Analyzing the potato abiotic stress transcriptome using expressed sequence tags

Willem Rensink, Amy Hart, Jia Liu, Shu Ouyang, Victoria Zismann, and C. Robin Buell

Abstract: To further increase our understanding of responses in potato to abiotic stress and the potato transcriptome in general, we generated 20 756 expressed sequence tags (ESTs) from a cDNA library constructed by pooling mRNA from heat-, cold-, salt-, and drought-stressed potato leaves and roots. These ESTs were clustered and assembled into a collection of 5240 unique sequences with 3344 contigs and 1896 singleton ESTs. Assignment of gene ontology terms (GOSlim/Plant) to the sequences revealed that 8101 assignments could be made with a total of 3863 molecular function assignments. Alignment to a set of 78 825 ESTs from other potato cDNA libraries derived from root, leaf, stolon, tuber, germinating eye, and callus tissues revealed 1476 sequences unique to abiotic stressed potato leaf and root tissue. Sequences present within the 5240 sequence set had similarity to genes known to be involved in abiotic stress responses in other plant species such as transcription factors, stress response genes, and signal transduction processes. In addition, we identified a number of genes unique to the abiotic stress library with unknown function, providing new candidate genes for investigation of abiotic stress responses in potato.

Key words: potato, Solanacaeae, abiotic stress.

Résumé : Afin d'accroître notre compréhension des réponses de la pomme de terre aux stress abiotiques ainsi que le transcriptome de la pomme de terre en général, les auteurs ont produit 20 756 EST (étiquettes de gènes exprimés) issus d'une banque d'ADNc préparée à partir d'ARNm d'origines diverses (feuilles et racines de pommes de terre exposées à la chaleur, au froid, au sel ou à la sécheresse). Ces EST ont été groupés pour former une collection de 5 240 séquences distinctes comprenant 3 344 contigs et 1 896 EST uniques. L'attribution des termes d'ontologie génique (GOSlim/Plant) à ces séquences a révélé que 8 101 attributions ont pu être faites pour un total de 3 863 attributions de fonction moléculaire. La comparaison avec une collection de 78 825 EST provenant d'autres banques d'ADNc préparées à partir de racines, feuilles, stolons, tubercules, yeux en germination ou cals a montré que 1 476 séquences étaient uniques aux tissus foliaires ou racinaires stressés. Des séquences parmi la collection de 5 240 séquences montraient de la similitude avec des gènes connus comme étant impliqués dans la réponse aux stress abiotiques chez d'autres plantes; ceux-ci comprenaient des facteurs de transcription, des gènes de réponse aux stress et des gènes impliqués dans les voies de signalisation. De plus, les auteurs ont identifié nombre de gènes uniques à la banque de stress abiotiques et de fonction inconnue qui constituent des gènes candidats pour l'étude des réponses aux stress abiotiques chez la pomme de terre.

Mots clés : pomme de terre, Solanacées, stress abiotique.

[Traduit par la Rédaction]

Introduction

Plants are exposed to numerous stresses, abiotic and biotic. Both of these stresses have substantial impacts on plant growth and development. With the growing need for increased agricultural output in the 21st century, gains in yield will need to be obtained through several approaches. Clearly, reducing yield loss due to abiotic and biotic stress factors would be one such approach. A 2nd approach would be to increase the amount of arable land through engineering increased tolerance of crop species to stressful conditions such as salinity and drought.

Biotic stress responses have been well investigated in plants with a central theme of resistance genes mediating recognition of the pathogen in which a signal transduction pathway is activated that leads to manifestation of resistance. The genetic and molecular basis of biotic stress responses has been amenable to researchers as manipulation of single genes, through either over-expression or mutation, has allowed for a systematic investigation of a wide range of components involved in the biotic stress response from the sensor to the signal transduction components to the actual genes involved in production of antimicrobial factors (for review, see Hammond-Kosack and Parker 2003). Abiotic stress responses have also been addressed in plants. Although the responses appear to be more complex, that is, polygenic in nature, a set of signal transduction events, coupled with production of stress response factors, appears to be

Received 11 November 2004. Accepted 12 April 2005. Published on the NRC Research Press Web site at http://genome.nrc.ca on 6 August 2005.

Corresponding Editor: F. Belzile.

W. Rensink, A. Hart, J. Liu, S. Ouyang, V. Zismann, and C.R. Buell.[1] The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA.

[1]Corresponding author (e-mail: rbuell@tigr.org).

 doi: 10.1139/G05-034

shared in the response to multiple abiotic stressors (for review see Hazen et al. 2003; Schinozaki et al. 2003; Zhu 2001). The 1st step in abiotic stress responses is perception of the stress. This is followed by signal transduction events involving 2nd messengers, for example, cytosolic calcium leading to activation of transcription factors. It is through the action of the transcription factors, which bind to and activate stress-response genes, that results in the synthesis of stress-related gene products and manifestation of stress tolerance.

