



Improving the Identification and Coverage of Plant Transmembrane Proteins in *Medicago* Using Bottom–Up Proteomics

Han Chung Lee¹, Adam Carroll², Ben Crossett³, Angela Connolly³, Amani Batarseh⁴ and Michael A. Djordjevic^{1*}

¹ Division of Plant Sciences, Research School of Biology, College of Medicine, Biology and the Environment, The Australian National University, Canberra, ACT, Australia, ² ANU Joint Mass Spectrometry Facility, Research School of Chemistry, College of Science, The Australian National University, Canberra, ACT, Australia, ³ Sydney Mass Spectrometry, The University of Sydney, Sydney, NSW, Australia, ⁴ BCAL Diagnostics, National Innovation Centre, Eveleigh, NSW, Australia

OPEN ACCESS

Edited by:

Wei Wang,
Henan Agricultural University, China

Reviewed by:

Rui-Cai Long,
Chinese Academy of Agricultural
Sciences, China
Pitter F. Huesgen,
Helmholtz-Verband Deutscher
Forschungszentren (HZ), Germany

*Correspondence:

Michael A. Djordjevic
michael.djordjevic@anu.edu.au

Specialty section:

This article was submitted to
Plant Proteomics and Protein
Structural Biology,
a section of the journal
Frontiers in Plant Science

Received: 17 August 2020

Accepted: 12 November 2020

Published: 18 December 2020

Citation:

Lee HC, Carroll A, Crossett B,
Connolly A, Batarseh A and
Djordjevic MA (2020) Improving
the Identification and Coverage
of Plant Transmembrane Proteins
in *Medicago* Using Bottom–Up
Proteomics.
Front. Plant Sci. 11:595726.
doi: 10.3389/fpls.2020.595726

Plant transmembrane proteins (TMPs) are essential for normal cellular homeostasis, nutrient exchange, and responses to environmental cues. Commonly used bottom–up proteomic approaches fail to identify a broad coverage of peptide fragments derived from TMPs. Here, we used mass spectrometry (MS) to compare the effectiveness of two solubilization and protein cleavage methods to identify shoot-derived TMPs from the legume *Medicago*. We compared a urea solubilization, trypsin Lys-C (UR-TLC) cleavage method to a formic acid solubilization, cyanogen bromide and trypsin Lys-C (FA-CTLC) cleavage method. We assessed the effectiveness of these methods by (i) comparing total protein identifications, (ii) determining how many TMPs were identified, and (iii) defining how many peptides incorporate all, or part, of transmembrane domains (TMD) sequences. The results show that the FA-CTLC method identified nine-fold more TMDs, and enriched more hydrophobic TMPs than the UR-TLC method. FA-CTLC identified more TMPs, particularly transporters, whereas UR-TLC preferentially identified TMPs with one TMD, particularly signaling proteins. The results suggest that combining plant membrane purification techniques with both the FA-CTLC and UR-TLC methods will achieve a more complete identification and coverage of TMPs.

Keywords: transmembrane protein, transmembrane domain, liquid chromatography, mass spectrometry, detergent-free purification, *Medicago truncatula*, TRAMDOMI algorithm

INTRODUCTION

Transmembrane proteins (TMPs) play critical roles in the function of all living organisms. Approximately 20 to 30% of the sequenced genomes from microbial or eukaryotic organisms encode TMPs (Komatsu et al., 2007; Schey et al., 2013; Almeida et al., 2017). In humans, roughly 50% of all known drug targets are TMPs (Hopkins and Groom, 2002; Lappano and Maggiolini, 2011; Tautermann, 2014), and many human diseases are caused by malfunctioning TMPs (Hardy, 2017; Hasegawa et al., 2017; Hattori et al., 2017; Szablewski, 2017). In plants, TMPs play important roles in solute transport (Pellizzaro et al., 2015), signal recognition and transduction (Kemmerling et al., 2011; Mohd-Radzman et al., 2015; Song et al., 2015; Imin et al., 2018), growth and

development (Clark et al., 1997; van der Knaap et al., 1999; Osakabe et al., 2005; ten Hove et al., 2011), and photosynthesis (Liu and Last, 2015; Okumura et al., 2016).

TMPs can be difficult to identify by proteomic strategies due to their high hydrophobicity (Seddon et al., 2004; Carpenter et al., 2008; Rawlings, 2016) and low abundance (Vit and Petrak, 2017). All proteomic strategies aimed at identifying TMPs begin with membrane enrichment that involves the purification of microsomal, organelle, or plasma membranes (Mirza et al., 2007; Ogawa et al., 2008; Huang et al., 2013; Abdallah et al., 2014; Guillier et al., 2014; Avila et al., 2015; Aloui et al., 2018). The effectiveness of the proteomic identification of TMPs from these membrane-enriched fractions, however, is compromised by contaminating cytoplasmic or membrane-associated proteins without transmembrane domains (TMDs). High ionic strength buffers can remove some of these contaminants with varied degrees of success (Marx et al., 2016; Vit and Petrak, 2017). Typically, TMPs represent roughly 20% of proteomic identifications (Chen et al., 2008). Although trypsin's high cleavage specificity and efficiency makes it the gold standard enzymatic method for MS-based bottom up proteomics (Olsen et al., 2004), its use typically enables sequence coverage limited to the soluble loops and terminal tails of TMPs (Schey et al., 2013). In addition, TMPs are difficult to solubilize and digest using standard urea solubilization and trypsin Lys-C-based procedures. Various detergents, chaotropic agents, organic solvents as well as proteinases and chemical cleavage reagents have been used to help solubilize and/or digest TMPs (Bennett et al., 1992; Newby et al., 2009; Guillier et al., 2014). There are few reports, however, comparing the efficiency of urea and acid-based procedures that aim to identify plant TMPs using mass spectrometry (MS).

To determine the protein makeup of a given sample using bottom-up proteomics requires maximal peptide coverage of the sample. TMPs with a high content of TMDs are underrepresented in MS identification since the most used protein purification method for bottom-up proteomics uses 8 M urea to solubilize the sample. The poor solubilization and thus poor trypsin digestion (UR-TLC) reduces the identification rate for proteins with a large number of TMPs (Long et al., 2018). The inability of urea to dissolve the membrane most likely contributes to trypsin Lys-C failing to cleave sites including TMDs. To address the poor ability of 8 M urea to solubilize hydrophobic TMPs, many proteomic studies use certain kinds of detergents to provide solubility (Kar et al., 2017). It is well accepted, however, that different detergents selectively solubilize certain proteins and not others (Churchward et al., 2005; Chen et al., 2007; Arachea et al., 2012; Laganowsky et al., 2013; Tanca et al., 2013). For example, hydrophobic membrane proteins in lipid rafts are poorly solubilized by detergents (Morel et al., 2006; Casem, 2016; Kusumi et al., 2020). In addition, detergents are not MS friendly and must be removed prior to MS analysis (Zhang and Li, 2004; Yeung et al., 2008). A previous *Medicago truncatula* study utilized an 8 M urea-based approach to solubilize and identify a wide range of proteins from different tissues (Marx et al., 2016). Here, we established a non-detergent proteomic approach by using FA to substitute for detergents and 8 M

urea (Zhao et al., 2013) for the solubilization of membrane samples of *Medicago*.

To address the shortcomings of urea and detergent-based approaches, we designed a non-detergent-based purification and identification strategy to analyze TMPs in *Medicago* using MS. *Medicago* is an important nitrogen-fixing agricultural crop, and this study augments the prior proteomic analyses of this plant (Natera et al., 2000; Djordjevic et al., 2003, 2007; Djordjevic, 2004; Zhang et al., 2006; Kusumawati et al., 2008; Lee et al., 2013; Marx et al., 2016). We used microsomal membrane preparations from *Medicago* shoot tissue as a common starting material. All the proteins from microsomal membrane preparation were precipitated by trichloroacetic acid (TCA). The protein samples were divided and then subjected to (i) the popular urea solubilization and trypsin Lys-C digestion-based method or (ii) a method for improving the cleavage of TMPs, which utilizes formic acid (FA) solubilization followed by cyanogen bromide (CNBr) cleavage and then trypsin Lys-C digestion (Quach et al., 2003; Girolamo et al., 2010; Vit and Petrak, 2017). It is well recognized that urea solubilization followed by enzymatic cleavage has a low efficiency in solubilizing and cleaving TMPs (Seddon et al., 2004; Carpenter et al., 2008; Rawlings, 2016). The cleaved peptides derived from the two methods were separated by high-pH reversed-phase peptide fractionation, and identified by orbitrap-based MS. The effectiveness of the two methods was assessed by (i) comparing total protein identifications, (ii) determining how many TMPs were identified, and (iii) defining how many peptides incorporate all, or part, of the TMD sequences using a new algorithm. In addition, we determined the subcellular location and predicted the function of the proteins identified. The objective of this study was to establish a detergent-free and effective strategy to identify plant TMPs and improve the overall identification and coverage of these proteins using MS.

MATERIALS AND METHODS

Plant Growth

Surface-sterilized *M. truncatula* cv. Jemalong A17 seeds (Imin et al., 2013) were germinated and grown on Fåhræus medium plates (Mohd-Radzman et al., 2015). Eight seedlings per plate were grown for 14 days in a Conviron growth chamber at 25°C with a 16-h photoperiod and a photon flux density of 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (Imin et al., 2013). Shoots were harvested separately and frozen in liquid nitrogen for immediate extraction or stored at -80°C before use. Three independent batches of *Medicago* shoot samples were collected to achieve independent biological replicates and enable an assessment of significance.

