Highlights

Title: Identification of vicilin, legumin and antimicrobial peptide 2a as macadamia nut allergens

- This is the first study on the identification and purification of macadamia nut allergens.
- Four novel IgE-reactive proteins have been identified in macadamia nut.
- Macadamia nut vicilin and legumin are high molecular weight allergens.
- Macadamia antimicrobial peptide 2a and nsLTP are low molecular weight allergens.
- The use of these allergens will improve diagnosis of macadamia nut allergy.

Identification of vicilin, legumin and antimicrobial peptide 2a as macadamia nut allergens

Stefan Kabasser[a], Kunal Pratap[b, c, d], Sandip Kamath[b, c, d], Aya C. Taki[e], Thanh Dang[f], Jennifer Koplin[f], Kirsten Perrett[f], Karin Hummel[g], Christian Radauer[a], Heimo Breiteneder[a], Andreas L. Lopata[b, c, d,*], Merima Bublin[a,*]

[a]Institute of Pathophysiology and Allergy Research, Center of Pathophysiology, Infectiology and Immunology, Medical University of Vienna, Vienna, Austria

[b]Molecular Allergy Research Laboratory, College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, QLD, Australia

[c]Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, QLD, Australia

[d]Center for Molecular Therapeutics, James Cook University, Townsville, QLD, Australia

[e]School of Veterinary Biosciences, Melbourne Veterinary School, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, VIC, Australia

[f]Department of Paediatrics, Murdoch Children's Research Institute, The University of Melbourne, Flemington Road, Parkville, VIC, Australia

[g]VetCore Facility for Research, University of Veterinary Medicine, Vienna, Austria

***Corresponding authors:**

Merima Bublin, PhD

Institute of Pathophysiology and Allergy Research

Center of Pathophysiology, Infectiology and Immunology

Medical University of Vienna

Waehringer Guertel 18-20, 1090 Vienna, Austria

E-mail: merima.bublin@meduniwien.ac.at

Andreas L. Lopata, PhD

Pharmacy and Medical Research Building

James Cook University

Townsville, QLD, Australia

Email: andreas.lopata @jcu.edu.au

Abstract

Macadamia nut is an increasingly popular food item of a healthy diet. However, macadamia nut is also a potent allergenic food. To date, there is little information about the allergenic proteins involved. In this study, using sera from macadamia nut allergic individuals, four IgE-binding proteins were detected. Their identities were determined by tandem mass spectrometry with de novo sequencing. Three IgE-reactive proteins, the vicilin Mac i 1, the legumin Mac i 2 and the antimicrobial peptide 2a/Mac i 1 (28-76) were purified from the nut while the non-specific lipid transfer protein was produced as a recombinant in *Pichia pastoris*. IgE-binding assays using sera from well-characterized groups of tree nut and/or peanut allergic patients revealed that the allergens were mainly recognized by sera from macadamia nut allergic individuals. Hence, these newly discovered allergens will enable molecular diagnostics to identify patients at high risk of macadamia nut allergy.

Keywords: allergen, food allergens, food allergy, legumin, macadamia allergy, macadamia nut, vicilin, tree nut allergy

1 Introduction

The macadamia tree *(Macadamia integrifolia)* is native to Australia and belongs to the family of Proteaceae. Within the last ten years, the global production of macadamia nuts has more than doubled (International Nut & Dried Fruit Council, 2020) and consequently, roasted macadamia nut (Buthelezi, Magwaza, & Tesfay, 2019) has become a popular ingredient in a variety of food products such as snacks, biscuits and cakes (Center for the Promotion of Imports (CBI), 2021).

While macadamia nut consumption is associated with beneficial health effects such as lowering of plasma total and LDL cholesterol levels (Alasalvar, Salvadó, & Ros, 2020; Garg, Blake, & Wills, 2003), it may pose a significant health risk to atopic patients. Epidemiological data from Australia suggest that clinically confirmed macadamia nut allergy affects approximately 0.2% of children (McWilliam et al., 2019) and adolescents (Sasaki et al., 2018). According to the epidemiological study by Brough et al., in Europe, the prevalence of macadamia nut allergy among tree nut and seed allergic children ranges from 10-17%, depending on the geographic region (Brough et al., 2020).

As for many other tree nuts, allergic reactions to macadamia nut can range from mild oral symptoms to potentially life-threatening anaphylaxis (De Knop, Hagendorens, & Bridts, 2010; Ehlers et al., 2020; Herbst, Wahl, & Frosch, 2010; McWilliam et al., 2018; Sutherland, O'Hehir, Czarny, & Suphioglu, 1999; Yoshida et al., 2021). Macadamia nut allergy diagnosis is based on patients' clinical history in combination with evidence of sensitization or, in unclear cases, an oral food challenge (OFC). While the measurement of extract-specific IgE often provided false negative results, skin-prick test (SPT) using macadamia nut extract and prick-to-prick test results often correlate well with clinical symptoms (De Knop et al., 2010; Ekbote, Hayman, & Bansal, 2010; Herbst et al., 2010; Sutherland et al., 1999; Yoshida et al., 2021). Molecular diagnosis

using individual allergens to quantify specific IgE (sIgE) levels in patients with suspected tree nut allergy has proved to be helpful to elucidate distinct sensitization phenotypes and to predict clinical reactivity, replacing the need for extract-based SPTs and avoiding resource-intensive OFCs (Ballmer-Weber et al., 2019).

The main protein families involved in tree nut allergy include 2S albumins, 7S globulins (vicilins), 11S globulins (legumins) and non-specific lipid transfer proteins (nsLTP) (Geiselhart et al., 2018). In contrast to other tree nuts such as hazelnut or walnut, knowledge of macadamia nut allergens is insufficient for improved diagnosis. Recently, IgE sensitization to vicilin-like antimicrobial peptides 2-1, 2-2 and 2-3 has been suggested as a potential indicator for systemic reactions to macadamia nuts (Ehlers et al., 2020). The vicilin-like antimicrobial peptides 2-3 was found to be the most abundant protein in macadamia nut extracts and displays high sequence similarity with the N- terminal part of the walnut allergen Jug r 2, implying a potential cross-reactivity (Rost, Muralidharan, & Lee, 2020). In earlier publications, three IgE-binding macadamia nut proteins (of apparent mW 12, 17 and 45 kDa) were observed but not identified in detail (Herbst et al., 2010; Sutherland et al., 1999). One case studies reported that a protein with an apparent mW of 17 kDa exhibited a low degree of cross-reactivity with hazelnut, but not with peanut (Sutherland et al., 1999). The mass-spectrometric analysis of the macadamia nut proteome provided evidence for a potential 11S legumin homologue, whereas the presence of a 2S albumin remains questionable (Rost et al., 2020). Currently, studies focusing on the identification and characterization of IgE-reactive macadamia nut proteins are limited, thus the value of individual components for molecular diagnosis is largely unknown.

The present study describes the purification and characterization of natural and recombinant IgE-binding macadamia nut proteins using various biochemical and immunological methods.

Together, the isolated IgE-binding proteins form a panel of allergens which may be useful in future studies to improve the accuracy of macadamia nut allergy diagnosis.

2 Materials and Methods

2.1. Chemicals and reagents

All reagents were purchased from Sigma-Aldrich (Saint Louis, MS, USA) unless stated otherwise.

2.2. Preparation of macadamia nut protein extract

Roasted macadamia nuts (species *M. integrifolia*) were obtained from a local supermarket. The nuts (120 g) were ground and defatted three times by stirring (1 h, at room temperature) in 720 mL of n-hexane (1:6, w/v). After drying, protein extract was prepared by adding 30 g of defatted macadamia nut powder in 150 mL of PBS (1:5, w/v), containing 3% polyvinyl polypyrrolydone and protease inhibitor cocktail tablets (1 tablet/50 mL Roche Molecular Biochemicals, Mannheim, Germany). The extract was stirred for 30 min, at 4 °C, centrifuged (40 000 × g, at 4 °C for 1 h), and the supernatant was filtered using a 0.45 µm filter (Sarstedt, Nümbrecht, Germany). An overview of protein extraction and subsequent purification of allergens is provided in Figure S1.

2.3. Purification of macadamia nut antimicrobial peptides

The macadamia nut protein extract (100 mL) was cooled to 4 °C and ice-cold methanol was added to a final concentration of 60% (v/v) to precipitate globulins. After stirring at 4 °C for 30 min, the extract was centrifuged (3000 x g, at 4 °C for 45 min). The supernatant was lyophilized and dried proteins resuspended in 10 mL of ddH_2O before dialyzing (24 h, at 4 °C) against 20 mM Tris/HCl, pH 8.0, using a dialysis tubing with an exclusion limit of 1 kDa (Spectrum Laboratories, Gardena, CA, USA). Then, the dialysate was loaded onto a self-packed Q Sepharose Fast Flow (GE Healthcare, Uppsala, Sweden) chromatography column (1.0 x 9.0 cm), pre- equilibrated with 20 mM Tris/HCl, pH 8.0. Column-bound proteins were eluted at a flow

rate of 1 mL/min by a linear gradient of NaCl from 0 to 0.5 M within 50 minutes (fraction size: 2 mL). Eluent was monitored for protein by following the absorbance at 280 nm. After analysis by 15% SDS-PAGE, fractions eluting between 0.3-0.4 M NaCl, assumed to contain MiAMP2a, were pooled.

2.4. Purification of Mac i 1 (7S globulin)

Macadamia nut protein extract was prepared as described above (section 2.2) with the difference that 20 mM Tris/HCl, pH 7.5, containing high salt concentration (1 M NaCl) was used in order to increase solubility of globulins. The filtered extract (8 mL) was loaded onto a HiPrep 26/60 Sephacryl S-200 HR size exclusion chromatography (SEC) column (GE Healthcare, Uppsala, Sweden), pre- equilibrated with 20 mM Tris/HCl, pH 7.5, 1 M NaCl. The column void volume was determined with Blue Dextran and the column was calibrated with proteins from a preparation of gel filtration standards (Sigma-Aldrich; Saint Louis, MS, USA) including β-amylase (200 kDa), alcohol dehydrogenase (150 kDa) and bovine serum albumin (66 kDa). In addition, a low molecular weight protein (lysozyme, 14 kDa) (Merck, Darmstadt, Germany) was added to the standard preparation. Proteins were eluted at a flow rate of 1.3 mL/min and 5 mL fractions were collected. Protein was monitored at 280 nm. After SDS-PAGE assessment, the fractions showed an enriched 50 kDa band (assumed to be Mac i 1). Thus, fractions eluting between 120-130 min (estimated to be 150-66 kDa from SEC calibrants and Blue Dextran loading) were pooled and dialyzed (3.5 kDa MWCO dialysis tubing, Spectrum Laboratories, Gardena, CA, USA), against 20 mM Tris/HCl, pH 8.0, for 24 h at 4 °C. Subsequently, the sample (20 mL) was loaded onto a Mono Q 5/50 GL column (GE Healthcare, Uppsala, Sweden), pre- equilibrated with 20 mM Tris/HCl, pH 8.0. Proteins were eluted at a flow rate of 1 mL/min over a time of 30 min by a linear NaCl gradient (0-0.4 M) and fractions of 1 mL were collected.

