MEDWIN PUBLISHERS
Committed to create value for Researchers

Bioinformatics & Proteomics Open Access Journal

Plant Cell Wall Proteomics: An Assessment Twenty Years after Launching

Jamet E*

Plant Science Research Laboratory, University of Toulouse, CNRS, UPS, Castanet-Tolosan, France

***Corresponding author:** Elisabeth Jamet, Plant Science Research Laboratory, University of Toulouse, CNRS, UPS, Castanet Tolosan, France, Email: jamet@lrsv.ups-tlse.fr

Mini Review

Volume 1 Issue 2

Received Date: September 27, 2017

Published Date: October 20, 2017

Abstract

Cell walls are complex structures surrounding plant cells. They provide not only mechanical support and protection against environmental changes, but also a mean for cell-to-cell communication. They are mainly constituted of polysaccharides (about 90% of their mass) and proteins. Cell wall proteins (CWPs) play critical roles because they contribute to the plasticity of the cell wall architecture during development and in response to biotic and abiotic environmental changes. Their systematic identification has started in the 2000's with the sequencing of the genome of the *Arabidopsis thaliana* model plant and the development of adapted mass spectrometry (MS) technologies. Since then, many other plants have been studied among which plants of agronomical interest. The description of cell wall proteomes has fully benefited not only from the improvement of MS technologies, but also from better sample preparation and peptide separation prior to MS analysis. Bioinformatics has also played critical roles by designing software allowing protein identification, annotation and quantification, as well creating MS data repositories.

Keywords: Cell wall proteins; Mass Spectrometry; Polysaccharides; *Arabidopsis thaliana*

Introduction

Plant cell walls constitute an extracellular compartment playing many roles during development and adaptation to environmental biotic and abiotic changes [1-4]. They are named primary cell walls as long as cells are growing, and they become secondary cell walls when cells differentiate to specific functions. The major components of primary cell walls are polysaccharides which fall into three categories: cellulose, hemicelluloses and pectins [5]. Secondary cell walls may contain additional polymers such as lignins [6]. All these compounds are present in different proportions depending on plant tissues and on plant species and their structure can be modified thanks to cell wall proteins (CWPs). Although present in minor amounts in cell walls, CWPs play major roles in cell wall structure, plasticity and signaling [7-9]. CWPs can degrade, ligate or even modify polysaccharides, e.g. polygalacturonases which are able to degrade pectic homogalacturonans [10], xyloglucan endo-transglycosyl hydrolases (XTHs) which cut and religate hemicellulosic xyloglucans [11], and pectin methylesterases (PMEs) which demethylate homogalacturonans [12]. These modifications have consequences on the properties of cell walls. For a long time, cell walls models only included the so-called

structural proteins which were assumed to form covalently linked networks giving rigidity to cell walls and protecting cells from pathogen invasion [5,13]. Besides, a few CWP families were characterized from biochemical studies. However, a clear picture of the cell wall proteome was not available, thus leading to an under-estimation of the physiological roles of cell walls.

By the end of the 1990's, proteomics studies started to develop thanks to impressive progresses in mass spectrometry (MS) technologies and to the availability of newly sequenced genomes. The first plant genome to be sequenced was that of the model plant *Arabidopsis thaliana* [14]. Then, the genomes of several plants of economical interest became available, such as those of rice [15], tomato [16], *Medicago truncatula* [17], and linum [18]. Numerous proteomics studies have then published among which those restricted to organelles, such as chloroplasts [19], mitochondria [20] and nuclei [21]. The case of cell walls was more puzzling because cell walls are an open compartment difficult to purify [22]. It contains minor amount of proteins and can be easily contaminated by intracellular proteins including the very abundant photosynthesis proteins. This mini-review aims at summarizing the main steps of the successful story of plant cell wall proteomics which started about twenty years ago with the first description of a small cell wall proteome by Robertson et al. (1997) [23].

