Identification of genes required for cold acclimation using next generation sequencing of sensitive to freezing mutants

KENT, OLIVIA, VICTORIA, CONSTANCE

How to cite:
KENT, OLIVIA, VICTORIA, CONSTANCE (2014) Identification of genes required for cold acclimation using next generation sequencing of sensitive to freezing mutants, Durham theses, Durham University. Available at Durham E-Theses Online: http://theses.dur.ac.uk/9455/

Use policy
The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.
Please consult the full Durham E-Theses policy for further details.
Identification of genes required for cold acclimation using next generation sequencing of sensitive to freezing mutants

Olivia Victoria Constance Kent

School of Biological and Biomedical Sciences
Durham University

Submitted for the Degree of Master of Science by Research
2013
ABSTRACT

Wild type Arabidopsis plants are able to survive temperatures below 0°C, providing they have been subjected to low positive temperatures for several days prior to freezing. This brief period of low temperature is known as cold acclimation and results in the activation of a number of pathways that bring about physical and biochemical changes necessary to withstand freezing stress. However, when the sensitive to freezing (sfr) mutants are subjected to temperatures that would promote acclimation in wild type plants, they fail to become more freezing-tolerant. This suggests that the mutation lies in a gene intrinsic to the acclimation process.

The sensitive to freezing mutants were discovered as part of an ethyl methanesulfonate (EMS) chemical mutagenesis screen, which induces single nucleotide polymorphisms (SNPs) into the genome. As a result, for each mutant the deficiency in freezing tolerance is caused by a SNP in an unknown gene. sfr4, sfr5, sfr8 and sfr9 were the focus of this investigation. Mapping intervals had been previously determined for each of the four sfr mutants using classical recombination-based techniques, giving a specific region of the genome in which to search for the SNP responsible for freezing sensitivity.

Full Illumina genome sequences were produced for sfr4, sfr5, sfr8 and sfr9, and these were mapped against the TAIR 10 Arabidopsis genome. Using the Integrative Genomics Viewer software, the mapping interval was scanned and all SNPs of interest were recorded. Additionally, a command line-based method of identifying SNPs was also applied to the sequencing data by collaborators at the University of Liverpool. SNPs found by both methods were verified to exist in the DNA of the mutant they corresponded to via direct DNA sequencing. This provided a number of candidate genes for each sfr mutant line. T-DNA insertional mutants were obtained for each of the genes in question, and these were genotyped via PCR. Homozygotes were tested for freezing tolerance, comparing them to the respective sfr mutant to the knockouts/knockdowns in each potential gene candidate.

Additional phenotypic data was collected for each of the sfr mutants, as they had not received much work prior to the investigation. At the end of this investigation, several candidates remained as potential causes of freezing tolerance for sfr5, sfr8 and sfr9; however no candidates were isolated for sfr4.
CONTENTS

Acknowledgements 1
Declaration of copyright 2
List of figures 3
List of tables 4

1. INTRODUCTION 5
   1.1 Chilling and freezing tolerance 5
   1.2 Freezing damage 6
   1.3 How membranes are protected 9
   1.4 Cold acclimation 13
   1.5 Cold sensing 14
   1.6 C-repeat binding factors (CBFs) 16
   1.7 Regulation of the CBFs 18
   1.8 Abscisic acid (ABA) 19
     1.8.1 ABA-independent and ABA-dependent cold acclimation 20
   1.9 Cross-talk between abiotic stress tolerance pathways 23
   1.10 The sfr mutants 24
     1.10.1 sfr2 25
     1.10.2 sfr3 26
     1.10.3 sfr6 27
   1.11 Experimental aims 29

2. MATERIALS AND METHODS 30
   2.1 Chemicals 30
   2.2 Plant tissue 30
     2.2.1 sensitive to freezing mutant seeds 30
     2.2.2 T-DNA insertional mutants 30
     2.2.3 Seed sterilisation 31
   2.3 Growth media 31
     2.3.1 Murashige and Skoog media 31
     2.3.2 Mannitol media 32
   2.4 Growth conditions 32
     2.4.1 Standard growth conditions 32
   2.5 Freezing treatment conditions 33
     2.5.1 Standard adult plant freezing assay 33
2.5.2 T-DNA insertional mutant adult plant freezing assay 34
2.5.3 Petri dish freezing assay 34
2.5.4 Cold treatment of seedlings for assessment of cold-inducible gene expression 34

2.6 Other stress growth conditions
2.6.1 Osmotic stress tolerance: seedling emergence 35
2.6.2 Osmotic stress tolerance: whole seedlings 35
2.6.3 Measurement of sucrose and light-regulated gene expression 36

2.7 Observational studies
2.7.1 Flowering time assay 36

2.8 Nucleic acid extraction
2.8.1 Extraction of genomic DNA for next generation sequencing 37
2.8.2 Extraction of genomic DNA for PCR 38
2.8.3 Extraction of RNA 39
2.8.4 cDNA synthesis for qRT-PCR 40

2.9 DNA
2.9.1 Primer design and synthesis 41
2.9.2 Polymerase chain reaction
2.9.2.1 Genotyping PCR 42
2.9.3 Gel electrophoresis 43
2.9.4 Gel extraction 43
2.9.5 Direct DNA sequencing 44

2.10 Relative quantification of transcripts by quantitative real time PCR 44
2.10.1 Testing cold-inducibility of candidate genes 47
2.10.2 Gene transcripts of T-DNA mutant lines 47

2.11 Next generation sequencing and bioinformatics
2.11.1 Illumina next generation sequencing 48
2.11.2 Analysis of next generation sequencing data via the Galaxy method 48
2.11.3 Integrative Genomics Viewer (IGV) 48
2.11.4 Selection criteria for mutations
2.11.4.1 Where mutation is found 50
2.11.4.2 Number of reads in which the mutation is present 51
2.11.4.3 Read direction 51
2.11.4.4 The type of mutation 52
2.11.4.5 The result of the mutation 52
2.11.4.6 The presence of the mutation in the DNA of the mutant plant 52
2.11.5 Analysis of next generation sequencing data via the command line method 53
2.11.6 The Arabidopsis Information Resource (TAIR) 53
2.11.7 NCBI Basic Local Alignment Search Tool (BLAST) 54

3. RESULTS 55

3.1 Examination of the CBF-COR gene pathway in sfr mutants 55
3.2 Downstream targets of the CBF genes 57
3.3 Next generation sequencing of sfr mutant genomic DNA 60
    3.3.1 Analysis methods 61
    3.3.2 Identification of SNPs 62
    3.3.3 Data quality and how this relates to mapping 63
3.4 Sequencing data 64
    3.4.1 sfr4 64
    3.4.2 sfr5 65
    3.4.3 sfr8 68
    3.4.4 sfr9 70
3.5 Investigating candidate SNPs 71
    3.5.1 Transcript levels of candidate genes under cold and ambient conditions 72
    3.5.2 Identification and validation of insertional mutants in candidate genes 72
    3.5.3 Adult plant freezing assay 75
3.6 sfr4 candidates 77
3.7 sfr5 candidates 77
    3.7.1 sfr5 candidate 1: At1g15690 (AVP1) 78
    3.7.2 Insert lines for AVP1 79
    3.7.3 Freezing assay for AVP1 insert lines 80
    3.7.4 sfr5 candidate 2: At1g15410 82
    3.7.5 Insert lines for At1g15410 82
3.8 sfr8 candidates 83
    3.8.1 sfr8 candidate 1: At3g50910 83
    3.8.2 Insert lines for At3g50910 84
    3.8.3 Freezing assay for At3g50910 insert lines 86
    3.8.4 sfr8 candidate 2: At3g51160 (MUR1) 88
    3.8.5 Insert lines for MUR1 88
3.8.6 Available *mur1* mutants
3.8.7 Freezing assay for *mur1*
3.8.8 *sfr8* candidate 3: At3g56040 (*UGP3*)
3.8.9 Insert lines for *UGP3*
3.8.10 Freezing assay for *UGP3* insert lines
3.9 *sfr9* candidates
3.9.1 *sfr9* candidate 1: At5g62680 (*GTR2*)
3.9.2 Insert lines for *GTR2*
3.10 Summary of results
3.11 Further investigations
   3.11.1 Effect of osmotic stress on seedling emergence
   3.11.2 Effect of osmotic stress on seedlings
   3.11.3 Development of a seedling freezing assay
   3.11.4 Dark/light regulated gene expression
   3.11.5 Flowering time in *sfr8*
   3.11.6 Phenotypic results summary
4. DISCUSSION
   4.1 Analysis approach
      4.1.1 Sequencing and bioinformatics
      4.1.2 Methods used for bioinformatics analysis
      4.1.3 T-DNA Insert lines
      4.1.4 Plant freezing assays
   4.2 Results from each of the mutants
      4.2.1 *sfr9*
      4.2.2 *sfr4*
      4.2.3 *sfr5*
      4.2.4 *sfr8*
   4.3 Final summary
5. APPENDICES
   5.1 Primer sequences
   5.2 Primer verification
6. REFERENCES
ACKNOWLEDGEMENTS

First of all I would like to thank my supervisor Dr. Heather Knight for giving me the opportunity to pursue this project, and for allowing me such free reign in its direction. Her help, guidance and unwavering support have been invaluable. I would also like to say thank you to my second supervisor Prof. Marc Knight; his support and advice throughout this project, particularly with the bioinformatics, were a huge help. I am eternally grateful for the encouragement they have both shown me.

Additional thanks need to go to the Bioinformatics department at the University of Liverpool, specifically Prof. Neil Hall, Dr. Anthony Hall and Laura Gardiner; without their input and data analysis (which went far beyond my capabilities!) this project would not have been possible. Mapping interval data provided by Dr. Glenn Thorlby was also very much appreciated.

I would also like to thank Dr. Piers Hemsley for his extraordinary patience with my atrocious grasp of molar calculations, Dr. Mags Pullen for her PCR optimisation tips and for keeping me company during evening lab sessions, and Beccy Lamb for showing me the ropes and never once being cross when I forgot to wash the Aracon tubes. Thanks also go out to the other MSc and PhD students in Lab 19 for always making things interesting, with a special mention for Alice Rowland for helping me stick to all my deadlines.

Thank you to Jack Lee and Charlotte Hall for making my MSc such an enjoyable experience outside of the lab – I’ll always remember the great times we had at 39 Wakenshaw Road. I’d also like to thank my wonderful proofreaders, Martin Dixon, Anna Mudge and Matt Smith – hopefully your excellent attention to detail has eradicated (almost) all of my mistakes.

Finally I would like to thank my Mum and Dad, without whom I would not have been able to complete this project.
DECLARATION OF COPYRIGHT

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.
LIST OF FIGURES

1.1. Theorised pathway of cold acclimation
1.2. Cross-talk between abiotic stress tolerances
2.1. Screenshot of a SNP on the Integrative Genomics Viewer software
2.2. Example electrophoresis gel for genotyping PCR
3.1. Expression levels of CBF1-3 under ambient and cold conditions in sfr4, sfr5, sfr8 and Col-0
3.2. Expression levels of KIN2 under ambient and cold conditions in sfr4, sfr5, sfr8 and Col-0
3.3. Expression levels of GOLS3 under ambient and cold conditions in sfr4, sfr5, sfr8 and Col-0
3.4. Level of genome coverage achieved for sfr4 next generation sequencing data
3.5. Level of genome coverage achieved for sfr9 next generation sequencing data
3.6. Position of the mutation within the sfr9 mapping interval in relation to the nearest gene
3.7. Example freezing assay for sfr4, sfr5 sfr8 and Col-0
3.8. Expression levels of AVP1 in Col-0 under ambient and cold conditions
3.9. Expression levels of AVP1 in the T-DNA insert line GK_596C07
3.10. Freezing assay comparing survival of Col-0, sfr5 and GK_596C07
3.11. Expression levels of At1g15410 under ambient and cold conditions in Col-0
3.12. Expression levels of At3g50910 under ambient and cold conditions in Col-0
3.13. Expression levels of At3g50910 in the T-DNA insert lines SALK_074693C and SALK_124555C
3.14. Expression levels of At3g50910 in the T-DNA insert line SALK_132810C
3.15. Freezing assay comparing survival of Col-0, SALK_074693C and SALK_132810C
3.16. Expression levels of MUR1 under ambient and cold conditions in Col-0
3.17. Freezing assay comparing survival of Col-0, sfr8, mur1-1, mur1-2 and mur1-3
3.18. Expression levels of UGP3 under ambient and cold temperatures in Col-0
3.19. Expression levels of UGP3 in the T-DNA insert line SALK_020654C
3.20. Expression levels of GTR2 under ambient and cold temperatures in Col-0
3.21. Expression levels of GTR2 in the T-DNA insert line SALK_052178C
3.22. Mannitol supplementation of Col-0, sfr4, sfr5 and sfr8 seedlings
3.23. Mannitol supplementation of Col-0, sfr4, sfr5 and sfr8 seedlings
3.24. Mannitol seedling-float for sfr4, sfr5, sfr8 and Col-0
3.25. Petri dish freezing assay involving sfr6 and Col-0
3.26. Expression levels of DIN6 under dark, light, sucrose and mannitol in sfr4, sfr8 and Col-0.
3.27. Flowering time assay involving sfr8 and Col-0
LIST OF TABLES

3.1. Genes identified as potential candidates for the freezing sensitivity seen in *sfr5*
3.2. Genes identified as potential candidates for the freezing sensitivity seen in *sfr8*
3.3. T-DNA insertional mutants used to create new alleles for potential *sfr* candidates
3.4. Flowering time assay for *sfr8* and Col-0
1. INTRODUCTION

As sessile organisms, plants must react to the environmental conditions to which they are subjected, and one condition that can vary greatly is ambient temperature. While drought is commonly perceived as the greatest limiter of crop production, it is important to note that approximately two thirds of the Earth’s land mass experiences temperatures below 0°C at some point during every year, and half is subjected to temperatures that drop lower than –20°C (Larcher 1994). When even chilling temperatures are able to cause damage to crops (Hetherington et al. 1989), freezing temperatures will seriously limit the area in which certain plants can be cultivated.

Tolerance of frost is extremely important to many crop species such as Triticum aestivum (winter wheat), however in recent years climate change has resulted in frosts in areas of the world where this would not normally be expected, causing damage to crops that would not usually experience such low temperatures (Gu et al. 2008, Warmund et al. 2008). Understanding how plants sense and respond to cold temperatures will be crucial for the production of plants engineered to tolerate freezing temperatures.

1.1. Chilling and freezing tolerance

The degree to which different plant species can tolerate low temperatures varies greatly. Some plants are unable to tolerate low positive temperatures, and temperatures between 10°C and 12°C can cause damage to them (Levitt 1980). These plants are known as chilling sensitive and are usually either native tropical plants, or annuals, such as rice and maize, which undergo only one breeding season. Plants that can withstand low positive
temperatures are known as chilling resistant, and can be sub-divided into two further categories; freezing sensitive and freezing tolerant. Freezing sensitive plants die when subjected to temperatures below 0°C regardless of any other conditions. Freezing tolerant species are able to survive sub-zero temperatures; some species have constitutively active defences whereas others have defences that first need to be activated. This activation process is known as cold acclimation. In *Arabidopsis thaliana* (hereafter referred to as Arabidopsis), it has been suggested that over 1000 genes are induced by cold (Kilian *et al.* 2007), and of these more than 170 encode transcription factors, however some estimates are much higher (Hannah *et al.* 2005). Some transcription factors are known to control a wide range of genes; the most prominent example being the C-repeat binding factor (CBF) transcription factors (Stockinger *et al.* 1997, Jaglo-Ottosen *et al.* 1998, Medina *et al.* 1999), which mediate the expression of a large number of genes that constitute the inducible freezing tolerance of plants.

1.2. Freezing damage

There are various ways in which plants are damaged when exposed to freezing temperatures. Woody species are particularly prone to embolisms in the xylem, in which gas bubbles are forced out of solution when freezing occurs (Sperry and Sullivan 1992). These gas bubbles break the continuous water column (an event known as cavitation), which disrupts the flow of water up a plant (Van den Honert 1948). Cavitation can persist for months and can both limit growth and result in shoot dieback (Sperry *et al.* 1993, Tyree and Cochard 1996). It has also been shown that ice formation within plant tissue can cause cellular damage which can lead to cell rupture (Ristic and Ashworth 1993, Uemura *et al.* 1995). However, it has been recognised for over 100 years, that injury to the plant cell
membranes has the most severe effect on plant survival of freezing, and that the plasma membranes are most damaged by cellular dehydration (Levitt 1980).

The intercellular fluid of plant cells has a much lower solute concentration than the intracellular fluid, and as a result, freezes at a higher temperature. Due to this formation of ice outside the cells, the water potential of the extracellular space is reduced, and this results in water being drawn out from the cytosol, down the concentration gradient (Uemura and Steponkus 1989). Due to these differences in water potential across the plasma membrane, damage can occur. It is this dehydration damage which draws parallels between freezing and drought stress, and the positive responses of plants to mediate drought damage use similar signalling components to those of freezing (Nakashima et al. 2000).

Freezing experiments carried out on protoplasts have demonstrated that changes occur to the organisation of the lipid bi-layer upon freezing (Webb and Steponkus 1993), however, it must be noted that whilst it is likely that protoplasts do portray an accurate representation of what is observed in nature, it has not yet been verified that this is what occurs in whole cells. In protoplasts the cell wall is absent, which presents an obvious limitation; it is quite likely that the cell wall will provide the plasma membrane with protection from ice crystal formation, and without the cell wall the damage protoplasts experience may be more severe (Minami et al. 2009).

Due to the loss of water from plant cells when they are frozen, vesicles of the plasma membrane bud off, disrupting the structure. During the thawing process water moves into
the cell, but due to the newly reduced size of the plasma membrane the cell bursts; this is known as ‘Expansion-induced lysis’ (Webb et al. 1994, Uemura and Steponkus 2003). This type of damage is only seen in non-acclimated protoplasts. The electrolytes that leak out from these membrane lesions can be measured quantitatively, and can consequently be used to provide data that demonstrate the level of damage caused to a plant by freezing.

Another type of damage that is seen in response to freezing in non-acclimated protoplasts is the ‘Hexagonal II phase’ membrane lesion. This occurs between approximately -2°C and -4°C in Arabidopsis (Uemura et al. 1995) and is characterised by plasma membrane destabilisation, causing membrane lipids to aggregate together forming tubes within the membrane, known as the ‘Hexagonal II Phase’. This occurs when two membranes are close together, such as when the plasma membrane and chloroplast membrane are in close proximity (Uemura et al. 1995), which is likely to occur if the cell has lost a large proportion of its liquid. Both ‘Expansion-induced lysis’ and ‘Hexagonal II phase’ membrane lesions are potentially eliminated by acclimation (Webb et al. 1994, Uemura et al. 1995).

There is one type of lesion, similar to ‘Hexagonal II Phase’, which is found exclusively in acclimated protoplasts; the ‘Fracture Jump’ lesion. While ‘Hexagonal II Phase’ lesions are never seen in acclimated protoplasts, these lesions are (Uemura et al. 1995). The ‘Fracture Jump’ lesion occurs when protoplasts (taken from acclimated plants) are frozen to -25°C, and appears as a localised deviation from the fracture plane of the plasma membrane (Webb and Steponkus 1993). Despite the potential differences that may occur between the response of living plant cells to those of protoplasts, it is known that the membrane is the
predominant site of injury in freezing, and protecting the membrane from destabilisation is key.

1.3. How membranes are protected

The plasma membrane (Gordon-Kamm and Steponkus 1984, Uemura et al. 1995), thylakoid membranes (Hincha et al. 1984), and the chloroplast envelope (Krause and Santarius 1975) are all severely damaged by freezing. As a result, many of the changes that occur during cold acclimation involve protecting cellular membranes, particularly in counteracting changes in cell and organelle volume caused by the desiccation that occurs upon freezing.

In some aspects the membrane is able to protect itself; alterations in the lipid composition of the membrane has been linked to an increase in freezing tolerance (Uemura and Steponkus 1994, Uemura et al. 1995). The fatty acid chains that are incorporated into the plasma membrane can be of varying lengths and can display varying degrees of saturation. Shorter chain lipids are less viscous and as result reduce rigidification of the membrane on exposure to cold. A higher number of double bonds in a fatty acid chain increase fluidity; these chains are unsaturated with hydrogens and as a result do not tessellate as readily as more saturated chains (which display a lower number of double bonds); as a result a membrane made up predominantly of unsaturated chains requires less heat to remain fluid (Mironov et al. 2012). Some of the enzymes involved in making membranes more tolerant to freezing are those that alter the saturation of these chains, such as FAD2, which is discussed in detail in section 1.5 (Okuley et al. 1994).
It has been shown that after as little as six hours of cold acclimation, increases in freezing tolerance are evident before changes in the plasma membrane structure can be seen (Ristic and Ashworth 1993, Uemura et al. 1995, Wanner and Juntila 1999). Storage lipids, specifically long-chain unsaturated triacylglycerides are greatly accumulated in the plasma membrane during cold acclimation, providing another aspect of protection (Degenkolbe et al. 2012). The changes that occur within the cellular membranes can be so specific that certain lipids can be used as a marker for increased freezing tolerance. If a plant has a higher proportion of these lipids it will have a higher tolerance to freezing (Degenkolbe et al. 2012).

In the chloroplast envelope it has been shown that acclimation results in changes in the activity or expression of enzymes that result in altered membrane composition. These enzymes alter the chemical groups attached to membrane lipids, and have the effect of creating lipids that are more amenable to forming a bilayer. This prevents non-bilayer structures such as ‘Hexagonal II phase’ (section 1.2) from forming, which would result in and a loss of membrane integrity. SFR2 (discussed in detail in section 1.10.1) is one such enzyme (Moellering et al. 2010). However, these defences alone do not do enough to keep membranes functional and extra mechanisms must be in place.

As well as membrane reorganisation, there are a number of mechanisms involved in preventing membrane rupture. The accumulation of various compatible solutes, which are metabolically inert compounds with no charge, such as proline, occurs to combat the solute potential caused by freezing in intracellular spaces (Wanner and Juntila 1999, Kaplan et al. 2004). The eskimo 1 (esk1) mutant, a constitutively active freezing mutant, accumulates proline even at ambient temperatures, 30-fold higher than non-acclimated wild type Arabidopsis (Xin and Browse 1998).
The sensitive to freezing mutant *sfr4* lacks the ability to accumulate sucrose and displays a complete lack of cold acclimation; this lack of sucrose has been found to be causative of the freezing tolerance deficiency (Uemura and Steponkus 2003). There are several theories to explain the mechanism by which sucrose may protect plant cells from freezing damage; one such method could be to prevent the dehydration caused by freezing stress, by keeping water molecules within the cell (Steponkus 1984) thereby potentially promoting membrane stability (Lineberger and Steponkus 1980). It may also act as a cryo-protectant for enzymes (Carpenter *et al.* 1986). It has been shown that sucrose, along with trehalose, a disaccharide, is extremely effective at providing protection for the membranes (Anchordoguy *et al.* 1987). Increased levels of trehalose have been shown to increase drought tolerance in rice, thus highlighting the potential importance of trehalose and other soluble sugars in the protection of cells against the desiccation that occurs during freezing (Garg *et al.* 2002). However, as with the majority of plant defences against freezing, the ability to accumulate soluble sugars alone does not lead to freezing tolerance; the majority *sfr* mutants are able to accumulate sucrose but are completely deficient in freezing tolerance (McKown *et al.* 1996).