The individual fractions were analysed by 15% SDS-PAGE and fractions eluting between 0.17-0.23 M NaCl enriched with a 50 kDa protein, were pooled. Subsequently, pooled fractions (15 mL) were subjected to a 5 mL HiTrap Con A Sepharose-4B column (GE Healthcare, Uppsala, Sweden), pre- equilibrated in 20 mM Tris/HCl, pH 7.5, containing 0.5 M NaCl. Mac i 1 was eluted in one fraction by adding 0.5 M alpha-methyl mannopyranoside (Sigma Aldrich, St. Louis, MS, USA).

2.5. Purification of Mac i 2 (11S globulin)

For the purification of Mac i 2, protein extraction and SEC were carried out as described in sections 2.2 and 2.4. Fractions eluting between 90-110 min (estimated to be >150 kDa from SEC calibrants and Blue Dextran loading) were pooled and applied to a 5 mL Con A Sepharose-4B column to remove any residual Mac i 1. Purified Mac i 2 was present in the flow-through, since it is not glycosylated and does not bind to ConA Sepharose-4B.

2.6. Expression and purification of recombinant macadamia nsLTP

The DNA sequence of mature macadamia nsLTP (ENA accession no: LR861101; nucleotide positions 79 to 375) was used for recombinant protein expression. Codon optimization, gene synthesis and subcloning of the sequence to the vector pPICZαA were performed by Thermofisher Scientific GeneArt GmbH (Regensburg, Germany). Recombinant Mac i nsLTP was expressed in the *Pichia pastoris* strain GS115, as described previously (Dubiela et al., 2017). The expressed protein was purified from the culture supernatant. Briefly, 800 mL of culture supernatant were lyophilized, redissolved in 20 mM sodium acetate buffer, pH 6.0 and dialyzed against the same buffer. The sample was loaded onto a SP Sepharose column, equilibrated with 20 mM sodium acetate buffer, pH 6.0. Column-bound proteins were eluted by a linear gradient of NaCl from 0-0.2 M.

2.7. Protein quantification and purity analysis

The BCA protein assay kit (Pierce, Cheshire, UK) was used to determine protein concentrations of the prepared macadamia protein extract and the purified samples. Purity of proteins was estimated by Coomassie-stained 15% SDS-PAGE (section 2.8) and tandem mass spectrometry (sections 2.10-2.11).

2.8. Electrophoresis

Macadamia nut extract (10 µg/lane) and purified proteins (3 µg/lane) were separated by 15% SDS-PAGE and 2D-PAGE under reducing and non-reducing conditions as previously described (Bublin et al., 2008). After electrophoresis, separated proteins were visualized by staining with Coomassie Brilliant Blue R or transferred to a nitrocellulose membrane.

2.9. Circular dichroism (CD) spectroscopy

For CD spectroscopy, macadamia antimicrobial peptide 2a and Mac i nsLTP (0.2 mg/mL) were dialyzed against 10 mM sodium phosphate buffer, pH 7.5. Mac i 1 and Mac i 2 were dialyzed against the same buffer containing 0.5 M NaF. CD spectra were recorded in the range of 190–250 nm at room temperature on a J-810S spectropolarimeter (Jasco International Co., Tokyo, Japan) using a 1 mm path length quartz cell. Spectra represent the average of three accumulations, recorded at 100 nm/min with a 2 s time constant, 1.0 nm resolution, and sensitivity of ±100 mdeg.

2.10. Protein analysis by Nano-LC ESI Orbitrap MS/MS

Protein identification by Nano-LC ESI Orbitrap MS/MS was performed at the VetCore Facility for Research (Veterinary University of Vienna, Vienna, Austria). According to the sample type suitable sample preparation protocols were applied: Solved pre-purified proteins were digested

directly using a standard protocol for in-solution digestion (Kumar et al., 2016) followed by peptide clean-up with C18 spin tips according to the manufacturer's instructions (Thermo Scientific). If proteins had been separated by SDS-PAGE or 2D-PAGE, a standard in-gel sample preparation protocol was used (Gutiérrez et al., 2019). More complex protein samples were prepared with a filter-aided sample preparation protocol (FASP) (Kumar et al., 2015). Resulting dried peptides of all sample preparation methods were subsequently analysed by Nano-LC ESI Orbitrap MS/MS as described previously (Gutiérrez et al., 2019).

2.11. Database search in ENA-macadamia with MASCOT

Raw data was searched using an in-house MASCOT server (version 2.4.1) with following parameters: enzyme trypsin; up to 2 missed cleavages; fixed modification carbamidomethyl (C); variable modifications deamidated (NQ), Gln->pyro-Glu (N-term Q), oxidation (M); MS peptide tolerance 10 ppm; MS/MS tolerance 0.02 Da; peptide charge 2+, 3+ and 4+. The database used was downloaded 5th Oct. 2018 from https://www.ebi.ac.uk/ena/browser/view/FLKO01000000.1 and additionally contained the sequence of Mac i nsLTP available in ENA as LR861101.1 in project PRJEB39358.

2.12. Database search and de novo sequencing with PEAKS X+

Database search and automated de novo sequencing were performed using the software PEAKS X+ (Bioinformatics Solutions Inc) (Ma et al., 2003). The following parameters were applied for database search: Parent mass error tolerance 10.0 ppm, fragment mass error tolerance 0.05 Da, enzyme trypsin, maximum missed cleavages 3, digest mode specific, fixed modification carbamidomethylation (C) (+57.02 Da), variable modifications deamidation (NQ) (+0.98 Da), oxidation (M) (+15.99 Da), pyro-Glu from Q (-17.03 Da), maximum variable post-translational modifications per peptide 3, database ENA_Macadamia, contaminant database cRAP. Protein

and peptide results were filtered according to the following criteria: peptide -10lgP threshold 0.1% FDR, protein -10lgP ≥ 20 and ≥ 2 unique peptides with significant peptides. The same parameters were applied for de novo sequencing. Resulting de novo peptides were filtered for a de novo score (Average Localized Confidence) ≥ 80.

2.13. Serum samples

The Australian HealthNuts study is a comprehensive population-based study of food allergy consisting of a cohort of 5276 children enrolled at age 1 and followed up to 6 years of age (Koplin et al., 2015; Osborne et al., 2011). For our study, a subset of patients recruited at 4-6 years with macadamia nut outcomes was investigated (Table 1). Patients' sera were grouped into patients with macadamia nut allergy (n=5) based on SPT to macadamia nuts ≥8 mm and one of the following; a) history of objective reaction >12 months ago consistent with OFC criteria or b) parent-reported avoiding food due to allergy. To compare allergen recognition profiles between macadamia nut allergic and tree nut allergic but macadamia tolerant patients, additional sera from tree nut allergic patients without macadamia nut allergy were used in this study and grouped into: (1) macadamia nut tolerant-sensitized individuals (n=8) who had a macadamia nut SPT 3-7 mm and parent reported ingestion history (eaten >1 time since age 4); (2) macadamia nut tolerant individuals without evidence of sensitization to macadamia nut (n=14), as defined by SPT ≤2 mm. Approval for the HealthNuts study was obtained from the Victorian State Government Office for Children (reference number CDF/07/492), the Victorian State Government Department of Human Services (reference number 10/07), and the Royal Children's Hospital Human Research Ethics Committee (reference number 27047). Informed consent was obtained from parents or guardians of all participants.

2.14. IgE ELISA

Wells of 96-well plates (Maxisorp; Nalge Nunc International, Roskilde, Denmark) were coated with 0.2 µg of pure protein or 1 µg of protein extract, diluted in coating buffer (50 mM Na-carbonate, pH 9.6). The plates were blocked for 2 h at room temperature with TBST containing 3% (w/v) BSA. Subsequently, patients' sera (diluted 1:10 in TBST containing 1% (w/v) BSA and 100 µg/mL horseradish peroxidase to block antibodies specific for cross-reactive carbohydrate determinants) were applied in duplicates overnight at 4°C. Detection of bound IgE was performed with an alkaline phosphatase (AP)-conjugated mouse anti-human IgE antibody (BD Pharmingen, San Jose, Ca, USA) followed by incubation with Sigma FAST p-nitrophenyl phosphate tablets (Sigma-Aldrich, St Louis, Mo, USA). Absorbance at 450 nm was measured. Sera of four non-atopic donors served as negative controls. Sera were regarded as positive if their OD exceeded the mean OD value of the four healthy controls plus three times their standard deviation.

2.15. IgE immunoblotting

IgE immunoblotting was performed as previously described (Kabasser et al., 2021) with the following modifications: membrane strips containing macadamia nut extract were blocked with low-fat powdered milk (5 % (w/v)) diluted in TBST. Subsequently, the strips were incubated overnight at 4 °C with pooled sera of macadamia nut allergic individuals or non-atopic controls (diluted 1:10-1:20). Before adding to the strips, pooled sera of macadamia nut allergic patients were incubated for 2 h at room temperature with 100 µg/mL horseradish peroxidase (Sigma-Aldrich, St Louis, Mo, USA). Bound IgE was detected using AP-conjugated antihuman-IgE (BD Pharmingen, San Diego, USA).

3 Results

3.1 Immunodetection of IgE-reactive proteins in macadamia nut extract

To identify IgE-reactive proteins in macadamia nut extract, an immunoblot using a serum pool from macadamia nut allergic patients (MA 1, 2 and 4) was performed (Figure 1a and Table 1). IgE-binding proteins were found in the low and high molecular weight range (12-70 kDa) under both reducing and non-reducing conditions (Figure 1a, lanes 1 and 3). Typically, the protein range between 10 and 70 kDa comprises IgE-reactive 2S, 7S and 11S seed storage proteins, as has been shown for other tree nuts (Geiselhart et al., 2018).

3.2 Purification and characterization of high molecular weight IgE-binding proteins

3.2.1 Mac i 1, a vicilin-like 7S globulin

Separation of macadamia nut protein extract by SEC resulted in three peaks containing high molecular weight proteins (Figure 2a). As visible in SDS-PAGE, a dominant 50 kDa band (referred to as Mac i 1) was found in the third peak corresponding to the 50 kDa IgE-binding protein in immunoblot. In anion exchange chromatography, Mac i 1 eluted through the addition of 0.2 M NaCl whilst other compounds required higher salt concentrations (Figure 2b). After the final step of Con A affinity chromatography, Mac i 1 was obtained with a purity of >90% (Figure 2c), as estimated from Coomassie-stained 15% SDS-PAGE.