In this founding article, cell suspension cultures were washed with salt solutions in order to elute CWPs from their walls without breaking their plasma membranes, thus using a so-called non-destructive method [23]. Twenty proteins were identified after separation by 1D-electrophoresis (1D-E) and Edman N-terminal sequencing. Later on, many other strategies were designed to increase the size of cell wall proteomes [24]. They include: (i) the improvement of cell wall purification procedures to limit the contamination by intracellular proteins in the so-called destructive methods [25]; (ii) the diversification of the salt solutions used to elute CWPs [26,27]; (iii) the introduction of affinity chromatography to separate proteins according to their charge or to their *N*-glycosylation status [28-32]; (iv) the use of 1D-E for protein separation prior to tryptic digestion because CWPs are mostly basic glycoproteins poorly separated by 2D-E [33]; (v) the use of combinatorial peptide ligand library (CPLL) chromatography to get access to minor CWPs [34,35]. More recently, shotgun analyses omitting the protein separation step have been successfully performed [36] and the first systematic quantitative analysis of CWPs has been published [37]. Altogether, many different types of organs have been analyzed including roots, leaves, inflorescences, fruits, seeds and cell suspension cultures (for an overview, see WallProtDB, www.polebio.lrsv.ups-tlse.fr/WallProtDB/). Plant cell wall proteomics studies have also greatly benefited from the improvement of peptide separation prior to MS analysis and the increase performance of mass spectrometers, from MALDI-TOF MS to LC-MS/MS. Finally, the careful bioinformatics annotation of the identified proteins has allowed (i) better distinguishing proteins predicted to be secreted from intracellular proteins, and (ii) assigning predicted functions to more than 85% of the identified proteins. In this purpose, two bioinformatics tools have been designed in our team: (i) ProtAnnDB (www.polebio.lrsv.ups-tlse.fr/ProtAnnDB/) is a pipeline of prediction of sub-cellular localization and functional domain using bioinformatics programs publicly available [38]; (ii) WallProtDB is a plant cell wall proteomics database collecting published cell wall proteomes after curated annotation of the identified proteins, based on the presence of predicted functional domains and on experimental work [39]. To facilitate the comparisons between cell wall proteomes, CWPs are grouped in nine functional classes in WallProtDB [33]: (i) proteins acting on polysaccharides like glycoside hydrolases [7,40]; (ii) oxido-reductases like class III peroxidases [41]; (iii) proteases [42]; (iv) proteins possibly related to lipid metabolism like non-specific lipid transfer proteins (ns-LTPs) [43]; (v) proteins possibly involved in signaling like arabinogalactan proteins (AGPs) [44]; (vi) proteins having interacting domains with proteins or polysaccharides like lectins [45]; (vii) structural proteins like extensins [46]; (viii) miscellaneous proteins; and (ix) proteins of yet unknown function. It should be mentioned that the functional class comprising the structural proteins is the smallest because of the difficulty to extract such proteins which are covalently linked to the other cell wall components [47,48]. This grouping helps getting an overview of newly described cell wall proteomes, but as all classifications, it has limitations and it has to evolve to take into account newly characterized proteins.

Nowadays, the size of newly described organ cell wall proteomes has been increased to 250 to 400 CWPs [27,34,36], i.e. at the most twenty times as large as the first described cell wall proteome. Presently, the larger plant cell wall proteome is that of *A. thaliana* which comprises more than 900 CWPs and covers about half of the predicted one (see WallProtDB). The second larger cell wall proteome is that of *Brachypodium distachyon*, which a monocot model plant, with nearly 600 CWPs.

CWPs route through the secretion pathway where they undergo post-translational modifications (PTMs) [49]. The best described are *N*-glycosylation, Pro hydroxylation and *O*-glycosylation. All these PTMs can be critical for protein structure and/or biological activity. Proteomics has also brought new information regarding this PTMs. Affinity chromatography on the concanavalin A (ConA) lectin has allowed the separation of *N*-glycoproteins and their identification. In some cases, the location and structure of *N*-glycans could also be described [29,30,50]. More recently, some studies have addressed the question of the localization of hydroxyproline (Hyp) residues resulting from the hydroxylation of Pro residues by prolyl-hydroxylation. Hyp residues were initially described in structural proteins such as extensins and in AGPs [51-53]. It seems that they are present in more protein families and that there is some variability in their distribution, so that it is yet difficult to propose a universal Pro hydroxylation code [34,54]. Finally, *O*-glycosylations are very difficult to describe because of the complexity of the structure of *O*-glycans. They require dedicated studies and thus cannot yet been included in omics strategies [55,56].