With the discovery that several of the original *sfr* mutants, including *sfr4*, showed a deficit in anthocyanin accumulation when subjected to cold acclimation (McKown *et al.* 1996), the role of anthocyanin in freezing tolerance was considered. Cold temperatures have been shown to up-regulate anthocyanin biosynthesis in Arabidopsis (Leyva *et al.* 1995, McKown *et al.* 1996), however no definitive evidence has been produce to suggest that this directly increases freezing tolerance. While anthocyanin itself is not likely to be a key factor in cold acclimation, it has been suggested that the anthocyanin biosynthesis pathway may share
some common regulatory elements with a pathway that is involved in freezing tolerance (McKown et al. 1996).

Group 2 late embyrogenesis (LEA) proteins, also known as dehydrins, have been shown to accumulate as part of cold acclimation, and protect plants from damage when freezing occurs (Close 2006). Dehydrins are hydrophilic and thermostable, and it has been suggested that they protect cells from dehydration (like compatible solutes) by adding stability to the plasma membrane. It has been suggested that they essentially act as a surfactant by preventing coagulation of membrane structures (Ismail et al. 1999). In Arabidopsis, it has been shown that Low-Temperature-Induced 30 (LTI30) and Cold-Regulated 47 (COR47) are dehydrins that accumulate specifically in response to cold temperature (Nylander et al. 2001), suggesting that their roles are not simply specific to the dehydration pathway and that they are likely to play a role in cold tolerance.

Altered gene expression (when measured as changes to levels of mRNA transcript and altered enzyme activity) can be seen within a few hours of cold conditions (Wanner and Junttila 1999), however longer-term changes to safeguard the plant against freezing damage such as decreased water content and structural changes take days or weeks to occur (Uemura et al. 1995).
1.4. Cold acclimation

Cold acclimation occurs when a plant has spent a period of time (days or weeks) at low positive temperatures, accepted to be around 5°C. In nature it corresponds to late autumn, allowing plants to prepare for sub-zero winter temperatures. This warning period allows it to activate mechanisms that facilitate the protection of the plant against subsequent freezing temperatures (Guy et al. 1987). It is known that as little as 12 hours of acclimation time provides Arabidopsis with some level of tolerance (Jaglo-Ottosen et al. 1998), however in most investigations is has been seen that a much longer time is required to provide the plant with fully functioning defences. Seven days has been found to be the minimum time required for complete cold acclimation to occur (Guy et al. 1987). Acclimation does not result in permanent protection from freezing; even when freezing tolerance has previously been activated in a plant it is lost after one to two days of exposure to warmer temperatures. This therefore suggests that, for species possessing inducible tolerance, a constitutively active freezing pathway would be too costly to maintain (Wanner and Junttila 1999).

A large number of the genes involved in the early stages of cold acclimation are transcription factors, and the discovery of the CBFs uncovered one of the major regulators of acclimation. The CBF transcription factors that are expressed in response to cold and activate a large number of genes that bring about increases in freezing tolerance (described in section 1.6). The CBFs bring about huge changes in what is being transcribed, as is shown when they are overexpressed (Gilmour et al. 2004). An example gene that is clearly induced by cold and the CBF transcription factors, Cold-Regulated 15a (COR15a), has been linked to protecting the chloroplast membrane from damage (Nakayama et al. 2007).
However, not all of the changes that occur during acclimation rely on changes in the expression of genes; as previously mentioned there are a number of physical changes that occur in the membranes that alter their structure and composition and negate the water loss that occurs during freezing. Likewise there are metabolic changes, and the accumulation of compatible solutes, soluble sugars and dehydrins to mediate dehydration (Chen and Murata 2002, Klotke et al. 2004).

1.5. Cold sensing

The process of cold acclimation relies on the assumption that plants have an ability to sense the low positive temperatures. There is a vast array of literature on the topic of moving between the initial drop in temperature and how the plant responds to this, and a number of pathways have been shown to be required for plants to respond to cold, however it has not been proven that any or all of these are needed for the acclimation process.

Due to the damage that cold temperatures can cause to membranes, it is not unsurprising that the plasma membrane has been theorised to be the first step in the pathway to sensing a temperature drop (Levitt 1980). In wild type Arabidopsis, at ambient temperatures, the cellular membranes are fluid. However, when temperatures decrease, the phospholipids within the membrane re-arrange their structure, to become less saturated, and as a result the membranes become more rigid (Alonso et al. 1997). The Arabidopsis fatty acid desaturase mutant (fad2) displays membrane rigidification at a higher temperature than the wild type plant, and lacks an enzyme essential for the synthesis of polyunsaturated lipids (Okuley et al. 1994). A combination of chemical treatments promoting membrane rigidification and blocking cytoskeletal rearrangement, allows Cold Acclimation-Specific 30
(CAS30) transcript accumulation. If membrane rigidification is blocked, accumulation of CAS30 is not seen. (Örvar et al. 2000).

A rapid calcium influx is seen early in response to cold in Arabidopsis (Knight et al. 1996), and calcium has been shown to be an important secondary messenger for the activation of some, if not all, cold up-regulated genes, and has also been shown to be required for cold acclimation (Tahtiharju et al. 1997). It has been suggested that cytoskeletal changes may lead to the opening of calcium channels, allowing calcium to move into the cell from extracellular stores. When chemical inhibitors were used to block cytoskeletal rearrangement, this resulted in reduced cold gene expression (Örvar et al. 2000). As of yet, these channels have not yet been identified, but it has been suggested that they are mechano-sensitive (Carpaneto et al. 2007). It has been suggested that when calcium enters cells it binds to calcium sensors, which allow it to initiate downstream effects (Huang et al. 2012). For further information reviews are available on this subject (Harper et al. 2004).

One proposed link between the increase in calcium and gene expression is a series of mitogen-activated protein kinases (MAPKs). It is not known if or how the calcium sensors interact with the MAPK cascade, however the Calcium/Calmodulin-Regulated Receptor-Like Kinase CRLK1 (which is bound by calcium/calmodulin to activate its kinase activity) has been shown to interact with MAPK/ERK Kinase Kinase 1 (MEKK1); moreover, the calcium cascade cannot proceed without it (Yang et al. 2010). There is also evidence of calcium-dependent protein kinases regulating freezing tolerance in Arabidopsis (Harper et al. 2004), which suggests that they may provide the link between calcium elevation and the kinase cascade.
It has been shown in Arabidopsis protoplasts that mitogen-activated protein kinase kinases (MAPKKs, specifically MEKK1) are induced in response to cold (Teige et al. 2004). MEKK1 specifically induces Mitogen-Activated Protein Kinase Kinase 2 (MAPKK2), and under cold conditions its major target is mitogen-activated protein kinase 4 (MAPK4). Phosphorylation of MAPK4 has been proven in leaf tissue. The mkk2 knockout mutant is unable to cold acclimate, proving that this is a vital stage in the lead-up to freezing tolerance (Teige et al. 2004).

1.6. C-repeate binding factors (CBFs)

Three CBF transcription factors have been identified in Arabidopsis as being important for freezing tolerance; these genes, CBF1, CBF2 and CBF3, are located in tandem array on chromosome IV (Jaglo-Ottosen et al. 1998). Ectopic expression of the CBFs has been shown to induce the cold acclimation process to occur even at ambient temperatures (Stockinger et al. 1997, Liu et al. 1998). Homologs of the CBFs have been found in Brassica napus (Jiang et al. 2011), barley (Morran et al. 2011), and even chilling-sensitive plants such as rice (Ito et al. 2006). It has been suggested that 478 cold-regulated genes are regulated via by the CBFs, and as a result will bring about a number of transcriptome changes in response to cold (Hannah et al. 2005).

CBF transcription factors act by binding to the CRT (C-Repeat); a sequence with a core repetition of the bases 5’-CCGAC-3’ that is found in the promoter region of COR (cold-regulated) genes (Baker et al. 1994, Stockinger et al. 1997). Experiments in yeast have shown CBF1 is a transcriptional activator of genes containing the CRT; the CBFs themselves
do not contain the CRT, however they are induced by cold. Over-expressing CBF1 has been shown to increase freezing tolerance by 3.3°C (Jaglo-Ottosen et al. 1998).

The CRT motif is also present in the dehydration-responsive element (DRE), which is a nine base repeat of 5'-TACCGACAT-3'. The DRE is found in the promoters of genes that are activated in drought conditions (Yamaguchi-Shinozaki and Shinozaki 1994). The CBF family are also known as the Dehydration-Response-Element-Binding 1 (DREB1) family, and the CBFs/DREB1s are activated in response to cold. Another family of DREB proteins, the DREB2s are activated by high-salinity and drought stresses (Liu et al. 1998, Nakashima et al. 2000) and these bind to the DRE. It has been suggested that there is cross-talk between DREB1s/CBFs and DREB2s, potentially via CBF4 (Haake et al. 2002). CBF4 is an apparent homolog of CBF1, and performs the equivalent role in drought tolerance. This suggests the two pathways may have evolved from the same transcription factors.

It has been shown in Arabidopsis that CBF transcription is not up-regulated in response to drought, and is specific to cold temperatures (Jaglo-Ottosen et al. 1998). Induction of the CBF genes has been seen to occur less than fifteen minutes after exposure to 4°C in Arabidopsis (Jaglo-Ottosen et al. 1998), with the induction of target genes of CBF transcription factors occurring two to three hours after transfer to cold conditions. Experiments have shown that the expression of single COR genes, such as COR15a, do not bring about enhanced freezing tolerance. This suggests that these CBFs control a complex regulatory pathway to bring about freezing tolerance. Regulation of the CBFs is obviously controlled as overexpression CBF1 has an adverse effect on the health of the plant; it is dwarfed, has low seed set, and is late flowering (Jaglo-Ottosen et al. 1998, Kasuga et al.
CBF1 and CBF3 function additively, whereas CBF2 acts differently (Novillo et al. 2007). The cbf2 null mutant results in increased expression of CBF1 and CBF3, and of the genes downstream from them, suggesting the CBF2 itself plays a part in the regulation of the other CBFs. The cbf2 null mutant shows increased freezing tolerance (Catalá et al. 2011). Recent work has shown that CBF2 is subjected to both positive and negative regulation. Five Altered CBF2 Expression (acex) mutants which were neither allelic to one another, nor constitutively active, were all seen to have altered expression of the CBFs, with one mutant displaying only reduced CBF2 expression and the remaining four displaying reduced expression of all CBFs (Novillo et al. 2012). The mutants which display reduced expression of CBF1 and CBF3 when only CBF2 has been affected supports the idea that they are themselves regulated by CBF2.

1.7. Regulation of the CBFs

Due to the fast induction of CBF gene expression in response to cold temperatures, it seemed likely that some factor must be present in cells at ambient temperatures, ready to activate CBF transcription. In 1998, Gilmour et al. proposed the existence of gene, which they named Inducer of CBF Expression (ICE) that was responsible for the activation of the CBF genes. Two ICE proteins have been identified, named ICE1 and ICE2 (Chinnusamy et al. 2003, Fursova et al. 2009). ICE1 was found via the screening of an EMS-mutagenised transgenic population of Arabidopsis plants, expressing a fusion of the firefly luciferase reporter gene to the promoter of CBF3, resulting in the photoluminescence of plants in response to cold temperatures; ice1 mutants were easily distinguishable by their lack of luminescence. ICE1 is a transcription factor that is able to recognise an element in the
promoter region of the CBF3 termed the ICE1 Box (Fursova et al. 2009), and ICE2 binds to CBF1 (Chinnusamy, Ohta et al. 2003; Fursova, Pogorelko et al. 2009). The ice1 mutant has been shown to block induction of CBF3, however it has little effect on CFB1, whereas the ice2 mutant affects CBF1 expression (Chinnusamy et al. 2003, Fursova et al. 2009). A similar protein has been suggested named ICE1-like, which may potentially activate the ICE1 Box of CBF1 and CBF2 (Chinnusamy et al. 2007).

ICE itself is regulated by the gene high expression of osmotically responsive genes 1, (HOS1) (Ishitani et al. 1998). The hos1 mutant was discovered via a mutant screen, and displayed increased induction of both CBFs and of downstream cold-induced genes under cold conditions. HOS1 shows homology to a RING-finger protein; many of these proteins act as E3 ubiquitin ligases, and it was suggested that this is the role of HOS1 (Dong et al. 2006). The HOS1 protein was found in the cytoplasm at normal growth temperatures (Ishitani et al. 1998), and was proposed to negatively regulate ICE1 by targeting it for degradation by the 26S proteasome (Lee et al. 2001). SIZ1 is the positive-regulator counterpart of HOS1, which sumoylates ICE1, stabilising it and allowing it to bind to the ICE1 box of CBF3, and activate gene expression (Miura et al. 2007).

**1.8 Abscisic acid (ABA)**

ABA is a phytohormone that is involved in a number of responses to abiotic stress; it has been suggested that it may mediate all environmental responses (Chen et al. 1983). An increase in ABA results in a number of physical changes to plant tissue, including the inhibition of growth and stomatal closure to reduce transpiration (Seo and Koshiba 2002).
Some groups have suggested that treatment of plants with ABA can lead to enhanced freezing tolerance (Lang et al. 1994), and a large number of genes have been found to be induced by both cold and ABA (Seki et al. 2002). However, it has also been noted that this rise in ABA on exposure to cold is transient, and within 48 hours it has returned to almost normal, whereas at this point the levels of freezing tolerance are still increasing (Lang et al. 1994). Both drought and freezing stress have been shown to lead to an increase in the cellular concentration of ABA (Mantyla et al. 1995).

The ABA deficient mutant, *aba-1*, is unable to produce ABA, and ABA-insensitive mutant (*abi1*), which can manufacture but not respond to ABA, both display reduced freezing tolerance (Gilmour and Thomashow 1991), which suggests that ABA does have a role in the acclimation process. However, Thomashow (2010) subsequently argued that this reduced freezing tolerance may be due to the general poor health of the *aba-1* and *abi1* mutants; their reduced cold acclimation is more likely to be a result of the numerous other effects of being either ABA-deficient or unable to respond to it.

### 1.8.1 ABA-independent and ABA-dependent cold acclimation

It has been previously suggested that cold acclimation occurs via an ABA-dependent or an ABA-independent pathway (Yamaguchi-Shinozaki and Shinozaki 1994). ABA-dependent cold acclimation is said to require the activation of the ABA response element (ABRE), a *cis*-element found in the promoter of certain genes (Guiltinan et al. 1990). AREBS (ABA-response element binding factors) bind to the ABRE, activating the genes which have this element (Uno et al. 2000).
The ABA-independent pathway for cold acclimation was previously defined as being mediated by the CBFs. Certain COR gene promoters have been found to contain ABREs, including COR15a (Baker et al. 1994), Low-Temperature Induced 78 (LTI78) (Nordin et al. 1993) and Protein Kinase 2 (KIN2) (Kurkela and Franck 1990). It was thought that these could be activated by ABA without the input of the CBFs. However, it has since been discovered that ABA can activate the CRT and this may occur via CBF1-3 (Knight et al. 2004). Studies have shown that whilst the CRT is activated by ABA, this does not occur at levels that alone induce cold acclimation (Knight et al. 2004). It is still open to debate as to how major a role ABA has in freezing tolerance; however, the role it plays in drought tolerance is well characterised (Seo and Koshiba 2002).

The induction and regulation of freezing tolerance is an extremely complex process. A great number of individual components have been proven to directly impact freezing tolerance in Arabidopsis and other species; however, there is a much lower level of knowledge concerning how each individual component links to another. In figure 1.1. both the known and postulated links between the main regulators of the freezing tolerance pathway have been highlighted.
Cold temperature is sensed by the membrane, which stabilises the cytoskeleton. This feeds back to the membrane and results in the opening of calcium channels. Calcium enters the cell and activates a MAPK cascade. How the cascade affects other aspects of the pathway currently remains unclear, but it may link to SIZ1. Sumoylation of ICE1 by SIZ1 allows ICE1 to activate the CBF genes (via the ICE1-BOX), which then brings about activation of COR genes, and with them, tolerance to freezing. When ICE1 is not required it is ubiquitinated by HOS1 and degraded. ABA has been shown to weakly induce expression of the CBFs, but also regulates freezing tolerance independently of the CBFs. However, it is important to note that alongside gene expression, there are other occurrences that bring about freezing tolerance, such as alterations in morphology, and the accumulation of compounds such as sucrose and compatible solutes.
1.9. Cross-talk between the abiotic stress tolerance pathways

There are various ways in which the abiotic stress tolerance pathways overlap. The signalling components for each pathway are often shared between drought, cold and high salinity, as seen in figure 1.2. Cold and drought stress essentially result in dehydration stress, so it is not unexpected that the two pathways may show cross-talk. Dehydrins, likewise, are present in response to drought (their role in freezing has been discussed previously) to protect cells from desiccation (Beck et al. 2007).

An increase in the level of cytosolic calcium is seen for drought, high salinity and cold (Knight 1999), similarly, the CDPKs previously mentioned have also been seen to be involved in the perception of other stresses (Urao et al. 1994). Reactive oxygen species (ROS) are molecules that are released due to stress and can cause damage to cells, however plants are able to use ROS as a form of signalling, and are capable of continually sensing the presence and levels of ROS (Miller et al. 2008).

Certain genes contain the ABRE and the CRT/DRE, an example of which, is the gene Responsive to Desiccation 29a (RD29a, also known as LTI78), proving that certain stress genes can be activated via different pathways, in the case of this gene, cold drought and high salinity (Yamaguchi-Shinozaki and Shinozaki 1994, Narusaka et al. 2003). The induction of one gene in response to a variety of stresses would indicate that plants mediate the damage caused by those stresses, to some extent, in the same way. In nature one stress is unlikely to occur alone, and as a result plants have to deal with the combination; the potential cross-talk between the different stress tolerance pathways may be a result of this.
1.10. The sensitive to freezing mutants

The original 1-7 sfr mutants were isolated from a population that had been subjected to ethymethylsulfamate (EMS) chemical mutagenesis (James and Dooner 1990). The sfr mutants were distinguished by the way of a cold acclimation screen; when acclimation had taken place, the sfr mutants showed different levels of freezing tolerance when compared to healthy wild type Columbia 0 (Col-0) plants (Warren et al. 1996). Two further mutants (sfr8 and sfr9) were identified at a later date, displaying the same reduced protection from freezing damage following acclimation (Thorlby et al. 1999). Classical mapping determined which chromosome each mutation was located on (McKown et al. 1996), and the mapping
intervals for each mutant were further refined at a later date (Thorlby et al. 1999). For each mutant only one allele is currently available; when sfr5 was initially isolated, there were two mutant alleles; however seeds for sfr5-2 are no longer available (Warren et al. 1996). Other than the noted sensitivity to freezing post-acclimation, there seem to be no pleiotropic effects of the mutations (McKown et al. 1996). Experimental data has shown that sfr4 does not accumulate sucrose, glucose or anthocyanin to the extent that wild type Arabidopsis does in response to cold acclimation (McKown et al. 1996). A causal relationship has been demonstrated between the reduced sugar content of the mutant and its lack of freezing tolerance; supplementing the plant with sucrose results in the restoration of freezing tolerance (Uemura and Steponkus 2003). For other mutants, such as sfr5, there is no discernible reason as to why it is sensitive to freezing, suggesting that it may be governed by a novel pathway.

1.10.1. sfr2

sfr2 showed no pleiotropic effects and did not lack tolerance for any other abiotic stress (McKown et al. 1996). The sensitive to freezing phenotype that is seen is particularly strong for the whole plant assay, but much weaker for the electrolyte leakage assay (McKown et al. 1996). The mutation was mapped to chromosome III (McKown et al. 1996). An 11kb region of Col-0 DNA where the mutation was mapped to was transformed into sfr2 plants; this was shown to restore freezing tolerance. Two genes were found in this region, and as a result these were sequenced, with a SNP being identified in one of the genes; a G to A transition that is most commonly seen with EMS mutagenesis, resulting in a missense mutation occurring at a very conserved amino acid within the gene At3g06510. The assumed loss-of-function was subsequently proven by the creation of other alleles (Thorlby et al. 2004).
The SFR2 protein was initially thought to be a constitutively expressed β-glucosidase (Thorlby et al. 2004), however this description was later altered to a galactolipid:galactolipid galactosyltransferase (Moellering et al. 2010). β-glucuronidase (GUS) staining localised SFR2 to predominantly green tissues, and further work has proven that it localises to the outer membrane of the chloroplast via the use of green fluorescent protein tagging and light microscopy, particularly in stomatal guard cells (Fourrier et al. 2008). The N-terminus has been proven to be responsible for targeting SFR2 to the outer chloroplast membrane. Immunoblotting was used to confirm that the C-terminus of the protein is located across the membrane. Whilst the gene responsible for the sfr2 mutation has been discovered, the mechanism involved in confirming freezing tolerance is currently unknown. It has been suggested that it may modify the envelope of the chloroplast somehow to increase freezing tolerance, however, this is just speculation (Browse 2010).

1.10.2. sfr3

sfr3 shows the greatest freezing sensitivity in young leaves, specifically those that have not yet fully expanded (McKown et al. 1996). It has been shown that they are not compromised in the expression of certain COR genes (Knight et al. 1999). sfr3, as with several of the other sfr mutants identified in this screen, also shows reduced anthocyanin accumulation during acclimation (Thorlby et al. 1999). Classical mapping placed the mutation to the centromeric region of chromosome I (Thorlby et al. 1999). Due to this there were problems in mapping the mutation to a fine resolution, however this region contained few genes; those that were pseudogenes or transposable element genes were discarded (Amid et al. 2012). T-DNA insertion lines were analysed for the genes which remained. Nine genes were sequenced...
using mutant *sfr3* DNA, and these were compared to the known annotated gene sequence. In one of the genes there was found to be a G to A mutation, which introduces a missense mutation into the coding region and creates a restriction site difference in gene *At1g36160* (Amid *et al.* 2012). This was the known gene *ACC1*, Acetyl CoA-carboxylase. This gene has been shown to be embryo lethal in a strong allele, however a weak allele (*glossyhead*) has been discovered that is not embryo lethal (Lü *et al.* 2011). This allele shows altered biosynthesis in the cuticular wax membrane. *sfr3* has been complemented with *ACC1* and this restored the freezing sensitivity entirely (Amid *et al.* 2012). Staining of plant leaf tissue revealed that *sfr3* plant tissue is damaged in response to freezing. It was also noted that wax crystals usually deposited on wild type stems when grown under cold conditions were not seen on *sfr3* inflorescences (Amid *et al.* 2012). This suggests that cuticular wax differences render Arabidopsis leaves unprotected under freezing conditions, resulting in the freezing injuries seen.

