ABRATES

Journal of Seed Science

ISSN 2317-1545

www.abrates.org.br/revista

ARTICLE

Journal of Seed Science, v.46, e202446010, 2024

http://dx.doi.org/10.1590/2317-1545v46281673

Peptidomics analysis reveals stress response proteins involved in the establishment of seed vigor in tobacco

Guoping Wang[1], Limeng Zhang[1], Yongzhi Niu[1], Ji Gu[1], Shuai Yang[1], Yongqi He[2], Zhoufei Wang[2]*, Yunye Zheng[1]*

ABSTRACT: Seed vigor of tobacco (*Nicotiana tabacum* L.) is established during seed development stage, while its regulators remain largely unknown. Here, a comparative peptidomics analysis of the developing seeds was conducted to reveal the regulators involving the establishment of tobacco seed vigor. The most significant difference of seed vigor was observed between seeds harvested at 20 and 30 days after pollination (DAP), and then the corresponding seeds were collected separately for peptidomics analysis. A total of 2932 and 2812 nonredundant peptides were identified in seeds harvested at 20 and 30 DAP, respectively. In which, 349 differentially expressed peptides (DEPs) were characterized. To explore the potential functions of these DEPs, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were further analyzed according to their precursor proteins. Most DEP precursor proteins were involved in response to abiotic stimulus, response to water, and protein processing in endoplasmic reticulum. Further, the peptides derived from the precursor proteins, such as late embryogenesis abundant (LEA) protein, heat shock protein, peroxiredoxin, and globulin proteins, may regulates the establishment of seed vigor by influencing reactive oxygen species. The results provide a foundation for further exploration of the peptides functions on the establishment of seed vigor in tobacco.

Index terms: *Nicotiana tabacum,* peptides, seed development, seed vigor.

RESUMO: O vigor das sementes de tabaco (*Nicotiana tabacum* L.) é estabelecido durante a fase de desenvolvimento das sementes, enquanto seus reguladores permanecem em grande parte desconhecidos. Foi realizada uma análise peptidômica comparativa das sementes em desenvolvimento para revelar os reguladores que envolvem o estabelecimento do vigor das sementes de tabaco. A diferença mais significativa no vigor das sementes foi observada entre as sementes colhidas aos 20 e 30 dias após a polinização (DAP), sendo então as sementes correspondentes coletadas separadamente para análise peptidômica. Um total de 2.932 e 2.812 peptídeos não redundantes foram identificados em sementes colhidas aos 20 e 30 DAP, respectivamente. Foram então caracterizados 349 peptídeos diferencialmente expressos (DEPs). Para explorar as funções potenciais desses DEPs, as análises da *Gene Ontology* (GO) e da *Kyoto Encyclopedia of Genes and Genomes* (KEGG) foram analisadas de acordo com suas proteínas precursoras. A maioria das proteínas precursoras da DEP estava envolvida na resposta ao estímulo abiótico, na resposta à água e no processamento de proteínas no retículo endoplasmático. Além disso, os peptídeos derivados das proteínas precursoras, como a *late embryogenesis abundant* (LEA), a proteína de choque térmico, a peroxirredoxina e as proteínas globulinas, podem regular o estabelecimento do vigor das sementes influenciando as espécies reativas de oxigênio. Os resultados fornecem uma base para uma maior exploração das funções dos peptídeos no estabelecimento do vigor das sementes de tabaco.

Termos de indexação: *Nicotiana tabacum*, peptídeos, desenvolvimento de sementes, vigor de sementes.

Corresponding author
wangzf@scau.edu.cn
zheng_yunye@126.com

Received: 12/21/2023.
Accepted: 04/08/2024.

[1]Yuxi Zhongyan Seed Company Ltd., Seed Engineering Technology Center of Yunnan Province, Yuxi 653100, China.

[2]The Laboratory of Seed Science and Technology, Guangdong Key Laboratory of Plant Molecular Breeding, South China Agricultural University, Guangzhou 510642, China.

INTRODUCTION

Tobacco (*Nicotiana tabacum* L.) as a cash crop is one of the most important crops in Yunnan Province of China. Cultivation area of tobacco is approximately 500,000 hectares in Yunnan Province, accounting 12.5% of global cultivation area. Seed vigor is an agronomical trait that determines the speed seed germination, vigorous seedling growth, and stress tolerant ability (Zhao et al., 2021). Seeds with high vigor will promote seed germination and seedling establishment in fields, which is critical for tobacco production. Seed vigor is established during seed development. Seed vigor of tobacco is high when harvested at seed maturity stage after 30 days pollination (Pan et al., 2016). However, the regulators involving the establishment of seed vigor are still unclear in tobacco.

Seed development can be divided into the embryogenesis, seed filling, and late drying maturation stages (Leprince et al., 2017). A lot of factors, such as reserves accumulation and desiccation tolerance, have been reported involving the establishment of seed vigor during seed development (Zhao et al., 2021). For example, the accumulated storage protein during development stage may affect seed germination and seedling establishment in soybean (Wei et al., 2020). Seeds can be divided into orthodox and recalcitrant seeds, and the orthodox seeds are desiccation tolerant and can be stored at low temperature (Kijak and Ratajczak, 2020). The acquisition of desiccation tolerance of orthodox seeds is included the accumulation of several diverse protective molecules, such as the LEA proteins, small heat shock proteins (sHSPs), universal stress proteins (USPs), antioxidative proteins, and dehydrins (Wang et al., 2015; Zhao et al., 2021). Tobacco seeds are orthodox seeds that generally gain their full desiccation tolerance at the last development stage. Whether peptides derived from the above-mentioned protective proteins involving the establishment of seed vigor is still unclear in tobacco.

Small peptides with less than 100 amino acids have been revealed involving both plant development and stress tolerance (Matsubayashi, 2014; Takahashi et al., 2018; Zeng et al., 2022). For example, the five-amino-acid PEPTIDE 1 (PEP1) encoded by *OsPEP1* regulates root development in rice (Xiang et al., 2021). Cysteine-rich peptides (CRPs) regulate the development of pollen, anther, and seed in *Arabidopsis* (Ge et al., 2010; Wuest et al. 2010). Several peptides, such as Clavata3/Embryo-surrounding Region-related 25 (CLE25) peptide (Christmann and Grill, 2018), endogenous Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 proteins (CAP)-derived peptide 1 (AtCAPE1) (Chien et al., 2015), and AtPep3 peptide (Nakaminami et al., 2018), regulates stress tolerance in *Arabidopsis*. Rice drought and salt stress response-1 (OsDSSR1) peptide (Cui et al., 2018) and drought tolerance 11 (OsDT11) peptide (Li et al., 2017) have been indicated to be associated with stress tolerance. Peptidomics provides a powerful method to identify peptides with the potential functional roles in plants (Ma et al., 2022), and it is now being used for bioactive peptide and biomarker discovery (Agyei et al., 2018; Trindade et al., 2021). However, peptidomics study on the establishment of seed vigor is limitedly in tobacco.

In this study, a comparative peptidomics was conducted in seeds harvested at two different developmental stages in tobacco. We observed that the majority differentially expressed peptides (DEPs) were derived from the precursor proteins such as LEA protein, heat shock protein, peroxiredoxin, and globulin proteins. The identified DEPs are hypothesized to be involved in response to desiccation tolerance. This work provides a foundation for further exploration of the peptide functions on the establishment of seed vigor in tobacco.

MATERIAL AND METHODS

Plant materials

Tobacco cultivar 'K326' was used in this study. The plants were grown in the experimental field of Xishuangbanna in Yunnan Province of China. Seeds were harvested at 15, 20, 25, 30, and 35 days after pollination (DAP). All seeds were dried to the moisture content approximately 4.5% for the following usage (Niu et al., 2023).

Seed germination

One hundred seeds per replicate were hydrated on the top of filter paper in Petri dishes with distilled water at 25 °C and light/dark for 12 hours conditions for 14 days. Germination ability was observed daily. Seeds were considered as germinated when the normal radicle protruded through seed coat and the cotyledon unfurled. The establishment of seedling was considered when the normal radicle and shoot observed. Germination percentage was tested after 7- and 14-days germination, respectively. Meanwhile, germination index (GI) was calculated as follows: $GI = \Sigma (Gt/t)$, where Gt is the number of the germinated seeds on day t (Niu et al., 2023). A total of 10 plants were randomly selected for the evaluation of root length. Three biological replications were performed.