The Solanaceae family contains a group of species that are a major source of vegetables for the human diet including potato, tomato, eggplant, and pepper. Other solanaecous species are important for ornamental purposes, for example *Petunia* and *Nicotiana* spp, whereas *Nicotiana tabacum* is grown for leaves for the tobacco industry. Thus, collectively, the Solanaceae family represents a major economic sector of agriculture. Potato, *Solanum tuberosum*, is the 5th largest crop produced in the world (http://apps.fao.org/default.jsp). It is grown for its tuber, a below ground modified stem tissue that has a large caloric content of primarily starch. Growth of potato and production of marketable tubers is affected by environmental conditions such as heat stress, cold stress, and drought stress (DeBlonde et al. 1999; Ewing 1981; Isherwood 1973; Lafta and Lorenzen 1995; Schwimmer et al. 1957; Veilleux et al. 1997).

At The Institute for Genomic Research, we have developed a structural and functional genomics resource platform for potato. This has involved generation of expressed sequence tags (ESTs) for potato (Ronning et al. 2003; http://www.tigr.org/tdb/potato/est2.shtml), construction of a gene index for potato representing a collection of unique transcripts for potato (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=potato), a ~10 000 clone potato cDNA microarray (http://www.tigr.org/tdb/potato/microarray_comp.shtml), and an expression profiling database for solanaceous species (Solanaceae Gene Expression Database; http://www.tigr.org/tdb/potato/SGED_index2.shtml). Our previous EST sequencing efforts focused on generation of ESTs from core tissues of potato (leaf, root, stolon, tuber, and germinating eyes) and biotic stressed leaves (late blight challenged, compatible, and incompatible), thereby, providing insight into potato development and biotic stress responses (Ronning et al. 2003). To further expand the representation of tissues and physiological responses in potato, we constructed a cDNA library from abiotic stressed potato leaves and roots and generated ~20 000 ESTs from this library. The cDNA library represents plants exposed to salt, heat, cold, or drought stress and thus represents a collective set of ESTs for abiotic stress responses in potato. In this study, we report on the sequence and analysis of this collection of ESTs that provides a large-scale molecular resource for understanding the abiotic stress transcriptome in potato.

Materials and methods

cDNA library construction and high-throughput sequencing

Solanum tuberosum var. *Kennebec* plants were grown from cuttings on a 16-h light and 8-h dark cycle at 25 °C for 3–4 weeks. Abiotic stress conditions were applied to 4 separate sets of plants. Set 1 involved saturation of the soil with 150 mmol NaCl/L. Tissues were harvested following application of the salt stress; leaves were collected at 2 h, 6 h, 12 h, 24 h, 48 h, and 96 h; and roots were collected at 2 h, 6 h, 12 h, and 48 h after stress initiation.

Set 2 was grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Leaves were harvested at 3 d, 5 d, and 7 d after cessation of watering. Roots were harvested 3 d and 5 d after cessation of watering.

Set 3 was grown under the standard conditions and then were cold stressed by placement at 4 °C. Cold stressed leaves were harvested at 2 h, 6 h, 12 h, 24 h, and 96 h after stress initiation. Roots were harvested at 2 h, 6 h, 12 h, 24 h, and 48 h after stress initiation.

Set 4 was grown under standard conditions and then heat stressed by placement at 35 °C. Heat stressed leaves were harvested at 2 h, 6 h, 12 h, 24 h, 48 h, and 96 h after stress initiation. Heat-stressed roots were harvested at 6 h, 12 h, 24 h, and 96 h after stress initiation.

RNA was isolated from all tissues using a modified Trizol (Invitrogen, Carlsbad, Calif.) method (Chomczynski and Sacchi 1987). Equal RNA from each tissue and stress-type was pooled to construct a cDNA library in the pCMVSport6.1 (http://www.invitrogen.com). The resulting library was termed POA for potato abiotic stress. DNA was isolated from cDNA clones using an alkaline lysis method and sequenced from both ends (5′ and 3′) using standard high-throughput sequencing methods with ABI Big Dyes (Applied Biosystems, Foster City, CA). Sequences were named by combining clone name (POAXXXX) with primer name (TP or TR (5′); TF or TV(3′)). Sequences were screened for *Escherichia coli* and vector sequences and trimmed to remove vector and low-quality regions. Sequences are available in the dbEST division of GenBank under the accession numbers CK259302–CK280059.

