An In-Silico Approach to Identify Ripening Related β-1,3-Glucanases and Their Role in Tomato Fruit †

Louisa Perry, Yoselin Benitez-Alfonso * and Candelas Paniagua *

Centre for Plant Sciences, Faculty of Biological Sciences, University of Leeds, Leeds, UK;
louisaperry@outlook.com

* Correspondence: y.benitez-alfonso@leeds.ac.uk (Y.B.-A.); C.Paniagua@leeds.ac.uk (C.P.);
  Tel.: 0113-34-32811

† Presented at the 1st International Electronic Conference on Plant Science, 1–15 December 2020; Available
  online: https://iecps2020.sciforum.net/.

Abstract: Tomato, *Solanum lycopersicum*, is one of the most cultivated fruits. However, between one quarter and half of the production is lost due to uncontrolled conditions during transport and storage. Understanding the biological components and processes that affect the physical properties and structure of tomato fruits during the post-harvest is essential to design new strategies to reduce these losses. The aim of this work is to identify specific cell wall β-1,3-glucanases expressed during tomato fruit ripening using a combination of phylogeny and transcriptomic analysis. Fifty β-1,3-glucanases (BG) were identified in-silico using bioinformatic tools. Phylogenetic analysis revealed that tomato genes are distributed in three clusters (α, β and γ) with evolutionary relations previously characterized in the model *Arabidopsis thaliana*. In Arabidopsis, cluster α comprises all enzymes identified to target intercellular channels named plasmodesmata whereas pathogenesis and stress responsive apoplastic enzymes are grouped in cluster β and γ. Analysis of tomato microarray data showed two types of regulation for BGs expressed in tomato fruit: enzymes in cluster α decreased their expression during the ripening, while enzymes in clusters β and γ increased their expression in mature fruits. The results suggest differential regulation for enzymes in different phylogenetic clusters suggesting evolutionary divergences in function. The analysis also provide new candidates for cell wall modifications that can be further studied as a possible targets in fruit improvement programmes.

Keywords: β-1,3-glucanases; shelf-life; tomato; cell wall

1. Introduction

The tomato fruit, *Solanum lycopersicum*, is, according to the Food and Agriculture Organization (FAO), the second most cultivated fruit worldwide with 243 million tonnes of fruit produced in 2018 [1]. Exports of tomato fruits from EU countries accounted for 17.7% of all vegetable outputs in the period 2015–2017 [2]. In spite of this, losses in production are considerable and are estimated to range between 24–50% in a single year [3]. Poor conditions during transport and storage from field to market are the main causes for these losses which are influenced by the lack of temperature control, physical injury, pests and disease [3]. Understanding the biological components and processes that affect the shelf-life of the tomato fruit is essential to design new strategies to reduce these losses.

One of the main processes that influences the shelf-life of tomato fruit is the rapid softening during ripening. Softening is a biological process, in which the dissasembly of cell wall (CW), cell turgor, fruit-water status, hydrostatic pressure and the accumulation and distribution of osmotically active solutes are involved [4]. Delaying this process is one of the major targets in fruit breeding

programmes. Research on this area focuses on introducing changes in the expression of specific cell wall modifying enzymes (CWME) that target major cell wall components (such as pectins and cellulose) [5]. However, the role of minor components, such as callose (β-1,3-glucan), in tomato ripening and post-harvest shelf-life remains unclear. Callose is a plant cell wall polysaccharide synthesised by callose synthases (Cals) and hydrolysed by β-1,3-glucanases (BG) [6]. BGs are a group of proteins belonging to the glycosyl hydrolase family 17 (GH17) that act in a wide variety of functions in plants [7]. Callose synthesis/degradation has been involved in disease response, in the control of intercellular transport and on the regulation of cell wall integrity and mechanical properties [8].

The objective of this study is to identify, via an in-silico approach, candidate BG enzymes participating in tomato ripening. To identify the candidate BGs genes, this study analysed predicted β-1,3-glucanase structural features (presence of protein sequence domains and catalytic residues); carried out a phylogenetic comparison with known and predicted *Arabidopsis thaliana* BG genes; analysed the expression of BG genes in tomato using publically available transcriptomic data; and used bioinformatic predictions for cellular and plant localization.

