Journal of Plant Biology, March 2005, 48(1) : 85-95

Low Expression Profiles of Heat Stress-Related Genes in *Capsicum annuum*

M. Ashrafuzzaman, S. June Oh[1], and Choo Bong Hong*
Department of Biological Sciences and Institute of Molecular Biology and Genetics,
Seoul National University, Seoul 151-742, Korea
[1]Department of Pharmacology, Pharmacogenomics Research Center, College of Medicine,
Inje University, Busan 614-735, Korea

A cDNA library was constructed for hot pepper plants that had been heat-shock-treated. We used a modified differential screening method, double negative screening, to isolate 500 cDNA clones that represented genes with low expression levels under conditions of high-temperature stress. Of those 500 clones, 200 were randomly selected for single-read sequencing from the 5' ends. After annotation with Blastx, the sequence was applied to InterProScan to scan for functional motifs of proteins. Among the cDNA clones analyzed, about 41% of the ESTs could not be functionally classified. However, of those that could be, the largest portion of the ESTs (15%) were assigned to the category of cell rescue and defense; genes involved in cell cycle/DNA processing constituted the smallest group, comprising 1% of the ESTs. Genes related to energy and protein fates constituted the second (10%) and third (9%) largest groups, respectively. Finally, 3% of the ESTs were assigned to transcription, and 2% to signal transduction. The high portion of unclassified ESTs probably resulted from the screening method, which was designed for low-expression messages. Likewise, the high number of ESTs for cell rescue and defense suggests that many genes with low levels of expression are associated with the stress response.

Keywords: cDNA library, double negative screening, high-temperature stress, hot pepper (*Capsicum annuum*)

Drought, salinity, extreme temperatures, and oxidative stress are often interconnected, and may induce similar cellular damage (Wang et al., 2003). The ability of plants to switch on signal transduction and adaptive responses to abiotic stresses is a critical step in determining survival and reproduction in adverse environments. Numerous stress-specific reactions as well as a sophisticated and complex network of adaptation all help to protect these plants from multiple environmental stresses (Chinnusamy et al., 2004).

Gene activation depends on a distinct set of transcription factors, whose expression imparts stress tolerance while leading to various physiological and biochemical actions. These transcription factors produce multiple phenotypic alterations, many of which are involved in stress responses (Pellegrineschi et al., 2001). Although individual member of the same transcription factor family often responds differently to various stimuli, some stress-responsive genes may share the same transcription factors, as indicated by significant overlap in their induced expression profiles (Bohnert et al., 2001; Seki et al., 2001; Chen et al., 2002; Fowler and Thomashow, 2002; Kreps et al., 2002). However, another level in the hierarchy of genetic control may have important bearing in regulating stress responses, namely, the components of signal transduction. Nevertheless, studies have been limited with regard to either signal transduction mechanisms in plants or abiotic stress interactions.

Genomics approaches have greatly facilitated the discovery of relevant plant genes (Asamizu et al., 2000; Ohlrogge and Benning, 2000; van der Hoeven et al., 2002). Large-scale cDNA sequencing projects have identified expressed sequence tags (ESTs), including those from wheat, maize, barley, soybean, *Arabidopsis*, rice, sugarcane, potato, tomato, tobacco, and hot pepper, all of which are now catalogued in the EST division (dbEST) of GenBank (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html). Out of a total of 3,148,436 entries, *Capsicum annuum* is represented by 30,149 entries. Ideally, the ESTs generated from cDNA libraries should denote all the genes expressed in a target organ and tissue, at a particular developmental stage, and/or in a specific environment. However, variations in expression levels among genes from a given tissue type yield mRNAs that differ in their abundance, making it difficult to capture rare mRNA in cDNA libraries. Therefore, to identify those

*Corresponding author; fax +82-2-875-1157
e-mail hcb@snu.ac.kr

rare genes by this approach, it is necessary to either sequence all the clones in the library or else prepare a normalized library. In theory, the latter approach would generate uniform abundances of cDNA classes. However, this requires large-scale sequencing of total cDNA libraries, which is accompanied by high costs for both resources and labor (Bonaldo et al., 1996; Carson and Botha, 2000; Reddy et al., 2002; Fernández et al., 2003).

Heat stress is one of the most serious constraints on crop production in tropical and subtropical regions. The severity of the situation, at least locally, worsens each year, probably because of global warming. Exposure to high temperatures causes an intricate set of changes in plant gene expression that can induce thermo-tolerance and cellular survival. This is known as the heat-shock response. At the molecular level, this reaction is a temporary re-programming of cellular activities, which involves the synthesis of heat-shock proteins (HSPs) and simultaneous cessation of normal protein synthesis (Chen et al., 2002).

To isolate the high-temperature stress-specific rare mRNA species in hot pepper from a relatively small number of analyzed sequences, we used a modified screening method -- double negative screening -- adopted from Cho et al. (2003). A cDNA library was constructed from hot pepper treated with physiologically high temperature. Through double negative screening of that cDNA library, we were able to isolate clones showing weak or negative signals by cDNA probes generated from heat-treated as well as normally grown plants. After the selected clones were sequenced, we used annotations of *Arabidopsis thaliana* as comparisons for putatively characterizing those rare transcript species from stressed hot pepper plants.

MATERIALS AND METHODS

Plant Material

Hot pepper (*C. annuum* cv. Bugang) plants were grown for 4 weeks (or to a height of 5 to 7 cm) in a chamber (16-h photoperiod, 25°C, 60% relative humidity, and 200 $\mu E/m^2s$ from white fluorescent lamps). The high-temperature treatment was applied by placing the pots in an incubator at 42°C for 30 min, or 1, 2, 3, or 4 h. Control plants were maintained at 25°C. Afterward, the collected seedlings were quickly frozen in liquid nitrogen and stored at –80°C.

