Balancing selection maintains diversity in a cold tolerance gene in broadly distributed Live Oaks

Journal:	<i>Genome</i>
Manuscript ID	gen-2016-0208.R2
Manuscript Type:	Article
Date Submitted by the Author:	27-Mar-2017
Complete List of Authors:	Meireles, Jose Eduardo; University of Minnesota Twin Cities, Ecology Evolution and Behaviour Beulke, Anne; University of Minnesota, Ecology, Evolution and Behavior Borkowski, Daniel; University of Notre Dame, Department of Biological Sciences Romero-Severson, Jeanne; University of Notre Dame, Department of Biological Sciences Cavender-Bares, Jeannine; University of Minnesota, Ecology, Evolution and Behavior
Is the invited manuscript for consideration in a Special Issue? :	Evolution of Tree Diversity
Keyword:	

SCHOLARONE™
Manuscripts

Balancing selection maintains diversity in a cold tolerance gene in broadly distributed Live Oaks

Jose Eduardo Meireles^{1*}, Anne Beulke¹, Daniel Borkowski², Jeanne Romero-Severson^{2**},
and Jeannine Cavender-Bares^{1***}

¹Ecology, Evolution, and Behavior, University of Minnesota

²Biological Sciences, University of Notre Dame

*meireles@umn.edu

**jromeros@nd.edu

***cavender@umn.edu

March 3, 2017

ABSTRACT

Cold poses major physiological challenges to plants, especially long lived trees. In trees occurring along variable temperature clines, the expected direction and consequences of selection on cold acclimation ability and freezing tolerance are not straightforward. Here we estimated selection in cold acclimation genes at two evolutionary timescales in all seven species of the American live oaks (*Quercus* subsection *Virentes*). Two cold response candidate genes were chosen: ICE1, a key gene in the cold acclimation pathway, and HOS1, which modulates cold response by negatively regulating ICE1. Two housekeeping genes, GAPDH and CHS1, were also analyzed. At the shallow evolutionary timescale, we demonstrate that HOS1 experienced recent balancing selection in the two most broadly distributed species, *Q. virginiana* and *Q. oleoides*. At a deeper evolutionary scale, a codon based model of evolution revealed the signature of negative selection in ICE1. In contrast, three positively selected codons have been identified in HOS1, possibly a signature of the diversification of *Virentes* into warmer climates from a freezing adapted lineage of oaks. Our findings indicate that evolution has favored diversity in cold tolerance modulation through balancing selection in HOS1 while maintaining core cold acclimation ability, as evidenced by purifying selection in ICE1.

INTRODUCTION

Cold poses major challenges to plant performance and survival and is known to be an important driver of distribution ranges of species and clades (MacArthur 1972; Donoghue 2008; Zanne et al. 2013). Chilling temperatures (15 – 0 °C, Guy et al. 2007) decrease enzymatic activity, metabolic rates, and impair cellular transport due to reduced membrane fluidity (Thomashow 1998; Chinnusamy, Zhu, and Zhu 2007; Knight and Knight 2012). Freezing may cause the intracellular formation of ice crystals that perforate membranes and kill the cell. Extracellular ice formation, while not lethal to cells, leads to dehydration stress since the water potential of ice is more negative than that of liquid water (Cavender-Bares 2005; Chinnusamy, Zhu, and Zhu 2007).

Many plants are able to acquire cold tolerance by going through a series of physiological changes in cell wall and membrane structure and accumulating cryoprotective molecules, in a process known as cold acclimation (Cavender-Bares 2005; Janska et al. 2010). These changes have been long hypothesized to incur physiological costs to the plant, ultimately leading to a trade-off between the degree of cold tolerance on one hand and growth and reproductive rates on the other (MacArthur 1972; Loehle 1998). It has been empirically demonstrated that freezing tolerance is negatively correlated with growth rates, and that the degree of cold tolerance often correlates with the temperatures experienced by different populations (Figure 1a; Darychuk, Hawkins, and Stoehr 2012; Koehler, Center, and Cavender-Bares 2012). However, the mechanisms underlying this trade-off may be more complex than simple differential resource allocation (Savage and Cavender-Bares 2013).

Species and populations that are broadly distributed, span steep temperature gradients, or inhabit locations of highly inconstant climate will be exposed to variable selective environments in space and time (Hedrick 2006; Savolainen, Pyhäjärvi, and Knürr 2007; Zhen and Ungerer 2008). For example, the same degree of cold tolerance would not be beneficial across the entire range of broadly distributed species because of temperature differences across geographic regions (Kelly 2006; Darychuk, Hawkins, and Stoehr 2012; Koehler, Center, and Cavender-Bares 2012). Similarly, the same population may experience different climates at varying temporal scales, such that different levels of cold tolerance may be favored in different years or decades (Figure 1; Hedrick 2006; Vitasse, Lenz, and Korner 2014). Furthermore, selection associated with yearly and decadal temperature variation may have stronger effects in seedlings given that they are more susceptible to freezing stress (Figure 1b; Boorse et al. 1998).

Given the stark consequences of cold damage for survival and fitness and the potential trade-offs between cold tolerance and growth, it is expected that genetic variation underlying cold acclimation should be under selection (Zhen and Ungerer 2008; Zhen, Dhakal, and Ungerer 2011; Preston and Sandve 2013). The specific form of selection depends on many factors, including the degree of climatic variability experienced, degree of gene flow between populations, the role of specific genes in response to cold and freezing stress and particularly on the evolutionary timescale in question.

The outcome of selection for cold acclimation at shallow evolutionary scales, related to population level processes, depends on its interplay with climatic variability (Zhen and Ungerer 2008; Zhen, Dhakal, and Ungerer 2011). In predictable climatic conditions, selection is generally expected to purge suboptimal alleles, favoring locally adapted populations

where the mean cold response matches the home site conditions (Hedrick 2006; Savolainen, Pyhäjärvi, and Knürr 2007). Alternatively, balancing selection may favor the maintenance of diversity in cold tolerance in populations that experience strong climatic fluctuations, if different levels of cold tolerance are advantageous in different years (Turelli and Barton 2004; Hedrick 2006; Vitasse, Lenz, and Korner 2014). Signatures of selective dynamics at this scale are generally inferred from population models that use allele frequency and genetic structure information (Nielsen 2005; Keller et al. 2012; Vitti, Grossman, and Sabeti 2013).

Molecular evolution analyses at deeper timescales help to uncover the history of long term selective pressures that have acted on cold tolerance genes. It is known, for example, that codons of functionally indispensable genes tend to be constrained and show stronger signatures of purifying (negative) selection whereas genes involved in regulation or non-essential functions are generally less constrained (Anisimova and Kosiol 2009; Vitti, Grossman, and Sabeti 2013). Additionally, footprints of processes such as shifts in climatic niche may in some cases be recovered from molecular models. For instance, the evolution of cold-intolerant alleles from a cold-tolerant clade may have been favored as a clade invades warmer climates, leaving a signature of positive selection in certain codons.

