Elucidating cold acclimation pathway in blueberry by transcriptome profiling

Jose V. Die, Lisa J. Rowland*

U.S. Department of Agriculture, Agricultural Research Service, Beltsville, MD, USA

* Corresponding Author:
Jeannine.Rowland@ars.usda.gov

Abstract

A fundamental goal of cold acclimation research is to understand the mechanisms responsible for the increase in freezing tolerance in response to environmental cues. Changes in gene expression underlie some of the biochemical and physiological changes that occur during cold acclimation. Detailed and comprehensive transcriptome annotation can be considered a prerequisite for effective analysis and a fast and cost-effective way to rapidly obtain information in the context of a given physiological condition. By computational predictions and manual curation, we have annotated 454 sequence assemblies from two blueberry cDNA libraries that represent flower buds in the first and second stages of cold acclimation. Gene ontology functional classification terms were retrieved for 4343 (80.0%) sequences. GO annotation files compatible with a commonly used annotation tool have been generated and are publicly available. By mining the dataset further, it was possible to associate presence of certain transcripts related to carbohydrate metabolism and lipid metabolism with different stages of cold acclimation. This was concomitant with differential presence of Zn finger functional domains and C3H-family transcription factors. The expression of a few selected genes was validated by quantitative real-time PCR assay. Results demonstrate that our transcriptome database is a rich resource for mining cold acclimation-responsive genes.

Keywords: annotation; GO terms; C3H; non-coding RNA; qPCR; transcription factors; *Vaccinium*

1. Introduction

To survive the winter, plants of the temperate zone start preparing as early as the late summer or early fall by a process known as cold acclimation. This, in woody perennials, is generally considered a two-step process, first triggered by shortening day-length and then, declining temperatures (Powell, 1987; Weiser, 1970). The specific mechanisms plants use to survive freezing temperatures vary somewhat depending on the plant species and, to a certain extent, the particular tissue/organ exposed to freezing (Fujikawas et al., 2009; Levitt, 1972). Genetic and molecular evidence shows that cold acclimation is a complex phenomenon involving the alteration of metabolism with synthesis of specific metabolites, proteins, lipids and carbohydrates, and changes in membrane composition (Guy, 1990; Janská et al., 2010). As a whole, these processes have the specific outcome of protection against

dehydration, oxidative damage and other stresses associated with freezing temperatures. With the advances in plant genomics research in recent years, fundamental knowledge of plant responses to cold is increasing at a rapid pace (Qin et al., 2011; Thomashow, 2010).

Blueberry (*Vaccinium* spp.) is an important fruit crop because of its high nutritional value, a rich source of antioxidants, and it is one of the major berry crops grown in the United States (USDA-NASS, 2013). Low temperatures are one of the critical environmental factors that limit its growth, survival and geographical distribution. Lack of cold hardiness, therefore, has been identified as one of the most important genetic limitations of current highbush blueberry (V. corymbosum) cultivars (Rowland et al., 2011). The use of *Arabidopsis* to understand the process of cold acclimation and general responses to cold temperature, while providing a wealth of information, has perhaps led to an over-extrapolation of the data, ignoring the context of the whole plant and its interaction with the environment (Gusta and Wisniewski, 2013). In addition, *Arabidopsis* and many herbaceous plants do not normally experience freezing temperatures and their cold acclimation capacity is generally less than that of woody species. Blueberry flower buds, on the other hand, can survive temperatures as low as ~ -30°C during the middle of winter yet suffer devastating damage from temperatures slightly below 0°C during early spring when deacclimating (Ehlenfeldt et al., 2012; Rowland et al., 2013). Consequently, cold acclimation is probably more complex in woody perennials and, within this context, blueberry has and continues to serve as a model system for studying cold tolerance (Rowland et al., 2011).

Over the past decade, our laboratory has been working toward increasing our understanding of the genetic control and regulation of cold hardiness in blueberry through a combination of genetic, molecular and physiological approaches. The ultimate goal is to integrate the knowledge obtained from these various approaches to develop more cold hardy cultivars for the industry. Particularly, genomics-based research has identified a number of potential genes associated with cold acclimation or contributing to cold hardiness in blueberry. The construction and analysis of cDNA libraries using RNA from cold-acclimated (CA) and nonacclimated (NA) floral buds

have revealed a marked difference in the transcriptional regulation under the two conditions (Dhanaraj et al., 2007; Dhanaraj et al., 2004). Microarray technology has allowed the identification of genes that had not been previously reported as cold-induced in *Arabidopsis*. In addition, it has provided evidence for differences between field and cold room-based acclimation (Dhanaraj et al., 2007; Rowland et al., 2008), reflecting the importance of the environmental context to the physiology, growth habit and life cycle of the plant and how it plays a relevant role in the elucidation of cold hardiness (Gusta and Wisniewski, 2013). Additionally, the generation of subtracted libraries has resulted in valuable supplemental data leading to the identification of many potential regulatory genes such as transcription factors (Naik et al., 2007).

One of the major advances in the past decade of research on plant survival during winter has been the discovery of the C-repeat/binding factor/dehydration-responsive element binding factor (CBF/DREB) gene family. These transcriptional activators have been shown in *Arabidopsis* to regulate a number of downstream genes associated with low temperature response in plants (Shinozaki et al., 2003; Thomashow et al., 2001). CBF existence and importance has also been established in many other plants, including woody species through heterologous gene expression studies (Wisnieswski et al., 2013). Cloning of a *V. corymbosum CBF* followed by overexpression in *Arabidopsis* resulted in the induction of *COR* (cold-regulated) gene expression and constitutive freezing tolerance in transgenic plants (Polashock et al., 2010). Overexpression of the same *CBF* gene under the control of the *CaMV 35S* promoter in transgenic blueberry lines has also been shown to result in an increase in freezing tolerance in non-acclimated plants, although not to the level found in cold acclimated plants (Walworth et al., 2012). This suggests that manipulation of the *CBF* system may be potentially useful for the improvement of freezing tolerance in woody fruit crops, although it is probably not the only pathway involved (Wisnieswski et al., 2013). Therefore, there is a growing interest in cold-responsive genes that are potentially under the control of these CBF activators and other transcriptional activators and collectively, these results provide evidence for an emerging picture of blueberry response to cold temperatures. Determining the identity of cold-responsive genes and their regulators in blueberry will be useful in understanding their function

and establishing their relative importance. Clearly such information may help in the development of strategies for the improvement of cold hardiness of woody perennials in general.

More and more large-scale, high throughput gene expression studies as well as the sequencing of entire transcriptomes and genomes, are being conducted today in actual crop species, because of the advent of affordable Next Generation Sequencing (NGS) technologies. Recently, the first blueberry transcriptome from 454 NGS has become publicly available making possible computational analyses of thousands of sequences within the blueberry and related species (Rowland et al., 2012a; Rowland et al., 2012b). The availability of these kinds of resources with new bioinformatics tools provide the ability to address a more comprehensive analysis of complex and multigenic traits and will help to integrate the vast amount of data generated (Die and Rowland, 2013a).

In this study, we aimed to mine the blueberry 454 transcriptome database more in depth than previously done in order to study global changes in gene expression and identify relevant pathways or functional groups of genes associated with first and second stages of cold acclimation. Our results show that this database is very useful for the ongoing cold acclimation research in blueberry to confirm previously reported changes, identify so far unreported genes, and establish which processes predominate during different stages of cold acclimation. Our dataset may be useful for comparative genomic studies among other woody perennial species as well.