2. Methods

2.1. Isolation of β-1,3-Glucanase Gene Sequences and Prediction of Domains and Catalytic Residues

Glycoside hydrolase family 17 (GH17) proteins of *Arabidopsis thaliana* and *Solanum lycopersicum* were downloaded from Phytozome v9.022 under the search term 'glyco_hydro_17' to identify genes which were likely to contain GH17 domains. Presence of a GH17 domain was confirmed in these sequences using the SMART [9], Conserved domain [10] and Interpro [11] prediction tools. To identify and remove sequence redundancy, the UGENE [12] bioinformatics software was used to align the obtained sequences using the MUSCLE [13] alignment function.

Further characteristics were identified which are associated with GH17 family proteins. Presence of signal peptides (SP) were predicted using Aramemnon [14], SignalP v5.0 [15], Spoctopus [16], Predisi [17], SMART [16] and Interpro [11] bioinformatic tools. Presence of GIP-anchor attachment sites were predicted using KohGPI [18], PredGPI [19] and BigP [20] bioinformatic tools. Presence of X8 domains were predicted using SMART [16], NCBI CDS [10] and Interpro [11] bioinformatic tools.

Conservation of glutamic acid residues E95 and E236 [21] were identified in Arabidopsis β-1,3-glucanases by Gaudioso-Pedraza and Benitez-Alfonso (2014) in 'Supplementary datasheet 1' [22]. The identified tomato β-1,3 glucanases were aligned with the Arabidopsis genes to determine the position of the catalytic residues in tomato.

2.2. Comparative Phylogenetic Analysis of Tomato and Arabidopsis glucanases

Arabidopsis thaliana GH17 genes identified in Gaudioso-Pedraza and Benitez-Alfonso (2014) [22] were aligned in UGENE [12] using the MUSCLE [13] alignment tool with the *Solanum lycopersicum* genes. Phylogenetic analysis was carried out in MEGA-X 10.1.840 software using the Jones-Taylor-Thornton (JTT) model. Conditions and restrictions were varied to produce a range of dendrograms with altered bootstrap values (between 100–1000), branch swap filter settings (none or very weak) and ML Heuristic Method settings (Nearest-neighbour-interchange (NNI)) or Subtree-pruning-regrafting-extensive (SPR level 5). Additionally, analysis was repeated using the same settings for full sequences and the domain only sequences. *Arabidopsis thaliana* GH17 genes with common names were annotated and homologous *Solanum lycopersicum* genes, which showed consistent close relations with the *Arabidopsis thaliana* GH17 genes across the different dendograms, were predicted to exhibit similar properties. Confidence of dendograms was related to the clustering of the known *Arabidopsis thaliana* genes based on Doxey et al. (2007) [23].

2.3. Gene Expression Profile Analysis

Microarrays for *Solanum lycopersicum* fruit development were downloaded from the ArrayExpress [24] database. Arrays were selected and normalised using an in-house R script,

developed by Sam Amsbury and Philip Kirk, to look for differentially expressed genes in tomato fruits. Pericarp expression and fruit development data was downloaded directly from the SGN Tomato Expression Atlas (TEA) [25] and TomExpress [26]. Expression trends were analysed and compared between TEA, TomExpress and Genevestigator [27] data to validate results. Gene intracellular localisations were predicted using PSORT [28] and WoLF [29] prediction tools and analysed to identify cell wall proteins. Tissue localisation expression data was downloaded from Genevestigator [27] to analyse fruit specific and non-specific genes and compared to predicted localisation using the ePlant BAR [30] database.

