os12g41980	triacylglycerol lipase, putative, expressed	12	c	26000435	26001831	OJ1327_A12
125	LOC_Os12g42010	triacylglycerol lipase, putative	12	c	26012052	26013206	OJ1327_A12

Supplemental excel sheet 2

Fragments Per Kilobase of transcript per Million (FPKM) expression values obtained from Rice Genome Annotation Project Database

Locus Name	Gene Product Name	Leaves-20 days (RPKM)	Seed-5 DAP (RPKM)	Embryo- 25 DAP (RPKM)	Endosperm- 25 DAP (RPKM)	Seed- 10 DAP (RPKM)
LOC_Os01g07760	phospholipase D, putative	79.4385	42.9071	163.821	4.2279	5.51822
LOC_Os01g07780	embryo-specific 3, putative	32.516	5.54351	3.66834	0	0
LOC_Os01g10250	hydrolase, alpha/beta fold	26.3544	10.9428	2.53505	0.467886	0.863978
LOC_Os01g11760	GDSL-like lipase/acylhydrolase	0	0	0	0	0
LOC_Os01g11790	GDSL-like lipase/acylhydrolase	0	0	0	0	0
LOC_Os01g14080	lipase class 3 family protein	1.02302	1.47081	0.70626	0	0.408576
LOC_Os01g15000	lipase, putative	7.98583	6.80384	8.70351	117.912	20.3993
LOC_Os01g20840	lipase class 3 family protein	15.5197	16.2984	11.6756	77.5955	44.0389
LOC_Os01g20860	phospholipase D.	4.66406	6.77368	7.4404	12.3492	16.4253
LOC_Os01g21560	esterase/lipase/thioesterase	6.00706	56.9157	11.3699	0.750207	1.21676
LOC_Os01g33784	lipase family protein, putative	22.5462	12.8077	42.3385	48.9816	27.1793
LOC_Os01g43140	lipase, putative,	0	8.99282	5.22409	0	0.578944
LOC_Os01g46090	GDSL-like lipase/acylhydrolase	0	0	0	0	0
LOC_Os01g46240	lipase class 3 family protein	0	4.91756	1.25961	0	0.410748
LOC_Os01g46250	lipase, putative,	0	0	0	0	0
LOC_Os01g46290	lipase, putative,	0	24.932	7.59201	0	0
LOC_Os01g46370	lipase class 3 family protein	0	0.961182	0	0	0
LOC_Os01g47610	lipase class 3 family protein	1.25409	1.42652	0	0	0
LOC_Os01g49380	lipase, putative,	2.27966	2.8529	3.51753	0.467075	0.683562
LOC_Os01g49510	OsPOP3 - Putative ProlylOligopeptidase homologue	4.23149	6.17908	13.4312	1.23195	1.50809
LOC_Os01g51360	lipase, putative, expressed	73.6142	12.8694	0.346511	0	0

LOC_Os01g52180	lipase, putative	0	0	0	0	0
LOC_Os01g55650	phospholipase, patatin family	4.13557	7.11762	16.2858	10.2068	9.27072
LOC_Os01g62010	hydrolase, alpha/beta fold	0	0	0	0	0
LOC_Os01g67420	lipase, putative,	0.887732	3.70417	16.9364	0.809191	1.4171
LOC_Os01g67430	lipase, putative	0	0	0	0	0
LOC_Os01g67450	lipase, putative	0	0	0	0	0
LOC_Os01g71010	lipase precursor, putative	12.9841	10.2192	5.81278	2.4346	3.94384
LOC_Os01g73740	lipase, putative,	0	0	0	0	0
LOC_Os02g03720	expressed protein	27.8964	1.0674	0.836389	0	0
LOC_Os02g09770	abhydrolasedomaincontaiprot	8.63385	4.07304	4.60147	0	0.31182
LOC_Os02g18480	lipase class 3 family protein	6.2418	6.72843	15.2319	1.06515	1.59942
LOC_Os02g18954	GDSL-like lipase/acylhydrolase	0	0	0	0	0
LOC_Os02g28040	lipase, putative	4.23283	7.27694	4.13493	3.30565	4.9645
LOC_Os02g31200	esterase/lipase/thioesterase	0	0	0	0	0
LOC_Os02g42170	phospholipase, putative	0	0	0	0	0
LOC_Os02g43700	triacylglycerol lipase like protein	35.3372	10.2236	8.47532	2.93983	1.51806
LOC_Os02g44860	GDSL-like lipase/acylhydrolase	1.28867	10.4032	22.2066	0	0.42688
LOC_Os02g50000	GDSL-like lipase/acylhydrolase	0	0	0.567951	18.4784	0
LOC_Os01g47610	lipase class 3 family protein	1.25409	1.42652	0	0	0
LOC_Os02g54010	lipase class 3 family protein	12.0197	5.4182	10.4777	7.6112	4.23027
LOC_Os02g55330	Putative ProlylOligopeptidase	1.7549	3.93577	2.76505	11.1881	2.51177
LOC_Os02g58500	phospholipase A2, putative	1.87906	2.54829	0	0	0
LOC_Os03g02740	phospholipase D PUTATIVE	0	1.24181	0	0	0
LOC_Os03g18000	phosphoinositide-specific phospholipase C	0	0	0	0	0
LOC_Os03g18010	phospholipase C	0.859184	0	0.43566	0	0
LOC_Os03g22670	triacylglycerol Lipase, putative	3.99163	0.546064	1.40303	0.182095	0.370746
LOC_Os03g25000	GDSL-like lipase/acylhydrolase	0	0	0	0	0
LOC_Os03g27370	phospholipase D,	0	0	0	0	0
LOC_Os03g30130	phospholipase C, putative	0.800239	6.38667	2.27426	10.9596	7.94104

LOC_Os03g50030	phospholipase A2	0	47.0225	39.6091	5.53607	7.45512
LOC_Os03g50410	lipase family protein	0	0	0	5.98631	0.670189
LOC_Os03g51010	hydrolase, alpha/beta fold	17.3812	7.1979	24.1679	6.38166	5.81158
LOC_Os03g61540	lipase class 3 family protein	8.16931	6.64639	1.4019	2.71138	2.75226
LOC_Os03g62410	phospholipase D,	0	0	0	0	0
LOC_Os04g21160	triacylglycerol lipase 1 precursor	0	0	0	0	0
LOC_Os04g35100	phospholipase C, putative	16.4286	7.40771	1.06805	23.2412	15.5446
LOC_Os04g41200	lipase, putative,	5.93978	5.47201	4.12582	3.7284	4.14086
LOC_Os04g43030	lipase class 3 family protein,	12.2021	2.13307	4.59723	0.936163	2.36832
LOC_Os04g56240	lipase, putative,	372.16	2.31104	13.2508	0	0
LOC_Os05g03610	phospholipase C,	4.72492	3.18645	0.333089	0	0
LOC_Os05g06140	lipase, putative,	0	0	0	0	0
LOC_Os05g07880	phospholipase D	2.09527	0	0	0	0
LOC_Os05g07890	embryo-specific 3, putative	4.5828	6.13125	9.99163	0	0
LOC_Os05g12330	uncharacterized protein KIAA1310	0	0	0	0	0
LOC_Os05g29050	phospholipase D p1	4.94835	2.86509	3.8679	1.69476	1.81776
LOC_Os05g29974	lipase, putative,	11.4361	4.86216	15.3722	2.58822	2.07142
LOC_Os05g30900	lipase precursor, putative	7.238	8.11416	3.6099	2.65177	3.11706
LOC_Os05g32380	phospholipase, putative	3.60566	0.74658	7.37712	0	0
LOC_Os05g33820	lipase, putative	14.8778	3.66512	0.821867	32.3585	10.7207
LOC_Os05g43080	The Gene Report Page only displays MSU Rice Genome Annotation Project Release 7 data .					
LOC_Os05g43080	The Gene Report Page only displays MSU Rice Genome Annotation Project Release 7 data .					
LOC_Os05g43110	The Gene Report Page only displays MSU Rice Genome Annotation Project Release 7 data .					
LOC_Os05g49830	lipase class 3 family protein	0	0	0	0	0
LOC_Os05g49840	phospholipase, putative	1.17572	8.83906	0	0	0
LOC_Os06g10850	lipase, putative,	0	0	0	0	0
LOC_Os06g40170	phospholipase D, putative	1.97795	0.632608	0.690201	0	0
LOC_Os06g40180	phospholipase D	3.61484	0.336105	0.131255	0	0
LOC_Os06g40190	phospholipase D, p	12.8215	14.3476	2.90251	0.225134	0.728486