Preparation of Poly(A)$^+$ RNA, Construction of cDNA Library, and Double Negative Screening

Total RNA was extracted from the hot pepper seedlings according to the method of Sambrook et al. (1989). Poly(A)$^+$RNA was purified via the PolyATtract mRNA isolation system (Promega, USA). From the purified poly(A)$^+$RNA, we constructed a unidirectional *Eco*RI/*Xho*I cDNA library, using a ZAP-cDNA synthesis kit and ZAP-cDNA GigapackIII gold packaging extracts (Stratagene, USA), according to the manufacturer's instructions. This cDNA library had a complexity of 2.2×10^7 pfu ml^{-1}, and >95% of the phages contained cDNA inserts (data not shown). For the double negative screening, after amplification and titering, duplicate plaques were lifted onto nylon membranes (Hybond-N, Amersham, USA) from 150-mm plates (about 5×10^3 pfu per plate). The plaque lifts were denatured, neutralized, rinsed, blotted, dried, and fixed by UV cross linking. Membranes were prehybridized, then hybridized with two probes for differential screening. Those probes were made using poly(A)$^+$RNA from either heat-shocked or unstressed hot pepper plants to the ^{32}P-labeled cDNA. The hybridized membranes were washed and exposed to X-ray film (Fuji, Japan). All the procedures described above essentially followed those of Sambrook et al (1989) and manufacturer's instructions. Five hundred clones were randomly picked up by precisely aligning both autoradiographs of the duplicate lifts that had reacted to the differential screening probes with the plate containing the original plaques. Isolated cDNA clones were *in vivo*-excised from the UniZAP-XR vector and subcloned into pBluescript SK(-) using the Exassist helper phage (M13, Stratagene).

Nucleotide Sequencing, Sequence Processing, and Functional Classification

Plasmid DNA was purified with an AccuPrep plasmid extrction kit (Bioneer, Korea). Sequence reactions were run on an automated sequencer (Model 3100; Applied Biosystems, USA). The nucleic acid sequences obtained for each cDNA clone were then converted into amino acid sequences for six different reading frames. Finally, we performed database searches at the National Center of Biotechnology Information server (http://www.ncbi.nlm.nih.gov) with Entrez, BLAST (Altschul et al., 1997). Sequences smaller than 200 nucleotides were not further processed. To assign the functioning of ESTs, hot pepper sequences were aligned to the GenBank nucleotide sequence data-

base using the BLASTX algorithm for comparison, with an E-value cut off at 10^{-5} or lower.

RESULTS AND DISCUSSION

cDNA Library of High Temperature-Stressed Hot Pepper

While accessing the transcription factors and the components of the signal transduction pathway in hot pepper plants, we were especially interested in ESTs with low expression profiles during high-temperature stress. In all, 500 clones were selected that showed weak or negative signals by the two cDNA probes synthesized from both treated and unstressed seedlings. From these, we randomly chose 215 single pass-sequenced cDNA clones that then gave rise to 200 high-quality ESTs. After vector-trimming and removal of low-quality sequences, the average read length of these ESTs was over 400 nucleotides; the average insert size of the corresponding ESTs was about 800 bp.

Functional Categorization of Hot Pepper ESTs

Among the 200 cDNA clones analyzed, approximately 59% of the ESTs were assigned a function by aligning them with the translated sequences of the GenBank nucleotide sequence database; the remaining 41% encoded proteins with insufficient similarity to proteins of known function for us to confidently assign any role other than "unknown". Genes of known function were sorted into 11 primary categories (Fig. 1). The largest set of genes (15%) was assigned to the category of cell defense or rescue, while those involved in cell cycling and DNA processing constituted the smallest group, or 1%. Genes that helped determine energy and protein fates formed the second (10%) and third (9%) largest groups, respectively. The fact that more genes were placed in the cell defense or rescue category suggests that hot pepper plants possess the ability to survive heat stress in their environment.

Of the 200 ESTs from the non-normalized hot pepper cDNA library, 169 were unique while 31 could be considered redundant. That is, 14 might have been represented twice and one, three times, although the redundancy here was presumptive due to the incomplete nature of the nucleotide sequences. ESTs were defined as redundant when they gave a BLASTX or BLASTN hit to the same accession number, or when

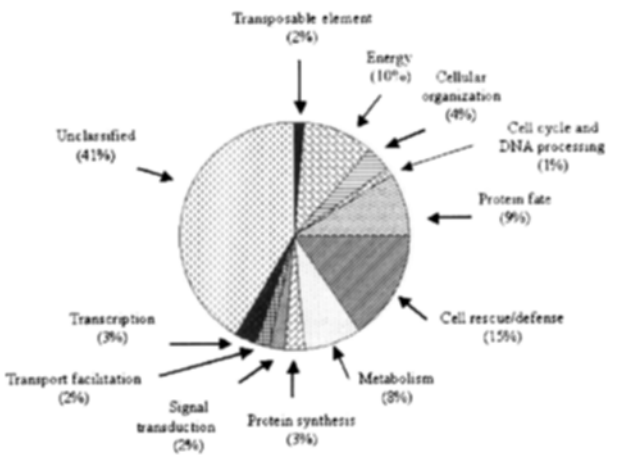

Figure 1. Pie chart showing fraction of high temperature stress-modulated genes in each functional category for cDNA library constructed from heat-stressed tissues of hot pepper.

they exhibited more than 95% identity over aligned regions and were assembled in a single contig. Overall, such redundancy in our cDNA library was less than about 15.5%, which is even lower than a normalized cDNA library when EST population size is considered (Carninci et al., 2000). Although oligonucleotide fingerprint (OFP) normalized cDNA library had very low redundancy and many powerful advantages, it cannot be adapted to small-scale EST research because of its high cost and technical intricacy (Clark et al., 2001). Our data also suggested that stressed plants possess a greater diversity of transcripts, likely because of the increased abundance of various transcripts that encode stress-adaptive determinants. These results are consistent with those previously reported for transcript profiling of salinity-stress responses by large-scale EST analysis (Kore-eda et al., 2004).