3. Results

3.1. Phylogeny Find Tomato Genes Closely Related to Known Arabidopsis BG Genes

In order to identify similar funcion of the BG from tomato plants with known Arabidopsis BGs genes, a phylogenetic tree was created (Figure 1). Identification of GH17 containing BGs in the tomato plant was done using Phytozome v12.122 database. A total of 51 genes were identified. Those genes were aligned using the UGENE [12] software to identify redundant sequences. Two genes, Solyc01g059980 and Solyc01g060020 were found to encode the same amino acid sequence, thus Solyc01g060020 was removed from the analysis.

The phylogenetic tree (Figure 1) showed that the BGs genes from tomato are categorised into three clusters, α, β and γ, as it has been shown for Arabidopsis in Doxey et al. (2007) [23]. In Arabidopsis, cluster α comprises all the proteins previously identified to localize at cell wall microdomains named plasmodesmata (PDBG), while enzymes in clusters γ include pathogenesis-induced proteins (BG1-BG5) [***Error! Bookmark not defined.***]. This suggests that tomato genes residing in the α cluster may function at plasmodesmata, thus controlling cell-to-cell signalling, while genes included in cluster γ could be related to pathogenesis. Phylogenetic analysis showed Plasmodesmal-localised BG 1 (PDBG1), PDBG2 and PDBG3 closely related to Solyc01g005830, Solyc04g007910 and Solyc08005000 tomato genes with probabilities between 90 to 70 bootstrap values. Pathogenesis related (PR) B-glucanase 1 (BG1), BG2, and BG3 genes showed to be closely related with tomato genes Solyc10g079860 and Solyc04g016470 (91 bootstrap value).

