BIOLOGICAL COMMUNICATIONS, vol. 66, issue 3, July–September, 2021 | https://doi.org/10.21638/spbu03.2021.303

SUPPLEMENTS

Protein glycation and drought response of pea (*Pisum sativum* L.) root nodule proteome: a proteomics approach

Julia Shumilina, Daria Gorbach, Veronika Popova, Alexander Tsarev, Alena Kusnetsova,
Maria Grashina, Mandy Dorn, Elena Lukasheva, Natalia Osmolovskaya,
Ekaterina Romanovskaya, Vladimir Zhukov, Christian Ihling, Tatiana Grishina,
Tatiana Bilova, and Andrej Frolov

Fig. S1-1. The electropherograms of the total protein preparations, isolated from *P. sativum* nodules, before (A) and after tryptic digestion (B). M denotes the molecular weight marker (Marker/Page Ruler #26619, Thermo fisher scientific, Bremen, Germany); 1–3: protein samples obtained from the root nodules of untreated SGE pea plants; 4–6: protein samples obtained from the root nodules of PEG-treated SGE pea plants (i.e. those subjected to osmotic stress). The gels were stained with colloidal Coomassie G-250.

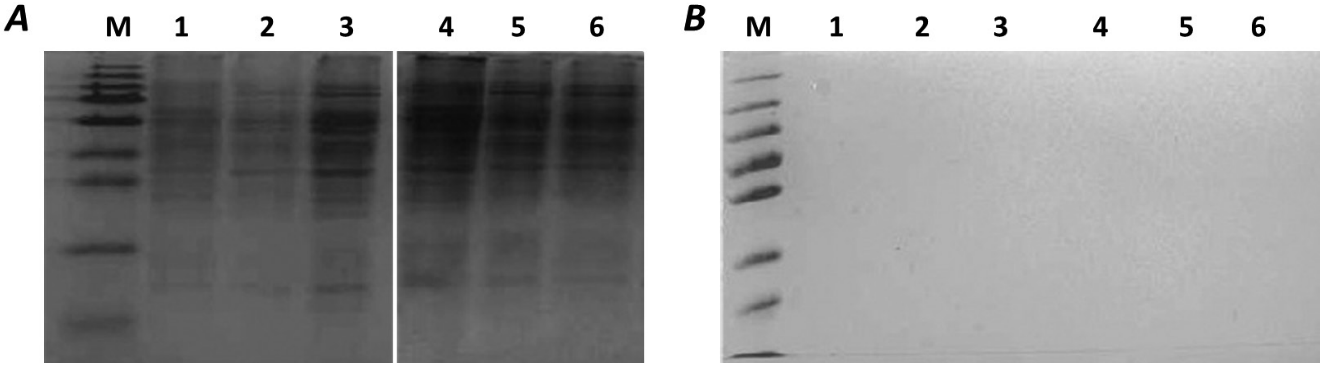


Fig. S1-2. The numbers of individual *Pisum sativum* (A) and *Rhizobium leguminosarum* (B) peptides, identified in the control (C) and drought-treated (stressed, S) groups. The data were acquired by nanoHPLC-LIT-Orbitrap-MS operated in the positive DDA mode. Identification of peptides and annotation of proteins relied on Sequest search against *P. sativum* (Kreplak et al. 2019) and *R. leguminosarum* (Afonin et al. 2017) sequence databases in Proteome Discoverer 2.1 software with implementation of characteristic modification-specific increments.

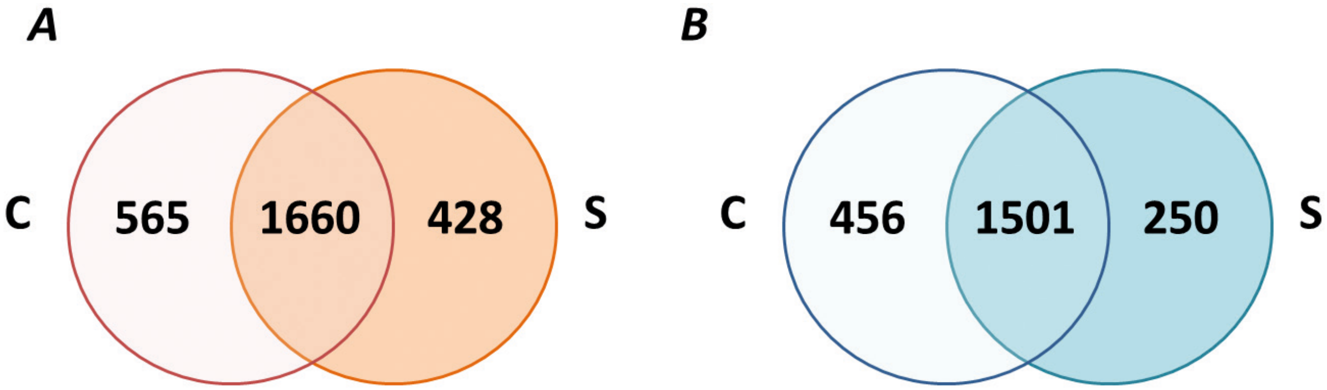


Fig. S1–3. The numbers of individual *Pisum sativum* (A) and *Rhizobium leguminosarum* (B) proteins, identified in the control (C) and drought-treated (stressed, S) groups. The data were acquired by nanoHPLC-LIT-Orbitrap-MS operated in the positive DDA mode. Identification of peptides and annotation of proteins relied on Sequest search against *P. sativum* (Kreplak et al. 2019) and *R. leguminosarum* (Afonin et al. 2017) sequence databases in Proteome Discoverer 2.1 software with implementation of characteristic modification-specific increments.

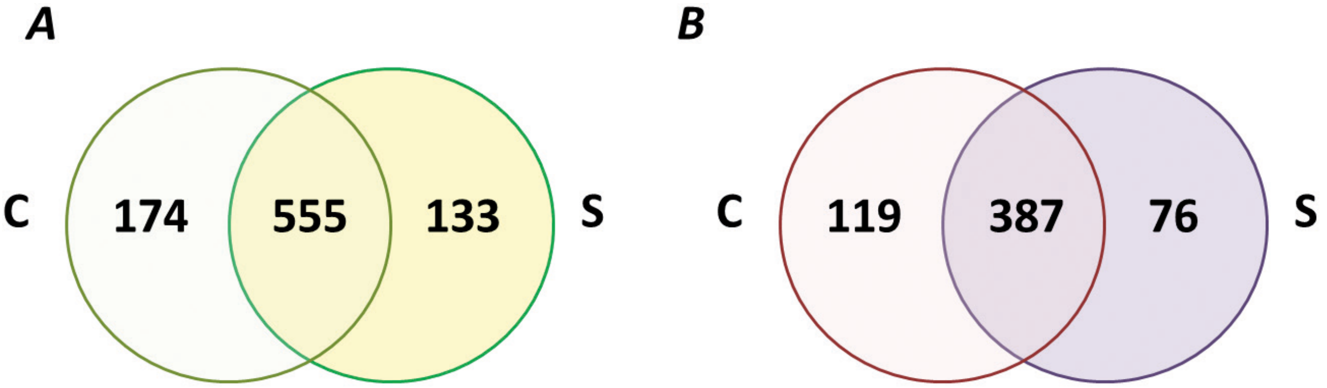


Fig. S1–4. Tandem mass spectra of the *m/z*, corresponding to the peptides of differentially expressed proteins, performed against a redundant legume sequence database. Spectra are given in order corresponding to the order of proteins in Table 2. Number of peptide spectrum matches (PSM) was calculated by Sequest engine in Proteome Discoverer 2.2 software.

Fig. S1–4.1. psat0s3599g0120.1(1 unique peptide(s), master protein) VQNIAADAVPDLLNYFK, (1890.99599 Da, m/z 945.99842), 3 PSMs

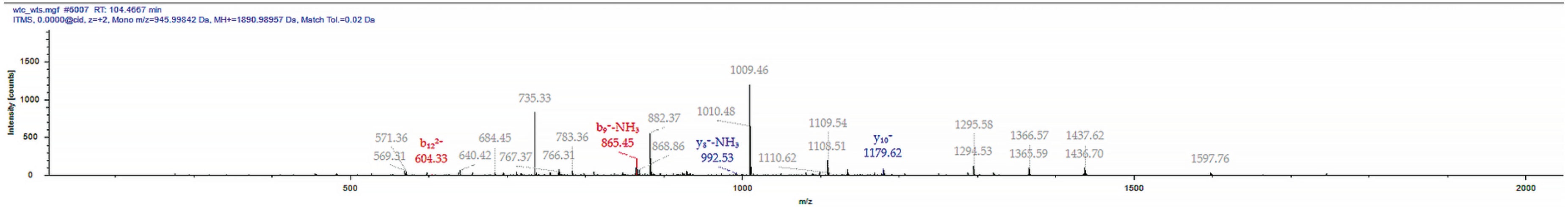


Fig. S1–4.2. psat5g100000.1(1 unique peptide(s), master protein) FSASFDMPGFLVSELK, (1774.87203 Da, *m/z*887.93672), 1 PSMs

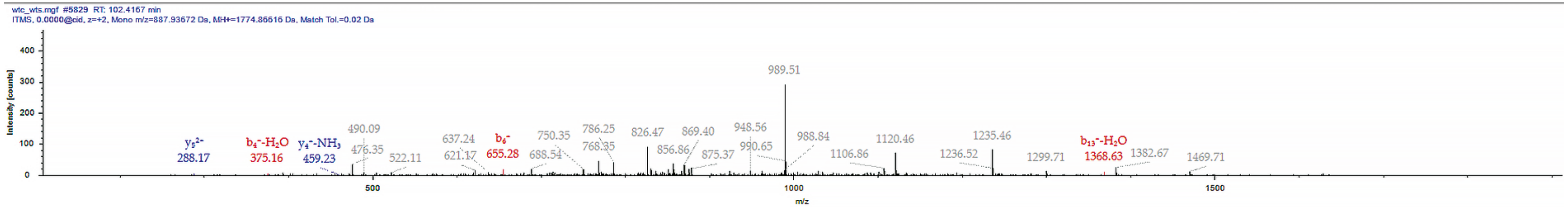


Fig. S1–4.3. psat0ss1196g0320.1(2 unique peptide(s), master protein) TNLVMVFGEITTK, (1452.77668 Da, *m/z*726.88960), 1 PSMs

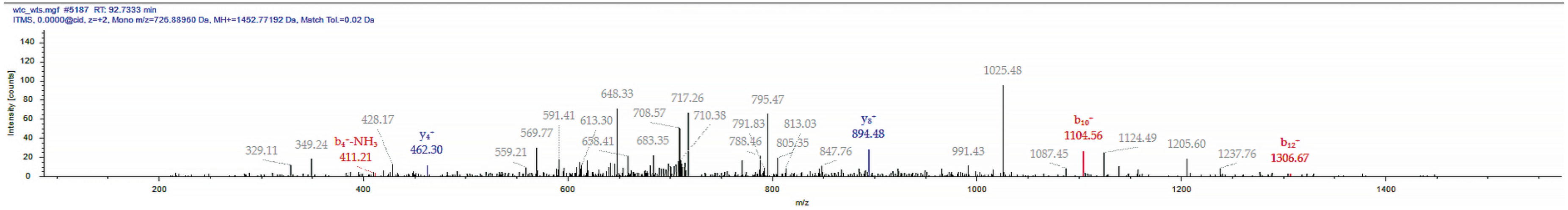


Fig. S1–4.4. psat0ss1196g0320.1 (2 unique peptide(s), master protein) VLVNIEQQSPDIAQGVHGHFSK, (2417.25718 Da, *m/z*605.06806), 1 PSMs

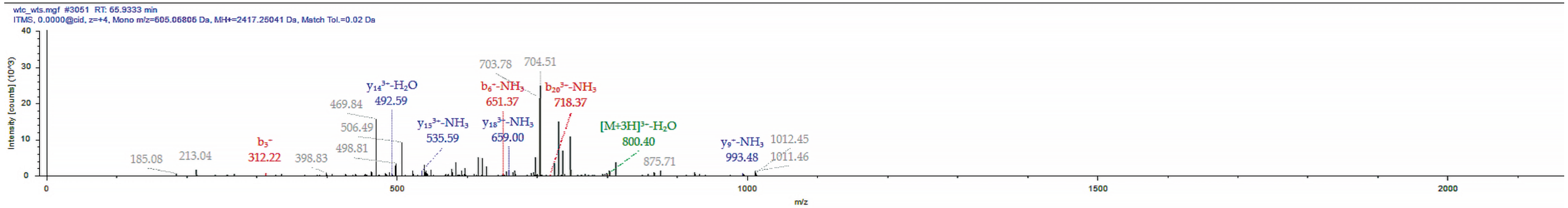


Fig. S1–4.5. psat6g152120.1 (1 unique peptide(s), master protein) FSASFDMPGFLVSELK, (1234.68304 Da, *m/z*617.84317), 3 PSMs

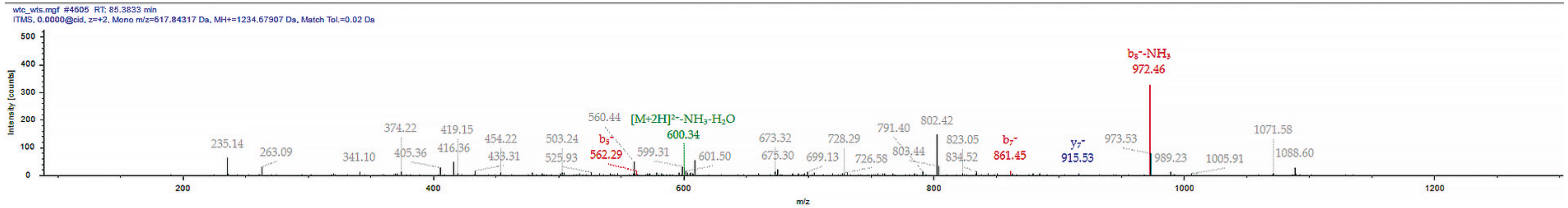


Fig. S1–4.6. psat0s241g0080.1(2 unique peptide(s), master protein) TKGVVVLGDATLGAIHIQK, (1920.12797 Da, *m/z*480.78606), 4 PSMs