Microsomal Membrane Preparation

Proteins were extracted from homogenized and ground tissue based on published methods (Marx et al., 2016) with slight modifications. In brief, in order to have approximately 100 mg of microsomal membrane (MM) material, 12 g of shoot tissues (leaves and cotyledons) from 14-day-old seedlings were ground into a fine powder in liquid nitrogen using a mortar and

pestle. After grinding, five volumes (*circa* 50 ml) of ice-cold extraction buffer [290 mM sucrose, 250 mM Tris (pH 7.6), 25 mM EDTA (pH 8.0), 10 mM KCl, 25 mM NaF, 50 mM sodium pyrophosphate, 1 mM ammonium molybdate, 1 mM PMSE, mini EDTA-free protease inhibitor (Roche)] was added to the ground plant tissue samples. The ground tissue was further homogenized by repeated probe sonication (MSE: Imgen technologies) (10 cycles of 1 min sonication on ice followed by a 30-s rest period on ice). The homogenized plant tissue was filtered through a 100- μ m filter (BD Falcon, Bedford, MA, United States) and subsequently centrifuged for 10 min (4,000 g, 4°C) to remove the remaining tissue debris. MMs were prepared by ultracentrifugation for 30 min at 100,000 g (4°C) to remove cytoplasmic proteins. After ultracentrifugation, the MM pellet was resuspended in 1 M Na₂CO₃ (pH 11) and incubated on ice for 5 min to remove weakly associated proteins (Huang et al., 2013). After incubation, the MMs were subject to ultracentrifugation (100,000 g, 4°C) for 30 min (Abas and Luschnig, 2010).

TCA Precipitation and Protein Solubilization

Medicago MM proteins were purified by TCA precipitation (Link and LaBaer, 2011) with a slight modification to remove the non-protein contamination. In brief, 500 μ l of 11% TCA was added into MM sample pellet and incubated on ice for 20 min. Another 500 μ l of ice-cold 10% TCA solution was added, and the sample was incubated at -20° C overnight. The solution was centrifuged at 20,000 g for 30 min to recover the precipitated protein and the supernatant discarded. The protein pellet was rinsed three times with 80% acetone (Marx et al., 2016), centrifuged at 20,000 g for 10 min and dried using a vacuum evaporator (VirTis, bench TopK). The protein sample was divided into two to assess the effectiveness of the two protein solubilization and cleavage protocols.

Urea Solubilization Followed by Trypsin-Lys-C Digestion (UR-TLC)

The dried protein pellet was re-solubilized in 1 ml of dissolving buffer: 8 M urea, 50 mM Tris-HCl (pH 8.0), 30 mM NaCl, 1 mM CaCl₂, 20 mM sodium butyrate, 10 mM nicotinamide, mini EDTA-free protease inhibitor (Roche). To improve protein solubility, protein samples were subjected to repeated probe sonication (10 times for 10 s of pulse and 10 s of rest on ice). Protein concentration was estimated using a Bradford assay (Bio-Rad). Proteins were reduced with 5 mM dithiothreitol at 60°C for 40 min. The reduced proteins were alkylated with 15 mM iodoacetamide in the dark at room temperature for 40 min. Alkylation was quenched by adding 5 mM dithiothreitol and incubated at room temperature for 15 min. The protein solution concentration was estimated by UV 280 absorbance, and 200 μ g of protein sample was enzymatically digested in a two-step process using a Trypsin-Lys-C mix (Promega). A 25:1 molar ratio of the enzyme was added to the protein solution and digested for 3 h at 37°C. After 3 h, the urea concentration was adjusted to 2 M by dilution with 50 mM

Tris, pH 8.0 and the reaction kept at 37°C overnight. After overnight digestion, the sample was run over a Sep-Pak C18 classic cartridge (Waters, Milford, MA, United States) to remove the salts, and the peptides were eluted using 100% acetonitrile (ACN).

Formic Acid Solubilization Followed by Initial CNBr Cleavage and Trypsin-Lys-C Digestion (FA-CTLC)

The dried protein pellet was dissolved in 500 μ l of 70% FA using probe sonication and chemically digested with CNBr (Sigma) with a 100-fold molar ratio excess to the amount of starting dried protein (Wong and King, 2015). The CNBr solution was prepared as described (Washburn et al., 2001; Crimmins et al., 2005). Essentially, CNBr crystal was dissolved in ACN to make a 5 M final concentration. After 24 h of incubation at room temperature in the dark, the supernatant was collected by centrifugation at 20,000 g for 10 min, and the FA and CNBr were safely removed by lyophilization using cold trap. The dried peptide was dissolved in 10% ACN and 25 mM ammonium bicarbonate. The sample was reduced, alkylated, and then digested with Trypsin Lys-C, before being lyophilized (see above). The effectiveness of each procedure was assessed by centrifuging (10 min, 10,000 g) after UR-TLC or FA-CTLC treatments and examining the residual pellet. The residual pellet that remained after the UR-TLC was subjected to a further FA-CTLC procedure to validate that the pellet contained poorly solubilized undigested protein. The remaining CNBr solution was destroyed by adding 5 volumes of 1 M sodium hydroxide (Lunn and Sansone, 1985) before being disposed into chemical waste containers.

Reverse-Phase High-pH Fractionation

The peptide digests were separated into eight fractions using the high-pH reversed-phase peptide fractionation Kit (Thermo Fisher Scientific, Waltham, MA, United States) (Kulak et al., 2017; Baldan-Martin et al., 2018). In brief, the C-18 spin column was equilibrated with 0.1% trifluoroacetic acid (TFA) before the digested peptides were loaded. Peptide samples (100 μ g) dissolved in 0.1% TFA were loaded onto the spin column and washed with MilliQ H₂O. The peptides were eluted into 16 fractions of increasing concentrations of ACN in 0.1% Triethylamine: 5% ACN (fraction 1), 7.5% ACN (fraction 2), 10% ACN (fraction 3), 12.5% ACN (fraction 4), 15% ACN (fraction 5), 17.5% ACN (fraction 6), 20% ACN (fraction 7), 25% ACN (fraction 8), 30% ACN (fraction 9), 35% ACN (fraction 10), 40% ACN (fraction 11), 45% ACN (fraction 12), 50% ACN (fraction 13), 60% ACN (fraction 14), 70% ACN (fraction 15), and 95% ACN (fraction 16). The 16 fractions were collected and then recombined into eight fractions. Final fraction 1 comprised RP fractions 1 and 16; final fraction 2 comprised RP fractions 2 and 15; and so on until final fraction 8 comprised RP fractions 8 and 9. The fractionated peptides were lyophilized to remove the solvent, re-dissolved into 100 μ l of 0.1% TFA, and cleaned up by C18 Ziptip (5 μ g loading capacity, Merck Millipore,

Burlington, MA, United States), and 1.7 μg of the sample was subjected to MS analysis.

Mass Spectrometry

The liquid chromatography (LC) was performed by using Thermo Scientific UltiMate™ 3,000 RSLCnano system with the setting at 60°C with customized columns. The columns were packed in-house using a laser puller and a pressure bomb, and the length of the columns were generally 35–40 cm with a 75- μm ID fused silica housing. The packing material used was Reprosil-Pur 120 C18-AQ, 1.9- μm particle size. The digested peptides were initially loaded onto the LC system with the mobile phases as 95% buffer A (0.1% formic acid/water) and 5% buffer B (0.1% formic acid/80% ACN/water). Samples were reconstituted in 10 μl of loading buffer (as above) and 3 μl directly injected for each run. Peptides were eluted with a 5–40% buffer B gradient for 90 min. The total acquisition time was 140 min, including a 95% ACN wash and re-equilibration. The LC was coupled to a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, United States). MS scans acquired in the Orbitrap (mass resolution was 70,000 at m/z 200; mass analyzer range was m/z 350–2,000). The 20 most intense ions with a charge state ≥ 1 were fragmented in the high-energy C-trap dissociation collision (HCD) cell, and subsequently, tandem mass spectra were acquired in the Orbitrap mass analyzer with a resolution of 35,000 at $m/z = 200$.

Data Analysis

All raw files generated by LC-MS/MS were processed by Proteome Discoverer 2.1 (Thermo Fisher Scientific) using the Sequest HT data analysis program to search against the *Medicago* protein sequence databases (UniProt, 2014.12.18.) (Bairoch et al., 2005). Database searching against the corresponding reversed database was also performed to evaluate the false discovery rate of peptide identification. The search parameters of Sequest HT were set as follows: precursor ion mass tolerance ± 10 ppm and product ion mass tolerance of 0.05 m/z units. The cleavage specificity was set up as Trypsin/LysC: C-terminal of arginine and lysine, and CNBR/Trypsin/LysC: C-terminal of methionine, arginine, and lysine. Standard peptide modification was as follows: carbamidomethylation (CAM) at cysteine residues was set as a fixed modification, while oxidation at methionine, lysine, and proline residues, C-terminal amidation and deamidation at asparagine and glutamine, as well as N-terminal glutamine to pyroglutamic acid were set as variable modifications. When CNBr was chosen as the cleavage agent, methionine was set to be homoserine (Met- > Hse) or homoserine lactone (Met- > Hsl) as a variable modification. The O-formylation at serine and threonine, which was caused by formic acid (Zheng and Doucette, 2016), was set to be a variable modification. The phosphorylation at serine, threonine, and tyrosine was also set to be variable modification. For all experiments, we used Pepecator with a strict cutoff (< 0.1) to determine the FDR of the peptides identified. Due to sequence redundancy, the proteins that had shared the same set of identified peptides were grouped into protein groups.

TMD Prediction

The TMD prediction was done by using the TMHMM Server, v. 2.0¹. TMHMM is a membrane protein topology prediction method based on a hidden Markov model (Sonnhammer et al., 1998). The web server-based search engine correctly predicts 97–98% of the transmembrane helices and can distinguish between soluble and membrane proteins with a specificity and sensitivity better than 99% (Krogh et al., 2001).

TMD (TRANDOMI, TRAns Membrane Domain Motif Identification)

The TMD analysis was annotated by a customized python script. In brief, two peptide sequence sets were prepared for TMD mapping. One was the detected peptide set, which was derived from the MS analysis of the samples. The second one was the complete *Medicago* TMD motif set, which was predicted by TMHMM server using the *M. truncatula* sequence databases (UniProt, 2014.12.18). TMD identification was done by mapping the detected peptide set to the *Medicago* TMD motif set. The mapping rules were defined as followed. Any detected peptide that satisfied one of the following rules was considered a hit:

1. The length of the detected peptide derived from MS encompassed the complete predicted *Medicago* TMD sequence or laid within the predicted TMD.
2. The detected peptide derived from MS extends from a position outside the TMD to within the TMD with a minimum of a two-amino acid overlap, or started within the TMD and extended to a position outside of the TMD with a minimum of two-TMD-amino acid overlap.