The live oaks (*Quercus* subsection *Virentes*) provide a unique opportunity to investigate selection on genes associated with cold tolerance. *Virentes* spans the wide temperature gradient across the Tropical-Temperate divide of North America and Mesoamerica (Figure 2, Cavender-Bares et al. 2015). Three of its species, *Q. oleoides*, *Q. virginiana* and *Q. fusiformis*, are broadly distributed, and their populations face strikingly different climatic conditions (Cavender-Bares et al. 2015). Broadly distributed *Virentes* also experience strong seasonal and inter-annual climatic variability such that different tolerance levels may be favored in

different years. Furthermore, Koehler, Center, and Cavender-Bares (2012) have identified variation in freezing tolerance both among and within species in *Virentes*, demonstrating that population home temperatures were negatively correlated with and significantly predict freezing tolerance after cold acclimation. Interestingly, some species within the *Virentes* that only rarely experience chilling or freezing have retained some ability to express freezing tolerance.

Cold acclimation in plants is governed by a complex transcriptional pathway with two key upstream genes, ICE1 and HOS1 (Chinnusamy, Zhu, and Zhu 2007; M. R. Knight and Knight 2012). The ICE1 (Inducer of CBF expression) gene encodes a transcription factor essential to initiate the cold acclimation cascade. It induces the transcription of CBF genes (Chinnusamy, Zhu, and Zhu 2007; Miura and Furumoto 2013), promoting the expression of COR (Cold Responsive) genes that are responsible for many cold responses (Hannah, Heyer, and Hinch 2005), for instance, membrane restructuring and stabilization (e.g. COR15a and DHN1, Steponkus et al. 1998; Koag 2003), enzyme stabilization during freezing (COR85 Kazuoka and Oeda 1994) and accumulation of intracellular polysaccharides (Miura et al. 2012). On the other hand, the HOS1 (High Expression of Osmotically Responsive) gene negatively regulates cold response expression by the ubiquitination of ICE1 transcription factor (Dong et al. 2006; Miura and Furumoto 2013).

Our major goals were to estimate the direction of selection in two candidate genes associated with cold tolerance (ICE1 and HOS1) at two different evolutionary timescales within the seven live oak species. We first evaluated whether cold acclimation genes have mostly experienced recent diversifying selection or, alternatively, whether allelic diversity within and between populations is maintained by balancing selection. Second, we tested whether

SNP variation was associated with the climatic clines experienced by populations of different species. Finally, we estimated the long term selection patterns of the two candidate cold tolerance genes using phylogenetically informed codon-based models of molecular evolution. Our inferences of selection at both evolutionary scales were compared to results from two putative housekeeping genes, GAPDB and CHR11.

METHODS

SAMPLING AND FREEZING TOLERANCE MEASUREMENT

All seven live oak (*Quercus* subsection *Virentes*) species were sampled: *Quercus virginiana*, *Q. oleoides*, *Q. geminata*, *Q. minima*, *Q. fusiformis*, *Q. sangreana*, and *Q. brandegeei*. The number of sampled specimens from each species varied from seven for *Q. geminata* to 51 for *Q. oleoides*.

Error! Reference source not found. summarizes our sampling and includes population assignment, coordinates of the collection site, and number of individuals sampled.

Populations have been delimited based on previous genetic studies using microsatellite and chloroplast data (Cavender-Bares et al. 2011; Gugger and Cavender-Bares 2013; Deacon and Cavender-Bares 2015; Cavender-Bares et al. 2015) as well as on geography, and the choice of populations to sample was designed to cover the geographic and climatic range of each species (Cavender-Bares et al. 2015). Leaves were collected either from the field or from trees grown in the greenhouse at the University of Minnesota for a previous study (Koehler, Center, and Cavender-Bares 2012). Leaf tissue was stored at -80°C whether collected in the field or from the greenhouse.

Stem freezing tolerance data was taken from Koehler, Center, and Cavender-Bares (2012) and Cavender-Bares et al. (2015). Freezing tolerance was estimated with the electrolyte leakage method, which uses variation in electrical conductivity to estimate the degree of cell lysis in response to different freezing treatments (Murray, Cape, and Fowler 1989). More details on how the measurement were conducted are given in Koehler, Center, and Cavender-Bares (2012).

Two key cold tolerance genes were sampled, ICE1 and HOS1. As stated before, ICE1 is a fundamental upstream gene in the CBF expression cascade that induces several cold acclimation responses, while HOS1 negatively regulates ICE1 transcripts, acting to decrease cold tolerance (reviews in Chinnusamy, Zhu, and Zhu 2007; Miura and Furumoto 2013). Two putative housekeeping genes were chosen for the purposes of comparison: GAPDB (Glyceraldehyde-3-phosphate dehydrogenase subunit B), essential for glycolysis, and CHR11 (Chromatin Remodeling Protein), an Imitation Switch (ISWI) protein in a family of ATP-dependent chromatin remodeling factors. CHR11 is needed for the regulation of cellular identity in the vegetative phase and for nuclear proliferation during female gametogenesis (Li et al. 2012).

DNA EXTRACTION, AMPLIFICATION AND SEQUENCING

Regions in four genes (HOS1, ICE1, GAPDB and CHR11) were amplified and sequenced to explore the genetic variation in cold tolerance. The frozen leaf tissue was ground and used for genomic DNA extractions following the standard CTAB extraction protocol (Doyle and Doyle 1987). Primers were designed to start and end in an exon while spanning an intron region of each candidate gene, using the whole-genome databases for *Prunus persica* and

Vitis vinifera (**Error! Reference source not found.**). They were tested in *Quercus* by PCR, sequencing, and finally BLASTing the results simply to confirm that the expected genomic regions were amplified. The primer annealing temperatures were then optimized for binding to the live oak genomic DNA by running gradient PCRs. The 20 μ L PCR reactions included: 1 μ L of template DNA (at about 10 ng μ L⁻¹), 14.5 μ L ddH₂O, 2 μ L of 10x PCR buffer containing MgCl₂, 0.4 μ L dNTPs (40 mM total, each dNTP at 10 mM), 1 μ L forward primer (at 10 μ M), 1 μ L reverse primer (at 10 μ M) and 0.1 μ L Taq DNA Polymerase (5 units/ μ L). The thermal cycler amplification program performed 40 cycles of (1) denaturing for 30 s at 94°C, (2) annealing for 45 s at 58 °C – 63°C and (3) extension for 2 min at 72°C, flanked by an initial 2 min denaturing step at 94°C and final 10 min extension step at 72°C. PCR products were then prepared for sequencing at University of Notre Dame and Beckman Coulter Genomics. Prior to Sanger sequencing, Notre Dame PCR products were purified and quantified with a NanoDrop 2000, whereas Beckman Coulter samples were quantified with the Qubit dsDNA HS Assay. Sequencer (v 4.10.1) was used to visualize chromatograms, assess sequence quality and base calls, and identify heterozygous sites, which were labeled with their corresponding IUPAC ambiguity code. The processed sequences were finally exported in FASTA format.

IDENTIFYING CODING REGIONS, SEQUENCE ALIGNMENT AND SNP EXTRACTION

Coding regions were identified by BLASTing a representative *Virentes* sequence of each gene against Genbank's protein database using the blastX algorithm. First, we ensured that the resulting blastX proteins matched the queried gene. The top 10 blastX hits were downloaded as DNA and their coding sequences (CDS) were added to the *Virentes* fasta. Subsequently, multiple alignment was performed in the augmented dataset using the FFT-NS-i algorithm

in MAFFT v7 (Kato and Standley 2013), allowing us to identify the coding sequences in each gene and annotate them as such. Finally, the CDSs from the blastX results were removed from the alignment. For each gene, both coding-sequence-only and SNP-only alignments were generated by subsetting the full alignment. The HOS1 sequences included exon 2 and part of its flanking introns (749 bp, **Error! Reference source not found.**), ICE1 sequences encompassed exon 2 and its 3' intron (458 bp, **Error! Reference source not found.**), GAPDB sequences included exon 3 and part of its flanking introns (578 bp, **Error! Reference source not found.**) and CHR11 sequences encompassed three small exons (1, 2, 3) of along with their neighboring introns (685 bp, **Error! Reference source not found.**).