During protein separation under reducing as well as non-reducing SDS-PAGE, Mac i 1 migrated as a single band (Figure 1b, lanes 6 and 11). In 2D-PAGE, multiple protein spots were visible (Figure S2a) which are assumed to be post-translationally processed forms of 7S globulin. Nano-LC ESI Orbitrap MS/MS analysis of in-gel tryptic digests of the 50 kDa protein (Figure 1b, lane

6) provided a 54% sequence coverage of UniProt entry Q9SPL4 (vicilin-like antimicrobial peptides 2-2). In total, 35 unique peptides were identified matching to the middle and C- terminal region (amino acid position 235–666) (Figure S3a and Table S1b). The calculated molecular mass of the fragment was 49.5 kDa (isoelectrical point: 6.5), which corresponds to the protein migration pattern (Figures 1b and S2a). Moreover, *in-silico* analysis of the fragment revealed two cupin domains and one possible N- linked glycosylation site at amino acid position 493. Hence, the purified protein represents a mature vicilin-like protein that originates from proteolytic processing of precursor Q9SPL4. The yield of purified protein was approximately 25 mg from 120 g of shelled macadamia nuts (see section 2.7).

3.2.2 Mac i 2, a legumin-like 11S globulin

Usually, legumin seed storage proteins present in tree nuts occur as ~350 kDa hexamers composed of six ~60 kDa monomers (Geiselhart et al., 2018), thus Mac i 2 was enriched in the first two peaks of the size exclusion chromatography (Figure 2a). Final purification was achieved by passage of the pooled peak fractions through a Con A Sepharose-4B column to remove contaminations with 7S globulin.

SDS-PAGE analysis under non-reducing conditions showed prominent bands in the range of 30 to 60 kDa (Figure 1b, lane 7). The bands are supposed to correspond to polymorphic 11S globulin monomers comprising basic and acidic polypeptide chains connected by at least one disulphide bridge. As shown for the legumins from other nut species (Müntz, 1998), the subunits dissociate into acidic chains of ~40 kDa and basic chains of ~20 kDa when reducing agent is added (Figure 1b, lane 12). In Coomassie-stained 2D-PAGE, the variety of individual protein spots in close proximity to each other indicate excessive post-translational processing of the respective polypeptides (Figure S2b).

Since there is currently no database entry for macadamia legumin, the isolated protein was identified by de novo sequencing. Through this process, 14 unique peptides were obtained, each composed of 6 to 34 amino acids (Table S1a). The identified peptides showed high amino acid sequence identity to conserved regions of legumins from *Asarum europaeum* (asarabacca), *Papaver somniferum* (opium poppy) and *Macleaya cordata* (plume poppy), hence it is evident that the purified protein represents the homologous legumin-like seed storage protein from *Macadamia integrifolia* (Figure S4). The protein sequence data of macadamia legumin is provided in the UniProt Knowledgebase under the accession number C0HLR7. The yield of pure protein was 118 mg from 120 g of shelled macadamia nuts (see section 2.7).

3.3 Purification and identification of low molecular weight IgE-binding proteins

3.3.1 MiAMP2a (Mac i 1 (28-76)), an antimicrobial peptide

For purification of the low molecular weight IgE-binding proteins in the range between 10 and 20 kDa (Figure 1) we used an already established protocol to isolate the alcohol-soluble prolamin fraction from the nut extract (Pfeifer et al., 2015). The addition of 60% (v/v) methanol to the extract resulted in the precipitation of globulins whilst macadamia antimicrobial peptides (MiAMPs) remained in solution. After anion exchange chromatography, three peaks containing low molecular weight proteins were obtained. Fractions of the first peak (Figure 2d) showed three bands of approximately 10, 13 and 17 kDa. In this fraction, in addition to MiAMP2a (10 kDa band; Figure S3b and Table S1c), also fragments of MiAMPs 2b-d were detected (13 and 17 kDa band; Table S1d-e), as assessed by in-gel tryptic digestion and subsequent mass-spectrometric analysis. The third fraction (Figure 2d, peak III) contained only MiAMP2a (Figure S3c and Table S1c). Twenty-one milligrams of MiAMP2a were obtained from 120 g of nuts (see section 2.7). Purified MiAMP2a exhibited one band of ~10 kDa under non-reducing conditions

(Figure 1b, lane 8) and dissociated into two bands when the reducing agent was present. It is plausible that the addition of DTT caused the reduction of disulphide bonds formed between several cysteine residues present in the sequence of MiAMP2a (Figures 1b, lane 13 and Figure S3b). Mass spectrometric analysis of corresponding spots in 2D-PAGE confirmed that MiAMP2a was present in the sample (Figure S2c and Table S1f-i). As known from earlier studies (Marcus, Green, Goulter, & Manners, 1999), the identified MiAMPs originate from proteolytic processing of a 666 amino acid long precursor protein (vicilin-like antimicrobial peptides 2-2) with database entry Q9SPL4, which we identified to be also the precursor of mature vicilin (section 3.2.1). High amino acid sequence identity between vicilin-like antimicrobial peptides 2-2, 2-3 (97%, Accession no: Q9SPL3) and 2-1 (97%, Accession no: Q9SPL5) indicate that the precursor exists in different isoforms contributing to highly polymorphic MiAMP and vicilin species present in macadamia nut.

3.3.2 Recombinant macadamia nsLTP

Peptides of macadamia nsLTP (ENA accession no: LR861101) were detected in the prolamin fraction of the extract by mass-spectrometric analysis (data not shown), but limited amounts of nsLTP complicated the process of purification attempts from the natural source. To overcome this problem, recombinant nsLTP was expressed as a soluble non-fusion protein in *P. pastoris*. The pure protein migrated as a single band at about 10 kDa in reducing and non-reducing 15% SDS-PAGE which is in good agreement with the theoretical mass of 9.3 kDa (Figure 1b, lanes 9 and 14; Figure 2e). Analysis of the protein by MS/MS yielded 59% sequence coverage of macadamia nsLTP (LR861101) (Figure S3c and Table S1f). The final yield was 2.8 mg per 800 mL *P. pastoris* culture (see section 2.7).

3.4 Secondary structure analysis

The folded state of the purified IgE-binding proteins was assessed by CD spectroscopy (Figure 3). The CD spectra of both Mac i 1 and Mac i 2 were typical for seed storage globulins, mainly consisting of beta-sheet structures. Therefore, the spectrum of Mac i 1 had a maximum at 193 nm and a minimum at 216 nm. Similarly, the spectrum of Mac i 2 had a maximum at 194 nm and a minimum at 212 nm. In contrast, the CD spectrum of MiAMP2a clearly showed an α-helical protein with two intense minima at 210 and 222 nm. Also, as anticipated, recombinant Mac i nsLTP provided a similar spectrum consistent with the established α-helical structure of nsLTP with minima at 210 and 221 nm.

3.5 Mac i 1, 2 and MiAMP2a represent relevant IgE-binding components in macadamia nut

The IgE-reactivity of macadamia nut extract and the purified proteins was assessed by IgE ELISA using individual sera from five macadamia nut allergic patients and eight macadamia nut sensitized but clinically tolerant volunteers (Figure 4). All five tested sera of allergic patients had IgE to macadamia extract and macadamia nut legumin (Mac i 2). Specific IgE to purified vicilin (Mac i 1) was detected in four and to MiAMP2a in three of five tested sera. Sensitization to recombinant Mac i nsLTP was observed in two of macadamia nut allergic patients. In the group of macadamia nut sensitized but clinically tolerant patients (n=8), extract was recognized by six (6/8), but Mac i 1 and MiAMP2a were weakly recognized by only two of the eight sera. IgE binding to Mac i 2 was seen in four and to Mac i nsLTP in two of macadamia nut tolerant patients. Notably, in the control group of 14 peanut and/or tree nut allergic but macadamia nut tolerant individuals only a very weak recognition of extract (5/14) was observed. The individual allergens were only weakly recognized by three of the tested patients (Figure 4).

4 Discussion

There has been substantial interest in the availability of pure and well-characterized allergen components for application in molecular diagnosis and immunotherapy in recent years. Up to date, knowledge of the culprit allergens of the potentially life-threatening macadamia nut allergy is lacking. In this study, we identified and characterized four IgE-binding proteins in macadamia nut which may be used as marker allergens to facilitate patient-tailored management of food allergy and to evaluate the risk of cross-reactivity with other tree nuts.

Our immunoblotting results of macadamia nut extract showed several IgE-binding components in the low and high molecular weight range (10-70 kDa). The overall observed pattern of IgE-reactive bands concurs with results from earlier case reports in which individual and pooled sera were used to demonstrate IgE reactivity to macadamia nut extract (Herbst et al., 2010; Sutherland et al., 1999).

Vicilin-like proteins are important allergens in tree nuts, and some of them (e.g. Cor a 11 from hazelnut) have been described as predictive markers for clinical reactivity (Masthoff et al., 2013). In this study, we purified a mature 50 kDa IgE-binding vicilin originating from a 666 aa precursor (Q9SPL4) which was previously also described as the precursor of a series of antimicrobial peptides (MiAMPs 2a-d) (Marcus et al., 1999). In silico analysis of the purified protein revealed cupin domains as well as a conserved glycosylation site, typical for members of the vicilin protein family. The amino acid sequence identity with IgE-binding vicilins from pecan, hazelnut and walnut is rather low (48%, 46% and 41%, respectively), which raises the question whether macadamia vicilin is involved in cross-reactivity with other allergenic foods.

Lately, two publications have highlighted the importance of IgE-binding antimicrobial peptides derived from proteolytic processing of vicilin precursors in peanut and walnut (Aalberse et al., 2020; Downs et al., 2014). In macadamia nut, IgE binding to the full-length precursor protein combining MiAMPs and mature vicilin was recently observed (Ehlers et al., 2020). In our study, we not only purified mature vicilin but also MiAMP2a derived from the N- terminal region of the precursor protein. As observed in other species, MiAMP2a contains paired C-X-X-X-C motifs enabling formation of disulfide bridges that contribute to a compact alpha-helical structure.

In addition to vicilins, the major protein constituents found in tree nuts are legumin-like 11S globulins. With regard to their clinical relevance, IgE sensitization to legumins from hazelnut (Cor a 9), almond (Pru du 6) and cashew (Ana o 2) have been identified as specific markers of tree nut allergy (Kabasser et al., 2021; Masthoff et al., 2013; van der Valk et al., 2017). Recently, peptides corresponding to a putative macadamia 11S globulin were identified by a shotgun mass-spectrometric approach. It was concluded that legumins, together with vicilin-like proteins and their processing products make up the most abundant protein species in macadamia nut (Rost et al., 2020). In line with these results, we identified an IgE-binding legumin-like seed storage protein in macadamia nut based on partial de novo sequencing. The sequenced peptides had >50% sequence identity with legumin-like proteins from other plants, namely *Macleaya cordata* (plume poppy) and *Papaver somniferum* (opium poppy). The degree of amino acid sequence identity shared with the IgE-binding legumins from Brazil nut (Ber e 2), hazelnut (Cor a 9), walnut (Jug r 4), cashew (Ana o 2), pistachio (Pis v 2), almond (Pru du 6), and peanut (Ara h 3) was between 35% and 55%. As for the vicilin, the degree of cross-reactivity between the

legumin and its allergenic homologues from other tree nuts is not established yet and warrants future investigation.