2. Material and Methods

2.1 Annotation, functional classification and bioinformatics tools

Blueberry cDNA libraries used in this study, from flower buds collected at 0 and 397 chill units (hours of exposure to temperatures from 0-7ºC) from northern highbush V. corymbosum cultivar 'Bluecrop', were constructed and described earlier (Rowland et al. 2012). The 0 and 397 chill unit time points corresponded to collection dates of 2006, September 9th and November 30th. Here, the Gene Ontology Functional Annotation Tool Blast2GO version 2.6.6 (Conesa et al., 2008; Götz et al., 2008) was

used to assign GO identities and enzyme commission numbers to contigs from assemblies of the 454 transcript sequences generated from these libraries. For the annotation, the following configuration settings were used: BLASTX against NCBI non-redundant (nr) protein database, *E*-value filter $\leq 10^{-6}$, HSP length cutoff of 33, maximum 20 BLAST hits per sequence and annotation cutoff of 50. Furthermore, to improve annotation ability, InterProScan was performed and results were merged to GO annotation.

The program Blast2GO was also used to assign biological functions, cellular components and cellular processes to the transcripts. Only sequences that were not successfully annotated were selected and re-annotated with more permissive parameters. Then, plant GOslim for all three independent GO categories were obtained from the AgBase database (http://www.agbase.msstate.edu/) to be further assigned to secondary categories and to be compared with those obtained in *Arabidopsis thaliana* (McCarthy et al., 2006). ATH GOslim *Arabidopsis* data were obtained from The *Arabidopsis* Information Resource (TAIR) database (http://www.arabidopsis.org/). The analysis of biological processes/pathways was carried out using the KEGG (Kyoto Encyclopedia of Genes and Genomes) map module supported by the Blast2GO bioinformatics tool. Blast2GO also enabled analysis related to over-representation of functional categories through the Gossip package (Blüthgen et al., 2005) for statistical assessment of annotation differences between two sets of sequences, using Fisher's exact test for each GO term. False discovery rate (FDR) controlled *P* values (FDR<0.01) were used for the assessment of differentially significant metabolic pathways. To identify transcription factors, BLASTX searches against a comprehensive *A. thaliana* transcription factor collection maintained at Plant Transcription Factor Database (PlantTFDB v3.0; http://planttfdb.cbi.pku.edu.cn/) were performed (Jin et al., 2013). Finally, the Coding Potential Calculator website (http://cpc.cbi.pku.edu.cn) was used to identify noncoding RNAs (Kong et al., 2007). Parameters for the website were set to use only the forward strand. The output data was analyzed, and a list of transcript IDs described as “noncoding” and “weakly noncoding” was created.

2.2 Plant material for qPCR

Flower buds were collected across two years at different times during development of

the highbush blueberry cultivar Bluecrop (*V. corymbosum*) and the southern rabbiteye cultivar Tifblue (*V. virgatum*), both grown at the USDA/ARS, Beltsville Agricultural Research Center, Beltsville, MD. The sample pools from each cultivar and each time point were made from a minimum of five plants. Flower buds constituting the first biological repetition were collected from field plants during the fall of 2005 with increasing exposure to chilling temperatures, measured as chill units (hours between 0-7°C). Buds were harvested at 0 (September 30th) and 436 (December 12th) chill units. The second biological repetition was comprised of flower buds collected from field plants during the fall of 2010, harvested at 0 (October 8th) and 371 (December 3rd) chill units. Samples were stored at -80ºC until used.

2.3 RNA extraction and quality controls

For total RNA isolation, bud samples were ground in liquid nitrogen and incubated at 65°C in a pre-warmed CTAB extraction buffer. Two chloroform:IAA (24:1) extractions were performed followed by overnight precipitation with LiCl (Chang et al., 1993). RNA pellets were resuspended in DEPC-treated water, precipitated again with ethanol and NaOAc, washed, and finally resuspended in 1 ml DEPC-treated water. RNA concentration was determined by measuring the optical density at 260 nm using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, USA). RNA quality was assessed by combining information from purity and integrity. First, purity was inferred from the absorption ratios using the NanoDrop. Only the RNA samples with A260/A280 ratios between 1.9 and 2.1 and A260/A230 greater than 2.0 were used in further analyses. Then, concerning integrity, RNA samples were visualized on 1% agarose gels stained with ethidium bromide. Finally, we amplified segments of the 5' and 3' regions of an ubiquitin carboxyl-terminal hydrolase gene across the cDNA samples by qPCR, as described below.

2.4 cDNA synthesis and quality controls

To eliminate any contaminating genomic DNA (gDNA), RNA extracts were treated with TURBO^{TM} DNase I (Life Technologies, USA), prior to cDNA synthesis. Two micrograms of DNase I-treated total RNA were used for the synthesis of cDNA. Complementary DNAs was synthesized by priming with oligo-dT$_{12-18}$ (Life Technologies, USA), using SuperScriptIII reverse transcriptase (Life Technologies,

USA) following the instructions of the provider. The cDNAs were diluted to a final volume of 100 µl. Each cDNA sample was then tested for presence of gDNA contamination by performing conventional PCR with a primer pair designed from two different exons of an alcohol dehydrogenase-like blueberry sequence (CF811586). The primer pair was designed to amplify a product of 1140 bp, which would span an intron, when using genomic DNA as template, or 528 bp when using cDNA as template (primers: NA799F, 5'-CCGCTGGTGATTGAAGAAGT-3'; NA799R, 5'-TTTCGCAACATTTAGCATGG-3'). In tests for gDNA contamination, the 1140 bp band was not amplified from any of the samples.

For evaluation of the integrity of the RNA, we used a 3':5' amplification cDNA ratio assessment on the grounds that the efficiency of cDNA synthesis is dependent on the intactness of mRNA (Nolan et al., 2006). This assay aimed at measuring the integrity of an ubiquitin carboxyl-terminal hydrolase blueberry transcript sequence (*UBP14*) by amplifying two 101 bp long cDNA fragments, one from the 5' and the other from the 3' region of the UBP14 gene (Die and Rowland, 2013b). The fragments are 1769 and 348 bp, respectively, from the 3' end of the cDNA. The 3':5' amplification ratios of the *UBP14* cDNA fragments were calculated from all samples using the comparative Cq method (Schmittgen and Livak 2008). All ratios fell within the range of 1.68–3.32 (2.61 ± 0.18; mean ± SEM). Only if ratios were > 4.43 would RNA quality be deemed inadequate (Die et al., 2011). Therefore, the cDNAs were judged to be suitable for qPCR analysis.

2.5 Primer design, secondary structure control and real-time qPCR assays

Primer pairs were designed to amplify genes that were expected to be differentially expressed between the libraries according to the bioinformatics analysis. All PCR primers were tested for specificity using NCBI's BLAST software (Altschul et al., 1990). Primers were designed using Primer3 software (Untergasser et al., 2012) and the following criteria: Tm of 60±1°C and PCR amplicon lengths of 80-120 bp, yielding primer sequences with lengths of 17-23 nucleotides and GC contents of 20-60%. For predicting the secondary structure of the amplicons, we used MFOLD v3.4 software with default settings of minimal free energy, 50 mM Na+, 3 mM Mg2+, and an annealing temperature of 60°C (Zuker, 2003). From these predictions, we chose primers that would yield amplicons with minimal secondary structures and melting

temperatures that would not hamper annealing. Designed primers meeting these criteria were synthesized by Integrated DNA Technologies (Coralville, IA, USA).