LOC_Os06g42730	Putative ProlylOligopeptidase	2.68362	4.02609	2.90142	0	0.280915
LOC_Os06g42860	triacylglycerol lipase precursor,	0	0	0	0	0
LOC_Os06g44060	phospholipase D. Active site motif	0.661016	3.31585	4.47033	0	0.22414
LOC_Os06g46350	PLA IIIA/PLP7, putative	0	0	2.56845	0	0
LOC_Os07g15680	phospholipase D,	0.220455	0.211105	1.22363	0	0
LOC_Os07g28250	lipase class 3 family protein,	0	0	0	0	0
LOC_Os07g33670	patatin-like phospholipase	4.44099	4.47085	11.874	0.933041	1.8613
LOC_Os07g34400	lipase class 3 family protein	0	0	0	0	0
LOC_Os07g34420	lipase class 3 family protein	0	0	0	0	0
LOC_Os07g34440	lipase class 3 family protein,	0	0	0	0	0
LOC_Os07g37840	lipase, putative,	0.84263	0.0348405	0.154198	0	0.0134988
LOC_Os07g39810	lipase class 3 family protein,	0	4.43906	0	0	0
LOC_Os07g47250	lipase precursor, putative,	8.0431	6.95328	12.6135	33.9673	19.4373
LOC_Os07g49330	phospholipase C,	55.8207	21.5555	28.6414	41.7476	17.872
LOC_Os08g04800	triacylglycerol lipase like protein,	0	0	0	0	0
LOC_Os08g06420	abhydrolase domain-containing protein	3.27259	91.2441	20.0967	32.5477	68.974
LOC_Os08g31060	phospholipase D alpha 1,	0	0	0	0	0
LOC_Os08g38092	abhydrolase domain-containing	3.75089	0	0	0	0
LOC_Os08g41780	triacylglycerol lipase precursor,	0.48748	2.43531	16.2682	1.50965	1.23572
LOC_Os09g01590	triacylglycerol lipase 1 precursor	9.71571	8.41525	11.4893	1.8458	3.15063
LOC_Os09g23150	monoglyceride lipase,	15.143	3.61829	3.5044	0.755203	0.600608
LOC_Os09g25390	phospholipase D, putative	0	0.539446	0	0	0
LOC_Os09g31050	cytosolic phospholipase A2 beta	0.991337	6.01568	3.06237	1.12928	1.33637
LOC_Os09g33820	lecithine cholesterol acyltransferase	5.09937	3.69697	9.27084	1.4199	1.66039
LOC_Os09g37100	phospholipase D,	7.41456	6.42872	32.2304	1.66058	2.47104
LOC_Os09g39790	lipase, putative	3.25072	8.17241	14.4234	1.48457	4.93306
LOC_Os10g38060	phospholipase D,	3.33576	10.2325	6.09642	6.44537	10.0046

LOC_Os10g41270	triacylglycerol lipase like protein	1.01152	0	0	0	0
LOC_Os11g01040	monoglyceride lipase	10.602	1.49997	3.43896	0	0
LOC_Os11g03520	GDSL-like lipase/acylhydrolase	0	0	0	0	0
LOC_Os11g19290	The Gene Report Page only displays MSU Rice Genome Annotation Project Release 7 data .					
LOC_Os11g19340	lipase, putative,	0	0	0	0	0
LOC_Os11g34440	phospholipase A2, putative	0	0	0	0	0
LOC_Os11g43510	lipase precursor, putative	19.4888	5.37924	3.36689	1.13504	1.28295
LOC_Os11g43760	lipase class 3 family protein	0	0	0	0	0
LOC_Os12g01030	monoglyceride lipase, putative	13.0296	1.81721	3.93153	0	0
LOC_Os12g16180	The Gene Report Page only displays MSU Rice Genome Annotation Project Release 7 data .					
LOC_Os12g17570	GDSL-like lipase/acylhydrolase	0	0.398696	1.31996	0	0
LOC_Os12g18860	Putative ProlylOligopeptidase homologue	12.2698	15.1787	24.7183	10.4498	7.38176
LOC_Os12g36770	lipase class 3 family protein	2.49291	2.16159	1.11654	0	0.301969
LOC_Os12g37560	phospholipase C, putative	3.24145	5.11899	4.44572	2.07673	2.62916
LOC_Os12g37630	domain of unknown function family protein	5.52688	10.0284	14.0254	1.29491	2.74427
LOC_Os12g37910	The Gene Report Page only displays MSU Rice Genome Annotation Project Release 7 data .					
LOC_Os12g41720	The Gene Report Page only displays MSU Rice Genome Annotation Project Release 7 data .					
LOC_Os12g41970	lipase class 3 family protein	0	0	0	0	0
LOC_Os12g41980	lipase class 3 family protein	0	0	0	0	0
LOC_Os12g42010	lipase class 3 family protein	0	0	0	0	0

Chapter 3

			Trancriptomics expression data in RPKM values								
BD192 Embryo vs leaf											
Feature ID	t-test: BD1	BD192_Em	BD192_Em	BD192_Em	BD192 Em	BD192_Le	BD192_Le	BD192_Le	BD192 Leaf RPKM - Means		
Os01g0817700	6.05E-03	48.305	95.151	77.855	73.77	0.75	0.911	1.75	1.137		
IDB Embryo vs leaf											
Feature ID	t-test: IDB	IDB_Embr	IDB_Embr	IDB_Embr	IDB Embr	IDB_Leaf1	IDB_Leaf2	IDB_Leaf3	IDB Leaf RPKM - Means		
Os01g0817700	5.27E-04	83.722	78.924	106.541	89.729	0.709	0.752	4.732	2.064		

Table 1. Transcriptomics RNA-Seq data comparing expression of Os01g0817700 in embryo and leaf in two rice lines (BD192 and IDB).

Primer name	Primer sequence
F.P.	5’-CCAACGCCGACAAGTACAAC-3’
R.P.	5’-ACCTCTCGGGAGCACCATT-3’
OsTuba1-F	5’-TACCCACTCCCTCCTTGAGC-3’
OsTuba1-R	5’-AGGCACTGTTGGTGATCTCG-3’

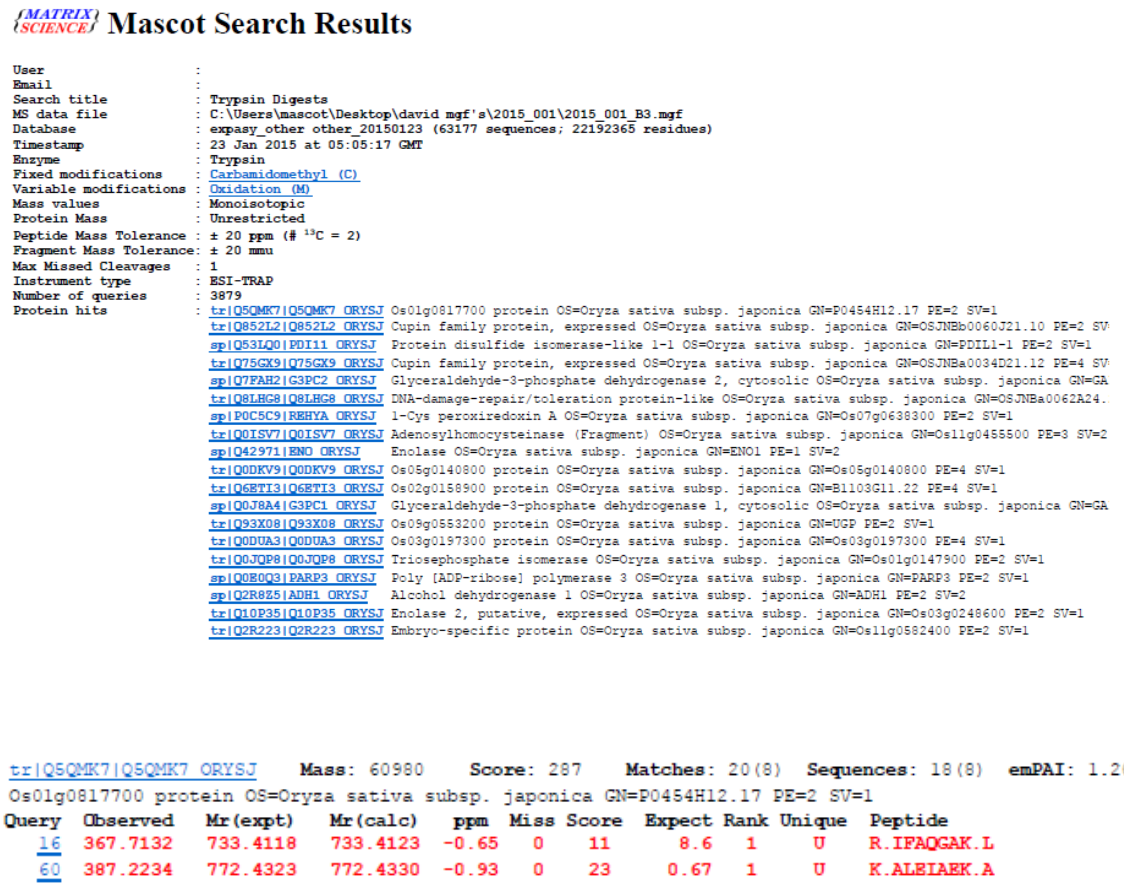
Table 2. Primers used for PCR expression analysis

