Don’t Cry for Me: Evolutionary and Functional Analysis of Two Rice Pollen Allergens, Ory s 1 and Ory s 12

By

Deborah Devis

Thesis by a combination of conventional and publication formats submitted to The University of Adelaide for the degree of Doctor of Philosophy

School of Agriculture, Food and Wine

The University of Adelaide

August 2019

Table of Contents

Abstract…………………………………………………………………………………..i

Declaration………………………………………………………………………………..iii

Acknowledgements…………………………………………………………...…………..iv

STATEMENT OF CONTRIBUTIONS OF JOINTLY AUTHORED MANUSCRIPT……...1

Chapter 1 – Introduction..……..3

1.1 Overview of Thesis…………………………………………………………………..3

1.2 Molecular Features of Grass Allergens and Development of Biotechnological Approaches for Allergy Prevention………………………………….……………………….……5

STATEMENT OF CONTRIBUTIONS OF JOINTLY AUTHORED MANUSCRIPT…….19

Chapter 2 – Evolutionary analysis of two rice pollen allergens, Ory s 1 and Ory s 12 ..21

2.1 Abstract……………………………………………………….……………………21

2.2 Introduction…………………………………………………….….……………..21

2.3 Methods………………………………………………………….…….…………23

2.3.1 Identification of gene homologs……………………………….……….…………2

2.3.2 Construction of phylogenetic trees, secondary structure prediction and expression analysis……………………….…………………………………………………………24

2.3.3 Synteny Analysis ………………………………….…………………….…………….25

2.3.4 Domain identification and 3D modelling……………………………………………25

2.4 Results…………………………………………………………………………………25

2.4.1 Evolutionary analysis of *OsEXPB* homologs……………………………...…………..25

2.4.2 The evolution of *OsEXPBs* in rice………………………………………...…………27

2.4.3 Structural analysis of OsEXPB homologs……………………………….…...……29

2.4.4 Evolutionary analysis of OsPRFA homologs………………………………..…….30

2.4.5 Structural analysis of OsPRF homologs………………………………………..31

2.5 Discussion………………………………………………………….’…..………...33

2.5.1 The evolution of *OsEXPB* involved both ancient and recent duplications……..….…..33

2.5.2 Copy number of *OsEXPB* is not a result of domestication……………………….……34

2.5.3 Loss of introns may promote translational efficiency…………………………......35

2.5.4 *OsEXPB1/10* are expressed specifically in reproductive tissue……………….….…36

2.5.5 Reproductive and vegetative profilins arose prior to monocot/dicot divergence….…..37

2.5.6 Putative IgE interaction sites appear on the surface of Profilins……………….……...38

2.5.7 Homologs of Ory s 1 and Ory s 12 showed varied evolutionary patterns…………..…40

STATEMENT OF CONTRIBUTIONS OF JOINTLY AUTHORED MANUSCRIPT……..41

Chapter 3 – Creation of mutants of rice pollen allergen genes using the genome editing system, CRISPR/Cas9……………………………………………………………43

3.1 Abstract...………………………………………………………………………...43

3.2 Introduction……………………………………………………………………43

3.3 Materials and Methods……………………………………………………………..……45

3.3.1 CRISPR target sequence design………………………………………………..…...45

3.3.2 CRISPR-Cas9 plasmid construction ………………………………………..…...46

3.3.3 Induction of calli…………………………………………………………..……..47

3.3.4 Cocultivation of calli with CRISPR/Cas9 infected Agrobacterium……………..……47

3.3.5 Selection and regeneration……………………………….……………..…..48

3.3.6 Growth conditions……………………………………………………..……48

3.3.7 Identification and genotyping of transgenic lines……………………………..….49

3.4 Results and discussion……………………………………………………………48

3.4.1 Creation of DEBV *OsEXPB* mutants generated with CRISPR/Cas9…………………. 48

3.4.2 Generation of *OsPRFA* mutants through CRISPR/Cas9 deletions……………..…...50

3.4.3 Pollen viability of T0 mutants………………………………………...….52

3.4.4 No roots were regenerated after *OsPRFLP04* transformation………………………….53

3.5 Conclusion ………………………………………………………………………….53

STATEMENT OF CONTRIBUTIONS OF JOINTLY AUTHORED MANUSCRIPT……..55

Chapter 4 – Functional assessment of two rice pollen genes, *OsEXPB* and *OsPRFA*

4.1 Abstract…………………………………………………………………………………...57

4.2 Introduction……………………………………………………………………………57

4.3 Materials and Methods……………………………………………………..…………60

4.3.1 Growth conditions……………………………………………………………..……60

4.3.2 Pollen and seed fertility assessment ……………………………………...…………60

4.3.3 Pollen tube growth……………………………………………………………………...60

4.3.4 Scanning electron microscopy…………………………………………………….61

4.3.5 RNA extraction and cDNA synthesis…………………………………………………..61

4.3.6 Expression analysis and qRT-PCR…………………………………………….……….61

4.3.7 Transient expression of GFP…………………………………………...………62

4.3.8 Actin binding Assay……………………………………………………...……….63

4.3.9 Protein expression………………………………………………………...………63

4.3.10 *In situ* hybridization………………………………………………………...……...64

4.3.11 Protein extraction from plants……………………………………………...……64

4.3.12 Immunoassays…………………………………………………………………...65

4.4 Results……………………………………………………………………………………66

4.4.1 *OsEXPB* is expressed in late stage pollen development but does not affect pollen fertility……………………………………………………………………………………..66

4.4.2 The pollen tube germination of *expb* was delayed……………………………..……67

4.4.3 Mutant *expb* had lower seed-setting rates…………………………………...……69

4.4.4 OsEXPB10-GFP is localised to the pollen tube tip…………………………..……..70

4.4.5 OsPRFA is expressed in late stage pollen development but does not affect pollen fertility……………………………………………………………………………………..71

4.4.6 Mutant *prf* showed aberrant and delayed pollen tube growth……………………..…...71

4.4.7 *prfa* had low seed setting rates…………………………………………………..…..73

4.4.8 OsPRFAa-GFP was evenly distributed in the pollen grain and tube, but prevented tube growth……………………………………………………………………………...74

4.4.9 Actin filaments in *prf* were over-polymerised in pollen………………………………75

4.4.10 *Osexpb-1* mutant pollen showed decreased binding to a pool of allergic human sera………………………………………………………………...……………....76

4.4.11 *Osexpb-1* mutant pollen showed decreased binding to a pool of allergic human sera .76

4.5 Discussion……………………………………………………………………….………..78

4.5.1 *OsEXPB1/10* function is similar to maize homologs…………………………………..78

4.5.2 *OsEXPB1/10* mutants set seeds, but were more susceptible to environmental conditions……………………………………………………………………..…..……..78

4.5.3 Generation of functioning *osexpb* triple homozygous rice plants…………………..……80

4.5.4 Potential function of the Pollen Allergen 1 domain……………………………..…..…80

4.5.5 Function of *OsPRFA* is similar to other profilins……………………………….…..81

4.5.6 Overexpression of *PRF* may also lead to aberrant pollen tube growth ……………..…82

4.5.7 Mutants of *OsEXPB* and *OsPRFA* show similar pollen tube germination phenotypes but not localisation or overexpression ………………………………………………..……….83

4.5.8 Pollen proteins are functionally important and may not have evolved to be allergenic………………………………………………………...………….……….83

4.5.9 Towards the generation of viable hypoallergenic rice plants…………………..………83

4.5.10 Conclusion……………………………………………………………………….85

Chapter 5 – Summary of Thesis…………………………………………………….87

5.1 Introduction…………………………………………………………………….87

5.2 Outcome of Aims………………………………...…………………………………87

5.3 Future Directions………………………………………………………………….89

Glossary………………………………………………………………….………..91

Supplementary Figures………………………………………………………………..92

Supplementary Tables……………………………………………………………….95

References from Chapters 2-5…………………………………………………….117

Thesis Abstract

Hay fever is a globally relevant disease that is characterised by sneezing, itchiness of eyes and general discomfort. Some proteins in pollen react with human Immunoglobulin E (IgE) to trigger an immune response, and these proteins are called allergens. Allergens are abundant in grass pollen and, due to grass being wind pollinated, is difficult to avoid. Despite the high prevalence of hay fever and grass pollen allergy, little is known about how these allergen genes evolved, what the specific function of these genes are and how they pertain to pollen development, and whether there is potential for the development of a hypoallergenic derivative in grasses.

To investigate how allergen genes evolved in grasses, I conducted a phylogenetic and structural analysis of allergens in the model grass, rice (*Oryza sativa*). Two pollen allergens and allergen families were identified; Ory s 1 belonged to the group 1 grass pollen allergen family homologous to β-expansins, and Ory s 12 belonged to the group 12 allergen family homologous to profilins. Group 1 genes were found only in monocots and appeared to have evolved recently and rapidly following speciation, and most grass species contained 3-4 copied of these genes. However, group 12 genes arose from an ancient duplication that resulted in reproductively expressed and vegetatively expressed copies that were found in most monocots and dicots. No correlation was found between the evolutionary patterns of Ory s 1 and Ory s 12, suggesting that these two gene families evolved allergenicity independently of each other.

To understand the function of allergen genes, 3 mutant Ory s 1 and 3 mutant Ory s 12 lines were generated using the CRISPR-Cas9 system. These mutant lines were assessed for pollen fertility, pollen tube germination and growth rates, and seed fertility. All mutant lines had normal pollen fertility, but showed low pollen tube germination rates and slow pollen tube

development, leading to lowered seed setting. To understand this further, wild type protein was transiently over-expressed with GFP (green fluorescent protein) in pollen grains. Ory s 1 caused the pollen tubes to grow rapidly compared to the wild type, suggesting Ory s 1 is involved in efficient pollen tube growth to promote rapid fertilisation of the ovule. Overexpression of Ory s 12 lead to aborted pollen tubes, similar to the mutant phenotype, suggesting that this profilin may be involved in preventing over-polymerization or under-polymerization of actin filaments, both of which may prevent efficient pollen tube growth. Interestingly, these two allergens have highly specific, but different, roles in pollen tube growth, suggesting allergenicity is not a result of shared function but of shared abundance in the mature pollen grain.

To investigate the potential for a hypoallergenic derivative, the mutant lines were assessed for seed fertility and assayed against a pool of pollen allergy patients. Ory s 12 mutants showed low seed setting, and are unlikely to be suitable as a semi-viable plant. However, homozygous mutant Ory s 1 plants showed higher seed setting rates under some environmental conditions and maintained some, if not all, viability. The homozygous mutant pollen proteins were assayed against human allergen patients from Queensland, and showed almost no binding to human sera, unlike pollen proteins from the wild type and mutant Ory s 12 lines, suggesting the Ory s 1 mutant is both a viable and hypoallergenic rice derivative.

In summary, Ory s 1 and Ory s 12 have evolved independently of each other, share similar localisation in the mature pollen grain and have highly specific roles in pollen tube development. However, a homozygous Ory s 1 mutant showed decreased binding to human allergy patient sera, suggesting that the construction of viable hypoallergenic grass derivatives is possible.

Thesis Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the jointaward of this degree. I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time. I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Deborah Devis 29/07/2019

Acknowledgements

I wish to acknowledge the wonderful help of my supervisors, Dabing Zhang, Gang li, Janet Davies and Iain Searle. Dabing, your presence was instrumental in helping me to develop into a capable and independent researcher. Gang, your help with all my specific little problems in the lab and helping me with techniques and analysis. Janet, your lovely research group and warm welcome to Queensland was the best possible conclusion to my PhD. Iain, your unending emotional support and willingness to talk through thought experiments kept me sane.

I would like to thank my lab group for their support. Cindy, thank you for always being there to share our woes and curries. Xiujuan, Reyhaneh, Hui and Carrie, for always brightening my day. Nico, thank you for introducing me to D&D. Ann, for being a rock and keeping me on track. I also appreciate the help of all the people around this campus. Many thanks to Kenneth Chambers, Gwenda Mayo, Matthew Tucker and all my wonderful plant genomics friends. To Shanghai Jiao Tong for giving me a wonderful research experience. Special thanks to Bianca Warnock for literally showing me my passion.

A very bright thank you to Callie, Bashir, Nidus, Stoneflow and the teller of our story. You will never understand how instrumental our adventures were to bringing me out of the dark and making me the best bard I can be. Matt, thank you for helping me become the true Prof Rimu, PhD. Rachael and Skye, for always listening to me whine.

Finally, thank you to my parents, siblings and in-laws who have helped me to see beyond the data and keep me going. Most of all to my husband, for your unceasing support and love when

I couldn't hold myself up. Your resilience and intelligence are as much a part of this thesis as I am.

Statement of Authorship

Title of Paper	Molecular features of grass allergens and development of biotechnological approaches for allergy prevention
Publication Status	☑ PublishedAccepted for ☐ Publication ☐ Unpublished and ☐ Unsubmitted work written in Submitted for Publicationmanuscript style
Publication Details	Devis, D.L., Davies, J.M., Zhang, D.B., 2017. Molecular features of grass allergens and development of biotechnological approaches for allergy prevention. Biotechnology Advances. 35, 545-56.

Principal Author

Name of Principal Author (Candidate)	Deborah Devis		
Contribution to the Paper	Wrote, researched edited paper		
Overall percentage (%)	80		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25/07/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Janet Davies		
Contribution to the Paper	Helped with writing and framing of paper, edited		
Signature		Date	19.06.2019

Contribution to the Paper	Helped with writing and framing of paper, edited		
Signature		Date	09/08/2019

Chapter 1 – Introduction

1.1 Overview of Thesis

Hay fever is a prevalent disease caused by pollen allergy, which triggers an immune response resulting in coughing, sneezing, itchiness of the eyes and impaired sleep. Grass pollen can often contain potent allergens, proteins that trigger an immune response, and cannot be avoided when airborne. Little is currently know about the evolution and function of these allergens in grass pollen, but further understanding may lead to the development of hypoallergenic plants, easier identification of putative allergens, or improved selective breeding.

Rice is a globally important grass crop and is also an excellent grass model. Rice was selected for this project as a model organism, but rice pollen allergy has also been reported. Likewise, many grass pollen allergens are cross reactive and allergies to one grass species will often result in allergies to other grass species. Two rice pollen allergens have been published on World Health Organization/International Union of Immunological Societies (WHO/IUIS), and these two allergens were chosen for further studies.

The analysis of literature lead to the following research questions:

1. How did Ory s 1 and Ory s 12 evolve and did they develop allergenicity or potency during domestication of rice?
2. Do allergen genes from the same family maintain sequences or structures that may cause allergenicity?
3. Can pollen allergen genes be modified to develop a fertile, hypoallergenic variant?
4. What is the functional role of pollen allergens in pollen, and are they essential?

The aims of this project had 4 main objectives:

1. **Sequence analysis and evolutionary analysis of rice pollen allergens Ory s 1 and Ory s 12**

a. Phylogenetic analysis

b. Protein Modelling

c. Expression analysis

d. Model development

2. **Mutation of Ory s 1 and Ory s 12 using CRISPR/Cas9**

a. Gene function knockout

b. Transformation into rice using agrobacterium

c. Generation of mutant line

3. **Functional analysis of mutant *ory s 1* and *ory s 12* pollen and anthers**

a. Assessment of pollen viability and fertility

b. Assessment of role in pollen tube development

4. **Allergenic assessment of mutant lines**

a. Analysis of protein abundance in mutants

b. Assessment of mutant pollen binding to allergy patient sera

This Thesis contains 5 chapters

Chapter 1: Introduction

Chapter 2: Evolutionary analysis of two rice pollen allergens, Ory s 1 and Ory s 12

Chapter 3: CRISPR/Cas9 mutation and transformation of rice pollen allergen homologs, *OsEXPB* and *OsPRF*

Chapter 4: Functional analysis of two rice pollen allergens, Ory s 1 and Ory s 12

Chapter 5: Conclusions and future applications

ARTICLE IN PRESS

Biotechnology Advances xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Biotechnology Advances

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

Research review paper

Molecular features of grass allergens and development of biotechnological approaches for allergy prevention

Deborah L. Devis[a], Janet M. Davies[b], Dabing Zhang[a,c,*]

[a] Joint International Research Laboratory of Metabolic & Developmental Sciences, University of Adelaide-Shanghai Jiao Tong University Joint Centre for Agriculture and Health, School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Urrbrae, South Australia 5064, Australia [b] School of BioMedical Sciences, Institute of Health and Biomedical Innovation, Queensland University of Technology, South Brisbane, Queensland 4101, Australia [c] School of Life Sciences and Biotechnology,

Shanghai Jiao Tong University, China

ARTICLEINFO

Keywords:
Poaceae
Allergy
Grain allergen Pollen allergen
Allergen-specific immunotherapy
Allergen evolution Allergen function Transgenic grass
Genetic engineering

ABSTRACT

Allergic diseases are characterized by elevated allergen-specific IgE and excessive inflammatory cell responses. Among the reported plant allergens, grass pollen and grain allergens, derived from agriculturally important members of the Poaceae family such as rice, wheat and barley, are the most dominant and difficult to prevent. Although many allergen homologs have been predicted from species such as wheat and timothy grass, fundamental aspects such as the evolution and function of plant pollen allergens remain largely unclear. With the development of genetic engineering and genomics, more primary sequences, functions and structures of plant allergens have been uncovered, and molecular component-based allergen-specific immunotherapies are being developed. In this review, we aim to provide an update on (i) the distribution and importance of pollen and grain allergens of the Poaceae family, (ii) the origin and evolution, and functional aspects of plant pollen allergens, (iii) developments of allergen-specific immunotherapy for pollen allergy using biotechnology and (iv) development of less allergenic plants using gene engineering techniques. We also discuss future trends in revealing fundamental aspects of grass pollen allergens and possible biotechnological approaches to reduce the amount of pollen allergens in grasses.

1. Introduction

Allergies are global diseases that affect a worldwide population. Among the allergic diseases, hay fever is a widespread allergic upper respiratory condition triggered by airborne pollen allergens that activate an Immunoglobulin E (IgE) response, leading to immediate release of inflammatory mediators such as histamines. Pollen allergen exposure can also exacerbate asthma in susceptible individuals (Aalberse, 2000; Bousquet et al., 2008; Davies, 2014; Greiner et al., 2011; Scala et al., 2010). Allergic diseases affect up to 20–30% of the world population, including children as young as 3, and are associated with disrupted sleep, impaired work performance and lower life quality (Andersson and Lidholm, 2003; Bousquet et al., 2008; Dykewicz and Hamilos, 2010; Suphioglu et al., 1992; Wheatley and Togias, 2015). The severity of allergic rhinitis varies between locations and seasons due to environmental effects and pollen production time (Beggs et al., 2015; Buters et al., 2015; Davies et al., 2012; Dykewicz and Hamilos, 2010; Nony et al., 2015). Genetic and environmental variations between individual plants, as well as differences between allergen proteins in various plants species, particularly grasses (Poaceae family), impact allergen exposure, sensitivity to allergens and symptoms of patients (Aiubi et al., 2015; Amardip et al., 2014; Buters et al., 2015; Y. Chen et al., 2016; Davies, 2014).

Allergic sensitization to aeroallergens frequently leads to the production of allergen-specific IgE antibodies in susceptible people (Moneret-Vautrin, 1997; Moneret-Vautrin and Kanny, 2007). IgE typically recognises motifs or conformational epitope structures on the surface of allergen proteins, which contain diverse and complex motifs, making it difficult to attribute IgE binding to any single linear peptide sequence (Adachi et al., 1993). Upon secondary exposure to allergens in sensitized individuals, IgE-mediates an immediate allergic reaction, prompting symptoms ranging from itching of the eyes, sneezing, and rashes, to severe anaphylaxis (Ramesh, 2008). Allergen abundance can be influenced by the environment and genetic predisposition of the source organism, and severity of symptoms can also be influenced by the genetic predisposition of the patient (Campbell and Mehr, 2015; Y. Chen et al., 2016; Fuhrmann et al., 2016). Allergens can be present in a variety of animal byproducts, cereal grains, and pollen grains (Y. Chen et al., 2016;

Abbreviations: AIT, allergen-specific immunotherapy; IgE, Immunoglobulin E; IgG, Immunoglobulin G; RNAi, RNA interference

* Corresponding author at: School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, China. E-mail address: dabing.zhang@adelaide.edu.au (D. Zhang).

http://dx.doi.org/10.1016/j.biotechadv.2017.05.005

Received 16 September 2016; Received in revised form 28 April 2017; Accepted 19 May 2017 0734-9750/ © 2017 Published by Elsevier Inc.

Please cite this article as: Devis, D.L., Biotechnology Advances (2017), http://dx.doi.org/10.1016/j.biotechadv.2017.05.005

Sandrini et al., 2015). Food allergens have been extensively reviewed, especially in the context of common or severe allergies to dairy and nuts that cause gastrointestinal or systemic anaphylactic reactions and/or atopic dermatitis (Campbell and Mehr, 2015; Patel and Volcheck, 2015; Pomés, 2008; Skypala and VliegBoerstra, 2014).

Grasses comprise one of the largest angiosperm families and contain many cultivated and widely distributed species that produce allergenic pollen and grains. Several reviews summarized respiratory diseases caused by grass, weed or tree pollen allergens (Andersson and Lidholm, 2003; Davies, 2014). In this review we focus on our understanding of the evolutionary events, common motifs and function of putative allergens in grass pollens, as well as efforts to alleviate allergies using biotechnological tools, and provide future research perspectives.

2. Global distribution of allergy-inducing grasses

Poaceae species are globally distributed and are the source of many pollen allergens. D'Amato et al. (2007) reported that airborne pollens from trees, weeds and grasses are associated with allergies across Europe. Smith et al. (2014) also proposed that the dominant species producing airborne pollens are grasses. Likewise, along with pollens of the ever-greens, grass pollens are dominant in the air in Calcutta, India (Mandal et al., 2008). In Lebanon, 17.75% of the allergy patients were allergic to grass pollen (Irani et al., 2013). Clinically important allergyinducing grass species have been recorded across the USA (Fuhrmann et al., 2016; White and Bernstein, 2003). Pollen and grain allergies have been well documented among both adults and children on the Australian continent as pollen allergies are more prevalent in Australia than other countries, and airborne grass pollen is a major contributor to allergic rhinitis in Australia and New Zealand (Asher et al., 2006; Australian Institute of Health and Welfare, 2011; Davies et al., 2015; Robertson et al., 1998; Woods et al., 2002) (Fig. 1). In Australia, the levels of airborne grass pollens varies greatly between location and climate (Beggs et al., 2015) and the distribution of grass types varies with latitude (Davies et al., 2015; Medek et al., 2016). Bermuda grass and Bahia grass are among the major pollen contributors to allergic disease in subtropical Queensland, and Ryegrass is one of the major contributors in temperate Melbourne (Davies et al., 2011, 2012).

The Poaceae family, which contains approximately 780 genera and around 12,000 species, includes many important grain-producing agricultural species such as rice (Oryza sativa), wheat (Triticum spp.) barley (Hordeum vulgare), and maize (Zea mays). In Poaceae, allergen proteins are often found in pollen or grains. In higher plants, pollen grains are produced via meiosis and mitosis within the male organ stamen, which contains the anther and supporting filament (Fig. 2) (Shi et al., 2015; D.S. Zhang et al., 2010; Zhang et al., 2013). Each pollen produced within the stamen contains two sperm cells and one vegetative cell and is enclosed by specialised pollen walls. Pollen grains fertilise ovules, which later develop into seeds and grains. Cereal grains are small, hard, and dry seeds and are a major global food source that contains a substantial amount of starch, a type of carbohydrate that provides dietary energy and nutrition (H. Zhang et al., 2010; Zhang et al., 2013). Grass pollens contain many proteins, only some of which are allergenic and cause an IgE response (Abou Chakra et al., 2012; Andersson and Lidholm, 2003). Currently at least 20 grass species such as timothy grass (Phleum pratense) and wheat have been reported to produce pollen proteins associated with the triggering of an IgE response, and the majority of these species are prevalent in temperate climates (Andersson and Lidholm, 2003). Rice, maize and wheat are crops that are grown and consumed on every continent excluding Antarctica. Therefore investigation of pollen allergens in agricultural Poaceae species is of great importance to the prevention of pollen allergies.

3. Allergens in Poaceae species

Allergen proteins are found in both pollens and grains and the number of allergen families varies between species. Wheat for example, contains up to 27 allergen families as characterized by the IUIS/WHO guidelines (Radauer et al., 2012, 2014). A number of allergen families are reported in grass pollens and/or grains, including β-expansin proteins Phl p 1 from timothy grass and Zea m 1 from maize, the lipid-transfer protein (LTP) Zea m 14 from maize, and the gliadin proteins Tri a 19, Tri a 20 and Tri a 21 from wheat (Table 1) (Anderson et al., 1989; Golias et al., 2013; Lehto et al., 2003; Matsuo et al., 2005; Pasini et al., 2002; Pastorello et al., 2013; Russell et al., 2008; Staiger et al., 1993; Trcka et al., 2012). Expansins are cell wall-loosening proteins that at least partially mediate the pH-dependent cell wall extension and cell growth. Expansins include two families, α-expansins and β-expansins, and are reported to be involved in modulating a variety of plant developmental events besides cell expansion, including organ morphogenesis, softening of fruits, and grass pollen tube growth (Artzi et al., 2016; Cho and Cosgrove, 2002; Grobe et al., 1999; Li et al., 2003; Marowa et al., 2016; Sampedro and Cosgrove, 2005; Valdivia et al., 2003, 2009; Wu et al., 2001; Yoo et al., 2003). Plant LTPs belong to a small and abundant lipid-binding protein group that exchange lipids between membranes, and have functions in various processes such as anther development and pollen formation (D.S. Zhang et al., 2010; Zhang et al., 2008). Gluten is a mixture of hundreds of related but distinct wheat storage proteins, mainly gliadin and glutenin (Anderson et al., 1989; Gil-Humanes et al., 2008; Gil-Humanes et al., 2011; Shewry et al., 2002; Weegels et al., 1996). Gliadin proteins in wheat grains promote dough elasticity and strength, and can also trigger an IgE response in allergy patients (Matsuo et al., 2005; Shewry et al., 2002). Similar storage proteins in other species, such as secalin from rye, hordein from barley, and avenin from oats, have also been collectively called gluten (Biesiekierski, 2017). Due to the prevalence of gluten in wheat and other grains, patients need to avoid wheat-based foods. However, because it is difficult and expensive to find appropriate dietary alternatives to allergenic grains, especially in countries where grain-based diets are the primary staple, it is important to understand the fundamental aspects of grass allergens and develop tools to alleviate food allergies.

Several studies in Poaceae species have identified allergen genes based on investigations using transcriptomics and proteomics as well as sequence comparison or IgE binding assays (Campbell et al., 2015; Golias et al., 2013; Li et al., 2012; Sekhar et al., 2015). Some putative allergens have been identified in DNA or transcriptome databases based on sequence similarity, but currently little is known about the specific structural or functional properties of proteins encoded by these sequences. Moreover, in many cases the expression of such protein sequences in relevant plant tissue and the capacity of these putative allergens to bind IgE or to elicit an allergic reaction de novo have not been demonstrated. Major groups of grass allergens discussed in this review include the α-amylase/trypsin inhibitors, β-expansin-like proteins expressed in pollen, and profilins expressed in both pollen and grain (Aalberse, 2000; Radauer et al., 2008, 2012, 2014).

4. Structure and conserved features of grass allergens

Putative allergen proteins can be predicted based on sequence homology compared with previously known allergens or known IgEbinding epitopes, but must be confirmed though immunological tests such as serological IgE assays or skin prick tests (Radauer et al., 2012, 2014). IgE binding epitopes can be comprised of continuous linear or discontinuous stretches of amino acids, but these motifs can be highly varied in amino acid sequences (Matsuo et al., 2015). Understanding how these allergen encoding genes evolved may be helpful in revealing allergenic properties and for designing future allergy prevention strategies.

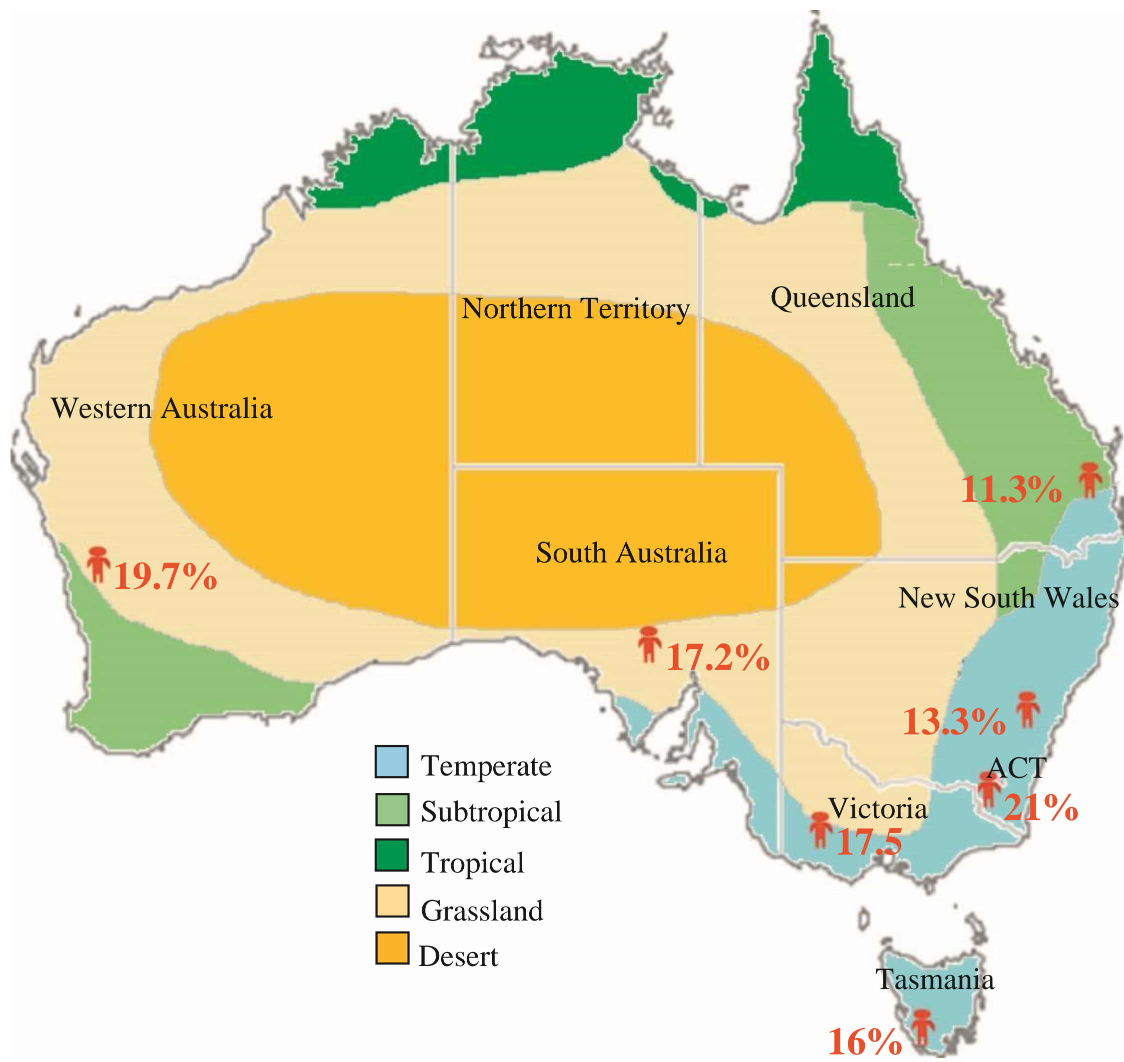


Fig. 1. Climate and prevalence of hay fever in Australia. Approximate percentage of state population with hay fever symptoms during peak season (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) Adapted from Australian Institute of Health and Welfare (2011).

7

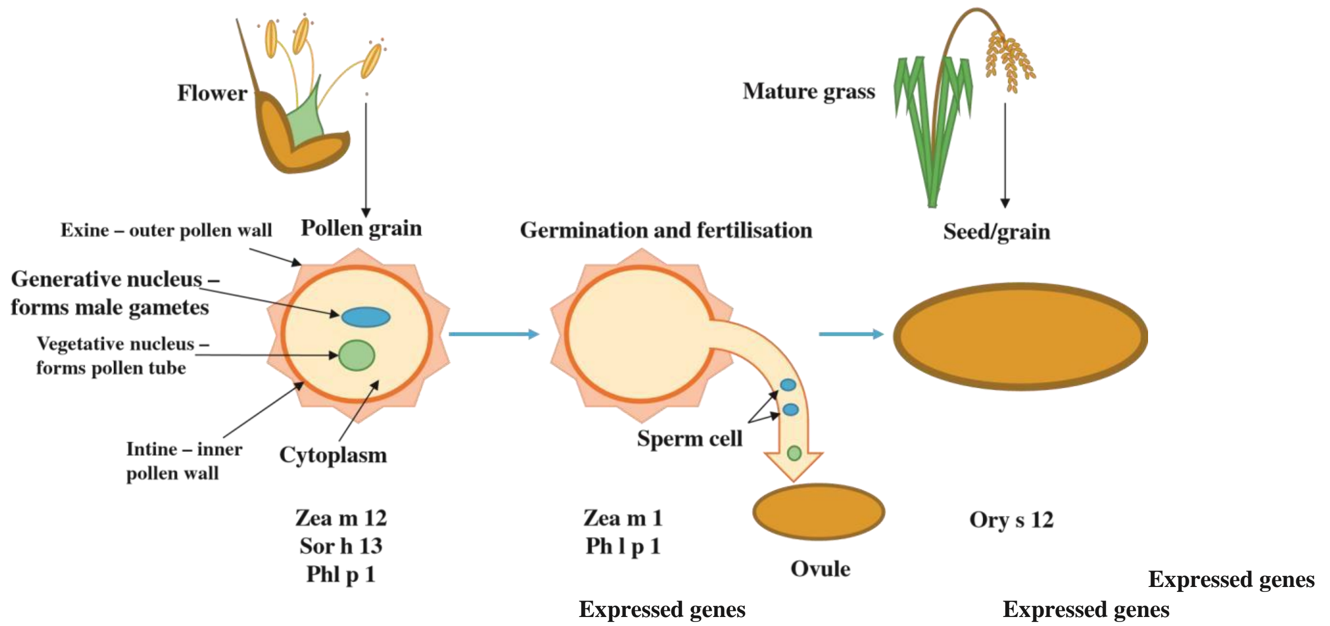


Fig. 2. Pollen germination and fertilisation.

Table1
PutativeallergenscharacterizedbyWHO/IUISguidelines − adaptedfrom M.L.Chenetal.(2016) .AllergensfromagriculturalPoaceaespeciesandtheirputativefunctions.Biologicalfunctionsthathavenotbeenexperimentallyconfirmedare markedwithanasterisk.

Species	Putativegenefamily/homolog	Allergenname	Biologicalfunction	Reference
Hordeumvulgare (barley)	Profilin	Horv12	Actin-bindingprotein*	Astwoodetal.(1995)
Oryzasativa (rice)	β Expansin	Orys1	Promotescellwallexpansion*	Xuetal.(1995)
	Profilin	Orys12	Actin-bindingprotein*	Yeetal.(2001)
Phleumpratense (timothy grass)	PollenOlee1allergen(10)	Phlp11	Accumulatesinpollentubecellwalls,possiblytoaidingerminationcontrolorpollen tubeemergence	deDiosAlchéetal.(2004), MarknellDeWittetal.(2002)
	Polygalacturonase	Phlp13	Pollenexpressedpolygalacturonasehasaroleinpollenmaturationandpollentube growth*	Sucketal.(2000)
	Profilin	Phlp12	Actin-bindingprotein*	Valentaetal.(1994)
	β Expansin	Phlp1	Promotescellwallexpansion*	Petersenetal.(1995), Valentaetal.(1992)
Sorghumhalepense (johnsongrass)	Polygalacturonase	Sorh13	Pollenexpressedpolygalacturonasehasaroleinpollenmaturationandpollentube growth*	Campbelletal.(2015)
	β Expansin	Sorh1	Promotescellwallexpansion*	Bashiretal.(2013), Campbelletal.(2015)
Triticumaestivum (wheat)	Gliadin	Tria19,Tria20,Tria21	Nutrientreservoiractivity*	Matsuoetal.(2005), Sanderetal.(2011)
	Glutenin	Tria26	Accumulatesintheendospermandformsaproteinmatrix.Thismatrixmayprovide carbon,nitrogenandsulphurstorageandaidinseedlinggrowth.	Andersonetal.(1989), Shewryetal.(2002), Weegelsetal.(1996)
	Lipidtransferprotein	Tria14	Lipidtransportation*	Sanderetal.(2011)
	Profilin	Tria12	Actin-bindingprotein*	Rihsetal.(1994)
	Profilin	Tria27	Providesenzymaticoxygenreductioninmultipletissues.Thispreventsoxidativestress inbothreproductiveandvegetativetissue,andcanalsomovebetweencells,allowing themtoindirectlycontrolmultipleprocesses	Maedaetal.(2006), Mengetal.(2010), Sanderetal.(2011)
	α Amylaseinhibitor	Tria28,Tria29,Tria30,Tria40	Directlyinhibits α amylasefrominsectsalivatopreventamylasedamagetoleaves	García-Marotoetal.(1990), Lullienetal.(1991), Sanderetal.(2011)
Zeamays (maize)	Polygalacturonase	Zeam13	Pollenexpressedpolygalacturonasehasaroleinpollenmaturationandpollentube growth	Allen andLonsdale(1992), Petersenetal.(2001)
	Lipidtransferprotein	Zeam14	Lipidtransportation*	Pastorelloetal.(2000)
	Profilin	Zeam12	Actin-bindingprotein*	Gibbonetal.(1998), Staigeretal.(1993)
	β Expansin	Zeam1	Inducesirreversiblecellwallexpansionandpromotespollentubegrowthand penetrationofovulebypollentube.Ovulepenetrationinfluencespollen-pollen competition.Expansinswithcellulose-binding-likedomainsarespecifictomonocots.	Broadwateretal.(1993), Cosgroveetal.(1997), SampedroandCosgrove(2005), Sampedroetal.(2015), Valdiviaetal.(2007)

Understanding the structure of allergens in the Poaceae family will help to uncover common features of allergens and identify targets for manipulation by gene technology. Particularly, allergens have varied high-order structures between protein families but conserved tertiary structure between homologs, raising the question of what causes a protein to have allergenic properties (Andersson and Lidholm, 2003; Petersen et al., 2001). One explanation of IgE cross-reactivity between grass allergens is that cross-reactivity may be due to structural similarities; thus, understanding structural differences and similarities among grass allergens will also clarify the evolutionary events of allergen genes in grasses (Hrabina et al., 2008; Marth et al., 2004). As more grass genome sequences and gene expression data become available, bioinformatics research promises to uncover more candidate allergenic proteins in grasses.

Different allergen families differ in their protein tertiary structures. For example, there are four main structures in the 40 allergens that represent all types of allergen structures: α-helices, α-helix and β-sheet associated structures, structures of α-helices and β-sheets that are not associated, and anti-parallel βsheets (Aalberse, 2000). The varied structures of allergens suggest that allergens may not have a specific common domain, but that any antigen can be allergenic if they have an ability to stimulate IgE and allergenic responses (Aalberse, 2000). Potentially, IgE recognises smaller structures of few amino acids in length. Two profilins, Hev b 8 from rubber tree and Zea m 12 from maize, were shown to contain a side loop primarily responsible for most IgE binding activity (Mares-Mejia et al., 2016). Furthermore, IgE binding was highest when Hev b 8 was in monomeric form, suggesting that some IgE epitopes are no longer exposed when Hev b 8 is in the dimeric form (Mares-Mejia et al., 2016). Therefore, understanding how structure relates to IgE binding avidity is essential for producing novel recombinant proteins for therapy.

Interestingly, Poaceae allergens share conserved cysteine between allergen family homologs in different taxa. Ten conserved cysteine residues in allergenic α-amylase/trypsin inhibitors were identified some allergens from rice, wheat and barley (Nakamura and Matsuda, 1996; Nakase et al., 1998). Likewise, Wang et al. (2014) revealed two cysteine rich motifs in allergenic αamylase/trypsin inhibitors but they were not present in non-allergenic homologs. Phylogenetic analysis showed that these α-amylase/trypsin inhibitor genes in rice were grouped in the same clade, suggesting that the cysteine residues are preferentially retained in the same evolutionary lineage (Wang et al., 2014). Another study analysed the structure of multiple wheat flour allergens and showed conservation of cysteine residues and disulphide bridges, suggesting that these structural features may play an important role in the function of allergens (Wang et al., 2013). Furthermore, cysteine residues are found in α-amylase/trypsin proteins, profilin proteins and β-expansin grass allergens (Fig. 4). However, how these residues are linked to protein function and allergenicity, and the functional consequences of disrupting disulphide bridges remain unclear.

5. α-Amylase inhibitors and trypsin inhibitors

α-Amylase is an enzyme existing in many organisms that is able to hydrolyse α-bonds of large, α-linked polysaccharides including starch and glycogen, forming glucose and maltose. Trypsin inhibitors are a type of serine protease inhibitor with the ability to reduce the biological activity of trypsin, an enzyme involved in the breakdown of a great variety of proteins. There have been reports that some genes encoding allergens show high homology to both α-amylase and trypsin inhibitor genes, suggesting that protein products of these genes may have multiple functions (Alvarez et al., 1995). Plant α-amylase inhibitors can inhibit the activity of digestive α-amylase in insect saliva, thus playing a protective role in plants (Franco et al., 2002; Kaur et al., 2014; Major and Constabel, 2008; Mehrabadi et al., 2012; Pytelkova et al., 2009). Trypsin inhibitors are also involved in plant defence through protein degradation (Rogozhin et al., 2012). The expression of α-amylase/trypsin inhibitors is frequently seen in seeds of cereals such as barley and wheat and the αamylase/trypsin inhibitor genes may have arisen during evolution following grass diversification (Adachi et al., 1993, Iulek et al., 2000, Wang et al., 2014). Notably, sequencebased prediction of α-amylase inhibitor-and trypsin inhibitor- types of allergens may not be accurate. For instance, in rye, α-amylase inhibitors also acted as a tripsin inhibitor proteins, but Iulek et al. (2000) showed that homologs in wheat and barley were unable to inhibit trypsin like in rye. Rice contains at least 7 α-amylase/trypsin inhibitor genes, but their function is currently unknown (Wang et al., 2014). As grain expressed allergens, αamylase and trypsin inhibitor proteins are likely to be denatured during cooking, losing their allergenicity easily.

6. Profilins

Present in both monocots and dicots, profilins control the growth of actin microfilaments and cytoskeleton restructuring (Gibbon et al., 1998; Sun et al., 2013; Wang et al., 2006). Profilins have sequence homologs that are expressed in vegetative tissues and/or reproductive organs, but the function of the putative allergenic profilins remains unknown. In Arabidopsis thaliana (Arabidopsis), profilins were shown to be partially redundant in function, but also have specific expression pattern and functions in determining leaf morphology and root growth (Müssar et al., 2015). In maize, there may be at least two functionally distinct profilin groups, one group inhibiting the hydrolysis of phospholipids and the other group involved in the dynamics of actin filament (Kovar et al., 2000). A proteomic study found the maize profilin-like putative allergen, Zea m 12, in the pollen coat, suggesting that this protein may play a specific role in this tissue (Wu et al., 2015). However, not all profilins are allergenic. Very low IgE reactivity was found for Zea m 12, indicating that it may not be a dominant allergen in grasses (Petersen et al., 2006; Radauer and Breiteneder, 2006). There is fairly low sequence similarity between allergenic and non-allergenic profilins (30–40%), but the allergenic properties have not been well characterized in grasses and the relationship between their function and allergenicity remains unclear (Santos and Van Ree, 2011).

7. β-Expansins

β-Expansins are cell wall expansion proteins expressed in multiple plant tissues. This family of allergens includes the Group 1 grass specific pollen allergens, which are the major allergens in grass pollen that account for the majority of IgE binding in patient sera and to which there is currently no prevention (Hirano et al., 2013). Although Group 1 allergens are β-expansin proteins in sequence, they exist in multiple isoforms within a single species and may have multiple functions (Cosgrove et al., 1997; Sampedro and Cosgrove, 2005). Expansins promote fruit softening in fruit-producing plants and root hair initiation in Arabidopsis, but these expansins may differ in function from grass specific β-expansin proteins (Cho and Cosgrove, 2002; Kalamaki et al., 2003; Yoo et al., 2003). Many allergenic β-expansin genes are expressed only in pollen and are involved in cell-wall loosening to promote pollen tube growth towards the ovary (Valdivia et al., 2009). In contrast, nonallergenic β-expansin genes are expressed in vegetative tissue, suggesting that allergenic β-expansin genes may have obtained diversified functions during evolution (Lee and Kende, 2001; Sampedro and Cosgrove, 2005; Sampedro et al., 2015; Schipper et al., 2002; Valdivia et al., 2009; Wu et al., 2001). One maize allergen, Zea m 1, was shown to play a role in pollen-pollen competition by influencing ovule access (Valdivia et al., 2007). In addition, mutations of some βexpansin allergen genes caused disrupted pollen formation, pollen dispersal and overall reproductive fitness, so β-expansin proteins likely play a crucial role in pollen development (Valdivia et al., 2009).