The importance of bioinformatics in proteomics flowcharts needs to be stressed (Figure 1). First of all, the access to annotated genomes with well-predicted open reading frames is critical. Prediction of exon-intron junctions has to be accurate and, if possible, curated with sequenced RNAs as for RefSeq sequences [57]. Indeed, MS-based proteomics is based on the comparison between mass lists of tryptic peptides, which are fragmented or not depending on the MS instrument, and theoretical mass lists calculated from the protein sequences present in databases [58,59]. Then, as stated above, when proteins are identified, it can be important to predict both their sub-cellular localization, especially to check the quality of sub-cellular proteomics experiments, and their functional domains [60,61]. This information allows assessing the biological role of proteins of interest and designing relevant experiments to demonstrate it [38,62]. Finally, public data repositories have been set up to permit sharing MS data [63].

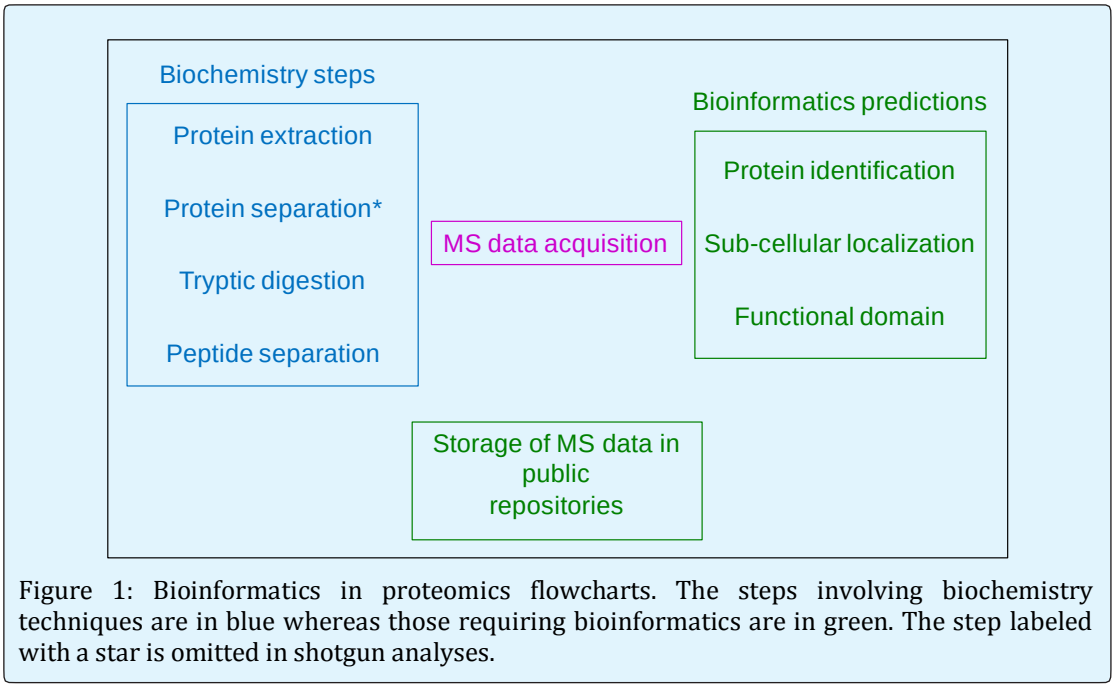

Figure 1: Bioinformatics in proteomics flowcharts. The steps involving biochemistry techniques are in blue whereas those requiring bioinformatics are in green. The step labeled with a star is omitted in shotgun analyses.

Altogether, cell wall proteomics studies have brought a large amount of information during the twenty last years and they have provided an overview of cell wall proteomes. In particular, many more proteins families playing roles in polysaccharide modifications are included in the present description of cell walls [7]. In the mean time, a lot of new functions have been described for CWPs thanks to genetics studies. For example, the role of CWPs in cell-to-cell communication has been illustrated by several examples including the release of signaling peptides by extracellular proteases to induce cell differentiation or plant defense mechanisms [64]. Next important issues in plant cell wall proteomics concern the identification of more structural CWPs, a better characterization of CWP PTMs, advances in the description of the extracellular peptidome and the

systematic quantification of proteins. Finally, the question of the existence of alternative non-conventional routes of protein secretion also needs to be solved [65]. Indeed, all the published cell wall proteomes mention the presence of proteins predicted to be intracellular [24]. However, to our knowledge, only a sunflower jacalin devoid of predicted signal peptide, has been shown to be secreted through the release of exosomes [66]. Some alternative mechanisms have been described in animal cells [67], but remain to be established for plant cells.