ANALYSES

Allelic diversity (A) and expected heterozygosity (H(E)) is provided for each gene for all seven species of *Virentes*. For each locus, expected heterozygosity H(E) was calculated using $H(E) = 1 - \sum_{i=1}^k p_i^2$ where p_i is the frequency of the i^{th} of k alleles. The analysis of recent selection was conducted with a Bayesian multinomial dirichlet model implemented in Bayescan (Foll and Gaggiotti 2008). In this model, population allele frequency differentiation from a common gene pool is measured with the F_{st} coefficient, which is decomposed with a logistic regression into a population-specific parameter β and a locus-specific component α . Significantly positive values of α suggest diversifying selection, whereas significantly negative α indicates balancing or purifying selection. Statistical significance for the model with selection was tested using Bayes Factors and evaluated in light of Jeffrey's interpretation (Kass and Raftery 1995). Input files for Bayescan were produced from the SNP alignments using a custom R script. Each of the seven species was analyzed separately, and populations were delimited as shown in **Error! Reference source not found.**. The Bayescan

MCMC chain was first tuned by running 20 independent pilot runs of length 5000, and then sampled for 100000 iterations. The first 50000 samples were discarded as burn-in and the remaining samples were thinned to every 10th iteration. The default Bayescan prior odds for the neutral model, with no selection, is 10. We have set prior odds 20 instead to minimize the chance of false positives.

The association between SNPs and freezing tolerance and temperature cline was investigated through generalized linear models (GLM) performed in R. For each individual, the minimum temperature at the coldest month (bioclim6) was extracted from the bioclim raster (Hijmans et al. 2005) at the 30arcsec resolution. Freezing tolerances and minimum temperature values were taken to be the dependent variable to be explained by each SNP, which entered the model as the independent variables. Population latitude was used as a covariate to minimize the chance of false positives, given that neutral SNPs may covary with environment because of population structure (Yu et al. 2005, Heerwaarden et al. 2015), especially north-south population differentiation (Eckert et al. 2010). Model selection using stepwise AIC was performed and t-values associated with each SNP were recorded and p-values lower than 0.05 were considered statistically significant.

Selective dynamics at deeper evolutionary time was inferred at the gene level for all species jointly using a phylogenetic model of codon evolution implemented in FUBAR (Fast Unconstrained Bayesian Approximation) developed by Murrell et al. (2013). It infers the rates of non-synonymous (dN) and synonymous (dS) substitutions for each codon, which are expected to be the same under neutral evolution. Positive selection acts to increase dN , leading to a positive $dN - dS$, whereas negative selection increases synonymous rates (dS), yielding negative $dN - dS$ values. The coding sequences of each candidate gene were

subjected to a FUBAR analysis performed with HyPhy in the datamonkey server (<http://www.datamonkey.org>). FUBAR's default neighbor joining trees used in the analyses were calculated with a HKY model of DNA evolution. We have also subjected our datasets to the Branch-Site REL model (Kosakovsky Pond et al 2011) which, unlike FUBAR, also identifies which lineages account for the inferred selection patterns. The Branch-Site REL was also run with HyPhy in the datamonkey server.

RESULTS

RECENT SELECTION ESTIMATED FROM ALLELE FREQUENCIES

Allelic diversity patterns were strikingly different among the four sampled genes. Across all species, we found 17 alleles of ICE1, 32 alleles of GAPDB, 55 alleles of CHR11 and 110 alleles of HOS1. The disparity in allelic diversity between ICE1 and HOS1 was also found within the two most widespread species: *Q. virginiana* had 7 alleles of ICE1 and 24 of HOS1 (**Error! Reference source not found.**), while 6 alleles of ICE1 and 29 alleles of HOS1 were found in *Q. oleoides* (**Error! Reference source not found.**). It should be noted that polymorphisms in non-coding regions of the genes were also considered when identifying alleles.

Bayescan detected very strong evidence for recent balancing selection in HOS1 in the two most widely distributed species: *Q. oleoides* had an α of -2.191 and $\log_{10}$ Bayes Factor (logBF) of 3.3 while *Q. virginiana*'s estimated α was -1.935 and LogBF of 2.0 (Figure 3, Table 1). No evidence for selection was found in any other species, including the broadly distributed *Q. fusiformis*. Bayescan did not detect significant evidence for selection in the ICE1 gene nor in any of the housekeeping genes (Figure 3, Table 1).

GENE LEVEL SELECTION AT PHYLOGENETIC TIMESCALES

We considered codons to significantly deviate from neutral evolution in the FUBAR (Murrell et al. 2013) analyses whenever the posterior probability for either positive or negative selection were higher than 0.85 and empirical Bayes factors (logBF) were at least substantial (Kass and Raftery 1995). Three HOS1 codons, 31, 145 and 146, showed significant (logBF 0.95, 1.21 and 1.58) signatures of positive selection as indicated by positive $dN - dS$ values, while three others, 30, 101 and 103, showed negative selection (Figure 4, logBF 2.17, 0.93 and 0.82). In contrast, FUBAR only estimated significantly negative $dN - dS$ values in ICE1, specifically on codons 10, 65 and 133 (logBF 1.16, 1.1 and 0.99). FUBAR inferred pervasive negative selection (10 codons with negative $dN - dS$) in GAPDB (Figure 4), but found weak evidence for selection in the putative housekeeping gene CHR11, revealing one codon under positive and another under negative selection. The Branch-Site REL model (Kosakovsky Pond et al 2011) did not identify significant differences in dN/dS rates across lineages in any of the genes according to a stepwise likelihood ratio test and Bonferroni corrected p-values.

ASSOCIATIONS BETWEEN CODING REGION VARIATION AND FREEZING TOLERANCE

The generalized linear models (GLMs) uncovered significant associations between HOS1 polymorphisms (SNPs) and both individual freezing tolerance (Figure 5, **Error! Reference source not found.** and **Error! Reference source not found.**) and minimum temperature at the collection locations (**Error! Reference source not found.** and **Error! Reference source not found.**) in *Q. oleoides* and *Q. virginiana*. The effect of a SNP on freezing tolerance was never significant in both species. For example, the A allele in SNP 282 was significantly (p-value 0.007) associated with less freezing tolerance in *Q. oleoides* but not in *Q. virginiana*. The only

ICE1 polymorphism significantly associated with the freezing tolerance and minimum temperatures was SNP 402 in *Q. oleoides* (p-value: 0.002).

DISCUSSION

We have found evidence that the diversity in HOS1, a key regulatory gene of cold acclimation, has been maintained through balancing selection (Table 1, Figure 3). This finding indicates that evolution has favored high standing genetic diversity in *Virentes*, which are distributed broadly across the tropical temperate divide and experience high climatic variability. These results are compatible with empirical observations of wide variation in cold tolerance within and among those species (Koehler, Center, and Cavender-Bares 2012; Cavender-Bares et al. 2015). Our analysis of codon evolution demonstrates that HOS1 has experienced positive selection in the past (Figure 4), possibly a signature of the diversification of *Virentes* into warmer climates from a freezing adapted lineage of oaks.

