Genetics & Applications

Institute for Genetic Engineering and Biotechnology University of Sarajevo

Research Article Open access

RNA Sequencing indicates gene expression changes in *Silene sendtneri* seeds after seed priming with silicic acid

Erna Karalija[1,2], Arnela Demir[1], Jelena Samardzic[3], Adisa Paric[1], Sabina Dahija[1], Felice Contaldi[2], Federico Martinelli[2]*

[1] Department of Biology, Faculty of Science, University of Sarajevo, Zmaja od Bosne 33-35; 71 000 Sarajevo, Bosnia and Herzegovina
[2] Department of Biology, University of Florence, Via Madonna del Piano, 6, 50019 Sesto Fiorentino, Florence, Italy
[3] Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Vojvode Stepe 444a, 11042 Belgrade, Serbia

DOI: 10.31383/ga.vol6iss2ga08

***Correspondence**
E-mail:
Federico.martinelli@unifi.it

Received
November, 2022
Accepted
December, 2022
Published
December, 2022

Abstract

To improve our understanding of the molecular mechanisms underlaying seed priming, RNA transcriptome analysis was performed using primed and non-primed seeds of *Silene sendtneri*. Seed priming was performed by submergence in 1% silicic acid for 24h at 4°C, followed by rinsing with sterilised water and desiccation to original moisture content. *Silene sendtneri* is a species with no sequenced genome and annotation of *de novo* assembly of transcriptome was done against several species. Gene ontology (GO) analysis indicated that genes related to heavy metal transporters and heat shock proteins are differentially expressed after priming with silicic acid. Within these gene categories, genes such as heavy metal-associated isoprenylated plant protein 26-like (log2fold -8.79) were downregulated, while others such as heavy metal ATPase 5 (log2fold 6.46), heat shock factor protein HSF30-like isoform X1 (log2fold 5.98) were upregulated.

Keywords

seed priming, RNA seq, heavy metals, pumilio proteins

Introduction

Silene sendtneri is an endemic plant species with high capacity to accumulate cadmium and can be classified as a Cd hyperaccumulator (Karalija et al., 2021). Heavy metal presence in the environment, especially in agricultural land, can lead to harmful effects on human health and different strategies are employed to alleviate such effects through soil remediation. Phytoremediation is one of the approaches, but recently other options are being investigated, such as changes in root absorption rate through exclusion of heavy metals. Seed priming is known as an effective method to improve plant traits including their tolerance to different biotic and abiotic stressors, but the mechanisms of this "primed" memory is still not well understood (Mladenov et al., 2020). Germination as a process, represents probably the most critical stage of plants life and consists out of three phases; phase 1 – water absorption – imbibition; phase 2 – activation of metabolic processes; phase 3 – growth processes – radicle protrusion. Seed priming exploits the processes of seed germination specifically events of phase II of germination influencing metabolic processes and gene expression through exposure to short stress (priming agent). Desiccation of the seeds post-priming must be performed before the commence of radicle growth and protrusion to preserve germination capacity of the seeds creating "primed" memory. Through desiccation process seed "memorises" the priming-induced changes enabling primed seeds to perform better under stress conditions. One of most prominent effects of seed priming is synchronisation of germination but it can affect the whole plant life cycle (Srivastava et al., 2021). From plant memory side, the most prominent events commence during signalling events in the phase II of germination process under priming conditions. Triggered signalling pathways and consequent genome-level changes are integrated into "primed" memory. Most investigated priming agents are selenium, salicylic acid, polyethylene glycol (PEG), $CaCl_2$ and thiourea (Srivastava et al., 2021). The priming process triggers changes at mRNA level as well as protein level in correspondence to specific priming agent. Several priming agents have been associated to upregulation of genes such as genes encoding different antioxidants (Paul et al., 2022), while others have been associated with downregulation of genes, such as PEG where downregulation of genes encoding antioxidants are identified in rice (Lei et al., 2021). The objective of presented study was to investigate molecular mechanisms underlying seed priming memory initiated by seed priming using silicic acid through transcriptome analysis of of primed/non-primed seeds. The study contributes to further elucidation of primed memory and identification of candidate genes responsible for primed memory and enhanced performance of primed plants.