	Leaf					
Contigs	**BD192**			**IDB**		
	Sample1	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**
Counted contigs	2,503,685	2,060,799	2,414,893	1,987,471	2,653,290	3,747,481
Unique contigs	2,408,745	1,999,226	2,350,016	1,941,017	2,590,278	3,610,811
Non-S contigs	94,940	61,573	64,877	46,454	63,012	136,670
Un-C contigs	824,288	686,116	654,523	590,945	763,779	1,339,752
Total contigs	3,327,973	2,746,915	3,069,416	2,578,416	3,417,069	5,087,233
Counted contigs (%)	75.23	75.02	78.68	77.08	77.65	73.66
Unique contigs (%)	72.78	72.78	76.56	75.28	75.80	70.98
Non-S contigs (%)	2.85	2.24	2.11	1.80	1.84	2.69
Un-C contigs (%)	24.77	24.98	21.32	22.92	22.35	26.34

	Embryo					
Contigs	**BD192**			**IDB**		
	Sample1	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**
Counted contigs	1,487,045	2,171,009	3,273,871	2,029,962	1,766,719	1,700,952
Unique contigs	1,451,236	2,128,279	3,261,305	1,987,144	1,732,409	1,661,849
Non-S contigs	35,809	43,730	12,566	42,818	34,310	39,103
Un-C contigs	303,390	384,119	891,683	313,886	435,925	480,036
Total contigs	1,790,435	2,555,128	4,165,554	2,343,848	2,202,644	2,180,988
Counted contigs (%)	83.05	84.97	78.59	89.87	80.21	77.99
Unique contigs (%)	81.05	83.29	78.29	79.13	78.65	76.20
Non-S contigs (%)	2	1.67	0.30	10.74	1.56	1.79
Un-C contigs (%)	16.95	15.03	21.41	10.13	19.79	22.01

Non-S contigs: Non-specific contigs; Un-C contigs: Un-counted contigs.

Table 3. Mapped contig results of RNA-Seq reads from BD192 and IDB rice lines in leaf and embryo tissues



Mascot Search Results

User	:
Email	:
Search title	: Trypsin Digests
MS data file	: C:\Users\mascot\Desktop\david mgf's\2015_001\2015_001_B3.mgf
Database	: expasy_other other_20150123 (63177 sequences; 22192365 residues)
Timestamp	: 23 Jan 2015 at 05:05:17 GMT
Enzyme	: Trypsin
Fixed modifications	: Carbamidomethyl (C)
Variable modifications	: Oxidation (M)
Mass values	: Monoisotopic
Protein Mass	: Unrestricted
Peptide Mass Tolerance	: ± 20 ppm (# ^{13}C = 2)
Fragment Mass Tolerance	: ± 20 mmu
Max Missed Cleavages	: 1
Instrument type	: ESI-TRAP
Number of queries	: 3879
Protein hits	: tr|Q5QMK7|Q5QMK7_ORYSJ Os01g0817700 protein OS=Oryza sativa subsp. japonica GN=P0454H12.17 PE=2 SV=1
	tr|Q852L2|Q852L2_ORYSJ Cupin family protein, expressed OS=Oryza sativa subsp. japonica GN=OSJNBb0060J21.10 PE=2 SV
	sp|Q53LQ0|PDI11_ORYSJ Protein disulfide isomerase-like 1-1 OS=Oryza sativa subsp. japonica GN=PDIL1-1 PE=2 SV=1
	tr|Q75GX9|Q75GX9_ORYSJ Cupin family protein, expressed OS=Oryza sativa subsp. japonica GN=OSJNBa0034D21.12 PE=4 SV
	sp|Q7FAH2|G3PC2_ORYSJ Glyceraldehyde-3-phosphate dehydrogenase 2, cytosolic OS=Oryza sativa subsp. japonica GN=GA
	tr|Q8LHG8|Q8LHG8_ORYSJ DNA-damage-repair/toleration protein-like OS=Oryza sativa subsp. japonica GN=OSJNBa0062A24.
	sp|P0C5C9|REHYA_ORYSJ 1-Cys peroxiredoxin A OS=Oryza sativa subsp. japonica GN=Os07g0638300 PE=2 SV=1
	tr|Q0ISV7|Q0ISV7_ORYSJ Adenosylhomocysteinase (Fragment) OS=Oryza sativa subsp. japonica GN=Os11g0455500 PE=3 SV=2
	sp|Q42971|ENO_ORYSJ Enolase OS=Oryza sativa subsp. japonica GN=ENO1 PE=1 SV=2
	tr|Q0DKV9|Q0DKV9_ORYSJ Os05g0140800 protein OS=Oryza sativa subsp. japonica GN=Os05g0140800 PE=4 SV=1
	tr|Q6ETI3|Q6ETI3_ORYSJ Os02g0158900 protein OS=Oryza sativa subsp. japonica GN=B1103G11.22 PE=4 SV=1
	sp|Q0J8A4|G3PC1_ORYSJ Glyceraldehyde-3-phosphate dehydrogenase 1, cytosolic OS=Oryza sativa subsp. japonica GN=GA
	tr|Q93X08|Q93X08_ORYSJ Os09g0553200 protein OS=Oryza sativa subsp. japonica GN=UGP PE=2 SV=1
	tr|Q0DUA3|Q0DUA3_ORYSJ Os03g0197300 protein OS=Oryza sativa subsp. japonica GN=Os03g0197300 PE=4 SV=1
	tr|Q0JQP8|Q0JQP8_ORYSJ Triosephosphate isomerase OS=Oryza sativa subsp. japonica GN=Os01g0147900 PE=2 SV=1
	sp|Q0E0Q3|PARP3_ORYSJ Poly [ADP-ribose] polymerase 3 OS=Oryza sativa subsp. japonica GN=PARP3 PE=2 SV=1
	sp|Q2R8Z5|ADH1_ORYSJ Alcohol dehydrogenase 1 OS=Oryza sativa subsp. japonica GN=ADH1 PE=2 SV=2
	tr|Q10P35|Q10P35_ORYSJ Enolase 2, putative, expressed OS=Oryza sativa subsp. japonica GN=Os03g0248600 PE=2 SV=1
	tr|Q2R223|Q2R223_ORYSJ Embryo-specific protein OS=Oryza sativa subsp. japonica GN=Os11g0582400 PE=2 SV=1

1. tr|Q5QMK7|Q5QMK7_ORYSJ **Mass:** 60980 **Score:** 287 **Matches:** 20(8) **Sequences:** 18(8) **emPAI:** 1.20

Os01g0817700 protein OS=Oryza sativa subsp. japonica GN=P0454H12.17 PE=2 SV=1

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
16	367.7132	733.4118	733.4123	-0.65	0	11	8.6	1	U	R.IFAQGAK.L
60	387.2234	772.4323	772.4330	-0.93	0	23	0.67	1	U	K.ALEIAEK.A

Fig.1. MASCOT analysis of LC/MS results obtained from lipase activity band after in gel lipase assay using MUF-butyrate.