Among the low molecular weight proteins present in tree nuts and peanut, members of the prolamin superfamily, including 2S albumins and nsLTPs play an important role in allergic disease (Geiselhart et al., 2018). Especially the 2S albumins from hazelnut (Cor a 14), walnut (Jug r 1), cashew (Ana o 3), and from peanut (Ara h 2), have been shown to correlate with the severity of allergic reactions (Ballmer-Weber et al., 2019; Blazowski, Majak, Kurzawa, Kuna, & Jerzynska, 2019; Garnier, Massip, Viel, Bienvenu, & Bienvenu, 2014; Kukkonen, Pelkonen, Mäkinen-Kiljunen, Voutilainen, & Mäkelä, 2015). Sensitization to nsLTPs is associated with mild to severe symptoms and mostly restricted to distinct geographical regions (Ruano-Zaragoza et al., 2020). The analysis of our aqueus protein extract indicated low-level expression of nsLTP in macadamia nut, however the extract was prepared by defatting macadamia flour which could cause the loss of some lipophilic compounds. Interestingly, all low molecular weight IgE-binding proteins (10-20 kDa) were identified as different post-translationally processed forms of MiAMPs. There was no evidence indicating the presence of 2S albumin, similar to what has been observed for almond (Kabasser et al., 2021). This is in line with the observations reported by Rost et al. who previously reported not having identified any peptides specific for 2S albumin by mass-spectrometric analysis of the macadamia nut proteome (Rost et al., 2020). However, this analysis depends on the availability of a 2S albumin sequence from macadamia or a of close homologue in the database.

The prepared macadamia nut extract generated in this study and individual purified proteins were analysed for IgE reactivity. In our tested cohort, all of patients with confirmed macadamia nut allergy had sIgE to the whole protein extract indicating that most important IgE-binding

components were present in-solution. All five sera from macadamia allergic patients had IgE specific to Mac i 2 and four had IgE specific to Mac i 1. In a previous publication, the full-length vicilin precursor was expressed, and results from immunological assays indicated an IgE-binding frequency of 30% among macadamia nut allergic patients. A positive correlation between the severity of allergic reactions and specific IgE levels to the protein was reported (Ehlers et al., 2020). Instead of using the recombinant full-length precursor, we tested our patient cohort with purified post-translationally processed mature vicilin from the natural source. The higher sensitization frequency in our cohort may result from relevant conformational IgE epitopes formed during processing and folding of the mature protein. Three of the four patients with a positive IgE reaction to mature vicilin were co-sensitized to MiAMP2a suggesting an important IgE-binding role of this specific processing product. Based on our findings, the vicilin (Mac i 1) and the legumin (Mac i 2) were designated novel allergens by the WHO/IUIS Allergen Nomenclature Sub-Committee. MiAMP2a was additionally listed as Mac i 1.0101 (28-76) as an individual IgE-binding moiety derived from the vicilin precursor (www.allergen.org).

In contrast to the above, the IgE-sensitization rate to recombinant Mac i nsLTP was lower among macadamia nut allergic patients. This observation might be explained by the fact that patients' sera from an Australian cohort were used in this study. In general, nsLTP sensitization is mainly associated with the Mediterranean area (Asero, Piantanida, Pinter, & Pravettoni, 2018). In order to evaluate the relevance of sensitization to macadamia nsLTP, further studies are required to be conducted in Mediterranean region.

Finally, our data suggest that MiAMP2a is recognized explicitly by macadamia nut allergic patients, as reflected by the overall reduced measured IgE-reactivity signal ($OD405_{nm} < 0.1$) within the control groups of macadamia nut tolerant patients. Especially for tolerant patients with

a negative SPT to macadamia nut, IgE reactivity was below 10%. These results are in good accordance with the earlier study from Ehlers et al. showing that macadamia nut tolerant patients had almost negligible sIgE titers to the full-length precursor (Ehlers et al., 2020). Therefore, MiAMP2a may represent a marker of macadamia nut allergy with possible application in molecular diagnosis. However, given the monocentric study design and the relatively small number of serum samples available for macadamia nut allergic individuals, our findings need to be confirmed in future investigations including other populations and larger cohorts. Future studies will be required to assess the extent of cross-reactivity with other tree nut species and whether different sensitization patterns correlate with the severity of clinical manifestations of macadamia nut allergy. The well-characterized macadamia nut allergens described in this study and registered with the WHO/IUIS Allergen Nomenclature Sub-Committee will help in food allergy diagnosis and the development of patient-specific dietary recommendations.

CRediT authorship contribution statement

Stefan Kabasser: Conceptualization, Investigation, Visualization, Writing – Original Draft

Kunal Pratap: Methodology, Investigation

Sandip Kamath: Resources, Writing – Reviewing & Editing

Aya C Taki, Thanh Dang, Kirsten Perrett, Jennifer Koplin: Resources

Karin Hummel: Data curation, Formal analysis

Christian Radauer: Formal analysis; Writing – Review & Editing

Heimo Breiteneder: Writing – Review & Editing

Andreas L Lopata: Writing – Review & Editing

Merima Bublin: Supervision, Project administration, Funding acquisition, Writing – Review & Editing

Declaration of Competing Interest

The authors have no conflicts of interest to declare.

Acknowledgements

This work was supported by funds of the Oesterreichische Nationalbank (Austrian Central Bank), Anniversary Fund, project number: 17560, the Austrian Science Fund (FWF; Doctoral Program W1248-B30 [MCCA] and grant P 30936-B30), Sandip Kamath is an NHMRC Peter Doherty Early Career Research Fellow (GNT1124143), and Kunal Pratap receives PhD scholarship from Australian Institute of Tropical Health and Medicine (AITHM). This research was supported using resources of the VetCore Facility (Proteomics) of the University of Veterinary Medicine Vienna.

5 References

Aalberse, R. C., Mueller, G. A., Derksen, N. I. L., Aalberse, J. A., Edwards, L. L., Pomés, A., … Briza, P. (2020). Identification of the amino-terminal fragment of Ara h 1 as a major target of the IgE-binding activity in the basic peanut protein fraction. *Clinical and Experimental Allergy*, *50*(3), 401–405. https://doi.org/10.1111/cea.13554

Alasalvar, C., Salvadó, J. S., & Ros, E. (2020). Bioactives and health benefits of nuts and dried fruits. *Food Chemistry*, *314*, Article 126192. https://doi.org/10.1016/j.foodchem.2020.126192

Asero, R., Piantanida, M., Pinter, E., & Pravettoni, V. (2018). The clinical relevance of lipid transfer protein. *Clinical and Experimental Allergy*, *48*(1), 6–12. https://doi.org/10.1111/cea.13053

Ballmer-Weber, B. K., Lidholm, J., Lange, L., Pascal, M., Lang, C., Gernert, S., … Vieths, S. (2019). Allergen recognition patterns in walnut allergy are age dependent and correlate with the severity of allergic reactions. *Journal of Allergy and Clinical Immunology: In Practice*, *7*(5), 1560-1567. https://doi.org/10.1016/j.jaip.2019.01.029

Blazowski, L., Majak, P., Kurzawa, R., Kuna, P., & Jerzynska, J. (2019). Food allergy endotype with high risk of severe anaphylaxis in children—Monosensitization to cashew 2S albumin Ana o 3. *Allergy: European Journal of Allergy and Clinical Immunology*, *74*(10), 1945–1955. https://doi.org/10.1111/all.13810

Brough, H. A., Caubet, J. C., Mazon, A., Haddad, D., Bergmann, M. M., Wassenberg, J., … Eigenmann, P. A. (2020). Defining challenge-proven coexistent nut and sesame seed allergy: A prospective multicenter European study. *Journal of Allergy and Clinical*

Immunology, *145*(4), 1231–1239. https://doi.org/10.1016/j.jaci.2019.09.036

Bublin, M., Radauer, C., Knulst, A., Wagner, S., Scheiner, O., Mackie, A. R., … Breiteneder, H. (2008). Effects of gastrointestinal digestion and heating on the allergenicity of the kiwi allergens Act d 1, actinidin, and Act d 2, a thaumatin-like protein. *Molecular Nutrition and Food Research*, *52*(10), 1130–1139. https://doi.org/10.1002/mnfr.200700167

Buthelezi, N. M., Magwaza, L., & Tesfay, S. (2019). Postharvest pre-storage processing improves antioxidants, nutritional and sensory quality of macadamia nuts. *Scientia Horticulturae*, *251*, 197–208. https://doi.org/10.1016/j.scienta.2019.03.026

Center for the Promotion of Imports (CBI). (2021). The European market potential for macadamia nuts. Retrieved from CBI website: https://www.cbi.eu/market-information/processed-fruit-vegetables-edible-nuts/macadamia-nuts/market-potential. Accessed August 6, 2021

De Knop, K. J., Hagendorens, M. M., & Bridts, C. H. (2010). Macadamia nut allergy: 2 Case reports and a review of the literature. *Acta Clinica Belgica*, *65*(2), 129–132. https://doi.org/10.1179/acb.2010.026

Downs, M. L., Semic-Jusufagic, A., Simpson, A., Bartra, J., Fernandez-Rivas, M., Rigby, N. M., … Mills, E. N. C. (2014). Characterization of low molecular weight allergens from English walnut (*Juglans regia*). *Journal of Agricultural and Food Chemistry*, *62*(48), 11767–11775. https://doi.org/10.1021/jf504672m

Dubiela, P., Aina, R., Polak, D., Geiselhart, S., Humeniuk, P., Bohle, B., … Borowski, T. (2017). Enhanced Pru p 3 IgE-binding activity by selective free fatty acid-interaction. *Journal of Allergy and Clinical Immunology*, *140*(6), 1728-1731.e10.

https://doi.org/10.1016/j.jaci.2017.06.016

Ehlers, A. M., Rohwer, S., Otten, H. G., Brix, B., Le, T. M., Suer, W., & Knulst, A. C. (2020). IgE-binding to vicilin-like antimicrobial peptides is associated with systemic reactions to macadamia nut. *Clinical and Translational Allergy*, *10*(1), 1–5. https://doi.org/10.1186/s13601-020-00364-5

Ekbote, A., Hayman, G., & Bansal, A. (2010). Macadamia nut allergy: Potentially misleading specific IgE results. *Allergy: European Journal of Allergy and Clinical Immunology*, *65*(10), 1345. https://doi.org/10.1111/j.1398-9995.2010.02354.x

Garg, M. L., Blake, R. J., & Wills, R. B. H. (2003). Macadamia nut consumption lowers plasma total and LDL cholesterol levels in hypercholesterolemic men. *Journal of Nutrition*, *133*(4), 1060–1063. https://doi.org/10.1093/jn/133.4.1060

Garnier, L., Massip, C., Viel, S., Bienvenu, J., & Bienvenu, F. (2014). Sensitisation to Cor a 14 and Cor a 9 is a risk marker for severe hazelnut allergy in children. *Clinical and Translational Allergy*, *4*(S2), O15. https://doi.org/10.1186/2045-7022-4-s2-o15

Geiselhart, S., Hoffmann-Sommergruber, K., & Bublin, M. (2018). Tree nut allergens. *Molecular Immunology*, *100*, 71–81. https://doi.org/10.1016/j.molimm.2018.03.011