Predicted gene functioning, scores, e-values, and accession numbers are summarized in Table 1. The largest group, i.e., those classified as cell rescue and defense genes, are mainly small heat-shock proteins that function as ATP-independent chaperones to prevent irreversible protein aggregation and facilitate subsequent protein renaturation in cooperation with ATP-dependent chaperones (Basha et al., 2004). Plant sHSPs respond to a wide range of environmental stimuli, including heat, cold, drought, salinity, and oxidative stresses. Increasing data suggest a strong correlation between sHSP accumulation and stress tolerance (Park and Hong, 2001; Sun et al., 2002; Wang et al., 2003, 2004). These HSPs likely protect a large set of proteins, with diverse cellular functions, against heat-induced damage. Other cell defense-related proteins

include the dehydrin-like protein (Borovskii et al., 2002), the Bax inhibitor (Huckelhoven, 2004), and Prf (Salmeron et al., 1996), as well as proteins for the antioxidant response (such as glutathione S-transferase), their encoding genes are also present in our collection.

Genes involved in energy functioning constitute the second largest group among our classified ESTs, and comprise those related to photosynthesis. It is well known that the photosynthetic systems in higher plants are most sensitive to high temperatures, or drought and salt treatments (Falk et al., 1996; Tezara et al., 1999; Seki et al., 2002; Kore-eda et al., 2004). Following exposure to abiotic stresses, ESTs that encode large and small subunits of rubisco, chlorophyll a/b binding proteins of the photosystem I and II light-harvesting complexes, photosystem I and II reaction center subunits, oxygen evolving complexes, and components of the mitochondrial ATP synthase and plastidic photosynthetic electron transport chain complex components showed marked declines in abundance. As our approach is designed for the detecting low-abundance genes, it seems to be reasonable to have these photosynthesis-related genes from double negative screening.

Genes in the protein fate category have a role in moving, modifying, storing, and degrading proteins; they constitute our third largest group, with nearly half of them being involved in proteolysis. This proteolysis of regulatory proteins is a key aspect of cellular regulation in eukaryotes (Wei et al., 1994; del Pozo and Estelle, 1999; Schaller, 2004). Thus, the ubiquitin-proteasome pathway is significantly implicated in the plant defense response (Azevedo et al., 2002; Liu et al., 2002). Directly related to this, we identified from our isolated cDNA collection, several regulatory subunits of that pathway (Table 1).

Genes involved in metabolism make up our smallest group of classified ESTs. These include many enzymes, especially those related to photosynthesis, e.g., rubisco activase, sedoheptulose-1,7-isophosphatase, and PEP carboxy kinase; cellular metabolism-related enzymes like malate dehydrogenase; and carbohydrate metabolism-related enzymes, such as beta-amylases, beta-galactosidases, and glyceraldehydes-3-phosphate dehydrogenase. As found while profiling via a microarray approach, most of these enzymes decrease their abundance of transcripts under abiotic stress (Seki et al., 2002). While 3% of our classified ESTs encode for elongation factors and 40S ribosomal proteins that are required for protein synthesis, 2% are considered transposable elements, including a gag-pol poly protein.

Genes for a variety of transcription factors that contain typical DNA binding motifs, such as MYB, AP2/ERF, ring-finger protein, and zinc fingers, have been reported as stress-inducible in rice and *Arabidopsis* (Shinozaki and Yamaguchi-Shinozaki, 1999; Rabbani et al., 2003). We also identified ESTs that code for these putative DNA binding motifs, namely two ring zinc finger family members, one MYB family member, one AP2/ERF, one PHD finger family member, and one bromodomain-containing protein. The MYB family is one of the largest families of transcriptional factors characterized in plants. MYB-related transcriptional activators are involved in regulating secondary metabolism, cellular morphogenesis, meristem formation, and the cell cycle, especially ABA- and gibberellin (GA)-signaling during seed germination (Jin and Martin, 1999; Gubler et al., 2002). Zinc finger proteins also can act as transcription factors (Margolin et al., 1994; Cook et al., 1999). The limited data from 200 ESTs in our non-normalized high temperature-stressed cDNA library also revealed expression from one protein kinase gene, one receptor-like protein kinase gene, one adenylate protein kinase gene, and one ADP-ribosylation factor 1. These are classified as components of cellular communication and signal transduction. Regulatory proteins in rice further control various functional genes under stress conditions, for example, adenylate kinase as an abiotic stress-responsive gene, or a receptor-like protein kinase gene (Rabbani et al., 2003).

Many genes with homologues in other organisms that have not been assigned a function (designated as "unclear classification" in Table 1) and have no recognizable homologue in any other organisms (designated as "unclassified" in Table 1) were highly represented in our EST collection. Although most of their sequences currently correspond to hypothetical proteins with unknown function, it will be important to determine their potential roles because these genes may be making unique contributions to thermo-tolerance in plants.

General Connections between Different Abiotic Stresses

Several of our ESTs show significant sequence similarity to genes affected by ABA, drought, and other environmental stresses. In fact, a significant number are up-regulated by drought, cold, or high-salinity conditions in *Arabidopsis* and rice, as demonstrated from previous microarray data (Kawasaki et al., 2001;

Table 1. Functional annotation of 200 ESTs from heat-stressed hot pepper.