Material and methods

Plant material

Seeds of *Silene sendtneri* were collected from natural growing population at locality Pjeskovita ravan (43.9128° N, 18.4628° E). Seeds were separated from fruits; dried and voucher specimen was deposited at Laboratory for Plant physiology, Faculty of Science, University of Sarajevo. Priming of seeds was performed using 1% silicic acid for 24 h followed by rinsing in sterile distilled water and desiccation to original water content at room temperature. Primed seeds were stored at 4 °C till use.

RNA isolation and transcriptome sequencing

Total RNA was extracted from 20 uniformed healthy primed or non-primed seeds using a RNAprep Pure Plant Plus Kit (Tiangen) according to the manufacturer instructions. The concentration of total RNA was measured using Qubit and appropriate Qubit kit. The purity and integrity of RNA was checked by Agilend Bioanalyser 2100 system. The cDNA library construction and RNA sequencing were performed in NovaGene (UK). cDNA libraries were constructed using Ilumina TruSeqTM RNA Sample Preparation Kit following manufacturers instruction. All samples were sequenced using Ilumina system HiSeq2500 following standard procedure.

Software and parameters

Analysis	Software	Version	Parameter	Remark
Assembly	Trinity	2.6.6	minKmerCov=3 min_glue=4	-
	Corset	4.6	-f ture, Default, -m 10	remove redundancy
	BUSCO	3.0.2	-m tran	-
Gene Functional Annotation (Buchfink et al., 2014; Chepelev et al., 2009; Götz et al., 2008; Kanehisa et al., 2007; Moriya et al., 2007)	Diamond	0.8.22	e-value = 1e-5	NR, KOG/COG, Swiss-Prot
	Diamond, KAAS	0.8.22	e-value = 1e-5	KEGG Annotation
	NCBI blast	2.9.0	e-value = 1e-5	NT Annotation
	hmmscan	HMMER 3.1	e-value = 0.01	Pfam Annotation
	blast2go	b2g4pipe_v2.5	e-value = 1e-6	GO Annotation
Mapping and Quantification	RSEM	1.2.28	--estimate-rspd --mismatch-rate 0.3	mapping to Corset filtered transcriptome
Mutation	SAMtools/BCFtools	1.9	bcftools view varFilter -Q 20 -d 1 -D 100	-
SSR Analysis	MISA, primer3	primer3-2.3.4	SSR: 1-10 2-6 3-5 4-5 5-5 6-5	Misa detect SSR, primer3 Primer Design
Differential Expression Analysis (Anders and Huber, 2010)	DESeq2	1.26.0	padj < 0.05 padj	For sample with bio-replicate using DESeq2, samples without bio-replicate using EdgeR.
	edgeR	3.28.0	padj < 0.005 & $\|\log_2(\text{foldchange})\| > 1$	
GO Enrichment	GOSeq, topGO	1.32.0, 2.32.0	Corrected P-Value < 0.05	-
KEGG Enrichment	KOBAS	v3.0	Corrected P-Value < 0.05	-
Protein-Protein Interaction Analysis	NCBI blast 29.0	v2.2.28+	e-value = 1e-10	Using blast, String database.

Read mapping, Gene Annotation and Analysis of Gene Expression Level

Sequencing raw reads were processed by filtering out reads with adaptor contamination, reads when uncertain nucleotides constitute more than 10 percent of either read (N > 10%), reads when low quality nucleotides (Base Quality less than 5) constitute more than 50 percent of the read.

To achieve comprehensive gene functional annotation using BLAST program (Altschul et al., 1997), seven databases were applied by Novogene: Nr (NCBI non-redundant protein sequences), Nt (NCBI nucleotide sequences); Pfam (Protein family; Finn et al., 2008); KOG/COG; Swiss-Prot; KEGG (Kyoto Encyclopedia of Genes and Genome); GO (Gene Ontology; Young et al., 2010).