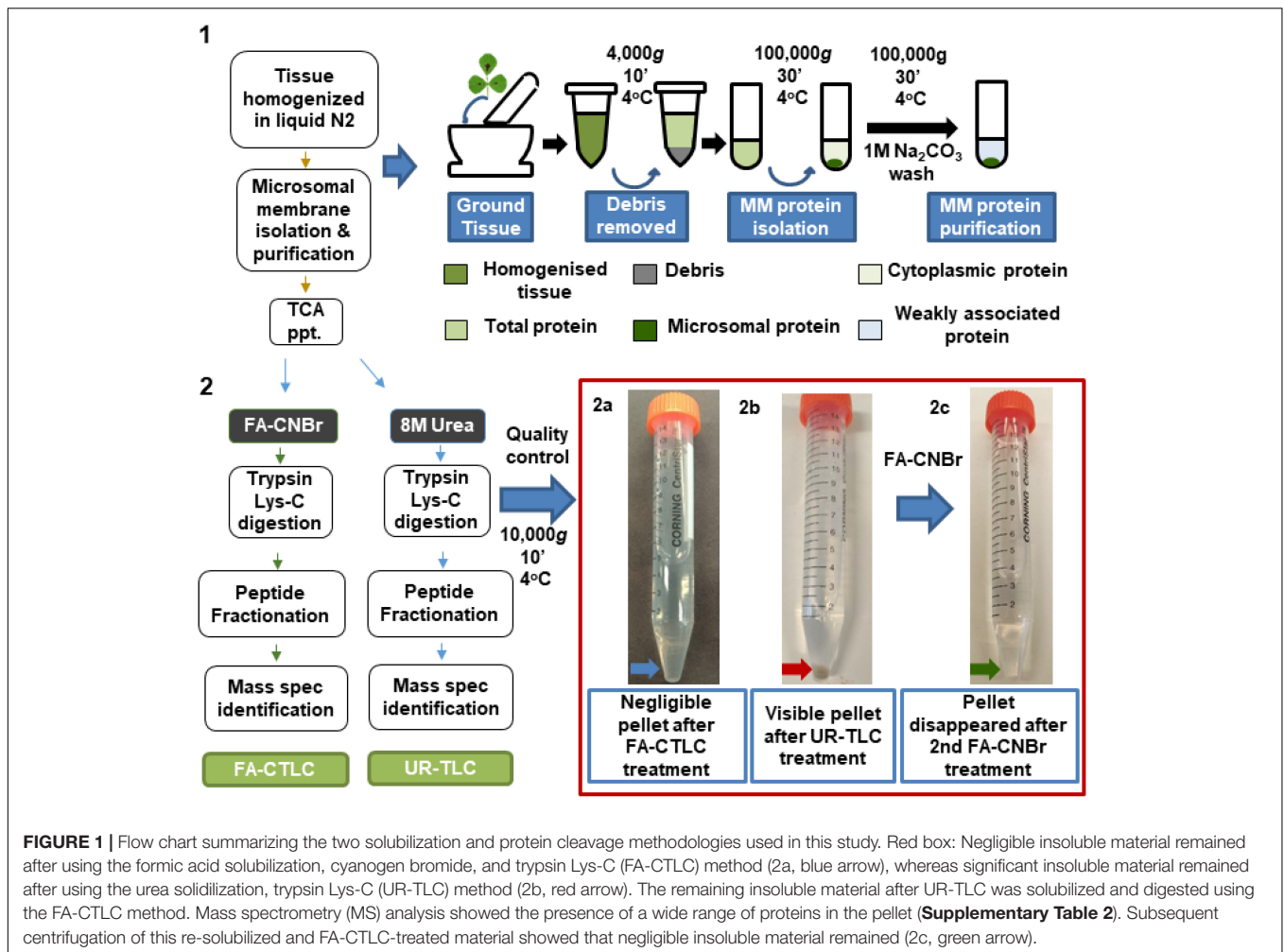
Calculate the Grand Average of Hydropathy Value for Protein Sequences

The Grand Average of Hydropathy (GRAVY) value is calculated by the sum of hydropathy values of all amino acids divided by the protein length (Kyte and Doolittle, 1982). Hydrophobicity score (arbitrary unit) below 0 is more likely cytoplasmic protein (hydrophilic protein), while scores above 0 are more likely TMPs (hydrophobic) (Magdeldin et al., 2012).

Subcellular Location Prediction

The subcellular protein location prediction was done by using LOCALIZER, a machine learning method for predicting subcellular protein localization in plant cells and is available at <http://localizer.csiro.au/>. It identifies proteins localized to chloroplasts and mitochondria by identifying the presence of transit peptides, and nucleus by using a collection of nuclear localization signals. It can achieve a prediction accuracy of over 90% for chloroplast and mitochondria, and 73% for nuclear proteins (Sperschneider et al., 2017). The queries of protein sequence were submitted directly to the server, and full plant sequences were chosen to perform the prediction.

¹<http://www.cbs.dtu.dk/services/TMHMM/>



Protein Functional Annotation

The functional annotation was done using Mercator: <http://mapman.gabipd.org/web/guest/app/Mercator>. Mercator is a web based annotation application that achieves accuracies above 90% in predicted functional annotations when compared to manual annotation (Lohse et al., 2014). The queries of protein sequence were submitted directly to the server, searched against the database including TAIR Release 10 and SwissProt/UniProt plant proteins database, and classified into functional plant categories according to MapMan BINs (Thimm et al., 2004).

RESULTS

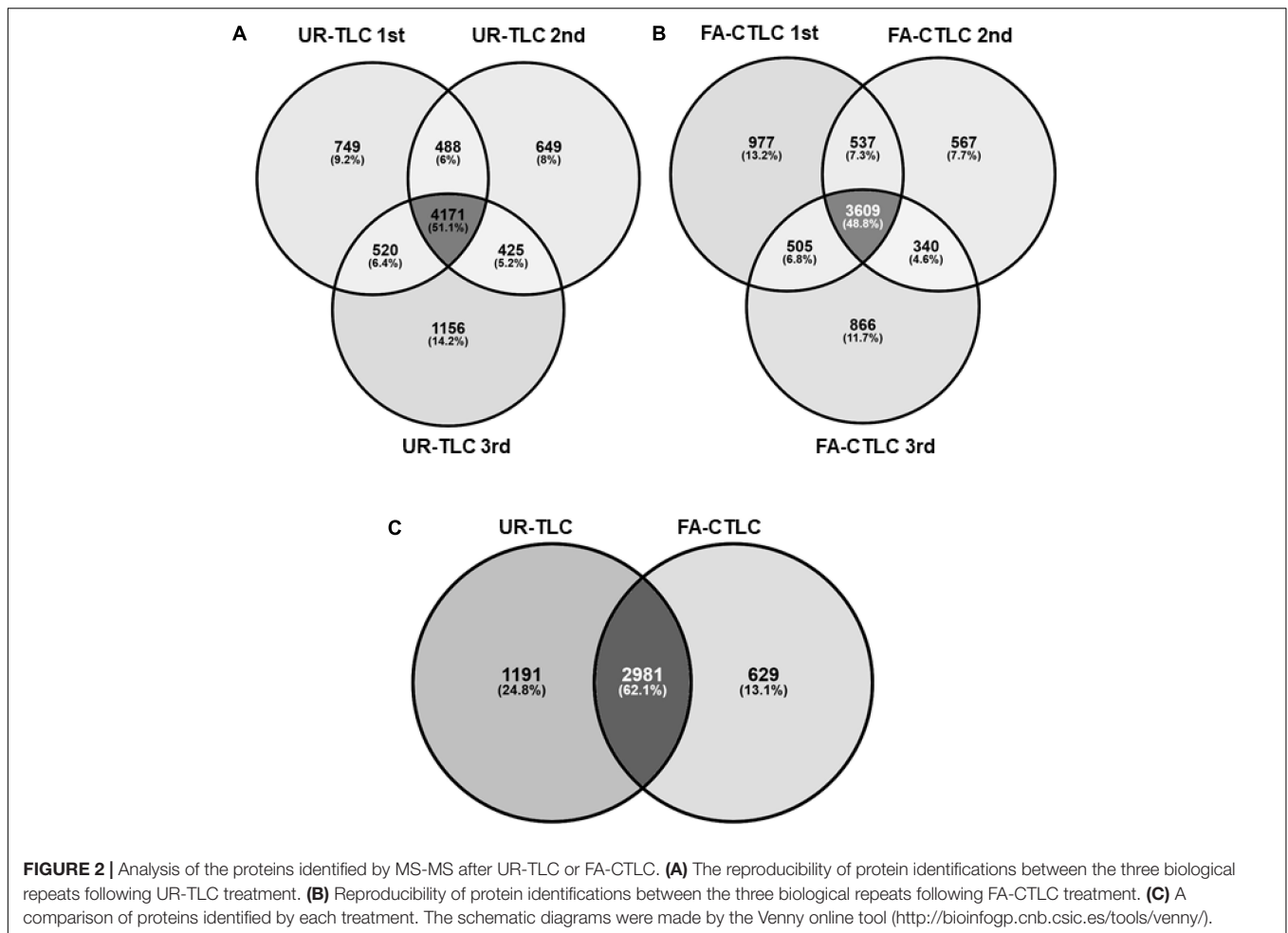
Establishment of a FA-CTLC Method for Improving Bottom-Up Proteomics and Membrane Protein Identification

As a preliminary assessment step, we determined the effectiveness of FA solubilization combined with CNBr treatment to validate that CNBr cleaves to the C-terminal side of the comparatively rare methionine (Met) residues to generate large peptide fragments, since methionine occurs, on average, at every 50 amino acids. The

MS results confirmed that FA solubilization followed by CNBr cleavage alone resulted in the production of large peptides with C-terminal Met residues, as expected (**Supplementary Table 1**). Peptides of large molecular mass with an uncharged, C-terminal Met do not ionize well and give poor-quality MS/MS spectra. As expected, this resulted in poor coverage of the proteome; only 2,566 protein groups were observed using the CNBr cleavage method only (at a 1% FDR). This was remedied by following the CNBr cleavage with trypsin Lys-C digestions.