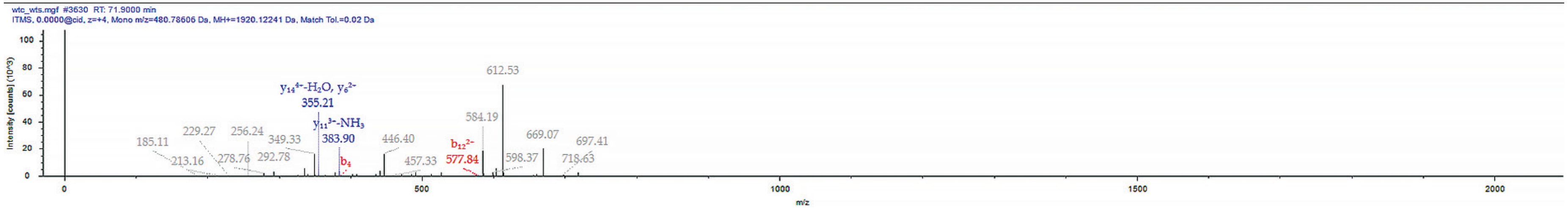


Fig. S1–4.7. psat0s241g0080.1(2 unique peptide(s), master protein) GVVVLGDATLGAIHIQK, (1690.98503 Da, *m/z*564.33144), 2 PSMs

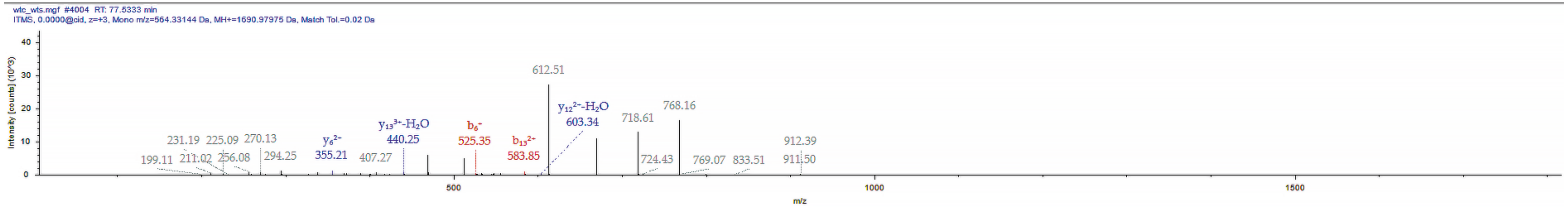


Fig. S1–4.8. psat5g086440.1 (1 unique peptide(s), master protein) FVEESSIPLVTIFNNEPSNHPFVVK,(2843.46142 Da, *m/z*711.61872), 1 PSMs

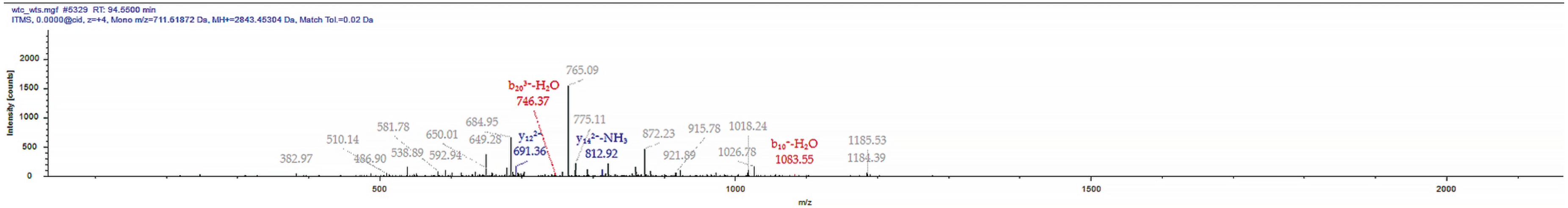


Fig. S1–4.9. psat6g198160.1(1 unique peptide(s), master protein) SGPYGQIFRPDNFVFGQSGAGNNWAK, (2814.33829 Da, *m/z*938.78134), 3 PSMs

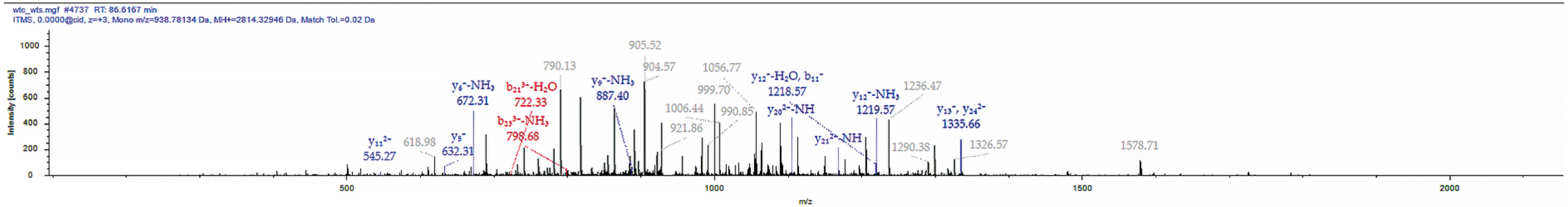


Fig. S1–4.10. psat6g079280.1(1 unique peptide(s), master protein) FANISVENTEGNRR, (1606.79321 Da, *m/z*803.89837), 3 PSMs

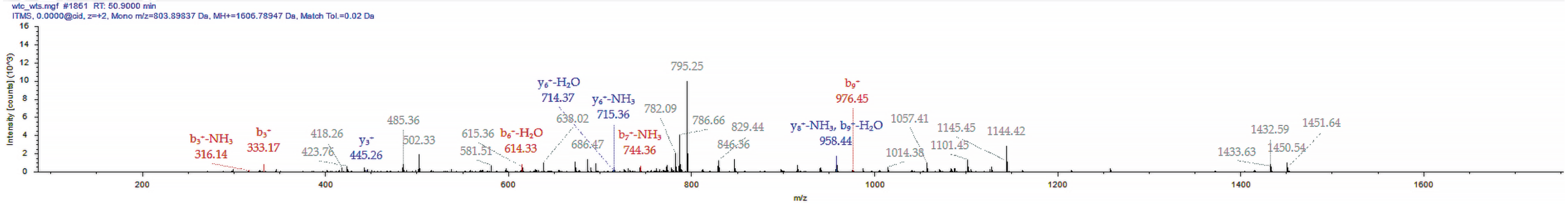


Fig. S1–4.11. psat1g143040.3 (1 unique peptide(s), master protein) IALTDNSIIEQVLGK, (1872.90479 Da, *m/z*807.45613), 4 PSMs

Fig. S1–4.12. psat6g104960.1 (1 unique peptide(s), master protein) VLVTGAAGQIGYALVPMIAR, (2016.13104 Da, *m/z*672.71354), 1 PSMs

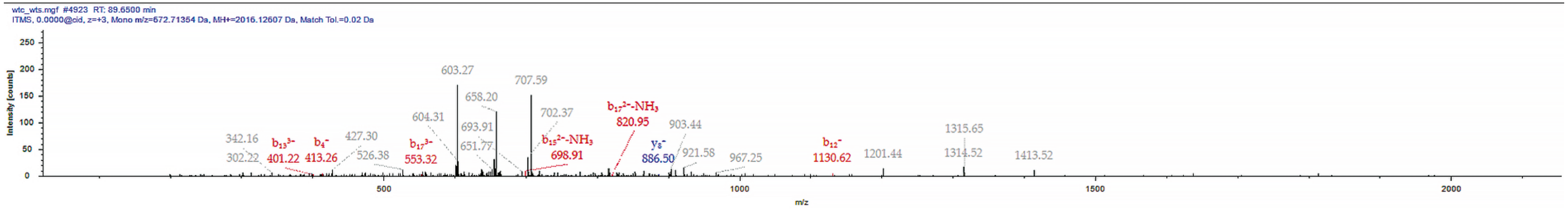


Fig. S1–4.13. psat6g147200.1 (1 unique peptide(s), master protein) STGFGLIYDSVENAK,(1600.78533 Da, *m/z*800.89406), 1 PSMs

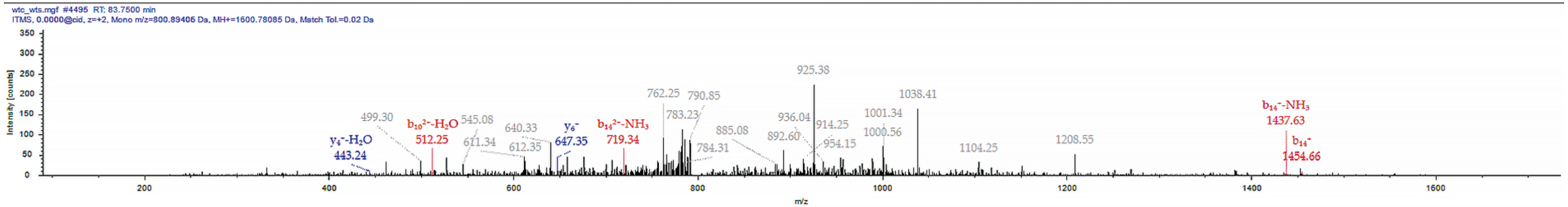


Fig. S1–4.14. psat0s3139g0040.1 (1 unique peptide(s), master protein) YHTKGDAALSDAVRDETKAK, (2257.18229 Da, *m/z*565.04935), 2 PSMs

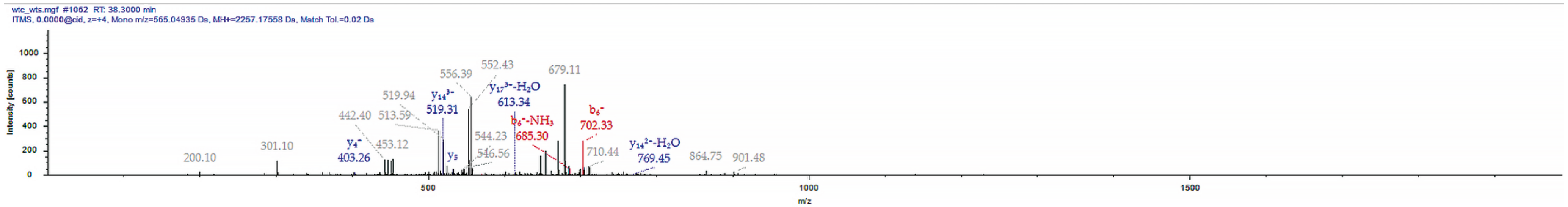


Fig. S1–4.15. psat2g039640.1(3 unique peptide(s), master protein) GMFSFLKDSAGVVDSPK, (1872.90479 Da, *m/z*624.97172), 4 PSMs

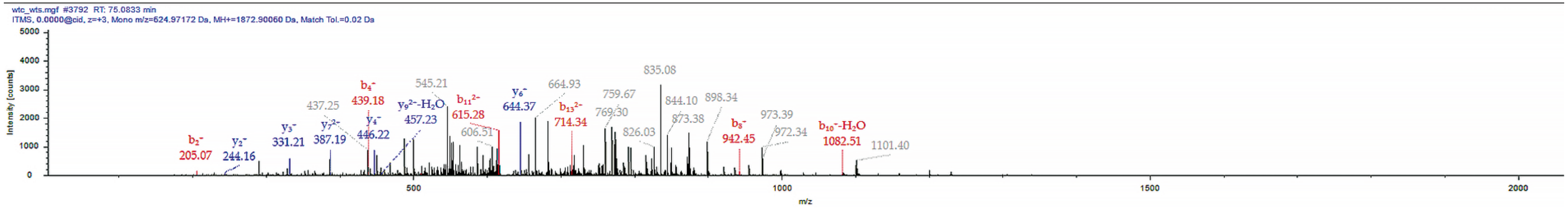


Fig. S1–4.16. psat2g039640.1(3 unique peptide(s), master protein) ASGEVVVGDATLGAIHIQK, (1865.01270 Da, *m/z*622.34078), 11 PSMs

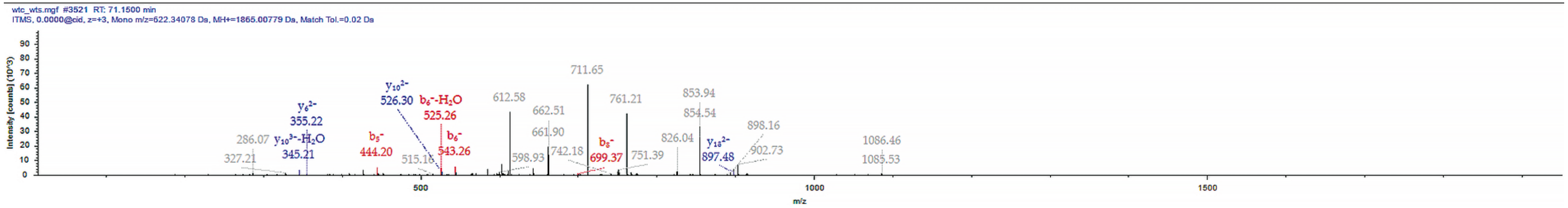


Fig. S1–4.17. psat0s241g0080.1(2 unique peptide(s), master protein) TKGVVVLGDATLGAIHIQK, (1920.12797 Da, *m/z*480.78606), 4 PSMs

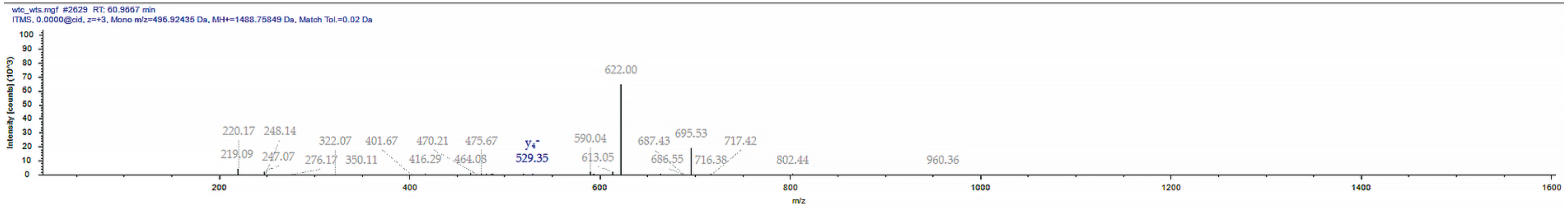


Fig. S1–5. Tandem mass spectra of the *m/z*, corresponding to the peptides identified in a search, performed against a redundant legume sequence database. Spectra are given in order corresponding to the order of proteins in Table 2. Number of peptide spectrum matches (PSM) was calculated by Sequest engine in Proteome Discoverer 2.2 software.