Identification of Differentially Expressed Genes (DEGs)

After GO annotation, the successfully annotated genes were grouped into three main GO domains: Biological Process (BP), Cellular Component (CC), Molecular Function (MF). (Result Directory: Result/3.Annotation/GO classification). Genes with fold change (FC) of expression levels >8 and q-value (adjusted P-value) <0.01 were considered as significantly expressed genes (DEGs) (Wang et al., 2010).

Results and Discussion

Seed priming is correlated to significant improvements of stress tolerance in plants, but mechanisms by which primed memory is established are still elusive (Mladenov et al., 2021). In our previous experiments we have described how silicic acid seed priming can contribute to better cadmium tolerance in *Silene sendtneri* (Karalija et al., 2021). Through transcriptome analysis of current study, we aimed to investigate the molecular mechanisms in primed state that subsequently influences seedling performance under stress conditions in plants developing from primed seed. Collected data reflect the RNA-seq data of primed state in SiA primed seed of *Silene sendneri* in correlation to non-primed seeds.

Transcriptome sequencing

After removing the adaptor sequences, low quality, and short reads, 27899385 and 2548320 clean reads were obtained for non-primed and SiA primed seeds, respectively. The Q30 percentage exceeded 90% and CG content (guanine-cytosine) was over 45% which suggests highly accurate and reliable sequencing (Table 1). Transcriptome assembly was performed using methods for species with no reference genome (Grabherr et al., 2011).

Analysis of new genes and differentially expressed genes (DEGs)

In the present study we used high-throughput RNA-sequencing (RNA-seq) to identify differentially expressed genes (DEGs) in the primed seeds with 1% silicic acid in relation to non-primed seeds. A total of 76 944 and 65 218 gene transcripts were assembled for non-primed and SiA-primed seeds, respectively.

By comparing the genes that show differential expression in SiA-primed seeds compared to non-primed seeds we found a total of 221 DEGs using log2FC≥4 and q value ≤ 0.01.

The heatmap carried out on the 221 common genes, show clearly it is possible split the DEGs analysed in two different cluster based on their different expression in SiA-primed seeds and non-primed seeds, respectively.

Table 1. Comparison of the raw transcriptome data for non-primed and SiA primed seeds

e	Raw reads	Clean reads	Raw bases	Clean bases	Error (%)	Q20 (%)[a]	Q30 (%)[b]	GC (%)	Total Mapped
NP	28184221	27899385	8.5	8.4	0.03	97.43	92.94	45.13	74.97%
SiA	25743678	25483202	7.7	7.6	0.02	98.21	94.76	45.45	73.39%

[a] Q20indicates a quality score of 20, a 1% chance of error and 99% confidence
[b] Q30indicates a quality score of 30, a 0.1% chance of error and 99.9% confidence

Venn diagram was constructed to determine the number of DEGs unique and in common between each of the two comparisons (Figure 1). A total of 221 differentially expressed genes (DEGs) were analysed. Among them, 117 genes unique to SiA-primed seeds and 102 to non-primed seeds respectively. Only two DEGs results in common among the two different experimental conditions analysed. DEGs in common were also analysed with MultiExperiment viewer (MeV).

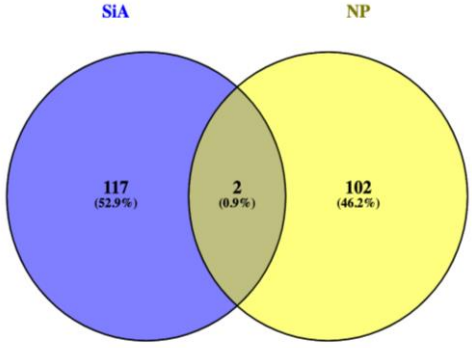

Figure 1. Venn diagram - Differentially expressed genes (DEGs) identified through RNA-Seq analysis in SiA-primed seeds in relation to non-primed seeds. Figure represents a total number of up- and down-regulated DEGs with log2FC≥4.