Additionally, the 50 identified tomato protein sequences were screened for the common BG domains; signal peptide (SP), the callose binding X8, and glycosylphosphatidylinositol (GPI) anchor, as described in Materials and Methods, to predict the structure and localisation of the proteins (Figure 1). Based on the data from the prediction tools, all identified genes contained a SP, with the exception of the γ cluster genes Solyc11g065300, Solyc11g065280, Solyc01g060010 and Solyc00g202560, and the α cluster gene Solyc06g073710. Research has identified two conserved glutamate (E) residues in GH17 proteins that characterize the catalytic domain of glycosyl hydrolases [***Error! Bookmark not defined.***]. These glutamate residue were identified within the tomato sequence by comparison to the catalytic residue positions in Arabidopsis identified by Gaudioso-Pedraza and Benitez-Alfonso (2014) [22]. Some of the genes lack one or both of the catalytic residues required to hydrolyse callose. Of the 50 tomato sequences, 36 showed the presence of both residues, 13 showed the presence of only one of the two, and 1 that showed no presence of either residue.
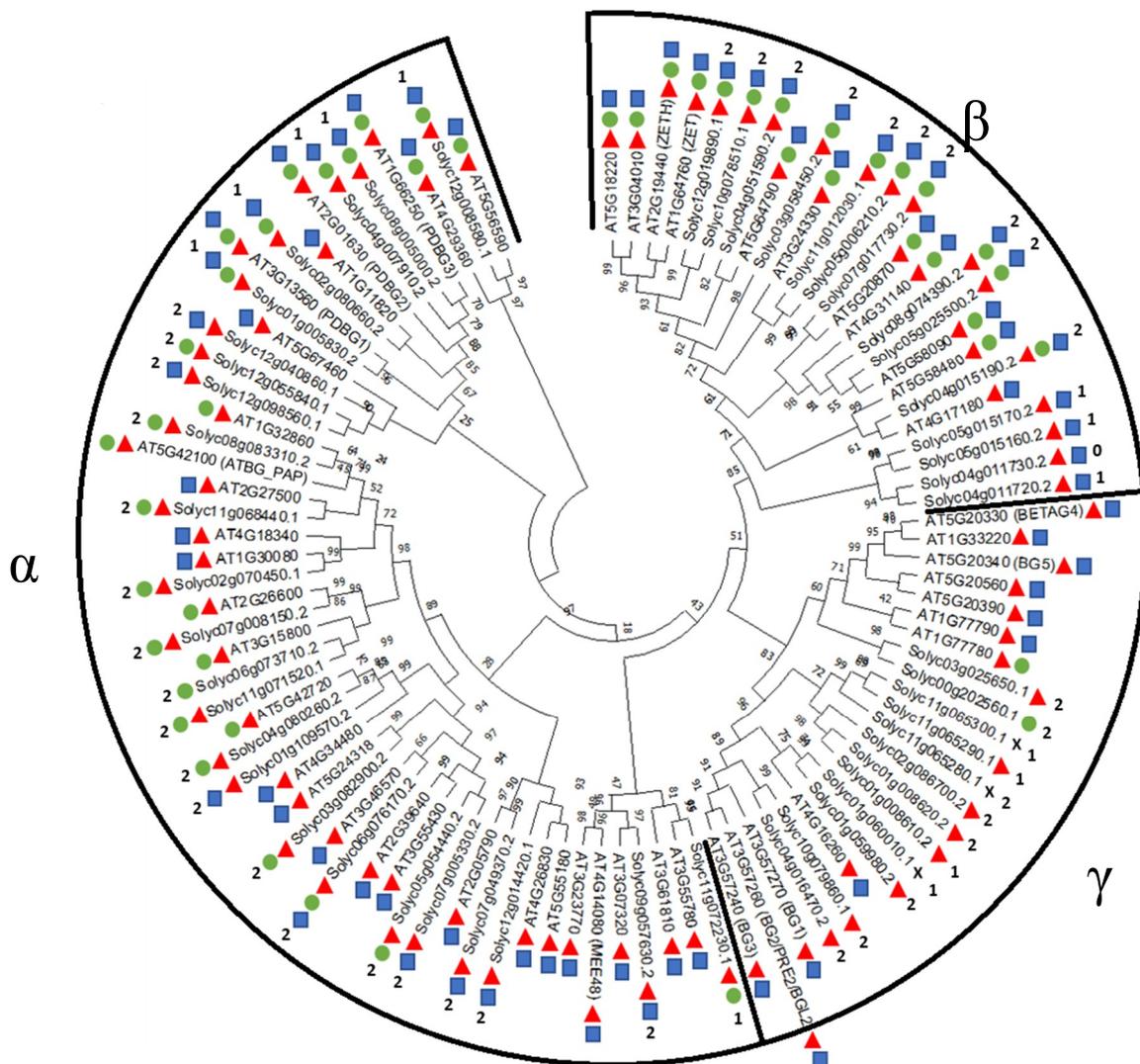
Figure 1. Phylogenetic relationships of β-1,3-glucanase genes in tomato and Arabidopsis. Amino acid sequences of predicted β-1,3-glucanase genes were aligned and a phylogenetic tree created using MEGA-X40 (using the WAG model and a bootstrap value of 1000). The tree is separated into clades (α, β and γ) as defined in Doxey et al. (2007) [23] for Arabidopsis. Predicted domains for each gene are indicated by the coloured spots: red/triangle = signal peptide, green/circle = GPI-anchor, blue/square = X8 domain, 'x' = no domains predicted. Presence of catalytic residue is indicated with number of catalic residues (0, 1 or 2) near to the coloured spots.

3.2. Identified Tomato BG Genes Show Varying Levels of Expression across Different Tissues and Organs

Tissue expression data for the 50 identified tomato BGs genes was extracted from the Genevestigator [27] database. The results were verified by comparing expression data in the ePlant BAR [30] database. The expression of these enzymes is shown to vary across tissues and organs (Data not showed). Based on the expression data from Genevestigator [27], one gene, Solyc04g016470, is found to be expressed specifically in tomato fruit. The seven genes: Solyc05g015160, Solyc05g015170, Solyc04g011720, Solyc04g011730, Solyc03g058450, Solyc04g051590 and Solyc05g054440, show similar tissue expression patterns with expression in flower and reproductive tissues and low expression elsewhere in the plant. In contrast, expression in the flower is uncommon for the other genes.