Suppl. Table S1 shows the overall mean real time PCR amplification efficiency of each primer pair (E) estimated from the data obtained from the exponential phase of each individual amplification plot and the equation $(1 + E) = 10^{slope}$ using LinReg software and the criteria of including three-five fluorescent data points with $R^2 \geq 0.998$ to define a linear regression line (Ramakers et al., 2003). PCR reactions were carried out in an IQ5 (Bio-Rad, Hercules, CA, USA) thermal cycler using iQ™ SYBR® Green Supermix (Bio-Rad, Hercules, CA, USA) to monitor dsDNA synthesis. Reactions contained 1 µl of the diluted cDNA as a template and 0.150 µM of each primer in a total volume reaction of 20 µl. Master mix was prepared and dispensed into individual wells using electronic Eppendorf Xplorer® multipipettes (Eppendorf AG, Germany). The following standard thermal profile was used for all PCRs: polymerase activation (95°C for 3 min), amplification and quantification cycles repeated 40 times (95°C for 30 sec, 60°C for 1 min). The specificity of the primer pairs was checked by melting-curve analysis performed by the PCR machine after 40 amplification cycles (60 to 95 ºC) and is shown in Suppl. Fig. S1. Fluorescence was analyzed using iQ5 2.1 standard optical system analysis software v2.1 (Bio-Rad). All amplification plots were analyzed using a baseline threshold of 30 relative fluorescence units (RFU) to obtain Cq (quantification cycle) values for each gene-cDNA combination.

2.6 qPCR data analysis

Calculations were performed using the advanced quantification model with efficiency correction, multiple reference gene normalization, and use of error propagation rules described by (Hellemans et al., 2007). Normalization was performed using three stably expressed blueberry reference genes identified previously: RNA helicase-like (*RH8*), clathrin adapter complexes medium subunit (*CACSa*) and the sequence encoding a hypothetical protein *Vc4g16320* (Die and Rowland, 2013b). PCR E of the references was: E_{RH8}=0.954±0.008; E_{CACSa}=0.870±0.015; $E_{4g16320}$=0.885±0.022 (mean ± SEM). To confirm the stable expression of the references in our material, we calculated the mean of their normalized relative quantities for the 2005 (1.022) and 2010 (1.046) data. Maximum fold expression difference between samples from the

mean was 0.416 and 1.036 for the 2005 and 2010 seasons, respectively.

3. Results

The goal of this work was to use the *V. corymbosum* 454 transcriptome database (http://bioinformatics.towson.edu/BBGD454/) to gain insight into changes in gene expression during the first and second stages of cold acclimation. From the website, we downloaded the files corresponding to the transcriptome sequence library assemblies from floral buds collected at the first and second stages of cold acclimation (0' and 397' of chill unit accumulation, respectively) and wrote a macro Excel© function to compare the contigs that comprised both libraries (5426 contigs; 2675 contigs in Lib 0' and 2751 contigs in Lib 397'). A total of 1460 contigs in Lib 0' (1460/2675 or 54.6%) and 1536 contigs in Lib 397' (1536/2751 or 55.8%) were differentially represented (present in one library but absent in the other). The number of contigs shared between the two floral bud libraries was 1215 (45.4% in the Lib 0'; 44.2% in the Lib 397').

3.1 Functional annotation and classification of the blueberry dataset

We annotated the contigs from both libraries using Blast2GO based on comprehensive information obtained from sequence similarity searches against the NCBI non-redundant (nr) protein database, InterProScan results, and KEGG mapping (Conesa et al., 2008; Götz et al., 2008). Overall, 90.6% of the contigs had significant similarity to known sequences in the protein database and around 26.7% (1312/5426 contigs) had highest matches to proteins of *Vitis vinifera*, the most closely-related species to blueberry with a sequenced genome (Suppl. Fig. 2). Among other non-Ericaceae family species, best matches were for *Theobroma cacao* (9.3%, 503 contigs), *Solanum lycopersicum* (7.3%, 394 contigs), *Populus trichocarpa* (6.7%, 363 contigs) and *Prunus persica* (6.7%, 362 contigs). The availability of the whole genome sequence and large number of predicted proteins from these species, along with limited available sequence information from blueberry, cranberry, lingonberry, rhododendron, etc., in the nr database explains why the blueberry sequences had highest homology to these non-Ericaceae genomes. Some of them were also woody plant species. Only 0.4% (21 contigs) matched with sequences from *Vaccinium* spp.,

indicating the scarcity of Vaccinium sequences in the nr database of GenBank, relative to some other species. Interestingly, a notable number of the contigs (9.4%, 509 contigs) lacked homology to any publicly available sequences.

3.2 Gene ontology annotation

A total of 4343 (80.0% of 5426) contigs were functionally classified into one or more ontologies; 3498 (64.5%) of the contigs were classified according to biological process, 3490 (64.3%) were classified by molecular functions, 3178 (58.6%) were classified in terms of cellular components, and 2344 (43.2%) were classified in all three ontologies. Among the 1083 contigs that could not be annotated, 509 lacked significant homology to other sequences, 180 had homology but could not be assigned GO terms, and 394 contigs had homology and could be assigned GO terms, but still could not be fully annotated.

To identify which biological processes might be implicated in the first and second stages of cold acclimation, the percentage of contigs assigned to the various biological processes were compared between Lib 0' and Lib 397' (Suppl. Fig. 3). The vast majority of contigs in both libraries were assigned to the cellular process (GO:0009987, 25.8% in Lib 0' vs 26.8% in Lib 397') and metabolic process (GO:0008152, 25.8% vs 26.9%) categories. The response to stimulus (GO:0050896) class was the third most highly represented in each library (14.7% vs 14.1%). Within this category (data not shown), response to stress (GO:0006950), response to chemical stimulus (GO:0042221) and response to abiotic stimulus (GO:0009628) were the most highly represented in both libraries. We did not detect any clear differences between the libraries, regarding the percentages of the GO terms, at this stage of the analysis.

We then compared the distribution of gene ontology annotations of our data to that of the *A. thaliana* genome. The contigs from both the blueberry 454 transcriptome data and the TAIR *A. thaliana* annotated database were mapped to respective GO plant slim categories using AgBase and *GOSlim Viewer* (McCarthy et al., 2006). A similarly large number of contigs were assigned to a wide range of gene ontology categories indicating similar distribution, while some other terms were clearly under-

or over-represented in the blueberry libraries in all three independent GO categories. For example, kinase and transferase activities and gene product localization to plastid or nucleus were generally under-represented in our libraries. On the contrary, terms associated with hydrolase activity, response to stress, response to biotic or abiotic stimuli, or gene products localized to the cell wall or plasma membrane were over-represented in the blueberry data (Fig. 1).

3.3 Annotation augmentation using InterProScan.

An InterProScan search against all European Bioinformatics Institute (EBI) databases resulted in the recovery of domain/motif functional information for 3270 contigs. Of these contigs, 2386 had homology to 2681 recognized InterPro families. The domains P-loop (IPR027417) and NAD(P)-binding (IPR016040) were the most frequently observed families, found in 66 contigs from both libraries. The most common families present in both libraries are shown in Suppl. Table 2. The InterProScan analysis also revealed the presence of some families that were specific to one of the libraries. Thus, the cytochrome P450, conserved site (IPR017972) was detected only in Lib 0' (8 contigs) whereas the Zinc finger, RING-type, conserved site (IPR017907) was observed only in Lib 397' (6 contigs). The domains (IPR codes and descriptions) specifically detected in only one of the two libraries are shown in Table 1.