Polypeptide extraction

Freshly harvested seeds at 20 and 30 DAP were used for polypeptide extraction. Firstly, the cracking solution (Roche Ltd. Basel, Switzerland) was added into the cracked sample for shaking and grinding for 400 s three times, and then was cracked on ice for 30 min. The supernatant was collected after the centrifuge at 4 °C for 15,000 rpm for 15 min. Polypeptide filtration was conducted using Ultrafiltration Spin Columns (Millipore, Billerica, USA) to centrifuge at 4 °C for 8000 g for 30 min to remove high molecular weight protein. Enzymatic hydrolysis of 3 to 10 KD polypeptide was conducted as follows: concentrated by centrifugal concentrator, drained, and redissolved with 100 µL 100 mM TEAB; add 5 µL trypsin to each sample, and carry out enzymolysis at 37 °C overnight. The C18 Zip Tip (MonoSpin C18, GL) was used to desalinate samples. The eluted peptide was drained by vacuum concentrator and stored at 80 °C for mass spectrometry detection. The polypeptide extraction was conducted by Gene Denovo Biotechnology Co. (Guangzhou, China). Three biological replications were performed.

Nano LC-MS/MS analysis

Samples were analyzed by on-line nanospray LC-MS/MS on an Thermo Scientific™ Orbitrap Fusion Lumos ™ coupled to an EASY-nano-LC 1200 system (Thermo Fisher Scientific, MA, USA). 4 µL peptide was loaded (analytical column: Acclaim PepMap C18, 75 µm x 25 cm) and separated with a 120 min linear gradient, from 6% B (B: 0.1% formic acid in 80% ACN) to 36% B. The column flow rate was maintained at 400 nL/min with the column temperature of 40 °C. The electrospray voltage of 2 kV versus the inlet of the mass spectrometer was used. The mass spectrometer was run under data dependent acquisition mode, and automatically switched between MS and MS/MS mode. The parameters were: (1) MS: scan range (m/z) =100–1500; resolution=120,000; AGC target=4e5; maximum injection time=50 ms; include charge states=2-7; (2) HCD-MS/MS: resolution=15,000; isolation window=3; AGC target=5e4; maximum injection time=35 ms; collision energy=27,32,3. The nano LC-MS/MS analysis was conducted by Gene Denovo Biotechnology Co. (Guangzhou, China). Three biological replications were performed.

Database search and analysis

Raw data were processed and analyzed by Spectronaut X (Biognosys AG, Switzerland) with default settings to generate an initial target list. Spectronaut was set up to search the database of tobacco along with contaminant database assuming trypsin as the digestion enzyme. Carbamidomethyl (C) was specified as the fixed modification. Oxidation (M) was specified as the variable modifications. The false discovery rate (FDR) Q value cutoff on precursor and protein level was applied 1%. Retention time prediction type was set to dynamic iRT. Data extraction was determined by Spectronaut X based on the extensive mass calibration. Spectronaut Pulsar will determine the ideal extraction window dynamically depending on iRT calibrati on and gradient stability (Kim et al., 2018). Q value cutoff on precursor and protein level was applied 1%. Decoy generation was set to mutated which similar to scrambled but will only apply a random number of AA position swamps (min=2, max=length/2). All selected precursors passing the filters were used for quantification. The average top 3 filtered peptides which passed the 1% Q value cutoff were used to calculate the major group quantities.

Differentially expressed peptide (DEPs) analysis and functional annotation

Differentially expressed peptide were analyzed with Student's *t* Test and Benjamini and Hochberg (BH). After that, different expressed proteins were filtered with the selection criteria of fold change > 1.2 and *P* value < 0.05. For all the identified polypeptides, use the proteins to which they belong. The protein functions and classification were analyzed based on searches against the following databases: Gene Ontology (EuKaryotic Orthologous Groups (KOG)/ Cluster of Orthologous Groups of Proteins (COG), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Okuda et al., 2008).

Statistical analysis

Data were analyzed using GraphPad Prism 9 (GraphPad Prism, San Diego, CA, USA). Significant differences among samples were compared using Student's *t*-test.

RESULTS

Characteristics of seed vigor were tested during seed development in tobacco. Generally, the traits of seed vigor were gradually increased with the process of seed development in tobacco. Germination percentage (7 d) of seeds harvested at 15, 20, 25, 30, and 35 DAP was approximately 0%, 39%, 86%, 86%, and 90%, respectively (Figure 1a).

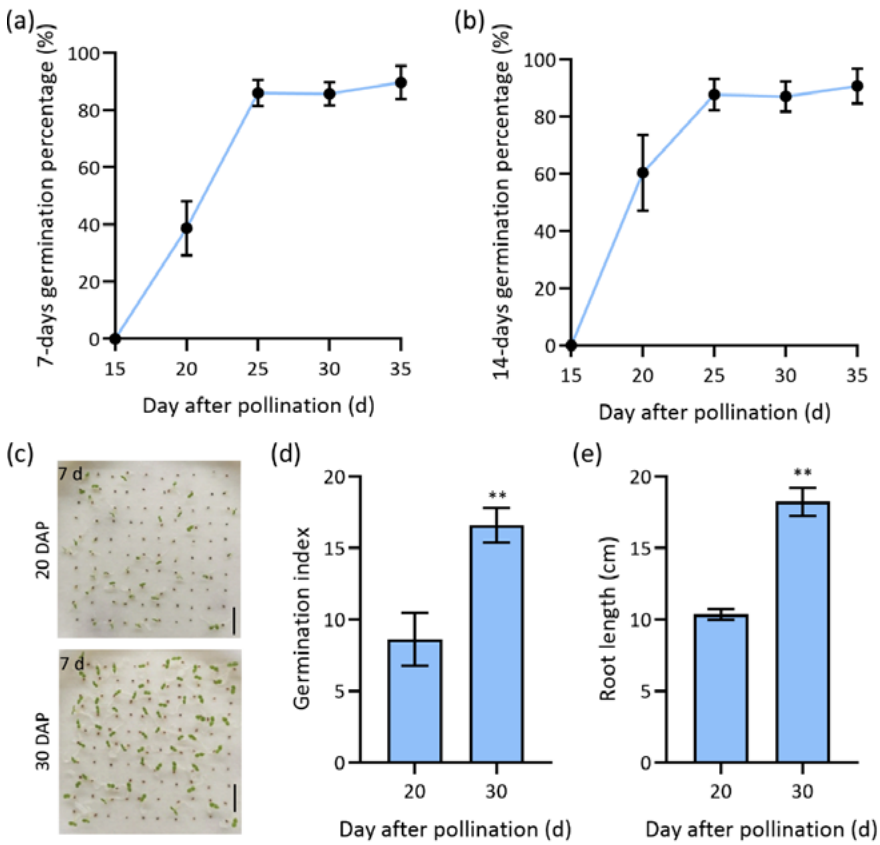

Figure 1. Characteristics of seed vigor during seed development in tobacco. (a) 7-days germination percentage and (b) 14-days germination percentage in seeds harvested at 15, 20, 25, 30, and 35 days after pollination (DAP). (c) Representative images of seed germination for 7 days in seeds harvested at 20 DAP and 30 DAP. Scale bars represent 10 mm. Comparisons of (d) germination index and (e) root length between seeds harvested at20 DAP and 30 DAP. Data were presented as mean ± SD, *n* = 3 biological replicates. In (d), (e) significant difference was determined by two-tailed Student's *t*-tests (**$P < 0.01$).

Germination percentage (14 d) of seeds harvested at 15, 20, 25, 30, and 35 DAP was approximately 0%, 60%, 88%, 87%, and 91%, respectively (Figure 1b). Similarly, the significant difference of seedling growth was observed between the early (20 DAP) and late (30 DAP) harvested seeds (Figure 1c). Germination index and root length were significantly higher in seeds harvested at 30 DAP compared with those of seeds harvested at 20 DAP (Figure 1de). These data suggest that the initial period for the establishment of seed vigor is from 20 DAP in tobacco.

To reveal the peptides involved in the establishment of tobacco seed vigor, seeds harvested at 20 DAP and 30 DAP were used for peptidomics analysis. To investigate the relationships among samples, principal component analysis (PCA) and Pearson correlation analysis were performed for the peptide expression. PCA indicated that PC1 and PC2 explained 58.8% and 22.8% of the peptide expression variation in all samples, respectively (Figure 2a). Pearson correlations between the replicates of 20 DAP samples (VS1-1, VS1-2, and VS1-3) and 30 DAP samples (VS2-1, VS2-2, and VS2-3) were larger than 0.9 (Figure 2b). The correlation results suggest that peptide expression was consistent among samples. Peptidomics analysis identified 2932 and 2812 nonredundant peptides derived from seeds of 20 DAP and 30 DAP, respectively (Figure 2c). In which, 983 endogenous peptides had fewer than 10 amino acids, 1549 had 10–15 amino acids, and 819 had more than 15 amino acids (Figure 2d).