Bioinformatic analyses

Sequences were assembled using CAP3 (Huang and Madam 1999) to reduce redundancy and identify unique sequences within the POA library. The resulting contig and singleton EST sequences were transitively annotated by selecting the top match obtained from a BLASTX search of the POA sequences against an in-house nonidentical amino acid database with a minimum expect value cutoff of $\leq 1\times10^{-5}$. EST sequences from other potato cDNA libraries (Ronning et al. 2003) were also clustered using CAP3 to generate a set of nonredundant sequences for each library and tissue. Gene ontologies (GOSlim/Plant; http://www.geneontology.org) were transitively annotated by searching with BLASTN (expect value cutoff $< 1\times10^{-10}$) against a curated set of coding sequences (CDS) of *Arabidopsis thaliana* GO-annotated proteins (Wortman et al. 2003).

Results and discussion

Sequencing and annotation

A total of 11 904 clones were sequenced from the 5′ and 3′ ends resulting in 20 756 good sequences (11 409 sequences from the 5′ end and 9347 sequences from the 3′ end) of which 8944 clones had sequences from both ends. The raw

Table 1. Abundant sequences in the potato abiotic stress cDNA library.

Contig No.[a]	No. of sequences in contig	No. of clones in contig	Tentative annotation	Organism	Annotated GenBank match[b]	*E* value
2816	202	108	Tomato ripening associated membrane protein	*Lycopersicon esculentum*	X73848.1	0
379	138	69	Plastidic aldolase	*Solanum tuberosum*	Y10380.1	0
1344	112	87	S-adenosyl-L-methionine synthetase	*Solanum brevidens*	AY635050.1	0
3176	110	60	Glyceraldehyde 3-phosphate dehydrogenase	*Solanum tuberosum*	AF527779.1	0
609	108	62	Hypothetical protein	*Solanum tuberosum*	AJ225024.1	0
11	105	55	Type II chlorophyll a/b-binding protein	*Lycopersicon esculentum*	X14036.1	0
118	92	76	Elongation factor 1-alpha	*Solanum tuberosum*	AB061263.1	0
1490	85	43	Polyubiquitin	*Petroselinum crispum*	X64344.1	0
624	83	46	Plastidic aldolase	*Nicotiana paniculata*	AB027001.1	0
788	83	69	Elongation factor 1-alpha	*Solanum tuberosum*	AB061263.1	0
1086	82	47	Cysteine proteinase	*Lycopersicon esculentum*	Z14028.1	0
6	80	43	Oxygen-evolving complex	*Lycopersicon esculentum*	Z11999.1	0
296	77	41	Chlorophyll a/b-binding protein CP29	*Vigna radiata*	AF139466.2	5×10^{-69}
1526	76	41	Heat shock protein	*Lycopersicon esculentum*	X54030.1	0
1508	70	40	Ribosomal protein L3	*Lycopersicon esculentum*	AY456411.1	0
2819	70	40	Major intrinsic protein 2	*Solanum tuberosum*	Y18312.1	0
1033	65	36	Alpha-tubulin (*tubA3* gene)	*Nicotiana tabacum*	AJ421413.1	0
2081	65	61	Elongation factor 1-alpha	*Solanum tuberosum*	AB061263.1	0
1015	62	36	Elongation factor 1-alpha	*Nicotiana paniculata*	AB019427.1	0
2612	44	39	Chloroplast carbonic anhydrase gene	*Lycopersicon esculentum*	AJ849376.1	0

[a]All clones are present in other cDNA libraries examined in this study.
[b]Sequences were searched against the nonredundant amino acid database at NCBI. Only matches to annotated sequences are reported.

sequences were trimmed to remove vector and low-quality sequences; after trimming the average length of the sequences was 795 nucleotides. In total, 6567 clones had paired end sequences and aligned on 1 contig, enabling estimation of an average insert size of 1.26 kb for the cDNA library. Not all paired end sequences from a clone could be aligned. This could be due to many factors such as short read lengths, lower quality sequences at the sequence termini, or large cDNA insert size. A nonredundant set of sequences was generated by the assembly of the 20 756 sequences using CAP3 to yield 5240 unique sequences with 3344 contigs and 1896 singletons. The sequences were annotated based on sequence similarity with entries in a nonidentical amino acid database using a minimum expect value cut-off of 1×10^{-5}. In total, 4916 of the 5240 sequences matched an entry in the database, leaving only 324 sequences novel to potato.

The top 20 most abundant sequences within the POA library are listed in Table 1 and represent known abundant "housekeeping genes" such as elongation factor, glyceraldehyde 3-phosphate dehydrogenase, photosynthetic-related proteins, and tubulin. Indeed, all of the top 20 abundant sequences were also detected in other potato cDNA libraries (see subsequent text) suggesting that they are not necessarily unique to the response to abiotic stress in the tissues sampled. However, some of these genes have been documented previously as being induced by stress. For example, the most abundant sequence (Contig 2816; 202 sequences from 108 clones) en-