EST ID	Description of best data match plus database accession number for homologous gene	Score	E-value
	Metabolism		
HTS10	gi|10720247|sp|O49074|RCA_LYCPN rubisco activase	485	e-136
HTS27	gi|3024121|sp|P93254| S-adenosylmethionine synthetase	372	e-102
HTS35	gi|22329419|ref|NP_172325.2| cysteine desulfurase, putative	127	3e-28
HTS55	gi|15229546|ref|NP_189036.1| dehydratase family	393	e-10
HTS147	gi|16950587|gb|AAG01894.2| PEP carboxykinase	392	e-108
HTS275	gi|7431231|pir||T06401 malate dehydrogenase	319	3e-86
HTS367	gi|12081917|dbj|BAB20861.1| cytosolic cysteine synthase	372	e-106
HTS447	gi|7431231|pir||T06401 malate dehydrogenase	397	e-109
HTS140	gi|15233426|ref|NP_193819.1| fatty acid hydroxylase, putative	248	1e-64
HTS205	gi|22329419|ref|NP_172325.2| cysteine desulfurase, putative	127	2e-35
HTS78	gi|27804768|gb|AAO22558.1| sedoheptulose-1,7-isphosphatase.	405	e-112
HTS385	gi|120676|sp|P09094 glyceraldehyde3 phosphate dehydrogenase	472	e-132
HTS375	gi|2071947|gb|AAB53629.1| beta-galactosidase	135	2e-30
HTS390	gi|5031285|gb|AAD38148.1| beta-amylase [*Prunus armeniaca*]	361	2e-98
HTS217	gi|18694346|emb|CAC85287.1| uroporphyrinogen III synthase	233	6e-60
HTS255	gi|10177106|dbj|BAB10440.1| pyruvate kinase [*Arabidopsis thaliana*]	221	3e-56
	Energy		
HTS11	gi|34921349|sp|Q9ZTS2| ferredoxin, chloroplast precursor	256	7e-67
HTS12	gi|34921349|sp|Q9ZTS2| ferredoxin, chloroplast precursor	233	4e-60
HTS21	gi|6899972|emb|CAB71293.1| chloroplast ferredoxin-NADP+	535	e-151
HTS29	gi|34787117|emb|CAD89270.1| putative PSI-D subunit precursor	369	e-101
HTS38	gi|131166|sp|P12372| PSI reaction center subunit II	348	1e-94
HTS45	gi|28629385|gb|AAO49652.1| Photosystem I-N subunit	175	1e-42
HTS46	gi|115781|sp|P27492|CB21_TOBAC CAB-16	485	e-136
HTS47	gi|34921349|sp|Q9ZTS2| ferredoxin, chloroplast precursor	246	4e-64
HTS68	gi|100196|pir||S14305 CAB-11	454	e-126
HTS74	gi|82078|pir||S00443 CAB-6A	433	e-120
HTS76	gi|3036951|dbj|BAA25394.1| light-harvesting CAB protein	365	e-107
HTS177	gi|12585428|sp|O82702|VAG1_TOBAC V-ATPase G subunit 1	56	2e-15
HTS270	gi|34787117|emb|CAD89270.1| PSI-D subunit precursor	383	e-105
HTS277	gi|6093830|sp|P80470|PSBY_SPIOL PSII core complex protein	110	3e-23
HTS327	gi|7489133|pir||T01782 GDP dissociation inhibitor	442	e-129
HTS335	gi|115803|sp|P14278|CB24_LYCES CAB-4	406	e-112
HTS410	gi|693920|gb|AAA80593.1| chlorophyll a/b binding protein	424	e-117
HTS427	gi|3036944|dbj|BAA25389.1| light-harvesting CAB	424	e-117
HTS457	gi|130271|sp|P17340|PLAS_LYCES plastocyanin, chloroplast pr	222	1e-56
HTS66	gi|48927500|emb|CAA61241.2| glycogen (starch) synthase	535	e-151
	Protein synthesis		
HTS25	gi|119150|sp|P17786|EF1A_LYCES elongation factor 1-alpha	412	e-121
HTS48	gi|18391048|ref|NP_563848.1| elongation factor 1B-gamma	82	1e-14
HTS51	gi|6984222|gb|AAF34799.1| 40S ribosomal protein S16	270	3e-71
HTS437	gi|20139798|sp|Q9LTF2|R103_ARATH 40S ribosomal protein S	186	7e-46
HTS357	gi|29892963|emb|CAD60652.1| elongation factor	202	7e-59
HTS227	gi|1076678|pir||S42643 40S ribosomal protein S17	149	5e-35

Table 1. (continued).