Fig. S1–5.1. A0A1Q8HNR1 (1 unique peptide(s), master protein) DTALGFNFRPVTYYDGQGFMVR, (2554,21835Da, *m/z*852,07544), 4 PSMs

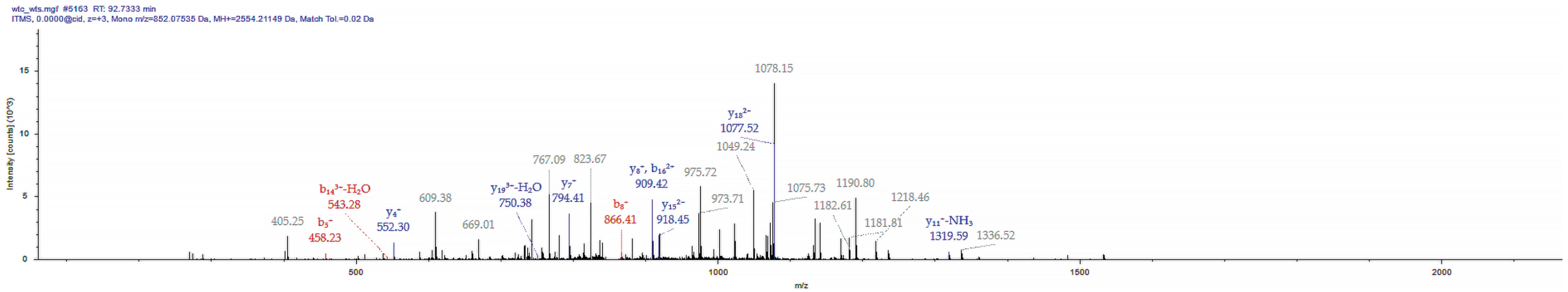


Fig. S1–5.2. A0A1Q8H6F7(3 unique peptide(s), master protein) DFDVPFAHTPAFVGSHVDGYDSMVK, (2738,25553 Da, *m/z*913,42084), 4 PSMs

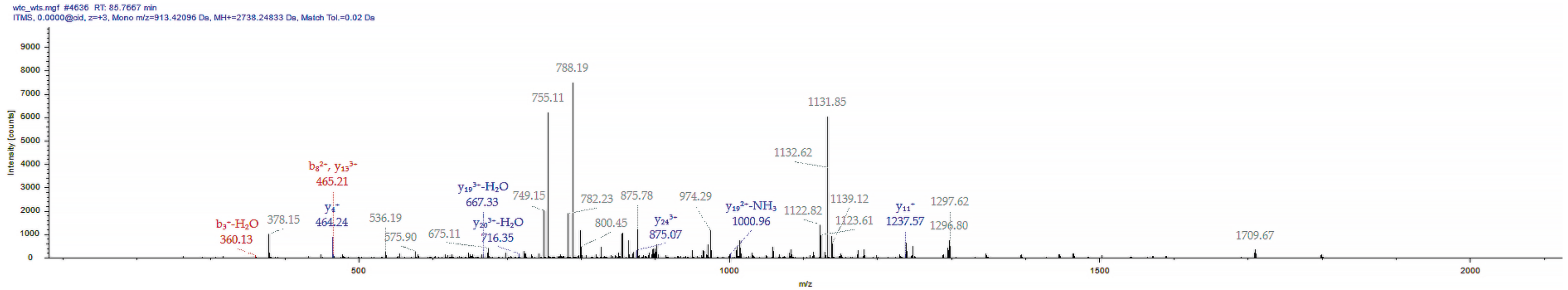


Fig. S1–5.3. A0A1Q8HIJ8 (1 unique peptide(s), master protein) LGELPDLAGFLLK, (1385,80388 Da, *m/z*693,40308), 3 PSMs

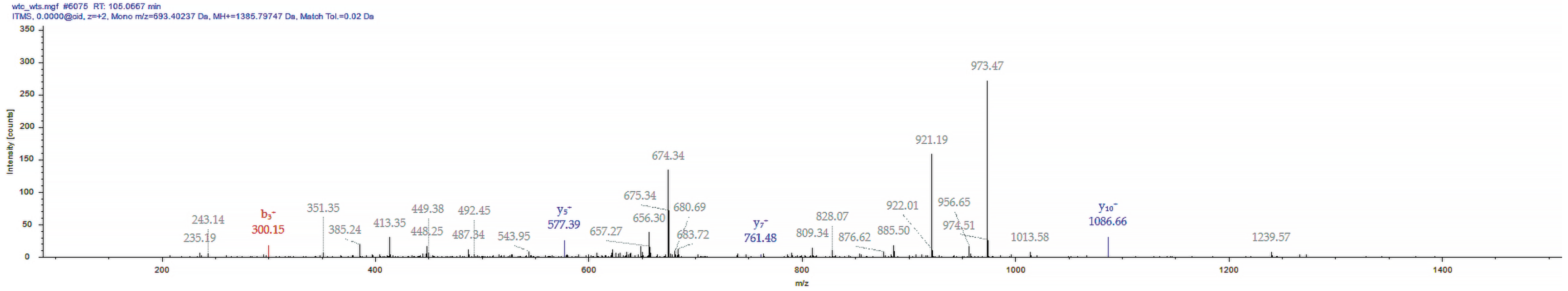


Fig. S1–5.4. A0A1Q8H9E9(1 unique peptide(s), master protein) TFDVLNPATGELLASLPDMGAAETR, (2589,28649 Da, *m/z*1295,14297), 4 PSMs

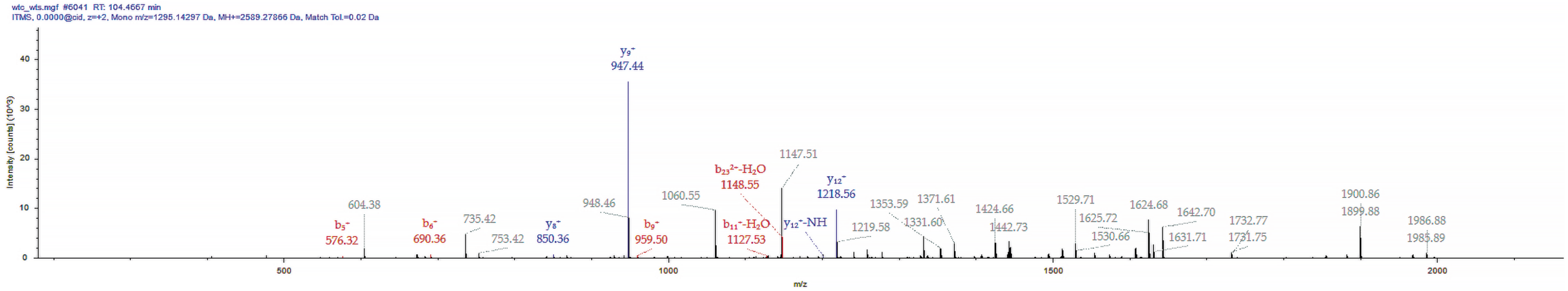


Fig. S1–5.5. A0A1Q8HHS1(1 unique peptide(s), master protein) VDVVYFGGYHPEGGLLAR, (1948,99157Da, *m/z*650,33386), 3 PSMs

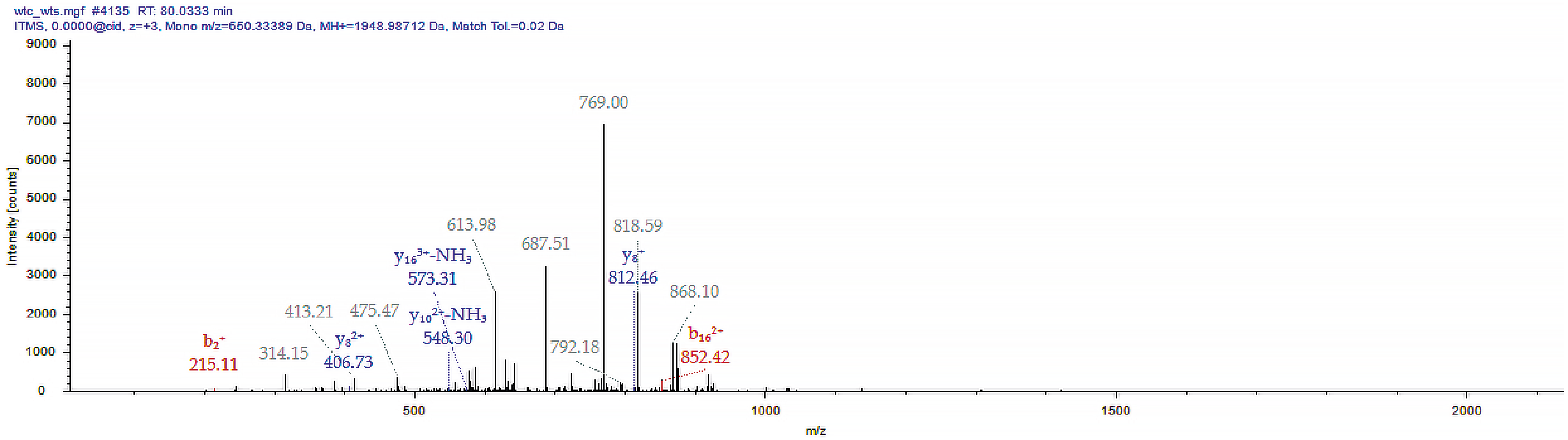


Fig. S1–5.6. A0A1Q8H9C5(1 unique peptide(s), master protein) VAAIIIEPVQGEGGFYSAPVAFMK, (2494,30504 Da, *m/z*1247,29754), 4 PSMs

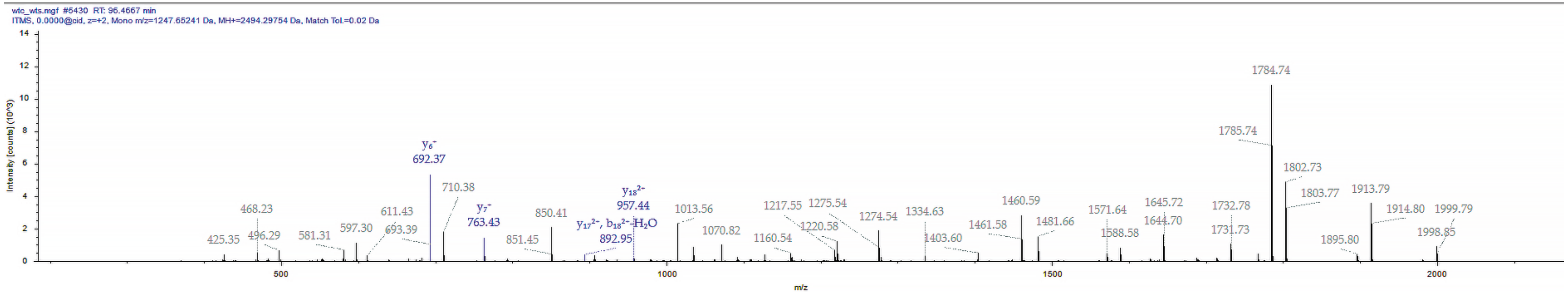


Fig. S1–5.7. A0A1Q8H6F8(1 unique peptide(s), master protein) DMDMAINNPVWGLYDAR, (1996,88916Da, *m/z*998,94672), 4 PSMs

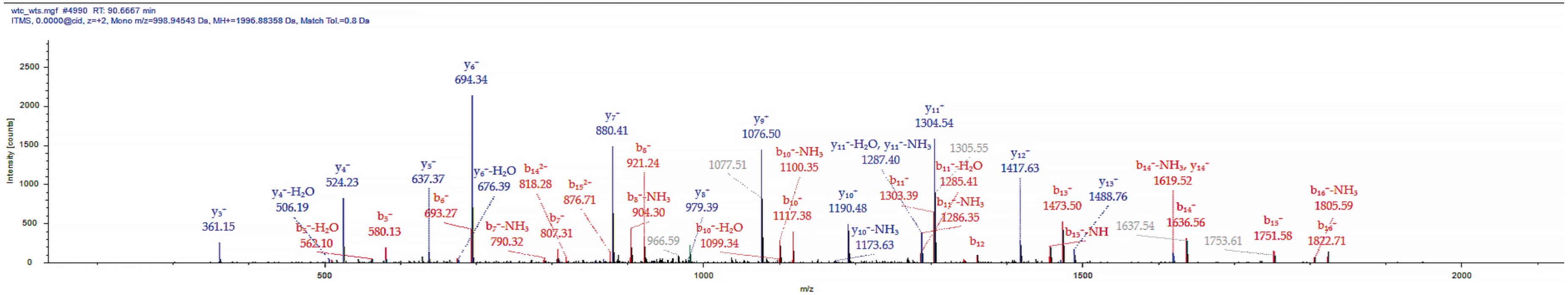


Fig. S1–5.8. A0A1Q8HEV3(3 unique peptide(s), master protein) ISLENIATGFAGSMDADYKKPGQLNNR, (2910,44143 Da, *m/z*970,81592), 4 PSMs

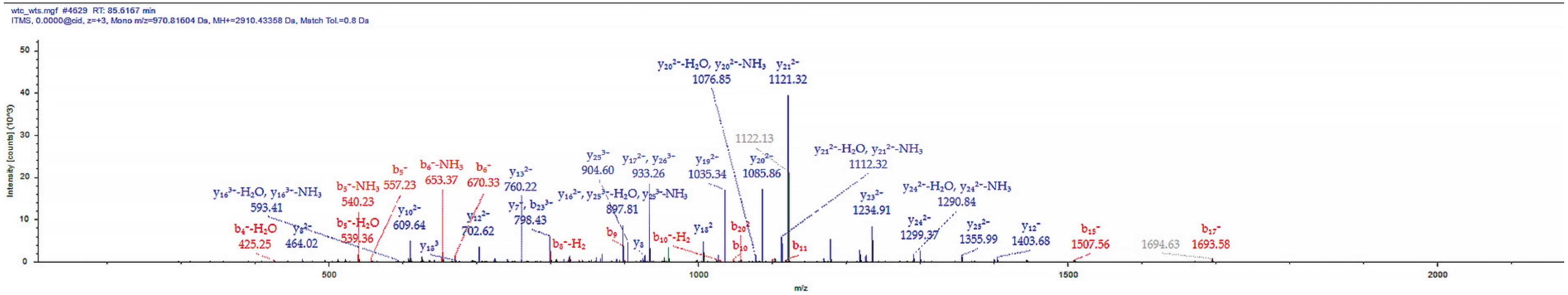


Fig. S1–5.9. A0A1Q8HBE8 (1 unique peptide(s), master protein) TREGNDLYYEMIESNVNKHGGGEGSK, (2884,31662Da, *m/z*721,83356), 4 PSMs