ARTICLE IN PRESS

Furthermore, Zea m 1 was also shown to have a role in pollen tube growth and ovule penetration, as mutated Zea m 1

exhibited aborted pollen tube growth, potentially due to reduced cell wall expansion and loosening activity (Cosgrove et al., 1997; Valdivia et al., 2009). Given these findings in maize, it is likely that β-expansins in other grasses may also be involved in pollen development and pollen tube growth. To this end, it will be important to investigate the full function within the plant of β-expansinlike allergen encoding genes in other Poaceae species.

7.1. Evolution and functional prediction of pollen allergens

Recently, we performed genome-wide analysis of sequences encoding putative pollen allergens, and predicted 145 and 107 candidate pollen allergens from rice and Arabidopsis, respectively (M. L. Chen et al., 2016; Cui et al., 2012). These pollen proteins are likely involved in metabolic processes such as cell wall metabolism and may play roles in stress responses based on gene expression data (M. L. Chen et al., 2016). Genes that encode these proteins appear to have evolved from large gene families and later diversified during evolution. Furthermore, sequence analysis among 25 plant species from green alga to angiosperms suggested that about 40% of putative pollen allergenic proteins, including profilins, are conserved from lower to higher plants, while other allergens such as β-expansins seemed to be formed later in evolution. Although a high frequency of gene duplication was observed among allergencoding genes, we predicted that these allergen-coding genes might have been selected by purifying selection during evolution. To explain the phylogenetic and evolutionary scenario of genes that encode putative pollen allergens, we proposed that many allergen gene families experienced both duplication and deletion events that resulted in highly variable degrees of conservation between species (Fig. 3). Genes originated from a putative allergen-coding gene may have experienced multiple duplications, resulting in increased copy number, and these genes may have either been maintained, deleted or modified to lose allergenicity or evolve tissue specific functions. Non-allergen encoding genes may have been duplicated and later modified to gain allergenicity, resulting in gene families that contain genes for both allergenic and non-allergenic proteins.

7.2. Allergens derived from ancient and recent duplications

Some allergen-encoding genes have been conserved in Poaceae species. A study of a putative rice allergen Ory s 2, showed a 33% sequence identity between the C-terminus of this gene and ryegrass (Lolium perenne) allergen genes, suggesting that conservation of allergen sequences occurs between grasses (Kerim et al., 2003). However, no immunological cross-reactivity was found between Ory s 2 and other grass allergens with similar protein structures (except wheat), suggesting that these conserved regions may not necessarily pertain to allergenicity (Kerim et al., 2003). However, another research group identified five wheat allergen cDNAs that encode proteins sharing IgE reactivity and cross-reactivity with other putative allergenic seed proteins in grains such as barley, rye and rice (Pahr et al., 2012). The cross-reactivity of these wheat allergens and seed proteins in other grasses suggests that these proteins may contain conserved motifs stemming from a common ancestor, raising the question as to whether the putative allergenicity is structurally or functionally significant. As such, structural and functional variation of putative allergens between grasses needs to be further studied to understand why the level of conservation varies.

Bioinformatics analysis suggested that Poaceae allergen genes may have arisen both prior to and following the divergence of grasses. For instance, several β-expansin genes may have evolved before the grass species diverged, but also underwent rapid evolution following divergence (Sampedro et al., 2015). β-expansin genes can be classified into two groups; one group existed before monocots and dicots diverged and has been conserved after the divergence, and the other group arose only in monocots, suggesting that the latter group may have a grass specific function (M. L. Chen et al., 2016). The rice β-expansin Ory s 1 is a major rice pollen allergen that is also conserved in maize, wheat, barley, timothy grass, johnson grass (Sorghum halepense), panic grass (Panicum hallii) and sorghum (Sorghum bicolor). These homologs are varied in gene copy number in different grass species, from 2 copies in timothy grass to 12 copies in maize. These evolutionary events may provide insight into the importance of copy number in grass species and aid in gene technology and gene silencing strategies to reduce allergens in transgenic plants.

7.3. Allergen genes duplicated within one single grass species

Phylogenetic analyses of all the allergen groups in multiple species could help to further understand the evolution and possibly the role of these genes in Poaceae. Understanding the evolution of allergenic and non-allergenic homologs in one species may also reveal why some homologous genes are allergenic and others are not, and why multiple allergen genes have been evolved. Allergen genes in Poaceae species were shown to be conserved within the genome of single species. Russell et al. (2008) showed that highly homologous allergen genes in rice often lack introns and have multiple copies of the gene located near each other on DNA. For example, seven αamylase/trypsin inhibitor allergen genes, all of which are expressed in seeds, are on chromosome 7 in rice (Wang et al., 2014). It was proposed that high copy numbers of rice pollen allergen genes were the result of an insertion of highly expressed genes that were modified to be more efficiently translated (Wang et al., 2014). M. L. Chen et al. (2016) proposed that these allergen gene families may be caused by frequent duplications, after which the allergen gene or allergenic feature of the gene has been retained. These findings raised the question of how efficient translation and gene copy number affect allergenicity and plant function.

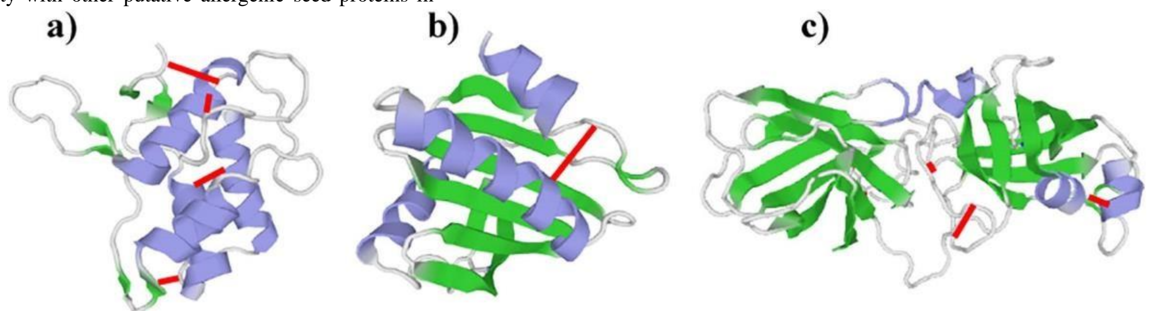


Fig. 3. Tertiary structure of Profilin, β-expansin and α-amylase/trypsin inhibitor. Tertiary structures of maize α-amylase/trypsin inhibitor (a), Arabidopsis thaliana profilin (b) and timothy grass β-expansin Phl p 1 (c) consist of α-helices (blue), β-sheets (green) and disulphide bridges (red bar). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Images were retrieved from SWISS-MODEL (Biasini et al., 2014).

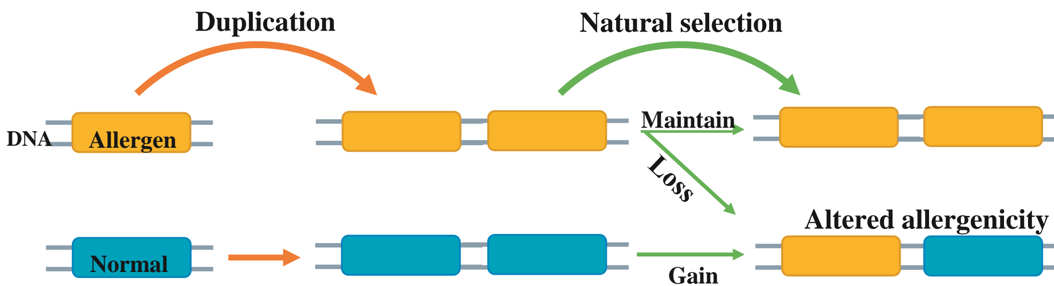


Fig. 4. Evolution of allergen genes. Allergen genes (yellow) and non-allergenic genes (blue) may experience duplications that lead to multiple gene copies. Through the course of natural selection, these copies can be maintained, lost or new allergens may be gained from non-allergenic genes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) Adapted from M. L. Chen et al. (2016).

8. Biotechnological approaches to prevent allergies to Poaceaeallergens

Although avoiding exposure to allergens is the most common form of prevention of allergic responses, post-harvest treatment is also important for alleviating allergenic properties (Dodo et al., 2008; Ogawa et al., 2000; Yu et al., 2013). Enzymatic treatment of harvested foods can decrease the concentration of allergen proteins in multiple foods, and transgenic foods with decreased level of allergen proteins resulted in lower IgE binding, decreased severity and prevalence of allergic reaction (Le et al., 2006a,b; Nakamura and Matsuda, 1996; Wakasa et al., 2011). However, due to the durable structure of pollen grains it is difficult to reduce allergenicity caused by pollen proteins.

9. Allergen-specific immunotherapy

Allergen-specific immunotherapy (AIT) has been used clinically for over 100 years to alleviate pollen allergies, but research into improving efficacy, efficiency, safety and specificity remains an active area of research. Research advances in AIT for grass pollen focus on temperate grass species and include the use of recombinant pollen allergen molecules (Nandy et al., 2015; Patel and Volcheck, 2015). A review paper by Calderon et al. (2012) summarized the safety and efficacy of AIT for grass pollen allergy prevention. Investigations on the use of AIT against timothy grass pollen allergy include (1) employment of different routes of administration such as subcutaneous or sublingual, which impact the cellular uptake of antigens by dendritic cell subtypes or B cells, (2) processing and presentation of allergens to induce tolerance by induction of T regulatory lymphocytes, and (3) deletion of allergenspecific T helper 2 (allergyprone) lymphocytes (Nandy et al., 2015; Rolland and O'Hehir, 1998; Sandrini et al., 2015; Valenta et al., 2016). Several new products for temperate grass pollen allergen immunotherapy are in clinical use as AIT (Cox et al., 2012; Dahl et al., 2006; Didier et al., 2009; Nelson, 2014; Senna et al., 2011; Sjölin et al., 2008).

Another promising method for relieving allergies through AIT is through modifying allergenic proteins. Focke-Tejkl et al. (2015) fused timothy grass pollen allergens with non-allergenic peptides for the production of hypoallergenic recombinant allergen derivatives. In murine models of AIT, Immunoglobulin G (IgG) antibodies were induced instead of IgE, and these antibodies could inhibit IgE-binding without causing allergic reactions, while still being able to induce immune tolerance (Focke-Tejkl and Valenta, 2012; Focke-Tejkl et al., 2015). Treatment with synthetic peptide immuno-regulatory epitopes has decreased symptoms of grass pollen allergies to both timothy grass and bermuda grass (Cynodon dactylon) (Ellis et al., 2014, 2015). Additionally the fusion of timothy grass pollen allergens with mannin, a plant polysaccharide, produced a hypoallergenic protein derivative that showed a significant reduction of IgE binding (Manzano et al., 2016). Furthermore, Devanaboyina et al. (2014) demonstrated that single point mutations in the predicted IgE binding sites on timothy grass pollen allergen Phl p 3 caused a 30% decrease in IgE binding activity compared with the wild type. Meanwhile this research also showed that a wide variety of the amino acid structures could be recognised by IgE (Devanaboyina et al., 2014). Revelations of how IgE recognises allergenic motif/structures will be very useful for the development of novel AIT strategies.

Pollen allergens elicit allergic sensitization via a respiratory mucosal route. However, allergens within the grain can cause symptoms as severe as anaphylaxis when ingested. Gluten is a mixture of grain proteins, some of which are major allergens. An AIT based on small peptide sequences of gluten was able to combat celiac disease caused by wheat and rye (Anderson, 2008; Nabavi and Hoseeinzadeh, 2009). Although this AIT has entered phase 1 clinical trials, an effective wheat allergen AIT has not yet become commercially available (Brown et al., 2011; Veeraraghavan et al., 2015).

AIT has been reported to cause a range of side effects, such as asthma, increased severity of allergic reactions, and anaphylactic shock (Calderon et al., 2012; Winther et al., 2006). For instance, one study on grass allergen therapy showed that about 30% of patients treated for grass allergy developed mild respiratory side effects after a three-year administration of injected allergen immunotherapy, and 4% of these patients discontinued treatment due to severe symptoms (Winther et al., 2006). Furthermore, 46% of the patients in the test group of a double-blind study of orally administered allergen therapy reported site-related side effects such as itching, compared to 4% in the placebo group (Dahl et al., 2006). Although these reactions were mostly mild and short-lived, 5 of the 316 tested patients withdrew due to continued side-effects related to the treatment (Dahl et al., 2006). It was explained that these side-effects are possibly due to variation in IgE sensitivity of individual patients to specific allergens (Calderon et al., 2012; FockeTejkl and Valenta, 2012). Furthermore, modified allergens with decreased IgE binding ability that were used in allergy therapies have been shown to retain immune T-cell response and induce immune tolerance (Kahlert et al., 2000). Further research into safer AITs is needed. Alternative therapies such as using transgenic hypoallergenic crops may provide a possible solution to these problems.

10. Genetic modification of plant allergen gene expression

Gene technology, which has been carried out with ryegrass and rice, provides the possibility of producing hypoallergenic plants that do not trigger an immune response. Possible methods to silence allergen genes include RNA interference (RNAi) and gene knockout approaches (Table 2). Studies in ryegrass used a pollen-specific rice promoter, Ory s 1 promoter, to control the expression of an anti-sense sequence of a ryegrass pollen allergen Lol p 5, which resulted in hypoallergenic ryegrass that exhibited low Lol p 5-specific IgE binding (Bhalla et al., 1999; Singh and Bhalla, 2008; Takagi et al., 2005). However, small amounts of remaining allergenic proteins could still trigger an immune response (Bhalla and Singh, 2004; Bhalla et al., 1999; Singh and Bhalla,

ARTICLE IN PRESS

Table 2
Comparison of biotechnological strategies for Poaceae allergen specific immunotherapy.

Species	Allergen source	Targeted allergen	Therapy	Method	Factor altered	Compared to wild type (%)	Reference
Rice	Seeds	14–16 kDa fragment	GMO plant	RNAi silencing	Transcript levels	20	Adachi et al. (1993)
	Pollen	RA17	GMO plant	RNAi silencing	Transcript levels	20–30	Tada et al. (1996)
	Unknown	AK107328, AK066092	GMO plant	RNAi silencing and gene knock out	IgE reactivity	10	Wakasa et al. (2011)
Ryegrass	Pollen	Lol p 5	GMO plant	RNAi silencing	IgE reactivity	0	Bhalla and Singh (2004)
	Pollen	Lol p 5	GMO plant	RNAi silencing	IgE reactivity	0	Bhalla and Singh (2004)
	Unknown	Lol p 1, Lol p 2	GMO plant	RNAi silencing	Protein levels, IgE reactivity	Unknown	Petrovska et al. (2004)
	Unknown	Lol p 5	Oral AIT		IgE reactivity	0	Takagi et al. (2005)
Timothy grass	Pollen	Phl p 3	AIT	Single nucleotide polymorphism	IgE reactivity	30	Devanaboyina et al. (2014)
	Pollen	Phl p 1, Phl p 2, Phl p 5, Phl p 6	AIT	Non-allergenic peptide fusion	IgE reactivity	0.5–60	Focke-Tejkl et al. (2015)
	Pollen	Unknown	AIT	Fusion of allergen protein with mannin	IgE reactivity	20–25	Manzano et al. (2016)

2008). Another study reported the down-regulation of the major ryegrass pollen allergens Lol p 1 and Lol p 2 by placing anti-sense Lol p 1 and Lol p 2 under the pollen-specific maize Zm13 promoter. The resultant transgenic plants showed decreased allergen levels and IgE reactivity (Petrovska et al., 2004). These studies were conducted over ten years ago, but no follow up studies have been reported on how to improve application of this technology. Technically, grass pollen and grain allergens are frequently encoded by multiple copies of genes, so a complete knock-out of the allergens is difficult. Furthermore, understanding the function of gene homologs in other grass species may aid in understanding whether modified hypoallergenic protein peptides can be used in place of wild type allergens in transgenic plants (Suphioglu et al., 1992; Valenta et al., 1991, 1992).

Development of wheat grains with reduced allergenic gliadin has been achieved through the use of RNAi (Gil-Humanes et al., 2008, 2010). Grains from these transgenic lines had reduced T-cell binding, so flour produced from these grains may be an alternative bread source for celiac patients, (GilHumanes et al., 2008, 2011, 2014). However, IgE binding analyses were not carried out in these studies, so it is unclear whether these transgenic lines will be suitable for patients with wheat allergy. Moreover, wheat contains other allergens that may still be present in these transgenic lines, and it is unclear whether biological function of these transgenic plants is affected.

Biotechnology has also been used to develop hypoallergenic rice lines. Transgenic plants containing allergen-specific antisense RNA showed approximately 20% of α-amylase/trypsin inhibitor allergens in grains compared with the wild type (Nakamura and Matsuda, 1996). Likewise, an RNAi approach decreased the transcript levels of αamylase/trypsin inhibitor allergens to 20%–30% in transgenic rice pollen (Tada et al., 1996). However, there were no analyses on the immune response of patients, so it is unclear whether the decreased transcript and protein levels would be effective in reducing allergic responses. These two studies were conducted in 1996, but there has been no follow-up report on using these rice lines nor is it clear whether biological function was affected in the transgenic plants. By knocking out a rice 26 kDa allergen (GbN-1) coupled with RNAi of another two allergens, α-globulin and β-glyoxalase, transgenic plants had decreased allergen levels in grains and a dramatic decrease (to about 10%) of IgE binding activity in sera of all 15 patients in the study compared to the control sera (Wakasa et al., 2011). These studies did not report on whether silencing allergen genes disrupted plant growth or function, so a functional analysis is needed to assess the viability and performance of hypoallergenic plants.

Gene knock-out or RNAi-mediated silencing of pollen-specific allergens may lead to disrupted pollen development and fertility or an insufficient decrease in allergen levels in transgenic grasses.
Therefore, it is necessary to explore gene-site-specific modification as an alternative method to reduce the level or alter the structure of pollen-specific allergen proteins, which can potentially reduce allergic responses while maintaining normal biological development and function of the plant. Genome editing technology such as CRISPR-Cas9 has been shown to be effective in many grasses, but has not yet been used to inhibit allergen activity (Feng et al., 2016; Kim et al., 2016; Li et al., 2016; Ma et al., 2015, 2016). This method may allow allergen families to be efficiently altered or reduced in expression in order to develop hypoallergenic plant derivatives that maintain biological function.

11. Challenges and future directions

Biotechnological approaches aimed at developing faster, safer and more targeted AIT has been heavily investigated in plants, particularly for grass pollen-specific AIT (Calderon et al., 2012; Passalacqua and Canonica, 2016). Focusing on reducing the expression or mutating the allergenic genes may provide a more convenient method to overcome pollen and grain induced allergic diseases compared with the need for prolonged and frequent AIT administrations. Transgenic plants may provide a relief for field workers during harvest and offer alternative consumables to overcome grain allergies, making these plants an economically desirable product. Ethical concerns about how transgenic grass allergens affect humans and the environment may prevent commercial distribution of hypoallergenic plants, but a better understanding of the function and role of allergen genes and comparative studies of transgenic and non-transgenic plants may help to ease some of these concerns (Graf et al., 2014; Vidal et al., 2015).

With the advance of bioinformatics and functional genomics, it is possible to investigate the biological function and the role of allergens in plant development. Future work on the history and molecular function of allergens and the knowledge applicable from one species to others is also critical. The structure and IgE binding efficiency of pollen and grain allergens varies greatly between plant species and cannot yet be attributed to a sole motif or protein structure (Aalberse, 2000). A better understanding of conserved structures and posttranscriptional or post-translational mechanisms that cause an immune response will be useful to determine why some antigens are allergenic and others are not. For instance, questions on how cysteine residues and disulphide bridges are associated with allergenicity or protein function are interesting to address. As many disulphide bridges are strongly conserved through different grass allergens despite lack of a clear common ancestral gene, it is unclear

how they pertain to a specific functional or structural role, or whether they were conserved due to a lack of selective pressure.

Research on the specific expression pattern and protein localisation

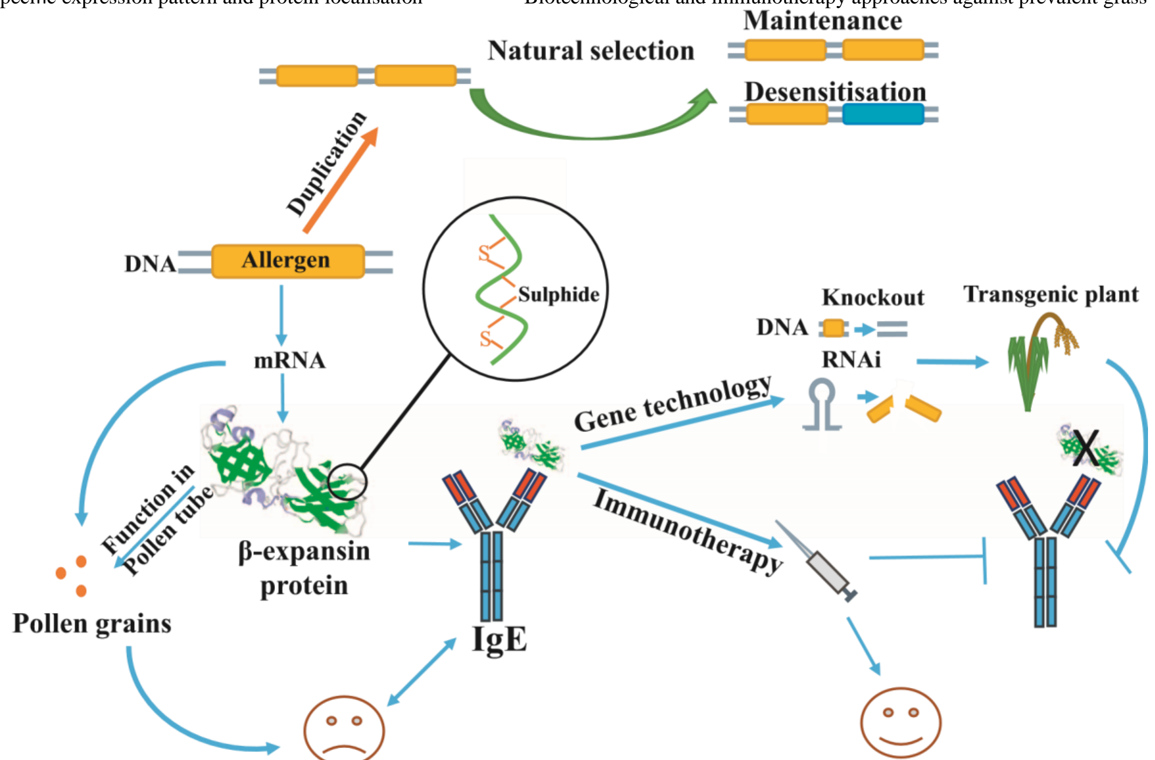


Fig. 5. Summary of Allergen biology and therapy. An Allergen gene (yellow) can evolve in multiple ways. The DNA is transcribed into mRNA and translated into a β-expansin protein, which has conserved cysteine bridges. The allergen may be expressed in pollen and cause an allergenic reaction in patients, confirmed through IgE binding tests. The allergen can be modified using gene technology methods to develop a transgenic plant, or modified for allergen immunotherapy use. The allergen derivative must then be tested for IgE and immunological tests to ensure no reaction occurs in allergy patients. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of allergen proteins will also be critical for allergen prevention. Lol p 1 and Lol p 5 in ryegrass localise to the intine wall, so localisation data could reveal whether localisation of allergens or proximity to the pollen grain surface plays a role in access of IgE to protein (Suphioglu et al., 1992). Furthermore, expression data many help answer whether some allergens vary in potency due to protein concentration or gene sequence, or whether IgE preferentially binds to proteins located in a particular tissue. This knowledge is essential to the efficient isolation of native allergens for immunotherapy.

The capacity for rational design of AIT could be further improved by understanding the relationship between different allergen structures and patient sensitivity. It was reported that adults with wheat grain allergies were more likely to be sensitized to wheat pollen, suggesting that some allergens within grain are also expressed in pollen or that grain and pollen allergens may share common features (Armentia et al., 2002). Potentially, these shared features could be used to identify and to isolate known food allergens and modifying these shared features them for pollen and grain allergen AIT or gene silencing.

Hypoallergenic plants may provide relief for patients suffering from pollen allergies, but it is necessary for these transgenic plants to retain normal biological function and support plant development in a manner comparable to wild type strains. Alternative and simpler methods such as CRISPR-Cas9 technology may aid in developing gene-site-specific mutations that allow for pollen growth while decreasing allergen-IgE reactivity in patients. The CRISPR-Cas9 system may provide exciting pathways into altering allergenicity of proteins in grasses while retaining normal function and development of the plant.

12. Summary

Biotechnological and immunotherapy approaches against prevalent grass induced allergic diseases have been extensively researched, but little is currently known about the evolution and function of allergens and this hinders the development of faster, safer and more targeted AIT and biologically viable transgenic hypoallergenic plants (Fig. 5). In this review, we summarized the prevalence and variation of Poaceae species and genes in the context of allergenicity, and pointed out the lack of knowledge on the evolution and function of genes encoding putative allergens in the Poaceae family. Beyond current clinical use of AIT for temperate grass pollen allergy, including new treatments based on whole allergen extracts, AIT for other common Poaceae allergens requires further research, as many species do not yet have an efficient AIT therapy. Finally, genetically modified hypoallergenic plants are promising but have not yet been developed to completely silence allergen genes in pollen and grains. Complete knowledge about allergen function and immune responses will be key to the development of new methods for AIT and transgenic hypoallergenic crops. Exciting new technologies such as genome editing may also provide a pathway for developing new hypoallergenic grasses for the reduction and prevention of allergenicity.

Funding

School of Agriculture, Food, and Wine, University of Adelaide (startup grant to D.Z.); Department of Industry, Innovation and Science, Australia for the Australia-China Joint Centre of Grains for Health (ACSRF481), Australian Research Council (DP170103352).

References

Aalberse, R.C., 2000. Structural biology of allergens. J. Allergy Clin. Immunol. 106, 228–238.
Abou Chakra, O.R., Sutra, J.P., Demey Thomas, E., Vinh, J., Lacroix, G., Poncet, P., et al., 2012. Proteomic analysis of major and minor allergens from isolated pollen cytoplasmic granules. J. Proteome Res. 11, 1208–1216.
Adachi, T., Izumi, H., Yamada, T., Tanaka, K., Takeuchi, S., Nakamura, R., et al., 1993. Gene structure and expression of rice seed allergenic proteins belonging to the αamylase/trypsin inhibitor family. Plant Mol. Biol. 21, 239–248. Aiubi, S., Kor, N.M., Azar, O.L., Rahmani, F.A., Berenjestanaki, Z.H., Zoraghi, E., 2015.

Causes of allergies and diagnosis and treatment. Int. J. Life Sci. 9, 18–24.
Allen, R.L., Lonsdale, D.M., 1992. Sequence analysis of three members of the maize polygalacturonase gene family expressed during pollen development. Plant Mol. Biol. 20, 343–345.
Alvarez, A.M., Adachi, T., Nakase, M., Aoki, N., Nakamura, R., Matsuda, T., 1995. Classification of rice allergenic protein cDNAs belonging to the α-amylase/trypsin inhibitor gene family. Biochim. Biophys. Acta 1251, 201–204.
Amardip, S., Poonam, Priti, 2014. Aeroallergens investigation with special reference to identification of diversity of allergenic plant species dispersing pollen grains. J. Environ. Res. Dev. 9, 454–464.
Anderson, R.P., 2008. Development of a Vaccine for Celiac Disease. S Karger AG, Basel, Switzerland.
Anderson, O.D., Greene, F.C., Yip, R.E., Halford, N.G., Shewry, P.R., Malpica-Romero, J.M., 1989. Nucleotide sequences of the two high-molecular-weight glutenin genes from the Dgenome of a hexaploid bread wheat, Triticum aestivum L. cv Cheyenne. Nucleic Acids Res. 17, 461–462.
Andersson, K., Lidholm, J., 2003. Characteristics and immunobiology of grass pollen allergens. Int. Arch. Allergy Immunol. 130, 87–107.
Armentia, A., Rodriguez, R., Callejo, A., Martin-Esteban, M., Martin-Santos, J.M., Salcedo, G., et al., 2002. Allergy after ingestion or inhalation of cereals involves similar allergens in different ages. Clin. Exp. Allergy 32, 1216–1222.
Artzi, L., Morag, E., Shamshoum, M., Bayer, E.A., 2016. Cellulosomal expansin: functionality and incorporation into the complex. Biotechnol. Biofuels 9, 61.
Asher, M.I., Montefort, S., Bjorksten, B., Lai, C.K.W., Strachan, D.P., Weiland, S.K., et al., 2006. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. Lancet 368, 733–743.
Astwood, J.D., Mohapatra, S.S., Ni, H., Hill, R.D., 1995. Pollen allergen homologues in barley and other crop species. Clin. Exp. Allergy 25, 66–72.
Australian Institute of Health and Welfare, 2011. Allergic Rhinitis ('Hay Fever') in Australia.: Cat. No. ACM 23. AIHW, Canberra.
Bashir, M.E.H., Ward, J.M., Cummings, M., Karrar, E.E., Root, M., Mohamed, A.B., et al., 2013. Dual function of novel pollen coat (surface) proteins: IgE-binding capacity and proteolytic activity disrupting the airway epithelial barrier. PLoS One 8, 15.
Beggs, P.J., Katelaris, C.H., Medek, D., Johnston, F.H., Burton, P.K., Campbell, B., et al., 2015. Differences in grass pollen allergen exposure across Australia. Aust. N. Z. J. Public Health 39, 51–55.
Bhalla, P.L., Singh, M.B., 2004. Knocking out expression of plant allergen genes. Methods 32, 340–345.
Bhalla, P.L., Swoboda, I., Singh, M.B., 1999. Antisense-mediated silencing of a gene encoding a major ryegrass pollen allergen. Proc. Natl. Acad. Sci. U. S. A. 96, 11676–11680.
Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 42, 252–258.
Biesiekierski, J.R., 2017. What is gluten? J. Gastroenterol. Hepatol. 32, 78–81.
Bousquet, J., Khaltaev, N., Cruz, A.A., Denburg, J., Fokkens, W.J., Togias, A., et al., 2008. Allergic rhinitis and its impact on asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). Allergy 63, 8–160.
Broadwater, A.H., Rubinstein, A.L., Chay, C.H., Klapper, D.G., Bedinger, P.A., 1993. Zea m I, the maize homolog of the allergen-encoding Lol p I gene of rye grass. Gene 131, 227–230.
Brown, G.J., Daveson, J., Marjason, J.K., Ffrench, R.A., Smith, D., Sullivan, M., et al., 2011. A phase 1 study to determine safety, tolerability and bioactivity of Nexvax2® in HLA DQ2+ volunteers with celiac disease following a long-term, strict gluten-free diet. Gastroenterology 140, S437–S438.
Buters, J., Prank, M., Sofiev, M., Pusch, G., Albertini, R., Annesi-Maesano, I., et al., 2015. Variation of the group 5 grass pollen allergen content of airborne pollen in relation to geographic location and time in season. J. Allergy Clin. Immunol. 136 (87–95), e6.
Calderon, M., Cardona, V., Demoly, P., 2012. One hundred years of allergen immunotherapy European Academy of Allergy and Clinical Immunology celebration: review of unanswered questions. Allergy 67, 462–476.
Campbell, D.E., Mehr, S., 2015. Fifty years of allergy: 1965–2015. J. Paediatr. Child Health 51, 91–93.
Campbell, B.C., Gilding, E.K., Timbrell, V., Guru, P., Loo, D., Zennaro, D., et al., 2015. Total transcriptome, proteome, and allergome of Johnson grass pollen, which is important for allergic rhinitis in subtropical regions. J. Allergy Clin. Immunol. 135, 133–142.
Chen, M.L., Xu, J., Devis, D., Shi, J.X., Ren, K., Searle, I., et al., 2016. Origin and functional prediction of pollen allergens in plants. Plant Physiol. 172, 341–357.
Chen, Y., Wong, G.W.K., Li, J., 2016. Environmental exposure and genetic predisposition as risk factor for asthma in China. Allergy, Asthma Immunol. Res. 8, 92–100.
Cho, H.T., Cosgrove, D.J., 2002. Regulation of root hair initiation and expansin gene expression in Arabidopsis. Plant Cell 14, 3237–3253.
Cosgrove, D.J., Bedinger, P., Durachko, D.M., 1997. Group I allergens of grass pollen as cell wall-loosening agents. Proc. Natl. Acad. Sci. U. S. A. 94, 6559–6564.
Cox, L.S., Casale, T.B., Nayak, A.S., Bernstein, D.I., Creticos, P.S., Ambroisine, L., et al., 2012. Clinical efficacy of 300IR 5-grass pollen sublingual tablet in a US study: the importance of allergen-specific serum IgE. J. Allergy Clin. Immunol. 130, 1327–1334.e1.
Cui, X., Wang, Q., Yin, W., Xu, H., Wilson, Z.A., Wei, C., et al., 2012. PMRD: a curated database for genes and mutants involved in plant male reproduction. BMC Plant Biol. 12, 215.
Dahl, R., Kapp, A., Colombo, G., deMonchy, J.G.R., Rak, S., Emminger, W., et al., 2006. Efficacy and safety of sublingual immunotherapy with grass allergen tablets for seasonal allergic rhinoconjunctivitis. J. Allergy Clin. Immunol. 118, 434–440.
D'Amato, G., Cecchi, L., Bonini, S., Nunes, C., Annesi-Maesano, I., Behrendt, H., et al., 2007. Allergenic pollen and pollen allergy in Europe. Allergy 62, 976–990.
Davies, J.M., 2014. Grass pollen allergens globally: the contribution of subtropical grasses to burden of allergic respiratory diseases. Clin. Exp. Allergy 44, 790–801.
Davies, J.M., Dang, T.D., Voskamp, A., Drew, A.C., Biondo, M., Phung, M., et al., 2011. Functional immunoglobulin E cross-reactivity between Pas n 1 of Bahia grass pollen and other group 1 grass pollen allergens. Clin. Exp. Allergy 41, 281–291.
Davies, J.M., Li, H., Green, M., Towers, M., Upham, J.W., 2012. Subtropical grass pollen allergens are important for allergic respiratory diseases in subtropical regions. Clin. Transl. Allergy 2, 1.
Davies, J.M., Beggs, P.J., Medek, D.E., Newnham, R.M., Erbas, B., Thibaudon, M., et al., 2015. Trans-disciplinary research in synthesis of grass pollen aerobiology and its importance for respiratory health in Australasia. Sci. Total Environ. 534, 85–96.
Devanaboyina, S.C., Cornelius, C., Lupinek, C., Fauland, K., Dall'Antonia, F., Nandy, A., et al., 2014. High-resolution crystal structure and IgE recognition of the major grass pollen allergen Phl p 3. Allergy 69, 1617–1628.
Didier, A., Melac, M., Montagut, A., Lheritier-Barrand, M., Tabar, A., Worm, M., 2009. Agreement of efficacy assessments for five-grass pollen sublingual tablet immunotherapy. Allergy 64, 166–171.
de Dios Alché, J., M'Rani-Alaoui, M., Castro, A.J., Rodríguez-García, M.I., 2004. Ole e 1, the major allergen from olive (Olea europaea L.) pollen, increases its expression and is released to the culture medium during in vitro germination. Plant Cell Physiol. 45, 1149–1157.
Dodo, H.W., Konan, K.N., Chen, F.C., Egnin, M., Viquez, O.M., 2008. Alleviating peanut allergy using genetic engineering: the silencing of the immunodominant allergen Ara h 2 leads to its significant reduction and a decrease in peanut allergenicity. Plant Biotechnol. J. 6, 135–145.
Dykewicz, M.S., Hamilos, D.L., 2010. Rhinitis and sinusitis. J. Allergy Clin. Immunol. 125, S103–S115.
Ellis, A.K., Frankish, C.W., Armstrong, K., Larche, M., Steacy, L., Hafner, R., 2014. Treatment with synthetic peptide immuno-regulatory epitopes derived from grass allergens leads to a substantial reduction in grass allergy symptoms in the environmental exposure unit. J. Allergy Clin. Immunol. 133, AB290-AB.
Ellis, A., Frankish, C.W., Armstrong, K., Larche, M., Steacy, L., Hafner, R., et al., 2015. Persistent treatment effect with grass synthetic peptide immuno-regulatory epitopes in grass allergy symptoms in an environmental exposure unit challenge after a second season of natural pollen exposure. J. Allergy Clin. Immunol. 135, AB158-AB.
Feng, C., Yuan, J., Wang, R., Liu, Y., Birchler, J.A., Han, F.P., 2016. Efficient targeted genome modification in maize using CRISPR/Cas9 system. J. Genet. Genomics 43, 37–43.
Focke-Tejkl, M., Valenta, R., 2012. Safety of engineered allergen-specific immunotherapy vaccines. J. Genet. Genomics 12, 555–563.
Focke-Tejkl, M., Weber, M., Niespodziana, K., Neubauer, A., Huber, H., Henning, R., et al., 2015. Development and characterization of a recombinant, hypoallergenic, peptide-based vaccine for grass pollen allergy. J. Allergy Clin. Immunol. 135 (1207U614).
Franco, O.L., Rigden, D.J., Melo, F.R., Grossi-de-Sa, M.F., 2002. Plant α-amylase inhibitors and their interaction with insect α-amylases - structure, function and potential for crop protection. Eur. J. Biochem. 269, 397–412.
Fuhrmann, C.M., Sugg, M.M., Konrad, C.E., 2016. Airborne pollen characteristics and the influence of temperature and precipitation in Raleigh, North Carolina, USA (1999–2012). Aerobiologia 32, 683–696.
García-Maroto, F., Maraña, C., Mena, M., Garcia-Olmedo, F., Carbonero, P., 1990. Cloning of cDNA and chromosomal location of genes encoding the three types of subunits of the wheat tetrameric inhibitor of insect α-amylase. Plant Mol. Biol. 14, 845–853.
Gibbon, B.C., Zonia, L.E., Kovar, D.R., Hussey, P.J., Staiger, C.J., 1998. Pollen profilin function depends on interaction with proline-rich motifs. Plant Cell 10, 981–993.
Gil-Humanes, J., Piston, F., Hernando, A., Alvarez, J.B., Shewry, P.R., Barro, F., 2008. Silencing of γ-gliadins by RNA interference (RNAi) in bread wheat. J. Cereal Sci. 48, 565–568.
Gil-Humanes, J., Piston, F., Tollefsen, S., Sollid, L.M., Barro, F., 2010. Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. Proc. Natl. Acad. Sci. U. S. A. 107, 17023–17028.
Gil-Humanes, J., Piston, F., Shewry, P.R., Tosi, P., Barro, F., 2011. Suppression of gliadins results in altered protein body morphology in wheat. J. Exp. Bot. 62, 4203–4213.
Gil-Humanes, J., Pistón, F., Barro, F., Rosell, C.M., 2014. The shutdown of celiac diseaserelated gliadin epitopes in bread wheat by RNAi provides flours with increased stability and better tolerance to over-mixing. PLoS One 9, 11.
Golias, J., Humlova, Z., Halada, P., Habova, V., Janatkova, I., Tuckova, L., 2013. Identification of rice proteins recognized by the IgE antibodies of patients with food allergies. J. Agric. Food Chem. 61, 8851–8860.
Graf, L., Hayder, H., Mueller, U., 2014. Endogenous allergens in the regulatory assessment of genetically engineered crops. Food Chem. Toxicol. 73, 17–20.
Greiner, A.N., Hellings, P.W., Rotiroti, G., Scadding, G.K., 2011. Allergic rhinitis. Lancet 378, 2112–2122.
Grobe, K., Becker, W.M., Schlaak, M., Petersen, A., 1999. Grass group I allergens (βexpansins) are novel, papain-related proteinases. Eur. J. Biochem. 263, 33–40.
Hirano, K., Hino, S., Oshima, K., Okajima, T., Nadano, D., Urisu, A., et al., 2013. Allergenic potential of rice-pollen proteins: expression, immuno-cross reactivity and IgE-binding. J. Biochem. 154, 195–205.
Hrabina, M., Jain, K., Gouyon, B., 2008. Cross-reactivity between pollen allergens from common Pooideae grasses and cultivated cereals. Clin. Exp. Allergy Rev. 8, 18–20.
Irani, C., Karam, M., Baz, Z., Maatouk, H., Zaitoun, F., 2013. Airborne pollen concentrations and the incidence of allergic asthma and rhinoconjunctivitis in Lebanon. Rev. Fr. Allergol. 53, 441–445.
Iulek, J., Franco, O.L., Silva, M., Slivinski, C.T., Bloch, C., Rigden, D.J., et al., 2000. Purification, biochemical characterisation and partial primary structure of a new αamylase inhibitor from Secale cereale (rye). Int. J. Biochem. Cell Biol. 32, 1195–1204.
Kahlert, H., Grage-Griebenow, E., Stüwe, H.-T., Cromwell, O., Fiebig, H., 2000. T-cell reactivity with allergoids: influence of the type of APC. J. Immunol. 165, 1807–1815.
Kalamaki, M.S., Powell, A.L.T., Struijs, K., Labavitch, J.M., Reid, D.S., Bennett, A.B., 2003. Transgenic overexpression of expansin influences particle size distribution and improves viscosity of tomato juice and paste. J. Agric. Food Chem. 51, 7465–7471.
Kaur, R., Kaur, N., Gupta, A.K., 2014. Structural features, substrate specificity, kinetic properties of insect α-amylase and specificity of plant alpha-amylase inhibitors. Pestic. Biochem. Physiol. 116, 83–93.
Kerim, T., Imin, N., Weinman, J.J., Rolfe, B.G., 2003. Proteomic analysis reveals developmentally expressed rice homologues of grass group II pollen allergens. Funct. Plant Biol. 30, 843–852.
Kim, H., Kim, S.T., Ryu, J., Choi, M.K., Kweon, J., Kang, B.C., et al., 2016. A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system. J. Integr. Plant Biol. 58, 705–712.