To reveal the DEPs in the developing seeds of tobacco, the peptide expression was compared between seeds harvested at 20 DAP and 30 DAP. A total of 349 DEPs were identified in seeds harvested at 20 DAP compared with that of seeds harvested at 30 DAP in tobacco (Figure 3a). Of them, 145 and 204 DEPs were up-regulated and down-regulated,

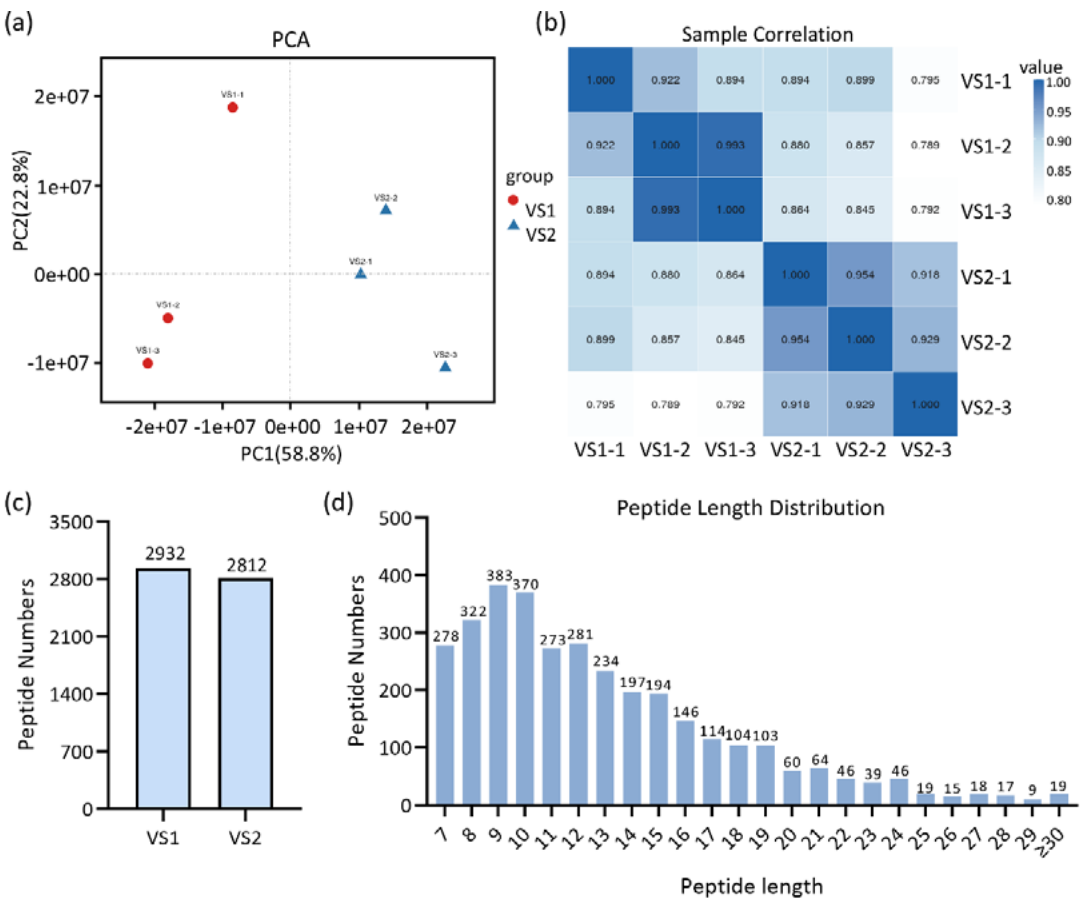

Figure 2. Identification of peptides in the developing seeds of tobacco. (a) The principal component analysis (PCA) and (b) Pearson correlation analysis among seed samples harvested at 20 DAP (VS1) and 30 DAP (VS2). (c) The number of identified peptides in seeds harvested at 20 DAP (VS1) and 30 DAP (VS2). (d) Distribution of the identified peptide lengths.

respectively (Figure 3b). To preliminarily explore the potential functions of these DEPs on the establishment of seed vigor, we performed GO and KEGG pathway analyses of their precursor proteins. We observed that these precursor proteins were mainly involved in response to abiotic stimulus and response to water (Figure 3c). KEGG analysis mapped that these precursor proteins were enriched in several pathways, the first one of which was protein processing in endoplasmic reticulum (Figure 3d). These data indicate that the stress responses may play important roles on the establishment of seed vigor during seed development in tobacco.

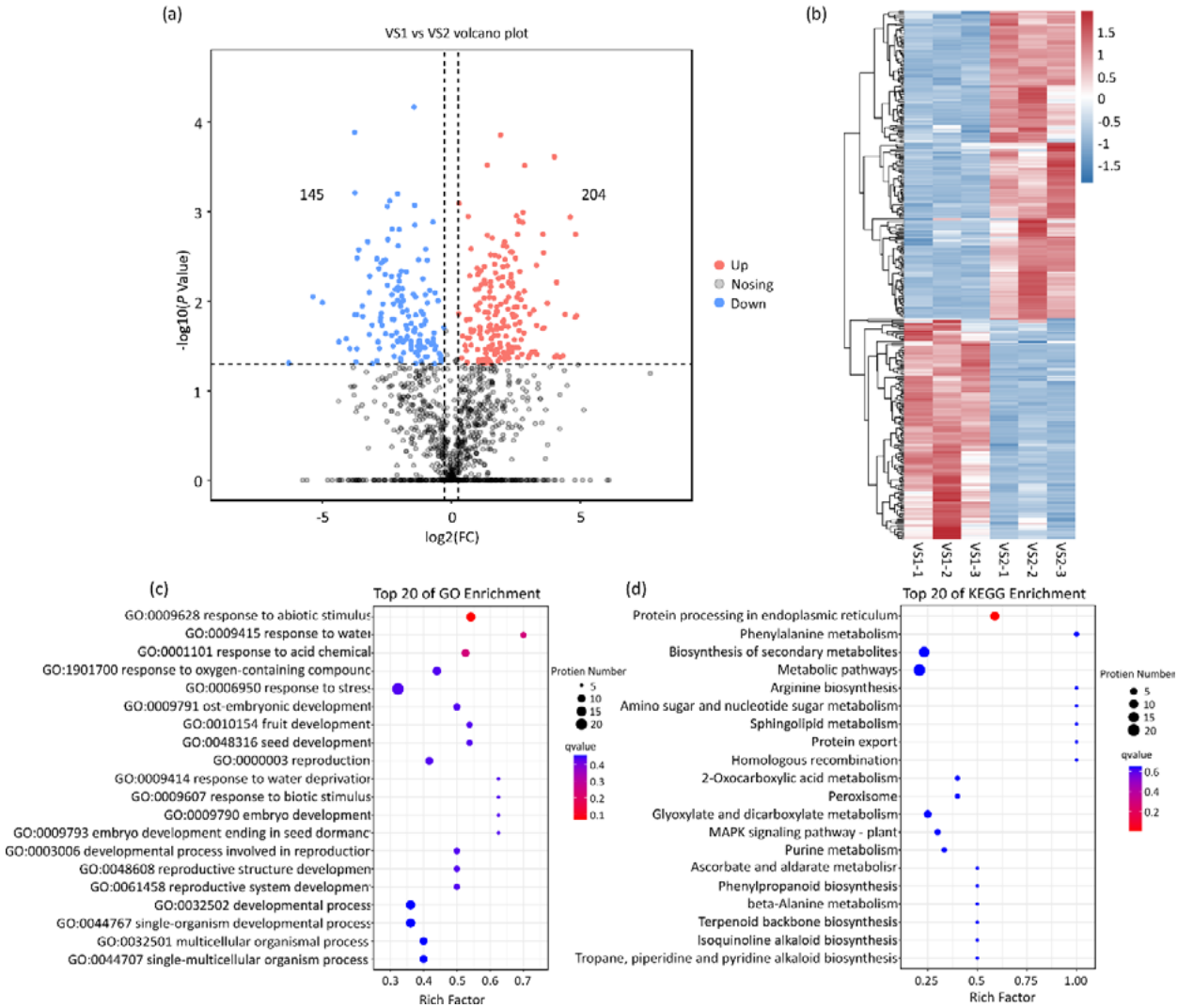

Figure 3. Characteristics of differentially expressed peptides in the developing seeds of tobacco. (a) Volcano plot of differentially expressed peptides. The horizontal axis represents the expressed fold change. The vertical axis represents the degree of statistical significance in differential expression. The higher -log10 (FDR) values represent greater differences. Black dots indicate no significant changes between seeds harvested at 20 DAP (VS1) and 30 DAP (VS2). The up-regulated peptides are represented by a red dot, down-regulated peptides by a blue dot. (b) Hierarchical clustering analysis of the differentially expressed peptides. Red, up regulation; Blue, down regulation. Values represent the log2 fold changes of peptide. (c) GO and (d) KEGG analysis of peptide precursor proteins. The horizontal axis represents the enrichment degree rich factor value, and the vertical axis represents the GO term or KEGG pathway information. The size of the circle represents the number of peptide precursor proteins in the pathway; the larger circle indicates the more number. The color of the circle represents the corrected *p*-value size; the darker the color, the smaller the *p*-value.