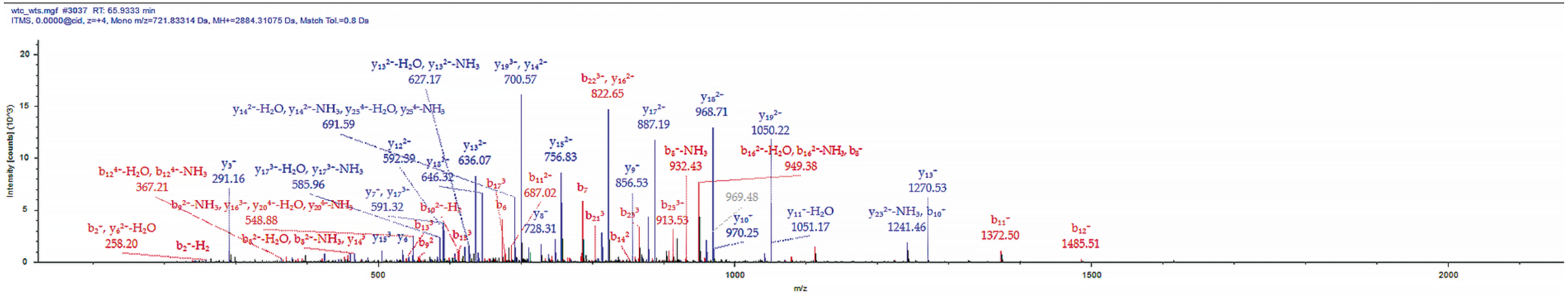


Fig. S1–5.10. A0A1Q8HEV3 (3 unique peptide(s), master protein) FVEFFGPGMDNMPLADR, (1942,88262 Da, *m/z*971,94238), 4 PSMs

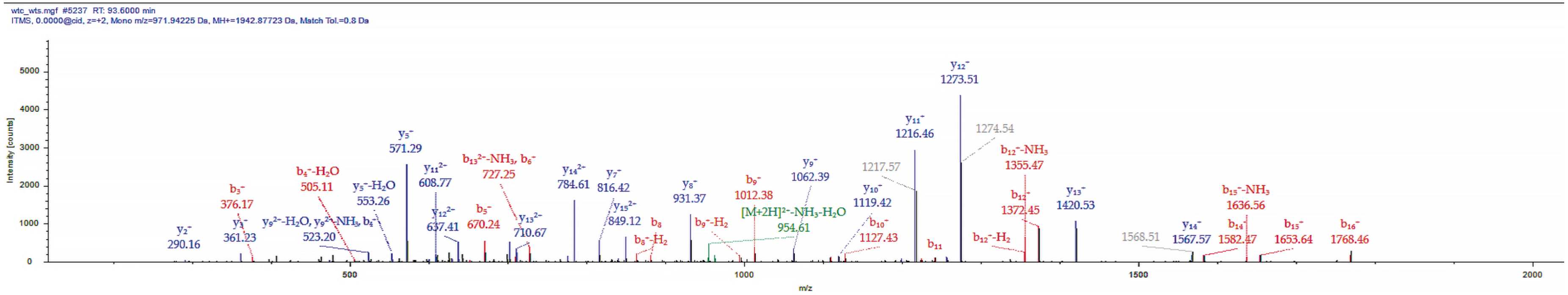


Fig. S1–5.11. A0A1Q8H6F7 (3 unique peptide(s), master protein) SEGSVPRDFDVPFAHTPAFVGSHVDGYDSMVK, (3450,60593 Da, *m/z*863,40436), 4 PSMs

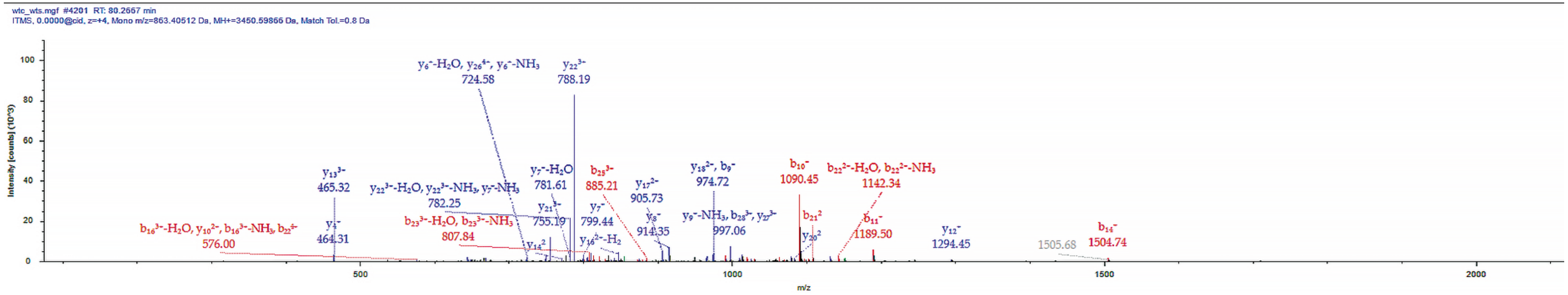


Fig. S1–5.12. A0A1Q8HEV3 (3 unique peptide(s), master protein) AASPAGAYLIDHDVAVADFNQYGTR, (2622,25830 Da, *m/z*874,75586), 4 PSMs

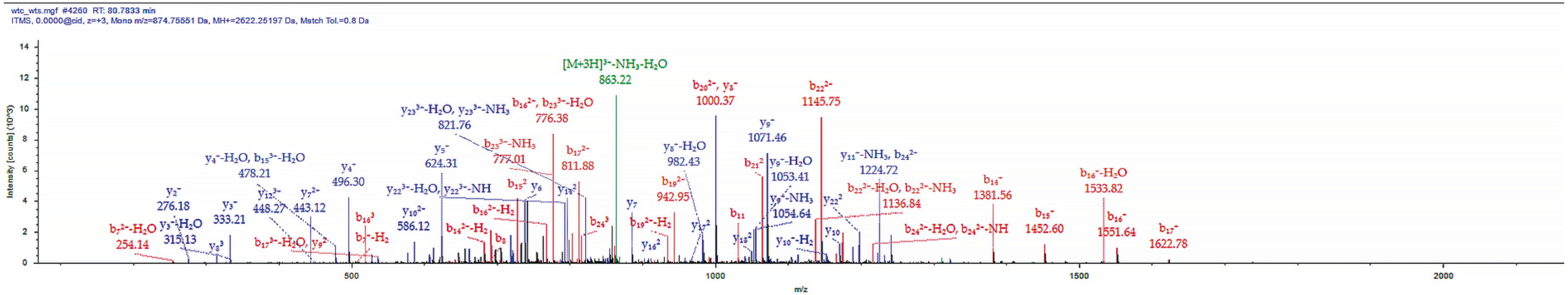


Fig. S1–5.13. A0A1Q8HBN9 (1 unique peptide(s), master protein) IIEIDLNKTEKEAFDK, (1906,01678 Da, *m/z*477,25819), 3 PSMs

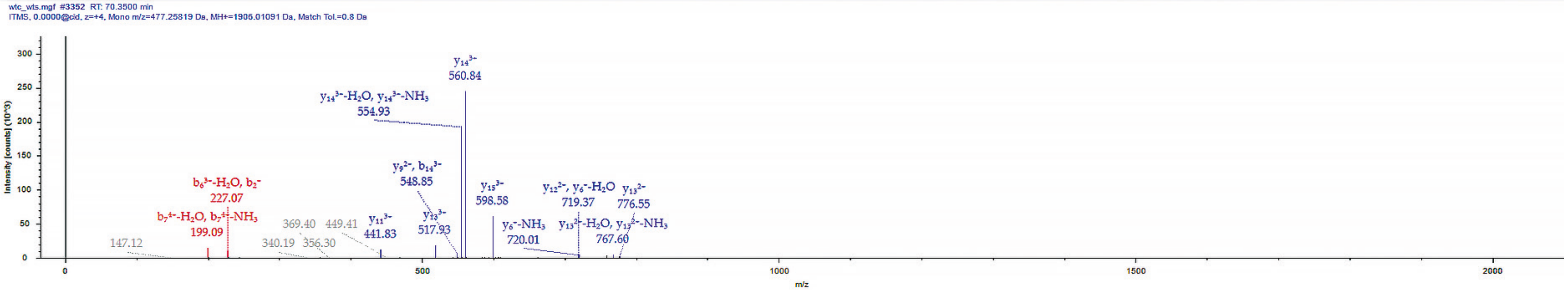


Fig. S1–5.14. A0A1Q8H6F7 (3 unique peptide(s), master protein) HFKEPSSVVSSSMTEDAAVFGGLK, (2510,22316Da, *m/z*837,41058), 4 PSMs

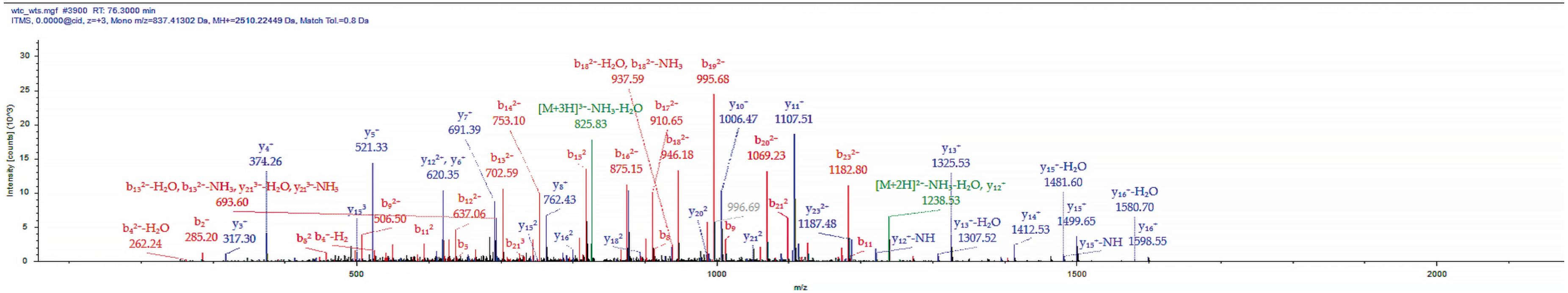


Fig. S1–5.15. A0A1Q8HJB2(1 unique peptide(s), master protein) AAAAGGHAVGVPFTPGRTDAAEAQTDAHSFAALKPR, (3516,77309 Da, *m/z*704,15924), 3 PSMs

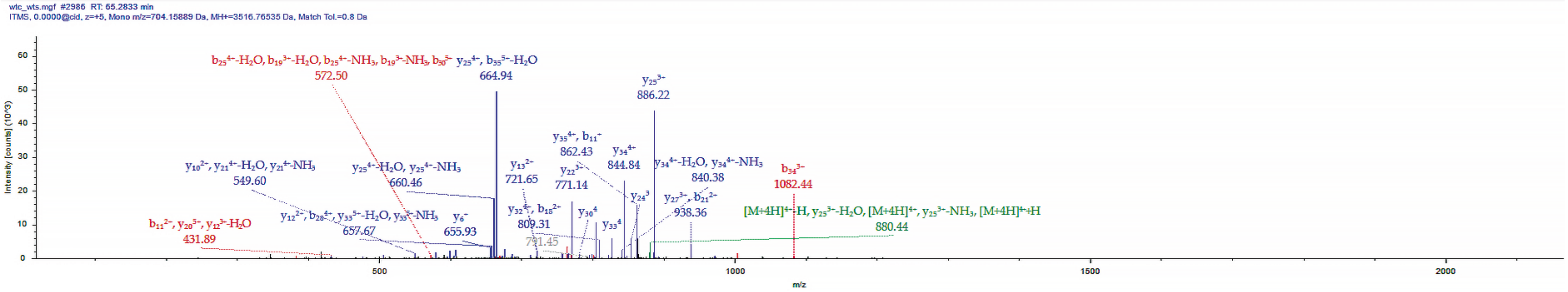


Fig. S1–5.16. A0A1Q8HLA8(1 unique peptide(s), master protein) AAVEEGILPGGGVALLRAVK, (1920,12241Da, *m/z*480,78606), 2 PSMs

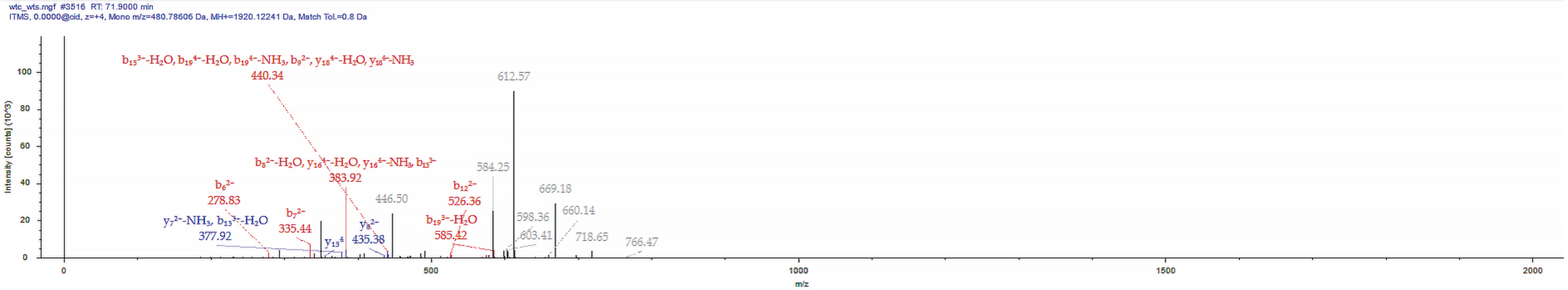


Fig. S1–5.17. A0A1Q8HKT2 (1 unique peptide(s), master protein) NREIMGATNPANADEGTIRK, (2174,06185 Da, *m/z*725,35730), 3 PSMs