1.10.3. *sfr6*

Identified in the same freezing tolerance screen as *sfr2* and *sfr3*, *sfr6* appeared to be somewhat different from the other *sfr* mutants, with distinguishable yellow leaves (McKown *et al.* 1996). Unlike the other members of this screen it was shown that *COR* gene expression was reduced; this was proven both by RNA blot analysis and quantitative real time PCR (qRT-PCR). It was also proven that this reduction in *COR* gene expression (specifically protein kinase 1 [KIN1]) was linked to the phenotype. The reduced gene expression, like the phenotype, was also found to be recessive (Knight *et al.* 1999). The defect in *COR* gene expression was found to occur downstream of CBF transcription factor expression and appeared to be associated specifically with the expression of genes
regulated via the CRT motif (Boyce et al. 2003). Transcripts of genes with the CRT in their promoters failed to accumulate correctly.

The gene responsible for the sfr6 mutation was identified by a combination of techniques (Knight et al. 2009) A mapping population crossed with Landsberg erecta (Ler) and coarse classical mapping mapped the mutation to a 44Mbp interval on chromosome IV, however, it was found to be close to the centromere which can cause problems due to low levels of recombination around centromeric regions. As a result, other methods were required to determine the identity of sfr6. Transfer DNA (T-DNA) insertion lines were obtained for genes between the two markers which sfr6 was found to exist. Due to the unusual appearance of sfr6 seedlings it was possible to visibly identify another allele of sfr6 from the T-DNA-tagged populations (Knight et al. 2009). Two lines were found containing an insert in sfr6 which included an insert in At1g04920 (Knight et al. 2009). The locus of this gene was sequenced from the sfr6 mutant, and a G to A mutation was located in this gene, which introduced a stop codon, truncating the protein significantly (Knight et al. 2009).

SFR6 was identified as the subunit Mediator16 (MED16) of the Mediator complex, a transcriptional co-activator (Bäckström et al. 2007). In eukaryotic organisms, Mediator forms a bridge between a transcription factor and RNA polymerase II, bringing positive or negative regulation of gene expression, depending upon the stimulus (Conaway and Conaway 2011). The ideas of SFR6 being part of a transcriptional regulatory complex explains the extra defects that sfr6 exhibits in addition to its cold-related phenotype; it displays deficiencies in gene expression for UV resistance and pathogen response, and also altered circadian rhythms (Knight et al. 2008, Wathugala et al. 2012).
1.11. Experimental aims

The aim of the work was to build upon the knowledge gleaned from the successful identification of sfr2, sfr3 and sfr6; I wished to find the gene responsible for the reduced freezing tolerance in the mutants sfr4, sfr5, sfr8 and sfr9. The advent of next generation sequencing technologies presented the opportunity to sequence entire genomes with relative ease. Compared to conventional mapping and crossing techniques, this vastly reduces the amount of time required to search for the SNPs which cause mutations in EMS mutants. This investigation specifically used Illumina next generation sequencing, and genome sequence data was compared to the known TAIR 10 reference genome. From this, SNPs were isolated within the known mapping regions for each mutant, and their potential impact on freezing tolerance was examined.
2. MATERIALS AND METHODS

2.1. Chemicals

All chemicals other than those mentioned were sourced from BDH Merck Ltd. (Dorset, UK) or SIGMA-Aldrich (Dorset, UK).

2.2 Plant tissue

2.2.1. sensitive to freezing mutant seeds

All four sfr mutants used in this study were originally isolated by screening an ethyl methanesulfonate (EMS)-mutagenised population of Columbia 0 (Col-0) Arabidopsis (James and Dooner 1990) for plants unable to cold acclimate to freezing temperatures (McKown et al. 1996, Warren et al. 1996, Thorlby et al. 1999).

2.2.2. T-DNA insertional mutants

SALK lines (Salk Institute for Biological Studies, San Diego, USA) (Alonso et al. 2003) and GABI-Kat lines (Genomanalyseimbiologischen System Pflanze - Max-Planck Institut für Molekulare Pflanzenphysiologie, Potsdam, Germany) (Kleinboelting et al. 2012) contain an insert of transfer DNA (T-DNA) within a gene, and this insert is potentially capable of disabling that gene. The T-DNA insert in SALK lines contained the plant-selectable marker NPTII (Neomycin phosphotransferase II), and was introduced via Agrobacterium tumefaciens. The NPTII marker introduced kanamycin resistance, and the T2 progeny of the original T1 transformants were selected for resistance to kanamycin (Azpiroz-Leehan and Feldmann 1997). GABI-Kat lines were created via a very similar method, but contain a SULr (sulfadiazine resistance) marker (Rosso et al. 2003). The Nottingham Arabidopsis Stock...
Centre (Nottingham, UK) dispatched segregating T₃ lines for all SALK and GABI-Kat T-DNA mutant lines required for this investigation. The genotyping of these lines is described in section 2.9.2.1.

2.2.3. Seed sterilisation

The required number of seeds was placed into a 1.5ml microcentrifuge tube and 70% ethanol was added to sterilise. The seeds and ethanol were vortexed in a low-speed bench-top vortex (Labnet, Oakham, UK) for approximately five minutes to aid surface sterilisation. Seeds were then pipetted onto 90mm sterile filter paper circles (Whatman PLC., Maidstone, UK) inside a laminar flow cabinet (AstecMicroflow, Hampshire, UK) and the ethanol was allowed to evaporate off. Once dried, the seeds were dispensed onto agar plates containing growth media.

2.3. Growth media

2.3.1. Murashige and Skoog media

Standard full strength Murashige and Skoog media (Murashige and Skoog 1962) containing vitamins (Duchefa Biochemie, Ipswich, UK) was used for initial plant growth. One litre of milli-Q water was combined with 4.4g of powdered media to give a 1x MS solution. The pH of the solution was brought to 5.8 by addition of 0.1M KOH. Plant tissue culture grade agar (SIGMA-Aldrich, Dorset, UK) was added to a final concentration of 0.8%, and the mixture was then autoclaved (121°C, 1.03 x 10⁵ Pₐ for 20 minutes). When cooled to approximately 50°C the media was poured into Petri dishes in the laminar flow hood. Lids were replaced as
the media was allowed to solidify. Two sizes of dish were used: 90mm (Scientific Laboratory Supplied Ltd, Hessle, UK) and 55mm (Fisher Scientific UK Ltd, Loughborough, UK).

### 2.3.2. Mannitol media

Media containing mannitol was made to the same specifications as in section 2.3.1, however varying amounts of mannitol (BDH Merck Ltd., Dorset, UK) were added to the MS solution before the pH was adjusted and the solution was autoclaved. Mannitol media was made up to 200mM 300mM, 400mM and 500mM mannitol concentrations for seedling emergence (section 2.6.1.).

### 2.4. Growth conditions

#### 2.4.1. Standard growth conditions

Seeds were sown onto Petri dishes containing MS media, and were allowed to stratify in a 5°C cold room for a minimum of four days. After stratification the lids were taped down with Micropore tape (3M United Kingdom PLC., Brackwell, UK) and the seeds were transferred to a Percival growth chamber (Percival Scientific Inc., Perry, USA), set to 20°C (+ or - 1°C) light levels at 150μE m⁻² s⁻¹ for 16 hours, followed by eight hours of darkness. Here seedlings were left to grow for seven days, before being transferred to peat plugs (LBS Horticulture Ltd., Lancashire, UK) of either 44mm or 41mm in size; these were placed in opaque trays to aid watering. Plants required for seed were transferred to a long day growth room (20°C [+ or - 2°C], 16 hours light, 8 hours darkness, 150 – 200μE m⁻² s⁻¹) to promote flowering. When plants began to flower the Arasystem (BETATECH bvba, Ghent, Belgium) was used to prevent cross-contamination of seeds between different lines. The plants were watered until no plant tissue remained green and tissue was allowed to desiccate prior to seed
collection. Once collected seeds were dried overnight at 37°C, and then stored at 5°C to preserve viability and germination rate. If plants were required for freezing assays they were grown under short day conditions, 20°C (+ or - 2°C), 12 hours light, 12 hours darkness and between 150 and 200 µE m⁻² s⁻¹ to encourage rosette growth and suppress flowering.

2.5. Freezing treatment conditions

2.5.1. Standard adult plant freezing assay

Seeds were sown onto Petri dishes containing MS media, and were grown under standard growth ‘short day’ growth conditions for five weeks. At this point they were relocated to a SANYO MLR-351 environmental test chamber (Sanyo, E&E Europe, BV, Biomedical division, UK) set to 5°C, 150 µE m⁻² s⁻¹, 10 hours light, 14 hours darkness, mimicking the conditions plants would experience during cold acclimation. After 10-14 days, half of the available the plants were transferred to a SANYO freezing chamber (MIR-254) (Sanyo, E&E Europe, BV, Biomedical division, UK), set at -7.5°C for 24 hours; the other half remained at 5°C. During this time plants were in darkness. After freezing, plants were transferred to a SANYO environmental test chamber set to 5°C for approximately 30 minutes before being returned to the 20°C ‘short day’ growth room where they were monitored for re-growth. Col-0 plants that were capable of cold acclimation were able to survive this treatment, whilst known sfr mutants showed poor survival rates. The re-growth of these plants was photographically documented.
2.5.2. T-DNA insertional mutant adult plant freezing assay

Post-genotyping (2.9.2.1), populations of homozygous T-DNA lines were tested for freezing sensitivity under the conditions listed above. Each T-DNA line was tested alongside the original EMS sfr mutant to which the T-DNA line corresponded, and with Col-0 plants as a baseline for freezing tolerance. If the T-DNA line responded to freezing conditions in a similar manner to the sfr mutant (i.e. little or no recovery after freezing) further work was conducted on these lines as the gene affected by the T-DNA insertion could be responsible for the freezing sensitivity seen in the sfr mutant.

2.5.3. Petri dish freezing assay

Col-0 and sfr6 seeds were sown onto Petri dishes containing MS media, and were grown under standard growth conditions for seven days. Seedlings were then transferred to a SANYO environmental test chamber set to 5°C, 150 μE m⁻² s⁻¹, 10 hours light, 14 hours darkness, for four days. At this point the seedlings were transferred to -7.5°C for 24 hours. During the freezing stage Petri dish lids were removed to aid airflow and ice nucleation. Post-freezing, seedlings were then returned to the 5°C SANYO chamber for 30 minutes, and then to the 20°C Percival chamber and were observed for re-growth after two days.

2.5.4. Cold treatment of seedlings for assessment of cold-inducible gene expression

Col-0, sfr4, sfr5 and sfr8 seeds were sown onto Petri dishes containing MS media; four petri dishes were sown for each seed line. Seedlings were grown up under standard growth conditions for seven days. At this point, two plates of each line were transferred to a SANYO chamber set at 5°C, and two to a SANYO chamber set at 20°C. The plants were subjected to light levels of 150 μE m⁻² s⁻¹ for the duration of their stay in the chamber. One plate of each
seed line was removed from each of the chambers after two hours, and approximately 30 seedlings were transferred to a 1.5ml microcentrifuge tube and were frozen in liquid nitrogen. The transfer from agar to liquid nitrogen was undertaken in less than one minute to preserve the RNA. The remaining plates were removed from the chambers after six hours, and were extracted from the agar and frozen in the same manner. The frozen plant tissue was then subjected to RNA extraction (section 2.8.3) and cDNA synthesis (section 2.8.4) for use in quantitative real time PCR, with the aim of measuring changes in CBFs, KIN2 and GOLS3 gene expression in the sfr mutants and Col-0.

2.6. Other stress growth conditions

2.6.1. Osmotic stress tolerance: seedling emergence

Col-0, sfr4, sfr5, and sfr8 seeds were sown onto Petri dishes containing MS agar supplemented with differing concentrations of mannitol: 0mM, 200mM, 300mM, 400mM and 500mM. The details of how this media was made are available in section 2.3.2. The seedlings were grown for seven days under standard growth conditions. For each mannitol concentration, there were five repeats for each seed line. At seven days of age, radicle emergence was monitored using a light microscope with a 40x lens, and percentage emergence was recorded for each plate recorded.

2.6.2. Osmotic stress tolerance: whole seedlings

Col-0, sfr4, sfr5 and sfr8 seeds were sown onto Petri dishes containing MS agar and were grown for seven days under standard growth conditions (section 2.4.1). At seven days of age the seedlings were extracted from the MS medium and transferred to a 24-well flat bottom suspension plate (Sarstedt, Newton, USA). Seedlings of each line were floated on 1.5 ml of
0mM, 330mM, 440mM and 550mM mannitol. Seedlings were transferred to a Percival growth chamber set to 20°C (+ or - 1°C) light levels at 150μE m² s⁻¹ and remained there for 72 hours. After this time the seedlings were photographed, and the level of cholorsis for each line was judged.

**2.6.3. Measurement of sucrose and light-regulated gene expression**

Col-0, *sfr4* and *sfr8* seedlings were sown onto Petri dishes containing MS agar and were grown under standard growth conditions for seven days. At this point they were extracted from the agar and were floated on two different solutions in both light and dark. 0.088M sucrose was the chosen test condition, with an iso-osmolar concentration of the non-metabolisable sugar mannitol (0.096M) to control for any osmotic stress caused by the increased level of sucrose. Half of the seedlings were kept in the light throughout the experiment, and the remainder were transferred to the dark for six hours, three hours after subjective dawn. Qualitative real time PCR measured the dark-induction and sucrose-repression of *DIN6*.

**2.7. Observational studies**

**2.7.1. Flowering time assay**

Col-0 and *sfr8* seedlings were sown on MS agar plates and grown under standard long day conditions in the Percival growth chamber. At seven days old the seedlings were transferred to peat plug and grown under long day conditions. Individual plants were monitored every day for signs of flowering; when a flower bolt was 1cm long the number of leaves on that
plant was noted, as was the day (Lee et al. 1993, Knight et al. 2008). The results of each individual were tabulated and statistically analysed using the Student’s t-test.

2.8. Nucleic acid extraction

2.8.1. Extraction of genomic DNA for next generation sequencing

The initial sequencing of the sfr mutant genomes was conducted using genomic DNA extracted via the Qiagen Plant Mini Kit (Qiagen Ltd., Crawley, UK). This was conducted before the commencement of this project by Dr. Heather Knight. For the re-sequencing of sfr4 a modified version of the CTAB (Cetyltrimethylammonium bromide) DNA extraction protocol (Richards et al. 2001) was used.

For the CTAB extraction method approximately 100mg of tissue was frozen in liquid nitrogen and ground in a pestle and mortar filled with liquid nitrogen. 800μl of CTAB extraction buffer (2% CTAB, 1.42M NaCl, 20mM EDTA, 100mM TrisHCl, 2% PVP 40, 5mM Ascorbic acid, 4mM DIECA [diethyldithiocarbamic acid]) was heated to 60°C and added to the plant tissue. The tissue was then transferred to a 1.5ml microcentrifuge tube and 3μl of β-mercaptoethanol was added. The tube was inverted and incubated at 60°C for 30 minutes. 500μl of chlorophorm:iso-amyl-alcohol (24:1) was added and mixed by inversion, before being centrifuged at 15,600g for five minutes. The supernatant was removed and the pellet allowed to air dry in a heated block set at 37°C for approximately 30 minutes. DNA was re-suspended in 100μl of TE buffer, and RNase A was added to a concentration of 10μg/μl. The sample was incubated at 30°C for 30 minutes. 500μl of cold isopropanol, was added, mixed by inversion and precipitated at -20°C for 30 minutes.
The sample was centrifuged at 15,600g for ten minutes and the supernatant was discarded. The DNA pellet was washed in 800μl of 70% ethanol for 20 minutes and spun at 15,600g for five minutes. The supernatant was removed and the sample was air dried at 37°C for 30 minutes. 100μl of TE buffer was used to re-suspend DNA and the sample was left overnight a 4°C. The sample was then sent to The Genome Analysis Centre (Norwich, UK) for further processing, including sonication and preparation for next generation sequencing.

2.8.2. Extraction of genomic DNA for PCR

This method is based on that detailed in Edwards et al. 1991, however, several alterations have been made. Approximately six seven day old seedlings (or one rosette leaf of an adult plant) were transferred to a 1.5ml microcentrifuge tube and were immediately frozen in liquid nitrogen. Leaf tissue was manually macerated using a micropestle. After approximately five seconds 400μl of Edward’s extraction buffer (250mM NaCl, 0.5% SDS, 25mM EDTA pH 8.0, 20mM Tris-HCl pH 7.5) was added. Once complete for all samples, they were spun for approximately one minute at 16,300g in a Genfuge 24D Microcentrifuge (Progen Scientific, London, UK). 300μl of supernatant was removed from the sample and transferred to a clean 1.5ml microcentrifuge tube, to which 300μl of isopropanol was added. The samples were mixed by inversion and a two-minute bench top incubation was carried out to aid the precipitation of DNA. The samples were spun for ten minutes at 16,300g and the supernatant was aspirated. All samples were spun under vacuum for approximately five minutes in an Eppendorf Concentrator 5301 (Eppendorf UK Limited, Stevenage, UK) to dry. 50μl of TE buffer (1mM EDTA, 10Mm, Tris-HCl pH8) was added to each sample, and all were
left overnight at 4°C to allow the DNA to re-dissolve. Samples were then transferred to -20°C for long-term storage before use.

2.8.3. Extraction of RNA

The Qiagen RNeasy Mini RNA extraction kit (Qiagen Ltd., Crawley, UK) was used to extract RNA from plant tissue. RNase-free filter tips and tubes were used throughout the procedure. Tissue samples were transferred directly from -80°C to a Dewar of liquid nitrogen. Each sample was individually removed from the Dewar and manually macerated using a micropestle (Anachem Ltd., Luton, UK). 450μl of RLT lysis buffer (with β-mercaptoethanol added) was added to each sample, and a motorised micropestle was used to homogenise the tissue. The sample was vortexed and heated to 56°C in a heat block (Labnet, Oakham, UK) for between one and three minutes, and were then transferred to ice for a minimum of five minutes. The lysate was transferred to a QIASHREDDER column to aid the shredding the plant tissue and allow separation of cell lysate from cell wall materials, and was then centrifuged for two minutes at 16,300g.

The purified lysate from the collection tube was transferred to a fresh collection tube, and combined with 225μl of 100% ethanol, before being transferred to an RNeasy spin column and associated collection tube. The new column was centrifuged for 30 seconds at 11,400g. Discarding the flow through, 350μl of RW1 wash buffer was added, and the column was centrifuged once again for 30 seconds at 11,400g, discarding the flow-through.

Per sample, 10μl DNase I stock solution and 70μl of Buffer RD were added to the column membrane and were incubated for 20 minutes, to remove genomic DNA which could have
contaminated the sample. 500μl of was buffer RPE was added to wash, and all columns were centrifuged for 30 seconds at 11,400g. This step was then repeated. The flow-through was discarded, and the column was spun at 11,400g for two minutes. RNA was eluted in 30μl of RNase-free water, centrifuged for one minute at 11,400g, and quality and yield of RNA were measured using a NanoDrop 1000 spectrophotometer (Nanodrop Products, Delaware, USA).

2.8.4. cDNA synthesis for qRT-PCR
The Applied Biosystems High Capacity cDNA synthesis kit (Applied Biosystems, Foster City, USA) was used to synthesise cDNA suitable for qRT-PCR. Reagents (2μl 10x RT buffer, 0.8μl 11mM dNTPs, 2μl 10x RT Random Primers, RNA) were thawed on ice. The amount of RNA was standardised between reactions to ensure 2μg was present in a 10μl volume. Each set of reactions was set up with a ‘No Template Control’, which contained water instead of RNA, and a ‘No Reverse Transcriptase’ control to assess genomic DNA contamination, (no cDNA is produced in this reaction therefore only contaminating genomic DNA is available as template for subsequent PCR amplification).

Samples were then transferred to a Thermal Cycler: 25°C for 10 minutes, 37°C for 120 minutes, 85°C for five seconds, with a final hold step at 4°C. Samples are then stored at -20°C for use in qRT-PCR, and before use were diluted 1:50.
2.9. DNA

2.9.1. Primer design and synthesis

Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/) was used to design primers for all PCR reactions including quantitative measurement of transcript levels, excluding those used for SALK and GABI-Kat genotyping. These were designed by T-DNA Express (http://signal.salk.edu/tdnaprimers.2.html). All oligonucleotide primers for PCR and qRT-PCR were ordered from Invitrogen (Life Technologies Ltd, Paisley, UK); all primer sequences are listed in the appendices.

2.9.2. Polymerase chain reaction

For all PCR applications, the DNA polymerase BioTaq RED Polymerase (Bioline, London, UK) was used; this product is a standard Taq polymerase without 3’ to 5’ exonuclease proofreading capability. Genomic DNA extracted using the above Edward’s method (2.8.2) was usually used as the template for PCR, and the standard reaction mixture for each sample included, 1μl of each the forward and reverse primers (50μM), 1μl dNTPs (10mM), 1.5μl of MgCl₂ (50mM), 1μl of BioTaq RED polymerase, 5μl of Bioline 10x buffer (Bioline, London, UK), 1μl of genomic DNA, and the reaction was made up to 50μl with water. For genotyping, reactions were often reduced to 20μl as extraction of the sample from a gel was not required.

The standard PCR cycle consisted of a first stage: 95°C for 5 minutes, 55°C for 5 minutes followed by 72°C for 5 minutes, which was repeated once, and a second stage consisting of: 95°C for 1 minute of denaturation, 55°C for 1 minute of annealing, and 72°C for 2 minutes of extension of the transcript. This second stage was repeated between 30 and 35 times.
depending on the template being amplified. There was a final stage of 72°C for 10 minutes, before the sample was held indefinitely at 4°C until further use.

2.9.2.1. Genotyping PCR

In SALK and GABI-Kat insertional mutant lines, a stretch of T-DNA has been inserted into a gene, and in doing so has disrupted the proper functioning of the gene, potentially knocking it out. T-DNA Express (http://signal.salk.edu/cgi-bin/tdnaexpress/) was used to identify available T-DNA insertion lines for the genes in question. These lines were grown up, and pre-designed primers from T-DNA Express were obtained for genotyping of individual plants. Genotyping PCR was conducted, with the expected results shown in figure 2.1.