SEARCH
Keyword search
Sequence search

BROWSE
Organisms
Taxonomy
Statistics
SCOP
Hierarchy
Ontologies
GO
EC
Phenotype

TOOLS
Compare genomes
Phylogenetic trees
Web services
Downloads

ABOUT
Description
Publications
Documentation

HELP
User support

Structural Classification | Genome Assignments | Sequence Alignments | Domain Combinations | Taxonomic Distribution

Alkaline phosphatase-like superfamily

SCOP classification

Root: SCOP hierarchy in SUPERFAMILY [SCOP 0] (11)
Class: Alpha and beta proteins (a/b) [SCOP 51349] (147)
Fold: Alkaline phosphatase-like [SCOP 53648]
Superfamily: Alkaline phosphatase-like [SCOP 53649] (5)
Families: 2,3-Bisphosphoglycerate-independent phosphoglycerate mutase, catalytic domain [SCOP 64162]
Alkaline phosphatase [SCOP 53650]
Arylsulfatase [SCOP 53653] (3)
Phosphonoacetate hydrolase [SCOP 102651]
DeoB catalytic domain-like [SCOP 142735]

Superfamily statistics

	Genomes (3,029)	*Uniprot* 2014_06	PDB chains (SCOP 1.75)
Domains	24,286	122,139	33
Proteins	23,852	121,688	33

Functional annotation

General category	Metabolism
Detailed category	Other enzymes

Gene Ontology (high-quality)

(show details)

	GO term	FDR (singleton)	FDR (all)	SDFO level	Annotation (direct or inherited)
Biological Process (BP)	developmental process	0.09397	0.8516	*Least Informative*	*Inherited*
Biological Process (BP)	multicellular organismal process	0.03448	0.2616	*Least Informative*	*Inherited*
Biological Process (BP)	response to stimulus	0.3946	1	*Least Informative*	*Inherited*
Biological Process (BP)	single-organism metabolic process	0.03639	0.00000000000004613	*Least Informative*	*Inherited*
Biological Process (BP)	protein metabolic process	0.1095	0.0165	*Least Informative*	*Inherited*
Biological Process (BP)	cellular macromolecule metabolic process	1	1	*Least Informative*	*Inherited*
Biological Process (BP)	nitrogen compound metabolic process	0.7573	1	*Least Informative*	*Inherited*
Biological Process (BP)	lipid metabolic process	0.000007227	0.00000000000004773	*Moderately Informative*	*Direct*
Biological Process (BP)	sulfur compound metabolic process	0.00003264	0	*Moderately Informative*	*Direct*
Biological Process (BP)	carbohydrate derivative metabolic process	0.00000001337	0	*Moderately Informative*	*Direct*
Biological Process (BP)	organ development	0.1019	0.06209	*Moderately Informative*	*Inherited*
Biological Process (BP)	anatomical structure morphogenesis	0.02653	0.05426	*Moderately Informative*	*Inherited*
Biological Process (BP)	cellular catabolic process	0.8371	0.9017	*Moderately Informative*	*Inherited*
Biological Process (BP)	cellular protein modification process	0.2059	0.2149	*Moderately Informative*	*Inherited*
Biological Process (BP)	organic hydroxy compound metabolic process	0.00004445	0.003651	*Moderately Informative*	*Inherited*
Biological Process (BP)	phosphate-containing compound metabolic process	0.6744	0.3561	*Moderately Informative*	*Inherited*
Biological Process (BP)	response to organic substance	0.00002781	0.3881	*Moderately Informative*	*Inherited*
Biological Process (BP)	organic substance catabolic process	0.7435	0.5658	*Moderately Informative*	*Inherited*
Biological Process (BP)	cellular biogenic amine metabolic process	0.00000002146	0.0000000001935	*Informative*	*Direct*
Biological Process (BP)	membrane lipid metabolic process	0.000003928	0.0000000000008566	*Informative*	*Direct*
Biological Process (BP)	organ morphogenesis	0.000001874	0.0000009375	*Informative*	*Direct*

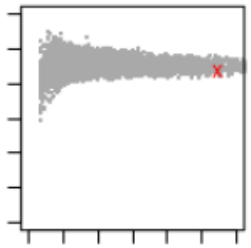
Fig.2. Superfamily search for Alkaline phosphatase –like domain

Global Model Quality Estimation:

QMEAN4 global scores:

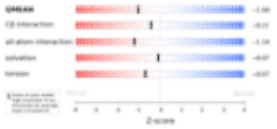
QMEANscore4: 0.693

Estimated absolute model quality:



Z-Score: -1.001

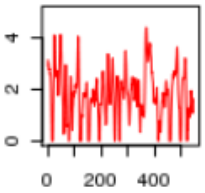
Score components:



Local scores:

Coloring by residue error:

Residue error plot:



QMEAN4 global scores:

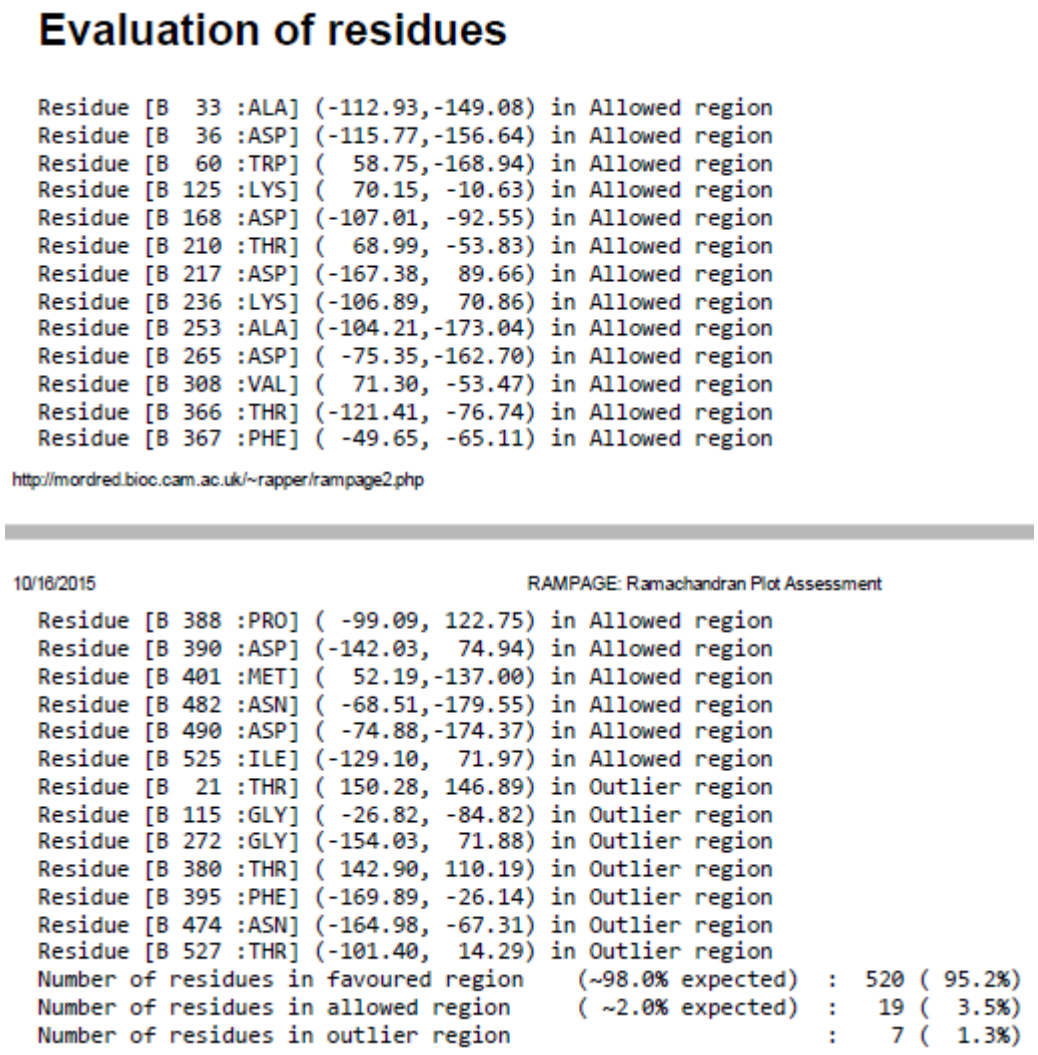
The QMEAN4 score is a composite score consisting of a linear combination of 4 statistical potential terms (estimated model reliability between 0-1). The pseudo-energies of the contributing terms are given below together with their Z-scores with respect to scores obtained for high-resolution experimental structures of similar size solved by X-ray crystallography:

Scoring function term	Raw score	Z-score
C_beta interaction energy	-169.70	-0.41
All-atom pairwise energy	-10407.03	-1.18
Solvation energy	-48.97	-0.07
Torsion angle energy	-128.18	-0.67
QMEAN4 score	0.693	-1.00

If you publish results from QMEAN, please cite the following paper:

Benkert P, Biasini M, Schwede T. (2011). "Toward the estimation of the absolute quality of individual protein structure models." Bioinformatics, 27(3):343-50.