3.4 KEGG pathway classification of transcripts

Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis supported by Blast2GO provided an alternative functional annotation for contigs based on association with biochemical pathways. In addition, we manually assigned annotated contigs to the KEGG Orthology (KO) system using the KEGG BRITE database (Kanehisa et al., 2012). Of the 4343 annotated contigs, 2137 were annotated with Enzyme Commission (EC) codes and mapped to 122 different KEGG pathways. This analysis resulted in the assignment of 678 EC codes to Lib 0' and 625 EC codes to Lib 397' and indicated that each of the libraries represented qualitatively comparable biochemical pathways. Table 2 shows the pathway category composition of each library. In both cases, metabolism was the predominant category (~98%). Although composition of most of the subsets of metabolism had similar values in the libraries, enzymes involved in carbohydrate metabolism (29.3% of metabolism) were over-

represented in Lib 397' whereas lipid metabolism enzymes were comparatively higher in Lib 0' (12.9% of metabolism). Next we analyzed the composition of these contrasting subsets. With regard to carbohydrate metabolism, the vast majority of KEGG pathways showed a higher presence in Lib 397' especially starch and sucrose metabolism and pyruvate metabolism pathways (Fig. 2A). The KEGG pathways associated with lipid metabolism, on the other hand, were under-represented in Lib 397' as compared to Lib 0', especially regarding enzymes involved in fatty acid elongation, glycerophospholipid metabolism and α-linolenic acid metabolism (Fig. 2B).

3.5 Gene ontology (GO) enrichment analysis

Next we used the Gossip module of Blast2GO to compare the abundance of GO terms associated with each library relative to the total number of annotated sequences. This module uses Fisher's Exact Test because of the multiple testing nature of the statistical comparison. For the analysis, the 1215 contigs that overlapped between the two libraries were set aside to minimize the redundancy. We considered expression ratios to be significantly different at a false discovery rate (FDR) of ≤0.01. Using this cutoff, no significant results were returned for Lib 0'. This hampered our ability to detect under-represented terms in Lib 397'; however, several GO terms were identified as being over-represented in Lib 397' (Suppl. Table S3). For example, translation and gene expression were the major terms that were over-represented in Lib 397' within the biological process category. Within the cellular component category, assignments to cytoplasm, chloroplast and membrane-bounded organelles were over-represented in Lib 397'. Finally, regarding molecular function, terms for RNA binding or nucleic acid binding were enriched in the Lib 397'.

3.6 Identification of genes involved in different stages of cold acclimation

We also used keywords associated with previously described cold acclimation-responsive genes to search both libraries in order to identify potentially differentially expressed genes between the two stages of cold acclimation. Cold acclimation-responsive genes can be divided into two groups: genes encoding components of the signal transduction pathways that regulate gene expression in response to stress and

are triggered by environmental cues and genes whose products directly participate in the protection of cells under stress conditions. Genes that were identified from these searches that might be expressed at higher levels in the first stage of cold acclimation (Lib 0') included several sensor genes possibly involved in cold response signaling such as Ca^{2+}-dependent protein kinase (CDPK, 1 contig), mitogen-activated protein kinase (MAPK, 3 contigs), calcineurin B-like protein (CBL, 1 contig), calmodulin-binding transcription activator (CAMTA, 2 contigs), phospholipase C (1 contig), and histidine kinase (HK, 1 contig). Moreover, we identified genes encoding 5 lipid-transfer protein (LTP) genes and 2 fatty acid desaturases with potential roles in altering the composition of proteins and lipids. Finally, genes encoding antifreeze proteins such as β-1, 3-glucanase-like proteins (GLP, 4 contigs), chitinase-like proteins (CLP, 1 contig), and late-embryogenesis-abundant proteins (LEA, 2 contigs) were also identified. Concerning Lib 397' and expression of genes that might increase during the second stage of cold acclimation, 9 contigs related to cold sensors or signal transduction genes (MAPK, phospholipase C and triphosphate inositol), 3 contigs involved in altering the composition of proteins and lipids (LTP and FAD), and 8 contigs encoding antifreeze proteins (β-1, 3-glucanase-like protein, thaumathin-like protein, and LEA) were identified. Moreover, we found the presence of genes encoding enzymes associated with metabolism of sugar alcohols (inositol, mannitol, sobitol) and early light-induced proteins (ELIP; 4 contigs) in Lib 397' but not in Lib 0'. Supp. Table 4 and 5 show the genes identified in this way, along with their associated GO terms for biological processes, molecular functions, and cellular components.

3.7 Computational identification and prediction of transcription factors

To identify putative transcription factors (TFs), we blasted the data from our two libraries against the plant TF database (PlantTFDB 3.0), which includes 2296 *Arabidopsis* TFs (Jin et al., 2013). Matches were considered significant at *E*-value $<10^{-5}$. Hits to 58 different *Arabidopsis* TF families were found; 36 were represented in Lib 0' and 37 in Lib 397'. Blastx results revealed that 169 and 157 contigs from each library were homologs of *Arabidopsis* TFs. This represents a TF frequency of 6.3% and 5.7% in Lib 0' and Lib 397', respectively. Although the total number of TFs was similar, the types differed between the libraries. For example, we found that the

basic/helix-loop-helix (bHLH) superfamily of TFs was over-represented in Lib 0′, whereas the C3H zinc finger and CO-like family was more predominant in Lib 397′ (Fig. 3). In Lib 0′, the most abundant TF family was WRKY (28 contigs, 16.6% of TFs) followed by the auxin response factors group (ARF; 14, 8.3%), the AP2/ERF superfamily (13, 7.7%) and bHLH group (13, 7.7%). In Lib 397', the high-frequency TF families were WRKY (27, 17.2%), C3H (14, 8.9%), ARF (12, 7.6%), and AP2/ERF (10, 6.4%). Suppl. Table 6 shows the total number of contigs identified per TF family and their distribution in percentages in each particular library.

3.8 Computational identification and prediction of noncoding RNAs

A number of contigs (509 contigs total from both libraries) lacked homology to any known publicly available sequence. One explanation for sequences lacking similarity to other known sequences is that they are too small to yield a significant score in the blast search; however, all of our contigs were filtered to be >200 bp. Furthermore, the 234 contigs from Lib 0' that lacked homology had an average size of 661 bp, and the 275 contigs from Lib 397' had an average length of 641 bp. Other explanations could be that these contigs are blueberry-specific or represent novel genes not previously characterized in woody perennials. Alternatively, they could be non-protein-coding sequences. To test this hypothesis, we used a web-based public tool that utilizes a support vector machine to distinguish between coding and noncoding RNAs (ncRNA; Kong et al. 2007). From this analysis, 223 (8.3%) of the contigs from Lib 0′ and 268 (9.7%) of the contigs from Lib 397' were described as ′noncoding′ or ′weakly noncoding′. A list of the transcript IDs categorized as ncRNA is shown in Suppl. Table 7.

3.9 Differential expression analysis by using qPCR

To validate the results of the bioinformatics analysis, the expression of certain contigs that were identified as likely to be differentially expressed was analyzed by quantitative real-time PCR (qPCR). Contigs associated with GOslim terms over-represented in blueberry libraries relative to those in *Arabidopsis* such as 'hydrolase activity', 'plasma membrane', 'cell wall' and 'response to biotic/abiotic stimulus' were selected from Lib 397'. Concerning the specific domain/motif elements

identified in the analysis, we selected contig00170 (IPR002119) and contig00449 (IPR017972) from Library 0' and contig00535 (IPR017907) from Lib 397'. Other contigs with potential roles in fatty acid elongation (contig00697 and contig02401), encoding LEA (contig00104) and ELIP (contig02750) proteins, having a Zn finger C3H-type domain (contig00315), and encoding a putative lncRNA (contig02743) were selected from library 397'. Fig. 4 shows the comparative expression analysis of selected contigs between 'Bluecrop' and the more cold-sensitive rabbiteye genotype 'Tifblue' during the natural cold acclimation process for a period ranging from end September/early October (weeks 39th -40th) to early/mid December (weeks 48th-50th) over the fall of 2005 and 2010. Fold changes > 2 (up- or down-regulation) were observed for 9 contigs as expected according to their differential presence in the libraries 0' and 397'. Transcripts that were induced in 'Bluecrop' were also induced in 'Tifblue' but at lower levels. Three other transcripts (contig00697, contig02585 and contig02231) did not show regulation under cold acclimation at the time points analyzed in 'Bluecrop'. However, even in these cases, it was possible to detect differences between cultivars. For example, the expression level of the transcript associated with plasma membrane in 'Bluecrop' was ~2-fold higher than that in 'Tifblue', and the expression level of the transcript associated with response to abiotic/biotic stimulus was almost negligible in 'Tifblue' in contrast to 'Bluecrop', resulting in clear genotype-dependent differences in expression.