ENVIRONMENTAL VARIABILITY AND THE MAINTENANCE OF GENETIC DIVERSITY THROUGH BALANCING SELECTION.

Our results suggest that variation in the cold tolerance gene HOS1 has been maintained by recent balancing selection in *Q. virginiana* and *Q. oleoides* (Table 1, Figure 3). Interestingly, they are the two most broadly distributed of the seven species of *Virentes* and span major variation in climate (Cavender-Bares et al. 2015). We did not detect a significant signature of selection in HOS1 in *Q. fusiformis*, another broadly distributed species, possibly because our sample size was limited for this species.

In principle, negative α parameters estimated from Bayescan may reflect two different evolutionary processes, balancing or possibly purifying selection, since both may act to decrease population differentiation (Foll and Gaggiotti 2008; Foll et al. 2010). However, because purifying selection acts to reduce genetic variability (T Mitchell-Olds, Willis, and Goldstein 2007), the high allelic diversity found in HOS1 strongly supports the interpretation that HOS1 has experienced balancing selection.

The inference of balancing selection in HOS1 provides a possible explanation for the observed variability in freezing tolerance in *Virentes*. Studies based on either experimental manipulations of saplings grown in controlled environments (Cavender-Bares 2007; Koehler, Center, and Cavender-Bares 2012) or under field conditions (other evergreen oaks, Cavender-Bares et al. 2005; Gimeno et al. 2008) have documented high variability in freezing tolerance within and between populations.

Balancing selection associated with environmental variability has been inferred in other plant and animal systems. For instance, Weinig et al. (2003) conducted a transplant experiment in *Arabidopsis thaliana* and suggested that allelic diversity observed in several QTLs were maintained by environmental heterogeneity and epistatic selection. Schmidt, Bertness, and Rand (2000) have shown that the maintenance of diversity in the *MPI* gene in the northern acorn barnacle was associated with varying levels of environmental stress. Balancing selection has also been detected in genetic variation underlying responses to biological stress. For example, that has been suggested that balancing selection has been responsible for the maintenance of polymorphisms in *R*-genes, related to pathogen resistance in *Arabidopsis* (Tian et al. 2002; Mitchell-Olds and Schmitt 2006). Similarly, a long term study of immune system genes (MHC) in sheep (Charbonnel and Pemberton 2005)

demonstrated that MHC was under balancing selection, likely in response to variation in parasite load.

MOLECULAR SELECTION AT DEEPER EVOLUTIONARY SCALES

In contrast to allele-frequency based methods, which capture recent evolution, codon-based models of molecular evolution such as FUBAR (Fast Unconstrained Bayesian Approximation, Murrell et al. 2013) uncover the long term selective dynamics that acted on coding regions of a gene (Anisimova and Kosiol 2009; Arenas 2015). Although our FUBAR analysis detected codons under selection in all four genes (Figure 4), the Branch-Site REL model was not able to identify which lineages accounted for the inferred selection. This result is not surprising, however, since the Branch-Site REL model is parameter rich and less likely to perform well in small datasets (Kosakovsky Pond et al 2011).

Our FUBAR molecular evolution analysis of the ICE1 gene identified three codons under strong purifying selection (significantly negative $dN - dS$, Figure 4), but no positive selection. Additionally, we found the allelic diversity of ICE1 to be very low when compared to HOS1. These findings suggest that ICE1 has been functionally constrained by selection over long evolutionary timescales, which was expected given the pivotal role that ICE1 plays in cold temperature sensing and initiation of the acclimation cascade in plants (Chinnusamy, Zhu, and Zhu 2007; Miura and Furumoto 2013). Deleterious mutations in ICE1 have been shown to impair the expression of CBFs and lead to significant decreases in acclimation potential and cold tolerance in *Arabidopsis* (Chinnusamy et al. 2003). The observation that ICE1's function is conserved across flowering plants, from rice (*Oryza sativa*) to tomatoes (*Solanum lycopersicum*) and aspen (*Populus tremula*), further supports the idea that it has been

under purifying selection on a deep evolutionary scale (Miura et al. 2012; Miura and Furumoto 2013).

On the other hand, we identified three HOS1 codons that experienced past positive selection (positive $dN - dS$, Figure 4) and three other codons that have undergone negative selection (negative $dN - dS$, Figure 4). This indicates that different HOS1 alleles have experienced directional selection at some point during the evolution of *Virentes*, possibly related to differences in cold tolerance. HOS1 negatively regulates cold response by physically interacting with ICE1 transcription factors and mediating their ubiquitination and degradation. Therefore, cold sensitivity and acclimation responses can be modulated by HOS1, whether through differential affinity to ICE1 by different HOS1 alleles or differential HOS1 expression. In any case, it is important to note that deleterious mutations in HOS1 lead to less stark consequences than the loss of ICE1 function. Dong et al. (2006) and others have shown that under-expression or loss of function in HOS1 leads to the accumulation of ICE1 products and consequently enhance cold acclimation responses (Chinnusamy, Zhu, and Zhu 2007). In contrast, over-expression of HOS1 leads to a reduction in cold acclimation and freezing tolerance, since ICE1 transcripts become degraded at higher rates (Dong et al. 2006; Chinnusamy, Zhu, and Zhu 2007; Miura and Furumoto 2013).

Certain HOS1 SNPs are associated with stem freezing tolerance, minimum temperature at the sample location, or both, even though population latitude was the most important explanatory covariate in all models (tables **Error! Reference source not found.** to **Error! Reference source not found.**). This result indicates that climatic variation has been a factor influencing genetic variation (Figure 5). Interestingly, the SNPs associated with freezing tolerance generally match codons showing a strong signature of selection (Figure 4),

providing additional evidence that the molecular evolution results are biologically meaningful and not false positives. Interpreting the direction of the effect of SNPs on freezing tolerance is not straightforward and out of the scope of this study. The freezing tolerance phenotype is the result of the interaction of different HOS1 mutations, ICE1 alleles and many other components not investigated here.

Our data cannot directly address which pressures the positively selected HOS1 sites were responding to during evolution. Nevertheless, we hypothesize that the positive selection in HOS1 is a consequence of the expansion of *Virentes* into warmer climates. The ancestral HOS1 allele was likely optimized to confer cold acclimation ability since live oaks likely evolved from freezing tolerant lineages (Cavender-Bares et al. 2015; Hipp et al. 2014) and the colonization of tropical and subtropical regions by *Q. oleoides* and *Q. virginiana* (Figure 2) is relatively recent (Cavender-Bares et al. 2015). The intrinsic trade-off between cold tolerance and growth rates in *Virentes* Koehler, Center, and Cavender-Bares (2012) may have favored phenotypes less responsive to cold but with faster growth in species inhabiting warmer climates. As a consequence, HOS1 alleles with higher binding affinity to ICE1 might have been positively selected.

Long-lived trees that span broad geographic ranges must cope with highly variable climatic conditions, such that selective pressures to enhance growth or stress tolerance vary through space and time. Our findings suggest that in broadly distributed live oak species, a resilient response to climate change occurs on a population level, by maintaining the standing genetic variation in key regulatory genes. The capacity for flexibility ensures survival of at least some individuals during episodic extreme events (e.g. severe cold spells or late spring freezes), buffering the species against extinction.