Fig.3. QMEAN server results showing validation of modeled three dimensional structure of new identified Os01g0817700 putative lipase gene



Evaluation of residues

```
Residue [B  33 :ALA] (-112.93,-149.08) in Allowed region
Residue [B  36 :ASP] (-115.77,-156.64) in Allowed region
Residue [B  60 :TRP] (  58.75,-168.94) in Allowed region
Residue [B 125 :LYS] (  70.15, -10.63) in Allowed region
Residue [B 168 :ASP] (-107.01, -92.55) in Allowed region
Residue [B 210 :THR] (  68.99, -53.83) in Allowed region
Residue [B 217 :ASP] (-167.38,  89.66) in Allowed region
Residue [B 236 :LYS] (-106.89,  70.86) in Allowed region
Residue [B 253 :ALA] (-104.21,-173.04) in Allowed region
Residue [B 265 :ASP] ( -75.35,-162.70) in Allowed region
Residue [B 308 :VAL] (  71.30, -53.47) in Allowed region
Residue [B 366 :THR] (-121.41, -76.74) in Allowed region
Residue [B 367 :PHE] ( -49.65, -65.11) in Allowed region
```

http://mordred.bioc.cam.ac.uk/~rapper/rampage2.php

10/16/2015 RAMPAGE: Ramachandran Plot Assessment

```
Residue [B 388 :PRO] ( -99.09, 122.75) in Allowed region
Residue [B 390 :ASP] (-142.03,  74.94) in Allowed region
Residue [B 401 :MET] (  52.19,-137.00) in Allowed region
Residue [B 482 :ASN] ( -68.51,-179.55) in Allowed region
Residue [B 490 :ASP] ( -74.88,-174.37) in Allowed region
Residue [B 525 :ILE] (-129.10,  71.97) in Allowed region
Residue [B  21 :THR] ( 150.28, 146.89) in Outlier region
Residue [B 115 :GLY] ( -26.82, -84.82) in Outlier region
Residue [B 272 :GLY] (-154.03,  71.88) in Outlier region
Residue [B 380 :THR] ( 142.90, 110.19) in Outlier region
Residue [B 395 :PHE] (-169.89, -26.14) in Outlier region
Residue [B 474 :ASN] (-164.98, -67.31) in Outlier region
Residue [B 527 :THR] (-101.40,  14.29) in Outlier region
Number of residues in favoured region    (~98.0% expected)  :  520 ( 95.2%)
Number of residues in allowed region     ( ~2.0% expected)  :   19 (  3.5%)
Number of residues in outlier region                        :    7 (  1.3%)
```

Fig.4. RAMPAGE analysis of modeled three dimensional structure of new identified Os01g0817700 putative lipase gene

Chapter 4

Additional file 2:

List of unigenes involved in Fatty acid biosynthesis and catabolism

Gene name	Gene abbreviation	Gene Annotation	Feature ID
Acetyl-coA carboxylase	ACC2	acetyl-CoA carboxylase, putative, expressed	LOC_Os05g22940
Acetyl-coA carboxylase	ACC1	acetyl-CoA carboxylase, putative, expressed	LOC_Os10g21910
Acyl-CoA Thioesterase	ACT2	acyl-CoA thioesterase 2, putative, expressed	LOC_Os04g47120
Acyl-CoA Thioesterase	ACT10	acyl-coenzyme A thioesterase 10, mitochondrial precursor, putative, expressed	LOC_Os09g34190
Acyl-CoA oxidase	ACX	acyl-CoA oxidase, putative, expressed	LOC_Os06g23760
Desaturase	SAD1	Stearoyl-ACP desaturase 1 , chloroplast precursor, putative, expressed	LOC_Os01g65830
Desaturase	SAD2	Stearoyl-ACP desaturase 2, chloroplast precursor, putative, expressed	LOC_Os01g69080
Desaturase	FAD2-1	fatty acid desaturase, putative, expressed	LOC_Os02g48560
Desaturase	FAD7	omega-3 fatty acid desaturase, chloroplast precursor, putative, expressed	LOC_Os03g18070
Desaturase	SAD4	Stearoyl-ACP desaturase 4 , chloroplast precursor, putative, expressed	LOC_Os03g30950
Desaturase		acyl-desaturase, chloroplast precursor, putative, expressed	LOC_Os03g53010
Desaturase	SAD5	Stearoyl-ACP desaturase 5, chloroplast precursor, putative, expressed	LOC_Os04g31070
Desaturase		acyl-desaturase, chloroplast precursor, putative,	LOC_Os06g30780

		expressed	
Desaturase		zeta-carotene desaturase, chloroplast/chromoplast precursor, putative, expressed	LOC_Os07g10490
Desaturase	FAD2-3	fatty acid desaturase, putative, expressed	LOC_Os07g23410
Desaturase	FAD2-2	fatty acid desaturase, putative, expressed	LOC_Os07g23430
Desaturase	FAD7	fatty acid desaturase, putative, expressed	LOC_Os07g49310
Desaturase	FAD2-4	fatty acid desaturase, putative, expressed	LOC_Os07g23390
Desaturase	SAD6	Stearoyl-ACP desaturase 6 , chloroplast precursor, putative, expressed	LOC_Os08g09950
Desaturase	SAD7	Stearoyl-ACP desaturase 7, chloroplast precursor, putative, expressed	LOC_Os08g10010
Desaturase	FAD6	fatty acid desaturase, putative, expressed	LOC_Os08g34220
Desaturase		desaturase/cytochrome b5 protein, putative, expressed	LOC_Os09g16920
Desaturase		omega-3 fatty acid desaturase	LOC_Os11g01340
Desaturase		omega-3 fatty acid desaturase	LOC_Os12g01370
Beta-ketoacyl-ACP reductase	EAR1	enoyl-acyl-carrier-protein reductase NADH, chloroplast precursor, expressed	LOC_Os08g23810
Beta-ketoacyl-ACP reductase	EAR2	enoyl-acyl-carrier-protein reductase NADH, chloroplast precursor, expressed	LOC_Os09g10600
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os01g47350
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os01g54860
Enoyl-CoA hydratase	ECH1	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os01g70090
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os02g43710
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os02g43720

Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os03g19680
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os05g45300
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os06g39344
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os10g40540
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os10g42220
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, expressed	LOC_Os11g17440
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, expressed	LOC_Os11g17580
Enoyl-CoA hydratase	ECH	enoyl-CoA hydratase/isomerase family protein, putative, expressed	LOC_Os12g16350
3s-hydroxyacyl-CoA dehydrogenase	HACDH	3-hydroxyacyl-CoA dehydrogenase, putative, expressed	LOC_Os01g24680
3s-hydroxyacyl-CoA dehydrogenase	HACDH	3-hydroxyacyl-CoA dehydrogenase, putative, expressed	LOC_Os02g17390
3s-hydroxyacyl-CoA dehydrogenase	HACDH	3-hydroxyacyl-CoA dehydrogenase, putative, expressed	LOC_Os05g06300
3s-hydroxyacyl-CoA dehydrogenase	HACDH	3-hydroxyacyl-CoA dehydrogenase, putative, expressed	LOC_Os05g29880
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os01g34560
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os02g11070

Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os02g49920
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os02g56860
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g06700
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g06705
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g08360
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g12030
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g13630
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g14170
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g26530
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os03g26620
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os04g02640
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os05g49290
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os05g49900
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os06g14810
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os06g15020
Beta-ketoacyl-	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os06g15170

ACP synthase			
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os06g15250
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os06g39750
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os07g15190
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os09g19650
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os09g34930
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os10g07010
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os10g28060
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os10g33370
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os11g37900
Beta-ketoacyl-ACP synthase	KAS	3-ketoacyl-CoA synthase, putative, expressed	LOC_Os04g36800
Ketoacyl-CoA Thiolase	KAT	3-ketoacyl-CoA thiolase, peroxisomal precursor, putative, expressed	LOC_Os02g57260
Ketoacyl-CoA Thiolase	KAT2	3-ketoacyl-CoA thiolase, peroxisomal precursor, putative, expressed	LOC_Os10g31950
Long-chain acyl-CoA synthetase	LACS	acyl-CoA synthetase protein, putative, expressed	LOC_Os05g25310
Long-chain acyl-CoA synthetase	LACS6	acyl-CoA synthetase protein, putative, expressed	LOC_Os11g04980
Long-chain acyl-CoA synthetase	LACS4	Long-chain-fatty-acid-CoA ligase 4, putative, expressed	LOC_Os11g06880

Long-chain acyl-CoA synthetase	LACS	acyl-CoA synthetase protein, putative, expressed	LOC_Os12g04990
Long-chain acyl-CoA synthetase	LACS9	acyl-CoA synthetase protein, putative, expressed	LOC_Os12g07110
Malonyl-CoA-ACP transacylase	MCMT	malonyl CoA-acyl carrier protein transacylase, mitochondrial precursor, putative, expressed	LOC_Os03g18590
acyl-ACPthioesterase		acyl-protein thioesterase, putative, expressed	LOC_Os01g07960
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os01g12910
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os01g12920
acyl-ACPthioesterase		acyl-ACP thioesterase, putative, expressed	LOC_Os01g31760
acyl-ACPthioesterase		OsPOP2 - Putative ProlylOligopeptidase homologue, expressed	LOC_Os01g42690
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os01g65950
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os02g32200
acyl-ACPthioesterase		myristoyl-acyl carrier protein thioesterase, chloroplast precursor, putative, expressed	LOC_Os02g43090
acyl-ACPthioesterase		palmitoyl-protein thioesterase 1 precursor, putative, expressed	LOC_Os03g01150
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os03g48480
acyl-ACPthioesterase		acyl-protein thioesterase, putative, expressed	LOC_Os04g09540
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os04g35590
acyl-		acyl-CoA thioesterase 2, putative, expressed	LOC_Os04g47120