Kovar, D.R., Drøbak, B.K., Staiger, C.J., 2000. Maize profilin isoforms are functionally distinct. Plant Cell 12, 583–598.
Le, L.Q., Lorenz, Y., Scheurer, S., Fotisch, K., Enrique, E., Bartra, J., et al., 2006a. Design of tomato fruits with reduced allergenicity by dsRNAi-mediated inhibition of ns-LTP (Lyc e 3) expression. Plant Biotechnol. J. 4, 231–242.
Le, L.Q., Mahler, V., Lorenz, Y., Scheurer, S., Biemelt, S., Vieths, S., et al., 2006b. Reduced allergenicity of tomato fruits harvested from Lyc e 1-silenced transgenic tomato plants. J. Allergy Clin. Immunol. 118, 1176–1183.
Lee, Y., Kende, H., 2001. Expression of β-expansins is correlated with internodal elongation in deepwater rice. Plant Physiol. 127, 645–654.
Lehto, M., Palosuo, K., Varjonen, E., Majuri, M.L., Andersson, U., Reunala, T., et al., 2003. Humoral and cellular responses to gliadin in wheat-dependent, exercise-induced anaphylaxis. Clin. Exp. Allergy 33, 90–95.
Li, L.C., Bedinger, P.A., Volk, C., Jones, A.D., Cosgrove, D.J., 2003. Purification and characterization of four β-expansins (Zea m 1 isoforms) from maize pollen. Plant Physiol. 132, 2073–2085.
Li, Y., He, Y., Zou, Z., 2012. Bioinformatics analysis on B cell epitopes of rice allergen RAG1. J. Agric. Sci. Technol. 13, 304–306.
Li, Q.L., Zhang, D.B., Chen, M.J., Liang, W.Q., Wei, J.J., Qi, Y.P., et al., 2016. Development of japonica photo-sensitive genic male sterile rice lines by editing carbon starved anther using CRISPR/Cas9. J. Genet. Genomics 43, 415–419.
Lullien, V., Alary, R., Guirao, A., Joudrier, P., Gautier, M.F., 1991. Isolation and nucleotide sequence of a cDNA clone encoding the bread wheat (Triticum aestivum L.) CM17 protein. Plant Mol. Biol. 17, 1081–1082.
Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., et al., 2015. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Mol. Plant 8, 1274–1284.
Ma, X.L., Zhu, Q.L., Chen, Y.L., Liu, Y.G., 2016. CRISPR/Cas9 platforms for genome editing in plants: developments and applications. Mol. Plant 9, 961–974.
Maeda, K., Hagglund, P., Finnie, C., Svensson, B., Henriksen, A., 2006. Structural basis for target protein recognition by the protein disulfide reductase thioredoxin. Structure 14, 1701–1710.
Major, I.T., Constabel, C.P., 2008. Functional analysis of the Kunitz trypsin inhibitor family in poplar reveals biochemical diversity and multiplicity in defense against herbivores. Plant Physiol. 146, 888–903.
Mandal, J., Chakraborty, P., Roy, I., Chatterjee, S., Gupta-Bhattacharya, S., 2008. Prevalence of allergenic pollen grains in the aerosol of the city of Calcutta, India: a two year study. Aerobiologia 24, 151–164.
Manzano, A.I., Javier Canada, F., Cases, B., Sirvent, S., Soria, I., Palomares, O., et al., 2016. Structural studies of novel glycoconjugates from polymerized allergens (allergoids) and mannans as allergy vaccines. Glycoconj. J. 33, 93–101.
Mares-Mejia, I., Martinez-Caballero, S., Garay-Canales, C., Cano-Sanchez, P., TorresLarios, A., Lara-Gonzalez, S., et al., 2016. Structural insights into the IgE-mediated responses induced by the allergens Hev b 8 and Zea m 12 in their dimeric forms. Sci. Rep. 6, 13.
Marknell DeWitt, A., Niederberger, V., Lehtonen, P., Spitzauer, S., Sperr, W.R., Valent, P., et al., 2002. Molecular and immunological characterization of a novel timothy grass (Phleum pratense) pollen allergen, Phl p 11. Clin. Exp. Allergy 32, 1329–1340.
Marowa, P., Ding, A., Kong, Y., 2016. Expansins: roles in plant growth and potential applications in crop improvement. Plant Cell Rep. 35, 949–965.
Marth, K., Focke, M., Flicker, S., Valenta, R., 2004. Human monoclonal antibody-based quantification of group 2 grass pollen allergens. J. Allergy Clin. Immunol. 113, 470–474.
Matsuo, H., Kohno, K., Morita, E., 2005. Molecular cloning, recombinant expression and IgEbinding epitope of omega-5 gliadin, a major allergen in wheat-dependent exercise-induced anaphylaxis. FEBS J. 272, 4431–4438.
Matsuo, H., Yokooji, T., Taogoshi, T., 2015. Common food allergens and their IgE-binding epitopes. Allergol. Int. 64, 332–343.
Medek, D.E., Beggs, P.J., Erbas, B., Jaggard, A.K., Campbell, B.C., Vicendese, D., et al., 2016. Regional and seasonal variation in airborne grass pollen levels between cities of Australia and New Zealand. Aerobiologia 32, 289–302.
Mehrabadi, M., Bandani, A.R., Mehrabadi, R., Alizadeh, H., 2012. Inhibitory activity of proteinaceous α-amylase inhibitors from Triticale seeds against Eurygaster integriceps salivary aα-amylases: interaction of the inhibitors and the insect digestive enzymes. Pestic. Biochem. Physiol. 102, 220–228.
Meng, L., Wong, J.H., Feldman, L.J., Lemaux, P.G., Buchanan, B.B., 2010. A membraneassociated thioredoxin required for plant growth moves from cell to cell, suggestive of a role in intercellular communication. Proc. Natl. Acad. Sci. U. S. A. 107, 3900–3905.
Moneret-Vautrin, D.A., 1997. Food allergens and their modifications by agro-food technology. Cah. Agric. 6, 21–29.
Moneret-Vautrin, D.A., Kanny, G., 2007. Future trends of research in food allergy: hypoallergenicity and vaccines. Rev. Fr. Allergol. 47, 305–313.
Müssar, K.J., Kandasamy, M.K., McKinney, E.C., Meagher, R.B., 2015. Arabidopsis plants deficient in constitutive class profilins reveal independent and quantitative genetic effects. BMC Plant Biol. 15, 16.
Nabavi, M., Hoseeinzadeh, Y., 2009. Anaphylaxis to wheat in an infant. Koomesh 11, Pe67–Pe70.
Nakamura, R., Matsuda, T., 1996. Rice allergenic protein and molecular-genetic approach for hypoallergenic rice. Biosci. Biotechnol. Biochem. 60, 1215–1221.
Nakase, M., Usui, Y., Alvarez-Nakase, A.M., Adachi, T., Urisu, A., Nakamura, R., et al., 1998. Cereal allergens: rice-seed allergens with structural similarity to wheat and barley allergens. Allergy 53, 55–57.
Nandy, A., Häfner, D., Klysner, S., 2015. Recombinant allergens in specific immunotherapy. Allergo J. Int. 24, 143–151.
Nelson, H.S., 2014. Oral/sublingual Phleum pretense grass tablet (Grazax/Grastek) to treat allergic rhinitis in the USA. Expert. Rev. Clin. Immunol. 10, 1437–1451.
Nony, E., Timbrell, V., Hrabina, M., Boutron, M., Solley, G., Moingeon, P., et al., 2015. Specific IgE recognition of pollen allergens from subtropic grasses in patients from the subtropics. Ann. Allergy Asthma Immunol. 114, 214–220.
Ogawa, T., Samoto, M., Takahashi, K., 2000. Soybean allergens and hypoallergenic soybean products. J. Nutr. Sci. Vitaminol. 46, 271–279.
Pahr, S., Constantin, C., Mari, A., Scheiblhofer, S., Thalhamer, J., Ebner, C., et al., 2012. Molecular characterization of wheat allergens specifically recognized by patients suffering from wheat-induced respiratory allergy. Clin. Exp. Allergy 42, 597–609.
Pasini, G., Simonato, B., Curioni, A., Vincenzi, S., Cristaudo, A., Santucci, B., et al., 2002. IgEmediated allergy to corn: a 50 kDa protein, belonging to the Reduced Soluble Proteins, is a major allergen. Allergy 57, 98–106.
Passalacqua, G., Canonica, G.W., 2016. Allergen immunotherapy history and future developments. Immunol. Allergy Clin. N. Am. 36, 1–12.
Pastorello, E.A., Farioli, L., Pravettoni, V., Ispano, M., Scibola, E., Trambaioli, C., et al., 2000. The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. J. Allergy Clin. Immunol. 106, 744–751.
Pastorello, E.A., Scibilia, J., Farioli, L., Primavesi, L., Giuffrida, M.G., Mascheri, A., et al., 2013. Rice allergy demonstrated by double-blind placebo-controlled food challenge in peachallergic patients is related to lipid transfer protein reactivity. Int. Arch. Allergy Immunol. 161, 265–273.
Patel, B.Y., Volcheck, G.W., 2015. Food allergy: common causes, diagnosis, and treatment. Mayo Clin. Proc. 90, 1411–1419.
Petersen, A., Schramm, G., Bufe, A., Schlaak, M., Becker, W.M., 1995. Structural investigations of the major allergen Phl p I on the complimentary-DNA and protein levels. J. Allergy Clin. Immunol. 95, 987–994.
Petersen, A., Suck, R., Hagen, S., Cromwell, O., Fiebig, H., Becker, W.M., 2001. Group 13 grass allergens: structural variability between different grass species and analysis of proteolytic stability. J. Allergy Clin. Immunol. 107, 856–862.
Petersen, A., Dresselhaus, T., Grobe, K., Becker, W.M., 2006. Proteome analysis of maize pollen for allergy-relevant components. Proteomics 6, 6317–6325.
Petrovska, N., Wu, X.L., Donato, R., Wang, Z.Y., Ong, E.K., Jones, E., et al., 2004. Transgenic ryegrasses (Lolium spp.) with down-regulation of main pollen allergens. Mol. Breed. 14, 489–501.
Pomés, A., 2008. Allergen structures and biologic functions: the cutting edge of allergy research. Curr. Allergy Asthma Rep. 8, 425–432.
Pytelkova, J., Hubert, J., Lepsik, M., Sobotnik, J., Sindelka, R., Krizkova, I., et al., 2009. Digestive α-amylases of the flour moth Ephestia kuehniella - adaptation to alkaline environment and plant inhibitors. FEBS J. 276, 3531–3546.
Radauer, C., Breiteneder, H., 2006. Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. J. Allergy Clin. Immunol. 117, 141–147.
Radauer, C., Bublin, M., Wagner, S., Mari, A., Breiteneder, H., 2008. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. J. Allergy Clin. Immunol. 121, 847–852.
Radauer, C., Pomes, A., Nandy, A., Ferreira, F., Rozynek, P., Raulf-Heimsoth, M., et al., 2012. Curation of the IUIS allergen database based on sequence similarity and protein family classification. Immunology 137, 442.
Radauer, C., Nandy, A., Ferreira, F., Goodman, R.E., Larsen, J.N., Lidholm, J., et al., 2014. Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. Allergy 69, 413–419.
Ramesh, S., 2008. Food allergy overview in children. Clin. Rev. Allergy Immunol. 34, 217–230.
Rihs, H.P., Rozynek, P., May-Taube, K., Welticke, B., Baur, X., 1994. Polymerase chain reaction based cDNA cloning of wheat profilin: a potential plant allergen. Int. Arch. Allergy Immunol. 105, 190–194.
Robertson, C.F., Dalton, M.F., Peat, J.K., Haby, M.M., Bauman, A., Kennedy, J.D., et al., 1998. Asthma and other atopic diseases in Australian children - Australian arm of the International Study of Asthma and Allergy in Childhood. Med. J. Aust. 168, 434–438.
Rogozhin, E.A., Ryazantsev, D.Y., Grishin, E.V., Egorov, T.A., Zavriev, S.K., 2012. Defense peptides from barnyard grass (Echinochloa crusgalli L.) seeds. Peptides 38, 33–40.
Rolland, J., O'Hehir, R., 1998. Immunotherapy of allergy: anergy, deletion, and immune deviation. Curr. Opin. Immunol. 10, 640–645.
Russell, S.D., Bhalla, P.L., Singh, M.B., 2008. Transcriptome-based examination of putative pollen allergens of rice (Oryza sativa ssp. japonica). Mol. Plant 1, 751–759.
Sampedro, J., Cosgrove, D.J., 2005. The expansin superfamily. Genome Biol. 6, 11.
Sampedro, J., Guttman, M., Li, L.C., Cosgrove, D.J., 2015. Evolutionary divergence of βexpansin structure and function in grasses parallels emergence of distinctive primary cell wall traits. Plant J. 81, 108–120.
Sander, I., Rozynek, P., Rihs, H.P., van Kampen, V., Chew, F.T., Lee, W.S., et al., 2011. Multiple wheat flour allergens and cross-reactive carbohydrate determinants bind IgE in baker's asthma. Allergy 66, 1208–1215.
Sandrini, A., Rolland, J.M., O'Hehir, R.E., 2015. Current developments for improving efficacy of allergy vaccines. Expert Rev. Vaccines 14, 1073–1087.
Santos, A., Van Ree, R., 2011. Profilins: mimickers of allergy or relevant allergens? Int. Arch. Allergy Immunol. 155, 191–204.
Scala, E., Alessandri, C., Bernardi, M.L., Ferrara, R., Palazzo, P., Pomponi, D., et al., 2010. Crosssectional survey on immunoglobulin E reactivity in 23 077 subjects using an allergenic molecule-based microarray detection system. Clin. Exp. Allergy 40, 911–921.
Schipper, O., Schaefer, D., Reski, R., Fleming, A., 2002. Expansins in the bryophyte Physcomitrella patens. Plant Mol. Biol. 50, 789–802.
Sekhar, B.C., Sachin, C., Raman, B.V., Bondili, J.S., 2015. Molecular characterization and In silico analysis of Sorghum Panallergens: Profilin and Polcalin. Indian J. Exp. Biol. 53, 726– 731.
Senna, G.E., Calderon, M., Milani, M., 2011. Allergy immunotherapy tablet: Grazax (R) for the treatment of grass pollen allergy. Expert. Rev. Clin. Immunol. 7, 21–27.
Shewry, P.R., Halford, N.G., Belton, P.S., Tatham, A.S., 2002. The structure and properties of gluten: an elastic protein from wheat grain. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 357, 133–142.
Shi, J.X., Cui, M.H., Yang, L., Kim, Y.J., Zhang, D.B., 2015. Genetic and biochemical mechanisms of pollen wall development. Trends Plant Sci. 20, 741–753.
Singh, M.B., Bhalla, P.L., 2008. Genetic engineering for removing food allergens from plants. Trends Plant Sci. 13, 257–260.
Sjölin, I., Haugaard, L., Kopp, T., Jansen, A., Brüning, H., Andersen, B.S., et al., 2008. High patient compliance on a once daily treatment regimen with the hay fever drug Grazax. J. Allergy Clin. Immunol. 121, S128-S.
Skypala, I., Vlieg-Boerstra, B., 2014. Food intolerance and allergy: increased incidence or contemporary inadequate diets? Curr. Opin. Clin. Nutr. Metab. Care 17, 442–447.
Smith, M., Jaeger, S., Berger, U., Sikoparija, B., Hallsdottir, M., Sauliene, I., et al., 2014.

Geographic and temporal variations in pollen exposure across Europe. Allergy 69, 913–923.
Staiger, C.J., Goodbody, K.C., Hussey, P.J., Valenta, R., DrØbak, B.K., Lloyd, C.W., 1993. The profilin multigene family of maize: differential expression of 3 isoforms. Plant J. 4, 631–641.
Suck, R., Petersen, A., Hagen, S., Cromwell, O., Becker, W.M., Fiebig, H., 2000. Complementary DNA cloning and expression of a newly recognized high molecular mass allergen Phl p 13 from timothy grass pollen (Phleum pratense). Clin. Exp. Allergy 30, 324–332.
Sun, T., Li, S., Ren, H., 2013. Profilin as a regulator of the membrane-actin cytoskeleton interface in plant cells. Front. Plant Sci. 4.
Suphioglu, C., Singh, M.B., Taylor, P., Bellomo, R., Holmes, P., Puy, R., et al., 1992. Mechanism of grass-pollen-induced asthma. Lancet 339, 569–572.
Tada, Y., Nakase, M., Adachi, T., Nakamura, R., Shimada, H., Takahashi, M., et al., 1996. Reduction of 14–16 kDa allergenic proteins in transgenic rice plants by antisense gene. FEBS Lett. 391, 341–345.
Takagi, H., Saito, S., Yang, L.J., Nagasaka, S., Nishizawa, N., Takaiwa, F., 2005. Oral immunotherapy against a pollen allergy using a seed-based peptide vaccine. Plant Biotechnol. J. 3, 521–533.
Trcka, J., Schaed, S.G., Scheurer, S., Conti, A., Vieths, S., Gross, G., et al., 2012. Riceinduced anaphylaxis: IgE-mediated allergy against a 56-kDa glycoprotein. Int. Arch. Allergy Immunol. 158, 9–17.
Valdivia, E.R., Durachko, D.M., Wagner, E., Stephenson, A.G., Cosgrove, D.J., 2003. βexpansins in maize pollen, role of Zea m 1 in development and fertilization. Plant Biol. 2003, 121.
Valdivia, E.R., Wu, Y., Li, L.-C., Cosgrove, D.J., Stephenson, A.G., 2007. A group-1 grass pollen allergen influences the outcome of pollen competition in maize. PLoS One 2.
Valdivia, E.R., Stephenson, A.G., Durachko, D.M., Cosgrove, D., 2009. Class B β-expansins are needed for pollen separation and stigma penetration. Sex. Plant Reprod. 22, 141–152.
Valenta, R., Duchene, M., Pettenburger, K., Sillaber, C., Valent, P., Bettelheim, P., et al., 1991. Identification of profilin as a novel pollen allergen; IgE autoreactivity in sensitized individuals. Science 253, 557–560.
Valenta, R., Vrtala, S., Ebner, C., Kraft, D., Scheiner, O., 1992. Diagnosis of grass pollen allergy with recombinant Timothy grass (Phleum pratense) pollen allergens. Int. Arch. Allergy Immunol. 97, 287–294.
Valenta, R., Ball, T., Vrtala, S., Duchene, M., Kraft, D., Scheiner, O., 1994. cDNA cloning and expression of Timothy grass (Phleum pratense) pollen profilin in Escherichia coli:

comparison with birch pollen profilin. Biochem. Biophys. Res. Commun. 199, 106–118.
Valenta, R., Campana, R., Focke-Tejkl, M., Niederberger, V., 2016. Vaccine development for allergen-specific immunotherapy based on recombinant allergens and synthetic allergen peptides: lessons from the past and novel mechanisms of action for the future. J. Allergy Clin. Immunol. 137, 351–357.
Veeraraghavan, G., Leffler, D.A., Kaswala, D.H., Mukherjee, R., 2015. Celiac disease 2015 update: new therapies. Expert Rev. Gastroenterol. Hepatol. 9, 913–927.
Vidal, N., Barbosa, H., Jacob, S., Arruda, M., 2015. Comparative study of transgenic and nontransgenic maize (Zea mays) flours commercialized in Brazil, focussing on proteomic analyses. Food Chem. 180, 288–294.
Wakasa, Y., Hirano, K., Urisu, A., Matsuda, T., Takaiwa, F., 2011. Generation of transgenic rice lines with reduced contents of multiple potential allergens using a null mutant in combination with an RNA silencing method. Plant Cell Physiol. 52, 2190–2199.
Wang, Z., Guo, C.-Q., Ren, H.-Y., 2006. Profilins in plant cells. Zhi Wu Sheng Li Yu Fen Zi Sheng Wu Xue Xue Bao 32, 261–270.
Wang, S.Y., Takahashi, H., Kajiura, H., Kawakatsu, T., Fujiyama, K., Takaiwa, F., 2013. Transgenic rice seeds accumulating recombinant hypoallergenic birch pollen allergen Bet v 1 generate giant protein bodies. Plant Cell Physiol. 54, 917–933.
Wang, J., Yang, L., Zhao, X., Li, J., Zhang, D., 2014. Characterization and phylogenetic analysis of allergenic Tryp_alpha_amyl protein family in plants. J. Agric. Food Chem. 62, 270–278.
Weegels, P.L., Hamer, R.J., Schofield, J.D., 1996. Functional properties of wheat glutenin. J. Cereal Sci. 23, 1–17.
Wheatley, L.M., Togias, A., 2015. Allergic rhinitis. N. Engl. J. Med. 372, 456–463.
White, J.F., Bernstein, D.I., 2003. Key pollen allergens in North America. Ann. Allergy Asthma Immunol. 91, 425–435.
Winther, L., Arnved, J., Malling, H.J., Nolte, H., Mosbech, H., 2006. Side-effects of allergenspecific immunotherapy. A prospective multi-centre study. Clin. Exp. Allergy 36, 254–260.
Woods, R.K., Stoney, R.M., Raven, J., Walters, E.H., Abramson, M., Thien, F.C.K., 2002. Reported adverse food reactions overestimate true food allergy in the community. Eur. J. Clin. Nutr. 56, 31–36.
Wu, Y.J., Meeley, R.B., Cosgrove, D.J., 2001. Analysis and expression of the α-expansin and βexpansin gene families in maize. Plant Physiol. 126, 222–232.
Wu, X.L., Cai, G., Gong, F.P., An, S.F., Cresti, M., Wang, W., 2015. Proteome profiling of maize pollen coats reveals novel protein components. Plant Mol. Biol. Report. 33, 975–986.
Xu, H.L., Theerakulpisut, P., Goulding, N., Suphioglu, C., Singh, M.B., Bhalla, P.L., 1995. Cloning, expression and immunoligical characterization of Ory s 1, the major allergen of rice pollen. Gene 164, 255–259.
Ye, Q., Xu, Y., Yan, F., Tang, L., Chen, F., 2001. Molecular cloning and characterization of rice pollen profilin. Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao 33, 452–456.
Yoo, S.D., Gao, Z.F., Cantini, C., Loescher, W.H., van Nocker, S., 2003. Fruit ripening in sour cherry: changes in expression of genes encoding expansins and other cell-wallmodifying enzymes. J. Am. Soc. Hortic. Sci. 128, 16–22.
Yu, J.M., Goktepe, I., Ahmedna, M., 2013. Enzymatic treatment of peanut butter to reduce the concentration of major peanut allergens. Int. J. Food Sci. Technol. 48, 1224–1234.
Zhang, D.S., Liang, W.Q., Yuan, Z., Li, N., Shi, J., Wang, J., et al., 2008. Tapetum Degeneration Retardation is critical for aliphatic metabolism and gene regulation during rice pollen development. Mol. Plant 1, 599–610.
Zhang, D.S., Liang, W.Q., Yin, C.S., Zong, J., Gu, F.W., Zhang, D.B., 2010. OsC6, encoding a lipid transfer protein, is required for postmeiotic anther development in rice. Plant Physiol. 154, 149–162.
Zhang, H., Liang, W.Q., Yang, X.J., Luo, X., Jiang, N., Ma, H., et al., 2010. Carbon Starved Anther encodes a MYB domain protein that regulates sugar partitioning required for rice pollen development. Plant Cell 22, 672–689.
Zhang, H., Xu, C.X., He, Y., Zong, J., Yang, X.J., Si, H.M., et al., 2013. Mutation in CSA creates a new photoperiod-sensitive genic male sterile line applicable for hybrid rice seed production.

Proc. Natl. Acad. Sci. U. S. A. 110, 76–81.

Statement of Authorship

Title of Paper	Chapter 2 – Evolutionary analysis of two rice pollen allergens, Ory s 1 and Ory s 12
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Deborah Devis		
Contribution to the Paper	Conducted research and analysis, Wrote, and edited paper		
Overall percentage (%)	90		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25/07/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

15/07/2019

Name of Co-Author	Gang Li		
Contribution to the Paper	Provided Editing feedback and technical advice		
Signature		Date	09/08/2019

Name of Co-Author	Dabing Zhang		
Contribution to the Paper	Helped with writing and framing of paper and editing		
Signature		Date	09/08/2019

Name of Co-Author	Janet Davies		
Contribution to the Paper	Provided Editing feedback		
Signature		Dat e	19/06/2019

Chapter 2 – Evolutionary analysis of two rice pollen allergens, Ory s 1 and Ory s 12

2.1 Abstract

Hay fever is a globally relevant disease that is caused by allergenic proteins, called allergens. Grass allergens belong to multiple gene families and little is known about who these seemingly unrelated gene families evolved allergenicity. In this chapter, I assessed the phylogenetic and structural relationship between two rice pollen allergen gene families, group 1 and group 12. I found that group 1 genes evolved multiple gene copies in most monocots following speciation and were abundant in pollen, but these genes were not present in dicots. The group 1 allergens are considered the major pollen allergen. However, Group 12 genes were present in most plant species and appeared to have arisen due to an ancient duplication in a profilin gene. These new profilins later diverged into vegetatively expressed and reproductively expressed profilins, that latter of which may contribute to grass pollen allergy. There did not appear to be a correlation between the evolutionary paths of Group 1 or Group 12 allergens, suggesting these two groups evolved allergenicity independently.

2.2 Introduction

Hay fever is a globally relevant disease caused by allergen-specific human Immunoglobulin E (IgE) reacting to proteins in pollen called allergens. Symptoms of hay fever manifest as sneezing, itching and reddening of eyes, disrupted sleep, and general discomfort, but this disease often cannot be avoided due to the abundance of pollen in the air during spring and summer. The most common family of plants that cause hay fever are Poaceae, many of which are grasses.

Poaceae include important agricultural species such as rice (*Oryza sativa*), wheat (*Triticum aestivum*), maize (*Zea mays)* and barley (*Hordeum vulgare)*, and also include wide-spread wild species that cover a huge amount of the Earths land area (Barker et al., 2001). Poaceae became

widespread during the cretaceous period (Prasad et al., 2005). Hay fever can be caused by many different grass species, but little is known about the evolution of allergen protein families in domesticated and agriculturally important species, such as rice. A high number of putative allergens are conserved in grasses, and 107 allergens were predicted to be in grass pollen alone (Chen et al., 2016). Many of these proteins may be involved in crucial functions, such as stress response and various metabolic processes, suggesting that these proteins play an important role in pollen development (Chen et al., 2016).

Domesticated rice cultivars contain two pollen allergens as defined by World Health Organization/International Union of Immunological Societies (WHO/IUIS) (Goodman and Breiteneder, 2019, Pomes et al., 2018): Ory s 1, a homolog of the β-expansin family, and Ory s 12 is a member of the ancient profilin family. Interestingly, these two allergens have different putative functions and evolutionary changes (Chen et al., 2016, Devis et al., 2017).

Ory s 1, or OsEXPB1 as it will be referred to in this paper, is part of the Group 1 allergens, which is a family homologous to pollen expressed β-expansins that are frequently involved in cell wall expansion (Cosgrove et al., 1997, Sampedro and Cosgrove, 2005). Homologous genes have high expression in maize and Brachypodium (*Brachypodium distachyon*) pollen, and Brachypodium contains 4 copies within the Group 1 family, suggesting that grasses may require a high number of these proteins for proper function

(Sharma et al., 2017, Valdivia et al., 2009). β-expansins appear to have arisen before grass species diverged from each other, but underwent rapid evolution following this event, so an understanding of the mechanisms of this rapid evolution and its causes may lead to greater management of allergen genes in the future (Chen et al., 2016, Sampedro et al., 2015).

Conversely, Ory s 12 is homologous to the OsPROFILIN A (OsPRFA), which belongs to the ancient profilin family. Profilins are associated with actin-binding, and have been conserved between higher and lower plants (Chen et al., 2016, Devis et al., 2017, Huang et al., 1996, Sharma et al., 2017). The

different evolutionary patterns of these two allergen-containing protein families may shed light onto how allergenicity is conserved and changed over time.

Due to the abundance of putative allergens in other species and the varied evolution of allergens within a single species, a solid understanding of grass pollen allergen evolution will help in identifying new allergens and understanding their functions and allergenicity. In this study, I analyse the evolutionary history and important structures in rice pollen allergens Ory s 1 and Ory s 12 in order to have a greater understanding of the origin of these genes and what sequences are associated with their allergenicity.

2.3 Materials and Methods

2.3.1 Identification of gene homologs

Sequences of OsEXPB and OsPRFA proteins were collected from the World Health Organization/International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Database, based on the sequences of Ory s 1.0101 and Ory s 12.0101 respectively (Radauer et al., 2014). Allergome was used to asses putative allergens that were not characterised by WHO/IUIS (Mari et al., 2006). Homologous protein sequences were collected using NCBI Basic local Alignment Search Tool (Altschul et al., 1990),
Phytozome (Altschul et al., 1990, Bairoch et al., 2007, Goodstein et al., 2012) and the
Universal Protein Resource (Bairoch et al., 2007), and sequences with amino acid identity more than 50% were used for further analysis. Gene names were assigned based on phylogenetic relationship, and designated gene names and corresponding unique locus identifiers are shown in supplementary tables 1 and 2.

2.3.2 Construction of phylogenetic trees, secondary structure prediction and expression analysis

Protein sequences were aligned using MUSCLE (Edgar, 2004). For OsEXPB, I used 59 protein sequences from Ryegrass (*Lolium perenne*), Timothy grass (*Phleum patense*), Wheat (*Triticum aestivum*), Barley (*Hordeum vulgare*), Brachypodium (*Brachypodium disachyon*), Maize (*Zea mays*), Johnson grass (*Sorghum halepense*), Sorghum (*Sorghum bicolor*), Panic grass (*Panicum hallii*), Palm (*Elaeis guineensis*), and Pineapple (*Ananas comosus*) to construct a Bayesian unrooted phylogenetic tree (Poisson model). The αexpansin OsEXPA4 was used as an outgroup. For OsPRFAa, I used PHYML to construct an unrooted maximum likelihood (1000 bootstrap replications) tree with 55 amino acid sequences from rice, bermuda grass (*Cynodon dactylon*), sorghum, maize, panic grass, timothy grass, barley, Brachypodium, Arabidopsis (*Arabidopsis thaliana*), wheat, pineapple, cherry (*Prunus avium*), hazelnut (*Corylus avellana*), Olive (*Olea europaea*), peanut (*Arachis hypogaea*), Birch (*Betula verrucosa*), sunflower (*Helianthus annuus*), tomato (*Solanum lycopersicum*) and moss (*Physcomitrella patens*). The moss protein was used as an outgroup.

Secondary structure of protein sequences was predicted using the Garnier Osguthorpe Robson algorithm provided by EMBOSS.

Gene expression for *OsEXPB* and *OsPRF* in root, leaf, shoot, inflorescence, anther and grain or fruit was predicted using online data from Phytozome, the Rice Genome Annotation project, Barlex, exVIP, PLEXdb, sunflowergenome.org, and Tomato Expression Atlas (Borrill et al.,

2016, Dash et al., 2011, Fernandez-Pozo et al., 2017, Ouyang et al., 2007). Gene expression analysis and visualisation was carried out using Multi experiment Viewer (Howe et al., 2011).

2.3.3 Synteny Analysis

Synteny analysis was carried out on CoGepedia using Os03g01640 as a reference for *OsEXPB* against *Z. mays* chromosomes 2,5 and 9, S. *bicolor* chromosome 1, *P. hallii* chromosome 1 and 3

(Lyons and Freeling, 2008, Lyons et al., 2008). Os10g17660 was used as a reference for *OsPRF* against *Z. mays* chromosomes 2, 6 and 8, *O. sativa* chromosome 6, *P. hallii* chromosomes 3, 4 and 5 and *S. bicolor* chromosomes 3, 9 and 10. Synteny analysis was carried out using 100kb upstream and downstream of the target reference.

2.3.4 Domain identification and 3D modelling

Domain identification was carried out using SMART and PFAM (El-Gebali et al., 2018, Letunic et al., 2015). OsPRFA tertiary structure was predicted with SWISS-PROT and modelled in Pymol.

3D structures were predicted using SWISS-MODEL and visualised is Pymol
(Biasini et al., 2014, Janson et al., 2016). OsEXPB was modelled against Zea m 1 (SMTL
ID : 2hcz.1). OsPRF was modelled against profilin homolog and pollen allergen Bet v 2 (SMTL ID : 5nzb.1). Sequence logo was created on Geneious using protein sequences from all sequences used in phylogenetic analysis except OsEXPA4, PHpPRFLP04 and PHpPRF2-like proteins (Kearse et al., 2012). Cysteine residues, Poly-L-Proline binding sites, PIP2 interaction sites and Actin interaction sites were annotated based on the CDD database (Marchler-Bauer et al., 2017). Putative IgE binding sites were taken from literature (Fedorov et al., 1997, Levin et al., 2013).

2.4 Results

2.4.1 Evolutionary analysis of *OsEXPB* homologs

To understand how *OsEXPBs* evolved, a BLAST was performed on an Ory s 1 sequence procured from WHO/IUIS Allergen Nomenclature database (allergenicity reference-7590339). Four homologs of Ory s 1 genes were found through sequence homology analysis in the rice genome and were named *OsEXPB10* (LOC_Os03g01610), *OsEXPB13* (LOC_Os03g01630),
OsEXPB1a (LOC_Os03g01640) and *OsEXPB1b* (LOC_Os03g01650) respectively. Group 1 β–expansin homologs were found only in monocot species. Phylogenetic analysis showed two clades of β–expansin homologs in monocots, which correlated with the groups Divergent Expansin B (Pollen)

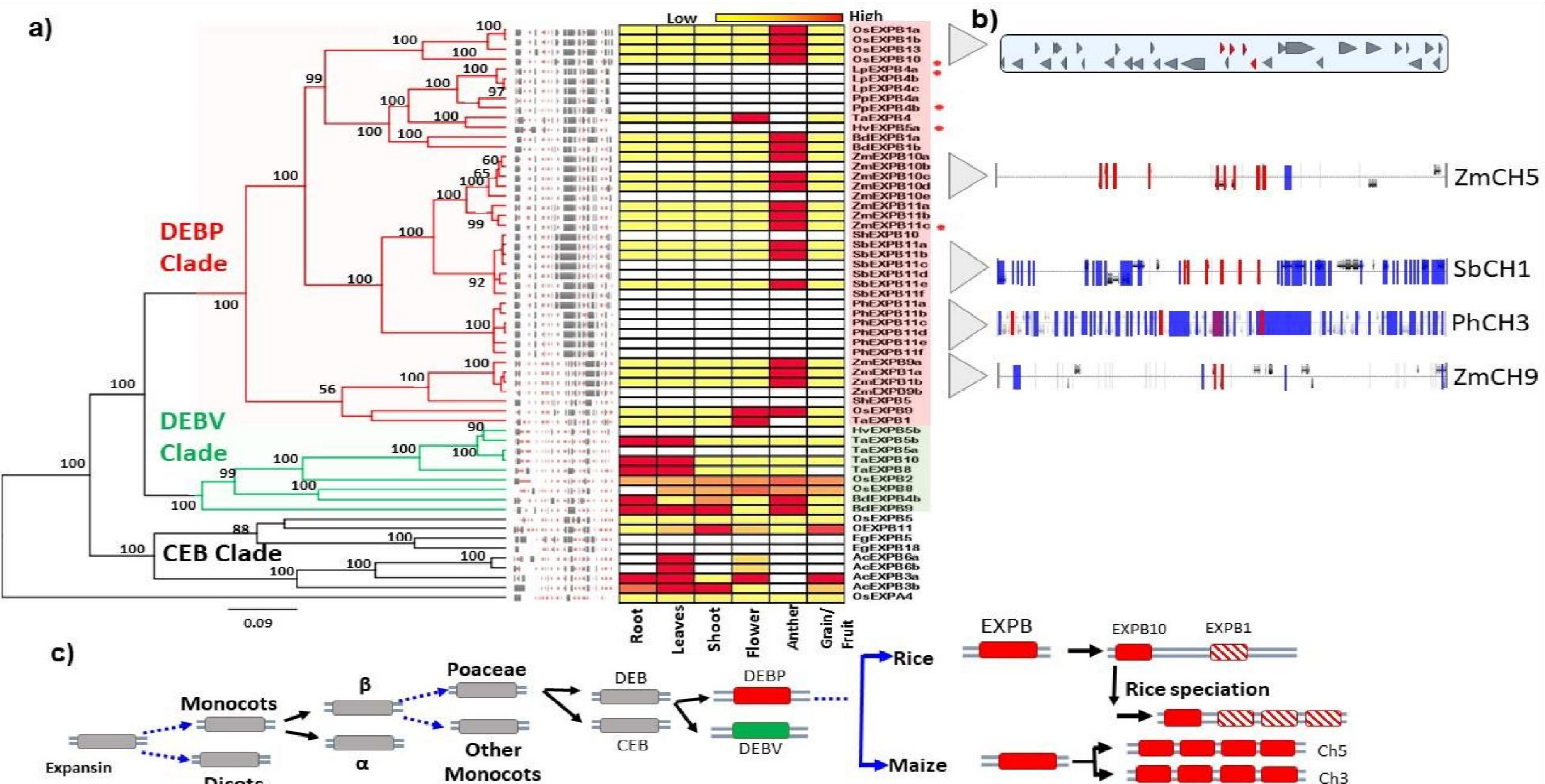


Figure 2.1 – Phylogenetic tree and model of expansin evolution

Baysean phylogenetic tree of OsEXPB1 homologs, including rice (*Oryza sativa*), Ryegrass (*Lolium perenne*), Timothy grass (*Phleum patense*), Wheat (*Triticum aestivum*), Barley (*Hordeum vulgare*), Brachypodium (*Brachypodium disachyon*), Maize (*Zea mays*), Johnson grass (*Sorghum halepense*), Sorghum (*Sorghum bicolor*), Panic grass (*Panicum hallii*), Palm (*Elaeis guineensis*), and Pineapple (*Ananas comosus*). Corresponding amino acid structure show α-helices (grey) and β-strands (pink). Relative expression of Root, Leaves, Shoot, Inflorescence, Anther and Grain/Fruit is shown in red (high expression) and yellow (low expression). Known allergens are marked with a red star. **b)** Synteny between 10,000kb region of rice chromosome 3 and regions surrounding Panic grass CH3, Sorghum Ch1, and Maize CH5 and Ch9. Syntenic regions are shown in blue, EXPBs are shown in Red, and genes are shown in grey. **c)** Model of expansin gene (grey box) evolution into DEBV (greeb box) and DEBP genes (red box). Blue arrows show speciation, black arrows show divergence in a single genome.

(**DEBP**), Divergent Expansin B (Vegetative) (**DEBV**) and Conservative Expansin B (**CEB**) groups (Figure 2.1). *OsEXPB10*, *13*, *1a* and *1b* genes were grouped together in the DEBP clade with other grass homologs that appeared to be strongly expressed in reproductive tissue, such as the anther and pollen (Figure 2.1), whereas genes in the DEBV group had lower expression in the anther but were expressed in other tissue such as shoots and roots. The vegetatively expressed CEB expansin homologs were grouped in a separate clade and were found in non-grass monocots. Rice, maize *(ZmEXPB*), sorghum (*SbEXPB* and *ShEXPB*), panic grass (*PhEXPB*), and ryegrass (*LpEXPB*) had multiple gene copies more closely related within the species than between the species. The Group 1 allergen group consisted of 4 *OsEXPB* genes that were found in tandem on chromosome 3. Synteny was found between rice chromosome 3, panic grass chromosome 3 and Sorghum chromosome 1 (Figure 2.1). These results indicated that the *OsEXPB* homologs are specifically present in monocot plants and the members show the various expression pattern between vegetative and reproductive tissues. Little synteny was found between rice chromosomes and maize chromosomes surrounding the *osEXPB* homologs, potentially due to the larger genome size of maize.

2.4.2 The evolution of *OsEXPBs* in rice

To understand the evolutionary events relating to *OsEXPBs* during rice domestication, phylogenetic analysis was performed on these genes from domesticated and wild rice species. Different gene copies of *OsEXPBs* were observed in rice species: four in *Oryza brachyantha* (wild rice), *Oryza glaberrima* (domesticated African rice), and *Oryza punctata* (red rice), and three in *Oryza officinalis* (Figure 2.2a). Rice genes fell into two clades and each species, except *O. brachyantha*, contained genes from both clades. *OsEXPB10* was grouped in a separate clade from *OsEXPB1a, OsEXPBb* and *OsEXPB13*, suggesting that a gene duplication occurred on chromosome 3 prior to grass species divergence, resulting in at least two *OsEXPB* genes, and two extra copies were accumulated following speciation. *O. brachyantha* contained three *EXPB* copies that fell into a separate clade. It is likely that *O. brachyantha* diverged before the the other rice species due to its unique placement on he phylogenetic tree.

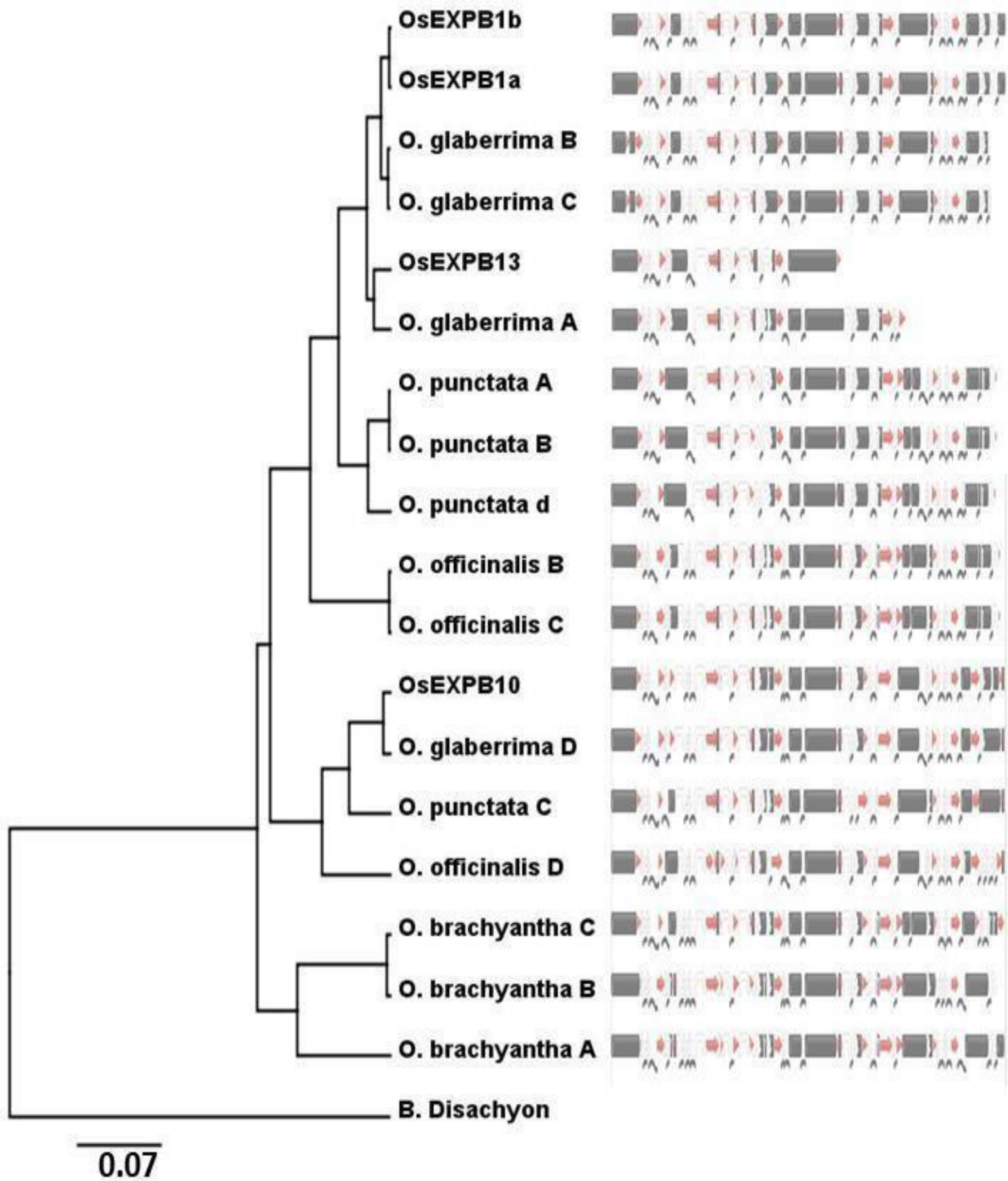


Figure 2.2 - Phylogenetic tree of domesticated rice cultivars and wild rice

Baysean phylogenetic tree containing domesticated *Oryza sativa, Oryza glaberrima,* and wild *Oryza punctate, Oryza officinalis* and *Oryza brachyantha*. Corresponding amino acid structure shows α-helices (grey) and β-strands (pink).

.

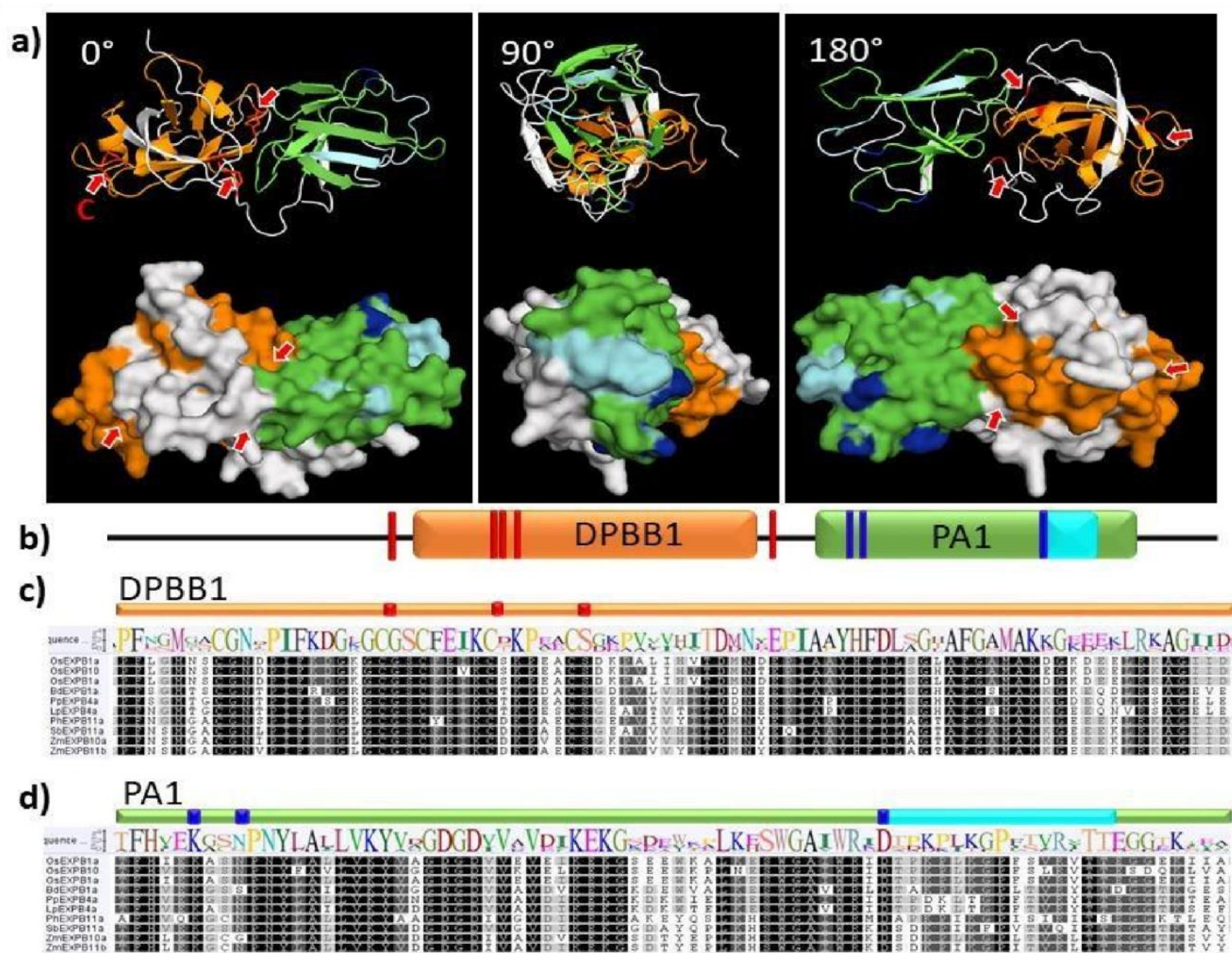


Figure 2.3-Tertiary structure of EXPB1

3D tertiary structure of EXPB1 displayed as ribbons or surface, modelled against Zea m 1. The protein contains the DPBB domain (orange), the Pollen allergen 1 domain (green), cysteine residues (red and red arrow) Phl p1 IgE binding hotspots (dark blue) and C yn d 1 predicted IgE binding site (cyan). b) Representation of the conserved domains between group 1 allergens. The protein contains the DPBB domain (orange), the Pollen allergen 1 domain (green), cysteine residues (red) Phl p1 IgE binding hotspots (dark blue) and Cyn d 1 predicted IgE binding site (cyan). Sequence alignment of the c) DPBB1 domain d) PA1 domain between expansin homologs OsEXPB10, OsEXPB1a, BdEXPB1a, PpEXPB4a, LPEXPB4a, PhEXPB11a, SbEXPB11a, ZmEXPB10a and ZmEXPB11b. Conserved amino acids are black, semi-conserved amino acids are grey. The amino acid logo above shows conserved amino acids as large letters and semi-conserve amino acids as small letters.