EST ID	Description of best data match plus database accession number for homologous gene	Score	E-value
	Protein fate (folding, modification, destination)		
HTS16	gi|15217581|ref|NP_174619.1| serine carboxypeptidase S10	310	4e-83
HTS20	gi|1332579|emb|CAA66667.1| polyubiquitin [*Pinus sylvestris*]	435	e-127
HTS61	gi|6671192|gb|AAF23126.1| cystatin [*Lycopersicon esculentum*]	366	e-101
HTS170	gi|118103|sp|P21568| peptidyl-prolyl cis-trans-trans isomerase	307	2e-82
HTS370	gi|1351030|sp|P21239| RUBISCO subunit binding protein alpha	368	e-101
HTS83	gi|1076678|pir||S42643 ubiquitin / ribosomal protein S27a	149	5e-35
HTS167	gi|209078|gb|AAA72893.1| alpha-peptide >gi|3603281|gb|A	70	4e-11
HTS235	gi|50907133|ref|XP_465055.1 putative ubiquitin ligase SINAT5	213	2e-70
HTS450	gi|118103|sp|P21568| peptidyl-prolyl cis-trans-trans isomerase	307	2e-82
HTS307	gi|18394416|ref|NP_564011.1| ubiquitin-conjugating enzyme	310	3e-83
HTS467	gi|15217581|ref|NP_174619.1| serine carboxypeptidase S10	280	2e-75
HTS495	gi|30693180|ref|NP_849748.1| ATP-dependent protease La	124	3e-27
HTS300	gi|18417611|ref|NP_568311.1| FtsH protease, putative	437	e-121
HTS435	gi|15076665|dbj|BAB62328.1| cyclophilin	268	2e-70
HTS230	gi|18076679|emb|CAC84774.1| P70 protein [*Nicotiana tabacum*]	254	3e-66
HTS475	gi|18394416|ref|NP_564011.1| ubiquitin-conjugating enzyme	310	3e-83
HTS430	gi|15217581|ref|NP_174619.1| serine carboxypeptidase S10	310	4e-83
HTS67	gi|18417611|ref|NP_568311.1| FtsH protease, putative	437	e-121
	Cell rescue, defense, and virulence		
HTS1	gi|3341464|emb|CAA12387.1| Hsp20.1 protein	92	2e-17
HTS3	gi|30697614|ref|NP_201008.2| heat-shock factor protein, putative	42	0.027
HTS4	gi|1708314|sp|P51819|HS83_PHANI heat-shock protein 83	444	e-126
HTS5	gi|27447206|gb|AAL92873.1| GST-like protein	381	e-104
HTS7	gi|30699467|ref|NP_178110.3| HSP 70, putative	391	e-108
HTS15	gi|7441328|pir||T07602 heat-shock protein 17.6 tomato	262	7e-69
HTS18	gi|38154485|gb|AAR12194.1| molecular chaperone Hsp90-2	431	e-119
HTS19	gi|37904866|gb|AAP57477.1| small heat-shock protein	367	e-100
HTS33	gi|2071947|gb|AAB53629.1| beta-galactosidase	135	2e-30
HTS39	gi|7441310|pir||T04316 heat-shock protein MTSHP precursor	221	2e-56
HTS50	gi|8547237|gb|AAF76312.1| Prf [*Lycopersicon esculentum*]	377	e-103
HTS53	gi|48474196|dbj|BAD22699.1| heat-shock protein 70	190	1e-6
HTS54	gi|10798648|emb|CAC12824.1| putative DNAJ protein	233	5e-61
HTS57	gi|7441328|pir||T07602 heat-shock protein 17.6	229	6e-59
HTS58	gi|7443855|pir||T07733 chaperonin 60 beta chain precursor	417	e-115
HTS70	gi|27447206|gb|AAL92873.1| GST-like protein	335	6e-91
HTS82	gi|100335|pir||S18181 dnaK-type molecular chaperone Nthsp70	350	3e-95
HTS190	gi|7443855|pir||T07733 chaperonin 60 beta chain precursor	279	2e-76
HTS195	gi|33591104|gb|AAQ23059.1| heat-shock factor RHSF5	232	8e-60
HTS250	gi|10697184|dbj|BAB16318.1| chaperonin-60 alpha subunit	367	e-101
HTS297	gi|28973653|gb|AAO64147.1| putative TPR-repeat protein	302	8e-81
HTS317	gi|38154489|gb|AAR12195.1| molecular chaperone Hsp90-1	419	e-116
HTS320	gi|7441310|pir||T04316 heat-shock protein MTSHP precursor	207	3e-52
HTS337	gi|3341464|emb|CAA12387.1| Hsp20.1 protein	268	1e-70
HTS135	gi|39579116|gb|AAR28754.1| Bax inhibitor	385	e-106
HTS377	gi|19813|emb|CAA42660.1| luminal binding protein (BiP)	276	3e-91
HTS400	gi|15219028|ref|NP_175665.1| 26.5 kDa class I sHSP	172	8e-42
HTS345	gi|7447302|pir||T06239 probable glutathione transferase	328	1e-88
HTS460	gi|10697184|dbj|BAB16318.1| chaperonin-60 alpha subunit	359	5e-98
HTS470	gi|37904866|gb|AAP57477.1| small heat-shock protein	360	2e-98
HTS487	gi|37905913|gb|AAO38853.1| dehydrin-like protein	183	5e-45

Table 1. (continued).

EST ID	Description of best data match plus database accession number for homologous gene	Score	E-value
	Transcription		
HTS9	gi|15224062|ref|NP_179958.1| zinc finger(C3HC4-type RING)	238	1e-61
HTS187	gi|30690620|ref|NP_850470.1| myb family transcription factor	197	1e-49
HTS397	gi|6862918|gb|AAF30307.1| putative RING zinc finger protein	146	7e-34
HTS14	gi|42561972|gb|AAS20427.1| ethylene-responsive factor-like	397	e-115
HTS185	gi|42563295|ref|NP_177903.4| PHD finger family protein	257	2e-67
HTS245	gi|55297001|dbj|BAD68476.1|bromodomain-containing protein	274	3e-72
	Transport facilitation		
HTS350	dbj|BAB40141.1|(AB035272)|plasmamembrane intrinsic protein	130	2e-11
HTS387	gi|12006187|gb|AAG44776.1| biotin carboxyl carrier protein	65	1e-11
HTS265	gi|50254435|gb|EAL17184.1| hypothetical protein CNBN0130	133	2e-33
HTS425	gi|6996562|emb|CAB75430.1|putative 16kDa membrane protein	250	3e-65
	Cellular communication and signal transduction		
HTS157	gi|30679085|ref|NP_195722.2| protein kinase family protein	190	4e-47
HTS13	gi|29376496|ref|NP_815650.1| GTP pyrophosphokinase	35	3.3
HTS445	gi|15235152|ref|NP_195118.1| protein phosphatase 2C family	234	9e-61
HTS480	gi|17104725|gb|AAL34251.1|putative ADP ribosylation factor	238	2e-65
	Cellular organization		
HTS43	gi|5327263|emb|CAB46351.1| aquaporin	432	e-120
HTS210	gi|3163946|emb|CAA06619.1| alpha-tubulin 1 [*Eleusine indica*]	390	e-124
HTS340	gi|8928432|sp|Q9ZRR5|TBA3_HORVU tubulin alpha-3 chain	426	e-131
HTS285	sp|P30175| actin-depolymerization (ADF)	210	2e-61
HTS325	sp|Po9469| vacuolar ATP synthase catalytic subunit A	183	6e-4
HTS227	gb|AAK96884.1|(AY009094) beta tubulin	146	7e-34
HTS450	gi|8928432|sp|Q9ZRR5|TBA3_HORVU tubulin alpha-3 chain	426	e-131
	Transposable elements; viral and plasmid proteins		
HTS355	gi|47825015|gb|AAT38786.1| putative gag-pol polyprotein	92	1e-32
HTS290	gi|7269781|emb|CAB77781.1|polyprotein of LTR transposon	173	6e-49
HTS6	gi|53689713|gb|AAU89728.1| retroelement pol polyprotein	156	3e-62
	Cell cycle and DNA processing		
HTS81	gi|22326839|ref|NP_197131.2| Rad21/Rec8-like family protein	174	4e-42
HTS260	gi|42408051|dbj|BAD09193.1| putative PrMC3 [*Oryza sativa*]	182	8e-45
	Classification not yet clear-cut (unknown)		
HTS22	gi|11994706|dbj|BAB02944.1| unnamed protein product	203	3e-84
HTS23	gi|30683790|ref|NP_567512.2| exostosin family protein	162	3e-08
HTS24	gi|18415962|ref|NP_568211.1| expressed protein	228	2e-58
HTS26	gi|21536807|gb|AAM61139.1| unknown	174	2e-12
HTS28	gi|8919877|emb|CAB96200.1| hypothetical protein	185	2e-15
HTS31	gi|32352168|dbj|BAC78577.1| hypothetical protein	161	2e-38
HTS330	gi|25513587|pir||E86289 T16N11.7 protein [*A. thaliana*]	136	3e-41
HTS34	gi|18402380|ref|NP_565700.1| expressed protein	133	3e-56
HTS59	gi|42570879|ref|NP_973513.1| expressed protein	189	2e-16
HTS64	gi|18405124|ref|NP_565909.1| radical SAM domain-containing	135	1e-30
HTS71	gi|18396280|ref|NP_564274.1| expressed protein	109	1e-22
HTS79	gi|15220924|ref|NP_176682.1| expressed protein [*A. thaliana*]	269	6e-71
HTS80	gi|27808548|gb|AAO24554.1| At1g61150 [*A. thaliana*]	342	6e-93
HTS84	gi|15234433|ref|NP_195371.1|hydrolase, alpha/beta fold protein	286	5e-76