Functional annotations

All DEGs identified in SiA-primed seeds were functionally analysed using gene ontology (GO) classification, COG (Cluster of Orthologous Group), KEGG (KyotoEncyclopedia of Genes and Genomes), Swiss-Prot and NR (non-redundant) databases using BLAST (Basic Local Alignment Search Tool) software (Table 2). Species utilised for the mapping were *Beta vulgaris* subsp. *vulgaris* (28.3%), *Chenopodium quiroa* (27.8%), *Spinacia oleracea* (18.6%), *Quercus suber* (3.4%), *Vitis vinifera* (0.8%), other species (21.2%).

For the GO classification analysis of DEGs, in the summary graph Figure 2A we can identify different subcategories with indicated significant change in expression, such as ribosome related genes, protein-containing complexes, processes related to rRNA binding, signal transduction and cytoplasm and cytosol.

Table 2. The Ratio of Successfully Annotated Genes

Statistical Items	Number of Unigenes	Percentage (%)
Annotated in NR	138921	60.34
Annotated in NT	95783	41.60
Annotated in KO	55556	24.13
Annotated in Swissport	105519	45.83
Annotated in PFAM	49496	21.50
Annotated in GO	67738	29.42
Annotated in KOG	38425	16.69
Annotated in all Databases	10406	4.52
Annotated in at least one Database	148714	64.59
Total Unigenes	230248	100

Furthermore, DEGs were assigned to three main categories (biological processes, molecular function, and cellular compartment) with different sub-categories where several DEGs were assigned to more than one sub-category. In category of biological processes (Figure 2B) subcategory translation and signal transduction were two categories with significantly affected differential gene expression. In category cellular compartment subcategories ribosome, cytoplasm, protein containing complex, intracellular, cell, and cytosol included statistically significant differentially

expressed genes after silicic acid priming compared to non-primed seeds (Figure 2C). The category molecular function included three subcategories with statistically different gene expression: structural constituent of ribosome, structural molecule activity and rRNA binding (Figure 2D). These results indicate that primary effect of seed priming using silicic acid results in different gene expression related to ribosome activity and other processes correlated to identified nine subcategories with statistically different gene expression. According to the KEGG mapping, genes related to ribosome structure and function have significantly changed expression patterns in seeds primed with SiA (Figure 3).

An analysis using KEGG database on biological pathways showed statistically significant change of gene expression levels in genes involved in ribosome function (Figure 3). There are several upregulated genes such as genes L3, L18, S13 and L13 (Figure 4). Upregulation of L3 could be related with early development considering that seed priming can stimulate early germination and faster seedling emergence. L3 if silenced can result in decrease of pre-rRNA affecting biogenesis, and decreased L3 levels can delay development (Popescu et al., 2004). L18 are involved in virus infections as identified in infections with cauliflower mosaic virus (CaMV) (Leh et al., 2000).