REFERENCES

- Anisimova, Maria, and Carolin Kosiol. 2009. "Investigating protein-coding sequence evolution with probabilistic codon substitution models." *Molecular Biology and Evolution* 26 (2): 255–71. doi:10.1093/molbev/msn232.
- Arenas, Miguel. 2015. "Trends in substitution models of molecular evolution." *Frontiers in Genetics* 6 (OCT). doi:10.3389/fgene.2015.00319.
- Boorse, Graham. C., Ewers, F. W., & Davis, S. D. 1998. "Response of chaparral shrubs to below-freezing temperatures: Acclimation, ecotypes, seedlings vs. adults". *American Journal of Botany*, 85(9): 1224–1230. doi.org/10.2307/2446631.
- Cavender-Bares, Jeannine. 2005. "Impacts of freezing on long-distance transport in woody plants." In *Vascular Transport in Plants*, edited by N Michelle Holbrook and Maciej A Zwieniecki, 401–24. doi:10.1016/B978-0-12-088457-5.50021-6.
- — —. 2007. "Chilling and freezing stress in live oaks (*Quercus* section *Virentes*): Intra- and inter-specific variation in PS II sensitivity corresponds to latitude of origin." *Photosynthesis Research* 94 (2-3): 437–53. doi:10.1007/s11120-007-9215-8.
- Cavender-Bares, Jeannine, P Cortes, S Rambal, R Joffre, B Miles, and A Rocheteau. 2005. "Summer and winter sensitivity of leaves and xylem to minimum freezing temperatures: A comparison of co-occurring Mediterranean oaks that differ in leaf lifespan." *New Phytologist* 168 (3): 597–612. doi:10.1111/j.1469-8137.2005.01555.x.
- Cavender-Bares, Jeannine, Antonio Gonzalez-Rodriguez, Pahlich, A., Koehler, K., and Deacon, Nicholas. 2011. "Phylogeography and climatic niche evolution in live oaks (*Quercus* series

Virentes) from the tropics to the temperate zone". *Journal of Biogeography* 38(5): 962–981.

doi.org/10.1111/j.1365-2699.2010.02451.x

Cavender-Bares, Jeannine, Antonio González-Rodríguez, Deren A R Eaton, Andrew L Hipp, Anne Beulke, and Paul S Manos. 2015. "Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach." *Molecular Ecology* 24 (14): 3668–87. doi:10.1111/mec.13269.

Charbonnel, N, and J Pemberton. 2005. "A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection." *Heredity* 95 (5): 377–88. doi:10.1038/sj.hdy.6800735.

Chinnusamy, Viswanathan, Masaru Ohta, Siddhartha Kanrar, Byeong-Ha Lee, Xuhui Hong, Manu Agarwal, and Jian-Kang Zhu. 2003. "ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*." *Genes & Development* 17 (8): 1043–54. doi:10.1101/gad.1077503.

Chinnusamy, Viswanathan, Jianhua Zhu, and Jian-Kang Zhu. 2007. "Cold stress regulation of gene expression in plants." *Trends in Plant Science* 12 (10): 444–51. doi:10.1016/j.tplants.2007.07.002.

Darychuk, N., B.J. Hawkins, and M. Stoehr. 2012. "Trade-offs between growth and cold and drought hardiness in subarctic Douglas-fir." *Canadian Journal of Forest Research* 42 (8): 1530–41. doi:10.1139/x2012-092.

Deacon, Nicholas J. and Cavender-Bares, Jeannine 2015. "Limited Pollen Dispersal Contributes to Population Genetic Structure but not Local Adaptation in *Quercus oleoides* Forests of Costa Rica". *PLOS ONE* 10(9): 25. doi.org/10.1371/journal.pone.0138783

- Dong, Chun-Hai, Manu Agarwal, Yiyue Zhang, Qi Xie, and J.-K. Zhu. 2006. "The negative regulator of plant cold responses, HOS1, is a RING E3 ligase that mediates the ubiquitination and degradation of ICE1." *Proceedings of the National Academy of Sciences* 103 (21): 8281–6. doi:10.1073/pnas.0602874103.
- Donoghue, M.J. 2008. "A phylogenetic perspective on the distribution of plant diversity." *Proceedings of the National Academy of Sciences* 105 (1): 11549–55. doi:10.1073/pnas.0801962105.
- Doyle, Jeff J, and Jane L Doyle. 1987. "A rapid DNA isolation procedure for small quantities of fresh leaf tissue." *Phytochemistry Bulletin* 19: 11–15.
- Eckert, Andrew J., van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., Gonzalez-Martinez, S. C. and Neale, D. B. 2010. "Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (*Pinus taeda* L., Pinaceae)". *Genetics* 185(3): 969–982. doi.org/10.1534/genetics.110.115543
- Foll, Matthieu, and Oscar E Gaggiotti. 2008. "A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective." *Genetics* 180 (2): 977–93. doi:10.1534/genetics.108.092221.
- Foll, Matthieu, Martin C Fischer, Gerald Heckel, and Laurent Excoffier. 2010. "Estimating population structure from AFLP amplification intensity." *Molecular Ecology* 19 (21): 4638–47. doi:10.1111/j.1365-294X.2010.04820.x.
- Gimeno, Teresa E, B. Pias, José P Lemos-Filho, and Fernando Valladares. 2008. "Plasticity and stress tolerance override local adaptation in the responses of Mediterranean holm oak seedlings to drought and cold." *Tree Physiology* 29 (1): 87–98. doi:10.1093/treephys/tpn007.

Gugger, Paul. F. and Cavender-Bares, Jeannine. 2013. "Molecular and morphological support for a Florida origin of the Cuban oak". *Journal of Biogeography* 40(4): 632–645. doi.org/10.1111/j.1365-2699.2011.02610.x

Guy, Charles, Fatma Kaplan, Joachim Kopka, Joachim Selbig, and Dirk K. Hinch. 2007. "Metabolomics of temperature stress." *Physiologia Plantarum* 132 (2): 071124124159002-??? doi:10.1111/j.1399-3054.2007.00999.x.

Hannah, Matthew A, Arnd G Heyer, and Dirk K Hinch. 2005. "A Global Survey of Gene Regulation during Cold Acclimation in *Arabidopsis thaliana*." *PLOS Genet* 1 (2): 893–902. doi:10.1371/JOURNAL.PGEN.0010026.

Hedrick, Philip W. 2006. "Genetic Polymorphism in Heterogeneous Environments: The Age of Genomics." *Annual Review of Ecology, Evolution, and Systematics* 37 (1). Annual Reviews: 67–93. doi:10.1146/annurev.ecolsys.37.091305.110132.

Heerwaarden, Joost., van Zanten, M. and Kruijer, W. 2015. "Genome-Wide Association Analysis of Adaptation Using Environmentally Predicted Traits". *PLOS Genetics* 11(10): 23. doi.org/10.1371/journal.pgen.1005594

Hijmans, Robert J, Susan E Cameron, Juan L Parra, Peter G Jones, and Andy Jarvis. 2005. "Very high resolution interpolated climate surfaces for global land areas." *International Journal of Climatology* 25 (15): 1965–78. doi:10.1002/joc.1276.