Table 1. (continued).

EST ID	Description of best data match plus database accession number for homologous gene	Score	E-value
HTS115	gi|584825|sp|P37707|B2_DAUCA B2 PROTEIN	120	8e-27
HTS117	gi|17065046|gb|AAL32677.1| unknown protein [*A. thaliana*]	360	2e-98
HTS127	gi|28874734|emb|CAC80137.1| progesterone 5-beta-reductase	259	4e-76
HTS207	gi|21593353|gb|AAM65302.1| unknown (*A. thaliana*)	300	3e-82
HTS237	gi|38344078|emb|CAE01738.2| OSJNBb0056F09.1	85	2e-15
HTS257	gi|15239148|ref|NP_196729.1| expressed protein [*A. thaliana*]	329	6e-89
HTS240	gi|6322674|ref|NP_012747.1| Ykl174cp	135	6e-23
HTS417	gi|9294516|dbj|BAB02778.1| unnamed protein product	325	1e-87
HTS455	gi|18396859|ref|NP_566224.1| expressed protein [*A. thaliana*]	232	1e-59
HTS477	gi|18398123|ref|NP_565388.1| expressed protein [*A. thaliana*]	195	7e-57
HTS485	gi|26452567|dbj|BAC43368.1| unknown protein (*A. thaliana*)	382	e-105
HTS365	gi|40736997|gb|AAR89010.1| expressed protein [*O. sativa*]	221	3e-59
HTS95	gi|21554147|gb|AAM63227.1| unknown [*A. thaliana*]	311	3e-96
HTS120	gi|18419954|ref|NP_568013.1| expressed protein [*A. thaliana*]	246	4e-64
HTS130	gi|15218701|ref|NP_171806.1| expressed protein [*A. thaliana*]	323	2e-87
	Unclassified protein and no hit		
HTS2	No significant similarity found.		
HTS100	gi|38603804|gb|AAR24647.1| At2g23330 (*A. thaliana*)	155	3e-61
HTS17	gi|15224172|ref|NP_179434.1| expressed protein	135	5e-53
HTS30	gi|34849885|gb|AAQ82839.1| At4g04790	133	4e-47
HTS37	gi|50080326|gb|AAT69660.1| unknown protein	146	2e-32
HTS40	gi|46911551|emb|CAG27615.1| putative leucine-rich repeat	141	1e-04
HTS41	gi|34902002|ref|NP_912347.1|ser/arg rich ribonucleoprotein	50	2e-10
HTS44	gi|18394220|ref|NP_563969.1| expressed protein	165	2e-09
HTS49	gi|31209269|ref|XP_313601.1| ENSANGP00000013268	142	2e-15
HTS52	No significant similarity found.		
HTS72	gi|6322674|ref|NP_012747.1| hypothetical ORF; Ykl174cp	135	2e-54
HTS73	gi|28373837|pdb|1N0R|A designed Ankyrin repeat protein	179	2e-13
HTS197	gi|47497554|dbj|BAD19626.1| hypothetical protein	133	1e-65
HTS200	gi|15239148|ref|NP_196729.1| expressed protein [*A. thaliana*]	329	6e-89
HTS215	gi|30694498|ref|NP_175343.2| expressed protein [*A. thaliana*]	172	9e-42
HTS247	No significant similarity found		
HTS280	gi|15642295|ref|NP_231928.1| conserved hypothetical protein	135	1e-77
HTS287	gi|5596352|dbj|BAA82607.1| sALK-7 [*Ephydatia fluviatilis*]	133	1e-08
HTS295	gi|42568382|ref|NP_199594.2| expressed protein [*A. thaliana*]	147	1e-04
HTS310	gi|46092516|dbj|BAD14378.1| hypothetical protein	131	2e-29
HTS315	gi|15233196|ref|NP_191076.1| expressed protein [*A. thaliana*]	129	1e-28
HTS347	gi|21428994|gb|AAM50216.1| GM13228p	136	1e-54
HTS440	gi|31211851|ref|XP_314910.1| ENSANGP00000012448	134	2e-77
HTS465	No significant similarity found		
HTS60	gi|50290053|ref|XP_447458.1| unnamed protein product	133	3e-28
HTS490	No significant similarity found		
HTS69	gi|25518436|pir||C86390 hypothetical protein T1K7.26	138	2e-31
HTS75	gi|50290053|ref|XP_447458.1| unnamed protein product	133	2e-24
HTS77	gi|50302195|ref|XP_451031.1| unnamed protein product	139	1e-77
HTS90	gi|42567155|ref|NP_194332.2| expressed protein [*A. thaliana*]	122	1e-26
HTS125	gi|18394407|ref|NP_564007.1| expressed protein [*A. thaliana*]	180	6e-14
HTS145	gi|25345695|pir||E86367 protein F26F24.22 [imported]	150	2e-09
HTS155	gi|9294484|dbj|BAB02703.1| unnamed protein product	201	4e-74

Table 1. (continued).