ACPthioesterase			
acyl-ACPthioesterase		acyl-protein thioesterase, putative, expressed	LOC_Os04g57370
acyl-ACPthioesterase		acyl-protein thioesterase, putative, expressed	LOC_Os04g57380
acyl-ACPthioesterase		acyl-protein thioesterase, putative, expressed	LOC_Os04g57390
acyl-ACPthioesterase		acyl-protein thioesterase, putative, expressed	LOC_Os05g51050
acyl-ACPthioesterase	FATB	myristoyl-acyl carrier protein thioesterase, chloroplast precursor, putative, expressed	LOC_Os06g05130
acyl-ACPthioesterase		myristoyl-acyl carrier protein thioesterase, chloroplast precursor, putative, expressed	LOC_Os06g39520
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os07g27870
acyl-ACPthioesterase		thioesterase family protein, putative, expressed	LOC_Os07g27960
acyl-ACPthioesterase	FATA	acyl-ACP thioesterase, putative, expressed	LOC_Os09g32760
acyl-ACPthioesterase		acyl-coenzyme A thioesterase 10, mitochondrial precursor, putative, expressed	LOC_Os09g34190
acyl-ACPthioesterase		palmitoyl-protein thioesterase 1 precursor, putative, expressed	LOC_Os10g41340
acyl-ACPthioesterase		myristoyl-acyl carrier protein thioesterase, chloroplast precursor, putative, expressed	LOC_Os11g43820

Additional file 3: List of unigenes involved in TAG acid biosynthesis and catabolism.

Gene name	Gene abbreviation	Gene Annotation	Feature ID
Caleosin		caleosin related protein, putative, expressed	LOC_Os02g50150
Caleosin		caleosin related protein, putative, expressed	LOC_Os02g50140
Caleosin		caleosin related protein, putative, expressed	LOC_Os02g50174
Caleosin		caleosin related protein, putative, expressed	LOC_Os03g12230
Caleosin	Clo 1	caleosin related protein, putative, expressed	LOC_Os04g43170
Caleosin	Clo 2	caleosin related protein, putative, expressed	LOC_Os04g43200
Caleosin		caleosin related protein, putative, expressed	LOC_Os06g14324
Caleosin		caleosin related protein, putative, expressed	LOC_Os06g14350
Caleosin		caleosin related protein, putative, expressed	LOC_Os06g14370
Diacylglycerol O-acyltransferase	DGAT1a	diacylglycerol O-acyltransferase, putative, expressed	LOC_Os02g48350
Diacylglycerol O-acyltransferase	DGAT1b	diacylglycerol O-acyltransferase, putative, expressed	LOC_Os06g36800
Diacylglycerol O-acyltransferase	DGAT2	diacylglycerol O-acyltransferase, putative, expressed	LOC_Os06g22080
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os01g57350
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os01g57420
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os02g54650
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os03g03400
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os03g31180
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os04g45800
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os04g54200
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os07g37580
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os08g05650
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os08g08110
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os08g15090
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os10g37280
Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os12g12260

Glycerol kinase (GK)	GK	diacylglycerol kinase, putative, expressed	LOC_Os12g38780
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os01g14900
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os01g19390
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os01g22560
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os01g22570
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os01g26000
Glycerol 3-phosphate O-acyltransferase	GPAT1	glycerol-3-phosphate acyltransferase 1, putative, expressed	LOC_Os01g44069
Glycerol 3-phosphate O-acyltransferase	GPAT2	glycerol-3-phosphate acyltransferase 2, putative, expressed	LOC_Os01g63580
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os02g02340
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os03g07060
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os03g52570
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os03g61720
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os05g20100
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os05g37600
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os05g38350
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os10g27330
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os10g41070
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os11g45400
Glycerol 3-phosphate O-acyltransferase	GPAT	glycerol-3-phosphate acyltransferase, putative, expressed	LOC_Os12g37600
Lipase		phospholipase D, putative, expressed	LOC_Os01g07760
Lipase		lipase/lipooxygenase, PLAT/LH2, putative, expressed	LOC_Os01g07780
Lipase		monoglyceride lipase, putative, expressed	LOC_Os01g10250
Lipase		GDSL-like Lipase/Acylhydrolase family protein, expressed	LOC_Os01g11760
Lipase		GDSL-like Lipase/Acylhydrolase family protein, expressed	LOC_Os01g11790
Lipase		triacylglycerol lipase, putative, expressed	LOC_Os01g14080
Lipase		triacylglycerol lipase, putative, expressed	LOC_Os01g15000
Lipase		triacylglycerol lipase, putative, expressed	LOC_Os01g20840
Lipase		phospholipase D. Active site motif family protein, expressed	LOC_Os01g20860
Lipase		monoglyceride lipase, putative, expressed	LOC_Os01g21560

Lipase	carboxylic ester hydrolase/ lipase, putative, expressed	LOC_Os01g33784
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g43140
Lipase	lipase 1, putative	LOC_Os01g46090
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g46240
Lipase	lipase-like protein, putative, expressed	LOC_Os01g46250
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g46290
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g46370
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g47610
Lipase	monoglyceride lipase, putative, expressed	LOC_Os01g49380
Lipase	esterase/lipase/thioesterase, putative, expressed	LOC_Os01g49510
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g51360
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g52180
Lipase	patatin-like phospholipase family protein, expressed	LOC_Os01g55650
Lipase	monoglyceride lipase, putative, expressed	LOC_Os01g62010
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g67420
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g67430
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os01g67450
Lipase	lipase precursor, putative, expressed	LOC_Os01g71010
Lipase	triacylglycerol lipase, putative	LOC_Os01g73740
Lipase	esterase/lipase/thioesterase family active site protein, putative, expressed	LOC_Os02g03720
Lipase	esterase/lipase/thioesterase, putative, expressed	LOC_Os02g09770
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os02g18480
Lipase	lipase/hydrolase, putative, expressed	LOC_Os02g18954
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os02g28040
Lipase	esterase/lipase/thioesterase, putative	LOC_Os02g31200
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os02g42170
Lipase	triacylglycerol lipase like protein, putative, expressed	LOC_Os02g43700
Lipase	GSDL-motif lipase, putative, expressed	LOC_Os02g44860

Lipase	GDSL-like Lipase/Acylhydrolase family protein, expressed	LOC_Os02g50000
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os02g52830
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os02g54010
Lipase	esterase/lipase/thioesterase, putative, expressed	LOC_Os02g55330
Lipase	phospholipase A2, putative, expressed	LOC_Os02g58500
Lipase	phospholipase D beta 1, putative, expressed	LOC_Os03g02740
Lipase	phosphoinositide-specific phospholipase C, putative	LOC_Os03g18000
Lipase	phosphoinositide-specific phospholipase C, putative, expressed	LOC_Os03g18010
Lipase	triacylglycerol Lipase, putative, expressed	LOC_Os03g22670
Lipase	lipase 2, putative	LOC_Os03g25000
Lipase	phospholipase D alpha 1, putative, expressed	LOC_Os03g27370
Lipase	phospholipase C, putative, expressed	LOC_Os03g30130
Lipase	phospholipase A2, putative, expressed	LOC_Os03g50030
Lipase	lipase family protein	LOC_Os03g50410
Lipase	monoglyceride lipase, putative, expressed	LOC_Os03g51010
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os03g61540
Lipase	phospholipase D gamma 3, putative, expressed	LOC_Os03g62410
Lipase	gastric triacylglycerol lipase precursor, putative	LOC_Os04g21160
Lipase	phospholipase C, putative, expressed	LOC_Os04g35100
Lipase	lipase family protein, expressed	LOC_Os04g41200
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os04g43030
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os04g56240
Lipase	phospholipase C, putative, expressed	LOC_Os05g03610
Lipase	lipase, putative, expressed	LOC_Os05g06140
Lipase	phospholipase D alpha 1 precursor, putative, expressed	LOC_Os05g07880
Lipase	lipase/lipooxygenase, PLAT/LH2, putative, expressed	LOC_Os05g07890
Lipase	esterase/lipase/thioesterase family protein, putative, expressed	LOC_Os05g12330
Lipase	phospholipase D p1, putative, expressed	LOC_Os05g29050
Lipase	monoglyceride lipase, putative, expressed	LOC_Os05g29974