codes a protein with similarity to a MIP class membrane channel protein from tomato. The tomato ripening-associated membrane protein (TRAMP) was reported to be expressed in numerous tissues, including roots, and was upregulated in tomato stems during water stress (Fray et al. 1994). The 3rd most abundant sequence (Contig 1344; 112 sequences from 87 clones) encodes S-adenosyl-L-methionine synthetase, which has been shown to be induced by several stresses including salt as well as mannitol and absicic acid (Espartero et al. 1994). The time course of accumulation of S-adenosyl-L-methionine synthetase transcripts reported previously overlaps well with the temporal period from which we collected tissue for the POA cDNA library construction. Two contigs encode putative plastidic aldolases (Contig 379, 138 sequences; Contig 624, 83 sequences). These aldolases have high identity with plastidic fructose-1,6-bisphosphate aldolases (NpAldP1 and NpAldP2) from *Nicotiana paniculata,* a water-deficit and stress-tolerant species (Yamada et al. 2000). NpAldP1 and NpAldP2 share 91% identity yet differ in their expression in response to salt stress. Whereas NpAldP2 is upregulated in response to salt stress, NpAldP1 is downregulated. Not surprisingly, 18 of the top 20 abundant sequences had the highest identity with annotated sequences previously identified from solanaceous species, which is consistent with the high degree of sequence conservation within the Solanaceae reported previously (Zamir and Tanksley 1988).

The nonredundant sequences were further analyzed by the assignment of gene ontology (GO) terms. Gene ontologies (The Gene Ontology Consortium 2000) are controlled vocabularies enabling an assessment and comparison of gene functions across species. Gene functions can be described at 3 levels based on molecular function, cellular component, or molecular function. GOSlim terms are a subset of the GO ontologies, which provide a higher level of annotation than the standard GO ontologies (http://www.geneontology.org), thereby allowing for a more global view of a genome or genomic dataset. Based on sequence similarity to *Arabidopsis*, a total of 8101 GOSlim assignments (molecular function (3863), biological process (3035), and cellular component (1203)) were made to the 5240 POA nonredundant sequences. A breakdown of the assignment to GOSlim molecular function terms is shown in Table 2. The most prevalent GOSlim categories (>10%) are that of catalytic activity (13.77%) followed by hydrolase (12.76%) and then transferase (10.87%), suggesting a high degree of basic metabolic activity in the stressed tissues. Less prevalent GOSlim categories (5%–10%) are protein binding (7.17%), "other" binding (6.68%), transporter activity (6.63%), nucleotide binding (6.39%), kinase (6.21%), and DNA binding (5.85%). These GO categories have been implicated in the general stress response pathway consisting of stress recognition, downstream signaling events, and adaptation responses.

Table 2. Classification of the nonredundant POA sequences to the molecular function gene ontologies.

GO term	Percent
Carbohydrate binding	0.60
Catalytic activity	13.77
Chaperone activity	1.16
DNA binding	5.85
Enzyme regulator activity	0.62
Hydrolase	12.76
Kinase	6.21
Lipid binding	0.65
Nuclease activity	0.54
Nucleic acid binding	3.00
Nucleotide binding	6.39
Other binding	6.68
Other molecular function	1.45
Oxygen binding	0.65
Protein binding	7.17
RNA binding	4.32
Signal transducer activity	0.85
Structural molecule activity	3.24
Transcription factor activity	4.01
Transcription regulator activity	1.19
Transferase	10.87
Translation factor activity, nucleic acid binding	1.37
Transporter activity	6.63

Note: The set of 5240 nonredundant POA sequences were searched against the CDS of the *Arabidopsis* proteome using BLASTN. Molecular function ontologies were transitively annotated to the POA sequences if the POA sequence matched an *Arabidopsis* protein with an expect value cutoff of $\leq 1\times10^{-10}$. The reduced representation of the molecular function ontologies were derived from goslim_plant.2003 available at http://www.geneontology.org/. GOSlim terms with fewer than 20 POA sequences were not included.

Comparative analysis with other potato ESTs

To identify sequences unique to the POA library and to identify sequences expressed in other potato tissues, we performed iterative BLASTN searches with the POA library against nonredundant sequence datasets constructed from ESTs from 9 other potato cDNA libraries. Table 3 summarizes the overlap of sequences between the potato libraries. The percent of POA sequences present in other libraries ranged from 16.6% (PPC, *Phytophthora infestans* challenged leaves-compatible) to 29.3% (PSE and PEY, sprouting eyes). This range is a reflection of the dataset size in these other cDNA libraries as well as the similarity of the transcriptomes; the deeper the library has been sequenced the more likely the libraries share more sequences. Because of the relative abundance of housekeeping genes and other general metabolic genes, such a result can be expected. The previously mentioned search revealed that the POA nonredundant sequence set shares 16.6%–29.3% of its sequences with any one of these cDNA libraries, yet a total of 1476 sequences were unique to the POA library and are not present in the other 9 cDNA libraries (Supplemental Data 1)[2], thus further expanding our knowledge of the potato transcriptome.