The FA-CTLC Is Superior at Solubilizing and Digesting Proteins From MM Preparations

A summary of the key steps of the two procedures is shown in **Figure 1**. To compare the effectiveness of the solubilization and digestion, MM preparations were split in half and subjected to either the UR-TLC or the FA-CTLC method. The results (**Figure 1**, red box; red arrow) showed that a considerable pellet of insoluble material remained after UR-TLC, but not after FA-CTLC. A second round of the UR-TLC was applied to the pellet, but after overnight digestion and subsequent re-centrifugation,



the pellet remained. By contrast, the application of the FA-CTLC method to the insoluble pellet that remained after using the UR-TLC method resulted in no observable pellet after centrifugation (Figure 1, red box; green arrow). An MS analysis of this FA-CTLC re-solubilized pellet material (red box; red arrow) identified 5,644 protein groups (at a 1% FDR) in the insoluble material that remained after UR-TLC (Supplementary Table 2). From the 5,644 protein groups identified by a re-extraction of the insoluble material, the UR-TLC failed to identify 979 protein groups (Supplementary Table 3). This demonstrated that considerable protein material remained in the UR-TLC pellet (red box; red arrow) and that the FA-CTLC is more effective in solubilizing and digesting the proteins from MM preparations as shown in the method flowchart (Figure 1).

Assessment of Inter-Sample Reproducibility and the Effectiveness of Protein Identification by MS After Using the FA-CTLC or the UR-TLC Method

We assessed the reproducibility of the two methods by examining the proteins identified using three biological repeats from each treatment (Figure 2). When considering proteins

with an FDR of <1%, we identified 4,171 protein groups common to all three biological repeats after UR-TLC and 3,609 protein groups common to all three biological repeats after FA-CTLC. The reproducibility of the proteins identified was 51.1 and 48.8% for the UR-TLC and FA-CTLC methods, respectively, (Figures 2A,B). We further compared the 4,171 protein groups common to all three biological repeats after UR-TLC (Figure 2A) to the 3,609 protein groups common to all three biological repeats after FA-CTLC (Figure 2B) and found that 2,981 groups of common proteins were identified by both methods (Figure 2C). Based on the results, 666 protein groups were unique to FA-CTLC identification method, while 1,523 protein groups were unique to UR-TLC. Supplementary Table 4 shows the complete list of protein groups identified in three biological repeats from both methods. The 1,523 UR-TLC-specific protein groups are shown in Supplementary Table 5 and the 666 FA-CTLC-specific protein groups identified after using the FA-CTLC are shown in Supplementary Table 6.

When considering proteins identified from three repeats, 7,946 protein groups were identified following UR-TLC. Protein groups (7,118) were identified following FA-CTLC (Supplementary Table 7). After combining the results

derived from the two methods, we identified 8,993 protein groups in total (**Supplementary Table 8**). To determine the effectiveness of the two methods at identifying TMPs, the identified proteins were analyzed for the presence of TMDs and hydrophobicity.

The FA-CTLC Method Preferentially Identifies TMPs With a Higher Number of TMDs

By using the TMHMM algorithm, 23.26% of the proteins in the *Medicago* Uniprot database were identified as TMPs. From the 7,946 protein groups identified after UR-TLC treatment, 2,817 protein groups (35.45%) contained at least one TMD, and of the 7,118 protein groups identified from the FA-CTLC treatment, 2,784 protein groups (39.11%) contained at least one TMD. Therefore, a higher percentage of TMPs can be identified by using the FA-CTLC method compared to using the UR-TLC method. Additional analysis showed that there were 5,129 protein groups identified from UR-TLC treatment and 4,334 protein groups identified from FA-CTLC treatment with 0 TMDs. This result suggests that the published procedures for removing non-membrane proteins (e.g., using sodium carbonate washes at pH 11) have poor efficacy. We further analyzed the TMP distribution in different biological repeats from each purification method using the TMHMM algorithm (**Figure 3**). About half of the TMD containing proteins identified using either method had only one TMD, and both purification methods gave no significant difference in distribution of TMDs to that predicted by analyzing the theoretical distribution of TMDs in all *Medicago* TMPs (inset of **Figure 3**, confirmed by Chi-Square Test, $p = 0.32$). This suggests that there is no major bias of either method in identifying TMPs and that the most abundant TMPs are likely to populate the lists of proteins identified. A significant difference between the proteins identified was that UR-TLC method preferentially identified proteins with only one TMD, whereas the FA-CTLC method preferentially identified more proteins with greater than four TMDs (**Figure 3**; $p < 0.05$). By combining the TMPs identified by UR-TLC and FA-CTLC, 3,289 TMP groups were identified, or 36.57% of all predicted TMPs.

The proteins identified after using UR-TLC or FA-CTLC methods were also examined using the GRAVY algorithm (Kyte and Doolittle, 1982). The GRAVY index indicates the hydrophobicity of the proteins, calculated by adding the hydrophobicity value for each residue and dividing by the length of the sequence. Proteins with a GRAVY scores above 0 are more likely to be hydrophobic proteins (Magdeldin et al., 2012). The GRAVY results (**Figure 3B**) showed that there were 1,420 (19.95%) protein groups identified by FA-CTLC and 1,313 (16.52%) proteins identified by UR-TLC, which displayed a GRAVY score greater than zero. These results indicated that the FA-CTLC method can preferably purify proteins that are hydrophobic.

We developed the TRAMDOMI algorithm to identify the peptides that contain all or part of TMD motifs within the TMPs. This algorithm enabled us to quantify the relative ability

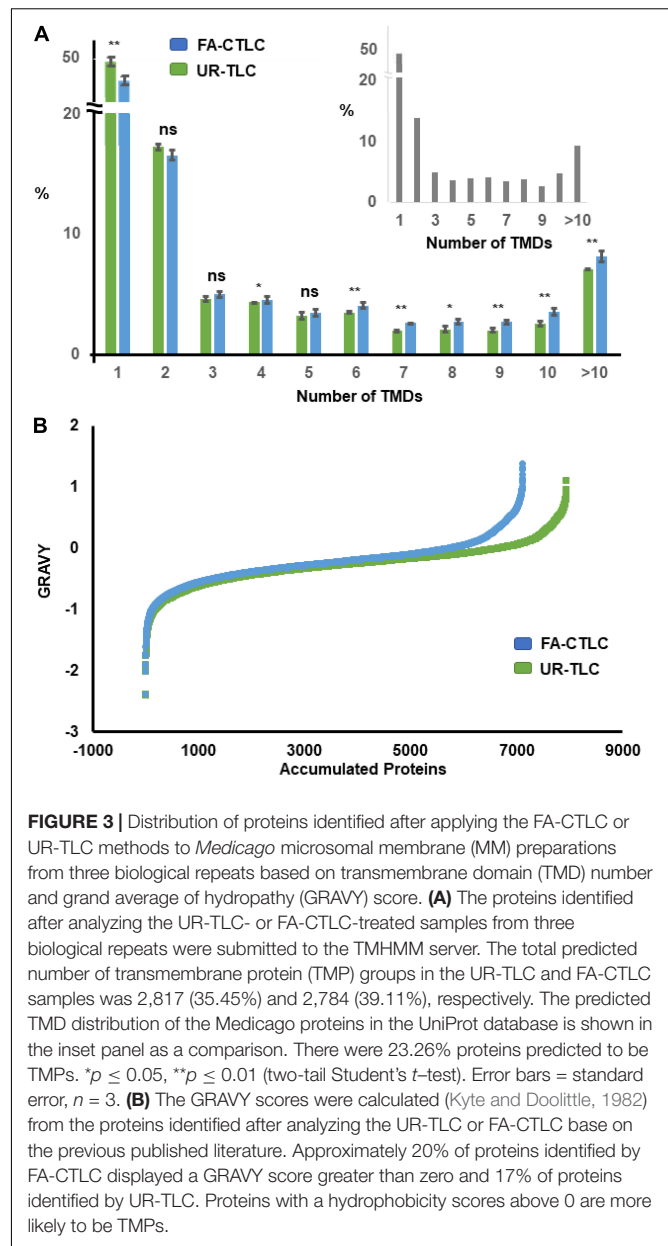


FIGURE 3 | Distribution of proteins identified after applying the FA-CTLC or UR-TLC methods to *Medicago* microsomal membrane (MM) preparations from three biological repeats based on transmembrane domain (TMD) number and grand average of hydropathy (GRAVY) score. **(A)** The proteins identified after analyzing the UR-TLC- or FA-CTLC-treated samples from three biological repeats were submitted to the TMHMM server. The total predicted number of transmembrane protein (TMP) groups in the UR-TLC and FA-CTLC samples was 2,817 (35.45%) and 2,784 (39.11%), respectively. The predicted TMD distribution of the *Medicago* proteins in the UniProt database is shown in the inset panel as a comparison. There were 23.26% proteins predicted to be TMPs. * $p \leq 0.05$, ** $p \leq 0.01$ (two-tail Student's t -test). Error bars = standard error, $n = 3$. **(B)** The GRAVY scores were calculated (Kyte and Doolittle, 1982) from the proteins identified after analyzing the UR-TLC or FA-CTLC base on the previous published literature. Approximately 20% of proteins identified by FA-CTLC displayed a GRAVY score greater than zero and 17% of proteins identified by UR-TLC. Proteins with a hydrophobicity scores above 0 are more likely to be TMPs.

of each method to identify peptides with TMD motifs. The search results showed that the FA-CTLC method can identify 9.36 times more TMD-containing peptides than the UR-TLC method (811 compared to 87; **Figure 4A**). Therefore, the results indicate that the FA-CTLC method is more effective at detecting peptides within TMPs that have TMD motifs, which boosts the number of TMPs identified. A list of identified TMPs and the TMDs identified using both purification methods is shown in **Supplementary Table 9**. To further illustrate the difference between the two methods, we compared the peptides identified for the MFS/sugar transporter (MTR_7g005910), which has 12 predicted TMDs (**Figures 4B,C**). Clearly, the FA-CTLC method identified more peptides within MTR_7g005910 with TMD motifs, whereas the UR-TLC method only identified