4. Discussion

Lack of mid-winter cold hardiness and susceptibility to spring frosts have been identified as two of the most important genetic limitations of highbush blueberry cultivars (Moore, 1994). These factors limit their productivity and geographical distribution. Investigation into the molecular basis of cold acclimation has been one of the major approaches to the study of plants' response to freezing temperatures and, therefore, to understanding one key component of winter survival (Arora and Rowland, 2011). Recently, results from the first large blueberry transcriptome sequencing project have become available. Four of the cDNA libraries sequenced in this project were developed from flower buds collected from (1) blueberry plants in the very early stages of acclimation (0 chill units), (2) plants approaching maximum cold hardiness (397 chill units), (3) plants after reaching maximum cold hardiness but

before deacclimation (789 chill units) and (4) deacclimating plants (1333 chill units). Therefore, the dataset generated represents a comprehensive view of the cold acclimation and deacclimation process in blueberry (Rowland et al., 2012a). In order to better elucidate the first and second stages of cold acclimation response pathways, here we have mined the data from the first two of these libraries (0 and 397 chill units) extensively by bioinformatics analysis and computational screenings. We have previously shown that CBF expression is elevated at 0 chill units relative to 397 chill units (Polashock et al., 2010) and that buds already have a cold hardiness level or LT_{50} of about -13ºC at this first time point indicating they are in the first stages of cold acclimation (Rowland et al., 2012a). Gene expression analyses have confirmed the involvement in cold response for some of the candidate genes identified as well.

Overlapping sequence information obtained from different libraries has been a useful tool to identify differentially expressed genes. Of the total number of contigs generated from the 0' and 397' libraries, ~45% from each were shared with the other library. One approach to increase the amount of differential transcriptomic information (including rare genes) in comparative studies would be to combine the generation of suppression subtractive hybridization libraries with NGS platforms. This novel strategy has been used successfully in plant research (Barros de Carvalho et al., 2013). In any case, the ~55% of unshared transcripts in our dataset allowed us to identify a large number of potentially differentially expressed transcripts and suggests an extensive remodeling of the transcriptome between the first and second stages of cold acclimation. This set of genes comprised a good starting point for further characterization.

To improve accessibility and the applicability of the GO terms generated from this study, the annotation files are provided in a format compatible with a publicly available tool, the functional annotation suite Blast2GO (Conesa et al., 2008; Götz et al., 2008) and they can be accessed via the Blueberry Genomics Database 454 transcriptome hosted on the Bioinformatics server at Towson University (BBGD454: http://bioinformatics.towson.edu/BBGD454/).

Comparing gene ontology annotations of a given transcriptome to that of a well-studied model species offers general information on the functional category

distributions. Therefore, contigs were grouped into GOslim categories to facilitate data interpretation. GOslim terms are a specified subset of higher-level ontology categories that provide a broad profile for genome to genome comparisons (Lomax, 2005). In our analysis, some GOslim classes mirrored those of *Arabidopsis* genes reflecting that a large diversity of blueberry transcripts is represented by these sequences. However, terms associated with hydrolase activity, response to stress, and response to stimuli, were over-represented in the blueberry data. This is not surprising considering the blueberry libraries were constructed from plants undergoing cold acclimation, and hydrolysis of sugars and lipids are known to be associated with plants' response to cold.

After completing the annotation of contigs from the two libraries, it was easy to search the library sequences by keywords using terms relevant to cold acclimation and identify certain types of sequences unique to each library. It is known that cold stress activates Ca^{2+} channels to increase the cytosolic Ca^{2+} level, and then triggers phospholipase C and D, producing inositol triphosphate and phosphatidic acid, respectively (Reddy et al., 2011). Next, a variety of signaling pathways are triggered, such as CDPK, MAPK, CBL, CAMTA and the activation of TF, all of which promote the production of cold-responsive proteins (Janská et al., 2010; Miura and Furumoto, 2013). In our study, we have identified a set of potential multiple primary sensors that are unique to Lib 0' (including 1 CDPK, 3 MAPK, 1 CBL, 2 CAMTA, 1 phospholipase and 1 HK genes) and thus, possibly involved in the signaling pathways during the first stage of cold acclimation. Furthermore we also have identified a set of multiple primary sensors that are present only in Lib 397' (5 MAPK, 2 phospholipases, 1 triphosphate inositol, 1 Ca^{2+}-mediated signaling and 1 Ca^{2+} channel genes). Thus, this may be part of the cold stress signal transduction pathway during the second stage of cold acclimation.

The plasma membrane, with its basic role in separating the internal from the external environment, is a primary site for the action of cold-responsive proteins. By altering the compositions of lipids, these proteins provide a means for the plant to stabilize the membrane structure (Welling and Palva, 2006). In our GOslim term comparisons with those from the *Arabidopsis* genome, an over-representation (~2.7%) was found in our libraries for terms for cell wall or plasma membrane localization. By searching the

libraries using “membrane”-related terms, we found several unique to both libraries LTP and FAD genes. On the other hand, in spite of the stabilization of the plasma membrane, some proteins function during cold exposure as inhibitors of ice nucleation. These proteins are so-called anti-freezing proteins (AFP). Within this group, we have detected several contigs unique to Lib 397’ that encode thaumatin-like protein (TLP), and late-embryogenesis-abundant protein (LEA).

The stabilization of the plasma membrane is also related to osmotic equilibrium. In order to maintain osmotic balance, lower the freezing point, and bind and protect specific molecules, plants accumulate a variety of cryoprotectants such as low molecular weight compounds or sugar alcohols (Janská et al., 2010; Uemura et al., 2006). Interestingly, several contigs encoding enzymes associated with the biosynthesis of mannitol, inositol and sorbitol were detected only in Lib 397’ suggesting these cryoprotectant molecules function in blueberry in the second stage of cold acclimation. Moreover, the accumulation of soluble sugars has been correlated with the development of freezing tolerance (Dahal et al., 2012; Hurry et al., 1994; Oquist et al., 1993). Also, energy is necessary to drive acclimation, and accumulation of sucrose is essential during this process (Wisnieswski et al., 2013). Contigs encoding enzymes involved in carbohydrate metabolism were over-represented in Lib 397’ (representing 29.3% of metabolism). Particularly, the most abundant contigs were mapped to starch and sucrose metabolism, glycolysis and pyruvate metabolism. These results suggest that certain carbohydrate metabolic pathways play an important role in the second stage of cold acclimation as well.