Gutiérrez, A. M., Sotillo, J., Schlosser, S., Hummel, K., & Miller, I. (2019). Towards understanding non-infectious growth-rate retardation in growing pigs. *Proteomes*, *7*(3), 31. https://doi.org/10.3390/PROTEOMES7030031

Herbst, R. A., Wahl, R., & Frosch, P. J. (2010). Specific IgE reactivity and identification of potential allergens in macadamia allergy. *Journal of the European Academy of Dermatology*

and Venereology, *24*(11), 1361–1363. https://doi.org/10.1111/j.1468-3083.2010.03642.x

International Nut & Dried Fruit Council. (2020). *Nuts & Dried Fruits - Statistical Yearbook 2019-2020*. Retrieved from https://www.nutfruit.org/files/tech/1587539172_INC_Statistical_Yearbook_2019-2020.pdf Accessed August 6, 2021

Kabasser, S., Hafner, C., Chinthrajah, S., Sindher, S. B., Kumar, D., Kost, L. E., … Bublin, M. (2021). Identification of Pru du 6 as a potential marker allergen for almond allergy. *Allergy: European Journal of Allergy and Clinical Immunology*, *76*(5), 1463–1472. https://doi.org/10.1111/all.14613

Koplin, J. J., Wake, M., Dharmage, S. C., Matheson, M., Tang, M. L. K., Gurrin, L. C., … Chhabra, S. (2015). Cohort Profile: The HealthNuts Study: Population prevalence and environmental/genetic predictors of food allergy. *International Journal of Epidemiology*, *44*(4), 1161–1171. https://doi.org/10.1093/ije/dyu261

Kukkonen, A. K., Pelkonen, A. S., Mäkinen-Kiljunen, S., Voutilainen, H., & Mäkelä, M. J. (2015). Ara h 2 and Ara 6 are the best predictors of severe peanut allergy: A double-blind placebo-controlled study. *Allergy: European Journal of Allergy and Clinical Immunology*, *70*(10), 1239–1245. https://doi.org/10.1111/all.12671

Kumar, G., Hummel, K., Ahrens, M., Menanteau-Ledouble, S., Welch, T. J., Eisenacher, M., … El-Matbouli, M. (2016). Shotgun proteomic analysis of *Yersinia ruckeri* strains under normal and iron-limited conditions. *Veterinary Research*, *47*(1), 1–13. https://doi.org/10.1186/s13567-016-0384-3

Kumar, G., Menanteau-Ledouble, S., Saleh, M., & El-Matbouli, M. (2015). *Yersinia ruckeri*, the

causative agent of enteric redmouth disease in fish. *Veterinary Research*, *46*(1), 103. https://doi.org/10.1186/s13567-015-0238-4

Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., & Lajoie, G. (2003). PEAKS: Powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *17*(20), 2337–2342. https://doi.org/10.1002/rcm.1196

Marcus, J. P., Green, J. L., Goulter, K. C., & Manners, J. M. (1999). A family of antimicrobial peptides is produced by processing of a 7S globulin protein in *Macadamia integrifolia* kernels. *Plant Journal*, *19*(6), 699–710. https://doi.org/10.1046/j.1365-313X.1999.00569.x

Masthoff, L. J. N., Mattsson, L., Zuidmeer-Jongejan, L., Lidholm, J., Andersson, K., Akkerdaas, J. H., … Pasmans, S. G. M. A. (2013). Sensitization to Cor a 9 and Cor a 14 is highly specific for a hazelnut allergy with objective symptoms in Dutch children and adults. *Journal of Allergy and Clinical Immunology*, *132*(2), 393–399. https://doi.org/10.1016/j.jaci.2013.02.024

McWilliam, Peters, R., Tang, M. L. K., Dharmage, S., Ponsonby, A. L., Gurrin, L., … Robertson, C. (2019). Patterns of tree nut sensitization and allergy in the first 6 years of life in a population-based cohort. *Journal of Allergy and Clinical Immunology*, *143*(2), 644-650. https://doi.org/10.1016/j.jaci.2018.07.038

McWilliam, V. L., Koplin, J. J., Field, M. J., Sasaki, M., Dharmage, S. C., Tang, M. L. K., … Allen, K. J. (2018). Self-reported adverse food reactions and anaphylaxis in the SchoolNuts study: A population-based study of adolescents. *Journal of Allergy and Clinical Immunology*, *141*(3), 982–990. https://doi.org/10.1016/j.jaci.2017.09.012

Müntz, K. (1998). Globulins from legume seeds: Structure and function during storage and reactivation. In *Plant Proteins from European Crops* (pp. 3–12). https://doi.org/10.1007/978-3-662-03720-1_1

Osborne, N. J., Koplin, J. J., Martin, P. E., Gurrin, L. C., Lowe, A. J., Matheson, M. C., … Allen, K. J. (2011). Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. *Journal of Allergy and Clinical Immunology*, *127*(3), 668-676. https://doi.org/10.1016/j.jaci.2011.01.039

Pfeifer, S., Bublin, M., Dubiela, P., Hummel, K., Wortmann, J., Hofer, G., … Hoffmann-Sommergruber, K. (2015). Cor a 14, the allergenic 2S albumin from hazelnut, is highly thermostable and resistant to gastrointestinal digestion. *Molecular Nutrition and Food Research*, *59*(10), 2077–2086. https://doi.org/10.1002/mnfr.201500071

Rost, J., Muralidharan, S., & Lee, N. A. (2020). A label-free shotgun proteomics analysis of macadamia nut. *Food Research International*, *129*, Article 108838. https://doi.org/10.1016/j.foodres.2019.108838

Ruano-Zaragoza, M., Somoza, M. L., Jiménez-Rodriguez, T. W., Soriano-Gomis, V., González-Delgado, P., Esteban-Rodriguez, A., … Fernández-Sánchez, J. (2021). Lipid transfer protein sensitization: Risk of anaphylaxis and molecular sensitization profile in Pru p 3-sensitized patients. *International Archives of Allergy and Immunology*, 182:425–432. https://doi.org/10.1159/000511977

Sasaki, M., Koplin, J. J., Dharmage, S. C., Field, M. J., Sawyer, S. M., McWilliam, V., … Allen, K. J. (2018). Prevalence of clinic-defined food allergy in early adolescence: The SchoolNuts study. *Journal of Allergy and Clinical Immunology*, *141*(1), 391-398.

https://doi.org/10.1016/j.jaci.2017.05.041

Sutherland, M. F., O'Hehir, R. E., Czarny, D., & Suphioglu, C. (1999). Macadamia nut anaphylaxis: Demonstration of specific IgE reactivity and partial cross-reactivity with hazelnut. *Journal of Allergy and Clinical Immunology*, *104*(4 I), 889–890. https://doi.org/10.1016/S0091-6749(99)70304-0

van der Valk, J. P. M., Gerth van Wijk, R., Vergouwe, Y., Steyerberg, E. W., Reitsma, M., Wichers, H. J., … de Jong, N. W. (2017). sIgE Ana o 1, 2 and 3 accurately distinguish tolerant from allergic children sensitized to cashew nuts. *Clinical and Experimental Allergy*, *47*(1), 113–120. https://doi.org/10.1111/cea.12794

Yoshida, K., Shirane, S., Kinoshita, K., Morikawa, E., Matsushita, S., Toda, M., … Narita, M. (2021). Macadamia nut allergy in children: Clinical features and cross-reactivity with walnut and hazelnut. *Pediatric Allergy and Immunology*, *32*(5), 1111–1114. https://doi.org/10.1111/pai.13469

Tables

Table 1. Characteristics of macadamia nut allergic and tolerant patients: clinical symptoms and IgE sensitization as determined by SPT

Patients no.	Age (yrs)/ sex	Macadamia nut allergy	Symptoms to macadamia nut (reported at 6 yrs)	Macadamia nut SPT wheal at 6 yrs (Ø mm)	Allergy to other foods at 4-6 yrs
Patients with macadamia nut allergy (MA 1-5)					
MA 1	6.1/M	yes	Current avoidance, last reaction >12months	8.0	Peanut, cashew, hazelnut, pistachio
MA 2	6.3/M	yes	Current avoidance, last reaction >12months	15.0	Hazelnut, walnut
MA 3	4.4/M	yes	Current avoidance, last reaction >12months	10.0	Cashew, pistachio
MA 4	6.1/F	yes	Current avoidance, last reaction >12months	11.0	Peanut, cashew, hazelnut, Brazil nut
MA 5	4.2/M	yes	Current avoidance, last reaction >12months	10.0	Peanut
Patients tolerant to macadamia nut with SPT 3-7 mm (MT 1-8)					
MT 1	6.1/F	no	AS	2.5	Peanut
MT 2	4.2/M	no	AS	5.0	Peanut, cashew, pistachio
MT 3	6.2/M	no	AS	4.0	Peanut
MT 4	4.1/F	no	AS	3.0	Cashew, pistachio
MT 5	6.2/M	no	AS	4.5	Peanut
MT 6	1.1/M	no	AS	4.0	Peanut, cashew, pistachio, walnut
MT 7	4.1/M	no	AS	3.0	Pecan, sesame
MT 8	1.2/F	no	AS	2.5	Hazelnut, pistachio, walnut
Patients tolerant to macadamia nut with SPT ≤ 2.0 mm (MT 9-22)					
MT 9	4.1/M	no	AS	0.0	Cashew
MT 10	4.1/M	no	AS	0.0	Peanut
MT 11	4.1/F	no	AS	0.0	Peanut, cashew, almond
MT 12	4.1/F	no	AS	0.0	No other food allergies
MT 13	6.1/M	no	AS	0.0	Cashew, walnut
MT 14	4.2/F	no	AS	0.0	Peanut, pistachio
MT 15	4.2/M	no	AS	0.0	Peanut, hazelnut, sesame
MT 16	4.1/F	no	AS	0.0	Hazelnut
MT 17	4.1/M	no	AS	0.0	No other food allergies
MT 18	4.1/M	no	AS	0.0	Cashew
MT 19	4.2/M	no	AS	0.0	Peanut, cashew, pistachio
MT 20	6.2/F	no	AS	0.0	Cashew
MT 21	4.1/M	no	AS	0.0	Peanut, almond, pistachio
MT 22	4.1/M	no	AS	0.0	Peanut, cashew, pistachio

*AS= asymptomatic

Figure Legends

Figure 1. IgE immunoblot (a) and Coomassie-stained SDS-PAGE of macadmia nut extract and purified proteins (b). a) IgE reactivity to extract was tested under non-reducing (-DTT) and reducing (+DTT) conditions using pooled sera (serum pool) from 3 macadamia nut allergic patients (MA 1, 2 and 4, Table 1) and pooled sera from 2 healthy control patients (NHS). b) Macadamia nut extract and purified proteins visualized by non-reducing and reducing Coomassie-stained 15% SDS-PAGE.