[2] Supplementary data for this article are available on the Web site or may be purchased from the Depository of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Ottawa, ON K1A 0S2, Canada. DUD 3695. For more information on obtaining material refer to http://cisti-icist.nrc-cnrc.gc.ca/irm/unpub_e.shtml.

Table 3. Detection of sequences from the potato abiotic stress sequences in 9 potato cDNA library EST collections.

cDNA library[a]	Tissue source	No. of ESTs[b]	No. of nonredundant sequences[c]	Sequences shared with POA[d]	Percent of POA library
PPC	*Phytophthora infestans* challenged leaves-compatible	5 062	3079	871	16.6
PTD	Tuber, dormant	5 049	2904	889	17.0
PTE	Tuber, microtuber	5 426	2718	901	17.2
POC	Callus	16 799	3550	977	18.6
PPI	*P. infestans* challenged leaves-incompatible	5 434	3454	1068	20.4
PST	Stolons	10 292	5196	1095	20.9
PSH	Healthy leaves	10 440	5348	1248	23.8
PRO	Root	10 190	5317	1490	28.4
PSE, PEY	Sprouting eyes	10 133	5483	1536	29.3
Total	—	78 825	—	—	—

[a]Detailed information on all the libraries except POC is available in Ronning et al. 2003. POC was constructed from *S. tuberosum* var. *Kennebec* callus tissue grown on solid media. All the libraries, except POC, were non-normalized.
[b]All ESTs are available in GenBank.
[c]Number of nonredundant sequences within the libraries were determined by assembly with CAP3.
[d]Shared sequences were defined by having >95% identity with the query sequence.

Abiotic stress-specific sequences

Within the collection of 1476 POA-unique sequences, 667 contigs and 809 singleton ESTs were present. Annotation for these sequences is available in the supplemental table (S1).[2] Sequences present in this collection had sequence identity with genes involved in signaling cascades, transport, membrane composition, photosynthesis, and transcription regulation. Multiple putative heat shock proteins, including 2 of the top 10 most abundant unique sequences (Table S1; Contig 3002, 29 sequences; Contig 161, 14 sequences), were present in the POA-unique dataset, consistent with a previous study showing expression of heat shock proteins in response to heat stress in potato (Ahn et al. 2003). Three other POA-unique sequences (Contig 209, 12 sequences; Contig 95, 11 sequences; Contig 177, 4 sequences) encode homologs of the *Nicotiana tabacum phi-1* gene that is induced in phosphate stress and application of abscisic acid (Sano et al. 1999; Sano and Nagata 2002), suggesting a role of this gene in multiple abiotic stress responses. Another contig (Contig 665, 4 sequences) encodes a homolog of an osmotin-like protein shown previously to be induced in response to salt, low temperature, and abscisic acid in the wild potato species, *Solanum commersonii* (Zhu et al. 1995).

Numerous POA-unique sequences share identity with genes primarily from *Arabidopsis* annotated as "unknown", "unnamed", or "expressed protein", and our finding that the potato homologs are expressed in abiotic stressed tissue provides a new layer of evidence for functional annotation for these genes. Furthermore, a number of the 1476 POA-unique sequences share identity with genes annotated as "hypothetical proteins" primarily from *Arabidopsis* providing evidence not only that the homolog in *Arabidopsis* encodes a real gene but also that it may be involved in abiotic stress responses. In addition, 324 sequences from the entire set of 5240 sequences did not have a match with entries in the nonidentical amino acid database.

Comparative analysis of the abiotic stress transcriptome with *Arabidopsis*

Whereas a few genes have been implicated in the abiotic stress response in potato, the response in *Arabidopsis* to abiotic stress has been studied in detail owing to the availability of a near-complete genome sequence and large collection of mutant lines. As a consequence, a number of components of the abiotic stress response are known (for review, see Schinozaki et al. 2003; Zhu 2001). To ascertain whether similar components are present in potato and potentially involved in the potato response to abiotic stress, we selected *Arabidopsis* genes, as well as gene family members, reported to be implicated in the abiotic stress response and searched the 5240 POA sequence dataset for putative orthologs and (or) paralogs. Using the most recent annotation of the *Arabidopsis* genome (Release 5, http://www.tigr.org/tdb/e2k1/ath1), we have identified 45 genes in *Arabidopsis* that have been implicated in abiotic stress responses (Table 4). Collectively, these *Arabidopsis* genes have been implicated in all 4 abiotic stresses (cold, heat, salt, and drought) that were sampled in the POA library used in this study. These *Arabidopsis* genes include sensors, transcription factors, and proteins involved in protective functions. Using TBLASTX, we searched the POA nonredundant sequences and identified 19 unique contigs and 8 singletons in the POA library with significant hits ($E < 1\times10^{-5}$) to 34 *Arabidopsis* abiotic stress sequences. The most abundant sequences in the POA library are RD19 (responsive to dessication), RD21, RD22, and RD28 with more than 20 sequences in the library. The singleton clones are LEA14 (late embryogenesis abundant protein), PLC1 (phospholipase C), CBL (cystathionine beta-lyase), ERD13 (early dehydration induced), ETR1 (ethylene response), LTI65 (low temperature induced), CBF1 (C-repeat / DRE binding factor) and ABA1 (ABA deficient 1).