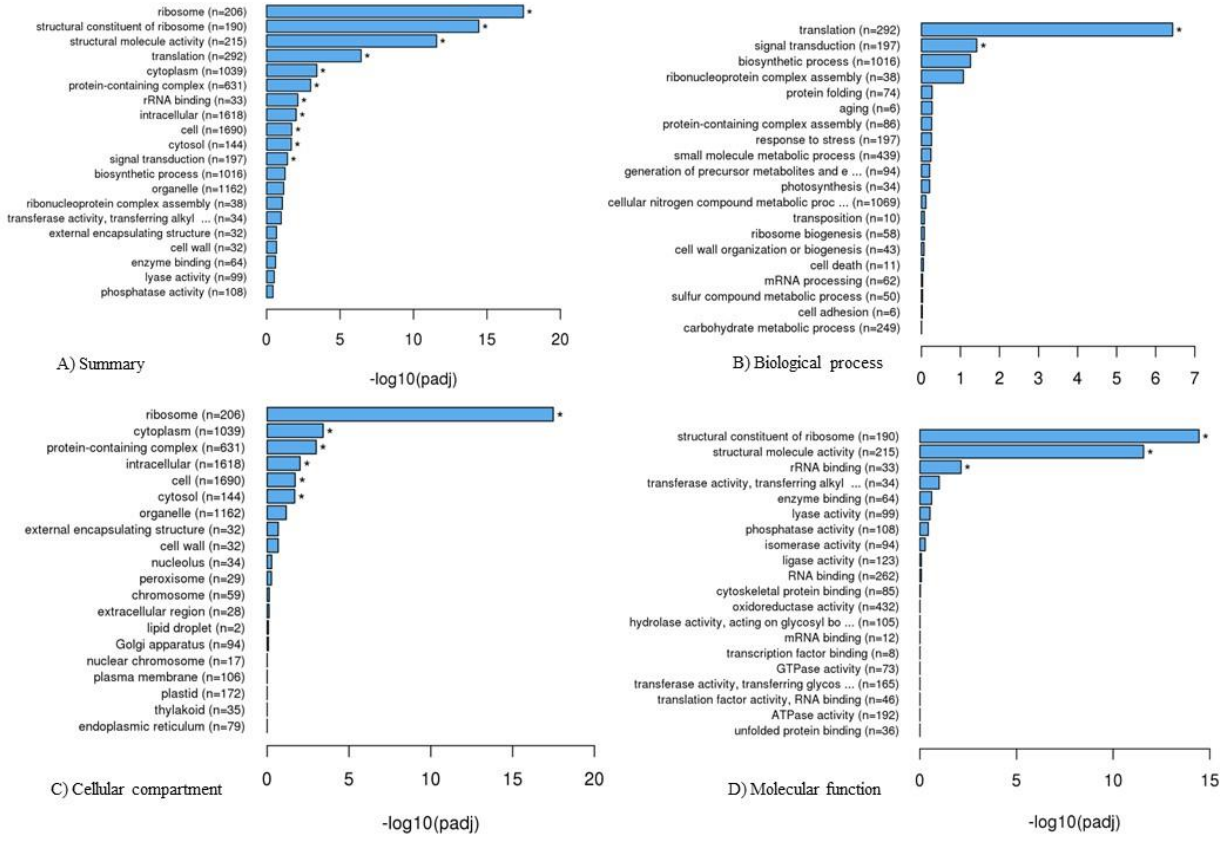

Figure 2. Functional annotation of differentially expressed genes in GO. A - summary of all DEGs; B- DEGs according to the biological function; C- DEGs according to cellular compartment and D – DEGs according to the molecular function. Asterix signifies statistically different gene expression in SiA primed seeds in relation to non-primed seeds at $p<0.05$ level.