![Figure 2.1](image)

*Figure 2.1. A representative image of genotyping PCR of T-DNA insertional lines. Wild type (WT), heterozygous (HET), and homozygous (HOM) banding patterns are indicated.*

PCR was carried out to verify the genotype of each plant, with the possible outcomes being wild type, heterozygous for the T-DNA insertion, or homozygous for the T-DNA insertion. The primers designed to verify that the genomic sequence was uninterrupted at the locus are referred to as the ‘genomic’ primers; when used in a PCR reaction these would produce a product of the expected size for a wild type or heterozygous individual. A ‘left border’ primer which anneals to the T-DNA insert was used with the ‘right’ genomic primer; if this combination gives a PCR band, it proves that an insertion of the T-DNA has taken place. If
the T-DNA reaction band is the only band present, the plant is a homozygote for the insertion, however if both a genomic band and a T-DNA band are present the plant is a heterozygote. Col-0 genomic DNA was used as a positive control for the genomic reaction and a negative control for the T-DNA insert.

2.9.3. Gel electrophoresis

PCR samples were run on ethidium bromide agarose gels, ranging from 0.8 to 1.2%, to verify that products of the correct size had been produced. The required amount of molecular grade agarose (Bioline, London, UK) was added to a flask, and was made up with the correct volume with 0.5xTBE buffer to give a 1% w/v gel (45mM Tris-borate, 1mM EDTA). The flask was heated until the agarose had completely dissolved. Ethidium bromide was added to the final concentration of 5µg/ml, and the gel was poured into a gel casting tray and allowed to set. 5µl of Hyperladder I DNA ladder (Bioline, London, UK) was run in one lane of every gel. The samples were run at 35mA for one hour, and were visualised using a UV gel documentation system (Uvitech Ltd., Cambridge, UK).

2.9.4. Gel extraction

The agarose gel was placed on a UV transluminator and a clean scalpel was used to excise each band, Each sample was transferred to a 1.5ml microcentrifuge tube and was weighed. The QIAquick gel extraction kit (Qiagen Ltd, Crawley, UK) was used to extract bands from ethidium bromide gels to be sent for direct DNA sequencing. Three volumes of gel solubilisation buffer QG (containing guanidine thiocyanate) were added per volume of gel (100mg = 100µl). The samples were incubated at 50°C for ten minutes, and one gel volume of isopropanol was added. Samples were transferred to QIAquick columns and collection
tubes, and centrifuged at 16,300g for one minute. 500μl of Buffer QG was added, and the spin was repeated. 750μl of Buffer PE was added to wash and prevent DNA from disassociating from the spin column, and samples were allowed to stand for between two and five minutes. The samples were centrifuged for one minute, the flow-through was discarded, and samples were centrifuged again to remove residual ethanol. DNA was then eluted in 30μl of buffer EB (10mM Tris.Cl pH8), allowing one minute standing before centrifuging. Concentrations of samples and contamination levels were then checked on a Nanodrop 1000 Spectrophotometer.

2.9.5. Direct DNA sequencing

All direct DNA sequencing was carried out using the Applied Biosystems 3730 sequencer, at the Durham DNA Sequencing and Fragment Analysis Facility (http://www.dur.ac.uk/biosciences/services/dna/).

2.10. Relative quantification of transcripts by quantitative real time PCR

Quantitative real time PCR was used to analyse the expression of genes of interest. The process involved the quantification of PCR product that had been amplified with primers specific to that gene, using cDNA as the template for PCR. This gave a relative quantification of the cDNA corresponding to the gene of interest, and therefore a relative quantification of the amount of transcript (mRNA) present in the original sample. Expression was normalised to expression of \textit{PEX4}, a constitutively expressed gene, to control for any differences in the amount of cDNA between samples (Wathugala et al. 2012). cDNA used in this assay was synthesised using reverse transcriptase (2.8.4).
Primers were designed to anneal to a ~100bp region of interest, and were used in the qRT-PCR assay. Where possible, they were designed to span an intron. An Applied Biosystems 7300 Real-Time PCR System was used for all assays. A 96-well plate (Starlab UK, Milton Keynes, UK) was used for the reaction, with each sample having three technical replicates, to assure that differences seen in gene expression were not as a result of pipetting inaccuracy. The number of biological replicates that were conducted for each experiment is indicated in the results; the optimum number of biological replicates for each experiment would be three, however due to time constraints it was not possible to conduct this number of separate experiments. This does have an effect on the statistical significance of these results, and for those experiments with only one biological replicate they can be considered no more than preliminary work. Those with two biological replicates that display the same expression pattern give results in which it is possible to have much more confidence, however, a third replicate would be the ideal; this is something that is suggested as further work.

Each well contained 7.5µl of SYBR Green 2X qPCRMastermix (PrimerDesign, Southampton, UK) 1µl of each primer (5µM), 1.5µl of water and 5µl of a 1:50 dilution of the initial cDNA reaction. The fluorescent dye SYBR green binds to newly synthesised double-stranded DNA and thus enables quantification of PCR products amplified during the assay. One set of cycling conditions were used for all qRT-PCR cycles conducted during this investigation. An initial 10 minute step at 95°C is followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute, followed by 95°C for 15 seconds, 60°C for one minute, and 95°C for 15 seconds. The final step, a dissociation step, allows the user to determine whether the primers are
amplifying the intended cDNA target or contaminant gDNA and indicate whether primer dimers have formed.

Results were analysed using the Applied Biosystems software associated with the machine (Version 1.4). The method of analysis used on qRT-PCR data was the $2^{-\Delta\Delta C_T}$ (delta delta $C_T$) method (Livak and Schmittgen 2001). The software outputs a $C_T$ (cycle threshold) value, which represents the number of cycles taken for the amount of fluorescent product to exceed an arbitrary threshold. Calculation of delta delta $C_T$ values for each sample gives a relative quantification of gene expression after normalisation to the endogenous control gene $PEX4$, referred to as the ‘RQ’ value. Statistics were performed in accordance with the Applied Biosystems user bulletin ‘Relative Quantification Algorithms in Applied Biosystems Real-Time PCR Systems Software’. In order to employ this method in which standardisation against a constitutively active endogenous control gene is used it must be assumed that the amplification efficiency of the control gene and the target gene are equal. This was verified via primer testing.

Before use, the amplification efficiency of new qRT-PCR primers must be estimated. $C_T$ values from the qRT-PCR reaction were plotted against a serial dilution of cDNA; if primers are functioning efficiently, a doubling of cDNA concentration should result in a shift of one $C_T$ unit. All primers used previous to this investigation have been verified to be successful; those designed in the course of this investigation were also verified, and the results of these verifications are included in the appendices. The results of qRT-PCR were displayed graphically, and error bars on histograms indicate RQ min and RQ max and constitute the acceptable error for a 95% confidence limit according to the Student $t$-test between these samples.
2.10.1. Testing cold-inducibility of candidate genes

qRT-PCR was used to see if the gene in question was inducible in response to cold. A candidate that did not appear to show cold-inducibility was not ruled out from further investigation, however a cold-inducible candidate may have suggested it was likely to be linked to the freezing sensitive phenotype.

Four wild type cDNA samples, (two hours at 20°C, six hours at 20°C, two hours at 5°C, six hours at 5°C) were tested with each set of real-time primers, and the level to which the transcript was detected at ambient and cold temperatures was compared.

2.10.2 Gene transcripts of T-DNA mutant lines

When a homozygous population of individuals was either received directly from the stock centre or isolated after genotyping (section 2.9.2.1) it was verified that the T-DNA insert was causing reduced levels of full-length transcript. This reduced transcript level would occur either due to the insertion occurring in the promoter region, hence reducing levels of normal transcript, or the insertion occurring in the coding region, interrupting and hence truncating the transcript. Insertions occurring in introns could also alter the splicing, affecting transcript levels.

Primers were designed for either side of the proposed site of T-DNA insertion, as to show that in a wild type sample the gene is functioning correctly, however in the insert line, the region of DNA would not be transcribed, as there would be a large insert in the middle.
2.11. Next generation sequencing and bioinformatics

2.11.1. Illumina next generation sequencing

Illumina next generation sequencing (NGS) involves the use of reversibly terminable fluorescent dyes, and produces a large number of short reads (Mardis 2008). It is idea for use in species where a reference genome has already been produced, and in this instance was used in the search for the single nucleotide polymorphisms (SNPs) responsible for each of the sfr mutations. This was carried out by The Genome Analysis Centre (Norwich, UK).

2.11.2. Analysis of next generation sequencing data via the Galaxy method

Galaxy (https://main.g2.bx.psu.edu/) is a free, open-source platform that allows the user to upload raw NGS data with the intention of mapping these data against a reference genome; The interface allows a variety of programs to conduct the mapping, and in this instance the mapping tool used was Bowtie for Illumina (https://main.g2.bx.psu.edu/). These data were mapped, using Bowtie’s default settings, to the latest Arabidopsis reference genome, TAIR 10 (Lamesch et al. 2012). This file-type that is outputted at the end of the process can be read on the Integrative Genomics Viewer (IGV) software (http://www.broadinstitute.org/software/igv/). This is a predominantly automated method of mapping, but does contain a level of customisability for the more advanced user. This method was used to create the .bam output files used in this investigation, and this was carried out by Professor Marc Knight.

2.11.3. Integrative Genomics Viewer (IGV)

The IGV software allows the resultant data from Illumina sequencing to be compiled and arranged into a format that allows the entire genome to be viewed and compared with the
most recent TAIR 10 Arabidopsis genome sequence (figure 2.2). The number of ‘reads’ conducted for each portion of the genome can also be seen, and it is evident when coverage is poor in areas of the genome, as the number available for a certain stretch of DNA is lacking. If a mutation is found in a high enough percentage of reads, the software will draw the user’s attention to it, allowing them to pick out mutations that have a highest confidence of being ‘true’ and found in the DNA. This threshold can be manually set, and for this particular study, was set at 0.7, as the mutations being searched for are homozygous, and hence should appear in 100% of reads, however, this would not account for possible sequencing errors, and setting the threshold at 1.0 could lead to mutations that are legitimate being ignored.

Figure 2.2. Example of two mutations viewed in the Integrative Genomics Viewer software. Bases that agree with the reference genome are coloured grey, however when a difference arises from the reference genome, the base is flagged up with a colour corresponding to which base it is.
As the mapping interval for each of the sfr mutants has been calculated (Thorlby et al. 1999), the software allows the user to manually scan through this specific region of the genome, noting down the position of potential causes of the mutation in each of the mutants. These mutations will be given an order of importance according to the following pre-determined set of criteria. It would be ideal to pursue all of the SNPs found by the software, even if they were deemed unlikely, however due to time constraints the number had to be reduced to the most likely candidates.

2.11.4. Selection criteria for mutations

2.11.4.1. Where the mutation is found

Mapping intervals for each of the sfr mutants had previously been determined, (Thorlby et al. 1999), and were subsequently updated in personal communications from Glen Thorlby. Mutations found outside these updated intervals were immediately discarded. The mapping intervals for each of the sfr mutants had been determined with a reasonable degree of accuracy, and it is highly likely that the mutations causing the sensitive to freezing phenotypes will be found within them.

Viewing the data on a finer scale, a mutation present in the coding region of a gene was treated with a much higher preference than one that is found elsewhere, as it is more likely to have a greater effect on the phenotype. It is possible that a mutation could occur in the promoter region, or a downstream UTR that controls transcript stability, and this could have a substantial effect on the gene, and while there could be the possibility of microRNAs – which are small, non-coding transcripts of RNA that bind to mRNA – having an impact on freezing tolerance, it is more likely that the cause of this mutation is a base change within a
gene. In the case of no mutations being found in genes within the interval, promoter regions would be the next step, as it can be argued that disrupting the promoter sequence of a gene will result in it not being activated.

2.11.4.2. Number of reads in which the mutation is present

If a candidate mutation is present in every read available for that region of genomic DNA, it is more likely that this mutation will be found in the DNA of the mutant plant, rather than being an error that occurred in the sequencing process. If a mutation is found in a region of the genome for which the data have only provided one read that has been repeatedly sequence there is a higher chance of this being a sequencing error, compared to reads that start and finish in different places. As a result, if, for one potential mutation, all of the reads are exactly the same length and have the same starting position they are counted as the ‘same’ read, and this greatly reduces the reliability of this SNP. While it should not be discounted, such mutations were put aside in favour of mutations which fulfilled more of the criteria, as this was only a one-year project.

2.11.4.3. Read direction

The direction of read is important; if a mutation is found in reads facing in only one direction, it is much more likely to be a sequencing error. If a mutation is found in various reads facing in both directions this increases the confidence that it is legitimate, however it also has to be take into account that it could just be co-incidence that reads facing in one direction do not show the mutation, and is not a strong enough criterion to reject a mutation purely on this basis.
2.11.4.4. The type of mutation

The original sfr mutants were generated by EMS, a chemical mutagen that introduces SNPs. It has been statistically proven that the vast majority of EMS mutations are from G or C to A or T, however the sfr mutation could be the result of a reversion event caused by the EMS. There is also the possibility that the mutation could have occurred independently of the EMS mutagenesis. As a result all base changes will be considered.

2.11.4.5. The result of the mutation

The SNP found in the IGV viewer and its position is located within the known coding DNA sequence of the gene. The effect that the mutation has on the protein sequence is shown by altering the amino acid in question and running it through the ExPASy online web application (http://web.expasy.org/translate/). The program gives the predicted translation for the sequence of the mutant DNA. This was compared to the translation that the un-mutated Col-0 translation.

2.11.4.6. The presence of the mutation in the DNA of the mutant plant

By this point the number of potential candidates for the mutation will be greatly reduced, and at this point PCR can be conducted to verify the existence of mutations within the genomic DNA of the mutant plants. Genomic DNA was extracted from the plant, and primers were designed to allow the amplification and sequencing of the portion of DNA which contained the putative mutation. Direct DNA sequencing took place, and using NCBI BLAST the resultant sequence was compared to the known DNA, and if there was a difference between the nucleotides of the wild type sequence and the mutant, this sample
was processed further. If the mutation was not present in the DNA, the potential mutant was discarded, as it was likely to have been an error during the Illumina sequencing.

One sfr8 candidate mutation was ruled out at this stage, as it was not found to be present in the DNA of the mutant plant. This highlights the importance of not relying on the software, as sequencing errors do occur and can appear to be very convincing.

2.11.5. Analysis of next generation sequencing data via the command line method

Some data – specifically sfr5 – did not map well using the Galaxy tool, and as a result was analysed in a different manner. Laura Gardiner of the University of Liverpool used a command line approach rather than an automated system to ‘call’ SNPs. This involved a number of Linux command line-operated bioinformatics programs, BWA (Burrows-Wheeler Aligner) (Li and Durbin 2009), SAMtools (Sequence Alignment/Map) (Li et al. 2009), GATK (Genome Analysis Toolkit) (McKenna et al. 2010, DePristo et al. 2011) and the ‘awk’ command. These allowed fine-tuning of the criteria each SNP had to abide by to be considered legitimate. The method outputs a list of SNPs, which are possible homozygotes and heterozygotes. For the purpose of this experiment, only homozygous SNPs are of interest.

2.11.6. The Arabidopsis Information Resource (TAIR)

Various tools provided by TAIR (http://www.arabidopsis.org/) were used; the gene database was used to verify if any of the candidate genes had been the subject of previous study. The SeqViewer was used to determine what area of the genes certain mutations existed within (promoter, etc). GBrowse was used to determine whether SNPs found via the command line method were within genes, as these data were supplied as genetic co-ordinates.
### 2.11.7. NCBI Basic Local Alignment Search Tool (BLAST)

BLAST, specifically nucleotide blast was used to compare the known sequence of a candidate gene with that sequenced from PCR product of the mutant DNA, to see if the reference allele of the Col-0 gene was different to that of the mutant gene.
3. RESULTS

3.1. Examination of the CBF-COR gene pathway in sfr mutants

The sensitive to freezing mutants studied in this investigation (sfr4, sfr5, sfr8 and sfr9) all resulted from ethyl methanesulfonate (EMS) chemical mutagenesis, which introduces single nucleotide polymorphisms (SNPs) into the genome. For each mutant, one of these SNPs rendered a gene crucial to the freezing tolerance pathway non-functional, as has been shown in previous literature (Warren et al. 1996). Although rare, it is possible for two SNPs to be introduced into the same gene; two base substitutions could potentially have an even greater effect on gene expression.

The most well-studied freezing tolerance pathway is that involving the C-repeat binding factors. These transcription factors regulate the cold regulated (COR) genes via the dehydration-responsive element/C-repeat (DRE/CRT) present in their promoters (Jaglo-Ottosen et al. 1998, Thomashow 2010). When the temperature drops, expression of the CBF genes is activated, and the CBF transcription factors activate COR genes. This pathway has been proven to be essential to the development of plant freezing tolerance (Stockinger et al. 1997, Liu et al. 1998, Thomashow 1999). When mutations occur within this pathway the effects on freezing tolerance are devastating, as is shown in the well-studied mutant sfr6 (Knight et al. 1999).

The four mutants studied in this investigation showed a less severe response to freezing when compared with sfr6 (Warren et al. 1996), however a disruption within the CBF pathway was still a potential cause for the phenotype seen in sfr4, sfr5 and sfr8. Due to the
importance of the CBF pathway to freezing tolerance, measuring the induction of CBF genes and two of their downstream COR gene targets, protein kinase 2 (KIN2) and galactinol synthase 3 (GOLS3) was crucial, either to eliminate the CBF pathway as an area of study, or concentrate on it.

![Graph](image)

**Figure 3.1.** Expression levels of CBFs 1-3 in wild type Col-0, sfr4, sfr5, and sfr8 following two hours of exposure to 20°C (ambient) or 5°C (cold).

Gene expression was investigated by the use of quantitative real time PCR (qRT-PCR), which measured relative levels of expression of gene transcript when plants had been exposed to low temperature. If transcript levels of each gene were similar in the mutant and Col-0, the CBF-controlled pathway was considered to be functioning correctly in the mutants and could be eliminated as a candidate for the freezing tolerance deficit.
Expression of the CBF genes leads to the activation and expression of COR genes, what is thought to be the major pathway that leads to freezing tolerance. As a result, primers were designed for qRT-PCR to generically target CBF expression. CBF1, CBF2 and CBF3 all show high levels of similarity to each other, show very similar expression patterns, and have been shown to be induced rapidly after exposure to low positive temperatures (Medina et al. 1999).

Negligible levels of CBF1-3 expression were observed in ambient temperature-treated Col-0 and sfr samples alike, but after two hours at 5°C there was an increase in the level of detected CBF expression in all samples (figure 3.1). These results were supported by an independent biological repeat (data from biological repeats are not shown). This shows that the expression of CBF1-3 is unaffected in the sfr mutants tested, suggesting that the SNP responsible for the freezing sensitivity of these mutants introduces the freezing tolerance defect downstream of the CBFs or in another pathway that leads to freezing tolerance.

3.2. Downstream targets of CBF genes

KIN2 is a so-called COR gene, the expression of which is activated by binding of the CBF transcription factors to the DRE/CRT element found in the promoter of all COR genes (Jaglo-Ottosen et al. 1998). Due to this, the rise in KIN2 expression occurred much later than in the CBFs, and as a result the time point chosen for this experiment was six hours (Knight et al. 1999).
There was evidently a low-level of KIN2 transcript already present at ambient temperatures (figure 3.2) but, when the temperature was dropped to 5°C, there was an increase in the expression of KIN2 in all lines tested. sfr4 induction of KIN2 was shown to be significantly lower than in Col-0 and the other sfr mutants, however an independent biological repeat showed sfr4 expression levels of KIN2 in line with the other sfr mutants, so this reduction may not be significant.

The expression of another CBF-regulated gene, GOLS3, was measured under cold conditions. The COR gene GOLS3 encodes a protein that protects the plant from oxidative damage when chilling and freezing occur (Nishizawa et al. 2008). Like KIN2, the time point at which measurements were taken was six hours. As seen in figure 3.3., extremely low levels were present at ambient temperatures; lower than those of KIN2. At 5°C, transcript levels
were greatly increased in all of the *sfr* mutants, similar to levels seen in Col-0, suggesting that each of the *sfr* lines tested express *GOLS3* normally. The same result was seen in an independent biological repeat.

![Image of GOLS3 expression levels](image)

*Figure 3.3. Expression levels of GOLS3 in wild type Col-0, sfr4, sfr5, and sfr8 following six hours of exposure to 20°C (ambient) or 5°C (cold).*

The results of these assays indicate that the CBF freezing tolerance pathway functions normally in *sfr4, sfr5* and *sfr8*. This would suggest that for these *sfr* mutants an alternative pathway might be the cause of their deficit in freezing tolerance. It also supports evidence that a fully functioning CBF pathway alone is insufficient for plants to achieve full freezing tolerance.
Gene transcript levels could not be confirmed for sfr9 due to the seeds no longer being viable. The mutant was isolated some years ago and, at the beginning of the study described in this thesis it was discovered that these seeds no longer germinated.

3.3. Next generation sequencing of sfr mutant genomic DNA

As the best-studied freezing tolerance pathway had been eliminated as the cause of the sfr4, sfr5 and sfr8 phenotypes, there remained a large number of other possible factors which could have caused a deficit in freezing tolerance. Mapping intervals had been previously determined for each of the mutations, indicating on which chromosome each of the mutations was found, and a map distance for each of the mutations was determined using known markers (Thorlby et al. 1999). Through personal communications with Glenn Thorlby, these mapping intervals were refined and translated into ‘Atg’ codes, giving a specific region of genes in which the SNP causing the mutation should exist.

Due to the recent increase in accessibility of next generation sequencing technologies, it was possible for full genome sequencing to be carried out for sfr4, sfr5, sfr8 and sfr9. Dr. Heather Knight extracted genomic DNA from sfr4, sfr5 and sfr8 using the Qiagen Plant Mini Kit, and from sfr9 using the CTAB method. All samples were sent to The Genome Analysis Centre (Norwich, UK) for sequencing via the Illumina method; a NGS technique that utilises reversible terminator dyes to create short 100bp ‘reads’ of the genome.

These reads were assembled onto the reference TAIR 10 genome (Lamesch et al. 2012), and SNP differences between the two genomes were noted. From this, SNPs that differed from
Col-0 could be isolated for each mutant and SNPs that occurred within the gene interval for each mutant were recorded.

### 3.3.1. Analysis methods

The initial analysis method that was applied to the data involved uploading raw NGS data files to the Galaxy (Hillman-Jackson et al. 2002) server (methods section 2.11.2), where they were automatically mapped using Bowtie (Langmead et al. 2009) on default settings; henceforth referred to as the Galaxy method. The output file for this method was a .bam file, which was viewable on the IGV software (Thorvaldsdóttir et al. 2012). This method was extremely user-friendly and required little prior knowledge of bioinformatics. The computing power required was low, and mapping the Illumina reads to the reference genome could be completed using a mid-level desktop computer.