Another aspect of winter survival that has not been very well studied is the cellular mechanism of photoprotection. In natural environments, winter survival encompasses more than just overcoming cold temperatures. Plants are subjected to a range of environmental conditions leading to CA. One factor that may become stressful in winter is light, which can result in photo-oxidative damage (Arora and Rowland, 2011). The Fisher’s test showed several GO terms related to chloroplast or chloroplast components to be over-represented in Lib 397’ suggesting a correlation between photosystems and the second stage of cold acclimation. Moreover, we also found several contigs encoding ELIPs to be differentially represented in that library. By December, under our conditions, highbush blueberry plants have already shed their

leaves. However, the flower buds collected for this work included the bud scales, which are specialized leaves. The decrease in activity of photosynthetic enzymes, as occurs during the winter, could potentially result in the absorption of light energy by chloroplasts in leaves (in evergreen plants) and in bud scales in excess of what can be processed. It has been proposed that ELIP family members play a photoprotective role under light stress conditions either by preventing the formation of free radicals and/or by participating in energy dissipation (Adamska et al., 2001; Montané and Kloppstech, 2000). ELIPs have been previously found to be among the most highly induced genes during in blueberry floral buds (Dhanaraj et al., 2004) and in rhododendron leaves (Peng et al., 2008; Wei et al., 2005; Wei et al., 2006) supporting the idea that adaptation mechanisms to high light stress are important in photosynthetic tissues in winter. Therefore, ELIPs may play an important protective role in the second stage of cold acclimation in blueberry floral buds.

The production of cold responsive proteins is regulated upstream by transcription factors (Chinnusamy et al., 2010; Singh et al., 2002). Based on our blast results, many contigs encoding putative transcription factors (TF) were identified. These belonged to different TF families and ranged from 5.71% to 6.32% of the contigs in each library. This is much less than the 17% of cold upregulated genes encoding TF reported in an *Arabidopsis* microarray analysis (Lee et al., 2005). Although *Arabidopsis* provides a wealth of information, differences in the tissues analyzed, the duration of the experiments, and conditions of natural or artificial cold acclimation environments, may result in limited direct applicability to woody species (Dhanaraj et al., 2007; Rowland et al., 2011; Wei et al., 2005; Wisniewski et al., 2006). In both libraries, the seven most highly represented TF families were WRKY, ARF, C3H, AP2/ERF, bHLH, C2H2, and NAC. Although these families play diverse roles in plant developmental processes and environmental responses, most of them have been reported to be linked to cold stress resistance in plants (Hu et al., 2008; Liu and Hu, 2013; Rushton et al., 2010; Şahin-Çevik and Moore, 2013). Another three families (bZIP, CO-like and MYB) accounted for 70% of the total number of TFs detected in Library 397′. CO-like TFs are important regulators of flowering. This family was absent from Lib 0′. It is important to note here that not all changes in gene expression that occur during the process of cold acclimation are directly involved in freezing tolerance. The capacity of woody perennial plants to survive winter freezes is

dependent on their entering a state of dormancy and developing cold hardiness (Powell, 1987). In blueberry, like in many other woody plants, dormancy and cold acclimation are regulated by similar environmental cues. As a result, development of endodormancy and cold acclimation are superimposed in floral bud meristems. It is also possible CA-responsive genes and endodormancy-responsive genes may interact and regulate each other. Consequently, it is not straightforward to dissect the mechanistic pathways specifically involving cold acclimation.

Notably, contigs encoding C3H-type zinc fingers were more predominant in Lib 397'. C3H-type zinc fingers are a type of plant cold shock domain protein (CSDP). Although their importance and biological functions are not fully established, experiments on *Arabidopsis* suggest that CSDPs may confer cold tolerance by functioning as RNA chaperones (Kim et al., 2007; Kim et al., 2009; Park et al., 2009). All of these results might indicate that CSDP are important players in blueberry response to cold stress.

Another significant outcome of this study was the identification of a number of contigs without homology to any previously described sequence. These genes may be non-protein sequences with regulatory functions involved in cold acclimation. Recent studies on mammalian transcriptomes and, to a lesser extent, on plant transcriptomes have pointed out a substantial portion of transcribed genes that have no protein-coding capacity. Recent research has focused on the role of non-coding RNAs as regulators of stress responses. Arbitrarily, ncRNAs >200 bases are commonly referred to as long ncRNAs (lncRNAs). Although biological functions and detailed mechanisms remain largely unknown, a growing body of evidence supports the idea that lncRNAs play important roles in multiple levels of gene regulation through transcriptional silencing mechanisms or are associated with post-transcriptional events including mRNA degradation (Zhang and Chen, 2013; Zhu and Wang, 2012). In total, 8.3% (Lib0′) and 9.7% (Lib397′) of the contigs were categorized as potential lncRNAs. A systematic identification of lncRNAs is limited to a few plant species (Boerner and McGinnis, 2012; Wu et al., 2013)) but a number of new regulators have been shown to be likely involved in stress responses (Heo and Sung, 2011; Liu et al., 2012; Swiezewski et al., 2009; Wu et al., 2012; Xin et al., 2011) including some cold-responsive lncRNAs (Heo and Sung, 2011; Swiezewski et al., 2009). It is possible our

results identify for the first time a set of potential lncRNAs that may play a role in regulation of cold acclimation in a woody perennial species.

Independent validation of the results observed from the *in silico* approach were obtained for several contigs using qPCR. Most of the genes analyzed showed a regulation pattern that was expected from the bioinformatic analysis. Moreover, we observed significant differential expression between the two cultivars with different cold hardiness levels (Arora and Rowland, 2011; Rowland et al., 2008). Genes showing significant differential expression between the two genotypes can be further explored as candidate genes for freezing tolerance using functional genomics approaches. As a whole, these results suggest that the transcriptome composition from the 454 sequencing project gives a rough estimation of the expression level of genes under the given conditions; thus, the approach of identifying CA-responsive genes through mining of this dataset followed by experimental analysis is useful and worth further investigation.

5. Concluding remarks

Information from comprehensive transcriptome sequencing is a critically important resource for many downstream analyses of gene expression including genome-wide expression profiling or comparisons using many candidate genes across tissues or conditions, even taxa, to help gain insights into the molecular differences that may underlie adaptive changes. To achieve these goals, the availability of a well annotated reference transcriptome is essential. Through annotation, transcriptome sequences are given biological meaning. The approach of coupling computational screening with experimental gene expression analysis provides an avenue for gene discovery and identification of differential expression. In general, CA-responsive genes can be divided into two groups: genes encoding components of the signal transduction pathways that regulate gene expression in response to stress and are triggered by environmental cues and genes whose products directly participate in the protection of the cells under stressful conditions. In our analysis we have identified a number of blueberry genes that belong to these two groups in both libraries. Under natural conditions, genes involved in cold acclimation are not limited to cold stress responses because plants are vulnerable to other abiotic stresses such as photooxidative damage.

Pathway analysis indicated that the carbohydrate metabolism pathway may play a vital role in blueberry during cold acclimation, particularly during the second stage. Upstream from the production of cold-responsive proteins, TFs play a key role in the regulation of gene expression. The abundance of bHLH in the 0 chill unit library and C3H-type Zn fingers in the 397 chill unit library is a starting point for further investigation into the regulatory mechanisms operating during the first and second stages of CA. The investigation of potential lncRNAs also offers a novel approach to the study of regulary mechanisms important in CA.

References

Web references

AgBase web server
http://www.agbase.msstate.edu/

Coding Potential Calculator
http://cpc.cbi.pku.edu.cn/

Kyoto Encyclopedia of Genes and Genomes (KEGG)
http://www.genome.jp/kegg/ or http://www.kegg.jp/

Plant Transcription Factor Database (PlantTFDB)
http://planttfdb.cbi.pku.edu.cn/

The Arabidopsis Information Resource (TAIR) database
http://www.arabidopsis.org/

V. corymbosum transcriptome database
http://bioinformatics.towson.edu/BBGD454/

Adamska, I., Kruse, E., Kloppstech, K., 2001. {Stable Insertion of the Early Light-induced Proteins into Etioplast Membranes Requires Chlorophyll a. Journal of Biological Chemistry 276, 8582-8587.
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403-410.
Arora, R., Rowland, L.J., 2011. Physiological research on winter-hardiness: deacclimation resistance, reacclimation ability, photoprotection strategies, and a cold acclimation protocol design. HortScience 46, 1070-1078.