2.4.3 Structural analysis of OsEXPB homologs

The group 1 *OsEXPBs* had 804 nucleotides in length and contained no introns, resulting in a protein of 267 amino acids in length that had two domains; the N-terminal doublepsi betabarrel fold (**DPBB**) domain and the C-terminal pollen allergen 1 (**PA1**) domain (Figure 2.3).

Six cysteine residues formed three cysteine bridges, which were found in the DPBB1 domain and buried in the core of the protein. No cysteine residues were found in the pollen allergen 1 domain (Figure 2.3b). The OsEXPB1a motif DTPKPLKGPFSVRVT shares

70% homology with a monoclonal antibody binding site in the Bermuda grass allergen, Cyn d 1 (Figure 2.3b) (Tiwari et al., 2009, Yuan et al., 2012). This sequence is also conserved in OsEXPB10 with the exception of a valine to leucine substitution of the first valine. However, valine and leucine are both hydrophobic and may have little or no effect on binding efficiency (Smith et al., 1994, Xu et al., 1995) (Figure 2.3b).This binding site was not well retained between grasses, with the exception of a proline at position 225 (P225) and a threonine at position 237 (T237) (Figure 2.3d).

Three Phl p 1 IgE-binding hotspot homologs were identified in rice at positions K176, N179 and D223 (Levin et al., 2013). All three residues were located in the PA1 domain and D223 was retained in all monocot β–expansins. N179 was retained in most grass β– expansins but was replaced with a glycine in maize Zea m 1 homologs located on Chromosome 5. The Cyn d 1 binding site and the Phl p 1 binding hotspots were located on the surface of the protein (Figure 2.3a).

OsEXPB13 had a polymorphism at nucleotide 466, resulting in an A to T substitution. This substitution resulted in a mutation from a glutamine to a stop codon and a truncated amino acid of 155 residues instead of 267 residues (Supplementary Figure 1).

2.4.4 Evolutionary analysis of OsPRFA homologs

Phylogenetic analysis showed *OsPRFA* grass homologs fell into two clades: One group contained 2 genes from rice, and genes from Bermuda grass, sorghum, maize, panic grass, timothy grass, brachypodium, and barley and some dicots. The second group contained *OsPRFc* and genes from both monocots and dicots (Figure 2.4a). Three *OsPRFA* homologs were found

with more than 85% identity at the amino acid level. Genes were named based on their grouping clade in a phylogenetic tree. *OsPRFAa* (Os10g17660) and *OsPRFAb* (Os10g17680) seemed to be arranged in tandem on chromosome 10 and shared 100% amino acid identity and 98.5% identity of

the nucleotide sequences of the coding region. While *OsPRFAc* (Os06g05880) was found on chromosome 6 and shared 85% amino acid identity with OsPRFAa and OsPRFAb . OsPRFAa and OsPRFAb were found to be expressed only in reproductive tissue, as did other grass genes that fell into the same clade (Figure 2.4a). However, *OsPRFc* and other members of the same clade were found to be expressed in vegetative tissue. Therefore, these monocot profilin genes fell into two clades; reproductively expressed monocot and vegetatively expressed profilins, which may have arisen and diverged before monocots speciated. Synteny was found between chromosomes 10 and 6, which contain *OsPRFAa* and *b*, and *OsPRFAc* respectively, suggesting that a chromosomal duplication lead to two copies of a profilin gene, which later diverged into reproductively and vegetatively expressed copies (Figure 2.4b).

2.4.5 Structural analysis of OsPRF homologs

OsPRFA homologs shared above 50% amino acid identity between species, suggesting profilin amino acid structure experiences high selective pressure to be retained, regardless of species or expression pattern. OsPRFAa was predicted to fold into a small protein consisting of two outer α-helices joined by one β-sheet (Figure 2.5a) Likewise, many actin binding sites and poly-Lproline sites (both involved in actin organisation) were retained, potentially contributing to a crucial actin interaction function of profilins (Figure 2.5b). The birch pollen profilin had two putative IgE binding sites; one binding site near the N terminus and one near the C terminus (Fedorov et al., 1997). These sites were also retained in rice OsPRFAa and OsPRFAb and are located near each other on the surface of the folded protein, due to a shared cysteine bridge present in all species. Rice poly L-proline binding sites were conserved between species, and any amino acid substitutions resulted in similar (e.g. hydrophobic to hydrophobic). The proline accumulates in pollen grains and also promotes stress tolerance (Verbruggen and Hermans, 2008).

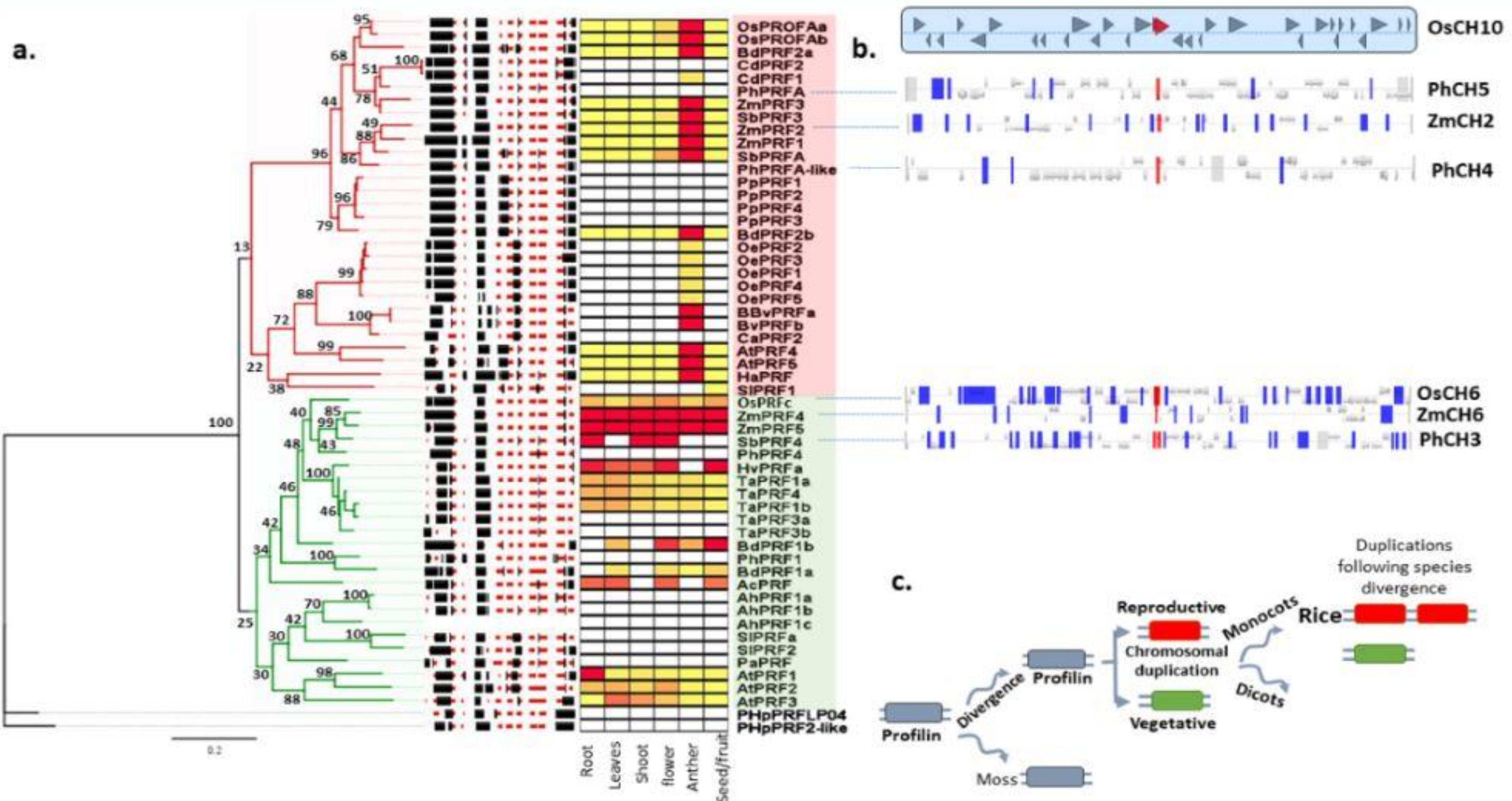


Figure 2.4. - Phylogenetic tree of Profilin evolution

Unrooted phylogenetic tree of *OsPRFA* homologs including rice, bermuda grass (Cd), sorghum (Sb), maize (Zm), panic grass (Ph), timothy grass (Pp), barley (Hv), Brachypodium (Bd), Arabidopsis (At) wheat (Ta), pinapple (Ac), cherry (Pa), hazelnut (Ca), Olive (Oe), peanut (Ah), Birch (bv), sunflower (Ha), tomato (Sl) and moss (PHp) in the reproductive clade (red) and Vegetative clade (green). Amino acid structure (right of phyotgenetic tree branches) shows α-helices (black box) and β-sheets (red pox). Expression heat map shows high expression in red and low expression in yellow. b) Synteny between 10,000kb region of rice chromosome 10 and regions surrounding *PhPRFA*, *ZmPRF2*, *PhPRFA-like*, *OsPRFc*, *ZmPRF4* and *SbPRF4* . Syntenic regions are shown in blue, Profilins are shown in Red, and genes are shown in grey. c) Model of profilin gene (Box and lines) evolution into reproductive (red) and vegetative (green) profilins. Grey arrows represent divergence.

Predicted IgE interaction sites from birch and rice were also seen in sunflower, where they were located on α- helices (Asturias et al., 2002, Fedorov et al., 1997). Six actin interaction sites were found in the IgE interaction region, and other actin interaction sites were found in close proximity at the surface of the protein. These amino acids were conserved in most species, and any amino acid changes between species resulted in an aliphatic to aliphatic substitution, except in the case of TaPRF1a and AtPRF4, which contained a Leucine instead of Met117.

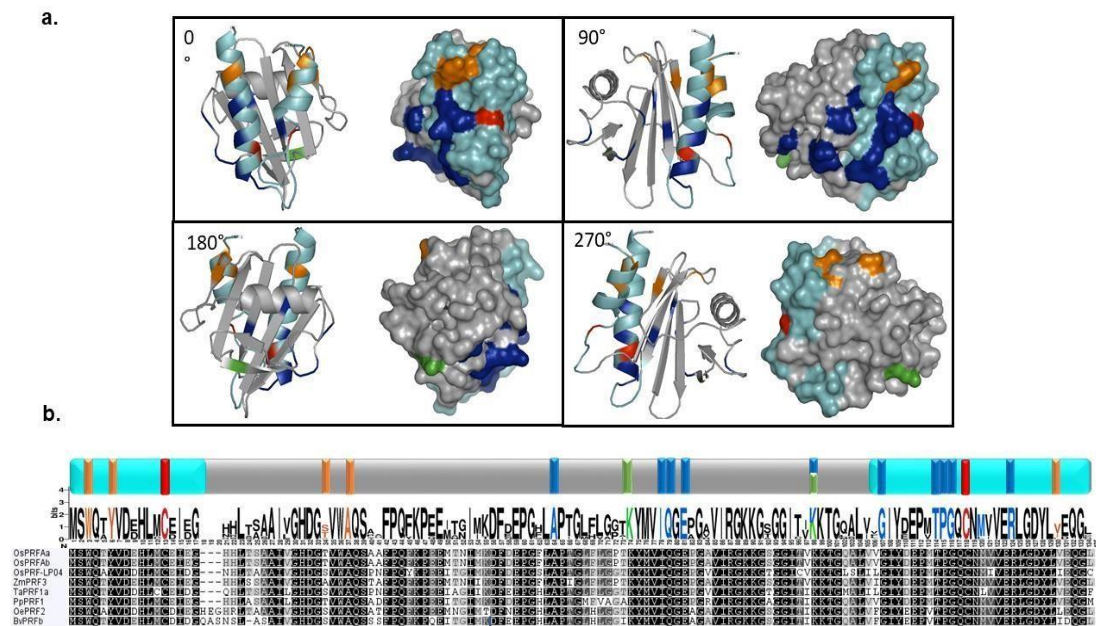


Figure 2.5a– Teriary Structure of profilins

Predicted Ory s 12 based on ZmPRF, with tertiary structure containing a-helices (Helix) and bsheets (Arrows) Pol L-proline binding sites (orange), cysteine residues (red), Actin interaction sites (dark blue), PIP2 interaction sites (green) and Birch profilin IgE interaction sites (cyan). b. Sequence logo of amino acid consensus between reproductive profilins, where height of letter correlates with amino acid conservation between species. Black shows 100% identity between profilin amino acids. c) Model of profilin gene (Box and lines) evolution into reproductive (red) and vegetative (green) profilins. Grey arrows represent divergence.

2.5 <u>Discussion</u>

2.5.1 The evolution of *OsEXPB* involved both ancient and recent duplications

The evolutionary history of the *OsEXPB* genes reveals that this family has undergone both ancient and modern duplications. Prior to divergence of monocots and dicots approximately 165 million years ago, an expansin gene arose that was likely to have been vegetatively expressed (Cheng et al.,

2012) (Figure 2.1), before diverging into αExpansins and β –Expansins (Sampedro and Cosgrove, 2005). Following this, monocots experienced a gene duplication, potentially the sigma whole genome duplication, which leads to an increased copy number of β –Expansin genes (Sampedro et al., 2015). Following grass divergence, a chromosomal duplication occurred and resulted in two chromosomes, each with a copy of the CEB expansin homolog. One of these genes evolved into a DEB gene, resulting in CEB and DEBs genes in grasses, but only CEB genes in non-grass monocots. DEBs differed from other expansions due to a deletion of a pair of cysteine residues near a binding groove, which may have remodelled the edges of the binding surface but not the central binding groove, leading to a grass specific function (Sampedro et al., 2015). During evolution, these DEB homologs diverged further into the DEBP and DEBV clades, after which DEBP genes experienced deletions of introns before speciation. Following speciation, small duplications in individual genomes resulted in high gene copy number, often in tandem, and ranged in number from 1 gene in Bermuda grass to 12 copies in Maize.

2.5.2 Copy number of *OsEXPB* is not a result of domestication

Most grass species contain multiple copies of DEBP genes, suggesting that high copy number may be beneficial for pollen development or function. The polymorphism in *OsEXPB13* lead to a truncated encoding protein, suggesting that this gene has either lost its function, or that the pollen allergen 1 domain is not essential to viability in the presence of redundant genes. However, *OsEXPB1a* or *OsEXPB1b* arose very recently in evolution, and may have arisen after the mutation occurred in *OsEXPB13*. If the newly duplicated *OsEXPB1* gene showed redundancy with the other *OsEXPB* genes, its function may have been retained in order to compensate for the loss or down regulation of *OsEXPB13*. Alternatively, an overabundance of genes may be detrimental to viability, or not experience selective pressure and eventually accumulate harmless mutations. Maize contains 9 β-expansin homologs, some of which also appeared to be pseudogenes, suggesting that these genes do not experience selective pressure to increase function gene copy above 3-4 copies. Potentially, *DEBP*

grass genes experienced increased copy numbers through rapid evolution due to environmental stimuli, but then experienced selective pressure that favoured a medium gene number.

Of the five rice species, two had been domesticated and three were wild varieties. All rice species had 3 or 4 DEBP *EXPB* gene copies, although it is unclear whether these genes are all functional. Due to the consistent copy number of genes between the rice genus, it is likely that the gene copy arose before human intervention. This suggests that *DEBP* gene copy number was not a result of domestication in rice, though it may have been unintentionally selected for in other domesticated species, such as maize. As more rice and grass species genomes become available, new information may shed light on whether allergenicity was influenced by domestication.

Brachypodium is not a domesticated species, but multiple *BdEXPB* genes occurred in a random pattern across four chromosomes (Sharma et al., 2017). In rice, the four *OsEXBP* genes appeared in tandem, further suggesting that gene copy number of these important pollen specific genes are duplicated later in rice development. The *BdEXPB* genes all appeared to be more highly expressed in the anther, but one copy had much lower expression than its homologs (Sharma et al., 2017). Both species contained at least three functional homologs, and four copies altogether, suggesting three as an ideal copy number regardless of domestication.

2.5.3 Loss of introns may promote translational efficiency

*OsEXPB*s and other DEBP homologs lacked or had reduced numbers of introns. This may promote high translational efficiency as no splicing would be necessary between transcription and translation, leading to lower translational regulation. Potentially, high translational efficiency of these genes is necessary because the life of pollen and, in particular, pollen tubes is very short. High efficiency of translation would allow a quick burst of *OsEXPB* translation, causing the pollen tube to germinate within minutes of pollination. In rice, pollen grain germination occurred 2 within minutes after the pollen was placed on the pistil, and the pollen tube was able to penetrate the style in less than 10 minutes, so high efficiency of translation of pollen tube growth genes, and therefore pollen tubes,

may also increase delivery of the male gametes to the embryo sac. Therefore, DEBP genes that had lost introns may have been selected as they increased fitness.

When assessing fitness, pollen that contain genes that promote high pollen tube growth efficiency may also have an advantage when fertilising an ovule. A pollen grain with a faster growing pollen tube will be more likely to reach the embryo sac than another grain that penetrated the stigma at the same time.

The mechanisms of intron loss are still unclear. One suggested model is the reverse transcriptase model, which suggests cDNA without introns recombines with genomic
DNA. This mechanism may also be correlated with the accumulation of Pseudogenes (Zhu and Niu, 2013). Some grass species, such as maize, appear to have a high copy number of DEBP pseudogenes, which may suggest that the reverse transcriptase model applies to intron loss in DEBPs. If this is the case, these genes may not be experiencing a high level of purifying selection and gene function loss, but instead accumulated pseudogenes from reverse transcriptase gene insertion. However, as all DEBP genes lacked introns, it is likely that the original intron loss occurred in a grass ancestor. Additionally, many of the DEBP genes are found in tandem along the same chromosome, so copy number after speciation may have occurred due to small gene duplications.

The pollen specific expression of beta expansins and the recent duplications in evolution suggests that the introns are lost from these DEB sequences, rather than gained in other expansins, before grasses diverged from other species.

2.5.4 *OsEXPB1/10* are expressed specifically in reproductive tissue

Grass species flower for a short period at a distinctive time each day, and all cereals are anemophilous (Friedman and Barrett, 2009). Expression data of DEBPs in grasses suggest that these genes are expressed exclusively in the mature anther and pollen (Figure 1, also see chapter 4). Some species showed high expression in the anther, potentially due to the mature anther containing pollen.

Many genes did not have pollen expression data available. However, Brachypodium DEBVs were expressed at low levels in most tissue, and one was expressed highly in the root.

The *OsEXPB1/10* promotor contained a pollen-specificity element (TGTGG) near the start codon, which may be responsible for pollen-specific expression (Azria and Bhalla, 2011). GUS under the Ory s 1 native promoter was expressed in anther and mature pollen, but no GUS signal was found in immature pollen or other reproductive tissues (Azria and Bhalla, 2011). This would suggest that the promotor experienced changes during evolution that pushed these genes to become pollen specific as opposed to polymorphisms in the coding region of the gene.

OsEXPB genes require little post transcriptional regulation, so the highly specified and quick expression and low maintenance may suggest that less resources need to be spent on translating *OsEXPB,* and instead pollen developmental mechanisms can focus on efficient pollen tube development.

2.5.5 Reproductive and vegetative profilins arose prior to monocot/dicot divergence

An ancient profilin gene arose before moss diverged from other plants and likely had a vegetative function. Before monocots and dicots diverged, a chromosomal duplication lead to two profilin copies in monocots. One profilin copy retained vegetative expression, but the other gene diverged to gain a reproductive tissue specific expression. Following species divergence, rice experienced a tandem duplication of the reproductive profilin.

Nucleotide sequence identity between *OsPRFAa* and *PHpPRF2-like* share 62% identity, suggesting that little sequence divergence has occurred over the profilin family in plants and fungi, but became highly diversified in vertebrates (Huang et al., 1996). The purpose of the IgE antibody class originally evolved to protect the body against parasites. After plant profilins diverged into DEBPs and DEBVs, profilins in pollen became airborne, more abundant and made more contact with body, profilin may have become capable of inducing allergic sensitisation. This may have been exacerbated

in some domesticated species. For example, in cotton, expression and transcript abundance of profilin genes increased due to domestication because of their crucial role in pollen development, leading to more potent pollen or fruit tissue (Pandey and Chaudhary, 2017). However, the allergenicity of single profilin protein encoded by particular genes may have been unaffected by domestication as they were probably a consequence of shared ancestry as opposed to newly evolved structures.

Profilin genes diverged into reproductive and vegetative prior to monocot speciation. In rice, reproductive profilins *OsPRFAa* and *OsPRFAb* may be involved in actin-binding, but their specific function is unknown (Devis et al., 2017). In Arabidopsis root and tobacco pollen, profilins localised to the cytoplasm, suggesting the role of these profilins has been conserved despite the difference in tissue specificity. However, the male gametophyte may have acted as a switch that initiated the expression of developmental profilins in male reproductive tissue as opposed to vegetative tissue (Kandasamy et al., 2002). In Arabidopsis, reduction in V-profilin transcripts lead to decreased lateral root growth and cell elongation, but it is unclear whether rice pollen expressed profilins will also contribute to cell elongation (Müssar et al., 2015)

It is unclear if there is a causative link between actin interaction sites and IgE binding sites, but point mutations at actin interaction sites may lead to understanding of whether allergenicity is influenced by these key amino acids.

2.5.6 Putative IgE interaction sites appear on the surface of Profilins

Putative IgE interaction sites in sun flower and birch profilin were located in close proximity at the tertiary level of the folded protein, suggesting this may be a single, discontinuous IgE epitope made by two separate chains that have been brought together during folding. A hypoallergenic derivative may be achieved through disrupting the binding of these two chains, but this is likely to lead to disrupted folding and therefore function. Alternatively, one chain could potentially be substituted with similar amino acids, causing the protein to fold properly but disrupt IgE binding. This would depend on the specificity of IgE binding to certain amino acids or amino acid sites, and provides an

exciting area of future study into the development of hypoallergenic proteins based on small amino acid chain substitution. This site was also found on the surface of the protein, potentially allowing IgE to recognise and bind to the protein with ease. Interestingly, the interaction sites were located on α-helices as opposed to β-sheets, potentially because β-sheets were found in the core of the protein and inaccessible to IgE. Alternatively, α-helices have side chains that may be more accessible to binding, as opposed to the flat nature of β-sheets. However, profilin appears to accommodate binding of three IgE molecules, despite its small size (Valenta et al., 1998)

Although six actin interaction sites were located in the region associated with an IgE binding site in birch profilin, it is unclear whether there is a similarity between actin interaction sites and IgE reactivity. Potentially, the surface the IgE interaction sites appear on are easily accessible for binding in general, and so is also the most efficient surface for actin binding as well as IgE binding.

Actin interaction sites have conserved amino acid sequences. Mutations that resulted in new amino acids within these sites generally were substituted by an amino acid of the same type, suggesting that function may have been maintained. TaPRF1a and AtPRF4 had an amino acid substitution of Met117 to leucine, but methionine and leucine are both hydrophobic molecules and actin interaction function may have been maintained. However, due to the close proximity of putative actin interaction sites on the surface of the protein, it is likely that interaction with actin is dependent on a group of amino acids, many of which are hydrophobic, as opposed to single amino acids.

Interestingly, putative IgE binding sites were not buried in reference to the Profilin-Actin complex, and IgE may be able to bind to the whole complex as well.

2.5.7 Homologs of Ory s 1 and Ory s 12 showed varied evolutionary patterns

Group 1 pollen allergens and profilins are both known to elicit allergenic responses, although Ory s 12 has not yet been properly characterised. *OsEXPB* and *OsPRF* appear to have evolved differently and have different structures, binding sites and conservation, so it is unclear why these unrelated

proteins both have allergenic potential. It is possible that IgE is able to recognise many different structures as antigens, and the allergenic relationship between Ory s 1 and Ory s 12 is not due to structure or evolution. Importantly, both *OsEXPB* and *OsPRF* are highly and specifically expressed in the anther and pollen, especially during late stage pollen development. This correlation suggests that allergenicity in these two distinct families may be due to other factors, such as expression and localisation of the proteins or even potentially functional similarities. Expansins are often involved in cell-wall expansion and may localise to the cell wall, making the protein easily accessible to IgE when the pollen ruptures upon inhalation (Sampedro and Cosgrove, 2005, Sampedro et al., 2015, Valdivia et al., 2009). However, profilins are often intracellular and may be buried inside the pollen instead of on the surface, although IgE may have access to profilin proteins when they are released from damaged and dead pollen (Swoboda et al., 2004). Besides the expression of these genes in pollen, there appears to be no evolutionary or structural similarities that may contribute to allergenicity.

Not all pollen expressed genes are putative allergens, but the differences in evolution and structure between Ory s 1 and Ory s 12 suggests there may be other factors contributing to allergenicity. Future studies on the function of these genes may lead to a better understanding of how some proteins become allergenic.

Statement of Authorship

Title of Paper	Chapter 4 – Functional assessment of two rice pollen genes, OsEXPB and OsPRFA		
Publication Status	☐ PublishedAccepted for ☐ Publication ☐ Unpublished and ☑ Unsubmitted work written in Submitted for Publicationmanuscript style		
Publication Details			

Principal Author

Name of Principal Author (Candidate)	Deborah Devis		
Contribution to the Paper	Conducted research, wrote and edited paper		
Overall percentage (%)	70		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25/07/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Xiujuan Yang		
Contribution to the Paper	Provided technical advice in CRISPR vector construction		
Signature		Date	05/07/2019

Name of Co-Author	Gang Li		
Contribution to the Paper	Provided Editing feedback and technical advice		
Signature		Date	15/07/2019
Name of Co-Author	Song Yu		
Contribution to the Paper	Aided in transformation and planting in China		
Signature		Date	20/09/2019
Name of Co-Author	Dabing Zhang		
Contribution to the Paper	Helped with writing and framing of paper and editing		
Signature		Date	09/08/2019

Chapter 3 – Creation of mutants of rice pollen allergen genes using the genome editing system, CRISPR/Cas9

3.1 Abstract

Functional analysis of pollen genes is essential to understanding the potential for a viable, hypoallergenic plant. In this chapter, I describe the generation of two mutant rice pollen allergen lines of Ory s 1 and Ory s 12. I designed specific CRISPR/Cas9 cassettes and transformed them into Nipponbare rice seeds using agrobacterium-mediated transformation. I was unable to obtain a mutant of vegetatively expressed Ory s 12, as regenerated plants did not develop roots. However, I obtained at least 3 viable mutant Ory s 1 lines and 3 viable mutant Ory s 12 lines that produced seeds ad were used for further functional analysis. CRISPR/Cas9 mutation of rice Ory s 2 and Ory s 12 was successful.

3.2 Introduction

Functional analysis of genes involved in pollen and seed development is essential to understanding the biological processes that lead to viability. Decreased function of important pollen genes may lead to compromised viability, so functional analysis of rice pollen allergens is necessary to assess the potential for a hypoallergenic plant. A common approach to analyse allergen gene function is to assess the resulting phenotype of mutants, where changed phenotype may be due to loss of gene function. Previously, hypoallergenic derivatives of rice allergens have employed RNA interference techniques to silence allergen genes in grasses (Bhalla and Singh, 2004, Petrovska et al., 2004, Tada et al., 1996, Valdivia et al., 2009, Wakasa et al., 2011). However, this process leads to knockdown of gene function, not complete knock out (Boettcher and McManus, 2015,
Rosa et al., 2018).

Alternative methods of site specific genome editing, such as the CRISPR/Cas9 system, provide a new path for efficient, highly specified and heritable genome editing that may lead to the generation of mutant plants with novel, hypoallergenic traits (Li et al., 2017, Ma et al., 2015, Ma et al., 2016,

Sedeek et al., 2019). Mutants derived with CRISPR/Cas9 may lead to loss of specific Ory s 1 or Ory s 12 gene function and shed light on the importance of these genes in pollen and the potential for a viable hypoallergenic plant.

CRISPR/Cas9 gene editing requires the transformation of the CRISPR cassette into an embryo. Agrobacterium-mediated transformation is often utilised to introduce genetic material to the single-cell seed, whereby vector transfer is mediated by the gram-negative soil bacterium *Agrobacterium tumefaciens* (Agrobacterium) (Nishimura et al., 2006). Rice seeds are placed on a media and generate calli, which are transfected with Agrobacterium containing the CRISPR cassette (Figure 3.1a-c). Calli are then selected

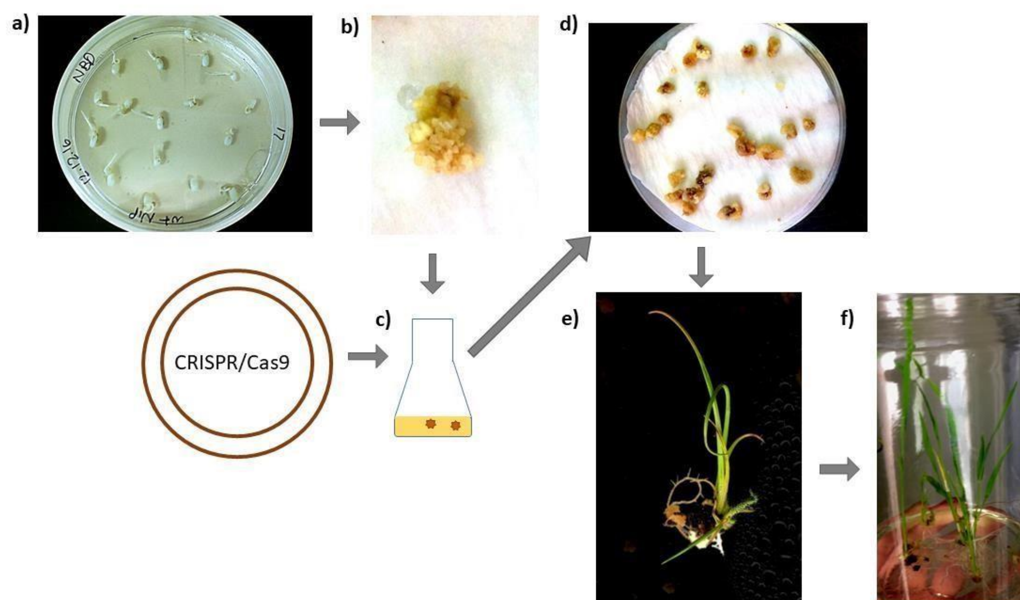


Figure 3.1 – Overview of vector transformation into rice

Process of vector transformation into rice **a)** Induction of calli from rice seeds, **b)** isolation and growth of calli, **c)** transfection of calli with CRISPR/Cas9 vector in Agrobacterium, **d)** selection of transformed calli on hygromycin and timentin, **e)** regeneration of shoots and **f)** regeneration of roots. Images taken from this experiment

with antibiotics, and leaves and roots are regenerated (Figure 3.1d-f). F2 generations can then be assessed for phenotype.

In this chapter, I generated rice mutants that contained mutated Ory s 1 and Ory s 12 by employing the CRISPR/Cas9 system to understand the function of Ory s 1 and Ory s 12.

3.3 Materials and Methods

3.3.1 CRISPR target sequence design

To uncover the function of allergen genes, I identified gene regions that could be used for target specific knock-down with the CRISPR/Cas9 system. Four homologous β-expansin genes; *OsEXPB1a, OsEXPB1b, OsEXPB10* and *OsEXPB13*, show high nucleotide identity (>70% identity) and all contained a DPBB1 domain and a PA1 domain (excluding *OsEXPB13*). I designed the CRISPR/Cas9 guide sequence to target upstream of the DPBB1 domain, so protein generated from the edited sequence would lack both functional domains. I chose one 20-bp target site to knock-down three genes: *OsEXPB1a,*

OsEXPB1b and *OsEXPB10* (5'GCCCCGACCGGACTGTTTCT3'). This sequence was

directly upstream of a protospacer adjacent motif (PAM) site (CGG). *OsEXPB13* has an early stop codon upstream of the PA1 domain, suggesting that the gene is non-functional. For this reason, I did not attempt to knock it out.

Three profilin homologs were found in rice: *OsPRFAa, OsPRFAb,* which are reproductively expressed, and *OsPRFLP04,* which is vegetatively expressed. *OsPRFAa* and *OsPRFAb* share 100% amino acid identity, 98% nucleotide identity, and approximately 79% nucleotide identity with *OsPRFLP04* each. *OsPRFAa* and *OsPRFAb* also shared high nucleotide identity in the promoter and coding sequence so no target sequence could be designed to target only one gene. Due to the conserved nature of profilins in plants and pollen, it was likely that complete knockout of profilin genes would lead to a sterile or unviable plant, so I designed two gRNAs 284bp downstream of the start codon, in order to retain some profilin function. To disrupt *OsPRFAa* and *OsPRFAb,*

I chose a 20bp target sequence between G284 and T303 (GCCCCGACCGGACTGTTTCT). To disrupt *OsPRFLP04*, I chose a 20bp target sequence between G284 and T303 (GCACCAACTGGTCTTTTCCT). Both sights were directly upstream of a PAM site (TGG).

3.3. 2 CRISPR-Cas9 plasmid construction

In order to introduce the CRISPR/Cas9 cassette into rice, an entry vector was constructed using the pBIN-sgR-U3::Cas9 vector . To anneal the primers, 10 nM of each primer was mixed with 8 µL annealing buffer (TE+50 mM NaCl)(NEB Biolabs) and slow cooled in a PCR machine (initial temperature 95^0C to final temperature 16^0C at with decrease of 0.1°C/s). I chose to use the CRISPR/Cas9 multi-targeting vector that has been proven to efficiently generate deletions in rice (Liang et al., 2016, Ma et al., 2015). pBIN-U3::Cas9 (Figure 3.2a) vector was digested with 25U BSAI (NEB) at 37°C for 2 hours. 10 ng annealed primer was used for ligation with 1 ng cut U3::Cas9 vector in 10X T4 ligase buffer (NEB) and 35U T4 ligase (NEB) in a volume of 10 µl. The ligation solution was incubated at 25^0C for 2 hour. Ligated plasmid was then transformed into *E. coli* DH5α competent

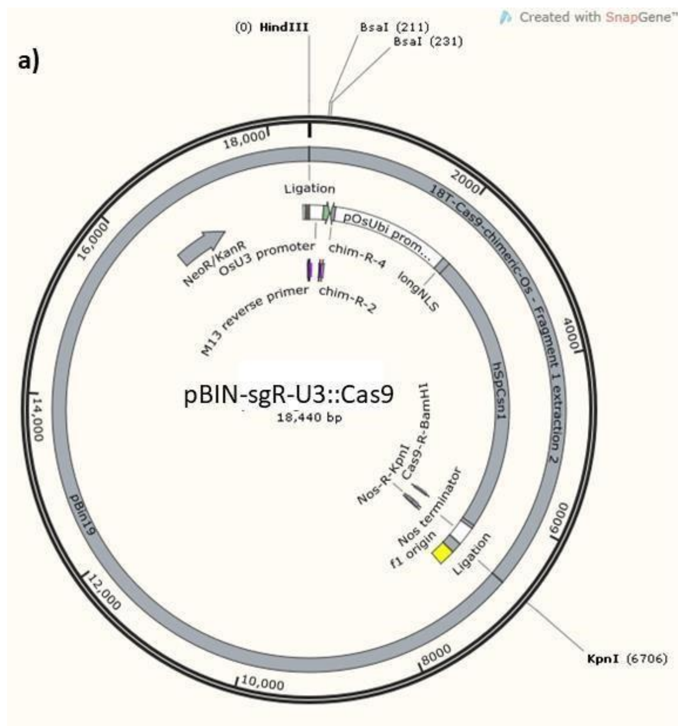


Figure 3.2- pBIN-sgR-U3::Cas9 vector for construction of the CRISPR/Cas9 cassette,

Including 18T Cas9-Chimeric for *Oryza sativa*, BSAI cut sites for insertion of sgRNA, pOsUbi3 promoter and kanamycin resistance

cells and grown over night on Lysogeny Broth (LB) with kanamycin (50µg/ml) at 37^0C. Colonies were screened through sequencing (AGRF).

3.3.3 Induction of calli

Wildtype Nipponbare seeds were used for transformation using the method described by Nushimura (Nishimura et al., 2006). All media used during transformation were shown in Supplementary Table 3. In a sterilised laminar flow, rice seeds were sterilized with 70% ethanol, then 30% sodium hypochlorite for 30 minutes, and finally washed five times with sterilised distilled water. Sterilised seeds were plated on NBD media (Supplementary Table 3) at a density of approximately 17 seeds/plate and incubated at 28^0C in the dark for 2 weeks. At 2 weeks, calli were cut from the seeds using a sterilised surgical blade and placed on fresh NBD. The calli were incubated at at 28^0C in the dark for 10 days.

3.3.4 Cocultivation of calli with CRISPR/Cas9 infected Agrobacterium

During this time, I prepared the agrobacterium for cocultivation. Vector positive Agrobacterium was grown in kanamycin+ liquid LB at 28°C for two nights. 2ml of this culture was added to 50ml kan+ YEB (Supplementary Table 3) and shaken at 28°C for 2 hours. Agrobacterium was then centrifuged and the supernatant was removed. The cells were resuspended in 50ml AAM.

At 3 weeks and 3 days, calli were placed in 50 ml AAM/Agrobacterium solution and shaken for 30 minutes (Nishimura et al., 2006). Calli were dried on sterile paper for 15 minutes and plated onto NBD plates that had two layers of filter paper and 0.5mM antibiotic.

Cocultivation was carried out by incubating the plates in the dark at 28°C for 3-4 days.

3.3.5 Selection and regeneration

At 4 weeks, cocultivated calli were dried on filter paper in a sterile laminar flow for 1 hour. Dried calli were placed onto fresh NBD media with 100 µg/ml hygromycin and 100 µg/ml timentin.

Selection plates were incubated in the dark at 28°C for 10 days and selection was repeated once more. Calli that had high levels of Agrobacterium after this time were discarded.

After 10 days, calli were then transferred to regeneration media (Supplementary Table 3) with 100 µg/ml of hygromycin and timentin. Plates were grown in light at 28°C for 24 weeks until green leaves developed. Regenerating calli were moved to fresh plates every 2 weeks.

Calli with regenerated shoots were transferred to sterile jars that contained root growth media (Supplementary Table 3) with 100 µg/ml hygromycin. Regenerating seedlings were transferred to new jars every two weeks. Seedlings were moved to pots between 46 weeks.

3.3.6 Growth conditions

Regenerated rice plants were grown in a growth chamber at the Waite Campus (University of Adelaide) at 28°C 12 hour day and 24°C 12 hour night, or in the paddy field in Shanghai as described previously (Li et al., 2018, Zhang et al., 2011). At least 10 independent F2 mutant plants were used to study phenotype across three lines. During flowering, the temperature in growth chambers was increased to 32°C and maximum humidity for two hours.

3.3.7 Identification and genotyping of transgenic lines

Genotyping of seedlings was carried out using 5 mm leaf tissue as a template with the Phyre direct PCR kit (Thermo Scientific) and using CRISPR confirm primers to amplify a region of 200 bp (Supplementary Table 4). Genes were sequenced through Sanger Sequencing (AGRF) and CRISPR deletions or insertions were analysed using TIDE (Brinkman et al., 2014).

3.4 Results and discussion

3.4.1 Creation of DEBV *OsEXPB* mutants generated with CRISPR/Cas9

One CRISPR target sequence was introduced into rice to generate mutations in *OsEXPB1* rice genes. This sequence (GTCGACAAGGCTCCCTTCCTCGGCT) was able to target *OsEXPB1a*

and *b* and *OsEXPB10* between G212 andT 232 (Figure 3.1a). As OsEXPB1a and b shared 100% nucleotide identity along the whole promotor, gene and 3' untranslated region, they were treated as a single gene and will be referred to as *OsEXPB1*.

The following 3 mutant lines were produced: *expb-1* was a double homozygous mutant, *expb-2* had a homozygous mutation in *OsEXPB1*, and *expb-3* had a double heterozygous mutation (Chapter 3, Figure 3.1c).

Mutations within the target sequence caused polymorphisms immediately upstream of the DPBB1 domain. The resulting mutations lead to frameshifts and early stop codons, leading to a truncated protein that lack both the DPBB1 and the PA1 domains.

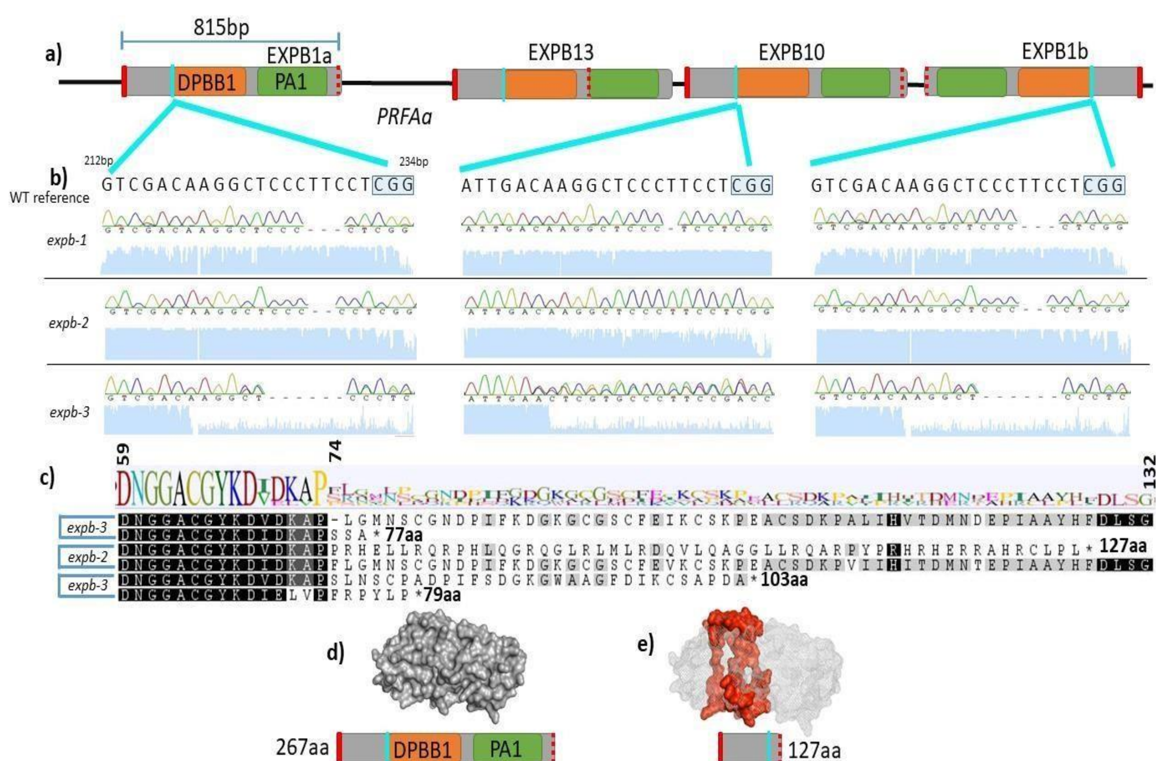


Figure 3.3- Location of CRISPR/Cas9 mutations in *expb* mutants

CRISPR mutations lead to truncated proteins. a) Target sequence (Cyan) in relation the functional domain DPBB1 (orange) and PA1 (green) domains, with start codon (red), stop codon (red dots) and genome (black line). b) CRISPR target sequence and resulting mutation in *Ory s 1b-21* in relation to gene structure (blue lines). c) Amino acid sequence alignment of lines 1, 2 and 3 showing Identical sequence (black), gaps (dash) and stop codon (star).d) WT protein and CDS. e) Truncated mutant protein (red) relative to the WT protein (grey net) and corresponding CDS.

At least 20 regenerated seedlings were obtained from the transformation of *expb* and *prf* CRISPR/Cas9 constructs into rice. Of these seedlings, 8 lines had edits in *expb* genes, and 3 lines contained edits in *prfa* genes. All regenerated seedlings grew poorly regardless of genotype; transformed plants grew between 1-4 stunted tillers, unlike

untransformed wild type plants which grew 5-7 tillers, and many plants died before maturity. Only 3 mutant *expb* mutants survived to maturity, but each plant showed very low seed setting and only 8 seeds were obtained between all three plants. Of the 8 seeds collected, 6 grew to maturity but only 2 contained an edited gene. This line was designated as *expb1-1* and contained a 2bp edit, leading to a frame shift and early stop codon (Figure 3.3a-b).

All three *prfa* edited transformed plants reached maturity, but none of these 3 set seeds. It is possible that low seed setting in mutant lines was due to a combination of stress from tissue culture regeneration and poor environmental conditions (Mohammed et al., 2019). This will be further elaborated upon in chapter 4.

Due to the low seed yield, transformation was repeated in Shanghai. After the second transformation, 2 independent mutant *expb1* lines were obtained for further functional analysis (Figure 3.3). Seed setting of these transformed plants was lower than wild type, but was greater than 30%. This may have been due to better growth conditions. Plants from the first round of transformation were grown in a growth chamber, but the second round of plants were grown in a field in Shanghai over summer. After seeing the seed setting success of the second round of transformations, I added a humidifier to the growth chamber in Adelaide and ran it at maximum during daylight hours.