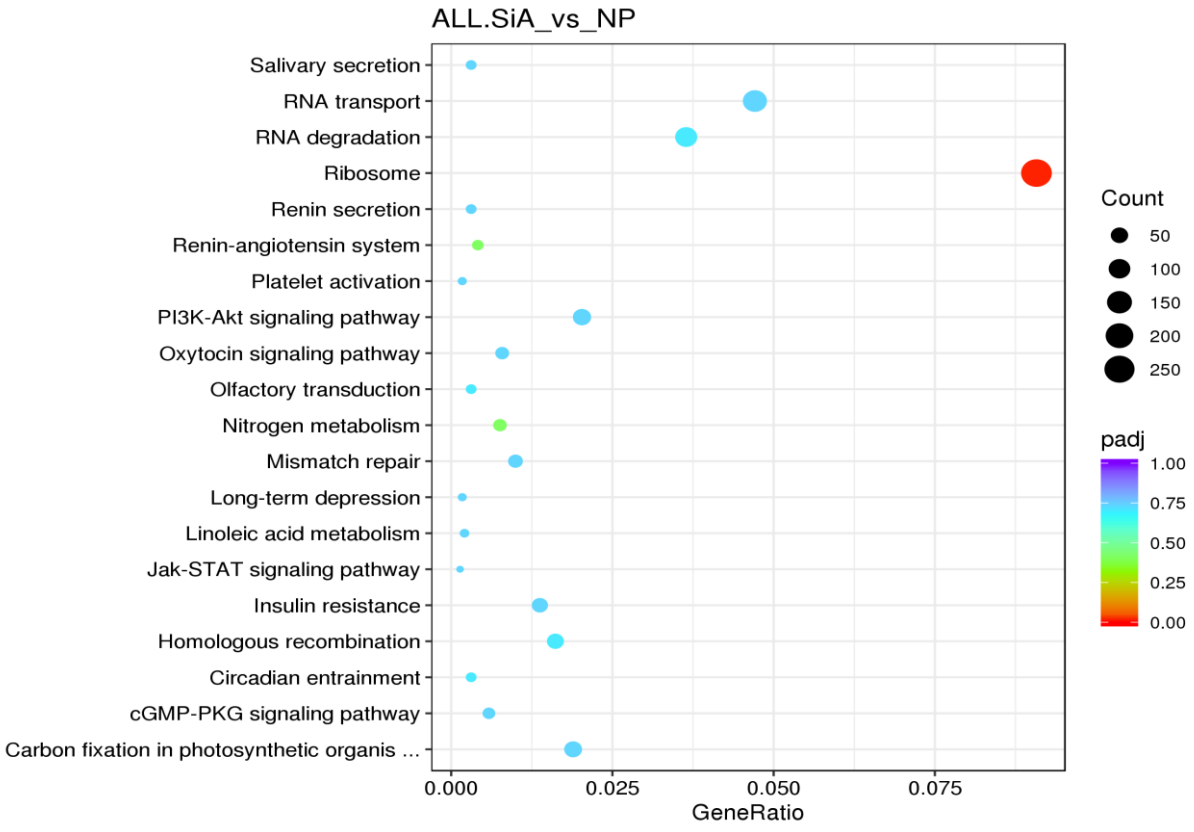

Figure 3. Diagram of KEGG enrichment analysis of DEGs in SiA primed seeds in relation to non-primed seeds.