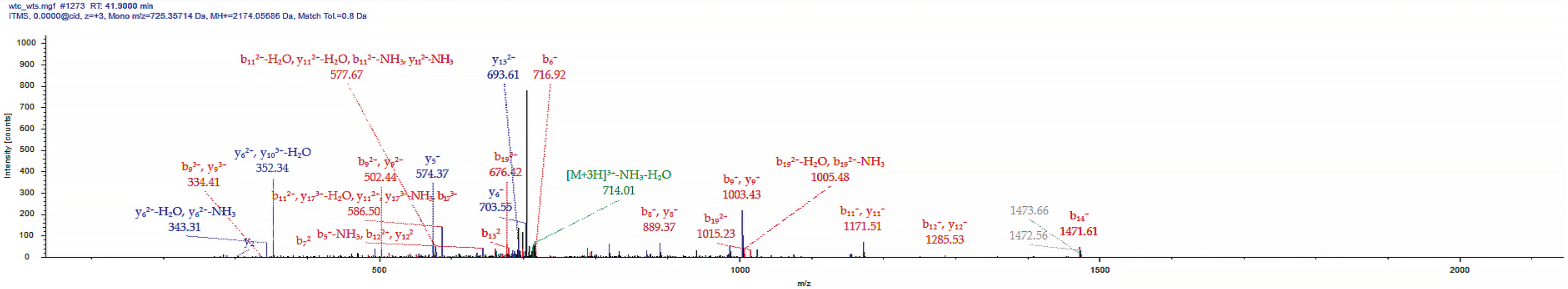
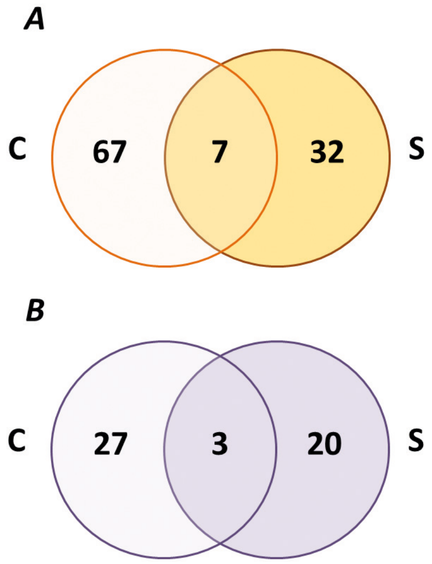


Fig. S1–6. Total numbers of oxidized (A) and glycated (B) peptides identified by nanoLC-MS-based proteomics analysis of pea (*P. sativum*) nodule protein tryptic digests obtained from control (C) and PEG-treated (stressed, S) plants. The data were acquired by nanoHPLC-LIT-Orbitrap-MS operated in the positive DDA mode. Identification of glycated peptides and annotation of glycated proteins relied on Sequest search against pea (*P. sativum*) sequence database (Kreplak et al 2019) in Proteome Discoverer 2.1 software with implementation of characteristic modification-specific increments.



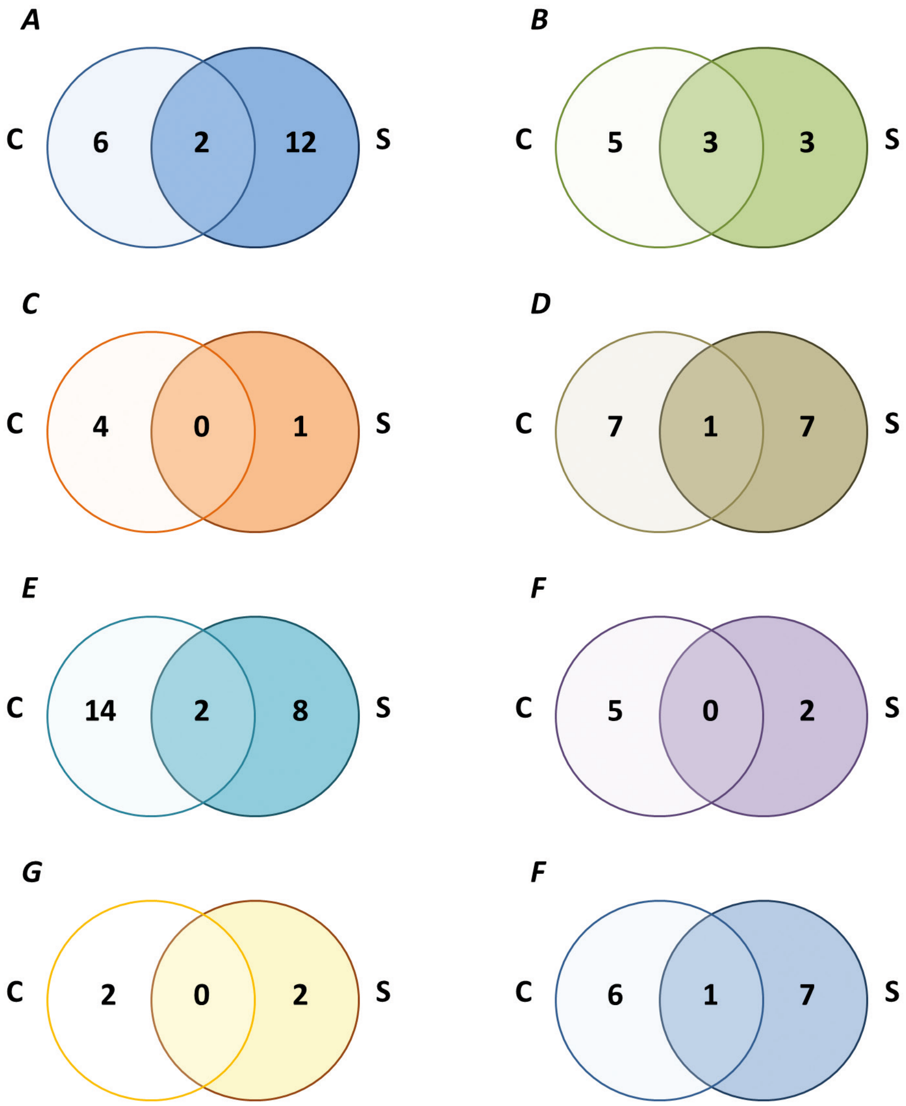


Fig. S1–7. Total numbers of peptides, identified as mono (A,D)-, doubly (B,E), triply (C,F) oxidized and kynurenin-containing (G,F) at tryptophan residues by nanoLC-MS-based proteomics analysis of *Pisim sativum* (A-D) and *Rhizobium leguminosarum* (E-H) nodule protein tryptic digests obtained from control (C) and PEG-treated (stressed, S) plants. The data were acquired by nanoHPLC-LIT-Orbitrap-MS operated in the positive DDA mode. Identification of glycated peptides and annotation of glycated proteins relied on Sequest search against pea (*P. sativum*, Kreplak et al 2019) and rhizobial (*R. leguminosarum*, Afonin et al. 2017) sequence database in Proteome Discoverer 2.1 software with implementation of characteristic modification-specific increments.

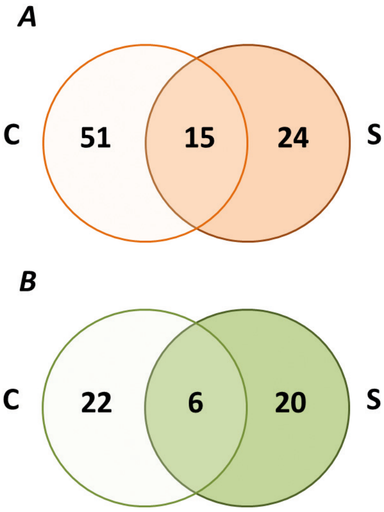
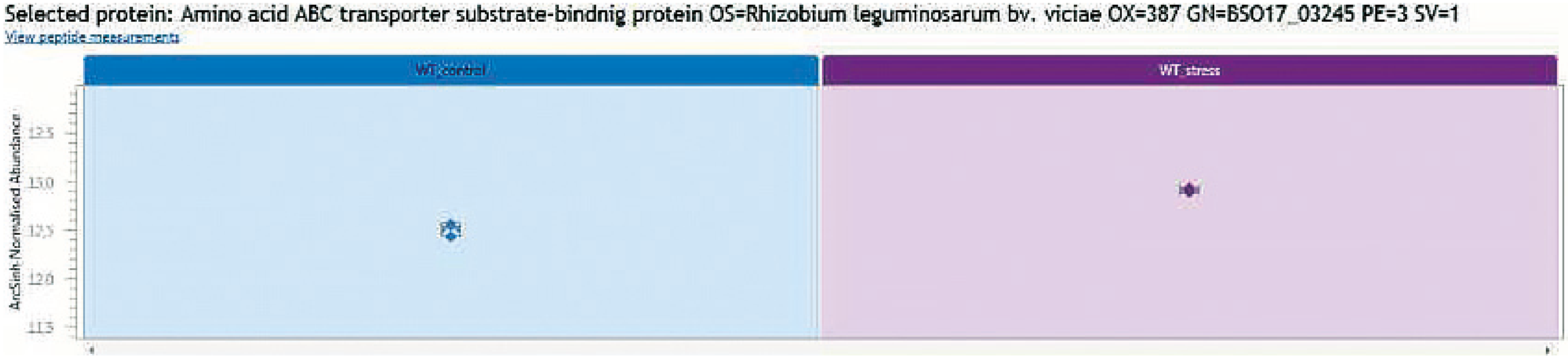


Fig. S1–8. Total numbers of oxidized (A) and glycated (B) peptides identified by nanoLC-MS-based proteomics analysis of rhizobial (*R. leguminosarum*) nodule protein tryptic digests obtained from control (C) and PEG-treated (stressed, S) plants. The data were acquired by nanoHPLC-LIT-Orbitrap-MS operated in the positive DDA mode. Identification of glycated peptides and annotation of glycated proteins relied on Sequest search against *R. leguminosarum* sequence database (Afonin et al. 2017) in Proteome Discoverer 2.1 software with implementation of characteristic modification-specific increments.

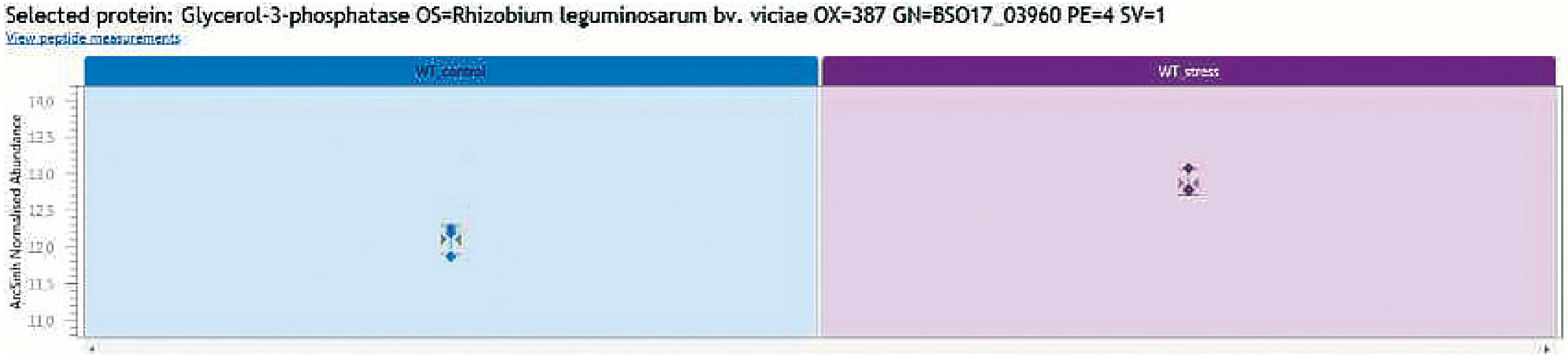
Fig. S1–9. Normalized abundance box plots corresponding to the rhizobial (*Rhizobium leguminosarum bv. viciae*) proteins identified as differentially regulated in presence of osmotic stress. Label-free quantification relied on the Progenesis QI software and ANOVA one-way dispersion analysis accomplished for integral protein content. The q values were obtained with one-way ANOVA dispersion analysis using Progenesis QI software.

Fig. S1–7.1. Amino acid ABC transporter substrate-bindnig protein OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_03245 PE=3 SV=1 q_value = 0,020485161, upregulated protein



Selected protein: Amino acid ABC transporter substrate-bindnig protein OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_03245 PE=3 SV=1

Fig. S1–7.2. Glycerol-3-phosphatase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_03960 PE=4 SV=1 q_value = 0,046527423, upregulated protein



Selected protein: Glycerol-3-phosphatase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_03960 PE=4 SV=1

Fig. S1–7.3. Succinate-semialdehyde dehydrogenase (NADP(+)) OS=Rhizobium leguminosarumbv. viciae OX=387 GN=gabD PE=3 SV=1, q_value = 0,046527423, downregulated protein

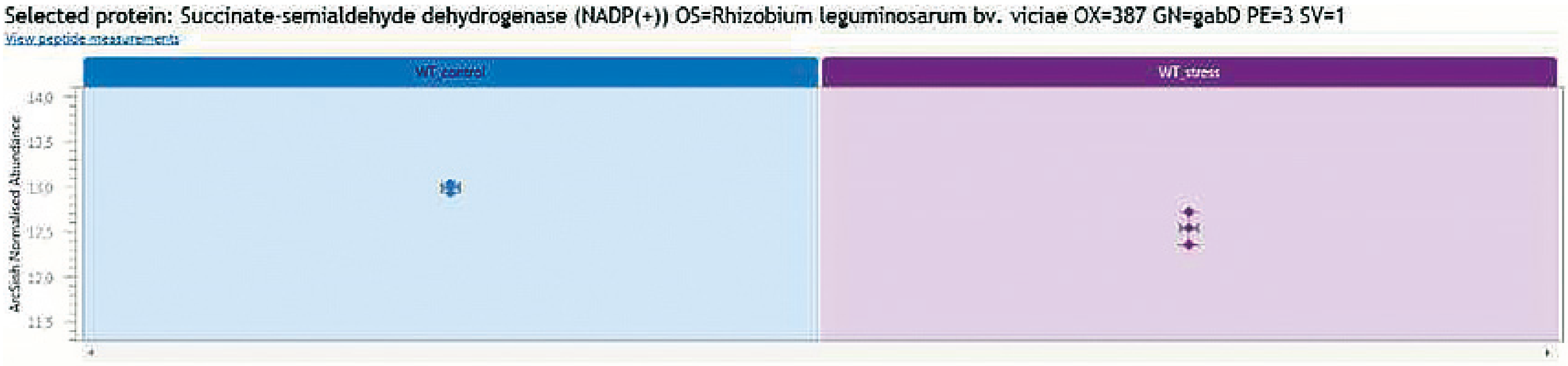

Selected protein: Succinate-semialdehyde dehydrogenase (NADP(+)) OS=Rhizobium leguminosarum bv. viciae OX=387 GN=gabD PE=3 SV=1