3.4.2 Generation of *OsPRFA* mutants through CRISPR/Cas9 deletions

As with *expb* mutants, two rounds of transformation were conducted to obtain *prf* mutants. The first round of transformations generated 17 plants, two of which appeared to have a heterozygous single nucleotide deletion in *PRFAa* G298, G299 and C302 respectively.

All plants grew to maturity but only C302 set seeds. This plant was designated as *prfa-1* (Figure3. 4). The second round of transformations yielded 23 plants, 4 of which showed deletions. Of these mutant lines, *prfa-2* had a mutation in *PRFAb* at position C302, *prfa3* had a deletion in *PRFAa* at position C303, *prfa-4* had a deletion in *PRFAa* T301 and *prfa-5* had a deletion in *PRFAb* at position C303 (Figure 3.4b). Mutants *prfa-2,3,4* and *5* grew to maturity and set seeds, but no mutant offspring were obtained from *prfa-4* or *prfa5*. Additionally, no homozygous mutants were obtained from any *prfa* line, suggesting mutant *prfa* pollen grains were unable to penetrate the ovule and were therefore outcompeted by wildtype pollen. This phenotype will be discussed further in Chapter 4.

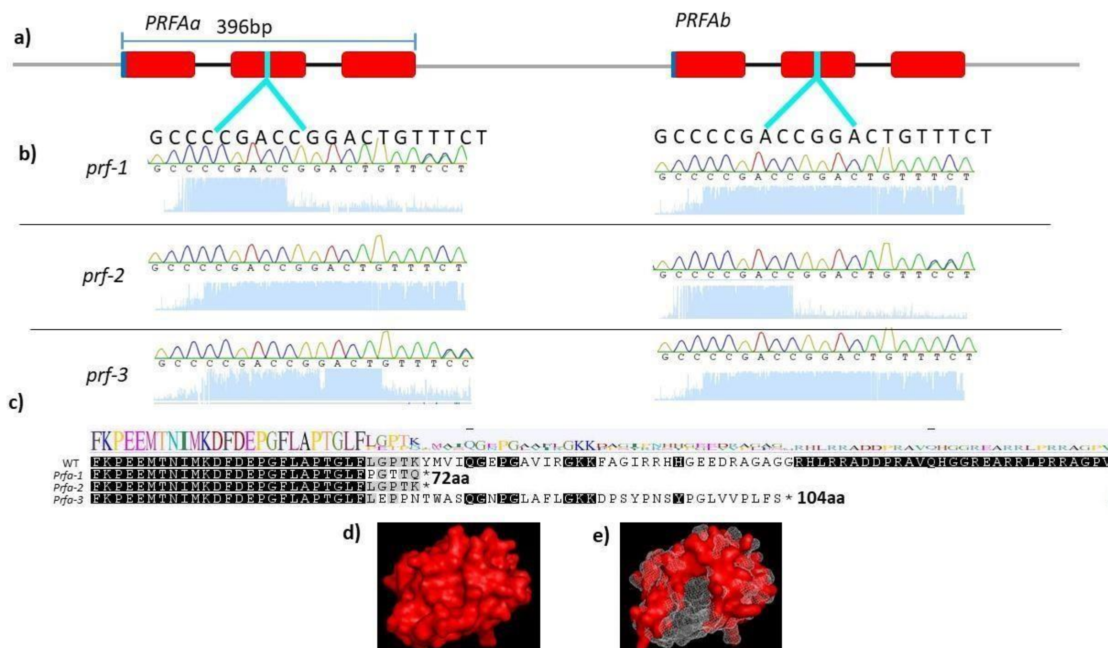


Figure 3.4- Location of CRISPR/Cas9 mutations in *prf* mutants

CRISPR mutations lead to truncated proteins. a) Target sequence (cyan) in relation the CDS (red box), Start codon (blue line) and genome/intron (black line). b) CRISPR target sequence and resulting mutation in *prfa-1, 2* and *3* and sequence quality of full genomic sequence (light blue, high shows high quality

homozygous sequence, low shows multiple heterozygous sequence) c) Amino acid sequence alignment of *prfa-1, 2* and *3* showing amino acid logo (tall letters show conservation), Identical sequence (black), and stop codon (star)d) WT protein. e) Truncated mutant protein (red) relative to the WT protein (grey net).

3.4.3 Pollen viability of T0 mutants

In order to assess pollen fertility and potential pollen phenotypes of the T0 generation, I used an iodine staining test to count fertile pollen grains. T0 plants from one mature mutant line grown in Adelaide showed low pollen fertility compared to wild type (89% fertile); *expb-2* had 53.4% viable pollen (8 tillers counted, $p=1.8x10^{-12}$) where 200 pollens were counted from 5 tillers. However, pollen from the T0 anthers of the second transformation were indistinguishable from wild type. *expb-2* was a single homozygous plant, but T0 *expb-1* was double homozygous, so the low pollen viability in *expb-2* is unlikely to have been caused solely by mutant EXPB. As this mutant line was the only plant to reach maturity, it was possible this defect was specific to this line or due to a personal error in how I cared for this line during growth.

Similarly, T0 *prfa-1* had 53.4% viable pollen, which was significantly less than the wild type (6 tillers counted, $p=4.03x10^{-25}$). The low number of tillers retrieved from *expb1* and *prfa-1* do not provide robust statistical analysis and may not reflect a true phenotype caused by the mutation. Likewise, the difference in pollen fertility between transformation rounds 1 and 2, and T0 and F2 suggest that the low pollen fertility is unlikely to be due to the mutation alone, but due to regeneration or growth stress. To test this assumption, 2 T0 plants that did not contain any edits were also assessed for pollen fertility. Both unedited regenerated plants showed low pollen fertility (57.1% and 53.2%) compared to untransformed wild type. No pollen fertility phenotypes were witnessed after this original round 1 T0 generation. Furthermore, both *expb-2* and *prf-1* were generated in the same round of transformations and were grown in the growth chamber in Adelaide, during which time the growth chamber experienced inconsistent temperature and humidity regulation. Low pollen viability was not observed after the growth chamber was fixed, and low pollen fertility was not observed in untransformed plants under inconsistent conditions,

suggesting the phenotype was not due to the mutation but due to transformation stress and/or growth conditions.

3.4.4 No roots were regenerated after *OsPRFLP04* transformation

I used 500 rice seeds during transformation for knock-down of *OsPRFLP04.* Leaves were regenerated from 25 of 100 selected calli, but roots were only regenerated on two seedlings, both of which contained no edits. This did not occur with the other *expb* or *prfa* mutants, which regenerated roots on most calli that were transferred to root media.

OsPRFLP04 is a vegetatively expressed profilin and is expressed in most tissue, but its particularly highly expressed in the root tip (Hruz et al., 2008). Similarly, tobacco profilins control tipping in both the pollen tube and root hair by regulating the organisation of actin (Swoboda et al., 2001). *OsPRFLP04* may also be essential for root tip development and loss of function may lead to aberrant or aborted root development, from which I was unable to attain any transformed seedlings. Interestingly, leaves did form from some calli. I was unable to genotype this growth to determine whether any edits had been made. Future studies may address the expression and localisation of *OsPRFLP04* in reference to the root tip using overexpression of OsPRFLP04-GFP, which may show the specificity of this protein during root development. Furthermore, a knockdown mutant, whereby sufficient protein is expressed for growth, may shed light on the function of this gene. As vegetative profilins were not the focus of my study, I did not continue with new transformations. I also assessed the occurrence of CRISPR/Cas9 efficiency in the T1 generation using TIDE (Brinkman et al., 2014). Some new edits were seen to have occurred in leaf tissue in the T1 generation, suggesting the CRISPR/Cas9 continued to make edits after the first generation. However, these edits were in a low occurrence (>16%) and may have occurred in non-reproductive material, as these mutations were not seen in the F2 generation and were therefore unlikely to elicit a phenotype in pollen.

3.5 Conclusion

In order to obtain mutants for functional analysis I constructed Ory s 1 and Ory s 12 CRISPR/Cas9 cassettes and transformed them into Nipponbare rice. Clearly, the process of obtaining mutants and gaining insights into pollen function is difficult and poses many obstacles, such as imperfect growth conditions, improper protein folding, and high sequence identity problems. However, I was successfully able to obtain 3 independent mutant lines of *expb* and *prfa* which were used for in-depth functional analysis as outlined in Chapter 4.

Statement of Authorship

Title of Paper	Chapter 4 – Functional assessment of two rice pollen genes, OsEXPB and OsPRFA
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Deborah Devis		
Contribution to the Paper	Conducted research, wrote and edited paper		
Overall percentage (%)	70		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	25/07/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Xiujuan Yang		
Contribution to the Paper	Provided technical advice in CRISPR vector construction and in situ hybridization		
Signature		Date	05/07/2019

Name of Co-Author	Gang Li

Contribution to the Paper	Provided Editing feedback and technical advice		
Signature		Date	06/07/2019

Name of Co-Author	Janet Davies		
Contribution to the Paper	Provided Laboratory space and antibodies		
Signature		Date	20/07/2019

Name of Co-Author	Dabing Zhang		
Contribution to the Paper	Helped with writing and framing of paper and editing		
Signature		Date	009/08/2019

Chapter 4 – Functional assessment of two rice pollen genes, *OsEXPB* and *OsPRFA*

4. 1 Abstract

Grains are a globally important food source produced by grasses such as rice, barley and wheat. Grass pollen is known to commonly cause hay fever due to an abundance of allergens in the pollen grain. However, little is known about the function of pollen allergens in pollen development or whether a hypoallergenic mutant plant would be viable. In this chapter, I found that two rice pollen allergens, Ory s 1 and Ory s 12, are involved in pollen tube development by promoting pollen germination and growth. Ory s 1 mutants had delayed pollen tube germination and aborted early in development, but over expression of Ory s 1 lead to rapid pollen tube growth compared to wild-type pollen. Regardless, Ory s 1 mutants were still able to set seeds. Ory s 12 also showed delayed and aberrant pollen tube germination, but over expression of Ory s 12 lead to short pollen tube growth also, potentially due to over polymerisation of actin. Finally, a homozygous Ory s 1 mutant, *osexpb-1,* showed decreased binding to Queensland grass pollen allergy patients, suggesting this new mutant line may be a partially viable, hypoallergenic derivative.

4.2 Introduction

Grains are a globally important staple food. Grasses such as rice, wheat and barley contribute to the global food source by providing abundant grains, or seeds, for consumption and grass reproduction. During reproduction, thousands of pollen grains are produced in the anther through meiosis and mitosis, and each pollen contains two sperm cells, which are delivered to the ovule through the pollen tube (Altschul et al., 1997, Shi et al., 2015, Zhang et al., 2010, Zhang and Wilson, 2009, Zhang et al., 2014). The pollen tube germinates from the pollen grain when a pollen

comes in contact with the pistil, which is the female reproductive organ of the flower that contains the ovule. Ovules fertilised by the sperm grow into hard, dry seeds that are eaten as grains. For this reason, pollen is a crucial aspect of plant reproduction, as grain production is directly related to pollination. On the another hand, many grass species release pollen grains that can also cause an allergic response in sensitised individuals, but due to the wind-pollinate nature of grasses, the removal of pollen grains from the air is difficult to achieve.

Allergies are recognised as an immune response triggered by a specific protein called an allergen. Many allergens are present in plants and come in contact with humans through fruit consumption or pollen inhalation. Unlike allergens found only in fruit, pollen allergens are difficult to avoid, and lead to hay fever symptoms such as itching of the eyes, sleep loss and sneezing (Beggs et al., 2015, Davies et al., 2012, Devis et al., 2017). In particular, grass pollens are known to constitute a large percentage of clinically significant allergenic pollen (Davies, 2014). While many different protein families can be allergens, it is currently unclear how some of these proteins evolved or what their biological function is. Allergens have varied evolution and function. Some allergen families but may have an abundance in pollen due to the important roles in pollen germination and tube growth (Devis et al., 2017, Sharma et al., 2017, Songnuan, 2013). Due to the prevalence of allergies caused by pollen, a better understanding of the role of these pollen allergens may shed light onto future therapies (Abou Chakra et al., 2012, Andersson and Lidholm, 2003).

One group of putative pollen allergens (Group 1 allergens) is a family homologous to β-expansins (Cosgrove et al., 1997, Sampedro and Cosgrove, 2005). Group 1 allergens are the major contributor to pollen allergenicity in grass pollens and account for the majority of IgE binding in patient sera (Hirano et al., 2013). The expansin superfamily can be categorised into α-expansins and β-expansins and act as cell wall expansion proteins expressed in most plant cells. In maize, β-expansins are expressed more abundantly than α-expansins and have varying degrees of tissue specificity (Wu et al., 2001). Two βexpansin families are expressed in maize reproductive

material; *EXPB1* is expressed in pollen and *EXPB4* is expressed in the husk (Wu et al., 2001). In rice, the maize *EXPB1* protein is homologous to rice expansins EXPB1 and EXPB10, which are also isoforms of the major rice pollen allergen, Ory s 1. While sequence homology suggests *OsEXPB1* may have a similar pollen function to that of *ZmExpB1*, the specific function in rice pollen remains to be elucidated.

Profilins are functionally important proteins that are conserved in plants and animals. In plants, profilins are actin-binding proteins with varied functions in vegetative and reproductive stages, and they can also bind to other molecules such as phosphatidynositides or long Poly-L-proline sites (Ren and Xiang, 2007). Profilin homologs appear to be conserved between grasses, and most grasses contained two pollenspecific profilin homologs (Sharma et al., 2017). Despite being heavily conserved between species, the wide variety of expression of profilins in various tissues suggest that profilins have specific roles in certain tissues. Additionally, the highly conserved nature of the genes suggests these proteins play an essential role in viability. In lily and tomato, mutated pollen specific profilins lead to aberrant pollen tube growth (Hirano et al., 2013, Yu and Parthasarathy, 2014). However, few studies have been conducted on the role of profilins in pollen development in grasses, so it is unclear if grass profilins also have this conserved function.

Studies on the allergic response to grass pollen frequently report the cases on the sensitivity to pollen profilin. For example, 30% of pollen allergy patients in Spain showed specific reactivity to Pho d 2, a date palm profilin (Asero et al., 2008). Interestingly, 58% of profilin- negative pollen allergy patients also showed sensitisation to grass pollen grains. While rice causes a low incidence of allergy, it does cause allergies to field workers, especially in Asian countries. For example, 0.8% of 1200 allergy patients (age 5-60) in Delhi showed rice specific allergic rhinitis or asthma (Kumar et al., 2007), suggesting that rice may contain similar protein homologs to other grasses, or that rice contains unique allergens not found in other grasses. However, unlike many common Australian grasses, such as Bahia and Timothy grass, rice has a fully sequenced genome and is

relatively easy to mutate, making it a promising model organism for studies. Mutation of rice proteins may lead to insights into the possibility of generating hypoallergenic plants. In this study we use rice as a model plant and assess the function of two putative rice pollen allergens, Ory s 1 and Ory s 12, and this research provides insight into the function of grass allergens.

4.3 Materials and Methods

4.3.1 Growth conditions

Nipponbare rice plants were grown in a growth chamber at the Waite Campus (University of Adelaide) at 28°C 12 hour day and 24°C 12 hour night, or as previously described in the paddy field in Shanghai. At least 10 independent mutant plants were used to study phenotype across three lines. During flowering, the temperature was increased to 32°C for two hours with maximum humidity.

4.3.2 Pollen and seed fertility assessment

Mature anthers were collected and placed into 1% iodine potassium iodide solution and agitated for one minute to break open the anthers. Stained pollen was observed using a

Nikon Optical microscope (Adelaide Microscopy).
Seed setting was assessed after drying by counting 15 inflorescences for each genotype.
Hard flowers were considered set, and flat, empty flowers were considered sterile.

4.3.3 Pollen tube growth

Wild-type flowers were used to grow pollen tubes on the pistil *in vivo*. Anthers were removed from the wild-type flower and mutant anthers were brushed onto wild-type pistils. Pollinated flowers were incubated at 32°C at maximum humidity for 15mins, 30 mins and 60 mins. Flowers were then removed and fixed in 3:1 ethanol/acetic acid solution for 2 hours. Flowers were washed with distilled water three times and moved into 1M

NaOH for 12 hours. Flowers were then placed in Decolourised Aniline Blue Solution (DABS, 0.1% analine blue in 1M NaOH, incubated at 48°C overnight) for 24 hours. Pistils were removed and washed in distilled water, and placed on a slide with 30% glycerol. The stain was excited at 402nm on a Leica Nikon Confocal microscope (Adelaide Microscopy).

4.3.4 Scanning electron microscopy

Mature anthers were collected and fixed in FAA overnight. Anthers were then dehydrated in ethanol using 2x 80% wash, 2x 90% and 3x 100% wash. Anthers were dehydrated using an EMS 3100 critical point dryer. Extracted pollen samples were plated with 10nm platinum and visualised on a Philips XL20.

4.3.5 RNA extraction and cDNA synthesis

RNA was extracted from crushed tissue using TRI reagent (Sigma Aldrich). DNA was precipitated using chloroform and isopropanol, and washed with 70% ethanol. DNA was degraded using 10 1X Dnase buffer and Dnase (BioLabs), and incubated at 37°C for 1.5 hours. RNA was precipitated with acid phenol and washed with 100% ethanol. cDNA was synthesised according to manufacturer's instructions using the iScript Select cDNA Synthesis Kit (BIO-RAD).

4.3.6 Expression analysis and qRT-PCR

To assess tissue specific expression, tissue was harvested from wild-type seedling, root, leaf, inflorescence, flower, anther and pollen stages 6-12. RNA was extracted from each tissue as described above, and specific primers were designed to amplify *OsEXPB10,*

OsEXPB1, OsPRFAa and *OsPRFAb* (Supplementary Table 1 and 2, Supplementary Figure 2). cDNA was used as a template in a qPCR reaction using FAST 2 x qPCR Master

Mix kit according to manufacturer's instructions (Primer Design) and 40X cycles of denaturation at 95°C for 3s and annealing/extension at 60°C for 20s.

4.3.7 Transient expression of GFP

To clone genes into transient expression vectors, the pCAMBRIA-1301-35s-GFP was used to clone *EXPB10* or *PRFAa* CDS, minus a stop codon, downstream of the CaMV 35s promotor and in frame with the GFP tag, leading to a fusion protein. *OsEXPB10* and *PRFAa* were amplified from cDNA template using Phusion taq polymerase (NEB) in a PCR according to manufacturer's instructions and purified using the PureLink Quick Gel Extraction and PCR Purification Combo kit (Invitrogen). PCR fragments and the pCAMBRIA-1301-35s-GFP vector were digested with *Bam*HI and *Spe*I at 37°C for 2 hours, and cleaned using the PureLink kit. I cloned *OsEXPB10* CDS and *PRFAa* CDS into pCAMBRIA-1301-35s-GFP with infusion ligation (Takara Bio). Ligated plasmid was transformed into competent *E. coli* DH5α cells and grown at 37°C overnight on ampicillin+ LB media. Colony PCR was used to confirm ligation.

Plasmids were attached to cold particles for bombardment. 4μg of plasmid was added to 25µl gold carrier (60mg/mL), vortexed for 2 minutes and incubated on ice for 10 minutes. 10µl of 0.1M spermadin was added. The solution was vortexed for 1 minute, then incubated on ice for 1 minute, and repeated 10 times. 25µl $CaCl_2$ was added and vortexed for 1 minute, then incubated on ice for 3 minutes. Solution was centrifuged at 13,000rpm for 5 seconds and the supernatant was removed. 100% ethanol was added to wash. Gold particles were resuspended in 70µl of 100% ethanol.

Onion segments were placed in HO Medium (36.5g D-sorbitol, 36.5g D-mannitol, 4.4g basal MS in 1L water, pH 5.8, 1.75g/L phytogel) for four hours with the epidermal lay facing up. The vector was then introduced into the onion cells using gold particle bombardment using a PDS-1000/He particle-delivery system (Bio-rad) at 11,000p.s.i, 25mm Hg vacuum, 1 cm gap between micro carrier and rupture disk and a distance of 6 cm between the micro carrier and tissue sample. The

onion cells were then grown overnight at 28°C and assessed on a Nikon confocal microscope under green UV light.

Tobacco pollen was collected in Tobacco pollen growth media, as previously described (Li et al., 2018). The vector was then introduced into the pollen using gold particle bombardment as described above. Tobacco pollen was grown in full light at 27°C and shaken on a petri dish at 8rpm for 4 hours and 8 hours.

Four replicates were assessed and protein-GFP was observed by excitation at 488nm using a Leica Nikon Confocal microscope (Adelaide Microscopy).

4.3.8 Actin binding Assay

To assess actin structure, mature wild-type anthers were placed in 100 µl PEM buffer (0.1M PIPES (pH 6.95), 2 mM EGTA, 1 mM $MgSO_4$) and 2 µl Phalloidin. Anthers were crushed to release pollen and Phalloidin was excited at 488nm on a Nikon confocal microscope under green UV light.

4.3.9 Protein expression

In order to have a wild-type protein reference, I designed two vectors to transiently express OsEXPB10 and OsPRFAa in *E. coli* BL21 competent cells. OsEXPB10 CDS was cloned into the pMAL-c2x vector (Biolabs) using BamHI and HindIII restriction sites to express a fusion protein with a Maltose Binding Protein (MBP) tag.

Vectors were transformed into BL21 competent *E. coli* using heat shock, grown on LB medium with 50µg/ml ampicillin and screened with PCR using the Dream taq polymerase (Thermo Scientific) (95°C for 15s, 57°C for 20s, 72°C for 1min/kb X35 cycles) (Primers in Supplementary Table 4) and sequenced (AGRF). Selected colonies of pMAL1::OsEXPB10-MBP were grown in Amp+ liquid LB overnight at 37°C while shaking. The next day, 1.5ml of the solution was added

to 100ml fresh LB (amp+) and grown at 37°C until the solution reached an optical density of 0.5. LacI production was induced using 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and cells were grown at 37°C. Samples were taken at 0, 2 and 4 and 8 hours from induction. Protein was run on a 12% SDS-PAGE at 80V for 20 minutes and 120V for 1 hour.

4.3.10 *In situ* hybridization

In order to understand the localisation of mRNA in the anther and pollen, I attempted an *in situ* hybridisation. I designed probes for *OsEXPB1a* and *b, OsEXPB10, OsPRFAa* and *OsPRFAb.* Due to the 100% nucleotide identity between *OsEXPB1a* and *OsEXPB1b,* I treated these two genes as one, because they would likely have the same function, properties and mRNA localisation. It was difficult to create probes that specifically targeted single genes, due to the high nucleotide identity between gene family homologs, so I designed primers from the 3'UTR of all genes. In order to synthesise probes, I used a DIG-labelling system to transcribe a cDNA template and label the RNA.

Tissue samples were taken from developing anther stages 6-12 and imbedded in paraffin. Paraffin was sectioned to 8mm and mounted on a glass slide. *In-situ* hybridization was carried out as previously described (Yang et al., 2018, Zeng et al., 2017) .

4.3.11 Protein extraction from plants

Protein was extracted from heading inflorescence, mature flower or 2 week root by adding tissue to extraction buffer (Tris.Cl, pH6.8, 100 mM, SDS 4%, DTT 200 mM) and incubating on ice for 1 hour and centrifuging for 20 minutes. Supernatant was stored at 4°C. To measure the concentration of native protein from pollen extractions, a BioRad Quick Start Bradford Protein Assay was performed with 1/20 and 1/50 dilutions.

To assess antigen binding and to target antigens, a 16-amino acid peptide of Ory s 1 (PDNWKANALYKSEIQVD) and a 16 amino-acid peptide of Ory s 12

(VKKTGQALVVGIYDEP) were synthesised by ABclonal Technology as immunological peptides to generate an antigen-specific antibody.

Extracted pollen proteins from wild-type rice, *expb1-1, expb1-2, prfa-1, prfa-2,* Bahia grass (*Paspalum notatum*), Johnson grass, Bermuda grass, Ryegrass and Timothy grass standards Bahia grass, were run on a 14% 1mm SDS-PAGE and transferred to a nitrocellulose membrane at 30 volts for 2.5 hours. The membrane was blocked with 5% skim milk in 1X PBS (150 mM NaCl , 2.5 mM $NaH_2PO_4.2H_2O$, 7.5 mM Na_2HPO_4, pH 7.4) for 30 minutes. Nitrocellulose membrane was incubated in primary antibody mAb ory s 1, rAb ory s 12, rabbit polyclonal grass pollen group 1 allergen specific antibody and FMC mAb A1 (as described in Smart, 1984) with binding buffer (PBST 5% skim milk) at 4°C overnight. The next day, the membrane was washed in binding buffer and incubated in Horseradish peroxidase mouse IgG (1/1000) or rabbit secondary antibody.

After washing, the membrane was visualised without substrate on an Odyssey according to manufacturer's instructions.

4.3.12 Immunoassays

To assess the potential of allergy patient sera in binding to the Ory s 1 or Ory s 12 antigen, an ELISA was performed. Bahia grass, wild-type rice, *osexpb1-1, osexpb1-2, prfa-1,* and *prfa-2* were tested against 30 allergy patient sera and 8 non-allergic patient sera. Wells were coated 5 µg/ml antigen in 0.05 M Carbonate Bicarbonate buffer pH 9.6 (0.019 M Na_2CO_3, 0.028 M $NaHCO_3$) and incubated over night at 4°C. Samples were blocked with 1% skim milk in PBST (0.05% Tween in PBS) at 37°C for 1.5 hours. Wells were washed 3-6 times with PBST between each antibody incubation. Samples were incubated in 1/10 primary human allergy or non-allergy patient sera for 90 1.5 hours, then 1/2000 rabbit anti-human IgE for 1 hour, and finally 1/2000 goat anti-rabbit IgG for 1 hour.

Development was carried out using the Sigma OPD Fast kit according to manufacturer's instructions and absorbance was measured with a spectrometer.

4.4 Results

4.4.1 *OsEXPB* is expressed in late stage pollen development but does not affect pollen fertility

To assess gene expression at different pollen stages, a qRT-PCR was conducted from pollen stages 6-12. Both *OsEXPB10* and *OsEXPB1* showed low relative expression in pollen stages 6-11, but high relative expression at pollen stage 12 (4.1a) suggesting that both genes may be involved in late stage pollen development processes, such as pollen tube germination.

To assess fertility, mature pollen grains were stained with iodine and counted. The F2 double homozygous *expb* line (*expb-1*) had 87.1% pollen fertility compared to wildtype (wt) pollen fertility(89.8%, n>600, p=0.8) (Figure 4.1b-d). Likewise, F2 homozygous expb-2 and expb-3 did not differ in pollen fertility compared to wild type,

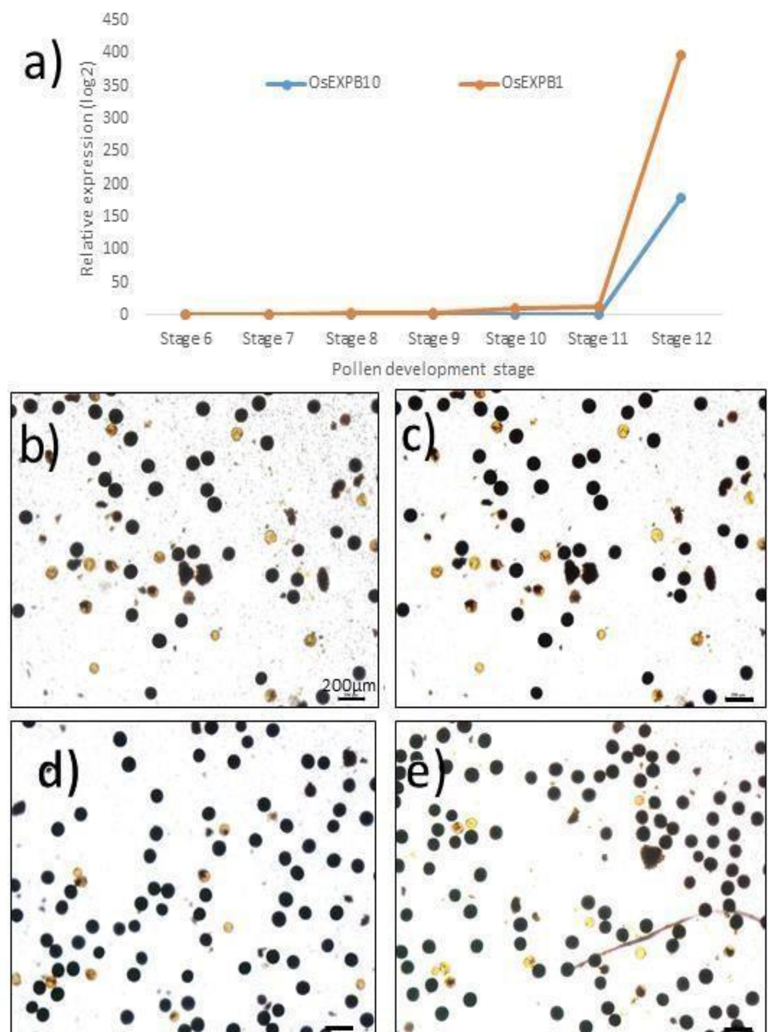


Figure 4.1 – pollen expression and and pollen fertility of *expb* mutants

a) Relative expression of OsEXPB10 (blue) and OsEXPB1 (orange) in pollen development stages. Actin was used as a control. Pollen viability of **b)** wild type, **c)** *expb-1,* **d)** *expb2,* **e)** *expb-3* (n>600 each genotype

showing 90.3% pollen fertility ($p=0.5$, $n>400$) and 90.69% ($n>400$ $p=0.4$) respectively. To understand whether changes to pollen structure occurred, pollen was assessed using SEM. All expb mutant pollen was indistinguishable from wild-type pollen grains (Supplementary Figure 4.3a-c, n=200). As both OsEXPB1 and OsEXPB10 were expressed in late stage pollen development, it is likely that these genes do not function during the crucial pollen development stages that could abort pollen development.

4.2 The pollen tube germination of *expb* was delayed

To understand whether mutant OsEXPB1/10 pollen tube germination was affected by the mutation, pollen grains were dusted onto wild-type pistils, which were sampled at 15 minutes, 30 minutes and 60 minutes (4.2a-d). Student t-tests were used for statistical analysis, and each experiment was replicated three times. No *expb-1* or *expb-2* pollen tubes germinated at 15 minutes, whereas *expb-3* pollen tubes germinated and average growth rate was 63.07 nm/s, which was significantly slower than a rate of 196.50 nm/s for wild type pollen tubes ($p= 1.36 \times 10^{-21}$)(4.2a-d, 4.3a). At 30 minutes, average mutant *expb* pollen tube growth was significantly lower than the wild type (253.77 nm/s); *expb-1* pollen tube growth rate was 68.75 nm/s ($n=52$, $p= 6.74 \times 10^{-10}$), *expb-2* growth rate was 96.75 nm/s ($n=52$, $p= 4.37 \times 10^{-11}$), and *expb-3* growth rate was 119.76 nm/s ($n=52$, $p=6.15 \times 10^{13}$) (Figure 4.2e-h, Figure 4.3a). Likewise, growth of *expb* mutant pollen tubes at 60 minutes was significantly lower than wild type (334.7nm/s); average *expb*-1 pollen tube growth rate was 97.61 nm/s ($n=52$, $p= 1.18 \times 10^{-05}$), *expb*-2 was 121.18 nm/s ($n=52$, $p= 1.27 \times 10^{-08}$) and *expb*-3 148.72 nm/s ($n=52$, $p= 4.22 \times 10^{-12}$) (Figure 4.2i-l, Figure 4.3a). Additionally, while no significantly different growth rate was seen between *expb-1* or *expb-2* and *expb-2* or *expb-3*, *expb-1* pollen tube growth was significantly lower than *expb-3* at 30 minutes ($p= 9.07 \times 10^{-05}$) and 60 minutes ($P= 3.39 \times 10^{-05}$). Average mutant *expb* pollen germination rates were lower than wild type (81.5%), where *expb-1* showed 41.95% germination ($p=0.004$, $n>100$) and *expb-2* showed 48.76% germination ($p=0.001$, $n>100$) and *expb-3* was 52.27% ($n>100$, $p=0.0008$) (Figure 4.3b).

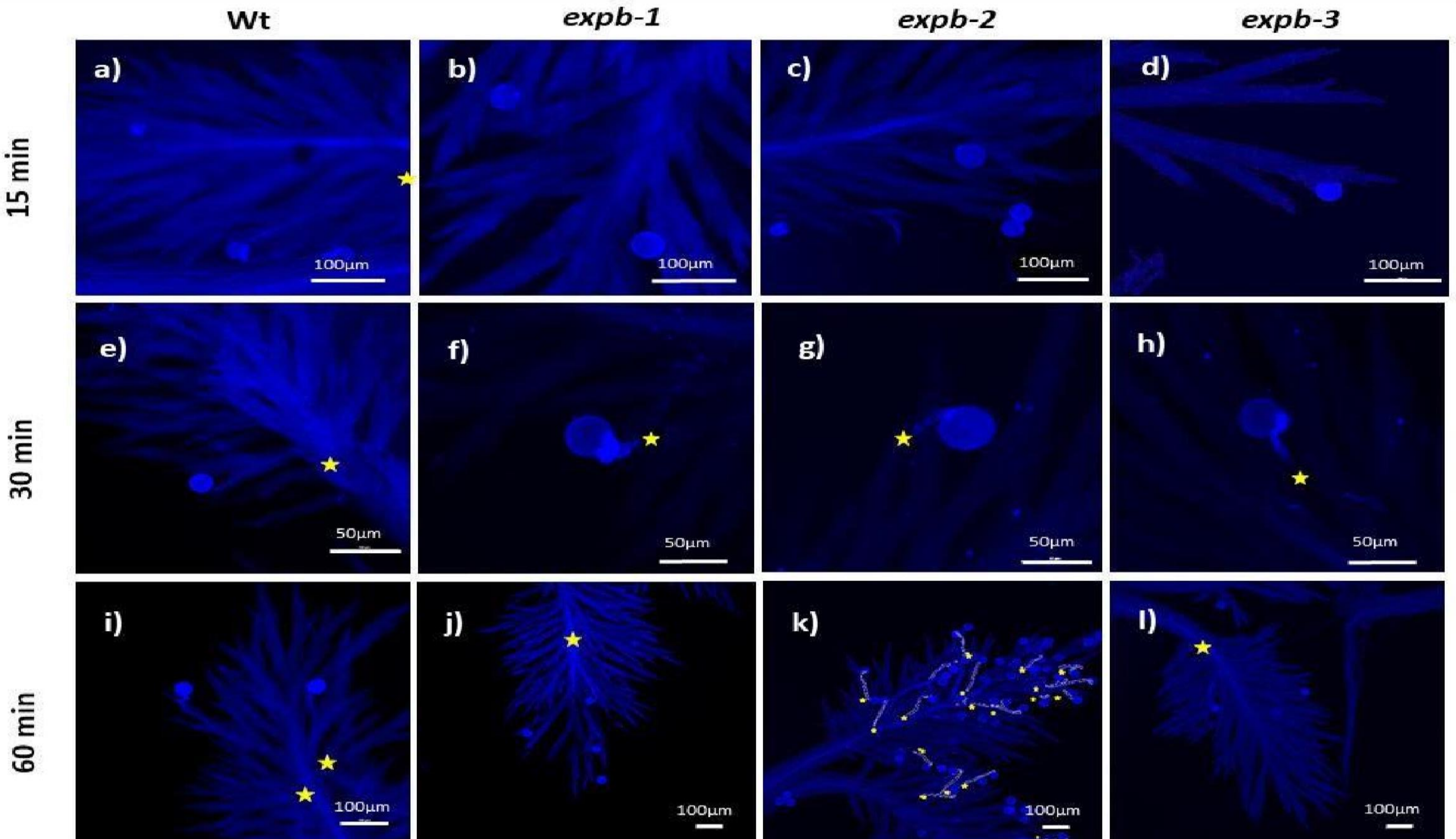


Figure 4.2 Aniline blue stained *expb* pollen tubes grown *In vivo*

Pollen tubes in wildtype (a,e,i), *expb-1* (b,f,j), *expb-2* (c,g,k) and *expb-3* (d,h,l) at 15 mins (a-d), 30 mins (e-h) and 60 mins (i-l) (n>50 each genotype). Tip of pollen tube is marked with a star.

However, there was no difference in pollen tube germination between mutant lines. No other phenotypic variation was seen betwe en wild type and mutant *osexpb* pollen tubes. It is likely that *OsEXPB* promotes pollen tube germination and growth, but some redundancy may be seen between genes, as heterozygous *expb-3* pollen tubes had a higher pollen tube growth rate than *expb-1*.

4. 4.3 Mutant *expb* had lower seed-setting rates

To assess grain fertility, seed setting rates were counted from 15 inflorescences from each line. Wild-type seed setting rate (85%) was significantly higher than F2 *exb-1* (41.7%, p=0.0003), *exb*-2 (56.7%, p= 0.0008) and *exb*-3 (44.3%, p=5.0x10^{-5}), suggesting that low pollen germination and tube growth rates may lead to the prevention of fertilization of the

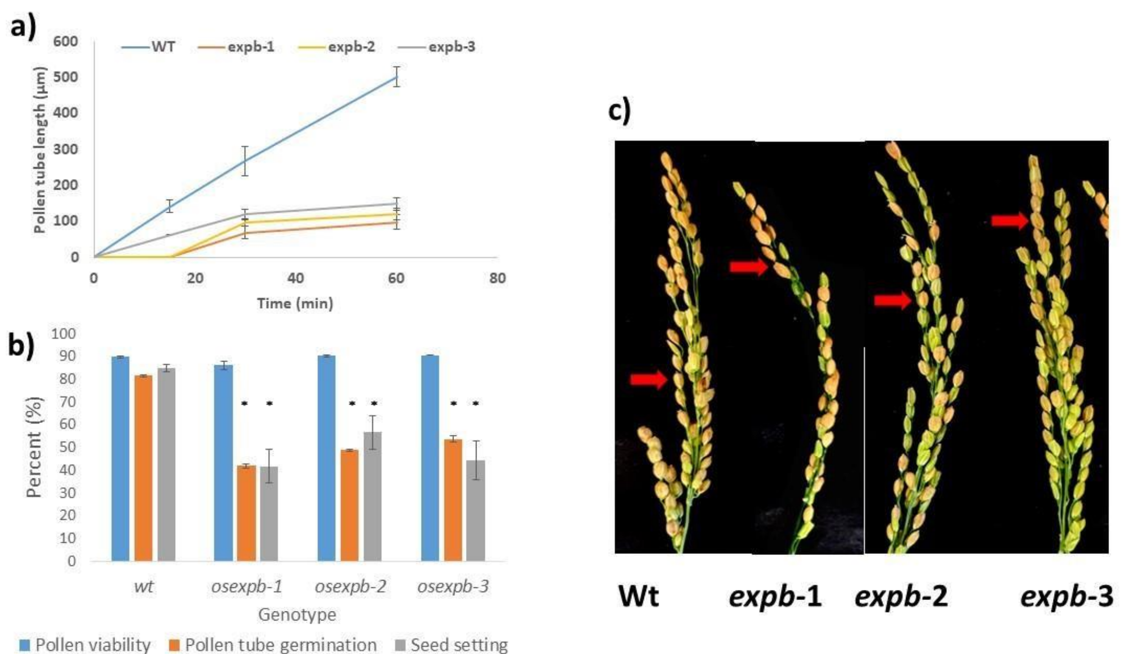


Figure 4.3 – Pollen tube growth rates and seed setting of *expb* mutants

a) Pollen tube growth rates (nm/min) of wildtype (blue), *expb*-1 (orange), *expb*-2 (grey), and *expb*-3 (yellow) (n>20). Error bars show standard error. b) Pollen viability (blue) (n>200 for each genotype), pollen germination (orange) (n>100 each genotype) and seed setting (grey) (n=15 panicles each genotype) of wildtype, *expb*-1, *expb*-2 and *expb*-3. Error bar shows standard error. Significant difference from wild type is marked with a star. ovule (Figure 4.3b-c). This may be controlled by fluctuations in environmental conditions, which further affect pollen tube growth rates.

4.4.4 OsEXPB10-GFP is localised to the pollen tube tip

To assess localisation of OsEXPB10, a transient protein expression of OsEXPB10-GFP was performed in tobacco pollen using a 35s::OsEXPB10-GFP vector. OsEXPB10 localised in the developing pollen tube near the pollen tube tip and had low expression in the pollen grain (Figure 4.7). Additionally, 35s::GFP pollen tubes had a growth rate of 30.23 nm/s at 4 hours growth and 32.62 nm/s at 8 hours (n=15), which was significantly lower than CaMV35s::OsEXPB10-GFP pollen tubes, which showed 133.86nm/s at 4 hours and 120.5 nm/s at 8 hours (n=15, $p=2.58 \times 10^{-7}$). It is likely that OsEXPB10 localises to the pollen tube tip in order to expand the cell wall longer, instead of wider, and that overexpression quickens this process of pollen tube growth.

Transient protein expression of OsEXPB10 in onion showed that OsEXPB did not localise in the nucleus, but did not localise elsewhere (Supplementary Figure 4.4a-b). It is likely that OsEXPB10 is specific to the pollen tube and was

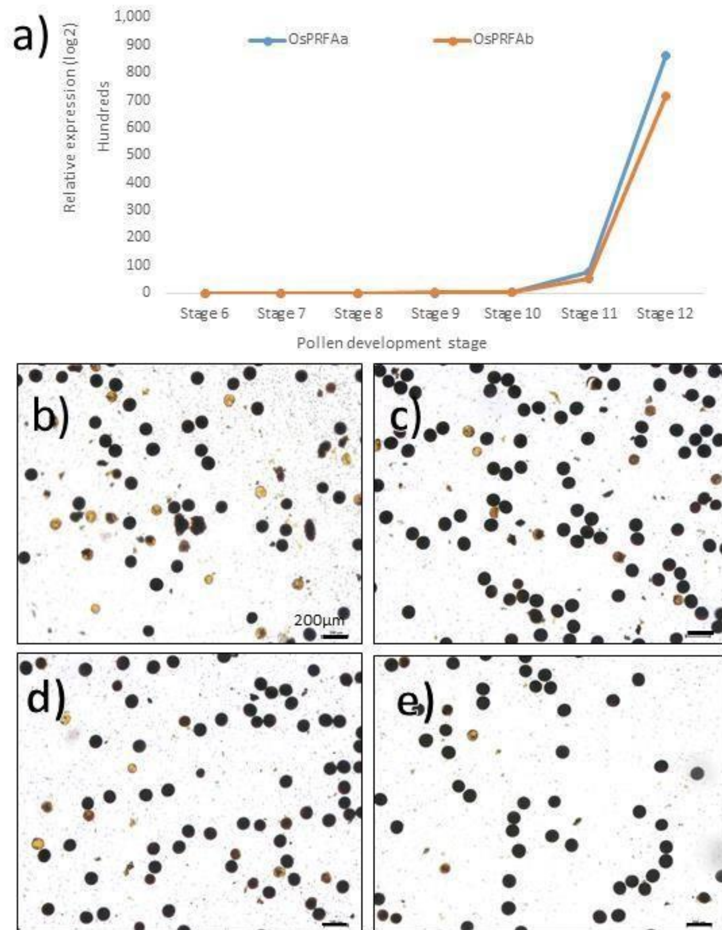


Figure 4.4 expression and pollen fertility of *prfa* mutants

a) Relative expression of OsPRFAa (blue) and OsPRFAb (orange) in pollen development stages. Actin was used as a control. Pollen viability of **b)** wild type, **c)** *prfa-1,* **d)** *prfa-2,* **e)** *prfa-3* (n>600 each genotype). Viable pollen are black and dead pollen are orange/brown. Scale bar is 200µm.

unable to localise anywhere in vegetative tissue.

4.4.5 OsPRFA is expressed in late stage pollen development but does not affect pollen fertility

OsPRFAa and *OsPRFAb* showed similar expression patterns. Both genes were slightly expressed at stage 11 and expressed highly at stage 12 (Figure 4.4a). This suggests that the genes may be functionally redundant and have a role in pollen tube germination. Like *expb* mutants, pollen viability of *prfa* CRISPR mutants was indistinguishable from wild type (89% n=800); *prfa-1* had 83.3% viability (n=500, p=0.8), *prfa-2* had 80.8% viability (n=400, p=0.12) and *prfa-3* had 83.1% viability (n=400, p=0.06) (Figure 4.4b-e). As both *OsPRFAa* and *OsPRFAb* were expressed in late stage pollen development, it is likely that these genes do not function during the crucial pollen development stages that could disrupt viability.

To understand whether changes to pollen structure occurred, pollen was assessed using SEM. All *prfa* mutant pollen was indistinguishable from wild-type pollen (Supplementary Figure 3d-e, n=200). As *OsPRFA* appears to be expressed in late stage pollen development, it is likely that pollen fertility was unaffected because protein may not have a function prior to stage 11 pollen development.