To further reveal the factors regulating the establishment of seed vigor, the details of DEPs involved in response to abiotic stimulus, response to water, and protein processing in endoplasmic reticulum were analyzed (Table 1; Table 2). In which, most DEPs derived from 11 kDa LEA-like and 1-Cys peroxiredoxin-like were involved in response to abiotic stimulus and response to water (Table 1). Meanwhile, most DEPs derived from 17.3 kDa class II heat shock protein-like and 17.9 kDa class II heat shock protein were involved in protein processing in endoplasmic reticulum (Table 2). Interestingly, the expressions of all DEPs derived 11 kDa LEA-like and the expressions of majority DEPs derived 17.3 kDa class II heat shock protein-like were significantly decreased in seeds harvested at 30 DAP compared with those of seeds harvested at 20 DAP, while the expressions of majority DEPs derived from 1-Cys peroxiredoxin-like were increased. The up- and down-regulated DEPs may participate in stress responses involving in the establishment of seed vigor in tobacco.

In the present data, we observed that a larger number of DEPs were derived from 11S globulin subunit beta-like proteins (Table 3). In which, more than one DEPs was derived from the same precursor protein, and many DEPs came from the same 11S globulin subunit beta-like proteins. Especially, more than 20 DEPs were derived from four types of 11S globulin subunit beta-like proteins (Nitab4.5_0000044g0350, Nitab4.5_0000436g0260, Nitab4.5_0001855g0030, and Nitab4.5_0004662g0050). Interestingly, the expressions of majority DEPs derived from 11S globulin were significantly increased in seeds harvested at 30 DAP compared with those of seeds harvested at 20 DAP seeds. These results indicate that the up-regulated DEPs derived from 11S globulin subunit beta-like proteins may participate in the establishment of seed vigor in tobacco.

Table 1. Details of peptides derived from the precursor proteins involved in stress responses in tobacco.

GO term	Protein_id	Peptide	log2 (fold change)	P Value	Description
GO:0009628 response to abiotic stimulus	Nitab4.5_0000062g0530	T.GVPISFL.L	1.91	0.0001	XP_009592595.1 PREDICTED: DNA polymerase delta catalytic subunit
	Nitab4.5_0000105g0110	P.PGTTTGT.G	-2.23	0.0160	XP_009604364.1 PREDICTED: 11 kDa late embryogenesis abundant protein-like
	Nitab4.5_0000105g0110	M(+42.01)QAAKEKA.A	-3.07	0.0184	XP_009604364.1 PREDICTED: 11 kDa late embryogenesis abundant protein-like
	Nitab4.5_0000105g0110	M(+42.01)QAAKEKAANVAA.S	-1.00	0.0417	XP_009604364.1 PREDICTED: 11 kDa late embryogenesis abundant protein-like
	Nitab4.5_0000204g0010	M.PGLTIGDDLPNLEVETTHGKM(+15.99).K	4.60	0.0012	XP_016505780.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000204g0010	M.PGLTIGDDLPNLEVETTHGKM.K	1.58	0.0020	XP_016505780.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000204g0010	P.GEPVVISPSVSN.E	0.93	0.0074	XP_016505780.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000204g0010	M.PGLTIGDDLPNLEVETTHGKMKLH.D	0.29	0.0137	XP_016505780.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000204g0010	T.ANLPSRKDYL.H	2.81	0.0227	XP_016505780.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000204g0010	N.MVDPDETDSSGQKVPSRA.L	-1.26	0.0281	XP_016505780.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000204g0010	Q.GYDTANLPSRKDYL.H	1.54	0.0497	XP_016505780.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000723g0150	M.S(+42.01)HYDNQFSAG.Q	-1.87	0.0098	XP_009624592.1 PREDICTED: dehydrin DHN1-like isoform X1
	Nitab4.5_0000723g0150	M.S(+42.01)HYDNQFSAGQA.L	-3.93	0.0337	XP_009624592.1 PREDICTED: dehydrin DHN1-like isoform X1
	Nitab4.5_0000781g0010	T.ANLPSGKDYLRFTNV	3.99	0.0002	XP_016490368.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000781g0010	Q.GYDTANLPSGKDYLRFTNV	2.29	0.0371	XP_016490368.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000781g0010	T.ANLPSGKDYL.R	4.21	0.0422	XP_016490368.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0003883g0030	M.S(+42.01)HYENQYSAGQA.L	-3.73	0.0142	XP_009773655.1 PREDICTED: late embryogenesis abundant protein-like
	Nitab4.5_0004359g0040	T.RLTADQKTGIK.K	1.17	0.0260	XP_009781784.1 PREDICTED: dnaJ homolog subfamily B member 13-like

Continue...

Table 1. Continuation.

GO term	Protein_id	Peptide	log2 (fold change)	P Value	Description
GO:0009628 response to abiotic stimulus	Nitab4.5_0005796g0030	M.A(+42.01)DLRDEHGNPIQ.L	-5.37	0.0089	XP_009765052.1 PREDICTED: late embryogenesis abundant protein
	Nitab4.5_0006582g0030	M.A(+42.01)KRLLPTLN.R	-0.64	0.0080	XP_009799257.1 PREDICTED: 10 kDa chaperonin-like
	Nitab4.5_0006582g0030	Y.RDEDILGTLHD	-0.29	0.0196	XP_009799257.1 PREDICTED: 10 kDa chaperonin-like
	Nitab4.5_0006582g0030	M.A(+42.01)KRLLPTLNRVLIE.K	0.53	0.0311	XP_009799257.1 PREDICTED: 10 kDa chaperonin-like
	Nitab4.5_0006960g0010	M(+42.01)DPYKYRPS.S	0.71	0.0052	XP_009794245.1 PREDICTED: catalase isozyme 1
	Nitab4.5_0007799g0080	K.RDIPEEKYSN.A	1.49	0.0078	XP_009758705.1 PREDICTED: putative lactoylglutathione lyase
	Nitab4.5_0007799g0080	A.FVKDPDGY.L	1.91	0.0107	XP_009758705.1 PREDICTED: putative lactoylglutathione lyase
	Nitab4.5_0007799g0080	L.LRKRDIPEEKYSN.A	0.43	0.0246	XP_009758705.1 PREDICTED: putative lactoylglutathione lyase
	Nitab4.5_0007799g0080	K.RDIPEEKY.S	2.60	0.0331	XP_009758705.1 PREDICTED: putative lactoylglutathione lyase
	Nitab4.5_0007799g0080	F.VKDPDGYL.F	1.81	0.0461	XP_009758705.1 PREDICTED: putative lactoylglutathione lyase
	Nitab4.5_0008428g0010	L.GGYPPSHGHI	-1.97	0.0103	XP_009601197.1 PREDICTED: 18 kDa seed maturation protein-like
	Nitab4.5_0010922g0010	M.A(+42.01)TIKIVKA.R	2.44	0.0271	XP_009764205.1 PREDICTED: enolase-like
GO:0009415 response to water	Nitab4.5_0000105g0110	P.PGTTTGT.G	-2.23	0.0160	XP_009604364.1 PREDICTED: 11 kDa late embryogenesis abundant protein-like
	Nitab4.5_0000105g0110	M(+42.01)QAAKEKA.A	-3.07	0.0184	XP_009604364.1 PREDICTED: 11 kDa late embryogenesis abundant protein-like
	Nitab4.5_0000105g0110	M(+42.01)QAAKEKAANVAA.S	-1.00	0.0417	XP_009604364.1 PREDICTED: 11 kDa late embryogenesis abundant protein-like
	Nitab4.5_0000723g0150	M.S(+42.01)HYDNQFSAG.Q	-1.87	0.0098	XP_009624592.1 PREDICTED: dehydrin DHN1-like isoform X1
	Nitab4.5_0000723g0150	M.S(+42.01)HYDNQFSAGQA.L	-3.93	0.0337	XP_009624592.1 PREDICTED: dehydrin DHN1-like isoform X1
	Nitab4.5_0000781g0010	T.ANLPSGKDYLRFTNV	3.99	0.0002	XP_016490368.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000781g0010	Q.GYDTANLPSGKDYLRFTNV	2.29	0.0371	XP_016490368.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0000781g0010	T.ANLPSGKDYL.R	4.21	0.0422	XP_016490368.1 PREDICTED: 1-Cys peroxiredoxin-like
	Nitab4.5_0003883g0030	M.S(+42.01)HYENQYSAGQA.L	-3.73	0.0142	XP_009773655.1 PREDICTED: late embryogenesis abundant protein-like
	Nitab4.5_0005796g0030	M.A(+42.01)DLRDEHGNPIQ.L	-5.37	0.0089	XP_009765052.1 PREDICTED: late embryogenesis abundant protein
	Nitab4.5_0008428g0010	L.GGYPPSHGHI	-1.97	0.0103	XP_009601197.1 PREDICTED: 18 kDa seed maturation protein-like