For the POA sequences (19 contigs and 8 singletons) with *Arabidopsis* homologs, we compared the transcript frequency of these sequences in the POA library with the transcript frequency in 8 other potato non-normalized cDNA libraries, which represent other stages of potato growth and development. Unlike most of the other libraries examined, the POA clones were sequenced from both ends and thus we determined frequency on a clone basis, rather than on a sequence basis. The potato abiotic stress homologs represented 1.53% of the clones within the POA library, which was substantially higher than the frequencies detected in the 8 other

Table 4. *Arabidopsis* genes implicated in abiotic stress responses present/absent in the potato abiotic stress library. EST, expressed sequence tag.

Locus name	Response	At gene name	At gene product	POA match	*E* value	No. of ESTs per contig
At4g39090	Abiotic stress	RD19	Responsive to dessication	Contig1086	3.40×10^{-134}	82
At2g37180	Abiotic stress	RD28	Responsive to dessication	Contig592	6.60×10^{-125}	54
At4g11310	Abiotic stress	RD21	Responsive to dessication	Contig1761	4.70×10^{-88}	37
At1g47128	Response to water deprivation	RD21	Responsive to dessication	Contig1761	1.50×10^{-161}	37
At5g25610	Abiotic stress	RD22	Responsive to dessication	Contig1284	1.20×10^{-97}	36
At5g51070	Dehydration	ERD1	Early dehydration induced	Contig411	7.80×10^{-82}	8
At3g11020	Response to water deprivation	DREB2B	DRE-binding protein	Contig131	1.30×10^{-16}	8
At5g05410	Response to water deprivation	DREB2A	DRE-binding protein	Contig1502	1.80×10^{-17}	8
At4g27410	Abiotic stress	RD26	Responsive to dessication	Contig1113	9.70×10^{-96}	7
At2g41430	Dehydration	ERD15	Early dehydration induced	Contig255	9.80×10^{-31}	6
At4g25470	Response to cold	CBF2/DREB1C	C-repeat / DRE binding factor	Contig2400	1.90×10^{-15}	6
At4g25480	Response to cold	CBF3/DREB1A	C-repeat / DRE binding factor	Contig2400	7.30×10^{-16}	6
At1g76180	Dehydration	ERD14	Early dehydration induced	Contig1389	6.30×10^{-26}	5
At1g20440	Response to cold	COR47/LEA	Cold regulated gene / late embryogenisis abundant protein	Contig1389	7.90×10^{-30}	5
At1g20450	Response to cold	LTI29/ERD10	Low temperature induced / early dehydration induced	Contig1389	8.10×10^{-20}	5
At1g08930	Dehydration	ERD6	Early dehydration induced	Contig3295	4.20×10^{-53}	4
At2g17840	Dehydration	ERD7	Early dehydration induced	Contig445	7.70×10^{-114}	4
At3g26744	Response to cold	ICE	Inducer of CBF expression	Contig1630	3.50×10^{-81}	3
At1g45249	Abiotic stress	ABF2	Abscisic acid responsive element-binding factor	Contig1374	8.40×10^{-69}	2
At1g49720	Abiotic stress	ABF1	Abscisic acid responsive element-binding factor	Contig1374	9.30×10^{-44}	2
At3g19290	Abiotic stress	ABF4	Abscisic acid responsive element-binding factor	Contig1374	2.10×10^{-51}	2
At4g34000	Abiotic stress	ABF3	Abscisic acid responsive element-binding factor	Contig1374	1.20×10^{-31}	2
At5g66400	Dehydration	RAB18	Responsive to ABA	Contig3169	8.60×10^{-22}	2
At5g51990	Response to cold	CBF4	C-repeat / DRE binding factor	Contig769	4.70×10^{-17}	2
At5g27150	Salinity response	NHX1	NA(+)/H(+) antiporter	Contig2638	2.70×10^{-61}	2
At2g47190	Salinity response	Myb2	Transcription factor	Contig3313	1.50×10^{-14}	2
At5g58670	Abiotic stress	PLC1	Phospholipase C	POADJ64TV	6.80×10^{-79}	1
At1g01470	Dehydration	LEA14	Late embryogenisis abundant	POABC40TV	2.80×10^{-25}	1
At2g30870	Dehydration	ERD13	Early dehydration induced	POAD391TV	2.10×10^{-15}	1
At1g66340	Response to cold	EIN1/ETR1	Ethylene response	POAD723TP	1.20×10^{-55}	1
At5g52300	Response to cold	LTI65	Low temperature induced	POADO59TV	0	1
At4g25490	Response to cold	CBF1/DREB1B	C-repeat / DRE binding factor	POADS38TP	1.70×10^{-17}	1
At5g67030	Response to cold	ABA1	ABA deficient 1	POAE482TP	5.20×10^{-115}	1
At4g17615	Response to osmotic stress	CBL	Cystathionine beta-lyase	POAC607TP	1.30×10^{-26}	1
At4g22260	Response to temperature	IM	Immutans	No hit	NA	0
At2g21620	Abiotic stress	RD2	Responsive to dessication	No hit	NA	0
At1g02930	Dehydration	ERD11	Early dehydration induced	No hit	NA	0
At3g30775	Dehydration	ERD5	Early dehydration induced	No hit	NA	0
At3g05880	Hyperosmotic salinity response	RCI2A	Rare-cold-inducible	No hit	NA	0
At5g67590	Response to cold	FRO1	Frostbite	No hit	NA	0
At2g42540	Response to cold	COR15A	Cold responsive protein	No hit	NA	0
At5g15960	Response to cold	KIN1	Stress responsive protein	No hit	NA	0
At5g15970	Response to cold	KIN2	Stress responsive protein	No hit	NA	0