Fig. S1–7.4. Uncharacterized protein OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_22220 PE=4 SV=1, q_value = 0,046527423, downregulated protein

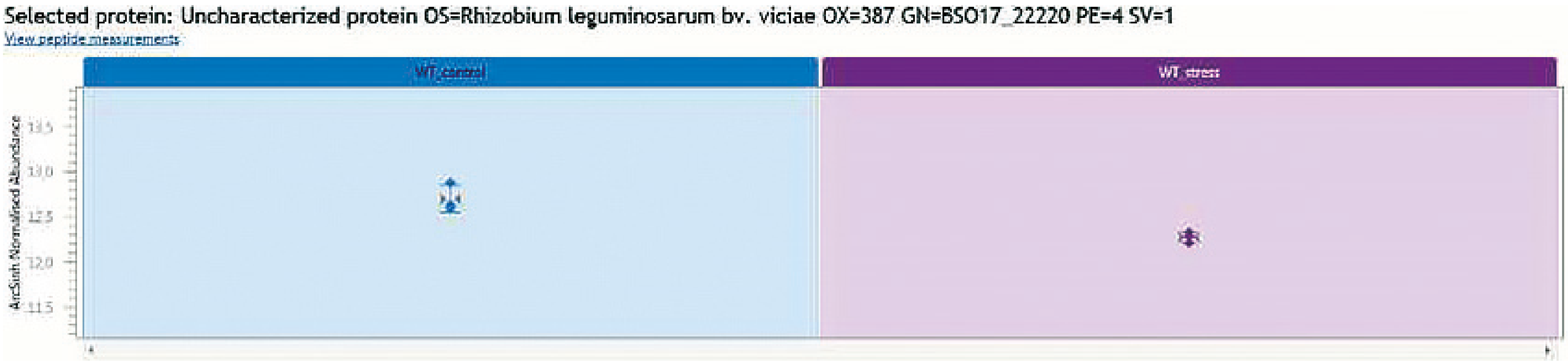
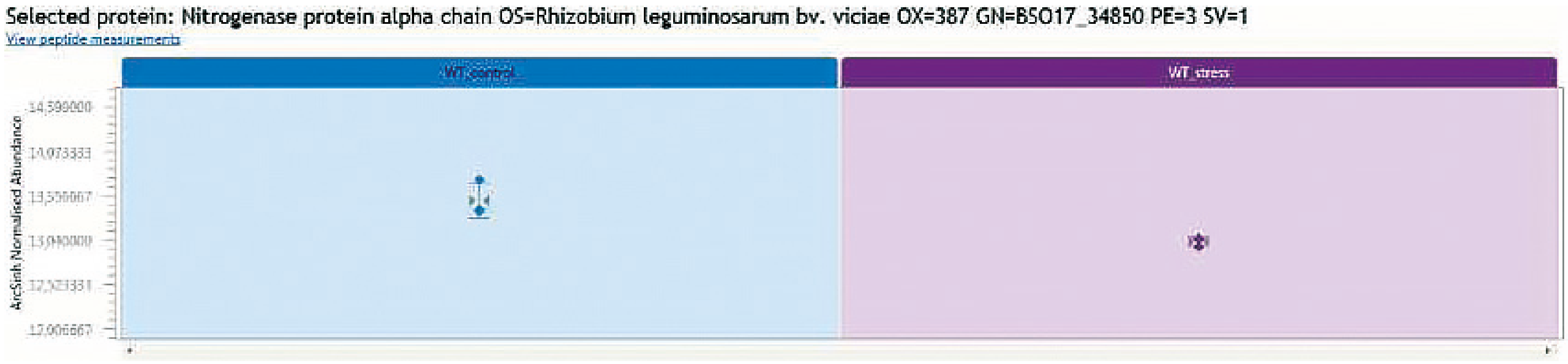
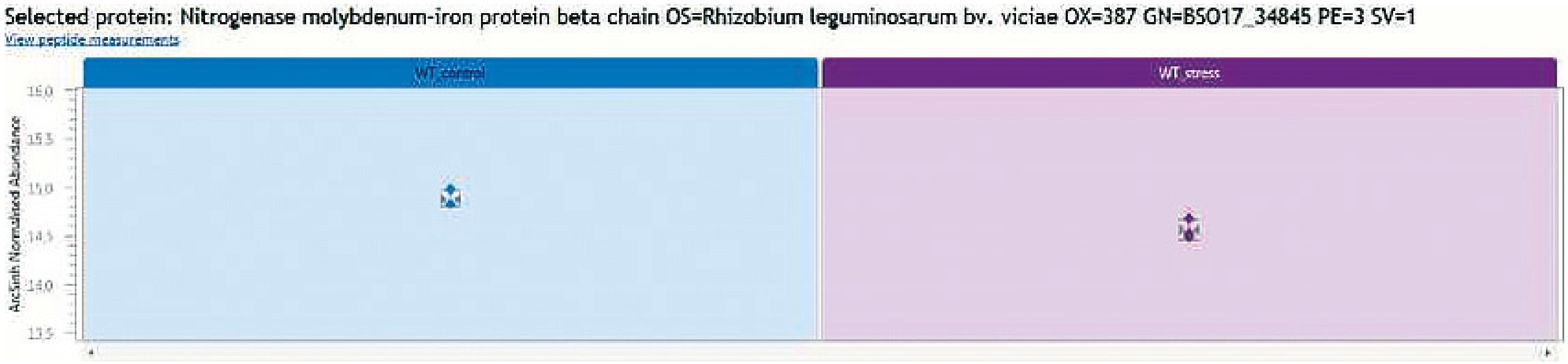

Selected protein: Uncharacterized protein OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_22220 PE=4 SV=1

Fig. S1–7.5. Nitrogenase protein alpha chain OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_34850 PE=3 SV=1, q_value = 0,046527423, downregulated protein



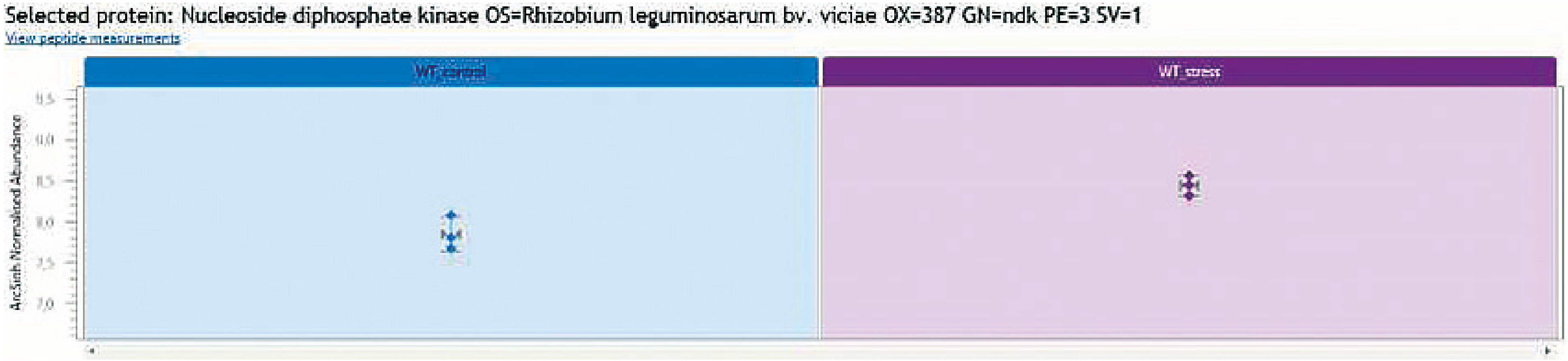
Selected protein: Nitrogenase protein alpha chain OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_34850 PE=3 SV=1

Fig. S1–7.6. Nitrogenase molybdenum-iron protein beta chain OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_34845 PE=3 SV=1, q_value = 0,046527423, downregulated protein



Selected protein: Nitrogenase molybdenum-iron protein beta chain OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_34845 PE=3 SV=1

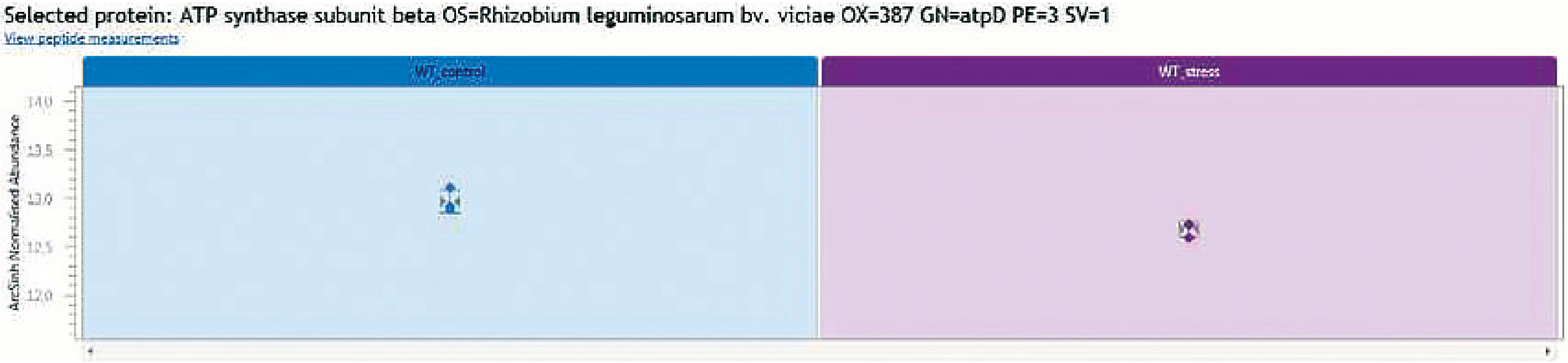
Fig. S1–7.7. Nucleoside diphosphate kinase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=ndk PE=3 SV=1, q_value = 0,046527423, upregulated protein



Selected protein: Nucleoside diphosphate kinase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=ndk PE=3 SV=1

View peptide measurements

Fig. S1–7.8. ATP synthase subunit beta OS=Rhizobium leguminosarum bv. viciae OX=387 GN=atpD PE=3 SV=1, q_value = 0,047946124, downregulated protein



Selected protein: ATP synthase subunit beta OS=Rhizobium leguminosarum bv. viciae OX=387 GN=atpD PE=3 SV=1

View peptide measurements

Fig. S1–7.9. Malate dehydrogenase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=mdh PE=3 SV=1, q_value = 0,047946124, upregulated protein

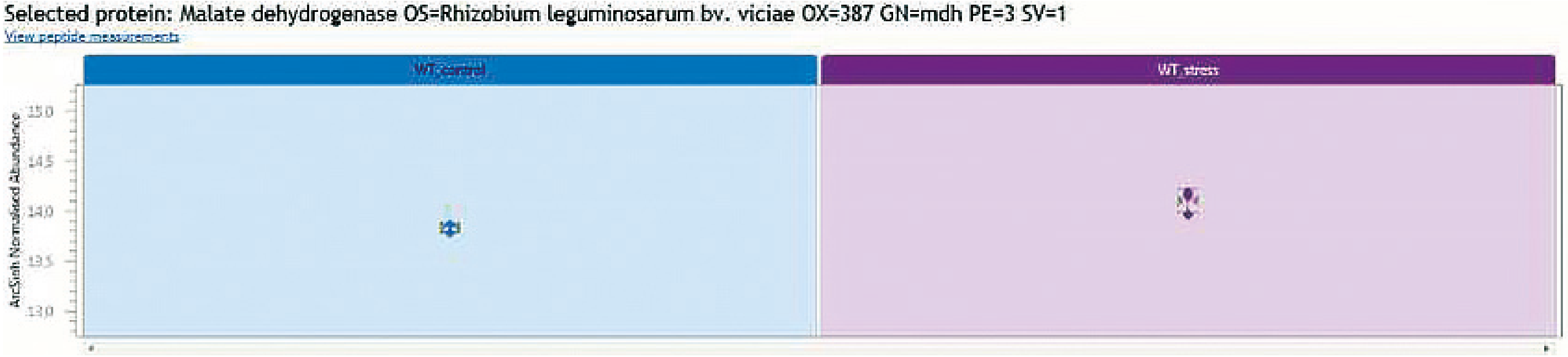

Selected protein: Malate dehydrogenase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=mdh PE=3 SV=1

Fig. S1–7.10. Glutamate–tRNA ligase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=gltX PE=3 SV=1, q_value = 0,053713792, downregulated protein