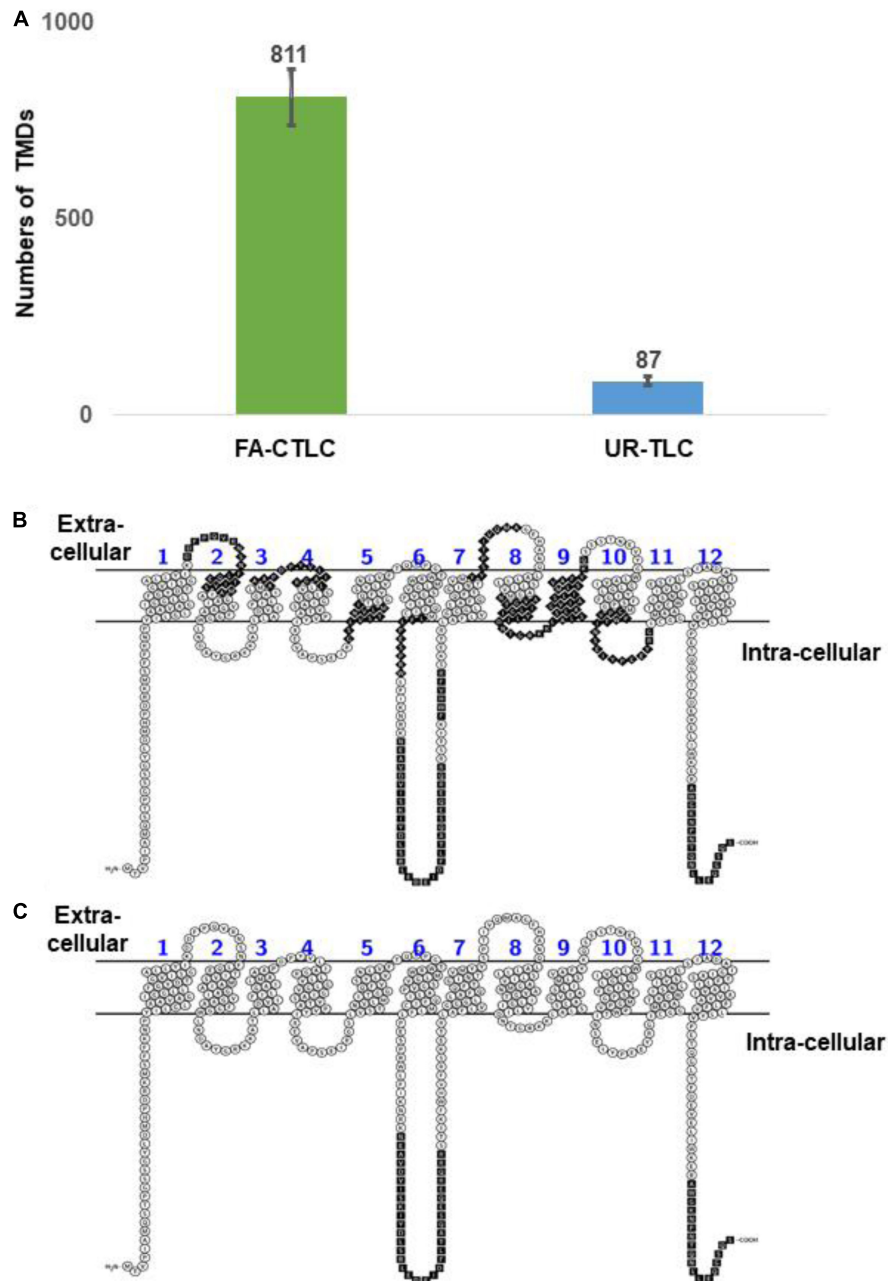


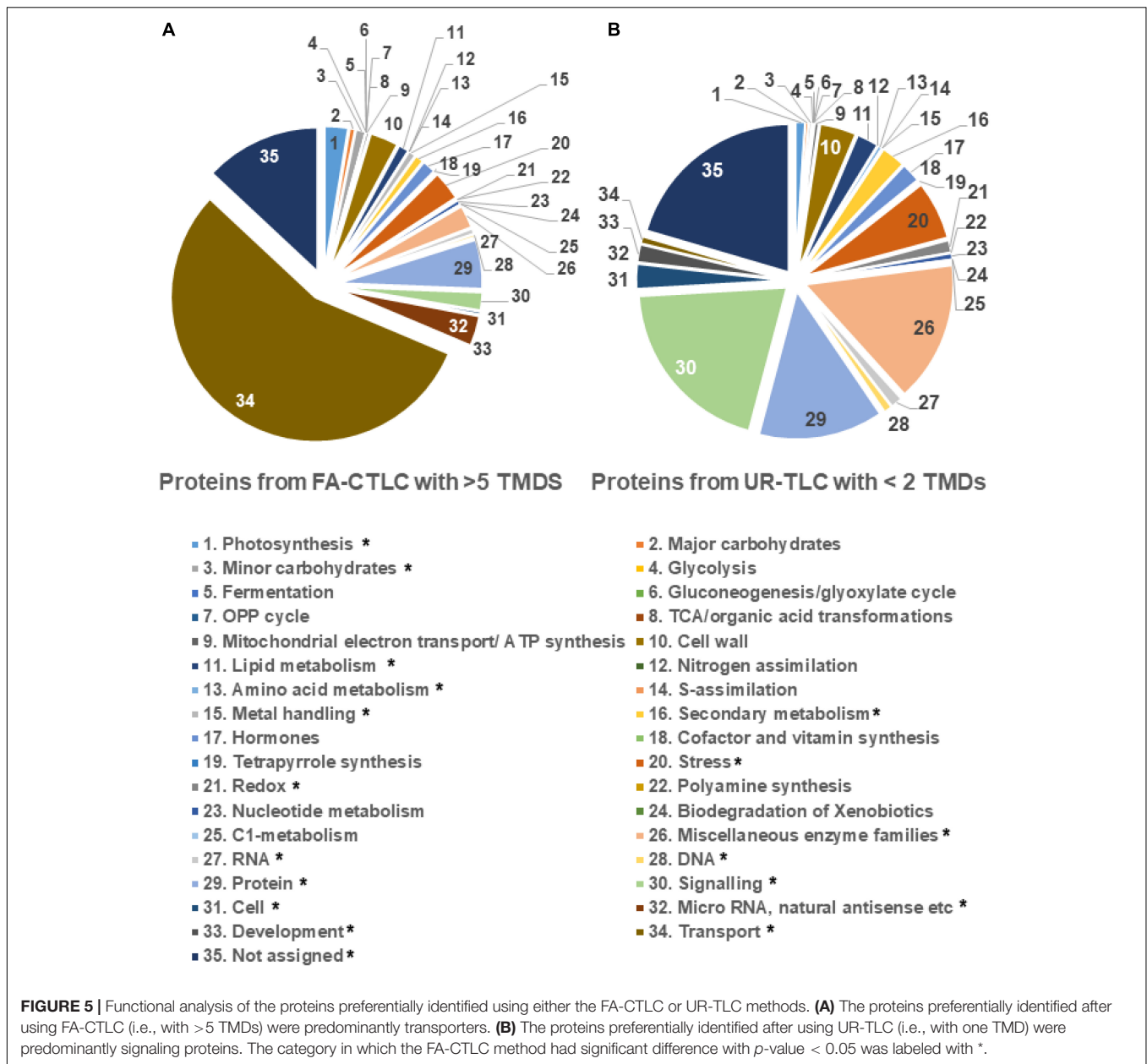
FIGURE 4 | The FA-CTLC method preferentially identifies peptides with TMD motifs. **(A)** A comparison of the total TMDs identified numbers from both purification methods. The FA-CTLC methods identified 9.3-fold more TMDs than the UR-TLC method. **(B,C)** The peptides identified in the MTR_7g005910 transporter of *Medicago* after using the FA-CTLC **(B)** or UR-TLC method **(C)**. The peptides that were identified by MS were colored in black. The schematic diagrams were made using the Protter online tool (Omasits et al., 2014) (<http://wlab.ethz.ch/protter/start/>).

MTR_7g005910 peptides predicted to loop into the cytoplasm, and none that contained TMD motifs.

Transporter Proteins Are Preferentially Identified by Using the FA-CTLC Method

Given that each solubilization and cleavage method identified distinct classes of peptides, a Mercator analysis was done to

determine if the two methods resulted in the enrichment of the identification of proteins with different functions. The results (**Figure 5**) show that the TMPs containing one TMD, which were preferentially identified by UR-TLC, were mostly functionally assigned as being signaling proteins (20.81%), and the proteins preferentially identified by FA-CTLC, which had four or greater TMDs were predominantly functionally assigned as being transporters (58.01%). We further examined the difference