Table 4 (*concluded*).

Locus name	Response	At gene name	At gene product	POA match	*E* value	No. of ESTs per contig
At5g52310	Response to cold	COR78 / RD29 / LTI140	Cold regulated gene	No hit	NA	0
At2g33380	Response to dessication	RD20	Responsive to dessication	No hit	NA	0

Table 5. Clone frequency in 9 potato non-normalized cDNA libraries of potato homologs of *Arabidopsis* abiotic stress-related genes.

cDNA library[a]	Tissue source	Clone frequency (%)
POA	Abiotic stress, leaves and roots	1.53
PPC	*Phytophthora infestans* challenged leaves-compatible	0.67
PPI	*P. infestans* challenged leaves-incompatible	1.03
PRO	Root	0.88
PSE, PEY	Sprouting eyes	0.37
PSH	Healthy leaves	0.46
PST	Stolons	0.29
PTD	Tuber, dormant	0.75
PTE	Tuber, microtuber	0.31

[a]Detailed information on all the libraries except POA is available in Ronning et al. 2003. POA is described in this study. All the libraries were non-normalized.

cDNA libraries (Table 5). The lowest frequency of these *Arabidopsis* homologs was in the stolon (0.29%) and microtuber libraries (0.31%), which are sequential stages in the development of tubers. The library with the 2nd highest frequency of these genes was the PPI library (1.03%), which was constructed from leaves challenged with an incompatible isolate of *P. infestans*, an oomycete pathogen of potato (Ronning et al. 2003), and consistent with previous reports of overlap in the signal transduction pathways of abiotic stress and disease resistance (Singh et al. 2002).

For some of the *Arabidopsis* stress response genes (11), no match could be found in the POA library. The following are several possible explanations for the lack of detection of putative orthologs of these stress-related genes. (*i*) We may not have sequenced the POA library to sufficient depth to capture these sequences. (*ii*) These genes may not be expressed in potato at the time points, tissues, or stress conditions that were used in the POA cDNA library construction and, therefore, would not be very abundant in the POA library, which was created by pooling multiple tissues, abiotic stresses, and timepoints. (*iii*) A subset of the *Arabidopsis* stress genes may have substantially diverged, preventing detection under our search criterion as we were unable to find putative potato homologs for 5 of the 11 *Arabidopsis* genes in the public databases. (*iv*) These genes may be absent in potato.

In conclusion, these analyses show that the POA library contains genes with a high likelihood of being involved in the abiotic stress response in potato. Furthermore, the sequencing of the POA library provides a valuable addition to the identification of the complete potato transcriptome that, to date, has been dominated by biotic stress libraries and organ-specific libraries. Also, the sequencing of the POA library provides new molecular resources for investigating abiotic stress in this crop species.

Acknowledgements

This work was supported through a grant from the National Science Foundation Plant Genome Research Program (DBI-9975866 and DBI-0218166). The assistance of the J. Craig Venter Joint Technology Center, the TIGR Sequencing Facility, and the TIGR IT Group are acknowledged. Sequences are available in the dbEST division of GenBank under the accession numbers CK259302–CK280059. All POA clones can be obtained from the University of Arizona Genomics Institute at http://www.genome.arizona.edu/.

References

Ahn, Y.-J., Claussen, K., and Zimmerman, J.L. 2003. Genotypic differences in the heat-shock response and thermotolerance in four potato cultivars. Plant Science, **166**: 901–911.