A second method known as the command line method was undertaken by collaborators at the University of Liverpool (Anthony Hall and Laura Gardiner) who are specialists in the analysis of bioinformatics data. This method involved the use of command line bioinformatics programs, (BWA Burrows-Wheeler Aligner) (Li and Durbin 2009), SAMtools (Sequence Alignment/Map) (Li et al. 2009), GATK (Genome Analysis Toolkit) (McKenna et al. 2010, DePristo et al. 2011) and the ‘awk’ command. This method required access to powerful supercomputers, and a high level of computer programming knowledge was required to execute the commands needed to analyze bioinformatics data in this way. When both the command line method and the Galaxy method yielded results on the same dataset, direct comparisons between the two methods were drawn in terms of how well each method identified SNPs.
3.3.2. Identification of SNPs

Using the Galaxy method, the raw sequence data were scanned in IGV within the mapping intervals (Thorlby et al. 1999), which were further refined via personal communications with Glenn Thorlby before the start of this investigation. When considering whether to investigate a SNP further as a possible cause of freezing sensitivity, the following criteria were applied; initially the coding region was searched for SNPs, however if there were no SNPs within the coding region for a given mapping interval, the promoter/3’ UTR was considered. Further to this, the SNP had to appear in at least 70% of reads available for that nucleotide.

The command line method did not require manual identification of SNPs as the selection criteria were part of the process, and the resultant output was SNP data for the whole genome sequence. A rating system listed them as likely to be homozygous or heterozygous, with heterozygous SNPs occurring in close to 50% of reads. Only SNPs within the mapping interval were considered as potential candidates. Some of these SNPs were not found in the coding sequence, and these were discarded as they were considered to be the less likely cause of the phenotype. This preference for SNPs in the coding sequence is a criterion that other groups have also used (Uchida et al. 2011).
3.3.3. Data quality and how this relates to mapping

For the four mutant Arabidopsis genomes tested, the mapped genome data varied greatly in quality. The level of mapping to a reference genome was extremely important; if a sample mapped ‘poorly’, it gave a low level of genome coverage; some areas were not replicated many times, and some areas were not sequenced at all.

All SNPs found in silico were also verified in the DNA by PCR. SNPs that are listed as ‘verified’ in the DNA in tables 3.1 and 3.2 were proven to exist in the genomic DNA by single read ABI sequencing of a PCR amplified amplicon spanning the SNP. The results of these sequence data were then compared to the known reference sequence using NCBI blast, and if the mutation was present, the SNP was investigated further.
3.4. Sequencing data

3.4.1. *sfr4*

![Figure 3.4. Screenshot of IGV showing a representative level of coverage seen in the original *sfr4* data. Each row is a level of sequence, and each bar is a paired-end read.](image)

The original Illumina run for *sfr4* yielded extremely low quality data which did not map well with the Galaxy method; the number of reads was low and there were large gaps between reads (figure 3.4). No confidence could be given to SNPs found in these data. The Illumina sequencing was repeated and, as a result, the command line was not using the original *sfr4* dataset. The genomic DNA used for the second sequencing attempt was extracted using a different method to that used previously, a modified version of the CTAB method (section 2.8.1).

The results from the second round of sequencing showed a much higher level of coverage for *sfr4*; software used to map the data via the command line method indicated that *sfr4*...
had an average coverage of approximately x100 across the entire genome. Despite this, none of the command line-generated SNPs corresponded to the mapping region (At1g48640 – At1g50200). Even when presented with a larger interval (At1g42460 – At1g50200) that, based on the mapping results, almost definitely contained the sfr4 mutation, still no SNPs were found. This sequencing data was made available too late to be mapped using the Galaxy method, and as a result there are no SNPs for this mutant.

3.4.2. sfr5

Similarly to sfr4, the sfr5 data were of insufficient quality to work well with the Galaxy method; this sequencing run gave an average coverage of 10-15x across the genome, but when these data were viewed on the Galaxy platform, there were areas within the mapping region that showed lower coverage than this. The command line method mapped these data to the reference genome with more success than the Galaxy method, and as a result did give SNP data for further analysis, however this meant that no comparison could be made between the two methods for this mutant.

The SNPs found via the command line method were compared to the known gene interval (At1g13490 - At1g17290) and several SNPs were found to exist within the interval, as shown in table 3.1. The system used in the command line method ranked the likelihood of SNPs to represent ‘homozygous’ or ‘heterozygous’ mutations in the genome; those ranked as heterozygous were expected to occur in 50% of the reads, and homozygous SNPs as close to 100%. Each SNP was also ranked on a scale from ‘least convincing’ to ‘highly convincing’ in terms of how likely it was to exist within the genome, rather than being a product of mapping or sequencing errors. All SNPs within the mapping interval that were located in the
coding regions of genes (bar the SNP listed as ‘least convincing’) were checked via PCR and sequencing to see if they were present in the DNA of sfr5. All of the SNPs listed as ‘less convincing’ were proven to be a product of sequencing and/or mapping errors, whereas the ‘highly convincing’ heterozygotes in genes were proven to exist.
Table 3.1. SNPs found via the command line method for sfr5. The ‘Likelihood of SNP’ column represents how likely it is that this SNP will appear in the genomic DNA of the mutant. All SNPs listed as ‘less convincing’ were proven to be a product of the sequencing of mapping processes; as a result the SNP listed as ‘Least convincing’ was not included.

<table>
<thead>
<tr>
<th>Position of SNP</th>
<th>Gene</th>
<th>Original base</th>
<th>Mutated base</th>
<th>Likelihood of SNP</th>
<th>Likely to be</th>
<th>Presence of mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr1:4644027</td>
<td>At1g13570</td>
<td>G</td>
<td>T</td>
<td>Less convincing</td>
<td>Heterozygote</td>
<td>Tested; not found in DNA</td>
</tr>
<tr>
<td>Chr1:4834566</td>
<td>At1g14120</td>
<td>T</td>
<td>A</td>
<td>Less convincing</td>
<td>Heterozygote</td>
<td>Tested; not found in DNA</td>
</tr>
<tr>
<td>Chr1:5205941</td>
<td>-</td>
<td>T</td>
<td>G</td>
<td>Highly convincing</td>
<td>Homozygote</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td><strong>Chr1:5300535</strong></td>
<td><strong>At1g15410</strong></td>
<td><strong>A</strong></td>
<td><strong>G</strong></td>
<td><strong>Highly convincing</strong></td>
<td><strong>Homozygote</strong></td>
<td><strong>Tested; found in DNA</strong></td>
</tr>
<tr>
<td><strong>Chr1:5402039</strong></td>
<td><strong>At1g15690</strong></td>
<td><strong>A</strong></td>
<td><strong>G</strong></td>
<td><strong>Highly convincing</strong></td>
<td><strong>Homozygote</strong></td>
<td><strong>Tested; found in DNA</strong></td>
</tr>
<tr>
<td>Chr1:5519332</td>
<td>At1g16110</td>
<td>G</td>
<td>C</td>
<td>Less convincing</td>
<td>Heterozygote</td>
<td>Tested; not found in DNA</td>
</tr>
<tr>
<td>Chr1:5519333</td>
<td>At1g16110</td>
<td>G</td>
<td>A</td>
<td>Less convincing</td>
<td>Heterozygote</td>
<td>Tested; not found in DNA</td>
</tr>
<tr>
<td>Chr1:5807864</td>
<td>At1g16980</td>
<td>C</td>
<td>A</td>
<td>Least convincing</td>
<td>Heterozygote</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr1:5819658</td>
<td>-</td>
<td>A</td>
<td>C</td>
<td>Highly convincing</td>
<td>Homozygote</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr1:5872831</td>
<td>At1g17180</td>
<td>A</td>
<td>C</td>
<td>Less convincing</td>
<td>Heterozygote</td>
<td>Tested; not found in DNA</td>
</tr>
</tbody>
</table>
3.4.3. sfr8

The data for sfr8 mapped well with the Galaxy method. Bowtie aligned reads to a large percentage of the genome, with very few areas showing gaps between reads or low numbers of reads over a certain area. These data were also mapped using the command line method, which predicted an average of 20x coverage across the genome.

The gene interval in which the SNP causing sfr8 was proposed to exist ran from At3g48750 to At3g57270. In table 3.2 all SNPs found for sfr8 using both analysis methods are listed, indicating which method(s) identified the SNP. The direct comparison of these two methods on the same data yielded results as to how well the two methods succeed in identifying SNPs in this type of data, and whether the same SNPs were identified. This comparison also included verifying if the SNP was present in the DNA of the mutant plant, as this gave real proof as to whether a method had successfully detected a ‘real’ SNP. Four SNPs in three different genes were found via both methods, with the majority of the SNPs found via the Galaxy method falling below the threshold coverage levels used by the command line method. One of these SNPs, in At3g56590, was found too late to be worked on further, and as result is a suggestion for further work.
Table 3.2. SNP data for sfr8, identifying which SNPs found via the Galaxy method, were also found by the command line method. No SNPs were found by the command line method alone. In the gene column, * indicates that the mutation was not found in a gene, but the gene recorded is the one it is closest to. **indicates that a mutation was found within multiple copies of the ‘same’ read (see section 2.11.4.).

<table>
<thead>
<tr>
<th>Position of SNP</th>
<th>Gene</th>
<th>Original base</th>
<th>Mutated base</th>
<th>% reads with mutation, (total no.)</th>
<th>Found by command line method?</th>
<th>Presence of mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr3:18110990</td>
<td>-</td>
<td>T</td>
<td>G</td>
<td>100 (4)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:18544327</td>
<td>-</td>
<td>G</td>
<td>T</td>
<td>100 (5)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:18544328</td>
<td>-</td>
<td>T</td>
<td>C</td>
<td>100 (3)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:18684521</td>
<td>-</td>
<td>T</td>
<td>C</td>
<td>100 (2)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:18715665</td>
<td>-</td>
<td>C</td>
<td>T</td>
<td>100 (4) *</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:18738148</td>
<td>-</td>
<td>G</td>
<td>T</td>
<td>100 (9)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:18841182</td>
<td>AT3G50700</td>
<td>C</td>
<td>T</td>
<td>100 (3)**</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:18841183</td>
<td>AT3G50700</td>
<td>A</td>
<td>T</td>
<td>100 (3)*</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3: 18920586</td>
<td>AT3G50910</td>
<td>C</td>
<td>T</td>
<td>100 (16)</td>
<td>Yes</td>
<td>Proven in DNA</td>
</tr>
<tr>
<td>Chr3: 19007726</td>
<td>AT3G51160</td>
<td>G</td>
<td>A</td>
<td>83 (6)</td>
<td>No</td>
<td>Proven in DNA</td>
</tr>
<tr>
<td>Chr3:19206875</td>
<td>AT3G51780</td>
<td>A</td>
<td>T</td>
<td>100 (10)</td>
<td>Yes</td>
<td>Tested; not found in DNA</td>
</tr>
<tr>
<td>Chr3:19206876</td>
<td>AT3G51780</td>
<td>A</td>
<td>T</td>
<td>100 (10)</td>
<td>Yes</td>
<td>Tested; not found in DNA</td>
</tr>
<tr>
<td>Chr3:20648053</td>
<td>AT3G55650</td>
<td>C</td>
<td>T</td>
<td>78 (9)</td>
<td>No</td>
<td>Not tested; sense mutation</td>
</tr>
<tr>
<td>Chr3:20795011</td>
<td>AT3G56040</td>
<td>A</td>
<td>T</td>
<td>100 (9)</td>
<td>Yes</td>
<td>Proven in DNA</td>
</tr>
<tr>
<td>Chr3:20795012</td>
<td>AT3G56040</td>
<td>T</td>
<td>C</td>
<td>100 (8)</td>
<td>Yes</td>
<td>Proven in DNA</td>
</tr>
<tr>
<td>Chr3:19082704</td>
<td>-</td>
<td>C</td>
<td>T</td>
<td>75 (4)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:19437232</td>
<td>-</td>
<td>T</td>
<td>C</td>
<td>100 (21)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:19778523</td>
<td>-</td>
<td>T</td>
<td>C</td>
<td>100 (4)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:20852578</td>
<td>AT3G56210 (3'UTR)</td>
<td>C</td>
<td>G</td>
<td>83 (6)</td>
<td>No</td>
<td>Not tested in DNA</td>
</tr>
<tr>
<td>Chr3:20966136</td>
<td>AT3G56590</td>
<td>C</td>
<td>T</td>
<td>100 (25)</td>
<td>Yes</td>
<td>Not tested in DNA</td>
</tr>
</tbody>
</table>
3.4.4. sfr9

![Screenshot of IGV showing a representative level of coverage seen for sfr9 Illumina sequence data. Each row is a level of sequence, and each bar is a paired-end read. Shown in the image is only a fraction of the genome coverage; this extended almost three times beyond what is seen.](image)

Like sfr8, sfr9 mapped well using both methods (figure 3.5), allowing direct comparisons between the two. Two intervals were supplied; an extremely refined interval which was not certain to contain the gene (At5g62730 – At5g63640) and a larger interval, in which there was much more confidence. The larger interval ranged from gene At5g61850 - At5g64470.

Coverage was excellent for this mutant, with complete coverage of the mapping interval; average coverage of the genome was estimated at 70x. However only one SNP was found
within the region using Galaxy; none were found using command line method. This one candidate was in the potential promoter region of At5g62680, glucosinolate transporter 2 (GTR2), which is involved in the accumulation glucosinolates in seeds (Nour-Eldin et al. 2012) and has been reported to be induced by sucrose (Lejay et al. 2008).

![Figure 3.6. Position of the mutation in the sfr9 mapping interval within the Arabidopsis genome. Indicated in blue is the start codon (reverse complement). In red is the 5' UTR of the gene GTR2. The highlighted base is the base changed by the SNP, from a G to a C.]

The position of the mutation places it some way from the start codon of the gene, however the gene was still investigated as it was potentially close enough to be affecting the promoter region of the gene, and if disruption to the promoter was sufficiently severe, this could be responsible for the phenotype seen in sfr9.

3.5. Investigating candidate SNPs

Once the SNPs had been identified within each mapping interval as described above, experiments were conducted to investigate whether or not any of the genes in which the SNPs occurred were linked to freezing tolerance. This analysis involved testing whether or not the gene was cold-responsive, followed by attempting to identify loss-of-function mutants for that gene and finally, testing those loss-of-function mutants for reduced freezing tolerance, in order
to be able to link the gene to the phenotype. In order to be able to test freezing tolerance, an appropriate assay was required, the development of which is described below. Unfortunately, the second set of sfr4 sequence data was received at the end of this MSc project and therefore sfr4 could not be subjected to this analysis.

3.5.1. Transcript levels of candidate genes under cold and ambient conditions

Candidate genes (genes that each contained a SNP and were present in the given gene interval for each mutant) were tested for up-regulation under cold conditions in Col-0 seedlings. Up-regulation under cold conditions might indicate a role in cold acclimation, as is seen in the CBFs, KIN2 and GOLS3 (Medina et al. 1999, Thomashow 2010). A gene not being up-regulated in cold would not, however, be discarded as a potential candidate for being linked to freezing sensitivity, as there are examples of genes being involved in the freezing tolerance pathway that are constitutively expressed. The most well-known example is ICE1 (Inducer of CBF Expression) which regulates the CBFs (Chinnusamy et al. 2003).

3.5.2. Identification and validation of insertional mutants in candidate genes

For a SNP to be successfully linked to a freezing sensitive phenotype, further mutant alleles that displayed the same phenotype as the sfr mutant needed to be generated for each prospective gene candidate. For each of the candidate genes, T-DNA insert lines were obtained and genotyped as described in the methods, section 2.9.2.1. When homozygous T-DNA insertion lines were isolated, it was vital to check whether the insertion was actually causing any
disruption to the levels of transcript expressed for each candidate gene. qRT-PCR was used to quantify relative transcript levels of the gene.

T-DNA Express (http://signal.salk.edu/cgi-bin/tdnaexpress) was used to determine the insert site, and Primer3Plus (https://www.bioinformatics.nl/primer3plus) was used to design primers to span the predicted insert site (one upstream of the putative insertion site and one downstream) so that, when the primers were used in conjunction with wild type cDNA normal expression of the particular gene would be seen. However, if the insertion disrupted the transcript, extremely reduced or no gene expression would be seen in the putative knockout line. These primers were verified to work efficiently, as described in section 2.10 of materials and methods. If normal wild type transcript levels were measured in the insert line, this line would not be useful in this investigation as expression of the candidate gene would not be impaired.
Table 3.3. T-DNA insertional mutants and their corresponding gene/sfr mutant. The outcome of the mutant (whether it was successful or not) is listed.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Gene</th>
<th>NASC name</th>
<th>Supplied as hom/het?</th>
<th>Confirmed by PCR</th>
<th>How many plants tested (total no.)</th>
<th>Problems encountered</th>
<th>Final status</th>
</tr>
</thead>
<tbody>
<tr>
<td>sfr5</td>
<td>At1g15690</td>
<td>GK596C07.01</td>
<td>Het</td>
<td>Hom</td>
<td>8 (8)</td>
<td>None</td>
<td>Successful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.02</td>
<td>Het</td>
<td>Hom</td>
<td>8 (16)</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.03</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.04</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.05</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.06</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.07</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.08</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.09</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.10</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.11</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GK596C07.12</td>
<td>Het</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>At1g15410</td>
<td>None available</td>
<td></td>
<td></td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>sfr8</td>
<td>At3g50910</td>
<td>SALK_074693C</td>
<td>Hom</td>
<td>Hom</td>
<td>Pool of seedlings</td>
<td>None</td>
<td>Successful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SALK_12455C</td>
<td>Hom</td>
<td>Hom</td>
<td>Pool of seedlings</td>
<td>None</td>
<td>Unsuccessful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SALK_132810C</td>
<td>Hom</td>
<td>Hom</td>
<td>8 (16)</td>
<td>None</td>
<td>Successful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SALK_150964</td>
<td>Het</td>
<td>Problem</td>
<td>8 (8)</td>
<td>No T-DNA bands</td>
<td>Unsuccessful</td>
</tr>
<tr>
<td>sfr9</td>
<td>At3g51160</td>
<td>SALK_027379</td>
<td>Het</td>
<td>Problem</td>
<td>15 (15)</td>
<td>Only one het;</td>
<td>Abandoned - turned to NASC-approved mur1 lines</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SALK_027387</td>
<td>Het</td>
<td>Problem</td>
<td>16 (16)</td>
<td>No T-DNA bands</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SALK_057153</td>
<td>Het</td>
<td>Het</td>
<td>8 (13)</td>
<td>Plants extremely ill</td>
<td></td>
</tr>
<tr>
<td></td>
<td>At3g56040</td>
<td>SALK_020654C</td>
<td>Hom</td>
<td>Hom</td>
<td>Pool of seedlings</td>
<td>Extremely poor</td>
<td>Successful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SALK_02178C</td>
<td>Hom</td>
<td>Hom</td>
<td>Pool of seedlings</td>
<td>None</td>
<td>Unsuccessful</td>
</tr>
</tbody>
</table>

No longer needed due to success of GK596C07.01
3.5.3. Adult plant freezing assay

The sensitive to freezing mutants do not recover from damage sustained during freezing, even when subjected to freezing temperatures post-acclimation. This phenotype is demonstrated by a freezing assay, which is consistent with previous findings (Warren et al. 1996). Adult plants between four and five weeks of age, of Col-0, sfr4, sfr5 and sfr8 mutants were subjected to the standard freezing assay used throughout this investigation. The conditions of two weeks at 5°C, 24 hours at -7.5°C, followed by three days of recuperation at 20°C were sufficient to allow Col-0 plants with functional freezing tolerance to recover, while the sfr mutants did not (Figure 3.7). When frozen, to counteract edge/insulation effects of the freezing chamber, all plants tested were randomised within trays. Flags that did not interfere with the experiment were used to identify which line each individual plant belonged to.

If plants that contained a T-DNA insert were subjected to these exact assay conditions and were unable to recover, this suggested that the affected by the T-DNA insertion in was linked to freezing tolerance and therefore potentially the same gene that had been affected by the original EMS screen in the sfr mutant.
Figure 3.7. A representative sample of a standard freezing assay. All plants were grown under standard short day conditions, and were then transferred to 5°C for two weeks to allow cold acclimation to occur fully. Those marked as 'Acclimated + frozen' were at this point transferred to -7.5°C for 24 hours; those marked 'Acclimated + not frozen' remained at 5°C for this 24-hour period. After the freezing event, all plants were transferred to a 20°C short day growth room to allow them to recover. Photographs were taken three days after transferral to 20°C.
3.6. sfr4 candidates

No sfr4 candidates were identified within the specified gene interval and as a result no further work could be carried out on this mutant. For the given gene interval, all of the genes present were inspected to see what they controlled. None of these genes seemed to be directly related to sugar, which was unexpected, as previous work on sfr4 has suggested that it fails to accumulate sugar during the cold acclimation process (McKown et al. 1996, Uemura et al. 2003).

3.7. sfr5 candidates

The command line analysis method resulted in the identification of two candidate genes for the sfr5 mutation: At1g15690 and At1g15410 (as shown in table 3.1). There were initially other candidates (listed in the table), but these were not found in the DNA of the mutant and as a result were eliminated from the search. At1g15690 encodes AVP1, a H⁺ pyrophosphotase which maintains vacuolar pH and controls auxin transport (Li et al. 2005). It is known to be involved in the response to water deficit stress (Pasapula et al. 2011). Over expression of this H⁺ pump has been shown to increase drought and salt tolerance when over-expressed in Arabidopsis. (Gaxiola et al. 2001, Li et al. 2005, Pasapula et al. 2011). Previous work on this gene has shown that the avp1-1 mutation has an extreme effect on both root and shoot development, rendering the mutant plants much smaller than the Col-0 controls (Li et al. 2005). Flowering is also affected; only 30% of avp1-1 null plants initiated flower development, with none of these developing a full flower, rendering the mutant sterile (Li et al. 2005).
sfr5 has never shown any difficulties when flowering, and produces viable seed when the mutation is homozygous. However, the mutant described in this literature is a complete null, whereas sfr5 is the result of a SNP. A single base change could potentially cause a much smaller change to the resultant protein and, if the mutation occurs some way downstream of the start codon, the protein may be truncated but still functional to a limited degree. The Arabidopsis gene At1g15410 was identified in part of a large study and has not received further work since then. It was predicted to be involved in racemase and epimerase activity, however this has not been biologically verified (Chen et al. 2007).

3.7.1. sfr5 candidate 1: At1g15690 (AVP1)

AVP1 was tested for cold-inducibility via qRT PCR but showed no difference in expression after either two hours or six hours of cold treatment, when compared to ambient temperatures. (Figure 3.8). This experiment underwent an independent biological repeat, which supported these results.
3.7.2. Insert lines for AVP1

No homozygous SALK lines were found for this candidate; however, one heterozygous GABI population (GK_596C07) was identified for AVP1. 12 T3 seed-sets were dispatched; each set was the progeny of a T2 plant harbouring at least one copy of the T-DNA insert within the gene. One of the T3 lines was identified as being homozygous, and this was followed up. Seeds were collected from the homozygous individual, and seedlings were grown. These seedlings were once again tested to verify the homozygous status of the mutant before progression to freezing assays and qRT-PCR. As shown in figure 3.9, GK_596C07 showed severely reduced expression of the full-length transcript, and as a result GK_596C07 plants were deemed suitable for the freezing assay. No independent biological repeat were available for this experiment. While not ideal, this was deemed to be acceptable as the
reduced expression of this gene was not expected to be specific to environmental conditions.