Lipase	lipase precursor, putative, expressed	LOC_Os05g30900
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os05g32380
Lipase	lipase, putative, expressed	LOC_Os05g33820
Lipase	GDSL-motif lipase/hydrolase-like protein, putative	LOC_Os05g43110
Lipase	GDSL-like Lipase/Acylhydrolase family protein, expressed	LOC_Os05g49830
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os05g49840
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os06g10850
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os06g40170
Lipase	phospholipase D alpha 2, putative, expressed	LOC_Os06g40180
Lipase	phospholipase D alpha 2, putative, expressed	LOC_Os06g40190
Lipase	phospholipase D alpha 2, putative, expressed	LOC_Os06g42730
Lipase	esterase/lipase/thioesterase, putative, expressed	LOC_Os06g42860
Lipase	gastric triacylglycerol lipase precursor, putative, expressed	LOC_Os06g44060
Lipase	patatin-like phospholipase family protein, expressed	LOC_Os06g46350
Lipase	phospholipase D beta 1, putative, expressed	LOC_Os07g15680
Lipase	triacylglycerol lipase, putative	LOC_Os07g28250
Lipase	patatin-like phospholipase family protein, expressed	LOC_Os07g33670
Lipase	triacylglycerol lipase, putative	LOC_Os07g34400
Lipase	triacylglycerol lipase, putative	LOC_Os07g34420
Lipase	triacylglycerol lipase, putative	LOC_Os07g34440
Lipase	monoglyceride lipase, putative, expressed	LOC_Os07g37840
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os07g39810
Lipase	lipase precursor, putative, expressed	LOC_Os07g47250
Lipase	phosphoinositide-specific phospholipase C, putative, expressed	LOC_Os07g49330
Lipase	triacylglycerol lipase like protein, putative, expressed	LOC_Os08g04800
Lipase	esterase/lipase/thioesterase, putative, expressed	LOC_Os08g06420
Lipase	phospholipase D alpha 1, putative, expressed	LOC_Os08g31060
Lipase	esterase/lipase/thioesterase, putative, expressed	LOC_Os08g38092
Lipase	gastric triacylglycerol lipase precursor, putative, expressed	LOC_Os08g41780

Lipase	gastric triacylglycerol lipase precursor, putative, expressed	LOC_Os09g01590
Lipase	monoglyceride lipase, putative, expressed	LOC_Os09g23150
Lipase	phospholipase D alpha 1, putative, expressed	LOC_Os09g25390
Lipase	phospholipase A2, group IVB isoform 6, putative, expressed	LOC_Os09g31050
Lipase	phospholipase A1, putative, expressed	LOC_Os09g33820
Lipase	phospholipase D delta, putative, expressed	LOC_Os09g37100
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os09g39790
Lipase	phospholipase D beta 1, putative, expressed	LOC_Os10g38060
Lipase	triacylglycerol lipase like protein, putative	LOC_Os10g41270
Lipase	monoglyceride lipase, putative, expressed	LOC_Os11g01040
Lipase	acyltransferase/ carboxylic ester hydrolase/ lipase, putative, expressed	LOC_Os11g03520
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os11g19290
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os11g19340
Lipase	phospholipase A2, putative, expressed	LOC_Os11g34440
Lipase	lipase precursor, putative, expressed	LOC_Os11g43510
Lipase	triacylglycerol lipase, putative	LOC_Os11g43760
Lipase	monoglyceride lipase, putative	LOC_Os12g01030
Lipase	phospholipase, putative, expressed	LOC_Os12g16180
Lipase	lipase/hydrolase, putative, expressed	LOC_Os12g17570
Lipase	esterase/lipase/thioesterase, putative, expressed	LOC_Os12g18860
Lipase	triacylglycerol lipase, putative, expressed	LOC_Os12g36770
Lipase	phospholipase C, putative, expressed	LOC_Os12g37560
Lipase	phospholipase/Carboxylesterase family protein, expressed	LOC_Os12g37630
Lipase	GDSL-motif lipase/hydrolase-like protein, putative, expressed	LOC_Os12g37910
Lipase	patatin-like phospholipase family protein, expressed	LOC_Os12g41720
Lipase	triacylglycerol lipase, putative	LOC_Os12g41970
Lipase	triacylglycerol lipase, putative	LOC_Os12g41980
Lipase	triacylglycerol lipase, putative	LOC_Os12g42010
1-Acyl-sn- glycerol-3-phosphate O-acyltransferase (LPAAT)	1-acyl-sn-glycerol-3-phosphate acyltransferase theta, putative,	LOC_Os02g24340

		expressed	
1-Acyl-sn- glycerol-3-phosphate O-acyltransferase (LPAAT)		1-acyl-sn-glycerol-3-phosphate acyltransferase theta, putative, expressed	LOC_Os01g57360
Oleosin	OLE 18	oleosin, putative, expressed	LOC_Os03g49190
Oleosin		oleosin, putative, expressed	LOC_Os04g32070
Oleosin		oleosin, putative, expressed	LOC_Os04g32080
Oleosin	OLE16	oleosin, putative, expressed	LOC_Os04g46200
Oleosin	OLE5	oleosin, putative, expressed	LOC_Os05g50110
Oleosin	OLE3	oleosin, putative, expressed	LOC_Os06g27910
Oleosin	OLE4	oleosin, putative, expressed	LOC_Os09g15520
Phospholipid:diacylglycerol acyltransferase	PDAT	lecithin cholesterol acyltransferase, putative	LOC_Os01g71800
Phospholipid:diacylglycerol acyltransferase	PDAT	lecithin cholesterol acyltransferase, putative	LOC_Os02g56910
Phospholipid:diacylglycerol acyltransferase	PDAT	lecithin:cholesterol acyltransferase, putative, expressed	LOC_Os02g37654
Phospholipid:diacylglycerol acyltransferase	PDAT	lecithin:cholesterol acyltransferase, putative, expressed	LOC_Os02g37750
Phospholipid:diacylglycerol acyltransferase	PDAT	lecithin:cholesterol acyltransferase, putative, expressed	LOC_Os03g13030
Phospholipid:diacylglycerol acyltransferase	PDAT	lecithin:cholesterol acyltransferase, putative, expressed	LOC_Os09g33820
Phospholipid:diacylglycerol acyltransferase	PDAT	lecithin:cholesterol acyltransferase, putative, expressed	LOC_Os10g08026
Phospholipid:diacylglycerol acyltransferase	PDAT1	lecithin:cholesterol acyltransferase, putative, expressed	LOC_Os09g27210
Phosphotidylcholine:diacylglycerol cholinephosphotransferase	PDCT	expressed protein	LOC_Os06g40500
Putative phosphatidic acid phosphatase	PAP1	phosphatidic acid phosphatase-related, putative, expressed	LOC_Os01g63060
Putative phosphatidic acid phosphatase	PAP2	phosphatidic acid phosphatase-related, putative, expressed	LOC_Os05g21180
Putative phosphatidic acid phosphatase	PAP3	phosphatidic acid phosphatase-related, putative, expressed	LOC_Os05g37910
Steroleosin		steroleosin, putative, expressed	LOC_Os02g30690
Steroleosin		dehydrogenase/reductase, putative, expressed	LOC_Os04g32070
Steroleosin		11-beta-hydroxysteroid dehydrogenase, putative, expressed	LOC_Os04g32080
Steroleosin		steroleosin, putative, expressed	LOC_Os02g30690
Transcription factor	LEC1	Leafy cotyledon1	LOC_Os02g49410

Transcription factor	OsLFL1	O. sativa LEC2 and FUSCA3 Like 1	LOC_Os01g51610
Transcription factor	ABI3	Abscisic acid-insensitive	LOC_Os01g68370
Transcription factor	WRI1	Wrinkled 1	LOC_Os11g03540

Additional file 4: Table S1.