Acknowledgments

The author is thankful to the Paul Sabatier-Toulouse 3 University and to the Centre National de la Recherche Scientifique (CNRS) for their financial support. This work was also granted by the IdEx project (ANR-11-IDEX-0002-02). The author wishes to thank her present and former students and colleagues for fruitful collaborations. She also apologizes for the many articles of interest which could not be quoted in this short article.

References

1. Le Gall H, Philippe F, Domon J, Gillet F, Pelloux J, et al. (2015) Cell wall metabolism in response to abiotic stress. Plants 4(1): 112-166.

2. Tenhaken R (2015) Cell wall remodeling under biotic stress. Front Plant Sci 5: 771.

3. Largo-Gosens A, Hernández-Altamirano M, García-Calvo L, Alonso-Simón A, Alvarez J, et al. (2014) Fourier transform mid infrared spectroscopy applications for monitoring the structural plasticity of plant cell walls. Front Plant Sci 5: 303.

4. Somssich M, Khan G, Persson S (2016) Cell wall heterogeneity in root development of Arabidopsis. Front Plant Sci 7: 1242.

5. Carpita N, Gibeaut D (1993) Structural models of primary cell walls in flowering plants, consistency of molecular structure with the physical properties of the walls during growth. Plant J 3(1): 1-30.

6. Zhao Q (2016) Lignification: Flexibility, Biosynthesis and Regulation. Trends Plant Sci 21(8): 713-721.

7. Franková L, Fry J (2013) Biochemistry and physiological roles of enzymes that 'cut and paste' plant cell-wall polysaccharides. J Exp Bot 64(12): 3519-3550.

8. Cosgrove D (2015) Plant cell wall extensibility: connecting plant cell growth with cell wall structure, mechanics, and the action of wall-modifying enzymes. J Exp Bot 67(2): 463-476.

9. Voxeur A, Höfte H (2016) Cell wall integrity signaling in plants: "To grow or not to grow that's the question". Glycobiology 26(9): 950-960

10. Sénéchal F, Wattier C, Rustérucci C, Pelloux J (2014) Homogalacturonan-modifying enzymes: structure, expression, and roles in plants. J Exp Bot 65(18): 5125-5160.

11. Cosgrove D (2016) Catalysts of plant cell wall loosening. F1000Res2016 5(F1000 Faculty Rev): 119.

12. Pelloux J, Rustérucci C, Mellerowicz E (2017) New insights into pectin methylesterase structure and function. Trends Plant Sci 12(6): 267-277.

13. Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol 6(11): 850-861.

14. The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408(6814): 796-805.

15. International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436 (7052): 793-800.

16. Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485(7400): 635-641.

17. Cannon S, Crow J, Heuer M, Wang X, Cannon E, et al. (2004) Databases and information integration for the Medicago truncatula genome and transcriptome. Plant Physiol 138(1): 38-46.

18. Wang Z, Hobson N, Galindo L, Zhu S, Shi D, et al. (2012) The genome of flax (*Linum usitatissimum*) assembled de novo from short shotgun sequence reads. Plant J 72(3): 461-473.

19. Salvi D, Rolland N, Joyard J, Ferro M (2008) Purification and proteomic analysis of chloroplasts and their sub-organellar compartments. Methods Mol Biol 432: 19-36.

20. Millar A, Liddell A, Leaver C (2007) Isolation and sub fractionation of mitochondria from plants. Methods Cell Biol 80: 65-90.

21. Petrovská B, Šebela M, Doležel J (2015) Inside a plant nucleus: discovering the proteins. J Exp Bot 66(6): 1627-1640.

22. Canut H, Albenne C, Jamet E (2017) Isolation of the cell wall. Methods Mol Biol 1511: 171-185.

23. Robertson D, Mitchell GP, Gilroy JS, Gerrish C, Bolwell GP, et al. (1997) Differential extraction and protein sequencing reveals major differences in patterns of primary cell wall proteins from plants. J. Biol. Chem. 272(25): 15841-15848.