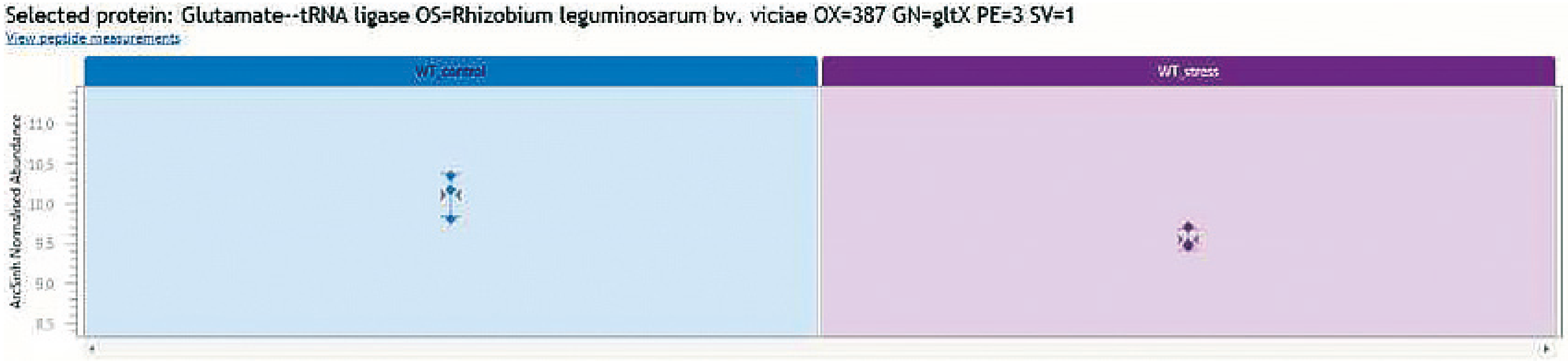
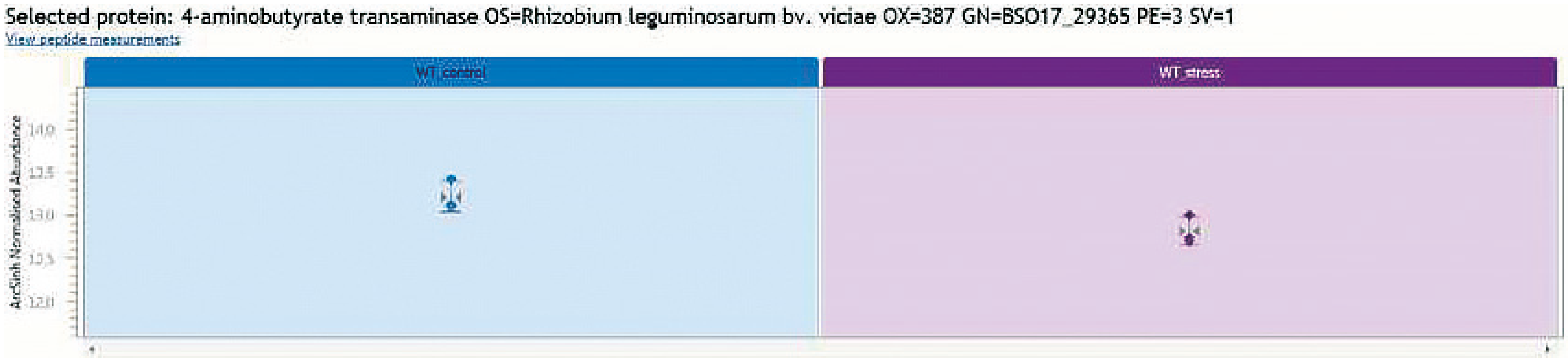
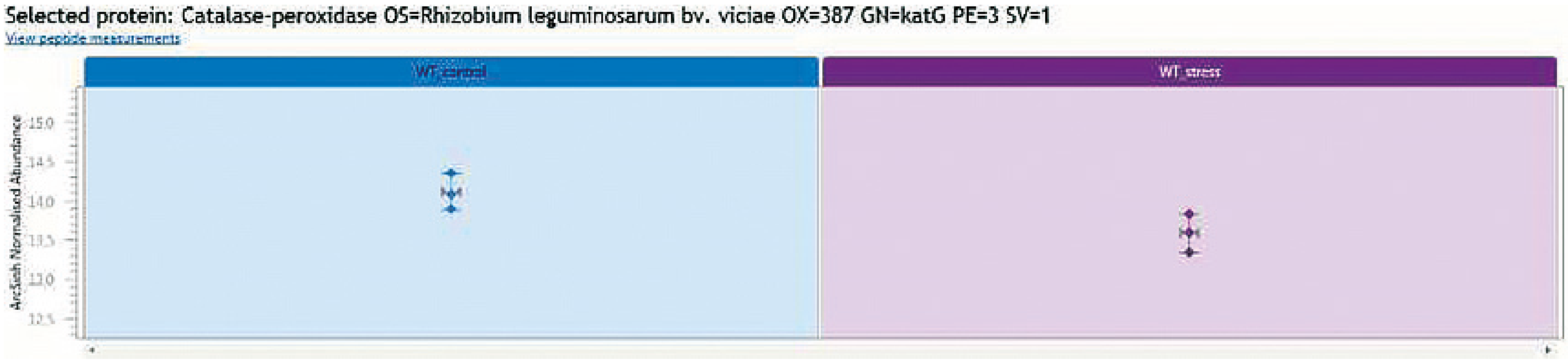

Selected protein: Glutamate--tRNA ligase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=gltX PE=3 SV=1

Fig. S1–7.11. 4-aminobutyrate transaminase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_29365 PE=3 SV=1, q_value = 0,06251984, downregulated protein



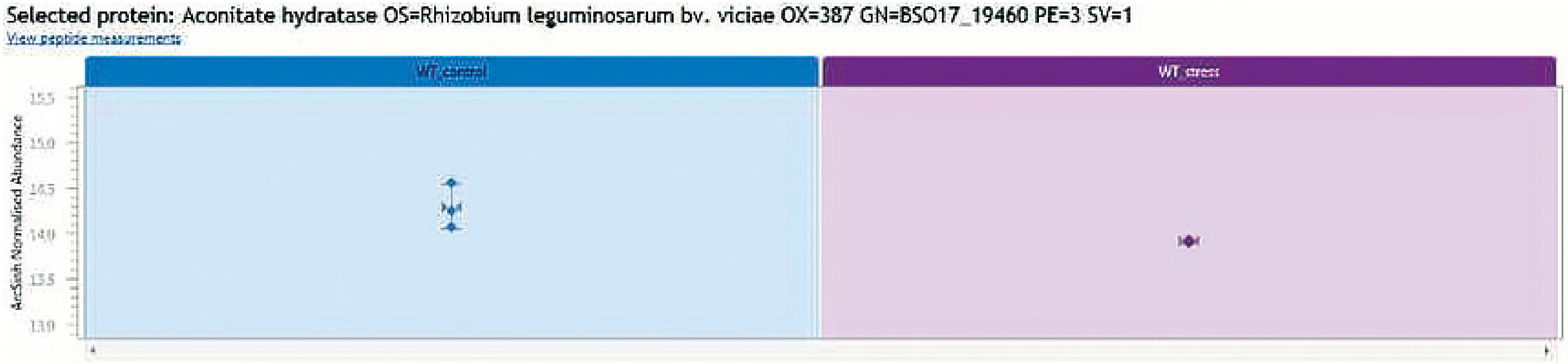
Selected protein: 4-aminobutyrate transaminase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_29365 PE=3 SV=1

Fig. S1–7.12. Catalase-peroxidase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=katG PE=3 SV=1, q_value = 0,067777334, downregulated protein



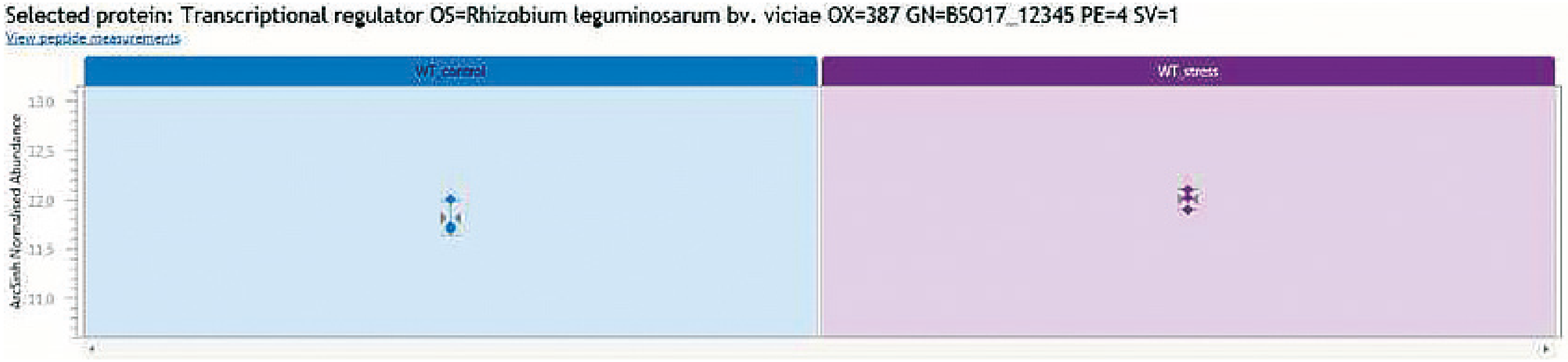
Selected protein: Catalase-peroxidase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=katG PE=3 SV=1

Fig. S1–7.13. Aconitatehydratase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_19460 PE=3 SV=1, q_value = 0,069089611, downregulated protein



Selected protein: Aconitate hydratase OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_19460 PE=3 SV=1

Fig. S1–7.14. Transcriptional regulator OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_12345 PE=4 SV=1, q_value = 0,16282851, upregulated protein



Selected protein: Transcriptional regulator OS=Rhizobium leguminosarum bv. viciae OX=387 GN=BSO17_12345 PE=4 SV=1

Fig. S1–10. Normalized abundance box plots corresponding to the pea (*Pisum sativum*) proteins identified as differentially regulated in presence of osmotic stress. Label-free quantification relied on the Progenesis QI software and ANOVA one-way dispersion analysis accomplished for integral protein content. The q values were obtained with one-way ANOVA dispersion analysis using Progenesis QI software.

Fig. S1–8.1. K4LMW7 Pathogenesis-related protein 10b, q_value = 0,015940083, downregulated protein

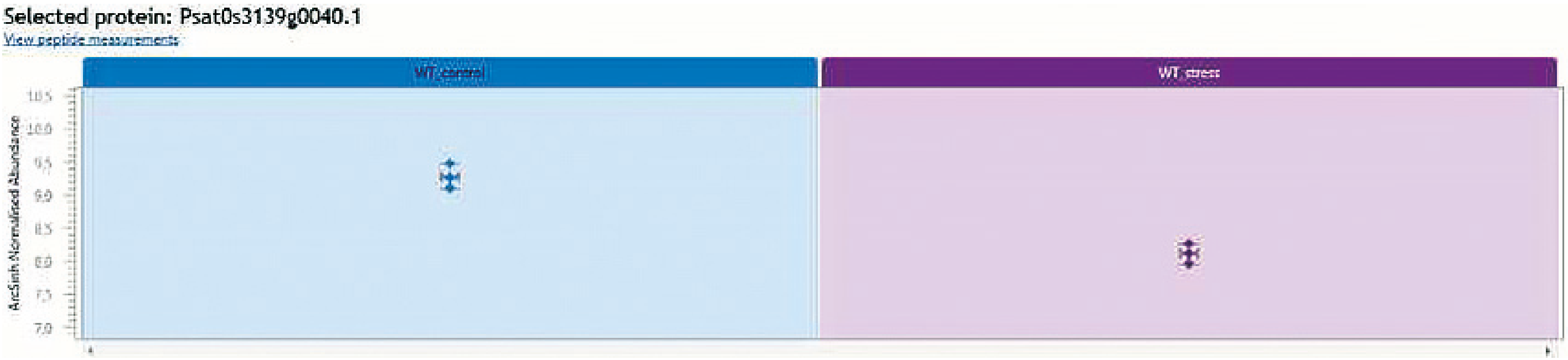


Fig. S1–8.2. A0A2Z6NNA7 Fructose-bisphosphatealdolase (EC 4.1.2.13) q_value = 0,054992063, upregulated protein

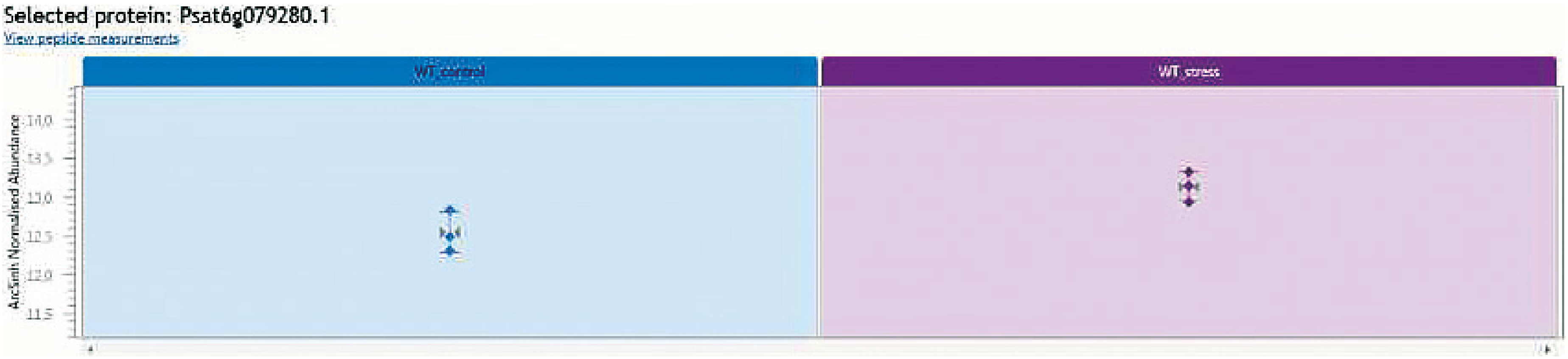


Fig. S1–8.3. A0A2K3MWH0 S-adenosylmethionine synthase (EC 2.5.1.6) q_value = 0,054992063, downregulated protein

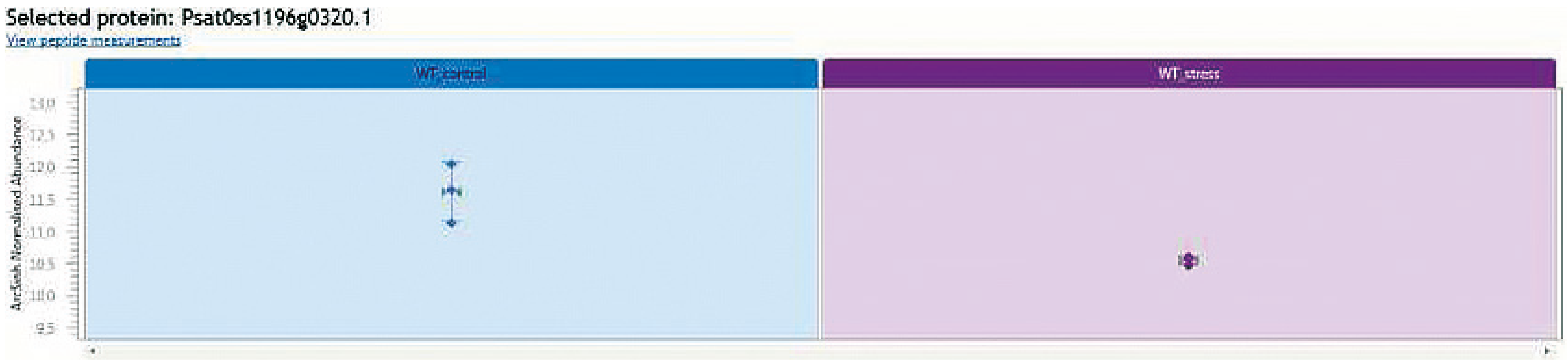


Fig. S1–8.4. P93851 Leghemoglobin K, q_value = 0,054992063, upregulated protein

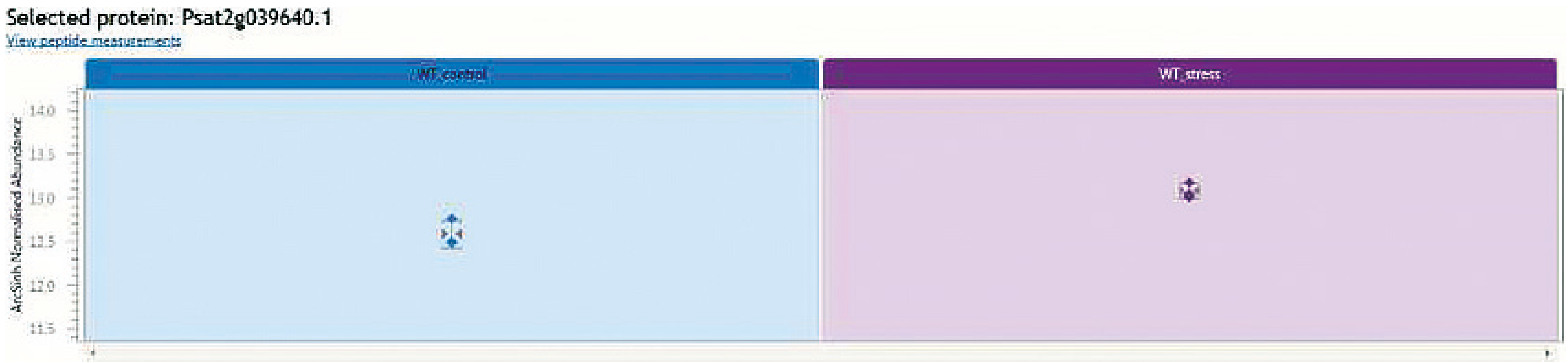


Fig. S1–8.5. A0A1S2YB61 Serine hydroxymethyltransferase (EC 2.1.2.1) q_value = 0,054992063, downregulated protein