4.4.6 Mutant *prf* showed aberrant and delayed pollen tube growth

To understand the role of profilins in pollen tube development, pollen tubes were grown *in vivo* for 15, 30 and 60 minutes (Figure 4.5). At 15 minutes, average wild-type pollen tube length was 141.8nm (n=20), but pollen from *prfa*-1, *prfa*-2 and *prfa*-3 had not germinated (Figure 4.5a-d, Figure 4.6a). At 30 minutes, wild-type pollen tubes were 268.8 nm (n=30) on average, which was significantly longer than pollen tubes from *prfa1* at 85.6nm (n>20, p=9.7 x10^{-9}), *prfa-2* at 88.5nm (n>20, p=3 x10^{-11}), and *prfa-3* at 80.6nm (n>20, p=4.8 x10^{-10}) (Figure 5e-h, Figure 4.6a). At 60 minutes, the pollen tubes were significantly shorter than wild-type pollen tubes (503nm, n=20): *prfa-1* grew to 111.9nm

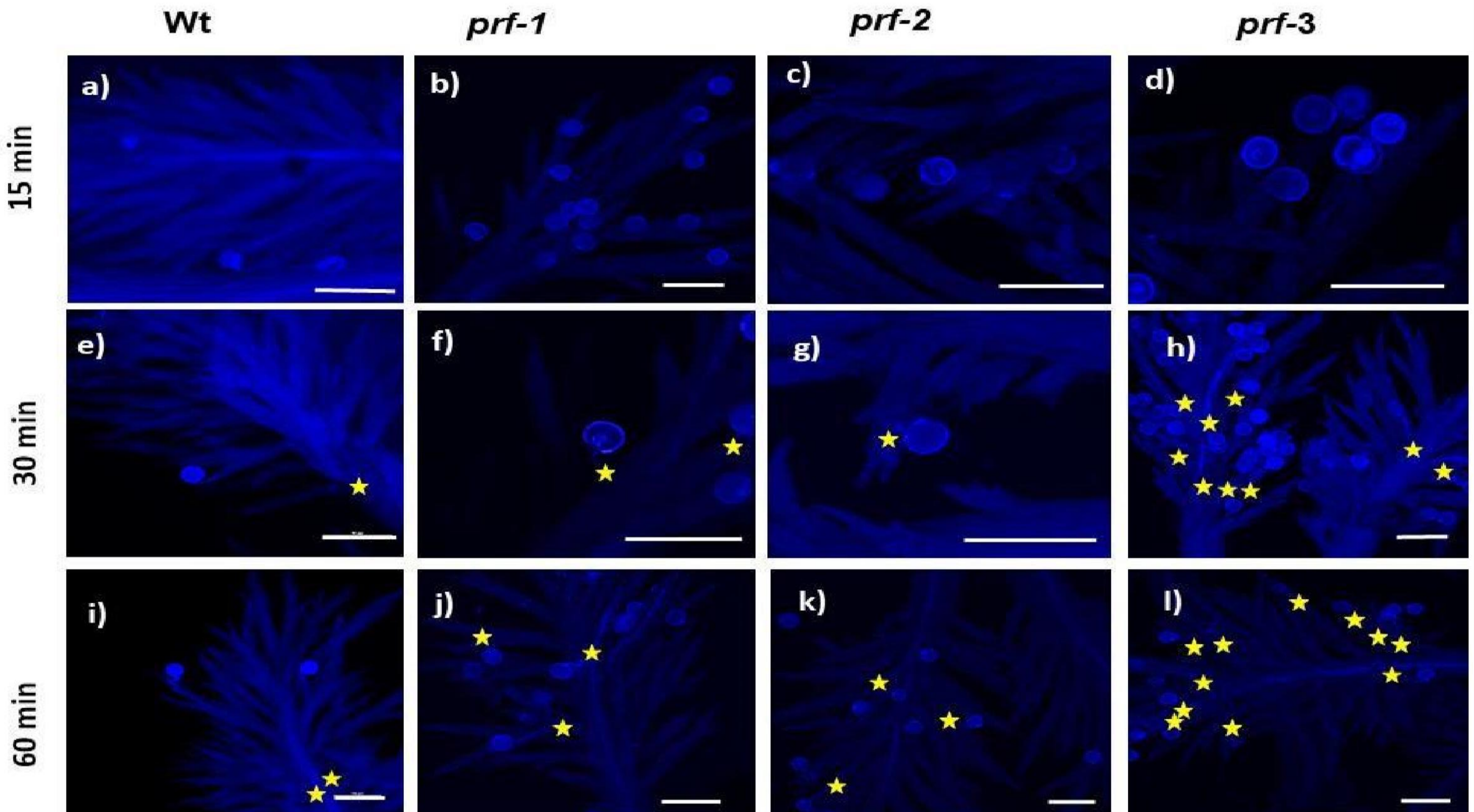


Figure 4.5- Aniline blue stained *prfa* pollen tubes grown *In vivo*

Aniline blue stained Pollen tubes grown *In vivo* from wildtype (**a,e,i**), *prfa-1* (**b,f,j**), *prfa2* (**c,g,k** and *prfa3* (**d,h,l**) at 15 mins (**a-d**, 30 mins **e-h** and 60 mins **i-l** (n>50 each genotype). Tip of pollen tube is marked with a star.

(n>20, $p=2.2 \times 10^{-21}$), *prfa-2* grew to 117.3nm (n>20, $p=2.2 \times 10^{-35}$), and *prfa-3* grew to 104.8nm (n>20, $p=5.3 \times 10^{-23}$) (Figure 4.5i-l, Figure 4.6a). This suggests that *OsPRFa* has an important function in pollen tube elongation and germination. The pollen tube length of mutants was not significantly different between 30 minutes and 60 minutes for *prfa*-1 (p=0.21), *prfa*-2 (0.07) and *prfa*-3 (p=0.21), but wild-type pollen tubes had an average growth rate of 148.9nm/s over the full 60 minutes, with no significant changes in growth rate for the duration, suggesting mutant pollen tube growth aborted around 30 minutes (Figure 4.6a). Mutant lines had a significantly slower growth rate than wild type; *prfa-1* pollen tubes grew at 39.3nm/min (p=0.001), *prfa-2* pollen tubes grew at 40.8nm/min (0.015) and *prfA-3* grew at 36.9nm/min (5.3×10^{-23}) (Figure 4.6a). There was no significant difference in growth rates between *prfa*-1, *prfa*-2 and *prfa*-3 (p=0.07, 0.07 and 0.06 respectively), suggesting there may be redundancy between *OsPRFAa* and *OsPRFAb*. Average mutant *prf* pollen germination rates were lower than wild type (81.5%), where *prfa-1* showed 37.9% germination (n>100, $p=6.63 \times 10^{-5}$), *prfa-2* showed 48.76% germination (n>90, $p= 3.13 \times 10^{-05}$) and *prfa-3* showed 45.32% (n>100, p=0.004) (Figure 4.6b). However, there was no difference in pollen tube germination between mutant lines.

4.4.7 *prfa* had low seed setting rates

To assess reproductive viability of *prfa* lines, seed setting was quantified. Wild-type plants showed seed setting rate of 83.5%, which was significantly higher than *prfa-1* at 34.86% ($p=11.1^{-16}$), *prfa-2* at 32.36% ($p=2.3^{-22}$) and *prfa-3* at 36.42% ($p=11.2^{-12}$) (Figure 4.6b-c). Potentially, low seed setting was due to aberrant pollen tube growth. No significant difference was seen between seed setting rates of *prfa-1*, *prfa-2* or *prfa-3*, likely due to similar pollen tube growth rates between mutants.

4.4.8 OsPRFAa-GFP was evenly distributed in the pollen grain and tube, but prevented tube growth

To assess localisation of OsPRFA, a transient protein expression of OsPRFAa-GFP was performed in tobacco using a pCAMBRIA 1301 35s::OsPRFAa-GFP vector. OsPRFAa-GFP didnot appear to localise to a specific part of the pollen grain or pollen tube, and was evenly distributed throughout (Figure 4.7d). Likewise, transient expression of OsPRFAa-GFP in onion cells showed an even distribution of protein across the whole cell (Supplementary Figure 4c). However, growth rate of over expression pollen tubes at 8 hours was 1.5nm/s (n=10), which was significantly lower than wild-type growth rate of 32.62 nm/s ($p=2.29x10^{-5}$). Likewise, this was significantly lower than the 8 hour growth rate of OsEXPB10-GFP (120.5 nm/s, $p=3.16x10^{-9}$), suggesting that an over expression of profilin leads to aberrant pollen tube growth in a similar manner to under-expression.

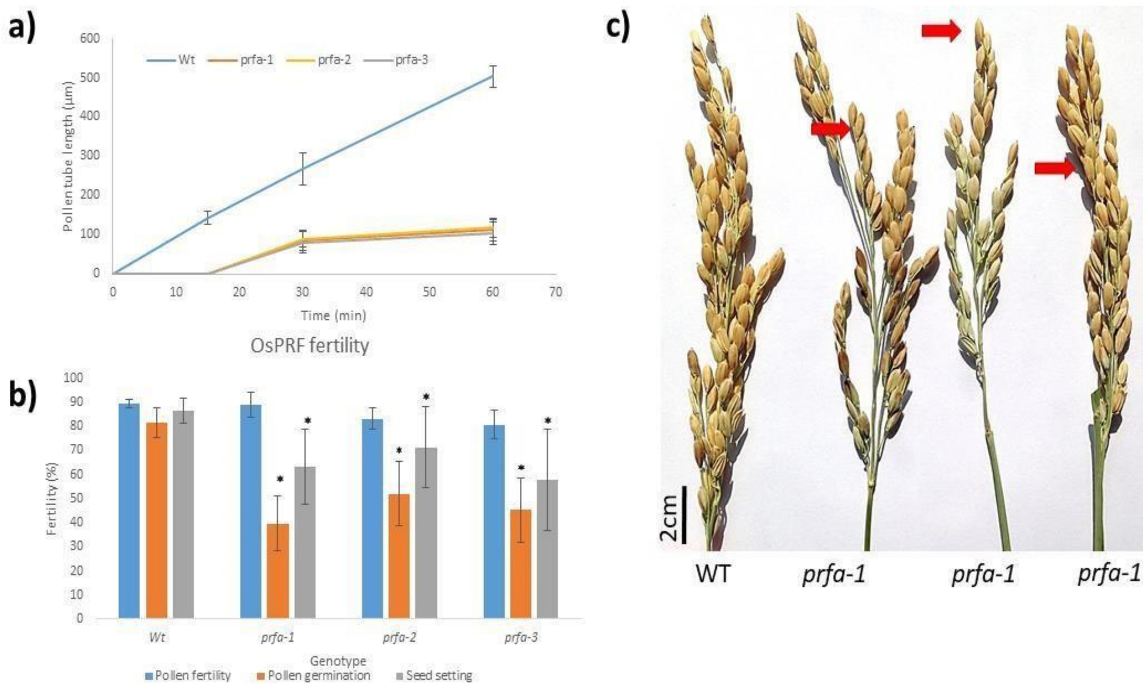


Figure 4.6 - Pollen tube growth rates and seed setting of *expb* mutants

a) Pollen tube growth rates (nm/min) of wildtype (blue), *prfa-1* (orange), *prfa-2* (grey), and *prfa-2* (yellow) (n>20). Error bars show standard error. **b)** Pollen viability (blue) (n>200 for each genotype), pollen germination (orange) (n>100 each genotype) and seed setting (grey) (n=15 panicles each genotype) of wildtype, *expb*1, *expb*-2 and *expb*-3. Error bar shows standard error. Significant difference from wild type is marked with a star.

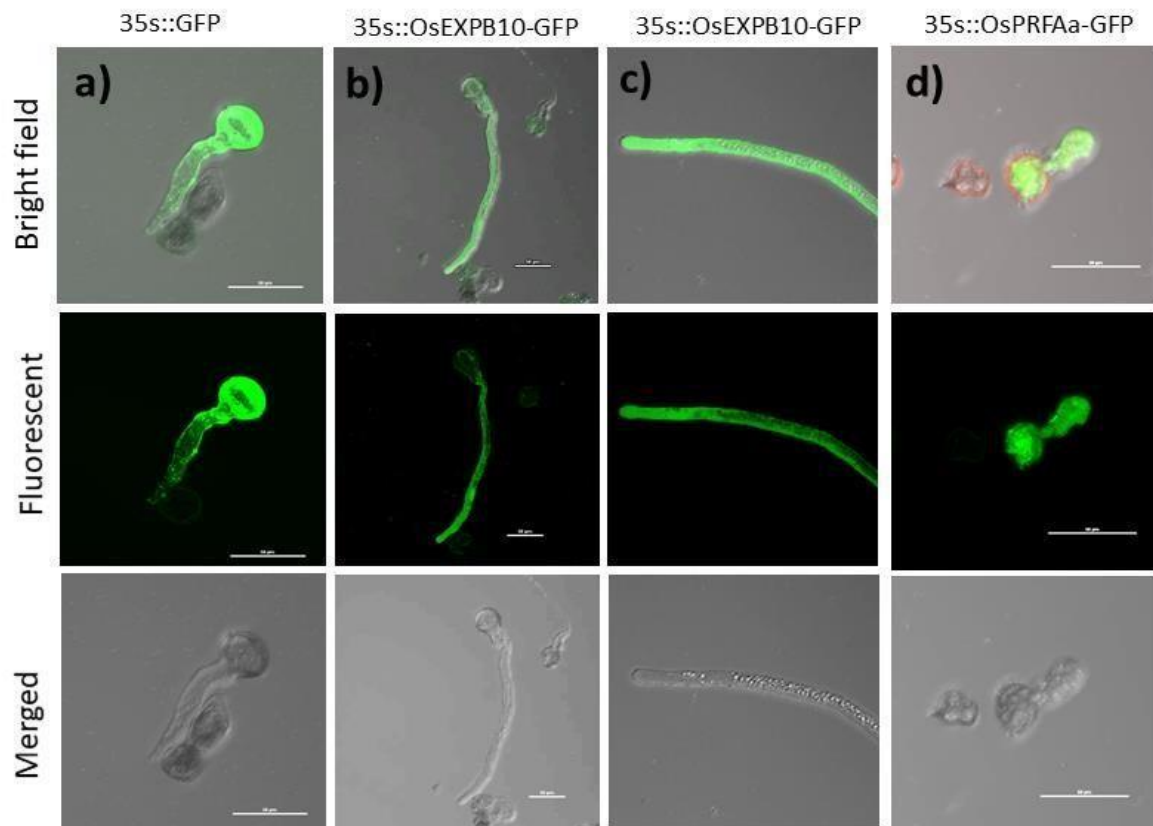


Figure 4.7 – Transient expression of GFP in tobacco pollen

Transient expression of a) 35s::GFP, b-c) 35s::OsEXPB10-GFP and d) 35s::OsPRFAa-GFP in tobacco pollen tubes after 4 hours of growth. Protein is localised to the pollen tip in 35s::OsEXPB10-GFP. Scale bars is 50µm

4.4.9 Actin filaments in *prf* were over-polymerised in pollen

To assess whether actin structure was affected in *prf* mutants, pollen grains were stained with phalloiden. Wild-type actin appeared as flecks that pointed toward the developing pollen tube, but actin in *prf-1* and *prf-2* pollen did not appear in any specific orientation. This may be because I was unable to germinate any mutant pollen tubes during this experiment due to not

ideal conditions, and future studies are necessary to highlight actin polymerization in the pollen tube.

4.4.10 Expression of protein and *in situ* hybridization were unsuccessful

In situ hybridization probe signal was shown to be present in stage 6 of pollen development, but not in later stages (Supplementary Figure 5). This is in contrast to the qPCR results, which showed higher expression in stages 11 and 12. Potentially, this was due to lack of healthy mature pollen on the slides. Mature pollen is very static and movable when dry, and pollen was easily lost from the slides during the *in situ* hybridization washes. It was therefore very difficult to retain enough pollen to properly assess the hybridization.

Alternatively, the probes I designed were not accurate or could not properly hybridize to the mRNA. Due to the high identity between gene family homologs, I was unable to design different unique probes. For this reason and the problems stated earlier, I did not consider this data in the functional analysis.

Similarly, I was unable to obtain any transiently expressed protein. I attempted to express the protein at 4°C overnight, at 28°C for 8 hours, and with 0.5mM IPTG and 1.5mM IPTG, but was still unable to obtain EXPB protein. Potentially, EXPB10 was unable to fold when expressed transiently. Due to the difficulty of this project and my focus on biological function, I did not continue with this experiment.

4.4.11 *Osexpb-1* mutant pollen showed decreased binding to a pool of allergic human sera

To assess whether protein decreased in mutant plants, 2.5 µg of total protein was run on an SDS-PAGE. Expansin protein in *expb-1* appeared to be lower in concentration than *expb-2* or wild-type. However, wild-type profilin did not appear to have a stronger band than *prfa-1* or *prfa-2* (Supplementary Figure 6). With specific antibodies, *expb-1* and *expb-2* showed bands that were faint compared to wild-type, *prf*

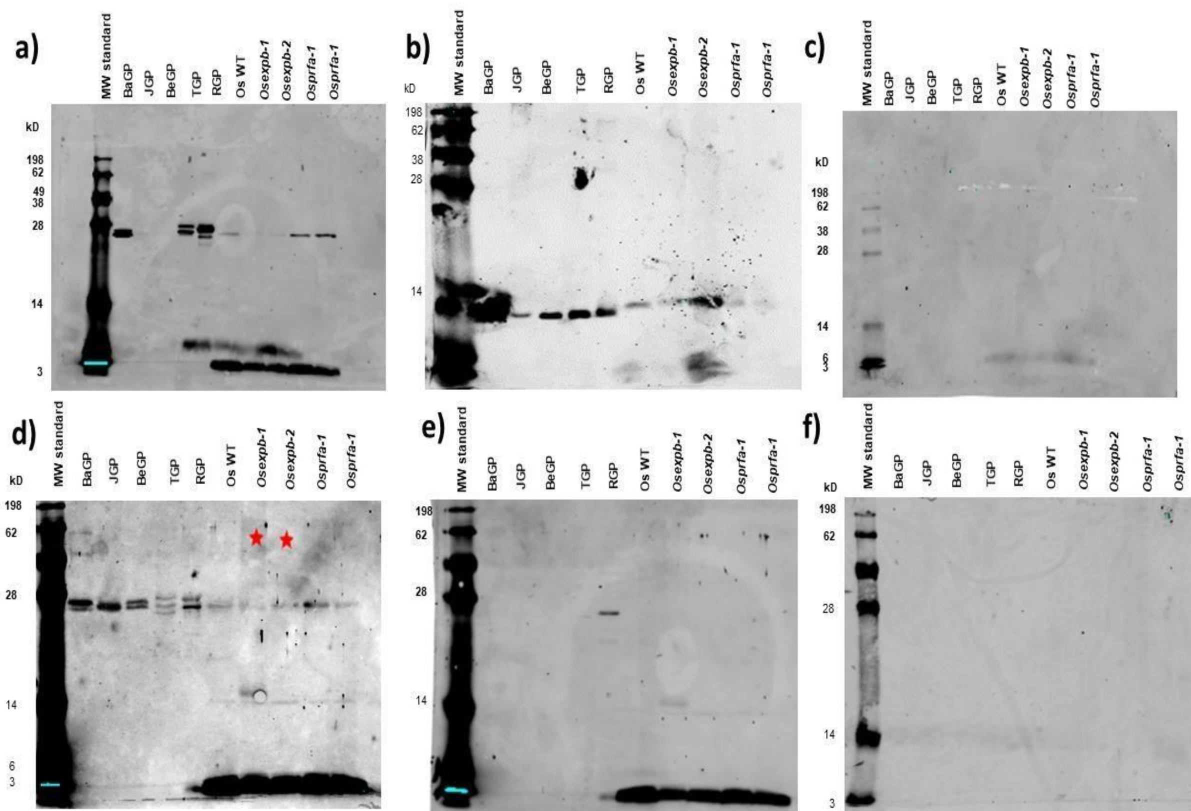


Figure 4.8 – Immunoassays of grass pollen proteins

Immunoblots of grass pollen with primary antibodies **a)** expansin A1 (Smith 1994), **b)** Ory s 12, **c)** negative control, **d)** pool of human allergic sera, **e)** pool of human nonallergic sera and **f)** human negative control. Antigen grass pollen protein samples including BaGP (Bahia grass pollen), JGP (Johnson grass pollen), BeGP (Bermuda grass pollen), TPG (Timothy grass pollen), RGP (Ryegrass pollen), Os WT (wildtype rice), *osexpb-1, osexpb-2, osprfa-1* and *osprf-2.*

2µg of protein was loaded in each well. Anti-expansin antibody also bound to one rice protein in *osexpb-1* near the 14kD marker, but this is likely to have been non-specific binding and is unlikely to reflect an allergen. No other binding to any rice proteins was seen in the human non-allergic patient sera or the negative control (Figure 4.8e-f).

mutants and other grasses, suggesting that Ory s 1 protein in these two mutants was greatly decreased (Figure 4.8a).

To understand whether mutant pollen showed decreased reactivity, the proteins were assayed against a pool of allergic human sera. For rice, patient sera bound to wild-type and *prfa-1* and

prfa-2, but not as strongly to *expb-1* or *expb-2* (Figure 4.8d). Binding to *expb-1* was extremely faint, suggesting *expb-2* and, particularly, *expb-1* showed lower reactivity with human sera and potentially lower allergenicity.

4.5 Discussion

4.5.1 *OsEXPB1/10* function is similar to maize homologs

Grasses share many similarities, and the function of OsEXPB1 may be conserved in other grass species. For example, maize β-expansin ortholog *ZmEXPB1* showed a similar phenotype to *expb1* mutants when mutated (Valdivia et al., 2009). Pollen tubes of Zm*expb1* plants showed delayed germination like *Osexpb1/10*, which lead to lower seed setting and reproductive success (Valdivia et al., 2009). However, Zm*expb1* mutants showed phenotypes not seen in OsEXPB1/10 mutant lines; Zm*expb1* mutant pollen tubes looped around each other and pollen aggregated together, unlike pollen from rice mutants.

This suggests that *OsEXPB1/10* may have a more specialised role in rice pollen than *ZmEXPB,* or that *ZmEXPB* has an extra role not present in rice. Potentially, expansins from each grass species have evolved individual roles beyond that of pollen germination since speciation. Further studies to find the specific functions of pollen expressed expansins in other grasses may answer this question.

4.5.2 *OsEXPB1/10* mutants set seeds, but were more susceptible to environmental conditions

OsEXPB1/10 mutants showed varied seed setting rates, ranging from 0%-70%. This seed setting rate varied widely, potentially due to environmental conditions. *OsEXPB1/10* mutant pollens were able to germinate and grow pollen tubes, but at a lower rate than wild-type pollen grains, despite no difference in mature pollen structure. Similarly, Zea m 1 caused aberrant pollen tube growth in maize, but did not affect the pollen wall structure, suggesting these expansins do not

have a roll in pollen development prior to germination (Valdivia et al., 2003, Valdivia et al., 2009). Rice pollen tubes grow rapidly but are highly susceptible to environmental changes, and more environmental changes are possible over a longer period of time (Coast et al., 2016, Thakur et al., 2010) . For example, pollen may reach the embryo in only 40 minutes and relatively little change in the environment may be evident in this short time period to inhibit pollen tube growth, but temperature and humidity may fluctuate more over a period of 1-2 hours and slowly growing mutant pollen tubes may abort due to this stress (Chen et al., 2008, Hedhly, 2011). In the case of *expb* mutants, pollen tubes were still short when grown for 60 minutes and did not reach the ovule. Environmental changes after this time may lead to aborted growth of these already short tubes, or prevent the tubes from continuing to grow.

When overexpressed, *OsEXPB10* caused rapid pollen tube growth in tobacco. This suggests that the crucial role of *OsEXPB1/10* in pollen tube development is to cause rapid growth of pollen tubes, potentially in order to overcome external stresses that may abort pollen tube growth. When comparing genotypes, *expb-3* plants performed better than *expb*-1 plants, and wild type performed the best. In the case of *expb*-3, no homozygous *expb1* was acquired, suggesting that, of *OsEXPB1a* and *OsEXPB1b* that had been treated as a single gene, one had not been mutated. This suggests that *OsEXPB1* and *OsEXPB10* are not fully redundant, but a higher number of functioning copies of *OsEXPB* lead to more efficient pollen tube growth.

Interestingly, most group 1 genes in grasses appeared to accumulate gene copy number following individual speciation, suggesting there may have been some simultaneous environmental event that lead to each species undergoing rapid evolution in order to compensate for fluctuating changes. High copy number may have evolved after sorghum and maize split but before rice species split. Potentially, a high copy number of group 1 genes evolved as recently as 10 million years ago.

The high potential of these genes to promote efficient pollen tube growth may provide interesting insights into how rice and other grasses overcome environmental changes. This may specifically be relevant in the face of changing global climate conditions, and provides a new area of study, whereby adding more β-expansin genes to grasses may help overcome pollen tube abortion due to changing external temperature.

4.5.3 Generation of functioning *osexpb* triple homozygous rice plants

The mutant *osexpb-1* contained a homozygous deletion in EXPB1a, EXPB1b and EXPB10, but still produced seeds. This suggests that EXPB proteins may not be essential in pollen tube development, despite promoting quick and efficient growth. Mutant *expb1* showed lowered fertility but not complete loss, which provides an exciting potential for hypoallergenic derivatives in non-agricultural grasses that do not rely on high seed yield.

Alternatively, EXPB13 retains a function in the absence of the other three homologs.

4.5.4 Potential function of the Pollen Allergen 1 domain

Mutant *osexpb* plants retained some pollen tube growth potential, even in the homozygous triple mutant. Mutant *osexpb* genes were generated by deleting a nucleotide upstream of the DBPP1 and PA1 domains, causing a truncated protein that lost both domains. Interestingly, in the case of double mutant *expb1-1,* pollen tubes were still able to germinate and grow, albeit at a slower rate. It is possible that *OsEXPB* promotes pollen tube growth and germination but is not necessary for germination. For example, *Osprf* mutants also showed decreased pollen tube germination rates, and functional profilin may be sufficient to promote germination in the absence of EXPB. Alternatively, *OsEXPB13* may retain some function, despite having a putative early stop codon that leads to a truncated protein containing only a DPBB1 domain and no PA1 domain. If the latter is the case,

OsEXPB13 would be able to function and show redundancy despite containing only a DPPB1 domain, suggesting that pollen tube germination is promoted by DPPB1. Potentially, the DPPB1 domain has a greater involvement in pollen tube germination, and the PA1 domain has a greater involvement in pollen tube growth rates. Future studies on *OsEXPB* mutants that contain only DPPB1 domains may shed light on whether the PA1 domain has a separate or specific function.

4.5.5 Function of *OsPRFA* is similar to other profilins

All mutations in *OsPRFA* showed similar phenotypes, suggesting the two genes are functionally redundant. PRFA mutant lines showed pollen fertility that was indistinguishable from wild-type, but pollen grains germinated at a lower frequency and pollen tubes grew more slowly, suggesting that PRFAs are both specifically involved in pollen tube germination and growth, instead of having a general role in actin-binding in other tissues like PRF-LP04 does. This phenotype is similar to tomato profilin *LePRO* antisense mutants, which showed low pollen germination and lead to slow pollen tube growth when grown *in vitro* and *in vivo* (Yu and Parthasarathy, 2014). This lower pollen tube germination and pollen tube growth lead to decreased seed setting in both rice and tomato mutants (Yu and Parthasarathy, 2014). Interestingly, there was no significant difference between mutant pollen tube length at 30 minutes and 60 minutes, suggesting that *prfa* pollen tube growth may abort between 30 and 60 minutes of growth on the pistil. No homozygous mutants were obtained during the experiment, suggesting mutant pollen tubes were unable to grow long enough to fertilize the ovule, leading to fertilization being carried out by wild type pollen tubes.

In Arabidopsis, mutants of pollen profilins 4 and 5 showed defective pollen tube growth and disorganised actin in the pollen tube (Liu et al., 2015). Double mutants showed less actin organisation than single mutants, suggesting that profilin 4 and 5 may have some redundancy (Liu et al., 2015). I was unable to obtain any double mutants, but this may be due to other factors.

4.5.6 Overexpression of *PRF* may also lead to aberrant pollen tube growth

Tobacco pollen tubes that transiently expressed 35s::OsPRFAa-GFP grew short pollen tube compared to wild-type. Similarly in lily, increased concentrations of PRF introduced via microinjection lead to thinner pollen tubes and erratic growth, suggesting that profilins can cause aberrant pollen tube growth and both high and low concentrations (McKenna et al., 2004, Vidali et al., 2001). Because profilins are involved in actin polymerization, lack of profilin may lead to low actin polymerization, and overexpression of profilin may lead to over-polymerization (Gibbon et al., 1997, Jimenez-Lopez et al., 2013, Sun et al., 2013, Wang et al., 2006, Yu and Parthasarathy, 2014). Mutant *prf* had disrupted putative actin binding sites, but functional redundancy may have overcome this, suggesting that 2 copies of this gene is functionally ideal to prevent low germination and pollen tubes from under or overexpression of reproductive profilins.

It is also possible that the localisation of *OsPRFAa-GFP* is not a reflection of the true localisation patterns of *OsPRFAa.* In tomato, *in situ* hybridisation showed that pollen profilin, *LePro1*, evenly localised in the pollen tube (Yu and Parthasarathy, 2014). However, if overexpression of *OsPRFAa* prevents normal pollen tube germination and causes pollen tubes to abort quickly, the pollen tube may have insufficient length to show true localisation of *OsPRFAa*. In my attempt to perform an *in situ* hybridization, I was unable to retain healthy mature pollen grains on the slide as they often dispersed during the washing process. Likewise, I was unable to obtain pollen tubes to imbed, because growth conditions were insufficient to germinate pollen tubes *in vitro.* However, future studies may reveal the true localisation of *OsPRFA* where better facilities are available.

4.5.7 Mutants of *OsEXPB* and *OsPRFA* show similar pollen tube germination phenotypes but not localisation or overexpression

Despite the different structures and evolution of *OsEXPB* and *OsPRFA* (as shown in chapter 2), these two allergen genes are expressed at similar stages and share a role in pollen tube germination and development. However, the specific functions of these genes differ. *OsEXPB* appears to be involved in rapid pollen tip growth, and *OsPRFA* helps organise actin within the pollen grain and tip, but may not accelerate tip growth. Mutant *expb* lines were also highly susceptible to environmental changes, suggesting *OsEXPB* protects pollen tube development from a fluctuating environment, whereas *prf* mutants did not appear to be affected in the same manner, and overexpression of PRFAa lead to slowly growing pollen tubes. Potentially, these two genes appear as pollen allergens predominantly because of their localisation to the pollen grain.

4.5.8 Pollen proteins are functionally important and may not have evolved to be allergenic

Both *OsEXPB* and *OsPRFA* appear to have important roles in pollen development, and knockout of these genes causes aberrant pollen tube growth and decreased fertility.

Profilins are an ancient gene that has been heavily retained, and β-expansins appear to experience selective pressure to retain copy number, despite evolving relatively recently (Chen et al., 2016, Devis et al., 2017, Sharma et al., 2017). Potentially, plant proteins did not evolve to have a role in plant defence against humans, but that they evolved important functions that promote viability. This suggests that allergenicity is not a result of shared evolution or function, and that allergic response evolved in humans independently of plant evolution.

4.5.9 Towards the generation of viable hypoallergenic rice plants

While little is known about hypoallergenic pollen, mutant grasses have been generated using RNAi that had decreased potential grain allergens (14-16kDa, 26kDa and 33kDa) and binding

to patient sera (Wakasa et al., 2011). With grain allergens, patients reacted to different mutant allergen combinations differently, suggesting that allergic reactions to a pool of proteins is both highly specified and may be caused by more than one protein. Therefore, it is unlikely that allergenicity will be decreased for all patients when generating hypoallergenic pollen as patients may react to multiple allergens at once. However, a ryegrass mutant with low Lol p 1, a homolog of Ory s 1 in ryegrass, showed decreased binding to allergy patient IgE, demonstrating that decrease of group 1 allergens can be achieved in other grasses (Petrovska et al., 2004). Similarly, mutation of pollen allergen Lol p 5 in ryegrass lead to decreased accumulation of this protein in pollen (Bhalla and Singh, 2004). However, no functional analysis of the *lol p 5* mutant plant was provided, so the viability of this plant is unclear. Potential hypoallergenic plant derivatives have been developed that do not affect viability, but have not been tested to show decreased patient sera binding. For example, mutants were generated in the 14-

16kDa protein, likely the α-amylase/trypsin inhibitor, that did not show significant disruption to plant viability, but these were not immunologically tested and may not show reduced binding to patient sera due to the individual nature of allergies (Nakamura and Matsuda, 1996, Tada et al., 1996). While mutant *expb-1* and *expb-2* showed lower binding to human allergy patient sera, viability of the rice plant was compromised, suggesting that a fully fertile mutant Ory s 1 rice plant is not possible. However, the high number of similar, potentially non-allergenic rice expansins could provide opportunities to replace the pollen specific, allergenic expansins with a hypoallergenic gene under the native Ory s 1 promoter. Alternatively, truncated proteins may lead to a lower occurrence of allergenicity whilst maintaining viability. If a plant contained truncated versions of OsEXPB1a,b and OsEXPB10, but retained near normal growth, this would lead to exciting questions about the viability of a truncated functional protein that contains a DPBB1 domain without an allergenic PA1 domain. Furthermore, site-specific mutation of

known IgE hotspots may lead to generation of a hypoallergenic plant, but no biological data is present about this possibility in plants (Levin et al., 2013)

I was unable to obtain a viable profilin mutant that showed decreased reactivity, but other technologies may target this problem. Testing for hypoallergenic profilin variants in tomato was achieved by screening against yeast profilin variants, and this technique could be employed to determine profilin deficient pollen (Paulus et al., 2012). However, this would still require functional analysis of these potentially hypoallergenic profilin derivatives, as *Osprfa* mutants showed compromised viability with the loss of only one allele. Furthermore, I was unable to confirm that Ory s 12 bound to allergic sera of Queensland's patients, but this may not be a true representation of profilin allergic patients. Therefore, testing of Ory s 12 against a tropical patient sample is necessary to determine the significance of this allergen. This highlights the importance of understanding both functional molecular genetics of and immunological response to allergens when generating hypoallergenic derivatives.

The final aim of this project was to assess the possibility of removing or decreasing allergenicity in pollen. Pollen from two *expb* mutants showed lower binding with human allergic sera, suggesting that these two mutants do not illicit strong allergenicity in respect to Queensland patients allergic to group 1 grass pollen allergens. Decreased allergenicity in a viable plant provides an exciting future for pollen allergy prevention.

4.5.10 Conclusion

Despite eliciting allergenicity, two rice pollen allergens, Ory s 1 and Ory s 12, had varied and specific roles in pollen tube development. It is likely that patients have been sensitised to these proteins due to their abundance in mature pollen, as opposed to a shared function. Despite the role of Ory s 1 in pollen tube development, a homozygous derivative, *expb1,* retained seed setting and showed decreased binding to human pollen allergy patient sera, suggesting the new

line is both viable and hypoallergenic. This provides exciting information towards generating hypoallergenic derivatives of potent grass pollen in the

future.

Chapter 5 – Summary of Thesis

5.1 Introduction

Allergens are proteins that cause allergies such as hay fever, and they are particularly prevalent in grass pollen. While many studies investigate the human response to allergens, little is known about the evolution and biological function of these proteins in plants. In this thesis, I attempted to determine whether there was an evolutionary or functional correlation between two rice pollen allergens, Ory s 1 and Ory s 12.

Firstly, I used phylogenetic analysis to understand how these genes evolved in grass pollen. I also assessed whether there were any structural similarities between these families by assessing secondary and tertiary structure of a range of homologs.

Secondly, I used CRISPR/Cas9 to construct rice plants that contained a mutant copy or copies of Ory s 1 or Ory s 12. I assessed the offspring of these plants to determine the viability of a knockout or knockdown of allergenic genes.

Finally, I determined the function of these genes by assessing pollen fertility, pollen tube germination and growth efficiency, seed setting, gene expression and transient allergen-GFP localisation. I also assessed whether allergy patient sera would bind to mutant protein in order to determine whether the mutant plants were hypoallergenic.

5.2 Outcome of Aims:

1. Sequence analysis and evolutionary analysis of rice pollen allergens Ory s 1 and Ory s 12

As outlined in chapter 2, I analysed the evolution of Ory s 1 and Ory s 12 using phylogenetics and protein modelling. Two gene families were identified as homologs; Ory s 1 was homologous to β-expansins and Ory s 12 was homologous to profilins. I found that these two gene families had different evolutionary paths, despite their common allergenic property. βexpansins were

found only in monocots and appeared to evolve multiple gene copies following genus divergence. However, profilin copy number was retained from an ancient gene duplication that occurred prior to monocot/dicot divergence and experienced little change
since.

Protein modelling showed that Ory s 1 and Ory s 12 did not share common protein structures, with the exception of containing disulphide bridges. Based on this evidence, I surmised that it is unlikely that allergenicity was a result of shared evolution or structure.

2. Mutation of Ory s 1 and Ory s 12 using CRISPR/Cas9

As outlined in Chapter 3, I attempted to generate mutant rice lines in order to understand the function of rice allergens. I identified homologous genes to target in order to generate CRISPR mutations; OsEXPB1 and OsEXPB10 for Ory s 1, and OsPRFAa and OsPRFAb for Ory s 12. These genes are specifically expressed in pollen, so complete knockout may have led to a sterile plant. To overcome this, I generated plants that may have a partially functioning protein that did not contain known IgE recognised sites and assessed heterozygous plants that retained some viability. I was successfully able to transform the plants and obtained 6 unique mutant lies.

3. Functional analysis of mutant *ory s 1* and *ory s 12* pollen and anthers

Chapter 4 outlines the functional analysis of the mutant lines generated to fulfil Aim 2. I analysed pollen and pollen tube phenotypes of *OsEXPB* and *OsPRFA* mutants and found that pollen fertility was unaffected, but both genes showed aberrant pollen tube germination and growth, which lead to low seed setting and reproductive fitness. However, the two phenotypes differed slightly, as overexpression of *OsEXPB10* lead to quick pollen tube growth, whereas overexpression of *OsPRFAa* lead to underdeveloped pollen tubes and low germination rates.

Despite the similarity of expression in the mature pollen grain and correlations between pollen tube growth, it is possible that allergenicity is not linked to function. Potentially, allergenicity to these two families develops in humans do the exposure through pollen inhalation.

4. Allergenic assessment of mutant lines

Chapter 4 addresses the possibility of a viable hypoallergenic rice derivative. In this project, I was successfully able to generate a semi-viable Ory s 1 rice mutant, *osexpb-1,* that showed decreased binding to Queensland pollen-allergic patient sera. This shows the exciting potentially for the generation of hypoallergenic grasses in the future.

5.3.1 Future Directions

Hay fever is a highly individual disease that differs from person to person. Due to this, it would be nearly impossible to completely knockout every gene with allergenic potential from every plant. However, addressing the major contributing gene families may provide relief for a large body of sufferers. In rice, I was able to produce a functional plant that contained a mutant copy of all *OsEXPB* alleles, and this could potentially occur in other, more potent, grasses such as wheat and barley. Furthermore, the future holds exciting potential for targeted mutation and transformation in grasses such as Bermuda grass, Timothy grass and Johnson grass.

As more genomes are sequenced, thorough phylogenetic and structural analysis of Australia grasses may be carried out, to identify whether there are any key differences in evolution or sequence between temperate and tropical grasses and whether they may have a link to allergenicity.

During phylogenetic analysis, I identified another class of potential allergens, Ory s 2, that appear to be homologous to the pollen allergen domain of Ory s 1 and timothy grass allergen Phl p 2 (Abou Chakra et al., 2012; Marth, Focke, Flicker, & Valenta, 2004). These short proteins may have allergenic potential, and immunological assessment could be carried out to determine

a difference in IgE-binding efficiency between the two families. This may reveal if IgE binding is influenced by the PA1 domain alone or Ory s 1 as a whole. Furthermore, functional analysis of this family may shed light onto whether these genes hold a function in the plant and whether the pollen allergen 1 domain has a specific, individual function in or outside of the pollen grain.

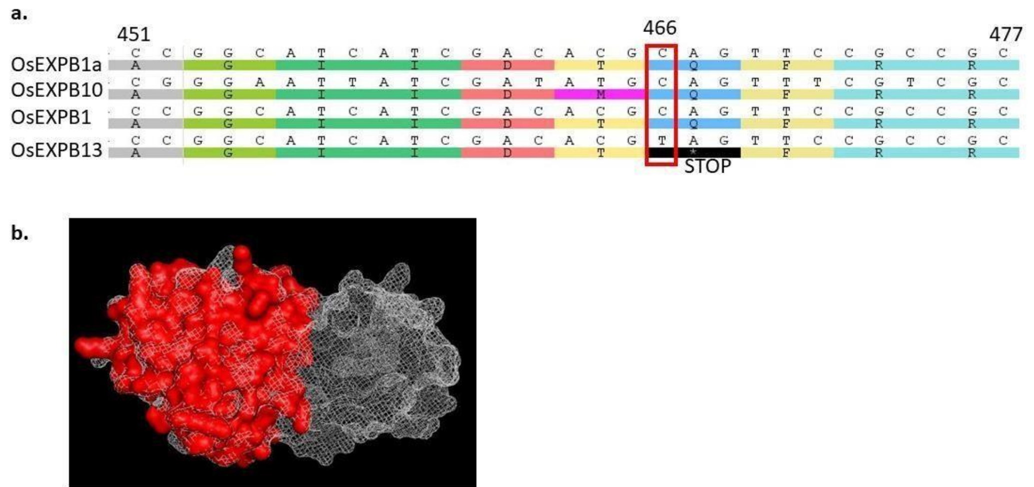
Knockout of EXPB1 and EXPB10 lead to varied fertility. Empirically, I noticed that fertility of *expb* mutants increased under high heat and humidity but this was not quantified. Future studies could assess the plausibility of an inducible sterile line, whereby pollen tube germination is inhibited by temperature or humidity without the need for transgenics, and this may be useful for selective and hybrid plant breeding (Rao, Deveshwar, Sharma, Kapoor, & Rao, 2018; Shukla et al., 2017; Wan et al., 2019).

Alternative methods of decreasing allergenicity in pollen to produce a viable, hypoallergenic pollen derivative still face the challenge of application. If a dominant mutation is generated that does not affect viability, it is possible that this plant could breed with other plants of the same species, whereby the whole population could take up mutant pollen and incorporate this trait in future generations. However, further research would need to be conducted to assess the affect such a method would have on the local and native ecosystems.

Cross-disciplinary research is essential to fully understand how allergens evolved, their role in plants and how this affects humans. This thesis sheds light on two rice pollen allergens, but many more grass allergens have yet to be functionally characterised. As genomic data and genetic modification technology improve, the future holds exciting opportunities to generate novel ways of preventing hay fever.

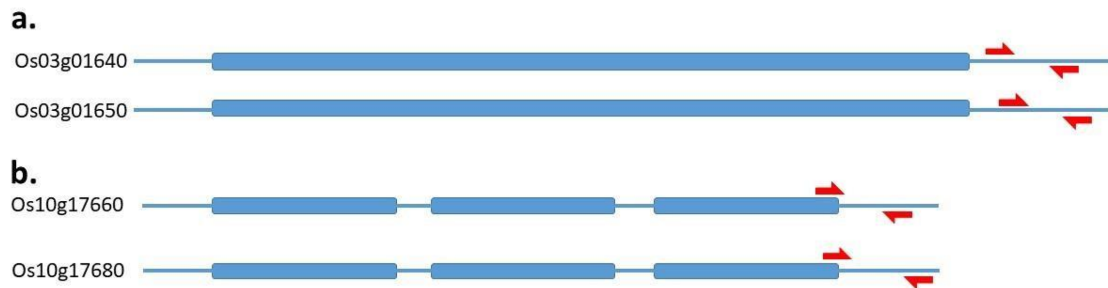
AIT	Allergen-Specific immunotherapy
RNAi	RNA interference
IgG	Immunoglobulin G
DEBP	Divergent Expansin β - Pollen
DEBV	Divergent Expansin β - Vegetative
CEB	Conservative Expansin β
IgE	Immunoglobulin E
PCR	Polymerase Chain Reaction
DPBB	5’ double-psi beta-barrel
PRF	Profilin
EXPB	β-expansin
SEM	Scanning Electron Microscopy

Supplementary Figures



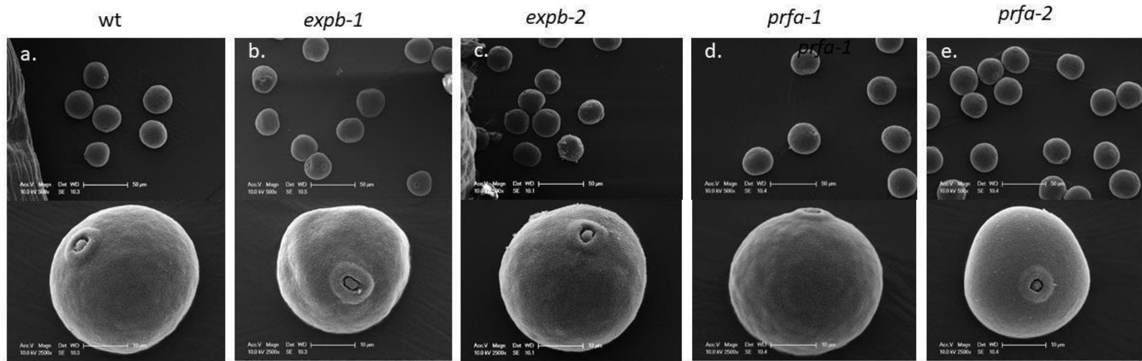
Supplementary Figure 1 – Comparison of expansin amino acid sequences to OsEXPB13

a) Alignment of nucleotide sequence and amino acid sequence showing a substitution mutation at 466 C->T in OsEXPB13 (Red box). **b)** Predicted tertiary structure of truncated *OsEXPB13 (red)* compared to predicted structure of *OsEXPB1* (grey mesh).