Table 2. Details of peptides derived from the precursor proteins involved in protein processing in endoplasmic reticulum in tobacco.

KEGG Pathway	Protein_id	Peptide	log2 (fold change)	P Value	Description
Protein processing in endoplasmic reticulum	Nitab4.5_0000286g0040	M.GIDAPLFHTL.Q	-1.75	0.0267	XP_009626559.1 PREDICTED: 17.3 kDa class II heat shock protein-like
	Nitab4.5_0000286g0050	K.LPPPEPKKPKTIE.I	0.66	0.0011	TKY58025.1 17.9 kDa class II heat shock protein
	Nitab4.5_0000286g0050	M(+15.99)(+42.01)DLALRDL.G	2.56	0.0427	TKY58025.1 17.9 kDa class II heat shock protein
	Nitab4.5_0000286g0050	K.RPPPEPKKPKVIE.V	-0.40	0.0459	TKY58025.1 17.9 kDa class II heat shock protein

Continue...

Table 2. Continuation.

KEGG Pathway	Protein_id	Peptide	log2 (fold change)	P Value	Description
Protein processing in endoplasmic reticulum	Nitab4.5_0000286g0060	M.GIDTPLLHTLQ.H	-0.99	0.0026	XP_016472899.1 PREDICTED: 17.3 kDa class II heat shock protein-like isoform X1
	Nitab4.5_0000286g0060	M.GIDTPLLH.T	-1.25	0.0354	XP_016472899.1 PREDICTED: 17.3 kDa class II heat shock protein-like isoform X1
	Nitab4.5_0000433g0150	C.GLEHGVLTVNVPKKETQEVPR-NVR.A	3.30	0.0039	XP_016488118.1 PREDICTED: 16.9 kDa class I heat shock protein 1-like
	Nitab4.5_0002443g0040	M.DVPGIKKEEV.K	2.07	0.0444	XP_009771699.1 PREDICTED: 17.6 kDa class I heat shock protein 3-like
	Nitab4.5_0004078g0050	M.A(+42.01)GKGEGPAIG.I	4.81	0.0018	XP_009777579.1 PREDICTED: heat shock cognate 70 kDa protein 2-like
	Nitab4.5_0004078g0050	M.A(+42.01)GKGEGPAIGIDLGT.T	4.08	0.0062	XP_009777579.1 PREDICTED: heat shock cognate 70 kDa protein 2-like
	Nitab4.5_0004078g0050	M.A(+42.01)GKGEGPAIGID.L	2.36	0.0139	XP_009777579.1 PREDICTED: heat shock cognate 70 kDa protein 2-like
Protein processing in endoplasmic reticulum	Nitab4.5_0004541g0060	C.GLEHGVLTV.N	3.02	0.0401	XP_016515155.1 PREDICTED: 16.9 kDa class I heat shock protein 1-like
	Nitab4.5_0005771g0020	A.VNPERTIFD.V	4.80	0.0152	XP_009592769.1 PREDICTED: luminal-binding protein 5
	Nitab4.5_0008173g0030	E.ITDPDAKKPEDWD.D	-2.03	0.0016	XP_009804593.1 PREDICTED: calreticulin isoform X1
	Nitab4.5_0008173g0030	E.ITDPDAKKP.E	-4.08	0.0264	XP_009804593.1 PREDICTED: calreticulin isoform X1
	Nitab4.5_0008173g0030	F.IDDPEDKKP.E	-4.36	0.0283	XP_009804593.1 PREDICTED: calreticulin isoform X1
	Nitab4.5_0011739g0010	K.LPPPEPKKPKTIE.V	0.66	0.0011	XP_009786753.1 PREDICTED: 17.3 kDa class II heat shock protein-like
	Nitab4.5_0011739g0010	K.LPPPEPKKP.K	-0.72	0.0013	XP_009786753.1 PREDICTED: 17.3 kDa class II heat shock protein-like
	Nitab4.5_0011739g0010	K.LPPPEPKKPKT.I	-2.81	0.0043	XP_009786753.1 PREDICTED: 17.3 kDa class II heat shock protein-like
	Nitab4.5_0011739g0010	M.GIDTPLFHTIH.H	-0.70	0.0263	XP_009786753.1 PREDICTED: 17.3 kDa class II heat shock protein-like
	Nitab4.5_0011739g0010	M.GIDTPLF.H	-1.57	0.0307	XP_009786753.1 PREDICTED: 17.3 kDa class II heat shock protein-like
	Nitab4.5_0011739g0010	Q.KLPPPEPKKP.K	-3.67	0.0338	XP_009786753.1 PREDICTED: 17.3 kDa class II heat shock protein-like

Table 3. Details of peptides derived from the precursor proteins globulin involved in seed vigor in tobacco.

Peptide	Protein_id	log2 (fold change)	P Value	Description
S.LSLPILNFL.Q	Nitab4.5_0000044g0350	2.77	0.0013	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.PILNFLQLSAERGTLYRNA.I	Nitab4.5_0000044g0350	3.57	0.0029	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.QLSAERGTLYRN.A	Nitab4.5_0000044g0350	2.55	0.0116	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
N.QLLIVPQ.N	Nitab4.5_0000044g0350	-1.38	0.0143	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
V.NSLSLPIL.N	Nitab4.5_0000044g0350	1.43	0.0155	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
S.ALRAMPEEVLMNSYQISRQEARSLKYN.R	Nitab4.5_0000044g0350	2.03	0.0163	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
N.SLSLPILNFL.Q	Nitab4.5_0000044g0350	2.65	0.0172	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like

Continue...

Table 3. Continuation.

Peptide	Protein_id	log2 (fold change)	P Value	Description
F.LQLSAERGTLYRNA.I	Nitab4.5_0000044g0350	1.94	0.0200	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.QLSAERGTLY.R	Nitab4.5_0000044g0350	1.59	0.0305	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.QLSAERGTLYRNA.I	Nitab4.5_0000044g0350	1.90	0.0323	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
N.PRGGRVSTVNSL.S	Nitab4.5_0000044g0350	2.49	0.0329	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.KLRENIGHPSRSD.V	Nitab4.5_0000044g0350	2.23	0.0340	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.RAMPEEVLMN.S	Nitab4.5_0000044g0350	1.01	0.0345	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
A.IVAPHWNMNAH.S	Nitab4.5_0000044g0350	2.35	0.0367	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.KLRENIGHPSRSDVYN.P	Nitab4.5_0000044g0350	3.12	0.0370	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
F.KTNDNAMVNPL.A	Nitab4.5_0000044g0350	1.68	0.0393	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.RAMPEEVLMNSYQISRQEARSLKYN.R	Nitab4.5_0000044g0350	1.90	0.0429	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
Y.TNTPMLM.Y	Nitab4.5_0000044g0350	1.38	0.0435	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
R.AMPEEVLMNSYQISRQEARSLKYN.R	Nitab4.5_0000044g0350	1.46	0.0458	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
S.RQEARSLKYN.R	Nitab4.5_0000044g0350	1.35	0.0485	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
T.NDNAMVNPL.A	Nitab4.5_0000044g0350	1.53	0.0486	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
L.LIVPQNFAIVK.R	Nitab4.5_0000044g0350	1.22	0.0488	XP_009603582.1 PREDICTED: 11S globulin subunit beta-like
M.RLKENLGQPSRAD.V	Nitab4.5_0000044g0370	1.07	0.0048	XP_016439381.1 PREDICTED: 11S globulin subunit beta-like
T.LPVLNFL.Q	Nitab4.5_0000436g0260	1.83	0.0056	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
N.RDELNVFGPGAKSGRQQY.A	Nitab4.5_0000436g0260	1.53	0.0068	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
F.FLAGNPQRGVQQQM.L	Nitab4.5_0000436g0260	-2.07	0.0069	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
R.GEQERGGNILSGFDAQLL.S	Nitab4.5_0000436g0260	-1.90	0.0072	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
N.GLTLPVLNFL.Q	Nitab4.5_0000436g0260	2.60	0.0082	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
K.YNRDELNVFGPGAKSGRQQY.A	Nitab4.5_0000436g0260	2.43	0.0117	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
Y.NRDELNVFGPGAKSGRQQY.A	Nitab4.5_0000436g0260	1.31	0.0133	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
F.NVDPEIIRR.L	Nitab4.5_0000436g0260	1.30	0.0133	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
S.RQEAKSLKYNRDELNVFGPGAKSGRQQY.A	Nitab4.5_0000436g0260	3.03	0.0138	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
L.DLTFRKF.F	Nitab4.5_0000436g0260	1.70	0.0143	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like

Continue...