![Graph showing expression levels of AVP1 in Col-0 and GK_596C07.](image)

**Figure 3.9.** Expression levels of AVP1 in Col-0 and GK_596C07.

### 3.7.3. Freezing assay for AVP1 insert lines

As shown in figure 3.10, the GK_0596C07 plants survived to a much higher level than *sfr5*, and are possibly healthier than the Col-0 plants. As the gene was knocked out completely in the insertional mutant, this would suggest that this gene is not responsible for the *sfr5* mutation.
Figure 3.10. Freezing assay conducted on Col-0, sfr5 and GK_596C07 plants. All plants were grown under standard short day conditions, and were then transferred to 5°C for two weeks. Those marked as ‘Acclimated + frozen’ were at this point transferred to -7.5°C for approximately 24 hours; those marked as ‘Acclimated + not frozen’ remained at 5°C for this 24-hour period. After the freezing event, all plants were transferred to a 20°C short day growth room to allow them to recover. Photographs were taken three days after transferral to 20°C.
3.7.4. sfr5 candidate 2: At1g15410

At1g15410 was tested for cold inducibility. As shown in figure 3.11, there was no evidence of cold-induction after two or six hours at 5°C. There was, however, a difference in expression between the two time points, suggesting that this gene may be either circadian or diurnally regulated. An independent biological repeat also displayed this time-regulated pattern of expression.

![Expression levels of At1g15410 in wild type Col-0 after exposure to 20°C (ambient) or 5°C (cold) at two time points, two hours and six hours.](image)

3.7.5. Insert lines for At1g15410

No suitable insertional mutants were available for this candidate gene. Due to time constraints, no further investigations into this candidate could be conducted. Had time permitted, an RNAi knockdown construct would have been created, using a pHELLSGATE vector. Col-0 plants would have been transformed with a plasmid containing the knocked-down gene, and expression levels would be tested.
3.8. \textit{sfr8} candidates

Three candidate genes were initially identified for \textit{sfr8} using the Galaxy method; \textit{At3g50910} is an unknown protein that has received little work, \textit{At3g51160}, Murus 1 (\textit{MUR1}) encodes an enzyme that catalyzes the first step in the de novo synthesis of GDP-L-fucose, mutants of which have been shown to display a distinct phenotype; rounded leaves, weak inflorescences, and reduced tolerance of drought. \textit{At3g56040} encodes UDP-glucose pyrophosphorylase 3 (\textit{UGP3}), which is required for biosynthesis of sulpholipids, a constituent part of photosynthetic membranes in the chloroplast (Okazaki \textit{et al.} 2009).

3.8.1. \textit{sfr8} candidate 1: \textit{At3g50910}

qRT-PCR was used to measure the expression of \textit{At3g50910} and indicated that this gene was slightly up-regulated in cold after both two hours and six hours. However this result is less accurate for two hours, due to the error bars overlapping. The fact that this gene is mildly up-regulated by low temperature may be consistent with a role in low temperature tolerance. Extremely similar results were seen for an independent biological repeat, which could neither confirm nor deny these findings.
The three homozygous SALK lines (SALK_074693C, SALK_124555C, SALK_132810C) for the gene At3g50910 were all successfully proven via PCR to be homozygotes. One heterozygous line was also obtained, but repeated attempts at genotyping failed to demonstrate a T-DNA insertion in the gene. As a result, this line was discarded due to there being three confirmed homozygotes available for this gene. The three remaining SALK lines were tested to verify whether the T-DNA insertion reduced the expression of the gene targeted. For the first two homozygous SALK lines (SALK_074693C and SALK_124555C) the insertion sites were so close together that the same set of primers were used to check expression of the transcript in both lines.

**Figure 3.12.** Expression levels of At3g50910 in wild type Col-0 after exposure to 20°C (ambient) or 5°C (cold) at two time points, two hours and six hours.

### 3.8.2 Insert lines for At3g50910

The three homozygous SALK lines (SALK_074693C, SALK_124555C, SALK_132810C) for the gene At3g50910 were all successfully proven via PCR to be homozygotes. One heterozygous line was also obtained, but repeated attempts at genotyping failed to demonstrate a T-DNA insertion in the gene. As a result, this line was discarded due to there being three confirmed homozygotes available for this gene. The three remaining SALK lines were tested to verify whether the T-DNA insertion reduced the expression of the gene targeted. For the first two homozygous SALK lines (SALK_074693C and SALK_124555C) the insertion sites were so close together that the same set of primers were used to check expression of the transcript in both lines.
Figure 3.13. Expression levels of At3g5910 in Col-0, SALK_074693C and SALK_124555C.

Figure 3.14. Expression levels of At3g5910 in Col-0 and SALK_132810C.
While SALK_074693C shows reduced expression of At3g50910 compared to the Col-0 sample, SALK_124555C actually expressed the gene more effectively than the wild type, figure 3.13, which is not uncommon in this sort of insert line; the 35S promoter that is present in the T-DNA insert can result in over-expression of the Col-0 gene (Ülker et al. 2008). SALK_074693C was tested in the freezing assay as expression is reduced. SALK_124555C, however, was not included in the freezing assay. There is no doubt that SALK_132810C (figure 3.14) displays severely reduced gene expression in comparison to the Col-0 control. This sample was confidently included in freezing assays as a knockout mutant for this gene. However, it must be noted that these experiments were completed only once and no biological repeats are available, resulting in this being a suggestion for further work.

3.8.3. Freezing assay for At3g50910 insert lines

Twenty plants for each of the two confirmed homozygous SALK lines for At3g50910 (SALK_074693C and SALK_132810C) were grown up to five weeks, and subjected to the freezing test protocol. The plants were allowed to recover for four days and were then photographed. The recovery of these plants was compared to that seen for sfr8 (figure 3.15) and the Col-0 control with which they were grown; if any of the SALK lines showed a loss of freezing tolerance, similar to the original EMS mutant, this constituted evidence of linkage. Both SALK lines for this gene candidate, At3g50910, showed recovery levels close to that of the Col-0 (approximately 50%), which would suggest that this gene is not responsible for the freezing sensitivity seen in sfr8.
Figure 3.15. Freezing assay conducted on Col-0, SALK_132810C, and SALK_074693C plants. All plants were grown under standard short day conditions, and were then transferred to 5°C for two weeks. Those marked as 'Acclimated + frozen' were at this point transferred to -7.5°C for approximately 24 hours; those marked as 'Acclimated + not frozen' remained at 5°C for this 24-hour period. After the freezing event, all plants were transferred to a 20°C short day growth room to allow them to recover. Photographs were taken three days after transferral to 20°C.
3.8.4. *sfr8* candidate 2: At3g51160 (*MUR1*)

*MUR1* was tested to verify whether it was cold inducible. As seen in figure 3.16 it is unlikely that *MUR1* is cold inducible, however it is possible that there is some circadian regulation occurring due to both six-hour time points show reduced levels of transcript compared to the two-hour. A similar pattern was seen in an independent biological repeat.

![MUR1 graph](image)

*Figure 3.16. Expression levels of MUR1 in wild type Col-0 after exposure to 20°C (ambient) or 5°C (cold) at two time points, two hours and six hours.*

3.8.5. Insert lines for *MUR1*

Three heterozygous SALK lines were available for the gene At3g51160. Genotyping of these proved difficult; for SALK_027379 16 plants were individually genotyped, and all plants showed an undisrupted Col-0 band and only one showed a T-DNA band (indicating it was a heterozygote). From a second seed batch the results were more promising; two
heterozygotes were identified, and one putative homozygote. This homozygote was grown up but was consistently unhealthy and subsequently died.

Genotyping the heterozygous SALK_027387 line was attempted several times, and both the original seed batch and a replacement seed batch failed to produce a T-DNA band, whilst a band indicating an uninterrupted wild type gene was observed for all samples. As a result this SALK line was abandoned.

The third heterozygous SALK_057153 also proved hard to genotype, with no heterozygous or homozygous individuals being discovered from the first batch of 16 plants. From the second, much like SALK_027279, these heterozygous individuals were extremely weak, and as a result no seed was obtained with enough time to investigate this SALK line further.

Work on this candidate gene proved that heterozygous SALK lines were often problematic, as insertions could not always be confirmed in the material sent out by NASC. The homozygous lines proved to be far more useful to this investigation, and in the case of MUR1, there were other available EMS mutants lines that had been verified by ABRC.

3.8.6. Available mur1 mutants

Three mur1 mutant lines were available from NASC, and these were used in freezing assays. Reduced expression tests were not needed as the lines supplied were created by EMS mutagenesis (and therefore represented single point mutations rather than complete disruption of the transcript) and had been published previously (Bonin et al. 1997). mur1-1
and mur1-2 are the most severely affected, showing only 2% of the Col-0 levels of L-fucose. 
mur1-3 is less severely affected, showing ~33% of Col-0 levels of L-fucose.

The two most severely affected lines, mur1-1 and mur1-2 show the same response as sfr8 to
freezing treatment; 0% survival, while mur1-3 showed equivalent survival to wild type. The
two severely affected mutants which had been suggested to show dwarfism, appear to be a
similar size to sfr8, while the Col-0 plants are considerably larger. As the reduction in L-
fucose is not that great in mur1-3, it is extremely likely that MUR1 is still a candidate for the
freezing tolerance deficit in sfr8.

3.8.7. Freezing assay for mur1 mutants

As shown in figure 3.17, mur1-1, mur1-2 both reacted in the same way as sfr8. mur1-3 did
not, however the L-fucose levels in this mutant are much higher than those in mur1-1 and
mur1-2.
Figure 3.17. A freezing assay conducted on Col-0, sfr8, mur1-1, mur1-2, and mur1-3 plants. All plants were grown under standard short day conditions, and were then transferred to 5°C for two weeks. Those marked as ‘Acclimated + frozen’ were at this point transferred to -7.5°C for approximately 24 hours; those marked as ‘Acclimated + not frozen’ remained at 5°C for this 24-hour period. After the freezing event, all plants were transferred to a 20°C short day growth room to allow them to recover. Photographs were taken three days after transferral to 20°C.
3.8.8. *sfr8* candidate 3: At3g56040 (*UGP3*)

*UGP3* was tested to verify whether it was cold inducible. As shown in figure 3.18, *UGP3* was shown to display no cold inducibility and no real difference between time points. An independent biological repeat supported this result.

![Graph showing expression levels of *UGP3*](image)

*Figure 3.18. Expression levels of *UGP3* in wild type Col-0 after exposure to 20°C (ambient) or 5°C (cold) at two time points, two hours and six hours.*

3.8.9. Insert lines for *UGP3*

One homozygous SALK line, SALK_020654C, was available and from the supplied seed seedlings were grown up. The germination rate of this line was extremely low; from the entire batch very few seedlings survived. Transcript levels of *UGP3* were verified in SALK_020654C and found to very low (Figure 3.19), however this experiment does not have any biological repeats, and would need to be replicated before this work is considered
anything more than preliminary. However, due to these results this mutant was used in subsequent freezing assays.

![Graph showing Relative expression vs. UGP3 for Col-0 and SALK_020654C]

*Figure 3.19. Expression levels of At3g5910 in Col-0 and SALK_020654C.*

### 3. 8.10. Freezing assay for UGP3 insert lines

Due to the low germination rate experienced with this mutant, only one adult plant could be frozen. This plant subsequently died, however there is very little evidence to suggest that this SALK line was or was not tolerant to freezing.
3.9. sfr9 candidates

3.9.1. sfr9 candidate 1: At5g62680 (GTR2)

Only one candidate gene was considered for this mutant, At5g62680, GTR2. This was the gene that corresponded to the only SNP found within the mapping interval for this mutant, and whilst the SNP in the original sfr9 DNA was not in the coding sequence of the gene, it is theorised that it may be within the promoter region. It is some distance from the 5′ UTR, as shown previously (Figure 3.6). As previously mentioned, sfr9 seeds were no longer viable and did not germinate, so no comparison of freezing tolerance could be conducted between insertion lines and the mutant. The gene was, however, tested for cold induction.

![Figure 3.20. Expression levels of GTR2 in wild type Col-0 after exposure to 20°C (ambient) or 5°C (cold) at two time points, two hours and six hours.](image_url)
As shown in figure 3.20, after two hours at 5°C, there seems to be no difference in transcript levels when compared to the Col-0 sample. After six hours, however, there appears to be slightly reduced transcript levels. This shows that the gene is in no way cold inducible. An independent biological repeat confirmed this to be the case.

3.9.2. Insert lines for GTR2

The one homozygous SALK line, SALK_052178C, was grown up and genotyped. The expression of GTR2 was examined in SALK_052178C. As shown in figure 3.21, the expression of the functional Col-0 gene is higher in the insertional mutant than in the Col-0 sample, resulting in this line being unsuitable for use in freezing tests. No biological repeats are available for this experiment; this result would need to be replicated before this candidate is truly disregarded.

![Bar chart showing expression levels of GTR2 in Col-0 and SALK_052178C.](image)

*Figure 3.21. Expression levels of GTR2 in Col-0 and SALK_052178C.*
3.10. Summary of results

Throughout this investigation, the T-DNA insert line available for one candidate for *sfr5*, *AVP1*, responded in the same way as Col-0 plants. However, it has not yet been ruled out due to phenotypic discrepancies seen between the *avp1* mutant and the T-DNA line, potentially indicating there was a problem with this line that has gone undetected. As a result this candidate warrants further work. Further work on this remaining candidate, At1g15410, is also suggested.

One candidate was successfully eliminated for *sfr8*, At3g50910, through the use of SALK mutants in freezing tests. The survival of these SALKs under freezing conditions would suggest that this gene is not responsible for the freezing tolerance deficit. The second candidate for *sfr8*, *MUR1*, still remains as a strong candidate, but cannot be fully accepted and requires more work. For *UGP3*, the only available homozygous SALK failed to reduce expression of the *UGP3* transcript, and thus no further investigations could be carried out on this mutant.

For the one *sfr9* candidate, the one available homozygous SALK did not produce enough viable seedlings to allow proper freezing testing. As a result, this candidate can neither be favoured nor rejected.
3.11. Further investigations

Little is known about the four sensitive to freezing mutants being investigated; of the four, only sfr4 has been the subject of further work, and it has been shown that the deficit in freezing tolerance in sfr4 is due to an inability to accumulate sugars during cold acclimation (McKown et al. 1996, Uemura et al. 2003).

While not mentioned in the literature, during this investigation it was noticed that sfr8 flowers earlier than any of the other mutants or Col-0. As a result of this, a number of phenotypic experiments were conducted to learn more about the other effects of the sfr mutations. There are other phenotypes (such as susceptibility to drought) which are often linked to freezing tolerance (Kasuga et al. 1999). Because of this sfr4, sfr5 and sfr8 were tested for alterations in osmotic stress tolerance.

3.11.1. Effect of osmotic stress on seedling emergence

Col-0, sfr4, sfr5 and sfr8 seeds were sown onto Petri dishes containing MS media supplemented with differing concentrations of mannitol; 0mM, 200mM, 300mM and 400mM. For each seed line at each concentration, there were five technical replicates. After seven days of growth under standard growth conditions, the number of seedlings that had scored ‘yes’ or ‘no’ for radicle emergence was counted under a microscope. The 0mM mannitol plates were used as a baseline level of germination. It was obvious that mannitol affected development, as there was a visible size difference between seedlings grown on MS media and those on media including mannitol (figure 3.22). However the number of seedlings that had undergone radicle emergence was extremely similar for all concentrations of mannitol. Seedlings that had germinated on the highest concentration of
mannitol were very stunted, and several developmental stages behind those grown on just MS agar, but radicle emergence was unchanged.

These results differ from those shown by (Boyce et al. 2003, Wathugala et al. 2011), who, when using these conditions, found 400mM mannitol to be enough to reduce radicle emergence to practically nil. However, three biological replicates, each with five technical replicates were conducted for this assay, and the results were consistent between all. As a result the experiment was repeated, with 0mM compared to 500mM mannitol.

Even at 500mM mannitol (figure 3.23), the emergence between the seed lines was not greatly altered (data not shown). Germination on MS media was at around 90-95% for all lines; this dropped to ~20% for all lines at 500mM. This would suggest that none of the sfr mutants are affected at the stage at which they first put out a radicle, however this does not eliminate the possibility that adult sfr plants are deficient in their ability to tolerate osmotic stress.
Figure 3.22. Petri dishes containing MS agar supplemented with four different concentrations of mannitol; 0mM, 200mM, 300mM, and 400mM. All seedlings were grown for seven days at 20°C in a Percival growth chamber, and were then viewed under a light microscope. Col-0 seedlings are used as a representative of what was seen for all seedlings tested.
Figure 2.23. Petri dishes containing MS agar supplemented with 500mM mannitol. All seedlings were grown for seven days at 20°C in a Percival growth chamber, and were viewed under a light microscope. Col-0 seedlings are used as a representative of what was seen for all seedlings tested.
3.11.2. Effect of osmotic stress on seedlings

Col-0, sfr4, sfr5 and sfr8 were tested for osmotic tolerance when floated on various concentrations of mannitol (0mM, 330mM, 400mM and 500mM) at 20°C under the standard long day conditions for 72 hours (figure 3.24). These were conditions used in previous studies (Knight et al. 2008, Wathugala et al. 2011).

None of the lines showed any signs of chlorosis when in water, as was expected. sfr4 showed a similar level of chlorosis to Col-0 with increasing concentrations of mannitol, and sfr5 seemed to perform better than Col-0, appearing much greener even on 500mM mannitol. sfr8 is significantly more chlorotic than the Col-0, indicating that sfr8 may have reduced osmotic tolerance compared to the Col-0.

Figure 3.24. Col-0, sfr4, sfr5 and sfr8 seedlings were floated on the indicated range of concentrations of mannitol for 72 hours and levels of chlorosis were judged against the Col-0 control.
3.11.3. Development of a seedling freezing assay

Freezing assays on whole plants require over two months to complete; conducting freezing assays on seedlings would reduce the time required to verify if an insert line or other mutant is freezing sensitive. Traditionally, freezing assays have been conducted on adult plants (Warren et al. 1996) however freezing assays conducted on seedlings have been previously trialled (Xin and Browse 1998). As a result, this method was developed and compared to results seen in adult plants. If the number of seedlings that survive freezing is a) comparable to that seen in mature Col-0 plants, and b) replicable, it would suggest that this method is a reliable source of information, and could potentially be used as a trial to determine whether a full whole-plant freezing assay was worthwhile. The third criterion was to test whether the method could distinguish a known sfr mutant from the Col-0.

Seedlings were sown on MS agar plates and grown in a Percival growth chamber for seven days. They were then transferred to 5°C for seven days, before being transferred to -7.5°C for 24 hours. When frozen, Petri dish lids were removed as to allow cold air to circulate. The seedlings were then returned to 5°C and then 20°C and were observed for re-growth. The plant samples this method was tested upon were Col-0 and sfr6; sfr6 has a very severe freezing sensitive phenotype and its use in this experiment was intended to determine whether both normal and deficient seedlings respond in the same way to freezing as adult plants.
Preliminary experiments suggested that Col-0 seedlings were able to recover from freezing to a degree that may be comparable to adult plants, that cold acclimation on plates is possible, and that the assay was capable of distinguishing between a freezing sensitive mutant and Col-0 (figure 3.25). sfr6 is naturally a much lighter colour than normal Col-0 seedlings; it is not only affected in its tolerance of freezing but in many other pathways (Knight et al. 2009). This difference in colouration makes no difference to the outcome of this experiment, but is just an example of what a ‘healthy’ sfr6 seedling looks like. It should be noted that freezing the agar plates changed the consistency of the agar on thawing, as can be seen in the figure. However, it did not seem to have an effect on the survival of the seedlings.

Figure 3.25. sfr6 and Col-0 seedlings were grown at 20°C for seven days, acclimated at 5°C for four days, and then frozen for 24 hours at -7.5°C. After freezing they were transferred to 20°C. Photographs were taken two days after the freezing event.
3.11.4. Dark/light regulated gene expression

It is known that sfr4 has reduced sugar accumulation (Uemura et al. 2003), and it is likely that sfr8 may have an altered perception of sugars, or an altered response to sugar. The sugar-responsive gene dark inducible 6 (DIN6) was used to see how these two sfr mutants responded to sugar. DIN6 is inducible in the dark and repressed by light or sugars. Seven day old Col-0, sfr4 and sfr8 seedlings were floated on either 5ml of 0.088M sucrose or 0.096M mannitol and subjected to either six hours in the light or six hours in the dark, both at 20°C. DIN6 gene transcript levels for each of the treatments were measured via qRT-PCR (figure 3.26).

![Graph showing expression levels of DIN6 in Col-0, sfr4, sfr8, with and without sucrose.](image)

*Figure 3.26. Expression levels of DIN6 in Col-0, sfr4, sfr8, when exposed to light or darkness, and mannitol or sucrose. Samples were subjected to each of the treatments for six hours, before samples were taken.*
is expressed in dark-treated plant tissue and is repressed by sucrose (Fujiki et al. 2000) as seen in the Col-0 control. Without sucrose, both sfr4 and sfr8 showed the expected response in both the dark (increased DIN6 expression) and in the light (reduced DIN6 expression). sfr4 shows higher levels of both dark and light expression than Col-0, particularly in the absence of sucrose, where DIN6 is evidently less repressed by sucrose than it is in Col-0. This could be as a result of the known sugar deficiency. sfr8 shows reduced expression under all conditions compared to Col-0. This would suggest that both sfr4 and sfr8 show an altered sensitivity to sugar, and this is potentially worth further investigation. As this experiment has no independent biological repeats, this result would need to be replicated two further times for these trends to be considered legitimate.

3.11.5. Flowering time in sfr8

Throughout this investigation it was noted that sfr8 repeatedly displayed an early flowering phenotype; however this was something that had not been reported in the literature. To test this observation, 20 individual sfr8 and Col-0 plants were grown up alongside each other under standard long day conditions (as described in the methods) and both the number of rosette leaves present on the plant when the flower bolt was 1cm long and the date that flowering occurred were recorded for each individual. The number of leaves present on the plant is an indication of the developmental stage the plant was in when flowering occurred (Lee et al. 1993, Knight et al. 2008)
There is an evident difference in both the number of leaves and the day by which flowering occurs; there was a one week difference between when all sfr8 plants and all Col-0 plants had flowered, which considering the age of these plants represent extremely different developmental stages. In figure 3.27, it is obvious that the sfr8 are much smaller and have a lower number of leaves than the Col-0, however all sfr8 plants had flowered by this point. Significant differences in flowering time between Col-0 and sfr8 were seen in two independent biological repeats. For the Student’s t-test, the P value was $2.12 \times 10^{-20}$ on a sample size of $n = 20$, suggesting there is a statistically significant difference between the number of leaves present when the flower bolt was 1cm long.