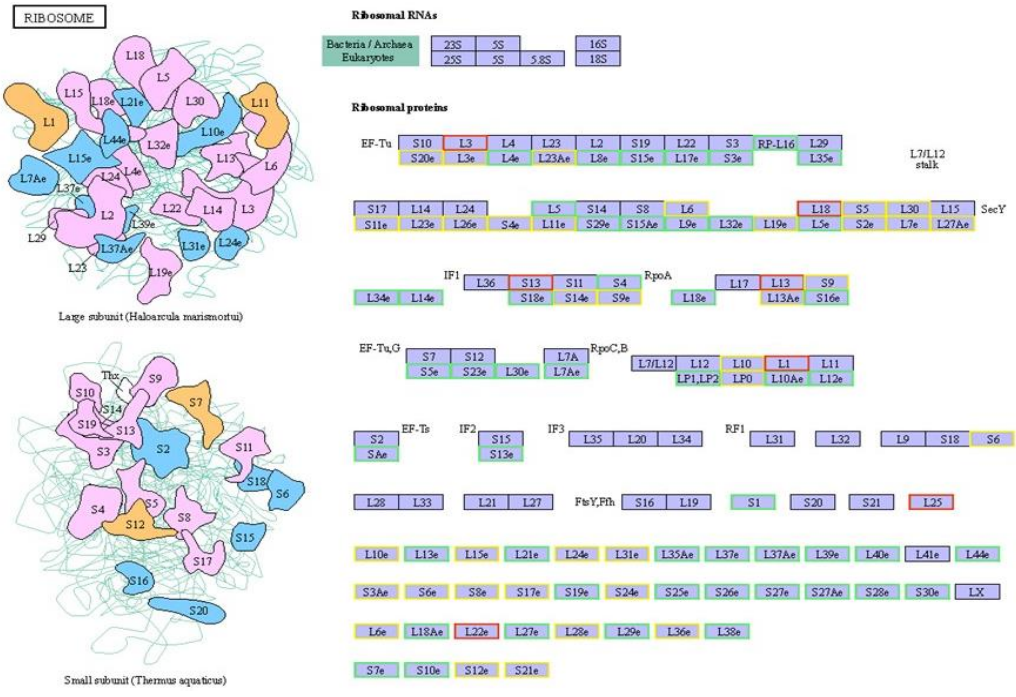

Figure 4. Visualisation of the most enriched pathways/genes (highlighted in green) related to ribosome and RNA degradation. Statistical method: hypergeometric test; FDR correction method: Benjamini and Hochberg Red marked genes (significantly upregulated genes) Green marked genes (significantly downregulated genes). Yellow (not significantly changed expression).

On the other hand, multiple genes related to ribosomal function show downregulation patterns in SiA primed seeds, such as S4, S1, L5 and L5e (Figure 4). Ribosomal protein L5 is also known to bind specifically to 5S rRNA and is involved in its nucleocytoplasmic transport (Rosorius et al., 2000), and the nuclear re-entry of 5S rRNA is mediated exclusively by the ribosomal protein L5 (Murdoch et al., 1996). Downregulation of this gene can affect rRNA transport. S4 protein also plays a role for the cytoplasm–nuclear translocation (Yi et al., 2002).

Identification of candidate genes involved in primed memory after SiA priming

Seed priming is a known mechanism for enhancement of plant stress response and tolerance. Analysis of annotated genes revealed several stress related genes up- and downregulated after SiA priming (Supplementary table 1). Presented study gives information on genes identified for the first time in *Silene sendtneri* such as heavy metal-associated isoprenylated plant protein and some predicted genes such as probable metal-nicotianamine transporter YSL6 which are connected to plants ability to detoxify toxic metals (Divol et al., 2013). These genes could be also responsible for the Cd accumulation abilities of *S. sentneri*. SiA treatment downregulated large number of heavy metals related as well as some other stress response genes such as heat shock protein genes. In the group of upregulated genes, members of pumilio (Pum)/Puf family RNA binding proteins genes, were upregulated after SiA priming. Members of this gene family are mostly related to plants response to stress, biotic as well as abiotic stress. In biotic stress expression of Pum genes is related to plants immunity (Huh, 2021) and synthesis of antibacterial peptides (Gerber et al., 2006). Expression of Pum genes is also changed under salt, cold or drought stress and upregulated expression of these genes can enhance heat stress tolerance (Nyiko et al., 2019). Considering DEGs identified after SiA seed priming we can postulate that SiA seed priming could be suitable for enhancement of plant tolerance toward biotic and abiotic stress, but it could reduce plant's ability to tolerate, accumulate and detoxify heavy metals. Further research is necessary to verify gene expression patterns in developing plant under stress.

Conclusions

Seed priming is a widely used technique but molecular mechanisms underlaying behind primed memory are still unclear. Silicic acid is a priming agent with known beneficial effects on plants tolerance of different stresses through induction of primed state in primed seeds. Large number of genes are affected by seed priming with SiA mostly affecting ribosome function and expression of genes correlated with plants response to biotic and abiotic stress.

Based on obtained results from RNA-Seq selected subset of differentially expressed genes will be further validated by quantitative real-time polymerase chain reaction (RT-qPCR).

Acknowledgement

This research has been supported by Ministry for education, science and Youth, Canton Sarajevo, through the project *Increase of heavy metal accumulation in selected hyperaccumulating plant species using seed priming* financed in 2019 and Ministry of Education, Science and Technological Development of the Republic of Serbia (Grant No. number 451-03-68/2022-14/200042).