Table 3. Continuation.

Peptide	Protein_id	log2 (fold change)	P Value	Description
F.RKFFLAGNPQRGVQQQ.M	Nitab4.5_0000436g0260	-2.75	0.0166	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
F.LAGNPQRGVQQQM.L	Nitab4.5_0000436g0260	-2.69	0.0185	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
K.YNRDELNVFGPGAKSGRQQ.Y	Nitab4.5_0000436g0260	1.60	0.0209	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
N.VDPEIIRR.L	Nitab4.5_0000436g0260	-0.67	0.0226	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
L.IDTSNQANQLDLTFRKFFLAGNPQRGVQQQ.M	Nitab4.5_0000436g0260	0.76	0.0309	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
G.EQERGGNILSGFDAQLL.S	Nitab4.5_0000436g0260	-1.55	0.0340	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
L.IDTSNQANQLDLTFRKFF.L	Nitab4.5_0000436g0260	-1.24	0.0348	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
A.IVAPHWNMNAHA.V	Nitab4.5_0000436g0260	-1.39	0.0373	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
L.IDTSNQANQLDLTFRKFFLA.G	Nitab4.5_0000436g0260	-0.94	0.0414	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
L.IDTSNQANQLDLTFRKFFLAGNPQRGVQQQM.L	Nitab4.5_0000436g0260	-0.66	0.0499	XP_016490392.1 PREDICTED: 11S globulin subunit beta-like
F.YLAGGVPESGRQQTQ.E	Nitab4.5_0000470g0070	-3.45	0.0113	XP_016454980.1 PREDICTED: 11S globulin seed storage protein 2-like
A.FYLAGGVPESGRQQTQ.E	Nitab4.5_0000470g0070	-3.43	0.0148	XP_016454980.1 PREDICTED: 11S globulin seed storage protein 2-like
F.YLAGGVPESGRQQ.T	Nitab4.5_0001855g0030	-2.48	0.0009	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
F.YLAGGVPESGRQ.Q	Nitab4.5_0001855g0030	-2.46	0.0020	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
R.AFYLAGGVPESGR.Q	Nitab4.5_0001855g0030	-3.24	0.0022	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
T.NRGGESFLLSPQ.R	Nitab4.5_0001855g0030	1.00	0.0043	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
N.RGGESFLLSPQ.R	Nitab4.5_0001855g0030	1.15	0.0045	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
Q.SRQRFQNIFRAF.D	Nitab4.5_0001855g0030	-2.32	0.0053	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
F.YLAGGVPESGR.Q	Nitab4.5_0001855g0030	-2.02	0.0059	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
F.YLAGGVPESGRQQT.Q	Nitab4.5_0001855g0030	-2.47	0.0066	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
T.QADIFSRQA.G	Nitab4.5_0001855g0030	2.32	0.0067	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
N.AYQISPNEAQRLKTNRGGESFLLSPQ.R	Nitab4.5_0001855g0030	-1.39	0.0090	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
F.YLAGGVPESGRQQTQ.A	Nitab4.5_0001855g0030	-3.45	0.0113	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
L.AGYTSVIRAMPVEVL.S	Nitab4.5_0001855g0030	2.20	0.0122	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
Y.TSVIRAMPVEVL.S	Nitab4.5_0001855g0030	1.75	0.0129	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
Q.SRQRFQNIF.R	Nitab4.5_0001855g0030	-1.89	0.0131	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like

Continue...

Table 3. Continuation.

Peptide	Protein_id	log2 (fold change)	P Value	Description
A.FYLAGGVPESGRQQTQ.A	Nitab4.5_0001855g0030	-3.43	0.0148	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
F.YLAGGVPESG.R	Nitab4.5_0001855g0030	-1.76	0.0185	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
E.EEGEFEEEQGRPRRGQQW.W	Nitab4.5_0001855g0030	0.80	0.0212	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
N.RGGESFLLSPQRRS.F	Nitab4.5_0001855g0030	0.83	0.0219	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
F.YLAGGVPESGRQQTQAGQRLQ.S	Nitab4.5_0001855g0030	-2.12	0.0223	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
N.DLNHRSNQLDQNLRAF.Y	Nitab4.5_0001855g0030	-2.79	0.0337	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
A.FYLAGGVPESGRQQ.T	Nitab4.5_0001855g0030	-1.29	0.0403	XP_016503441.1 PREDICTED: 11S globulin seed storage protein 2-like
M.NRGGESFLLSPQ.R	Nitab4.5_0001855g0060	1.00	0.0043	XP_009761684.1 PREDICTED: 11S globulin seed storage protein 2-like
K.MNRGGESFLLSPQ.R	Nitab4.5_0001855g0060	-0.93	0.0104	XP_009761684.1 PREDICTED: 11S globulin seed storage protein 2-like
L.AGYTSVIRAMPVEVL.T	Nitab4.5_0001855g0060	2.20	0.0122	XP_009761684.1 PREDICTED: 11S globulin seed storage protein 2-like
Y.TSVIRAMPVEVL.T	Nitab4.5_0001855g0060	1.75	0.0129	XP_009761684.1 PREDICTED: 11S globulin seed storage protein 2-like
Y.MDMSASRGTLYPN.A	Nitab4.5_0001855g0060	0.58	0.0161	XP_009761684.1 PREDICTED: 11S globulin seed storage protein 2-like
N.RGGESFLLSPQRRS.I	Nitab4.5_0001855g0060	0.83	0.0219	XP_009761684.1 PREDICTED: 11S globulin seed storage protein 2-like
N.REEATVFAGRRSGGYSTRA.F	Nitab4.5_0004662g0050	2.02	0.0024	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
Y.NNAPQLL.Y	Nitab4.5_0004662g0050	2.14	0.0024	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
M.RLGENLGRPSRAD.V	Nitab4.5_0004662g0050	2.30	0.0054	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
K.YNREEATVFAGRRSGGYSTRA.F	Nitab4.5_0004662g0050	2.22	0.0062	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
A.ERGVLYQ.A	Nitab4.5_0004662g0050	2.82	0.0077	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
N.REEATVFAGRRSGGYST.R	Nitab4.5_0004662g0050	1.51	0.0113	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
N.REEATVFAGRRSGGY.S	Nitab4.5_0004662g0050	3.18	0.0126	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
N.AVMAPYWNMN.A	Nitab4.5_0004662g0050	1.78	0.0130	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
L.NSHKLPILNWL.Q	Nitab4.5_0004662g0050	4.85	0.0145	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
F.KTNDQAMTSPL.A	Nitab4.5_0004662g0050	2.41	0.0151	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
M.RLGENLGRPS.R	Nitab4.5_0004662g0050	1.99	0.0171	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
L.KYNREEATVFAGRRSGGYSTRA.F	Nitab4.5_0004662g0050	1.51	0.0182	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
N.REEATVFAGRRSGGYS.T	Nitab4.5_0004662g0050	1.95	0.0200	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like

Continue...

Table 3. Continuation.