Hipp, Andrew L, Deren A R Eaton, Jeannine Cavender-Bares, Elisabeth Fitzek, Rick Nipper, and Paul S Manos. 2014. "A Framework Phylogeny of the American Oak Clade Based on Sequenced RAD Data." *PLoS ONE* 9 (4): e93975. doi:10.1371/journal.pone.0093975.

- Janska, A, P Marsik, S Zelenkova, and J Ovesna. 2010. "Cold stress and acclimation - what is important for metabolic adjustment?" *Plant Biology* 12 (3): 395–405. doi:10.1111/j.1438-8677.2009.00299.x.
- Kass, Robert E, and Adrian E Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90 (430): 773–95. doi:10.1080/01621459.1995.10476572.
- Katoh, Kazutaka, and Daron M Standley. 2013. "MAFFT multiple sequence alignment software version 7: Improvements in performance and usability." *Molecular Biology and Evolution* 30 (4): 772–80. doi:10.1093/molbev/mst010.
- Kazuoka, Toru, and Kenji Oeda. 1994. "Purification and Characterization of COR85-Oligomeric Complex from Cold-Acclimated Spinach." Oxford University Press.
<http://pcp.oxfordjournals.org/content/35/4/601.full.pdf+html>.
- Keller, Stephen, Levsen, Nicholas, Olson, Matthew and Tiffin, Peter. 2012. "Local adaptation in the flowering-time gene network of balsam poplar, *Populus balsamifera* L." *Molecular Biology and Evolution* 29(10): 3143–3152. doi.org/10.1093/molbev/mss121
- Kelly, John K. 2006. "Geographical Variation in Selection, from Phenotypes to Molecules." *The American Naturalist* 167 (4): 481–95. doi:10.1086/501167.
- Knight, Marc R., and Heather Knight. 2012. "Low-temperature perception leading to gene expression and cold tolerance in higher plants." *New Phytologist* 195 (4): 737–51.
doi:10.1111/j.1469-8137.2012.04239.x.

- Koag, M.-C. 2003. "The binding of Maize DHN1 to Lipid Vesicles. Gain of Structure and Lipid Specificity." *Plant Physiology* 131 (1). American Society of Plant Biologists: 309–16.
doi:10.1104/pp.011171.
- Koehler, Kari, Alyson Center, and Jeannine Cavender-Bares. 2012. "Evidence for a freezing tolerance-growth rate trade-off in the live oaks (*Quercus* series *Virentes*) across the tropical-temperate divide." *New Phytologist* 193 (3): 730–44. doi:10.1111/j.1469-8137.2011.03992.x.
- Kosakovsky Pond, Sergei., Murrell, B., Fourment, M., Frost, S. D. W., Delport, W., & Scheffler, K. 2011. "A Random Effects Branch-Site Model for Detecting Episodic Diversifying Selection". *Molecular Biology and Evolution* 28(11): 3033–3043. doi.org/10.1093/molbev/msr125
- Li, Guang, Jiawei Zhang, Jiqin Li, Zhongnan Yang, Hai Huang, and Lin Xu. 2012. "Imitation Switch chromatin remodeling factors and their interacting RINGLET proteins act together in controlling the plant vegetative phase in *Arabidopsis*." *The Plant Journal* 72 (2): 261–70.
doi:10.1111/j.1365-313X.2012.05074.x.
- Loehle, Craig. 1998. "Height growth rate tradeoffs determine northern and southern range limits for trees." *Journal of Biogeography* 25 (4): 735–42. doi:10.1046/j.1365-2699.1998.2540735.x.
- MacArthur, Robert. 1972. *Geographical Ecology: Patterns in the Distribution of Species*. New York: Harper.
- Mitchell-Olds, T, J H Willis, and D B Goldstein. 2007. "Which evolutionary processes influence natural genetic variation for phenotypic traits?" *Nature Reviews Genetics* 8 (11): 845–56.
doi:10.1038/nrg2207.

- Mitchell-Olds, Thomas, and Johanna Schmitt. 2006. "Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*." *Nature* 441 (7096): 947–52.
doi:10.1038/nature04878.
- Miura, Kenji, and Tsuyoshi Furumoto. 2013. "Cold signaling and cold response in plants." *International Journal of Molecular Sciences* 14 (3): 5312–37. doi:10.3390/ijms14035312.
- Miura, Kenji, Hayato Shiba, Masaru Ohta, Seung Won Kang, Ayaka Sato, Takashi Yuasa, Mari Iwaya-Inoue, Hiroshi Kamada, and Hiroshi Ezura. 2012. "SlICE1 encoding a MYC-type transcription factor controls cold tolerance in tomato, *Solanum lycopersicum*." *Plant Biotechnology* 29 (3): 253–60. doi:10.5511/plantbiotechnology.12.0303a.
- Murray, M B, J N Cape, and D Fowler. 1989. "Quantification of frost damage in plant tissues by rates of electrolyte leakage." *New Phytologist* 113: 307–11. doi:10.1111/j.1469-8137.1989.tb02408.x.
- Murrell, Ben, Sasha Moola, Amandla Mabona, Thomas Weighill, Daniel Sheward, Sergei L Kosakovsky Pond, and Konrad Scheffler. 2013. "FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection." *Molecular Biology and Evolution* 30 (5): 1196–1205.
doi:10.1093/molbev/mst030.
- Nielsen, Rasmus. 2005. "Molecular Signatures of Natural Selection." *Annual Review of Genetics* 39 (1): 197–218. doi:10.1146/annurev.genet.39.073003.112420.
- Preston, Jill C, and Simen R Sandve. 2013. "Adaptation to seasonality and the winter freeze." *Frontiers in Plant Science* 4 (12): 3–1. doi:10.3389/fpls.2013.00167.

- Savage, Jessica A, and Jeannine Cavender-Bares. 2013. "Phenological cues drive an apparent trade-off between freezing tolerance and growth in the family Salicaceae." *Ecology* 94 (8): 1708–17. doi:10.1890/12-1779.1.
- Savolainen, Outi, Tanja Pyhäjärvi, and Timo Knürr. 2007. "Gene Flow and Local Adaptation in Trees." *Annual Review of Ecology, Evolution, and Systematics* 38 (1). Annual Reviews: 595–619. doi:10.1146/annurev.ecolsys.38.091206.095646.
- Schmidt, Paul S., Mark D. Bertness, and David M. Rand. 2000. "Environmental heterogeneity and balancing selection in the acorn barnacle *Semibalanus balanoides*." *Proceedings of the Royal Society of London B: Biological Sciences* 267 (1441).
- Steponkus, Peter L, Matsuo Uemura, Raymond A Joseph, Sarah J Gilmour, and Michael F Thomashow. 1998. "Mode of action of the COR15a gene on the freezing tolerance of *Arabidopsis thaliana*." *Proceedings of the National Academy of Sciences* 95 (24): 14570–5. doi:10.1073/pnas.95.24.14570.
- Thomashow, Michael F. 1998. "Role of cold-responsive genes in plant freezing tolerance." *Plant Physiology* 118 (1): 1–8. doi:10.1104/PP.118.1.1.
- Tian, Dacheng, Hitoshi Araki, Eli Stahl, Joy Bergelson, and Martin Kreitman. 2002. "Signature of balancing selection in *Arabidopsis*." *Proceedings of the National Academy of Sciences* 99 (17): 11525–30. doi:10.1073/pnas.172203599.
- Turelli, Michael, and N. H. Barton. 2004. "Polygenic Variation Maintained by Balancing Selection: Pleiotropy, Sex-Dependent Allelic Effects and G x E Interactions." *Genetics* 166 (2): 1053–79. doi:10.1534/genetics.166.2.1053.