Figure 2. Purification of macadamia nut allergens. a) Size exclusion chromatography of crude macadamia nut extract. Peak fractions indicated by roman numbers were collected and pooled. b) and c) Purification of Mac i 1. The third pool (III) from SEC was separated by anion exchange chromatography (b). The first pool (I) from anion exchange chromatography was further subjected to Con A affinity chromatography (c). d) Purification of MiAMPs by anion exchange chromatography. e) Purification of rMac i nsLTP by cation exchange chromatography. All fractions were analysed by Coomassie-stained 15% SDS-PAGE under reducing conditions. Bidirectional arrows indicate fractions used for further purification and/or protein characterization. Italic letters (a-d) with vertical bars indicate retention times of Blue Dextran and protein standards (a: Blue Dextran, b: β-amylase, c: alcohol dehydrogenase, d: bovine serum albumin).

Figure 3. Far-UV circular dichroism (CD) spectroscopy of Mac i 1, Mac i 2, MiAMP2a and recombinant macadamia nsLTP.

Figure 4. ELISA analysis of IgE binding to macadamia nut extract (a), Mac i 1 (b), Mac i 2 (c), MiAMP2a/Mac i 1 (28-76) (d), and Mac i nsLTP (e). IgE binding to purified components was

analysed using sera from macadamia nut allergic (black bars), macadamia nut tolerant but sensitized (dark-grey bars) and peanut/tree nut allergic individuals tolerant and not sensitized to macadamia nut (light-grey bars) individuals. Sera were counted positive if they exceeded the mean OD value of four healthy controls plus three times their standard deviation.

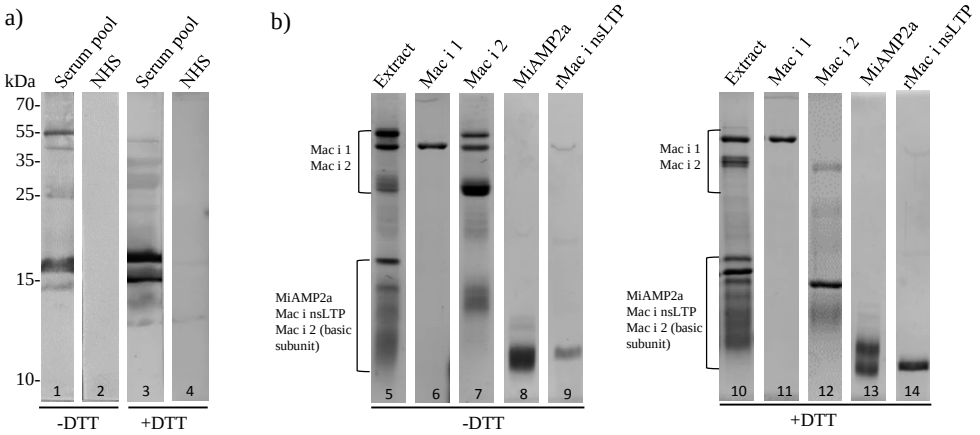

Figure 1

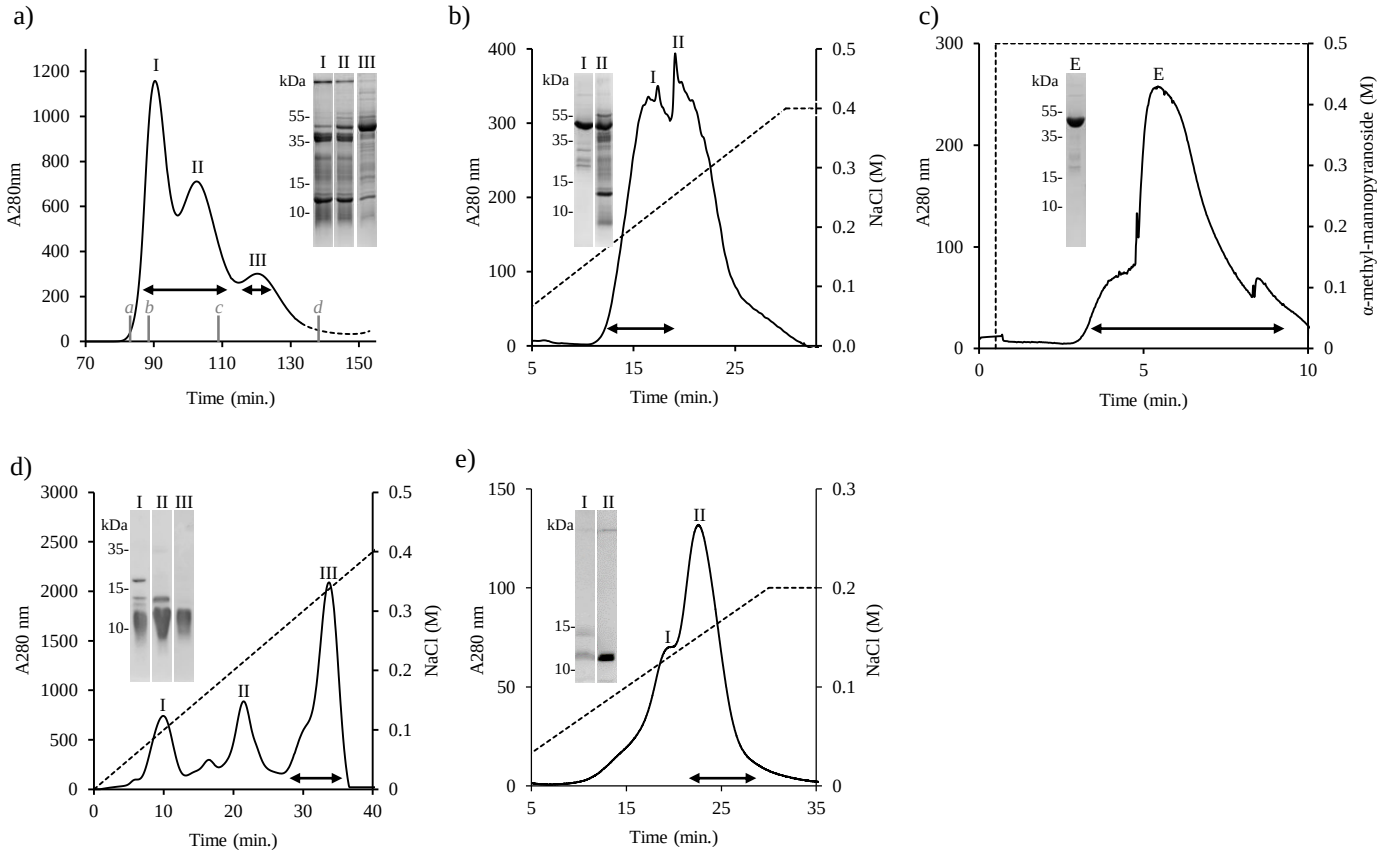

Figure 2

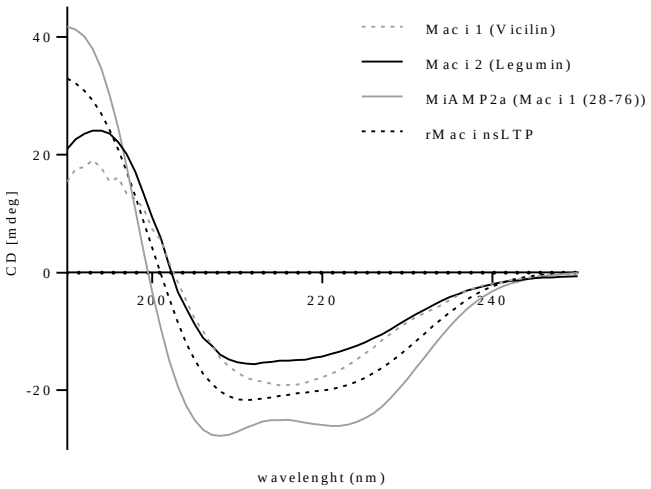

Figure 3

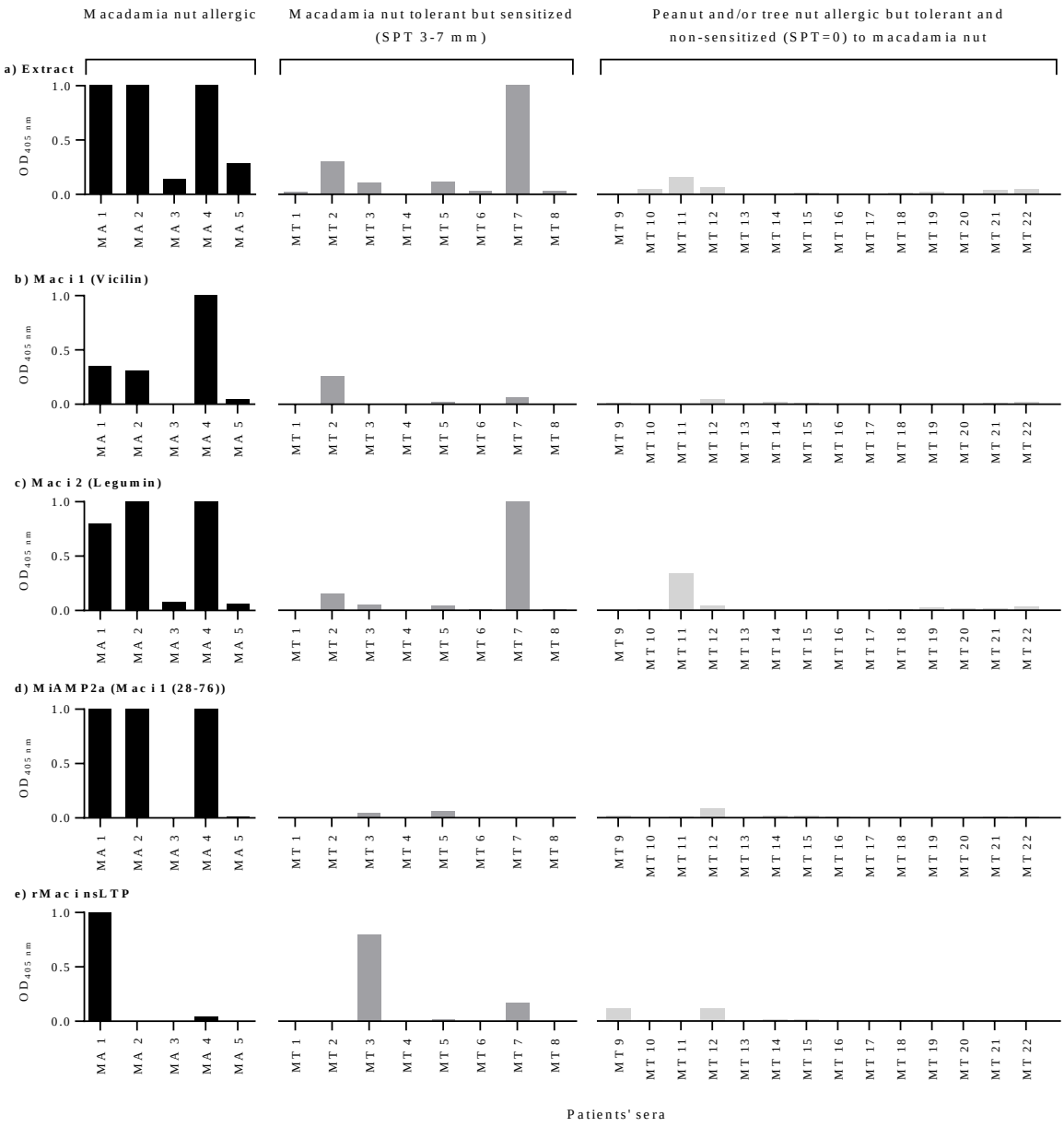

Figure 4

Supplementary material

Identification of vicilin, legumin and antimicrobial peptide 2a as macadamia nut allergens