Peptide	Protein_id	log2 (fold change)	P Value	Description
Y.NREEATVFAGRRSGGY.S	Nitab4.5_0004662g0050	3.55	0.0202	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
K.YNREEATVFAGRRSGGYS.T	Nitab4.5_0004662g0050	2.12	0.0247	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
L.SAIRAMPEEVLMNSYQIS.R	Nitab4.5_0004662g0050	2.20	0.0306	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
L.SAIRAMPEEVLMNSY.Q	Nitab4.5_0004662g0050	1.77	0.0342	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
I.RAMPEEVLMN.S	Nitab4.5_0004662g0050	1.01	0.0345	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
Y.NREEATVFAGRRSGGYS.T	Nitab4.5_0004662g0050	2.14	0.0376	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
Y.NREEATVFAGRRSGGYST.R	Nitab4.5_0004662g0050	2.18	0.0411	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
L.KYNREEATVFAGRRSGGY.S	Nitab4.5_0004662g0050	2.99	0.0429	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
I.RAMPEEVLMNSYQISRQEARSLKYN.R	Nitab4.5_0004662g0050	1.90	0.0429	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
Y.NREEATVFAGRRSGGYSTRA.F	Nitab4.5_0004662g0050	2.23	0.0430	XP_016489177.1 PREDICTED: 11S globulin subunit beta-like
D.EEEGEFEEEQGRSRRGQQW.W	Nitab4.5_0006457g0010	0.30	0.0008	XP_009766463.1 PREDICTED: 11S globulin seed storage protein 2-like

DISCUSSION

Orthodox seeds acquire desiccation tolerance at the maturation stage, and a large set of genes are involved in seed desiccation tolerance (Kamble and Majee, 2022). In this study, we observed that desiccation phase begins at the period from 20 DAP in tobacco. Proteins associated with seed vigor have been extensively studied by comparative proteomics methods in plants (Wang et al., 2015). There is substantial evidence that peptides play critical roles in plant growth and stress responses (Matsubayashi, 2014; Takahashi et al., 2018; Zeng et al., 2022). Therefore, a comparative peptidomics analysis was conducted to identify the peptides that related with the establishment of tobacco seed vigor in this study. Proteolytic processing from precursor proteins is an essential source of peptides (Chen et al., 2020). Our data showed that many DEPs were derived from the precursor proteins, such as LEA protein, heat shock protein, peroxiredoxin, and globulin protein, involving stress responses in tobacco. Our identified peptides may play important roles on the establishment of seed vigor in tobacco during seed development.

The analysis of precursor proteins can elucidate the potential functions of DEPs. In this study, we observed that the functions of precursor proteins were mainly associated with stress responses. Desiccation tolerance is a key issue for the establishment of seed vigor in plants. Several protective mechanisms have been proposed for seed desiccation tolerance, including the metabolic 'switch off', structural stabilization, accumulation of protective molecules and removal of reactive oxygen species (ROS) (Wang et al., 2015). LEA proteins are well characterized as the protective molecules against desiccation stress through replacing water, sequestering ions, and removing ROS and so on (Battaglia et al., 2008; Wang et al., 2015). In this study, we observed that the DEPs derived LEA proteins were observed at the initial desiccation phase (20 DAP) during seed development in tobacco. Further, many DEPs in the developing seed were derived from the peroxiredoxin precursor proteins and the precursor protein globulin in tobacco. Globulin is involved in the regulation of seed vigor by influencing hydrogen peroxide (H_2O_2) levels in rice (Peng et al., 2022). Seed dehydration will lead to the production of ROS during the mature stage (Bailly, 2004; Kranner and Birtic, 2005; Berjak

and Pammenter, 2008). We thus propose that the potential functions of DEPs derived from the LEA, peroxiredoxin, and globulin proteins may promote desiccation tolerance by increasing the ability to remove ROS for the establishment of seed vigor in tobacco.

Plant peptides can be mainly divided into the precursor-derived and nonprecursor-derived peptides (Tavormina et al., 2015). Most of the plant peptides already studied are derived from precursor proteins (Brito et al., 2018). Similarly, we observed that the most of peptides were derived from the precursor proteins in tobacco in this study. Meanwhile, peptides can also be directly translated from their corresponding transcripts, such as the nonprecursor-derived group of small open reading frames (sORF) (Tavormina et al., 2015). Several studies have confirmed that some lncRNAs have sORF coding the short peptides with key biological functions (Choi et al., 2019). The peptides belong to the nonprecursor-derived sORF group that involved in the establishment of tobacco seed vigor need to be further investigated in the future. Moreover, peptides are involved in signaling and cell-cell communication during seed development (Qu et al., 2015). It has been reported that the storage proteins in the starchy endosperm are hydrolyzed into a mixture of small peptides and free amino acids for seed germination (Salmenkallio and Sopanen, 1989). It would be interesting to assess whether the peptides act as signals involved in the establishment of seed vigor during seed development in tobacco, and to reveal whether the peptides involved in the regulation of seed germination.

CONCLUSIONS

Peptidomics analysis was used to identify the peptides involved in the establishment of seed vigor during seed development of tobacco in this study. Many candidate peptides derived from stress response-related precursor proteins, such as LEA protein, heat shock protein, peroxiredoxin, and globulin proteins, may involve in the response of desiccation tolerance to regulate the establishment of seed vigor in tobacco. These findings could make a significant contribution to understand the biological functions of peptides in the establishment of seed vigor in tobacco.

ACKNOWLEDGEMENTS

The research was supported by the Science and Technology Project of China National Tobacco Corporation Yunnan Company (No. 2021530000242033; 2023530000241007).

REFERENCES

AGYEI, D.; TSOPMO, A.; UDENIGWE C.C. Bioinformatics and peptidomics approaches to the discovery and analysis of food-derived bioactive peptides. *Analytical and Bioanalytical Chemistry*, v.410, n.15, p.3463–3472, 2018. https://doi.org/10.1007/s00216-018-0974-1

BAILLY, C. Active oxygen species and antioxidants in seed biology. *Seed Science Research*, v.14, n.2, p.93–107, 2004. https://doi.org/10.1079/SSR2004159

BATTAGLIA, M.; OLVERA-CARRILLO, Y.; GARCIARRUBIO, A.; CAMPOS, F.; COVARRUBIAS, A.A. The enigmatic LEA proteins and other hydrophilins. *Plant Physiology*, v.148, n.1, p.6–24, 2008. https://doi.org/10.1104/pp.108.120725

BERJAK, P.; PAMMENTER, N.W. From *Avicennia* to *Zizania*: seed recalcitrance in perspective. *Annals of Botany*, v.101, n.2, p.213–228, 2008. https://doi.org/10.1093/aob/mcm168

BRITO, M.S.; DEPAOLI, H.C.; COSSALTER, V.; AVANCI, N.C.; FERREIRA, P.B.; AZEVEDO, M.S.; STRINI, E.J.; QUIAPIM, A.C.; GOLDMAN, G.H.; PERES, L.E.P.; GOLDMAN, M.H.S. A novel cysteine-rich peptide regulates cell expansion in the tobacco pistil and influences its final size. *Plant Science*, v.277, p.55–67, 2018. https://doi.org/10.1016/j.plantsci.2018.09.002

CHEN, Y.L.; FAN, K.T.; HUNG, S.C.; CHEN, Y.R. The role of peptides cleaved from protein precursors in eliciting plant stress reactions. *New Phytologist*, v.225, n.6, p.2267–2282, 2020. https://doi.org/10.1111/nph.16241

CHIEN, P.S.; NAM, H.G.; CHEN, Y.R. A salt-regulated peptide derived from the CAP superfamily protein negatively regulates salt-stress tolerance in Arabidopsis. *Journal of Experimental Botany*, v.66, n.17, p.5301–5313, 2015. https://doi.org/10.1093/jxb/erv263

CHOI, S.W.; KIM, H.W.; NAM, J.W. The small peptide world in long noncoding RNAs. *Briefings in Bioinformatics*, v.20, n.5, p.1853–1864, 2019. https://doi.org/10.1093/bib/bby055

CHRISTMANN, A.; GRILL, E. Peptide signal alerts plants to drought. *Nature*, v.556, n.7700, p.178–179, 2018. https://doi.org/10.1038/d41586-018-03872-4

CUI, Y.; LI, M.; YIN, X.; SONG, S.; XU, G.; WANG, M.; LI, C.; PENG, C.; XIA, X. OsDSSR1, a novel small peptide, enhances drought tolerance in transgenic rice. *Plant Science*, v.270, p.85–96, 2018. https://doi.org/10.1016/j.plantsci.2018.02.015

GE, X.; CHANG, F.; MA, H. Signaling and transcriptional control of reproductive development in *Arabidopsis*. *Current Biology*, v.20, n.22, p.R988–R997, 2010. https://doi.org/10.1016/j.cub.2010.09.040