Supplementary Table 1. List of main enzymes involved in fatty acid biosynthesis and degradation

Enzyme	Symbol	EC Number
Fatty acid biosynthesis		
Acetyl-CoA carboxylase	ACC	6.4.1.2
malonyl-CoA-ACP transacylase	MCMT	2.3.1.3.9
Beta-ketoacyl-ACP synthase I	KASI	2.3.1.4.1
Beta-ketoacyl-ACP synthase II	KASII	2.3.1.179
Beta-ketoacyl-ACP synthase III	KASIII	2.3.1.180
Enoyl-ACP reductase (NADH)	EAR	1.3.1.9
Acyl-ACP thioesterase A	Fat A	3.1.2.14
Acyl-ACP thioesterase B	Fat B	3.1.2.14
Fatty acid desaturation		
Δ^{12} (ω6)-Desaturase	FAD	1.14.19.6
Stearoyl-ACP Desaturase	SAD	1.14.19.2
Fatty acid catabolism		
Long-chain acyl-CoA synthetase	LACS	6.2.1.3
Acyl-CoA oxidase	ACX	1.3.3.6
Enoyl-CoA hydratase	ECH	4.2.1.17
3s-hydroxyacyl-CoA dehydrogenase	HACDH	1.1.1.35
Ketoacyl-CoA Thiolase	KAT	2.3.1.16
Acyl-CoA Thioesterase	ACT	3.1.2.2
Enoyl-CoA isomerase	Isom	5.3.3.8

Supplementary Table 2. Expression of four FAD2 genes in six different tissues from rice *Nipponbare* expressed as reads per million reads

Gene name	RNAseq from 20 days leaves	RNAseq from post-emergence inflorescence	RNAseq from anther	RNAseq from 10 daa seed	RNAseq from 25 daa embryo	RNAseq from 25 daa endosperm
FAD2-1	849	655	339	187	725	47
FAD2-2	0	0	0	0	0	0
FAD2-3	37	8.5	3	1	4	0
FAD2-4	0	0	0	0	0	0

daa –days after anthesis

Additional file 5: Gene expression (RPKM) values of affected genes in Lipid biosynthesis pathway.

Feature ID	Fold change	P-value	RPKM (WT1)	RPKM (WT2)	RPKM (WT3)	Means(WT)	RPKM(RNAi1)	RPKM(RNAi2)	RPKM(RNAi3)	Means(RNAi)	Gene abbreviation	DAA
LOC_Os06g22080	-1.31	0.45	51.7	19.92	21.82	31.15	26.17	27.29	17.65	23.7	DGAT2	10
LOC_Os06g22080	-1.71	7.73E-3	25.97	38.03	45.09	36.37	12.48	25.95	16.99	18.47	DGAT2	15
LOC_Os06g22080	-1.16	0.28	45.59	34.62	33.29	37.83	38.45	29.92	32.46	33.61	DGAT2	20
LOC_Os02g50174	1.52	0.27	62.24	41.11	27.04	43.46	60.49	109.79	33.36	67.88	Caleosin	10
LOC_Os02g50174	-1.33	0.04	55.02	70.92	92.81	72.92	38.06	60.01	43.35	47.14	Caleosin	15
LOC_Os02g50174	-1.97	5.02E-3	122.76	109.12	56.37	96.08	62.25	39.86	45.46	49.19	Caleosin	20
LOC_Os03g12230	-1.14	0.42	111.47	72.49	68.1	84.02	76.75	81.91	63.09	73.92	Caleosin	10
LOC_Os03g12230	-1.58	6.60E-3	95.35	88.53	108.89	97.59	31.94	67.8	61.14	53.63	Caleosin	15
LOC_Os03g12230	-1.27	0.5	53.1	65.17	49.21	55.83	29.92	24.02	81.38	45.11	Caleosin	20
LOC_Os04g32080	-1.36	0.42	195.32	87.63	62.73	115.23	83.6	105.9	62.99	84.16	STEROLEOSIN	10
LOC_Os04g32080	-1.36	0.03	170.83	231.1	298.03	233.32	130.15	197.45	117.38	148.33	STEROLEOSIN	15
LOC_Os06g22080	-1.16	0.28	45.59	34.62	33.29	37.83	38.45	29.92	32.46	33.61	STEROLEOSIN	20
LOC_Os02g49410	-1.23	0.42	275.43	131.95	145.74	184.37	160.36	183.05	107.35	150.25	LEC1	10
LOC_Os02g49410	-1.66	3.91E-3	114.09	140.21	190.38	148.22	58.66	116.18	59.47	78.1	LEC1	15
LOC_Os02g49410	-1.44	0.12	177.91	159.5	188.49	175.3	95.62	83.43	195.44	124.83	LEC1	20

Additional file 6: Table S2.

Table S2. Expression of four FAD2 genes in six different tissues from rice *Nipponbare* expressed as reads per million reads

Gene name	RNAseq from 20 days leaves	RNAseq from post-emergence inflorescence	RNAseq from anther	RNAseq from 10 daa seed	RNAseq from 25 daa embryo	RNAseq from 25 daa endosperm
FAD2-1	849	655	339	187	725	47
FAD2-2	0	0	0	0	0	0
FAD2-3	37	8.5	3	1	4	0
FAD2-4	0	0	0	0	0	0

daa –days after anthesis

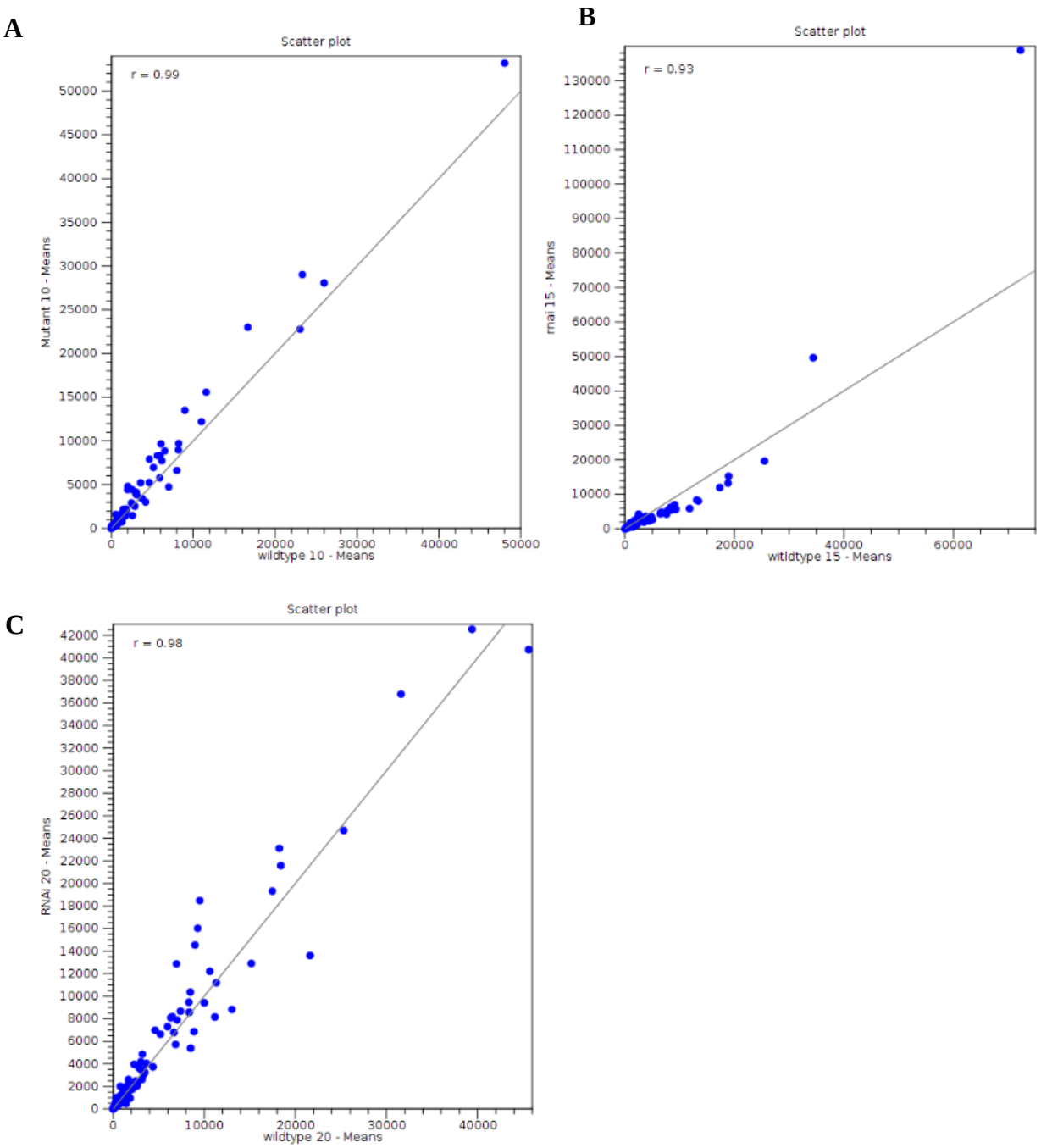


Figure 1. Scatter plot of the expression levels of all genes. Comparison between null segregates (wild type) and *OsFAD2-1* RNAi lines at 10 daa (A), 15 daa (B) and 20 daa (C) stages. The values on the y-axis represent means of the expression levels of *OsFAD2-1* RNAi lines, the values on the x-axis represent the means of the expression levels of null segregates.