EST ID	Description of best data match plus database accession number for homologous gene	Score	E-value
HTS165	gi|15237534|ref|NP_196003.1| expressed protein [*A. thaliana*]	130	1e-12
HTS167	gi|209078|gb|AAA72893.1| alpha-peptide >gi|3603281|gb|A	70	4e-11
HTS62	gi|42567155|ref|NP_194332.2| expressed protein	122	2e-26
HTS225	gi|13172242|gb|AAK14060.1| major latex-like protein	135	1e-30
HTS305	gi|37532246|ref|NP_920425.1| hydroxyproline-rich glycoprotein	276	4e-73
HTS8	gi|30682618|ref|NP_187941.2| expressed protein	268	1e-70
HTS105	gi|20453247|gb|AAM19862.1| AT4g03420/F9H3_4(*A. thaliana*)	267	2e-70
HTS110	gi|42568382|ref|NP_199594.2| expressed protein [*A. thaliana*]	159	1e-37
HTS160	gi|30694478|ref|NP_191218.2| hydroxyproline-rich glycoprotein	156	6e-37
HTS56	gi|12005328|gb|AAG44394.1| unknown [*Hevea brasiliensis*]	186	1e-15
HTS175	gi|34903328|ref|NP_913011.1| unnamed protein product	299	e-106
HTS360	gi|26452567|dbj|BAC43368.1| unknown protein (*A. thaliana*)	344	1e-93
HTS497	gi|21553932|gb|AAM63013.1| unknown [*A. thaliana*]	146	4e-47
HTS180	gi|23198154|gb|AAN15604.1| putative protein [*A. thaliana*]	374	e-102
HTS420	gi|34902002|ref|NP_912347.1| unnamed protein	150	2e-10
HTS32	gi|37202090|gb|AAQ89660.1| At1g70760 (*A. thaliana*)	183	6e-45
HTS267	gi|53850511|gb|AAU95432.1| At1g04635 (*A. thaliana*)	127	2e-28
HTS380	gi|25402836|pir||H86313 protein F2H15.10 [imported]	276	1e-73
HTS395	gi|12642910|gb|AAK00397.1| unknown protein (*A. thaliana*)	107	4e-22
HTS405	gi|20465579|gb|AAM20272.1| unknown protein (*A. thaliana*)	237	3e-61
HTS220	gi|30793919|gb|AAP40412.1| unknown protein [*A. thaliana*]	244	2e-63

Seki et al., 2002). These genes include homologues of glutathione S-transferase, S-adenosylmethionine synthetase, receptor-like protein kinase, protein kinase, and protein phosphatase 2C-like proteins. Similarly, a comparison with gene expression profiles from *Arabidopsis* under drought and cold stresses has also revealed a number of common genes, such as those for ascorbate peroxidase, a glycine-rich protein, and an ethylene-responsive element binding protein (Seki et al., 2001).

Large-scale EST projects are often accompanied by high redundancy; thus, an increased input of effort for novel gene discovery is required. The limited data from 200 ESTs of our cDNA library showed 6 ESTs of putative transcription factors and 4 ESTs of putative signal transduction pathway component genes, numbers that are even higher than previously reported for normalized rice cDNA. Reddy et al. (2002) used a large-scale EST approach to isolate drought-responsive genes from a normalized cDNA library, but could locate only 1% of the ESTs in the categories of transcription factor genes and signal transduction pathway component genes. Thus, our modified differential screening "double negative screening" approach, though using only a relatively small number of analyzed ESTs, seemed to enable us to detect low-abundance transcripts with high similarity to expressed proteins, unknown protein products, and novel sequences.

ACKNOWLEDGEMENTS

This work was supported by a research grant from the Crop Functional Genomic Center of the 21st Century Frontier Research Program to C.B. Hong (CG1434), funded by the Ministry of Science and Technology of Korea. M. Ashrafuzzaman was partially supported by the Doctoral Scholarship for Foreign Student (DSFS) program from Seoul National University, Korea.

Received November 29, 2004; accepted January 3, 2005.

LITERATURE CITED

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

Asamizu E, Nakamura Y, Sato S, Tabata S (2000) A large scale analysis of cDNA in *Arabidopsis thaliana*: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries. DNA Res 7: 175-180

Azevedo C, Sadanandom A, Kitagawa K, Freialdenhoven A, Shirasu K, Schulze-Lefert P (2002) The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance. Science 295: 2073-2076

Basha E, Lee GJ, Breci LA, Hausrath AC, Buan NR, Giese KC, Vierling E (2004) The identity of proteins associated with a small heat shock potein during heat stress *in vivo*

indicates that these chaperones protect a wide range of cellular functions. J Biol Chem 279: 7566-7575

Bohnert HJ, Ayoubi P, Borchert C, Bressan RA, Burnap RL, Cushman JC, Cushman MA, Deyholos M, Fischer R, Galbraith DW (2001) A genomics approach towards salt stress tolerance. Plant Physiol Biochem 39: 295-311

Bonaldo MF, Lennon G, Soares MB (1996) Normalization and subtraction: Two approaches to facilitate gene discovery. Genome Res 6: 791-806

Borovskii GB, Stupnikova IV, Antipina AI, Vladimirova SV, Voinikov VK (2002) Accumulation of dehydrin-like proteins in the mitochondria of cereals in response to cold, freezing, drought and ABA treatment. BMC Plant Biol 2: 5-12

Carninci P, Shibata Y, Hayatsu N, Sugahara Y, Shibata K, Itoh M, Konno H, Okazaki Y, Muramatsu M, Hayashizaki Y (2000) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res 10: 1617-1630

Carson DL, Botha FC (2000) Preliminary analysis of expressed sequence tags for sugarcane. Crop Sci 40: 1769-1779

Chen W, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan SH, Zou G, Whitham SA, Budworth PR, Tao Y, Xie Z, Chen X, Lam S, Kreps JA, Harper JF, Si-Ammour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn T, Dangl JL, Wang X, Zhu T (2002) Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell 14: 559-574