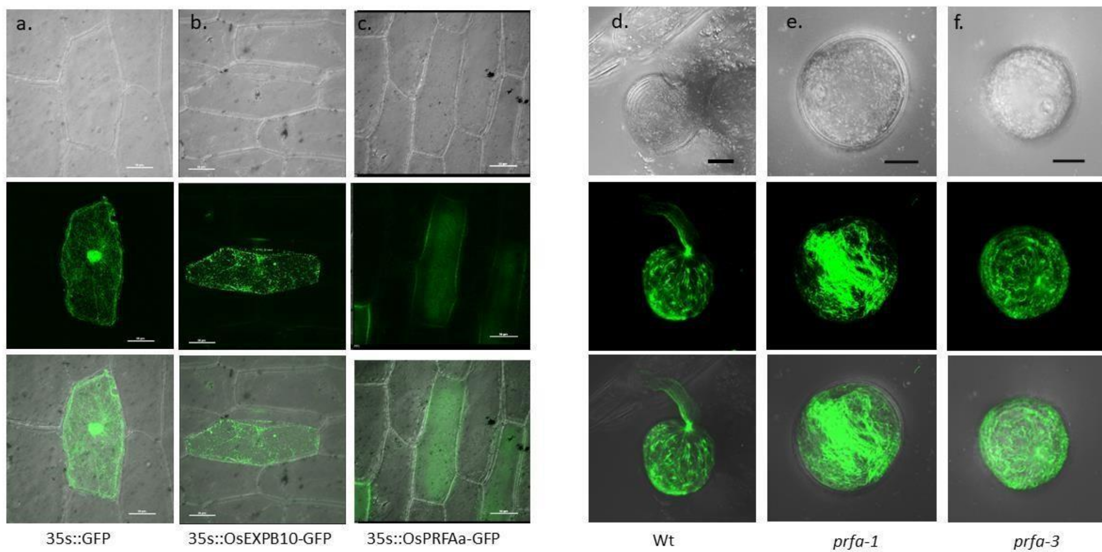
Supplementary Figure 2 – Regions amplified during qRT-PCR

a) Structure and primers of Os03g01640 and Os03g01640 **b)** Structure and primers of Os10g17660 and Os10g17680. Untranslated regions are represented by a line, Exons are shown as a box, and primers are shown as red arrows.



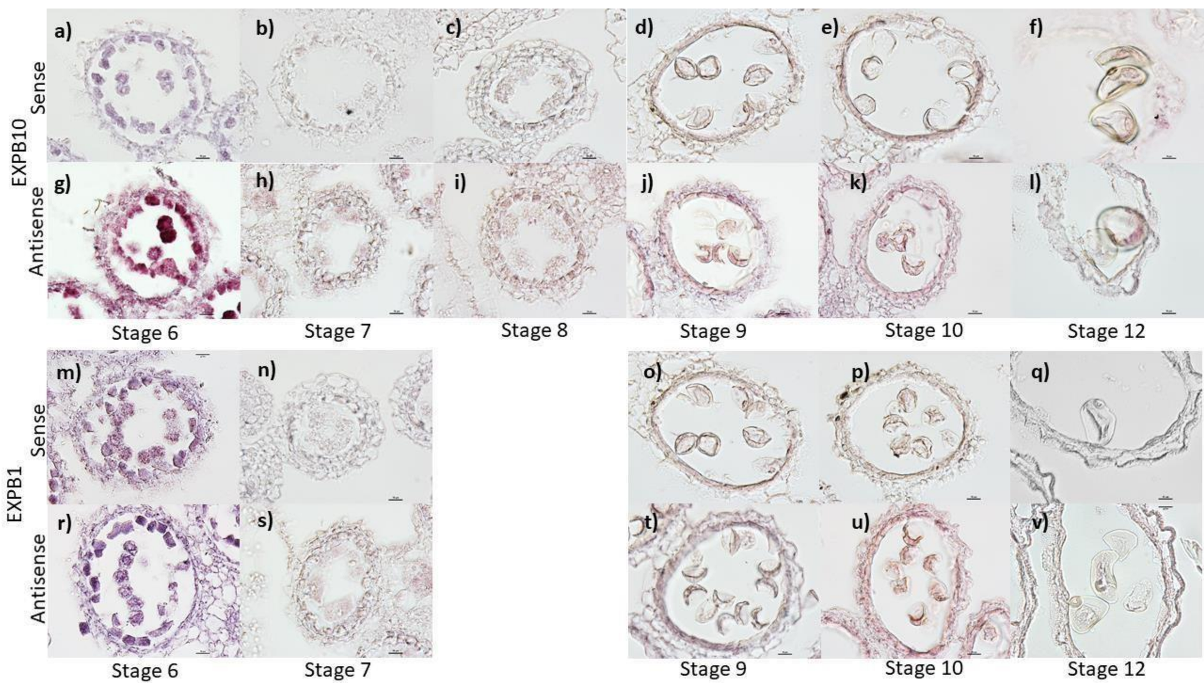
Supplementary Figure 3 – SEM images of mutant pollen

Scanning electron microscopy of mature pollen from a) wt b) *expb-1* c) *expb-2* d) *prfa-1* and e) *prfa-2*. Scale bar shows 50µm



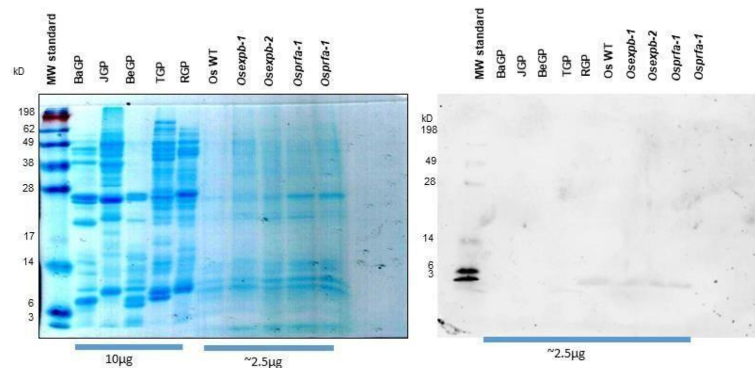
Supplementary Figure 4 – SEM images of mutant pollen

Scanning electron microscopy of mature pollen from a) wt b) *expb-1* c) *expb-2* d) *prfa-1* and e) *prfa-2*. Scale bar shows 50µm



5 – *in situ* hybridization of OsEXPB1 and OsEXPB10 during pollen development

OsEXPB10 sense (a-f) and antisense probe (g-l) in stages 6-12 of pollen development. OsEXPB1 sense (m-q) and antisense (r-v) probes in stages 6-12 of pollen development (excluding stage 8). Darker pink in the antisense shows binding of probe to mRNA.



Supplementary Figure 6 – SDS-PAGE and negative immunoblot of grass pollen proteins

a) SDS-PAGE of protein extracted from Bahia grass (BaGP), Johnson grass (JGP), Bemuda Grass (BeGP), Timothy Grass (TGP), Ryegrass (RGP) and rice pollen.10µg of non-rice grasses pollen protein and 2.5µg of rice pollen protein was loaded in each well. b) Negative control immunoblot incubated with rabbit secondary antibody but no Ory s 12 primary antibody. This immunoblot was done concurrently with Figure 4.8b

Supplementary Table 1 – EXPB homologs

Species Name	Designated Gene name	Allergen/Other	Genbank ascention	Phytozome ascention	Chromosome	Protein Sequence	Protein length	Protein similarity to OsEXPB10
za sative **(Rice)**	OsEXPB10			Os03g01640	3	MASSCLLLACVVAAAMVSAVSCGPPKVPPGPNITAAY GKQWLEARGTWYGKPKGAGPDDNGGACGYKDIDKA PFLGMNSCGNDPIFKDGKGCGSCFEVKCSKPEACSD KPVIIHITDMNTEPIAAYHFDLSGHAFGAMAKEGKDEEL RKAGIIDMQFRRVRCKYPGETKVTFHVEKGSNPNYFA VLVKYVGGDGDVVKVELKEKGSEEWKPLNESWGAIW RIDTPKPLKGPFSLRVTTESDQKLVANDVIPDNWKANA LYKSEIQVD*	267	
	OsEXPB1a	Ory s 1	AAA86533	Os03g01610	3	MASSSLLLACVVVAAMVSAVSCGPPKVPPGPNITTSY GDKWLEAKATWYGAPKGAGPKDNGGACGYKDVDKA PFLGMNSCGNDPIFKDGKGCGSCFEIKCSKPEACSDK PALIHVTDMNDEPIAAYHFDLSGLAFGAMAKDGKDEEL RKAGIIDTQFRRVKCKYPADTKITFHIEKASNPNYLALLV KYVAGDGDVVEVEIKEKGSEEWKALKESWGAIWRIDT PKPLKGPFSVRVTTEGGEKIIAEDAIPDGWKADSVYKS NVQAK	267	82
	OsEXPB13			Os03g01630	3	MASSSLLLASVVVAAMVSAVSCGPPKVPPGPNITASY GDKWLEARATWYGAAKGAGRKDNSGACGYKDVDKA PFLGMNSCGNDPIFKDGKGCGSCFEIKCSKPKACSDK PVLIHVTDMNDEPIAAYHFDLFGLAFGAMAKDGKDEEL RKAGIIDT*FRRFKCKYPADTKITFHIEKASNPNYLALLV KYVAGDGDVVEVEIKEKGSEEWKALKESWGAIWRIDT PKPLKGPFSVRVTTEGGEKIIAEDAIPDGWKADSVYKS NVQAK	267	70
	OsEXPB1b	Ory s 1		Os03g01650	3	MASSSLLLACVVVAAMVSAVSCGPPKVPPGPNITTSY GDKWLEAKATWYGAPKGAGPKDNGGACGYKDVDKA PFLGMNSCGNDPIFKDGKGCGSCFEIKCSKPEACSDK PALIHVTDMNDEPIAAYHFDLSGLAFGAMAKDGKDEEL RKAGIIDTQFRRVKCKYPADTKITFHIEKASNPNYLALLV KYVAGDGDVVEVEIKEKGSEEWKALKESWGAIWRIDT PKPLKGPFSVRVTTEGGEKIIAEDAIPDGWKADSVYKS NVQAK*	267	82

Ananas comosus **(Pineapple)**	AcEXPB6a	Aco012814.1	OAY68019.1		MDTALVTWHICARSEQKHSNFSADAASQWGPASATWYGDPHGAGPADNGGACGFKNVNLPPFSSMTSCGNPALFKSGKGCGACYQVKCTSHPACSGNPSTLVITDACLGGVCLDAPFHFDMSGTTFGSMAKPGLDDELRHAGKIPIQFTRVPCNYPGLNIAFHVENGSNPFYFAILIEYENGDGDLTAVDLMEGGSHGQGVWTPMRESWGAIWRLDSNHPLQGPFSVRLTTLTSGKTLVAADVIPANWKPLTTYRSVVNFSN	276
	AcEXPB6b		XP_020085129.1		MALRKGFGFCTFYFIFAFLSSFSICARSEQKHSNFSADAASQWGPASATWYGDPHGAGPADNGGACGFKNVNLPPFSSMTSCGNPALFKSGKGCGACYQVKCTSHPACSGNPSTLVITDACLGGVCLDAPFHFDMSGTTFGSMAKPGLDDELRHAGKIPIQFTRVPCNYPGLNIAFHVENGSNPFYFAILIEYENGDGDLTAVDLMEGGSHGQGVWTPMRESWGAIWRLDSNHPLQGPFSVRLTTLTSGKTLVAADVIPANWKPLTTYRSVVNFSN	276
	AcEXPB3b	Aco011715.1	XP_020094823		MAFSLSGSSSVVLLVSLMFVCPCFCFKSKSFNFSTSEAALNWNTASATWYGSPTGAGPDDNGGGCGYKNVNLPPFSSMTSCGNPALFKSGEGCGSCYKVKCNSHPACSGKPSTLIITDACLGGVCLDKTYHFDMSGTLFGSMAYPGKEDVLRHAGILPVQFARVPCEYPGRNIAFHVEEGSNPNYLAILVEYEDGDGDLRSVDIMESSSSAQSTYSNAWTPMRESWGAIWRLDTSRPLRGPFSIRLTTDTSDRTLVARNVIPANWQPNQVYRSIVNYST	279
	AcEXPB3a	Aco012813.1	XP_020083725.1		MALPLHKQASSIVYYIALLALISLFNSCSCSEYKRLNYSDAPSSWGPASVTWYGAPTGAGPDDNGGGCGFKNVNLPPFSSMTACGNPTIFKSGKGCGACYQVKCTSHPACSGKPSTITITDACLDGVCLDAPVHFDMSGTSFGSMAQPGREDELRHAGKLQIQYTRVPCNYPGLNIAFHVEEGSNPVYFAILIEYEDGDGDLASVDLMEGAGPTAGTWTPMRESWGAIWRLDSDHALQGPFSIRLSTLTSGKTLVASNVIPADWQPLSTYRSIVNYSN	278
Brachypodiu m disachyon	BdEXPB1a	XP_010229338.1	Bradi1g78120	1	MASSSSSTLLLAASILATLASSAHGIPKVPPGPNITATYGDKWLDAKSTWYGKPTGAGPKDNGGACGYKDVDKEPFSGMTSCGNTPIFRDGRGCGSCFEIKCTKPDACSGDPVLVHITDDNEEPIAAYHFDLSGHAFGSMAKKGKEQDLRSAGEVEIQFRRVKCKYPEGTKVTFHVEKGSSPNYLAILVKYVGGDGDVVAVDVKEKGKDEWVALKESWGAVWRLDTAKPLKGPLTVRYTTDGGTKGESEDVIPEDWKPDTMYESK*	266

	BdEXPB1b		Bradi4g00360	4	MAASSSQVMLAMALLAALLSLAHGIPKVPPGPNITATY NGKWLDAKSTWYGRPEGAGPKDNGGACGYKDVDKP PFNGMTSCGNTPIFRDGRGCGSCFEIKCDKPAEFCSG QPVLVHITDDNEEPIAAYHFDLSGKAFGSMAKKGQEQ KLRGCGEVEIQFRRVKCYYPLGTKVTYHVEKGSNPNY LALLVKFVGGDGDVVAVEVQEKGKYNWIPLKESWGAV WRIDTAKPLKGPLSVRYTTDGGTKAVSPDVIPEKWKP DTMYVAKY	266	64
	BdEXPB9	XP_01022934 7.1|	Bradi1g78340	1	MASSSSSSVSAALLLCLLLAFHGVSCAAKAKHGSGSK KTHHSPPKPHAPSPLPAPPATIIPPPVATNGSSNSSSP GAGDEGWMDARATWYGAPNGAGPDDNGGACGFKD VNLPPFSAMTSCGNEPLFKDGKGCGSCYQIRCLSRM HPACSGVPETVIITDMNYYPVSRFHFDLSGTAFGAMAK DQRNDELRHAGIIDMQFKRVACQYPGLTVTFHVEHGS NPYYMAILVEYENGDGDVDQVDIMESTPDGSGEPTGQ WVPMKESWGSIWRMDTRRPMHGPFSLRITNESGQTL VADQVIPADWEPNAIYSSIIQFD*	316	
	BdEXPB4b	XP_00355903 0.1	Bradi1g78350	1	MAKHLSLLLVAAPPLLMLLLFSSVYGSSAGGGPNLNAS AVSFGQSGVARATWYGAPNGAGPYDNGGACGFKNV NKYPFMAMTSCGNQPLFKDGKGCGACYKIKCTKHKA CSGRTETVVITDMNYYPVAPYHFDLSGTAFGKLAKPG RNDELRHAGIIDIQFTRVPCEFPGLKVGFHVEEGSNAV YMAILVEYENGDGDVVQVDLMESGRGGGRWTRMKE SWGSIWRLDSNHRLQAPFSLRIRNESGKTLVARNVIPS NWRPNTFYRSIVQYS	272	
Elaeis guineensis **(Palm)**	EgEXPB5		XP_019705836 .1		MATSLLATFLSLSCLFTLSFCFQTKRFDLPMAGRDWLP AGATWYGSPRGAGSEGGACGYGNAVDQAPYSAMIS AGSTTIFNKGKGCGSCYQVMCTTHKACSGKPVTVVIA DYSCPGTVCQRESAHFDFSGTAFGAMAKPGMADKLR SAGYIYIQYSRVACEYPGRTVTFHVEHGSNPNYFAVVV EFEDGDGDLSAVEIKEGSGEWRAMRESWGAVWRLD TATPVHGPFSIRLTTQYSGQKLVANSVIPEGWQAGATY RSFVNYS	264	
	EgEXPB18		XP_008803824 .1		MATSLLATFLSLSCLFTLSFCFHTKHFALPMAGHDWSP AGATWYGSAHGAGSDGGACGYRNAVDQAPFSAMIAA GSPSLFKSGKGCGACYQVMCTTHEACSRKPVTVVITD ECPGGVCLKESAHFDLSGTAFGAMAKPGMADELRNA GILPIQYSRVACEYPGKTITFHVDSGSNPYYMAVVVEF EDGDGDLSAVELKDGSAEWRAMAQSWGAVWRLD SGSTLHGPLSIQLTSQYSGKKLVANNVIPAGWKPGATY RSFVNYS	264	

***Iordeum vulgare* (Barley)**	HvEXPB5a	Hor v 1	AK368047	BAJ99250.1	MAGVSTNAIALVALLSVLFTSVRSAVNYDTAMARSYNS GWLPAKATWYGAPNGPGPDDNGGACGFKNVNQYPI SSMGACGNEPIYADGEGCGMCYEIKCEYNNNPSCSG QPRRIVIIDMNYYPVARYHLDLSGTAFGSMARYGLNEQ LRHAGIIDMQFRRVRCNFPGMKITFHVQRGSNPNYLA VLVEYLNVDGTVVRMELMQNMNGRPTGYWQDMRRS WGSVWRMDTNRPLQGPFSIRITTDTGKMLVANNAIPA YWQQAKAYWSNIQFY	271	68
	HvEXPB5b		AK375220	BAK06415.1	MASSSALLVAAVLAVVVCGAHGIAKVPCGPNITATYVS EWKDAKTTWYGKPTGAGPKDNGGACGYKEVDKAPF FGMTSCGNVPIFKDGRGCGSCFEIKCTKPEACSGEPTI VTITDKNEEPIAAYHFDLSGHAFGTMAKKGQEQKLRDA GEVEIKFRRVKCKYPEGTKVNFHVEKGSSPNYLALVIK FLQGDGDVVAVDIKPKGKDKWIELKESWGAVWRIDTP DKLIGPFSVRYTTEGGTKTVAEDVIPKGWKPDTSYETK	262	
***Lolium perenne* (ryegrass)**	LpEXPB4a	Lol p 1	M57474		MASSSSVLLVVALFAVFLGSAHGIAKVPPGPNITAEYG DKWLDAKSTWYGKPTGAGPKDNGGACGYKNVDKAP FNGMTGCGNTPIFKDGRGCGSCFEIKCTKPESCSGEA VTVTITDDNEEPIAPYHFDLSGHAFGSMAKKGEEQNVR SAGELELQFRRVKCKYPDDTKPTFHVEKASNPNYLAIL VKYVDGDGDVVAVDIKEKGKDKWTELKESWGAVWRI DTPDKLTGPFTVRYTTEGGTKSEFEDVIPEGWKADTS YSA	263	
	LpEXPB4b		M57476		MASSSSVLLVVALFAVFLGSAHGIAKVPPGPNITAEYG DKWLDAKSTWYGKPTGAGPKDNGGACGYKNVDKAP FNGMTGCGNTPIFKDGRGCGSCFEIKCTKPESCSGEA VTVTITDDNEEPIAPYHFDLSGHAFGSMAKKGEEQNVR SAGELELQFRRVKCKYPDDTKPTFHVEKASNPNYLAIL VKYVDGDGDVVAVDIKEKGKDKWIELKESWGAVWRID TPDKLTGPFTVRYTTEGGTKSEFEDVIPEGWKADTSY SAK	263	52
	LpEXPB4c		X57678		MASSSSVLLVVALFAVFLGSAHGIAKVPPGPNITAEYG DKWLDAKSTWYGKPTGAGPKDNGGACGYKDVDKAP FNGMTGCGNTPIFKDGRGCGSCFEIKCTKPESCSGEA VTVTITDDNEEPIAPYHFDLSGHAFGSMAKKGEEQKLR SAGELELQFRRVKCKYPDGTKPTFHVEKASNPNYLAIL VKYVDGDGDVVAVDIKEKGKDKWIELKESWGAVWRID TPDKLTGPFTVRYTTEGGTKSEVEDVIPEGWKADTSY SAK	263	69

***Panicum hallii* (Panic grass)**	PhEXPB11a	Pahal.C04489	3	MAAASTHLVAVAVVLAALVGGAWCGPPKVPPGKNISA DCDGKWLEAKATWYGKPTGAGPDDNGGACGYKEVN KAPFNGMGACGNSPIFKDGLGCGSCYEIKCDKPAECS GEPVIVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEEE KLRKAGIIDMQFRRVKCKYPADTKIAFHVQKGCNPNYL ALLVKYAAGDGDIVGVDIKEKGAKEYQSLKHSWGAIW RMDAPKPIKGPISIRITSEGGKTLEQEDVIPEGWKPDTL YPSKLQF*	268	71
	PhEXPB11b	Pahal.C04487	3	MAAASTHLVAVAVVLAALVGGAWCGPPKVPPGKNISA DCDGKWLEAKATWYGKPTGAGPDDNGGACGYKEVN KAPFNGMGACGNSPIFKDGLGCGSCYEIKCDKPAECS GEPIIVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEEEK LRKAGIIDMQFRRVKCKYPADTKIAFHVEKGCNPNYLA LLVKYAAGDGDIVGVDMKEKGAKEYQSLKHSWGAIW RMEAPKPIKGPISIRITSEGGKTLEQEDVIPEGWKPDTL YPSKLQF*	268	
	PhEXPB11c	Pahal.I00040	9	MAAASTHLVAVAVVLAALVGGAWCGPPKVPPGKNISA DCDGKWLEAKATWYGKPTGAGPDDNGGACGYKEVN KAPFNSMGACGNSPIFKDGLGCGSCYEIKCDKPAECS GEPVIVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEEE KLRKVGIIDMQFRRVKCKYPADTKIAFHIEKGCNPNYLA LLVKYAAGDGDIVGVDIKEKGAKEYQSLKHSWGAIWR MDTPKPIKGPISIRITSEGGKTLEQEDVIPEGWKPDTLY PSKLQF*	268	
	PhEXPB11d	Pahal.I00042	9	MAAASTHLVAVAVVLAALVGGAWCGPPKVPLGKNISA DCDGKWLEAKATWYGKPTGAGPDDNGGACGYKEVN NAPFNSMGACGNSPIFKDGLGCGSCYEIKCDKPAECS GEPVIVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEEE KLRKAGIIDMQFRRVKCKYPADTKIAFHVEKGCNPNYL ALLVKYAAGDGDIVDVDIKEKGAKEYQSLKHSWGAIW RMDTPKPIKGPISIRITSEGGKTLEQEDVIPEGWKPDTL YPSKLQF*	268	
	PhEXPB11e	Pahal.J00957	9	MAAASTHLVAVAVVLAALVGGAWCGPPKVPPGKNISA DCDGKWLEAKATWYGKPTGAGPDDNGGACGYKEVN KAPFNSMGACGNSPIFKDGLGCGSCYEIKCDKPAECS GEPVIVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEEE KLRKAGIIDMQFRRVKCKYPADTKIAFHVEKGCNPNYL ALLLKYAAGDGDIVGVDIKEKGAKEYQSLKHSWGAIWR MDTPKPIKGPISIRITSEGGKTLEQEDVIPEGWKPDTLY PSKLQF*	268	

	PhEXPB11f		Pahal.I00041	9	MAAASTHLVAVAVVLAALVGGAWCGPPKVPPGKNISA DCDGKWLEAKATWYGKPTGAGPDDNGGACGYKEVN KAPFNSMGACGNSPIFKDGLGCGSCYEIKCDKPAECS GEPVIVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEEE KLRKAGIIDMQFRRVKCKYPADTKIAFHVEKGCNPNYL ALLLKYAAGDGDIVGVDIKEKGAKEYQSLKHSWGAIWR MDTPKPIKGPISIRITSEGGKTLEQEDVIPEGWKPDTLY PSKLQF*	268	
***Phleum pratense* (Timothy grass)**	PpEXPB4a		P43213		MASSSSVLLVVVLFAVFLGSAYGIPKVPPGPNITATYGD KWLDAKSTWYGKPTGAGPKDNGGACGYKDVDKPPF SGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPV VVHITDDNEEPIAPYHFDLSGHAFGAMAKKGDEQKLR SAGELELQFRRVKCKYPEGTKVTFHVEKGSNPNYLAL LVKYVNGDGDVVAVDIKEKGKDKWIELKESWGAIWRI DTPDKLTGPFTVRYTTEGGTKTEAEDVIPEGWKADTS YESK	263	
	PpEXPB4b	Phl p 1	Q40967		MASSSSVLLVVALFAVFLGSAHGIPKVPPGPNITATYG DKWLDAKSTWYGKPTAAGPKDNGGACGYKDVDKPP FSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEP VVVHITDDNEEPIAAYHFDLSGIAFGSMAKKGDEQKLR SAGEVEIQFRRVKCKYPEGTKVTFHVEKGSNPNYLALL VKFVAGDGDVVAVDIKEKGKDKWIALKESWGAIWRIDT PEVLKGPFTVRYTTEGGTKGEAKDVIPEGWKADTAYE SK	263	57
***orghum bicolor* (Sorghum)**	SbEXPB11a		Sb01g050180 Sb01G540000	1 1	MGVNMMSWSMKVALVVALVFLVGGAWCGPPKVAPG KNITATYGSDWLEAKATWYGKPTGAGPDDNGGACGY KDVNKAPFNSMGACGNLPIFKDGLGCGSCFEIKCDKP AECSGEAVVVHITDMNYEQIAAYHFDLAGTAFGAMAK KGEEEKLRKAGIIDMQFRRVKCKYGEKVTFHVEKGSN PNYLALLVKYVDGDGDVVGVDIKEKGGEAYQPLKHSW GAIWRKDSDKPIKFPVTIQITTEGGTKTAYEDVIPEGWK ADTTYTAK*	266 266	67
	SbEXPB11b				MGVNMMSWSMQVVLVVALAFLVGGAWCGPPKVAPG		
	SbEXPB11c		Sb01g050230	1	MGVNMMSWSMQVAVVVALAFLVGGAWCGPPKVAPG KNITATYGSDWLEAKATWYGKPTGAGPDDNGGACGY KDVNKAPFNSMGACGNLPIFKDGLGCGSCFEIKCNKP AECSGEAVVVHITDMNYEQIAAYHFDLAGTAFGAMAK KGEEEKLRKAGIIDMQFRRVKCKYGEKVTFHVEKGSN PNYLALLVKYVDGDGDVVGVDIKEKGGDAYQSLKHSW GAVWRKDSDKPIKFPVTVQITTEGGTKTAYEDVIPEGW KADTTYTAK*	266	

	SbEXPB11d	Sb01g050210	1	MGVNMMSWSMQVALVVALAFLVGGAWCGPPKVAPG KNITATYGSDWLEAKATWYGKPTGAGPDDNGGACGY KDVNKAPFNSMGACGNLPIFKDGLGCGSCFEIKCDKP AECSGEAVVVHITDMNYEQIAAYHFDLAGTAFGAMAK KGEEEKLRKAGIIDMKFRRVKCKYGEKVTFHVEKGSN PNYLALLVKYVDGDGDVVGVDIKEKGGDAYQPLKHSW GAIWRKDSDKPIKFPVTVQITTEGGTKTAYEDVIPEGW KADTTYTAK*	266	
	SbEXPB11e	Sb01g050220	1	MMSWSMQVALVVALAFLVGGAWCGPPKVAPGKNITA TYGSDWLEAKATWYGKPTGAGPDDNGGACGYKDVN KAPFNSMGACGNLPIFKDGLGCGSCFEIKCDKPAECS GEAVVVHITDMNYEQIAAYHFDLAGTAFGAMAKKGEE EKLRKAGIIDMKFRRVKCKYGEKVTFHVEKGSNPNYLA LLVKYVDGDGDVVGVDIKEKGGDAYQPLKHSWGAIW RKDSDKPIKFPVTVQITTEGGTKTAYEDVIPEGWKADT TYTAK*	262	
	SbEXPB11f	Sb01g050190	1	MGANMMSWSMQVALVVALAFLVGGAWCGPPKVAPG KNITATYGSDWLEAKATWYGKPTGAGPDDNGGACGY KDVNKAPFNSMGACGNLPIFKDGLGCGSCFEIKCDKP AECSGEAVVVHITDMNYEQIAAYHFDLAGTAFGAMAK KGEEEKLRKAGIIDMKFRRVKCKYGEKVTFHVEKGSN PNYLALLVKYVDGDGDVVGVDIKEKGGDAYQPLKHSW GAIWRKDSDKPIKFPVTVQITTEGGTKTAYEDVIPEGW KADTTYTAK*	266	
***Sorghum halepense* (johnson)**	ShEXPB10	KF887425		MGVNMMSWSMQVALVVALAFLVGGAWCGPPKVAPG KNITATYGSDWLEAKATWYGKPTGAGPDDNGGACGY KDVNKAPFNSMGACGNLPIFKDGLGCGSCFEIKCDKP AECSGEAVVVHITDMNYEQIAAYHFDLAGTAFGAMAK KGEEEKLRKAGIIDMKFRRVKCKYGEKVTFHVEKGSN PNYLALLVKYVDGDGDVVGVDIKEKGGDAYQPLKHSW GAIWRKDSDKPIKFPVTVQITTEGGTKTAYEDVIPEGW KADTTYTAK	266	
	ShEXPB5	KF887426		MGSLANKIVAMAAVLAALVTGGSCAPKKFPPGPNITTN YNGQWLSARATWYGQPNGAGPDDNGGACGIKNVNL PPYNGFTACGNVPIFKDGKGCGSCYEVRCKEMPECS GNPITVFITDMNYEPIAPYHFDFSGKAFGSLAKPGLNDK LRHCGIMNVEFRRVRCKLGGKIMFHVEKGSNPNYLAV LVKNVADDGNIVLMELEDKASPGFKPMKQSWGAVWR FDTPKPVKGPFSIRLTSESGKKLVAPNVIPATWKPDTL YNSNIQF	266	0

Triticum aestivum **(Wheat)**	TaEXPB5a	AY589582	MAGVSSNAIALVALLSVLFTGVRSAVNYDTAVARSYNS GWLPAKATWYGAPNGAGPDDNGGACGFKNVNQYPI SSMGACGNEPIFAGGEGCGMCYEIKCDYSNNPSCSG QPRRIVITDMNYYPVARYHLDLSGTAFGSMARYGLND RLRHAGIIDMQFRRVPCNFPGMKVTFHVQRGSNPNYL AVLVEYANVDGTVVRMELMQTINGRPTGYYQDMRRS WGSIWRMDTNRPLQGPFAIRITSDTGKTLVANNAIPAY WQPDHAYWSNLQFY	271	52
	TaEXPB8	AY543542	MAGVSTNAIALVLVTLLSVLVTSVRSAANYDTAAARSY NSGWLPAKATWYGAPTGAGPMDNGGACGFKNVNKY PFSSMTSCGNEPLFDGGAGCGSCYEIRCIAANNPSCS GQPRTVVITDMNYYPVARYHFELSGTAFGAMAKYGLN DKLRHAGIIDMQFRRVRCNFPGMKVTFHVQRGSNPNY LAVLVEYANIDGTVVRMELMQTRNGRPTGSWEPMRR SWGSIWRMDTSRPLQGPFSMRITSDSGKTLVANNVIP AYWRPDNAYWSNVQFY	273	68
	TaEXPB5b	U91981	MAGVSSNAIALVALLSMLFTSVRSAVNYDTAVARSYTS GWLPAKATWYGAPNGAGPDDNGGACGFKNVNQYPI SSMGACGNEPIFAGGEGCGMCYEIKCEYSNNPSCSG QPRRIVITDMNYYPVARYHLDLSGTAFGSMAKYGLND RLRHAGIIDMQFRRVPCNFPGMKVTFHVQRGSNPNYL AVLVEYVNVDGTVVRMELMQTMNGRPTGYWQDMRR SWGSVWRMDTNRPVQGPFSIRITSDTGKTLVANNAIP AYWQQDHAYWSNLQFY	271	
	TaEXPB10	AY543544	MAGVSSNAIAFVALLSVLFTGVRSAVNYDTAVARSYNS GWLPAKATWYGAPNGAGPDDNGGACGFKNVNQYPI SSMGACGNEPIFAGGEGCGMCYEIKCDYSNNPSCSG QPRRIVITDMNYYPVARYHLDLSGTAFGSMARYGLND RLRHAGIIDMQFRRVPCNFPGMKVTFHVQRGSNPNYL AVLVEYANVDGTVVRMELMQTINGRPTGYYQDMRRS WGSIWRMDTNRPLQGPFAIRITSDTGKTLVANNAIPAY WQPDHAYWSNLQFY	271	55
	TaEXPB4	AY543539	MASSSSSVLLVAAVLAAVVCGAHGIAKVPPGPNITASPTSYGNKWLDAKTTWYGKPTGAGPKDNGGACGYKEV DKAPFHGMTSCGNIPIFKDGRGCGSCFELKCTKPEAC SGEPTMVTITDKNEEPIAPYHFDLSGHAFGSMAKKGEE QKLRDAGEVEIKFRRVKCKYPAGTKVNFHVEKSSNENYLALVIKFLQGDGDVVGVDIKQKGEDKWTELNESWGA VWRIDTPHKLIGPFSVRYTTEGGTKTVVDDVIPKGWKP DTSYEAKGGY	270	

	TaEXPB1	AY533101			MGSVSYVLAAAVLAALVSGGACIPKVPPGPNITTNYNN QWLSAKATWYGRPTGSGPKDNGGACGIKDVNLAPYN GMIACGNVPIFKDGKGCGSCYEIKCQKPSPCSDKPITIF ITDKNYEPIAPYHIDLSGTAFGAMATPGKEQTLRSFGEL ELQFRRVRCKYAPGTKITFHVEKGSNPNYLAVLVKFVS DDGDVVQMDIQESKSPAWIPLTLSWGAIWRWDGAKP LKGPFSIRVTSESGKKLIAKDVIPANWKADTVYTSNVQF	265	
***ea mays* (Maize)**	ZmEXPB10a		GRMZM2G110 025	5	MRAQVAMAVVLVFLVSGAWCGPPKVPPGKNITATYG KDWLDAKATWYGKPTGAGPDDNGGGCGYKDVNKPP FNSMGACGNIPIFKDGLGCGSCFEIKCDKPVECSGKPV VVHITDMNYEPIAAYHFDLAGTAFGAMAKKGEEEKLRK AGIIDMQFRRVKCKYDSKVTFHLEKGCGPNYLALLVKY VDGDGDIVAVDVKEKGSDTYEPLKHSWGAIWRKDSD KPLKGPLTVRLTTEGGTKSVYDDVIPANWKANTAYTAK *	263	
	ZmEXPB10b		GRMZM2G144 856	5	MRAHVAMAVALVFLVSGAWCGPPKVPPGKNITATYGK DWLDAKATWYGKPTGAGPDDNGGGCGYKDVNKPPF NSMGACGNIPIFKDGLGCGSCFEIKCDKPVECSGKPV VVHITDMNYEPIAAYHFDLAGTAFGAMAKKGEEEKLRK AGIIDMQFRRVKCKYDSKVTFHLEKGCGPNYLALLVKY VDGDGDIVAVDVKEKGSDTYEPLKHSWGAIWRKDSD KPLKGPLTVRLTTEGGTKSVYDDVIPANWKANTAYTAK *	263	67
	ZmEXPB10c		GRMZM2G127 106	5	MRAQVAMVVALVFLVRGAWCGPPKVPPGKNITATYG KDWLDAKATWYGKPTGAGPDDNGGGCGYKDVNKPP FNSMGACGNIPIFKDGLGCGSCFEIKCDKPVECSGKPV VVHITDMNYEPIAAYHFDLAGTAFGAMAKKGEEEKLRK AGIIDMQFRRVKCKYDSKVTFHLEKGCGPNYLALLVKY VDGDGDIVAVDVKEKGSDTYEPLKHSWGAIWRKDSD KPLKGPLTVRLTTEGGTKSVYDDVIPANWKANTAYTAK *	263	
	ZmEXPB10d		GRMZM2G144 898	5	MRAQVAMEEALVFLVSGAWCGPPKVPPGKNITATYG KDWLDAKATWYGKPTGAGPDDNGGGCGYKDVNKPP FNSMGACGNIPIFKDGLGCGSCFEIKCDKPVECSGKPV VVHITDMNYEPIAAYHFDLAGTAFGAMAKKGEEEKLRK AGIIDMQFRRVKCKYDSKVTFHLEKGCGPNYLALLVKY VDGDGDIVAVDVKEKGSDTYEPLKHSWGAIWRKDSD KPLKGPLTVRLTTEGGTKSVYDDVIPANWKAQHRLHR QIINFSADNTLS*	270	

	ZmEXPB10e			GRMZM2G109973	5	MQAQVAMVVALSFLVSGAWCGPPKVPPGKNITATYG KDWLDAKATWYGKPTGAGPDDNGGGCGYKDVNKPP FNSMGACGNIPIFKDGLGCGSCFEIKCDKPVECSGKPV VVHITDMNYEPIAAYHFDLAGTAFGAMAKKGEEEKLRK AGIIDMQFRRVKCKYGSKVTFHLEKGCGPNYLALLVKY VDGDGDIVAVDIKEKGSDTYEPLKHSWGAIWRKDSDK PLKGPLTVRLTTEGGTKTVYDDVIPANWKANTAYTAK*	263	68
	ZmEXPB11a			GRMZM2G089699	9	MWSLVQVQVLVAVALAFLVGGAWCGPPKVPPGKNIT AKYGSDWLDAKATWYGKPTGAGPDDNGGGCGYKDV NKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAEC SGKPVVVVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEE EKLRKAGIIDMQFRRVKCKYGSKVTFHLEKGCNPNYLA LLVKYVDGDGDIVAVDIKEKGSDTYEPLKHSWGAIWRK DSDKPIKGPITVQLTTEGGTKTVYDDVIPAGWKPNTAY TAK*	270	68
	ZmEXPB11b		XM_008675755	GRMZM2G007685	3	MWSLVQVQVLVAVALSFLVGGAWCGPPKVPPGKNIT ANYGSDWLDAKATWYGKPTGAGPDDNGGGCGYKDV NKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAEC SGKPVVVVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEE EKLRKAGIIDMQFRRVKCKYGSKVTFHLEKGCNPNYLA LLVKYVDGDGDIVAVDIKEKGSDTYEPLKHSWGAIWRK DSDKPIKGPITVRLTTEGGTKTVYDDVIPTDWKPNTAYT TK*	263	67
	ZmEXPB11c	Zea m 1		GRMZM2G164785	3	MWSLVQVQVLAAMALAFLVGGAWCGPPKVPPGKNIT ANYGSDWLDAKATWYGKPTGAGPDDNGGGCGYKDV NKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAEC SGKPVVVVYITDMNYEPIAAYHFDLAGTAFGAMAKKGEE EKLRKAGIIDMQFRRVKCKYGSKVTFHLEKGCNPNYLA LLVKYVDGDGDIVAVDIKEKGSDTYEPLKHSWGAIWRK DSDKPIKGPITVRLTTEGGTKTVYDDVIPAEWKPNTAYT T*	263	
	ZmEXPB9a			GRMZM2G072886	9	MGSLANNIMVVGAVLAALVVGGSCGPPKVPPGPNITT NYNGKWLTARATWYGQPNGAGAPDNGGACGIKNVN LPPYSGMTACGNVPIFKDGKGCGSCYEVRCKEKPECS GNPVTVFITDMNYEPIAPYHFDLSGKAFGSLAKPGLND KLRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGCNPNYL AVLVKFVADDGDIVLMEIQDKLSAEWKPMKLSWGAIW RMDTAKALKGPFSIRLTSESGKKVIAKDIIPANWRPDAV YTSNVQFY*	269	

	ZmEXPB9b		GRMZM2G146551	9	MGSLVNNIMVVGAVLAALVAGGSCGPPKVPPGPNITTNYNGKWLTARATWYGQPNGAGAPDNGGACGIKNVNLPPYSGMTACGNVPIFKDGKGCGSCYEVRCKEKPECSGNPVTVYITDMNYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGCNPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPFSIRLTSESGKKVIAKDVIPANWRPDAVYTSNVQF	269	
	ZmEXPB1a		GRMZM2G181202	9	MGSLANNIMVVGAVLAALVAGGSCGPPKVPPGPNITTNYNGKWLTARATWYGQPNGAGAPDNGGACGIKNVNLPPYSGMTACGNVPIFKDGKGCGSCYEVRCKEKPECSGNPVTVYITDMNYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGCNPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPFSIRLTSESGKKVIAKDVIPPNWRPDAVYTSNVQFY*	269	
	ZmEXPB1b		GRMZM2G020852	9	MGSLANNIMVVGVVLAALVAGGSCGPPKVPPGPNITTNYNGKWLTARATWYGQPNGAGAPDNGGACGIKNVNLPPYSGMTACGNVPIFKDGKGCGSCYEVRCKEKPECSGNPVTVYITDMNYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGCNPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPFSIRLTSESGKKVIAKDVIPPNWRPDAVYTSNVQFY*	269	

Supplementary Table 2 – PRFA homologs

Species tree	Name	on	Chroms	Phytozome	Uniprot accession	Genbank accession	Protein	Length	Proteimtity to Os 'OF1Aa (%)
***Oryza sativa* (Rice)**	OsPRFAa	Ory s 12	10	Os10g17660	AC074354	XP_015613779.1	MSWQTYVDEHLMCEIEGHHLTSAAIVGHDGTVWAQSAAFPQFKPEEMTNIMKDFDEPGFLAPTGLFLGPTKYMVIQGEPGAVIRGKKGSGGITVKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGL	131	
	OsPRFAb		10	Os10g17680	DP000086	XP_015613779.1	MSWQTYVDEHLMCEIEGHHLTSAAIVGHDGTVWAQSAAFPQFKPEEMTNIMKDFDEPGFLAPTGLFLGPTKYMVIQGEPGAVIRGKKGSGGITVKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGL	131	100
	OsPRFc	OsPRFLP04	6	Os06g05880	Q5VMJ3	XP_015642483.1	MSWQAYVDDHLMCEIDGNHLTAAAIVGHDGSVWAQSPNFPQYKPEEITGIMKDFDEPGSLAPTGLFLGGTKYMVIQGEPGVVIRGKKGTGGICVKKTGLSLILGIYDEPMTPGQCNMIVERLGDYLIEQGC	131	83
***Ananas comosus* (Pineapple)**	AcPRF	Ana c 1	10	Aco009958.1	Q94JN2	AAK54835.1	MSWQAYVDDHLMCEIDGQHLSSAAILGHDSTVWAQSPNFPQFKPEEISAILNDFENPGSLAPTGLYLGGTKYMVIQGEPGVVIRGKKGTGGITVKKTNLALIIGVYDEPMTPGQCNMVVERLGDYLLEQGF	131	79
***Arabidopsis thaliana*(Arabidopsis**	AtPRF1		2	At2g19760	Q42449		MSWQSYVDDHLMCDVEGNHLTAAAILGQDGSVWAQSAKFPQLKPQEIDGIKKDFEEPGFLAPTGLFLGGEKYMVIQGEQGAVIRGKKGPGGVTIKKTNQALVFGFYDEPMTGGQCNLVVERLGDYLIESEL*	131	76
	AtPRF2		4	AT4G29350	Q42418	AAL67046.1	MSWQSYVDDHLMCEVEGNHLTHAAIFGQDGSVWAQSSAFPQLKPAEIAGINKDFEEAGHLAPTGLFLGGEKYMVVQGEAGAVIRGKKGPGGVTIKKTTQALVFGIYDEPMTGGQCNLVVERLGDYLIESGL	131	76
	AtPRF3		5	AT5G56600	Q9FE63	NP_200471	MSWQTYVDDHLMCDVAGNRLTAAAILGQDGSVWAQSNNFPQVKPEEIQGIKDDFTTPGTLAPTGLFLGGNKYMVIQGEPNAVIRGKKGAGGVTIKKTTLALVFGIYDEPMTPGQCNMVVENLGEYLIESGL	131	74