3.3. BGs Expression Change When Comparing Post-Anthesis and Ripened Fruits

Microarray data of fruit pericarp genes was downloaded from the SGN Tomato Expression Atlas (TEA) [25]. Expression values and trends were analysed and, of the 50 tomato genes, 16 were

identified as displaying significant expression in developing fruit from anthesis to red ripe stage (Figure 2). The data showed different patterns of expression, some genes were found to increase in expression during fruit ripening and some were found to decrease in expression during fruit ripening. The expression of these genes correlated well with their evolutionary relations. Expression of enzymes grouped in cluster α decreased during the ripening, while enzymes in clusters β and γ increased. Interpretation of this result in light of the activities identified in Arabidopsis suggest that BGs acting at plasmodesmata are required to improve cellular communication in early developing fruits whereas defense enzymes are required later in mature fruits suceptible to pathogen attack. The expression trends were supported by comparison to TomExpress [**Error! Bookmark not defined.**] and Genevestigator [27] expression data.
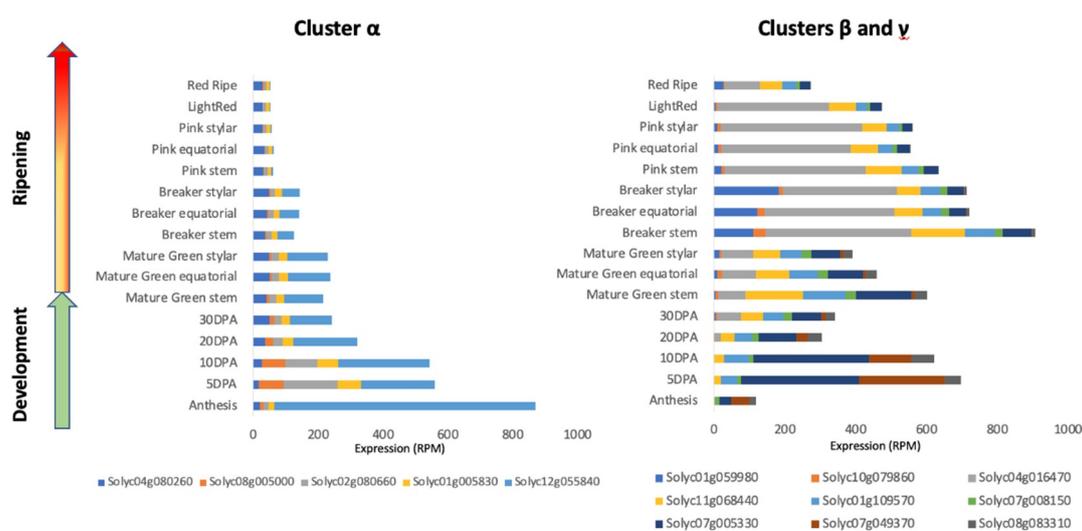
Figure 2. Expression profile of diverse tomato genes in pericarp ripening. Expression levels (in RPM) of different β-1,3-glucanases during ripening, from fruit anthesis to red ripe (DPA = days past anthesis), based on data extracted from the SGN-TEA database [25]. Left panel shows genes with decreased expression during pericarp ripening belonging to the cluster α. Rightt panel shows the genes with increased expression during pericarp ripening.

5. Conclusions

The aim of this publication was to explore the family of β-1,3-glucanases genes in tomato fruit using a combination of phylogeny and transcriptomic analysis, in order to propose new targets to modify cell walls (and tomato softening) for fruit breeding programmers. 50 callose degrading enzymes (β-1,3-glucanases) were identified in-silico using bioinformatic tools. Phylogenetic analysis revealed tomato genes being distributed in three cluster (α, β and γ), expression of enzymes in cluster α (which comprises all previously localized plasmodesmata proteins identified in Arabidopsis) decrease during the ripening, while enzymes in clusters β and γ (including Arabidopsis pathogenesis-induced proteins) increase. This result allows the identification of specific β-1,3-glucanases in tomato fruit that can influence fruit ripening (Figure 2) and that could be used as a target to delay softening thus to improve tomato shelf life.

Author Contributions: C.P. and Y.B.-A. conceived and designed the experiments and the project; L.P. performed the experiments; L.P., C.P. and Y.B.-A. analyzed the data; L.P., C.P. and Y.B.-A. wrote the paper.