Chinnusamy V, Schumaker K, Zhu JK (2004) Molecular genetic perspectives on cross-talk and specificity in abiotic stress signaling in plants. J Exp Bot 55: 225-236

Cho YB, Jeon HJ, Hong CB (2003) Cloning and functional annotation of rare mRNA species from drought-stressed hot pepper (*Capsicum annuum*). J Plant Biol 46: 83-89

Clark MD, Hennig S, Herwig R, Clifton SW, Marra MA, Lehrach H, Johnson SL, Group tW, WU-GSCnEST Group (2001) An oligonucleotide fingerprint normalized and expressed sequence tag characterized zebrafish cDNA library. Genome Res 11: 1594-1602

Cook T, Gebelein B, Belal M, Mesa K, Urrutia R (1999) Three conserved transcriptional repressor domains are a defining feature of the TIEG subfamily of Sp1-like zinc finger proteins. J Biol Chem 274: 29500-29504

del Pozo JC, Estelle M (1999) Function of the ubiquitin-proteasome pathway in auxin response. Trends Plant Sci 4: 107-112

Falk S, Maxwell DP, Laudenbach DE, Huner NPA (1996) Photosynthetic adjustment to temperature. *In* NR Baker, ed, Photosynthesis and the Environment. Kluwer Academic, Dordrecht, pp 367-385

Fernández P, Paniego N, Lew S, Hopp HE, Heinz RA (2003) Differential representation of sunflower ESTs in enriched organ-specific cDNA libraries in a small scale sequencing project. BMC Genomics 4: 40-49

Fowler S, Thomashow MF (2002) *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14: 1675-1690

Gubler F, Chandler PM, White RG, Llewellyn DJ, Jacobsen JV (2002) Gibberellin signaling in barley aleurone cells. Control of SLN1 and GAMYB expression. Plant Physiol 129: 191-200

Huckelhoven R (2004) Bax inhibitor-1, an ancient cell death suppressor in animals and plants with prokaryotic relatives. Apoptosis 9: 299-307

Jin H, Martin C (1999) Multifunctionality and diversity within the plant MYB-gene family. Plant Mol Biol 41: 577-585

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell 13: 889-905

Kore-eda S, Cushman MA, Akselrod I, Bufford D, Fredrickson M, Clark E, Cushman JC (2004) Transcript profiling of salinity stress responses by large-scale expressed sequence tag analysis in *Mesembryanthemum crystallinum*. Gene 341: 83-92

Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper JF (2002) Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. Plant Physiol 130: 2129-2141

Liu Y, Schiff M, Serino G, Deng XW, Dinesh-Kumar SP (2002) Role of SCF ubiquitin-ligase and the COP9 signalosome in the N gene-mediated resistance response to tobacco mosaic virus. Plant Cell 14:1483-1496

Margolin JF, Friedman JR, Meyer WK, Vissing H, Thiesen HJ, Rauscher FJR (1994) Kruppel-associated boxes are potent transcriptional activation system for regulated gene expression in transgenic plants. Proc Natl Acad Sci USA 91: 4509-4513

Ohlrogge J, Benning C (2000) Unraveling plant metabolism by EST analysis. Curr Opin Plant Biol 3: 224-228

Park SM, Hong CB (2001) Class I small heat-shock protein gives thermotolerance in tobacco. J Plant Physiol 159: 25-30

Pellegrineschi A, Reynolds M, Pacheco M, Brito RM, Almeraya R, Yamaguchi-Peng Z, Staub JM, Serino G, Kwok SF, Kurepa J, Bruce BD, Vierstra RD, Wei N, Deng XW (2001) The cellular level of PR500, a protein complex related to the 19S regulatory particle of the proteasome, is regulated in response to stresses in plants. Mol Biol Cell 12: 383-392

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. Plant Physiol 133: 1755-1767

Reddy AR, Ramakrishna AC, Sekhar AC, Nagablushana I, Ravindra Babu P, Bonaldo MF, Soares MB, Bennetzen

JL (2002) Novel genes are enriched in normalized cDNA libraries from drought-stressed seedlings of rice (*Oryza sativa* L. subsp. *indica* cv. Nagina 22). Genome 45: 204-211

Salmeron JM, Oldroyd GED, Rommens CMT, Scofield SR, Kim H-S, Lavelle DT, Dahlbeck D, Staskawicz BJ (1996) Tomato *Prf* is a member of the leucine-rich-repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. Cell 86: 123-133

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, New York

Schaller A (2004) A cut above the rest: The regulatory function of plant proteases. Planta 220: 183-197

Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K (2001) Monitoring the expression pattern of 1,300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell 13: 61-72

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, Satou M, Akiyama K, Taji T, Yamaguchi-Shinozaki K, Carninci P, Kawai J, Hayashizaki Y, Shinozaki K (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J 31: 279-292

Shinozaki K, Yamaguchi-Shinozaki K (1999) Molecular responses to drought stress. *In* K Shinozaki, K Yamaguchi-Shinozaki, eds, Molecular Responses to Cold, Drought, Heat and Salt Stress in Higher Plants. RG Landes, Austin, pp 11-28

Sun W, van Montagu M, Verbruggen N (2002) Small heat shock proteins and stress tolerance in plants. Biochim Biophys Acta 1577: 1-9

Tezara W, Mitchell VJ, Driscoll SD, Lawlor DW (1999) Water stress inhibits plant photosynthesis by decreasing coupling factor and ATP. Nature 401: 914-917

van der Hoeven RV, Ronning C, Giovannoni JJ, Martin G, Tanksley SD (2002) Deductions about the number, organization and evolution of genes in the tomato genome based analysis of a large EST collection and selective genomic sequencing. Plant Cell 14: 1441-1456

Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: Towards genetic engineering for stress tolerance. Planta 218: 1-14

Wang W, Vinocur B, Shoseyov OI, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci 9: 244-253

Wei Y, Waltz DA, Rao N, Drummond RJ, Rosenberg S, Chapman HA (1994) Identification of the urokinase receptor as an adhesion receptor for vitronectin. J Biol Chem 269: 32380-32388