24. Albenne C, Canut H, Jamet E (2013) Plant cell wall proteomics: the leadership of *Arabidopsis thaliana*. Front Plant Sci 4: 111.

25. Feiz L, Irshad M, Pont-Lezica RF, Canut H, Jamet E (2006) Evaluation of cell wall preparations for proteomics: a new procedure for purifying cell walls from Arabidopsis hypocotyls. Plant Methods 2: 10

26. Boudart G, Jamet E, Rossignol M, Lafitte C, Borderies G, et al. (2005) Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: Identification by mass spectrometry and bioinformatics. Proteomics 5(1): 212-221.

27. Printz B, Dos Santos Morais R, Wienkoop S, Sergeant K, et al. (2015) An improved protocol to study the plant cell wall proteome. Front Plant Sci 6: 237.

28. Irshad M, Canut H, Borderies G, Pont-Lezica R, Jamet E (2008) A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. BMC Plant Biol 8: 94.

29. Zhang Y, Giboulot A, Zivy M, Valot B, Jamet E, et al. (2011) Combining various strategies to increase the coverage of the plant cell wall glycoproteome. Phytochemistry 72(10): 1109-1123.

30. Ruiz-May E, Hucko S, Howe K, Zhang S, Sherwood R, et al. (2014) A comparative study of lectin affinity based plant *N*-glycoproteome profiling using tomato fruit as a model. Mol Cell Proteomics 13(2): 566-579.

31. Minic Z, Jamet E, Negroni L, der Garabedian PA, Zivy M, et al. (2007) A sub-proteome of *Arabidopsis thaliana* trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. J Exp Bot 58(10): 2503-2512.

32. Xu S, Medzihradszky K, Wang Z, Burlingame A, Chalkley R (2016) *N*-Glycopeptide profiling in Arabidopsis inflorescence. Mol Cell Proteomics 15(6): 2048-2054.

33. Jamet E, Albenne C, Boudart G, Irshad M, Canut H, et al. (2008) Recent advances in plant cell wall proteomics. Proteomics 8(4): 893-908.

34. Nguyen-Kim H, San Clemente H, Balliau T, Zivy M, Dunand C, et al. (2016) *Arabidopsis thaliana* root cell wall proteomics: Increasing the proteome coverage using a combinatorial peptide ligand library and description of unexpected Hyp in peroxidase amino acid sequences. Proteomics 16(3): 491-503.

35. Righetti PG, Boschetti E (2013) Combinatorial peptide libraries to overcome the classical affinity-enrichment methods in proteomics. Amino Acids 45(2): 219-229.

36. Hervé V, Duruflé H, San Clemente H, Albenne C, Balliau T, et al. (2016) An enlarged cell wall proteome of *Arabidopsis thaliana* rosettes. Proteomics 16(24): 3183-3187.

37. Duruflé H, Hervé V, Ranocha P, Balliau T, Zivy M, et al. (2017) Cell wall adaptation of two contrasted ecotypes of *Arabidopsis thaliana*, Col and Sha, to sub-optimal growth conditions: an integrative study. Plant Sci 263: 183-193.

38. San Clemente H, Pont-Lezica R, Jamet E (2009) Bioinformatics as a tool for assessing the quality of sub-cellular proteomic strategies and inferring functions of proteins: plant cell wall proteomics as a test case. Bioinform Biol Insights 3: 15-28

39. San Clemente H, Jamet E (2015) Wall Prot DB, a database resource for plant cell wall proteomics. Plant Methods 11(1): 2.

40. Lombard V, Golaconda Ramulu H, Drula E, Coutinho P, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42(Database issue): D490-D495.

41. Francoz E, Ranocha P, Nguyen-Kim H, Jamet E, Burlat V, et al. (2015) Roles of cell wall peroxidases in plant development. Phytochemistry 112: 15-21.