Conflicts of Interest: "The authors declare no conflict of interest."

References

1. Food and Agricultural Organisation of the United Nations. FAOSTAT Statistical Database. 1997. Available online: http://www.fao.org/faostat/en/ (accessed on 7 September 2020).
2. Rossi, R. The EU fruit and vegetable sector: Main features, challenges and prospects. European Parliament 2019.
3. Burden, J.; Wills, R.; Toet, A. *Prevention of Post-Harvest Food Losses Fruits, Vegetables and Root Crops a Training Manual*; FAO: Rome, Italy, 1989; Volume 17.2. Available online: http://www.fao.org/3/t0073e/T0073E00.htm.
4. Saladié, M.; Matas, A.J.; Isaacson, T.; Jenks, M.A.; Goodwin, S.M.; Niklas, K.J.; Xiaolin, R.; Labavitch, J.M.; Shackel, K.A.; Fernie, A.R.; et al. A Reevaluation of the Key Factors That Influence Tomato Fruit Softening and Integrity. *Plant Physiol.* **2007**, *144*, 1012–1028, doi:10.1104/pp.107.097477.
5. Goulao, L.F.; Oliveira, C. Cell wall modifications during fruit ripening: When a fruit is not the fruit. *Trends Food Sci. Technol.* **2008**, *19*, 4–25, doi:10.1016/j.tifs.2007.07.002.
6. Han, X.; Huang, L.-J.; Feng, D.; Jiang, W.; Miu, W.; Li, N. Plasmodesmata-Related Structural and Functional Proteins: The Long Sought-After Secrets of a Cytoplasmic Channel in Plant Cell Walls. *Int. J. Mol. Sci.* **2019**, *20*, 2946.
7. . Ma, S.; Gross, K.C.; Wasserman, B.P. Developmental Regulation of the (1,3)-β-Glucan (Callose) Synthase from Tomato. *Plant Physiol.* **1991**, *96*, 664–667, doi:10.1104/pp.96.2.664.
8. Levy, A.; Erlanger, M.; Rosenthal, M.; Epel, B. A Plasmodesmata-Associated β-1,3-Glucanase in Arabidopsis: A Plasmodesmal β-1,3-Glucanase. *Plant J.* **2007**, *49*, 669–682.
9. Schultz, J. SMART: A web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* **2000**, *28*, 231–234.
10. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268, doi:10.1093/nar/gkz991.
11. Mitchell, A.L.; Attwood, T.K.; Babbitt, P.C.; Blum, M.; Bork, P.; Bridge, A.; Brown, S.D.; Chang, H.-Y.; El-Gebali, S.; Fraser, M.I.; et al. InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **2019**, *47*, D351–D360, doi:10.1093/nar/gky1100.
12. Okonechnikov, K.; Golosova, O.; Fursov, M. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167, doi:10.1093/bioinformatics/bts091.
13. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797, doi:10.1093/nar/gkh340.
14. Schwacke, R.; Schneider, A.; van der Graaff, E.; Fischer, K.; Catoni, E.; Desimone, M.; Frommer, W.; Flügge, U.I.; Kunze, R. ARAMEMNON, a Novel Database for Arabidopsis Integral Membrane Proteins. *Plant Physiol.* **2003**, *131*, 16–26, doi:10.1104/pp.011577.
15. Armenteros, J.J.A.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; Von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37*, 420–423, doi:10.1038/s41587-019-0036-z.
16. Viklund, H.; Bernsel, A.; Skwark, M.; Elofsson, A. SPOCTOPUS: A Combined Predictor of Signal Peptides and Membrane Protein Topology. *Bioinformatics* **2008**, *24*, 2928–2929, doi:10.1093/bioinformatics/btn550.
17. Hiller, K.; Grote, A.; Scheer, M.; Munch, R.; Jahn, D. PrediSi: Prediction of Signal Peptides and Their Cleavage Positions. *Nucleic Acids Res.* **2004**, *32*, 375–379, doi:10.1093/nar/gkh378.
18. Fankhauser, N.; Maser, P. Identification of GPI-anchor Attachment Signals by a Kohonen Self-Organizing Map. *Bioinformatics* **2005**, *21*, 1846–1852, doi:10.1093/bioinformatics/bti299.
19. Pierleoni, A.; Martelli, P.; Casadio, R. PredGPI: A GPI-Anchor Predictor. *BMC Bioinform.* **2008**, *9*, 392, doi:10.1186/1471-2105-9-392.
20. Eisenhaber, B.; Bork, P.; Eisenhaber, F. Prediction of Potential GPI-modification Sites in Proprotein Sequences. *J. Mol. Biol.* **1999**, *292*, 741–758, doi:10.1006/jmbi.1999.3069.
21. Wojtkowiak, A.; Witek, K.; Hennig, J.; Jaskolski, M. Structures of an active-site mutant of a plant 1,3-β-glucanase in complex with oligosaccharide products of hydrolysis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *69*, 52–62, doi:10.1107/S0907444912042175.
22. Gaudioso-Pedraza, R.; Benitez-Alfonso, Y. A phylogenetic approach to study the origin and evolution of plasmodesmata-localized glycosyl hydrolases family 17. *Front. Plant Sci.* **2014**, *5*, 5, doi:10.3389/fpls.2014.00212.