	AtPRF4		4	AT4G29340	Q38904		MSWQTYVDEHLMCDVGDGQGHHLTAAAIVGHDGS VWAQSANFPQFKGQEFSDIMKDFDEPGHLAPTGLF MAGAKYMVIQGEPGAVIRGKKGAGGITIKKTGQSCV FGIYEEPVTPGQCNMVVERLGDYLLEQGL*	134	80
	AtPRF5		2	AT2G19770	Q38905	AAM62866.1	MSWQAYVDEHLMCDVGDGQGHHLTAAAIIGHDGS VWAQSANFPQFKPQEITDIMKDFDEPGHLAPTGMF LAGLKYMVIQGEPNAVIRGKKGAGGITIKKTGQSMV FGLYEEPVTPGQCNMVVERLGDYLIEQGL	134	78
***Arachis hypogaea*(Peanut)**	AhPRF1a	Ara h 5			Q9SQI9	AGA84056.1	MSWQTYVDNHLLCEIEGDHLSSAAILGQDGGVWA QSSHFPQFKPEEITAIMNDFAEPGSLAPTGLYLGGT KYMVIQGEPGAIIPGKKGPGGVTIEKTNQALIIGIYDK PMTPGQCNMIVERLGDYLIDTGL	131	79
	AhPRF1b	Ara h 5			D3K177	AAD55587	MSWQTYVDNHLLCEIEGNHLSSAAILGQDGSVWAQ SSNFPQFKPEEITAIMNDFAEPGSLAPTGLYLGGTK YMVIQGEPGAVIRGKKGPGGVTIKKTNQALIIGIYDE PMTPGQCNMIVERLGDYLIDTGL	131	80
	AhPRF1c	Ara h 5			Q5XXQ5	AY726606	MSWQAYVDEHLICDIEGNQLTSAAILGQDGSVWAQ SSNFPQFKPEEITAIMNDFAEPGTLAPTGLYLGGTK YMVIQGEPGAVIQXXKGPGGVTIKKTNQALIIGIYDEP MTPGQCNMVVERLGDYLIE	128	79
Betula Verrosca	BvPRFa	Bet v 2			P25816	AAA16522.1.	MSWQTYVDEHLMCDIDGQASNSLASAIVGHDGSV WAQSSSFPQFKPQEITGIMKDFEEPGHLAPTGLHL GGIKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVF GIYEEPVTPGQCNMVVERLGDYLIDQGL	133	80
	BvPRFb				A4K9Z8	ABG48509.1.	MSWQTYVDEHLMCDIDGQGQQLAASAIVGHDGSV WAQSSSFPQFKPQEITGIMKDFEEPGHLAPTGLHL GGIKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVF GIYEEPVTPGQCNMVVERLGDYLIDQGL	133	80
***achypo dium distachyo n*(Brachyp**	BdPRF2a		2	Bradi2g19360	KQK05308	XP_00356807 1	MSWQTYVDEHLMCEIEGHHLGSAAILGHDGTVWA QSAAFPAFEPKEMTDIMKDFDEPGHLAPTGMFLGG AKYMVIAGEPGAVIRGKKGSGGITIKKTGQALVIGIY DEPMTPGQCNMVVERLGDYLVEQGM	131	89

	BdPRF2b		2	Bradi2g49340	KQK09648	XP_00356976 5	MSWQTYVDEHLMCDIEGHHLVSAAILGHDGTVWAQSADFPSFKPEEMTNIMKDFDEPGTLAPTGLFLASAKYMVIQGEPGAVIRGKKGSGGITLKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGM	131	92
	BdPRF1b		1	Bradi1g49270	KQK19606	XP_00356432 7	MSWQAYVDEHLMCEIEGNHLTAAAILGQDGSVWSQSDNFPQFKPEQITAIMKDFDEPGTLAPTGLFFGSEKYMVIQGEPGVVIRGKKGPGGITIKKTGQALLVGIYDEPMTPGQCNMIVERLGDYLVEQGM	131	83
	BdPRF1a		2	Bradi2g05690	KQK03116.1.	XP_00356468 2	MSWQAYVDEHLLCDIDGQRLTAAAILGHDGSVWAQSESFPQVKPEEVTAVMNDFNEPGSLAPTGLYLAGTKYMVIQGEPGAVIRGKKGPGGVTIKKTTLAIIIGIYEEPMTPGQCNMVVERLGDYLLEQGF	131	76
***Corylus avellana* (Hazelnut)**	CaPRF2	Cor a 2			Q9AXH5	AAK01236	MSWQAYVDEHLMCDIDGQGQQLAASAIVGHDGSVWAQSSSFPQLKPEEITGIMKDFDEPGHLAPTGLHLGGTKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVFGIYEEPVTPGQCNMVVERLGDYLAEQGL	133	81
***Cynodon dactylon*(Bermuda Grass)**	CdPRF2				O04725	CAA69669.1	MSWQAYVDDHLMCEIEGHHLTSAAIIGHDGTVWAQSAAFPAFKPEEMANIMKDFDEPGFLAPTGLFLGPTKYMVIQGEPGAVIRGKKGSGGVTVKKTGQALVIGIYDEPMTPGQCNMVIEKLGDYLIEQGM	131	92
	CdPRF1	Cyn d 12			O04725	CAA69670.1	MSWQAYVDDHLMCEIEGHHLTSAAIIGHDGTVWAQSAAFPAFKPEEMANIMKDFDEPGFLAPTGLFLGPTKYMVIQGEPGAVIRGKKGSGGVTVKKTGQALVIGIYDEPMTPGQCNMVIEKLGDYLIEQGM	131	92
elianthus annuus	HaPRF	Hel a 2	3		O81982	Ha3_0003445 9	MSWQAYVDEHLMCDIEGTGQHLTSAAILGLDGTVWAQSAKFPQFKPEEMKGIIKEFDEAGTLAPTGMFIAGAKYMVLQGEPGAVIRGKKGAGGICIKKTGQAMIMGIYDEPVAPGQCNMVVERLGDYLLEQGM	133	77
***ordeum vulgare* (Barley)**	HvPRFa	Hor v 12	7	HORVU7Hr1G	P52184 022570	AAA92503.1.	MSWQTYVDDHLCCEIDGQHLTSAAILGHDGRVWVQSPNFPQFKPEEIAGIIKDFDEPGHLAPTGLFLGGTKYMVIQGEPGVVIRGKKGTGGITIKKTGMPLILGIYDEPMTPGQCNLVVERLGDYLVEQGF	131	82

Species	Profilin	Alias		UniProt	GenBank	Sequence	Length	Identity
	HvPRFb			F2CT70		MSWQTYVDEHLMCDIEGHHLASAAILGHDGTVWAQSADFPSFKPEEMTNIMKDFDEPGTLAPTGLLLGSAKYMVIQGEPGAVIRGKKGSGGITLKKTGQALVIGIYDEPMTPGQCNMVVERLGDYLVEQGM	131	91
	HvPRFd		1	F2EK80	AK376557.1	MSWQTYVDEHLMCDIEGHHLASAAILGHDGTVWAQSADFPQFGPNEITGIMKDFDEPGYLAPTGMFIATAKYMVIQGEPGAVIRGKKGAGGITIKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGM	131	87
	HvPRFc			F2EGC6		MSWQTYVDEQLLCDIDGQRLAAAAILGHDGAVWAQSEPFPEVKPEEITAVINDFDEPGSLAPTGLFLGGTKYMVIQGEPGAVIRGKKGSGGVTIKKTSLAIIGIYEEPMTPGQCNMVVERLGDYLLEQG	131	77

Species	Profilin	Alias		UniProt	GenBank	Sequence	Length	Identity
***Olea europaea* (Olive)**	OePRF2			O24170	Y12429	MSWQAYVDDHLMCDIEGHEGHRLTAAAIVGHDGSVWAQSATFPQFKPEEMNGIMTDFNEPGHLAPTGLHLGGTKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVFGIYEEPVTPGQCNMVVERLGDYLLEQGL	134	83
	OePRF3		O24171		Y12430	MSWQAYVDDHLMCDIEGHEGHRLTAAAIVGHDGSVWAQSATFPQFKPEEMNGIMTDFNEPGHLAPTGLHLGGTKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVFGIYEEPVTPGQCNMVAERLGDYLLEQGL	134	82
	OePRF1	PRFilin 1		CAA73035.1		MSWQAYVDDHLMCDIEGHEDHRLTAAAIVGHDGSVWAQSATFPQFKPEEMNGIMTDFNEPGHLAPTGLHLGGTKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVFGIYEEPVTPGQCNMVVERLGDYLVEQGM	134	83
	OePRF4		P0DKE4		AAZ30395	MSWQAYVDDHLMCDIEGHEDHRLTAAAIVGHDGSVWAQSATFPQFKPEEMNGIMTDFNEPGHLAPTGLHLGGTKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVFGIYEEPVTPGQCNMVVERLGDYLLEQGL	134	83
	OePRF5				AAZ30396.1	MSWQTYVDDHLMCDIEGHEDHRLTAAAIVGHDGSVWAQSATFPQFRPEEMNGIMTDFNEPGHLAPTGLHLGGTKYMVIQGEAGAVIRGKKGSGGITIKKTGQALVFGIYEEPVTPGQCNMVVERLGDYLLEQGL	134	83

nicum hallii ell's Panic Grass)	PhPRFA	5	Pahal.E01712	A0A2T8IK63	XP_025817063	MSWQTYVDEHLMCETEGHHLTCAAIIGHDGTVWAQSAAFPSFKPEEMSNIMKDFDEPGFLAPTGLFLGPTKYMVIQGEPGAVIRGKKGSGGITVKKTGQALVIGIYEEPMTPGQCNMVVERLGDYLVEQGL	131	95
	PhPRFA-like	3	Pahal.C02260	A0A2S3HAD5	XP_025807755.1	MSWQTYVDEHLMCESEGHHLTSAAIVGHDGTVWAQSAAFPQFKPEEMANIMKDFDEPGHLAPTGLFLGPTKYMVIQGEPGAVIRGKKGSGGITVKKTGQALIIGIYDEPMTPGQCNLVVERLGDYLIEQGM	131	94
	PhPRF4	4	Pahal.J00160	K3XZY5	XP_004964829	MSWQTYVDEHLMCEIEGQQLTSAAIVGHDGSIWAQSPNFPQYKPEEIAAIMKDFDDPGTLAPTGLFLGGTKYMVIQGEPGVVIRGKKGTGGITVKKTTLALIIGIYDEPMTPGQCNMIVERLGDYLVEQGF	131	84
	PhPRF1		Pahal.E04285	A0A2S3HYJ1	XP_025819452.1	MSWQAYVDDHLLCDIDGQRLSAAAILGHDGAVWAQSDAFPQVKPEEITAIMNDFNEPGSLAPTGLYLGGSKYMVIQGEPGVVIRGKKGPGGITIKKTNLAIIIGIYEEPMAPGQCNMVVERLGDYLVDQGF	131	75

***Phleum pratense*(Timothy Grass)**	PpPRF1	P35079.1	CAA54686.1.	MSWQTYVDEHLMCEIEGHHLASAAILGHDGTVWAQSADFPQFKPEEITGIMKDFDEPGHLAPTGMFVAGAKYMVIQGEPGRVIRGKKGAGGITIKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGM	131	89
	PpPRF2	O24650.1	CAA70608.1.	MSWQTYVDEHLMCEIEGHHLASAAILGHDGTVWAQSADFPQFKPEEITGIMKDFDEPGHLAPTGMFVAGAKYMVIQGEPGAVIRGKKGAGGITIKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGM	131	89
	PpPRF4	O24650	CAA70610	MSWQTYVDEHLMCEIEGHHLASAAILGHDGTVWAQSADFPQFKPEEITGIMKDFDEPGHLAPTGMFVAGAKYMVIQGEPGAVIRGKKGAGGITIKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGM	131	89
	PpPRF3	O24282.1	CAA70609.1.	MSWQTYVDEHLMCEIEGHHLASAAIFGHDGTVWAQSADFPQFKPEEITGIMKDLDEPGHLAPTGMFVAAAKYMVIQGEPGAVIRGKKGAGGITIKKTGQALVVGIYDEPMTPGQCNMVVERLGDYLVEQGM	131	89

***Physcomitrella patens* (Moss)**	PHpPRFLP04	PRFilin LP04		Pp3c18_17930 V3.2	A9TUU0	XM_001782313	MSWQSYIDDHLMYEISPGHSLAAAAIIGHNGSVWA QSENFPQLSPEEVDKLLNGFEENSPLAQNGLFLGG SKYMVLQGDPGIVIRGKKGPGGCTIRKTNSAFVIGIY DEPCTPGECNIAVEKLGEYLFEQGL	132	63
	PHpPRF2-like	PRFilin 2-like		Pp3c18_15960 V3	A9RDI7	XM_001752228	MSWQSYIDDHLMYEISEGHSLTSAAIVGHDGSVWA QSSSFPQLSPVEVEKLLDGFEENSSLPSNGLFLGG AKYMVLQGDPGIVIRGKKGPGGCTIRKTLSAFVIGLY DEPCTPGECNIAVEKLGEYLYEQGI	132	64
***Prunus avium* (Cherry)**	PaPRF	Pru a 2			Q9XF39	AF129425	MSWQAYVDDHLMCDIDGNRLTAAAILGQDGSVWS QSATFPAFKPEEIAAILKDLDQPGTLAPTGLFLGGTK YMVIQGEAGAVIRGKKGSGGITVKKTNQALIIGIYDE PLTPGQCNMIVERLGDYLIEQGL	131	79
***Solanum lycopersicum* (Tomato)**	SlPRF1		12	Solyc12g0446	Q41344 30	NP_00123386 9	MSWQTYVDDHLMCDIEGTGHHLSSAAILGFDGSV WAQSPNFPKFKAEEITNIMKDFDEPGHLAPTGLFLA GTKYMVIQGEPGAVIRGKKGPGGITIKKTAQALIFGV YEEPVTPGQCNMVVEKIGDYLVDQGY	133	80
	SlPRFa		11	Solyc11g0701 30	Q8VWR0	AAL29690.1	MSWQTYVDDHLMCDIEGNHLTSAAIIGQDGSVWAQ SANFPQFKPEEITAIMNDFAEPGTLAPTGLHLGGTK YMVIQGEAGAVIRGKKGAGGITVKKTNQALIIGIYDE PMTPGQCNMIVERLGDYIIEQGL	131	83
	SlPRF2		8	Solyc08g0661 10	Q93YG7	CAD10377	MSWQTYVDEHLLCENEGNHLTSAAIIGQDGTVWAQ SANFPQFKPEEITGIMNDFAVPGTLAPTGLYLGGTK YMVIQGEPEAVIRGKKGPGGITIKKTNQALIIGIYDEP MTPGQCNMIVERLGDYLIEQSL	131	82

***Sorghum bicolour* (Sorghum)**	SbPRF3		3	Sb03g034110	C5XJ77	XP_00245633 2.	MSWQTYVDEHLMCEIEGHHLTSAAIIGHDGTVWAQ STAFPQFKPEEMTNIMKDFDEPGFLAPTGLFLGPTK YMVIQGEPGAVIRGKKGSGGITVKKTGQALVIGIYDE PMTPGQCNMVVERLGDYLVEQGL*	131	98
	SbPRFA		9	Sb09g026120	C5Z1D7	XP_00244140 8.1	MSWQTYVDEHLMCEIEGHHLTSAAIAGHDGAVWA QSATFPEFKPEDMTNIMKDFDEPGHLAPTGLFLGAT KYMVIQGEPGAVIRGKKGSGGITVKKTGQALIIGIYD EPMTPGQCNMVVERLGDYLVEQGM*	131	92
	SbPRF4		10	Sb10g003670	C5Z4B6	XP_00243649 3.1	MSWQTYVDDHLMCEIDGQHLSAAAIFGLDGSLWA QSTGFPQLKPEEVAAITKDFDEPGTLAPTGLFVGGT KYMVIQGEPGAVIRGKKGTGGITIKKTGMALIIGIYDE PMTPGQCNMVVERLGDYLVEQG	131	82
***Triticum aestivum* (Wheat)**	TaPRF1a				ADK35122.1.	ADK35122	MSWQTYVDDHLCCEIDGQHLTSAAILGHDGSVWA QSPNFPQFKPEEIAGIIKDFEEPGHLAPTGLFLGGTK YMVIQGEPGVVIRGKKGTGGITIKKTGMALILGIYDE PMTPGQCNLVVERLGDYLVEQGF	131	82
	TaPRF4	Tri a 12.0104			B6EF35	CAQ57979	MSWKAYVDDHLCCEIDGQHLTSAAILGHDGSVWA QSPNFPQFKPEEIAGIVKDFEEPGHLAPTGLFLGGT KYMVIQGEPGVVIRGKKGTGGITIKKTGMALILGIYD EPMTPGQCNLVVERLGDYLIDQGY	131	80
	TaPRF1b	Tri a 12.0101			P49232	CAA61943.2.	MSWQTYVDDHLCCEIDGQHLTSAAILGHDGSVWTE SPNFPKFKPEEIAGIVKDFEEPGHLAPTGLFLGGTKY MVIQGEPGVVIRGKKGTGGITIKKTGMALILGIYDEP MTPGQCNLVVERLGDYLIDQGY	131	79
	TaPRF3a	Tri a 12.0102	P49233			CAA61944.2.	MSWQAYVDDHLCCEIDGQHLTSAAILGHDGSVWA ESPNFPKFKPEEIAGIVKDFEEPGHLAPTGLFLGGTK YMVIQGEPGVVIRGKKGTGGITIKKTGMALILGIYDE PMTPGQCNLVVERLGDYLIDQGY	131	79

	TaPRF3b	Tri a 12.0103		P49234	CAA61945.2.	MSWKAYVDDHLCCEIDGQNLTSAAILGHDGSVWA QSPNFPQFKPEENAGIVKDFEEPGHLAPTGLFLGG TKYMVIQGEPGVVIRGKKGTGGITIKKTGMALILGIYD EPMTPGQCNLVVERLGDYLIDQGY	131	79

***Zea mays*(Maize)**	ZmPRF3	Zea m 12.0103	5	GRMZM5G876 P35083 285	CAA51720.1.	MSWQTYVDEHLMCEIEGHHLSSAAIVGHDGAVWA QSTAFPQFKPEEMTNIIKDFDEPGFLAPIGLFLGPTK YMVIQGEPGAVIRGKKGSGGITVKKTGQALVIGIYDE PMTPGQCNMVVERLGDYLVEQGL	131	95
	ZmPRF2	Zea m 12.0102	2	GRMZM2G109 P35082 842	CAA51719	MSWQAYVDEHLMCEIEGHHLAAAAIVGHDGAAWA QSTAFPEFKTEDMANIMKDFDEPGHLAPTGLFLGPT KYMVIQGEPGAVIRGKKGSGGITVKKTGQALVVGIY DEPMTPGQCNMVVERLGDYLLEQGM	131	90
	ZmPRF1	Zea m 12.0101	2	GRMZM2G074 P35081 361	CAA51718	MSWQTYVDEHLMCEIEGHHLTSAAIVGHDGATWA QSTAFPEFKPEEMAAIMKDFDEPGHLAPTGLILGGT KYMVIQGEPGAVIRGKKGSGGITVKKTGQSLIIGIYD EPMTPGQCNLVVERLGDYLLEQGM	131	89
	ZmPRF4	Zea m 12.0104	6	GRMZM2G108 O22655 780	AAB86960.1.	MSWQAYVDEHLMCEIEGQHLSAAAIVGHDGSVWA QSESFPELKPEEVAGIIKDFDEPGTLAPTGLFVGGT KYMVIQGEPGVVIRGKKGTGGITIKKTGMSLIIGVYD EPMTPGQCNMVVERLGDYLIEQGF	131	80
	ZmPRF5	Zea m 12.0105	9	GRMZM5G877 Q9FR39 388	AAG35601.1.	MSWQAYVDDHLLCDIEGQHLSAAAIVGHDGSVWA QSENFPELKPEEVAGMIKDFDEPGTLAPTGLFVGG TKYMVIQGEPGVVIRGKKGTGGITIKKTGMSLIIGIYD EPMTPGQCNMVVERLGDYLIEQGF	131	78

Supplementary Table 3 – Transformation media

Stock solutions (1L)	Chemical	Grams/Litre	pH
N6-1 macro：20×	KNO3	56.6	
	(NH4)2SO4	9.26	
	KH2PO4	8	
N6-2 macro：40×	CaCl2·2H2O	6.64	
N6-3 macro：40×	MgSO4·7H2O	7.4	
B5 microI：100×			
	MnSO4·H2O	0.781	
	ZnSO4·7H2O	0.2	
	H3BO3	0.3	
	KI	0.075	
B5 microII：1000×	Na2MoO4·2H2O	0.125	
	CuSO4·5H2O	0.0125	
	CoCl2·6H2O	0.0125	
B5 Vitamin I：100×	thiamine hydrochloride	0.5	
	pyridoxine hydrochloride	0.05	
	nicotinic acid	0.05	
B5 vitamin II：100×		0.1	
	Glycine		
	Final volume		
2,4 D：100×2,4 D	2,4 D	0.2	
	Ethanol		
200×iron salt	Na2-EDTA	3.73	
	FeSO4·7H2O	2.78	
Shoot regeneration Media	Murashige and Skoog basal salts	4.4	5.8
	6-ba	0.002	
	Kt	0.0005	

	NAA	0.0005	
	Sucrose	30	
	Sorbitol	15	
	Phytogel	2	
Root regeneration media	Murashige and Skoog basal salts	2.2g	5.6
	Sucrose	30	

		surement per litre	
NBD2 Medium These solutions are made as previously shown	N6-1	50 ml	5.8
	N6-2	25 ml	
	N6-3	25 ml	
	B5 micro I	10 ml	
	B5 micro II	1ml	
	B5 vitamin I	10ml	
	B5 vitamin II	10ml	
	200×iron salt	7.5 ml	
	2-4D	15ml	
	sucrose	30g	
	myo-inositol	0.1g	
	L- proline	0.5g	
	L-glutamine	0.5g	
	casamino acid	0.5g	
	pH	5.8	
	Gellan Gum	3.5g	
Selection	Hygromycin B	100 µg/ml	
	Timentin	100 µg/ml	
YEB Media	Beef Extract Yeast Extract Peptone Sucrose $MgSO_4.H_2O$		7

Supplementary Table 4 – Primers

Purpose	Target gene	Primer Orientation	Primer sequence	Target Vector
CRISPR knockout	Os03g01610, XX50, XX40	Forward	TGGCGTCGACAAGGCTCCCTTCCT	pCBSG vector
	Os03g01610, XX50, XX40	Reverse	AAACAGGAAGGGAGCCTTGTCGAC	pCBSG vector pCBSG vector
	Os10g17660, XX80	Forward Reverse	TGGCGCCCCGACCGGACTGTTTCT	pCBSG vector
	Os10g17660, XX80		AAACAGAAACAGTCCGGTCGGGGC	
pLAT52 vector construction	Os03g01610	Forward	TACCATGGAGTCTAGAATGGCATCCTCCTCCCTTCTA	
	Os03g01610	Reverse	ATATCTCCTTGGATCCACTTGGCCTGGACGTTGGA	
	Os10g17660	Forward Reverse	TACCATGGAGTCTAGAATGTCGTGGCAGACGTACGT	
	Os10g17660		ATATCTCCTTGGATCCACAGGCCCTGCTCTACGAGGT	
qRT-PCR	Os03g01610	Forward	CCATTTGGAGGATAGACACC	
	Os03g01610	Reverse	GCTTAAAATGCTCACTTGG	
	Os03g01640	Forward	GTGTCACCACCGAGTCTGAC CCCCTGTGTGCATAGGTAGC	
	Os03g01640	Reverse	GGTCGAGAGGCTCGGCGA ACAATCCCCATGCATGCA	
	Os10g17660	Forward	GGTTGAGAGGCTCGGTGA	
	Os10g17661	Reverse	GATCTGCAATGTGTGGTGT	
	Os10g17680	Forward Reverse		
	Os10g17681			
bridization probe Sense	Os03g01650, XX10	Forward	TAATACGACTCACTATAGGGATATGTGTTGCTGCATGCTT	
	Os03g01650, XX10	Reverse	TAAATAAATAGCGCAGTGCA	
	Os03g01640	Forward	TAATACGACTCACTATAGGGCGTCATCCCCGATAATTGGAA	
	Os03g01640	Reverse	AGATACTCCCTATTATTTT	
	Os10g17660	Forward	TAATACGACTCACTATAGGGTGTAGATCACCGTCGCCATC	
	Os10g17660	Reverse	TTTTAGCGGCGTCAAAGGA	
	Os10g17680	Forward Reverse	TAATACGACTCACTATAGGGCCTGTAGATCGATCGATCA	
	Os10g17680		TGTTCCAGATTCACAACTCA	
ybridization probe Antisense	Os03g01650, XX10	Forward	ATATGTGTTGCTGCATGCTT	
	Os03g01650, XX10	Reverse	TAATACGACTCACTATAGGGGACCAATCCCAGCCAGATTTG	
	Os03g01640	Forward	CTAACGTCATCCCCGATAATTGGAA	
	Os03g01640	Reverse	TAATACGACTCACTATAGGGAGATACTCCCTATTATTTT	
	Os10g17660	Forward	TGTAGATCACCGTCGCCATC	
	Os10g17660	Reverse	TAATACGACTCACTATAGGGTTTTAGCGGCGTCAAAGGA	
	Os10g17680	Forward Reverse	CCTGTAGATCGATCGATCAC	
	Os10g17680		TAATACGACTCACTATAGGGTGTTCCAGATTCACAACTCA	
pMAL-c2X vector construction	Os03g01610	Forward	TTCAGAATTCGGATCCATGGCATCCTCTTGTCTCCTTC	pBI221-pLAT52::()-mGFP
	Os03g01610	Reverse	GGCCAGTGCCAAGCTTGTCGACTTGGATCTCTGATTTG	pBI221-pLAT52::()-mGFP
	Os10g17660	Forward Reverse	TACCATGGAGTCTAGAATGTCGTGGCAGACGTACGT	pBI221-pLAT52::()-mGFP
	Os10g17660		ATATCTCCTTGGATCCACAGGCCCTGCTCTACGAGGT	pBI221-pLAT52::()-mGFP
CRISPR sequence confirm	Os03g01640 Os03g01640	Forward	TTGTCTCCTTCTGGCCTGT	
	Os03g01610	Reverse	GCGGACGCGACGAAACTGCA GTGTTGTGGTGGCGGCTA	
	Os03g01610	Forward	GGCCTTACGGAGCTCTTCG	
	Os10g17660	Reverse	TTGGGCAGCAGGAGGCGGCG CTGACGAATTGAACTGTGG	
	Os10g17660	Forward	GCAGGCAGATCGAGCAGGCG TGAGAAACTGCGGACCCAT	
	Os10g17680	Reverse		
	Os10g17680	Forward Reverse		

REFERENCES FROM CHAPTERS 2-5

Abou Chakra, O.R., Sutra, J.P., Demey Thomas, E., Vinh, J., Lacroix, G., Poncet, P., et al., 2012. Proteomic analysis of major and minor allergens from isolated pollen cytoplasmic granules. J Proteome Res. 11, 1208-16.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic Local Alignment Search Tool. Journal of Molecular Biology. 215, 403-10.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25.

Andersson, K., Lidholm, J., 2003. Characteristics and immunobiology of grass pollen allergens. Int Arch Allergy Immunol. 130, 87-107.

Asero, R., Monsalve, R., Barber, D., 2008. Profilin sensitization detected in the office by skin prick test: a study of prevalence and clinical relevance of profilin as a plant food allergen. Clinical and Experimental Allergy. 38, 1033-7.

Asturias, J.A., Gomez-Bayon, N., Arilla, M.C., Sanchez-Pulido, L., Valencia, A., Martinez, A., 2002. Molecular and structural analysis of the panallergen profilin B cell epitopes defined by monoclonal antibodies. International Immunology. 14, 993-1001.

Azria, D., Bhalla, P.L., 2011. Agrobacterium-mediated transformation of Australian rice varieties and promoter analysis of major pollen allergen gene, Ory s 1. Plant Cell Reports. 30, 1673-81.

Bairoch, A., Bougueleret, L., Altairac, S., Amendolia, V., Auchincloss, A., Puy, G.A., et al., 2007. The universal protein resource (UniProt). Nucleic Acids Research. 35, D193-D7.

Barker, N.P., Clark, L.G., Davis, J.I., Duvall, M.R., Guala, G.F., Hsiao, C., et al., 2001.

Phylogeny and subfamilial classification of the grasses (Poaceae). Annals of the Missouri Botanical Garden. 88, 373-457.

Beggs, P.J., Katelaris, C.H., Medek, D., Johnston, F.H., Burton, P.K., Campbell, B., et al., 2015. Differences in grass pollen allergen exposure across Australia. Aust N Z J Public Health. 39, 51-5.

Bhalla, P.L., Singh, M.B., 2004. Knocking out expression of plant allergen genes. Methods. 32, 340-5.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research. 42, 252-8.

Boettcher, M., McManus, M.T., 2015. Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. Molecular Cell. 58, 575-85.

Borrill, P., Ramirez-Gonzalez, R., Uauy, C., 2016. expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform. Plant Physiology. 170, 2172.

Brinkman, E.K., Chen, T., Amendola, M., van Steensel, B., 2014. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research. 42, e168-e.

Chen, M.L., Xu, J., Devis, D., Shi, J.X., Ren, K., Searle, I., et al., 2016. Origin and functional prediction of pollen allergens in plants. Plant Physiology. 172, 341-57.

Chen, S.-q., Zhong, W., Liu, M.-x., Xie, Z.-w., Wang, H.-h., 2008. Pollen grain germination and pollen tube growth in pistil of rice. Rice Science. 15, 125-30.

Cheng, J., Khan, A., Wenming, Q., Li, J., Zhou, H., Zhang, Q., et al., 2012 Diversification of genes encoding granule-bound starch synthase in monocots and dicots is marked by multiple genome-wide duplication events. Public Library of Science One 7:e30088.

Coast, O., Murdoch, A.J., Ellis, R.H., Hay, F.R., Jagadish, K.S.V., 2016. Resilience of rice (Oryza spp.) pollen germination and tube growth to temperature stress. Plant Cell and Environment. 39, 26-37.

Cosgrove, D.J., Bedinger, P., Durachko, D.M., 1997. Group I allergens of grass pollen as cell wall-loosening agents. Proceedings of the National Academy of Sciences of the United States of America. 94, 6559-64.

Dash, S., Van Hemert, J., Hong, L., Wise, R.P., Dickerson, J.A., 2011. PLEXdb: gene expression resources for plants and plant pathogens. Nucleic Acids Research. 40, D1194-D201.

Davies, J.M., 2014. Grass pollen allergens globally: the contribution of subtropical grasses to burden of allergic respiratory diseases. Clin Exp Allergy. 44, 790-801.

Davies, J.M., Li, H., Green, M., Towers, M., Upham, J.W., 2012. Subtropical grass pollen allergens are important for allergic respiratory diseases in subtropical regions. Clin Transl Allergy. 2, 1.

Devis, D.L., Davies, J.M., Zhang, D.B., 2017. Molecular features of grass allergens and development of biotechnological approaches for allergy prevention. Biotechnology Advances. 35, 545-56.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., et al., 2018. The Pfam protein families database in 2019. Nucleic Acids Research. 47, D427-D32.

Fedorov, A.A., Ball, T., Mahoney, N.M., Valenta, R., Almo, S.C., 1997. The molecular basis for allergen cross-reactivity: crystal structure and IgE-epitope mapping of birch pollen profilin. Structure. 5, 33-45.

Fernandez-Pozo, N., Zheng, Y., Snyder, S.I., Nicolas, P., Shinozaki, Y., Fei, Z., et al., 2017. The Tomato Expression Atlas. Bioinformatics. 33, 2397-8.

Friedman, J., Barrett, S.C.H., 2009. Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. Annals of Botany. 103, 1515-
27.

Gibbon, B.C., Ren, H., Staiger, C.J., 1997. Characterization of maize (*Zea mays*) pollen profilin function in vitro and in live cells. Biochemical Journal. 327, 909-15.

Goodman, R.E., Breiteneder, H., 2019. The WHO/IUIS Allergen Nomenclature. Allergy.
74, 429-31.

Goodstein, D.M., Shu, S.Q., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., et al., 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research. 40, D1178-D86.

Hedhly, A., 2011. Sensitivity of flowering plant gametophytes to temperature fluctuations. Environmental and Experimental Botany. 74, 9-16.

Hirano, K., Hino, S., Oshima, K., Okajima, T., Nadano, D., Urisu, A., et al., 2013. Allergenic potential of rice-pollen proteins: expression, immuno-cross reactivity and IgE binding. Journal of Biochemistry. 154, 195-205.

Howe, E.A., Sinha, R., Schlauch, D., Quackenbush, J., 2011. RNA-Seq analysis in MeV. Bioinformatics (Oxford, England). 27, 3209-10.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al., 2008. Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. Advances in Bioinformatics. 2008, 5.

Huang, S.R., McDowell, J.M., Weise, M.J., Meagher, R.B., 1996. The Arabidopsis profilin gene family - Evidence for an ancient split between constitutive and pollen-specific profilin genes. Plant Physiology. 111, 115-26.

Janson, G., Zhang, C., Prado, M.G., Paiardini, A., 2016. PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL. Bioinformatics. 33, 444-6.

Jimenez-Lopez, J.C., Rodriguez-Garcia, M.I., Alche, J.D., 2013. Analysis of the effects of polymorphism on pollen profilin structural functionality and the generation of conformational, T- and B-cell epitopes. PLoS One. 8, e76066.

Kandasamy, M.K., McKinney, E.C., Meagher, R.B., 2002. Plant profilin isovariants are distinctly regulated in vegetative and reproductive tissues. Cell Motil Cytoskeleton. 52, 22-32.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28, 1647-9.

Kumar, R., Srivastava, P., Kumari, D., Fakhr, H., Sridhara, S., Arora, N., et al., 2007. Rice (*Oryza sativa*) allergy in rhinitis and asthma patients: A clinico-immunological study. Immunobiology. 212, 141-7.

Letunic, I., Doerks, T., Bork, P., 2015. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 43, D257-60.

Levin, M., Rydnert, F., Kallstrom, E., Tan, L.W., Wormald, P.J., Lindstedt, M., et al., 2013. Phl p 1-specific human monoclonal IgE and design of a hypoallergenic group 1 grass pollen allergen fragment. J Immunol. 191, 551-60.

Li, G., Yang, X.J., Zhang, X.Q., Song, Y., Liang, W.Q., Zhang, D.B., 2018. Rice morphology determinant-mediated actin filament organization contributes to pollen tube growth. Plant Physiology. 177, 255-70.

Li, J.Y., Sun, Y.W., Du, J.L., Zhao, Y.D., Xia, L.Q., 2017. Generation of targeted point mutations in rice by a modified CRISPR/Cas9 system. Molecular Plant. 10, 526-9.

Liang, G., Zhang, H.M., Lou, D.J., Yu, D.Q., 2016. Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. Scientific Reports. 6, 8.

Liu, X., Qu, X., Jiang, Y., Chang, M., Zhang, R., Wu, Y., et al., 2015. Profilin Regulates Apical Actin Polymerization to Control Polarized Pollen Tube Growth. Molecular Plant. 8, 1694-709.

Lyons, E., Freeling, M., 2008. How to usefully compare homologous plant genes and chromosomes as DNA sequences. Plant Journal. 53, 661-73.

Lyons, E., Pedersen, B., Kane, J., Freeling, M., 2008. The value of nonmodel genomes and an example using SynMap within CoGe to dissect the hexaploidy that predates the rosids. Tropical Plant Biology. 1, 181-90.

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., et al., 2015. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Molecular Plant. 8, 1274-84.

Ma, X.L., Zhu, Q.L., Chen, Y.L., Liu, Y.G., 2016. CRISPR/Cas9 Platforms for genome editing in plants: developments and applications. Molecular Plant. 9, 961-74.

Marchler-Bauer, A., Bo, Y., Han, L.Y., He, J.E., Lanczycki, C.J., Lu, S.N., et al., 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Research. 45, D200-D3.

Mari, A., Scala, E., Palazzo, P., Ridolfi, S., Zennaro, D., Carabella, G., 2006. Bioinformatics applied to allergy: Allergen databases, from collecting sequence information to data integration. The Allergome platform as a model. Cellular Immunology. 244, 97-100.

Marth, K., Focke, M., Flicker, S., Valenta, R., 2004. Human monoclonal antibody-based quantification of group 2 grass pollen allergens. J Allergy Clin Immunol. 113, 470-4.

McKenna, S.T., Vidali, L., Hepler, P.K., 2004. Profilin inhibits pollen tube growth through actin-binding, but not poly-L-proline-binding. Planta. 218, 906-15.

Mohammed, S., Samad, A.A., Rahmat, Z., 2019. *Agrobacterium*-mediated transformation of rice: constraints and possible solutions. Rice Science. 26, 133-46.

Müssar, K.J., Kandasamy, M.K., McKinney, E.C., Meagher, R.B., 2015. Arabidopsis plants deficient in constitutive class profilins reveal independent and quantitative genetic effects. BMC Plant Biol. 15, 16.

Nakamura, R., Matsuda, T., 1996. Rice allergenic protein and molecular-genetic approach for hypoallergenic rice. Biosci Biotechnol Biochem. 60, 1215-21.

Nishimura, A., Aichi, I., Matsuoka, M., 2006. A protocol for Agrobacterium-mediated transformation in rice. Nature Protocols. 1, 2796-802.

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., et al., 2007. The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res.
35.

Pandey, D.K., Chaudhary, B., 2017. Evolutionary expansion and structural functionalism of the ancient family of profilin proteins. Gene. 626, 70-86.

Paulus, K.E., Schmid, B., Zajic, D., Schafer, A., Mahler, V., Sonnewald, U., 2012. Hypoallergenic profilin - a new way to identify allergenic determinants. FEBS Journal. 279, 2727-36.

Petrovska, N., Wu, X.L., Donato, R., Wang, Z.Y., Ong, E.K., Jones, E., et al., 2004. Transgenic ryegrasses (Lolium spp.) with down-regulation of main pollen allergens. Mol Breeding. 14, 489-501.

Pomes, A., Davies, J.M., Gadermaier, G., Hilger, C., Holzhauser, T., Lidholm, J., et al., 2018. WHO/IUIS Allergen Nomenclature: Providing a common language. Molecular Immunology. 100, 3-13.

Prasad, V., Stromberg, C.A., Alimohammadian, H., Sahni, A., 2005. Dinosaur coprolites and the early evolution of grasses and grazers. Science. 310, 1177-80.

Radauer, C., Nandy, A., Ferreira, F., Goodman, R.E., Larsen, J.N., Lidholm, J., et al., 2014. Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. Allergy. 69, 413-9.

Rao, G.S., Deveshwar, P., Sharma, M., Kapoor, S., Rao, K.V., 2018. Evolvement of transgenic male-sterility and fertility-restoration system in rice for production of hybrid varieties. Plant Molecular Biology. 96, 35-51.

Ren, H.Y., Xiang, Y., 2007. The function of actin-binding proteins in pollen tube growth. Protoplasma. 230, 171-82.

Rosa, C., Kuo, Y.-W., Wuriyanghan, H., Falk, B.W., 2018. RNA interference mechanisms and applications in plant pathology. Annual Review of Phytopathology. 56, 581-610.

Sampedro, J., Cosgrove, D.J., 2005. The expansin superfamily. Genome Biol. 6, 11.

Sampedro, J., Guttman, M., Li, L.C., Cosgrove, D.J., 2015. Evolutionary divergence of βexpansin structure and function in grasses parallels emergence of distinctive primary cell wall traits. Plant J. 81, 108-20.

Sedeek, K.E.M., Mahas, A., Mahfouz, M., 2019. Plant Genome Engineering for Targeted Improvement of Crop Traits. Frontiers in Plant Science. 10, 16.

Sharma, A., Sharma, N., Bhalla, P., Singh, M., 2017. Comparative and evolutionary analysis of grass pollen allergens using *Brachypodium distachyon* as a model system. PLoS ONE. 12, e0169686.

Shi, J.X., Cui, M.H., Yang, L., Kim, Y.J., Zhang, D.B., 2015. Genetic and biochemical mechanisms of pollen wall development. Trends Plant Sci. 20, 741-53.

Shukla, P., Singh, N.K., Gautam, R., Ahmed, I., Yadav, D., Sharma, A., et al., 2017. Molecular Approaches for Manipulating Male Sterility and Strategies for Fertility Restoration in Plants. Molecular Biotechnology. 59, 445-57.

Smith, P.M., Avjioglu, A., Ward, L.R., Simpson, R.J., Knox, R.B., Singh, M.B., 1994. Isolation and characterization of group-I isoallergens from Bermuda grass pollen. Int Arch Allergy Immunol. 104, 57-64.

Songnuan, W., 2013. Wind-pollination and the roles of pollen allergenic proteins. Asian Pacific Journal of Allergy and Immunology. 31, 261-70.

Sun, T., Li, S., Ren, H., 2013. Profilin as a regulator of the membrane-actin cytoskeleton interface in plant cells. Frontiers in Plant Science. 4, 512.

Swoboda, I., Bhalla, P.L., Xu, H., Zhang, Y., Mittermann, I., Valenta, R., et al., 2001. Identification of pronp1, a tobacco profilin gene activated in tip-growing cells. Plant Mol Biol. 46, 531-8.

Tada, Y., Nakase, M., Adachi, T., Nakamura, R., Shimada, H., Takahashi, M., et al., 1996. Reduction of 14-16 kDa allergenic proteins in transgenic rice plants by antisense gene. FEBS Lett. 391, 341-5.

Thakur, P., Kumar, S., Malik, J.A., Berger, J.D., Nayyar, H., 2010. Cold stress effects on reproductive development in grain crops: An overview. Environmental and Experimental Botany. 67, 429-43. 12

Tiwari, R., Bhalla, P.L., Singh, M.B., 2009. Mapping of IgE-binding regions on recombinant Cyn d 1, a major allergen from Bermuda Grass Pollen (BGP). Clinical and molecular allergy : CMA. 7, 3-.

Valdivia, E.R., Durachko, D.M., Wagner, E., Stephenson, A.G., Cosgrove, D.J., 2003. βexpansins in maize pollen, role of Zea m 1 in development and fertilization. Plant Biology 2003, 121-.

Valdivia, E.R., Stephenson, A.G., Durachko, D.M., Cosgrove, D., 2009. Class B βexpansins are needed for pollen separation and stigma penetration. Sex Plant Reprod. 22, 141-52.

Valenta, R., Almo, S., Ball, T., Dolecek, C., Steinberger, P., Laffer, S., et al., 1998. The immunoglobulin E-allergen interaction: a target for therapy of type I allergic diseases. Int Arch Allergy Immunol. 116, 167-76.

Verbruggen, N., Hermans, C., 2008. Proline accumulation in plants: a review. Amino Acids. 35, 753-9.

Vidali, L., T. McKenna, S., Hepler, P.,2001 Actin polymerization Is essential for pollen tube growth. Molecular Biology Cell **12,** 2534-2545

Wakasa, Y., Hirano, K., Urisu, A., Matsuda, T., Takaiwa, F., 2011. Generation of transgenic rice lines with reduced contents of multiple potential allergens using a null mutant in combination with an RNA silencing method. Plant Cell Physiol. 52, 2190-9.

Wan, X.Y., Wu, S.W., Li, Z.W., Dong, Z.Y., An, X.L., Ma, B., et al., 2019. Maize genic male-sterility genes and their applications in hybrid breeding: progress and perspectives. Molecular Plant. 12, 321-42.

Wang, Z., Guo, C.-Q., Ren, H.-Y., 2006. Profilins in plant cells. Zhiwu Shengli yu Fenzi Shengwuxue Xuebao. 32, 261-70.

Wu, Y., Meeley, R.B., Cosgrove, D.J., 2001. Analysis and expression of the α-expansin and β-expansin gene families in maize. Plant Physiol. 126, 222-32.

Xu, H.L., Theerakulpisut, P., Goulding, N., Suphioglu, C., Singh, M.B., Bhalla, P.L., 1995. Cloning, expression and immunological characterization of Ory s 1, the major allergen of rice pollen. Gene. 164, 255-9.

Yang, X., Li, G., Tian, Y., Song, Y., Liang, W., Zhang, D., 2018. A gice glutamyl-tRNA synthetase modulates early anther cell division and patterning. Plant Physiology. 177, 728.

Yu, L.X., Parthasarathy, M.V., 2014. Molecular and cellular characterization of the tomato pollen profilin, LePro1. Plos One. 9, 14.

Yuan, H.C., Wu, K.G., Chen, C.J., Su, S.N., Shen, H.D., Chen, Y.J., et al., 2012. Mapping of IgE and IgG(4) antibody-binding epitopes in Cyn d 1, the major allergen of bermuda grass pollen. International Archives of Allergy and Immunology. 157, 125-35.

Zeng, J., Dong, Z., Wu, H., Tian, Z., Zhao, Z., 2017. Redox regulation of plant stem cell fate. EMBO J. 36, 2844-55.

Zhang, D., Liang, W., Yin, C., Zong, J., Gu, F., Zhang, D., 2010. OsC6, encoding a lipid transfer protein, is required for postmeiotic anther development in rice. Plant Physiology. 154, 149-62.

Zhang, D.B., Wilson, Z.A., 2009. Stamen specification and anther development in rice. Chin Sci Bull. 54.

Zhang, Y.C., Liao, J.Y., Li, Z.Y., Yu, Y., Zhang, J.P., Li, Q.F., et al., 2014. Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. Genome Biology. 15.

Zhang, Z., Zhang, Y., Tan, H.X., Wang, Y., Li, G., Liang, W.Q., et al., 2011. *RICE MORPHOLOGY DETERMINANT* encodes the Type II Formin FH5 and regulates rice morphogenesis. Plant Cell. 23, 681-700.

Zhu, T., Niu, D.-K., 2013. Frequency of intron loss correlates with processed pseudogene abundance: a novel strategy to test the reverse transcriptase model of intron loss. BMC Biology. 11, 23-.