Fig. S1–8.6. B7FMK4 40S ribosomal protein S24 q_value = 0,054992063, downregulated protein

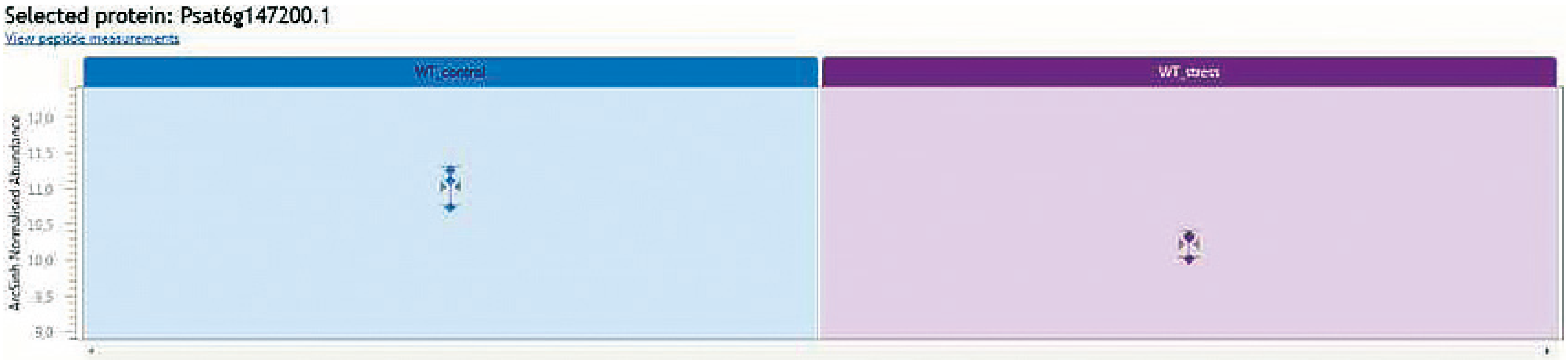


Fig. S1–8.7. A0A396JGV3 Putative ribosomal protein L7 q_value = 0,054992063, down-regulated protein

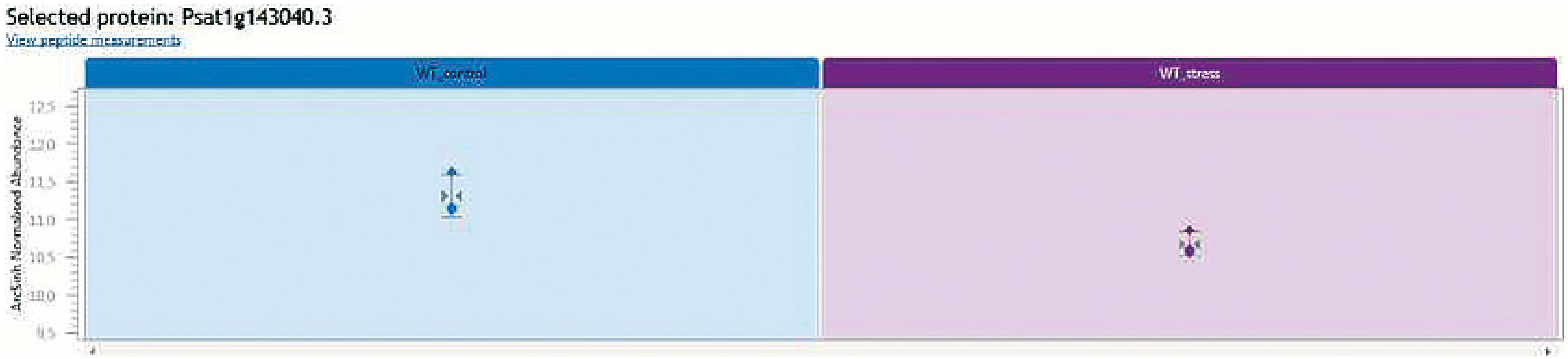
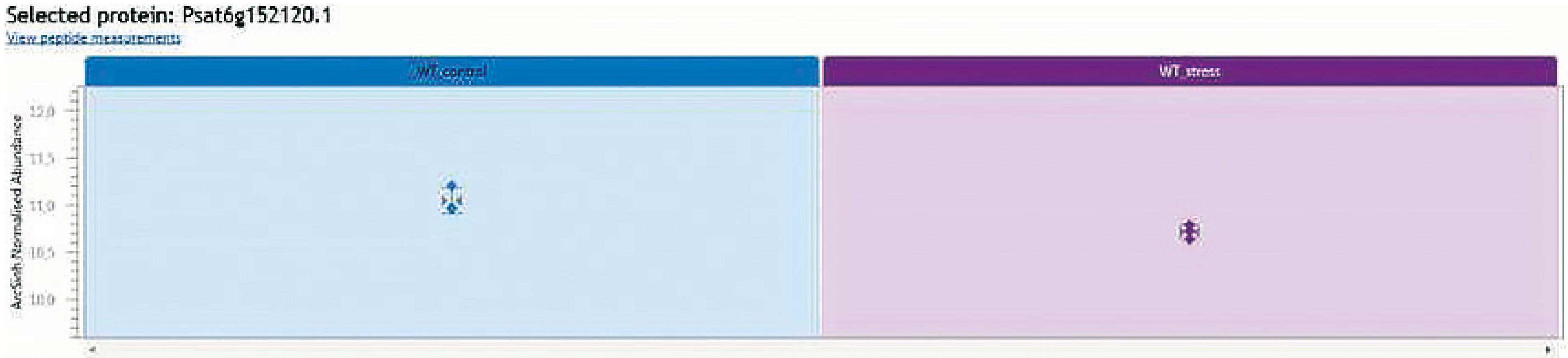


Fig. S1–8.8. A0A1S2Y9D8 Aminopeptidase (EC 3.4.11.-) q_value = 0,054992063, downregulated protein

Fig. S1–8.9. A0A2K3NQW3 Calreticulin q_value = 0,054992063, downregulated protein



Selected protein: Psat6g152120.1

Fig. S1–8.10. A0A2P4IHJ4 Deleted. KeggNOG-mapper: Tubulin is the major constituent of microtubules. It binds two moles of GTP, one at an exchangeable site on the beta chain and one at a non-exchangeable site on the alpha chain q_value = 0,054992063, downregulated protein

Selected protein: Psat6g198160.1

Fig. S1–8.11. A0A2K3NST0 Leghemoglobin q_value = 0,069087391, upregulated protein

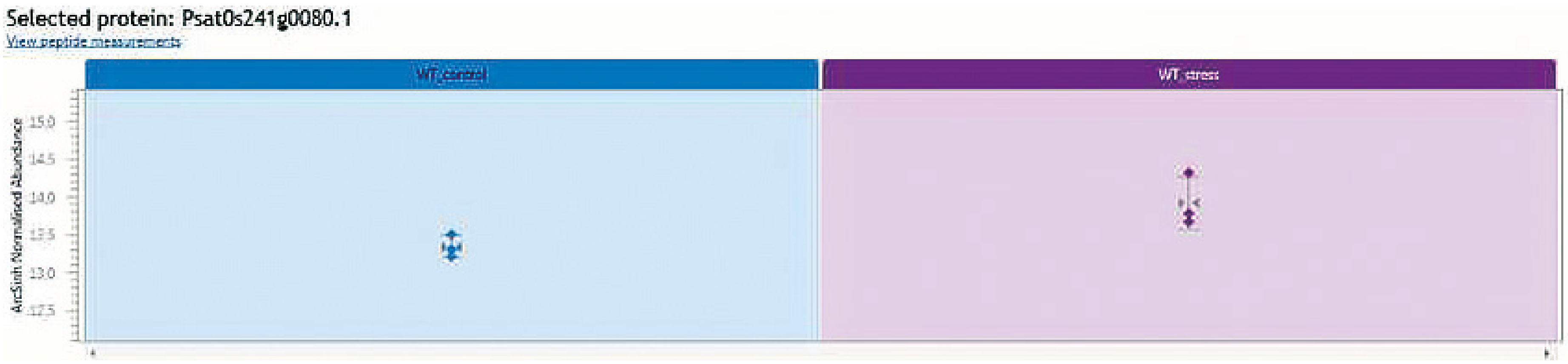


Fig. S1–8.12. B7FJQ4 Malate dehydrogenase (EC 1.1.1.37) q_value = 0,070454934, upregulated protein

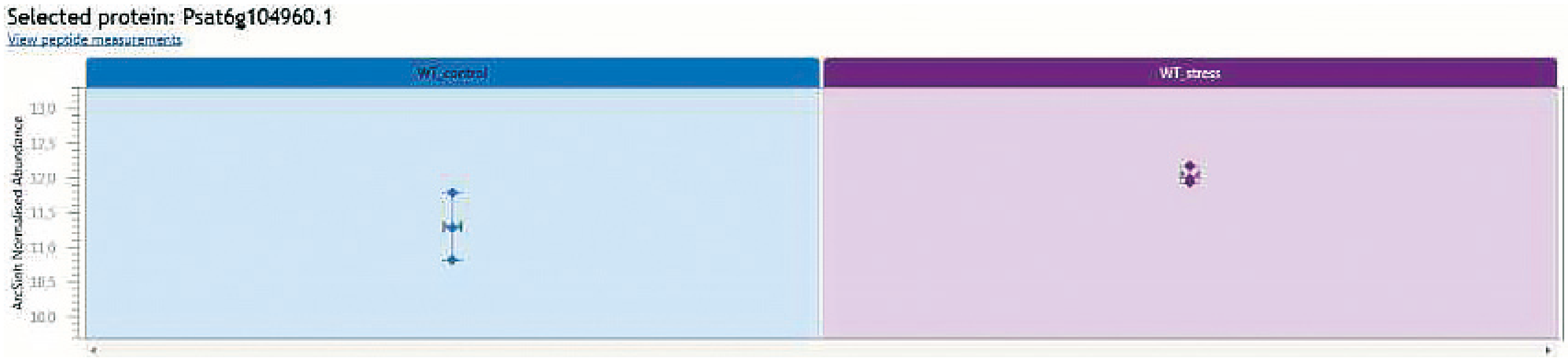
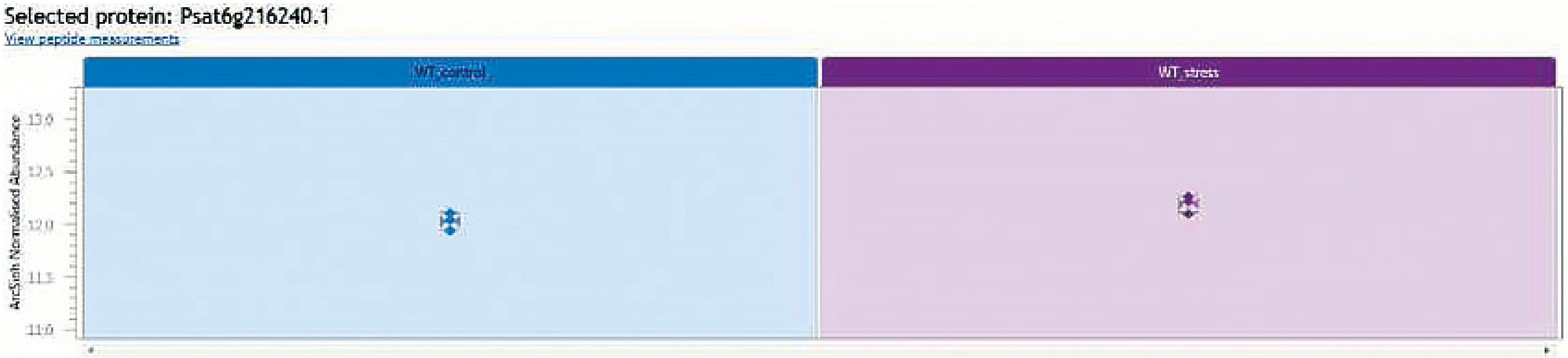


Fig. S1–8.13. A0A2K3NWF8 Protein disulfide-isomerase q_value = 0,070454934, downregulated protein

Selected protein: Psat5g086440.1

Fig. S1–8.14. A0A2K3MYZ6 Golgin subfamily A member 4-like protein q_value = 0,070454934, upregulated protein



Selected protein: Psat6g216240.1

Table S1–1. Protein concentrations and yields acquired for the isolates obtained from the root nodules of the control and drought-treated pea plants

	Name	Concentration (mg/ml)	SD	%	Freshweight (g)	Yields of the protein extraction (mg/g)
Control	1wt	2.27	0.54	23.84	0.073	31.2
	2wt	1.84	0.38	20.42	0.069	26.6
	3wt	1.37	0.15	11.21	0.056	24.4
Drought stress	1wt	2.62	1.00	38.37	0.084	31.2
	2wt	1.72	0.46	26.95	0.0108	159.3
	3wt	0.87	0.13	14.49	0.095	9.2