**Table 3.4: Flowering time of sfr8 compared to Col-0. The number of leaves on each plant when the flower bolt was 1cm long was measured, to the nearest leaf. Raw data was used for statistical tests. Day 0 = day seeds were transferred from the cold room to growth chamber.**

<table>
<thead>
<tr>
<th></th>
<th>Col-0</th>
<th>sfr8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average no. leaves per plant when flower bolt was 1cm long (to nearest whole leaf)</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>Day all plants had flowered by</td>
<td>34</td>
<td>28</td>
</tr>
</tbody>
</table>
Figure 3.27. Comparison of sfr8 (left) and Col-0 (right) in time taken to flower. Both sets of plants were sown at the same time and were grown under the same conditions. sfr8 plants are much smaller than Col-0 plants, and all sfr8 plants have established inflorescences, whereas much fewer Col-0 plants have yet to put up a flower stalk. Photograph was taken on day 28, with day 0 being the day that seeds were transferred into the Percival growth chamber.
3.12. Phenotypic results summary

For all three mutants, germination on mannitol did not vary from that seen in the Col-0 control. When floated on different concentrations of mannitol, there were slight differences seen between the responses of the different mutants. sfr5 seemed less injured than the Col-0 control, whereas sfr8 showed higher levels of damage at a lower mannitol concentration. sfr4 showed a similar level of tolerance to Col-0.

sfr8 and sfr4 both showed results that differed from the Col-0 control when DIN6 gene transcript levels were measured. sfr8 showed a potential hypersensitivity to sucrose, whereas sfr4 showed a potential lack of sensitivity to sucrose. The difference in flowering time seen between sfr8 and Col-0 was seen to be statistically significant, however it remains unknown if this is linked to the sensitive to freezing phenotype.
4. DISCUSSION

The original screen for sensitive to freezing (sfr) mutants was conducted in 1996, on a population of ethyl methanesulfonate (EMS) mutagenized Columbia 0 (Col-0) Arabidopsis (Warren et al. 1996). Due to the mutagenesis method, single nucleotide polymorphisms (SNPs) were the focus of this investigation. The screen was for reduced freezing tolerance post-acclimation, rather than innate freezing tolerance; all of the sfr mutants identified via this screen behave in the same way as Col-0 plants if frozen without an acclimation step. It is when the plants are acclimated that the difference between the freezing tolerance of the mutant and the Col-0 is seen. Arabidopsis is an ideal candidate for use in freezing tolerance screens as so much is known about how it acclimates and responds to freezing (Thomashow 2010). The use of forward genetics (the search for a genotype using a phenotype) works well in Arabidopsis, primarily due to the wealth of genetic information that is available, particularly for the Col-0 ecotype. Forward genetics studies have evidently been successful in the field of freezing tolerance; the genes responsible for the freezing tolerance deficit in sfr2, sfr3 and sfr6 have all been identified via this method (Knight et al. 1999, Kim et al. 2004, Thorlby et al. 2004, Amid et al. 2012).

The search for plants with disrupted post-acclimation freezing tolerance required a very precise freezing assay; a plant must be frozen to the point of being noticeably more damaged than the healthy Col-0 control, but the individual plant must ultimately survive and set seed to allow further investigation. This was negated in the isolation of the sfr mutants by the generation of sibling plants that displayed the same genotype, allowing plants to be thoroughly frozen (and thus die) without the genotype being lost (Warren et al.)
The aim of this study was to identify genes that were potentially responsible for the freezing sensitivity seen in the mutants \textit{sfr4}, \textit{sfr5} \textit{sfr8} and \textit{sfr9}. Quantitative real time PCR (qRT-PCR) experiments measured expression of the C-repeat binding factors (CBFs) and the cold regulated (\textit{COR}) genes protein kinase 2 (\textit{KIN2}) and galactosyl synthase 3 (\textit{GOLS3}) in cDNA from \textit{sfr} seedlings subjected to cold conditions. These experiments showed that the CBF pathway is not disrupted in \textit{sfr4}, \textit{sfr5} or \textit{sfr8}, indicating that all of these mutants are deficient in a different pathway leading to freezing tolerance, however a third independent biological repeat would be required to truly confirm this. There are a number of pathways that result in drought tolerance that also have cross talk with the freezing tolerance pathways, one such example is the dehydration-responsive-element-binding protein (DREB2)-mediated induction of the same \textit{COR} genes activated by the CBFs, discussed in more detail later. Abiotic stress tolerance networks are known to converge, and it could be possible that any of the \textit{sfr} mutants studied may have deficiencies in osmotic tolerance. As a result, this was something that was investigated.

\textbf{4.1. Analysis approach}

The original mapping interval of each of the \textit{sfr} mutants (Thorlby \textit{et al}. 1999), was further refined via personal communication with Glenn Thorlby (unpublished data). The basis of mapping relies on the fact that when a marker is close to the mutation (in terms of physical distance), the likelihood of a recombination event occurring between the two is reduced (Jones \textit{et al}. 1997). To determine the position between the two markers and the mutation, there must be two meiotic recombination events; one between the left marker and the mutation, and one between the right marker and the mutation. It is only when this occurs that it is possible to map a mutation to a certain region.
An Illumina genome sequence was conducted for each of the sfr mutants. These genome sequence data were then analysed in two different ways; which method was used for which sample depended upon how successful the genome sequencing run had been, and how well the data had mapped to the reference genome. One method, the command line method, required supercomputing power, an advanced knowledge of computer programming and various specific bioinformatics software packages. The resultant dataset was a list of SNPs and their position in the genome. The second method, the Galaxy method, required very little knowledge of bioinformatics and was an almost completely automated process; the output a file that was easily viewable on the Integrative Genomics Viewer (IGV) software package, and a manual comparison of the next generation sequencing (NGS) data and the reference genome to search for SNPs was required. A low frequency of mutations was found within some areas of the genome studied, which was not expected. Often, plants mutagenised via EMS experience a large number of SNPs, however a large number of these SNPs will introduce sense mutations that have no effect on phenotype. Mutations found by both methods were verified to exist within the genomic DNA of the respective sfr mutant by PCR and Sanger sequencing.

A collection of mutants have been developed in which a transfer-DNA (T-DNA) insert has been introduced into the genome; the insertion is random, however most genes have several of these lines available (Alonso et al. 2003). For the candidate genes for each of the sfr mutants, further mutant alleles were obtained in the form of these T-DNA insertional mutants, if available. In these lines, the expression levels of the gene potentially knocked-out/down by the T-DNA insertion was tested. If expression levels were significantly reduced,
the new mutant allele was used in a freezing assay alongside a Col-0 control, to test whether it responded in the same manner to acclimation and freezing. It was also compared to the sfr mutant it potentially corresponded to.

The method used to identify the genes responsible for each of the sfr mutations studied in this investigation involved the use of classical mapping combined with modern techniques; next generation sequencing and SNP calling. There are advantages and disadvantages to both classical mapping and in silico analysis of genome sequence in a search for mutations, and also to the other methods that are available for the identification of genes from a phenotype. The advantages of this method were that a large amount of the data analyses were conducted in silico reducing the amount of time taken; for classical mapping to be continued, an F₂ population of Landsberg erecta (Ler) x Col-0 would be grown up and genotyped to refine the interval for each of the sfr mutants. The large number of genetic differences that exist between the Col-0 and Ler ecotypes could be used as markers to identify a region in which each of the sfr mutations is found.

The downside of this mapping technique is that it can take an extremely long time. It can take many years to get from the EMS mutant to the gene in which the mutation occurs, even when working with the short generation time of Arabidopsis, which speeds up the genotyping of F₁ and F₂ generations. This method would not be possible for sfr9 as the only available resource for this mutant is genomic DNA. It has also been previously shown that in areas of low recombination (as were seen with sfr6), classical mapping can only go so far as to determine an interval in which a mutant exists and fine mapping is virtually impossible (Knight et al. 2009).
Another potential technique, which blends classical mapping with the use of next generation sequencing, is SHOREmapping. Like traditional mapping, it requires the original mutant to be crossed into another Arabidopsis accession (i.e. Ler), an F₂ population must be obtained, and the genomic DNA of 500 pooled F₂ plants that segregate for freezing sensitivity are subjected to bulk sequencing via Illumina next generation sequencing (Schneeberger et al. 2009). The read data is compared with the reference sequence, and the data are virtually mapped using the known markers existing between Col-0 and Ler. Using these data, the SHOREmap software finds an interval in which only Col-0 DNA is present, meaning that the mutation causing the phenotype will be found within this interval. SHOREmapping software is used to refine this interval, and the final result is an extremely small number of SNPs that may be responsible for the mutant phenotype (Schneeberger et al. 2009).

While SHOREmapping does still require the basic level of mapping, (mutant plants must be grown and crossed with Ler plants), when an F₂ population is achieved the process transitions into mapping-by-sequencing. However, conducting freezing tests on 2000 individual F₂ plants in order to find the one quarter that were freezing sensitive would be no trivial matter. Much like continuing with classical mapping, this technique would not be suitable for use with sfr9.

Were two alleles present for each of the sfr mutants involved in this study, next generation sequencing runs could be conducted for the two separate alleles, and the gene which displayed two different mutations would be pulled out as that responsible for the freezing
sensitivity. This is not something that could be conducted for the sfr mutants, as all only have one available allele.

4.1.1. Sequencing and bioinformatics

Next generation sequencing data were obtained, involving the preparation of a genomic DNA library, which was then sequenced via the Illumina method. An entire genome sequence was produced for each of sfr4, sfr5, sfr8 and sfr9. Due to the EMS mutagenesis treatment each of the sfr mutants was subjected to, it was known that a higher number of SNPs would be present within the mapping interval for each sfr mutant than would be causing the sensitive to freezing phenotype. Likewise, there would also be benign, naturally accumulated SNPs that were found in all individuals of the specific Col-0 population that was used to generate the sfr mutants. Only one of these SNPs would be responsible for the freezing sensitivity, however there would be a great number of silent mutations introduced into the genome. There were a large number of false positives in this investigation, with very little way of determining what was a ‘legitimate’ candidate and what was not without using PCR and Sanger sequencing. A number of criteria (methods section 2.11.4) were applied to the two mapping methods in an attempt to remove as many of these false positives in silico. Additionally, a SNP that occurred in more than one of the sfr mutants could be discounted as it was extremely likely to be a population-specific SNP.

The sequence data received were of various different qualities, and it became evident that Illimuna sequencing cannot always provide adequate sequencing depth required for the Galaxy method. Regions near centromeres, tandem repeats, and sequences that occur more than once in different parts of the genome can all reduce the efficiency of the alignment of
sequence data to the reference genome (Treangen and Salzberg 2011). It could be that these areas occur over region to which a particular sfr mutation has been mapped. In this instance other techniques may have to be adopted, such as sequence capture, where oligos are designed to a certain region rather than sequencing the whole genome, and this region is enriched and given a deeper level of coverage (Hoppman-Chaney et al. 2010).

4.1.2. Methods used for bioinformatics analysis

The Galaxy method is widely used and accepted as a feasible way of pulling SNPs out of the Illumina data (Hillman-Jackson et al. 2002). The method is time-consuming, and can be liable to human error in the mis-recording of SNPs, however it does give a visual representation of the genome meaning it is very obvious if a region has not mapped well to the reference genome. The original dataset received for sfr4 was a good example of this advantage. With this method it is possible to tailor the mapping software Bowtie to the requirements of the user, and this is something that may yield better mapping results if repeated multiple times with slight changes to the settings (Langmead et al. 2009).

Human error is a negligible factor using the command line method, however the method employed was not dedicated to this specific type of SNP-calling; which entailed locating purely homozygous SNPs. When the sfr5 data were analysed, there were seven (table 3.1) heterozygous SNPs within genes. All were proven to not exist within the genomic DNA of sfr5 in the DNA, bar one which was not tested on the grounds of how low the likelihood of it existing was. Likewise, the criteria of the command line method are biased towards high levels of genome coverage; as part of the criteria used in the command line method, areas of the genome must have 10x coverage for a SNP to be considered legitimate. This resulted
in SNPs that were found via the Galaxy method and were proven to exist within the DNA (e.g., the \textit{sfr8} SNP in \textit{murus 1} \textit{[MUR1]}) failing to be identified via the command-line method. Just because an area of genome displays low coverage, this does not mean that any mutations found within these regions are not legitimate; all of these must then be retested within the genomic DNA to determine whether they are legitimate.

The comparison of the same data using both methods is the most interesting with regards to this; it was of interest to see if both methods yielded the same SNPs as likely candidates. The overall number of SNPs the command line picked up on was lower than those found via the Galaxy method; for \textit{sfr8}, two of the candidates genes that were investigated were found via the command line method, however \textit{MUR1} was not. The region of the genome in which \textit{MUR1} exists did not display as much sequence depth as other regions, and as a result probably did not meet the criterion for genome coverage that was present in the command line method. As a result, a suggestion in modifying the command line method for this region of the genome would be to lower the stringency, and accept regions of lower coverage. Likewise, the mutation found in the \textit{sfr9} genome sequence was not identified by the command line method; in fact, the command line method did not detect any potential candidates for \textit{sfr9}. This was also the case for the re-sequenced \textit{sfr4} data.

While the command line method has produced fewer SNPs for the two Illumina datasets where the data were of sufficient quality to be analysed by both methods, it should not be forgotten that for \textit{sfr5}, the command line method was the only method that yielded SNPs. The data for \textit{sfr5} did not map well with Galaxy, and it was not possible to pull SNPs out of the data it provided. If possible both SNP identification methods should be used, as the
command line method is potentially better at mapping the raw NGS data to the genome. With reduced stringency levels, it could be even more successful.

Providing the raw data are of high enough quality, the Galaxy and Bowtie online platform allow a user with very little experience of bioinformatics to successfully identify legitimate candidates genes for mutations caused by SNPs. Having proven that the Galaxy method is able to successfully identify SNPs is extremely useful, as next generation sequencing is becoming an extremely useful tool, and the ability for an end-user to do their own data analysis is crucial.

4.1.3. T-DNA Insert lines

SALK and GABI-Kat T-DNA lines are a cheap and relatively efficient method of knocking out or down the potential candidates from the SNPs identified from next generation sequencing data (Alonso et al. 2003). This project depended heavily upon the use of SALK and GABI-Kat T-DNA insertional lines, and throughout the process it was discovered that there were various problems that could occur and hence hinder progress. Even after the process of genotyping an insert line (which in itself can be time consuming, and isolating a homozygote can take several months) there is no guarantee that the insertion will cause a change in expression of the transcript. This was a problem with some of the lines that were available. Insertions were in extremely poor positions in the gene, including intragenic regions, too far into the untranslated regions, or in introns. Throughout this investigation there were several cases when a SALK line failed to reduce the expression of the gene, thus rendering it useless. Conversely, it was sometimes seen that the T-DNA insert could bring about increased levels of expression for a given gene; this is a known problem due to the CamV 35s promoter used
to insert the T-DNA (Ülker et al. 2008). Another problem was that for some genes, there was no available T-DNA insert line, as was the case for the sfr5 candidate, At1g15410.

Even when insert lines were available, several lines displayed weak germination and survival rates, meaning that only one qRT-PCR experiment could be conducted when testing the effect of the T-DNA insert on the expression of the gene. Due to time limitations it was not possible to allow these plants to self-pollinate and grow up enough progeny to allow repeats of the qRT-PCR tests, however due to the nature of the knockdown mutant, i.e. the gene has been inactivated by an insertion into the gene which should be present regardless of what abiotic conditions the plant is subjected to, I believe these result are sufficient to make an informed judgement on the efficiency of the inactivation of the gene.

It may not always be preferable to completely eradicate a gene transcript; difficulties were encountered when genotyping the insertional mutants for the sfr8 candidate, MUR1. No homozygote individuals were successfully obtained. The known mur-1 and mur-2 mutants have extremely low levels of L-fucose (Zablockis et al. 1996), but production is not completely knocked out; it is possible that a complete knockout (as may be caused by a T-DNA insert line) may be lethal.

If only one insertional mutant is available and it successfully knocks down the gene, this is not sufficient to confirm linkage and it is preferable if the mutant can be complimented with a wild type sequence. As a result, to be completely confident in the result of an insertional mutant, it may be preferable to supplement the results using an RNAi knockdown. The results of RNAi are often not as extreme as T-DNA insertional lines; the gene in question is
often knocked-down rather than out, however this can be advantageous when a complete knockout is lethal, and is a valuable addition if no T-DNA insert lines are available. RNAi is, however, a more labour intensive method. A fragment to be focussed on must be chosen, cloned into the appropriate vector and infiltrated into the plant tissue. Certain fragments do not successfully transfer into vectors on the first attempt, which can result in the process becoming time consuming. This method is not particularly feasible for a large number of candidates, and it was not possible to create RNAi clones and successfully transform wild type plants on the timescale of this investigation.

4.1.4. Plant freezing assays

While extremely useful, the standard freezing assay is slow and laborious, and can sometimes yield unexpected results, including inconsistencies between repeats. There are an extremely large number of variables that can affect the results of a freezing assay and minimising variability is the key to producing replicable results. Conventional freezing assays on adult plants take a number of months to complete, and as a result finding a quick and easy of way of verifying the freezing sensitivity of plants while at the seedling stage would have been very beneficial. An assay that did just this was developed, however this occurred at the very end of the period of study and as a result was not used. It may, however, prove useful in future investigations of this sort.

One particular problem with freezing assays is choosing a freezing temperature that will give reproducible results, i.e. a temperature that will always result in the death of freezing sensitive plants. Likewise, the position and number of individual plants within the freezing chamber can have an extreme effect on the survival of plants. sfr5 is known to display a less
severe phenotype than other mutants tested (McKown et al. 1996), which makes freezing assays more challenging, as it was seen on one occasion that sfr5 survived freezing conditions that had been previously shown to be lethal. This occurred when the growth chamber contained a larger number of plants than usual, suggesting that overcrowding resulted in raised the temperature around the plants, or protected plants from ice formation. In initial investigations, plants were transferred from a 5°C growth chamber to a freezing chamber, pre-cooled to -7.5°C. This is an extremely unnatural situation, and it is known that Arabidopsis plants growth differs between environmental growth chambers and an outdoor environment (Mishra et al. 2012). The sudden shock of such a substantial drop could potentially cause more damage than a gradual cooling to this temperature; this is something that has been shown in other species (Finkle et al. 1974). Transferring the plant from the 5°C growth chamber to the -7.5°C freezing chamber when the temperature is decreasing could be an option.

Edge effects and differences in air flow within the freezing chamber were a concern, and attempts were made to control these to some degree by randomising plants of different lines within the tray. Placing the peat plugs on trays that allow air circulation may be an alternative method to prevent insulation problems.

By the time adult rosette plants are suitable for freezing tests, it is often the case that they have become root-bound by their peat plugs, and that roots have managed to penetrate the outer casing of the plug, leaving them exposed when subjected to freezing temperatures. This potential root exposure is not a situation that the plants would experience in nature, and as a result the damage that occurs to the roots may have a significant effect upon their
recovery. Somehow, protecting the roots of adult plants while they are frozen may produce more accurate, reproducible results.

Throughout this investigation it was discovered that plants of different ages showed that their responses to freezing might differ. Plants that were four weeks old when they were put into acclimation were less likely to recover from a freezing event than plants that were five weeks old, suggesting that further investigations into optimal plant age are required.

The results of a freezing assay could be supported by electrolyte leakage assays (Jaglo-Ottenisen et al. 1998). The one advantage of electrolyte leakage assay is that leaves are tested over a range of freezing temperatures, so it is not crucial to find an optimum freezing assay temperature. However, this is not something that could have been conducted on all of the candidate mutants due to live plants being unavailable for sfr9.

4.2. Results from each of the mutants

4.2.1. sfr9

Due to the fact that the sfr9 mutant no longer germinates, nothing can be done in the way of phenotypic investigations. Genomic DNA used for the genome sequence was extracted from seeds. The one mutation found for this mutant (via the Galaxy method) is between two genes and is potentially within the promoter region of glucosinolate transporter 2 (GTR2), however, it may be too far away from the closest gene to have any effect upon its expression. This area of the genome was checked for ESTs and none were found, and it was verified that throughout the mapping interval for this mutant, the genome sequencing coverage was complete and without gaps. As the mutation causing the sfr9 phenotype must
be within this interval, there is very little that could be done with this mutant. One suggestion would be to re-sequence this region using the sequence capture method discussed previously, focussing specifically upon the mapping interval.

4.2.2. sfr4

The deficit in freezing tolerance in sfr4 has been shown to be due to a significantly reduced accumulation of sucrose during cold acclimation when compared to Col-0 plants (Uemura et al. 2003). When an sfr4 plant has been frozen without acclimation, levels of sucrose are not thought to differ from those seen in Col-0. sfr4 has also been shown to not accumulate anthocyanin to the same levels as the Col-0 in response to cold (McKown et al. 1996, Uemura et al. 2003). Anthocyanins, which are antioxidants that protect plants from damage by reactive oxygen species (Nagata et al. 2003), have been shown to accumulate in response to increased sucrose levels. Therefore, this lack of anthocyanin in sfr4 could be the product of the sucrose deficiency (Teng et al. 2005). The dark/sucrose gene expression experiment (3.26) suggests that either the perception of sucrose, or the cellular levels of sucrose have been altered in sfr4, which is consistent with previous findings.

sfr4 showed no unusual germination response compared to Col-0 when grown on mannitol-supplemented media, nor when seedlings were floated on various concentrations of mannitol. These results indicate that the freezing-sensitive phenotype of sfr4 is not accompanied by an osmotic sensitivity (Warren 1996). The use of compatible solutes to prevent water loss from cells occurs in both droughted conditions and when freezing of plant tissue occurs. As a result, this would suggest that when plants have been grown under
ambient temperature conditions, the freezing and osmotic tolerance of $sfr4$ should be equivalent to that of Col-0.

It has been previously proven in protoplasts that $sfr4$ has a deficit in sucrose, and that supplementing the mutant with sucrose results in the restoration of freezing tolerance (McKown et al. 1996). Using protoplasts is not always an accurate representation of the response on a whole plant scale, and as a result the intention was to use the Petri dish freezing assay, with sucrose-supplemented media, to see if $sfr4$ seedlings showed the same freezing deficit when sucrose was readily available. However, this was not performed due to the late development of the assay, and is a suggestion for further work.