- Vitasse, Yann, Armando Lenz, and Christian Korner. 2014. "The interaction between freezing tolerance and phenology in temperate deciduous trees." *Frontiers in Plant Science* 5 (October): 541. doi:10.3389/fpls.2014.00541.
- Vitti, Joseph J, Sharon R Grossman, and Pardis C Sabeti. 2013. "Detecting Natural Selection in Genomic Data." *Annual Review of Genetics* 47 (1): 97–120. doi:10.1146/annurev-genet-111212-133526.
- Weinig, Cynthia, Lisa A. Dorn, Nolan C. Kane, Zachary M. German, Solveig S. Halldorsdottir, Mark C. Ungerer, Yuko Toyonaga, Trudy F. C. Mackay, Michael D. Purugganan, and Johanna Schmitt. 2003. "Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*." *Genetics* 329 (September): 321–29. doi:10.1086/509587.
- Yu, Jianming, Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., McMullen, M., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S. and Buckler, E.S. 2006. "A unified mixed-model method for association mapping that accounts for multiple levels of relatedness". *Nature Genetics* 38(2): 203–208. doi.org/10.1038/ng1702
- Zanne, Amy E, David C Tank, William K Cornwell, Jonathan M Eastman, Stephen A Smith, Richard G FitzJohn, Daniel J McGlinn, et al. 2013. "Three keys to the radiation of angiosperms into freezing environments." *Nature* 506 (7486): 89–92. doi:10.1038/nature12872.
- Zhen, Ying, and Mark C Ungerer. 2008. "Relaxed Selection on the CBF/DREB1 Regulatory Genes and Reduced Freezing Tolerance in the Southern Range of *Arabidopsis thaliana*." *Molecular Biology and Evolution* 25 (12): 2547–55. doi:10.1093/molbev/msn196.

Zhen, Ying, Preeti Dhakal, and Mark C. Ungerer. 2011. "Fitness Benefits and Costs of Cold Acclimation in *Arabidopsis thaliana*." *The American Naturalist* 178 (1): 44–52.
doi:10.1086/660282.

Draft

Table 1. Bayescan results for two housekeeping and two candidate genes in the three most broadly distributed live oaks. Negative values of the selection parameter α indicate balancing selection. Models with $\log_{10}(BF) > 1.0$ were considered significant and are highlighted in grey.

Gene	Species	LogBF	alpha
CHR11	<i>Q. fusiformis</i>	-1.03	-0.103
CHR11	<i>Q. oleoides</i>	-1.11	-0.057
CHR11	<i>Q. virginiana</i>	-0.93	-0.118
GAPDB	<i>Q. fusiformis</i>	-1.18	-0.042
GAPDB	<i>Q. oleoides</i>	-1.28	-0.031
GAPDB	<i>Q. virginiana</i>	0.53	-1.367
HOS1	<i>Q. fusiformis</i>	-1.17	-0.041
HOS1	<i>Q. oleoides</i>	3.3	-2.191
HOS1	<i>Q. virginiana</i>	2	-1.935
ICE1	<i>Q. fusiformis</i>	-1.26	-0.037
ICE1	<i>Q. oleoides</i>	-1.3	-0.028
ICE1	<i>Q. virginiana</i>	-1.19	-0.037

Figure 1. Conceptual model of how (A) the cold tolerance / growth rate tradeoff may lead to balancing selection provided that (B) alleles confer different tolerances and that seedlings are more susceptible to cold damage than adults. C) A warm year selects against seedlings carrying the green allele (high freezing tolerance, slow growth) more strongly than the adults. However, the green allele persists to a cold year in the adults, when it has an advantage over low freezing tolerance (orange allele) individuals. This process results in the maintenance both alleles in the population.

Figure 2. Geographic location of populations sampled in this study, broadly covering the ranges of all seven species of live oaks (*Virentes*). Map colors show the minimum temperature in the coldest month and highlight the temperature variation experienced by distributed species. The population delimitation for each species is provided in **Error! Reference source not found.**

Figure 3. Bayescan results pooled across the independent analyses for each of the seven *Virentes* species. Negative alpha estimates (y axis) indicate balancing selection whereas positive values would suggest positive selection. Log10 Bayes factors (LogBF) greater than 0.5 (grey box) indicates a significant deviation from a null model with no selection. See **Error! Reference source not found.** for details on population delimitation for each species.

Figure 4. Codon selection profiles across the candidate genes HOS1 and ICE1 (first column) and putative housekeeping genes GAPDB and CHR11 (second column) inferred from a FUBAR analysis in HyPhy. Sites that experienced significant negative selection ($dN - dS < 0$) are highlighted in orange whereas sites sites that experienced significant positive selection ($dN - dS > 0$) are shown in blue. HOS1 codons significantly associated with freezing tolerance or minimum temperature (tables **Error! Reference source not found.**, **Error! Reference source not found.**, **Error! Reference source not found.**, and **Error! Reference source not found.**) are marked with red asterisks.

Figure 5. Examples of the association between different HOS1 SNPs in two broadly distributed live oak species, *Q. virginiana* (green) and *Q. oleoides* (purple) and freezing tolerance measured with the electrolyte leakage method. Significant effects in intraspecific comparisons are marked with an asterisk. Statistical details from the GLMs are shown in **Error! Reference source not found.** and **Error! Reference source not found.**

Figure 1. Conceptual model of how (A) the cold tolerance / growth rate tradeoff may lead to balancing selection provided that (B) alleles confer different tolerances and that seedlings are more susceptible to cold damage than adults. C) A warm year selects against seedlings carrying the green allele (high freezing tolerance, slow growth) more strongly than the adults. However, the green allele persists to a cold year in the adults, when it has an advantage over low freezing tolerance (orange allele) individuals. This process results in the maintenance both alleles in the population.

136x135mm (300 x 300 DPI)

Figure 2. Geographic location of populations sampled in this study, broadly covering the ranges of all seven species of live oaks (Virentes). Map colors show the minimum temperature in the coldest month and highlight the temperature variation experienced by distributed species. The population delimitation for each species is provided in Table 2.

101x72mm (600 x 600 DPI)

Figure 3. Bayescan results pooled across the independent analyses for each of the seven Virentes species. Negative alpha estimates (y axis) indicate balancing selection whereas positive values would suggest positive selection. Log10 Bayes factors (LogBF) greater than 0.5 (grey box) indicates a significant deviation from a null model with no selection. See Table 2 for details on population delimitation for each species.

123x100mm (600 x 600 DPI)

Figure 4. Codon selection profiles across the candidate genes HOS1 and ICE1 (first column) and putative housekeeping genes GAPDB and CHR11 (second column) inferred from a FUBAR analysis in HyPhy. Sites that experienced significant negative selection (●) are highlighted in orange whereas sites that experienced significant positive selection (●) are shown in blue. HOS1 codons significantly associated with freezing tolerance or minimum temperature (tables Table 5, Table 6, Table 7 and Table 8) are marked with red asterisks.

177x124mm (300 x 300 DPI)

Figure 5. Examples of the association between different HOS1 SNPs in two broadly distributed live oak species, *Q. virginiana* (green) and *Q. oleoides* (purple) and freezing tolerance measured with the electrolyte leakage method. Significant effects in intraspecific comparisons are marked with an asterisk. Statistical details from the GLMs are shown in Table 5 and Table 6.