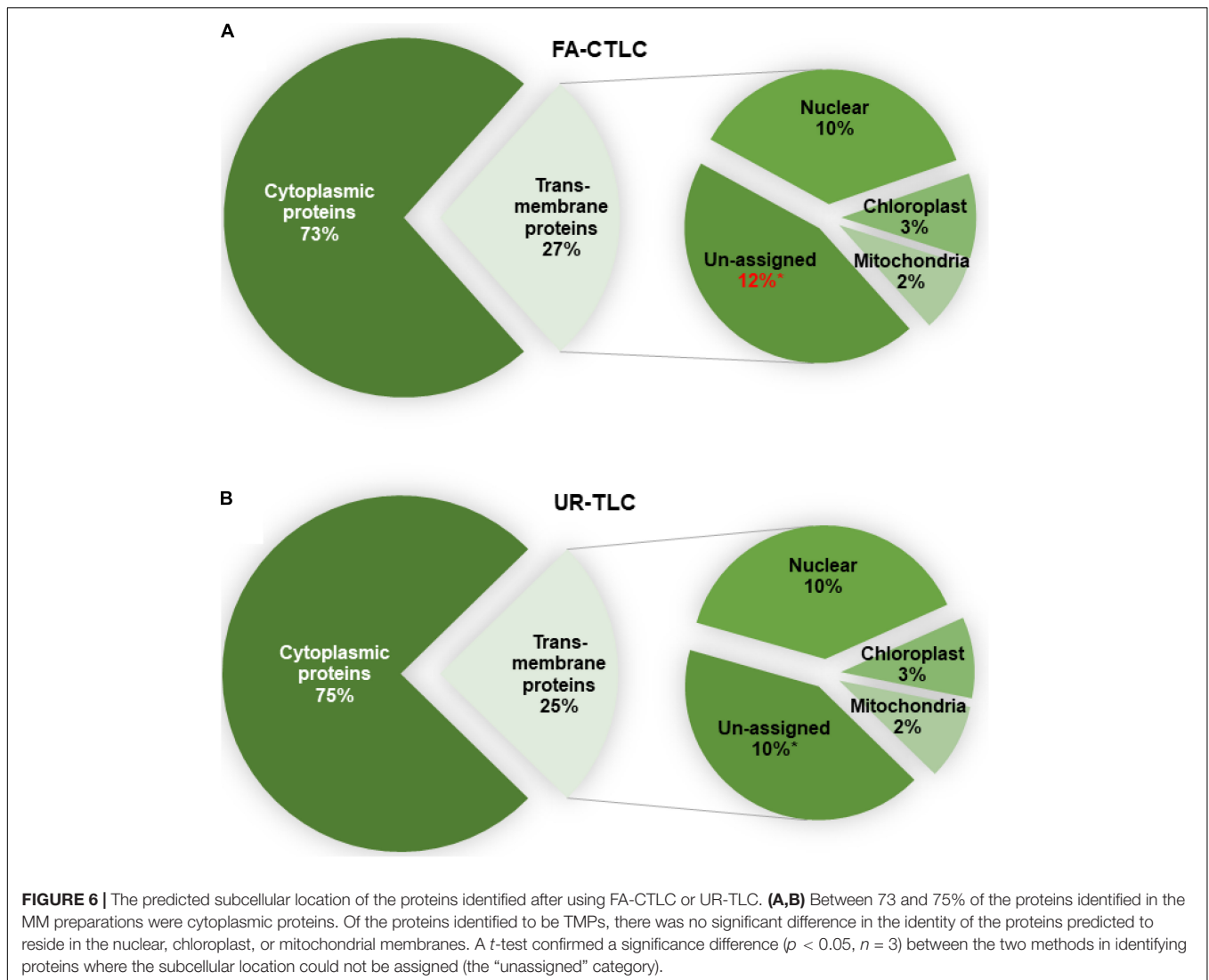
between two data sets by a binomial test. The results showed that the proteins identified by the FA-CLTC method were significantly different in 18 categories when compared to the UR-TLC method (Figure 5). The complete protein functional analysis list is shown in Supplementary Table 10, and the binomial test results are shown in Supplementary Table 11.

To determine the likely membrane where the TMPs identified reside, all proteins were analyzed for their subcellular location using LOCALIZER (Figure 6). Irrespective of the method used, an analysis of the TMPs identified showed that there was a similar distribution of proteins predicted to reside in the membranes of the nucleus, chloroplast, or mitochondria. For both methods, the FA-CTLC method identified significantly more TMPs where the subcellular location could not be assigned to an organelle

($p = 0.044$, $n = 3$). The complete subcellular location prediction list is shown in Supplementary Table 12.

DISCUSSION AND CONCLUSION

The results showed that the FA-CTLC method was superior at solubilizing and digesting more hydrophobic proteins from MM preparations. Of the 57,065 proteins in the MT data base, 13,274 (23.26%) are predicted to be TMPs. The combined output of the two methods identified 3,289 TMP groups representing 36.57% of all TMPs, which is 1.5-fold more protein identifications achieved in a recent quantitative proteomic analysis of young Medicago seedlings (Long et al., 2018) and more comparable to the



number of *Medicago* proteins identified using a similar sampling and bioinformatics procedure and similar instrumentation (Marx et al., 2016). In this study, we achieved a comparable number of identifications by a cost-effective method with less fractions and MS runs. About 50% of the TMPs identified using either method had only one TMD, but reassuringly, both purification methods gave no significant difference in the distribution of TMDs to that predicted by analyzing the theoretical distribution of TMDs in all *Medicago* TMPs using the THMMM algorithm. Therefore, this result suggests that there was no major bias of either method in identifying TMPs. To compare the TMD identification efficiency between the two methods utilized, we customized the TRAMDOMI algorithm to reveal how many TMDs were purified and identified from each purification method. The TRAMDOMI algorithm is the first python script designed for matching TMDs with peptides identified by MS. By using the TRAMDOMI algorithm, we identified a significant benefit of using the FA-CTLC method: this method preferentially identifies TMPs with a significantly

higher number (9.4-fold) of TMDs than the UR-TLC method. In addition, each method identified partially non-overlapping TMP cohorts. Each purification method still had its preference, since 666 protein groups are unique to FA-CTLC identification method, while 1,523 protein groups are unique to UR-TLC. This result was validated by the FA-CTLC method identifying more transporter proteins, which have >8 TMDs, whereas the UR-TLC method preferentially enriched signaling proteins, which contain one TMD. The results implied that TMPs that were buried in cell membrane were difficult to denature or solubilize using 8 M urea. Therefore, UR-TLC method most likely shaves the exposed extra- and intracellular domains that loop away from the TMP regions imbedded inside the membrane. This deficiency leads to lower protein sequence coverage for proteins with a higher number of TMDs. By contrast, the UR-TLC method gave a better identification of TMPs with one TMD. TMPs from *Medicago* have a variable number of TMDs that range from 1 to over 30. Therefore, TMDs constitute a variable percentage of the composition of TMPs. TMDs are poorly

represented in bottom-up MS (Kar et al., 2017), and the ability of a TMP to be detected by MS depends on its subcellular location, tissue specificity, natural abundance, the methodology used for fractionation, and the sensitivity and accuracy of the instrumentation (Bausch-Fluck et al., 2015; Itzhak et al., 2016; Reinke et al., 2017). Therefore, any MS-based method designed to improve the identification and coverage of TMPs should identify peptides from those parts of the TMPs that include the TMDs.

It is unclear if the acid-based solubilization or the preliminary cleavage at Met residues followed by the trypsin/LysC digestion is the basis for the improved TMD coverage in this work. Recently, Sun et al. (2020) used VAILase cleavage of purified proteins to marginally improve TMD coverage, although VAILase is not currently commercially available. Therefore, it is possible that using proteases such as VAILase, which cut at aliphatic amino acids (Val, Ala, Ile, Leu, and Thr), may improve TMD coverage (Sun et al., 2020).

Increasing the protein sequence coverage of TMPs is known to benefit quantitative proteomics (Ishihama et al., 2008; Millionni et al., 2011; Koziol et al., 2013). Therefore, since the FA-CTLC method can provide a higher sequence coverage of proteins with a higher number of TMDs, it has the potential to provide superior data for quantitative proteomics. We recommend that combining the two methods should achieve better TMP identification and a better coverage of TMP peptides. The reproducibility among the biological repeats could be further improved by employing label-free quantification (Mosley et al., 2011; Müller et al., 2018; Barkovits et al., 2020), which may further reveal the differences in TMP abundance between the two solubilization procedures. After combining the results derived from the two methods, we identified 8,993 protein groups and 3,289 TMPs in young shoot tissues. Therefore, if more tissues were examined and more extensive, membrane fractionation techniques applied, the number of TMPs and their coverage would be expected to increase.

REFERENCES

- Abas, L., and Luschnig, C. (2010). Maximum yields of microsomal-type membranes from small amounts of plant material without requiring ultracentrifugation. *Anal. Biochem.* 401, 217–227. doi: 10.1016/j.ab.2010.02.030
- Abdallah, C., Valot, B., Guillier, C., Mounier, A., Balliau, T., Zivy, M., et al. (2014). The membrane proteome of Medicago truncatula roots displays qualitative and quantitative changes in response to arbuscular mycorrhizal symbiosis. *J. Proteom.* 108, 354–368. doi: 10.1016/j.jprot.2014.05.028
- Almeida, J. G., Preto, A. J., Koukos, P. I., Bonvin, A., and Moreira, I. S. (2017). Membrane proteins structures: A review on computational modeling tools. *Biochim. Biophys. Acta Biomembr.* 1859, 2021–2039. doi: 10.1016/j.bbamem.2017.07.008
- Aloui, A., Recorbet, G., Lemaitre-Guillier, C., Mounier, A., Balliau, T., Zivy, M., et al. (2018). The plasma membrane proteome of Medicago truncatula roots as modified by arbuscular mycorrhizal symbiosis. *Mycorrhiza* 28, 1–16. doi: 10.1007/s00572-017-0789785
- Arachea, B. T., Sun, Z., Potente, N., Malik, R., Isailovic, D., and Viola, R. E. (2012). Detergent selection for enhanced extraction of membrane proteins. *Protein. Expr. Purif.* 86, 12–20. doi: 10.1016/j.pep.2012.08.016
- Avila, J. R., Lee, J. S., and Torii, K. U. (2015). Co-Immunoprecipitation of Membrane-Bound Receptors. *Arabid. Book* 13:e0180. doi: 10.1199/tab.0180

DATA AVAILABILITY STATEMENT

The raw mass spectrometry data was submitted via ProteomeXchange (Vizcaíno et al., 2014) to the PRIDE repository (PXD022299) (Vizcaíno et al., 2012). TRANDOMI script is available on public server Github (<https://github.com/HanChung-Lee/TRANDOMI.git>) for downloading.

AUTHOR CONTRIBUTIONS

HL and MD designed the research. HL performed the research. HL, ACa, AB, and MD wrote the manuscript. ACa wrote the python scripts and analyzed the data. BC and ACo performed the MS. All authors contributed to the article and approved the submitted version.

FUNDING

HL was supported by a Taiwan-ANU Scholarship. An Australian Research Council grant to MD (DP150104250) supported this work.