KAMBLE, N.U.; MAJEE, M. ABI transcription factors and PROTEIN L-ISOASPARTYL METHYLTRANSFERASE module mediate seed desiccation tolerance and longevity in *Oryza sativa*. *Development*, v.149, n.11, dev200600, 2022. https://doi.org/10.1242/dev.200600

KIJAK, H.; RATAJCZAK, E. What do we know about the genetic basis of seed desiccation tolerance and longevity? *International Journal of Molecular Sciences*, v.21, n.10, p.3612, 2020. https://doi.org/10.3390/ijms21103612

KIM, Y.J.; CHAMBERS, A.G.; CECCHI, F.; HEMBROUGH, T. Targeted data independent acquisition for mass spectrometric detection of RAS mutations in formalin-fixed, paraffin-embedded tumor biopsies. *Journal of Proteomics*, v.189, p.91–96, 2018. https://doi.org/10.1016/j.jprot.2018.04.022

KRANNER, I.; BIRTIC, S. A modulating role for antioxidants in desiccation tolerance. *Integrative and Comparative Biology*, v.45, n.5, p.734–740, 2005. https://doi.org/10.1093/icb/45.5.734

LEPRINCE, O.; PELLIZZARO, A.; BERRIRI, S.; BUITINK, J. Late seed maturation: Drying without dying. *Journal of Experimental Botany*, v.68, n.4, p.827–841, 2017. https://doi.org/10.1093/jxb/erw363

LI, X.; HAN, H.; CHEN, M.; YANG, W.; LIU, L.; LI, N.; DING, X.; CHU, Z. Overexpression of *OsDT11*, which encodes a novel cysteine-rich peptide, enhances drought tolerance and increases ABA concentration in rice. *Plant Molecular Biology*, v.93, n.1-2, p.21–34, 2017. https://doi.org/10.1007/s11103-016-0544-x

MA, W.; ZHANG, C.; ZHANG, W.; SHENG, P.; XU, M.; NI, Y.; CHEN, M.; CHENG, B., ZHANG, X. TMT-based comparative peptidomics analysis of rice seedlings under salt stress: an accessible method to explore plant stress-tolerance processing. *Journal of Proteome Research*, v.21, n.12, p.2905–2919, 2022. https://doi.org/10.1021/acs.jproteome.2c00318

MATSUBAYASHI, Y. Posttranslationally modified small-peptide signals in plants. *Annual Review of Plant Biology*, v.65, p.385–413, 2014. https://doi.org/10.1146/annurev-arplant-050312-120122

NAKAMINAMI, K.; OKAMOTO, M.; HIGUCHI-TAKEUCHI, M.; YOSHIZUMI, T.; YAMAGUCHI, Y.; FUKAO, Y.; SHIMIZU, M.; OHASHI, C.; TANAKA, M.; MATSUI, M.; SHINOZAKI, K.; SEKI, M.; HANADA, K. AtPep3 is a hormone-like peptide that plays a role in the salinity stress tolerance of plants. *Proceedings of the National Academy of Sciences of the United States of America*, v.115, n.22, p.5810–5815, 2018. https://doi.org/10.1073/pnas.1719491115

NIU, Y.; WANG, C.; SUO, W.; WANG, G.; ZHAO, J.; WANG, Z.; ZHENG, Y. Isopropylmalate synthase *NtIPMS* as a potential molecular marker for seed vigor in tobacco. *Plant Biotechnology*, v.40, n.1, p.43–49, 2023. https://doi.org/10.5511/plantbiotechnology.23.0118a

OKUDA, S.; YAMADA, T.; HAMAJIMA, M.; ITPH, M.; KATAYAMA, T.; BORK, P.; GOTO, S.; KANEHISA, M. KEGG Atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Research*, v.36, p.W423–W426, 2008. https://doi.org/10.1093/nar/gkn282

PAN, W.; MA, W.G.; ZHENG, J.Y.; GENG, S.B. Chlorophyll fluorescence of tobacco seeds as marker of seed maturity. *Southwest China Journal of Agricultural Sciences*, v.29, n.4, p.966–969, 2016. https://doi.org/10.16213/j.cnki.scjas.2016.04.042

PENG, L.; LU, H.; CHEN, J.; WU, Z.; XIAO, Z.; QING, X.; SONG, J.; WANG, Z.; ZHAO, J. Characteristics of seed vigor in rice varieties with different globulin accumulations. *International Journal of Molecular Sciences*, v.23, n.17, p.9717, 2022. https://doi.org/10.3390/ijms23179717

PENG, L.; SUN, S.; YANG, B.; ZHAO, J.; LI, W.; HUANG, Z.; LI, Z.; HE, Y.; WANG, Z. Genome-wide association study reveals that the cupin domain protein OsCDP3.10 regulates seed vigour in rice. *Plant Biotechnology Journal*, v.20, n.3, p.485–498, 2022. https://doi.org/10.1111/pbi.13731

QU, L.J.; LI, L.; LAN, Z.; DRESSElHAUS, T. Peptide signalling during the pollen tube journey and double fertilization. *Journal of Experimental Botany*, v.66, n.17, p.5139–5150, 2015. https://doi.org/10.1093/jxb/erv275

SALMENKALLIO, M.; SOPANEN, T. Amino acid and peptide uptake in the scutella of germinating grains of barley, wheat, rice, and maize. *Plant Physiology*, v.89, n.4, p.1285–1291, 1989. https://doi.org/10.1104/pp.89.4.1285

TAKAHASHI, F.; SUZUKI, T.; OSAKABE, Y.; BETSUYAKU, S.; KONDO, Y.; DOHMAE, N.; FUKUDA, H.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. A small peptide modulates stomatal control via abscisic acid in long-distance signalling. *Nature*, v.556, n.7700, p.235–238, 2018. https://doi.org/10.1038/s41586-018-0009-2

TAVORMINA, P.; DE CONINCK, B.; NIKONOROVA, N.; DE SMET, I.; CAMMUE, B.P. The plant peptidome: an expanding repertoire of structural features and biological functions. *Plant Cell*, v.27, n.8, p.2095–2118, 2015. https://doi.org/10.1105/tpc.15.00440

TRINDADE, F.; BARROS, A.S.; SILVA, J.; VLAHOU, A.; FALCAO-PIRES, I.; GUEDES, S.; VITORINO, C.; FERREIRA, R.; LEITE-MOREIRA, A.; AMADO, F.; VITORINO, R. Mining the biomarker potential of the urine peptidome: from amino acids properties to proteases. *International Journal of Molecular Sciences*, v.22, n.11, p.5940, 2021. https://doi.org/10.3390/ijms22115940

WANG, W.Q.; LIU, S.J.; SONG, S.Q.; MOLLER, I.M. Proteomics of seed development, desiccation tolerance, germination and vigor. *Plant Physiology and Biochemistry*, v.86, p.1–15, 2015. https://doi.org/10.1016/j.plaphy.2014.11.003

WEI, X.; KIM, W.S.; SONG, B.; OEHRLE, N.W.; LIU, S.; KRISHNAN H.B. Soybean mutants lacking abundant seed storage proteins are impaired in mobilization of storage reserves and germination. *ACS Omega*, v.5, n.14, p.8065–8075, 2020. https://doi.org/10.1021/acsomega.0c00128

WUEST, S.E.; VIJVERBERG, K.; SCHMIDT, A.; WEISS, M.; GHEYSELINCK, J.; LOHR, M.; WELLMER, F.; RAHNENFUHRER, J.; VON MERING, C.; GROSSNIKLAUS, U. *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. *Current Biology*, v.20, n.6, p.506–512, 2010. https://doi.org/10.1016/j.cub.2010.01.051

XIANG, D.; MENG, F.; WANG, A.; WU, Y.; WANG, Z.; ZHENG, S.; MAO, C. Root-secreted peptide OsPEP1 regulates primary root elongation in rice. *Plant Journal*, v.107, n.2, p.480–492, 2021. https://doi.org/10.1111/tpj.15303

ZENG, Y.; TANG, Y.; SHEN, S.; ZHANG, M.; CHEN, L.; YE, D.; ZHANG, X. Plant-specific small peptide AtZSP1 interacts with ROCK1 to regulate organ size in Arabidopsis. *New Phytologist*, v.234, n.5, p.1696–1713, 2022. https://doi.org/10.1111/nph.18093

ZHAO, J.; HE, Y.; HUANG, S.; WANG, Z. Advances in the identification of quantitative trait loci and genes involved in seed vigor in rice. *Frontiers in Plant Science*, v.12, p.659307, 2021. https://doi.org/10.3389/fpls.2021.659307