Conflict of interest

Authors declare no conflict of interest.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Nature Precedings 1-1.

Buchfink B, Xie C, Huson DH (2014) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12: 59-60.

Chepelev I, Wei G, Tang Q, Zhao K (2009) Detection of single nucleotide variations in expressed exons of the human genome using RNA-Seq. Nucleic Acids Res 37: 106-106.

Divol F, Couch D, Conéjéro G, Roschzttardtz H, Mari S, Curie C (2013) The Arabidopsis YELLOW STRIPE LIKE4 and 6 transporters control iron release from the chloroplast. Plant Cell 25: 1040-1055.

Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A (2008) The Pfam protein families database. Nucleic Acids Res 36: 281–288.

Gerber AP, Luschnig S, Krasnow MA, Brown PO, Herschlag D (2006) Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in Drosophila melanogaster. Proc Natl Acad Sci USA 103: 4487–4492.

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M. Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36: 3420-3435.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644-652.

Huh SU (2021) The role of pumilio RNA binding protein in plants. Biomolecules 11: 1851.

Karalija E, Selović A, Dahija S, Demir A, Samardžić J, Vrobel O, Zeljković SĆ, Parić A (2021) Use of seed priming to improve Cd accumulation and tolerance in *Silene sendtneri*, novel Cd hyper-accumulator. Ecotoxicol Environ Saf 210: 111882.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2007) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36: 480-484.

Leh V, Yot P, Keller M (2000) The cauliflower mosaic virus translational transactivator interacts with the 60S ribosomal subunit protein L18 of *Arabidopsis thaliana*. Virology 266: 1-7.

Lei C, Bagavathiannan M, Wang H, Sharpe SM, Meng W, Yu J (2021) Osmopriming with Polyethylene Glycol (PEG) for Abiotic Stress Tolerance in Germinating Crop Seeds: A Review. Agronomy 11: 2194.

Li B, Dewey C (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:1-6.

Mao X, Cai T, Olyarchuk JG, Wei L (2005) Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics 21: 3787-3793.

Mladenov V, Fotopoulos V, Kaiserli E, Karalija E, Maury S, Baranek M, Segal NA, Testillano PS, Vassileva V, Pinto G, Nagel M (2021) Deciphering the epigenetic alphabet involved in transgenerational stress memory in crops. Int J Mol Sci 22: 7118.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35: 182-185.

Murdoch KJ, Allison LA (1996) A role for ribosomal protein L5 in the nuclear import of 5S rRNA in *Xenopus* oocytes. Exp Cell Res 227: 332–343.

Nyiko T, Auber A, Bucher E (2019) Functional and molecular characterization of the conserved Arabidopsis PUMILIO protein, APUM9. Plant Mol Biol 100: 199–214.

Paul S, Dey S, Kundu, R. (2022) Seed priming: an emerging tool towards sustainable agriculture. Plant Growth Regul 97: 215–234

Popescu SC, Tumer NE (2004) Silencing of ribosomal protein L3 genes in N. tabacum reveals coordinate expression and significant alterations in plant growth, development and ribosome biogenesis. Plant J 39:29-44.

Rosorius O, Fries B, Stauber RH, Hirschmann N, Bevec D, Hauber J (2000) Human ribosomal protein L5 contains defined nuclear localization and export signals. J. Biol. Chem 275: 12061–12068.

Srivastava AK, Suresh Kumar J, Suprasanna P (2021) Seed 'primeomics': plants memorize their germination under stress. Biol Rev 96: 1723-1743.

Wang L, Feng Z, Wang X, Wang X, Zhang X (2010) DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26: 136-138.

Yi L, Qu LJ, Chang S, Su Y, Gu H, Chen Z (2002) Two nuclear localization signals required for the nuclear localization of rice ribosomal protein S4. Plant Sci 162: 251-256.

Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis or RNA-seq: accounting for selection bias. Genome Biol 11: 1-12.