ACKNOWLEDGMENTS

We thank Neha Patel for providing technical support on protein purification and Steve Binos for providing technical advice on MS and PD search.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.595726/full#supplementary-material>

- Bairoch, A., Apweiler, R., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., et al. (2005). The Universal Protein Resource (UniProt). *Nucleic Acids Res.* 33, D154–D159. doi: 10.1093/nar/gki070
- Baldan-Martin, M., Lopez, J. A., Corbacho-Alonso, N., Martinez, P. J., Rodriguez-Sanchez, E., Mourino-Alvarez, L., et al. (2018). Potential role of new molecular plasma signatures on cardiovascular risk stratification in asymptomatic individuals. *Sci. Rep.* 8:4802. doi: 10.1038/s41598-018-2303723037
- Barkovits, K., Pacharra, S., Pfeiffer, K., Steinbach, S., Eisenacher, M., Marcus, K., et al. (2020). Reproducibility, Specificity and Accuracy of Relative Quantification Using Spectral Library-based Data-independent Acquisition. *Mole. Cell. Proteom.* 19, 181–197. doi: 10.1074/mcp.RA119.001714
- Bausch-Fluck, D., Hofmann, A., Bock, T., Frei, A. P., Cerciello, F., Jacobs, A., et al. (2015). A mass spectrometric-derived cell surface protein atlas. *PLoS One* 10:e0121314–e0121314. doi: 10.1371/journal.pone.0121314
- Bennett, M. K., Calakos, N., Kreiner, T., and Scheller, R. H. (1992). Synaptic vesicle membrane proteins interact to form a multimeric complex. *J. Cell Biol.* 116:761775.
- Carpenter, E. P., Beis, K., Cameron, A. D., and Iwata, S. (2008). Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.* 18, 581–586. doi: 10.1016/j.sbi.2008.07.001
- Casem, M. L. (2016). “Chapter 5 - Membranes and Membrane Transport,” in *Case Studies in Cell Biology*, ed. M. L. Casem (Boston: Academic Press), 105–125.

- Chen, E. I., Cociorva, D., Norris, J. L., and Yates, J. R. III (2007). Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *J. Proteome Res.* 6, 2529–2538. doi: 10.1021/pr060682a
- Chen, E. I., McClatchy, D., Park, S. K., and Yates, J. R. III (2008). Comparisons of mass spectrometry compatible surfactants for global analysis of the mammalian brain proteome. *Anal. Chem.* 80, 8694–8701. doi: 10.1021/ac800606w
- Churchward, M. A., Butt, R. H., Lang, J. C., Hsu, K. K., and Coorsen, J. R. (2005). Enhanced detergent extraction for analysis of membrane proteomes by two-dimensional gel electrophoresis. *Proteome Sci.* 3:5. doi: 10.1186/1477-5956-35
- Clark, S. E., Williams, R. W., and Meyerowitz, E. M. (1997). The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell* 89, 575–585. doi: 10.1016/s0092-8674(00)80239-1
- Crimmins, D. L., Mische, S. M., and Denslow, N. D. (2005). Chemical cleavage of proteins in solution. *Curr. Protoc. Protein Sci.* 11:14. doi: 10.1002/0471140864.ps1104s40
- Djordjevic, M. A. (2004). Sinorhizobium meliloti metabolism in the root nodule: a proteomic perspective. *Proteomics* 4, 1859–1872. doi: 10.1002/pmic.200300802
- Djordjevic, M. A., Chen, H. C., Natera, S., Van Noorden, G., Menzel, C., Taylor, S., et al. (2003). A global analysis of protein expression profiles in Sinorhizobium meliloti: discovery of new genes for nodule occupancy and stress adaptation. *Mol. Plant Microbe Interact.* 16, 508–524. doi: 10.1094/mpmi.2003.16.6.508
- Djordjevic, M. A., Oakes, M., Li, D. X., Hwang, C. H., Hocart, C. H., and Gresshoff, P. M. (2007). The Glycine max xylem sap and apoplast proteome. *J. Proteome Res.* 6, 3771–3779. doi: 10.1021/pr0606833
- Girolamo, F. D., Ponzi, M., Crescenzi, M., Alessandrini, J., and Guadagni, F. (2010). A simple and effective method to analyze membrane proteins by SDS-PAGE and MALDI mass spectrometry. *Anticancer. Res.* 30, 1121–1129.
- Guillier, C., Cacas, J. L., Recorbet, G., Depretre, N., Mounier, A., Mongrand, S., et al. (2014). Direct purification of detergent-insoluble membranes from Medicago truncatula root microsomes: comparison between floatation and sedimentation. *BMC Plant Biol.* 14:255. doi: 10.1186/s12870-014-0255-x
- Hardy, J. (2017). Membrane damage is at the core of Alzheimer's disease. *Lancet Neurol.* 16:342. doi: 10.1016/s1474-4422(17)3009130091
- Hasegawa, T., Sugeno, N., Kikuchi, A., Baba, T., and Aoki, M. (2017). Membrane Trafficking Illuminates a Path to Parkinson's Disease. *Tohoku J. Exp. Med.* 242, 63–76. doi: 10.1620/tjem.242.63
- Hattori, N., Arano, T., Hatano, T., Mori, A., and Imai, Y. (2017). Mitochondrial-Associated Membranes in Parkinson's Disease. *Adv. Exp. Med. Biol.* 997, 157–169. doi: 10.1007/978-981-10-4567-7_12
- Hopkins, A. L., and Groom, C. R. (2002). The druggable genome. *Nat. Rev. Drug Discov.* 1, 727–730. doi: 10.1038/nrd892
- Huang, T.-K., Han, C.-L., Lin, S.-I., Chen, Y.-J., Tsai, Y.-C., Chen, Y.-R., et al. (2013). Identification of Downstream Components of Ubiquitin-Conjugating Enzyme PHOSPHATE2 by Quantitative Membrane Proteomics in Arabidopsis Roots. *Plant Cell* 25, 4044–4060. doi: 10.1105/tpc.113.115998
- Imin, N., Mohd-Radzman, N. A., Ogilvie, H. A., and Djordjevic, M. A. (2013). The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in Medicago truncatula. *J. Exp. Bot.* 64, 5395–5409. doi: 10.1093/jxb/ert369
- Imin, N., Patel, N., Corcilius, L., Payne, R. J., and Djordjevic, M. A. (2018). CLE peptide tri-arabinylation and peptide domain sequence composition are essential for SUNN-dependent autoregulation of nodulation in Medicago truncatula. *New Phytol.* 218, 73–80. doi: 10.1111/nph.15019
- Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F. U., Kerner, M. J., et al. (2008). Protein abundance profiling of the Escherichia coli cytosol. *BMC Genom.* 9:102. doi: 10.1186/1471-2164-9102
- Itzhak, D. N., Tyanova, S., Cox, J., and Borner, G. H. (2016). Global, quantitative and dynamic mapping of protein subcellular localization. *eLife* 5:e16950. doi: 10.7554/eLife.16950
- Kar, U. K., Simonian, M., and Whitelegge, J. P. (2017). Integral membrane proteins: bottom-up, top-down and structural proteomics. *Exp. Rev. Proteom.* 14, 715–723. doi: 10.1080/14789450.2017.1359545
- Kemmerling, B., Halter, T., Mazzotta, S., Mosher, S., and Nurnberger, T. (2011). A genome-wide survey for Arabidopsis leucine-rich repeat receptor kinases implicated in plant immunity. *Front. Plant Sci.* 2:88. doi: 10.3389/fpls.2011.00088
- Komatsu, S., Konishi, H., and Hashimoto, M. (2007). The proteomics of plant cell membranes. *J. Exp. Bot.* 58, 103–112. doi: 10.1093/jxb/erj209
- Kozioł, J., Griffin, N., Long, F., Li, Y., Latterich, M., and Schnitzer, J. (2013). On protein abundance distributions in complex mixtures. *Proteome Sci.* 11, 5–5. doi: 10.1186/1477-5956-1115
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Kulak, N. A., Geyer, P. E., and Mann, M. (2017). Loss-less Nano-fractionator for High Sensitivity, High Coverage Proteomics. *Mol. Cell. Proteom.* 16, 694–705. doi: 10.1074/mcp.O116.065136
- Kusumawati, L., Imin, N., and Djordjevic, M. A. (2008). Characterization of the secretome of suspension cultures of Medicago species reveals proteins important for defense and development. *J. Proteome Res.* 7, 4508–4520. doi: 10.1021/pr800291z
- Kusumi, A., Fujiwara, T. K., Tsunoyama, T. A., Kasai, R. S., Liu, A. A., Hirosawa, K. M., et al. (2020). Defining raft domains in the plasma membrane. *Traffic* 21, 106–137. doi: 10.1111/tra.12718
- Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105–132. doi: 10.1016/0022-2836(82)9051590510
- Laganowsky, A., Reading, E., Hopper, J. T. S., and Robinson, C. V. (2013). Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* 8, 639–651. doi: 10.1038/nprot.2013.024
- Lappano, R., and Maggolini, M. (2011). G protein-coupled receptors: novel targets for drug discovery in cancer. *Nat. Rev. Drug Discov.* 10, 47–60. doi: 10.1038/nrd3320
- Lee, J., Lei, Z., Watson, B. S., and Sumner, L. W. (2013). Sub-cellular proteomics of Medicago truncatula. *Front. Plant Sci.* 4:112. doi: 10.3389/fpls.2013.00112
- Link, A. J., and LaBaer, J. (2011). Trichloroacetic acid (TCA) precipitation of proteins. *Cold Spring Harb. Protoc.* 2011, 993–994. doi: 10.1101/pdb.prot5651
- Liu, J., and Last, R. L. (2015). A land plant-specific thylakoid membrane protein contributes to photosystem II maintenance in Arabidopsis thaliana. *Plant J.* 82, 731–743. doi: 10.1111/tj.12845
- Lohse, M., Nagel, A., Herter, T., May, P., Schroda, M., Zrenner, R., et al. (2014). Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. *Plant Cell Environ.* 37, 1250–1258. doi: 10.1111/pce.12231
- Long, R., Gao, Y., Sun, H., Zhang, T., Li, X., Li, M., et al. (2018). Quantitative proteomic analysis using iTRAQ to identify salt-responsive proteins during the germination stage of two Medicago species. *Sci. Rep.* 8:9553. doi: 10.1038/s41598-018-2793527938
- Lunn, G., and Sansone, E. B. (1985). Destruction of cyanogen bromide and inorganic cyanides. *Anal. Biochem.* 147, 245–250. doi: 10.1016/0003-2697(85)90034-x
- Magdeldin, S., Yoshida, Y., Li, H., Maeda, Y., Yokoyama, M., Enany, S., et al. (2012). Murine colon proteome and characterization of the protein pathways. *BioData Mining* 5:11. doi: 10.1186/1756-0381-511
- Marx, H., Minogue, C. E., Jayaraman, D., Richards, A. L., Kwiecien, N. W., Siahpirani, A. F., et al. (2016). A proteomic atlas of the legume Medicago truncatula and its nitrogen-fixing endosymbiont Sinorhizobium meliloti. *Nat. Biotechnol.* 34, 1198–1205. doi: 10.1038/nbt.3681
- Millioni, R., Tolin, S., Puricelli, L., Sbrignadello, S., Fadini, G. P., Tessari, P., et al. (2011). High abundance proteins depletion vs low abundance proteins enrichment: comparison of methods to reduce the plasma proteome complexity. *PLoS One* 6:e19603–e19603. doi: 10.1371/journal.pone.0019603
- Mirza, S. P., Halligan, B. D., Greene, A. S., and Olivier, M. (2007). Improved method for the analysis of membrane proteins by mass spectrometry. *Physiol. Genom.* 30, 89–94. doi: 10.1152/physiolgenomics.00279.2006
- Mohd-Radzman, N. A., Binos, S., Truong, T. T., Imin, N., Mariani, M., and Djordjevic, M. A. (2015). Novel MtCEP1 peptides produced in vivo differentially regulate root development in Medicago truncatula. *J. Exp. Bot.* 66, 5289–5300. doi: 10.1093/jxb/erv008
- Morel, J., Claverol, S., Mongrand, S., Furt, F., Fromentin, J., Bessoule, J., et al. (2006). Proteomics of Plant Detergent-resistant Membranes. *Mol. Cell. Proteom.* 5, 1396–1411. doi: 10.1074/mcp.M600044-MC P200
- Mosley, A. L., Sardi, M. E., Pattenden, S. G., Workman, J. L., Florens, L., and Washburn, M. P. (2011). Highly reproducible label free quantitative proteomic