Stefan Kabasser[a], Kunal Pratap[b, c, d], Sandip Kamath[b, c, d], Aya C Taki[e], Thanh Dang[f], Jennifer Koplin[f], Kirsten Perrett[f], Karin Hummel[g], Christian Radauer[a], Heimo Breiteneder[a], Andreas L Lopata[b, c, d,*], Merima Bublin[a,*]

[a]Institute of Pathophysiology and Allergy Research, Center of Pathophysiology, Infectiology and Immunology, Medical University of Vienna, Vienna, Austria

[b]Molecular Allergy Research Laboratory, College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, QLD, Australia

[c]Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, QLD, Australia

[d]Center for Molecular Therapeutics, James Cook University, Townsville, QLD, Australia

[e]School of Veterinary Biosciences, Melbourne Veterinary School, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, VIC, Australia

[f]Department of Paediatrics, Murdoch Children's Research Institute, The University of Melbourne, Flemington Road, Parkville, VIC, Australia

[g]VetCore Facility for Research, University of Veterinary Medicine, Vienna, Austria

***Corresponding authors:**

Merima Bublin, PhD

Institute of Pathophysiology and Allergy Research

Center of Pathophysiology, Infectiology and Immunology

Medical University of Vienna

Waehringer Guertel 18-20, 1090 Vienna, Austria

E-mail: merima.bublin@meduniwien.ac.at

Andreas L. Lopata, PhD

Pharmacy and Medical Research Building

James Cook University

Townsville, QLD, Australia

Email: andreas.lopata @jcu.edu.au

Table S1. List of *M. integrifolia* peptides identified by *de-novo* sequencing and Nano-LC ESI Orbitrap MS/MS

a) *De-novo* sequencing of *M. integrifolia* legumin (Mac i 2) (Figure 1b, lane 7)

No.	Identified Peptide
1	LLVAQCR
2	LQQISVSQPR
3	QRIQSEGGVTEFWDENEDQFQCTGVAAMR
4	NIIQPNSLSLPNYSPSPR
5	LVYIER
6	GLLGVTFPGCPETYQSSRDEQSYR
7	APGKMLVVLPAGVAHWCLNDGK
8	EDLVAVSVNNLNNQANQLNQK
9	SYYLAGSQNQESQR
10	LCHGSWSNTYQNLLSPFNQNLMADALNVDVETVR
11	LHHYLDNPR
12	LDGGPQLHAGPHWLMNAHSLFYLTR
13	LKDANVFVVPR
14	MGPAQVLAQSYKSFAGEAQNLK

b) Mac i 1: protein band (Figure 1b, lane 6) identified as Q9SPL4 (aa 181-661)

No.	Identified Peptide
1	DNRESYNLECGDVIR
2	EAIVVPVGHPVVFVSSGNENLLLFAFGINAQNNHENFLAGR
3	EFFPAGGQNPEPYLSTFSK
4	EGVIISASQEQIR
5	EILEAALNTQAER
6	ELTRDDSESRR
7	ERNVLQQIEPQAMELAFAAPR
8	ESYNLECGDVIR
9	EVEELFNSQDESIFFPGPR
10	FLQTISTPGQYK
11	FLQTISTPGQYKEFFPAGGQNPEPYLSTFSK
12	FRTEEGHISVLENFYGR
13	GGESSRGPYNLFNK
14	GGSGRYEEGEEKQSDNPYYFDER
15	GPYNLFNK
16	GPYNLFNKRPLYSNK
17	KEVEELFNSQDESIFFPGPR
18	LHIAKFLQTISTPGQYK
19	IPAGTTFYLINR

20	LRGVLGQQR
21	LVLLEANPNAFVLPTHLDADAILLVTGGR
22	NVLQQIEPQAMELAFAAPR
23	NVLQQIEPQAMELAFAAPRK
24	NYRLVLLEANPNAFVLPTHLDADAILLVTGGR
25	QSDNPYYFDER
26	REAIVVPVGHPVVFVSSGNENLLLFAFGINAQNNHENFLAGR
27	RGGESSRGPYNLFNK
28	RHEEEEDVHYEQVK
29	RHEEEEDVHYEQVKAR
30	RPLYSNK
31	TEEGHISVLENFYGR
32	VVVVASGEADVEMACPHLSGR
33	YEEGEEKQSDNPYYFDER
34	YGQAYEVKPEDYR
35	DNRESYNLECGDVIR

c) MiAMP2a: protein band (Figure 1b, lane 8 and Figure 2d, peak III) identified as Q9SPL4 (aa 28-76)

No.	Identified Peptide
1	FEEDIDWSK
2	QCMQLETSGQMR
3	QEYEECKR
4	RCVSQCDKR
5	RFEEDIDWSK
6	RQCMQLETSGQMR

d) MiAMP2a-d: protein band (Figure 2d, peak I, 17 kDa band) identified as Q9SPL3 (aa 1-179)

No.	Identified Peptide
1	QCMQLETSGQMR
2	DPQQQYEQCQK
3	LQYQCQR

e) MiAMP2a-d: protein band (Figure 2d, peak I, 13 kDa band) identified as Q9SPL3 (aa 1-179)

No.	Identified Peptide
1	QCMQLETSGQMR
2	FEEDIDWSK
3	QQQYCQR

4	CKEICEEEEEYNRQRDPQQQYEQCQK
5	LQYQCQR

f) MiAMP2a: protein spot A (Figure S2c) identified as Q9SPL4 (aa 28-76)

No.	Identified Peptide
1	FEEDIDWSK
2	QCMQLETSGQMR
3	QEYEECKR

g) MiAMP2a: protein spot B (Figure S2c) identified as Q9SPL4 (aa 28-76)

No.	Identified Peptide
1	FEEDIDWSK
2	QCMQLETSGQMR
3	QEYEECKR
4	RCVSQCDKR
5	RFEEDIDWSK
6	RQCMQLETSGQMR

h) MiAMP2a: protein spot C (Figure S2c) identified as Q9SPL4 (aa 28-76)

No.	Identified Peptide
1	QCMQLETSGQMR

i) protein spot D (Figure S2c): no peptide detected

j) Recombinant nsLTP: protein band (Figure 1b, lane 9) identified as LR861101

No.	Identified Peptide
1	CGVNLPYK
2	LAPCLTYLR
3	NAYNSISGINAAYAGGLPAK
4	NAYNSISGINAAYAGGLPAKCGVNLPYK
5	QTACGCLKNAYNSISGINAAYAGGLPAK
6	SGGAVPGTCCNAVK
7	SGGAVPGTCCNAVKNLNNSAK
8	TTPDRQTACGCLK