76x28mm (600 x 600 DPI)

Draft

SUPPLEMENTARY MATERIAL

Table 1. Summary of our sampling and population delimitation.

Species	Country	Population	Lat	Lon	N
<i>Q. brandegeei</i>	Mexico	San Baratolo	23.766	-110.0152	4
<i>Q. brandegeei</i>	Mexico	Sierra la Laguna	23.5001	-110.0635	4
<i>Q. brandegeei</i>	Mexico	Valle perdido	23.7048	-110.1388	4
<i>Q. fusiformis</i>	Mexico	Cerro del Diente	22.591389	-97.906389	2
<i>Q. fusiformis</i>	Mexico	Monterrey Saltillo	25.285833	-99.946667	8
<i>Q. fusiformis</i>	Mexico	San Carlos	24.35192	-98.57465	4
<i>Q. fusiformis</i>	Mexico	Victoria Jaumave	23.37551	-99.12281	8
<i>Q. fusiformis</i>	USA	Granbury	32.2814	-97.7342	6
<i>Q. geminata</i>	USA	Carolina Beach	34.035	-77.894	3
<i>Q. geminata</i>	USA	Suwannee	30.3867	-83.1749	4
<i>Q. minima</i>	USA	Jonathon Dickinson State Park	27.01117884	-80.12183223	5
<i>Q. minima</i>	USA	San Felasco	29.65276899	-82.28045862	6
<i>Q. oleoides</i>	Belize	Berril Boom	17.6283	-88.3676	9
<i>Q. oleoides</i>	Belize	Mountain Pine Ridge	16.9796	-88.9864	2
<i>Q. oleoides</i>	Costa Rica	Bagaces	10.612	-85.4	5
<i>Q. oleoides</i>	Costa Rica	Ranch Mari	11.014	-85.637	4
<i>Q. oleoides</i>	Honduras	Danli	14.03333282	-86.58333588	2
<i>Q. oleoides</i>	Honduras	Olanchito	15.48099995	-86.56900024	2
<i>Q. oleoides</i>	Honduras	Zamorano	14.54831052	-87.23838906	3
<i>Q. oleoides</i>	Mexico	Nautla	20.18777767	-96.8525	4
<i>Q. oleoides</i>	Mexico	Rascon	21.96	-99.25	2
<i>Q. oleoides</i>	Mexico	Sayula	17.93048416	-94.14837852	3
<i>Q. oleoides</i>	Mexico	San Luis Potosi - Tampico	22.08514	-98.58318	2
<i>Q. oleoides</i>	Mexico	Tamaulipas - Tampico	23.30164	-97.51994	12
<i>Q. oleoides</i>	Mexico	Verracruz - Tampico	21.6006667	-97.85148133	6
<i>Q. sagraeana</i>	Cuba	Cabezas	22.281153	-83.523505	3
<i>Q. sagraeana</i>	Cuba	La Palma	23.024656	-82.214388	5
<i>Q. sagraeana</i>	Cuba	Minas de Matahambre	22.353153	-83.565257	3
<i>Q. sagraeana</i>	Cuba	Vinales	22.365991	-83.430021	3
<i>Q. virginiana</i>	USA	Barnacle Historic State Park	26.251	-81.804	6
<i>Q. virginiana</i>	USA	Gainesville	29.6456	-82.3362	8
<i>Q. virginiana</i>	USA	Houston	31.057	-96.9504	5

<i>Q. virginiana</i>	USA	La Combe	30.4477	-88.8827	8
<i>Q. virginiana</i>	USA	SE	28.3761	-82.1996	3
<i>Q. virginiana</i>	USA	Wilmington	34.035	-77.894	7

Draft

Table 2. Primers used to amplify and sequence the candidate and housekeeping genes used in this study. The lengths of both the full sequences and coding region (CDS) is provided in base pairs.

Gene	Forward primer	Reverse primer	Length	
			Full	CDS
ICE1	AAGATCCAGGAGGG GTTGAC	GGGAAAGAAGAAGG GTCTGC	458	453
GAPDB	AGGAACTTCCTCCGA TGCTG	ATGCCAAGTTCAGCC CAAG	578	423
CHR11	ACAGCTCACTCAGC CTTTGAG	TCTGAATACTTTAAG CAAACAATGC	685	186
HOS1	TGGATGAAAGTGCTG TGTC	GGTGTAAGTCATGAA GCTGTGC	749	534

Draft

Table 3. Alleleic diversity (A) and expected heterozygosity (H(E)) for different genes in each of the seven species of *Virentes*.

Species	HOS1		ICE		CHR		GAP	
	A	H(E)	A	H(E)	A	H(E)	A	H(E)
<i>Q. brandegeei</i>	7	0.719	1	0	1	0	3	0.656
<i>Q. fusiformis</i>	6	0.833	4	0.656	2	0.1528	3	0.5
<i>Q. geminata</i>	7	0.844	1	0	1	0	1	0
<i>Q. minima</i>	13	0.891	1	0	1	0	2	0.375
<i>Q. oleoides</i>	29	0.922	6	0.69	6	0.5833	10	0.692
<i>Q. sagraeana</i>	12	0.878	1	0	2	0.4734	2	0.5
<i>Q. virginiana</i>	24	0.873	7	0.436	2	0.0364	10	0.855

Draft

Table 4. GLM examining the association between HOS1 SNPs and latitude (independent variables) and stem freezing tolerance in *Q. oleoides*. Model adjusted R-squared was 0.957.

SNP	Estimate	Std. error	t-value	p-value
Latitude	3.8351	0.2115	18.132	0
282A	-2.3761	0.8177	-2.906	0.0068
420T	2.3761	0.8807	2.698	0.0113
493G	0.6801	0.3551	1.915	0.065
651C	-1.3544	0.3322	-4.077	0.0003
655N	-0.8052	0.3967	-2.03	0.0513

Draft

Table 5. GLM examining the association between HOS1 SNPs and latitude (independent variables) and stem freezing tolerance in *Q. virginiana*. Model adjusted R-squared was 0.991.

SNP	Estimate	Std. Error	t-value	p-value
Latitude	2.4871	0.0363	68.576	0
681G	-0.2076	0.1471	-1.412	0.1653

Draft

Table 6. GLM examining the association between HOS1 SNPs and latitude (independent variables) and minimum temperature at the coldest month (bioclim 6) in *Q. oleoides*. Model adjusted R-squared was 0.901.

SNP	Estimate	Std. error	t-value	p-value
Latitude	8.407	0.7698	10.922	0
211A	3.277	1.3774	2.379	0.021
420T	-3.635	1.3298	-2.734	0.0085
493G	2.675	1.3986	1.912	0.0612
630G	1.893	0.8248	2.295	0.0257

Draft

Table 7. GLM examining the association between HOS1 SNPs and latitude (independent variables) and minimum temperature at the coldest month (bioclim 6) in *Q. virginiana*. Model adjusted R-squared was 0.747.

SNP	Estimate	Std. error	t-value	p-value
Latitude	2.2615	0.5264	4.296	0.0001
346A	-0.749	0.2802	-2.673	0.0108
493G	0.4144	0.2457	1.686	0.0995
681T	-1.0446	0.5086	-2.054	0.0466

Draft