Barros de Carvalho, G.A., Batista, J.S.S., Marcelino-Guimaraes, F.C., Costa do Nascimento, L., Hungria, M., 2013. Transcriptional analysis of genes involved in nodulation in soybean roots inoculated with Bradyrhizobium japonicum strain CPAC 15. BMC Genomics 14, 153.

Blüthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H., Dieter, B., 2005. Biological profiling of genes groups utilizing gene ontology. Genome Informatics 16, 106-115.

Boerner, S., McGinnis, K.M., 2012. Computational identification and functional predictions of long noncoding RNA in Zea mays. PLoS ONE 7, e43047.

Chang, S., Puryear, J., Cairney, J., 1993. A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter 11, 113-116.

Chinnusamy, V., Zhu, J.-K., Sunkar, R., 2010. Gene regulation during cold stress acclimation in plants, Plant Stress Tolerance. Springer, pp 39-55.

Conesa, A., #xF6, tz, S., 2008. Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics.

Dahal, K., Kane, K., Gadapati, W., Webb, E., Savitch, L.V., Singh, J., Sharma, P., Sarhan, F., Longstaffe, F.J., Grodzinski, B., 2012. The effects of phenotypic plasticity on photosynthetic performance in winter rye, winter wheat and Brassica napus. Physiologia Plantarum 144, 169-188.

Dhanaraj, A., Alkharouf, N., Beard, H., Chouikha, I., Matthews, B., Wei, H., Arora, R., Rowland, L., 2007. Major differences observed in transcript profiles of blueberry during cold acclimation under field and cold room conditions. Planta 225, 735-751.

Dhanaraj, A.L., Slovin, J.P., Rowland, L.J., 2004. Analysis of gene expression associated with cold acclimation in blueberry floral buds using expressed sequence tags. Plant Science 166, 863-872.

Die, J.V., Obrero, A.n., González-Verdejo, C.I., Román, B., 2011. Characterization of the 3':5' ratio for reliable determination of RNA quality. Analytical Biochemistry 419, 336-338.

Die, J.V., Rowland, L.J., 2013a. Advents of genomics in blueberry. Molecular Breeding.

Die, J.V., Rowland, L.J., 2013b. Superior Cross-Species Reference Genes: A Blueberry Case Study. PLoS ONE 8, e73354.

Ehlenfeldt, M.K., Rowland, L.J., Ogden, E.L., Vinyard, B.T., 2012. Cold-hardiness, acclimation, and deacclimation among diverse blueberry genotypes. Journal of the American Society for Horticultural Science 137, 31-37.

Fujikawas, S., Kasuga, J., Takata, N., Arakawa, K., 2009. Factors related to change of deep supercooling capability in xylem parenchima cells of tress, in: Gusta, L., Wisnieswski, M., KK, T. (Eds.), Plant Cold Hardiness from the Laboratory to the Field. CABI Press, Wallingford, pp 29-42.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Research 36, 3420-3435.

Gusta, L.V., Wisniewski, M., 2013. Understanding plant cold hardiness: an opinion. Physiologia Plantarum 147, 4-14.

Guy, C.L., 1990. Cold Accelimation and Freezing Stress Tolerance: Role of Protein Metabolism. Annual Review of Plant Physiology and Plant Molecular Biology 41, 187-223.

Hellemans, J., Mortier, G., De Paepe, A.D., Speleman, F., Vandesompele, J., 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology 8, R19.

Heo, J.B., Sung, S., 2011. Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. Science 331, 76-79.

Hu, H., You, J., Fang, Y., Zhu, X., Qi, Z., Xiong, L., 2008. Characterization of transcription factor gene SNAC2 conferring cold and salt tolerance in rice. Plant Molecular Biology 67, 169-181.

Hurry, V.M., Malmberg, G., Gardestrom, P., Oquist, G., 1994. Effects of a short-term shift to low temperature and of long-term cold hardening on photosynthesis and ribulose-1, 5-bisphosphate carboxylase/oxygenase and sucrose phosphate synthase activity in leaves of winter rye (Secale cereale L.). Plant Physiology 106, 983-990.

Janská, A., Maršík, P., Zelenková, S., Ovesná, J., 2010. Cold stress and acclimation – what is important for metabolic adjustment? Plant Biology 12, 395-405.

Jin, J., Zhang, H., Kong, L., Gao, G., Luo, J., 2013. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Research.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M., 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Research 40, D109-D114.

Kim, J.Y., Park, S.J., Jang, B., Jung, C.H., Ahn, S.J., Goh, C.H., Cho, K., Han, O., Kang, H., 2007. Functional characterization of a glycine-rich RNA-binding protein 2 in Arabidopsis thaliana under abiotic stress conditions. The Plant Journal 50, 439-451.

Kim, M.-H., Sasaki, K., Imai, R., 2009. Cold shock domain protein 3 regulates freezing tolerance in Arabidopsis thaliana. Journal of Biological Chemistry 284, 23454-23460.

Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., Gao, G., 2007. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Research 35, W345-W349.

Lee, B.-h., Henderson, D.A., Zhu, J.-K., 2005. The Arabidopsis Cold-Responsive Transcriptome and Its Regulation by ICE1. THE PLANT CELL ONLINE 17, 3155-3175.

Levitt, J., 1972. Responses of plants to environmental stresses. Academic Press.

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., Arenas-Huertero, C., Chua, N.-H., 2012. Genome-Wide Analysis Uncovers Regulation of Long Intergenic Noncoding RNAs in Arabidopsis. THE PLANT CELL ONLINE 24, 4333-4345.

Liu, S., Hu, L., 2013. Genome-wide analysis of the auxin-response-factor gene family in cucumber. Genetics and molecular research: GMR 12.

Lomax, J., 2005. Get ready to GO! A biologist's guide to the Gene Ontology. Briefings in Bioinformatics 6, 298-304.

McCarthy, F., Wang, N., Magee, G.B., Nanduri, B., Lawrence, M., Camon, E., Barrell, D., Hill, D., Dolan, M., Williams, W.P., Luthe, D., Bridges, S., Burgess, S., 2006. AgBase: a functional genomics resource for agriculture. BMC Genomics 7, 229.

Miura, K., Furumoto, T., 2013. Cold Signaling and Cold Response in Plants. International journal of molecular sciences 14, 5312-5337.

Montané, M.-H., Kloppstech, K., 2000. The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? Gene 258, 1-8.

Moore, J.N., 1994. The blueberry industry of North America. HortTechnology 4, 96-102.

Naik, D., Dhanaraj, A.L., Arora, R., Rowland, L.J., 2007. Identification of genes associated with cold acclimation in blueberry (Vaccinium corymbosum L.) using a subtractive hybridization approach. Plant Science 173, 213-222.

Nolan, T., Hands, R.E., Bustin, S.A., 2006. Quantification of mRNA using real-time RT-PCR. Nature Protocols 1, 1559-1582.

Oquist, G., Hurry, V.M., Huner, N.P., 1993. Low-temperature effects on photosynthesis and correlation with freezing tolerance in spring and winter cultivars of wheat and rye. Plant Physiology 101, 245-250.