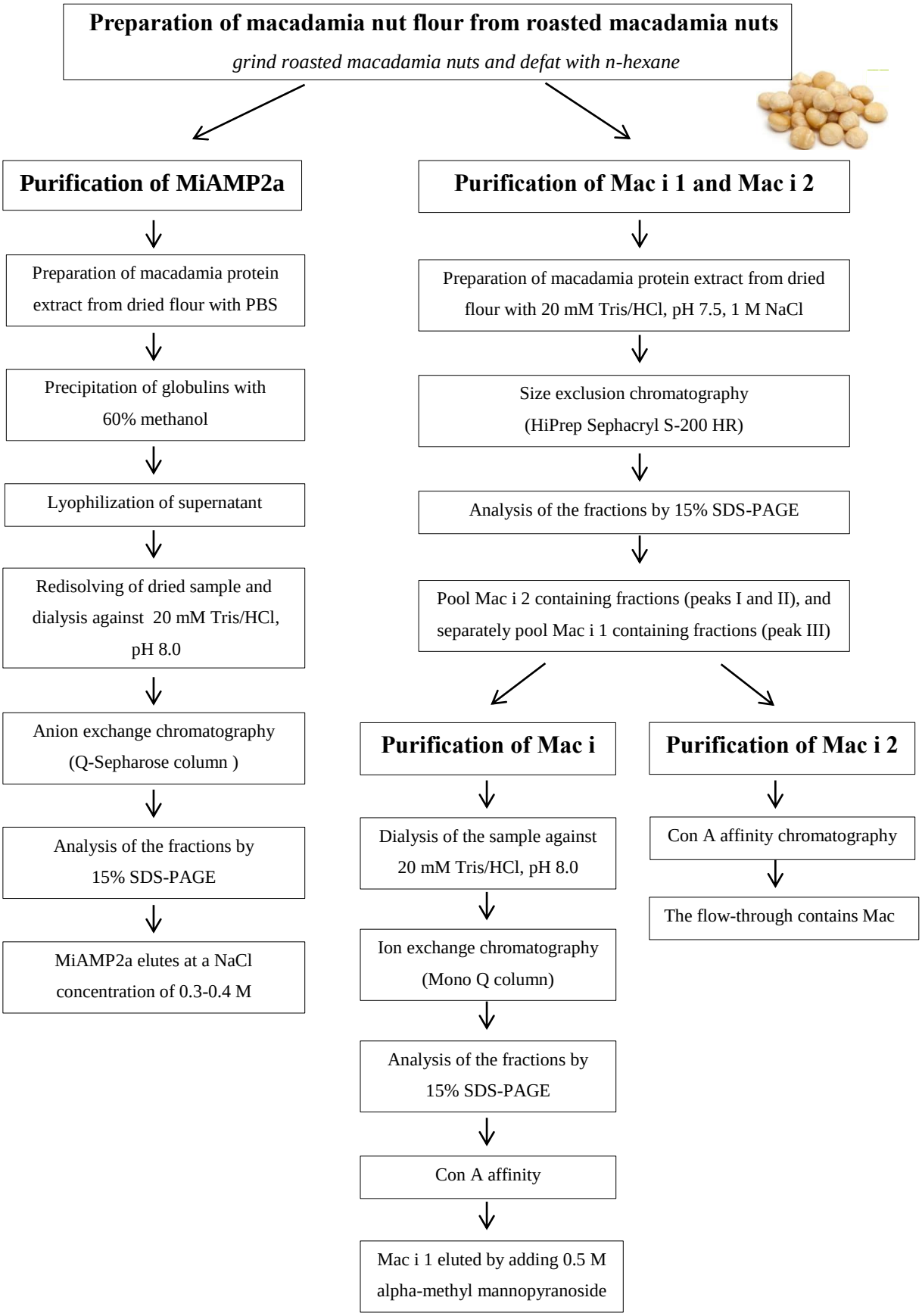

Figure S1. Flow chart of macadamia nut allergen extraction and purification

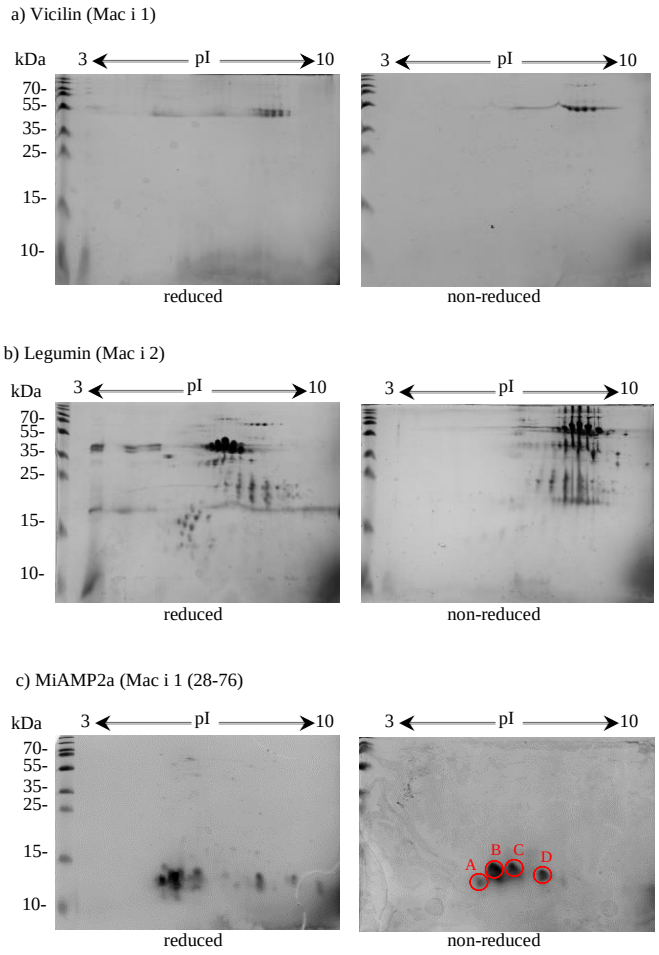

Figure S2. Coomassie-stained 2D-PAGE of purified macadamia proteins analysed under reducing (left) and non-reducing (right) conditions. Protein spots encircled in red were analysed by tandem mass spectrometry.

(a) Mac i 1 (Q9SPL4; aa 181-661)

181 KEEDNKRDPQ QREYEDCRRR CEQQEPRQQY QCQRRCREQQ RQHGRGGDLI NPQR**GGSGRY**

241 **EEGEEKQSDN PYYFDERSLS TRFRTEEGHI SVLENFYGRS KLLRALKNYR LVLLEANPNA**

301 **FVLPTHLDAD AILLVTGGR**G ALKMIHR**DNR ESYNLECGDV IRIPAGTTFY LINRDNNERL**

361 **HIAKFLQTIS TPGQYKEFFP AGGQNPEPYL STFSKEILEA ALNTQAERLR GVLGQQREGV**

421 **IISASQEQIR ELTRDDSESR** RWHIR**RGGES SRGPYNLFNK RPLYSNKYGQ AYEVKPEDYR**

481 QLQDMDVSVF IA[N]ITQGSMM GPFFNTR**STK VVVVASGEAD VEMACPHLSG RHGGRRGGKR**

541 **HEEEEDVHYE QVKAR**LSK**RE AIVVPVGHPV VFVSSGNENL LLFAFGINAQ NNHENFLAGR**

601 **ERNVLQQIEP QAMELAFAAP RKEVEELFNS QDESIFFPGP R**QHQQQSSRS TKQQQPLVSI

661 LDFVGF

(b) MiAMP2a (Mac i 1 (28-76)) (Q9SPL4; aa 1-240)

1 MAINTSNLCS LLFLLSLFLL STTVSLAESE FDR**QEYEECK RQCMQLETSG QMRRCVSQCD**
AMP-2a

61 **KRFEEDIDWS K**YDNQDDPQT DCQQCQRRCR QQESGPRQQQ YCQRRCKEIC EEEEEYNRQR
AMP-2b

121 DPQQQYEQCQ ERCQRHETEP RHMQTCQQRC ERRYEKEKRK QQKRYEEQQR EDEEKYEERM
AMP-2c

181 KEEDNKRDPQ QREYEDCRRR CEQQEPRQQY QCQRRCREQQ RQHGRGGDLI NPQRGGSGRY

(c) Mac i nsLTP (LR861101)

1 MANSGVMKLV CLVLACMVVA APLAEAAITC GQVVSK**LAPC LTYLRSGGAV PGTCCNAVKN**
Signal peptide

61 **LNNSAKTTPD RQTACGCLKN AYNSISGINA AYAGGLPAKC GVNLPYK**ISP SINCATYTLS

121 LYNF

Figure S3. Mass spectrometric identification of macadamia proteins. Bold letters indicate peptides identified by Nano-LC ESI Orbitrap MS/MS. Macadamia antimicrobial peptides (MiAMPs) encoded in Q9SPL4 are underlined. Amino acids in boxes represent possible N-linked glycosylation sites identified using the *NetNGlyc 1.0 Server*. The signal peptide of Mac i nsLTP was predicted with the *SignalP-5.0 Server*.

```
M. integrifolia peptides ------------------------------------------LLVAQCRLQQISVSQPRQ 18
A. europaeum legumin     ---MARSLLLLPITLSLLLLSAYA------IGEEQQYSHRRLRDAHQCRIQHLSASRPSQ 51
P. somniferum legumin    ---MAKAIPFVTLIIFFLVSTAFAQYQTGQQGQQYQQQQIQQQEARQCRIQRITPSQPTQ 57
M. cordata legumin       MAITKKPLPLI-ILAFFLISVAVAQRETS------PSQQTRLQEARQCRLQRLTASQPNQ 53
                                                          . ***:*::: *:* *

M. integrifolia peptides RIQSEGGVTEFWDENEDQFQCTGVAAMRNIIQPNSLSLPNYSPSPRLVYIER--GLLGVT 76
A. europaeum legumin     RIESEGGVTELWNQNEEQFQCAGVAAMRNIIEPNSLSLPNYSPSPRLVYIQQGKGLLGMS 111
P. somniferum legumin    RIQSEGGVTEMWNEFDDQFQCAGVAALRNIIQPNCLSLPNFSPSPRLVYIQQGQGLIGMS 117
M. cordata legumin       RIESEGGVTELWNEYEDQFQCAGVAAIRNIIQPNSLSLPNFSPSPRLVYIQQGQGLLGLS 113
                         **:*******:*:: ::****:****:****:**.*****:*********::  **:*::

M. integrifolia peptides FPGCPETYQSSRDEQSY----------------------RAPGKMLVVLPAGVAHWCLN 113
A. europaeum legumin     YPGCAEAYHSSGRPSYQI-----VRGEQQQQRGEQHQKVHRIRQGDIVALPAGVAYWCYN 166
P. somniferum legumin    IPGCAETYHSGQQSVR---GGSIRTR--QGEQSDQHQKVHRVRQGDIVAVPAGVAHWCYN 172
M. cordata legumin       QPGCAETYHSGQQYKSPREGSSPRGQQGQQRQRDQHQKVHRIRQGDIVALPAGVAHWCYN 173
                          *** *:*:*.                          *     :*.:*****:** *

M. integrifolia peptides DGKEDLVAVSVNNLNNQANQLNQK--SYYLAGSQNQESQR----LCHGSWSNTYQNLLSP 167
A. europaeum legumin     DGDEKLVALSITDLSSQANQLNQMPRSFYLAGGQPKQSAG--SIHTRGSQDYNAENILRA 224
P. somniferum legumin    DGNEELVAVSVNDLNNNANQLEQKLRSFYLAGGHQQIYSSVQSPQQQQQQKETFQNVFRA 232
M. cordata legumin       DGNEQLVAVSVSDFKNNANQLEQKLRSFYLAGGQPQR--SMQSQQQQAQQRETFQNVFRA 231
                         **.*.***:*:.::..:****:*   *:****.: :          : .   . :*::

M. integrifolia peptides FNQNLMADALNVDVETVR------------------------------------------ 185
A. europaeum legumin     FDANIMAEAFDVSMDTVRKMQREDERGFIVKADREAMSMIRPDEEEEE-------YKNGM 277
P. somniferum legumin    FDENLMAEAFNLPVEVIRRMQQEDQRGFIVRVTGEGMRMIRPEEDEEYEQRYQSRRENGM 292
M. cordata legumin       FDENLMAEAFNIPVDIVRKMQQEDERGLIVKVKGEGMRMIRPEEEEEYEHSYQLRRQNGL 291
                         *: *:**:*::: :: :*

M. integrifolia peptides --------LHHYLDNPRLDG--------------------------------GPQLHAGP 205
A. europaeum legumin     EEAYCNLRINQYLDNPREADVYSRQAGRLNSVNMKKLPILRYLEMRDEKGSLYQNSIFAP 337
P. somniferum legumin    EELYCNLKIRQYLDNPKEADIYTEQAGRINHANGQKLQILNYMDMSAEKGNLYPNAMHAP 352
M. cordata legumin       EEAYCNLKIRQYLDNPKEADVYSRQAGRLNIVNSQKLQILNYMGMSAEKGNLYPNAMHAP 351
                                :.:*****:  .                                   :   .*

M. integrifolia peptides HWLMNAHSLFYLTRLKD-------------------ANVFVVPRMG-------------- 232
A. europaeum legumin     HWTMNAHTIHYVTRGRGRVQVVGHDGEKVLDARVNEGDMFVVPQYFAVMKQADKNGLEWV 397
P. somniferum legumin    HWTMNAHSVFYVTRGDAHVEIVGSNGQTVYYDHVKQGELFVVPQQFVSSMRAGNNGFEYV 412
M. cordata legumin       HWSMNAHSIVYVTRGEAHVQVVGSNGQTVLDDKVNQGDLFVVPQYFVSTLRAGNNGFEYV 411
                         ** ****:: *:**                     .::****:

M. integrifolia peptides ----------------------PAQVLAQSYKSFAGEAQNLK----------------- 252
A. europaeum legumin     SIKTSALPMRSPLVGSTSAIKGMPIQVLTNSYRISKSEAQNLKYNRERHVMLLPPPTTRS 457
P. somniferum legumin    AFKTSSQPMKSPLVGYTSAFKAMPIQVLANSFQISTQEAQSLKYNREHQTMLLPLRNVRS 472
M. cordata legumin       AFKTSGQPMKSPLVGYTSAFKAMPIQVLTNSFQISSQEAQSLKYNREHHTMLLPPRTGSS 471
                                       * ***::*::    ***.**

M. integrifolia peptides -
A. europaeum legumin     S 458
P. somniferum legumin    S 473
M. cordata legumin       S 472
```

Figure S4. Sequence alignment of legumin seed storage proteins from different species. Accession numbers: CAA64761.1 *(Asarum europaeum legumin)*, XP_026460168.1 *(Papaver somniferum legumin)*, OVA02407.1 *(Macleaya cordata legumin)*. Peptides of *Macadamia integrifolia* legumin were obtained by de-novo sequencing. Note that amino acids Leu/Iso cannot be distinguished by de-novo sequencing.