42. Van der Hoorn R (2008) Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol 59: 191-223.

43. Jacq A, Burlat V, Jamet E (2017) Plant cell wall proteomics as a strategy to reveal candidate proteins involved in extracellular lipid metabolism. Curr Protein Pept Sci in press

44. Seifert GJ, Roberts K (2007) The biology of arabinogalactan proteins. Annu Rev Plant Biol 58: 137-161.

45. Bellande K, Bono J-J, Savelli B, Jamet E, Canut H (2017) Plant lectins and lectin receptor-like kinases: How do they sense the outside? Int J Mol Sci 18(6): 1164.

46. Showalter AM, Keppler B, Lichtenberg J, Gu D, Welch LR (2010) A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. Plant Physiol 153(2): 485-513.

47. Chen Y, Ye D, Held MA, Cannon MC, Tui R, et al. (2015) Identification of the abundant hydroxyproline-rich glycoproteins in the root walls of wild-type Arabidopsis, an *ext3* mutant line, and its phenotypic revertant. Plants 4(1): 85-111.

48. Tan L, Eberhard S, Pattathil S, Warder C, Glushka J, et al. (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. Plant Cell 25(1): 270-287.

49. Kim S, Brandizzi F (2016) The plant secretory pathway for the trafficking of cell wall polysaccharides and glycoproteins. Glycobiology 26(9): 940-949.

50. Zhang M, Chen G, Lv D, Li X, Yan Y (2015) *N*-linked glycoproteome profiling of seedling leaf in *Brachypodium distachyon* L. J Proteome Res 14(4): 1727-1738.

51. Kieliszewski MJ (2001) The latest hype on Hyp-*O*-glycosylation codes. Phytochemistry 57(3): 319-323.

52. Kieliszewski MJ, Lamport DTA (1994) Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. Plant J 5(2): 157-172.

53. Tan L, Leykam JF, Kieliszewski MJ (2003) Glycosylation motifs that direct arabinogalactan addition to arabinogalactan-proteins. Plant Physiol 132(3): 1362-1369.

54. Canut H, Albenne C, Jamet E (2016) Post-translational modifications of plant cell wall proteins and peptides: A survey from a proteomics point of view. Biochim Biophys Acta 1864(8): 983-990.

55. Tan L, Varnai P, Lamport DT, Yuan C, Xu J, et al. (2010) Plant *O*-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short bifurcated side chains. J Biol Chem 285: 24575-24583.

56. Léonard R, Petersen BO, Himly M, Kaar W, Wopfner N, et al. (2005) Two novel types of *O*-glycans on the mugwort pollen allergen Art v 1 and their role in antibody binding. J Biol Chem 280(9): 7932-7940.

57. O'Leary N, Wright M, Brister J, Ciufo S, Haddad D, et al. (2016) Reference sequence (Ref Seq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44(D1): D733-D745.

58. Perkins D, Pappin D, Creasy D, Cottrell J (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20(18): 3551-3567.

59. Craig R, Cortens J, Beavis R (2004) Open source system for analyzing, validating, and storing protein identification data. J Proteome Res 3(6): 1234-1242.

60. Emanuelsson O, Brunak S, Von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2(4): 953-971.

61. Jones P, Binns D, Chang HY, Fraser M, Li W, et al. (2014) Inter Pro Scan 5: genome-scale protein function classification. Bioinformatics 30(9): 1236-1240.

62. Jamet E (2004) Bioinformatics as a critical prerequisite to transcriptome and proteome studies. J Exp Bot 55(406): 197-1979.

63. Perez-Riverol Y, Alpi E, Wang R, Hermjakob H, Vizcaíno J (2015) Making proteomics data accessible and reusable: current state of proteomics databases and repositories. Proteomics 15(5-6): 930-949.

64. Tavormina P, De Coninck B, Nikonorova N, De Smet I, Cammue B (2015) The plant peptidome: An expanding repertoire of structural features and biological functions. Plant Cell 27(8): 2095-2118.

65. Rose JKC, Lee S-J (2010) Straying off the highway: Trafficking of secreted plant proteins and complexity in the plant cell wall proteome. Plant Physiol. Biochem 153(2): 433-436.

66. Regente M, Pinedo M, Elizalde M, De la Canal L (2012) Apoplastic exosome-like vesicles: a new way of protein secretion in plants. Plant Signal Behav 7(5): 544-546.

67. Nickel W, Seedorf M (2008) Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. Ann Rev Cell Dev Biol 24: 287-308.