23. Doxey, A.C.; Yaish, M.W.F.; Moffatt, B.A.; Griffith, M.; McConkey, B.J. Functional Divergence in the Arabidopsis -1,3-Glucanase Gene Family Inferred by Phylogenetic Reconstruction of Expression States. *Mol. Biol. Evol.* **2007**, *24*, 1045–1055, doi:10.1093/molbev/msm024.

24. Athar, A.; Füllgrabe, A.; George, N.; Iqbal, H.; Huerta, L.; Ali, A.; Snow, C.; Fonseca, N.; Petryszak, R.; Papatheodorou, I.; et al. ArrayExpress Update—From Bulk to Single-Cell Expression Data. *Nucleic Acids Res.* **2019**, *47*, 711–715, doi:10.1093/nar/gky964.

25. Fernandez-Pozo, N.; Zheng, Y.; Snyder, S.I.; Nicolas, P.; Shinozaki, Y.; Fei, Z.; Catala, C.; Giovannoni, J.J.; Rose, J.K.C.; Mueller, L.A. The Tomato Expression Atlas. *Bioinformatics* **2017**, *33*, 2397–2398, doi:10.1093/bioinformatics/btx190.

26. Zouine, M.; Maza, E.; Djari, A.; Lauvernier, M.; Frasse, P.; Smouni, A.; Pirrello, J.; Bouzayen, M. TomExpress, a Unified Tomato RNA-Seq Platform for Visualization of Expression Data, Clustering and Correlation Networks. *Plant J.* **2017**, *92*, 727–735, doi:10.1111/tpj.13711.

27. Hruz, T.; Laule, O.; Szabo, G.; Wessendorp, F.; Bleuler, S.; Oertle, L.; Widmayer, P.; Gruissem, W.; Zimmermann, P. Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Adv. Bioinform.* **2008**, *2008*, 1–5, doi:10.1155/2008/420747.

28. Nakai, K.; Kanehisa, M. Expert system for predicting protein localization sites in gram-negative bacteria. *Proteins* **1991**, *11*, 95–110, doi:10.1002/prot.340110203.

29. Horton, P.; Park, K.-J.; Obayashi, T.; Fujita, N.; Harada, H.; Adams-Collier, C.; Nakai, K. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **2007**, *35*, W585–W587, doi:10.1093/nar/gkm259.

30. Waese, J.; Fan, J.; Pasha, A.; Yu, H.; Fucile, G.; Shi, R.; Cumming, M.; Kelley, L.A.; Sternberg, M.J.E.; Krishnakumar, V.; et al. ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* **2017**, *29*, 1806–1821, doi:10.1105/tpc.17.00073.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.