Chomczynski, P., and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**: 156–159.

DeBlonde, P.M.K., Haverkort, A.J., and Ledent, J.F. 1999. Responses of early and late potato cultivars to moderate drought conditions: agronomic parameters and carbon isotope discrimination. Eur. J. Agronomy, **11**: 91–105.

Espartero, J., Pintor-Toro, J.A., and Pardo, J.M. 1994. Differential accumulation of S-adenosylmethionine synthetase transcripts in response to salt stress. Plant Mol. Biol. **25**: 217–227.

Ewing, E.E. 1981. Heat stress and the tuberization stimulus. Am. Potato J. **58**: 31–49.

Fray, R.G., Wallace, A., Grierson, D., and Lycett, G.W. 1994. Nucleotide sequence and expression of a ripening and water stress-related cDNA from tomato with homology to the MIP class of membrane channel proteins. Plant Mol. Biol. **24**: 539–543.

Hammond-Kosack, K.E., and Parker, J.E. 2003. Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. Curr. Opin. Biotechnol. **14**: 177–193.

Hazen, S.P., Wu, Y., and Kreps, J.A. 2003. Gene expression profiling of plant responses to abiotic stress. Funct. Integr. Genomics, **3**: 105–111.

Huang, X., and Madan, A. 1999. CAP3: a DNA sequence assembly program. Genome Res. **9**: 868–877.

Isherwood, F.A. 1973. Starch-sugar interconversion in *Solanum tuberosum*. Phytochemistry, **12**: 2579–2591.

Lafta, A.M., and Lorenzen, J.H. 1995. Effect of high temperature on plant growth and carbohydrate metabolism in potato. Plant Physiol. **109**: 637–643.

Ronning, C.M., Stegalkina, S., Ascenzi, R., Bougri, O., Hart, A., Utterback, T., Vanaken, S., Reidmuller, S., White, J., Cho, J., Karamycheva, S., Pertea, G.M., Lee, Y., Karamycheva, S., Sultana, R., Tsai, J., Quackenbush, J., Griffiths, H.M., Restrepo, S., Smart, C.D., Fry, W.E., van der Hoeven, R., Tanksley, S., Zhang, P., Jin, H., Yamamoto, M., Baker, B.J., and Buell, C.R. 2003. Comparative analyses of potato expressed sequence tag libraries. Plant Physiology, **131**: 419–429.

Sano, T., and Nagata, T. 2002. The possible involvement of a phosphate-induced transcription factor encoded by *phi-2* gene from tobacco in ABA-signaling pathways. Plant Cell Physiol. **43**: 12–20.

Sano, T., Kuraya, Y., Amino, S., and Nagata, T. 1999. Phosphate as a limiting factor for the cell division of tobacco BY-2 cells. Plant Cell Physiol. **40**: 1–8.

Schinozaki, K., Yamaguchi-Schinozaki, K., and Seki, M. 2003. Regulatory network of gene expression in the drought and cold stress responses. Curr. Opin. Plant Biol. **6**: 410–417.

Schwimmer, S., Hendel, C.E., Harrington, W.O., and Olson, R.L. 1957. Interrelation among measurements of browning of processed potatoes and sugar components. Am. Potato J. **34**: 119–132.

Singh, K., Foley, R.C., and Onate-Sanchez, L. 2002. Transcription factors in plant defense and stress responses. Curr. Opin. Plant Biol. **5**: 430–436.

The Gene Ontology Consortium. 2000. Gene ontology: tool for the unification of biology. Nature Genetics, **25**: 25–29.

Veilleux, R.E., Paz, M.M., and Levy, D. 1997. Potato germplasm development for warm climates: genetic enhancement of tolerance to heat stress. Euphytica, **98**: 83–92.

Wortman, J.R., Haas, B.J., Hannick, L.I., Smith, R.K., Maiti, R., Ronning, C.M., Chan, A.P., Yu, C., Ayele, M., Whitelaw, C.A., et al. 2003. Annotation of the *Arabidopsis* genome. Plant Physiology, **132**: 461–468.

Yamada, S., Komori, T., Hashimoto, A., Kuwata, S., Imaseki, H., and Kubo, T. 2000. Differential expression of plastidic aldolase genes in *Nicotiana* plants under salt stress. Plant Science, **154**: 61–69.

Zamir, D., and Tanksley, S.D. 1988. Tomato genome is comprised of fast-evolving, low-copy-number sequences. Mol. Gen. Genet. **213**: 254–261.

Zhu, J.-K. 2001. Plant cell tolerance. Trends Plant Sci. **6**: 66–71.

Zhu, B., Chen, T.H., and Li, P.H. 1995. Expression of three osmotin-like protein genes in response to osmotic stress and fungal infection in potato. Plant Mol. Biol. **28**: 17–26.