Although $sfr4$ is the most well studied of the four $sfr$ mutants involved in this investigation, and as a result was expected to be the easiest to identify, no candidate genes were isolated. The initial sequencing run provided data that were of low coverage to the point of being unusable, however, even when the Illumina sequencing was repeated, the command line method was unable to find any candidate genes from the new data. Due to the late availability of the new data, they were not analysed via the Galaxy method. Even when the mapping interval for $sfr4$ was extended, as the original interval was based on only one recombinant for one of the flanking markers, no further candidates were found. It is possible that, though the overall genome sequence coverage was high, regions within the mapping interval that contained $sfr4$ were areas that were under-represented in the sequencing, and as a result the SNP causing the $sfr4$ mutation could simply not have sequenced, or had so few reads that it was not able to be taken as a confident SNP.
Despite the lack of candidates, when every gene in the sfr4 interval was surveyed, none seemed to have any obvious links to sucrose, even though the mutation is present within this region. This suggests that the pathway that is disrupted in sfr4 may be more complex than simply governing the accumulation of sucrose.

4.2.3. sfr5

sfr5 has no reported phenotype other than its freezing deficiency, and of the sfr mutants it shows the weakest deficiency in freezing tolerance; the sensitivity phenotype is sometimes variable (McKown et al. 1996). When originally isolated, two alleles for sfr5 were identified and were found to be co-dominant (Warren et al. 1996); one allele sfr5-1 is the allele focused upon in this experiment, and it is homozygous. Unfortunately sfr5-2 seeds were no longer available at the start of this investigation; if both alleles had undergone genome sequencing only gene(s) which contained SNPs in both genomes could have been focussed upon. This would have greatly limited the number of candidate SNPs that may have been responsible for the phenotype.

At the seedling emergence stage sfr5 appeared to be affected to the same extent as Col-0 when grown on mannitol-supplemented media, but no more. When seven days old, seedlings responded better than Col-0 when floated in mannitol, and suffered less chlorosis. sfr5 appears more tolerant of osmotic stress than Col-0. This was not the result expected, and could potentially suggest a disrupted osmolyte pathway. The freezing sensitive phenotype seen in sfr5 was the weakest of the mutants tested, and under certain conditions could be seen to disappear; this is discussed in section 4.1.4.
Due to the poor sequence data quality for this mutant, these data were only analysed using the command line method. The two potential genes found for this mutation were At1g15410 and AVP1, a H⁺ translocating inorganic pyrophosphotase. Through the use of one confirmed homozygous GABI-Kat line, it was found that an avp1 mutant responded to freezing stress in a similar way to Col-0 plants, and as a result it would appear that a mutation in AVP1 was not responsible for the freezing sensitivity. Due to the severity of the homozygous avp1-1 mutation, this is not wholly unexpected; avp1-1 is unable to flower and hence cannot set seed (Li et al. 2005), and this is something that has never been seen with sfr5. It is possible that while AVP1 expression had been knocked out in the GABI-Kat mutant, there was enough protein remaining for it to be functional, as none of the phenotypes associated with avp1-1 were seen with the T-DNA avp1 mutant. As a result this would suggest that further work be carried out upon this candidate.

The remaining candidate for sfr5 was not investigated due to unavailability of SALK or GABI-Kat insert lines. Of those data supplied via the command-line method, At1g15410 is currently the only other prospective candidate for this mutant; all weaker mutations were proven not to exist in the DNA of the sfr5 plant. The next course of action, had time permitted, would be to make an RNAi knock down of this gene. At1g15410 is a hypothetical protein that has been suggested to be located in the chloroplasts, and has been proposed to be part of an aspartate-glutamate racemase family. Not much is known about the function of this gene, however it has been found that aspartate racemases and glutamate racemases are enzymes which have been found to show high levels of similarity to each other in bacteria, and that they do not require a co-factor to function. (Gallo and Knowles 1993) Glutamate racemase catalyses the conversion of D-glutamate to L-glutamate, and vice-versa
(Gallo and Knowles 1993), and they do not need a co-factor to function. In bacteria, glutamate racemases are essential for cell biosynthesis, and D-glutamate is necessary for the synthesis of peptidoglycan found in the cell wall (Fisch 2009). Likewise, aspartate racemases convert D-aspartate to L-aspartate, and also does not require a co-enzyme (Yamauchi et al. 1992). It is possible that the involvement of these proteins in bacterial cell walls could infer an involvement of At1g15410 in the plant cell wall.

4.2.4. sfr8

There was no prior knowledge as to the cause of the freezing tolerance deficit in sfr8, other than that the mutation was recessive and caused a loss-of-function. The CBF pathway was shown to be functioning as normal (figure 3.1), and the osmotic tolerance assay for those both grown on plates and floated on solution showed similar results to the Col-0 control (figure 3.24).

During the investigation it was noted that sfr8 flowered much earlier than Col-0; in a sample size of twenty, sfr8 plants flowered an average of seven days earlier than Col-0 under long day conditions. This phenotype was seen under ambient conditions, with no potential causes of stress. Flowering and freezing sensitivity phenotypes often appear in the same mutant; sfr6 is known to flower later than Col-0, due to reduced expression of genes involved in the photoperiodic flowering pathway (Knight et al. 2008). It is likely that there are links between the flowering time pathway and freezing tolerance. One gene, Long Vegetative Phase 1 (LOV1), links freezing tolerance to flowering time and, when mutated, results in delayed flowering under long day conditions (Yoo et al. 2007). LOV1 is a plant-
specific NAC-domain transcription factor, which has been shown to negatively regulate 
Constans (CO) and Flowering Locus T (FT), two integral components of the flowering 
pathway (Turck et al. 2008). Alongside this regulation of CO, it has also been shown to be 
involved in the cold response. The lov1-4 null mutant was shown to display reduced freezing 
tolerance, whereas overexpression of LOV1 caused increased expression of the genes Cold-
Regulated 15a (COR15A) and KIN1. In the lov1-4 mutant, COR15A was down-regulated, 
however KIN1 was unaffected, suggesting LOV1 is not required for KIN1 expression. It was 
theorised that LOV1 may bind to the CRT/DRE of COR genes. Another late flowering mutant 
is High Expression of Osmotically Responsive Genes 9 (HOS9). HOS9 encodes a 
homeodomain transcription factor, and it is suggested to be responsible for constitutive 
freezing tolerance, as hos1 mutants are more sensitive to freezing before and after 
acclimation (Zhu et al. 2004).

Of the sfr8 candidate genes, At3g50910 was proven to not be the cause of the sfr8 
mutation. Two confirmed homozygous SALK lines were tested for expression levels and 
were proven to successfully knock out the expression of the Col-0 gene. These SALK lines 
were tested via a freezing test, that even when expression of the gene containing the T-DNA 
was significantly reduced, the T-DNA mutant plants responded to freezing in a similar way to 
Col-0. The gene UDP-Glucose Pyrophosphorylase 3 (UGP3) was not rejected as a candidate, 
however due to germination problems with the T-DNA insert line used, its sensitivity to 
freezing when the gene was knocked out was not freezing tested.

The remaining sfr8 candidate of those tested, MUR1, encodes the enzyme GDP-D-mannose-
4,6-dehydratase, which is responsible for catalysing the first step in the de novo synthesis of
GDP-L-fucose, a precursor of L-fucose (Bonin et al. 1997). L-fucose is present in
glycoproteins and polysaccharides found in the plant cell wall. Previously generated MUR1
EMS mutations, mur1-1, mur1-2 and mur1-3 were used in this investigation (Reiter et al.
1993). mur1-1 and mur1-2 are extremely similar; both show less than 2% of the L-fucose
seen in Col-0 plants in aerial tissues, but only a reduction to 40% of Col-0 levels in roots
(Bonin et al. 1997). Both are slightly dwarfed, mur1-2 slightly more so than mur1-1. This
dwarfism compared to the Col-0 control can be seen in the freezing assay conducted upon
them (fig 3.17). mur1-3 is less badly affected, with ~33% L-fucose in aerial tissues (Reiter et
al. 1997).

In mur-1 and mur1-2, the 98% reduction in L-fucose levels has been shown to be non-lethal,
as L-fucose is substituted with L-galactose in the mutant plants (Zablackis et al. 1996). While
this substitution may function under normal conditions, under freezing conditions the
structural differences between the two compounds may cause the cell wall to weaken and
fail. mur1-3 reacts to freezing in a similar way to Col-0, suggesting that if the plant is able to
synthesize some L-fucose but not to the normal level seen in Col-0 plants, the cell wall is
still functional under normal conditions.

The missense mutation introduced into MUR1 in sfr8 is G to A, which is characteristic of
EMS mutations. The SNP corresponds to an already existing mur1 mutant, mur1-4, which
has been previously described as having an extremely low L-fucose content compared to the
Col-0 control, proving that this exact amino acid substitution has a detrimental effect on the
function of this protein. The area of the gene in which the mutation occurs is extremely
conserved, and as a result must be important to the correct functioning of MUR1 (Bonin et al. 1997).

There are a number of experiments that could be conducted to either confirm or remove MUR1 as a candidate for sfr8. The most obvious of these would be to supplement the growth media of sfr8 seedlings throughout their growth, then subsequently subject them to a freezing assay. If this restored freezing tolerance in the adult plants, it would implicates that MUR1 is sfr8. There are also a number of phenotypes that are associated with the mur1 mutants that have not yet been verified to exist within sfr8. The freezing sensitivity has been confirmed, but would need to be repeated for true reliability; furthermore, electrolyte leakage tests would be a useful addition in comparing the two. The flowering phenotype that is seen in sfr8 could be measured in mur1 mutants, however this phenotype has not previously been reported for mur1 mutants, and as a result could be due to a SNP induced by the EMS that is independent of the gene causing the freezing sensitivity.

Two strong alleles of mur1; mur1-1 and mur1-2 have been tested and, as no homozygotes were successfully obtained for the T-DNA insertional lines, possibly due to the lethal nature of a complete MUR1 knockout (Bonin et al. 1997), an RNAi knockdown may be more appropriate. Due to the deficit of L-fucose caused by the mur1 mutation, growing on media containing fucose restores the phenotype of the mur1 mutants (Freshour et al. 2003), and this is something that could be attempted with sfr8 and could strengthen the case of this candidate being responsible for the freezing tolerance deficit in sfr8. mur1-1 and mur1-2 mutants have slightly weaker inflorescences and stems than normal plants; this was not something that was confirmed in sfr8 due to the fact that when plants
began to flower, they were fitted with the Aracon system to prevent cross-contamination of seed. There is also a greater internode length in *mur1* plants compared to Col-0. This is something that could warrant further investigation. Another observation previously made about *mur1* mutants, was that root cell elongation is somewhat impaired, even though the fucose deficiency is not as severe in the roots as the shoots. This is something that could be measured in *sfr8* via microscopy (Van Hengel and Roberts 2002). Conducting the petri dish freezing assay on both *mur1* and *sfr8* but with L-fucose-supplemented media is an option for further work; if the dwarf phenotype and cell wall sensitivity were removed, this would give evidence that *MUR1* is *sfr8*.

The results for *DIN6* expression suggest that *sfr8* is hypersensitive to sugars (figure 3.26), which could correspond with the hypersensitivity sucrose signals that have previously been recorded in *mur1* (Li *et al.* 2007), providing more evidence that *sfr8* may be caused by a *mur1* mutation. When sucrose is added, the gene expression is somewhat more repressed in *sfr8* than in Col-0. Similarly, when sugar is not added, *sfr8* acts in the way Col-0 acts when sugar has been added. It is a possibility that *sfr8* is responding to the small amount of sugar that will already be present in the plant, acting in the way that a supplemented Col-0 plant would (Li *et al.* 2007).
4.3. Final summary

A number of gene candidates were identified for each of the SFR genes except for sfr4, and were verified to exist within the DNA of the sfr mutant to which they corresponded. Throughout the course of this investigation, more information has been gathered about the responses of several of the sfr mutants to different abiotic factors. For one, it has been confirmed that none of the three mutations tested affect the CBF pathway. While it could be argued that a gene downstream of the CBFs could be knocked-out in the sfr mutants, it is unlikely that one defective COR gene would result in the complete loss of freezing tolerance seen in the sfr mutants. The CBFs are the best-known factor involved in freezing tolerance, however they may only control 12% of the cold induced transcriptome (Fowler and Thomashow 2002). It is evident that a functioning CBF pathway alone cannot confer freezing tolerance, and as a result there must be a number of other factors involved.

In terms of candidate genes for each of the sfr lines, both sfr4 and sfr9 have resulted in no candidates. The lack of candidates for sfr4 may be due to the higher stringency levels of the command line method, and may be able to produce SNPs if the stringency is lowered, however, it is unknown how a candidate will be found for sfr9, as there appears to be no apparent reason as to why one has not been discovered. sfr5 is still potentially left with two candidates that required further investigation. However, due to the fact that the sfr5 data could not be analysed with both methods, it is possible that there are SNPs that the Galaxy method would have picked up on had the data quality been better.

sfr8 was the most successful of all the sfr mutants studied, with three candidates still remaining. The freezing sensitivity of plants lacking UGP3 remains unknown, and as a result it cannot be ruled out. There is also At3g56590, which was identified too late to be of any
use in this investigation which warrants further work. However, at this point *MUR1* appears to be the strongest candidate to be causing the freezing sensitivity, with a known knockout proven to exist at the exact same base change as that seen in *sfr8*; further tests on the known *mur1* EMS mutants with regard to their freezing sensitivity is strongly recommended.
5. APPENDICES

5.1. Primer sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primers looking for <em>sfr8</em> mutations</strong></td>
<td></td>
</tr>
<tr>
<td>CCATGCCTA CTTT TTCTGCT</td>
<td>Forward primer to verify that there was a mutation in the gene At3g50910</td>
</tr>
<tr>
<td>AAGCATGACAAGGCGCTACT</td>
<td>Reverse primer to verify that there was a mutation in the gene At3g50910</td>
</tr>
<tr>
<td>TTCTCGGAAAGGCTACGAA</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At3g51160</td>
</tr>
<tr>
<td>TTGCTGCTTACGATCCATGT</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At3g51160</td>
</tr>
<tr>
<td>TGGAAAAAGCTATTGCTGCTG</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At3g51780</td>
</tr>
<tr>
<td>TCTCCAAAACCTCGTAGCCATAA</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At3g51780</td>
</tr>
<tr>
<td>CTGGGTGTGATTGCAAATGG</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At3g56040</td>
</tr>
<tr>
<td>TGATTTTGCACTCGAAGCCTG</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At3g56040</td>
</tr>
<tr>
<td>TTGAAACCTGTGGAGTTCAAGG</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At3g56590</td>
</tr>
<tr>
<td>TTCCAAACAAAGGATTTCAG</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At3g56590</td>
</tr>
<tr>
<td>Sequence</td>
<td>Function</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Primers looking for sfr5 mutations</strong></td>
<td></td>
</tr>
<tr>
<td>ACCGTCTTCAGTCTCTTTGC</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At1g15410</td>
</tr>
<tr>
<td>TCCATTGGATCAACACACTTCT</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At1g15410</td>
</tr>
<tr>
<td>GGCCTTGCTCTTGGTTACAA</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At1g15690</td>
</tr>
<tr>
<td>CGAAGAAGGGAGCAAAGACA</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At1g15690</td>
</tr>
<tr>
<td>TGAATCCGGTTCATAAGAAAA</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At1g17180</td>
</tr>
<tr>
<td>TTCCGAAAACAAGGAAGAAACA</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At1g17180</td>
</tr>
<tr>
<td>CTAGCTTTGGCGTGATTCTCG</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At1g13570</td>
</tr>
<tr>
<td>GCGGAACAAGGGAGACAAAAAA</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At1g13570</td>
</tr>
<tr>
<td>GAGATCTTGGGTTGTGAATCG</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At1g16110</td>
</tr>
<tr>
<td>CAAGCACCTGTTCATGCT</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At1g16110</td>
</tr>
<tr>
<td>TGATGATCTTGGGCGGCGTCT</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At1g14120</td>
</tr>
<tr>
<td>AATCTATCTTTGGGACGTACCTT</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At1g14120</td>
</tr>
<tr>
<td><strong>Primers looking for sfr9 mutations</strong></td>
<td></td>
</tr>
<tr>
<td>TGGTGTTGGTGGTCTAGGT</td>
<td>Forward primer to verify presence of a SNP mutation in the gene At5g62680</td>
</tr>
<tr>
<td>AGAGGATGGTGGAGATGTCG</td>
<td>Reverse primer to verify presence of a SNP mutation in the gene At5g62680</td>
</tr>
<tr>
<td>Sequence</td>
<td>Function</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ATCTTCCCTCAGCTTTTCTTCCG</td>
<td>Right primer for genotyping of insertional mutants SALK_074693C and SALK_124555C</td>
</tr>
<tr>
<td>GCTGATTCCACCTGGGTCTAG</td>
<td>Left primer for genotyping of insertional mutants SALK_074693C and SALK_124555C</td>
</tr>
<tr>
<td>TGTGTTTCGGTGAAATTCTGG</td>
<td>Right primer for genotyping of insertional mutant SALK_132810C</td>
</tr>
<tr>
<td>CATCCCAAGCATCATAAAC</td>
<td>Left primer for genotyping of insertional mutant SALK_132810C</td>
</tr>
<tr>
<td>TTTTCGTGCGAGAATTGAATC</td>
<td>Right primer for genotyping of insertional mutant SALK_150964</td>
</tr>
<tr>
<td>ATGACAAAGGCCTACTGCTGC</td>
<td>Left primer for genotyping of insertional mutant SALK_150964</td>
</tr>
<tr>
<td>ACAATGGGGTCAGAGAACAAC</td>
<td>Right primer for genotyping of insertional mutant SALK_027379</td>
</tr>
<tr>
<td>AGCATCCATGTATCCAGCATC</td>
<td>Left primer for genotyping of insertional mutant SALK_027379</td>
</tr>
<tr>
<td>TTTCTTCTGGAAAAGGCTAC</td>
<td>Right primer for genotyping of insertional mutant SALK_027387</td>
</tr>
<tr>
<td>TCATCCTCAACTCATTGGAGG</td>
<td>Left primer for genotyping of insertional mutant SALK_027387</td>
</tr>
<tr>
<td>AGCATCCATGTATCCAGCATC</td>
<td>Right primer for genotyping of insertional mutant SALK_057153</td>
</tr>
<tr>
<td>ACAATGGGGTCAGAGAACAAC</td>
<td>Left primer for genotyping of insertional mutant SALK_057153</td>
</tr>
<tr>
<td>GTGCGTTTCTTCAGTAGCTTC</td>
<td>Right primer for genotyping of insertional mutant SALK_020645C</td>
</tr>
<tr>
<td>GGCTTTAATTCGCGTAGATC</td>
<td>Left primer for genotyping of insertional mutant SALK_020654C</td>
</tr>
<tr>
<td>TAACAGAGCCACGTATAAGG</td>
<td>Right primer for genotyping of insertional mutant SALK_052178C</td>
</tr>
<tr>
<td>AACACGAAAACGTGGGTGAC</td>
<td>Left primer for genotyping of insertional mutant SALK_052178C</td>
</tr>
<tr>
<td>ATTTTGGCGATTTCGGAAC</td>
<td>SALK LBB 1.3 (Left Border primer)</td>
</tr>
<tr>
<td>CATTGTTGCAACGTAGGTCTG</td>
<td>Forward primer for genotyping of insertional mutant GK596C07</td>
</tr>
<tr>
<td>GAGCATAACAGAGACAGCAAC</td>
<td>Reverse primer for genotyping of insertional mutant GK596C07</td>
</tr>
<tr>
<td>GACAGACTGCCTAGCATTGGAGT</td>
<td>GABI-Kat Left Border primer</td>
</tr>
<tr>
<td>Sequence</td>
<td>Function</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CTGAAATCCAGATCCGAAA</td>
<td>Forward primer to measure transcript levels of <em>MUR1</em> in Col-0</td>
</tr>
<tr>
<td>ATCTTCCTCGGTTCAACGAC</td>
<td>Reverse primer to measure transcript levels of <em>MUR1</em> in Col-0</td>
</tr>
<tr>
<td>AGAGCTGAGAGGACGAGGAG</td>
<td>Forward primer to measure transcript levels of <em>UGP3</em> in Col-0 and SALK_020654C</td>
</tr>
<tr>
<td>ACCATGGCTAATCCTCAAGC</td>
<td>Reverse primer to measure transcript levels of <em>UGP3</em> in Col-0 and SALK_020654C</td>
</tr>
<tr>
<td>GAAAGTCGTTTATAGGGCTGGA</td>
<td>Forward primer to measure transcript levels of <em>GTR2</em> in Col-0 and SALK_052178C</td>
</tr>
<tr>
<td>GTGTTCAATGATCCCAAGC</td>
<td>Reverse primer to measure transcript levels of <em>GTR2</em> in Col-0 and SALK_052178C</td>
</tr>
</tbody>
</table>
5.2. Primer verification graphs

Primer validation for AVP1 qRT-PCR primers

\[ y = 0.8422x + 14.396 \]
\[ R^2 = 0.9932 \]
Primer validation for At1g15410 qRT-PCR primers

\[ y = 1.0004x + 21.162 \]

\[ R^2 = 0.9894 \]
Primer validation or MUR1 qRT-PCR primers

\[ y = 0.9962x + 17.912 \]

\[ R^2 = 0.9967 \]
Primer validation for UGP3 qRT-PCR primers

\[ y = 0.9504x + 18.233 \]

\[ R^2 = 0.9973 \]
Primervalidation for GTR2 qRT-PCR primers

\[ y = 1.0024x + 18.796 \]

\[ R^2 = 0.9992 \]
6. REFERENCES


dependent signal transduction pathway under drought and high-salinity conditions."

Urao, T., Katagiri, T., Mizoguchi, T., Yamaguchi-Shinozaki, K., Hayashida, N. and Shinozaki, K.
(1994). "Two genes that encode Ca$^{2+}$-dependent protein kinases are induced by drought and
high-salt stresses in Arabidopsis thaliana." Molecular and General Genetics MGG 244(4):
331-340.


Van Hengel, A. J. and Roberts, K. (2002). "Fucosylated arabinogalactan-proteins are required

Physiology 120(2): 391-400.

small fruit crops across the eastern United States associated with the April 2007 freeze”.

affecting the development of freezing tolerance in Arabidopsis thaliana (L.) Heynh." Plant
Physiology 111(4): 1011-1019.


rice orthologue of SENSITIVE TO FREEZING-6 and can act as a regulator of COR gene
expression, osmotic stress and freezing tolerance in Arabidopsis." New Phytologist 191(4):
984-995.


oat and rye: two cereals at the extremes of freezing tolerance." Plant Physiology 104(2):
467-478.
Xin, Z. and Browse, J. (1998). "eskimo1 mutants of Arabidopsis are constitutively freezing

Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt 

"Properties of aspartate racemase, a pyridoxal 5'-phosphate-independent amino acid 

regulated receptor-like kinase CRLK1 interacts with MEKK1 in plants." Plant Signaling and 


Zablackis, E., York, W. S., Pauly, M., Hantus, S., Reiter, W.-D., Chapple, C., Albersheim, P. and 
Science 272(5269): 1808.

mediates cold tolerance through a CBF-independent pathway." Proceedings of the National 