Park, S.J., Kwak, K.J., Oh, T.R., Kim, Y.O., Kang, H., 2009. Cold shock domain proteins affect seed germination and growth of Arabidopsis thaliana under abiotic stress conditions. Plant and Cell Physiology 50, 869-878.

Peng, Y., Lin, W., Wei, H., Krebs, S.L., Arora, R., 2008. Phylogenetic analysis and seasonal cold acclimation-associated expression of early light-induced protein genes of Rhododendron catawbiense. Physiologia Plantarum 132, 44-52.

Polashock, J.J., Arora, R., Peng, Y., Naik, D., Rowland, L.J., 2010. Functional Identification of a C-repeat Binding Factor Transcriptional Activator from Blueberry Associated with Cold Acclimation and Freezing Tolerance. Journal of the American Society for Horticultural Science 135, 40-48.

Powell, L., 1987. The Hormonal Control of Bud and Seed Dormancy in Woody Plants, in: Davies, P. (Ed.), Plant Hormones and their Role in Plant Growth and Development. Springer Netherlands, pp 539-552.

Qin, F., Shinozaki, K., Yamaguchi-Shinozaki, K., 2011. Achievements and Challenges in Understanding Plant Abiotic Stress Responses and Tolerance. Plant and Cell Physiology 52, 1569-1582.

Ramakers, C., Ruijter, J.M., Deprez, R.H.L., Moorman, A.F.M., 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neuroscience Letters.

Reddy, A.S.N., Ali, G.S., Celesnik, H., Day, I.S., 2011. Coping with Stresses: Roles of Calcium- and Calcium/Calmodulin-Regulated Gene Expression. THE PLANT CELL ONLINE 23, 2010-2032.

Rowland, L., Alkharouf, N., Darwish, O., Ogden, E., Polashock, J., Bassil, N., Main, D., 2012a. Generation and analysis of blueberry transcriptome sequences from leaves, developing fruit, and flower buds from cold acclimation through deacclimation. BMC Plant Biology 12, 46.

Rowland, L.J., Bell, D.J., Alkharouf, N., Bassil, N.V., Drummond, F.A., Beers, L., Buck, E.J., Finn, C.E., Graham, J., McCallum, S., Hancock, J.F., Polashock, J.J., Olmstead, J.W., Main, D., 2012b. Generating Genomic Tools for Blueberry Improvement. International Journal of Fruit Science 12, 276-287.

Rowland, L.J., Dhanaraj, A.L., Naik, D., Alkharouf, N., Matthews, B., Arora, R., 2008. Study of Cold Tolerance in Blueberry Using EST Libraries, cDNA Microarrays, and Subtractive Hybridization. HortScience 43, 1975-1981.

Rowland, L.J., Hancock, J.F., Bassil, N.V., 2011. Blueberry, in: Folta, K.M., Kole, C. (Eds.), Genetics, Genomics and Breeding of Berries. Science Publishers, Enfield, NH, USA, pp 1-40.

Rowland, L.J., Ogden, E.L., Takeda, F., Glenn, D.M., Ehlenfeldt, M.K., Vinyard, B.T., 2013. Variation among Highbush Blueberry Cultivars for Frost Tolerance of Open Flowers. HortScience 48, 692-695.

Rushton, P.J., Somssich, I.E., Ringler, P., Shen, Q.J., 2010. WRKY transcription factors. Trends in Plant Science 15, 247-258.

Şahin-Çevik, M., Moore, G.A., 2013. Identification of a drought-and cold-stress inducible WRKY gene in the cold-hardy Citrus relative Poncirus trifoliata. New Zealand Journal of Crop and Horticultural Science 41, 57-68.

Shinozaki, K., Yamaguchi-Shinozaki, K., Seki, M., 2003. Regulatory network of gene expression in the drought and cold stress responses. Current Opinion in Plant Biology 6, 410-417.

Singh, K.B., Foley, R.C., Oñate-Sánchez, L., 2002. Transcription factors in plant defense and stress responses. Current Opinion in Plant Biology 5, 430-436.

Swiezewski, S., Liu, F., Magusin, A., Dean, C., 2009. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature 462, 799-802.

Thomashow, M.F., 2010. Molecular Basis of Plant Cold Acclimation: Insights Gained from Studying the CBF Cold Response Pathway. Plant Physiology 154, 571-577.

Thomashow, M.F., Gilmour, S.J., Stockinger, E.J., Jaglo-Ottosen, K.R., Zarka, D.G., 2001. Role of the Arabidopsis CBF transcriptional activators in cold acclimation. Physiologia Plantarum 112, 171-175.

Uemura, M., Tominaga, Y., Nakagawara, C., Shigematsu, S., Minami, A., Kawamura, Y., 2006. Responses of the plasma membrane to low temperatures. Physiologia Plantarum 126, 81-89.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3—new capabilities and interfaces. Nucleic Acids Research 40, e115.

USDA-NASS, 2013. Noncitrus Fruits and Nuts 2012 Preliminary Summary.
Walworth, A., Rowland, L., Polashock, J., Hancock, J., Song, G.-q., 2012. Overexpression of a blueberry-derived CBF gene enhances cold tolerance in a southern highbush blueberry cultivar. Molecular Breeding 30, 1313-1323.
Wei, H., Dhanaraj, A., Rowland, L., Fu, Y., Krebs, S., Arora, R., 2005. Comparative analysis of expressed sequence tags from cold-acclimated and non-acclimated leaves of Rhododendron catawbiense Michx. Planta 221, 406-416.
Wei, H.U.I., Dhanaraj, A.L., Arora, R., Rowland, L.J., Fu, Y.A.N., Sun, L.I., 2006. Identification of cold acclimation-responsive Rhododendron genes for lipid metabolism, membrane transport and lignin biosynthesis: importance of moderately abundant ESTs in genomic studies. Plant, Cell & Environment 29, 558-570.
Weiser, C.J., 1970. Cold Resistance and Injury in Woody Plants: Knowledge of hardy plant adaptations to freezing stress may help us to reduce winter damage. Science 169, 1269-1278.
Welling, A., Palva, E.T., 2006. Molecular control of cold acclimation in trees. Physiologia Plantarum 127, 167-181.
Wisnieswski, M., Nassuth, A., Teulières, C., Marque, C., Rowland, L.J., Cao, P.B., Brown, A., 2013. Genomics of cold hardiness in woody plants. Critical Reviews in Plant Sciences, in press.
Wisniewski, M.E., Bassett, C.L., Renaut, J., Farrell, R., Tworkoski, T., Artlip, T.S., 2006. Differential regulation of two dehydrin genes from peach (Prunus persica) by photoperiod, low temperature and water deficit. Tree Physiology 26, 575-584.
Wu, H.-J., Wang, Z.-M., Wang, M., Wang, X.-J., 2013. Widespread Long Noncoding RNAs as Endogenous Target Mimics for MicroRNAs in Plants. Plant Physiology 161, 1875-1884.
Wu, J., Okada, T., Fukushima, T., Tsudzuki, T., Sugiura, M., Yukawa, Y., 2012. A novel hypoxic stress-responsive long non-coding RNA transcribed by RNA polymerase III in Arabidopsis. RNA Biology 9, 302-313.
Xin, M., Wang, Y., Yao, Y., Song, N., Hu, Z., Qin, D., Xie, C., Peng, H., Ni, Z., Sun, Q., 2011. Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. BMC Plant Biology 11, 61.
Zhang, Y.C., Chen, Y.Q., 2013. Long noncoding RNAs: new regulators in plant development. Biochemical and biophysical research communications 436, 111-114.
Zhu, Q.-H., Wang, M.-B., 2012. Molecular functions of long non-coding RNAs in plants. Genes 3, 176-190.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31, 3406-3415.