- analysis of RNA polymerase complexes. *Mol. Cell. Proteom.* 10:M110.000687. doi: 10.1074/mcp.M110.000687
- Müller, F., Fischer, L., Chen, Z. A., Auchynnikava, T., and Rappsilber, J. (2018). On the Reproducibility of Label-Free Quantitative Cross-Linking/Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 29, 405–412. doi: 10.1007/s13361-017-18371832
- Natera, S. H., Guerreiro, N., and Djordjevic, M. A. (2000). Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis. *Mol. Plant Microbe Interact.* 13, 995–1009. doi: 10.1094/mpmi.2000.13.9.995
- Newby, Z. E. R., O'Connell, J. D., Gruswitz, F., Hays, F. A., Harries, W. E. C., Harwood, I. M., et al. (2009). A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nat. Protoc.* 4, 619–637. doi: 10.1038/nprot.2009.27
- Ogawa, M., Shinohara, H., Sakagami, Y., and Matsubayashi, Y. (2008). Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. *Science* 319:294. doi: 10.1126/science.1150083
- Okumura, M., Inoue, S., Kuwata, K., and Kinoshita, T. (2016). Photosynthesis Activates Plasma Membrane H⁺-ATPase via Sugar Accumulation. *Plant Physiol.* 171, 580–589. doi: 10.1104/pp.16.00355
- Olsen, J. V., Ong, S. E., and Mann, M. (2004). Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteom.* 3, 608–614. doi: 10.1074/mcp.T400003-MCP200
- Omasits, U., Ahrens, C. H., Muller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30, 884–886. doi: 10.1093/bioinformatics/btt607
- Osakabe, Y., Maruyama, K., Seki, M., Satou, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2005). Leucine-Rich Repeat Receptor-Like Kinase1 Is a Key Membrane-Bound Regulator of Abscisic Acid Early Signaling in Arabidopsis. *Plant Cell* 17, 1105–1119. doi: 10.1105/tpc.104.027474
- Pellizzaro, A., Clochard, T., Planchet, E., Limami, A. M., and Morere-Le Paven, M. C. (2015). Identification and molecular characterization of Medicago truncatula NRT2 and NAR2 families. *Physiol. Plant* 154, 256–269. doi: 10.1111/ppl.12314
- Quach, T. T., Li, N., Richards, D. P., Zheng, J., Keller, B. O., and Li, L. (2003). Development and applications of in-gel CNBr/tryptic digestion combined with mass spectrometry for the analysis of membrane proteins. *J. Proteome Res.* 2, 543–552. doi: 10.1021/pr0340126
- Rawlings, A. E. (2016). Membrane proteins: always an insoluble problem? *Biochem. Soc. Transact.* 44, 790–795. doi: 10.1042/BST20160025
- Reinke, A. W., Mak, R., Troemel, E. R., and Bennett, E. J. (2017). In vivo mapping of tissue- and subcellular-specific proteomes in *Caenorhabditis elegans*. *Sci. Adv.* 3, e1602426–e1602426. doi: 10.1126/sciadv.1602426
- Schey, K. L., Grey, A. C., and Nicklay, J. J. (2013). Mass spectrometry of membrane proteins: a focus on aquaporins. *Biochemistry* 52, 3807–3817. doi: 10.1021/bi301604j
- Seddon, A. M., Curnow, P., and Booth, P. J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochim. Biophys. Acta* 1666, 105–117. doi: 10.1016/j.bbmem.2004.04.011
- Song, W., Wang, B., Li, X., Wei, J., Chen, L., Zhang, D., et al. (2015). Identification of Immune Related LRR-Containing Genes in Maize (*Zea mays* L.) by *Genome-Wide Sequence Analysis*. *Int. J. Genom.* 2015:231358. doi: 10.1155/2015/231358
- Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 6, 175–182.
- Sperschneider, J., Catanzariti, A. M., DeBoer, K., Petre, B., Gardiner, D. M., Singh, K. B., et al. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci. Rep.* 7:44598. doi: 10.1038/srep44598
- Sun, B., Liu, Z., Fang, Z., Dong, W., Yu, Y., Ye, M., et al. (2020). Probing the Proteomics Dark Regions by VAILase Cleavage at Aliphatic Amino Acids. *Anal. Chem.* 92, 2770–2777. doi: 10.1021/acs.analchem.9b05048
- Szablewski, L. (2017). Glucose Transporters in Brain: In Health and in Alzheimer's Disease. *J. Alzheimers Dis.* 55, 1307–1320. doi: 10.3233/jad-160841
- Tanca, A., Biosa, G., Pagnozzi, D., Addis, M. F., and Uzzau, S. (2013). Comparison of detergent-based sample preparation workflows for LTQ-Orbitrap analysis of the *Escherichia coli*. *Proteomics* 13, 2597–2607. doi: 10.1002/pmic.2012.00478
- Tautermann, C. S. (2014). GPCR structures in drug design, emerging opportunities with new structures. *Bioorg. Med. Chem. Lett.* 24, 4073–4079. doi: 10.1016/j.bmcl.2014.07.009
- ten Hove, C. A., Bochdanovits, Z., Jansweijer, V. M. A., Koning, F. G., Berke, L., Sanchez-Perez, G. F., et al. (2011). Probing the roles of LRR RLK genes in Arabidopsis thaliana roots using a custom T-DNA insertion set. *Plant Mol. Biol.* 76, 69–83. doi: 10.1007/s11103-011-9769-x
- Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37, 914–939. doi: 10.1111/j.1365-313x.2004.02016.x
- van der Knaap, E., Song, W. Y., Ruan, D. L., Sauter, M., Ronald, P. C., and Kende, H. (1999). Expression of a gibberellin-induced leucine-rich repeat receptor-like protein kinase in deepwater rice and its interaction with kinase-associated protein phosphatase. *Plant Physiol.* 120, 559–570. doi: 10.1104/pp.120.2.559
- Vit, O., and Petrak, J. (2017). Integral membrane proteins in proteomics. How to break open the black box? *J. Proteomics* 153, 8–20. doi: 10.1016/j.jpro.2016.08.006
- Vizcaino, J. A., Côté, R. G., Csordas, A., Dianas, J. A., Fabregat, A., Foster, J. M., et al. (2012). The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 41, D1063–D1069. doi: 10.1093/nar/gks1262
- Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., et al. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* 32, 223–226. doi: 10.1038/nbt.2839
- Washburn, M. P., Wolters, D., and Yates, J. R. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19, 242–247. doi: 10.1038/85686
- Wong, S. H., and King, C. Y. (2015). Amino Acid Proximities in Two Sup35 Prion Strains Revealed by Chemical Cross-linking. *J. Biol. Chem.* 290, 25062–25071. doi: 10.1074/jbc.M115.676379
- Yeung, Y. G., Nieves, E., Angeletti, R. H., and Stanley, E. R. (2008). Removal of detergents from protein digests for mass spectrometry analysis. *Anal. Biochem.* 382, 135–137. doi: 10.1016/j.ab.2008.07.034
- Zhang, K., McKinlay, C., Hocart, C. H., and Djordjevic, M. A. (2006). The Medicago truncatula small protein proteome and peptidome. *J. Proteome Res.* 5, 3355–3367. doi: 10.1021/pr060336t
- Zhang, N., and Li, L. (2004). Effects of common surfactants on protein digestion and matrix-assisted laser desorption/ionization mass spectrometric analysis of the digested peptides using two-layer sample preparation. *Rap. Commun. Mass Spectrom.* 18, 889–896. doi: 10.1002/rcm.1423
- Zhao, Q., Liang, Y., Yuan, H., Sui, Z., Wu, Q., Liang, Z., et al. (2013). Biphasic microreactor for efficient membrane protein pretreatment with a combination of formic acid assisted solubilization, on-column pH adjustment, reduction, alkylation, and tryptic digestion. *Anal. Chem.* 85, 8507–8512. doi: 10.1021/ac402076u
- Zheng, S., and Doucette, A. A. (2016). Preventing N- and O-formylation of proteins when incubated in concentrated formic acid. *Proteomics* 16, 1059–1068. doi: 10.1002/pmic.201500366

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Lee, Carroll, Crossett, Connolly, Batarseh and Djordjevic. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.